Sheffield Hallam University | College of Business, Technology and Engineering

MSc Dissertation Report

"Prediction of Brake Failure in Scania Trucks using Neural Networks”.

A dissertation submitted in partial fulfilment of the requirements of Sheffield Hallam University for the degree of Master of Science in **Big Data Analytics**

Student Name	Madhankumar Duraisamy
Student ID	C2070486
Supervisor	Maria Luisa Davila Garcia
Date of Submission	23/01/2023 (Extended Deadline)

INCLUDE EITHER STATEMENT 1 **OR** STATEMENT 2:

STATEMENT 1:

This dissertation does NOT contain confidential material and thus can be made available to staff and students via the library.

STATEMENT 2:

This dissertation **IS CONFIDENTIAL** and circulation should be restricted to those involved in its assessment only.

ACKNOWLEDGEMENT

I want to express my heartfelt gratitude and thanks to my supervisor Dr. Maria Luisa, Davila Garcia for invaluable guidance and support throughout my dissertation journey. Her insights and feedback given greatly contributed to the quality of my work. Our weekly meetings were essential in keeping me focused and on schedule. I deeply appreciate your guidance and the additional tips you provided to enhance my project.

I'm also grateful to my classmates for their encouragement and positive spirit, especially during the challenging final stages of my dissertation. Their support was a key factor in maintaining my motivation.

Lastly, a huge thank you to my family and partner for their unwavering support during the tough and stressful times. Their presence and encouragement were crucial in helping me deliver my best work.

Sincere appreciations from the bottom of my heart.

ABSTRACT

This dissertation presents an in-depth investigation into the development and application of neural network models for predicting brake failures in Scania trucks. The study begins with a comprehensive analysis of the importance of brake systems in heavy-duty trucks and the potential of neural networks for predictive maintenance. Extensive literature review explores the underlying causes of brake failures and the role of neural networks in various predictive applications, identifying a significant research gap in brake failure prediction for heavy-duty trucks.

The research methodology adopts a quantitative approach, utilizing data from the UCI Machine Learning repository and employing the CRISP-DM methodology for neural network model development. Three neural network models - Multi-Layer Perceptron, 1D Convolutional Neural Network, and Recurrent Neural Networks with Long Short-Term Memory - are developed and rigorously tested. The Multi-Layer Perceptron demonstrates considerable accuracy, the 1D CNN shows promising results in handling time-dependent characteristics of the data, and the RNN-LSTM excels in capturing long-term dependencies crucial for pattern recognition in brake failure scenarios. The results demonstrate the models' varying degrees of efficacy in accurately predicting brake failures, providing vital insights into their practical application in the field of heavy-duty truck maintenance.

The discussion and analysis of results compare these models within the broader context of predictive maintenance for heavy-duty trucks. The study concludes with a synthesis of findings, addressing limitations and outlining future research directions. This research makes a significant contribution to the field of predictive maintenance for heavy-duty trucks, demonstrating the potential of neural networks in enhancing safety and operational efficiency. The models developed offer valuable insights into the practical application of advanced machine learning techniques in a critical aspect of vehicle safety.

Table of Contents

LIST OF FIGURES:

LIST OF TABLES

CHAPTER 1: INTRODUCTION

1.1 Research Background

Brakes are considered to be the most essential component that ensures safety by giving the vehicles the ability to decelerate or slow down when needed. Brakes in any particular vehicle must operate dependably since it is critical to ensuring the safety of the drivers and other people and properties on or near the roads (Rawat, 2020).
Since heavy-duty trucks, like those made by Scania, are essential to the global movement of commodities and resources, the trucking sector in particular depends largely on the reliable functioning of braking systems (Gondek et al., 2016). Therefore, it is crucial to keep the functionality of the braking system intact since brake failure can have disastrous results, including injuries, crashes, and possibly fatalities. Faults in braking systems can result due to several reasons, such as issues concerning hydraulic systems and/or any types of electrical or mechanical wear and tear (Kafunah et al., 2021). The implications of a brake system malfunction can be severe, and such breakdowns might be unpredictable.

In the modern era of technological developments, especially in the area of artificial intelligence, it becomes possible to better forecast and avoid problems in heavy-duty truck braking systems (Oh & Lee, 2020). A particular type of artificial intelligence called neural networks has shown remarkable potential for decision assistance and predictive modelling across a range of applications. The development of neural networks to forecast braking problems in Scania trucks is the main subject of the present dissertation.

Furthermore, tragic incidents are possible outcomes of commercial vehicles' brake failure, which endangers road safety. Commercial trucking fleets may be kept in good repair and dependability with the use of early warning systems that can identify and forecast when brakes will fail. Results from using AI methods, especially neural networks, for predictive maintenance have been encouraging as of late (Raveendran et al. 2018). Training and pattern recognition are capabilities of neural networks, which are computer models modelled after the anatomy and physiology of the human brain. It is feasible to create a prediction model that can detect patterns suggesting brake failure before to its occurrence by building a neural network model using past data linked to truck braking performance. Arena et al. (2022) found that this might lead to proactive maintenance measures, which in turn could improve trucking activities' security and effectiveness. Neural networks have been used for predictive maintenance in several fields with successful results. One example is the work of Tercan and Meisen (2022), who used a deep neural network to forecast when industrial equipment would break down, proving that these kinds of models can cut down on wasted time and money. In a similar vein, Wu et al. (2019) suggested a method based on recurrent neural networks for forecasting aviation engine failures, which boosted security and enhanced maintenance scheduling.

Predictive maintenance with neural networks may be applied to Scania trucks by utilising data gathered from many vehicle sensors, including temperature, wear indications, and breaks. Unusual patterns of temperature, pressure variations, or anomalous braking performance are some of the characteristics that a neural network model might learn to recognise through long-term analysis of these sensor information. Following training, the model can keep an eye on data from sensors in real time and send out alarms or warnings if it sees something that might cause the brakes to fail. There are several benefits of using neural networks to forecast when Scania trucks' brakes would fail. Primarily, it facilitates pre-emptive maintenance planning, enabling the arrangement of fixes or substitutions before to a significant malfunction. Additionally, it reduces the likelihood of accidents triggered by brake failure, which is a major improvement to road safety. Finally, it helps save money by reducing unscheduled downtime and optimising repair and maintenance activities (Rawat, 2020). Therefore, there is significant potential to optimise maintenance procedures and increase road safety through the use of neural networks to forecast when Scania trucks' brakes would fail. These models may learn to interpret patterns in sensor data that indicate when the brakes are about to fail, allowing for preventative maintenance to be performed at the last possible moment.

1.2 Research Question

The fundamental questions which have been addressed throughout this research are:

How neural network model can be used to predict brake failure in Scania trucks?

How does the neural network model effectively prevent unnecessary breakdowns, and optimize operational efficiency in the context of the trucking industry?

1.3 Research Aim

This research aims to develop a cutting-edge neural network model capable of accurately classifying and predicting brake failures in Scania trucks before they lead to failure, thereby enhancing vehicular safety, reducing downtime, and potentially saving lives by preempting on-road breakdowns and accidents.

1.4 Research Objectives:

Table 1 Research Objectives and Measures

S.No	Objective	Measure
1	Initiate the research by conducting an exhaustive literature review to understand brake failure and also the current state of neural network applications in predictive maintenance, particularly within the heavy vehicle sector.	Completion of a comprehensive literature review document that covers the specified topics
2	Gather a detailed collection of data points relating to brake parameters and historical failure incidents from Scania trucks, ensuring a mix of both normal operation and failure cases for a balanced study.	Collection of datasets of brake parameters and historical failure incidents from relevant sources
3	Engage in meticulous data preparation and exploration, employing techniques to clean, normalize, and analyze the data, thereby setting the stage for effective machine learning processes and gaining insights into the underlying patterns and correlations	Data preprocessing and exploratory analysis completed with a report on data quality, statistical summaries, and identified patterns and correlations ready for modelling.
4	Design and implement three neural network models, and carry out the training process using the training subset. Validate the performance of each model with the validation subset to ensure accuracy and reliability before external testing.	Three developed neural network models with documented training and validation result including performance metrics.
5	Examine all three neural network models to an evaluation against a separate test dataset, representative of real-world operational data, to assess their predictive capabilities.	Test Completion and documentation detailing the performance metrics of the three models against the test dataset.
6	Conduct an in-depth comparison of the three models post-external testing, analyzing their performance to identify the model with the highest predictive proficiency and operational applicability.	Completion of comparative analysis of three neural network models outlining the performance and applicability of each model, with a selection of the best-performing model.
7	Critically analyze the selected model’s results, draw conclusions on its efficacy in operational environments, and formulate recommendations for its implementation and further research.	A final report summarizing the chosen model's efficacy, with a set of practical recommendations for implementation and future research directions.

1.5 Key Deliverables

Development of reliable, accurate and effective predictive model by using neural network to detect brake failures in Scania trucks, along with comprehensive recommendations for its implementation.

1.6 Scope and Significance of the Research

The prediction of braking system failures in Scania trucks is the exclusive subject of this study, which also considers potential contributing variables. The creation and use of a prediction model based on neural networks are included in the scope. The goal of the study is to develop a model for averting accidents due to brake failure problems before they happen, rather than delving into their repair or rehabilitation.

Wide-ranging effects may result from the effective creation and application of a neural network-based prediction model for braking system failures in Scania vehicles. This research may contribute to a decrease in accidents, injuries, and deaths as well as an improvement in the commercial viability of trucking processes by lowering interruptions and by strengthening the safety and dependability of heavy-duty vehicles.

This study help to improve safety and dependability in other fields, as well as a contribution to the broader topic of predictive maintenance in the automobile industry.

1.7 Research Structure

The current research has been categorized into different chapters, each of which focuses on a distinct area of investigation. The overview of the subsequent chapters has been briefly discussed below:

Chapter 1: Introduction

This chapter sets the stage for the study, highlighting the critical role of brake systems in heavy-duty trucks and the potential of neural networks in predicting brake failures. It outlines the key research questions, aims, and objectives, focusing on the development of a neural network model to enhance safety and reduce downtime in Scania trucks.

Chapter 2: Literature Review

Literature review critically examines existing research on brake failures in heavy-duty trucks and the application of neural networks in predictive maintenance. It also discusses the causes, impacts, and technological solutions for brake failures, highlighting the role of neural networks in analyzing data and predicting equipment failure. This chapter identifies gaps in current research, particularly in the application of neural networks for predicting brake system failures in heavy-duty trucks like Scania.

Chapter 3: Research Design

This chapter outlines the research design and methodologies adopted for the study. It also discusses the application of the CRISP-DM methodology for neural network model development, focusing on data collection and the use of various data analysis techniques.

Chapter 4: Model Development and Results

Describes the development of three neural network models . It details the implementation of the CRISP-DM methodology, covering data preparation, model building, and evaluation. This chapter presents the results obtained from these models, analyzing their performance in predicting brake failures.

Chapter 5: Discussion and Analysis

This chapter discusses the results obtained from the neural network models, analyzes their implications, and assesses their efficacy in real-world scenarios. It offers a critical evaluation of each model's performance and its practical applications in the context of predictive maintenance for heavy-duty trucks.

Chapter 6: Conclusion and Future Work

The final chapter summarizes the research findings, discuss the limitations of the study, and suggest directions for future research. It concludes on the potential of neural networks in enhancing safety and operational efficiency in the trucking industry, particularly in predicting brake failures in heavy-duty trucks.

CHAPTER 2: LITERATURE REVIEW

2.1 Introduction

The transportation industry is of greater importance as it plays a major role in the global economy. It helps in goods movement and also people in vast networks. In this complicated environment, the reliability and safety of heavy-duty trucks are considered critical. Brake failure performed in commercial vehicles is increasing at a greater pace and poses a major threat to the safety of the road (Patrick & Andre, 2020). There is a need for certain innovative approaches to monitor and prevent such failures. Road safety is a major concern for industry stakeholders and also the public. Brake failure in heavy-duty trucks could result in drastic results which include accidents, economic loss, and injuries. Researchers and engineers are turning their attention increasingly to advanced technologies like neural networks by recognizing the centrality of braking systems to ensure safer and more reliable transportation. The development of predictive strategies for maintenance could help in solving risks linked with brake failure (Ak & et. al., 2013). In literature review aims to predict brake failure in Scania Trucks using Neural Networks.

2.2 Brake failure in heavy duty trucks

Brake failure in heavyProv-duty trucks is primarily caused by overheating during prolonged braking, especially on steep downgrades, leading to brake fade and failure (Bowman & Coleman, 1989). The mechanical condition of trucks, particularly defects in the brake system, significantly increases the likelihood of crash involvement (Blower, Green, & Matteson, 2010). Furthermore, the design and material of brake discs are crucial factors affecting their longevity and effectiveness, impacting the safety and reliability of these vehicles (Collignon et al., 2013).

The failure of brakes in heavy-duty trucks leads to severe safety risks, including increased chances of crashes involving drivers, passengers, and pedestrians (Ramarathnam et al., 2009). These incidents often result in significant economic losses due to damage to vehicles, goods, and infrastructure, in addition to the environmental impact, particularly when hazardous cargo is involved, leading to contamination and costly cleanup procedures (Kim & Jeong, 2013); (Rapanić, 1996).

Preventing brake failure in heavy-duty trucks involves routine maintenance and inspections to ensure the braking system is in optimal condition (Kim & Jeong, 2013). Driver education on safe braking procedures and recognizing early warning signs is also essential. Additionally, the implementation of advanced technologies like anti-lock braking systems (ABS) and electronic stability control (ESC) enhances braking performance and reliability, thereby reducing the likelihood of failure (Tadesse & Lei, 2020); (Scott et al., 2003).

2.3 Neural networks in predictive maintenances

As per Agarwal & et al. (2016), Neural networks are a subsection of machine learning that is designed to review patterns, make intelligent decisions, and learn through data (Limpert, 2019). It serves as an advanced computational model that possesses the power of machine learning to review equipment failure and optimize the schedule of maintenance. Such networks are effective in analyzing larger data sets linked to industrial machinery performance (Dhar, 2010). Such a practice stance permits early identification of patterns through norms which allows mitigating issues before they escalate into a major failure.

In addition to that, Dhar(2010) stated that the application of neural networks is of greater importance as it is suitable for predictive maintenance. As commercial vehicles have to face diverse driving condition and different loads, the ability of neural networks to adapt and review changing patterns helps in enhancing its effectiveness in predicting failure mainly to every context. Such adaptability is mainly important in the transportation field which possesses higher operational demands upon vehicles (Prakash, Patil & Kalyani, 2013). The application of neural networks for predictive maintenance could lead to a decline in maintenance costs and downtime. Through enabling interventions timely linked to data-driven predictions, companies need to perform maintenance activities effectively decline disruptions for operations, and optimize critical components lifespans (Agarwal & et al., 2016). Such approaches not only help in vehicle safety but even improve the effectiveness of costs for strategies of maintenance.

Furthermore, the capacity of predictive maintenance to optimise maintenance times and decrease interruption has led to its considerable interest in industrial applications. The use of neural networks for the analysis of sensor data and the prediction of breakdowns in equipment has been extensively used in predictive maintenance programmes. In this context, according to Zhang et al. (2019), the Long Short-Term Memory (LSTM) network is a well-liked neural network design that is employed for predictive maintenance. Periodic analysis is an ideal match for LSTM networks because of its ability to capture temporal relationships in consecutive sensor data. Using sensor data, LSTM networks were used to estimate the gears' remaining usable life. High precision for forecasting was attained by the authors, allowing for prompt maintenance steps to be conducted prior to bearing problems. A Convolutional Neural Network

(CNN) is another type of neural networks that Silva and Capretz (2019) state is utilised for predictive maintenance. Extracting spatial characteristics from spectrograms or sensor pictures is a breeze using CNNs. By spotting outliers, the CNN allowed for preventative maintenance to head off disastrous breakdowns. Serradilla et al. (2020) also reports that hybrid neural network designs have been created specifically for predictive maintenance. Predictive maintenance has therefore found a useful use for neural networks, such as LSTM, CNN, and hybrid designs. They make it possible to anticipate failures with reasonable accuracy and to implement preventative maintenance plans. This research shows that neural networks can be used for predictive maintenance in many different ways.

