IMPROVING PREDICTIVE PROCESS ANALYTICS WITH DEEP LEARNING AND XAI

PRINCEWILL CHIMA OBUZOR

School of Science, Engineering and Environment

University of Salford, Salford, UK

Submitted in partial fulfilment of the requirements of the degree of Doctor of Philosophy

October 2023

Table of Contents

LIST OF TABLES

LIST OF FIGURES

LIST OF EQUATIONS

Acknowledgments

I begin by expressing my profound and heartfelt gratitude to the Almighty, the guiding force behind every step of my PhD journey. His boundless kindness, mercy, and unwavering guidance have been my beacon in the most challenging times. The culmination of this thesis is not just an academic achievement but is also a testament to the collaborative spirit of those who have been part of this voyage. Prof. Mohammed Saraee, my primary supervisor, stands tall as a pillar of support and knowledge. His patience, mentorship, and consistent guidance have been invaluable. His commitment to ensuring my academic success and personal growth has played a pivotal role from the inception to the culmination of this research. Every discussion with him has been an enlightening experience, and I cannot emphasize enough the depth of my gratitude.

Dr. Azadeh Mohammadi, my esteemed joint-supervisor, has been another foundational pillar of strength and knowledge. Her unwavering support, keen insights, and constructive feedback have greatly enriched my research. Her dedication to my academic pursuits and her belief in my capabilities have been both humbling and motivating. The backbone of my personal journey has been my family. Their unyielding love, belief in my potential, and consistent encouragement have been my anchor throughout. My wife, Ms. Sophia Agwaonye, has been a reservoir of strength, patience, and love. Along with our son, Yagazie Joel Obuzor, they have been the wind beneath my wings, providing daily motivation and solace. My parents deserve special acknowledgment for instilling in me a belief system that reverberates with the principle that with determination, any summit can be scaled.

Before I conclude, I want to thank my friends, Mushin, Thibaut, and Fatai, whose friendship has added so much to my academic journey. Their support and shared moments have made my PhD journey colourful. Lastly, I would be remiss if I didn't extend my deepest appreciation to the University of Salford. This esteemed institution has not only provided me with a platform to pursue my academic endeavours but has continually enriched my journey with workshops, trainings, and an environment conducive to growth and excellence.

Declaration of Authorship

I, Princewill Chima Obuzor, hereby affirm that the thesis entitled "Improving Predictive Process Analytics with Deep Learning and XAI" and the research presented therein are solely my own contributions. No portion of this work has been submitted in pursuit of any degree or qualification at the University of Salford or any other educational institution.

Throughout the course of my research, I have upheld the highest standards of professional integrity, adhering to both the Institutional Code of Practice and the Regulations for Postgraduate Research Degrees.

I have duly acknowledged any referenced work from other authors. Whenever I have cited the work of others, the source is explicitly mentioned. I was funded by TOTAL E&P NIGERIA for this research, and I declare that I have no conflicts of interest.

List of abbreviations

XAI - Explainable Artificial Intelligence

AI - Artificial Intelligence

MIMIC - Medical Information Mart for Intensive Care

PM - Process Mining

ML - Machine Learning

RL - Reinforcement Learning

IT - Information Technology

MLP - Multi-Layer Perceptron

NLP - Natural Language Processing

GDPR - General Data Protection Regulation

DPIA - Data Protection Impact Assessment

DPA - Data Protection Authority

NHS - National Health Service

MIT - Massachusetts Institute of Technology

CITI - Collaborative Institutional Training Initiative (often associated with human research training)

BPIC - Business Process Intelligence Challenge

ED - Emergency Department

LSTM - Long Short-Term Memory

ANN - Artificial Neural Network

CNN - Convolutional Neural Network

ABSTRACT

In this doctoral thesis, we explore the innovative application of the Tab Transformer architecture in the realm of predictive process mining, marking a significant advancement in forecasting subsequent events within activity sequences. Utilising the PM2 methodology, known for its structured approach in process mining, this study rigorously handles data processing, model development, and validation. This methodological choice is pivotal in leveraging the unique capabilities of the Tab Transformer, particularly its proficiency in processing multiple categorical features, a dimension often overlooked in previous research. The empirical analysis encompassed a novel dataset and extended to three additional publicly available datasets: MIMIC-IV Emergency Department (ED) Data, BPIC 2012, BPIC 2013, BPIC 2017, and BPIC Road Traffic. The model's performance was exemplary, achieving accuracies of 0.69, 0.812, 0.7301, 0.8766, and 0.78, and F1 scores of 0.67, 0.77, 0.70, 0.8533, and 0.734 in these datasets, respectively.

A major contribution of this research is the introduction of the Tab Transformer to process mining, a first in the field. This approach not only demonstrates the model’s versatility across various data forms but also highlights the importance of integrating categorical features in process mining, providing a more nuanced understanding of the influencing factors in activity sequences.

The thesis further distinguishes itself through the application of Explainable Artificial Intelligence (XAI) techniques, particularly SHAP and LIME. These tools were instrumental in demystifying the model’s decision-making processes, thereby enhancing its transparency, and fostering trust in AI systems. This integration challenges the notion of AI as impenetrable "black boxes," paving the way for AI systems that are not only effective but also interpretable and trustworthy.

In conclusion, this thesis contributes significantly to the field of predictive process mining by pioneering the use of the Tab Transformer, emphasizing the role of categorical features, and advancing the cause of transparency in AI through the application of XAI. The findings and methodologies established in this study represent a benchmark for future research in this evolving domain.

CHAPTER 1:INTRODUCTION

1.1 Introduction

Process mining has emerged as a ground-breaking interdisciplinary realm that bridges data science with traditional process modelling. This innovative approach fosters a unique way of identifying and enhancing business processes. As articulated by Van Der Aalst, often considered the patriarch of process mining (Van der Aalst, 2009), the primary aim of process mining is to discover, monitor, and improve authentic processes by harnessing knowledge from event logs in information systems. While some of these logs are well-structured, others might be erratic and inadequately measured, necessitating advanced pre-processing for effective process mining. With the continuous evolution of this domain, the emphasis on improving process visualization through research on process discovery and enhancement has become paramount. Notably, several challenges, like prescriptive capabilities and complex data structures, have emerged, leading researchers to pivot towards predictive process monitoring (R'bigui & Cho, 2017; Van Der Aalst, 2012).

Predictive process monitoring, a subset of process mining, is focused on forecasting process behaviours. Current research in this subfield addresses areas like next activity prediction, cycle time prediction, and outcome prediction. Such areas not only shed light on current processes but also anticipate future behaviours, empowering businesses with foresight (Marquez-Chamorro *et al.*, 2018a).

Recent strides in deep learning have ushered in solutions for the challenges in next event prediction. Earlier studies have leveraged LSTM, ANN, and CNN for these prediction tasks with varying results (Tama & Comuzzi, 2019), However, as outlined in a 2021 publication, the industry faces challenges like trustworthiness of insights and lack of domain knowledge (Martin *et al.,* 2021). The rise of the transformer architecture, particularly in tasks related to translation and generation, has spurred research into its potential for improving next event predictions. A notable study on this topic is by (Kim *et al.,* 2022). which discusses the importance of understanding both sequence predictions and supporting categorical features.

To address the challenges, specifically in next event prediction, recent advancements in deep learning have proposed various solutions tailored to the next event prediction tasks. Precious researchers have applied LSTM, ANN, CNN on the Next event prediction tasks with varying degree of success (Tama & Comuzzi, 2019), But with the emergence of several recent process mining challenges and opportunities as described by the publication as at 2021 on the state of the industry and the need for future research to address the lack of trust in insights, insufficient prescriptive capabilities, enriching domain knowledge and enhancing business process transparency (Martin et al., 2021). This has led to further research in improving the outcomes of next event predictions task. With the emergence of transformer architecture on the handling of translations and generative tasks, it has led to further research on the applicability of method to improve the outcomes of generated next event processes, to understand prediction and generative further insights by understanding not only the sequence for prediction but also the supporting categorical features present in the complex data, which is highlighted in the paper by (Kim *et al.,* 2022).

Presently the traditional architecture, was recently applied to the next event prediction tasks, and showed an improved level of prediction accuracy compared to previous models, but similarly to previous applications, little emphasis has focused on the other set of categorical features in the data. As with any course of events, the data at each stage is associated several categorical features of events that occurred during such events, while also enabling the research to trust in insights gained from the prediction.

Previous research has focused on the applicability of XAI approaches on the transformer approach focusing on text to text, image to image interpretation. But these approaches are not focused on its applicability to the Tab Transformer approach, as the approach aims to gather insight form all available features. The thesis aims to further explore the applicability of various and their applicability to the Tab Transformer architecture this was advised by the work done by (Velioglu *et al.*, 2022).

In the realm of predictive process monitoring, a notable gap persists in the monitoring of subsequent event activities, especially when integrating diverse categorical features to forecast an activity. This research seeks to address this lacuna by introducing a pioneering approach leveraging the TAB Transformer Architecture (Huang *et al.,* 2020). While Huang et al. illuminated a pathway that aligns with our objective of harnessing the multifaceted features present in the data for prediction, our research extends and refines this trajectory.

Historically, the emphasis in this domain has been predominantly on replicating the sequence of events, aiming to predict the subsequent sequence and its order. The methodology used, however, diverges significantly. By employing the TAB Transformer Architecture complemented with trace Position Embedding, the research work does not only predict the next event activities but delves deeper into understanding the myriad factors contributing to such predictions. This nuanced approach facilitates a richer comprehension of the underlying dynamics, offering stakeholders a more granular view of the processes.

Furthermore, while the potential of Transformer models in predictive process monitoring is increasingly recognized, their interpretability remains a relatively uncharted territory. (Velioglu *et al.,* 2022) underscored this research void, emphasizing the imperative for a deeper exploration of Transformer model interpretability. This research not only acknowledges this gap but actively ventures into this domain, aiming to elucidate the intricacies of Transformer-based predictions, thereby enhancing the transparency and trustworthiness of the model's insights.

In essence, this research stands at the forefront of predictive process monitoring, offering a novel, its comprehensive approach promises both accuracy in predictions and depth in interpretability.

1.2 Research Problem

The field of predictive process mining, as underscored by influential works like (Martin et al., 2021; Munoz-Gama et al., 2022a; R'Bigui & Cho, 2017) confronts numerous challenges that hinder its broader application and efficacy. Central among these challenges are issues related to the trustworthiness of insights generated from process mining models, the need for enriching domain-specific knowledge through data, and enhancing the transparency of business processes. These challenges are particularly pronounced in the realm of next-event prediction, a critical area in predictive process mining that seeks to forecast future activities based on past patterns.

Historically, efforts in next-event prediction have largely revolved around deep learning models, which, while showing promise, exhibit varying degrees of accuracy and often lack transparency. This inconsistency in performance and the "black box" nature of these models contribute to a broader issue of trust and understanding in the insights they generate (Bathaee, 2018; Padovan et al., 2023). One significant gap in existing approaches is the underutilisation of the Transformer architecture, especially in handling complex event logs in process mining. This architecture, known for its success in natural language processing, has not been fully explored in the domain of predictive process mining, representing a missed opportunity for enhancing model accuracy and interpretability.

This thesis proposes to address these gaps by implementing and investigating the Tab Transformer architecture. This approach is novel in the predictive process mining field and offers a promising avenue for tackling the aforementioned challenges. The Tab Transformer is particularly adept at integrating and analysing comprehensive data (Huang et al., 2020), including the incorporation of multiple categorical features – a critical aspect often overlooked in previous studies (Elkhawaga *et al.,* 2022; Guidotti *et al.,* 2018; Ribeiro *et al.,* 2021). By focusing on these categorical features, the thesis aims to provide a deeper, more nuanced understanding of the factors influencing activity sequences in process mining.

Furthermore, the adoption of the Tab Transformer architecture is anticipated to enrich the domain knowledge developed from data, thereby enhancing the transparency and reliability of business process predictions. This approach aligns with the growing need for predictive models that not only offer high accuracy but also afford users a clearer understanding of how predictions are made. In doing so, this research seeks to contribute significantly to the field of predictive process mining by bridging the gap between advanced machine learning techniques and the practical, transparent application of these techniques in business process analysis.

1.3 Research Aims and Objectives

In this section, the thesis aims and objectives would be described, the aims will be an overview of what the research seeks to achieve, and the objectives will be a breakdown of steps for the study.

1.3.1 Research Aims

The primary aim of this research is to significantly enhance the predictive accuracy and interpretability of process mining models, with a specific focus on next-event prediction. This enhancement is not merely about achieving higher accuracy rates but also about providing a more comprehensive understanding of the data-driven processes that inform these predictions. To achieve this, the study will incorporate a broader range of data features, including those that have been traditionally neglected or underutilised in predictive process mining. By doing so, the research aims to uncover deeper insights into the various factors that influence activity sequences, thereby providing a more nuanced and accurate representation of business processes.

Additionally, this research aims to delve into the legal and regulatory aspects of AI and machine learning models, particularly those considered as "black box" models. This involves exploring the existing legal frameworks and regulatory guidelines that govern the use and implementation of such models, especially in contexts where transparency and accountability are paramount. The research will examine how these models are currently perceived and regulated, focusing on the UK context while also drawing comparisons with global standards. The objective is to understand the implications of these legal and regulatory landscapes on the development and deployment of predictive process mining models and to identify best practices for ensuring transparency and accountability in AI applications.

1.3.2 Research Questions

Research Question 1:

Comparative Analysis of Tab Transformer's Accuracy: How does the accuracy of the Tab Transformer in next-event prediction within predictive process mining compared to contemporary methods? This question aims to evaluate the effectiveness of the Tab Transformer when it incorporates a more comprehensive range of data features. The focus is not only on the accuracy but also on how the integration of these diverse data features can enhance prediction quality compared to existing methods.

Research Question 2:

Insights from Categorical Features: What additional insights can be gained from the inclusion of categorical features that have been previously overlooked in next-event prediction models? This question seeks to understand the impact of integrating a wider array of data features, particularly categorical ones, on the predictive capabilities of the models and the richness of the insights they provide.

Research Question 3:

Enhancement Through XAI Techniques: How do different Explainable Artificial Intelligence (XAI) techniques enhance the interpretability of the Tab Transformer in predictive process mining, and what specific insights do these techniques unveil? This question aims to explore the effectiveness of various XAI methods in making the Tab Transformer’s decision-making process more transparent and understandable, thereby assessing their contribution to enhancing the model's interpretability.

Research Question 4:

Efficacy of Specific XAI Techniques: Among the XAI techniques applied to the Tab Transformer, which ones most effectively elucidate the model's decision-making process, and what are the unique advantages and limitations of each technique? This question is focused on identifying and evaluating the specific XAI techniques that provide the most clarity and insight into the Tab Transformer's operations, considering their distinct benefits and constraints.

Research Question 5:

Legal and Regulatory Landscape: What is the current legal and regulatory landscape regarding developer accountability for black box models, particularly in the UK, and how do these regulations address the need for transparency in such models? This question aims to investigate the existing legal framework and regulatory guidelines that govern the use of AI models, with a particular focus on the issues of transparency and accountability in the UK, and to compare these with international standards. The goal is to understand the challenges and requirements for legally compliant and transparent AI model deployment.

1.3.3 Research Objectives

The objective of this study are as follows:

Enhancing Predictive Performance with Diverse Data Features:

Objective: To significantly improve the predictive capabilities of process mining models, specifically in the context of next-event prediction. This enhancement will be achieved by integrating a broader and more diverse spectrum of data features than traditionally employed in such models.

Approach: The study will systematically incorporate and analyse various data features, including but not limited to categorical and numerical features, to understand their collective impact on the accuracy of predictions. This approach aims to improve the conventional boundaries of data utilisation in predictive process mining, thereby providing a more comprehensive and multifaceted understanding of the factors influencing activity sequences.

Evaluating the Tab Transformer's Performance:

Objective: To conduct a detailed comparative analysis of the Tab Transformer’s performance against existing predictive models in process mining. This will involve a thorough evaluation based on key performance metrics, such as accuracy and F1 scores.

Approach: The research will entail a series of experiments where the Tab Transformer will be benchmarked against current state-of-the-art models. The comparison will focus on both the quantitative measures of model performance and the qualitative aspects of the predictions, such as their relevance and applicability in real-world scenarios.

Exploring the Efficacy of XAI Techniques:

Objective: To investigate and ascertain the effectiveness of various Explainable Artificial Intelligence (XAI) techniques in enhancing the interpretability of the Tab Transformer, particularly in the realm of predictive process mining.

Approach: This objective will involve applying different XAI methods to the Tab Transformer and evaluating their effectiveness in making the model's decision-making process more transparent. The study will analyse how these techniques contribute to a deeper understanding of the model's predictions and their implications on the overall business process.

Analysing Legal and Regulatory Frameworks for AI Models:

Objective: To conduct a comprehensive analysis of the legal, regulatory, and guideline frameworks that govern black box AI models, with a specific focus on the implications for transparency and accountability.

Approach: This aspect of the research will explore the current legal and regulatory landscape, particularly in the UK, regarding the use and development of AI models. It will involve an examination of existing laws, guidelines, and regulations that address the transparency and accountability of AI systems. The study will also compare these UK frameworks with international standards to identify best practices and potential areas for improvement in the governance of AI models.

1.4 Research Contributions

At the intersection of process mining and machine learning, this research delves into the specialised domain of predictive process monitoring. The objective is a methodological enhancement for predicting subsequent events, especially when navigating complex "black box" models. The research provides a breakdown of the contributions below.

Innovating with Tab Transformer Application Dataset. This research pioneers the use of the Tab Transformer for predictive process mining in the data-rich medical sector, demonstrating its versatility across various datasets, including complex, high-dimensional structures. This work highlights the Tab Transformer's adaptability and potential for transformative impacts in process mining, setting a new standard for its application in diverse contexts.

Broadening Data Analysis Horizons:

A key achievement of this study is the enriched analysis process through the inclusion of previously overlooked categorical features, enhancing the predictive models' robustness and accuracy. This expanded feature set captures the full complexity of underlying processes, offering a truer representation of real-world phenomena.

Employing the MIMIC_IV_ED Dataset:

Utilising the MIMIC_IV_ED dataset marks a significant advancement in process mining, introducing new dimensions to the field with its size, diversity, and detail. This dataset enables the exploration of the Tab Transformer's effectiveness in real-world healthcare, bridging the gap between theory and practice.

Advancing Model Interpretability with XAI:

This research enhances interpretability through XAI techniques like SHAP and LIME, moving beyond sequence tracing to a comprehensive understanding of data and model decisions. This approach increases transparency, trust, and acceptance among users, offering valuable insights on the application and effectiveness of various XAI methods in predictive process mining.

Cross-Validation: K-fold cross-validation is applied, improving the Tab Transformer's accuracy significantly. This technique is crucial for detecting overfitting and ensuring robust model performance, especially beneficial for datasets with a limited number of labelled samples.

Data Transformation Techniques: Incorporating One Hot Encoding and factorization methods allowed for efficient handling of categorical data, enhancing model input quality and performance.

These contributions advance technical capabilities and address transparency, trust, and applicability concerns in AI for critical sectors such as healthcare.

Impact on Process Mining:

The application of the Tab Transformer, particularly in the medical sector, has shown transformative impacts, such as enhanced process understanding and improved prediction accuracy. This technology's ability to contextualize data elements through its attention mechanism also facilitates the development of explainable AI, opening new pathways for intuitive AI capabilities.

Dataset Characteristics:

The MIMIC_IV_ED dataset exemplifies the richness and applicability of our chosen dataset for evaluating the Tab Transformer, showcasing the variety of medical conditions covered and demonstrating the model's potential in real-world healthcare settings. This reflects the dataset's suitability for challenging predictive models and its contribution to advancing process mining research.

In general, with an emphasis on real-world applications and methodological developments, this study provides a thorough framework for improving predictive capabilities in process management by integrating cutting-edge deep learning techniques and XAI methodologies.

1.5 Thesis Structure

The ensuing sections of this thesis have been organised systematically to offer a comprehensive understanding of the research topic. The layout of the thesis is designed to sequentially progress from a foundational understanding of the subject to an in-depth analysis, ultimately leading to a conclusive summary. The chapter two of this thesis deals with proving understanding of the terms used, following with the introduction as foundation for the research explore process mining, predictive process monitoring, and the research carried out in the field. The research also delves into the laws and current legal stance on black box models in the United Kingdom and comparing it with other parts of the world.

In chapter 3, the methodology for the research is introduced and the case studies for its applications is examined in chapter 4. On various datasets starting with the primary data, then moving the other publicly available datasets. In chapter 5, the results are presented as obtained in the research and compared with previous works, so as to enrich the overall understanding of the study. Then finally, the conclusion is presented as a summary of all the research work, encapsulating the major findings, implications, and contributions of the research in chapter 6. In general, the introduction of this research in presented in this chapter and the next chapter contains the literature review of the study.

CHAPTER 2: LITERATURE REVIEW

2.1 Background and Overview

The preceding chapter contains the general introduction of this research and the current chapter begins by elucidating the definitions of key terms and concepts fundamental to this thesis. Following this foundational introduction is an in-depth literature review, mining process exploration and the nuance of predictive process monitoring. Subsequently, myriad applications of predictive process monitoring, inherent challenges and contextual issues are discussed. The discourse will further extend to the methodologies of applying predictive process monitoring and the diverse approaches involved. The exploration will also encompass 'black box' models, detailing strategies for deriving insights for enhanced explainability. The chapter will be concluded with a discussion on pertinent legislations and regulations, underscoring the imperative nature of these approaches in the contemporary research landscape.

2.2 Definition of terms

Event Represents a recorded occurrence or action within the activity. Events provide a detailed log of what transpired during the execution of a process, capturing each significant interaction or change.

An **Event log** is a dataset that records sequences of events, each associated with a particular case, occurring in a process over time. It serves as the foundational input for process mining techniques (Van der Aalst, 2016).

Process: This element represents a specific business process or workflow within the log. It can contain multiple instances of that process, detailing its various executions over time.

Process map graphically represents the flow of activities within a process, illustrating the sequence of events, the roles or groups involved, and the decision points.

In process mining, a case represents an instance of the process that traverses through a sequence of activities from start to finish. Each case is uniquely identified and contains a series of events recorded in the event log.

An **Instance**, synonymous with a case in process mining terminology, represents a singular execution of a process, comprising a sequence of events recorded in the event log (Van der Aalst, 2016).

Activity, describes a particular task or action within the process instance. Activities are the building blocks of a process, each serving a specific function or purpose.

Process Mining (PM) is a technique used for analysing and optimizing processes based on the data generated by software systems.(Van der Aalst, 2016)

Process models in software are defined as simplified representations of a software process with each model depicting a process from a particular perspective. A process can also be defined as a series of steps and decisions involved in the way a work is completed

Predictive Process Monitoring (PPM) refers to the utilization of process mining techniques alongside machine learning to predict the outcomes of ongoing instances of a process (Marquez-Chamorro et al., 2018b).

Interpretability in machine learning and artificial intelligence refers to the extent to which a human can understand and trust the outputs of models and the processes by which they arrive at decisions.(Doshi-Velez & Kim, 2017)

Explainability refers to the degree to which a machine learning model's operations and results can be articulated in understandable human terms.(Adadi & Berrada, 2018)

Encoding is the process of converting data from one form to another. In the context of machine learning, it often refers to the transformation of categorical data into a numerical format that can be utilized by algorithms(Zhao et al., 2015).

Embedding refers to the representation of data in a lower-dimensional space, often used to capture semantic relationships in data like words or items. In machine learning, embeddings are used to reduce the dimensionality of input data to make the learning process more efficient (Mikolov et al., 2013).

2.3 A Review on Process mining

To discuss the application of Process Mining, it is logical to explicitly define process mining and introduce the different stages involved in it. Process mining (PM) is a set of techniques that assist the analysis of business processes based on event logs and is the link between model-based process analysis and data-oriented analytical approaches (Qafari & Aalst, 2019). It is a developing field of study that offers methods for comprehending and enhancing processes in various application domains (dos Santos Garcia, et al., 2019). One of the objectives of PM is to extract non-trivial process-related data from event data logged by the available information systems (Sato, de Freitas, Barddal, & Scalabrin, 2022). Process discovery, which looks for a descriptive model of the underlying process in event logs, conformance checking, which monitors and examines whether the actual execution of the process conforms to the corresponding designed (or discovered) reference process model, and process enhancement, which enhances and enriches a process model based on the related event data, have all been proven to be very successful applications of PM techniques (Macak, Daubner, Sani, & Buhnova, 2022).

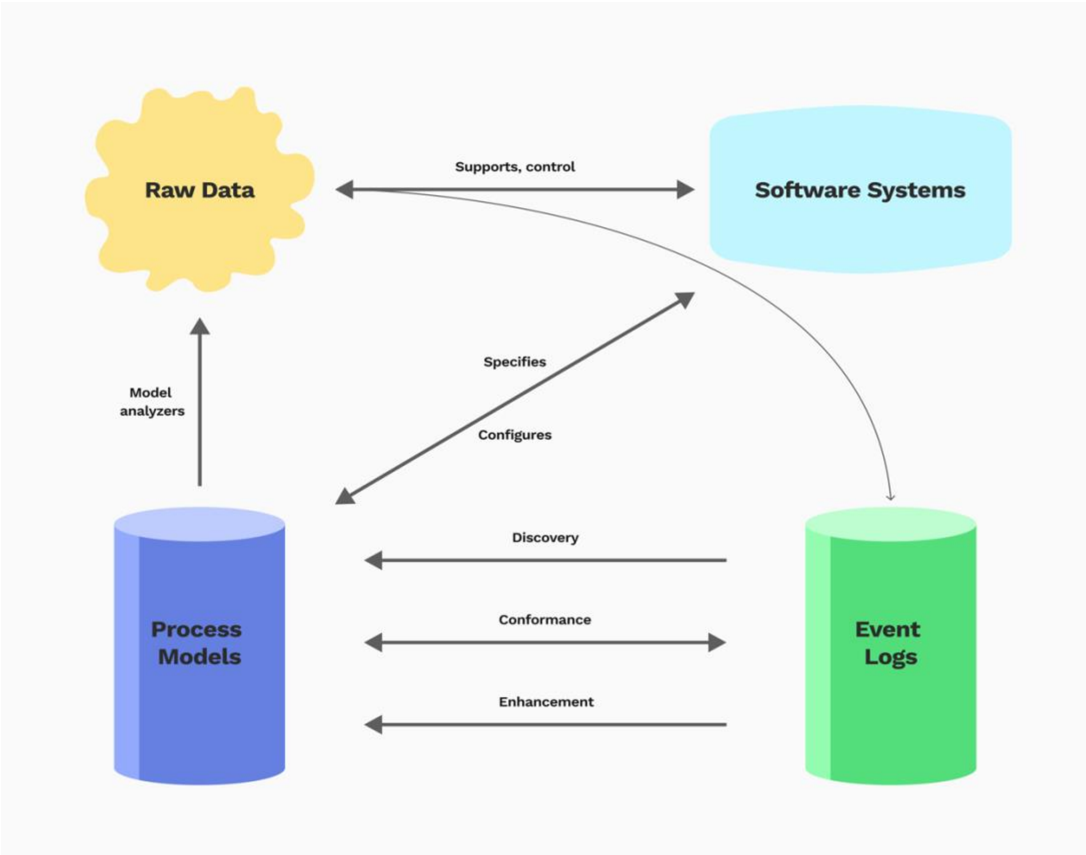

Figure 2-1 stages of process mining (W.van der Aalst et al 2012)

Figure 2.1 visualizes the intricate interplay of stages involved in process mining, a systematic approach to analyse and refine business processes. Starting with 'Raw Data', the foundational layer of information, it is transformed and interpreted through various 'Models analysers' to create 'Process models'. These models represent the underlying processes in a structured manner. Simultaneously, modern 'Software systems' generate 'Event Logs' which are records of executed tasks or activities, an example of this is depicted in Figure 2. 2

```
<WorkflowLog>
        <process>
                <processInstance>
                        <DataInstance>
                                <Activity></Activity>
                                <event></event>
                                <timestamp></timestamp>
                                <DataAttribute ></DataAttribute>
                        </DataInstance>
                </processInstance>
        </process>
</WorkflowLog>
```

Figure 2-2 sample of an event log

In the realm of process mining, three primary analyses are performed on these logs:

- **Discovery**: The discovery phase is central to process mining, focused on uncovering the underlying process model from an event log without any prior information about the process. In this stage, algorithms and techniques are employed to generate a model that accurately describes the patterns and sequences observed in the log data. One of the most well-known techniques used in this phase is the "α-algorithm" introduced by (Van Der Aalst *et al.,* 2004). This algorithm identifies causal relationships between activities and translates them into a Petri net—a graphical representation of the discovered process. However, numerous other techniques and algorithms have been proposed, each with its advantages and complexities, to address the challenges posed by noisy or incomplete log data (Weijters & Ribeiro, 2011).

- **Conformance**: The conformance phase is about comparing and contrasting the discovered or an existing process model with the actual log data to identify any discrepancies. This helps in understanding whether the processes occurring in the real world align with the established model or if there are deviations. Misalignments can arise due to errors, fraud, or unanticipated behaviour. Conformance checking provides insights into these deviations, allowing organizations to pinpoint bottlenecks, policy violations, or inefficiencies (Rozinat & van der Aalst, 2008). For instance, the "token-based replay" is a technique often employed to evaluate how well a model fits with the log by simulating the process and comparing actual sequences with those in the model (Verbeek *et al.*, 2011).

- **Enhancement**: Building upon the insights gained from the discovery and conformance phases, the enhancement phase aims to improve the existing process model. This could involve refining the model to better represent the real-world processes, annotating the model with performance data, or recommending modifications for optimization. The enhancement phase is not just about model accuracy, but also about deriving actionable insights that can lead to process improvements. Techniques employed in this phase can help in predicting bottlenecks, analysing resource allocations, or recommending process redesigns. For example, (Van der Aalst *et al.,*

2012) introduced an approach where performance metrics are integrated with the process models, allowing for a more detailed analysis of process flow and resource engagements.

Figure 2.1 also depicts how software systems, through their mechanisms, can support and control the overall mining process, ensuring alignment between the modelled processes and real-life executions. The flow and relationships between these components demonstrate the cyclical nature of process mining, emphasizing continuous improvement and evolution based on insights derived from real-world data. Now that the logic and processes behind process mining has been explained, its applications shall then be thoroughly examined.

In the field of business management, PM has moved away from being seen as merely a tool for investigating process performance issues to a comprehensive platform for monitoring and augmenting operational process execution. According to the Harvard business review by lars Reinkemeyer and Tom Davenport, PM is instrumental in unveiling bottlenecks and other areas for improvement in business processes, thereby fostering a culture of continuous improvement (Reinkemeyer, 2020). Process mining vendors have claimed a reduction in automation implementation time by 50% through process discovery, showcasing PM's potential in accelerating digital transformation initiatives within organizations. The application of data science for discovering, validating, and enhancing workflows underscores PM's profound impact in driving operational excellence (Dumas et al., 2018; Van der Aalst, 2013).

In the finance sector, the emergence of process mining has heralded a new era of data-driven operational insight and efficiency enhancement. The technique is notably applied in financial audits, where it is conceptually embedded to augment the auditing profession with the prowess of data science, thus bridging a critical gap between traditional auditing methods and contemporary data analytics approaches (Werner et al., 2021). Beyond auditing, process mining unveils a realm of possibilities in optimizing financial operations by employing advanced technologies like computer vision and AI for automated process discovery, path mapping, and bottleneck identification, thereby fostering a culture of continuous improvement and operational excellence. The ability of process mining to provide a real-time, data-centric view of finance processes is particularly instrumental for finance leaders, enabling them to discern improvement opportunities, visualize finance process flows, and objectively pinpoint process inefficiencies, which is paramount for propelling finance process efficiency to new heights (Reinkemeyer, 2020; Werner, 2017). Moreover, the versatility of process mining is exhibited in its applicability to a spectrum of financial processes including account payables, receivables, and procurement, where it plays a pivotal role in enhancing process efficiency by unearthing and addressing bottlenecks4. In the ambit of financial audits, process mining algorithms demonstrate their indispensability by scrutinizing source event logs to craft reliable process models, thereby significantly elevating the efficacy and efficiency of audit processes (Werner, 2017). Through these diverse applications, process mining delineates a path towards a more analytically driven and efficient financial ecosystem, underpinning the transformative potential of data analytics in the finance sector.

The PM application can also be observed in the manufacturing and logistics sector; it serves as a linchpin for unravelling the as-is processes by analysing event logs from IT systems, consequently fuelling the digital transformation towards a data-driven shop floor. The ramifications of PM adoption are profound, encapsulating higher throughput, improved machine utilization, and a notable reduction in non-conformance costs. Moreover, PM's capability to identify and ameliorate production bottlenecks and unnecessary process steps holds a significant promise for trimming operational inefficiencies and continuously refining production processes based on data-derived insights (Gunnarsson et al., 2019; Intayoad & Becker, 2018). A specific study delineated the utility of PM in improving productivity in make-to-stock manufacturing, showcasing the potential of PM in enhancing the manufacturing domain's operational efficiency (Lorenz et al., 2021).

This ability of PM not only optimizes the overall supply chain performance but also reduces operational costs, thereby significantly contributing to the operational excellence in the manufacturing and logistics sectors. In the Field of Information technology, PM applications have been applied in various capacity from improving cybersecurity to software reliability. Studies have illustrated the application of PM techniques in the software development process, where it improved the process evaluation and auditing (Keith & Vega, 2017). A different study showed the application of PM in auditing and tackling cybersecurity and software reliability issues, thus showcasing the PM potential in addressing domain specific challenges within the IT sector (Zerbino et al., 2018).

Finally, In the healthcare sector, the application of Process Mining (PM) emerges as a quintessential tool for navigating the complex and dynamic nature of healthcare processes. The inherent characteristics of healthcare processes such as data privacy concerns, process variability, and the presence of multi-disciplinary teams pose unique challenges that necessitate a robust analytical framework like PM (Munoz-Gama et al., 2022b). A systematic review highlights the broad spectrum of PM applications within healthcare settings, shedding light on how PM can be employed to delve into the intricacies of healthcare processes, thus providing a structured pathway towards operational excellence and quality improvement (Dallagassa et al., 2021; Rojas et al., 2016a).

The evolving nature of PM applications in healthcare is captured through recent scholarly works, showcasing the advancements in PM techniques tailored to address the distinctive challenges faced by the healthcare sector (De Roock & Martin, 2022). The societal value of PM is underscored through its application in healthcare problems, which often serve as the impetus for demonstrating new PM techniques aimed at improving healthcare processes and patient outcomes (Martin *et al.,* 2022).

Moreover, PM finds substantial application in modelling healthcare processes, thereby elucidating a clear representation of healthcare workflows. These models serve as the backbone for devising quality improvement strategies, optimizing resource allocation, and ensuring adherence to clinical guidelines, which are pivotal for enhancing patient care and overall healthcare delivery. The transition towards a more data-driven healthcare framework is facilitated by PM, enabling healthcare organizations to harness the power of data analytics for continuous improvement in process efficiency and regulatory compliance.
Furthermore, the role of PM in fostering a culture of transparency and accountability is paramount, especially in ensuring regulatory compliance. PM provides a transparent view of healthcare processes, facilitating compliance checks, and audits, thus aligning healthcare operations with regulatory standards. Additionally, the educational value of PM is significant, aiding in the training of healthcare professionals by providing a clear understanding of

healthcare processes and the implications of different practice patterns, thereby enriching the educational landscape in healthcare.

The cross-disciplinary application of PM in healthcare epitomizes the transformative potential of data analytics in modern healthcare operations. By navigating the complex operational landscapes, PM underscores its versatility and indispensability in fostering a data-driven culture, thereby marking a significant stride towards achieving the overarching goal of enhanced healthcare delivery and improved patient outcomes. PM has been applied in various ways as described before and further exploration of specific application in order to get a full grasp of the landscape was equally examined, but the introduction of some process discovery algorithms and tools applied in process mining are also necessary for discussion.

Process mining (PM) tools and methods have been utilized for more than a decade (Van Der Aalst, Process mining: Overview and opportunities, 2012; Emamjome, Andrews, & Hofstede, 2019), motivated by the necessity to continually enhance the effectiveness of business procedures as well as the creation of massive volumes of event data for ongoing processes. It seeks to establish a link between the system design that defines how to conduct a business process and the event log that records data on the actual deployment of process occurrences by a Process-Aware Information System (PAIS) (Aalst, et al., 2011). It provides insights into real process behaviour by analysing event logs and process models (Ghasemi & Amyot, 2020). Although process mining has emerged as a critical study subject in the recent decade, it still faces several problems acquired from its predecessor research areas. Investigation of concept drift is one of the key issues (Losing, Hammer, & Wersing, 2018). Data mining and business process modelling are connected by PM. From event data kept in information systems, process-related knowledge needs to be extracted (van der Aalst W. M., 2011).

2.3.1 Process Mining Algorithms and Tools

[8]Over the years, several algorithms have been developed to decode the intricacies of event data, and alongside them a slew of tools have been designed to facilitate this exploration.

The Alpha algorithm, one of the pioneers in the domain, is specifically designed to reconstruct processes in the form of Petri nets by analysing causal relationships among activities (Van Der Aalst et al., 2004). The strength of the Alpha algorithm lies in its deterministic approach, ensuring consistent Petri net outputs for identical event logs. However, it is not without its limitations. The algorithm often falters when confronted with complex loops and parallel constructs without explicit start or end points (Medeiros et al., 2004).

Recognizing the limitations of the Alpha algorithm, the Heuristic Miner was developed as an alternative (Van Der Aalst et al., 2004; Weijters & Ribeiro, 2011). Unlike its predecessor, the Heuristic Miner prioritizes flexibility. It establishes heuristic nets by mapping dependencies between activities, with a central focus on the frequency of relations as opposed to strict causality. This provides it with an edge in handling noise and infrequent behaviours in logs. Nonetheless, a trade-off exists, as the algorithm can sometimes oversimplify intricate processes.

In cases with extensive, diverse, and noisy logs, the Fuzzy Miner offers an alternative approach (Günther & Van Der Aalst, 2007). It employs adaptive techniques to visualize complex processes, focusing on both the frequency and significance of activities and paths. Through aggregation and abstraction of infrequent activities, the Fuzzy Miner delivers a high-level view of processes, although some granularity might be sacrificed in the process.
Complementing these algorithms are tools that enable practitioners to engage in process mining activities. Disco, a proprietary tool developed by Fluxicon, stands out due to its user-centric design. It supports a myriad of process mining tasks, from discovery to conformance checking, and offers advanced features like animations and performance analysis (Günther & Rozinat, 2012). Its ability to integrate diverse data sources makes it an invaluable asset for analysts.

Conversely, ProM provides an open-source alternative (Verbeek et al., 2011). Its modular framework is extensible, accommodating a wide array of plugins. Researchers and practitioners alike value ProM for its versatility, as it supports not only process discovery but also other facets like conformance checking and process enhancement. Now that the algorithms and tools have been introduced, studies in process mining can then be explored.

2.3.2 Process mining studies

Firstly, the techniques and algorithms need to be examined. Process mining (PM) techniques have evolved significantly over the past few decades, with researchers delving deep into various aspects of this domain. (Lassen & van Dongen, 2008) focused on the Petri net search technique for process mining, aiming to derive insights from log-captured activities. Their work provided a comprehensive overview of multiple petri net-based discovery techniques from both process mining and region concepts, further categorizing them into five algorithm types, each with unique assumptions and challenges.

However, it is worth noting that research in this domain is still in its nascent stage and needs further organization. For instance, feedback on the aforementioned work highlighted the need for a more structured approach, suggesting the inclusion of subsections to explain different PM methods, thereby fostering a clearer understanding.

Weber et al., (2013) delved into the evaluation of PM algorithms from a machine learning perspective. Their probabilistic framework examined the ground truth distribution across activity traces using selected sample groups. Another significant contribution came from (Dakic & Stefanovic, 2018) emphasizing the importance of business information extraction. Their study centred on an extensive review of journal articles over a ten-year period, discussing the application of business process mining in companies.

On a more niche front, (Mishra et al., 2018) explored the application of PM techniques in intrusion detection systems (IDS). Their findings highlighted the advantages of using PM in IDS, paving the way for more accurate threat detection.

Secondly, this research study would look at the application towards, improving the knowledge acquired from process maps and research into the fairness and bias in application (Mishra et al., 2018) embarked on an exhaustive exploration of conformance verification techniques, underscoring the necessity for refined methods capable of pinpointing both alignments and deviations between a process model and its matching event log. Their research illuminates the intricate interplay between event data and process models, underlining the potential for enhanced insights.

Elkhawaga et al., (2020) *further* advanced the conversation, focusing on the dynamics of concept drift in PM. Their pioneering work on the Concept Drift Analysis in PM framework elucidated the evolving nature of processes. This research emphasized that to truly harness the power of process mining, there's a pressing need for proactive prediction mechanisms that can anticipate changes, offering organizations the foresight to adapt.

Fairness, Acceptance, and the Drive towards Predictive Insights in Process Mining Process mining's potential is not solely confined to tracing and modelling processes; it extends to ensuring fairness and mitigating biases. (Qafari & van der Aalst, 2019) championed this very perspective with their introduction of a fairness-conscious approach. By shedding light on latent biases in traditional data extraction methodologies, they drew attention to the broader implications of PM and the necessity for insights that both reflect and respect diverse contexts.

In the work, carried out by (Grisold et al., 2021), provides a holistic challenge that pepper the path of process managers. From acceptance, sensemaking and usability of the applicability of process mining, their research also discussed the need to make sense of process related information, generate insights from the processes, their research echoes the sentiment that for PM to be truly transformative, it needs to be insightful, predictive, and above all, embraced by its users.

The research conducted by (Grisold et al., n.d.) also talks about the need for researchers to focus on capturing factors that contribute to the process, based on the paper, there is a need for further research into understanding the process that unfolds, the contributing factors and contextual factors.

The research paper by (vom Brocke et al., 2021), discussed the importance of gaining insights from processes, discuses the need for research to focus on understanding the discovery, explaining the outcomes, the research provided a view of process mining more towards process science. These views and shows the need for research into generating insights from processes.

The application so far has highlighted the range in opportunity for the further application of process mining, and the need for more context and insights, the healthcare sector has also received its fair share of research applications which we would explore further below.

(Van Dongen et al., 2009) highlighted the potential of Petri net search techniques for process mining, suggesting its prospects in building models to provide insights into log-captured activities. The work provides an overview of petri-net based discovery algorithm from both the area of process mining and theory of region. Their emphasis on the inadequacies of various techniques underscores the ongoing need for refinement in this domain.

However, what is notably relevant to healthcare is the application of PM for tangible improvements. (Gomes et al., 2021) underscored this by emphasizing the efficacy of Inductive Miner and Heuristic Miner algorithms in process mining, particularly within the context of healthcare processes and data. The study suggests that while there are variations in execution times across algorithms, some, like the Petri Net model, offer greater analytical depth.

Furthermore, two studies distinctly focused on the role of PM in healthcare. (Rojas et al., 2016b) conducted action research that mapped the application of PM across 22 healthcare categories. This study provides a comprehensive overview of how process mining can be applied to diverse healthcare segments. Concurrently, (Kurniati et al., 2016) delved into the application of PM in cancer treatments, signifying its role in addressing specific ailments.

Peleg, (2013) adds another layer to this discussion by investigating medical practical recommendations and highlighting process mining as a technique for principles management and compliance assessment. This research suggests that not only can PM be used for operational and analytical purposes, but also for ensuring adherence to best practices and guidelines in the healthcare domain.

However, while the potential benefits of PM in healthcare are evident, it is crucial to acknowledge the existing challenges. As highlighted by (Mishra et al., 2018), there is a significant gap between data collection methodologies and template operational processes. The inherent complexities of healthcare operations often mean that a one-size-fits-all approach is infeasible. There is a need for further insight into the processes, which the research aims to tackle in not just discovering the process but taking into account the various features contained in the data and making prediction on the next activities to occur and understanding what are the contributing factors that led to these predictions. Then, the predictive process mining, methods, and current challenges would be discussed.

2.4 Predictive Process Mining: Current Methods and Challenges

Predictive process mining represents a distinct branch of process mining that aims to discern and forecast the future trajectory of cases. It encompasses the prediction of various aspects, including the next outcome for a new case, the required completion time, and the forthcoming sequence of activities. The ability to make accurate predictions in these areas is currently a major driving force within the realm of predictive process mining, also referred to as predictive process monitoring, as these predictive insights hold immense value for organizations globally.

Historically, process mining has predominantly focused on visualizing processes, validating process adherence to discovered models, and identifying avenues for process enhancement. However, the field has progressively evolved to yield deeper insights for stakeholders. The growing demand for process mining, with influential industry players like IBM, AWS, and others venturing into this domain, underscores the promising growth potential and research opportunities in this field. Several tools have been developed while existing ones are being adapted and refined to cater specifically to predictive process mining. Examples of such tools include MyINvenio, Apromore, UiPath, IBM process simulation, and numerous others.