To summarise, the application of neural networks upon predictive maintenance presents a major shift in the reliability of industry equipment. Taking commercial vehicles into consideration, the usage of neural networks helps increase efficiency and safety of operation by giving accurate predictions of failures with implications that extend to the particular context of brake systems in Scania trucks (Ljung & Glad, 2016). As technology and research have continued to evolve, the neural network's role in predictive maintenance is poised to be integral increasingly for the optimization of industrial processes and also the assurance of standards of safety

2.3.1 Application of artificial neural networks for the prediction of performance and exhaust emissions

The study by Kiani et al. (2010) focused on using artificial neural networks (ANN) to predict engine performance and exhaust emissions in a spark ignition engine fuelled by ethanol-gasoline blends. The research involved training the ANN with data from a four-cylinder, four-stroke engine operated with different ethanol-gasoline blend ratios and at various engine speeds and loads. The model, based on a standard back-propagation algorithm, showed high accuracy in predicting engine parameters such as brake power, output torque, and emission indices for CO, CO2, HC, and NOx, demonstrating correlation coefficients ranging from 0.71 to 0.99. This result highlights the potential of ANN in accurately modelling engine performance and emissions, indicating its effectiveness in automotive applications.

However, the study's focus on ethanol-gasoline blends and a specific engine type suggests a research gap for other fuel types and engine models, including heavy-duty vehicles like Scania trucks. Moreover, while the ANN model effectively predicted engine performance, it did not address brake failure prediction, which is crucial for vehicle safety. The research, conducted in a controlled environment, leaves room for real-world application and testing, including integration with vehicle diagnostic systems. Future research could explore adapting these ANN methodologies for a broader range of vehicles and specifically focus on predicting brake system performance and potential failures, thereby enhancing predictive maintenance and safety in the automotive sector.

2.3.2 Application of artificial neural networks Prediction Models for Defective Sensor:

A. Shekar et al. (2018) discussed about Artificial Neural Network (ANN) framework that developed to predict the health of automotive components, particularly focusing on brake pads. This approach was designed to address the challenges posed by noisy and defective sensor data, which are common in complex dynamic systems like automobiles. The methodology involved using high-dimensional data from a variety of sensors, ensuring a comprehensive analysis of component health. To enhance the model's accuracy in the presence of noisy signals, the study introduced a data augmentation approach. This method enabled the ANN to make reliable predictions despite the quality issues often found in sensor data.

The findings of the study were significant for the field of predictive maintenance in vehicles. The ANN model demonstrated effective prediction capabilities even when faced with noisy or faulty sensor data, highlighting its robustness. It could compensate for the failure of individual sensors by leveraging data from an array of sensors, thus enhancing the reliability of its predictions. The practical application of this model was validated in an industrial setting at Robert Bosch GmbH, proving its effectiveness in real-world scenarios. However, the research had certain limitations. It did not specifically focus on the prediction of brake malfunctions in truck models such as Scania, nor did it extensively cover a wide range of vehicle models. Additionally, the study did not provide extensive data on the long-term reliability and performance of the ANN framework under diverse and changing real-world conditions, indicating areas for further research.

2.3.3 Application of Machine Learning Approaches in Intelligent Braking Systems

Jianhao Zhou et al. (2019) employed a composite machine learning approach to predict driver brake intention and intensity. This research collected various driving data from the Controller Area Network (CAN) bus in real driving scenarios, primarily in urban and rural settings. For preprocessing, the ReliefF and RReliefF algorithms were used for feature subset selection. The data, which might include mislabelled instances, was trained using a random forest algorithm, an ensemble machine learning method. Additionally, the study utilized an open-loop nonlinear autoregressive with external input (NARX) network for the online recognition and prediction of brake intensity, demonstrating its capability to learn long-term dependencies in the data.

The key findings of the study were quite significant in the context of predictive driver behavior. It was discovered that brake intention could be accurately predicted 0.5 seconds before the actual braking event. Moreover, the NARX network showed proficiency in predicting brake intensity 1 second in advance, indicating its potential in real-time application scenarios. These findings highlight the system's redundancy and fault tolerance, suggesting its suitability in scenarios involving sensor failure or loss of CAN messages. However, the research primarily focused on the analysis of driver behavior and did not delve into the mechanical aspects of brake malfunction or predictive maintenance. This marks a notable gap, as the study does not address the direct prediction or prevention of mechanical brake malfunctions, a crucial aspect for vehicle safety and maintenance.

2.4 Research Gap

2.4.1 Limited Application in Heavy Vehicles

Vodovozov et al. (2021) represents a notable trend in current research where the development of neural network-based control systems primarily focuses on green energy vehicles rather than heavy-duty trucks. This study showcases the advancement in neural network applications for controlling blended braking systems, addressing both intensive and gradual deceleration scenarios. Their approach includes a torque gradient control without a tire model, offering energy recovery during braking processes on varying road surfaces. The results from their research demonstrate the effectiveness of neural networks in adapting to dynamic behaviours, which could not be fully captured by traditional physics-based models. This advancement is significant as it points towards more efficient, responsive, and adaptive braking systems in vehicles.

However, this research also highlights a significant gap in the application of neural networks in the automotive sector, particularly concerning heavy-duty trucks. While the study's outcomes are promising for green energy vehicles, their direct application to heavy-duty trucks, which have distinct operational and mechanical characteristics, remains unexplored. Heavy-duty trucks, owing to their size, weight, and different usage patterns, present unique challenges that require tailored neural network solutions. This gap in research underscores the need for focused investigations into neural network applications that cater specifically to the complex requirements of heavy-duty truck braking systems. Addressing this gap could lead to enhanced safety, efficiency, and predictive maintenance capabilities in a sector that plays a crucial role in global logistics and transportation.

2.4.2 Complexity in Braking Systems of Large Vehicles

Dobaj (2022) demonstrated valuable insights into the use of neural networks for vehicle braking system diagnostics. It highlights the potential of neural networks in understanding and predicting the behavior of braking systems, which is crucial for ensuring vehicle safety and reliability. However, the study primarily focuses on general vehicle braking systems without delving into the specific complexities associated with heavy-duty trucks. This oversight is significant because the braking systems of large vehicles like heavy-duty trucks are inherently more complex due to factors such as their larger size, heavier weight, and varied cargo characteristics. These factors can significantly influence braking dynamics, making the development of neural network models for these systems more challenging.

The research gap identified in Dobaj's work points to an unmet need in the automotive industry for specialized neural network applications that cater to the unique requirements of heavy-duty truck braking systems. This gap presents an opportunity for future research to develop neural network models that are specifically tailored to address the complexities of heavy-duty trucks. Such research would not only enhance the predictive accuracy and reliability of braking system diagnostics but also contribute significantly to the safety and efficiency of these large vehicles. Addressing this research gap is crucial for advancing the field of automotive safety technology, particularly in an era where heavy-duty trucks play a critical role in global transportation and logistics.

2.4.3 Heavy Duty Vehicle Fuel Consumption Modeling Using Artificial Neural Networks

Wysocki et al., (2019) represented a significant advancement in the application of Artificial Neural Networks (ANNs) to model fuel consumption in heavy-duty vehicles. This research utilizes data collected via the CAN bus, a vehicle's internal communication system, to feed the ANN model. The primary objective of the study was to develop a more accurate and reliable model for predicting fuel consumption in heavy-duty vehicles, which are known for their substantial fuel usage and consequential environmental impact. The ANN approach is particularly beneficial in this context due to its ability to process and analyze large datasets, capturing complex, non-linear relationships that are typical in vehicle operational data. The researchers compared the ANN model with conventional polynomial regression models, demonstrating that the ANN approach could more effectively represent the fuel consumption patterns, especially in transient states of the vehicle's operation.

However, the research also uncovers a gap in the broader application of ANNs within the heavy-duty vehicle sector, particularly in braking systems. While the study effectively demonstrates the potential of neural networks in interpreting complex data for fuel consumption, it does not extend this application to other critical aspects of vehicle operation, such as braking systems. The unique challenges posed by heavy-duty truck braking systems - such as their response to different loads, driving conditions, and wear and tear - require specialized neural network models for predictive maintenance and efficient operation. This gap highlights an opportunity for future research to explore how the principles and methodologies applied in fuel consumption modelling can be adapted to enhance the safety, efficiency, and predictive maintenance capabilities of braking systems in heavy-duty vehicles. Such advancements could lead to significant improvements in the overall performance and environmental impact of these vehicles.

2.5 Prediction model's efficacy and precision in a practical situation

As per Jarrett & Clark (2012), The precision and efficacy of the prediction model in the present situation showcase consideration in predictive maintenance for braking systems in heavy-duty trucks. The success of the prediction model is dependent upon the ability to change theoretical accuracy in real-world usage (Eriksson & Nielsen, 2014). Taking the practical situation into consideration, the efficacy of the model is impacted by its ability to regularly and accurately showcase impending braking system failure under various operational conditions. The model needs to demonstrate adaptability to the dynamic nature of heavy-duty operation of trucks, accommodating changes in terrain, driving patterns, and load conditions. According to Zhang et al. (2021), in real-world applications, prediction model is vital for improving the efficiency and accuracy of braking systems. Using complex algorithms and past data, these models may optimise braking techniques, increase overall safety, and anticipate brake performance with high accuracy. As an example of an area where prediction modelling has been successful, consider brake distance predictions. Models can predict how far a car needs to stop by looking at things like the car's speed, the road's circumstance, and the specifications of the braking system. In order to create brake systems that are both safe and operate to their full potential, this data is crucial.

According to Rajaram & Subramanian (2015), The prediction model's effectiveness centres upon its real-time responsiveness to changing conditions that are faced by heavy-duty trucks. There are various factors in variables like environmental change, the dynamic nature of components of braking, and system wear and tear. Regular testing and validation in different ranges of scenarios are important to assess to robustness of the model (Kirches, Bock, Schloder & Sager, 2013). Moreover, the application of the feedback mechanism permits the model to learn and adapt which is linked with actual results that increase its efficacy over time. Westerhof & Kalakos(2017) have stated that prediction model precision is tied closely to the ability to differentiate between normal variations in the behaviour of the system and anomalies that indicate impending failure. The false positive which predicts failure incorrectly could result in unwanted maintenance intervention which enhances the costs of operations. On the contrary, false negatives that fail to predict actual failure face major safety risks (Mohammadpour, Franchek & Grigoriadis, 2012). It is important to strike a balance between continuous refinement and optimisation which is important to ensure the practical utility of the model.

A study conducted by Smith et al. (2018) showcased the practical uses of a model for prediction by demonstrating its efficacy in calculating braking distances according to several real-world conditions. Predicting when brake pads will wear is another critical component of an effective braking mechanism. It is essential to precisely estimate the remaining longevity of brake pads in order to schedule upkeep and avoid unanticipated brake failures. Brake pads wear out gradually over time. Using variables including mileage, driving habits, and climatic conditions, Li et al. (2019) created a prediction model that used machine learning approaches to properly forecast when the brake pads would wear. By lowering the likelihood of braking system failures and allowing for prompt replacement of brake pads, this model proved to be very accurate and dependable.

To summarise, braking systems that use prediction model have shown to be both effective and accurate in real-world applications. These models aid in improved safety, optimum braking efficiency, and preventative maintenance preparation by precisely forecasting braking lengths and calculating brake pad wear. Meanwhile, the precision and efficacy of the prediction model for the failure of braking systems in heavy-duty trucks need to be evaluated in real-world scenarios. Its adaptation to various conditions which is responsible for dynamic factors of operation and ability to decline false positives and negatives is found to determine success in increasing road safety, declining downtime ad also optimise the costs of maintenance (Westerhof & Kalakos, 2017). An assessment of the performance of a model under practical conditions is important in instilling confidence in applicability and reliability in the complex and demanding context of heavy-duty operations of trucks.

2.6 Advantage of predictive maintenance regarding braking systems of Scania vehicles

As per Fei & et.al. (2016), the evaluation of the financial and security benefits of predictive maintenance concern for the braking systems of Scania vehicles has led to strategic benefits of applying advanced strategies of maintenance in the commercial sector of transportation. Financially, predicted maintenance adoption for the braking systems of Scania had resulted in cost savings through resolving risks linked with unwanted failure. By forecasting the issues accurately before they escalate, the fleet could play and effectively build schedule activities of maintenance, declining downtime, and optimizing resource utilization (Yazdinejad, Parizi, Dehghantanha & Choo, 2020). Such practice approaches not only decline direct expenses of unscheduled repairs but even increase the operational effectiveness of Scania vehicles. In addition to that, predictive maintenance leads to the enhancement of the lifespan of components of braking. By acknowledging and mitigating problems at an early stage, the tear and wear upon parts like brake pads, callipers, and discs could be effectively managed. Such extension of component life results in a declined frequency of replacement which leads to direct cost savings for the operation of the fleet (Fei & et al., 2016). Moreover, the increased reliability of braking systems leads to a decline in the chances of accidents and associated costs like claims of insurance, legal expenses, and damage to cargo (Dunn, Guenther & Radlinski, 2019). Through the point of security into consideration, Westerhof & Kalakos (2017) have stated that predictive maintenance plays a major role in ensuring the safety of Scania vehicles and their surroundings. When the braking system failure is detected in the early stage it decreases the chances of accidents which are caused due to malfunctions and thus, prevents truck occupants and other users of the road. Such preventive approaches build alignment with the commitment of Scania toward safety and lead to the improvement of road safety standards (Rajaram & Subramanian, 2015). The usage of predictive maintenance technology like neural network prediction models helps in increasing the security of heavy-duty trucks by giving reliable insights into safety component conditions.

It is further stated by Dunn, Guenther & Radlinski (2019) that the advantage of security is found to extend the protection of valuable cargo and prevent environmental hazards. A well-maintained system of braking decreases the chances of accidents which could lead to damage to cargo mainly in case of hazardous stuff (Turri, 2015). It not only safeguards the interest of finance of fleet operators but even leads to enhanced public and environmental objectives of safety. To summarise, the reviewing of security and financial advantage of predictive maintenance for Scania vehicles' braking systems emphasizes cost savings, increased safety, and efficiency of operation (Yazdinejad, Parizi, Dehghantanha, Srivastava, Mohan & Rababah, 2020). Through applying the maintenance strategies, Scania could place itself at the initial position by giving priority to economic viability and also road users, drivers, and also environment.

2.7 Other Related works and Research Gap:

The literature reviews comprise information on prediction models, ML and AI in general for forecast and neural network is treated as a possible approach that can be used for forecast of braking failure. Some reviews are found to use neural models but cover only narrow application in automobile areas and not in the braking system of large vehicles like trucks. The following table provides an overview of related work and gap identified.