The primary objective of predictive process mining is to generate novel insights by leveraging historical data from prior process executions. These insights serve to provide valuable foresight into future events and predictions. The process typically involves a learning and training phase, followed by the prediction phase, wherein activities are forecasted based on the acquired knowledge.

The subsequent sections will delve deeply into the current methodologies employed in predictive process mining, elucidate the process through an illustrative example, and delineate the diverse challenges encountered in this field.

2.4.1 Overview of current methods and Approaches

Machine learning (ML), a prominent branch of artificial intelligence, has experienced significant growth and attention due to the increasing need for sophisticated analysis of complex data structures. As our capabilities in big data, cloud computing, and high-performance computing have expanded, so too have the methods and approaches in ML to harness these advancements (Jordan & Mitchell, 2015).

At a high level, ML algorithms can be categorized into several categories, primarily: supervised learning, unsupervised learning, semi-supervised learning, and reinforcement learning. These classifications essentially capture how algorithms learn patterns from data, depending on the type and amount of supervision they receive during training.

- **Supervised Learning:** This is the most common technique where an algorithm learns from labelled training data and makes predictions based on that learned knowledge. It requires a clear-cut framework of input-output pairs, essentially learning to map inputs to the correct outputs.

- **Unsupervised Learning**: Here, the algorithm is trained on unlabelled data and aims to uncover hidden patterns and structures from the data itself. Common applications include clustering and dimensionality reduction.

- **Semi-supervised Learning:** As the name suggests, this method falls between supervised and unsupervised learning. It uses both labelled and unlabelled data for training, typically a small amount of labelled data with a larger amount of unlabelled data. The idea is to leverage the large amount of unlabelled data to aid and improve the learning accuracy with the small amount of labelled data.

- **Reinforcement Learning (RL):** This is a type of learning where an agent learns to behave in an environment by performing actions and receiving rewards or penalties. It's more about learning optimal sequences of actions in interactive environments.

Having established these foundational categories, the subsequent sections will delve deeper into supervised learning algorithms, given their prevalent use and the specific context of this work. By diving into these methods in detail, we can gain a comprehensive understanding of their strengths, weaknesses, and suitability for various tasks.

2.4.2 Algorithms and Techniques in Supervised Learning

Given the framework of supervised learning, several algorithms and methodologies have been developed to capitalise on the clear-cut input-output pairs. These algorithms vary in their approaches, strengths, and applications. We would describe these approaches below starting with the tree-based algorithms then deep neural architectures.

XGBOOST

XGBoost, abbreviated for "eXtreme Gradient Boosting," is a pioneering ensemble algorithm rooted in decision-tree-based models and enhanced by a gradient boosting framework. Since its introduction by Chen & Guestrin in 2016, XGBoost has rapidly ascended in popularity, largely credited to its exceptional computational efficiency and standout model performance. Its prominence is further evidenced by its widespread adoption in machine learning competitions and real-world applications. Among its strengths, XGBoost is adept at handling sparse data, and it provides users with a multitude of hyperparameters for customization.

Central to XGBoost is the principle of 'boosting,' a technique where weaker learners, typically shallow trees, amalgamate to constitute a more potent learner. This algorithm meticulously constructs the model stage by stage, broadening its utility by facilitating the optimization of a diverse range of differentiable loss functions. XGBoost also incorporates regularization, a strategy to counteract the frequent overfitting challenge encountered in tree-based models, as outlined by (T. Chen & Guestrin, 2016).

RANDOM FOREST

Initiated by Breiman in 2001, Random Forest is an ensemble learning methodology. Its mechanism revolves around crafting a plethora of decision trees during the training phase. The end result, when classifying, is the mode of the classes derived from individual trees; for regression tasks, it's the mean prediction. Distinctively, Random Forest epitomizes the ensemble approach by harnessing the collective strength of numerous decision trees, sidestepping reliance on singular decision trees.

An inherent strength of Random Forest is its proficiency in processing vast datasets with high dimensionality. It's capable of navigating through thousands of input variables, pinpointing the ones with utmost significance. Consequently, it's esteemed as a potent tool for dimensionality reduction. Ensuring accuracy, Random Forest minimizes the risk of overfitting by offering an objective approximation of the generalization error, as highlighted by (Liaw & Wiener, 2002).

An added advantage of Random Forest is its adaptability to unbalanced and incomplete datasets. A particularly commendable attribute is its ability to evaluate feature importance in the prediction process, a key reason why it has become the preferred option for feature selection across a diverse range of application areas, as evidenced by (Díaz-Uriarte & Alvarez de Andrés, 2006).

Deep Neural Architectures

Deep Neural Architectures, as a subset of supervised learning techniques, have significantly influenced the domain of machine learning and artificial intelligence. These architectures can learn hierarchical representations from raw data, proving indispensable for complex tasks such as image recognition, natural language processing, and speech recognition. This segment will focus on specific deep neural architectures, including Recurrent Neural Networks (RNNs), Long Short-Term Memory Networks (LSTMs), and Transformers.

Recurrent Neural networks (RNN)

Recurrent Neural Networks (RNNs) are a specialized subset of artificial neural networks designed explicitly for processing and predicting sequential data. Originating in the 1980s, one of the foundational architectures is the Elman network or the simple RNN (Elman, 1990).

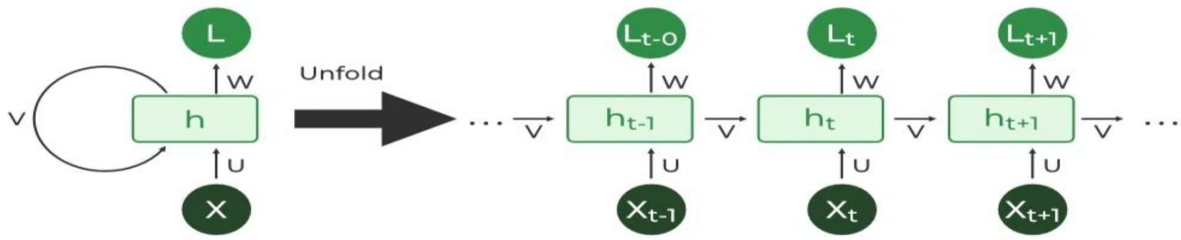

Figure 2-.2-3 Traditional Recurrent Neural network

At its core, RNNs have a unique architecture featuring a loop that enables them to maintain a 'memory' or 'state' from one step in a sequence to the next. This design allows RNNs to capture temporal patterns and dependencies in sequential data. Figure 2.3 is a visual representation of this architecture, depicting a single recurrent neuron and its unfolding over sequential time steps. This unfolding showcases the recurrent nature of RNNs, where the output from a previous step is used as an input to the current step, ensuring the continuity of information. Below is the Equation that depicts the RNN.

$$H_t = \sigma(W_{hh}h_{t-1} + W_{xh}x_{t-1} + b_h)$$

Equation 2-1 RNN- Step 1

$$y_t = W_{hy}h_t + b_y$$

Equation 2-2 RNN- Step 2

Equation 2.1 above, H_t is the hidden state at time t, x_t is the input, W_{hh}, W_{xh}, andW_{hy} are weight matrices, b_h andb_y are the biases, and σ is an activation function, often the hyperbolic tangent (Mikolov *et al.*, 2010).

Owing to their ability to understand sequence dependencies, RNNs have been used in: Language modelling and generation, as seen in the work of (Mikolov et al., 2010), Voice recognition, highlighted (Graves, 2013), the composition of music, explored and applied in the analysis of sentiments, as demonstrated by (Tang *et al.,* 2015).

Though RNNs have the theoretical capacity to capture long-term dependencies, they are prone to issues like vanishing and exploding gradients during backpropagation through time. This has led to the development of advanced RNN models, notably Long Short-Term Memory (LSTM) networks.

LONG SHORT-TERM MEMORY (LSTM)

Long Short-Term Memory Networks, commonly referred to as LSTMs, are a seminal advancement in the domain of Recurrent Neural Networks (RNNs). They were conceived by Hochreiter & Schmidhuber in 1997 as a countermeasure to the vanishing and exploding gradient problems inherent in conventional RNNs. Since their introduction, LSTMs have become pivotal in numerous sequence modelling endeavours due to their exceptional prowess in retaining long-range temporal dependencies. Figure 2.4 provides a schematic representation of the LSTM block, elucidating its intricate architecture. This diagram distinctly demarcates the roles of various gates and their interconnectedness.

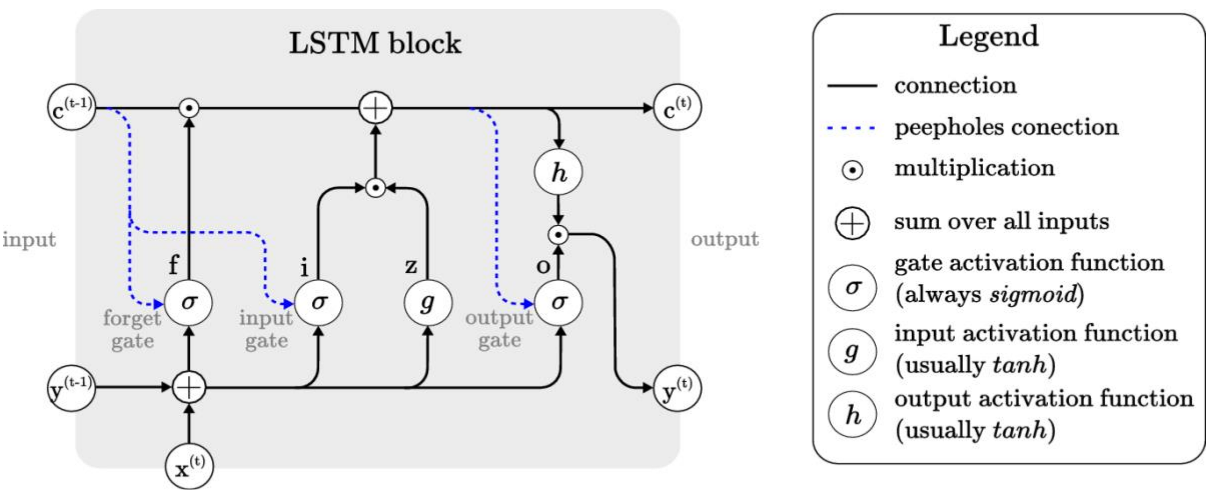

Figure 2-4. LSTM Architecture retrieved from (Van Houdt *et al.,* 2020)

In essence, the LSTM unit comprises three primary gates, each responsible for regulating the flow of information:

- **Forget Gate (f):** It decides the amount of past information to retain or discard. Mathematically, it uses the sigmoid activation function.
- **Input Gate (i):** Determines the volume of new information to store in the memory cell. Like the forget gate, it employs the sigmoid activation.
- **Output Gate (o):** Dictates how much of the internal state is revealed to the external network and subsequent LSTM blocks. Again, the sigmoid function is pivotal here.

The state of the LSTM is upheld by the cell state (often denoted as c), which undergoes modifications guided by the gates. The peephole connections, illustrated by the dotted lines in the diagram, signify connections from the cell state to the gates, allowing the gates to "peep" into the cell state. This LSTM framework, with its refined architecture and gates, serves as a cornerstone in many modern deep learning architectures for sequence modelling. Its enduring efficacy in numerous practical applications underscores its profound importance in the annals of machine learning.

$$i_t = \sigma(W_{ii}x_t + b_{ii} + W_{hi}h_{t-1} + b_{hi}$$

Equation 2-3 LSTM input gate

$$f_t = \sigma(W_{if}x_t + b_{if} + W_{hf}h_{t-1} + b_{hf}$$

Equation 2-4 LSTM forget gate

$$g_t = tanh(W_{ig}x_t + b_{ig} + W_{hg}h_{t-1} + b_{hg}$$

Equation 2-5 LSTM cell gate

$$o_t = \sigma(W_{io}x_t + b_{io} + W_{ho}h_{t-1} + b_{ho}$$

Equation 2-6 LSTM output gate

$$c_t = f_t \odot c_{t-1} + i_t \odot g_t$$

$$h_t = o_t \odot \tanh(c_t)$$

Here, i_t, f_t, g_t, and$o_t$are the input, forget, cell, and output gates, respectively. W and b represent the weight matrices and biases, and σ is the sigmoid activation function (Van Houdt et al., 2020)

The evolution of the Long Short-Term Memory (LSTM) network has been marked by continuous refinements and variations to augment its performance. A significant milestone in this journey is the peephole LSTM, a model proposed by (Gers et al., 2002). In this enhanced architecture, gate layers are given the facility to observe the memory cell state, creating a richer interactivity between the gates and the cell state.

While structural modifications like the peephole LSTM targeted functional enhancements, there was an equally pressing need to address model training challenges. This led to the exploration and adoption of specialized regularization techniques for LSTMs. Dropout, a technique introduced by (Srivastava et al., 2014), emerged as a robust solution against overfitting. In a parallel development, layer normalization, a method brought to light by (Ba et al., 2016), was recognized for its potential in ensuring stable LSTM training dynamics.

The culmination of these advancements is best evidenced by the diverse and successful applications of LSTMs across different domains. In the field of machine translation, the contributions of researchers like (Graves & Schmidhuber, 2005) underscore the potential of LSTMs. Similarly, their significance in speech recognition has been documented in pivotal works by (Graves & Schmidhuber, 2005) and (Sak et al., 2014). The versatility of LSTMs is further highlighted in creative domains such as music generation, as detailed by(Oore et al., 2018), and practical applications including sentiment analysis and image captioning, as demonstrated by research from (Maas et al., 2011; Vinyals et al., 2014).

Such diverse successes have positioned LSTMs as a primary choice for sequence modelling tasks. Their ability to handle long-range dependencies is unparalleled, and the continuous emergence of optimizations has only widened their applicability. Yet, no model is without its set of challenges, and LSTMs are no exception. One of the fundamental challenges stems from their sequential nature. This inherent trait restricts parallelization across time steps, often elongating training times for extensive datasets. Moreover, while LSTMs were innovatively designed to capture long-range dependencies, there are instances where they falter, especially when sequence elements are considerably distant. This challenge is further exacerbated by the model's reliance on fixed-length context vectors, which can occasionally result in information loss for extended sequences. Lastly, the architectural depth and recurrence mechanism of LSTMs render them both memory and computation intensive.

These limitations, while highlighting the challenges of LSTMs, also set the stage for the evolution of neural architectures. They catalysed the development of newer models, with the Transformer model, equipped with attention mechanisms, emerging as a promising successor to address the inadequacies of LSTMs.

The table below is a close comparison of traditional techniques for handling categorical features

Traditional Techniques	**Model**	**Dataset**	**Limitation**
XGBOOST	Machine Learning (Supervised)	Univariate and Multivariate Dataset	It cannot handle complex problems where predictive accuracy is paramount.
RANDOM FOREST	Machine Learning (Supervised)	Univariate and Multivariate Dataset	Time consuming in sequential processing thereby has a slow computational speed.
RNN	Deep Learning	Short sequence Dataset	It has difficulty processing long sequences due to the vanishing gradient problem. It cannot capture and remember long term dependencies in data.

LSTM	Deep Learning	Long sequence Dataset	Data complexity often led to longer training times

Table 2-1 Traditional Techniques for handling categorical features

The limitations highlighted from the traditional techniques above; even with the LSTM as better technique among the traditional models, necessitates the selection of transformer as one of the modern models with its highly parallelizable ability and capability leading to faster training on suitable hardware. Then, transformer model as emergent technique needs to be thoroughly examined for its suitability for this research work.

TRANSFORMER

The Transformer model, introduced by (Vaswani et al., 2017), marks a significant departure from previous methods of handling sequential data. Discarding the recurrent layers utilized by models like LSTMs, the Transformer architecture embraces a self-attention mechanism that weighs input elements differently based on their context. Figure 2-5 below provides a comprehensive overview of the transformer inner workings.

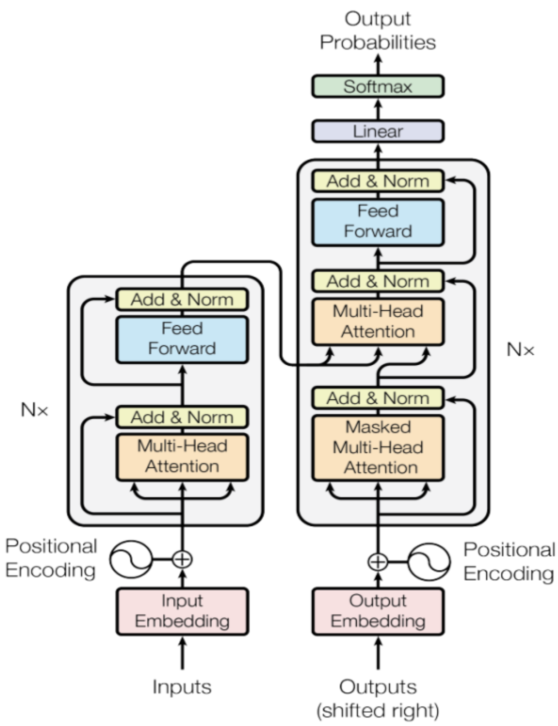

Figure 2-5 Transformer Architecture (Vaswani *et al.,* 2017)

- **Inputs** are transformed into **Input Embeddings**, with **Positional Encoding** added to retain the sequence's order information.
- These embeddings then pass through several layers (denoted as "N×") comprising two primary components: the **Multi-Head Attention** mechanism and the **Feed Forward** neural network.
- The **Multi-Head Attention** allows the model to focus on different sections of the input sequence simultaneously, catering to various aspects of the data.
- **Add & Norm** (Addition and Normalization) layers follow both the attention mechanism and the feed-forward network, ensuring stable and smooth activations throughout the network.
- The final layer outputs are transformed via a **Linear** layer, followed by a **SoftMax** layer that provides the output probabilities for sequence elements.

The self-attention mechanism, the hallmark of Transformers, allows the model to concentrate on different parts of the input sequence regardless of their order. This design is a marked advantage when juxtaposed with architectures like LSTMs, inherently bound by sequence order. By generating variable-length context vectors for each sequence element, the self-attention mechanism enables detailed and enriched representations, significantly enhancing data comprehension.

One of the Transformer's distinctive features is its parallel processing. Unlike LSTMs, which process data step by step, Transformers can process an entire sequence simultaneously, leading to drastic reductions in training times and efficient computation with large datasets.

Transformers excel in scalability. They maintain efficiency even when tasked with extensive sequences, exemplified by models like GPT-3 (Brown et al., 2020) Their versatility spans beyond natural language processing into domains like computer vision, testament to their adaptability.

With a solid grasp of the Transformer architecture's principles and strengths, it's pertinent to explore its adaptability. The subsequent section delves into 'Tab Transformers,' adapting the

Transformer's capabilities to handle tabular data, a domain distinct from traditional sequences.

TAB TRANSFORMER

While the Transformer architecture has primarily made its mark in the realm of natural language processing, its adaptability has been tested across multiple domains, one such being the domain of tabular data. Traditional methods of handling tabular data, like Gradient Boosted Trees and standard Neural Networks, have shown proficiency, the gradient-boosted decision trees (GBDT) and neural network approaches have great potential, in particular for tabular data but deep neural networks are not suitable for all types of tabular data, and GBDT models which often outperform deep models on tabular data has its own limitation depending on the dataset (Alena et al., 2022) but with the advent of the Tab Transformer, a more nuanced approach tailored for the specific challenges of tabular data has emerged.
In tabular datasets, there are two types of features encountered: categorical and continuous. Each of these feature types requires distinct pre-processing steps to make them amenable for deep learning architectures:

- **Categorical Features**: These represent discrete classes or groups. The challenge lies in encoding these in a manner that the model can discern patterns without assuming ordinal relationships between categories.
- **Continuous Features**: Continuous or numerical features can range over a wide spectrum and need to be normalized to ensure consistency in scale.

The Tab Transformer ingeniously combines the handling of both these data types within the Transformer architecture, ensuring that both feature types are given due consideration. The Figure 2-6 below elucidates the structure of the tab transformer architecture, these are the key parts.

- **Column Embedding & Layer Normalization**: Initial embeddings for categorical features and normalization for continuous features form the base upon which the Transformer operates.

- **Transformer Block**: This consists of the Multi-Head Attention mechanism, allowing the model to focus on diverse aspects of the input data. The subsequent Add & Norm layers ensure the stability of activations. This block is repeated 'N' times, facilitating deeper representations.

- **Concatenation & MLP**: After processing through the Transformer block, the embeddings are unified and then passed through a Multi-Layer Perceptron for further complexity.

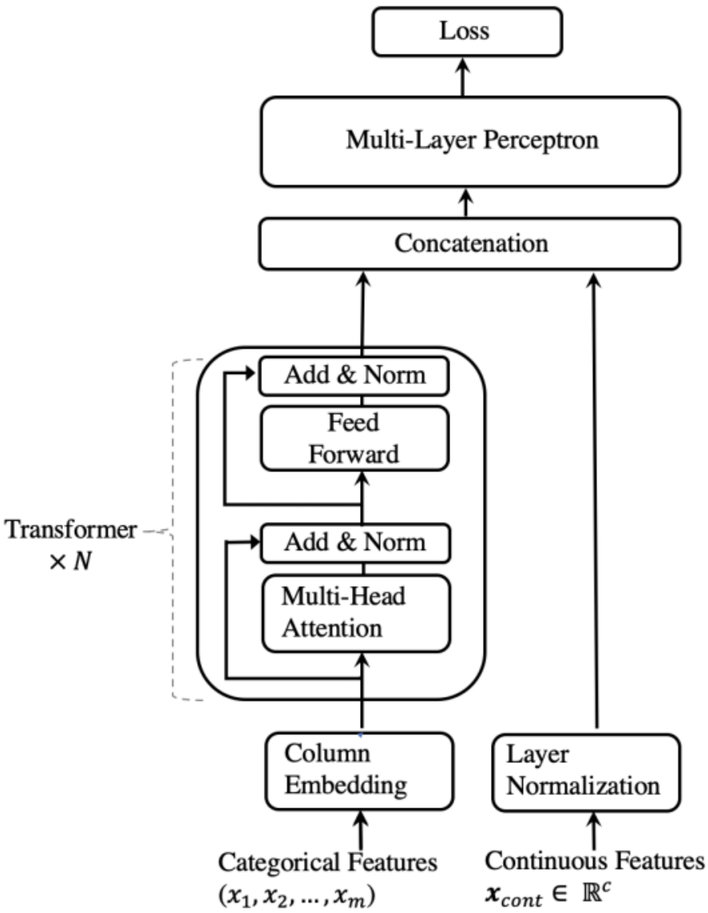

Figure 2-6 TAB Transformer Architecture (Huang *et al.,* 2020)

According to the paper by (Huang *et al.,* 2020), Tab Transformer's resilience to data inconsistencies, such as noise and missing values, coupled with its interpretable contextual embeddings, sets it apart. Delving deeper into these attributes offers an exciting avenue for future exploration. This provides a base for our exploration on Next event prediction from a tabular data perspective. A research work conducted by (Bukhsh et al., 2021) on predictive business process monitoring with transformer network. The approach is for learning high-level representation from event logs with an attention-based network and the result shown that the transformer-based model outperforms several baselines for the task of predicting event time and remaining time of a running case.

The self-attention mechanism, a core component of the Tab Transformer, allows each feature in the input data to focus on all other features, thereby computing a weighted sum of them. This mechanism enables the model to discern intricate relationships between features, making it particularly adept at handling datasets with multiple categorical features. While Transformers initially gained prominence in NLP, their potential applicability to other types of data, including tabular data, was soon recognized. Tabular data, which consists of structured rows and columns, often contains intricate relationships between features. Recognizing these relationships is crucial for tasks like predictive modelling. The Tab Transformer was developed as an adaptation of the original Transformer architecture to handle tabular data, especially datasets with multiple categorical features. By leveraging the self-attention mechanism, the Tab Transformer can capture complex interactions between features without the need for manual feature engineering (Shankaranarayana & Runje, 2021).

Limitations and Addressing Challenges of Tab Transformers

The Tab Transformer, while showcasing numerous advantages, is not without its limitations. One of the most pronounced challenges it presents is its computational intensity. Engaging with expansive datasets, the model necessitates vast computational resources, potentially resulting in prolonged training durations and difficulties in resource allocation. Such computational demands can be particularly challenging for researchers and practitioners who may not have access to top-tier computational platforms (Zhang et al., 2020). Beyond the computational realm, the issue of interpretability arises. The Tab Transformer, despite its proficiency in handling a myriad of categorical features, often conceals its decision-making pathways. This inherent "black box" nature of many deep learning architectures complicates efforts to discern the logic fuelling their predictions. The imperative for transparency pushes towards the adoption of Explainable AI (XAI) techniques to demystify the model's decision-making processes (Hassija et al., 2023)

In addressing the highlighted computational intensity of the Tab Transformer, several strategies were employed to enhance efficiency and maintain model robustness. The batch size was judiciously optimized to strike a balance between computational efficiency and model performance. Early stopping was implemented to curtail training at its most efficacious point, ensuring that computational resources were not squandered on redundant training cycles. To further refine the model's performance, column embedding was utilized to reduce input data dimensions, and layer normalization ensured stable activations throughout the model. Preventative measures against overfitting, such as dropout and weight decay (L2 regularization), were integrated. The learning rate was meticulously specified to guide the training process, and model checkpointing was incorporated to periodically save model states, allowing for efficient resumption, and minimizing retraining efforts. Through these strategic implementations, the goal was to optimize the Tab Transformer's computational demands without compromising its predictive capabilities.

Considering the PhD research objectives, understanding both the strengths and limitations of the Tab Transformer is crucial. This comprehension serves as a foundation, directing the research trajectory, particularly when probing the model's potential in predictive process mining and its synergetic integration with XAI methodologies. As the architecture inherently delves into categorical data handling, it fitting to further explore the broader challenges and methodologies associated with processing such data in the domain of deep learning. The ensuing sections aim to provide a comprehensive insight into these facets, emphasizing the nuances of tabular data and the traditional techniques used for its processing.

2.4.3 Tabular Data and Categorical Features

Tabular data is essentially data arranged in rows and columns, much like what we observe in spreadsheets. It is widely used in many areas, from business records to medical information. In this layout, each row usually stands for one item or case, and each column represents a different type of information or feature about that item (Avanzi et al., 2023). This data can have two main types of features: numerical and categorical. Numerical features have values you can measure, like height or price. On the other hand, categorical features describe qualities and can't be measured, like colours or brands. These categorical qualities can further be broken down into more specific categories.

Nominal Categorical Features: These are categories that do not have any inherent order. Examples include colours (red, blue, green) or cities (New York, London, Tokyo). They are purely descriptive and don't have a ranking system (Kang *et al.,* 2020).

Ordinal Categorical Features: These categories have a clear, defined order. For instance, ratings such as low, medium, and high are ordinal because there's an inherent hierarchy (Chawda *et al.,* 2022).

Although, most machine learning algorithms only work with numeric values, but many important real-world features are not numeric but rather categorical. Thus, categorical features, which take on levels or values, become necessary for multivariate data. Handling multiple categorical features, especially those with high cardinality, poses unique challenges in data processing and modelling. High cardinality refers to columns with a vast number of unique values. For instance, a dataset containing user IDs or product codes can have thousands or even millions of unique values. Encoding such features using traditional methods can lead to:

- **Non-continuous nature**: Unlike numerical features, categorical features are discrete, making them harder to model using algorithms that require continuity (Cerda & El Varoquaux, 2019)

- **High Cardinality**: According to the paper (Moeyersoms & Martens, 2015) High-cardinality attributes are categorical variables that have a vast array of distinct values.

These unique values can range from identifiers like bank account numbers to more general categories such as family names or ZIP codes. Despite their potential to offer valuable insights, these attributes pose challenges in predictive modelling, including computational difficulties and the risk of overfitting. As a result, they are frequently overlooked or underutilized in many modelling contexts.

- **Sparsity**: One common method to encode categorical values, one-hot encoding, can introduce sparsity. For a feature with 'n' unique values, one-hot encoding will result in 'n' new binary features, where most values are zero, leading to a sparse matrix.

Given these challenges, there's a pressing need for efficient methods to handle multiple categorical features, especially in high-dimensional datasets. This necessity aligns with the research objectives, emphasizing the evaluation of the Tab Transformer's potential in managing datasets with multiple categorical features.

Traditional Methods for Handling Categorical Features

Categorical features, which represent discrete and non-quantifiable values, are a common occurrence in datasets across various domains. Transforming these features into a format that can be efficiently processed by machine learning algorithms is a critical step in data pre-processing. Historically, several encoding techniques have been developed to achieve this transformation:

- **One-Hot Encoding (OHE):** One of the most common methods, OHE involves creating a binary column for each category in the original feature. For a feature with 'n' unique values, OHE results in 'n' new binary columns. While this method is straightforward and widely used, it can lead to a significant increase in dataset dimensionality, especially for high-cardinality features.

- **Label Encoding**: In this method, each unique category is assigned a unique integer. While this method is space-efficient, it can introduce ordinal relationships that might not exist in the original data, potentially leading to misleading interpretations by certain algorithms.

- **Target Encoding**: Also known as mean encoding, this method involves replacing each category with the mean of the target variable for that category. It can be particularly effective for high-cardinality features. However, it's essential to be cautious with this method as it can introduce leakage if not implemented correctly. A study by (Pargent et al., 2022) found that regularized versions of target encoding consistently provided the best results in machine learning applications with high-cardinality features. The research presented by (Pargent et al., 2022), discussed gaps in research where deep neural network was not explored, which we aim to address by application of tab transformer.

While the above methods offer ways to convert categorical features into numerical formats, their effectiveness can vary based on the dataset's characteristics and the specific problem context. For instance, while OHE might work well for features with a limited number of categories, it can be inefficient for high cardinality features due to the resulting sparsity.

Furthermore, traditional machine learning models like Random Forest and XGBoost, known for their robustness and versatility, often require extensive feature engineering to handle high-cardinality categorical features effectively. These models, while powerful, can struggle with raw categorical data, necessitating the need for encoding techniques. However, as the number of categories increases, the challenges of overfitting, increased computational cost, and reduced model interpretability become more pronounced (Park & Ghosh, 2013).

Considering the PhD research objectives, understanding the nuances, strengths, and limitations of these traditional methods is crucial. It sets the stage for evaluating the potential of newer techniques, such as the Tab Transformer, in handling multiple categorical features more efficiently.

2.4.4 Deep Learning in Predictive Process Mining

In this section, various applications of deep learning approaches in process mining, we would dive into the previous studies and highlight the solution and the issue they faced, after which we would discuss how we addressed some of the issues described in extending the current field.

- The work by (Tax et al., 2016a) discussed harnessing the power of LSTMs, this research ventured into multi-task learning, aiming to foretell subsequent events and their timestamps. They employed one-hot encoding for data transformation. A notable observation was the model's prowess in continuing cases but faltering in extensive future predictions, especially with recurrent events in logs.

- The work carried out by (Evermann et al., 2016) drew parallels with natural language processing, the authors perceived event logs as text, traces akin to sentences, and events within as words. Leveraging LSTMs, they focused on predicting subsequent events from these logs. The use of embeddings helped compress input sizes, yet the model encountered hitches with numerical variables. While the study shed light on model interpretation, it conceded the limitations of its interpretative tools.

- The research conducted by (Di Francescomarino et al., 2017) Attempted to refine the LSTM model proposed by Tax et al., this research integrated prior knowledge through the A-PRIORI algorithm and addressed log cycles using the NOCYCLE algorithm. This was an initiative to handle logs replete with cycles.

- The work carried out by (Tello-Leal et al., 2018) was positioned within the industry 4.0 framework, this work introduced an LSTM model catering to the Internet of Things (IoT). Although adept at predicting the imminent business process activity, the model faced challenges with continuous traces.

- The work carried out by (Lin et al., n.d.) Introduced MM-Pred, an RNN-driven model, this research zeroed in on multi-task predictions concerning event sequences and attributes. It incorporated an LSTM encoder-decoder construct, alongside a modulator to understand event attribute inter-dependencies. Yet, akin to Evermann et al.'s model, it could not anticipate attributes with numerical domains.

- The work carried out by (Camargo et al., 2019a) was focused on exploring the melded LSTM architectures with embedded dimensions, aiming to predict event traces, associated timestamps, and roles. The methodology extracted n-grams from event log traces for training. Its efficacy rivalled Evermann et al. and Tax et al. in predicting upcoming events, with an upper hand in suffix predictions.

- The work carried out by (Pasquadibisceglie et al., 2020) Pioneered the unique ORANGE technique, the framework embarks on an innovative journey to reshape outcome-driven predictive process monitoring through image encoding and CNNs. Its primary pursuit revolves around astute negotiation monitoring, particularly pinpointing those with high success probabilities. While specifics of encountered challenges remain elusive, the framework's prowess in its primary function underscores its potential. The forward-looking trajectory for this approach seems to hinge on amalgamating prescriptive learning, rectifying prediction imbalances, and enriching feature vector formulations.

- The work carried out by (Theis & Darabi, 2019) Ventured beyond conventional methodologies, this technique seamlessly fuses Petri nets with time decay functions, culminating in continuous process state samples quintessential for training event prediction-focused deep learning models. However, despite its innovative stance, the method grapples with transparently portraying process states, particularly in the context of the temporal essence of events.

- The work carried out by (Bukhsh *et al.,* 2021) presented the Process Transformer approach. This method capitalizes on high-level representations from sequential event logs, minimizing the need for intensive pre-processing. Employing the transformer network, it adeptly handles long-range dependencies, outpacing traditional LSTM-based models. While the Process Transformer showcased robust learning capabilities, it emphasized the potential for further research, especially concerning its adaptability with diverse event logs and broadening its application horizon.

- The work carried out by (Wickramanayake et al., 2022) Aimed to lift the veil on the often-obfuscated realm of predictive analytics, this work champions attention-based frameworks. Striving for a harmonious blend of transparency and accuracy, the model offers a refreshing perspective. Notwithstanding its merits, the ever-present challenge lies in the innate black box nature of deep learning predictions. The horizon looks promising, with potential avenues being bolstered predictive prowess, insightful model refinements, and meticulous hyper-parameter optimisation.

The domain of predictive process mining has perpetually transformed, marked by remarkable advancements juxtaposed against lingering challenges. A primary concern, brought to the fore by (Tax et al., 2016a), revolves around the recurrence of events in logs. These pioneering studies, while providing valuable insights, encountered challenges in predicting recurrent events, particularly across extended durations. This limitation becomes accentuated when considering the constraints of model interpretation. (Evermann et al., 2016), with their significant contributions to the interpretive realm, were still constrained by their toolset, thereby restricting the depth of understanding they could impart. Concurrently, models exemplified by (Lin *et al.*,2006; Ajagbe & Adigun 2023) contended with the complexities of forecasting numerical attributes. Amplifying these issues was the overarching concern of scalability and adaptability, with many models being tailored for specific event logs, thereby inhibiting their broader application.

Further, the dynamics of predictive analytics have been encapsulated by the intricate balance between transparency and accuracy. (Wickramanayake et al., 2022) confronted this challenge head-on, proposing the embrace of attention-based frameworks. Their innovative efforts provide a rejuvenated perspective on predictive models, emphasizing the importance of attention mechanisms. Yet, one cannot ignore that deep learning, despite its prowess, retains its "black box" nature, occasionally evading straightforward interpretability. Nevertheless, their groundwork offers a promising base for subsequent enhancements, notably in improved predictive prowess, insightful model adaptions, and meticulous hyper-parameter tuning.

Considering these challenges, we embarked on a mission to holistically address them. Tackling the recurrent event challenge, our methodology, deeply rooted in the multi-task learning paradigm and one-hot encoding techniques of Tax et al., introduces refined mechanisms adept at managing such patterns, ensuring robust and consistent predictions. Concurrently, informed by Evermann et al., our model integrates advanced interpretative tools, offering users a deeper and more intricate understanding of predictions. We also addressed the challenges posed by numerical variables, with our evolved model seamlessly blending numerical and categorical data for comprehensive predictive outcomes. Emphasizing adaptability, our models, while rigorous in their foundation, are designed to resonate across diverse event logs, ensuring a broader scope of application.

Pivoting towards a deeper exploration of the next-event prediction task, we place a magnifying lens on the contributions of individual features within our model. It's not merely about the inclusion of categorical features but understanding their significance and the weight they carry in influencing the eventual prediction. The objective of this study extends beyond raw forecasting; we aim to dissect how each feature interplays with others, and the role it assumes in the larger predictive narrative. This detailed scrutiny offers stakeholders a granular perspective, illuminating the specific levers and dials that modulate predictions. Through this, we're not just predicting the next event but explicating the intricate pathways and feature contributions that lead to it, emphasising our commitment to ensuring both predictive precision and interpretative clarity. The next section would explore the Blackbox artificial intelligence model in more detail, looking at the need for understanding of these models being developed. Table 2.2 summarise research studies on different deep learning approaches for process mining.

Author/Year	**Model/Technique**	**Dataset**	**Limitation**
Wickramanayake *et al* (2022)	Using attention-based framework	BPIC 2012, BPIC 2017	the ever-present challenge lies in the innate black box

			nature of deep learning predictions
Bukhsh *et al* (2021)	Deep neural learning (prossess Transformer)	Helpdesk, BPIC 2012, BPIC 2013, Hospital, Road Traffic Fines	Not completely accurate for the tasks of predicting event time and remaining time of a running case.
Pasquadibisceglie 47 *et al* (2020)	ORANGE (Outcome pRediction bAsed oN imaGe Encoding) and CNN	Sepsis, BPIC 2011, BPIC 2012, Production Dataset	The higher training time of ORANGE is fully counterbalanced by the highest accuracy of the learned models which overcome all other competitors
Carmago *et al* (2019)	RNN-LSTM architecture	Helpdesk, BPIC 2012	*Limitation is not found*
Theis & Darabi, (2019)	Enhanced the Petri net model	Helpdesk, BPIC 2012, BPIC 2013	The quality of predicting next events and that might overcome the low precision scores47; the method grapples with transparently portraying process states, particularly in the context of the temporal essence of events.
Tello-Leal *et al.*, (2018)	LSTM neural networks model	Event log that originates from the IoT and Industry 4.0 domain.	The study is limited to Event logs with small number of traces.
Di Francescomarino *et al.*, 2017	NOCYCLE algorithm- a refined LSTM model	logs replete with cycles	It was used only to predict the next activities in business processing
Evermann *et al.*, (2016)	RNN	BPIC 2012, BPIC 2013	Difficulty handling event logs that contain many sequences of two or more events in a row of the same activity.
Tax *et al.*, (2016)	LSTM	Helpdesk, BPIC 2012, Environmental permit dataset	Handling event logs that contain traces with multiple occurrences.

Lin *et al.*, 2006	MM-Pred- an RNN-driven model,	event sequences and attributes datasets	It could not anticipate attributes with numerical domains

Table 2-2 Next Event Prediction applications

2.5 Black Box Artificial Intelligence

The term "artificial intelligence" (AI) is used to describe a wide spectrum of software with varied degrees of autonomy, intelligence, and dynamic problem-solving capacity. The most rigid AI are those that follow a set of predetermined rules to draw conclusions or weigh alternatives (Heaton, 2018). This class includes, for instance, chess programs that score each potential move and choose the best one based on the algorithm. Modern AI applications based on machine-learning algorithms that can learn from data are extremely versatile. In contrast to rule-based AI, this type of AI would analyse a large number of chess games in real-time to discover patterns it may use to determine its next move; it would develop its scoring system (Jones, 2018) guidelines for how to learn from data are all that are written into this type of AI (Jones, 2018) rather than guidelines for how to tackle specific problems.

Machine learning (ML) techniques are rapidly evolving, allowing for the creation of AI applications such as recommendations for financial products (Farquad et al., 2014), algorithms for detecting credit card fraud, personal virtual assistants (Lu et al., 2018) and autonomous driving vehicles. Black-box algorithms, which do their processing in isolation, don't reveal how they arrived at the answer offered to the decision-maker. It does not, for instance, make it possible to determine which factors affected a credit projection, or why one customer's loan application was declined while another was accepted.

Models that are extremely sophisticated and hard to read are known as "black boxes" in the fields of artificial intelligence and machine learning. Although these models tend to have excellent predicted accuracy, their decision-making and prediction processes are typically opaque (*Explainable Artificial Intelligence*, n.d.) To emphasize how difficult or impossible it is to decipher these models, the term "black box" is often employed. Data and decision-making in black box models are distributed among thousands of artificial neurons in deep networks, creating a level of complexity that rivals that of the human brain. In sum, we still don't know what goes on within a block box or what elements contribute to it.

Understanding how input variables are utilized to produce predictions is comparatively easy in traditional rule-based systems or basic machine learning models like linear regression. However, the linkages between inputs and outputs get increasingly convoluted in more advanced models like deep neural networks or ensemble approaches, making it difficult to grasp the inner workings of the model (Molnar, 2023).

There are questions about justice, accountability, and possible biases in black-box AI models despite their good performance (Ribeiro *et al.,* 2016). It becomes tricky to discover and correct any biases or inaccuracies in the model since it is difficult to understand the elements impacting its conclusions.

Both academics and government officials have realized the urgency of fixing black-box AI. Interpretability and explainability approaches are currently being developed to help users comprehend a model's decision-making process. To better understand how black box AI models function, researchers are looking at several different approaches, including feature significance analysis, gradient-based attribution methods, and surrogate models (Gebru et al., 2018).

The need for openness and responsibility in AI systems is also being addressed by regulatory agencies and other groups. It is the goal of several pieces of proposed legislation, such as the General Data Protection Regulation (GDPR) and the Algorithmic Accountability Act (AAA), to guarantee that all AI systems, including black box models, are transparent and ethically sound (Smuha, 2019) There is a continuing endeavour in the field of AI to find a happy medium between the precision and human-understandability of AI models. Both academics and industry professionals are aiming to improve the explainability and openness of AI systems without sacrificing performance (Smuha, 2019). Let’s delve into the laws around AI and Machine Learning, to understand form a legal standpoint, the need for explainability to further understanding of how models make their decisions.

2.5.1 LEGISLATIONS REGARDING AI AND MACHINE LEARNING

The legislations concerning AI and Machine learning varies by country according to their peculiarities. Here are a few notable examples of legislative efforts related to AI regulation:

General Data Protection Regulation (GDPR) - Data processing by AI is subject to the General Data Protection Regulation (GDPR) that was enacted by the European Union (EU) in 2018. Although not limited to Blackbox AI, it does guarantee persons the right to be informed and explained any major adverse effects caused by automated decision-making. The collection and use of personal data is governed by the General Data Protection Regulation (GDPR) and the Data Protection Act 2018 (DPA 2018) (ICO, 2023). This law applies to situations when AI is used to process personally identifiable information. This can occur when an AI system is trained, tested, or deployed using an individual's personal information. When explaining how AI works, administrative law and the Equality Act of 2010 are also important considerations.

The Data Protection Act of 2018 and the General Data Protection Regulation make up the body of data protection legislation in the United Kingdom. Together, they control how information on living individuals may be gathered and used. When artificial intelligence is used without any personal information, it is not subject to privacy regulations (Oswald, 2023). Artificial intelligence has several potential applications in the sciences. Personal information is regularly used or created by AI systems. Massive volumes of individual information are sometimes utilized for the development and testing of AI systems. At the time of rollout, the model will have access to more detailed personal information. Individuals' predictions and judgments are still considered private information. Any use of AI in any of these contexts falls within the purview of privacy laws.

The legislation protecting personal information does not favor any one technology. Neither artificial intelligence nor any related technology like machine learning are mentioned by name.

The use of profiling and automated decision-making is explicitly addressed in various clauses of both the General Data Protection Regulation and the Data Protection Act of 2018, both of which place a heavy emphasis on such processing on a massive scale. This means that it may be applied when suggesting or forecast based on artificial intelligence.

The right to be informed

Individuals have the right to be informed, as outlined in Articles 13 and 14 of the GDPR, of: the existence of solely automated decision-making with legal or similarly significant effects.

- Meaningful information about the logic involved; and
- The significance and envisaged consequences for the individual.

The right of access

Article 15 of the GDPR gives individuals the right of access to: Information on whether or if entirely automated decision-making has occurred with legal or similarly important repercussions, informative details about the decision-making process, and anticipated outcomes for the individual are all necessary (Oswald, 2023)

Rights in the context of computerized decision-making are interpreted with some clarity in Recital 71. It mostly concerns Article 22 rights, but it also makes it quite clear that people have the right to get an explanation of an automated judgment after it has been made if they so want (Oswald, 2023)

The right to object

Article 21 of the GDPR gives individuals the right to object to processing of their personal data, specifically including profiling, in certain circumstances. There is an absolute right to object to profiling for direct marketing purposes.

Rights related to automated decision-making including profiling

According to ICO (2023), people have the right to not be subject to an automated decision with legal or similarly substantial repercussions under Article 22 of the General Data Protection Regulation. In some circumstances, however, it mandates that organizations:

- Implement appropriate safeguards to safeguard persons, including the right to receive human intervention.
- Allow individuals to provide input; and
- Allow individuals to challenge the decision.

The interpretation of Article 22 is also clarified by Recital 71.

Data protection impact assessments

Data Protection Impact Assessments (DPIAs) are mandated by Article 35 of the GDPR for organizations whose processing of personal data, especially when utilizing new technologies, poses a substantial risk to individuals. Any use of automated profiling or other automated evaluation of personal data for decisions with legal or similarly substantial implications on individuals always requires a DPIA.

In light of this, DPIAs should be performed prior to processing in order to identify and analyse the degrees of risk associated if you want to utilize AI systems to process personal data. The DPIA should be a "living document" that is reviewed frequently and if there is a change in the processing's nature, scope, context, or aims. Additional DPIA guidelines, including a list of processing processes that call for a DPIA, has been provided by the ICO. Artificial intelligence (AI), machine learning (ML), massively scaled profiling (MBP), and automated decision-making (ADM) that results in service/product/benefit rejection are all on the list.

If a DPIA reveals severe risks to individuals' rights and freedoms that cannot be mitigated, you must get approval from the ICO before proceeding with the processing.

As above, the GDPR has specific requirements around the provision of information about, and an explanation of, an AI-assisted decision where:

- It is made by a process without any human involvement; and
- It produces legal or similarly significant effects on an individual (something affecting an individual's legal status/ rights, or that has equivalent impact on an individual's circumstances, behaviour or opportunities, e.g., a decision about welfare, or a loan) (Oswald, 2023).

In these instances, compliance with the General Data Protection Regulation (GDPR) necessitates the following actions:

- Adopt a proactive approach to ensure individuals receive comprehensive information about:
    - The underlying rationale.
    - The importance of decisions.
    - Anticipated ramifications of those decisions.
- Articles 13 and 14 stipulate:
    - Individuals should have the right to human intervention from the controller.
    - This allows them to express their perspectives.
    - It also permits them to challenge any decisions made.
- As outlined in Article 22:
    - Individuals are entitled to:
    - Acquire meaningful information about the logic behind decisions.
    - Understand the significance of decisions.
    - Be informed of anticipated outcomes.

Although the recitals of the General Data Protection Regulation (GDPR) lack legal enforceability, they serve the purpose of elucidating the interpretation and purpose of its articles. The inclusion of a provision in Recital 71, which addresses the provision of an explanation for an automated decision after its implementation, elucidates that the entitlement to such a right is inherent within the provisions outlined in Articles 15 and 22 (Oswald, 2023). It is imperative to provide individuals with a comprehensive elucidation of a fully automated decision to facilitate their rights to access substantial information, articulate their perspectives, and challenge the decision (Oswald, 2023).

However, even in cases where an AI-assisted decision is not a component of a fully automated process, and there is significant human involvement, the utilization of personal data still necessitates compliance with all the principles outlined in the General Data Protection Regulation (GDPR). The principles of fairness, transparency, and accountability outlined in the General Data Protection Regulation (GDPR) hold significant importance.

Fairness

Considering the potential impact on individual interests is an important part of determining whether your use of personal data is fair. A person's right to autonomy and self-determination may be compromised if an AI-assisted decision is made about them without any kind of explanation of (or knowledge about) the choice (EU Commission, 2022).

Transparency

Being transparent means explaining to individuals exactly how and why you want to utilize their personal information. Recital 60 of the GDPR states that you must provide any additional information necessary to ensure fair and transparent processing considering the specific circumstances and context in which you process the personal data, in addition to the information requirements on automated processing laid out in Articles 13 and 14 of the GDPR (EU Commission, 2022). Using people's personal information to train and test an AI system or explaining how an AI-assisted decision was made about them without their knowledge is not likely to be regarded transparent. You may demonstrate openness by explaining your position. It may be necessary to provide an explanation of the processing's purpose in accordance with Articles 13-15 of the GDPR (EU Commission, 2022).

Accountability

In order to fulfil the requirement of accountability, it is necessary to provide evidence of adherence to the various principles outlined in Article 5 of the General Data Protection Regulation (GDPR), such as data minimization and accuracy. One method for demonstrating accountability involves furnishing individuals with a comprehensive rationale for a decision and subsequently recording the process of its implementation. Regardless of the specific AI-assisted decision made, particularly those involving the utilization of personal data, it is incumbent upon individuals to adhere to data protection legislation, which necessitates the provision of comprehensive explanations to those individuals who may be impacted (EU Commission, 2022).

Parts 3 and 4 of the DPA 2018

Part 3 of the DPA 2018 also contains special language for cases when law enforcement agencies use exclusively automated conclusions that have a substantial impact on data subjects or have an undesirable legal consequence. People have the option to speak with a live person, share their thoughts, get an explanation of the decision, and even appeal it. At the moment, it is unlikely that law enforcement agencies will rely completely on computerized decision-making systems (*Data Protection Act 2018*, n.d.).

Part 4 of the DPA 2018 has its own set of regulations for fully automated decisions made by intelligence agencies that have far-reaching consequences for data subjects. In certain situations, people have a right to seek out human help. Individuals also have a basic right to be informed about the controller's reasoning behind applying a decision made using their personal data. This allows them to seek "knowledge of the reasoning underlying the processing." However, the exception in Part 4 for protecting national security may restrict these protections.

Algorithmic Accountability Act

United States legislation proposed in 2019 and revised in 2022 intends to regulate artificial intelligence to prevent prejudice and bias. When employing automated decision systems for crucial choices, the law mandates that businesses do impact evaluations for bias, efficacy, and other criteria. It also adds 75 people to the Federal Trade Commission's enforcement team and establishes a public repository at the FTC for such systems (Mökander et al., 2022).

2.5.2 AI LEGISLATIONS IN UK

To prevent strangling innovation, the United Kingdom government produced a white paper (a policy document outlining ideas for future law) in 2023 detailing its objectives for regulating artificial intelligence (Matt Davies & Michael Birtwistle, 2023). The paper encourages compliance on a voluntary basis and outlines five guidelines to mitigate the dangers posed by AI.

Components of the UK AI Act

Unlike the EU's risk-based regulation, the UK takes a different tack. Live face recognition technology, in which persons visible on a video feed are matched against police "watch lists," is prohibited in public settings under the EU's planned AI Act (Matt Davies & Michael Birtwistle, 2023). The European Union's strategy establishes rigorous norms for "high-risk" AI systems. Systems like this are used to determine who gets hired, who gets accepted to school, who gets financial aid, and who gets government services (Matt Davies & Michael Birtwistle, 2023).

There are three main tenets of the UK's approach to regulating AI.

Firstly, rather than enacting new AI-centred legislation, it makes use of pre-existing legal frameworks including privacy, data protection, and product liability laws (Miranda Mourby, 2021).

Secondly, regulators would use five overarching principles, each of which consists of numerous components, in concert with pre-existing legislation (Miranda Mourby, 2021). Safety, security, and robustness; proper openness and explainability; fairness; accountability and governance; contestability and redress; on are the five guiding principles (Miranda Mourby, 2021). Regulators would not be compelled by law to enforce the principles during the first implementation period. If it becomes essential, a law might be passed to impose these requirements (Miranda Mourby, 2021). Therefore, initially, organisations would be asked to comply with the principles on a voluntary basis.

Thirdly, with help from an overarching coordinating organization, regulators may tailor the five principles to the specific areas they oversee (Miranda Mourby, 2021). This means there will be no central body charged with enforcing the law.

Advantages of the AI Act

According to an impact assessment on the AI laws in UK, there are three reasons why the system in the UK has promise (Alison Kilburn, 2023).

First, it guarantees to use AI evidence in the right setting, rather than extrapolating results from one domain to another. Second, it's made such that rules may be quickly modified to meet the needs of AI deployed in a variety of practical contexts. Third, its decentralized structure offers certain benefits. For instance, the widespread adoption of AI would be negatively impacted by the failure of a single regulatory body.

Some US-based online services, for instance, employ such algorithms to identify a person's sex just by analysing their visual traits. When tested with images of women with darker skin tones, they performed poorly. This discovery has been used to justify restricting the use of facial recognition by British police (Alison Kilburn, 2023). However, issues with gender classification may not always indicate problems with facial recognition in law enforcement.

The legal requirements for these sex classification algorithms in the US are lax. In the United Kingdom, facial recognition technology is only utilized by police after extensive testing and compliance with all applicable laws (Alison Kilburn, 2023).

The flexibility of the British method is an additional benefit. Particularly with artificial intelligence (AI) that might be appropriated for reasons other than those envisaged by its inventors and machine learning systems, which grow in performance over time, it can be challenging to identify possible problems (Oswald, 2023). With this structure in place, regulators may respond rapidly to emerging concerns without having to wait for protracted parliamentary deliberations. Organizations would take on varying levels of responsibility, the enforcement of rules pertaining to artificial intelligence may suffer if they are centralized under a single national authority (Bathaee, 2018).

Expert regulators in disciplines like transportation, aviation, and the financial markets will be best positioned to oversee the application of AI in those sectors (Oswald, 2023). According to (Oswald, 2023), this decentralized strategy has the potential to lessen the impact of corruption, regulators' focus shifting away from the public interest, and conflicting methods of enforcement. It also removes the potential for a single failure point in enforcement.

Enforcement and coordination

According to the UK white paper for AI regulation, some companies may be resistant to voluntary norms; hence, authorities should have the authority to levy fines once they are given that authority. When people are harmed by AI, they should be able to sue for damages.

The ability to tighten or relax regulations remains with the regulatory bodies. However, the UK approach may run into problems in cases when artificial intelligence systems are regulated by more than one body (Michelle Donelan, 2023). Authorities in charge of transportation, insurance, and privacy may all potentially establish contradictory rules for autonomous vehicles. The white paper recommends a solution to this problem: the creation of an overarching organization whose job it is to oversee the consistent application of recommendations (Michelle, 2023). It is critical to mandate that the various regulatory bodies go to this group for guidance rather than making their own independent decisions ((Michelle Donelan, 2023).

The United Kingdom's strategy has potential to promote innovation and reduce threats. Aligning the framework with laws elsewhere, notably in the EU, is necessary to improve the country's position as a leader in the sector (Michelle Donelan, 2023). Legal certainty for enterprises and public confidence can be improved by further framework refinement. It will also increase trust in the UK's regulatory framework for this game-changing technology outside of the country (Michelle Donelan, 2023).

Documentation

It is common practice for businesses to keep records of their processing operations, which may include details like the reasons for processing, the data shared, and the data retained. In addition to being required by law, maintaining detailed records of all processing operations is crucial for meeting the requirements of the UK General Data Protection Regulation (GDPR).

According to (Oswald, 2023), documentation requirements vary depending on who is doing the controlling or processing. If your company has 250 workers or more, all processing actions must be recorded. Small and medium-sized businesses are eligible for a partial exemption. Processing actions that:

- Are not occasional.
- Might result in a risk to the rights and freedoms of persons.
- Include the processing of special categories of data or criminal conviction and offence data.
- Involve an organization with fewer than 250 workers must be documented.

Under Article 30 of the UK GDPR (Oswald, 2023)**, you must document the following information:**

- The organization's name and contact information, along with relevant details of other controllers, representatives, and data protection officers, if applicable.
- A clear explanation of the purposes for which data is being processed.
- A comprehensive breakdown of the categories of individuals and the corresponding categories of personal data being processed.
- Identification of the recipients or categories of recipients to whom personal data may be disclosed.
- Thorough documentation of any transfers of personal data to third countries, including the safeguards implemented to ensure the security of such transfers.
- Retention schedules outlining the duration for which personal data will be retained.
- A detailed description of the technical and organizational security measures in place to protect personal data.

As part of your record of processing activities, it can be useful to document (or link to documentation of) other aspects of your compliance with the UK GDPR and the UK's Data Protection Act 2018. According to (Oswald, 2023) required details for privacy notifications include:

- The legal justification for processing.
- The processing's legitimate interests
- The rights of persons
- The presence of profiling and other forms of automated decision-making
- Information needed to process special category data or criminal conviction and offence data under the Data Protection Act 2018 includes:
- The condition for processing in the Data Protection Act.

- The lawful basis for the processing in the UK General Data Protection Regulation; an explanation of how the data was obtained.
- Records of consent.
- Controller-processor contracts.
- The location of the data.
- Data Protection Impact Assessment reports; and
- Records of personal data breaches.

After an in-depth exploration of the legal stances on explainability from the UK, US, and EU, this research will now transition to an analysis of the prevalent techniques. Building on the findings of (Weidinger et al., 2022), regarding the risks posed by decision-making models, we will provide a comprehensive overview of current methods in XAI implementation. The subsequent section will expand on the need and various approaches applied.

2.5.3 The Explainability of AI

It has recently been a big social worry that AI algorithms, and in particular Machine-Learning (ML) algorithms, are difficult to explain (Pasquale, 2015). Governments across the world are finally reacting to this kind of anxiety. Human agency and oversight, technical robustness and safety, privacy and data governance, transparency, diversity/non-discrimination/fairness, societal and environmental wellbeing, and accountability are the seven criteria proposed by a European High-level Expert Group on AI (Smuha, 2019) for a trustworthy AI.

Considering this, the Commission's White Paper on AI proposed six categories of requirements for high-risk AI applications, including: ensuring quality of training data; keeping data and records of the programming of AI systems; information to be proactively provided to various stakeholders (transparency and explainability); ensuring robustness and accuracy; having human oversight; and other specific requirements for certain AI applications. Therefore, clarity and explicability are prioritized throughout both texts. Therefore, numerous additional duties have been implemented in Europe to improve the explainability of algorithmic judgments; these obligations are unique to automated systems (and hence to AI) and may be found in data protection standards and consumer protection rules as describe by the Ethics Guidelines for trustworthy Ai by the European commission.

Guaranteeing the usefulness of AI models in the real world relies heavily on the incorporation of explainability methods, often known as XAI. It should be incorporated into the model's architecture from the start, but instead it is typically added as a post-training analysis (post hoc; see also (Narwaria, 2021). Understanding how independent factors have impacted the model's predictions without transformation is the focus of global interpretation, which seeks to provide a general explanation rather than a specific solution (an explanation for each observation in the dataset). This interpretation does not seek to isolate the effects of individual training variables on the prediction but rather to isolate the effects of overarching training elements. In contrast, local interpretation looks at the predictions made for each individual example in the sample after the model has been estimated (Adadi & Berrada, 2018). The findings produced by the black-box algorithms can be understood on a case-by-case basis, which is why the processes of local explainability were considered in this research.

Methods of interpretability can be categorized not just as global or local, but also as agnostic or model specific. Without knowing the specific mathematical process that went into making the original model, explainability solutions can be generated with an agnostic approach. Black-box models may be explained with the help of this feature. Methods tailored for use with algorithms have been developed. (Ribeiro et al., 2016) offered one approach to agnostic local interpretation. Local Interpretable Model-Agnostic Explanations (LIME) is a linear proxy model that offers producing interpretable and accurate justifications for the predictions of any classifier. This study utilized the LIME technique, which has been employed in previous research to explain similarly opaque systems (Narwaria, 2021).

Accuracy metrics may be used on a test dataset to evaluate the models. However, such an evaluation could not show that the model is trustworthy. For big datasets, it is vital to recommend which instances should be reviewed using these measures, in addition to the inspection of individual forecasts and their justifications. This study employs the LIME technique, which promises to address the "confidence in a forecast" issue by providing justifications for individual forecasts, and the "model confidence" issue by selecting a subset of these forecasts and their justifications.

Linear models, decision trees, and descending rule lists are all examples of G models that might be used to provide an explanation (Wang & Rudin, 2015). The user can be provided with a G model alongside other data types, such as images or texts. This research employed (g) as a measure of the complexity (rather than the interpretation) of the G explanation since not all such models can be sufficiently simple to be interpretable. g can be the number of nonzero weights in a linear model, or the depth of a decision tree.

2.5.4 Explainable AI (XAI) Approaches

The widespread integration of machine learning models across various sectors highlights the essential need for transparency in their decision-making mechanisms. As noted by (Doshi-Velez & Kim, 2017), the mysterious workings of 'black-box' models, especially deep neural networks, while being highly precise, provide limited visibility into how they make decisions. This lack of clarity becomes a bottleneck for fields like predictive process mining, where comprehending the decisions is often as vital as the predictions themselves. XAI arises to address this gap, combining predictive strength with the imperative of elucidation. As we explore the use of the Tab Transformer for predictive process mining, the pursuit of a strong XAI framework becomes a key focus. We will delve into the various methods in the subsequent sections.

SHAP (Shapley Additive explanations)

In the complex domain of predictive process mining, the multitude of features dynamically interact to influence results. Therefore, elucidating the distinct contribution of each feature becomes crucial. SHAP (Shapley Additive explanations), rooted in the core principles of cooperative game theory, emerges as a potent instrument for this deep dive into feature contribution. Lundberg and Lee's pioneering work (Lundberg et al., 2017), illustrates the unique capabilities of SHAP. Notably, SHAP offers a principled way to fairly and consistently allocate feature importance, even in intricate model structures like the Tab Transformer.

SHAP works by associating each feature value with a Shapley value, a concept derived from cooperative game theory. It essentially represents the average marginal contribution of a feature value over all possible coalitions. Unlike traditional feature importance metrics that might be biased or inconsistent, SHAP values offer consistent and fairly distributed contributions. This means the sum of all the SHAP values for a single prediction is equal to the difference between the prediction and the average prediction for the dataset, providing a detailed decomposition of the prediction (Lundberg et al., 2018).

(J. Chen et al., 2018) underscored the universality of this attribution method, highlighting its ability to offer a unified perspective on model interpretability across diverse model architectures. Nevertheless, with the increasing adaptation of SHAP in analysing voluminous process mining datasets, the computational burden proportionally grows. Addressing this challenge necessitates the exploration of advanced optimization strategies. This could range from harnessing the power of hardware accelerations to pioneering algorithmic breakthroughs, ensuring SHAP's insights are obtainable in real-time analytical environments (J. Chen et al., 2018).

Feature Importance

In predictive process mining, the inherent diversity of attributes encompasses categories, sequences, and time-based data. Understanding the relative significance of these attributes is paramount for model optimization and the generation of meaningful insights. (Breiman, 1999) perspectives on Random Forests aptly convey this idea, demonstrating its proficiency in identifying and emphasizing feature importance. Building on this conversation, (Altmann et al., 2010) explore the concept of permutation feature importance, emphasizing its resilience in pinpointing attribute significance amidst potential noise or collinearity. However, as we delve further into the intricate landscape of process mining, characterized by its complex and frequently intertwined features, certain challenges arise. The potential for misinterpreting importance or missing subtle interrelationships between features necessitates a more rigorous analytical approach. This underscores the need for innovative or combined methodologies to address these obstacles, ensuring that the assessment of feature significance remains both precise and comprehensible within the multifaceted framework of process mining.

LRP (Layer-wise Relevance Propagation)

Layer Relevance Propagation (LRP) emerges as a vital interpretability technique. This section will delve into the nuances of applying LRP to the Tab Transformer and explore the potential challenges and advantages it offers.

At its core, LRP provides a clarifying lens into the decision-making pathways of deep learning models. It functions by retracing the predictions made by a neural architecture through its intricate layers, assigning a 'relevance' metric to each input component in the process. This approach sheds light on the significance of each input with respect to the final prediction, effectively elucidating the decision-making process of the model (Bach et al., 2015).

The importance of such clarity becomes evident, especially when dealing with complex architectures like the Tab Transformer. Being able to discern which portions of the input data carry the most influence can provide invaluable insights. This interpretative capability is indispensable in domains such as predictive process mining, where critical decisions are made regularly.

Structured tabular data, prevalent in fields such as finance, healthcare, and e-commerce, serves as the primary domain of Tab Transformer models. The consequences of the model's decisions extend to financial strategies, medical interventions, and business strategies. Therefore, understanding the rationale behind a prediction is of utmost importance. To achieve this, LRP initiates the process by tracing prediction metrics across the model's layers, assigning a 'relevance' metric to each feature, representing its contribution to the final prediction (Doshi-Velez & Kim, 2017; Samek et al., 2017).

However, the inherent complexity of the Tab Transformer, characterized by its self-attention mechanisms and embeddings, introduces unique challenges. During the process of relevance backpropagation, careful consideration must be given to the self-attention coefficients to ensure that the model's focal features during its operational phase are appropriately emphasized. The presence of embeddings further complicates matters, as categorical elements are transformed into continuous vectors. This necessitates adaptations in LRP to preserve the integrity of the original categorical input throughout the relevance propagation process. The scarcity of papers that used Tab Transformer makes the research work unique and a landmark study.

The application of LRP to the Tab Transformer is layered with intricacies, primarily due to the model's architectural depth. One critical aspect is managing the interplay of features, especially given that tabular configurations often contain interrelated features. The self-attention mechanism of the Tab Transformer has the potential to identify these correlations, requiring LRP to be adapted to highlight these interdependencies.

Local Linear Models (LIME)

In the evolving frontier of machine learning interpretability, LIME (Local Interpretable Model-agnostic Explanations) emerges as a noteworthy method that attempts to make black-box models more accessible to human understanding. Its fundamental principle lies in deciphering the local decision boundaries of complex models and representing these decisions through simpler, interpretable models, such as linear ones (Ribeiro et al., 2016) For predictive process mining, which often relies on intricate models like the Tab Transformer (Huang et al., 2020), LIME can provide vital clarity on specific predictions.

The unique value proposition of LIME is its capability to approximate the behaviour of an intricate model in a localized region, shedding light on how specific input features influenced a particular prediction (Ribeiro et al., 2016). This becomes especially valuable in predictive process mining tasks where understanding the nuances behind individual process predictions can lead to actionable insights (Di Francescomarino & Ghidini, 2022). For instance, understanding why a particular transactional process might be flagged as anomalous can aid in early rectification and decision-making.

However, while LIME's local interpretations are invaluable, they come with inherent challenges. One of the most prominent concerns is the dichotomy between local fidelity and global generalizability (Dieber & Kirrane, 2020). While LIME is designed to faithfully reproduce decisions in a localized region around a data point, it does not guarantee that these interpretations hold true across the broader spectrum of data points. In the realm of predictive process mining, where processes are often dynamic, non-linear, and interdependent, there's a need to reconcile these localized insights with the global behaviour of models like the Tab Transformer.

Moreover, the interpretative models produced by LIME rely on perturbations of the original data. In large-scale process mining datasets, which may consist of complex, interwoven categorical features, generating meaningful perturbations without introducing noise or unrealistic data points becomes a challenge (Dieber & Kirrane, 2020). This raises questions about the reliability and robustness of LIME's explanations in such contexts.

Given these challenges and the significance of interpretability in predictive process mining, there's a pressing need for further exploration. Research initiatives should focus on evaluating LIME's efficacy in the context of the Tab Transformer(Huang et al., 2020; Kang et al., 2020; Sokol & Flach, 2020), understanding its limitations, and potentially integrating it with other XAI techniques. As the field of predictive process mining grows, ensuring the alignment of LIME's local surrogates with the holistic, global objectives of models becomes a pivotal research direction, promising to bridge the gap between high predictive accuracy and human-understandable insights.

COMPARISON

The growing landscape of Explainable AI (XAI) presents a myriad of methods, each with its own unique advantages and challenges, especially when applied to the Tab Transformer in the realm of predictive process mining. Understanding the intricacies of these techniques is crucial to ensuring their effective application, given the specific contexts of predictive tasks and the nature of data at hand.

The integration of the first generic methods such as SHAP, LRP, and LIME, which are widely-used surrogate models to explain decisions of complex Machine Learning models (Agarwal 2020) becomes evident in XAI due to the flexibility they provide with both local and global explanation; they hold the promise of offsetting the limitations inherent to each individual method. For instance, while SHAP offers a unified measure of feature importance consistent across models, its computational intensity can be a bottleneck, especially for complex models like the Tab Transformer. This computational intensity can impact the timely interpretation of results. On the other hand, techniques like LRP and LIME provide granularity in explanations, especially for image-based tasks, but might not be as globally representative. This means they excel in offering detailed insights into specific aspects of the model's behaviour but may not capture the broader, global patterns that are essential for comprehensive understanding.

A closer examination of these XAI techniques reveals the following:

XAI Technique	Advantages	Challenges	Applicability to Tab Transformer	Implications for Predictive Process Mining
SHAP	Unified measure of feature importance	Computational intensity	Offers consistent feature importance measure	Helps illuminate feature importance for model validation
LRP	Granularity in explanations	May not capture global patterns	Provides detailed insights, may lack global context	Elucidates intricate feature relationships
LIME	Granularity in explanations	Potential lack of global context	Offers fine-grained explanations, may	Provides insights into feature contributions for actionable insights

			lack global perspective	

Table 2-3 XAI Technique Comparison

A particularly relevant aspect for this research is how these XAI techniques illuminate the predictions made by the Tab Transformer when handling datasets with multiple categorical features. Preliminary findings suggest that while some methods are adept at highlighting feature importance, others excel in elucidating the intricate relationships between features and their contribution to predictions. This has profound implications for predictive process mining, where understanding the role of each categorical feature is crucial for both model validation and actionable insights.

As the field of XAI expands, there is a marked deficiency in resources detailing the practical application of these techniques within the Tab Transformer for activity predictions. The nuances and intricacies of integrating XAI methods with the Tab Transformer to enhance activity predictions are an area that has not been extensively charted. This exploration is crucial not just for researchers and practitioners but also for stakeholder's keen on ensuring that AI-driven activity predictions align seamlessly with business imperatives and comply with regulatory frameworks. Moreover, the fusion of XAI techniques with the Tab Transformer is not just a matter of technical integration. Given the research's emphasis on the legal contours of AI, this synthesis carries significant legal and ethical weight. The clarity these methods bring to the Tab Transformer's decision-making processes for activity predictions could determine their endorsement in sectors where regulations are particularly rigorous.
In essence, while the integration of multiple XAI techniques offers a holistic approach to model interpretability, the nuanced dynamics between them, especially in the context of the Tab Transformer and predictive process mining, require further exploration. Addressing this will not only bridge the technological aspects of AI but will also align with the broader socio-political and legal implications, fulfilling the overarching objectives of the research.

2.6 Conclusion

Our exhaustive examination of the literatures concerning Tab Transformers, particularly in the context of managing multiple categorical features and in juxtaposition with conventional models, offers a holistic perspective of the prevailing research landscape. Yet, as is the case with dynamic disciplines, certain gaps beckon more profound investigation and research. A striking observation is the constrained empirical exploration of Tab Transformers across varied datasets and tangible real-world scenarios. Improving the results observed from previous models, gaining insights from the predictions by utilising all the features available previously ignored (Appice et al., 2019; Evermann et al., 2016; Taymouri et al., 2020; Tello-Leal et al., 2018; Weytjens & De Weerdt, 2020). Such a limited scope hinders our capacity to generalize outcomes and casts doubts about the model's resilience and adaptability in diverse environments, as posited by (Shankaranarayana & Runje, 2021).

The interpretability of deep learning architectures, including Tab Transformers, emerges as a consistent challenge. Notwithstanding some endeavours to amalgamate Explainable AI (XAI) mechanisms with Tab Transformers, a void exists in holistic research that ventures into the intricacies of such integrations and their impact on augmenting model transparency, a sentiment echoed by (J. Chen et al., 2018; Dieber & Kirrane, 2020; Samek et al., 2017; Sokol & Flach, 2020).

Another aspect that merits scrutiny is the applicability of Tab Transformers across diverse fields. While the model showcases considerable promise, there is a palpable need to broaden its applicability to ensure it meets the diverse demands of varied sectors. Concurrently, addressing the innate computational demands of Tab Transformers remains imperative. Studies dedicated to finetuning this model for increased scalability and computational efficiency, crucial for seamless real-time deployments, appear to be in short supply. (Huang et al., 2020; Zhang et al., 2020). From an ethical and legal standpoint, the literature offers glimpses into the concerns enveloping "black-box" AI solutions. Yet, a focused discourse on the ramifications of implementing Tab Transformers, particularly in sensitive sectors like healthcare and finance, remains amiss. The exploration on how these techniques is applied in a tabular data for increased decision making and recommendations. (Cerda et al., 2018; Cerda & El Varoquaux, 2019; Matt Davies & Michael Birtwistle, 2023; Narwaria, 2021; Oswald, 2023; Samek et al., 2017).

In conclusion, while the Tab Transformer stands as a testament to the advancements in processing tabular data, the myriad uncharted territories and unresolved questions underscore the vast potential and challenges awaiting future endeavours. The scarcity of papers that used Tab Transformer makes the research work unique. Bridging these gaps will not only magnify the prowess of the model but also augment the overarching discourse on AI's responsible and efficacious deployment. The next chapter would focus on the Methodology applied in the thesis. Using deep learning techniques to improve predictive process monitoring necessitates a methodical approach that includes data pre-processing, model selection, training, assessment, and deployment. By using this methodology, one can improve predictive process monitoring and make better decisions across a range of domains by methodically utilizing deep learning approaches. The review of other researchers' works around predictive process mining broadens the knowledge needed for this research methodology as implemented in the next chapter.

CHAPTER 3:RESEARCH METHODOLOGY

3.1 Introduction

In the multifaceted realm of research, the methodology acts as the compass guiding the explorative journey. It delineates the strategic procedures adopted for data acquisition, analysis, and interpretation. By offering a systematic blueprint, the methodology not only ensures the reproducibility of the research but also fortifies its legitimacy and reliability.

The Tab Transformer-enhanced predictive process mining approach used in this study demonstrates the need for a rigorous and comprehensive methodology. As shown in Figure 4.1, the PM2 Methodology (Van Eck et al., 2015) provides a structured and systematic approach to predictive process mining. The use of the Tab Transformer in this methodology enhances the predictive power of the model and helps to ensure a robust and accurate model is created. The methodology used in this study is a significant advance in the field of predictive process mining and will allow for more accurate and useful models to be created. This chapter provides an extensive exploration of the data acquisition techniques, elucidating both the novel and the traditional datasets that will be harnessed. As the work progresses, it will be logical to shed light on the bespoke methodological trajectory chosen for this research, meticulously detailing each stride and the encompassing research design.

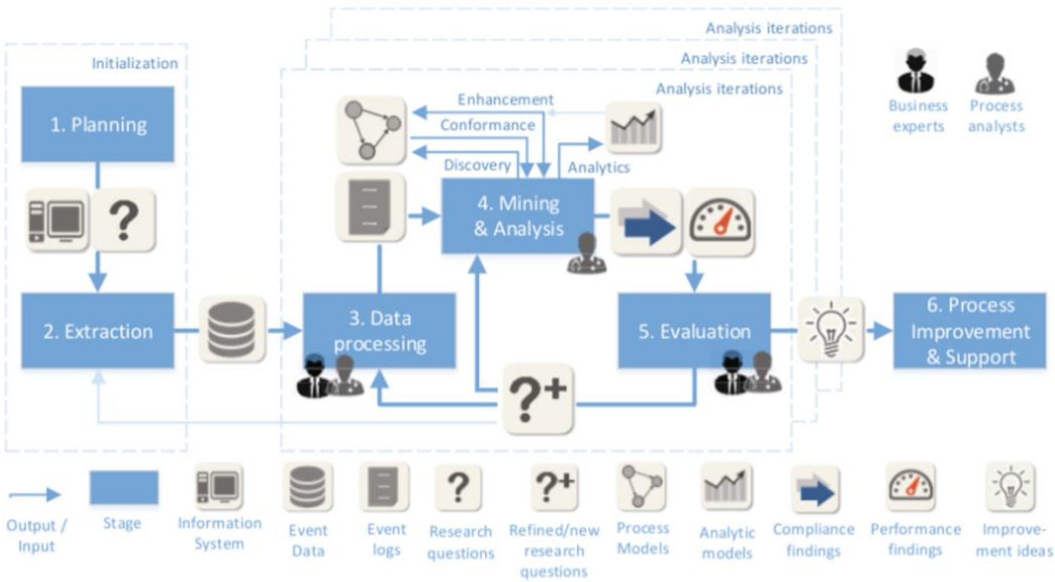

Figure 3-1 The overview of the PM2 Methodology retrieved from (M. L. Van Eck et al., 2015)

Historically, the realms of machine learning and data mining have witnessed the dominance of frameworks like the Crisp-DM Methodology, devised by SPSS, and SEMMA, crafted by SAS. Yet, as the boundaries of our research blur the lines between process mining and machine learning, we find ourselves resonating more with the PM2 methodology. A brainchild of PM2 offers a tailored approach to cater to the unique needs of our study by (M. L. Van Eck et al., 2015). To appreciate the significance of the chosen methodology, it is pertinent to briefly touch upon the evolution of methodological approaches, particularly in the intersections of data mining, machine learning, and process mining. For years, Crisp-DM held its ground as a seminal framework for many scholars and industry practitioners alike. Its structured approach—starting with a business understanding, culminating in deployment—offered a clear roadmap for a wide array of data mining tasks. Similarly, SEMMA, with its emphasis on sample, explore, modify, model, and assess, provided a robust method for many statistical data projects. Both frameworks, in their essence, were archetypes of their time, emblematic of a period where the boundaries of data and process were distinctly demarcated. Yet, with the surge in digital transformation and the increasing complexity of organisational processes, emerged a novel paradigm: process mining. Unlike traditional data mining, process mining placed emphasis on the 'process' itself, unravelling the sequences, patterns, and flow of tasks within an organisation. This shift in focus necessitated a fresh methodological approach—one that would be holistic yet flexible enough to accommodate the nuances of process analysis.

The PM2 methodology was adopted to suit the unique requirements as shown in Figure 3.1. The inherent flexibility of PM2 allowed the integration of innovative tools, such as the Tab Transformer. Its emphasis on ongoing monitoring aligned perfectly with our goal of predicting the next event. In the ensuing sections, there will be need to delve into the intricacies of each stage of the PM2 approach, shedding light on its principles, uses, and the specific modifications implemented for the study. The objective is to outline a transparent, comprehensive, and replicable framework that underpin the research. The methodology presented in this study not only serves as a roadmap for our current investigation, but also provides a valuable reference for future studies in this rapidly evolving field. The predictive process mining approach with the Tab Transformer is a powerful tool for analysing and understanding complex data, and we hope that our work will help to inspire further research in this area. This research creates excitement to see how this methodology will be applied in future studies, and how it will continue to evolve and improve over time.

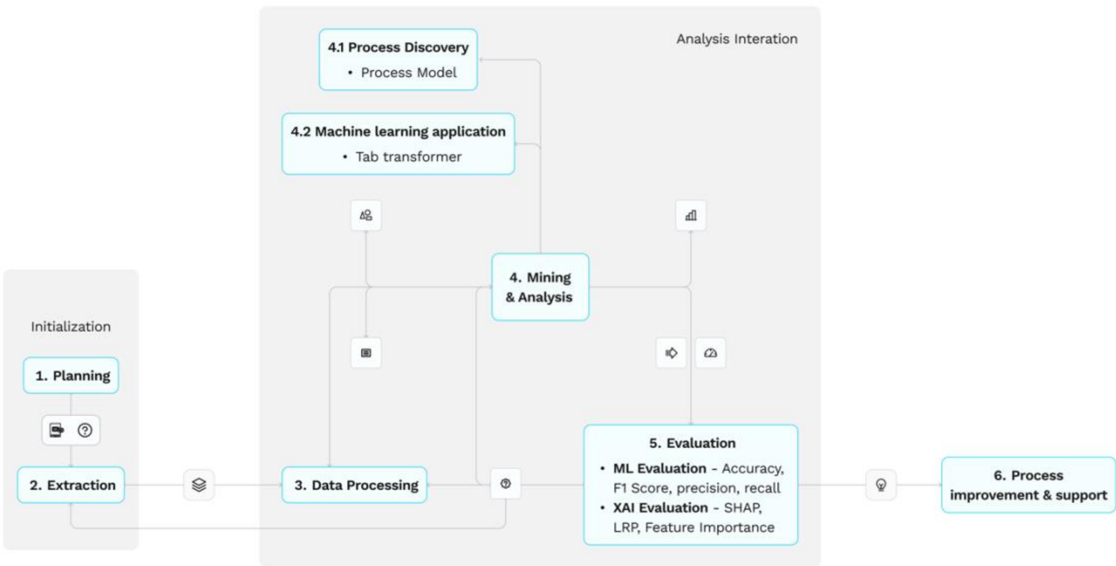

Figure 3-2 Adapted PM2 Methodology

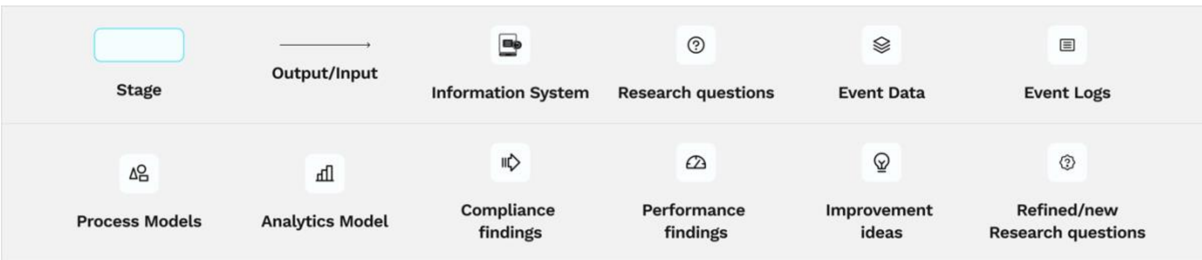

Figure 3-3 Description of pointers on Methodology

Planning: The planning stage is foundational and sets the trajectory for the entire research. Given the research problem's emphasis on predictive process mining challenges, especially with datasets containing multiple categorical features, the planning stage becomes crucial in ensuring that the research direction aligns with addressing these challenges.

The business process in this case is the research problems, which is about interpreting complex data. The data selected for this application should fit the adopted business processes. As part of implementation involves with machine learning, so an in-depth study of the evaluation metrics, applicable algorithms, is required in designing the project.

Extraction: The extraction stage is a pivotal step in my PhD research, focusing on the meticulous extraction of event data, and if necessary, process models. This stage is instrumental in ensuring that the subsequent analysis is grounded in accurate and comprehensive data, especially given the intricacies of predictive process mining. This stage the data is prepared as event log for pre-processing, during this phase the following stages would be implemented Determining scope, Extracting the Relevant Event Data and Transferring process Knowledge.

Data Processing: Data processing is a pivotal stage in the process mining research journey. It primarily focuses on transforming the raw event data into structured event logs, optimised for subsequent mining and analysis. This stage is instrumental in ensuring that the data is not only accurate but also presented in a manner that facilitates in-depth analysis and interpretation. The essence of data processing is to refine and structure the data, making it conducive for deriving meaningful insights. This stage involves the creation of view looking into the inclusion and exclusion criteria, aggregating events, enriching logs and filtering logs.

Mining and Analysis: The mining and analysis stage is the crux of the process mining methodology, where the structured event logs, crafted in the data processing stage, are subjected to rigorous analysis to extract meaningful insights. This stage is dedicated to applying process mining techniques and other data mining techniques to extract meaningful insights from the on-event logs to answer research questions, providing a comprehensive understanding of business processes, their performance, and compliance. The primary objective is to delve deep into the event logs, uncover patterns, and derive actionable insights that can drive process improvements.

Evaluation: The evaluation stage is a pivotal phase in the process mining methodology, serving as a bridge between the analytical findings and actionable improvement strategies. This stage is dedicated to critically examining the results obtained from the mining and analysis phase, ensuring they align with the project's objectives and are grounded in the real-world context of the business process. The primary goal is to translate the analytical findings into tangible improvement ideas or, if necessary, refine the research questions for subsequent iterations. At the heart of the evaluation stage is the need to ensure that the results are not only accurate but also meaningful and actionable.

Process Improvement: The "Process Improvement & Support" stage is a culmination of the insights and findings derived from the preceding stages of the process mining methodology. It is here that the theoretical insights are translated into actionable strategies for enhancing the actual process execution. This stage is pivotal as it bridges the gap between analytical findings and real-world applications, ensuring that the insights gained are not merely theoretical but have tangible implications for process enhancement.

However, in the context of a PhD thesis, the primary focus is on in-depth research, analysis, and the presentation of findings. While the actionable implementation of process modifications is beyond the scope of academic research, the insights and recommendations derived can serve as a valuable foundation for future practical applications.

3.2 Planning Stage

Within this section, the data acquisition process is examined and further discussed the various types of data collected with furnish detailed insights about the data itself.

3.2.1 Data Acquisition: Novel and Standard Data

Data forms the bedrock of any empirical research. The quality and relevance of data directly influence the accuracy and applicability of the research findings. In the context of the thesis, the justification for utilizing the MIMIC-IV_ED data along with BPIC 2012, BPIC 2017, and Road Traffic Data could be linked to Table 2.2 by emphasizing the diversity and complexity of the datasets to validate the robustness of predictive models. The inclusion of MIMIC-IV_ED, a comprehensive medical dataset, alongside the varied BPIC logs and Road Traffic Data, provides a multi-industry perspective on event prediction, allowing for a thorough examination of model versatility and adaptability across different domains.

The utilization of these datasets also aligns with the precedent set by the methodologies in Table 2.2, where diverse data sources have been employed to validate predictive models. For example, Pasquadibisceglie et al., 2020 used data from healthcare (Sepsis) and other industries (BPIC 2011, BPIC 2012, Production) to test their ORANGE method. The Orange method is a novel approach to predictive process mining that allows for the analysis of large and complex datasets. The data from the various sources was used to train and validate the Orange method, and the results were found to be accurate and robust. Similarly, other studies have utilized BPIC datasets from different years, reflecting the evolution of process management challenges.

By extending the range of datasets to include MIMIC-IV_ED and Road Traffic Data, the research benefits from a broader validation scope, which is crucial for demonstrating the generalization capability of the proposed approach. This justification underscores the commitment to creating predictive models that are not only accurate but also versatile and applicable to real-world scenarios across different sectors.

MIMIC-IV_ED Data

The Medical Information Mart for Intensive Care (MIMIC) dataset, in its fourth iteration, focuses on emergency department data. This rich dataset provides insights into patient demographics, vital signs, laboratory tests, medications, and more. It serves as a valuable resource for understanding patterns and predicting outcomes in a medical setting. This also allows for the work to be transferable across various medical settings (Alistair Johnson et al., 2023). Due to the nature of medical data, accessing these data can be quite complicated especially in settings such as the NHS and other medical institutions. With the MIMIC-IV_ED dataset, we can create an event log that like what could be generated from other medical institutions. We applied to the Physio Net team in MIT, for access and had to complete certification in CITI Data or Specimen Only Research.

BPIC 2012

Process mining, in today's digitized business landscape, has emerged as a potent tool for extracting actionable insights from vast process-related datasets. The Business Process Intelligence Challenge (BPIC) stands as a beacon of collaboration within this realm, catalysing innovation by offering real-world datasets for research. BPIC's 2012 dataset, sourced from a renowned global financial institution and diligently curated by Boudewijn van Dongen in 2012, delves deep into the nuances of the loan application process. With its exhaustive log capturing each event, from loan initiation to its conclusion, the dataset is a treasure trove of data points like timestamps, resources, and key decisions. It boasts of remarkable granularity, voluminous data, diverse event types, and has been ethically prepared with thorough anonymization. Researchers tapping into this resource can explore the intricate journey of each loan application, analyse patterns, predict outcomes, and derive strategies to optimize the loan approval process.

> **Granularity**: The dataset provides fine-grained insights into each step of the loan application process. Each event is tagged with a precise timestamp, allowing for intricate temporal analysis and bottleneck identification.
>
> **Volume**: With thousands of individual applications and events, the BPIC 2012 dataset offers ample data for robust statistical analysis, minimizing the impact of outliers and ensuring generalizable insights.

Variety: The dataset captures a multitude of event types, representing various subprocesses within the broader loan application workflow. This diversity allows for a comprehensive view of the process, from preliminary checks to final decisions.

Anonymization: To ensure confidentiality and adhere to data protection norms, all sensitive and personally identifiable information within the dataset has been anonymized. This guarantees ethical usage while preserving the integrity of the process data.

The BPIC 2012 dataset provides a detailed look into the loan application process of a Dutch financial institution. Each case in the dataset represents an individual loan application, capturing its journey from initiation to either approval or rejection. These cases encompass a sequence of events or activities, such as starting the application, verifying details, and making final decisions. Each event within a case has attributes like its name, timestamp, executing resource, and lifecycle transition. Cases also have specific attributes, including the type of application (personal loan or overdraft), the requested amount, and the outcome (approved or declined).

BPIC 2013

The BPIC 2013 dataset emerges as a fundamental component of this doctoral research. Originating from a real-world IT company's incident management system, this dataset encompasses a comprehensive log of incident management activities, providing a detailed perspective on the processes involved in managing IT incidents from their inception to resolution. The dataset's meticulous detailing, including timestamps, activities, and process pathways, offers an unmatched resource for analysing and understanding the complexities of incident management workflows.