Table 2 Other Related works and Research Gaps

Sr.	**Research paper**	**Authors**	**Model used**	**Gap**
1	Brake fault diagnosis through machine learning approaches - A review	Manghai & et. Al., 2017	Systematic review	No use of neural network in braking system
2	The failure prediction of a brake disc due to nonthermal or mechanical stresses	Chen & Kienhöfer 2021	Use of soid mechanics formula and stress concentration	No use of neural network model in Scania truck
3	Prediction Model and Experimental Study on Braking Distance under Emergency Braking with Heavy Load of Escalator	Li & et. Al., 2020	Use of mechanical motion and escalator system	Did not developed neutral model for prediction in braking system
4	Real Time Condition Monitoring on Brakes using Machine Learning Techniques	Jayakrishnan & et. Al., 2020	Use of condition monitoring method	No application of neural network in brake failure of trucks

5	Prediction of Brake Pad Wear Using Various Machine Learning Algorithms	Steffan & et. Al., 2022	Use of various ML algorithms to predict brake problem	No development of model by using neural network

2.8 Chapter Summary

The exploration of artificial neural networks (ANNs) in the context of heavy-duty vehicle systems, specifically in the domain of braking systems, reveals a nuanced landscape of both advancements and significant research gaps. The insights from studies like Kiani et al. (2010) and Wysocki et al. (2019) underscore the efficacy of ANNs in modeling complex vehicle behaviors such as engine performance and fuel consumption. These studies highlight the potential of neural networks to process and analyze intricate data patterns, thereby enhancing vehicle operational efficiency. Particularly, the application of ANNs for fuel consumption modeling in heavy-duty vehicles demonstrates their capacity to address the sophisticated dynamics of these larger vehicles, which are often subject to variable loads and diverse operating conditions.

However, there exists a marked research gap in extending these ANN applications to the specific nuances of heavy-duty truck braking systems. Studies like those by Vodovozov et al. (2021) and Dobaj (2022) indicate a limited focus on neural network-based control systems for heavy-duty trucks, particularly in their braking mechanisms. This is critical because the braking systems of such trucks are inherently more complex due to factors like size, weight, and cargo dynamics. This complexity not only presents challenges in developing ANN models but also underscores the need for tailored solutions. Hence, the focus of future research should be on developing specialized neural network models that can accurately predict and enhance the efficiency of braking systems in heavy-duty vehicles. Such research would not only fill a critical gap in the current landscape but also contribute significantly to the safety and operational efficiency of heavy-duty trucks, which are integral to global logistics and transportation networks. This focused approach in research can potentially lead to groundbreaking advancements in automotive safety technology, particularly in an era where heavy-duty trucks are pivotal in global trade and commerce.

CHAPTER 3: RESEARCH DESIGN

3.1 Research Philosophy

Research Philosophy refers to the underlying belief system that guides the research approach, encompassing the assumptions about the nature of knowledge. Positivism is chosen because it emphasizes the importance of objective measurements, quantifiable data, statistical analysis. (Saunders, Lewis, & Thornhill, 2019). Neural network models would be developed and validated through statistical measures of performance, such as accuracy, precision, recall, and the area under the ROC curve, all of which are in line with a positive approach.

3.2 Research Approach

Research Approach dictates whether the research is inductive, building theories through the collection of data, or deductive, testing hypotheses based on existing theories. Deduction is used for theory testing, which is characteristic of a positivism philosophy. A deductive approach is suitable for developing predictive models as it involves formulating hypotheses based on existing theory (neural networks) and then designing a study to test the hypotheses with data. This approach is appropriate for validating model predictions against observed outcomes (Bryman, 2012). Deduction is chosen over induction because this study involves testing a theory (neural networks) rather than building a theory from observations.

3.3 Research Methodology

Research Methodology represents the overarching strategy and rationale of research. A quantitative methodology relies on numerical data and statistical methods to understand phenomena and often results from a positivist philosophy.

The methodology should be quantitative, given the nature of neural networks require large volumes of numerical data for training and validation. In this study, the dataset is numerical and sizable, consisting of various parameters that influence brake performance. Quantitative methods are adept at handling such data, applying various statistical techniques to uncover patterns and train the predictive models. Quantitative research allows for the application of statistical analysis to validate the model's accuracy and reliability (Creswell & Creswell, 2017). The efficacy of neural networks is evaluated based on quantitative metrics such as accuracy, precision, recall, F1 score, and the area under the ROC curve. These metrics are derived from a quantitative analysis of the model's performance, making quantitative methodology the most appropriate approach for your research. Quantitative methods are more suitable than qualitative due to the mathematical nature of neural networks and large datasets.

3.4 Research Strategy

Research Strategy is the plan for how to conduct the research, including experimentation, surveys, case studies, etc. An experimental strategy is typically used in scientific research that seeks to establish causal relationships. The strategy of experimentation fits well with the development of a neural network model, as it will involve manipulating the model's parameters and structure to observe changes in performance. Experimentation is a powerful strategy to establish causal relationships, crucial for understanding the impact of predictors on brake failure (Fisher, 1935). While a case study could provide in-depth analysis of individual instances of brake failure, experimentation allows for broader generalization and validation of the neural network model.

3.5 Time Horizon

Time Horizons refers to the timeframe over which the data is collected. A cross-sectional time horizon involves collecting data at a single point in time, whereas a longitudinal approach involves collecting data over an extended period. Since this study involves analyzing data collected at a single point in time to develop the predictive model, a cross-sectional time horizon is most appropriate. It is efficient and suitable for this type of predictive modeling research (Robson & McCartan, 2016). A longitudinal study would be less practical due to time constraints and the nature of model development, which does not require tracking changes over time.

3.6 Data collection

Data collection is a critical portion of investigation consider which is critical in gathering information. In this consideration, the dataset for this extent has been collected through the UCI Machine learning repository (AB,2016). There were two information sets accessible training dataset and test wherein a training set had sixty thousand information focuses and testing sets had sixteen thousand information focuses (Creswell, 2013). In this consideration, fifty-nine thousand information focuses were labelled in training sets as negative class, and one thousand were named as a positive class. The names of traits were made anonymous to defend proprietary data.

3.7 Project Methodology

In this study, CRISP DM (Cross-Industry Standard Process for Data Mining) methodology is employed. The CRISP-DM methodology is a structured approach to planning and executing data mining projects (Wiener et al., 2019). It is highly regarded for its systematic framework that guides projects from understanding the business problem to deploying the solution effectively.

Schnell et al. (2019) highlighted the advantages of CRISP-DM in engineering applications, particularly for tasks like optimization and predictive maintenance, which are closely related to the objectives of this project.

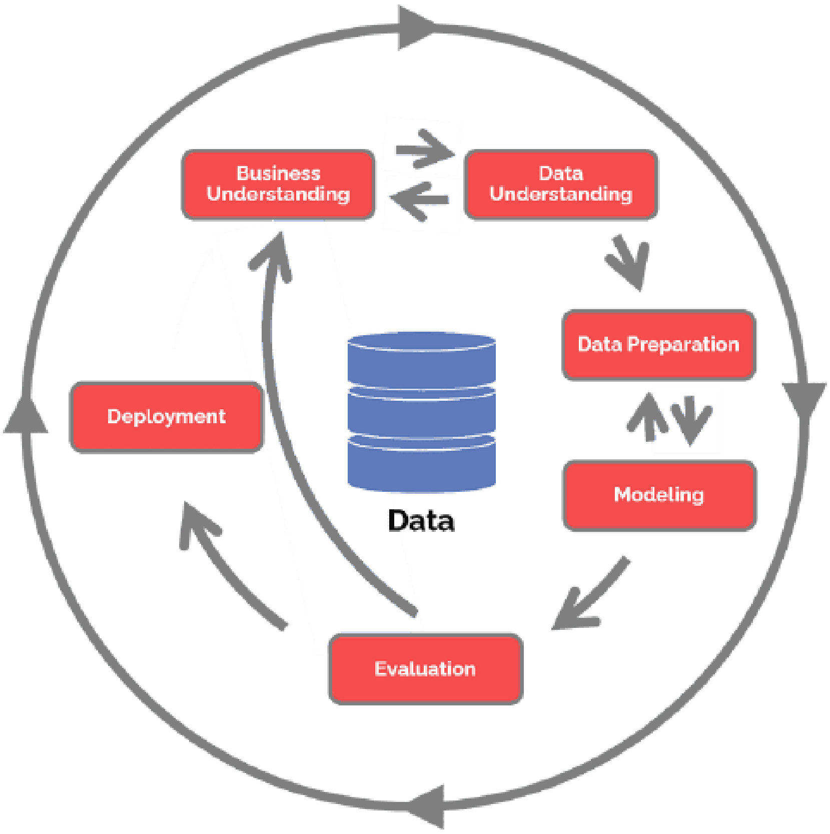

Figure 1 Phases in CRISP-DM, Hotz (2018)

In this study, CRISP-DM's phased approach is beneficial for managing the complexity of neural network projects, providing clear steps and deliverables that can be revisited as needed (Caetano et al., 2014). Unlike other methodologies that may be more rigid, CRISP-DM is known for its flexibility, making it adaptable to changes in project requirements, which is a common occurrence in data-driven projects (Studer et al., 2020). CRISP-DM supports the iterative nature of model building and validation, allowing for continuous improvement of neural network models as more data becomes available or as the understanding of the problem deepens (Studer et al., 2020).

Compared to other methodologies like SEMMA (Sample, Explore, Modify, Model, Assess) or KDD (Knowledge Discovery in Databases), CRISP-DM's broader stages, iterative loops, and focus on business understanding provide a more holistic approach to project management in complex data-driven initiatives like the development of neural network models for predictive maintenance.

3.8 Tools and Techniques

In the development of the neural network model for predicting brake failure in heavy-duty trucks, a comprehensive suite of methodologies and tools is employed, addressing every phase from initial data handling to final model evaluation. This approach encompasses data preprocessing, advanced neural network model development, and rigorous testing. The selection of tools, such as Python and TensorFlow, and techniques like Recursive Feature Elimination (RFE) and SMOTE for data balancing, is informed by extensive literature review. These choices are aligned with the positivism and data-driven research philosophy of the study, ensuring that the model is both scientifically robust and practically applicable in the realm of predictive maintenance for heavy-duty trucks.

Supervised machine learning will be employed in our project as we have control over both the input variables, referred to as features, and the output variable, which is labeled as "class" in our dataset.

3.9 Programming Languages

Python: Python is chosen for its prominence in the field of machine learning and neural network development. Renowned for its extensive libraries like TensorFlow and Keras, Python facilitates efficient model building and training. Its user-friendly nature and rich ecosystem enable seamless integration of various techniques necessary for our complex predictive maintenance model. The choice of Python, coupled with its powerful libraries, aligns perfectly with the sophisticated demands of predicting brake failures in heavy-duty trucks, ensuring a robust and versatile framework for this research.

3.10 Exploratory Data Analysis

Pandas: Pandas is a python library used for its powerful data manipulation and analysis capabilities, crucial for handling and preparing the brake failure dataset.

NumPy: NumPy supports Pandas with high-performance multidimensional array operations, which are essential for numerical computations in data analysis.

Seaborn: Seaborn is also a python library utilized for its advanced data visualization features. Seaborn is particularly effective for statistical graphics, enabling the visualization of complex data patterns and relationships.

Matplotlib: Matplotlib is utilized for its flexibility and wide range of plotting capabilities. It provides detailed and customizable plots, crucial for a deep and nuanced understanding of the dataset's characteristics and underlying trends.

These libraries are collectively chosen because it is robust toolkit, enhancing the efficacy of data processing, analysis, and visualization which supports to develop accurate neural network model

3.11 Deep Learning Frameworks

TensorFlow: TensorFlow is a comprehensive, open-source deep learning framework developed by the Google Brain team. It's used for creating and training machine learning models. It's highly flexible, allowing for easy model experimentation and deployment across various platforms.

Keras : Keras is an open-source software library operating on top of TensorFlow that provides a Python interface for artificial neural networks. Keras simplifies the creation and training of deep learning models with its high-level building blocks. Keras is especially useful for fast prototyping and experimentation.

These frameworks were chosen because it is well-suited for developing sophisticated deep learning models due to their flexibility, ease of use, and extensive community support.

3.12 Model Selection and Architecture

The selection of three specific models in this study is driven by their distinct capabilities and suitability for the task of predicting brake failures in Scania trucks.

- Multi-Layer Perceptron
- 1D Convolutional Neural Network
- Recurrent Neural Networks – Long Short-Term Memory

3.11.1 Multi-Layer Perceptron

An MLP (Multi-Layer Perceptron) is a type of neural network capturing complex patterns and relationships in data, making them suitable for predictive modeling tasks. It consists of multiple layers of nodes, each typically employing a nonlinear activation function.

The choice of MLP in this study is because of its appropriateness for the brake failure prediction task. MLPs are adept at handling the multivariate nature of the dataset and capable of modeling the non-linear relationships between various features and brake failure occurrences. Meyer (2021) study demonstrates the proven track record of MLP in similar predictive maintenance tasks and its efficiency in predicting the failures.

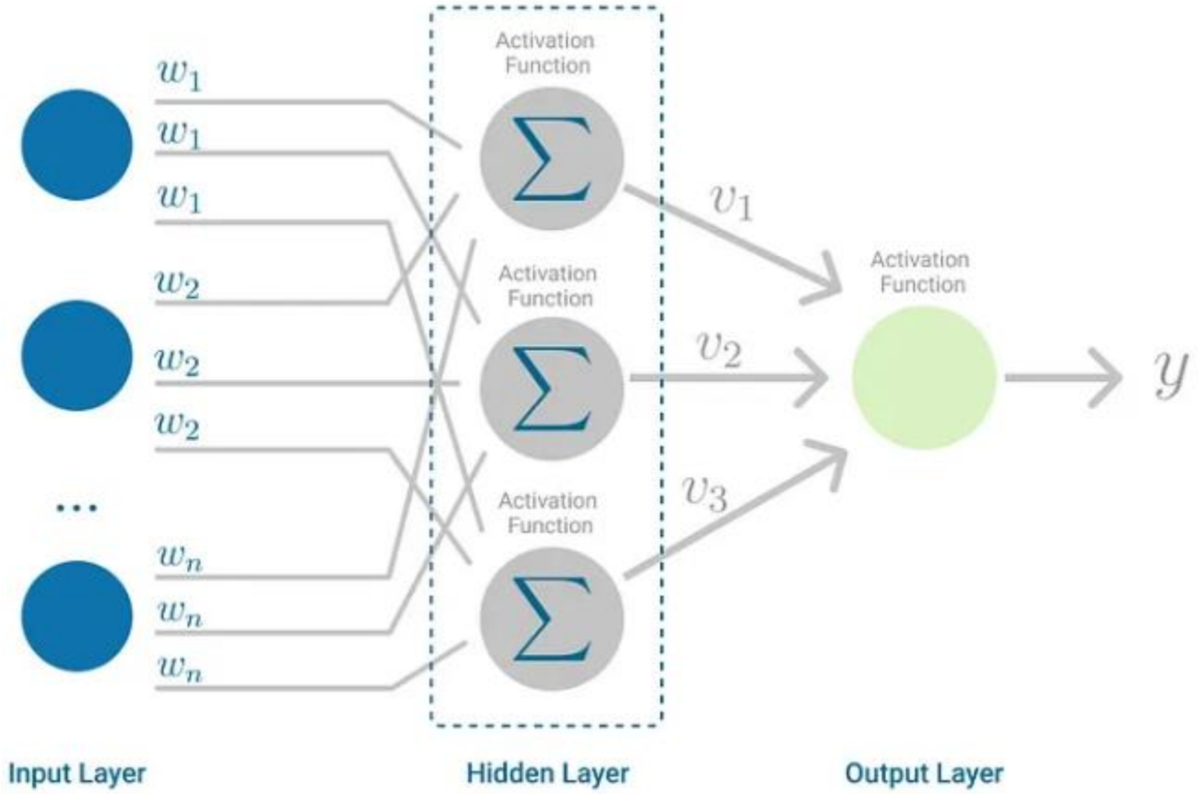

Figure 2 *Multi-Layer Perceptron Architecture*

3.11.2 1D Convolutional Neural Network

A 1D Convolutional Neural Network (1D CNN) is a type of neural network that excels in processing sequential data, such as time-series or sensor signals. It is effective in capturing patterns across time, making it ideal for analyzing data where temporal relationships are significant.

The choice of a 1D CNN in this study is because of its ability to process and learn from the time-dependent characteristics of brake failure data in heavy-duty trucks. This model is adept at handling sequences, such as those found in sensor data, by identifying trends and anomalies over time which are crucial for predicting brake failures. Kiangala and Wang (2020) highlighted the effectiveness of 1D CNNs in time-series analysis and predictive maintenance tasks is well-documented, showcasing its successful application in similar contexts

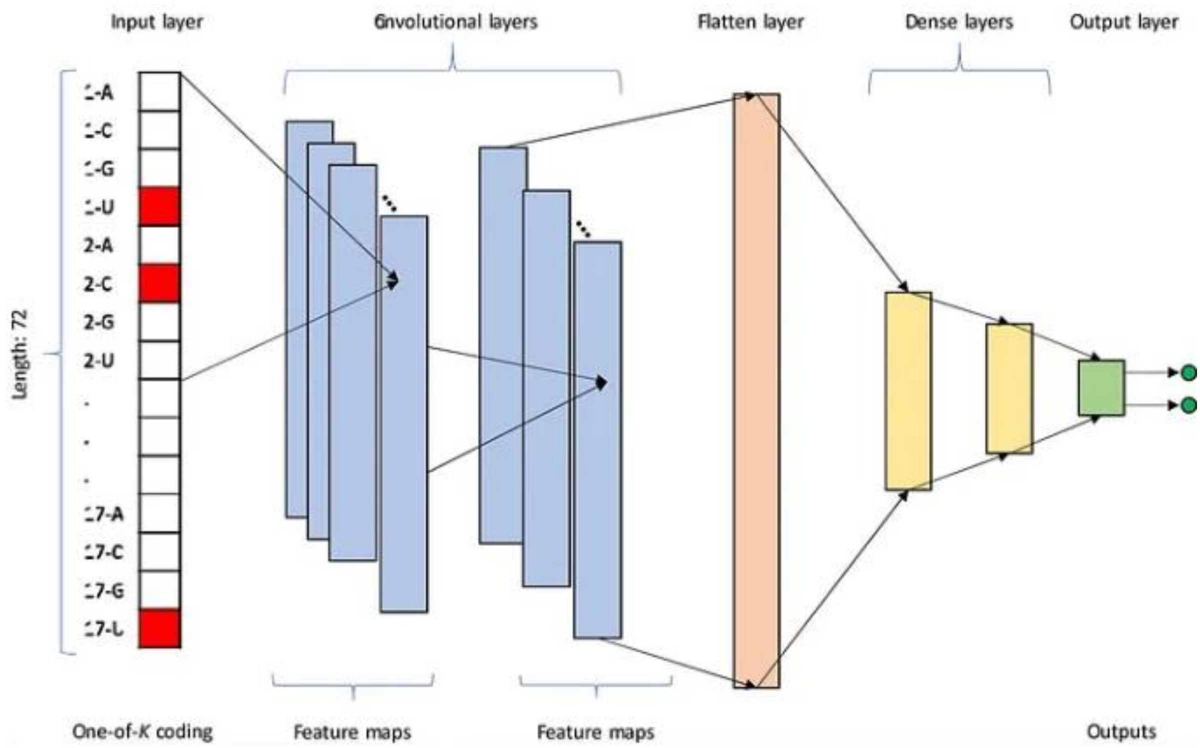

Figure 31D Convolutional Neural Network - Architecture

3.11.3 RNN with Long Short Term Memory and Bidirectional Twist

A Recurrent Neural Network (RNN) with Long Short-Term Memory (LSTM) and a bidirectional twist is a sophisticated model designed for sequential data analysis. LSTM units are capable of learning long-term dependencies, overcoming the limitation of traditional RNNs in handling long sequences. The bidirectional architecture allows the network to gather information from both past (backward) and future (forward) states, enhancing its predictive accuracy.

The selection of this model in this study is due to its ability to effectively process and learn from sequential data, such as time-series or sensor data common in brake failure scenarios in heavy-duty trucks. The LSTM's proficiency in remembering long-term dependencies is crucial for recognizing patterns that lead to brake failures, while the bidirectional approach ensures a comprehensive analysis of the sequential data. Lindemann et al. (2021) provided an overview of LSTM network architectures for time series prediction, noting the effectiveness of sequence-to-sequence networks, including bidirectional models, in fulfilling requirements for accurate time series prediction.

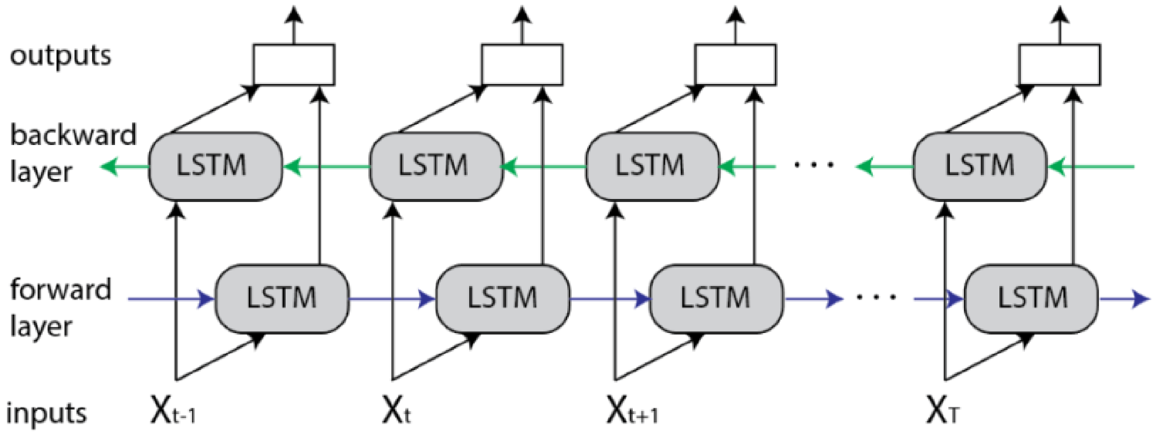

Figure 4 Long Short-Term Memory with Bidirectional Twist – Architecture

3.12 Model Evaluation

Metrics included in TensorFlow/Keras (tensorflow.keras.metrics) offer tools for evaluating model performance, encompassing metrics like precision, recall, binary accuracy, and AUC-ROC. These are crucial for evaluating the effectiveness of the constructed classification mode. (Ghandali et al., 2021)

3.13 Ethical Considerations

This research has been conducted in accordance with the ethical guidelines stipulated by Sheffield Hallam University (SHU, 2017), and a adhering to ethical principles at every stage, including data collection, analysis, and presentation. In this study, discussion and results are based on well-known literature ensuring their accuracy and impartiality. For data protection, all related documents were securely stored in a private folder, accessible exclusively to the researcher. Research activities was undertaken solely after acquiring the necessary authorization and ethical approval. This study has emphasized proper usage of proprietary data and respected for intellectual property rights. Adhering to a legal framework and ensuring transparent practices have been safeguarded against legal problems. efforts have been made to mitigate any bias in training data to safeguard unfair predictions and results. This ethical consideration is verified through the inclusion of the ethics form in Appendix B

CHAPTER 4 : MODEL DEVELOPMENT AND RESULTS

4.1 CRISP-DM Application and Methodology

This chapter is dedicated to the development of the neural network model for brake failure prediction in Scania trucks, aligning with the structured approach of the Cross-Industry Standard Process for Data Mining (CRISP-DM). It encompasses stages of data preparation, model building, evaluation, and deployment, aiming to create an effective predictive maintenance system. The CRISP-DM framework guides and ensure us the research process in a methodical and transparent manner, underpinned by both theoretical understanding and practical applicability.

4.1.1 Business Understanding

The initial stage of the CRISP-DM methodology in this dissertation involves a thorough understanding of the objectives, requirements, and limitations of our research. This phase focuses on comprehending the essential elements necessary for developing a reliable neural network model for predicting brake failures in heavy-duty trucks.

Project Scope and Objectives: The Research aim to create an accurate predictive model to identify potential brake failures, aiding in proactive maintenance decisions, operational efficiency which will improve safety and efficiency of heavy-duty vehicles.

Intended Users: The primary users are maintenance teams in the trucking industry and Scania's technical staff.

Expected Outcomes: Successful development of a reliable model with high accuracy, aiding in reducing maintenance downtime, improving operational efficiency and improving truck safety.

This phase sets the stage for the following CRISP-DM steps, aligning the development process with the research's practical goals, ensuring the model meets its intended purpose effectively.

4.1.2 Data Understanding

This section examines the APS Failure at Scania Trucks dataset, which forms the foundation of our analytical endeavour. The dataset is multivariate sourced from the UCI Machine Learning repository (centre for Machine Learning and Intelligent Systems) (AB, 2016). There were two data sets available training and test dataset wherein a training set possessed 60,000 data points and testing sets possessed 16,000 data points (Creswell, 2013). Each dataset encompasses 171 features, which include single numerical counters and histogram-based bins that capture different operational conditions.

The dataset is classified into two main categories: the positive class, which signifies component failures related to the APS system, and the negative class, representing failures not related to the APS system.

It is observed that 59,000 out of 60,000 data points in the training dataset belongs to the negative class and 1,000 to the positive class which creates significant imbalance. The test set contains 16,000 datapoints, presumably with a similar distribution.

Further, it comprises datatype of both integer and real numbers, indicative of various operational parameters and conditions. Seven of the 171 attributes are histogram variables, providing a detailed perspective on conditions such as temperature ranges, which is pivotal for understanding the nuances of the truck's operational environment. The names of attributes were made anonymous to safeguard proprietary data.

Importing the dataset

```
x = pd.read_csv("aps_failure_training_set.csv", skiprows=20, na_values=["na"])
np.random.seed(3482)
```

```
x.head(5)
```

[48]:

	class	aa_000	ab_000	ac_000	ad_000	ae_000	af_000	ag_000	ag_001	ag_002	...	ee_002	ee_003	ee_004	ee_005	ee_006	ee_007
0	neg	76698	NaN	2.130706e+09	280.0	0.0	0.0	0.0	0.0	0.0	...	1240520.0	493384.0	721044.0	469792.0	339156.0	157956.0
1	neg	33058	NaN	0.000000e+00	NaN	0.0	0.0	0.0	0.0	0.0	...	421400.0	178064.0	293306.0	245416.0	133654.0	81140.0
2	neg	41040	NaN	2.280000e+02	100.0	0.0	0.0	0.0	0.0	0.0	...	277378.0	159812.0	423992.0	409564.0	320746.0	158022.0
3	neg	12	0.0	7.000000e+01	66.0	0.0	10.0	0.0	0.0	0.0	...	240.0	46.0	58.0	44.0	10.0	0.0
4	neg	60874	NaN	1.368000e+03	458.0	0.0	0.0	0.0	0.0	0.0	...	622012.0	229790.0	405298.0	347188.0	286954.0	311560.0

5 rows × 171 columns

```
print("Shape of the trained dataset:", x.shape)
```

```
Shape of the trained dataset: (60000, 171)
```

Figure 5 Overview of Training dataset

Missing values are present and denoted by 'na', which will require strategic handling and preprocessing to ensure the robustness of the predictive modelling.

This detailed understanding is instrumental in guiding the data preparation and feature engineering processes. The anonymized nature of the attributes demands careful consideration of the underlying patterns that each attribute may represent.

4.1.2.1 Understanding Feature Categories – Histogram and Numerical Features

The dataset's feature information indicates that it includes both numeric values and histogram data, with the latter represented by bins that count different types of occurrences. The convention for naming this data includes an identifier followed by a bin number, formatted as 'Identifier_Bin'. In this naming convention, the segment before the underscore signifies the identifier, and the segment after it signifies the bin number.

To analyze this structure, a specialized function named 'Identifier_bin_count' was developed. This function leverages the Counter utility from Python's collections library, enabling the counting of elements within an iterable collection. The process involves separating the feature names at the underscore (_) and incrementally tallying the identifiers across the dataset. Upon applying the 'Identifier_bin_count' function, we discover that the dataset includes 7 distinct histogram identifiers, specifically "ag", "ay", "az", "ba", "cn", "cs", and "ee" as shown in below (Figure 6). These identifiers represent different operational conditions captured in the data. Each of these features is further divided into 10 bins, totaling 70 attributes that contain data from histogram bins whereas 100 attributes that contain numerical data.

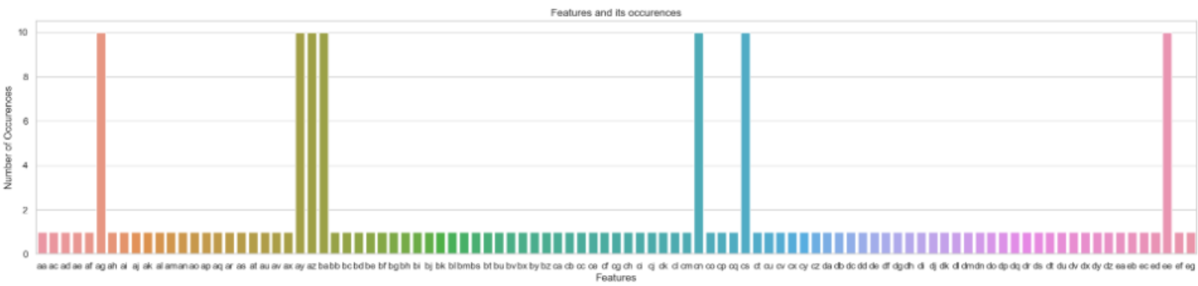

Figure 6 Histogram and Numerical bin identifiers

In preparation for in-depth feature analysis, the dataset's attributes are carefully partitioned. The histogram features, distinct in their bin-based structure, are separated from the numerical attributes. The analysis of these separated features, and how they contribute to the predictive modeling process, will be detailed in the following chapter of the study.

4.1.2.2 Data Exploration and Feature Analysis

Data Exploration is the process of viewing facts and figures in mitigating research issues. It is important to find out research questions and assist in the interpretation of data. It is an essential part of the research study that assists in the process of data interpretation. It is an essential part of the study that effectively assists in data building (Maher & et al., 2018). Quantitative analysis explores numerical data patterns, links, and trends. This possesses feature analysis, identification of correlation, and usage of machine learning techniques to predict the probabilities of brake failure.

A study demonstrated by Li et al. (2018) discusses the importance of histogram feature analysis and visualizing the attributes which helped in reviewing the attributes of datasets.

4.1.2.3 Feature Selection

Feature selection is a crucial process where the most relevant features (or attributes) are selected from a larger set of data. This is important for improving the performance of machine learning models and ensuring better class separability (Cai et al., 2018). One of the major effective feature selections is Recursive Feature Elimination (RFE). RFE is chosen feature selection methodology in this study.

Recursive Feature Elimination (RFE) is a technique used in machine learning to select the most relevant features for a model (Jeon & Oh, 2020). It works by repeatedly building a model and removing the least important feature at each step. This process continues until all features have been ranked according to their importance. Jeon and Oh (2020) presented RFE in combination with a machine learning algorithm called the Random Forest Classifier (RF). It was found that RF-RFE combined is effective in identifying the most important features.

In this study focusing on the development of a neural network model for Scania trucks, the concept of Recursive Feature Elimination (RFE) can be significantly beneficial to identify and select the top ‘n’ features and perform analysis on the feature.

Feature Category

In this study, the dataset’s information states that features are divided into histogram and numerical categories. The main motive is to get access to the importance of every feature in the link to dataset objectives, by examining how features are linked with target variables and impact other patterns, a thorough analysis of features is conducted to assess their significance.