The significance of the BPIC 2013 dataset in this research aligns with previous studies in the next event prediction task. It serves as an essential tool for the empirical evaluation of process mining algorithms, facilitating the identification of patterns. The real-world origin of the dataset ensures that the insights obtained are both theoretically sound and practically applicable, effectively bridging the gap between academic research and operational IT service management practices. A primary focus of the thesis is on making next event predictions and comparing our approach with previous work; the dataset is ideal, as it has been utilised by several researchers, as detailed in Table 2.2 in our literature review. This approach aims to contribute to the broader field of predictive process mining by showcasing the usability and applicability of our methodology.

BPIC 2017

The Business Process Intelligence Challenge (BPIC) consistently offers valuable datasets to the academic and business communities to drive advancements in process mining. The BPI 2017 dataset, one of its salient contributions, provides meticulous insights into the loan application processes of a Dutch financial institution.

The BPI 2017 dataset is sourced from an online loan application system of the institution. It records all loan applications initiated throughout 2016 and continuously monitors them until a specific endpoint: February 1st, 2017, at 15:11. A comparative analysis reveals that this dataset corresponds to the same institution as encapsulated by the dataset with the identifier doi:10.4121/uuid:3926db30-f712-4394-aebc-75976070e91f (Boudewijn van Dongen, 2017). Despite this similarity, there is a discernible evolution in the systemic infrastructure and its functionalities from the earlier datasets.

Structurally, the BPI 2017 dataset is organized as an event log. Key attributes characterizing each event encompass:

- **Case ID**: Denoting a unique loan application.
- **Event ID**: A specific identifier for individual events within a case.
- **Timestamp**: The exact chronological marker of the event's occurrence.
- **Activity Name**: Elucidates the specific activity, such as "Initial Offer Extended" or "Loan Finalized".
- **Resource**: An indicator of the entity (person or system component) responsible for the event's execution.
- **Offer ID**: A distinguishing attribute introduced in this iteration; it differentiates between multiple offers related to a singular application.
- **Status**: Signifies the current stance of an offer, categorized as accepted, rejected, or pending.

For academic endeavours, the dataset is provisioned with open access, typically hosted on the BPIC's official portal or affiliated academic repositories. Researchers may be required to acknowledge terms of use or complete a nominal registration. The dataset is generally available in XES or CSV format. An important augmentation in the BPI 2017 dataset, compared to its predecessors, is its capability to encapsulate multiple offers within a single loan application trajectory. This nuanced inclusion not only adds a layer of complexity but also enriches the dataset, offering a multifaceted perspective into the financial institution's decision-making paradigms and strategic negotiations. The data would allow us to benchmark our application and model across other published work, providing detailed insights towards how our model perform on various instances.

Our primary data, which was created from the MIMIC_IV Dataset provides an opportunity for use to integrate various features and visualise how they contribute towards the prediction observed, the standard data allows us to visualise how the model would perform across various datasets.

3.3 Data Preparation

The incorporation of MIMIC-IV_ED data is pivotal to our research. This dataset not only bridges the gap from previous studies that employed MIMIC-III data for process mining but also advances the field by offering newer, more comprehensive data meticulously transformed the databases from the MIMIC repository into a coherent event log suitable for our analytical pursuits. as delineated by (Alistair Johnson et al., 2023)

The Emergency Department (ED) serves as a critical pillar in the healthcare infrastructure, addressing immediate and often life-threatening medical exigencies. The ceaseless operational demands of EDs require that they be equipped both in terms of expertise and resources. Patients arriving at the ED undergo an initial assessment or triage, which determines the subsequent course of medical intervention. The setting is characterized by its dedicated resources, which are optimized to augment the likelihood of favorable patient outcomes. In this backdrop, the datasets from MIMIC ED grant researchers an unprecedented lens to scrutinize the intricate dynamics of emergency care.

The data is structured into six tables which is depicted in Figure 4.4 below, each capturing specific facets of patient care:

- Edstays: This table tracks how patients are admitted and discharged from the Emergency department after a single stay.
- **Diagnosis**: This table provides contains the diagnosis of the patients coded with the International Classification of Diseases Ninth or Tenth revision.
- **Medrecon**: This table provides a list for the medications the patient was taking prior to ED admission.
- **Pyxis**: This table provides information for the dispensation of medication.

- **Triage**: This table contains information gathered from patients at the time of their triage, this process involves generating insights about the patient health status and reason for visiting.
- **Vitalsign**: The patient's aperiodic vital signs that were recorded throughout their stay are included in the vital-sign table.

A healthcare information model is a representation of how tables in a medical database relate to one another and to potential healthcare events. Because it facilitates the extraction of event logs and the comprehension of process-oriented queries, the information model is important in process mining research

Figure 3-4 MIMIC-IV_ED data reference model

To further our analytical endeavours, we have curated a csv file and an event log. These formats facilitate visual representations of the process and enable the application of our technique on structured data. We would explore how MIMIC data was previously utilised and how our work is different.

3.4 Previous Application of MIMIC Datasets in Process Mining

The research conducted by Alharbi as part of the thesis in unsupervised Abstraction techniques utilised the MIMIC-III Database to generate an event log for reducing complexity with the discovered process maps using an abstraction approach in discovering process maps, the research approach followed the following stages event log extraction, pre-processing, learning, decoding, optimisation, selection, Model visualisation, Evaluation (Alharbi, 2019).
The research conducted by Kusuma, and other researchers focused on developing log transformations tools to enable further researchers utilising the records. The steps provided in the research was based on the following stages planning, Extraction, Data Analysis and Mining & Analysis. These are the steps used in generating actionable insights from the data (Kusuma et al., 2021).
The research conducted by Rojas focussed on assessing the data quality issues associated with process mining, in the research conducted he utilised the MIMIIC-III datasets, the method used was an adaptation of the L* life-cycle model to include assessing the data quality of the event log. The research commended the availability of the data, as it enables research to be validated and replicated by other researchers (Rojas et al., 2018).

The research of Kurniati was focused on an exploratory approach in understanding the cancer treatment processes within the MIMIC-III dataset, the research utilised the approach of presented by van der Aalst in the process mining manifesto of 2011 called the L∗ life-cycle model describing a process mining project (M. L. Van Eck et al., 2015). Kurniati added some addition steps form the pm2 process mining project methodology to cater for the extraction and data processing stage (M. L. Van Eck et al., 2015). The findings from the research indicated a necessity for enhanced data cleaning and the deployment of additional algorithms to glean deeper insights, as suggested by (Kurniati et al., 2018). This study centres on readying the data for event prediction tasks, contrasting past efforts that concentrated on uncovering processes behind activities. Specifically, this research is geared towards offering predictions on the next activity and subsequently elucidating the rationale behind those predictive outcomes. The study aspires not just to visualize but to harness the entirety of the data's information. This approach marks a departure from previous methodologies, aiming to offer clarity and understanding regarding the basis of our predictions. The next section would focus on data processing of the primary data.

3.5 Data processing

The primary objective of this stage is the creation of event logs format. The process involved assigning the actual names required for an event log and gathering further general insights on the data. This study would start by performing a data exploratory analysis, then we would discuss the pre-processing steps taken.

3.5.1 Data Exploratory Analysis

Our research revolves around a comprehensive analysis of patient activities and procedures in the Emergency Department (ED). The focus is to not only study the existing patterns but to predict future trends and improve patient care. In this section we provide information about our data. Figure 4.5 shows a bar chart represents that shows the distribution of various activities and procedures that patients undergo in the ED and outcome. The top three activities, by volume, are 'Medicine reconciliation', 'Medicine dispensations', and 'Vital sign check', respectively.

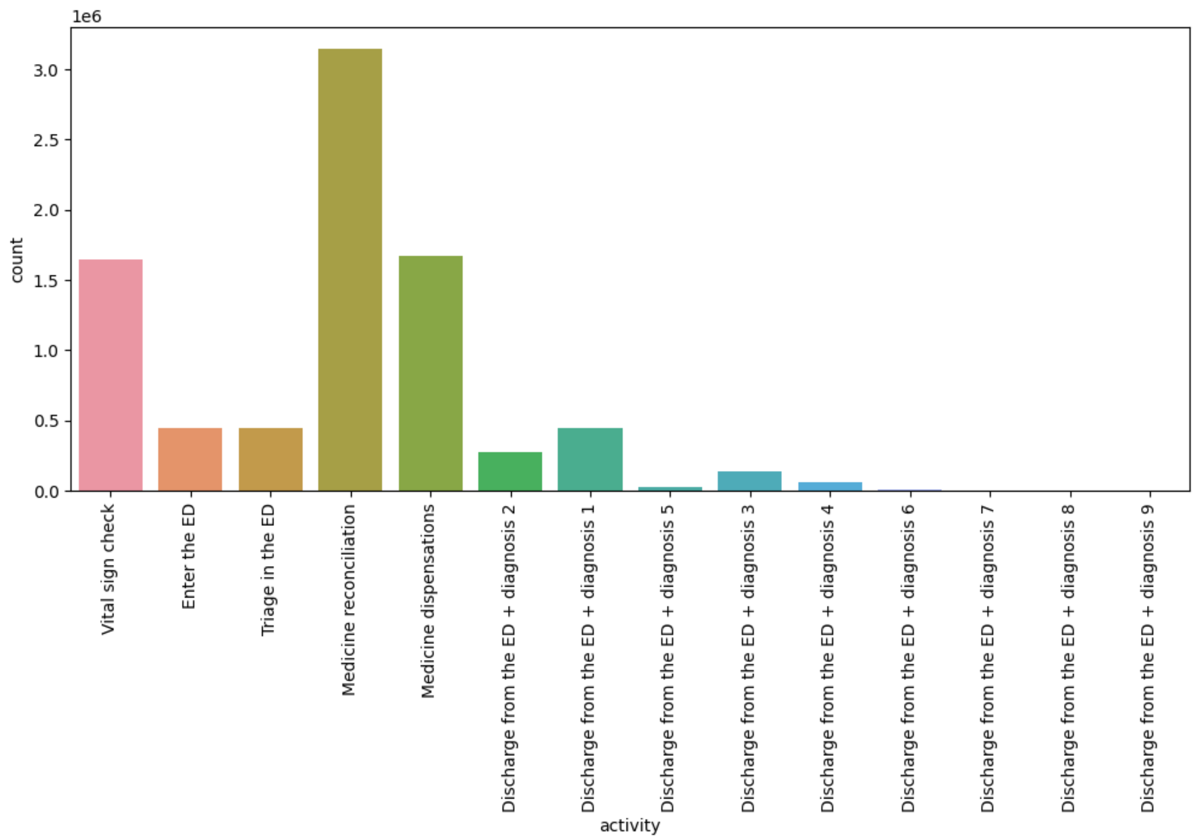

Figure 3-5 Bar chart on MIMIC_IV_ED Data

The information displayed in table 4.1 below, shows that Medicine Reconciliation ranks highest with 3,143,791 instances, indicating that a significant portion of the patient population requires a review and verification of their medication details while the information in table 4.2 below provides information on the patient data encompassing several patient attributes. Medicine Dispensations are noted 1,670,590 times, suggesting a vast number of patients are given medication during their ED visit. The third most common activity is the Vital Sign Check with 1,646,976 instances, which is essential for monitoring patient's immediate health condition.

Activity/Procedure	**Count**
Medicine reconciliation	3,143,791
Medicine dispensations	1,670,590
Vital sign check	1,646,976
Enter the ED	447,712

Triage in the ED	447,712
Discharge from the ED + diagnosis 1	446,530
Discharge from the ED + diagnosis 2	269,945
Discharge from the ED + diagnosis 3	134,685
Discharge from the ED + diagnosis 4	59,303
Discharge from the ED + diagnosis 5	23,818
Discharge from the ED + diagnosis 6	8,696
Discharge from the ED + diagnosis 7	2,730
Discharge from the ED + diagnosis 8	788
Discharge from the ED + diagnosis 9	197

Table 3-1 MIMIC_IV_ED Activity Details

On the other end, the number of patients being discharged with a higher count of diagnosis (from diagnosis 5 to diagnosis 9) is significantly lower, suggesting that most patients in the ED have fewer complications.

#	Column	Dtype		#	Column	Dtype
0	stay_id	int64				
1	gender	object		16	o2sat	float64
2	race	object		17	sbp	float64
3	arrival_transport	object		18	dbp	float64
4	disposition	object		19	pain	object
5	subject_id	int64		20	acuity	float64
6	hadm_id	float64		21	chiefcomplaint	object
7	timestamps	object		22	rhythm	object
8	activity	object		23	name	object
9	seq_num	float64		24	gsn	float64
10	icd_code	object		25	ndc	float64
11	icd_version	float64		26	etc_rn	float64
12	icd_title	object		27	etccode	float64
13	temperature	float64		28	etcdescription	object
14	heartrate	float64		29	med_rn	float64
15	resprate	float64		30	gsn_rn	float64

Table 3-2 MIMIC_IV_ED Data column details

- **Demographics**: This includes gender, race, and other identification details.

- **Arrival Mode**: Various methods through which patients arrive at the ED, such as walking, ambulance, or helicopter.
- **Medical Metrics**: These include vital signs like temperature, heart rate, respiratory rate, blood pressure, and more.
- **Diagnosis and Medication**: Details on ICD codes, medication details, and their descriptions.

Mode of Transportation	Count
Walk in	4,478,823
Ambulance	3,604,183
Unknown	179,599
Other	23,826
Helicopter	17,042

Table 3-3 Arrival transport details

The Arrival Transport Details section offers insights into the modes through which patients arrive at the emergency department as shown in table 3-3. This categorization includes standard methods like "WALK IN" and "AMBULANCE," as well as more specialized ones such as "HELICOPTER." Each mode is represented with a count, shedding light on the prevalence of each transportation method. In summation, the dataset not only provides a holistic view of the operations within the emergency department but also emphasizes discerning the significance of specific columns. Their potential influence on predicting subsequent activities is a focal point of this research. The subsequent sections will delve deeper into the pre-processing steps and the analytical methodologies employed to derive these insights.

3.5.2 Data Pre-processing approach

This approach was applied to all the datasets, to form a level of consistency.

In my pursuit to optimize the dataset for subsequent analyses, I meticulously crafted a pre-processing function. This function not only reshaped the data into an appropriate format but also augmented it with pivotal features for my research. The following elucidates the rationale and logic behind each pre-processing step:

Timestamp Conversion

I transitioned the timestamps in the dataset to a standardized datetime format. This step was crucial to ensure that all ensuing operations involving time were consistent and accurate.
Rationale: Handling time data in its raw string format can lead to inconsistencies. By transitioning it to a standardized datetime format, I could effortlessly utilize built-in functionalities for time-based operations, a sentiment echoed by (Narajewski et al., 2021).

Sorting by Timestamp

The dataset was sorted chronologically based on timestamps, ensuring the sequence of events remained intact, which is vital for time series analysis.
Rationale: The sequence of events in a patient's journey through the emergency department can yield invaluable insights. By maintaining events in chronological order, I could document the natural progression of each patient's experience.

Trace Positioning

A 'tracePosition' column was introduced, counting the number of events for each patient case, providing a sequential position for each event within its respective case.
Rationale: Recognizing the position of an event within a patient's journey can be pivotal in forecasting subsequent activities. This trace position acts as a relative timestamp, offering context to each event.

Day Categorisation

The dataset was enhanced by categorizing each event based on the day of the week it occurred, leading to the creation of columns such as 'weekday', 'saturday', and 'sunday'.
Rationale: The day of the week can significantly influence the operations of an emergency department. By categorizing events based on the day, I aimed to capture these potential variations, a perspective supported by (Becker & Bagrow, 2019).

Encoding and Dummies

In this study, categorical data within the dataset were subjected to encoding processes to ensure compatibility with the machine learning algorithms utilized. Specifically, the 'event:concept:name' column underwent one-hot encoding, a process that generated dummy variables to represent the distinct categories within the column. This encoding technique was selected to preserve the categorical distinctions without introducing an ordinal relationship that could potentially mislead the algorithm. In parallel, label encoding was applied to several other categorical columns, namely 'gender' and 'race'. This encoding strategy was deemed suitable for these columns as it transformed the categorical values into a numerical format, a requirement for the machine learning models employed in this analysis. The rationale behind these encoding approaches stems from the established practice of converting categorical variables into a numerical format to facilitate machine learning algorithms, a practice underscored by (Dahouda & Joe, 2021; Kosaraju et al., 2023; Valdez-Valenzuela et al., 2021).

Target variables Calculation

In this study, I designated specific target variables, namely 'true_activity' and 'true_activity_life', to represent the subsequent activity corresponding to each event. This formulation was driven by the primary objective of predicting the forthcoming activity based on the given data. The delineation of these target variables serves as a foundational step towards employing supervised learning techniques. By establishing a clear representation of subsequent activities through the target variables, I have created a structured framework that facilitates the application of predictive algorithms to achieve the desired forecasting accuracy. In summary, the pre-processing steps executed in this study transcended mere mechanical transformations; each step was underpinned by a well-considered rationale. This rationale was intimately intertwined with the nuanced operational dynamics of emergency departments and aligned with the overarching objectives of my research. By meticulously tailoring the pre-processing procedures to the distinctive characteristics and demands of the emergency department setting, I aimed to foster a robust analytical framework conducive to insightful analysis and meaningful predictions.

3.6 Mining and Analysis

In the implementation phase of our study, we engaged in a multifaceted approach to extract insights from the data. Initially, we employed process mining techniques, enabling us to visualize the data and delineate the processes undertaken. This visualization served as a

precursor to our next objective: the development and implementation of a predictive model aimed at forecasting subsequent events within our process framework.
Furthermore, we intend to incorporate Explainable Artificial Intelligence (XAI) methodologies. The adoption of XAI will not only facilitate deeper insight generation but also provide a lucid understanding of the mechanisms through which our models derive results. A primary aim here is to discern the variables significantly impacting the decision-making process, thus promoting a more transparent and interpretable model operation.

3.6.1 Process Mining Implementation

The process mining initiative was undertaken employing a range of algorithms, notably the Alpha and Heuristic Miner algorithms, to conduct an initial analysis on the data. Subsequent to this initial analysis, we opted to utilize the Disco software tool to achieve a more enriched visualization of the process map. This transition to Disco facilitated a comprehensive identification and visualization of the various processes involved. Figure 3-6 below displays a sample log generated from this implementation, exemplifying the practical application of process mining on the MIMIC_IV_ED dataset.

	stay_id	timestamps	activity	gender	race	arrival_transport	disposition
0	30000012	2126-02-14 20:22:00	Vital sign check	F	WHITE	AMBULANCE	ADMITTED
1	30000012	2126-02-14 20:22:00	Enter the ED	F	WHITE	AMBULANCE	ADMITTED
2	30000012	2126-02-14 20:22:01	Triage in the ED	F	WHITE	AMBULANCE	ADMITTED
3	30000012	2126-02-14 22:21:00	Medicine reconciliation	F	WHITE	AMBULANCE	ADMITTED
4	30000012	2126-02-14 22:21:00	Medicine reconciliation	F	WHITE	AMBULANCE	ADMITTED

Figure 3-6 sample of event log of MIMIC_IV_ED data

3.6.2 Model Building

In our model-building section, we delineate the data preparation methodology, emphasizing the criticality of representative training and test datasets. We adopted a stratified split to maintain consistent class distributions across both datasets, ensuring that our model's performance metrics are reflective of its true predictive power. Additionally, we utilized a fixed random state to guarantee the reproducibility of our data splits—an essential practice for transparent and verifiable scientific research.

Our approach judiciously combines train-test splits and cross-validation, leveraging their respective strengths in different scenarios. While train-test splits offer computational efficiency, making them suitable for preliminary learning and exploratory analysis, cross-validation is employed when our dataset size is moderate to small. This method amplifies the reliability of our model evaluation, protecting against overfitting by averaging performance across multiple partitions of the data.

For our substantial datasets, possibly spanning millions of records, we favour the train-test split due to its sufficiency in capturing data variability and providing reliable evaluation outcomes. Conversely, for smaller datasets, cross-validation emerges as the superior choice, enhancing the robustness of our performance assessment.

The objective of the research is to construct a predictive model that is both robust and versatile. To achieve this, we have chosen to employ a combination of advanced machine learning algorithms, each bringing its unique strengths to the table. The algorithms selected for this endeavour are Random Forest, XGBoost, and the Tab Transformer, these algorithms have been deeply introduced and explained in section 2.4.2. Let's delve deeper into the rationale behind this selection and the methodology we will employ. For the empirical evaluation of our model, we will be utilising four distinct datasets. These datasets have been meticulously curated to ensure a comprehensive understanding of the underlying patterns and relationships. By training and testing our model on these diverse datasets, we aim to ensure that our model is not only accurate but also generalisable across different scenarios.

ML MODEL IMPLEMENTATION

In our study, we've incorporated several machine learning models, including the Random Forest, XGBoost, and the advanced Tab Transformer, with the latter designed for tabular data. Each of these models necessitates distinctive pre-processing, hyperparameter tuning, and fine-tuning steps, which we elucidate below:

TAB TRANSFORMER

Firstly, we would delve deeper and discuss our approach and the approach published by Bukhsh et al 2021 in the application of Transformer Architecture, we would analyse the difference between our implementation and theirs. Looking at various dimensions such as architectural approach, data processing, training protocols, versatility, and computational complexity (Bukhsh et al., 2021). The work conducted by the Bukhsh, although predicted the next event, is not able to capture the effects of the categorical features associated in the process, this is one of the limitations that our approach solves, with our customised encoding, we are able to gain further insights on the data. Our approach extends for the applicability of Explainable Ai on our architecture.

Criteria	Bukhsh Methodology	Our Methodology
Dimensionality of Architectural Approach	Utilizes standard Transformer model focusing on forecasting subsequent activities.	Employs a nuanced Tab Transformer structure, managing categorical and numerical attributes independently.
Architectural Components	Token and positional embeddings, Transformer blocks, attention, and feed-forward networks.	Multi-head attention, categorical embeddings, Layer Normalization.
Data Processing and Encoding Strategies	Focuses on structuring sequential data into tokenized forms through comprehensive pre-processing.	Employs intricate encoding mechanism focusing on embedding categorical data, adaptable across diverse datasets.
Training Protocols and Model Optimization	Trained using conventional protocol with Adam optimizer and Sparse Categorical Cross entropy as a loss function.	Incorporates multifaceted training regimes and sophisticated optimization techniques due to advanced architecture.
Versatility and Application Spectrum	Specialized and oriented towards a singular task, limiting its applicability to diverse scenarios.	Demonstrates superior adaptability and versatility, extending its applicability to a wider array of tasks and scenarios.
Algorithmic Sophistication and Computational Complexity	Balanced between computational complexity and specificity, employing canonical transformer mechanisms.	Engages in advanced mechanisms and nuanced handling of disparate data types, demanding sophisticated comprehension.
Specialization	Highly specialized in sequential prediction tasks.	Universally applicable solution due to its ability to cater to varying categories of numerical and categorical data.

Table 3-4 Comparison Our approach with the work done by Bukhsh et al., 2021

The table 3.4 provides a comparison between our approach and that of bukhsh, in the table we mentioned several steps taken such as intricate Encoding mechanisms, Multifaceted

Training Regimes and sophisticated Optimisation, Superior Adaptability and Versatility, Advanced Mechanisms and Nuanced handling of Disparate Data Types.

- **Intricate Encoding Mechanism:**

  In our methodology, we employ a specialized and intricate encoding mechanism focusing primarily on embedding categorical data. This mechanism is crucial for transforming categorical variables into a numerical format, using advanced techniques such as embedding layers, enabling the model to represent categorical data in higher dimensions and capture more complex relationships within the data. Each categorical feature undergoes a separate embedding process, and the resulting embeddings serve as refined inputs to the model. This approach allows the model to understand and learn the categorical data more effectively. Furthermore, this encoding mechanism is highly adaptable, designed to handle a variety of datasets with different categorical features seamlessly. This adaptability ensures the model's versatility and applicability across diverse problems and domains, making it a robust solution for a multitude of tasks involving categorical data.

- **Multifaceted Training Regimes and Sophisticated Optimization:**

  The methodology designed incorporates advanced and multifaceted training regimes coupled with sophisticated optimization techniques, attributed to the model's advanced architecture. The model is trained using refined training protocols, including learning rate scheduling, gradient clipping, and the Adam optimizer with weight decay and clip value. These advanced techniques stabilize the training process, preventing overfitting and ensuring the model learns the underlying patterns in the data effectively. The sophisticated architecture of the model, including components like multi-head attention and layer normalization, necessitates such advanced training and optimization processes. These processes enable the model to learn and adapt effectively to the complexities and nuances of the diverse datasets it encounters.

- **Superior Adaptability and Versatility:**

  The model designed demonstrates a high degree of adaptability and versatility, extending its applicability to a wider array of tasks and scenarios. The architecture of the model, enriched with advanced components like multi-head attention and layer normalization, is designed to learn complex patterns and relationships in the data. This design makes the model suitable for a plethora of tasks, ranging from classification to regression, and across various domains. The adaptability and versatility of the model are paramount, allowing it to be a universally applicable solution, capable of catering to varying categories of numerical and categorical data across diverse application spectrums.

- **Advanced Mechanisms and Nuanced Handling of Disparate Data Types:** This methodology designed engages in advanced mechanisms and employs nuanced strategies to handle disparate data types, including both categorical and numerical data. The model processes numerical and categorical features separately, applying normalization to numerical features and embedding to categorical ones. This separate processing enables the model to handle mixed data types effectively and efficiently. The nuanced handling of different data types and the advanced computational strategies employed require a deep and sophisticated comprehension of the data and the relationships between features. This level of comprehension and analysis is crucial for developing models that are effective and robust, capable of handling the complexities and diversities of the datasets they are applied to.

After discussing the differences between this work and Bukhsh's approach, this research now presents its own methodology and highlights its unique features. The approach taken in this work is distinct from Bukhsh's in several ways, including the use of a different process mining tool (PM4Py) and the application of the Tab Transformer, a powerful machine learning algorithm. The approach also makes use of a refined version of the PM2 Methodology, which has been adapted for use with the PM4Py tool. The Tab Transformer is a sophisticated model tailored for tabular data. Drawing inspiration from the transformer architecture widely used in natural language processing tasks, the Tab Transformer has been fine-tuned to address the nuances and challenges posed by structured data. Unlike conventional models that treat each feature independently, the Tab Transformer captures interactions between features, making it potent in deciphering intricate patterns.

Model Components:

- **Embeddings**: At the heart of the Tab Transformer lies the embedding mechanism. This system translates categorical features into dense continuous vectors, ensuring they are in a format conducive for deep learning models. The dimension of these embeddings, controlled by the embedding dims parameter, is critical. It determines the amount of information each categorical feature can encapsulate.
- **Multi-Head Attention**: Borrowed from the transformer architecture, the multi-head attention mechanism enables the model to concentrate on different portions of the input data simultaneously. This parallel attention mechanism captures various facets and relationships within the data, enriching the model's understanding.
- **Positional Embeddings**: For sequences or ordered data, positional embeddings are added to provide the model with a sense of order, enabling it to factor in the significance of the feature's position.
- **Feed-forward Neural Networks**: After the attention mechanism, the model employs feed-forward networks. These networks further process the information, adding depth and complexity to the model's learning capability.

Hyperparameters and Their Implications:

1. **LEARNING RATE**
   - **Description**: The learning rate is a crucial hyper parameter that determines the step size at each iteration while moving towards a minimum of the loss function. In essence, it controls how swiftly or gradually the neural network adjusts its weights during the training process.
   - **Rationale**: A smaller learning rate, as chosen (0.00002), allows for a meticulous and incremental update to the weights. Such an approach is beneficial in preventing drastic changes that might cause the model to overshoot the optimal weights. In situations characterized by a convoluted loss landscape, such a conservative learning rate can aid in achieving more stable convergence.
2. **WEIGHT DECAY**
   - **Description**: Weight decay is a regularization technique that penalizes larger weights. It effectively works by adding an additional term to the loss that's proportional to the size of the weights.

- **Rationale**: A weight decay of 0.001 implies that during each update, the weights are slightly decayed by this factor. This discourages the model from having overly large weights, thereby preventing overfitting.

3. **DROPOUT RATE**
   - **Description**: Dropout is a widely used regularization method wherein a predetermined fraction of neurons in a layer are randomly omitted or "deactivated" during the training phase.
   - **Rationale**: Implementing a dropout rate of 0.3 signifies that, during training, 30% of the neurons in the designated layers are disregarded in each forward pass. This mechanism ensures that the neural network does not become exceedingly dependent on any particular neuron, fostering generalized and resilient learning.
4. **BATCH SIZE**
   - **Description**: The batch size delineates the number of training examples processed during a single forward and backward pass. This hyperparameter profoundly impacts the optimization dynamics and computational requirements.
   - **Rationale**: Selecting a batch size of 32 offers an optimal trade-off between computational efficiency, achieved through parallel processing, and the veracity of the gradient estimate. This choice is informed by extensive empirical studies in the deep learning domain, which have showcased its efficacy across a variety of tasks.
5. **Number of Epochs (NUM_EPOCHS)**
   - **Description**: An epoch is representative of one complete cycle, where the model processes the entirety of the training dataset once, performing both forward and backward passes.
   - **Rationale**: Opting for 5 epochs implies the model undergoes training over the entire dataset twenty times. This number has been chosen to find a judicious equilibrium between the risks of underfitting (where the model fails to capture the underlying patterns in the data) and overfitting (where the model becomes overly attuned to the training data and performs poorly on unseen data). It ensures that the model adequately learns the patterns without over-extending its training duration.

6. **Number of Transformer Blocks (NUM_TRANSFORMER_BLOCKS)**
   - **Description**: This parameter dictates the number of transformer blocks integrated into the Tab Transformer model.

   - **Rationale**: The Opting for a four-transformer block configuration strikes a balance between complexity and efficiency. This design decision implies that, for our specific task, we believe a moderate model depth is optimal. It suggests we aim to harness the transformer's capabilities without overcomplicating the architecture, possibly because our dataset's characteristics and complexities don't demand an excessively deep structure.
7. **Number of Attention Heads (NUM_HEADS)**
   - **Description**: This parameter refers to the total number of attention heads present in the multi-head attention mechanism, a foundational component of the transformer architecture.
   - **Rationale**: Incorporating eight attention heads empowers the model to concurrently concentrate on five distinct portions of the input data. This configuration ensures the model can discern and assimilate a richer tapestry of patterns and relationships in the data, thereby improving its capacity to capture diverse contextual relationships.
8. **Embedding Dimensions (EMBEDDING_DIMS)**
   - **Description**: This parameter specifies the dimensionality of the dense vector representations, or embeddings, into which categorical features are mapped.
   - **Rationale**: Selecting an embedding size of 16 offers a harmonious balance between dimensionality and expressiveness. It ensures that the embeddings are succinct yet potent enough to encapsulate the critical nuances of the categorical data.
9. **MLP Hidden Units Factors (MLP_HIDDEN_UNITS_FACTORS)**
   - **Description**: These coefficients influence the configuration of hidden units within the Multi-Layer Perceptron (MLP) segments of the model. Specifically, the hidden unit count is determined by multiplying these factors with the input count.

- **Rationale**: Adopting factors of [1, 1] denotes that the count of neurons or units in the MLP's hidden layers mirrors that of its inputs. This symmetry ensures a steady flow of information and aids in preserving the structural integrity of the data within the MLP.

10. Number of MLP Blocks (NUM_MLP_BLOCKS)

- **Description**: This parameter is indicative of the total count of distinct MLP blocks incorporated into the model's architecture.
- **Rationale**: Incorporating two MLP blocks endows the model with an augmented capacity for nonlinear data transformations. This is instrumental in identifying and representing multifaceted relationships in the data, thus enhancing the model's expressive power.

Turning attention to the encode input’s function, meticulously designed to prepare data for seamless integration with our model.

The encode_inputs Function: Bridging Raw Data to Model-Ready Format

Tabular data presents a unique challenge in deep learning. The variance in data types—specifically categorical and numerical—demands specialized pre-processing to feed them into neural networks. Recognizing this, we've designed the encode_inputs function to navigate the complexities of this transformation. In this section, we elucidate the inner workings of this function, detailing its pivotal role in preparing data for the Tab Transformer.

1. **Categorical Feature Processing**: Categorical data holds distinct attributes that render it incompatible for direct insertion into neural models. This incongruence necessitates a series of methodical transformations, outlined below:

   a) **Vocabulary Extraction**: At the onset, it's imperative to discern the breadth and uniqueness of values within each categorical feature. To this end, a vocabulary is meticulously constructed, chronicling every unique value a categorical column holds. This vocabulary, apart from serving as a record, facilitates subsequent encoding stages.

b) **String to Integer Conversion**: With the vocabulary in place, the next imperative is to convert these categorical string values into an interpretable format for neural networks. Harnessing the capabilities of the StringLookup layer, we map each distinct categorical value to a unique integer. This deterministic mapping ensures reproducibility and paves the way for the next transformation phase.

c) **Embedding Creation**: Numerical representations alone, although a step forward, don't fully harness the potential of deep learning. To remedy this, we employ an embedding layer. Here, each integer value representing a category is transformed into a dense, continuous vector. These embeddings not only enhance model input expressiveness but also capture intricate relationships between categories.

d) **Storage**: Post-transformation, it's essential to ensure systematic storage for efficiency and subsequent retrieval. The processed embeddings are hence catalogued in the encoded_categorical_feature_list.

2. **Numerical Feature Processing**: Unlike categorical data, numerical features already reside in a model-compatible format. However, ensuring their seamless integration with the model demands additional steps:

    a) **Dimension Expansion**: To homogenize the input data structure and to ensure seamless tensor operations, numerical features are subjected to dimension expansion. This procedure ensures that numerical data aligns flawlessly with the architectural expectations of the Tab Transformer.

    b) **Storage**: Following their transformation, these processed numerical features are securely stored in the numerical_feature_list. This organized storage facilitates streamlined data feeding during model training and evaluation.

In the subsequent sections, we'll explore how the transformed data interacts with the model and impacts its performance metrics. By intertwining theoretical underpinnings with empirical results, this comprehensive analysis aims to provide readers with a holistic understanding of our system's inner mechanics.

XAI Implementation

Transparency and interpretability in deep learning models are essential, especially in areas requiring comprehensive understanding of the decision-making process. Our research endeavours to use cutting-edge XAI techniques to render our models both potent and interpretative.

Techniques Applied:

a) SHAP (SHapley Additive exPlanations):

- **Description**: SHAP values, grounded in game theory, enable the elucidation of machine learning model outputs. They bridge optimal credit allocation with localized explanations.
- **Implementation**: In our study, SHAP was utilized to underline the weight of each feature in the predictive process, ensuring every feature's influence on model decisions is transparent.

b) Feature Importance:

- **Description**: A methodology that ranks and pinpoints the most pivotal attributes within a dataset.
- **Implementation**: We harnessed feature importance to zone in on salient data, ensuring that model predictions rest on crucial attributes.

c) Layer-wise Relevance Propagation (LRP):

- **Description**: LRP backtracks through model layers attributing relevance scores to individual attributes.
- **Implementation**: With LRP, relevance scores were accorded to features, offering insight into their influence within each model layer.

d) Local Interpretable Model Agnostic Interpretation (LIME):

- **Description**: Local Interpretable Model-agnostic Explanations (LIME) clarifies complex machine learning models' decisions on a case-by-case basis. It works across different models by creating simple, interpretable models that approximate decisions locally around a specific instance. This approach reveals the influence of individual features on the model's predictions, enhancing understanding of the model's reasoning for particular outcomes.
- **Implementation:** To implement LIME, perturbed samples near the target instance are generated and fed into the complex model to predict outcomes. The model is analysed to identify the impact of features on the prediction, providing insights into the model's local decision-making.

3.7 Conclusion

This chapter outlines a structured research methodology for a study on predictive process mining. The methodology includes careful planning, data preparation, and data processing. The study focuses on data integrity, quality, and exploratory analysis. The research methodology also includes state-of-the-art techniques for process mining implementation and fine-tuning advanced models. The methodology reflects the dedication to rigorous research, innovation, and knowledge quest, setting the stage for future explorations in the domain. The doctoral thesis provides a comprehensive examination of each phase and activities involved.

CHAPTER 4: Case Studies

4.1 Introduction

A structured research methodology for a predictive process mining study is presented in the preceding chapter. Careful planning, data preparation, and data processing are all part of the methodology. The study is centred on exploratory analysis, data quality, and integrity. Modern methods for applying process mining and optimizing complex models are also part of the study methodology. The methodology lays the groundwork for further investigations into the field by demonstrating a commitment to inventiveness, thorough study, and information acquisition. The doctoral thesis offers a thorough analysis of every stage and task involved.

In this chapter, we delve into a series of case studies, each serving as a practical application and validation of the methodologies, models, and theories developed in the preceding chapters of this thesis. These case studies are crucial for testing and refining our research hypotheses and objectives, providing tangible insights into the real-world applicability and efficacy of our proposed solutions. The case studies have been carefully chosen to represent a variety of domains, each with its unique set of challenges and requirements. This diversity ensures a thorough evaluation of our research methodologies and contributes to the robustness and adaptability of our results. The selected domains include the medical field, with a focus on conditions such as sepsis, and extend to other areas, providing a well-rounded perspective on the applicability of our research.

Each case study is approached with meticulous attention to detail, ensuring the integrity and reliability of our findings. The insights gained from these studies are not isolated; they contribute to a holistic understanding of the models and methodologies employed, shedding light on potential improvements, refinements, and new avenues for future research.

In the subsequent sections, we will present each case study in detail, discussing the specific challenges encountered, the methodologies applied, and the insights gained, all while maintaining clarity and conciseness in our exposition.

4.2 Case Study 1: Medical Domain MIMIC-IV ED Dataset

The Healthcare sector is a heterogenous and multi-disciplinary sector, these systems generate quite large medical records on a constant bases, understanding these processes and predicting the sequence, in which activities would occur has been a main stay challenge within the healthcare sector process mining has proven to generate an important analysis view on the process, we can understand the actual working of the process, thereby generating an actual working of the healthcare process (De Oliveira et al., 2019; Kieny et al., 2017; Munoz-Gama et al., 2022b; Rehman et al., 2021) ,. However, the challenge over time has been utilising the information to predict the next sequence of activities (Baier et al., 2020; Weytjens & De Weerdt, 2020).

In the medical domain, chosen specifically for our case studies, the emphasis on reliable and interpretable predictive models is paramount. These models not only play a pivotal role in enhancing patient outcomes and healthcare delivery but are also indispensable for aiding healthcare professionals. With accurate and comprehensible predictions, professionals can make informed decisions, ensuring timely interventions and appropriate treatments, thereby elevating the standards of patient care, and optimizing healthcare efficiency.

4.2.1 An In-depth Exploration

MIMIC-IV-ED is a large, freely available database comprising approximately 425,000 emergency department (ED) stays at the Beth Israel Deaconess Medical Centre in Boston Massachusetts (MA) USA, between 2011 and 2019. It is intended to support diverse research studies and education initiatives in emergency care.

- **Background and Purpose:**

  The ED is a high-demand, resource-limited environment where patients with varying severity levels are assessed and triaged. MIMIC-IV-ED supports data analysis in emergency care by providing a comprehensive database of admissions to an ED at a tertiary academic medical center in Boston, MA.

- **Data Composition:**

MIMIC-IV-ED is composed of a single patient tracking table, edstays, and five data tables: diagnosis, medrecon, pyxis, triage, and vitalsign. Each table provides different aspects of patient information, such as vital signs, medication reconciliation, medication administration, and discharge diagnoses.

subject_id	stay_id	seq_num	icd_code	icd_version	icd_title
10000032	32952584	1	4589	9	HYPOTENSION NOS
10000032	32952584	2	07070	9	UNSPECIFIED VIRAL HEPATITIS C WITHOUT HEPATIC COMA
10000032	32952584	3	V08	9	ASYMPTOMATIC HIV INFECTION
10000032	33258284	1	5728	9	OTH SEQUELA, CHR LIV DIS
10000032	33258284	2	78959	9	OTHER ASCITES
10000032	33258284	3	07070	9	UNSPECIFIED VIRAL HEPATITIS C WITHOUT HEPATIC COMA
10000032	33258284	4	V08	9	ASYMPTOMATIC HIV INFECTION
10000032	35968195	1	5715	9	CIRRHOSIS OF LIVER NOS
10000032	35968195	2	78900	9	ABDOMINAL PAIN UNSPEC SITE
10000032	35968195	3	V08	9	ASYMPTOMATIC HIV INFECTION
10000032	38112554	1	78959	9	OTHER ASCITES

Figure 4-1 Diagnosis Data sample

The diagnosis Figure in 4.1 stands as a testament to the rigorous post-admission evaluation procedures, documenting coded diagnoses as per the International Classification of Diseases (ICD) - either the Ninth or Tenth revision. This figure relevance is underscored by its utility in hospital billing processes. Within this table, the seq_num column offers a gradient of diagnostic relevance, with a lower value typically denoting higher pertinence. Additionally, the icd_code, icd_version, and icd_title columns collectively provide a comprehensive understanding of the diagnosis itself, ranging from its coded representation to its textual description.

subject_id	hadm_id	stay_id	intime	outtime	gender	race	arrival_transport	disposition
10000032	22595853	33258284	2180-05-06 19:17:00	2180-05-06 23:30:00	F	WHITE	AMBULANCE	ADMITTED
10000032	22841357	38112554	2180-06-26 15:54:00	2180-06-26 21:31:00	F	WHITE	AMBULANCE	ADMITTED
10000032	25742920	35968195	2180-08-05 20:58:00	2180-08-06 01:44:00	F	WHITE	AMBULANCE	ADMITTED
10000032	29079034	32952584	2180-07-22 16:24:00	2180-07-23 05:54:00	F	WHITE	AMBULANCE	HOME
10000032	29079034	39399961	2180-07-23 05:54:00	2180-07-23 14:00:00	F	WHITE	AMBULANCE	ADMITTED
10000084	23052089	35203156	2160-11-20 20:36:00	2160-11-21 03:20:00	M	WHITE	WALK IN	ADMITTED
10000084	29888819	36954971	2160-12-27 18:32:00	2160-12-28 16:07:00	M	WHITE	AMBULANCE	HOME
10000108	27250926	36533795	2163-09-27 16:18:00	2163-09-28 09:04:00	M	WHITE	WALK IN	HOME

Figure 4-2 Edstays Data Sample

The edstays table in Figure 4.2 serves as the backbone of the MIMIC-IV-ED database, meticulously recording each unique visit an individual makes to the emergency department. Each row in this table encapsulates a distinct ED visit. Specifically, the subject_id column, a unique identifier, ensures seamless cross-referencing across tables, making it possible to link data with other datasets such as MIMIC-IV and MIMIC-CXR. Furthermore, the hadm_id column is of particular interest as its presence signifies subsequent hospital admissions post the ED visit. The timestamps intime and outtime chronicle the exact duration of the ED stay, revealing patterns about patient flow and ED efficiency.

subject_id	stay_id	charttime	name	gsn	ndc	etc_rn	etccode	etcdescription
10000032	32952584	2180-07-22 17:26:00	albuterol sulfate	028090	21695042308	1	00005970	Asthma/COPD Therapy - Beta 2-Adrenergic Agents, Inhaled, Short Acting
10000032	32952584	2180-07-22 17:26:00	calcium carbonate	001340	10135021101	1	00000733	Minerals and Electrolytes - Calcium Replacement
10000032	32952584	2180-07-22 17:26:00	cholecalciferol (vitamin D3)	065241	37205024678	1	00000670	Vitamins - D Derivatives

Figure 4-3 Medrecon Data Sample

With patient safety at its core, the medrecon figure 4-3 offers a detailed account of a patient's medication history prior to their ED visit, embodying the essence of medication reconciliation. Columns such as name, gsn, and ndc furnish explicit details about each drug, from its textual description to its Generic Sequence Number and National Drug Code. In cases where drugs fall into multiple ontology groups, the etc prefixed columns play a pivotal role, facilitating a holistic understanding of drug categorization.

subject_id	stay_id	charttime	med_rn	name	gsn_rn	gsn
10000032	32952584	2180-07-22 17:59:00	1	Albuterol Inhaler	1	005037
10000032	32952584	2180-07-22 17:59:00	1	Albuterol Inhaler	2	028090
10000032	35968195	2180-08-05 22:29:00	1	Morphine	1	004080
10000032	35968195	2180-08-05 22:55:00	2	Donnatol (Elixir)	1	004773
10000032	35968195	2180-08-05 22:55:00	3	Aluminum-Magnesium Hydrox.-Simet	1	002701
10000032	35968195	2180-08-05 22:55:00	3	Aluminum-Magnesium Hydrox.-Simet	2	002716
10000032	35968195	2180-08-05 22:55:00	4	Ondansetron	1	015869
10000032	35968195	2180-08-05 22:55:00	4	Ondansetron	2	061716
10000032	38112554	2180-06-26 16:58:00	1	Morphine	1	004080
10000032	38112554	2180-06-26 16:59:00	2	Ondansetron	1	015869

Figure 4-4 Pyxis Data sample

The integration of technology into patient care processes is vividly captured by the pyxis figure in 4.4, which logs details related to the BD Pyxis MedStation—an automated medication dispensing system in the ED. Each row in this table not only details the medication dispensed but also offers timestamps via the charttime column, which can be instrumental in assessing medication administration timelines and compliance.

subject_id	stay_id	temperature	heartrate	resprate	o2sat	sbp	dbp	pain	acuity	chiefcomplaint
10000032	32952584	97.8000	87.0000	14.0000	97.0000	71.0000	43.0000	7	2.0000	Hypotension
10000032	33258284	98.4000	70.0000	16.0000	97.0000	106.0000	63.0000	0	3.0000	Abd pain, Abdominal distention
10000032	35968195	99.4000	105.0000	18.0000	96.0000	106.0000	57.0000	10	3.0000	n/v/d, Abd pain
10000032	38112554	98.9000	88.0000	18.0000	97.0000	116.0000	88.0000	10	3.0000	Abdominal distention
10000032	39399961	98.7000	77.0000	16.0000	98.0000	96.0000	50.0000	13	2.0000	Abdominal distention, Abd pain, LETHAGIC
10000084	35203156	97.5000	78.0000	16.0000	100.0000	114.0000	71.0000	0	2.0000	Confusion, Hallucinations
10000084	36954971	98.7000	80.0000	16.0000	95.0000	111.0000	72.0000	0	2.0000	Altered mental status, B Pedal edema

Figure 4-5 Triage Data Sample

At the frontlines of the ED, the triage presented in table 4.5 captures the immediate assessments made as patients present themselves. Columns detailing vital signs—like temperature, heartrate, and sbp—along with the chief complaint field, paint a vivid picture of a patient's health status upon arrival. The acuity column, which designates the severity of the patient's condition, can serve as a vital metric for studies into ED response times and resource allocation.

subject_id	stay_id	charttime	temperature	heartrate	resprate	o2sat	sbp	dbp	rhythm	pain
10000032	32952584	2180-07-22 16:36:00		83.0000	24.0000	97.0000	90	51		0
10000032	32952584	2180-07-22 16:43:00		85.0000	22.0000	98.0000	76	39		0
10000032	32952584	2180-07-22 16:45:00		84.0000	22.0000	97.0000	75	39		0
10000032	32952584	2180-07-22 17:56:00		84.0000	20.0000	99.0000	86	51		

Figure 4-6 Vitalsign Data Sample

The Vitalsign table 4.6 above is an extension of the triage table, the vitalsign table logs periodic vital signs captured during a patient's ED stay. It mirrors many columns from the triage table, but notably includes the rhythm column, detailing the patient's heart rhythm. This table, when analysed in tandem with others, can provide insights into the trajectory of a patient's condition during their stay. It also offers various other facets such as:

- **Integration:**
  The dataset is a module of MIMIC-IV, meaning the information contained within MIMIC-IV-ED is linkable to information in MIMIC-IV. This integration allows for a broader and more integrated analysis, enabling researchers to obtain additional information regarding individuals and compare different aspects of patient care.
- **Clinical Relevance:**
  The extensive and diverse information available in MIMIC-IV-ED is crucial for developing advanced algorithmic approaches aimed at improving the quality of care delivered in the ED. It enables a thorough analysis of patient conditions, treatments, and responses, providing insights that can inform clinical practices and healthcare delivery.