4.1.2.4 Histogram Feature Analysis

In this study, the dataset includes a comprehensive set of 70 histogram-based attributes. However, to streamline the analysis, it opted to focus on a subset of the most relevant features. For this purpose, Recursive Feature Elimination (RFE) technique is employed which is widely recognized for its simplicity in execution and for its proficiency in identifying crucial predictors within the training data that have a significant impact on the outcome variable. RFE requires us to specify the desired number of features to retain as well as the selection model.

In this project, we have chosen to concentrate on the 15 most influential features as determined by RFE in conjunction with the Random Forest Classifier (Jeon & Oh, 2020). The fifteen features identified as most important are "ay_005, ee_005, ay_006, ag_001, ay_008, cn_000, ag_002, cn_004, ag_003, ba_002, cs_002, ba_003, ba_004, cs_004, ee_003." Then it is required to construct a data frame comprising these fifteen attributes to facilitate the classification of the target outcome.

4.1.2.5 Univariate Analysis

In univariate analysis, one variable at a time is examined, observing its distribution and central tendency. This method allows us to understand the individual behavior of each feature.

For this, a function named 'plots' is developed which calculates the mean, standard deviation, and generates the Probability Density Function (PDF), Cumulative Distribution Function (CDF), and Box plot for each feature.

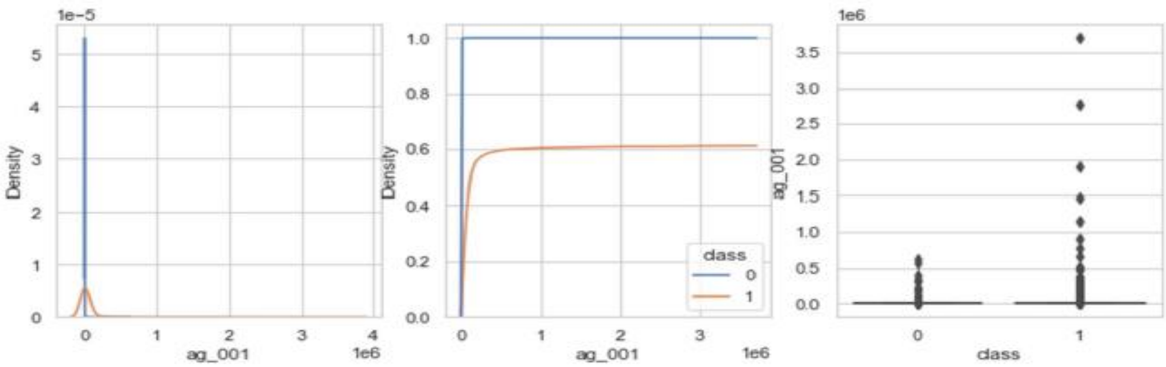

Figure 7 Univariate Analysis for histogram bin -Plot – ag_001 Feature

From the “ag_001” plot (Figure 7), we can understand that higher values are generally associated with APS system failures and lower values with non-failures, although there are exceptions.

For attributes such as “ag_003, ba_002, ay_008, ba_003, cn_004, ba_004, cs_002, ee_003, and cs_004,” lower readings typically suggest the APS component is functioning correctly, whereas higher readings may signify a failure. These attributes, in comparison to 'ag_001' and 'ay_005', present a greater likelihood of being linked to an APS component failure. The graphical representations in (Figure 8,9) illustrate the behavior of 'ay_008' and 'cs_004', respectively.

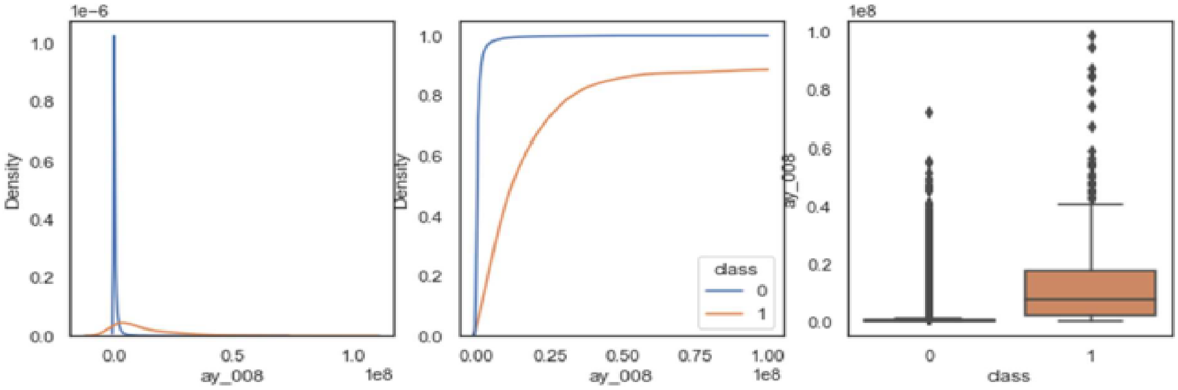

Figure 8 Univariate Analysis for histogram bin Plot - ay_008 Feature

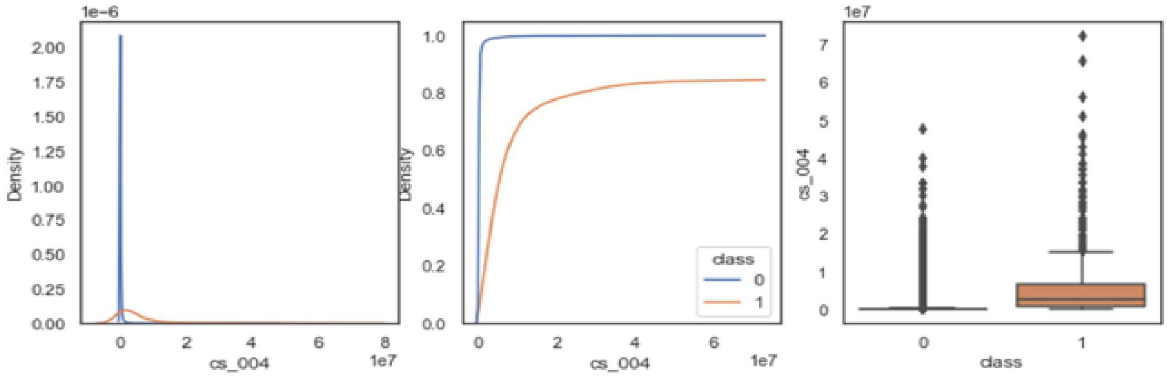

Figure 9 Univariate Analysis for histogram bin Plot – cs_004 Feature

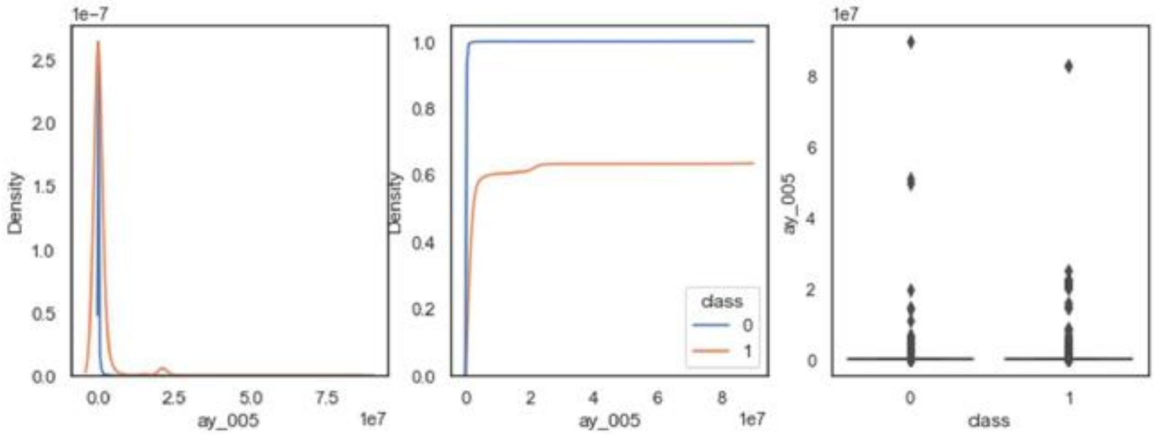

Figure 10 Univariate Analysis for histogram bin -Plot – ay_005 Feature

For instance, in the case of "ay_005", higher values do not always lead to APS failure, demonstrating the nuanced nature of predictive analysis. Additionally, the frequency of APS failure in these situations is minimal when compared with the frequency of non-APS failure.

Each feature's behavior and its implications on APS failure probability are illustrated through individual plots, providing a detailed view of the feature's impact on the system's performance.

4.1.2.6 Correlation between top features

Correlation is a key statistical tool that helps us understand how two variables are related linearly. It's measured using a correlation coefficient, which ranges from -1 to +1.

Pearson correlation coefficient is determined by comparing the product of the deviations of two variables from their respective means, divided by the product of the squared deviations of each variable from its mean. This helps us quantify the relationship between two variables (Medina-González et al., 2019).

To visualize these relationships, especially among the top histogram features of our data, we use a heat map from the seaborn module.

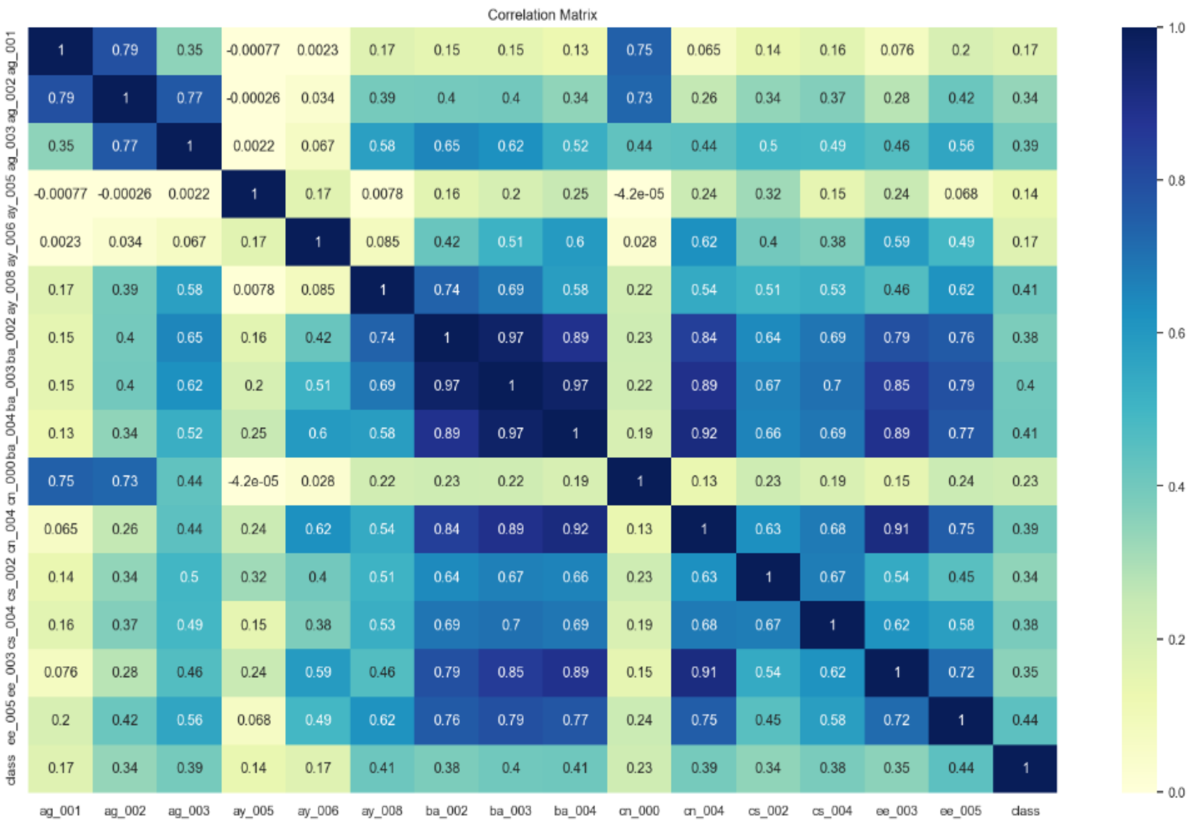

Figure 11 Correlation matrix - Histogram features

Figure 11 illustrates the correlation matrix between these top features. From this visualization, we notice that features like “ba_002, ba_003, and ba_004” have a strong positive correlation, almost close to 1, meaning they are closely related to each other. On the other hand, the feature "ay_005" stands out as the most uncorrelated among the group. It shows a weaker, negative correlation with features such as “ag_001, cn_000, and ag_002”, implying a divergent relationship with these attributes.

4.1.2.7 Bivariate Analysis

Bivariate analysis involves examining two variables simultaneously to understand their relationship. This analysis helps to determine whether there is an association between the two variables, the strength of this association, or if there are significant differences between them.

For instance, we've identified 'ay_005' as the feature with the lowest correlation coefficient. This prompts us to explore how changes in 'ay_005' correspond with changes in other attributes.

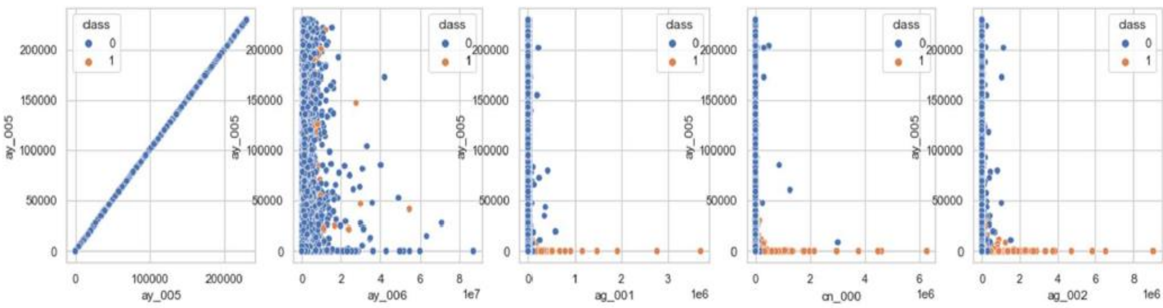

Figure 12 Bivariate Analysis for histogram bin -Plot – ay_005 and other Features

To conduct this analysis, scatter plots are particularly useful as they visually reveal patterns between two variables. An example can be seen in one of our figures, where we notice certain trends regarding the APS component failures. Specifically, when the value of 'ay_005' is nearly zero, failures in the APS component occur regardless of the values in “ag_002, cn_000, and ag_001”. However, as the value of 'ay_005' increases, the likelihood of failures in non-APS systems seems to rise. Additionally, these variables do not exhibit a linear relationship, reinforcing the idea that they are negatively correlated. This lack of linear relationship further indicates the unique and complex interactions between these features.

4.1.2.8 Numerical feature analysis

Numerical feature analysis is a process in data science and statistics where quantitative attributes or variables in a dataset are examined and analyzed. It involves exploring and understanding the characteristics and statistical properties of these numerical features, which can be either discrete (e.g., counts, integers) or continuous (e.g., measurements, floating-point numbers).

In the scope of numerical feature analysis, we are working with a total of 100 features. To manage this effectively, we plan to focus our exploratory data analysis on the top 'n' features.

For this selection process, we will apply the similar method of 'Recursive Feature Elimination' that we used for analyzing histogram features. Our aim is to identify the top 15 features using the Random Forest Classifier algorithm.

These selected top fifteen features include “am_0, aq_000, ap_000, al_000, bu_000, bv_000, bj_000, ci_000, cq_000, cj_000, dn_000, dg_000, dr_000, do_000, and dx_000”. To analyze these features in relation to our project's goals, we will compile them into a new data frame. This data frame will not only include these fifteen key numerical features but also the 'class' attribute as the target variable. This approach allows us to focus on the most influential features in our dataset, enhancing the effectiveness and efficiency of our analysis.

4.1.2.9 Univariate Analysis:

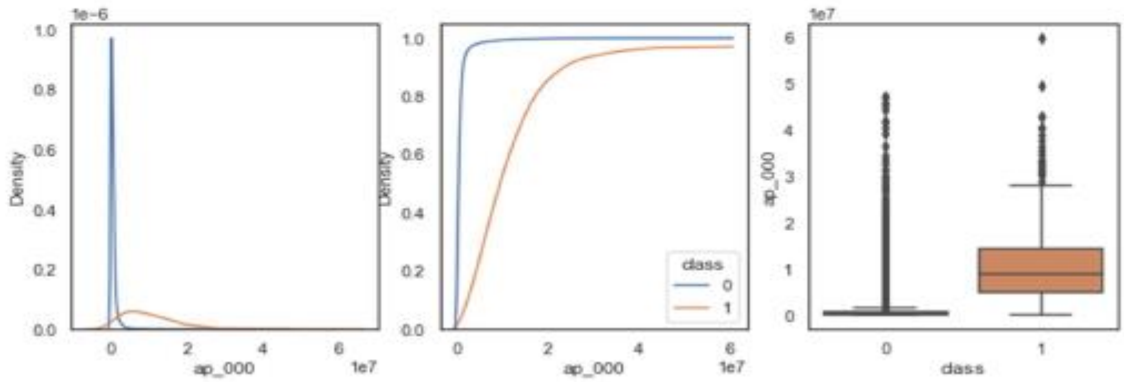

Figure 13 Univariate Analysis for Numerical Features -Plot – ap_000 Feature

In the above (Figure 13), it's evident that there is considerable overlap between the instances of class 0 (non-APS failure) and class 1 (APS failure). Interestingly, the probability of both APS and non-APS failures is almost identical, hovering around 1. When examining the values, we find that APS failures (Class 1) generally exhibit higher values than non-APS failures (Class 0). However, there are certain instances where unexpectedly high values do not lead to APS failures. Additionally, the data points for APS failures show a wider spread compared to non-APS failures.

This observation is consistent with other features like “aq_000, bu_000, bj_000,”, where a higher value typically indicates an APS component failure. However, in some cases, the values of the two classes are indistinguishable.

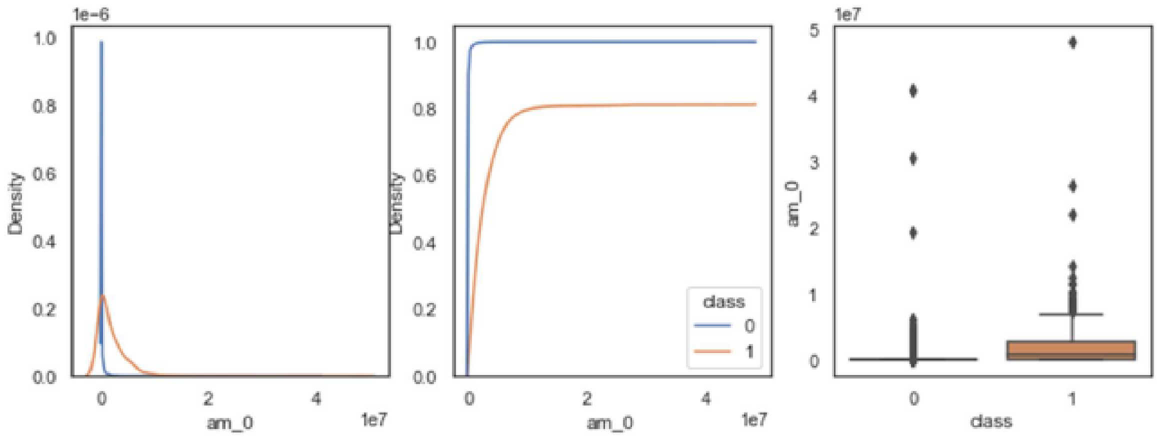

Figure 14 *Univariate Analysis for Numerical Features -Plot – am_0 Feature*

For example, Figure 14 pertaining to the feature 'am_0' illustrates this lack of separability, showing that values in instances of both APS and non-APS failures can be equally high, making it challenging to discern whether a high value will lead to an APS failure. The probability of APS failure in this context is noted to be high, around 0.8.

Similar patterns are observed in features like “al 000, dg 000 and cj 000”, where larger values are associated with failures in both classes. This is somewhat different from the feature 'dn_000', as shown in Figure 15, indicating that each feature has unique characteristics and influences on the outcome.

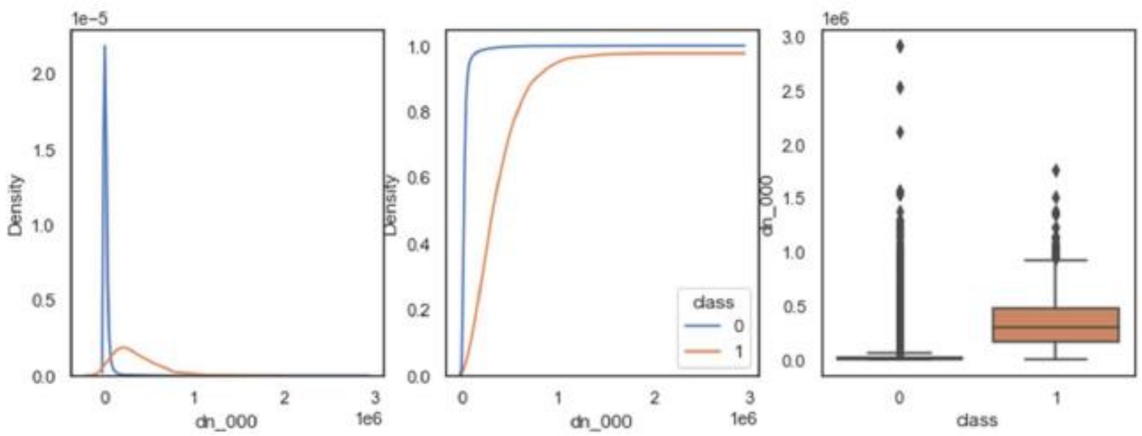

Figure 15 *Univariate Analysis for Numerical Features -Plot – dh_000 Feature*

A unique case is presented by the feature ‘dn_000’. Unlike other features where high values usually indicate APS failures, in this case, very high values are linked to non-APS system failures. This unusual pattern suggests that the data may not be standardized, as the probability of failure is high and almost equal for both classes. This pattern is also observed in feature ‘dx_000’, where exceptionally high values lead to non-APS system failures, diverging from the general trend seen in other features.

4.1.2.10 Correlation between top features

From (Figure 16), it is interesting to note a key difference in the behaviour of numerical features compared to histogram features. While histogram features sometimes show negative correlations, all the numerical features in our dataset are positively correlated. This means that as one feature increases, the other tends to increase as well.

Among these numerical features, 'dx_000' and 'dg_000' show the lowest correlation coefficient, indicating a weaker positive relationship compared to others. On the other hand, 'dn_000' and 'aq_000' stand out with the highest correlation coefficient of 0.97. This high value suggests that these two features are strongly linearly related to each other. Additionally, features like “bv_000, ci_000, bu_000,” also demonstrate a strong linear relationship. This kind of insight is crucial, as it helps in understanding how changes in one feature might be mirrored or influenced by changes in another within the dataset.

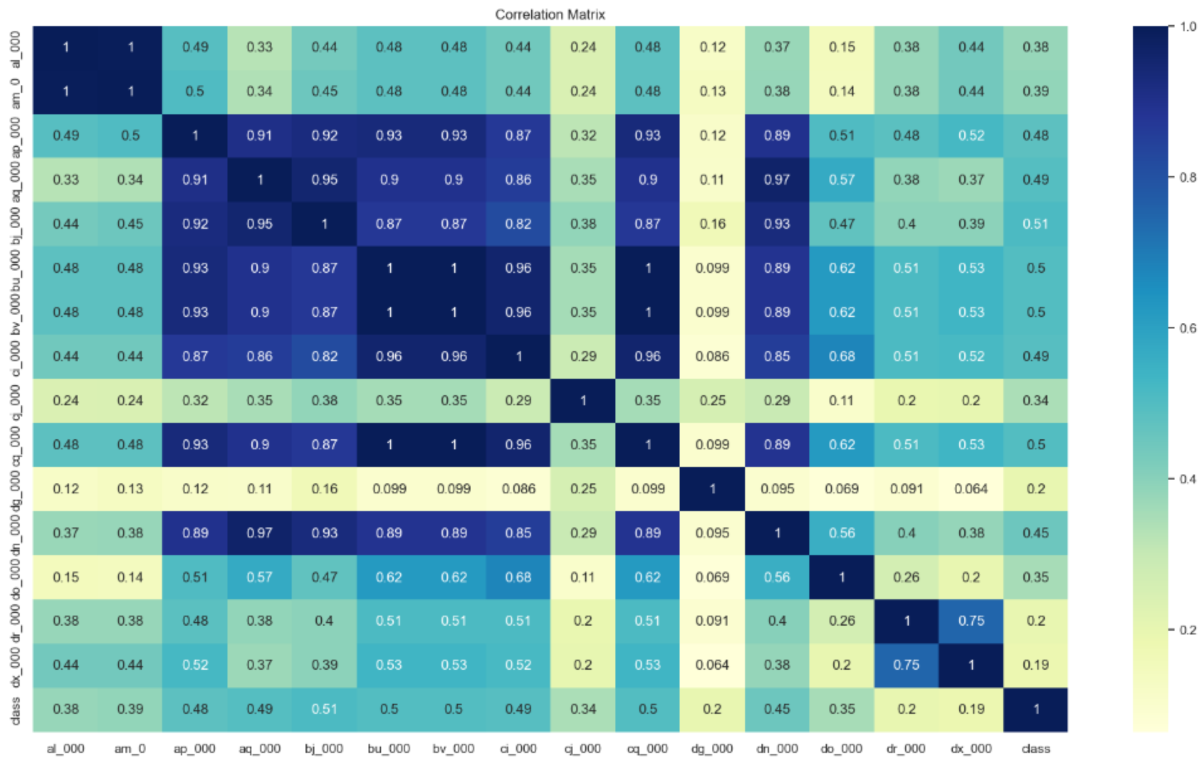

Figure 16 *Correlation matrix – Numerical Features*

4.1.2.11 Bivariate Analysis

Understanding the behaviour of 'dx_000', which is identified as the least correlated feature in our dataset, can be insightful. By conducting a bivariate analysis of 'dx_000' against other features, as shown in the below (Figure 17), we gain valuable insights into its variations.

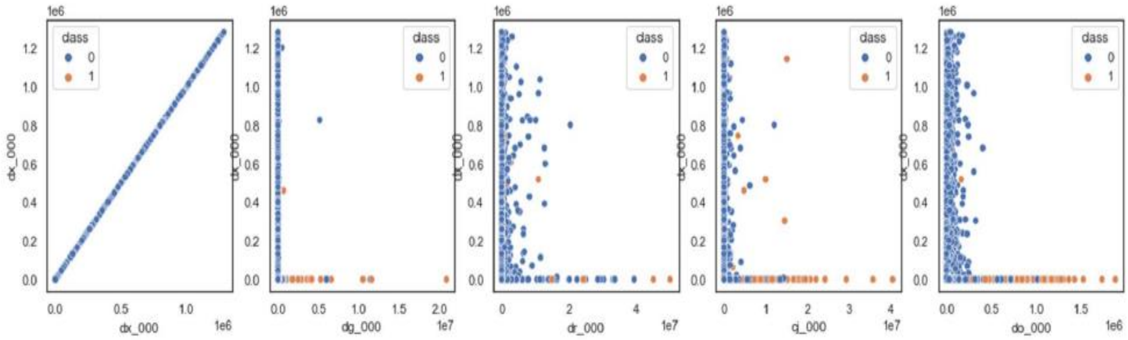

Figure 17 Bivariate Analysis for Numerical Features -Plot – dx_000 and other Features

The analysis reveals a specific pattern: when the value of 'dx_000' is nearly zero, it tends to signal an APS component failure, corresponding to class 1. Conversely, a high value in 'dx_000' often indicates a failure in non-APS components. This pattern is consistent across all features when considered in relation to 'dx_000'.

Interestingly, there's no clear linear relationship between 'dx_000' and these other features. This lack of a straightforward linear trend suggests that these characteristics are negatively correlated with each other. Such insights are crucial as they help in understanding the distinct role 'dx_000' plays in the dataset, particularly in its interaction with other features and its impact on determining APS and non-APS failures.

4.1.3 Data Preparation

Data Preparation phase in the CRISP-DM methodology is critical and includes necessary data pre-processing activities, as it involves transforming raw data into a suitable format for modeling. In the phase of data pre-processing, various tasks were undertaken .As a first step, mitigating missing values was a priority. Moreover, reviewing numerical attributes and attributes appropriate for generating histograms was important. The managing of an imbalanced dataset nature in which the negative class had a larger representation than the positive class has also played significant aspects. Such steps had formed collectively regarding the core of the data pre-processing process.

4.1.3.1 Dealing with Missing Values:

Handling missing values is crucial in statistical analysis and machine learning, as most algorithms require complete datasets. Inappropriate handling of missing data can lead to biased estimates, reduced statistical power, and invalid conclusions. Methods for addressing missing data include deletion, single imputation (like mean substitution), and more advanced techniques like multiple imputation or model-based methods, each appropriate under different circumstances and assumptions about the nature of the missingness (Pratama et al., 2016)

In this study, missing value is a major concern since only 2% of the instances are complete, with some features missing up to 80% of their data, hence handling missing data is critical and important.