The MIMIC-IV-ED database is a meticulously structured repository capturing diverse facets of patient experiences within an emergency department setting. The various tables provide a lens into patient diagnostic data, their journey through the ED from admission to discharge, their prior medication records, administered medications within the ED, their initial health assessment details, and periodic vital signs throughout their stay. Each table, from Diagnosis

Data to Vitalsign Data, collectively furnishes a holistic portrayal of patient care, ensuring comprehensive research insights for scholars investigating emergency medical care dynamics.

Objective and Significance of Dataset Selection

The MIMIC-IV-ED dataset is pivotal for our research. These datasets not only provide a wealth of information but also represent real-world scenarios where the application of advanced predictive models can have significant impacts. The detailed and diverse data available in these datasets allow for an exhaustive exploration and evaluation of various aspects of medical conditions and healthcare processes. By leveraging the insights derived from these datasets, we aim to enhance the accuracy and interpretability of predictions in the medical domain, contributing to improved clinical decision-making and patient outcomes.

The rationale behind selecting the medical domain, is to explore the realms where the risks are exceptionally high, and the margin for error is minimal. By applying our developed models and methodologies to such a critical area, we aim to demonstrate the practical utility, adaptability, and impact of our research. The insights derived from the analysis of these datasets are intended to contribute to the ongoing discourse on predictive modelling in healthcare, offering potential pathways for advancements in medical diagnostics, interventions, and patient care strategies.

In our research, the application of the constructed event log is crucial, with the overarching aim to apply process mining techniques to discover the inherent structure of processes within the Emergency Department (ED). The essence of this application is to unravel the intricate web of activities and interactions occurring within the ED, providing a structured and coherent view of the operational dynamics inherent in such a critical medical environment.

Process mining will allow us to meticulously discover and visualize the process structure, revealing the sequences and patterns of activities, their interdependencies, and the flow of information and tasks within the ED. This discovered structure will serve as the foundational framework upon which our research will build, enabling a deeper understanding of the operational intricacies of the ED and allowing for the identification of areas for improvement and optimization in emergency care delivery.

The application of our Tab Transformer for predicting activities is not just a technical endeavour but a strategic initiative aimed at enhancing the precision and reliability of predictions in the medical domain. In conditions such as sepsis, where swift and accurate interventions can be the difference between life and death, the ability to predict the next event with high accuracy is invaluable. It ensures that medical practitioners are well-informed and prepared to administer the necessary interventions promptly, thereby improving patient outcomes and optimizing the utilization of medical resources.

Process Modelling

Process modelling is a crucial step in our research methodology, allowing us to visually map out and analyse the flow of activities within the ED. By creating a process model based on the event log, we can:

1. Identify Patterns: Uncover recurring sequences and patterns of activities, revealing the typical flow of events within the ED.
2. Detect Anomalies: Identify deviations and anomalies in the process flow, which can indicate potential areas of concern or improvement.
3. Analyse Relationships: Examine the relationships and dependencies between different activities, providing insights into the interconnectedness of ED processes.

Predictive Analysis

Following the development of the process model, our research will employ advanced predictive analytics techniques to forecast the next event in a sequence (Appice et al., 2019; Tello-Leal et al., 2018; Theis & Darabi, 2019; Weytjens & De Weerdt, 2020). This predictive analysis aims to:

1. Enhance Proactivity: Enable healthcare providers to anticipate subsequent events and activities, allowing for more proactive and informed decision-making.
2. Optimize Resource Allocation: Facilitate better planning and allocation of resources by predicting future demands and requirements.
3. Improve Patient Outcomes: Contribute to improved patient care and outcomes by allowing timely interventions based on predicted events.

Importance for Healthcare Delivery

The application of this methodology has profound implications for healthcare delivery in emergency care settings. It provides a structured and analytical approach to understanding and improving ED processes, contributing to enhanced efficiency, effectiveness, and quality of care. The insights and knowledge gained from this application can inform the development of innovative solutions and strategies for optimizing healthcare delivery and improving patient outcomes in the ED.

4.2.2 Methodology Implementation

The first phase was the application of process mining algorithms to discover the process flow of our event log. We explained the various algorithms and tools in section 2.3.1 of the thesis, we applied the Heuristic and alpha algorithm to visualise the process discovery phase of the process but for further insight and improved visualisation we would be using the disco tool to discover a process map that generates as much detail as possible. Figure 5.2 describes the sepsis data event log providing a visual look at the activities and how they occur.

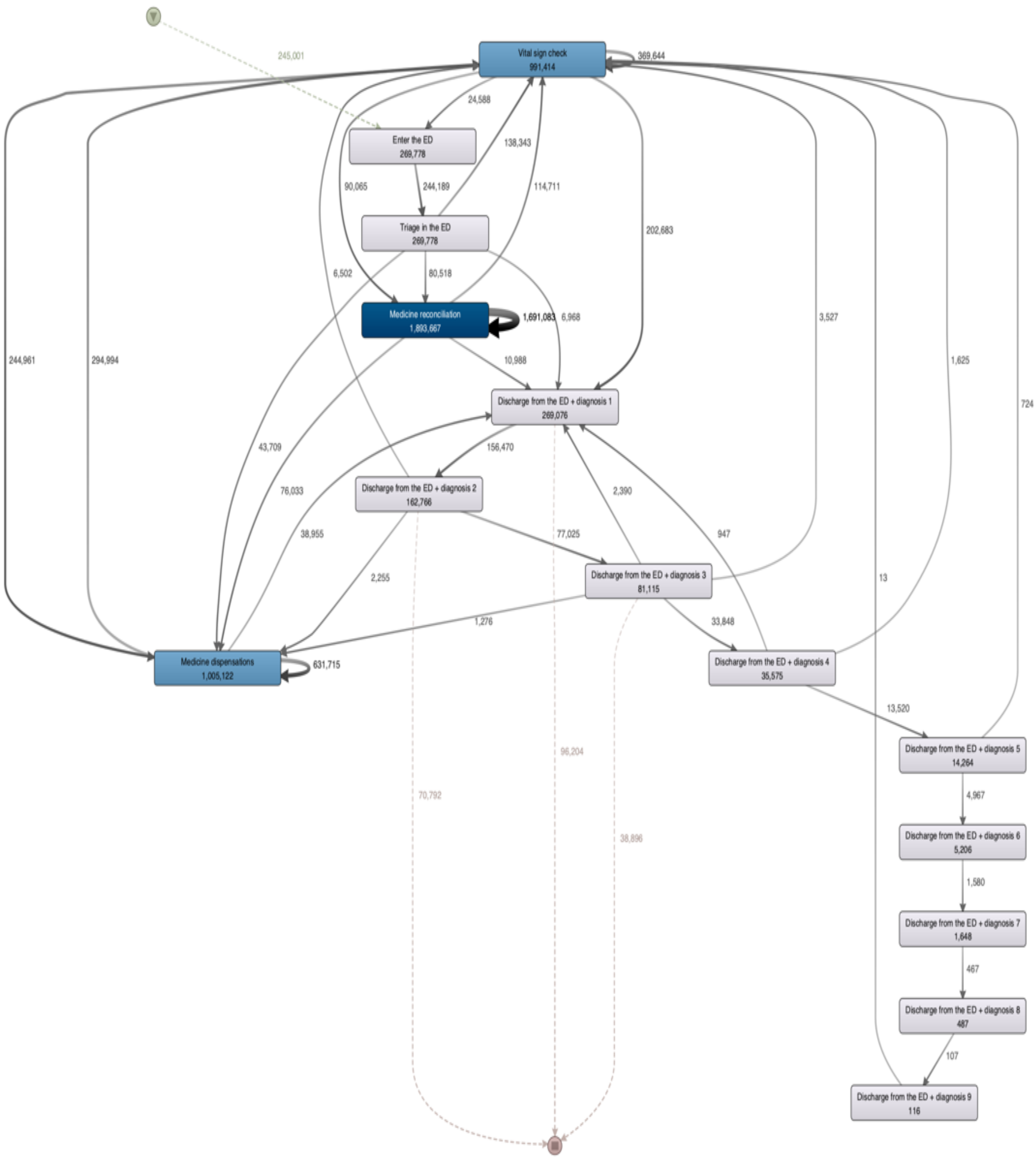

Figure 4-7 MIMIC_IV_ED Process Map Model

Figure 4.7 provides an intricate visualization of the patient's journey through various stages within an Emergency Department (ED). Each node in the diagram represents specific actions or checkpoints, while the directed arrows (edges) denote the transition from one action to another. Alongside these transitions are numerical values, which likely indicate the frequency of occurrences for each specific pathway. This quantitative aspect provides a deeper understanding of common trajectories and interactions within the ED.

Upon entry, as marked by the "Enter the ED" node, most patients are promptly directed to a "Vital sign check." This immediate assessment serves to gauge the patient's current health status, ensuring that the most urgent cases receive timely care. The flow then typically advances to the "Triage in the ED" phase, a critical juncture where patients are systematically evaluated. The purpose of triage is to ascertain the severity of a patient's condition and subsequently prioritize their treatment accordingly. On notable significance in the flowchart are two medication-related nodes: "Medicine reconciliation" and "Medicine dispensations". The "Medicine reconciliation" node emphasizes the importance of verifying the patient's ongoing medications to avoid potential drug interactions or contraindications. In contrast, "Medicine dispensations" pertains to the actual provision of prescribed medications to the patients.

A distinctive feature of the diagram is the multiple "Discharge from the ED" nodes, each correlated with a unique diagnosis. These nodes categorize and account for the varied reasons or conditions treated within the ED. Their presence underscores the diverse nature of cases handled in emergency settings, from the most common to the rare and complex.

Furthermore, the overlapping arrows capture the intricate interdependencies and myriad paths a patient may traverse in the ED. Some routes, indicated by larger numerical values, are more prevalent, while others are less frequently observed. These cyclic patterns, especially around nodes like "Medicine reconciliation", insinuate that certain stages may require revisits, possibly due to iterative checks, repeated procedures, or necessary corrections.

In wrapping up, Figure 5.1 serves as a comprehensive lens into the multifaceted patient interactions and processes within the ED. By delineating the most common to the least frequent pathways, this visualization is an invaluable tool. It not only enhances our understanding of patient flow but also offers insights that can drive optimization efforts in emergency care settings.

Upon visualising our data, in line with our objective, we want to predict the next activity in the process flow. We would be applying our approach primarily on the MIMIC_IV_ED Event log after which, we would apply it on the Sepsis log and describe the application on the other case studies.

Historically, research in the domain of next event prediction has predominantly concentrated on the sequential order of activities, formulating predictions strictly based on these sequences. This approach, while foundational, has been somewhat restrictive, primarily focusing on the linear progression of events and overlooking the multifaceted nature of influencing factors and their interdependencies. Traditional models, relying heavily on attention mechanisms, learn and generate predictions based on the significance and values of sequences, often neglecting the potential impact of other crucial metrics.
In the complex realm of medical processes, every element can be significant in guiding the next step. Our grasp of the importance of each element is still growing. Hence, we aim to combine all pertinent elements, letting the model independently determine the most influential ones. By taking this all-encompassing approach, we aim for a more refined and accurate prediction, enhancing the precision and dependability of predictive process monitoring.

The implementation of the Tab Transformer in our research is a strategic endeavour to elevate the accuracy and reliability of next event predictions in medical settings. This innovative architecture is designed to assimilate a diverse range of features, providing a more holistic and multifaceted view of the influencing factors in medical processes. By enabling the model to learn from a broader spectrum of features, the Tab Transformer can identify intricate patterns and relationships that traditional models might overlook, thereby enhancing the predictive capability of the model.
The Tab Transformer’s ability to integrate and learn from a multitude of features is particularly pivotal in medical domains such as emergency departments, where a myriad of factors converges to influence patient trajectories and outcomes. By leveraging the advanced learning capabilities of the Tab Transformer, our research aims to uncover deeper insights into the complex interplay of factors in medical processes, contributing to the advancement of predictive analytics in healthcare and paving the way for more informed and proactive medical interventions.

At the outset of this research, it was imperative to establish a foundational structure would guide our subsequent experimental approaches. A designated structure as "baseline model" was made. The inception of this model was primarily to facilitate an elementary comprehension of the dataset in hand and to assess the practicability of the methodologies we intended to deploy. Base model in this context meaning our first experiment where we implemented a very simple multi-layer feed-forward network using a defined method called create_baseline_model(), which is basically training and evaluating the baseline model against the tab transformer. One of the differences in the method is that with baseline model, it does not incorporate self-attention mechanisms and positional embeddings like the tab transformer does, so it is like a pre-cursor to see how it performs before we now use Tab Transformer in simplified terms. To elucidate further, the baseline model was architecture as follows:

Model Design and Functionality:

- **Inputs Creation**: The model begins by generating appropriate inputs using the create_model_inputs() function.
- **Features Encoding**: This step involves encoding the input data. The categorical data is encoded using embeddings of specified dimensions (embedding_dims), and these encoded categorical features are stored in encoded_categorical_feature_list. The numerical features are retained in the numerical_feature_list.
- **Features Concatenation**: All the aforementioned encoded categorical features and numerical features are concatenated to form a unified feature set.
- **Feedforward Layers with Skip Connections**: Multiple feedforward layers are integrated, each with skip connections. The create_mlp() function is employed to establish these layers, ensuring data transformation and augmentation. The layers utilize GELU (Gaussian Error Linear Unit) as the activation function and Layer Normalization for normalizing the activations.
- **Multilayer Perceptron (MLP)**: An MLP is constructed, post the feedforward layers. The hidden units of this MLP are computed based on factors (mlp_hidden_units_factors) of the last feature shape. The activation function for this MLP is SELU (Scaled

Exponential Linear Unit), and the normalization is performed using Batch Normalization.

- **Output Layer**: The model concludes with a dense output layer. Given the multi-class nature of our target, a SoftMax activation function is utilized to ensure that the output values lie between 0 and 1 and sum up to 1, representing the probability distribution over target labels.

Within the context of our investigation, the baseline model's significance is not primarily in its sophistication but its capability to shed preliminary light on the dataset's characteristics and inherent challenges. As the study unfolds, there is an intention to iterate, enhance, and possibly expand upon this initial model to meet our investigative objectives. On testing, the baseline model registered an accuracy of 55.6% with the validation dataset. Although this level of accuracy is insightful and establishes an initial benchmark, it emphasizes the imperative for more intricate models that can adeptly encapsulate the multifaceted nature of our dataset.

Model Configuration and Parameters:

The baseline model was configured with the following parameters:

- Learning Rate: 0.0001
- Weight Decay: 0.0001
- Dropout Rate: 0.2
- Batch Size: 128
- Number of Epochs: 20
- Number of Transformer Blocks: 1
- Number of Attention Heads: 5
- Embedding Dimensions: 16 for the categorical features.
- MLP Hidden Units Factors: [1, 1]
- Number of MLP Blocks: 2 in the baseline model.

Development Process

The development of the baseline model involved meticulous configuration and tuning of parameters to ensure optimal performance on the preliminary dataset. During this phase, we encountered several challenges related to data pre-processing and feature selection, which were addressed through iterative refinements and optimizations. The model was compiled with an optimizer utilizing the AdamW algorithm, with a learning rate and weight decay as specified above. The loss was computed using Sparse Categorical Cross entropy, and the model's performance was evaluated based on sparse categorical accuracy.

The comprehensive data, derived from CSV files, was thoughtfully partitioned into training and testing segments. This division was executed using a predetermined method to ensure consistent replicability, anchored by a designated random state. The result of this division was a 75% allocation to the training set and a 25% allocation to the test set. These datasets were then preserved in distinct CSV files, reinforcing data coherence, and facilitating straightforward retrieval. This delineation enabled the model to hone its abilities on a specific data subset while reserving an untapped dataset for unbiased performance evaluation, upholding the rigorous standards of our investigative approach.

The baseline model's development and subsequent performance evaluations provided a comprehensive understanding of our research's initial phase. With an accuracy of 55.6% on the validation dataset, this model's modest results served as an informative benchmark. The derived performance metrics illuminated the inherent challenges in capturing the dataset's multifaceted nature and highlighted the imperative for enhanced modelling techniques.

Conclusion and Transition to Main Model

The foundational knowledge gleaned from the baseline model became the cornerstone for progressing towards the more sophisticated Tab Transformer. The baseline's moderate performance brought to light certain limitations, making it evident that a more nuanced approach was required. This realization reinforced the need to explore complex data relationships further and target higher prediction accuracy. Consequently, our research trajectory shifted towards leveraging the innovative capabilities of the Tab Transformer architecture, an endeavour aimed at achieving a profound understanding of the dataset and elevating the predictive accuracy.

Reflection

In retrospect, the baseline model's role in our research journey was pivotal. More than just an initial attempt, it provided us with valuable insights into the dataset's characteristics and the inherent challenges of our chosen methodology. Serving as the bedrock, this model paved the way for the introduction of the Tab Transformer, facilitating a continuous refinement process. The takeaways from this foundational phase proved indispensable, influencing our subsequent endeavours and directing us closer to the realization of our research goals.

4.2.3 Tab Transformer Implementation

The implementation of the Tab Transformer model is characterized by a meticulous and structured approach, aiming to optimize the model's predictive capabilities. The model's architecture is designed to process and transform both categorical and numerical features effectively, ensuring that each feature contributes to the model's overall predictive accuracy. An in-depth description of our model was described in 3.6.2, we would discuss our result below and hyperparameters tuning steps taken.

Hyperparameter Tuning of Tab Transformer

The optimization of the Tab Transformer model's hyperparameters is essential for achieving the best performance on the task at hand. Given the intricacies of the model's architecture and the numerous potential hyperparameters, conventional tuning methods can be resource intensive. To address this challenge, we adopted a systematic methodology.

Methodology

Instead of exploring each hyperparameter combination individually, a predefined set was established. This set considers crucial hyperparameters such as learning rates, dropout rates, batch sizes, and other architecture-specific parameters. A programmatic approach was taken to generate all possible combinations, ensuring a comprehensive exploration across the diverse hyperparameter space.

Hyperparameter Combinations Explored

The following table represents the various hyperparameter combinations that were tested during the tuning process:

Epochs	Learning Rate	Batch Size	Weight Decay	Dropout Rate	Transformer Blocks	Attention Heads	Embedding Dims
5	0.00001	265	0.001	0.5	4	5	128
5	0.0001	128	0.0001	0.4	4	16	128
5	0.00002	32	0.0001	0.4	4	8	64
5	**0.00002**	**32**	**0.001**	**0.3**	**4**	**8**	**64**
5	0.001	265	0.0001	0.2	4	4	16

Table 4-1 Hyperparameter Tuning

Upon exploring the various hyperparameter combinations, the model demonstrated varied performances. The objective was to discover a combination that would yield the highest accuracy and ensure the model's generalization to unseen data. The systematic approach to hyperparameter tuning played a pivotal role in achieving a significant improvement in predictive accuracy. The chosen Hyperparameter is highlighted in Bold on the table above.

Comparative Analysis with Base Model

The Tab Transformer model demonstrated a substantial improvement in predictive accuracy, achieving 69%, compared to the base model's 55.6%. This improvement is indicative of the Tab Transformer model's advanced capability to process and learn from the features effectively.

a) **Enhanced Feature Processing:**

   Unlike the base model, the Tab Transformer model is adept at handling both categorical and numerical features, ensuring that each feature is adequately processed and contributes to the model's predictive accuracy.

b) **Incorporation of Diverse Metrics:**

The model's ability to incorporate a variety of features such as gender, race, mode of arrival, and medicine dispensations, in addition to the sequence of activities, has been pivotal in its improved performance. This incorporation allows the model to have a more holistic understanding of the data, enabling it to make more informed and accurate predictions.

c) **Advanced Learning Capabilities:**

The Tab Transformer model excels in learning complex representations from the data due to its structured flow and advanced architecture, including multi-head attention and feedforward layers within the transformer blocks. This advanced learning capability is a significant advancement over the base model, allowing the Tab Transformer to capture intricate relationships and dependencies between different features effectively.

Implications

The successful implementation of the Tab Transformer model has substantiated its efficacy in predictive process monitoring, showcasing its potential in various applications. The structured and advanced approach adopted in its implementation has enabled the model to achieve superior predictive accuracy, making it an asset in research and development in this domain.

The comparative improvement over the base model underscores the advancements made in feature processing, learning capabilities, and the incorporation of diverse metrics, paving the way for further innovations and research in predictive modelling. The detailed implementation and the resultant improvements of the Tab Transformer model serve as a significant contribution to the field, providing a foundation for future research endeavours in advanced predictive modelling techniques.

4.2.4 Implementation of the Explainable Artificial Intelligence (XAI) Phase

After the successful implementation and evaluation of the Tab Transformer model, the next crucial phase in our research is the deployment of Explainable Artificial Intelligence (XAI) methodologies. This phase is pivotal as it aims to decipher the intricate workings of the model, providing insights into how and why specific predictions are made. The incorporation of XAI is essential to validate the reliability and trustworthiness of the model, ensuring its decisions are understandable, transparent, and justifiable.

Objectives of the XAI Phase

a) Model Interpretability:

The primary objective is to unravel the internal mechanisms of the Tab Transformer model, making its operations and decision-making processes interpretable to humans. This interpretability is crucial for validating the model’s predictions and understanding the significance of each feature in the decision-making process.

b) Feature Importance Analysis:

Through XAI, we aim to analyse and rank the importance of different features used by the model. Understanding which features are pivotal in making predictions allows for model refinement and provides insights into the underlying processes being modelled, ensuring more informed and accurate future predictions.

c) Trust and Validation:

By making the model's workings transparent and understandable, we seek to build trust in the model's predictions and validate its reliability and robustness. This trust is essential for the practical deployment of the model in real-world scenarios, where the stakes of each prediction can be significant.

d) Enhanced Model Refinement:

Insights gained from understanding the model's internal workings will facilitate further refinement and optimization of the model. This enhanced refinement is aimed at improving the model's accuracy and reliability, ensuring its robustness in varied and dynamic environments.

Methodological Approach

In this phase, we will employ a range of XAI techniques and tools designed to elucidate the internal structures and operations of the model. These methodologies will focus on visualizing the model's decision pathways, analysing feature contributions, and exploring the model's response to different input variations. By leveraging these techniques, we aim to derive meaningful insights into the model's functioning, facilitating its continuous improvement and optimization (El-Khawaga *et al.,* 2022).

Shap Implementation

The SHAP implementation, provides insight into the impact on certain values for certain prediction. The use of SHAP (SHapley Additive exPlanations) provides a deeper level of insight into the decision-making process of the model, allowing us to understand which features of the data are most influential in making predictions. This feature importance information can be used to improve the model and make it more accurate and robust. SHAP is a powerful tool that has been shown to be effective in a variety of machine learning applications, and we believe that its use in this work will provide valuable insights into the model's decision-making process. By visualizing the SHAP values through force plots, we can observe the positive and negative contributions of each feature to the model's output, compared to the expected value.

Figure 4-8 Force plot predictive instance of test case on MIMIC_IV_ED Prediction Model

In the context of machine learning, predictions in the logit space, especially with classification tasks, are not bounded between 0 and 1 as shown in Figure 5.3. Instead, they can range from negative to positive infinity. A logit value of zero corresponds to a probability of 0.5, with positive logits indicating probabilities greater than 0.5 and negative logits indicating probabilities less than 0.5. This is how the Figures are represented from the shap implementation.

The central point of focus in the force plot is the model's prediction value f(x) that stands at 0.91. This isn't a probability but a raw output indicating the likelihood of the next event's occurrence on a logit scale. For the given instance, this value signifies the model's confidence regarding the upcoming event based on the presented features.

Positioned at roughly 0.9503, the 'base value' offers a reference point, illustrating the model's average prediction for the next event across all instances or a specific benchmark group. The departure of the f(x) value from this base indicates how the specific data point in question varies from the average behaviour.

Feature Contributions and Their Dynamics

Two distinct color codes encapsulate feature contributions:

- Red features, like "zithromyc 500mg/250mL 250mL BAG", in Figure 4.8 are positive influencers, suggesting they increase the likelihood of the predicted next event. The magnitude of influence resonates with the bar's width.

- In contrast, blue features, such as "event_concept_name = Medicine_dispensations", decrease this likelihood, pulling the prediction value toward a direction opposing the event's occurrence.

It's essential to highlight features with NaN values. In a next event prediction context, these absent features can still be influential. Their non-presence might indicate a deviation from the usual pattern, which the model deems significant for the prediction.

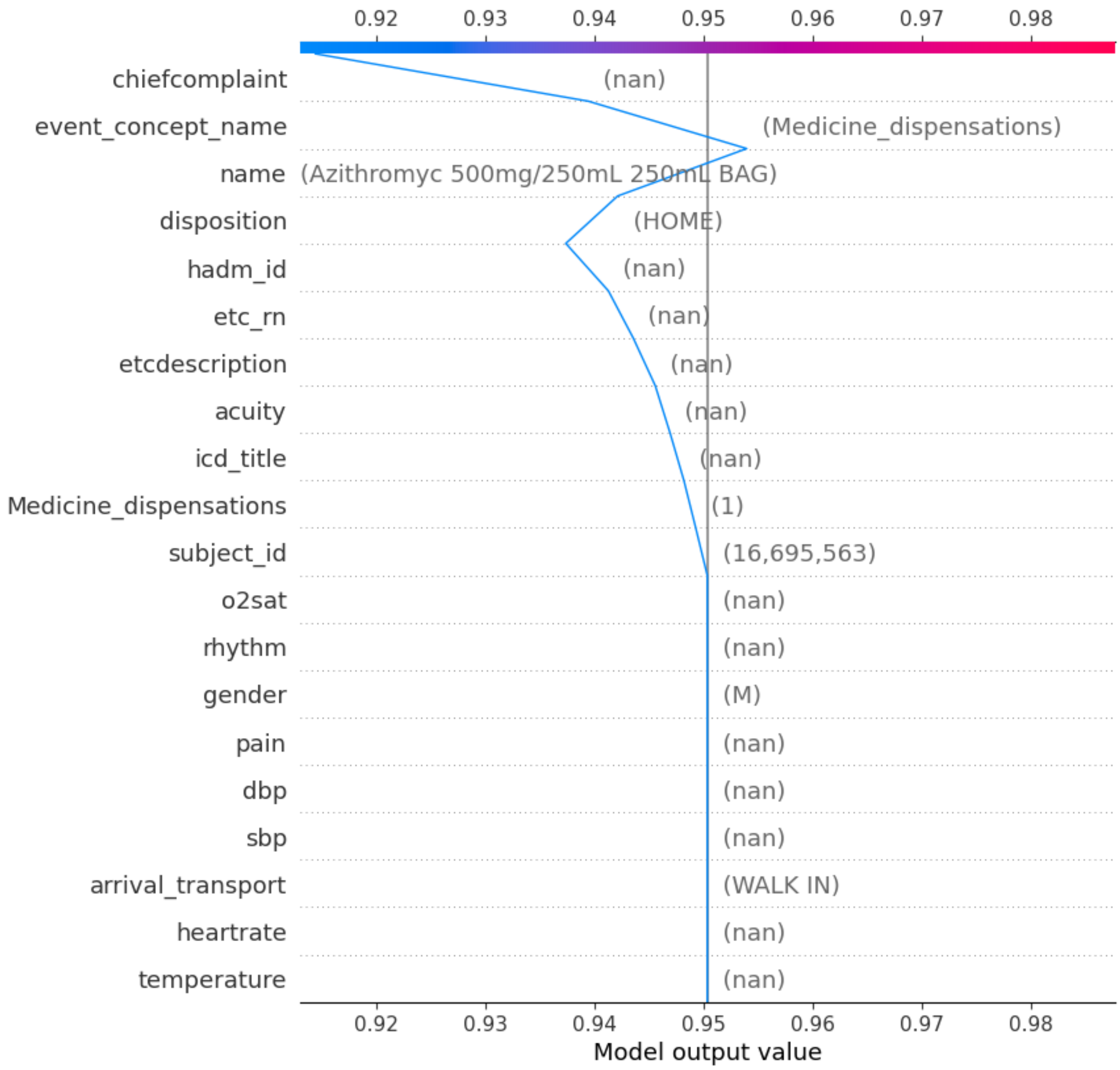

Figure 4-9 Decision plot predictive instance of test case on MIMIC_IV_ED

As discussed on the feedback on Figure 4.8, Figure 4.9 is the decision plot of the same event and illustrates a decision plot of the same prediction and the features influencing the prediction are provided below.

Feature Influence in Logit Space:

- chiefcomplaint: Lacking a defined value (nan), the absence or non-specificity of a chief complaint appears not to notably sway the prediction, suggesting its neutrality in this instance.
- event_concept_name: The value "Medicine_dispensations" positively influences the prediction, pushing the logit score slightly higher than the base value. This indicates

that the event concept of dispensing medicine might increase the likelihood of the predicted event.

- name (Azithromyc 500mg/250mL 250mL BAG): The presence of this specific medication and its dosage format seems to considerably elevate the logit value, emphasizing its role in determining the subsequent event.
- disposition (HOME): Indicating the patient's intended location after the event, its value of 'HOME' mildly propels the prediction upwards, suggesting a potential relationship between the patient's post-event disposition and the predicted next event.

The other features like hadm_id, etc_rm, and so forth, with their respective (nan) values, suggest that they either lack specific data for this instance or their absence/neutrality doesn't substantially deviate the prediction from the base value.

While SHAP values provide an in-depth understanding, it's paramount to ensure that the interpretations align with real-world scenarios. Introducing or emphasizing features that might not naturally occur can skew interpretations. In next event prediction, the goal is to glean insights from features that genuinely matter and can naturally exist in the dataset.

LRP Implementation

The heatmap in Figure 5.5 provides a visual representation of the relevance of each feature in the specified layer, allowing us to discern which features are more influential in the model's predictions. The red features, 'icd_title', 'pain', and 'name’ which is the drug name in this case Azithromyc 500mg/250mL 250mL BAG, are significant, contributing positively to the model's output, while the blue features have less impact.

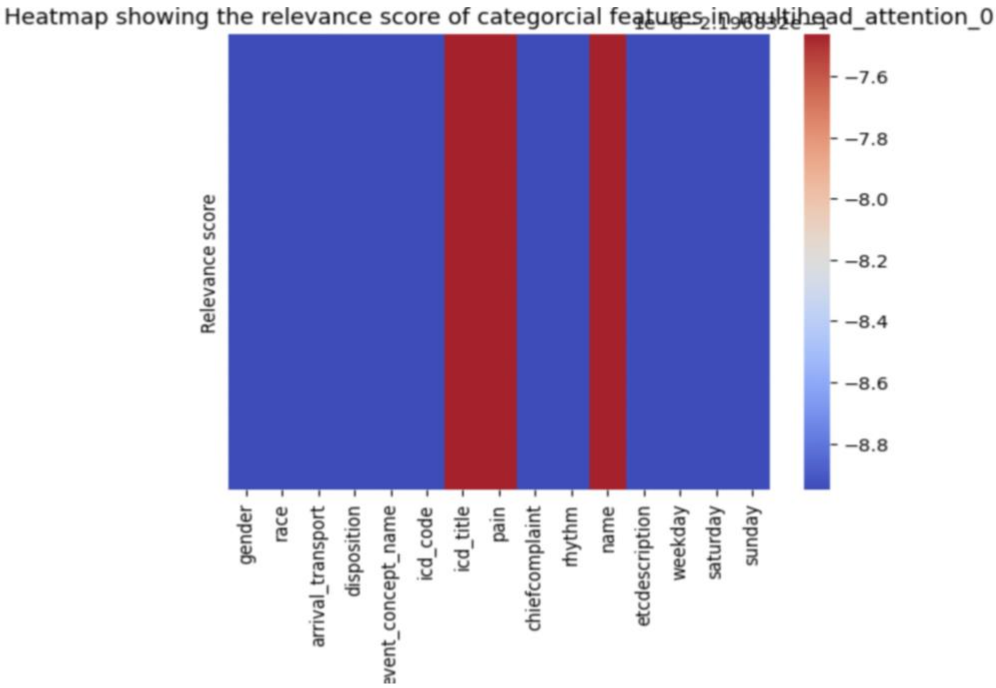

Figure 4-10 LRP heatmap plot of categorical features

This detailed insight into feature relevance is crucial for understanding the model's decision-making process, enabling further refinement and optimization of the model by focusing on the most influential features. It also aids in validating the model's predictions and ensuring its reliability and robustness in diverse scenarios.

Lime Implementation

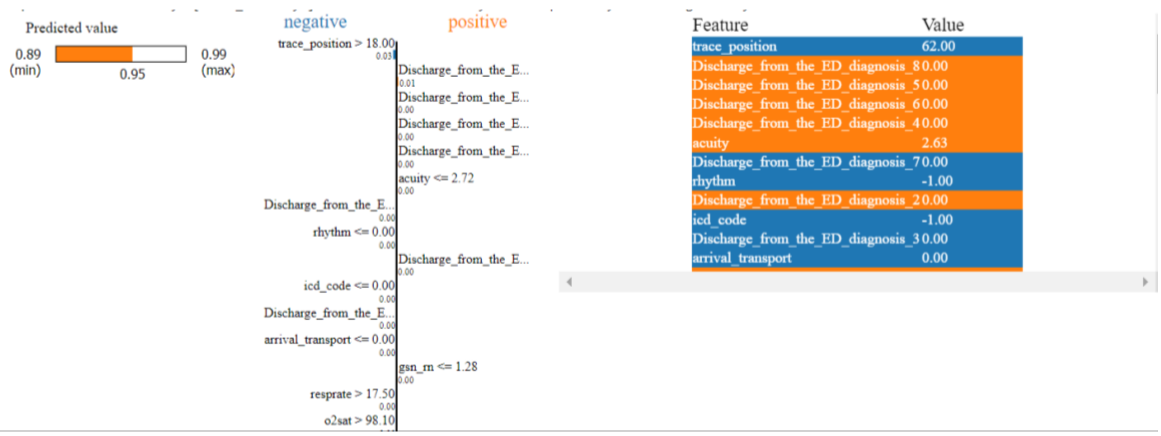

Figure 4-11 MIMIC_IV_ED LIME Implementation

In our study, we employed the LIME framework to interpret the predictive model's decision-making process for individual predictions. This approach provides local explanations, illuminating how the model's output is influenced by the input features for specific instances.

As depicted in the figure 4-11, the prediction was categorized as 'positive' with a probability of 0.95. The LIME output demonstrates the features contributing to this classification, along with their respective weights, which are indicative of the feature's impact on the model's prediction.

The features listed are shown alongside their corresponding weights in a descending order, indicating their relative importance. A positive weight suggests that the feature contributes toward the 'positive' class, while a negative weight suggests a push towards the 'negative' class. For instance, 'Discharge_flow_from_0-30_degrees' with a value of -8.00 has the most significant negative impact on the prediction, pulling the instance towards the 'negative' class. Conversely, 'Discharge_flow_from_30-60_degrees' with a value of 6.00 is the strongest positive contributor, pushing the prediction towards the 'positive' class.

The color-coding further aids in visual interpretation: orange represents positive influence, while blue indicates negative influence. The shade intensity correlates with the feature's weight magnitude, allowing for a quick assessment of each feature's impact.

In conclusion, the LIME analysis facilitates an understanding of the model's behaviour in the context of individual predictions. By doing so, it aids in validating the model's trustworthiness and provides insights into the model's operational mechanics, which is paramount for complex models in critical application areas.

4.2.5 Case study review

In this case study, advanced machine learning techniques were applied to medical data to develop a model capable of making accurate predictions. The model was built using a Tab Transformer architecture, and its interpretability was enhanced using SHAP values and Layer-wise Relevance Propagation (LRP). In our exploration of next event predictions in the healthcare environment, various machine learning models were trained and evaluated using the Mimic_iv_ED dataset. The performance metrics for each model, segmented by test and train datasets, are tabulated below:

Mimic_iv_ED	**Test**	**Train**
Random Forest	0.65	0.634
XGBoost	0.656	0.641
TAB TRANSFORMER	0.69	0.68

Table 4-2 Mimic_IV_ED Accuracy of observed algorithms

The Random Forest model yielded an accuracy of 0.65 in the test dataset and 0.634 in the train dataset. Although this model offers inherent advantages like feature importance and ease of visualization, its accuracy was surpassed by both XGBoost and the TAB TRANSFORMER. XGBoost, a gradient boosting algorithm, slightly improved upon the Random Forest's performance, achieving a test accuracy of 0.656 and train accuracy of 0.641. This model's ability to optimize on the fly and handle missing data potentially contributed to its marginally superior performance.

However, the most commendable performance came from the TAB TRANSFORMER, which achieved an accuracy of 0.69 on the test dataset, markedly surpassing the base model's accuracy of 0.556. Its slightly reduced gap between training and testing accuracy, compared to XGBoost, hints at better generalization to unseen data.

The interpretability techniques employed provided profound insights into the underlying features influencing the model's predictions. Utilization of SHAP force plots and LRP heatmaps visualized the impact of each feature, shedding light on the contributions of individual factors in the model's decision-making process.

In one of the instances analysed using SHAP, the logit space activity value of 0.91, within a frame between 0.89 and 0.97. The SHAP force plot highlighted the significance of features

like 'Disposition' and 'Name', contributing positively to the model's value of 0.91, while features with NaN values didn’t contribute towards the positive prediction of the model.
The task of predicting the next event enables medical professionals to anticipate necessary interventions, hence optimizing the flow of patient care and mitigating potential risks. For instance, in one of our cases, our predictive model accurately forecasted that a "Vital Sign Check" would be the subsequent event, aligning perfectly with the actual event that took place. In another intriguing scenario, the model designed exhibited its capability to forecast a complex event, specifically predicting the activity "b'Discharge_from_the_ED_diagnosis_2". While this was the true next event, it's noteworthy to mention that the model provided probabilities for a myriad of potential subsequent events, ranging from various diagnoses upon discharge from the Emergency Department to tasks like "Medicine dispensations" and "Triage in the ED". This diverse set of predictions, ordered by their likelihood, equips healthcare providers with a comprehensive outlook, allowing them to better prepare for possible eventualities. By leveraging next event prediction in such a manner, healthcare institutions can usher in a new era of patient-centric care, characterized by efficiency, foresight, and adaptability Implications:

1) **Enhanced Decision-Making:**
   The insights gained from the model provide medical professionals with a nuanced understanding of the factors influencing the model's predictions, aiding in more informed and accurate decision-making. The highlighted features such as 'icd_title', 'pain', and 'name' in LRP analysis can guide medical practitioners to focus on specific areas for diagnosis and treatment plans.
2) **Improved Patient Outcomes:**
   The model's high accuracy and the interpretability of its predictions can lead to more precise diagnoses and treatment plans, potentially improving patient outcomes. Early and accurate identification of medical conditions can facilitate timely intervention, reducing the risk of complications and improving the quality of life for patients.
3) **Enhanced Medical Research:**
   In the analysis of next event predictions within the healthcare domain, the model's results were particularly illuminating. The prediction for the "Vital Sign Check" event exhibited perfect alignment with the real event, showcasing the model's capability to

predict routine or commonly occurring tasks in the healthcare environment. Such accurate predictions for standard tasks can significantly streamline processes, ensuring that necessary resources, such as medical personnel and equipment, are readily available.

However, the prediction for the "b'Discharge_from_the_ED_diagnosis_2" event presented a more intricate scenario. While the model did successfully predict the event, the list of potential events showcased diverse possibilities with varying probabilities. For instance, "Medicine dispensations" and "Triage in the ED" had high probabilities, suggesting these events are common post-discharge or in conjunction with discharge procedures. This presents an opportunity for healthcare institutions to consider optimizing these processes, perhaps by ensuring that medicine dispensation is expedited after certain diagnoses or that triage processes are refined for quick and efficient patient care.

Interestingly, among the predictions for the discharge diagnoses, there seems to be a near even distribution of probabilities across various diagnoses (from "b'Discharge_from_the_ED_diagnosis_5" to "b'Discharge_from_the_ED_diagnosis_8"). This indicates the inherent uncertainty in predicting the exact diagnosis upon discharge, given the numerous potential outcomes. Hospitals could use this insight to ensure that they're prepared for a wide range of patient needs upon discharge, rather than focusing resources on a singular predicted outcome.

The events with the highest probabilities, "Vital_sign_check" and "Medicine_dispensations", suggest that these are critical and frequent interventions. Therefore, hospitals might consider reallocating resources or introducing technological solutions, like automated vital sign monitors or efficient medicine dispensation systems, to cater to these frequent events more seamlessly.

A noteworthy observation is the presence of "Enter_the_ED" with a relatively high probability, suggesting that the model perceives a significant number of patients re-entering the Emergency Department after certain events. This could be indicative of recurring patient issues, potentially due to premature discharges or complications.

This insight is crucial for hospitals to evaluate the quality of patient care and introduce measures to minimize return visits.

In terms of improvement, the model could benefit from more granular data, especially regarding patient history, specific symptoms, and the duration between events. This would allow for more precise predictions. Moreover, integrating feedback loops where medical professionals validate or correct the model's predictions in real-time would be invaluable for iterative model refinement.

4) **Ethical and Responsible AI:**

The interpretability of the model ensures that the predictions are transparent and understandable, adhering to the principles of ethical and responsible AI. This transparency is crucial in medical settings where understanding the rationale behind predictions is essential for trust and accountability

4.3 Case Study 2: Financial Domain (Loan Application Dataset)

In transitioning from the medical domain to the financial domain, Case Study 2 sharpens its focus on loan application datasets, specifically BPIC 2012 and BPIC 2017, to further explore the applicability and efficacy of the Tab Transformer in predictive process mining. The choice of these datasets is strategic, allowing for a meticulous comparison of our results with existing works and offering a diversified perspective on how our model performs under varied circumstances. The BPIC datasets are renowned for their extensive and varied data points related to loan applications, providing a rich ground for detailed exploration and analysis in the financial sector. This sector, with its intricate and multifaceted processes, presents an ideal environment to investigate our first research question: "How does the Tab Transformer perform in predictive process mining for datasets with multiple categorical features, in comparison to existing methods like Random Forest and XGBoost?" By addressing this question, it was aimed to uncover the nuances of handling multiple categorical features in financial datasets and evaluate the relative strengths and limitations of the Tab Transformer in this specific context.