A fundamental aspect of a feature's utility in a predictive model is its ability to help distinguish between the target classes. A feature with zero variance which mean it has same value across all data points and it does not differentiate between the classes in any way and therefore has no predictive power. In this study, we will remove the feature “cd_000” since all the values in data points are same, it explains us that standard deviation is zero. Hence after first elimination, the dataset will change from actual shape (60000, 171) to (60000, 170).

```
def constant_value(df):

    constant_value_feature = []
    info = df.describe()
    for i in df.columns:
        if info[i]['std']==0:
            constant_value_feature.append(i)
    df.drop(constant_value_feature,axis=1,inplace=True)
    return df,constant_value_feature

x , dropped_feature = constant_value(x)
print("The features that are dropped due to having a constant value (0 std. dev.) are: ",dropped_feature)
print("Shape of our feature set: ",x.shape)
```

```
The features that are dropped due to having a constant value (0 std. dev.) are:  ['cd_000']
Shape of our feature set:  (60000, 170)
```

Figure 18 Feature cd_000 with Zero standard deviation

Processing the missing data involves a few techniques which will help us in improving model accuracy. Below steps explains each step involved in handling missing data

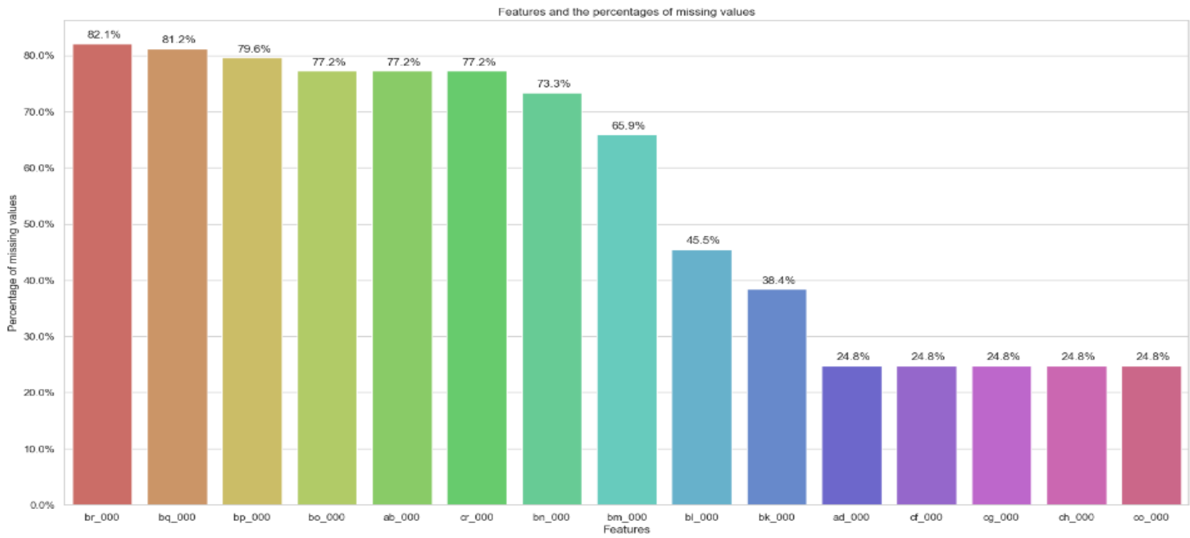

Figure 19 Features with Highest Missing Value

One common approach demonstrated by Toutenburg and Shalabh (2009) "Statistical Analysis with Missing Data" provides extensive guidance on handling missing data, including the recommendation to consider discarding variables with too many missing values. When a large portion of data is missing from a feature, it questions the reliability of that feature and excessive missing data can reduce model accuracy. This method facilitates removing the features with a high number of missing values. To understand the high number of missing values in our dataset, a bar plot is created using seaborn package. Therefore, as shown in the below (Figure 19) seven features such as “bn_000, bp_000, bq_000, bo_000, br_000, ab_000, cr_000,” has more than 70% of missing value which is high number of missing values, so it is removed. After removing the feature, our new dataset shape will be (60000, 163).

```
Number of features whose missing values are 15% to 70% imputed with MICE: 148

MICE Imputed Features: ['aa_000', 'ad_000', 'ae_000', 'af_000', 'ag_000', 'ag_001', 'ag_002', 'ag_003', 'ag_004', 'ag_005',
'ag_006', 'ag_007', 'ag_008', 'ag_009', 'ah_000', 'ai_000', 'aj_000', 'al_000', 'am_0', 'an_000', 'ao_000', 'ap_000', 'aq_00
0', 'ar_000', 'as_000', 'at_000', 'au_000', 'av_000', 'ax_000', 'ay_000', 'ay_001', 'ay_002', 'ay_003', 'ay_004', 'ay_005',
'ay_006', 'ay_007', 'ay_008', 'ay_009', 'az_000', 'az_001', 'az_002', 'az_003', 'az_004', 'az_005', 'az_006', 'az_007', 'az_
008', 'az_009', 'ba_000', 'ba_001', 'ba_002', 'ba_003', 'ba_004', 'ba_005', 'ba_006', 'ba_007', 'ba_008', 'ba_009', 'bb_00
0', 'bc_000', 'bd_000', 'be_000', 'bf_000', 'bg_000', 'bh_000', 'bi_000', 'bj_000', 'bk_000', 'bl_000', 'bm_000', 'bs_000',
'bt_000', 'bu_000', 'bv_000', 'by_000', 'bz_000', 'cb_000', 'ce_000', 'cf_000', 'cg_000', 'ch_000', 'ci_000', 'cj_000', 'ck_
000', 'cl_000', 'cm_000', 'cn_000', 'cn_001', 'cn_002', 'cn_003', 'cn_004', 'cn_005', 'cn_006', 'cn_007', 'cn_008', 'cn_00
9', 'co_000', 'cp_000', 'cq_000', 'cs_000', 'cs_001', 'cs_002', 'cs_003', 'cs_004', 'cs_005', 'cs_006', 'cs_007', 'cs_008',
'cs_009', 'ct_000', 'cu_000', 'cv_000', 'cx_000', 'cy_000', 'cz_000', 'da_000', 'db_000', 'dc_000', 'dd_000', 'de_000', 'dn_
000', 'do_000', 'dp_000', 'dq_000', 'dr_000', 'ds_000', 'dt_000', 'du_000', 'dv_000', 'dx_000', 'dy_000', 'dz_000', 'ea_00
0', 'ec_00', 'ed_000', 'ee_000', 'ee_001', 'ee_002', 'ee_003', 'ee_004', 'ee_005', 'ee_006', 'ee_007', 'ee_008', 'ee_009',
'ef_000', 'eg_000']
```

Figure 20 MICE Imputed Features with Missing ranging between 15% to 70%

Azur et al. (2011) exemplified about Multiple imputation by chained equations (MICE) focusing on practical aspects and challenges of using this method. Azur et al. (2011) study proves that MICE, as a model-based imputation method, uses the entire dataset to estimate missing values, potentially leading to more accurate and less biased imputations, especially when missingness is related to the observed data. The MICE (Multiple Imputation by Chained Equations) technique is more sophisticated and suitable for situations where a larger portion of data is missing. It iteratively models each feature with missing values as a function of other features, which can be a more accurate reflection of complex data relationships. This approach supports dealing features with 15% to 70% missing values in our dataset implementing the model-based imputation using the MICE technique. This involves using the Sklearn's iterative imputer module with a Ridge Regressor. Ridge Estimator is chosen because it works well with a large number of predictors and provides smaller imputation error compared to other estimators. The smaller imputation error implies that the predicted values for the missing data are likely to be more accurate. In our study, at each stage we identify attributes in the dataset with missing data and it is treated as the response variable (denoted as y) whereas other column in the dataset is treated as predictors (denoted as x). Ridge Regressor is applied to this set of predictor variable(x) and response variable (y) to predict the missing values of y. This is applied iteratively and repeatedly for a default of 10 rounds to predict missing values, which is then applied to both training and test datasets.

```
Number of features whose missing values are 5% to 15% imputed with mean: 14

Mean Imputed Features: ['ak_000', 'ca_000', 'dm_000', 'df_000', 'dg_000', 'dh_000', 'dl_000', 'dj_000', 'dk_000', 'eb_000',
'di_000', 'ac_000', 'bx_000', 'cc_000']
```

Figure 21 MEAN Imputed Features with Missing ranging between 5% to 15%

Kanchana, S., & Thanamani, A. S. (2014) focuses on identifying efficient imputation methods to handle fewer missing data in datasets. This study involves a comprehensive guide for researchers and data analysts in selecting suitable imputation methods for handling missing data. It highlights the importance of understanding the dataset's characteristics and the nature of the missing data to choose the most effective imputation technique. This detailed explanation of their study provides insights into the effectiveness of mean imputation relative to other methods and underscores the importance of method selection in data preprocessing. This approach makes us easier to implement mean imputation to fewer missing value in our dataset. We have identified 14 features which have missing values from the range of 5% to 15%. For attributes which have missing values from 5% to 15%, mean imputation is employed, leveraging the Simple Imputer from the Sklearn library.

Above structured approach to addressing missing data aids in refining the dataset for enhanced model performance.

4.1.3.2 Data Normalization

Normalizing data is essential as it facilitates the gradient descent process, enabling algorithms to find the cost function's minimum more effectively. Without appropriate normalization, models may disproportionately weigh features with higher numerical ranges, resulting in skewed results. For our project, we will use the 'MinMaxScaler' from the 'Preprocessing' module of the Sklearn library. This scaler modifies each feature to a specific range, usually between 0 and 1. By doing this, it ensures a consistent scale for all attributes within our training data. Such normalization is critical to improve the efficacy and precision of our predictive algorithms (Pedregosa et al., 2011).

Process of Data Normalization can be defined by

$$x_{std} = \frac{(x - x.\min(axis = 0))}{(x.\max(axis = 0) - x.\min(axis = 0)\ )}$$

$$x_{scaled} = x_{std}\ * (\max - \min\ ) + min$$

4.1.3.3 Handling Class Imbalance

The 'class' attribute in this dataset identifies data points as 'pos' for positive, indicating a failure associated with the APS system, or 'neg' for negative, indicating no relation to APS system failures. For modelling purposes, we convert these labels from strings to integers, assigning 1 to 'pos' and 0 to 'neg'. Despite this conversion, the class feature is highly imbalanced as shown in (Figure 22).

This dataset suffers from a notable imbalance explaining 98.3% of the values are labelled as negative, and only 1.67% are labelled as positive. This creates an imbalanced dataset, which can lead to misleadingly high accuracy in models due to the bias towards the majority class. Shape of the training data shape was (60000, 171) with 59000 negatives and 1000 positives.

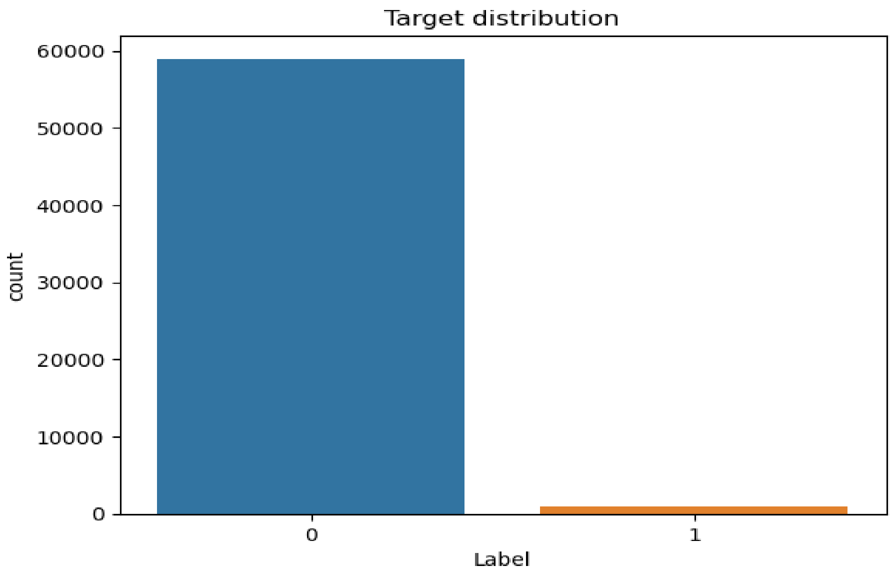

Figure 22 Distribution of Class Variable – Initial Training Dataset

Synthetic Minority Oversampling Technique (SMOTE) is a well-known preprocessing approach for handling imbalanced datasets. SMOTE helps to balance the class distribution by creating synthetic samples in the minority class and it allows the minority class to have a more significant presence in the dataset and also reduces the risk of overfitting (Hussein et al., 2019).

To address the high imbalance in this dataset, Synthetic Minority Over-sampling Technique (SMOTE) is employed. This sampling technique aims to balance the dataset by generating synthetic minority class samples, ensuring a more representative distribution of both classes.

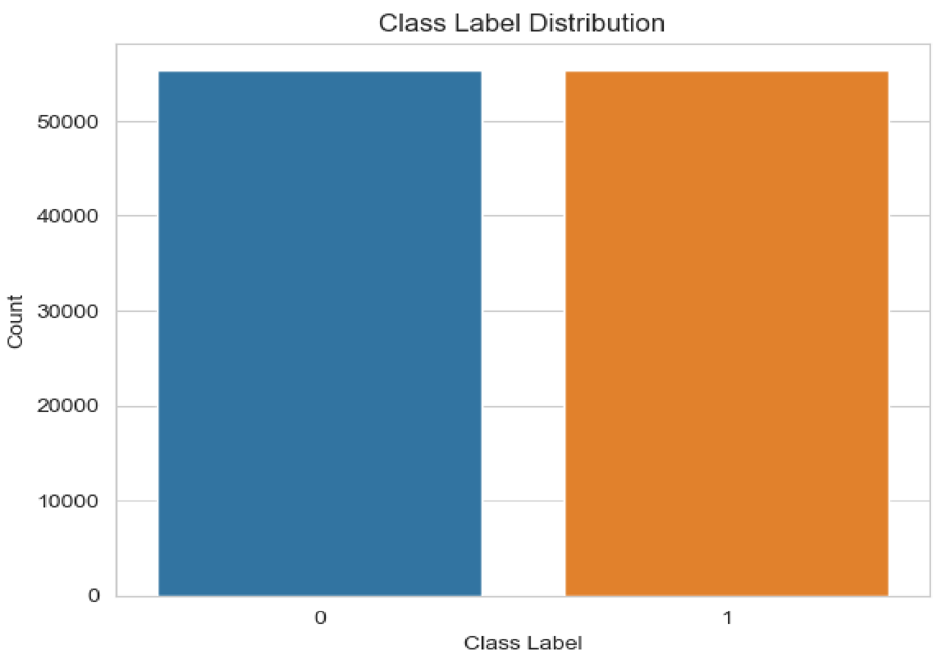

Figure 23 Distribution of Class Variable after SMOTE

Initially, the training data shape was (60000, 171) with 59000 negatives and 1000 positives. After applying this sampling techniques and handling missing values the shape changes to (110754, 163), with equal distribution of negatives and positives as shown in (Figure 23). This more equitable distribution allows for a more accurate and reliable model, as it better captures the nuances of both 'pos' and 'neg' class.

4.1.4 Model Building

4.1.4.1 Model 1: Multi-Layer Perceptron - Keras Sequential neural network

- Hidden Layers Count : 3
- Activation Function : ReLU and Sigmoid
- Regularization : L1
- Optimizer : Adam Optimizer
- Loss Function : Binary Cross entropy
- Learning Rate : 0.01
- Batch Size : 32 (Training)
- Epochs : 10
- Validation Split : 20%

Model was built utilizing a comparable framework to that used in previous studies for classification tasks (Zhao & Xiao, 2021). Keras Sequential neural network designed for binary classification, consisting of an input layer implicitly defined by the training data's feature size, followed by three hidden layers with 128, 64, and 32 neurons respectively, all using ReLU activation and L1 regularization to reduce overfitting. The output layer utilizes a sigmoid activation function, ideal for binary outcomes. It's compiled with Adam optimizer and binary cross entropy loss, including accuracy, precision, recall, and AUC-ROC as performance metrics. To enhance training efficiency and prevent overfitting, an Early Stopping callback is employed, monitoring validation loss with specific parameters for improvement threshold and patience. The model is trained on a specified training dataset with a 20% validation split, using a batch size of 32 and up to 10 epochs, though Early Stopping may truncate training if no improvement in validation loss is observed.

4.1.4.2 Model 2: 1D Convolutional Neural Network

- Hidden Layers Count : 4
- Activation Function : ReLU and sigmoid
- Optimizer : Adam Optimizer
- Loss Function : Binary Cross entropy
- Learning Rate : 0.01

- Batch Size : 32 (Training)
- Epochs : 10
- Validation Spilt : 20%

The constructed model is a 1D Convolutional Neural Network (CNN) using Keras, tailored for binary classification tasks. It begins with a 1D convolutional layer with 32 filters and a kernel size of 3, using 'relu' activation, which processes input data reshaped to fit its requirements. This is followed by a max pooling layer with a pool size of 2 to reduce the dimensionality. The model then adds another 1D convolutional layer with 64 filters and the same kernel size and activation, coupled with another max pooling layer. After these convolutional and pooling layers, the output is flattened to serve as input for the fully connected layers. It includes two densely connected layers with 128 and 64 neurons, respectively, both using 'relu' activation. The final layer is a dense layer with a single neuron and a 'sigmoid' activation function, suitable for binary classification. The model is compiled with the Adam optimizer with a learning rate of 0.01 and uses binary cross-entropy as the loss function. It includes metrics such as accuracy, precision, recall, and AUC for performance evaluation. Training is conducted over 10 epochs with a batch size of 32 and a validation split of 20%. Finally, the model is evaluated on the test data, providing insights into its loss and accuracy in real-world scenarios. This architecture leverages the strength of CNNs in feature extraction from sequence data, aiming for effective learning and classification performance.