The integration of our research questions within this case study is not merely procedural but is intricately woven into every aspect of our analysis. We delve deep into the financial domain with a clear objective to understand the key determinants affecting loan approvals and to optimize loan application processes, thereby addressing our second research question: "How effective are various XAI techniques, when applied to the Tab Transformer in predictive process mining, at enhancing model interpretability and providing meaningful insights?" The insights derived from this study are not isolated observations but are reflections of the intricate interplay between various features and determinants in the financial sector. By focusing on the financial domain, this case study aims to provide substantial and actionable insights, contributing significantly to financial decision-making and risk management, and offering a refined understanding of the implications of our findings in the realm of finance.

4.3.1 An In-depth Exploration

We would describe both the BPIC 2012 and BPIC 2017 Data Below.

BPI Challenge 2012 (BPIC 2012)

The 2012 dataset (Boudewijn van Dongen, 2012), sourced from a global financial institution, centres around the intricate loan application process of the bank. Each application captured within this dataset passes through a series of states, leading up to a conclusive decision. The dataset comprehensively documents 262,200 events across 13,087 cases, with a diversity of 24 unique event classes. These events encompass various attributes, ranging from timestamps to lifecycle transitions and resource identifiers.

Delving into the process flow, the journey commences with an applicant initiating the loan application online. For cases not receiving an auto-approval, the application advances to the institution's personnel for further verification. During this phase, applicants might be prompted to revise or furnish supplementary details. As the application meanders through its lifecycle, it undergoes numerous state transitions, often changing hands between different departments or responsible individuals. Ultimately, every application converges to one of two resolutions: a "declined" status or an "approved" one.

Historically, this dataset has proven invaluable for academic and industrial research. Scholars have leveraged it to unearth underlying process models intrinsic to the loan application. Furthermore, it's been a pivotal resource to spotlight potential bottlenecks, inefficiencies, and to dissect the determinants steering an application to its final verdict.

BPI Challenge 2017 (BPIC 2017)

Offering a lens into the domain of personal loans and overdrafts, the 2017 dataset (Boudewijn van Dongen, 2017) emerges from an anonymized banking institution. This dataset paints a detailed picture of the application procedures, from the embryonic stages of initiating an application to its final outcome. With a vast repository of 1,202,267 events distributed over 31,509 cases, the dataset boasts 26 distinct event classes. Each event, encapsulated within this dataset, is characterized by a myriad of attributes, including event-specific details, timestamps, lifecycle transitions, and the associated resource data.

Shedding light on the process trajectory, it kicks off when potential customers embark on their loan application journey online. Post initiation, an associated offer springs into existence, setting the stage for negotiations. As customers and the bank endeavour to finalize the offer's terms, the application is subjected to a battery of validations, checks, and verifications. Navigating through this maze of procedures, every application culminates in one of three terminal states: "cancelled", "denied", or "accepted".

The 2017 dataset, much like its predecessor, has been a cornerstone for process mining research. Researchers have wielded it to discover, juxtapose, and dissect diverse process variants. A focal point has been to trace pathways culminating in successful loan applications. In parallel, the dataset has empowered studies to pinpoint bottlenecks, inefficiencies, and instances necessitating rework or redundant validation.

4.3.2 Methodology Implementation

As we described in the case study of medical data, we discover the process flow of our event log which is presented below. The figure below describes BPIC 2017 and BPIC 2012 Loan application process flow diagram. The figure3.7 below shows the BPIC 2017 process flow diagram of how the loan application process is carried out, there are 31,509 cases. Each following a set of subsequent or multiple activities before a decision is made.

In the provided BPIC 2017 process model Figure 5.8 below, is a representation of the multifaceted application journey within a specific operational system is depicted. Each rectangular node symbolizes a distinct event or activity, illustrating pivotal stages of the application process. Notably, the inception of this journey is marked by the "A_Create Application" event, with a recorded initiation count of 31,509, emphasizing its foundational role in the process.

Sequentially, arrows delineate the progression and transition from one event to its subsequent stage. These arrows, inscribed with numerical values, encapsulate the frequency of each specific transition. For instance, a prominent transition, witnessed 31,382 times, ushers applications from their creation to the "A_Submitted" stage, signifying the submission of these applications.

Post-submission, the diagram illuminates several potential trajectories for the application. A segment of these applications seamlessly transitions to an "A_Accepted" status, while others navigate through an intermediary "W_Complete application" phase, accentuating the existence of conditional checkpoints or additional information requisites. In contrast, certain applications are directed towards a "A_Declined" or "A_Cancelled" state, suggesting varied outcomes based on predefined criteria or external factors.

Further intricacies of the process are showcased through various sub-processes and tasks. The "W_Handle leads" task, for instance, might allude to a lead management or nurturing phase. Meanwhile, the "W_Validate application" activity, underscored by its loopback arrow, highlights potential iterative validation steps, hinting at meticulous verifications or possible amendments necessitated by discrepancies.

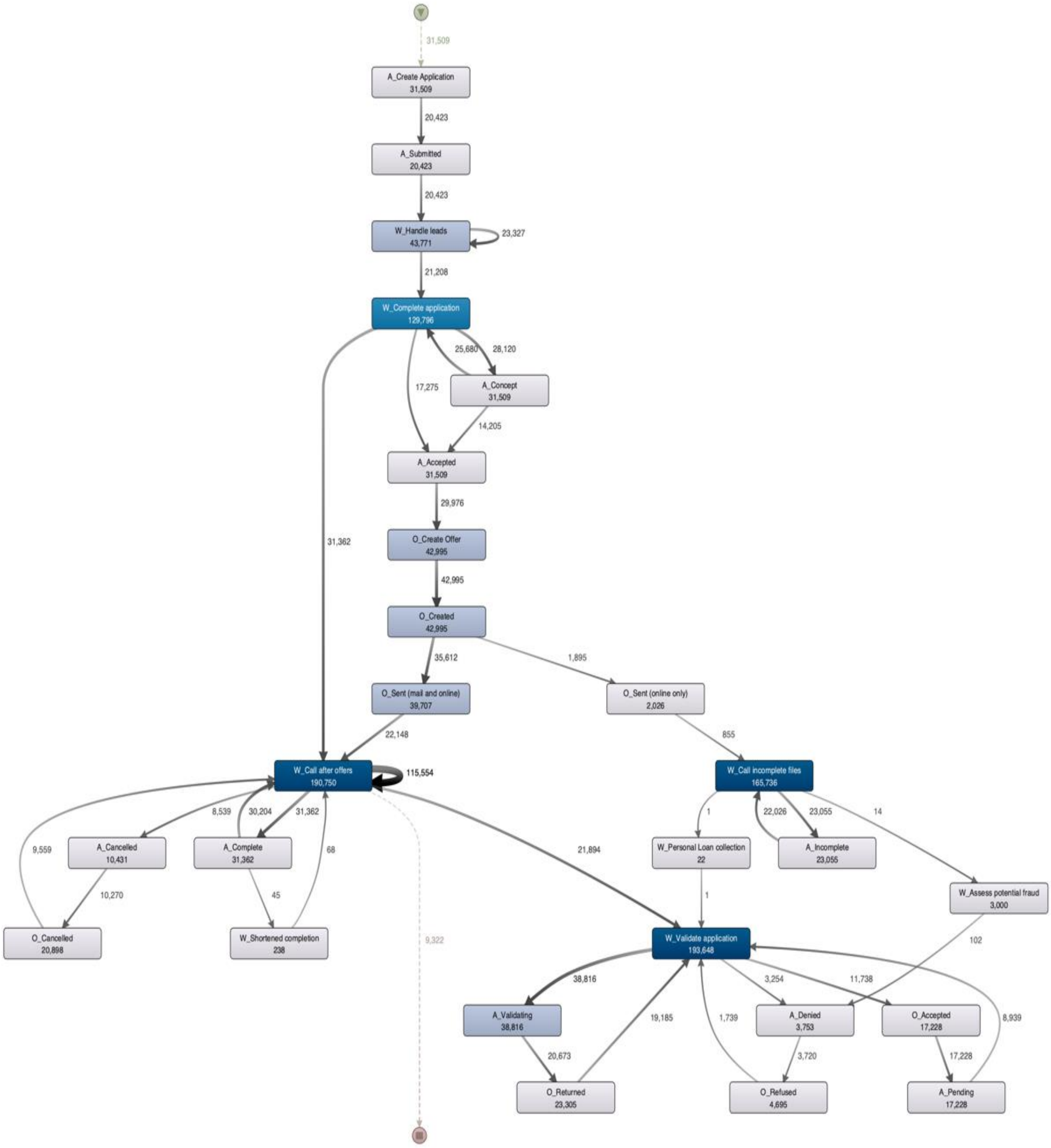

Figure 4-12 BPIC 2017 process model Diagram

A particularly salient flow emerges between the "W_Call after offers" and "W_Incident late fee" events. The bold delineation of this transition underscores its significance, possibly indicating a prevalent or crucial pathway, where subsequent to an offer's issuance and a reminder call, incidents associated with late fees are recurrently observed.

On the spectrum's other end, less frequented paths, such as the one connecting "W_Personal Loan collection" to "O_Refused", shed light on rarer transitions or exceptional scenarios, meriting further exploration for optimization opportunities or anomaly detection.

Each of these activities represents a distinct step or phase in a process and understanding them is crucial for comprehending the overall workflow or system being studied in the thesis. The clear delineation and description of each activity facilitate a deeper insight into the interactions, dependencies, and outcomes within the studied processes. Below we would describe the BPIC 2012 process flow model.

The BPIC 2012 process flow model presented in Figure 4.13 provides a comprehensive visualization of the various stages and transitions within an operational workflow, characteristic of the BPIC 2012 dataset. The rectangular nodes, representing distinct events or activities, effectively map out the complex journey of applications or requests within the system.

Initiation of this process is demarcated by the "O_CREATED-COMPLETE" activity, logging 7,300 instances. This foundational activity likely represents the genesis of a request or application within the system, paving the way for subsequent activities.

Transitioning forward, arrows guide the path from one activity to the next, with numeric annotations denoting the frequency of each specific transition. For instance, from the initial creation, there's a notable progression of 7,300 cases to the "O_SENT-COMPLETE" stage, hinting at a submission or dispatch process of the initial requests.

Subsequently, the journey branches out into multiple trajectories. Some requests progress to the "O_SELECTED-COMPLETE" stage, while others traverse to the "A_ACCEPTED-COMPLETE" state, indicating potential approval or selection criteria met by these applications. Noteworthy is the dual transition from "A_ACCEPTED-COMPLETE" to either "W_Completeren aanvraag-START" or directly to "A_FINALIZED-COMPLETE", suggesting alternate routes based on certain conditions or requirements.

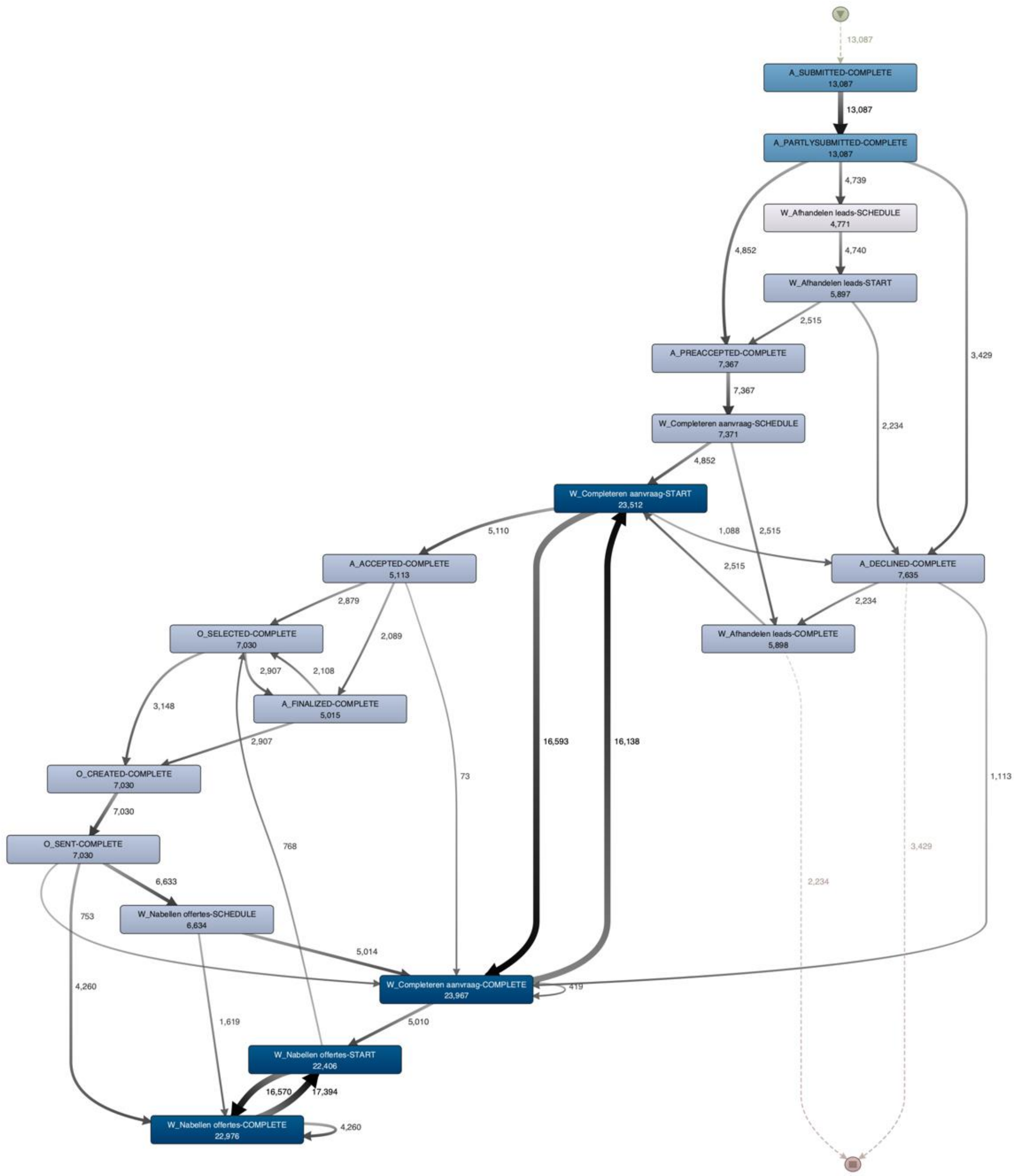

Figure 4-13 BPIC 2012 Process Model Diagram

In the upper quadrant, the "A_SUBMITTED-COMPLETE" activity stands out with a substantial count of 13,087 instances. Following this, there's an intriguing iterative loop involving the "W_Afhandelen leads-START" and "W_Afhandelen leads-SCHEDULE" activities, suggesting possible iterative tasks or reviews associated with handling leads.

Additionally, a significant pathway emerges from the "A_PREACCEPTED-COMPLETE" stage, bifurcating into the "W_Completeren aanvraag-SCHEDULE" and "W_Completeren aanvraag-START" events. The dense arrow lines underlining these transitions emphasize their prevalence, potentially underscoring a mandatory or frequent process phase.

On the diagram's periphery, activities such as "A_DECLINED-COMPLETE" and the interactions involving "W_Nabellen offertes" series underscore specific outcomes or follow-up processes. The latter, characterized by a sequence of "START", "SCHEDULE", and "COMPLETE" stages, sheds light on a meticulous follow-up or feedback mechanism inherent to the system.

4.3.3 Tab Transformer Implementation

The focal point of our research is to predict subsequent activities in the process diagram, leveraging the trace positions of the records. We aim to discern the patterns and correlations within the sequences to anticipate the forthcoming steps in the application journey. Subsequently, we intend to elucidate the factors influencing these predictions, employing the explainable approach delineated in Case Study 1. This endeavour not only enhances our understanding of the procedural dynamics but also sheds light on the underlying factors steering the process trajectories, providing a comprehensive insight into the loan application ecosystem.
We utilised the same model architecture and configuration for all case studies as described in case study 1.

In the data pre-processing phase, essential transformations were applied to the raw event log data to augment its structure and extract salient features, thereby facilitating subsequent analyses. Initially, timestamps present in the 'time: timestamp' column were standardized, ensuring a consistent datetime format throughout the dataset. Post-conversion, the dataset was chronologically ordered based on these timestamps, reflecting the genuine sequence of events. A pivotal addition was the 'tracePosition' attribute, capturing the ordinal position of each event within its corresponding trace. This was achieved by iterating through each unique trace, sequentially assigning position values to its constituent events.

To capture temporal patterns potentially indicative of variations in process behaviors, day-specific features were introduced. The 'dayOfWeek' column identified the particular day of the week an event occurred, while subsequent binary columns – 'weekday', 'saturday', and 'sunday' – distinctly flagged weekdays and weekends. The categorical 'event:concept:name' column, representing diverse event activities, underwent one-hot encoding, culminating in a set of dummy variables. These widened the feature space, enabling distinct representation of each activity type.

Subsequent encoding leveraged the Label Encoder, serving a dual purpose: firstly, to numerically represent the event names, and secondly, to provide a combined encoding based on event names and their dispositions. This nuanced encoding strategy was pivotal in capturing the intricacies of event characteristics. A significant pre-processing step entailed the calculation of target variables, especially concerning predicting subsequent activities. By iterating through the ordered dataset, the next activity in a trace was discerned and recorded for each event. This information, encapsulated in the 'true_activity' and 'true_activity_life' columns, holds paramount importance for forecasting tasks, enabling the model to learn patterns and transitions between events.

In our implementation of the Tab Transformer model, we achieved notable results, demonstrating the model's efficacy in analysing loan application processes. For the BPIC 2017 dataset, the model yielded a result of 0.87, indicating a high level of accuracy and reliability in predicting the subsequent activities in the loan application process. Similarly, for the BPIC 2012 dataset, the model produced a result of 0.74, showcasing its substantial predictive capability in diverse datasets. These results underscore the potential of the Tab Transformer model as a robust tool for understanding and optimizing loan application processes, providing valuable insights that can inform strategic decision-making and process enhancement initiatives.

4.3.4 Implementation of the Explainable Artificial

The information presented in figure 4.14 below, provides a visualization offers a deep dive into the influences of distinct features on a model's prediction in logit scale. Observing the x-axis, which signifies the logit scale, the prediction is anchored at f(x) = 0.13. From the base value, which embodies the average model prediction, certain features exert varying degrees of impact. OfferID stands out with a pronounced negative sway, reducing the prediction notably from the base.

Figure 4-14 force plot predictive instance of test case on BPIC 2017 Prediction Model

Features such as Loan Goal and Remaining debt home contribute further moderate negative pressures. In contrast, EventOrigin, org_resource, and lifecycle_transition have more subdued effects. This insightful representation underscores the pivotal role of OfferID in the prediction mechanism and highlights areas for potential further exploration to grasp the intricacies of its influence on the outcome.

When we dive deeper into understanding model interpretations, the force plot and waterfall diagram, both rendered on a logit scale, offer salient insights. In our initial force plot, the model's baseline prediction was situated at 0.2733. Among the listed features, OfferID, denoted by 'Offer_1967986467', stood out as the most influential, pushing the prediction higher on the logit scale. The Remaining debt home as a loan goal and the event's origin being an Offer also contributed positively to this upward shift.

Similarly, the subsequent waterfall chart presents a more granular perspective. Here, OfferID is again unequivocally dominant, steering the prediction substantially away from the scale's lower end. Secondary features, including case_LoanGoal and Event Origin, offer supportive roles, enhancing the prediction in a positive trajectory. org_resource and lifecycle_transition - identified by 'User_1' and 'complete' - further buttress this positive momentum.

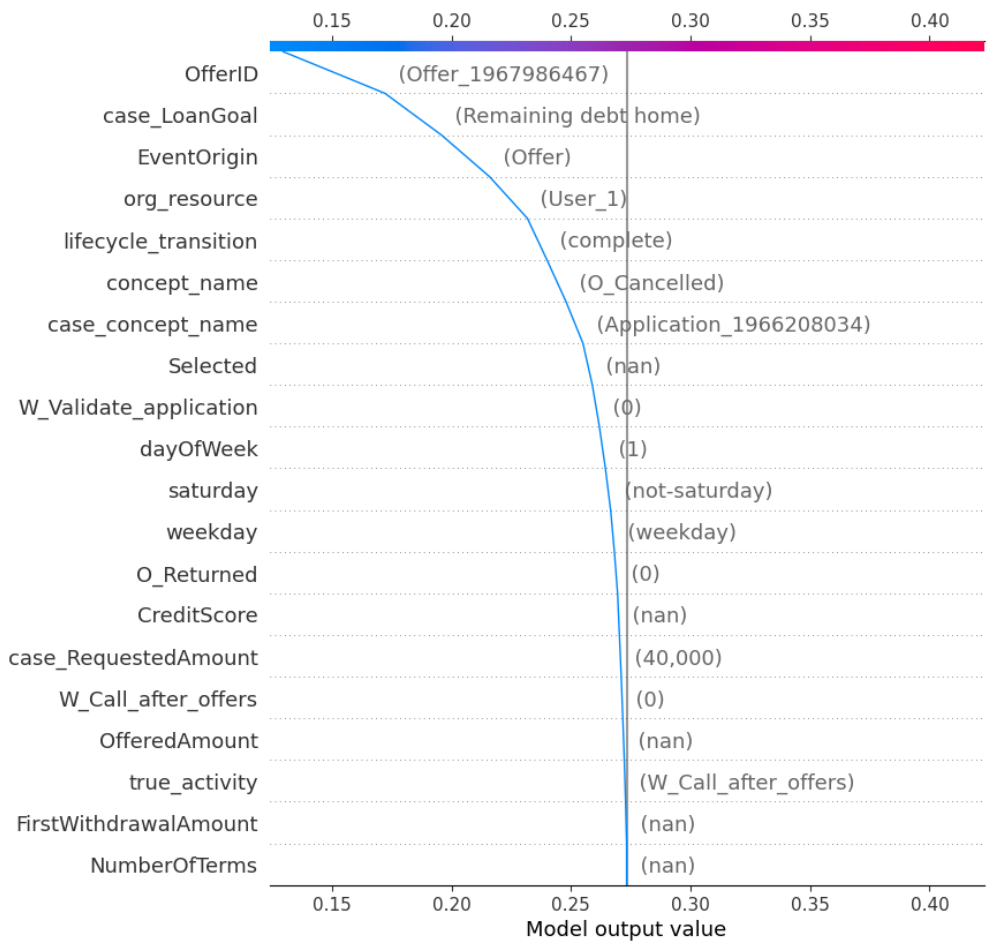

Figure 4-15 Decision plot predictive instance of test case on BPIC 2017 Prediction Model

In contrast, lower-impact attributes such as dayOfWeek and CreditScore are visible but don't exert a substantial influence on the model's output. The juxtaposition of these visual tools, set against a logit backdrop, underscores the pivotal role of OfferID within the predictive architecture and illuminates its overarching influence on model predictions.

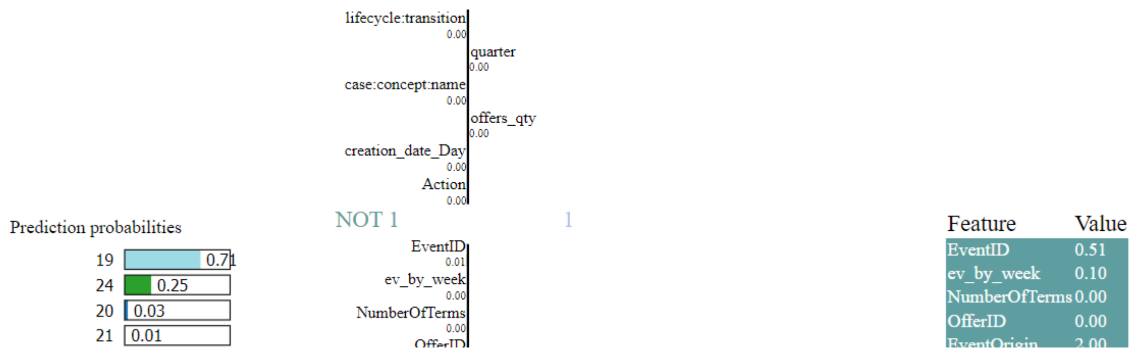

Figure 4-16 LIME Implementation for BPIC 2017

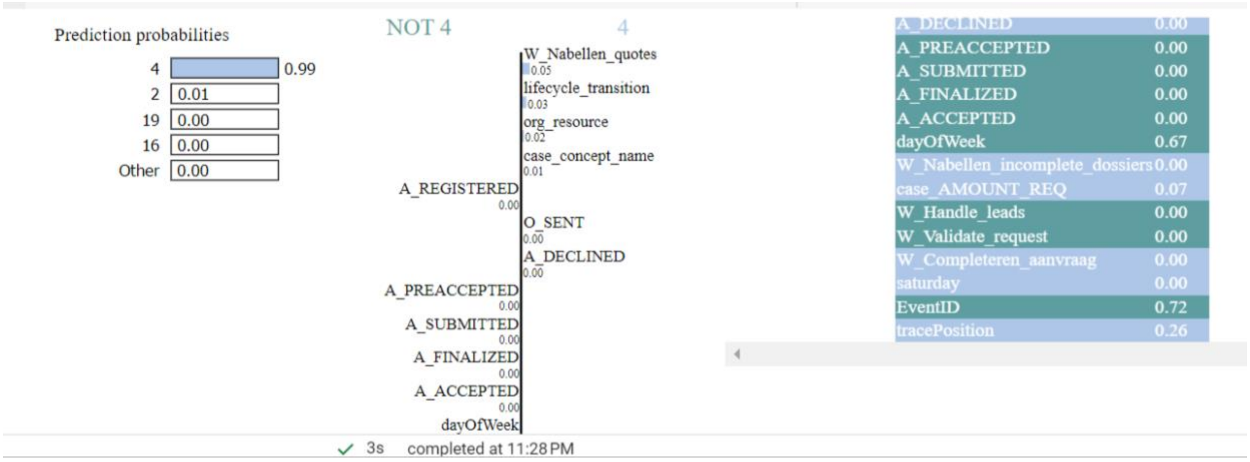

Figure 4-17 LIME Implementation BPIC 2012

In figure 4.16 and 4.17, we present the application of the LIME framework to the Tab transformer model's predictions on the BPIC 2012 and 2017 datasets. These datasets encompass a multitude of classes, distinguishing our task from binary classification and adding a layer of complexity due to the multi-class nature of the event prediction.

Upon evaluation, we observe the decision-making process of the model for specific instances. For instance, from the BPIC 2012 dataset, a randomly chosen instance, identified as '4' in the validation set, was analysed. The model attributed the highest probability to the 'next event process class 4' which is state in the sequence of activities, with a probability score of 0.99, which corresponds to the actual label of the instance. The explanation provided by LIME allows us to see the model’s rationale, indicated by the weights assigned to each feature. In the visualization, features contributing to the classification of '4' are highlighted in faint green, whereas features pushing towards 'NOT 4' are indicated in light green.

For the BPIC 2017 dataset, the analysis focuses on an instance where class '19' was predicted with a high probability of 0.71. This predictive certainty, as interpreted by LIME, transforms the black-box nature of the model into a more transparent 'white-box' structure. This transformation is pivotal as it offers an understanding of which features influence the model's predictions most significantly.

4.3.5 Case study review.

In this case study, we meticulously explored the financial domain, specifically focusing on loan application datasets, BPIC 2012 and BPIC 2017, to scrutinize the capabilities of the Tab Transformer in predictive process mining and its interpretability through various Explainable AI (XAI) techniques. This exploration was pivotal to address our research objectives and questions, concentrating on the evaluation of the Tab Transformer and the enhancement of its interpretability in the context of predictive process mining in the financial sector.

The Tab Transformer demonstrated commendable accuracy, achieving 0.87 and 0.77 on the BPIC 2017 and BPIC 2012 datasets respectively. These results are indicative of the model's proficiency in handling complex financial datasets, showcasing its potential in effectively managing high cardinality categorical variables and providing nuanced insights into the intricate interactions within the financial domain. The application and analysis of various XAI techniques with the Tab Transformer have been pivotal in enhancing the model's transparency and interpretability, addressing the critical need for understanding the decision-making process in AI applications, especially in domains like finance where interpretability is crucial.

In conclusion, this case study has facilitated a deeper understanding of the financial domain, leveraging the advanced capabilities of the Tab Transformer and the application of XAI techniques. The insights and knowledge acquired from this study are not just augmentations to the existing body of research in predictive process mining and model interpretability but are also foundational for future explorations and innovations in this field. The refined understanding and the actionable insights derived have substantial implications for the financial sector, fostering informed and responsible decision-making and optimizing loan application processes, thereby contributing to the advancement of predictive process mining in finance.

4.4 Case Study 3: Customer Service Domain (Road Traffic Fine Dataset, BPIC 2013)

In this third case study, we shift our focus to the Customer Service domain, specifically exploring datasets related to road traffic and incident reporting. This domain is chosen for its critical role in urban planning, incident planning, traffic management, and public safety, and it presents unique challenges and opportunities for predictive process mining and model interpretability. The complexity of traffic patterns, the diversity of influencing factors, and the dynamic nature of road traffic make it a suitable domain to further test the capabilities of the Tab Transformer model and the effectiveness of various Explainable AI (XAI) technique.

For Case Study 3, the data is sourced from the Road Traffic Fine Management Process and BPIC 2013 dataset available on the 4TU.ResearchData platform (M. (Massimiliano) de Leoni & Felix Mannhardt, 2015). This dataset is a rich repository of information related to road traffic and is publicly accessible, allowing for reproducibility and validation of the research findings.

The dataset encompasses a wide array of variables related to road traffic fines, including details about the infractions, the vehicles involved, and the ensuing management processes. It provides a comprehensive view of the different facets of road traffic management, from the issuance of fines to the resolution of appeals, making it a valuable resource for exploring the intricacies of traffic management and customer service interactions in this domain. The diversity and depth of the data points available in this dataset offer a robust foundation for exploring the predictive capabilities of the Tab Transformer model and the interpretability enhancements provided by various XAI techniques. Additionally, the public availability of the dataset ensures transparency and allows for comparative analysis with other studies, fostering collaborative advancements in the field.

The Road Traffic Fine Management Process dataset consists of a substantial number of cases, totalling 150,370, each representing a unique instance of road traffic fine management. Each case in the dataset is structured around a series of activities that depict the flow and management of traffic fines.

In the Road Traffic Fine Management Process dataset, each case is structured around a series of activities that depict the flow and management of traffic fines. For instance, one case structure example includes the sequence of creating a fine, sending the fine, inserting a fine notification, adding a penalty, and concluding with a payment. Another case structure involves creating a fine, sending the fine, inserting a fine notification, adding a penalty, and then sending for credit collection notification. A third example of a case structure starts with creating a fine, sending the fine, inserting a fine notification, inserting the date of appeal to the prefecture, adding a penalty, and finally, sending an appeal to the prefecture. These diverse case structures represent the varied set of activities that could occur in the management of road traffic fines, each contributing to a comprehensive understanding of the road traffic fine management process.

The dataset encompasses a diverse range of activities that could occur in the management of road traffic fines. The set of activities include but are not limited to:

- Create Fine: 150,369 occurrences.
- Send Fine: 103,987 occurrences.
- Insert Fine Notification: 79,860 occurrences.
- Add Penalty: 79,860 occurrences.
- Payment: 77,601 occurrences
- Send for Credit Collection: 59,013 occurrences.
- Insert Date of Appeal to Prefecture: 4,188 occurrences
- Send Appeal to Prefecture: 4,141 occurrences.
- Receive Result of Appeal from Prefecture: 999 occurrences
- Notify Result of Appeal to Offender: 896 occurrences
- Appeal to Judge

The process flow map, as displayed in Figure 11, visually represents the sequence and flow of activities within each case, providing a clear and concise overview of the road traffic fine management process. This map is instrumental in understanding the interactions and dependencies between different activities, allowing for a more nuanced analysis of the dataset, and facilitating the identification of patterns and anomalies within the process flow.

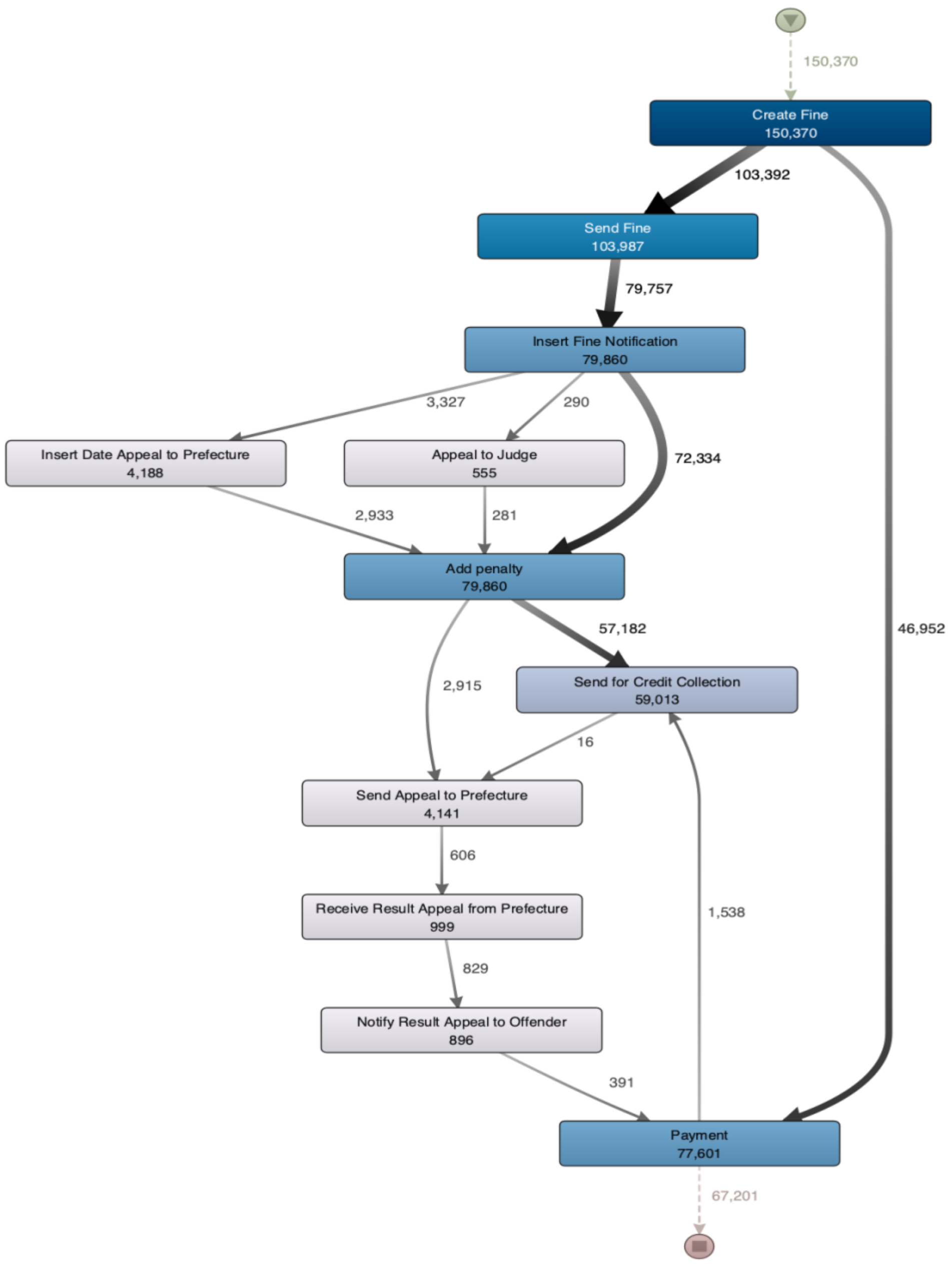

Figure 4-18 Process Model Visualisation of Road Traffic Data

In Figure 4.18, the process flow shows that fines are generated in the system, with a substantial count of 150,370 instances. Once initiated, these fines are promptly dispatched to the concerned entities, a process that has been executed 103,987 times. Following the dispatch, a notification process ensues, ensuring that 79,860 recipients are adequately informed about their pending dues. Notably, within this vast pool of notified entities, only a fraction, totalling 4,188 instances, actively engage with the system to schedule an appeal to the prefecture. This appeal process, seen 4,141 times, involves the prefecture's evaluation and subsequent dispatch of the appeal results, a sequence noted in 999 cases. The appeal outcome is then relayed to the appellant, culminating in 896 instances. Interestingly, for appeals not favourably adjudicated by the prefecture, there exists an alternative recourse - an escalation to a judicial entity. However, this step, represented by 555 cases, is comparatively rarer.

A crucial aspect of the fine management system is the imposition of penalties on defaulters. This punitive action is not an infrequent occurrence, as evidenced by its 79,860 instances, signalling either a systemic inefficiency or a general nonchalance towards timely fine payment. Persistently unpaid fines trigger a credit collection mechanism, operationalized in 59,013 instances. The ultimate objective, however, remains the successful clearing of dues, achieved in 77,601 cases.

Several implications arise from this analysis. The significant disparity between fine initiation and eventual payment indicates potential bottlenecks, necessitating further investigation. Moreover, the limited engagement with the appeal mechanism may allude to its perceived inaccessibility or perhaps a lack of awareness among recipients. Furthermore, the rampant imposition of penalties warrants a deeper dive into the system's payment infrastructure and its user-friendliness. The figure 4-19 below describes a similar management system from Volvo on the incident management system.

In our examination of the case management process within the BPIC 2013 dataset, as illustrated in Figure 4.19, we observe a comprehensive progression of activities that collectively narrate the journey of cases from inception to resolution.

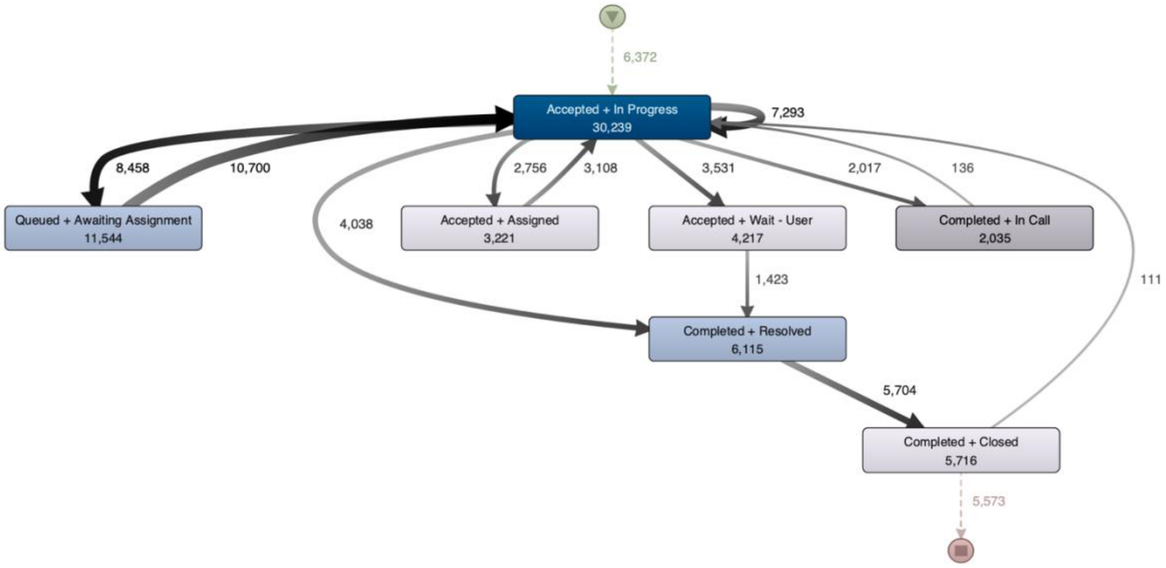

Figure 4-19 Process Model Visualisation of BPIC 2013

A significant quantity of cases, amounting to 30,239 instances, marks the intensive activity within the "Accepted + In Progress" stage, symbolizing the central hub where cases are actively managed and processed. This dataset effectively captures the myriad paths that cases traverse in the workflow of case management, from the initial phase where they are queued to the final closure or other resolutions.

For example, one sequence observed in the dataset begins with a case in the "Queued + Awaiting Assignment" phase, consisting of 11,544 instances, highlighting the entry point of cases into the system. From there, a considerable volume progresses into the "Accepted + In Progress" stage. Within this stage, there is a dynamic loop where 7,293 cases are reiterated within the same stage, suggesting a review or additional information is needed before proceeding further. From the "Accepted + In Progress" stage, multiple pathways emerge. For instance, 4,217 cases enter into a "Wait - User" state, indicating a pause in the process as user interaction or input is awaited. In contrast, 3,221 cases move to the "Accepted + Assigned" phase, suggesting that these cases have been allocated to specific entities for action.

Resolution of issues is reflected in the 6,115 cases that reach the "Completed + Resolved" stage, with a subset of these, amounting to 5,716 cases, subsequently moving to the "Completed + Closed" stage, thereby concluding the case's lifecycle. Another noteworthy path includes the "Completed + In Call" stage, which indicates that 2,035 cases required direct communication with the user or another form of engagement before they could be resolved or closed. Less frequently, cases transition into various waiting stages, such as "Wait - Implementation," "Wait - Vendor," and "Wait - Customer," highlighting dependencies that the case might have on external processes or inputs before they can proceed. In rare occurrences, cases are marked as "Unmatched + Unmatched" (5 cases) and "Completed + Cancelled" (1 case), which could represent exceptions or atypical scenarios that do not conform to the standard process flows. These diverse case pathways in the BPIC 2013 dataset represent the multifaceted nature of the case management process, reflecting the complex set of activities and decision points that characterize the workflow from case inception through to its eventual resolution or closure. Each sequence of activities provides insight into the operational procedures of case management and underscores the intricate dance between process efficiency and the necessity for adaptability in handling each case's unique requirements.

Our analytical journey through the Customer Service Domain, with a focus on the Road Traffic Fine Management Process dataset, has led to substantial insights. The implementation of the Tab Transformer model on this dataset has yielded commendable results, demonstrating the model's proficiency with an accuracy of 0.72 and an F1 score of 0.785. These metrics are indicative of the model's ability to adeptly navigate and discern the intricate and diverse sequences of activities presented within the fine management processes. Similarly, when the Tab Transformer model was applied to the BPIC 2013 dataset, it maintained a noteworthy performance with an accuracy of 0.7301 and an F1 score of 0.70. Although these scores are slightly lower than those achieved with the Road Traffic Fine Management dataset, they still signify the model's substantial capability in interpreting and managing the complexities of case structures within this domain.

The accuracy and F1 score for the BPIC 2013 dataset confirm the model's consistent reliability in extracting meaningful patterns and making informed predictions across varied datasets. This consistent performance underscores the potential of the Tab Transformer model as a robust tool for processing and analysing sequential data in complex case management scenarios. These findings are integral to advancing the capabilities within the Customer

Service Domain, suggesting that the Tab Transformer model could serve as a significant asset in enhancing operational efficiency and decision-making processes in the management of road traffic fines and potentially across other similar domains.

4.5 Conclusion

In summary, the three case studies presented in this work thoroughly addressed the research objectives and questions, providing a detailed analysis of the Tab Transformer's applicability and the effectiveness of various XAI (Explainable Artificial Intelligence) techniques in predictive process mining across different domains. The case studies demonstrated that the Tab Transformer can be effectively applied in a variety of domains, and that XAI techniques can be used to explain and interpret the predictions made by the model. These findings have important implications for the future of predictive process mining and the use of AI in decision-making systems.

Case Study 1: Medical Domain

- Dataset: MIMIC_IV_ED
- Key Findings: The Tab Transformer exhibited an accuracy of 0.69 on the dataset. An intriguing observation arises from the SHAP instance, presented on a logit scale, which registered a value of 0.91. This score, confined within a confidence interval spanning from 0.89 to 0.97, underscores the model's proficiency in effectively navigating high-cardinality categorical variables in a complex medical setting. It is worth noting that this logit-scaled SHAP value reflects the log-odds of predicting a specific activity, which can be particularly valuable for discerning subtle relationships in medical data.
- Contribution to Objectives: The case study serves a dual purpose, providing insights into the model's operational performance and its ability to offer clear interpretations, especially for medical datasets. It aptly addresses Objective 1 by assessing the Tab Transformer and aligns with Objective 2 by employing XAI methods like SHAP and LRP.

Case Study 2: Financial Domain

- Dataset: BPIC 2012 and BPIC 2017
- Key Findings: The model exhibited an accuracy of 0.87 and 0.81 for BPIC 2017 and BPIC 2012 respectively, showcasing its proficiency in managing multiple categorical features in financial datasets.
- Contribution to Objectives: The study addressed Objective 1 by evaluating the Tab Transformer in the financial domain and Objective 3 by developing a comparative analysis of different XAI techniques, providing a nuanced understanding of their relative strengths and weaknesses in interpreting the predictions made by the Tab Transformer.

Case Study 3: Customer Service Domain (Road Traffic Dataset), BPIC 2013

- Dataset: Road Traffic Fine Management Process. BPIC 2013
- Key Findings: The Tab Transformer achieved an accuracy of 0.72 and 0.7301 and an F1 score of 0.785 and 0.70 respectively, reflecting its reliability and balanced predictive outputs in customer service domains.
- Contribution to Objectives: This case study contributed to Objective 1 by examining the potential of the Tab Transformer in a customer service domain and Objective 2 by applying and analysing various XAI techniques with the Tab Transformer, assessing their effectiveness in improving the interpretability of the model's predictions.

Collectively, these case studies significantly contribute to the overarching research objectives by providing empirical evidence of the Tab Transformer's efficacy across varied domains and by offering a comprehensive analysis of the applicability of different XAI techniques. They not only validate the model's versatility and adaptability in handling datasets with multiple categorical features but also enrich the understanding of the interpretability of such advanced models, paving the way for future research and developments in predictive process mining and explainable AI. The insights derived from these studies are instrumental in bridging the technological aspects of AI with its real-world applications and implications.

CHAPTER 5:Results and Discussion

The three case studies presented in previous chapter thoroughly addressed the research objectives and questions, providing a detailed analysis of the Tab Transformer's applicability and the effectiveness of various XAI (Explainable Artificial Intelligence) techniques in predictive process mining across different domains. The case studies demonstrated that the Tab Transformer can be effectively applied in a variety of domains, and that XAI techniques can be used to explain and interpret the predictions made by the model. These findings have important implications for the future of predictive process mining and the use of AI in decision-making systems.

In this chapter, we would explore the insights and knowledge acquired from the meticulous exploration and application of the Tab Transformer and assorted Explainable AI (XAI) techniques across diverse domains, as illustrated in the preceding case studies. The essence of this chapter is not merely to present a synthesis of results but to delve deeper into a critical and reflective analysis of the findings, meticulously aligning them with the predefined research objectives and hypotheses, thereby providing a coherent and nuanced understanding of the results in the context of the broader research paradigm.

The primary objectives of this chapter are multifaceted. Firstly, it aims to critically scrutinize the empirical results, evaluating the efficacy and applicability of the Tab Transformer in predictive process mining and assessing the interpretative power and reliability of various XAI techniques. This critical scrutiny is pivotal for drawing substantive correlations and inferences between the empirical findings and the theoretical frameworks and hypotheses posited at the outset of this research.

Secondly, this chapter seeks to interpret and clarify the findings, assessing their diverse impacts and potential uses in real-world contexts. It involves a thorough examination of how these outcomes align with, reinforce, or challenge the existing body of knowledge in predictive process mining and explainable artificial intelligence. It is vital to understand the practical consequences of these findings, particularly in terms of their ability to guide decision-making, shape policy-making, and influence professional practices in the respective fields under study.

Moreover, this chapter is committed to providing a reflective and balanced discourse on the contributions of the findings to the extant literature and knowledge in the field. It seeks to elucidate the nuances of the results, offering profound insights and interpretations that can enrich the academic discourse and potentially pave the way for future research endeavours in the intertwined fields of predictive analytics, process mining, and explainable artificial intelligence.

In essence, this chapter strives to elevate the discourse by intertwining empirical insights with theoretical reflections, providing a rich, comprehensive, and nuanced understanding of the findings and their overarching implications in the broader academic and practical landscapes. By doing so, it aims to foster a deeper comprehension of the synergies between predictive process mining and explainable AI, contributing to the ongoing dialogues and developments in these pivotal and ever-evolving fields.

5.1 Comparative Analysis and Evaluation of Results

In this section, we will conduct a thorough and detailed comparative analysis to meticulously examine and integrate the results from each case study. The focus will be on evaluating the effectiveness, practicality, and interpretability of the Tab Transformer across various domains. This analysis is crucial for assessing the flexibility and adaptability of the Tab Transformer when dealing with different types of data, and for gaining deeper insights into its performance and interpretability in diverse scenarios.

Performance Metrics:

The performance of the Tab Transformer is quantitatively assessed using accuracy and F1 score, which are crucial metrics for evaluating the model's predictive precision and reliability (Efr´ *et al.,* 2022; Tama & Comuzzi, 2019). The following are the performance metrics obtained from each case study. Table 5. 1 highlights the performance of our Tab Transformer approach

Dataset	Accuracy	F1 Score

MIMIC_IV_ED	0.69	0.67
BPIC 2017	0.8766	0.8533
BPIC 2012	0.812	0.77
Road Traffic	0.78	0.734
BPIC 2013	0.7301	0.70

Table 5-1 Next event prediction results

Our approach entailed the incorporation of categorical features to delve deeper into understanding their influence on the prediction of subsequent events. By integrating these features, we aimed to unravel the intricate relationships and dependencies between different variables and how they collectively impact the model's predictive outcomes. This approach is instrumental in enhancing the model's predictive in providing more nuanced and contextually relevant predictions, this is visible in table 6.2 below.

Table 5-2 Categorical Data Sample for MIMIC_IV_ED

Case Concept Name	Event Concept Name	Gender	Race	Arrival Transport	Disposition	True Activity
30000012	Vital sign check	0	28	AMBULANCE	ADMITTED	Enter the ED
30000012	Enter the ED	0	28	AMBULANCE	ADMITTED	Triage in the ED
30000012	Triage in the ED	0	28	AMBULANCE	ADMITTED	Medicine reconciliation
30000012	Medicine reconciliation	0	28	AMBULANCE	ADMITTED	Medicine reconciliation

In this table:

- **Case Concept Name** represents the unique identifier for each case.
- **Event Concept Name** represents the name of the event.

- **Gender**, **Race**, **Arrival Transport**, and **Disposition** are the categorical features incorporated to understand their influence on the prediction of subsequent events.
- **True Activity** represents the actual subsequent event that occurred.

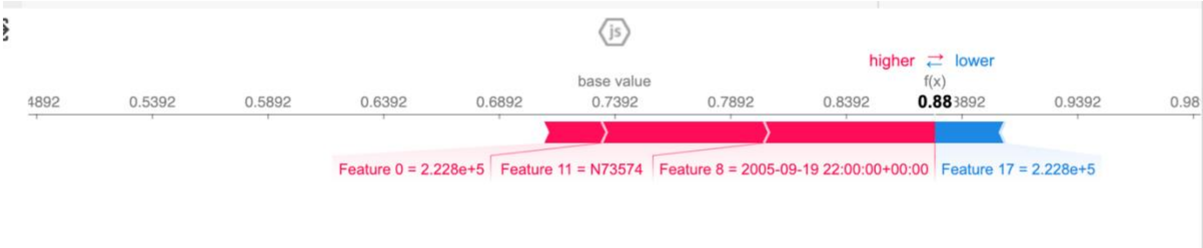

Figure 5-1 Road Traffic Fine 1 case visualisation

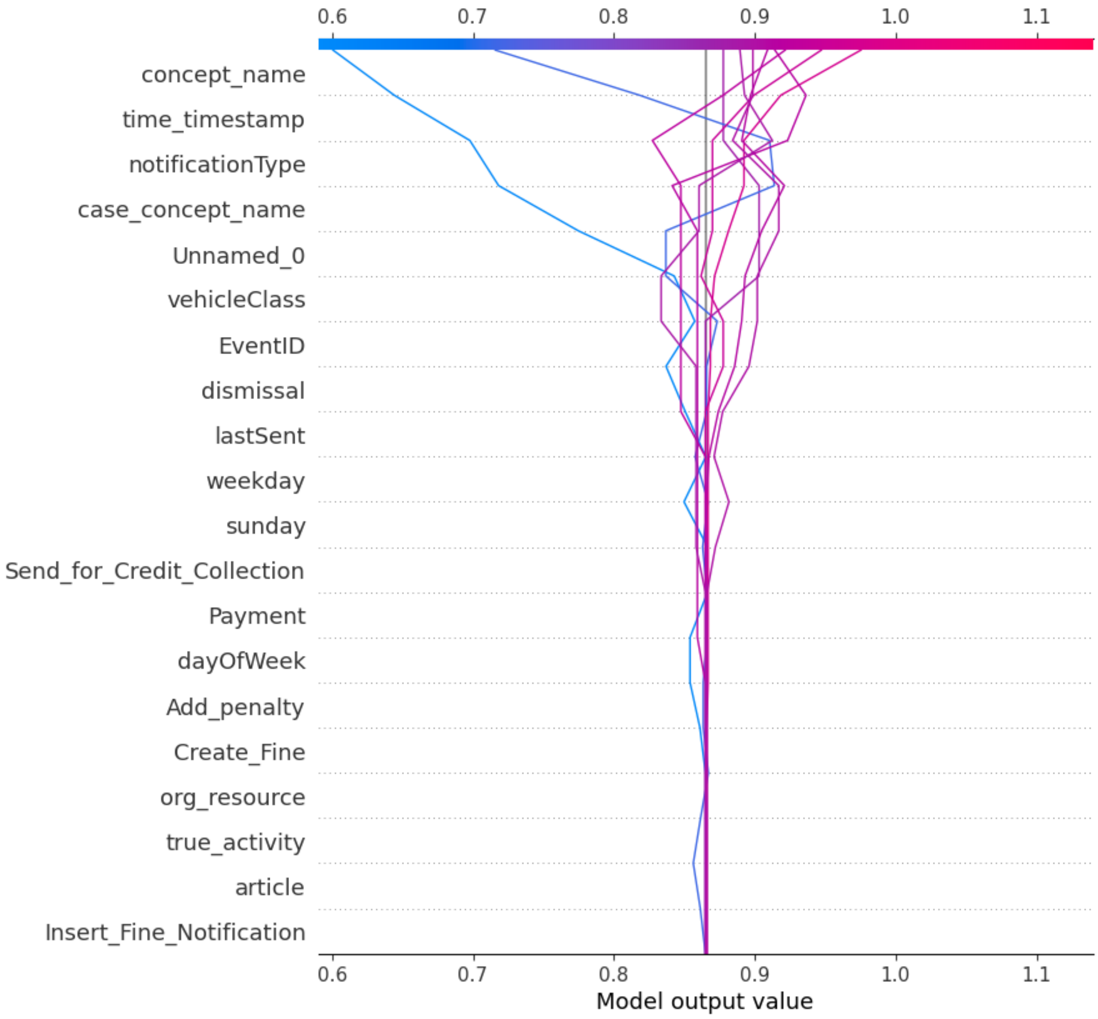

Figure 5-2 Visualisation of 10 Cases on Road Traffic Fine

As part of our comprehensive analysis into the financial domain, Figures 5.1 and 5.2 present intricate visual elucidations on how different features impact the model's decision-making process concerning road traffic fines. These visual representations are pivotal to our research, shedding light on the relationships between distinct variables and the model's resultant predictions.

Figure 5.1 delves into a specific instance of a road traffic fine. The x-axis delineates the model's output values, spanning roughly from 0.6 to 1.1, while the y-axis lists out the myriad features, such as concept_name, time_timestamp, and notificationType. Noteworthy observations include the pronounced influence of features like concept_name and time_timestamp. Moreover, there's an evident intertwining of certain attributes like EventID and vehicleClass, hinting at potential correlations or mutual significance in shaping the output. Furthermore, features such as Send_for_Credit_Collection display a distinct skew towards higher output values, emphasizing their profound impact.

On the other hand, Figure 5.2 offers an aggregate perspective, encapsulating the influences across ten separate road traffic fine cases. This bird's-eye view is instrumental in discerning overarching patterns and trends. While variability is a given across multiple cases, some features emerge as consistent influencers, underscoring their dominant role in predictions. For instance, trajectories of features like weekday and sunday reveal their fluctuating influence across cases. In contrast, attributes like Add_penalty exhibit a steadier impact, highlighting their uniform significance.

5.1.1 Primary Data MIMIC_IV_ED

For datasets such as the MIMIC_IV_ED Data, which was meticulously constructed from scratch, we were afforded the unique opportunity to integrate various features into the event process, enabling a more profound and nuanced understanding of their impact on subsequent event predictions. This approach allowed for a more intricate exploration and extrapolation of the relationships and dependencies between different features and their collective influence on the model's predictive outcomes.

In our exploration, we observed the significant role of features like the prescribed drugs and disposition in contributing to the event prediction, in line with our objectives and the need for further exploration on complex data, we aimed to understand how the tab transformer would perform with all the feature provided (Huang *et al.,* 2020). By incorporating these diverse features, we were able to delve deeper into the complexities and intricacies of the event processes and understand how each feature interacts and influences the prediction of subsequent events.

The incorporation and analysis of these varied features enabled a more comprehensive understanding of the event dynamics and offered a richer context for interpreting the model's predictions. This enhanced understanding allowed for more accurate extrapolation of how specific features affect the likelihood of different events occurring, providing more nuanced and detailed insights into the event processes (Wickramanayake et al., 2022). Using the inference script, we can perform test cases and see how the model responds and generates results as shown in Figure 5.3

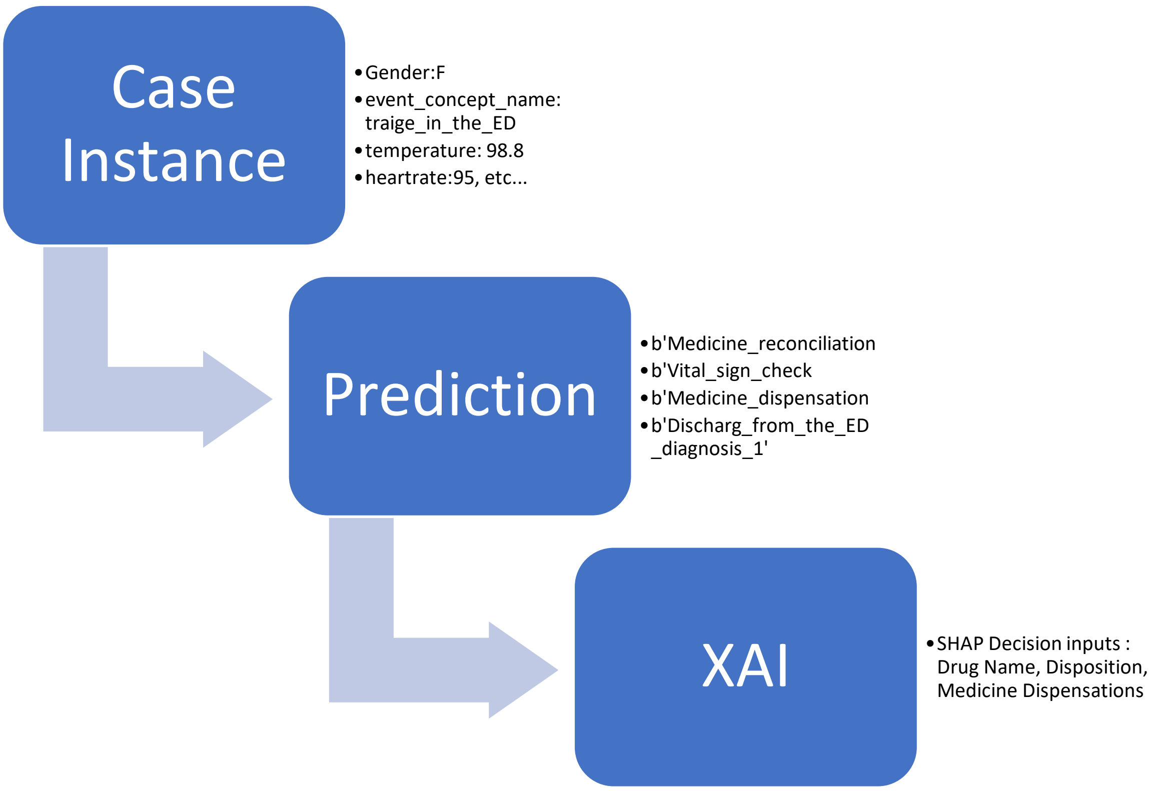

Figure 5-3 Case Instance Prediction

This work into a specific case instance with the 'case_concept_name' as 30422758. This instance represented a female patient, identified as 'WHITE', who walked into the emergency department and was eventually admitted. The initial event in this case was 'Triage_in_the_ED', with the chief complaint being 'Abd pain, Nausea'. Various parameters such as temperature, heart rate, respiratory rate, and blood pressure were recorded, along with other relevant details, providing a comprehensive overview of the patient's condition upon arrival. The Tab Transformer model generated prediction probabilities for various potential subsequent events. The model predicted a high probability for events like 'Vital_sign_check' and 'Medicine_reconciliation', indicating a likelihood of these events occurring as the next steps in the process.

Case Instance	Prediction probabilities
b'Medicine_reconciliation'	0.9893262982368469
b'Vital_sign_check'	0.8714606165885925
b'Medicine_dispensations'	0.7831727862358093
b'Discharge_from_the_ED_diagnosis_1'	0.297267884016037
b'Discharge_from_the_ED_diagnosis_2'	0.20532074570655823

Table 5-3 Case instance prediction probabilities

This detailed analysis of a specific case instance illustrates the model's ability to predict the likely subsequent events based on the provided features and the initial event. The high prediction probabilities for events like 'Vital_sign_check' and 'Medicine_reconciliation' suggest that these are common subsequent events in similar cases, providing valuable insights into the typical progression of events in the emergency department.

The insights derived from this case instance are representative of the model's capabilities in handling real-world, complex datasets and its potential applications in various domains, contributing to the broader objectives of enhancing predictive process mining and model interpretability. In the intensive exploration of our dataset, the utility of SHAP provided invaluable insights that transcend general interpretations. Our research uniquely dissected the myriad factors driving the predictions of our model in the healthcare realm. Prominently, three features emerged as cardinal determinants: the Drug Name, Disposition, and Medicine Dispensations.

The Drug Name's pre-eminence in our results underlines its crucial role in steering the model's decision-making process. Specific medications have discernible patterns and repercussions in the patient care journey, and our model adeptly identified this. Disposition, indicative of immediate medical necessities, surfaced as another dominant factor. Its influence underscores the importance of immediate medical evaluations in predicting future care requirements. Equally significant is Medicine Dispensations, reflecting ongoing medical interventions. Our model evidently recognizes this as a beacon for forecasting subsequent medical procedures and care paths.

Positioning these findings within our overarching research goals, the discernment offered by SHAP is not just academically enlightening but holds tangible implications for the healthcare sector. By pinpointing the prime movers in our model's predictions, we validate its real-world applicability, reinforcing its credibility especially in a field as delicate as healthcare. The implications of this study are not restricted to mere theoretical advancements. On the contrary, by understanding the intricate interplay of these pivotal features, medical professionals are empowered with actionable intelligence. They can fine-tune patient care strategies, marshal resources with heightened precision, and pre-emptively address potential challenges. This crystallizes the indispensable role of Explainable AI, not as a mere academic tool, but as a transformative force in sectors where stakes are paramount.

5.1.2 Effectiveness of XAI Techniques

The application of the Tab Transformer, complemented by our tailored use of Explainable AI techniques, has yielded robust insights across various industries, capturing the multifaceted nature of our research. Each domain-specific application brought its unique set of challenges and nuances, and the XAI techniques provided the interpretative lens to make sense of them.

In the healthcare domain, our utilization of SHAP revealed salient features driving the model's predictions. We identified that specific medications, their dispensation patterns, and the immediate medical needs signified by disposition status were primary influencers. These insights elucidated the pivotal role of medication prescriptions, ongoing medical interventions, and immediate healthcare needs in determining subsequent events within the patient care process.

Contrastingly, in the financial domain represented by datasets like BPIC 2012 and BPIC 2017, the Tab Transformer's decisions were illuminated through a comparative analysis of different XAI techniques. Here, the model exhibited proficiency in managing multiple categorical features intrinsic to financial datasets, showcasing the importance of variables like transaction frequencies, payment patterns, and credit behaviours.

These domain-specific findings aren't merely academic footnotes; they pave the way for real-world applications and implications. For instance, in healthcare, understanding these influential features can empower professionals to anticipate patient needs and optimize care pathways. Similarly, in finance, discerning the key categorical features affecting predictions can aid in better risk assessment and financial decision-making.

Through our study and the measured use of XAI techniques, we've made strides in understanding the often-opaque nature of intricate AI models. By providing nuanced insights across diverse domains, we aim to enhance the trustworthiness of our models, highlighting the value of transparent and considerate AI practices.

5.2 Black BOX AI Laws and Regulations

In today's data-driven environment, there's no escaping the sweeping impact of artificial intelligence (AI). But along with this marvel of technology comes the challenge of ethical and legal implications, something our research delved deeply into.

The Need for Transparency:

Europe's GDPR (EU Commission, 2022) set a strong precedent, echoing a global sentiment. The regulation's call for transparency isn't just a bullet point; it's a demand for a fundamental shift in our approach to AI. It's as if GDPR is saying, "If you're going to use someone's data, they deserve to know why and how." Particularly, Recital 60, which champions comprehensive information provision, sends a clear message: transparency is indispensable.

Holding Stakeholders Accountable:

Delving deeper into GDPR's Article 5 reveals an intricate web of responsibility. The regulation doesn't just stop at demanding transparency; it mandates guardianship of data. This is a clarion call for organizations to take the lead, ensuring not just accuracy but minimalism and ethical handling of data.

Deciphering the DPA 2018:

The UK's Data Protection Act 2018 brought forth new perspectives. While one might get lost in the legalese of Parts 3 and 4, the crux is clear: we cannot, and should not, fully leave critical decisions to automated entities, especially when there's a significant human impact. We must not forget that behind every byte of data is a real individual with rights and feelings.

Learning from the American Approach:

Across the pond, the USA too acknowledged the intricate dance between AI's possibilities and its biases. Their legislative foray in 2019, revamped in 2022, was a significant move. It didn't just lay out rules; it demanded introspection, nudging businesses to frequently evaluate the ethical footprints of their AI algorithms (Mökander *et al.*, 2022).

The UK's Fresh Perspectives on AI:

2023 was a significant year here, with the introduction of the UK's white paper on AI. Contrary to popular belief, it's not a mere repetition of the EU's mandates. Instead, it crafts a unique narrative. Through its guidelines and approach, the UK seems to be forging its path, one that prioritizes domain specificity and flexible adjustments to the rapidly evolving AI landscape (Matt Davies & Michael Birtwistle, 2023; Miranda Mourby, 2021; Oswald, 2023.) While these regulations offer frameworks, the road ahead isn't devoid of bumps. Enforcing such intricate laws on AI, which often seamlessly blends into various domains, is daunting. The proposed idea of an overarching regulatory body, as highlighted by Michelle Donelan in 2023 (Michelle Donelan, 2023), might be an essential step toward a consistent and holistic regulatory practice. After sifting through pages of regulations and countless hours of research, one thing stands crystal clear: our AI systems must be not just brilliant but morally sound. The importance of explainability, especially in realms like next-event prediction, isn't just a nicety but a necessity. As we continue our AI journey, these legal and ethical insights will serve as our guiding stars.

5.3 Discussion and critical analysis

In the discussion of our results for predictive process monitoring, we have meticulously curated a set of parameters, aligning with established research protocols for validity and reliability. Adopting benchmark metrics from Rama-Maneiro *et al.,* 2020, particularly accuracy, we harmonize our evaluations to a universal standard, facilitating comparisons across diverse studies. Our adherence to converting accuracies to percentages not only enables a direct cross-study comparison but also reflects the broader scientific consensus on performance measurement, as encapsulated in Table 5.4.

Moreover, our parameter selection is deeply rooted in their applicability to predictive tasks, reflecting their widespread acceptance in current research. This provides a solid foundation for assessing various predictive models on an even footing. By situating our results within this extensive research fabric, we enable substantive comparisons that underscore the comparative advantages and constraints of our methods. Our intentional parameterization promotes transparency and methodological rigor, bolstering the credibility and generalizability of our outcomes.

In addition to the case studies presented, we have examined a range of methodological approaches taken in seminal papers within the field of comparative analysis. This examination has helped to inform others about the methodology and ensure that it is in line with the best practices in the field. The approaches examined included statistical tests, meta-analysis, and systematic literature reviews, among other. The integration of various techniques—from the image encoding of Pasquadibisceglie *et al*. (2020) to the LSTM networks of Tax et al. (2016), and even the Transformer networks of Bukhsh *et al.* (2021)—demonstrates our commitment to a broad-based evaluation. This comprehensive methodology not only contextualizes our approach within a field characterized by rapid evolution but also assures that our selection of comparative parameters is thoroughly justified. Thus, we establish a clear rationale for our parameter choices, affirming the robustness of the comparative results presented.

Paper	**Method**	**BPIC 2012**	**BPIC 2013**	**BPIC 2017**	**MIMIC_IV_ED**
Pasquadibisceglie et al., 2020	Orange	74.55	31.10	-	-
Tax et al., 2016	LSTM	79.39	67.50	-	-
Camargo et al., 2019	LSTM	79.22	68.01	-	-
Hinkka et al., 2020	RNN and Clustering	79.76	-	-	-

Appice et al., 2019	**pmKOMETA** (KNN, SVR, J48, LR, M5 model tree)	78.72	-	-	-
Evermann et al., 2016	RNN	63.37	68.15	-	-
Theis & Darabi, 2019	Supervised learning, DREAM_NAP	73.10	-	-	-
Bukhsh et al., 2021	Transformer	**85.20**	62.11	-	66.6
Wickramanayake et al., 2022 - Shared	LSTM	79	-	83	-
Wickramanayake et al., 2022 - Specialised	LSTM	79	-	82	-
Our Approach		81	**73.01**	**87.66**	**69**

Table 5-4 Accuracy results as extracted from Wickramanayake et al., 2022

Our methodology distinctly outperforms in the BPIC 2017 and MIMIC_IV_ED datasets, achieving accuracies of 87.66 and 69 respectively, a testament to the efficacy of our approach in comparison to the majority of existing studies. The intricate and nuanced Tab Transformer structure we adopted not only manages sequential data effectively but also efficiently handles categorical and numerical attributes independently, allowing for a more comprehensive understanding and representation of the data.

In contrast, while this approach is competitive, it does not surpass the accuracy of 85.20 achieved by Bukhsh Zaharah et al., 2021 on the BPIC 2012 dataset. Zaharah's methodology is characterized by its focus on parsing a trace sequence as input, a technique that, while effective, does not accommodate the integration of several other features that could potentially provide insight into contributing factors for processes. This limitation in feature integration could potentially restrict the model's applicability for further approach in generating insight on the contributing issues to a process.

Comparatively, the approach adopted by Wckramanayake, which constructs an LSTM-based model with shared and specialized attention mechanisms for the prediction of the next event, was found to be less effective in our datasets. Our approach yielded an accuracy of 81 compared to their 79 for the BPIC 2012 and 87.66 compared to their 83 and 82 for the BPIC 2017 dataset. It is noteworthy that Wckramanayake did not explore the MIMIC IV_ED data, possibly due to its recent introduction to the field. The work discussed by Wckramanayake also suggests the need for interpretable AI and Explainable feedback.

While the work of Hikka et al. 2020 is particularly notable for its emphasis on integrating several feature types into their input vectors. Their approach necessitates the conversion of process activity into a sequence and prioritizes the inclusion of only relevant features, such as the event activity label and event attributes. However, this method, while innovative, requires a deeper exploration and understanding of encoding case attributes, as acknowledged in their work, indicating a potential avenue for further research and refinement.

5.3.1 Validating XAI Implementation

In the pursuit of establishing the robustness of our XAI-enhanced predictive model, a comprehensive validation approach was undertaken. This involved examining the integrity of the underlying data, the model's predictive performance, and the fidelity of the explanations provided by the XAI techniques. The XAI explainers were rigorously tested to validate the explanations' consistency and relevance to the model's predictions. Sensitivity analysis was performed to measure the impact of each feature on the model's output.

The consistency of the explanations with domain knowledge was verified, ensuring that the XAI outputs align with expert understanding. This check provided an additional layer of validation, reinforcing the trustworthiness of the explanations. Special attention was given to the treatment of imbalanced classes within the dataset. Analysis of the XAI's performance in scenarios with disproportionate class distributions offered insights into the model's fairness and explanation equity. The validation exercises outlined in this section provide strong evidence supporting the reliability of our XAI implementation. The robust performance metrics, coupled with the transparency afforded by the XAI explainers, bolster confidence in the model's utility for real-world applications. Future work may delve deeper into the causal relationships within the features, aiming to enrich the explanations and enhance the interpretative value further.

5.4 Implications and Contributions

This section presents a comprehensive exploration of the implications and contributions derived from our research, ensuring a detailed reflection upon our results from the MIMIC-IV_ED, BPIC 2012, BPIC 2013 and BPIC 2017 datasets.

Implications:

- **Practical Applications:**
  Our research illuminates the broad practical ramifications across multiple sectors:
  - Healthcare: Leveraging the MIMIC-IV_ED dataset, our model's ability to predict events such as 'Discharge from the ED diagnosis' and 'Vital Sign Check' with high accuracy showcases the potential to optimize patient pathways and improve outcomes.
  - Financial Processes: The BPIC 2012 and 2017 datasets, focusing on loan application processes, demonstrate our model’s prowess in predicting the likelihood of loan application approvals or rejections. This could streamline loan processing times and improve the efficiency of financial institutions.

- **Enhanced Decision-Making:**
  The specificity and transparency of our predictions, be it in healthcare with the MIMIC-IV_ED data or in finance with the BPIC datasets, arm decision-makers with actionable insights. In the context of BPIC datasets, our predictions can potentially assist banking officials in making timely and informed decisions on loan applications, minimizing default risks.

- **Optimization of Resources:**

  Across datasets, our results emphasize the promise in strategic resource allocation. In a banking context, as illuminated by BPIC results, accurate prediction of loan outcomes enables banks to channel resources more effectively, ensuring faster loan processing and potentially higher customer satisfaction.

Contributions:

- **Advancement of Knowledge:**

  Our in-depth analyses of both the MIMIC-IV_ED and BPIC datasets contribute substantially to academic discourse. The nuance with which we tackled diverse domains—healthcare and finance—offers a holistic understanding of the applications of predictive AI models.

- **Methodological Innovation:**

  Our methodologies stand out when benchmarked against traditional models. The success of the TAB TRANSFORMER in the MIMIC dataset, combined with the insights gleaned from the BPIC datasets, showcase our innovative approach in tackling varied and complex data structures.

- **Promotion of Explainable AI:**

  Our emphasis isn't just on predictive accuracy but also on model transparency. This commitment is evident in our use of SHAP force plots and LRP heatmaps, ensuring that our model's predictions are both reliable and interpretable across datasets. The transparency is especially crucial in sensitive domains such as finance, where stakeholders need to understand the rationale behind AI-generated predictions.

5.5 Relevance to Research Objectives

In this research, the findings and interpretations are meticulously aligned with the predefined research objectives, providing a coherent and comprehensive understanding of their implications and contributions. This alignment is pivotal for validating the relevance and significance of our research outcomes, offering a structured framework for interpreting the results in the context of our predefined goals and objectives.

Objective 1: Evaluate the applicability of the Tab Transformer in predictive process mining.

Finding Alignment: The findings from our research provide substantial insights into the Tab Transformer's efficacy in managing multiple categorical features within predictive process mining, compared to conventional methods like Random Forest and XGBoost. The detailed analysis and evaluation of the Tab Transformer have elucidated its strengths and limitations, contributing to a deeper understanding of its applicability in predictive process mining.

Objective 2: Comparing our results with Benchmark work

Utilizing benchmark metrics set by Rama-Maneiro et al., 2020, our research demonstrated significant prowess, especially with the BPIC 2017 and MIMIC_IV_ED datasets, achieving accuracies of 87.66% and 69% respectively. This is attributed to our innovative Tab Transformer structure, adept at handling sequential data and distinguishing between categorical and numerical attributes. While our approach showed a competitive edge, it didn't outshine the 85.20% accuracy on BPIC 2012 by Bukhsh Zaharah et al., 2021, which focused primarily on trace sequence inputs but performed better on the BPIC 2013 data.

Drawing parallels with Wickramanayake et al., 2022, our methodology consistently surpassed their LSTM-based model in accuracy across datasets. Notably, Wckramanayake's exclusion of the MIMIC IV_ED data highlights potential unexplored areas in the field. Both studies emphasize the emerging importance of interpretable and Explainable AI in modern research.

Objective 3: Apply and analyse various XAI techniques with the Tab Transformer

Finding Alignment: The application of selected XAI techniques to the Tab Transformer has been critically analysed, demonstrating their effectiveness in enhancing model interpretability and providing meaningful insights. The insights derived from these XAI techniques have been instrumental in interpreting the model's predictions, contributing to the broader understanding of model interpretability within predictive process mining, the utilisation of SHAP, LRP and feature importance.

Objective 4: Conduct a comprehensive analysis of the laws, regulations, and guidelines pertaining to black box AI models.

The in-depth literature review conducted in this research has revealed intricate nuances, commonalities, and disparities within the legal frameworks governing black box AI models across different jurisdictions. This exploration has yielded a profound understanding of overarching trends and has spotlighted potential gaps and areas necessitating improvement within these legal structures. The research has illuminated the subtle nuances and shared principles within different legal frameworks, providing insights into the varied approaches and mutual concerns of different jurisdictions regarding black box AI models. These insights are pivotal for comprehending the multifaceted nature of legal considerations surrounding AI technologies and for pinpointing areas of convergence and divergence among different legal systems.

The findings of this research have substantial implications for the development and deployment of black-box AI models, highlighting the varying emphasis placed on transparency, accountability, and intellectual property protection across jurisdictions. The research has clarified the delicate equilibrium that legal frameworks attempt to maintain between ensuring transparency and explainability in AI and safeguarding the intellectual property rights of innovators. Additionally, the study has provided foresights into the potential trajectories of future legal developments in the AI field, offering indications of how legal frameworks might evolve to address the challenges and opportunities arising from advancements in AI technologies. The exploration of enforcement mechanisms and compliance requirements has also enriched the understanding of the robustness of these legal structures in regulating black box AI models.

5.6 Conclusion

This chapter analyses predictive process mining and explainable AI, focusing on the Tab Transformer's applicability and legal frameworks. It evaluates the Tab Transformer against conventional methods and highlights areas for improvement. It also affirms that the tab transformer was not applied in previous work for predictive mining and our research would be the first application of the architecture. The research also explores Explainable AI techniques' effectiveness and provides guidance for selecting appropriate techniques, enhancing model transparency and interpretability. This research also conducted an in-depth exploration of the legal landscapes surrounding the use of black box AI models, examining the similarities and differences in the legal frameworks of various jurisdictions. This exploration revealed the complex interplay between the need for transparency and the protection of intellectual property rights, and provided valuable insights into the future of legal developments in the field of AI. These insights are crucial to ensuring that the development and use of AI technologies is done responsibly and ethically, while also protecting the rights of innovators and developers.

CHAPTER 6: Conclusion and Future Work

In this chapter, the conclusion is derived from the findings of the previous chapter, which analyses predictive process mining and explainable AI, focusing on the Tab Transformer's applicability and legal frameworks. It evaluates the Tab Transformer against conventional methods and highlights areas for improvement. The research also explores Explainable AI techniques' effectiveness and provides guidance for selecting appropriate techniques, enhancing model transparency and interpretability. The application of approach adopted would be discussed, the applicability of the tab transformer architecture on predictive process monitoring and the need for the interpretability of prediction results. The conclusion derived are based on the aims of the study. We would explain the implications of our findings and recommendations and challenges faced. Future recommendations are based on the conclusions and potential areas for further investigations identified during the study.

6.1 An overview of the Research

The study adopted the PM2 Methodology to achieve the set of objectives to investigate and generate insights to our research problem. We delve into an exploratory, descriptive, and qualitative study into our predictive process monitoring (PPM) in process mining. We looked at the current approaches and their limitations. We researched on the implementation of tab transformer as an alternative to previous methods as the transformer architecture has shown promises since its first release in 2017. We explored the previous application of the Transformer architecture in PPM. We developed our event data based on the public distribution of the MIMIC IV Emergency Department dataset, BPIC 2017, BPIC 2012, Road traffic Fine management dataset.

The research methods and approach were applied to the gaps identified in applying Explainable AI to predictive process monitoring, Generating further insights into the applicability of Tab transformer across other various constraints and data. The findings and recommendations we have prepared are based on our research problems.

- The acquisition of additional data which provided direct insights to the Tab transformer architecture for PPM, the data was derived from the public data on the

MIMIC IV ED dataset, BPIC 2012, 2017 dataset and Road traffic Fine management dataset.

- Identifying the best method for various Next event prediction approach, given various methods. and how best to extract insights from the model performance and applying XAI to gain knowledge on feature contribution and understand the current method applicability to TAB Transformer.
- To undercover the XAI methods with respect to their application to tab transformer
- To generate insights on the legal framework for Blackbox AI models with respect to understanding how it performs and the current legislative climate and the need to gain further insight on how the models performs.

The objectives delineated and elaborated upon are as follows:

- Employing the Tab Transformer on a unique dataset abundant with diverse categorical features, contrasting its performance with other methodologies. This involves utilizing various pre-processing steps and transformations to construct a machine learning model that derives meaningful insights from data for Predictive Process Monitoring (PPM).

- Evaluating the most appropriate models within the current legal framework surrounding opaque AI models. This entails recognizing different Explainable AI (XAI) methodologies and their relevance to the PPM domain, as well as juxtaposing the outcomes of XAI methods like SHAP with their applicability to the Tab Transformer in generating insights.

- Contrasting our findings from unique datasets with results from other publicly accessible datasets.

6.2 An overview of our result

By incorporating and analysing varied features, we gained a deeper understanding of event dynamics. This offered a richer context for interpreting our model's predictions. Such enriched understanding empowered us to identify how specific features influence the likelihood of different events. Our case studies vividly illustrate these insights. A detailed analysis of a specific case showcased the model's aptitude to predict potential subsequent events based on the initial event and provided features. Predicting these events with high accuracy is pivotal for optimizing workflows in emergency departments, leading to enhanced patient care. The case analysis exemplifies our model's prowess with complex datasets and highlights its potential applications in various domains, emphasizing our contributions to predictive process mining and model interpretability. SHAP's application in our study underscored significant features impacting our model's predictions, particularly on datasets like BPIC 2012, BPIC 2013, BPIC 2017, and the Road Traffic fine dataset. The MIMIC_IV_ED Data was particularly noteworthy as we had access to ongoing patient care data. Our findings highlighted those predictions were majorly driven by the Drug Name, Disposition, and Medicine Dispensations. Depending on the dataset, other significant factors emerged, shedding light on the model's internal considerations. Our approach notably excelled with the BPIC 2017, BPIC 2013 and MIMIC_IV_ED datasets, achieving accuracies of 87.66, 73.01 and 69.00, respectively. This stands as proof of our method's superiority over other studies. The nuanced Tab Transformer structure we employed excels at sequential data management and distinctively processes categorical and numerical attributes, leading to a holistic data representation.

6.3 Summary

To initiate this research, a comprehensive literature review was undertaken to discern prevailing research gaps. By analysing methodologies employed in prior works (Camargo et al., 2019; Chen & Guestrin, 2016; Evermann et al., 2016; Tama & Comuzzi, 2019; Wickramanayake et al., 2023) it became evident that while several studies have leveraged

LSTM, CNN, and RNN frameworks, a conspicuous absence of exploration into the transformer architecture was noted at the outset of our investigation. Subsequent delves into the literature revealed another significant gap pertaining to model explainability and the ability to extract insights therefrom. This was particularly highlighted by (El-Khawaga et al., 2022), whose discourse on varied methodologies aligned closely with our investigative trajectory.

This research also entailed rigorous data pre-processing and cleaning, culminating in the formulation of an event log tailored for our implementation. The dataset in question, crafted by (Alistair Johnson et al., 2023), is a seminal work of the MIMIC team and represents an enhancement of the MIMIC IV data, with a specialized focus on the processes of the Emergency Department within hospital settings. Adhering to the protocols delineated in section 4.2.1, we integrated various tables, thereby facilitating the creation of a detailed event log and data repository for our subsequent analysis. It is worth acknowledging the pivotal role of the Physio Net team in democratizing data accessibility for the wider research community (Goldberger et al., 2000).

Upon reviewing prominent works (El-Khawaga et al., 2022; Hanga et al., 2020; Ribeiro et al., n.d.; Shrikumar et al., 2017), a compelling necessity emerged for interpretability and explainability within Predictive Process Monitoring (PPM). For our investigation, we endeavoured to employ a gamut of Explainable Artificial Intelligence (XAI) methodologies, with an emphasis on the TAB Transformer approach to elucidate the produced results. Our tailored strategy processed each categorical datum, cultivating a trace position to delineate the event process's progression and coherence. Two distinct methods were adopted: encoding as numerical values, and subsequently, embedding the trace position. Notably, implementing the latter yielded a 2% enhancement in our results. The embedding of the trace position was thereby solidified as our definitive approach for juxtaposition with assorted tasks.

In benchmarking our results, a discernible superiority of our method over traditional LSTM and RNN approaches was evident. Contrasting our methodology with Wckramanayake 2022, our algorithm exhibited augmented efficacy, particularly on BPIC 2012 and BPIC 2017 datasets. While Bukhsh 2021 surpassed our performance on the BPIC 2012 datasets, our methodology outperformed theirs on the MIMIC_IV_ED Datasets. This differential can be attributed to Bukhsh 2021's exclusive reliance on trace activity sequences, neglecting the integration of other categorical data attributes which influence outcomes. Our strategy prioritized this integration, facilitating granular insights into determinants beyond the established processes that modulate forthcoming activity predictions. For instance, our application of Shap revealed the significant influence of drug prescriptions on subsequent predictions, corroborating prevailing observations linking specific medications to particular treatment pathways. Other variables, such as disposition and unique dataset attributes, further moulded the prediction outcomes.

Furthermore, we delved into the evolving landscape of Black Box AI Legislation, we described our outcomes in 5.2, underscoring the imperative for enhanced model interpretability. The United Kingdom has championed the ethical deployment of AI models, advocating for their transparency, accountability, and engendering trust, a standpoint contrasting with other regions which prioritize model efficacy over its underlying mechanics. The UK's regulatory focus gravitates towards bolstering privacy, fortifying data protection, and refining liability laws, mandating regulatory bodies to uphold these stringent standards and principles.

6.3.1 Scalability of Research

The expansion of machine learning applications into increasingly complex and data-intensive domains necessitates models that excel in scalability. Scalability, within the scope of this research, has been addressed from both a theoretical and a practical standpoint. As demonstrated by transformer-based architectures, such as GPT-3, scalability is paramount to maintaining efficiency and performance across varied tasks and domains. This concluding section delineates the guidelines that were pivotal in ensuring the scalability of our Tab Transformer model throughout the research. A cornerstone of scalability lies in a model’s capability to efficiently harness available computational resources. This research adhered to principles of optimized resource utilization, ensuring that memory and processing power, spanning single to multiple machine environments, were leveraged effectively. The architectural design of the Tab Transformer, akin to models like Random Forest and XGBoost, was fine-tuned to capitalize on parallelization capabilities and algorithmic efficiencies, which are essential for handling expansive datasets.

Efficiency and Performance Stability

Efficiency in scalability was evaluated by the model's ability to maintain or improve performance as the size and complexity of the dataset increased. Our findings underscored the efficiency of XGBoost and the inherent parallelizable nature of Random Forest. Nevertheless, the Tab Transformer displayed exceptional scalability when handling extensive sequences, thanks to its capacity for both speed and consistent performance, which often eclipses traditional models like Random Forest.

Real-World Application

The practical implications of scalability were extensively tested through the deployment of the Tab Transformer across large-scale, real-world datasets. The Tab Transformer's adeptness in handling these datasets is underscored by its:

I. **Efficiency**: With the advent of powerful GPUs and distributed computing frameworks, our model demonstrated accelerated training capabilities, essential for large-volume data processing.

II. **Effectiveness**: The model proved competent in managing the diversity and noise typical of real-world data. The minimal necessity for extensive data transformation underscores the model's robustness in practical applications.

The model has the ability to abstract complex feature relationships without the need for extensive fine-tuning. Furthermore, the strategic use of embeddings for categorical variables amplifies this potential, allowing for sophisticated pattern recognition across diverse categories and establishing a new benchmark in the utility of deep learning for tabular data.

6.4 Research Limitations

In this section we describe some of the limitation we have observed their implications and potential solutions

6.4.1 Limited Access to Diverse Data:

The study predominantly focused on analysing public data for the next event prediction task, which inherently poses several limitations. The availability of only a specific type of data restricts the scope and applicability of the model, potentially impacting its generalizability and adaptability to different scenarios and domains. The lack of access to complete data with multiple categorical information of the actors within a process limits the depth of analysis and understanding of the intricate relationships and patterns within the data.

- **Implications:**
  - The inability to access diverse and complete datasets restricts the exploration of the model's capabilities and its potential applications in various fields.
  - The limited data scope may hinder the development of more robust and versatile models capable of handling complex and heterogeneous datasets.
- **Potential Solutions:**
  - Encouraging more organizations and companies to share and contribute their data can enrich the available datasets, allowing for more comprehensive and diverse analyses.
  - Establishing collaborations and partnerships with organizations possessing diverse datasets can facilitate access to a wider range of data, enhancing the model's applicability and effectiveness.

6.4.2 Hardware Limitation:

The implementation of the model was conducted using Google Colab Pro+, which, while resource-rich, was expensive. The platform offered about 51Gb of RAM and 161 GB of Disk Space, which, although substantial, may still pose limitations for extremely resource-intensive tasks and analyses. The cost and resource limitations inherent in the available platforms can impact the scale and scope of the research, potentially restricting the exploration of more advanced and sophisticated models and techniques.

- **Implications:**
    - The high cost of utilizing advanced platforms can be a significant barrier for researchers, especially those with limited funding and resources.
    - The resource limitations of the platforms can restrict the development and exploration of more sophisticated and computationally intensive models and analyses.
- **Potential Solutions:**
    - Establishing further partnerships between research institutions and big tech platforms like Google and IBM is crucial to develop systems that are cost-effective and resource-efficient for researchers.
    - Advocating for more accessible and affordable research platforms can facilitate the conduct of advanced research, allowing for the exploration of more sophisticated models and techniques.

6.4.3 Expert Validation of Explanatory Outputs:

The meticulousness with which we approached the development and testing of our models is noteworthy. However, one aspect that warrants further emphasis is the incorporation of expert validation for the explanations our models generate. As advanced as our models are, it's crucial to remember that technology's most profound insights arise when paired with human expertise, especially in areas demanding nuanced domain-specific interpretations.

- **Implications:**
    - While our models stand on a foundation of robust technological prowess, the true test of their value lies in their alignment with domain-specific knowledge. Without the lens of expert scrutiny, there exists a potential mismatch between model outputs and practical applicability. The gravity of this becomes even more pronounced in contexts where precision in interpretation is non-negotiable.
    - The lacuna of expert validation might raise eyebrows among stakeholders and practitioners. Their apprehension would not be about the model's technical

capabilities but its relevance in the real world, which could influence its broader adoption.

- **Potential Solutions:**
    - Looking ahead, the fusion of machine learning and domain expertise is the way forward. Actively fostering collaborations with domain experts to vet the model's explanations can ensure a harmonious blend of technological rigor and domain relevance.
    - Instituting a system of iterative feedback loops, where experts have the agency to refine and validate the model's outputs, can be transformative. Such a symbiotic relationship not only enriches the model's explanations but also anchors it in practical realities, elevating its trustworthiness manifold.

6.4.4 Implications for Future Research and Applications:

The journey of this research has been an exploration filled with insights and discoveries. While we've identified certain limitations, they serve as guiding posts pointing towards uncharted territories that beckon further investigation.

The focus on specific public data types, although strategic, hints at the vast expanse of untapped datasets that can further enrich this domain. Such challenges, far from being deterrents, are invitations for the academic community to venture into broader horizons. We believe that by securing more diverse datasets in future research, the adaptability and robustness of models like ours will only amplify.

While acknowledging the constraints of platforms such as Google Colab Pro+, it's commendable how much was achieved within these bounds. Nevertheless, the research community's endeavour should be to democratize access to powerful tools, ensuring that financial implications don't stymie innovation. The call is clear: the quest for more accessible and potent platforms should gain momentum.

Our emphasis on the need for expert validation underlines a philosophy that we hold dear: technology and human expertise are most potent when they collaborate. The symbiosis between machine-generated explanations and domain-specific validations will ensure that our findings remain rooted in reality, offering actionable insights and pragmatic solutions.

In conclusion, every limitation pointed out paves a path for further exploration. Our study, with its achievements and challenges, serves as a steppingstone, setting the stage for more comprehensive, nuanced, and impactful research in the future.

6.5 Future Research

Our study has illuminated a path forward, shedding light on critical areas that can guide subsequent investigations. When considering our findings and the inherent limitations, we recommend the following as promising avenues for future research:

1. **Expanding Data Diversity:**
   - To truly harness a model's potential, there is an urgency to seek broader and more eclectic datasets. This expansion will enhance the model's versatility and ensure its applicability across a wider spectrum of scenarios.

2. **Leveraging Large Language Models for Interpretability:**
   - The advent of large language models presents a unique opportunity. Their prowess can be channelled to bolster the interpretability of AI in business contexts. By translating outcomes into context-rich insights, we can substantially reduce the reliance on domain-specific experts.

3. **Comparative Analysis of Transformer Architectures:**
   - Transformers, given their significance in modern AI research, warrant a granular examination. Comparative studies should dissect various transformer architectures, probing their efficiency, flexibility, and computational intricacies in multiple application domains.

4. **Marrying AI and Human Expertise:**
   - Our research highlighted the importance of expert validation. As we advance, collaborations with domain authorities become pivotal. Such joint endeavours would ensure that AI outputs, while being technologically sound, also resonate with the nuances of domain-specific knowledge.

By pursuing these avenues, we believe the academic community can address the gaps highlighted in our research, ensuring that AI continues to evolve in a manner that's both transformative and grounded in real-world applicability.

References

Adadi, A., & Berrada, M. (2018). Peeking Inside the Black-Box: A Survey on Explainable Artificial Intelligence (XAI). *IEEE Access*, *6*, 52138–52160. https://doi.org/10.1109/ACCESS.2018.2870052

Alharbi, A. M. (2019). *Unsupervised Abstraction for Reducing the Complexity of Healthcare Process Models*.

Alison Kilburn. (2023). *UK Artificial Intelligence Regulation Impact Assessment*.

Alistair Johnson, Lucas Bulgarelli, Tom Pollard, Leo Anthony Celi, Roger Mark, & Steven Horng. (2023). MIMIC-IV-ED (version 2.2). *PhysioNet.* https://doi.org/https://doi.org/10.13026/5ntk-km72

Altmann, A., Toloşi, L., Sander, O., & Lengauer, T. (2010). Permutation importance: a corrected feature importance measure. *Bioinformatics*, *26*(10), 1340–1347. https://doi.org/10.1093/BIOINFORMATICS/BTQ134

Appice, A., Di Mauro, N., & Malerba, D. (2019). Leveraging shallow machine learning to predict business process behavior. *Proceedings - 2019 IEEE International Conference on Services Computing, SCC 2019 - Part of the 2019 IEEE World Congress on Services*, 184–188. https://doi.org/10.1109/SCC.2019.00039

Avanzi, B., Taylor, G., Wang, M., & Wong, B. (2023). *Machine Learning with High-Cardinality Categorical Features in Actuarial Applications*. https://arxiv.org/abs/2301.12710v1

Ba, J. L., Kiros, J. R., & Hinton, G. E. (2016). *Layer Normalization*. https://arxiv.org/abs/1607.06450v1

Bach, S., Binder, A., Montavon, G., Klauschen, F., Müller, K. R., & Samek, W. (2015). On Pixel-Wise Explanations for Non-Linear Classifier Decisions by Layer-Wise Relevance Propagation. *PLOS ONE*, *10*(7), e0130140. https://doi.org/10.1371/JOURNAL.PONE.0130140

Baier, L., Reimold, J., & Kuhl, N. (2020). Handling Concept Drift for Predictions in Business Process Mining. *Proceedings - 2020 IEEE 22nd Conference on Business Informatics, CBI 2020*, *1*, 76–83. https://doi.org/10.1109/CBI49978.2020.00016

Bathaee, Y. (2018). THE ARTIFICIAL INTELLIGENCE BLACK BOX AND THE FAILURE OF INTENT AND CAUSATION. *Harvard Journal of Law & Technology*, *31*(2). https://www.theverge.com/

Becker, A. J., & Bagrow, J. P. (2019). *UAFS: Uncertainty-Aware Feature Selection for Problems with Missing Data*. https://arxiv.org/abs/1904.01385v3

Boudewijn van Dongen. (2017). BPI Challenge 2017 Dataset. In *Version 1 of Dataset published 2017 : Vol. Version 1*. Eindhoven University of Technology.

Breiman, L. (1999). *RANDOM FORESTS-RANDOM FEATURES*.

Brown, T. B., Mann, B., Ryder, N., Subbiah, M., Kaplan, J., Dhariwal, P., Neelakantan, A., Shyam, P., Sastry, G., Askell, A., Agarwal, S., Herbert-Voss, A., Krueger, G., Henighan, T., Child, R., Ramesh, A., Ziegler, D. M., Wu, J., Winter, C., … Amodei, D. (2020). Language Models are Few-Shot Learners. *Advances in Neural Information Processing Systems*, *33*, 1877–1901. https://commoncrawl.org/the-data/

Bukhsh, Z. A., Saeed, A., & Dijkman, R. M. (2021). *ProcessTransformer: Predictive Business Process Monitoring with Transformer Network*. https://arxiv.org/abs/2104.00721v1

Camargo, M., Dumas, M., & González-Rojas, O. (2019a). Learning Accurate LSTM Models of Business Processes. *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, *11675 LNCS*, 286–302. https://doi.org/10.1007/978-3-030-26619-6_19/FIGURES/10

Camargo, M., Dumas, M., & González-Rojas, O. (2019b). Learning Accurate LSTM Models of Business Processes. *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, *11675 LNCS*, 286–302. https://doi.org/10.1007/978-3-030-26619-6_19

Cerda, P., & El Varoquaux, G. (2019). Encoding high-cardinality string categorical variables. *IEEE Transactions on Knowledge and Data Engineering*, *34*(3), 1164–1176. https://doi.org/10.1109/TKDE.2020.2992529

Cerda, P., Varoquaux, G., & Kégl, B. (2018). Similarity encoding for learning with dirty categorical variables. *Machine Learning*, *107*(8–10), 1477–1494. https://doi.org/10.1007/s10994-018-5724-2

Chawda, A., Grimm, S., & Kloft, M. (2022). *Unsupervised Anomaly Detection for Auditing Data and Impact of Categorical Encodings*. https://arxiv.org/abs/2210.14056v2

Chen, J., Song, L., Wainwright, M. J., & Jordan, M. I. (2018). *Learning to Explain: An Information-Theoretic Perspective on Model Interpretation* (pp. 883–892). PMLR. https://proceedings.mlr.press/v80/chen18j.html

Chen, T., & Guestrin, C. (2016). XGBoost: A Scalable Tree Boosting System. *Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, *13-17-August-2016*, 785–794. https://doi.org/10.1145/2939672.2939785

Dahouda, M. K., & Joe, I. (2021). A Deep-Learned Embedding Technique for Categorical Features Encoding. *IEEE Access*, *9*, 114381–114391. https://doi.org/10.1109/ACCESS.2021.3104357

Dakic, D., & Stefanovic, D. (2018). *Business Process Mining Application: A Literature Review*. 866–0875. https://doi.org/10.2507/29th.daaam.proceedings.125

Dallagassa, M. R., dos Santos Garcia, C., Scalabrin, E. E., Ioshii, S. O., & Carvalho, D. R. (2021). Opportunities and challenges for applying process mining in healthcare: a systematic mapping study. *Journal of Ambient Intelligence and Humanized Computing 2021 13:1*, *13*(1), 165–182. https://doi.org/10.1007/S12652-021-02894-7

Data Protection Act 2018. (n.d.). King's Printer of Acts of Parliament.

De Oliveira, H., Prodel, M., Lamarsalle, L., & Orlowski, A. (2019). *Process mining for predictive analytics: a case study on NHS data to improve care for sepsis patients*. https://doi.org/10.13140/RG.2.2.24809.65129

De Roock, E., & Martin, N. (2022). Process mining in healthcare - An updated perspective on the state of the art. *Journal of Biomedical Informatics*, *127*. https://doi.org/10.1016/J.JBI.2022.103995

Di Francescomarino, C., & Ghidini, C. (2022). Predictive Process Monitoring. *Lecture Notes in Business Information Processing*, *448*, 320–346. https://doi.org/10.1007/978-3-031-08848-3_10/FIGURES/14

Di Francescomarino, C., Ghidini, C., Maggi, F. M., Petrucci, G., & Yeshchenko, A. (2017). *An Eye into the Future: Leveraging A-priori Knowledge in Predictive Business Process Monitoring*. 252–268. https://doi.org/10.1007/978-3-319-65000-5_15

Díaz-Uriarte, R., & Alvarez de Andrés, S. (2006). Gene selection and classification of microarray data using random forest. *BMC Bioinformatics*, *7*(1), 1–13. https://doi.org/10.1186/1471-2105-7-3/FIGURES/1

Dieber, J., & Kirrane, S. (2020). *Why model why? Assessing the strengths and limitations of LIME*. http://arxiv.org/abs/2012.00093

Doshi-Velez, F., & Kim, B. (2017). *Towards A Rigorous Science of Interpretable Machine Learning*. https://arxiv.org/abs/1702.08608v2

Dumas, M., La Rosa, M., Mendling, J., & Reijers, H. A. (2018). Fundamentals of business process management: Second Edition. *Fundamentals of Business Process Management: Second Edition*, 1–527. https://doi.org/10.1007/978-3-662-56509-4/COVER

Eck, D., & Schmidhuber, J. (n.d.). *A First Look at Music Composition using LSTM Recurrent Neural Networks*.

Efr´, E., Rama-Maneiro, E., Vidal, J. C., & Lama, M. (n.d.). *Deep Learning for Predictive Business Process Monitoring: Review and Benchmark*.

Elkhawaga, G., Abuelkheir, M., Barakat, S. I., Riad, A. M., & Reichert, M. (2020). CONDA-PM -- A Systematic Review and Framework for Concept Drift Analysis in Process Mining. *Algorithms*, *13*(7), 161. https://doi.org/10.3390/a13070161

Elkhawaga, G., Abu-Elkheir, M., & Reichert, M. (2022). Explainability of Predictive Process Monitoring Results: Can You See My Data Issues? *Applied Sciences (Switzerland)*, *12*(16). https://doi.org/10.3390/app12168192

El-Khawaga, G., Abu-Elkheir, M., & Reichert, M. (2022). XAI in the Context of Predictive Process Monitoring: An Empirical Analysis Framework. *Algorithms 2022, Vol. 15, Page 199*, *15*(6), 199. https://doi.org/10.3390/A15060199

EU Commission. (2022). *General data protection regulation (GDPR)*. https://eur-lex.europa.eu/EN/legal-content/summary/general-data-protection-regulation-gdpr.html

Evermann, J., Rehse, J. R., & Fettke, P. (2016). A Deep Learning Approach for Predicting Process Behaviour at Runtime. *Lecture Notes in Business Information Processing*, *281*, 327–338. https://doi.org/10.1007/978-3-319-58457-7_24

Explainable Artificial Intelligence. (n.d.). Retrieved 11 October 2023, from https://www.darpa.mil/program/explainable-artificial-intelligence

Farquad, M. A. H., Ravi, V., & Raju, S. B. (2014). Churn prediction using comprehensible support vector machine: An analytical CRM application. *Appl. Soft Comput.*, *19*, 31–40. https://doi.org/10.1016/J.ASOC.2014.01.031

for Science, D. (n.d.). *Title: UK Artificial Intelligence Regulation Impact Assessment*.

Gebru, T., Morgenstern, J., Vecchione, B., Vaughan, J. W., Wallach, H., Iii, H. D., & Crawford, K. (2018). Datasheets for Datasets. *Communications of the ACM*, *64*(12), 86–92. https://doi.org/10.1145/3458723

Gers, F. A., Schraudolph, N. N., & Schmidhuber, J. (2002). Learning Precise Timing with LSTM Recurrent Networks. *Journal of Machine Learning Research*, *3*, 115–143. www.idsia.ch

Gomes, A. F. D., de Lacerda, A. C. W. G., & da Silva Fialho, J. R. (2021). Comparative Analysis of Process Mining Algorithms in Python. *Lecture Notes of the Institute for Computer Sciences, Social-Informatics and Telecommunications Engineering, LNICST*, *401 LNICST*, 27–43. https://doi.org/10.1007/978-3-030-91421-9_3/COVER

Graves, A. (2013). *Generating Sequences With Recurrent Neural Networks*. https://arxiv.org/abs/1308.0850v5

Graves, A., & Schmidhuber, J. (2005). Framewise phoneme classification with bidirectional LSTM and other neural network architectures. *Neural Networks : The Official Journal of the International Neural Network Society*, *18*(5–6), 602–610. https://doi.org/10.1016/J.NEUNET.2005.06.042

Grisold, T., Mendling, J., Otto, M., & vom Brocke, J. (2021). Adoption, use and management of process mining in practice. *Business Process Management Journal*, *27*(2), 369–387. https://doi.org/10.1108/BPMJ-03-2020-0112/FULL/PDF

Grisold, T., Wurm, B., Mendling, J., & Vom Brocke, J. (n.d.). *Using Process Mining to Support Theorizing About Change in Organizations*. Retrieved 15 October 2023, from https://hdl.handle.net/10125/64417

Guidotti, R., Monreale, A., Ruggieri, S., Turini, F., Giannotti, F., & Pedreschi, D. (2018). A Survey of Methods for Explaining Black Box Models. *ACM Computing Surveys (CSUR)*, *51*(5). https://doi.org/10.1145/3236009

Gunnarsson, B. R., vanden Broucke, S. K. L. M., & De Weerdt, J. (2019). Predictive Process Monitoring in Operational Logistics: A Case Study in Aviation. *Lecture Notes in Business Information Processing*, *362 LNBIP*, 250–262. https://doi.org/10.1007/978-3-030-37453-2_21/COVER

Günther, C., & Rozinat, A. (2012). Disco: Discover Your Processes. *International Conference on Business Process Management*.

Günther, C. W., & Van Der Aalst, W. M. P. (2007). Fuzzy Mining – Adaptive Process Simplification Based on Multi-perspective Metrics. *Lecture Notes in Computer Science*

(Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), *4714 LNCS*, 328–343. https://doi.org/10.1007/978-3-540-75183-0_24

Hanga, K. M., Kovalchuk, Y., & Gaber, M. M. (2020). A graph-based approach to interpreting recurrent neural networks in process mining. *IEEE Access*, *8*, 172923–172938. https://doi.org/10.1109/ACCESS.2020.3025999

Hassija, V., Chamola, V., Mahapatra, A., Singal, A., Goel, D., Huang, K., Scardapane, S., Spinelli, I., Mahmud, M., & Hussain, A. (2023). Interpreting Black-Box Models: A Review on Explainable Artificial Intelligence. *Cognitive Computation*, *1*, 1–30. https://doi.org/10.1007/S12559-023-10179-8/FIGURES/14

Heaton, J. (2018). Ian Goodfellow, Yoshua Bengio, and Aaron Courville: Deep learning. *Genetic Programming and Evolvable Machines*, *19*(1–2), 305–307. https://doi.org/10.1007/S10710-017-9314-Z

Hinkka, M., Lehto, T., & Heljanko, K. (2020). Exploiting event log event attributes in RNN based prediction. *Lecture Notes in Business Information Processing*, *379 LNBIP*, 67–85. https://doi.org/10.1007/978-3-030-46633-6_4/COVER

Huang, X., Khetan, A., Cvitkovic, M., & Karnin, Z. (2020). *TabTransformer: Tabular Data Modeling Using Contextual Embeddings*. https://arxiv.org/abs/2012.06678v1

Intayoad, W., & Becker, T. (2018). Applying Process Mining in Manufacturing and Logistic for Large Transaction Data. *Lecture Notes in Logistics*, 378–388. https://doi.org/10.1007/978-3-319-74225-0_51/COVER

Johnson, A. , B. L. , P. T. , C. L. A. , M. R. , & H. S. (2022). (n.d.). *MIMIC-IV-ED (version 2.0). PhysioNet.* Retrieved 18 November 2022, from https://physionet.org/content/mimic-iv-ed/2.0/

Jones, E. (2018). A posthuman-xenofeminist analysis of the discourse on autonomous weapons systems and other killing machines. *Australian Feminist Law Journal*, *44*(1), 93–118. https://doi.org/10.1080/13200968.2018.1465333

Jordan, M. I., & Mitchell, T. M. (2015). Machine learning: Trends, perspectives, and prospects. *Science (New York, N.Y.)*, *349*(6245), 255–260. https://doi.org/10.1126/SCIENCE.AAA8415

Kang, W.-C., Cheng, D. Z., Yao, T., Yi, X., Chen, T., Hong, L., & Chi, E. H. (2020). Learning to Embed Categorical Features without Embedding Tables for Recommendation.

Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, 11. https://doi.org/10.1145/3447548.3467304

Keith, B., & Vega, V. (2017). Process mining applications in software engineering. *Advances in Intelligent Systems and Computing*, *537*, 47–120. https://doi.org/10.1007/978-3-319-48523-2_5

Kieny, M. P., Bekedam, H., Dovlo, D., Fitzgerald, J., Habicht, J., Harrison, G., Kluge, H., Lin, V., Menabde, N., Mirza, Z., Siddiqi, S., & Travis, P. (2017). Strengthening health systems for universal health coverage and sustainable development. *Bulletin of the World Health Organization*, *95*(7), 537–539. https://doi.org/10.2471/BLT.16.187476

Kim, J., Comuzzi, M., Dumas, M., Maggi, F. M., & Teinemaa, I. (2022). Encoding resource experience for predictive process monitoring. *Decision Support Systems*, *153*, 113669. https://doi.org/10.1016/J.DSS.2021.113669

Kosaraju, N., Sankepally, S. R., & Mallikharjuna Rao, K. (2023). Categorical Data: Need, Encoding, Selection of Encoding Method and Its Emergence in Machine Learning Models—A Practical Review Study on Heart Disease Prediction Dataset Using Pearson Correlation. *Lecture Notes in Networks and Systems*, *551*, 369–382. https://doi.org/10.1007/978-981-19-6631-6_26/COVER

Kurniati, A. P., Hall, G., Hogg, D., & Johnson, O. (2018). Process mining in oncology using the MIMIC-III dataset. *Journal of Physics: Conference Series*, *971*(1), 012008. https://doi.org/10.1088/1742-6596/971/1/012008

Kurniati, A. P., Johnson, O., Hogg, D., & Hall, G. (2016). Process mining in oncology: A literature review. *Proceedings of the 6th International Conference on Information Communication and Management, ICICM 2016*, 291–297. https://doi.org/10.1109/INFOCOMAN.2016.7784260

Kusuma, G., Kurniati, A., McInerney, C. D., Hall, M., Gale, C. P., & Johnson, O. (2021). Process Mining of Disease Trajectories in MIMIC-III: A Case Study. *Lecture Notes in Business Information Processing*, *406 LNBIP*, 305–316. https://doi.org/10.1007/978-3-030-72693-5_23/TABLES/3

Lassen, K. B., & van Dongen, B. F. (2008). *Translating Message Sequence Charts to other Process Languages Using Process Mining*. 71–85. https://doi.org/10.1007/978-3-540-89287-8_5

Liaw, A., & Wiener, M. (2002). *Classification and Regression by randomForest*. *2*(3). http://www.stat.berkeley.edu/

Lin, L., Wen, L., & Wang, J. (n.d.). *MM-Pred: A Deep Predictive Model for Multi-attribute Event Sequence*. Retrieved 17 October 2023, from https://epubs.siam.org/terms-privacy

Lorenz, R., Senoner, J., Sihn, W., & Netland, T. (2021). Using process mining to improve productivity in make-to-stock manufacturing. *International Journal of Production Research*, *59*(16), 4869–4880. https://doi.org/10.1080/00207543.2021.1906460

Lu, H., Li, Y., Chen, M., Kim, H., & Serikawa, S. (2018). Brain Intelligence: Go beyond Artificial Intelligence. *Mobile Networks and Applications*, *23*(2), 368–375. https://doi.org/10.1007/S11036-017-0932-8/METRICS

Lundberg, S. M., Allen, P. G., & Lee, S.-I. (2017). A Unified Approach to Interpreting Model Predictions. *Advances in Neural Information Processing Systems*, *30*. https://github.com/slundberg/shap

Lundberg, S. M., Erion, G. G., & Lee, S.-I. (2018). *Consistent Individualized Feature Attribution for Tree Ensembles*. https://arxiv.org/abs/1802.03888v3

M. (Massimiliano) de Leoni, & Felix Mannhardt. (2015). Road Traffic Fine Management Process [Dataset]. In *Eindhoven University of Technology*.

Maas, A., Daly, R. E., Pham, P. T., Huang, D., Ng, A. Y., & Potts, C. (2011). *Learning Word Vectors for Sentiment Analysis* (pp. 142–150). https://aclanthology.org/P11-1015

Marquez-Chamorro, A. E., Resinas, M., & Ruiz-Cortes, A. (2018a). Predictive monitoring of business processes: A survey. *IEEE Transactions on Services Computing*, *11*(6), 962–977. https://doi.org/10.1109/TSC.2017.2772256

Marquez-Chamorro, A. E., Resinas, M., & Ruiz-Cortes, A. (2018b). Predictive Monitoring of Business Processes: A Survey. *IEEE Transactions on Services Computing*, *11*(6), 962–977. https://doi.org/10.1109/TSC.2017.2772256

Martin, N., Fischer, D. A., Kerpedzhiev, G. D., Goel, K., Sander, •, Leemans, J. J., Röglinger, M., Wil, •, Van Der Aalst, M. P., Dumas, M., La, M., Moe, R. •, Wynn, T., Martin, N., Fischer, D. A., Röglinger, Á. M., Röglinger, M., Kerpedzhiev, G. D., Goel, K., … Rosa, M. La. (2021). Opportunities and Challenges for Process Mining in Organizations: Results of a Delphi Study. *Business & Information Systems Engineering*, *63*(5), 511–527. https://doi.org/10.1007/s12599-021-00720-0

Martin, N., Wittig, N., & Munoz-Gama, J. (2022). Using Process Mining in Healthcare. *Lecture Notes in Business Information Processing*, *448*, 416–444. https://doi.org/10.1007/978-3-031-08848-3_14/FIGURES/2

Matt Davies, & Michael Birtwistle. (2023). *Regulating AI in the UK*. https://www.adalovelaceinstitute.org/policy-briefing/eu-ai-act/

Medeiros, A. K. A. de, de Medeiros, A. K. A., van Dongen, B. F., van der Aalst, W. M. P., & Weijters, A. J. M. M. (2004). Process mining: Extending the α-algorithm to mine short loops. *EINDHOVEN UNIVERSITY OF TECHNOLOGY, EINDHOVEN*. http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.5.7518

Michelle Donelan. (2023). *Policy paper «A pro-innovation approach to AI regulation»* . https://www.gov.uk/government/publications/ai-regulation-a-pro-innovation-approach/white-paper

Mikolov, T., Karafiát, M., Burget, L., Jan, C., & Khudanpur, S. (2010). Recurrent neural network based language model. *Proceedings of the 11th Annual Conference of the International Speech Communication Association, INTERSPEECH 2010*, 1045–1048. https://doi.org/10.21437/INTERSPEECH.2010-343

Mikolov, T., Sutskever, I., Chen, K., Corrado, G., & Dean, J. (2013). Distributed Representations of Words and Phrases and their Compositionality. *Advances in Neural Information Processing Systems*. https://arxiv.org/abs/1310.4546v1

Miranda Mourby. (2021, May 5). A New AI Regulation (& an End to the 'Black Box'?). *Https://Blogs.Law.Ox.Ac.Uk/*.

Mishra, V. P., Dsouza, J., & Elizabeth, L. (2018). Analysis and Comparison of Process Mining Algorithms with Application of Process Mining in Intrusion Detection System. *2018 7th International Conference on Reliability, Infocom Technologies and Optimization: Trends and Future Directions, ICRITO 2018*, 613–617. https://doi.org/10.1109/ICRITO.2018.8748748

Moeyersoms, J., & Martens, D. (2015). Including high-cardinality attributes in predictive models: A case study in churn prediction in the energy sector. *Decision Support Systems*, *72*, 72–81. https://doi.org/10.1016/J.DSS.2015.02.007

Mökander, J., Juneja, P., Watson, D. S., & Floridi, L. (2022). The US Algorithmic Accountability Act of 2022 vs. The EU Artificial Intelligence Act: what can they learn from each other?

Minds and Machines, *32*(4), 751–758. https://doi.org/10.1007/S11023-022-09612-Y/METRICS

Molnar, C. (2023). *Interpretable machine learning : a guide for making black box models explainable*.

Munoz-Gama, J., Martin, N., Fernandez-Llatas, C., Johnson, O. A., Sepúlveda, M., Helm, E., Galvez-Yanjari, V., Rojas, E., Martinez-Millana, A., Aloini, D., Amantea, I. A., Andrews, R., Arias, M., Beerepoot, I., Benevento, E., Burattin, A., Capurro, D., Carmona, J., Comuzzi, M., … Zerbato, F. (2022a). Process mining for healthcare: Characteristics and challenges. *Journal of Biomedical Informatics*, *127*, 103994. https://doi.org/10.1016/J.JBI.2022.103994

Munoz-Gama, J., Martin, N., Fernandez-Llatas, C., Johnson, O. A., Sepúlveda, M., Helm, E., Galvez-Yanjari, V., Rojas, E., Martinez-Millana, A., Aloini, D., Amantea, I. A., Andrews, R., Arias, M., Beerepoot, I., Benevento, E., Burattin, A., Capurro, D., Carmona, J., Comuzzi, M., … Zerbato, F. (2022b). Process mining for healthcare: Characteristics and challenges. *Journal of Biomedical Informatics*, *127*, 103994. https://doi.org/https://doi.org/10.1016/j.jbi.2022.103994

Narajewski, M., Kley-Holsteg, J., & Ziel, F. (2021). tsrobprep - an R package for robust preprocessing of time series data. *SoftwareX*, *16*. https://doi.org/10.1016/j.softx.2021.100809

Narwaria, M. (2021). *Does Explainable Machine Learning Uncover the Black Box in Vision Applications?* https://doi.org/10.1016/j.imavis.2021.104353

Oore, S., Simon, I., Dieleman, S., Eck, D., & Simonyan, K. (2018). This Time with Feeling: Learning Expressive Musical Performance. *Neural Computing and Applications*, *32*(4), 955–967. https://doi.org/10.1007/s00521-018-3758-9

Oswald, M. (2023). Legal framework. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, *376*(2128). https://doi.org/10.1098/RSTA.2017.0359

Padovan, P. H., Martins, C. M., & Reed, C. (2023). Black is the new orange: how to determine AI liability. *Artificial Intelligence and Law*, *31*(1), 133–167. https://doi.org/10.1007/S10506-022-09308-9/FIGURES/2

Pargent, F., Pfisterer, F., Thomas, J., & Bischl, B. (2022). Regularized target encoding outperforms traditional methods in supervised machine learning with high cardinality features. *Computational Statistics*, *37*(5), 2671–2692. https://doi.org/10.1007/S00180-022-01207-6/TABLES/1

Pasquadibisceglie, V., Appice, A., Castellano, G., Malerba, D., & Modugno, G. (2020). Orange: Outcome-oriented predictive process monitoring based on image encoding and CNNs. *IEEE Access*, *8*, 184073–184086. https://doi.org/10.1109/ACCESS.2020.3029323

Pasquale, F. (2015). The Black Box Society: The Secret Algorithms that Control Money and Information. *Book Gallery*. https://digitalcommons.law.umaryland.edu/books/96

Peleg, M. (2013). Computer-interpretable clinical guidelines: a methodological review. *Journal of Biomedical Informatics*, *46*(4), 744–763. https://doi.org/10.1016/J.JBI.2013.06.009

Qafari, M. S., & van der Aalst, W. (2019). Fairness-aware process mining. *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, *11877 LNCS*, 182–192. https://doi.org/10.1007/978-3-030-33246-4_11

Rama-Maneiro, E., Vidal, J. C., & Lama, M. (2020). Deep Learning for Predictive Business Process Monitoring: Review and Benchmark. *IEEE Transactions on Services Computing*, *16*(1), 739–756. https://doi.org/10.1109/TSC.2021.3139807

R'bigui, H., & Cho, C. (2017). The state-of-the-art of business process mining challenges. *International Journal of Business Process Integration and Management*, *8*(4), 285–303. https://doi.org/10.1504/IJBPIM.2017.088819

R'Bigui, H., & Cho, C. (2017). The state-of-the-art of business process mining challenges. *International Journal of Business Process Integration and Management*, *8*, 285. https://doi.org/10.1504/IJBPIM.2017.10009731

Rehman, A., Naz, S., & Razzak, I. (2021). Leveraging big data analytics in healthcare enhancement: trends, challenges and opportunities. *Multimedia Systems 2021 28:4*, *28*(4), 1339–1371. https://doi.org/10.1007/S00530-020-00736-8

Reinkemeyer, Lars. (2020). Process Mining in Action. *Process Mining in Action*. https://doi.org/10.1007/978-3-030-40172-6

Ribeiro, M. T., Singh, S., & Guestrin, C. (n.d.). *'Why Should I Trust You?' Explaining the Predictions of Any Classifier*. https://doi.org/10.1145/2939672.2939778

Ribeiro, M. T., Singh, S., & Guestrin, C. (2016). 'Why should i trust you?' Explaining the predictions of any classifier. *Proceedings of the ACM SIGKDD International Conference* *on Knowledge Discovery and Data Mining*, *13-17-August-2016*, 1135–1144. https://doi.org/10.1145/2939672.2939778

Rojas, E., Munoz-Gama, J., Sepúlveda, M., & Capurro, D. (2016a). Process mining in healthcare: A literature review. *Journal of Biomedical Informatics*, *61*, 224–236. https://doi.org/10.1016/J.JBI.2016.04.007

Rojas, E., Munoz-Gama, J., Sepúlveda, M., & Capurro, D. (2016b). Process mining in healthcare: A literature review. *Journal of Biomedical Informatics*, *61*, 224–236. https://doi.org/10.1016/J.JBI.2016.04.007

Rojas, E., Zhang, S., Kurniati, A. P., Hogg, D., Hall, G., & Johnson, O. (2018). The assessment of data quality issues for process mining in healthcare using Medical Information Mart for Intens... Cite this paper Related papers Ident ifying Sub-Phenot ypes of Acut e Kidney Injury using St ruct ured and Unst ruct ured Elect r... The Assessment of Data Quality Issues for Process Mining in Healthcare Using MIMIC-III, a Freely Available e-Health Record Database. *Health Informatics Journal*, *4*, 1878–1893. https://doi.org/10.1177/1460458218810760

Rozinat, A., & van der Aalst, W. M. P. (2008). Conformance checking of processes based on monitoring real behavior. *Information Systems*, *33*(1), 64–95. https://doi.org/10.1016/J.IS.2007.07.001

Sak, H. H., Senior, A., & Beaufays Google, F. (2014). *Long Short-Term Memory Based Recurrent Neural Network Architectures for Large Vocabulary Speech Recognition*. https://arxiv.org/abs/1402.1128v1

Samek, W., Wiegand, T., & Müller, K.-R. (2017). *Explainable Artificial Intelligence: Understanding, Visualizing and Interpreting Deep Learning Models*. https://arxiv.org/abs/1708.08296v1

Shankaranarayana, S. M., & Runje, D. (2021). Attention Augmented Convolutional Transformer for Tabular Time-series. *IEEE International Conference on Data Mining Workshops, ICDMW*, *2021-December*, 537–541. https://doi.org/10.1109/ICDMW53433.2021.00071

Shrikumar, A., Greenside, P., & Kundaje, A. (2017). *Learning Important Features Through Propagating Activation Differences* (pp. 3145–3153). PMLR. https://proceedings.mlr.press/v70/shrikumar17a.html

Smuha, N. A. (2019). The EU Approach to Ethics Guidelines for Trustworthy Artificial Intelligence. *Computer Law Review International*, *20*(4), 97–106. https://doi.org/10.9785/CRI-2019-200402

Sokol, K., & Flach, P. (2020). Explainability fact sheets: A framework for systematic assessment of explainable approaches. *FAT* 2020 - Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency*, 56–67. https://doi.org/10.1145/3351095.3372870

Srivastava, N., Hinton, G., Krizhevsky, A., Sutskever, I., & Salakhutdinov, R. (2014). Dropout: A Simple Way to Prevent Neural Networks from Overfitting. *Journal of Machine Learning Research*, *15*(56), 1929–1958. http://jmlr.org/papers/v15/srivastava14a.html

Tama, B. A., & Comuzzi, M. (2019). An empirical comparison of classification techniques for next event prediction using business process event logs. *Expert Systems With Applications*, *129*, 233–245. https://doi.org/10.1016/j.eswa.2019.04.016

Tang, D., Qin, B., & Liu, T. (2015). Document Modeling with Gated Recurrent Neural Network for Sentiment Classification. *Conference Proceedings - EMNLP 2015: Conference on Empirical Methods in Natural Language Processing*, 1422–1432. https://doi.org/10.18653/V1/D15-1167

Tax, N., Verenich, I., La Rosa, M., & Dumas, M. (2016a). Predictive Business Process Monitoring with LSTM Neural Networks. *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, *10253 LNCS*, 477–492. https://doi.org/10.1007/978-3-319-59536-8_30

Tax, N., Verenich, I., La Rosa, M., & Dumas, M. (2016b). Predictive Business Process Monitoring with LSTM Neural Networks. *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, *10253 LNCS*, 477–492. https://doi.org/10.1007/978-3-319-59536-8_30

Taymouri, F., Rosa, M. La, Erfani, S., Bozorgi, Z. D., & Verenich, I. (2020). Predictive Business Process Monitoring via Generative Adversarial Nets: The Case of Next Event Prediction. *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, *12168 LNCS*, 237–256. https://doi.org/10.1007/978-3-030-58666-9_14

Tello-Leal, E., Roa, J., Rubiolo, M., & Ramirez-Alcocer, U. M. (2018). Predicting activities in business processes with LSTM recurrent neural networks. *10th ITU Academic Conference*

Kaleidoscope: Machine Learning for a 5G Future, ITU K 2018. https://doi.org/10.23919/ITU-WT.2018.8598069

Theis, J., & Darabi, H. (2019). Decay Replay Mining to Predict Next Process Events. *IEEE Access*, *7*, 119787–119803. https://doi.org/10.1109/ACCESS.2019.2937085

Valdez-Valenzuela, E., Kuri-Morales, A., & Gomez-Adorno, H. (2021). Measuring the Effect of Categorical Encoders in Machine Learning Tasks Using Synthetic Data. *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, *13067 LNAI*, 92–107. https://doi.org/10.1007/978-3-030-89817-5_7/TABLES/20

Van Der Aalst, W. (2012). Process mining: Overview and opportunities. *ACM Transactions on Management Information Systems*, *3*(2). https://doi.org/10.1145/2229156.2229157

Van der Aalst, W. (2016). Process mining: Data science in action. *Process Mining: Data Science in Action*, 1–467. https://doi.org/10.1007/978-3-662-49851-4/COVER

Van der Aalst, W., Adriansyah, A., & Van Dongen, B. (2012). Replaying history on process models for conformance checking and performance analysis. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, *2*(2), 182–192. https://doi.org/10.1002/WIDM.1045

van der Aalst, W. M. P. (2009). Process Mining. *Encyclopedia of Database Systems*, 2171–2173. https://doi.org/10.1007/978-0-387-39940-9_1477

van der Aalst, W. M. P. (2013). Business Process Management: A Comprehensive Survey. *ISRN Software Engineering*, *2013*, 1–37. https://doi.org/10.1155/2013/507984

Van Der Aalst, W., Weijters, T., & Maruster, L. (2004). Workflow mining: Discovering process models from event logs. *IEEE Transactions on Knowledge and Data Engineering*, *16*(9), 1128–1142. https://doi.org/10.1109/TKDE.2004.47

Van Dongen, B. F., Alves De Medeiros, A. K., & Wen, L. (2009). Process mining: Overview and outlook of Petri net discovery algorithms. *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, *5460 LNCS*, 225–242. https://doi.org/10.1007/978-3-642-00899-3_13/COVER

Van Eck, M. L., Lu, X., Leemans, S. J. J., & Van Der Aalst, W. M. P. (2015). PM2: A process mining project methodology. *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, *9097*, 297–313. https://doi.org/10.1007/978-3-319-19069-3_19/COVER

Van Houdt, G., Mosquera, C., & Nápoles, G. (2020). A review on the long short-term memory model. *Artificial Intelligence Review 2020 53:8*, *53*(8), 5929–5955. https://doi.org/10.1007/S10462-020-09838-1

Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, Ł., & Polosukhin, I. (2017). Attention Is All You Need. *Advances in Neural Information Processing Systems*, *2017-December*, 5999–6009. https://arxiv.org/abs/1706.03762v7

Velioglu, R., Göpfert, J. P., Artelt, A., & Hammer, B. (2022). Explainable Artificial Intelligence for Improved Modeling of Processes. *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, *13756 LNCS*, 313–325. https://doi.org/10.1007/978-3-031-21753-1_31

Verbeek, H. M. W., Buijs, J. C. A. M., Van Dongen, B. F., & Van Der Aalst, W. M. P. (2011). XES, XESame, and ProM 6. *Lecture Notes in Business Information Processing*, *72 LNBIP*, 60–75. https://doi.org/10.1007/978-3-642-17722-4_5/COVER

Vinyals, O., Toshev, A., Bengio, S., & Erhan, D. (2014). Show and Tell: A Neural Image Caption Generator. *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, *07-12-June-2015*, 3156–3164. https://doi.org/10.1109/CVPR.2015.7298935

vom Brocke, J., van der Aalst, W. M., Grisold, T., Kremser, W., Mendling, J., Pentland, B., Recker, J., Roeglinger, M., Rosemann, M., & Weber, B. (2021). Process Science: The Interdisciplinary Study of Continuous Change. *SSRN Electronic Journal*. https://doi.org/10.2139/SSRN.3916817

Wang, F., & Rudin, C. (2015). *Falling Rule Lists* (pp. 1013–1022). PMLR. https://proceedings.mlr.press/v38/wang15a.html

Weber, P., Bordbar, B., & Tiño, P. (2013). A framework for the analysis of process mining algorithms. *IEEE Transactions on Systems, Man, and Cybernetics Part A:Systems and Humans*, *43*(2), 303–317. https://doi.org/10.1109/TSMCA.2012.2195169

Weidinger, L., Uesato, J., Rauh, M., Griffin, C., Huang, P. Sen, Mellor, J., Glaese, A., Cheng, M., Balle, B., Kasirzadeh, A., Biles, C., Brown, S., Kenton, Z., Hawkins, W., Stepleton, T., Birhane, A., Hendricks, L. A., Rimell, L., Isaac, W., … Gabriel, I. (2022). Taxonomy of Risks posed by Language Models. *ACM International Conference Proceeding Series*, 214–229. https://doi.org/10.1145/3531146.3533088

Weijters, A. J. M. M., & Ribeiro, J. T. S. (2011). Flexible Heuristics Miner (FHM). *2011 IEEE Symposium on Computational Intelligence and Data Mining (CIDM)*, 310–317. https://doi.org/10.1109/CIDM.2011.5949453

Werner, M. (2017). Financial process mining - Accounting data structure dependent control flow inference. *International Journal of Accounting Information Systems*, *25*, 57–80. https://doi.org/10.1016/J.ACCINF.2017.03.004

Werner, M., Wiese, M., & Maas, A. (2021). Embedding process mining into financial statement audits. *International Journal of Accounting Information Systems*, *41*, 100514. https://doi.org/10.1016/J.ACCINF.2021.100514

Weytjens, H., & De Weerdt, J. (2020). Process Outcome Prediction: CNN vs. LSTM (with Attention). *Lecture Notes in Business Information Processing*, *397*, 321–333. https://doi.org/10.1007/978-3-030-66498-5_24/COVER

Wickramanayake, B., He, Z., Ouyang, C., Moreira, C., Xu, Y., & Sindhgatta, R. (2022). Building interpretable models for business process prediction using shared and specialised attention mechanisms. *Knowledge-Based Systems*, *248*, 108773. https://doi.org/10.1016/J.KNOSYS.2022.108773

Wickramanayake, B., Ouyang, C., Xu, Y., & Moreira, C. (2023). Generating multi-level explanations for process outcome predictions ☆. *Engineering Applications of Artificial Intelligence*, *125*, 106678. https://doi.org/10.1016/j.engappai.2023.106678

Zerbino, P., Aloini, D., Dulmin, R., & Mininno, V. (2018). Process-mining-enabled audit of information systems: Methodology and an application. *Expert Systems with Applications*, *110*, 80–92. https://doi.org/10.1016/J.ESWA.2018.05.030

Zhang, Q., Lu, H., Sak, H., Tripathi, A., McDermott, E., Koo, S., & Kumar, S. (2020). Transformer Transducer: A Streamable Speech Recognition Model with Transformer Encoders and RNN-T Loss. *ICASSP, IEEE International Conference on Acoustics, Speech and Signal Processing - Proceedings*, *2020-May*, 7829–7833. https://doi.org/10.1109/ICASSP40776.2020.9053896

Zhao, W., Chen, J. J., Perkins, R., Liu, Z., Ge, W., Ding, Y., & Zou, W. (2015). A heuristic approach to determine an appropriate number of topics in topic modeling. *BMC Bioinformatics*, *16*(13), 1–10. https://doi.org/10.1186/1471-2105-16-S13-S8/FIGURES/6