4.1.4.3 Model 3: Recurrent Neural Network (RNN) with a Bidirectional LSTM

- Hidden Layers Count : 3
- Activation Function : ReLU and Sigmoid
- Optimizer : Adam Optimizer
- Loss Function : Binary Cross entropy
- Learning Rate : 0.01
- Batch Size : 32 (Training)
- Epochs : 10
- Validation Split : 20%

The developed model is a Recurrent Neural Network (RNN) employing Long Short-Term Memory (LSTM) layers, designed specifically for binary classification. The model architecture is built using Keras and starts with a bidirectional LSTM layer consisting of 64 units, which allows the model to capture dependencies from both past and future information in the sequence. This layer is set to return sequences, enabling the stacking of another LSTM layer. The second LSTM layer is also bidirectional but comprises 32 units and does not return sequences, focusing on capturing higher-level temporal features. Following these recurrent layers, the model includes two densely connected layers with 128 and 64 neurons, respectively, each employing 'relu' activation to introduce non-linearity. The final layer of the model is a dense layer with a single neuron and a 'sigmoid' activation function, appropriate for binary output. The model is compiled using the Adam optimizer with a learning rate of 0.01, and the loss function is binary cross-entropy. Performance metrics such as accuracy, precision, recall, and AUC are included for comprehensive evaluation. The training process involves 10 epochs with a batch size of 32 and a validation split of 20%. Post-training, the model's effectiveness is assessed on test data, providing insights into its loss and accuracy metrics. This RNN architecture, leveraging bidirectional LSTMs and fully connected layers, is designed to efficiently process sequential data, making it well-suited for tasks involving time series or sequence prediction.

4.1.5 Model Training

4.1.5.1 Model 1: Multi-Layer Perceptron - Keras Sequential neural network

```
Final Training Results:
Loss: 0.1590
Accuracy: 96.19%
Recall: 95.90%
AUC-ROC: 98.89%

Final Validation Results:
Loss: 0.1619
Accuracy: 96.34%
Recall: 97.41%
AUC-ROC: 99.06%
```

Figure 24 MLP – Training and Validation Results

The training results show a loss of 0.1590, an accuracy of 96.19%, a recall of 95.0%, and an AUC-ROC of 98.89%. The validation results indicate a slightly higher loss of 0.1619, but improved metrics elsewhere with an accuracy of 96.34%, a recall of 97.41%, and an AUC-ROC of 99.06%. These results suggest the model is performing very well on both the training and validation datasets, with high accuracy and recall, and excellent AUC-ROC scores, indicating a strong ability to distinguish between the two classes. The training was conducted over 10 epochs with a batch size of 32 and a validation split of 20%

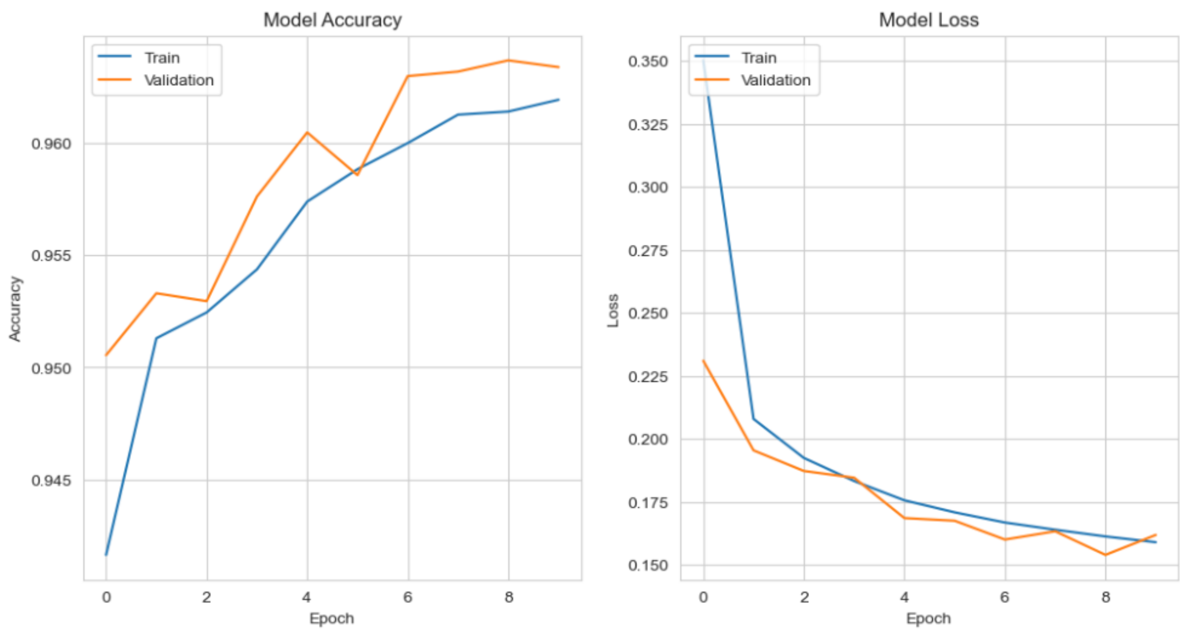

Figure 25 MLP - Training and Validation Loss/ Accuracy Curves

Figure (25) shows two line graphs depicting the model accuracy and model loss over epochs during training. The left graph shows the accuracy of the model, with both training and validation accuracy improving over time, and the validation accuracy slightly outperforming the training accuracy by the end of the training. The right graph shows the loss of the model, with a sharp decrease in the initial epochs followed by a gradual decline. Both training and validation losses decrease over time, with the validation loss slightly lower than the training loss at the end of the training process. These trends indicate a well-fitting model with no signs of overfitting, as the validation metrics are consistent with or slightly better than the training metrics.

4.1.5.2 Model 2: 1D Convolutional Neural Network

```
Final Training Results:
Loss: 0.0195
Accuracy: 99.35%
Recall: 99.24%
AUC-ROC: 99.95%

Final Validation Results:
Loss: 0.0197
Accuracy: 99.39%
Recall: 99.20%
AUC-ROC: 99.94%
```

Figure 26 1D CNN – Training and Validation Results

Above (Figure 26) depicts the final training and validation results for a 1D Convolutional Neural Network (CNN) model used for binary classification. The training results show an impressive loss of 0.0195, with an accuracy of 99.35%, a recall of 99.24%, and an AUC-ROC of 99.95%. The validation results are similarly outstanding, indicating a loss of 0.0197, an accuracy of 99.39%, a recall of 99.20%, and an AUC-ROC of 99.94%. These results suggest that the model is exceptionally well-tuned, with almost perfect accuracy and recall, and an excellent ability to distinguish between the positive and negative classes as evidenced by the high AUC-ROC scores. The close similarity between training and validation results implies that the model generalizes very well to new, unseen data and that there is no significant overfitting or underfitting. The model, which includes 1D convolutional layers, max pooling, flattening, and dense layers, was compiled with the Adam optimizer and trained with batch size of 32 and 10 epochs, with validation split of 20% achieving high performance in both training and validation.

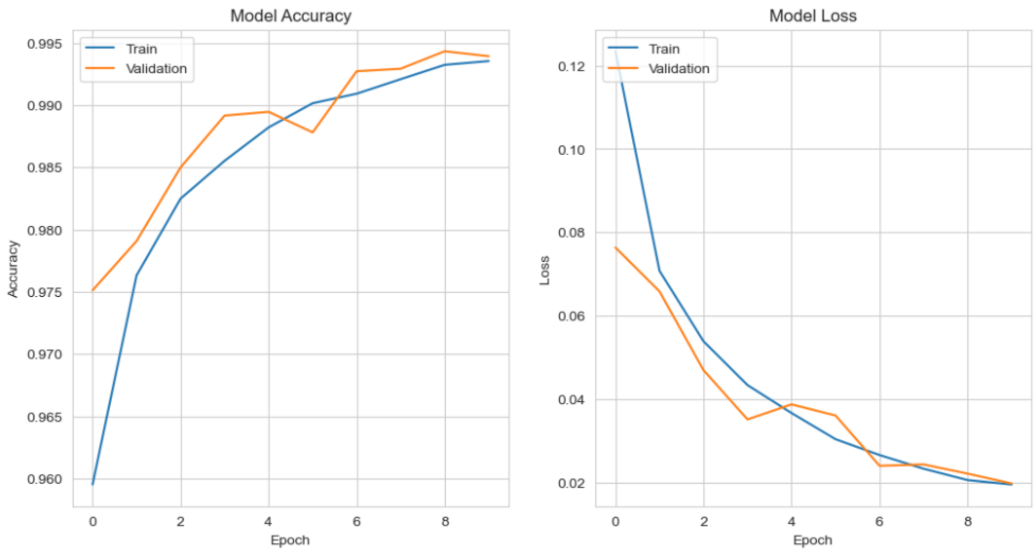

Figure 27 1D CNN - Training and Validation Loss/ Accuracy Curves

Above (Figure 27) shows two graphs, one for model accuracy and the other for model loss, across training epochs for a 1D Convolutional Neural Network model. In the accuracy graph, both training and validation accuracy increase sharply and then plateau, with validation accuracy slightly lagging behind but remaining close to the training accuracy, indicating good generalization. The loss graph depicts a sharp decrease in loss for both training and validation in the initial epochs, with the validation loss decreasing at a slower rate but closely following the training loss. By the end of the training, the model shows very high accuracy and very low loss, with both metrics demonstrating convergence, which suggests that the model has learned effectively and there is no significant overfitting, given the close alignment of the training and validation lines.

4.1.5.3 Model 3: Recurrent Neural Network (RNN) with a Bidirectional LSTM

```
Final Training Results:
Loss: 0.0793
Accuracy: 97.36%
Recall: 97.54%
AUC-ROC: 99.51%

Final Validation Results:
Loss: 0.0740
Accuracy: 97.67%
Recall: 97.75%
AUC-ROC: 99.53%
```

Figure 28 RNN - LSTM Bidirectional – Training and Validation Results

Above (Figure 28) depicts the final training and validation results from a bidirectional LSTM Recurrent Neural Network model. The training phase concluded with a loss of 0.0793, an accuracy of 97.36%, a recall of 97.54%, and an AUC-ROC of 99.51%. The validation phase yielded very similar results, with a loss of 0.0740, an accuracy of 97.67%, a recall of 97.75%, and an AUC-ROC of 99.53%. These outcomes are indicative of a highly effective model that is likely not overfitting, as evidenced by the validation metrics being on par with or slightly better than the training metrics. The high recall and AUC-ROC values suggest that the model is very capable of correctly identifying the positive class while also differentiating well between classes. The consistency between training and validation results also suggests good model generalizability.

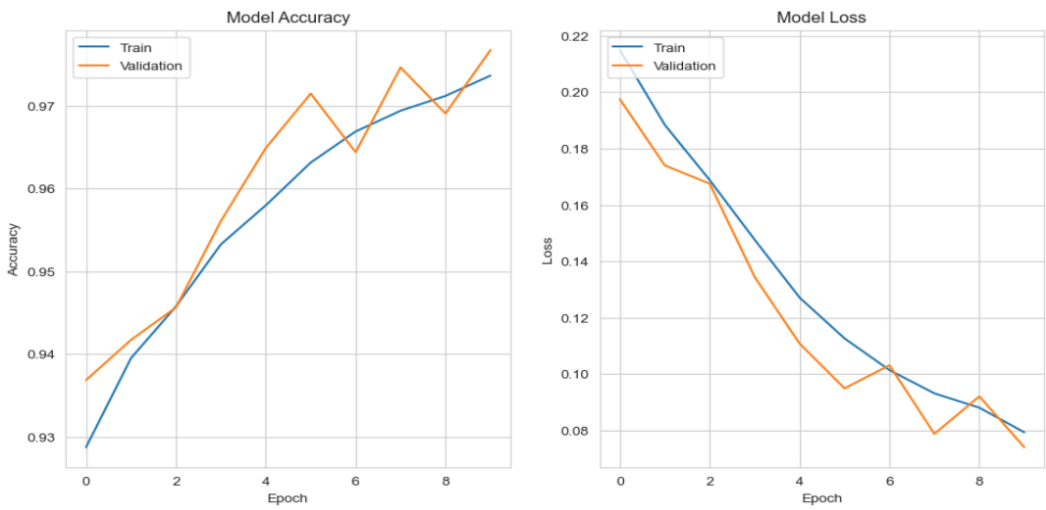

Figure 29 *Training and Validation Loss/ Accuracy Curves*

Above (Figure 29) illustrates the progression of model accuracy and loss over the course of training for a bidirectional LSTM model. The Model Accuracy graph shows a steady increase in accuracy for both the training and validation datasets, with accuracy rates ending slightly below 98%. The Training accuracy surpasses the Validation accuracy slightly by the final epoch. The Model Loss graph depicts a sharp decline in loss for both training and validation, with the validation loss initially higher than the training loss, but both converge to similar values by the end of training. The loss values reach just above 0.08 by the final epoch. Both graphs show a good fit without any apparent overfitting, as indicated by the convergence of training and validation lines.

4.1.6 Model Evaluation

Evaluation phase is a critical step in the CRISP-DM process, where the results of the modeling phase are thoroughly assessed. The primary objective is to evaluate how well the predictive model performs. In the context of the Evaluation phase, the use of a test dataset is particularly important. In this study the test dataset is a separate set of data that has not been used in the training phase of the model. This is crucial for an unbiased evaluation of the model's performance. The test dataset acts as a proxy for new, unseen data. It helps in assessing how the model will perform in real-world scenarios, where it will encounter data, it hasn't been trained on. Various metrics such as accuracy, recall, AUC score, etc., depending on the problem type are used to quantify the model's performance on the test dataset. Evaluating the model on a test dataset can also help in identifying issues like overfitting, where the model performs well on the training data but poorly on new data. It can also highlight other problems like biases in the model.

Loading the test Dataset

```
x_test = pd.read_csv("aps_failure_test_set.csv",skiprows=20, na_values=["na"])
np.random.seed(3482)
```

```
x_test.head()
```

	class	aa_000	ab_000	ac_000	ad_000	ae_000	af_000	ag_000	ag_001	ag_002	...	ee_002	ee_003	ee_004	ee_005	ee_006	ee_007	ee_008
0	neg	60	0.0	20.0	12.0	0.0	0.0	0.0	0.0	0.0	...	1098.0	138.0	412.0	654.0	78.0	88.0	0.0
1	neg	82	0.0	68.0	40.0	0.0	0.0	0.0	0.0	0.0	...	1068.0	276.0	1620.0	116.0	86.0	462.0	0.0
2	neg	66002	2.0	212.0	112.0	0.0	0.0	0.0	0.0	0.0	...	495076.0	380368.0	440134.0	269556.0	1315022.0	153680.0	516.0
3	neg	59816	NaN	1010.0	936.0	0.0	0.0	0.0	0.0	0.0	...	540820.0	243270.0	483302.0	485332.0	431376.0	210074.0	281662.0
4	neg	1814	NaN	156.0	140.0	0.0	0.0	0.0	0.0	0.0	...	7646.0	4144.0	18466.0	49782.0	3176.0	482.0	76.0

5 rows × 171 columns

```
print("shape of x_test:",x_test.shape)
```

```
shape of x_test: (16000, 171)
```

Figure 30 Overview of Test Dataset (Evaluation)

Above (Figure 30) shows that the test data (new data) consists of 16,000 datapoints with 171 features. Jenghara et al. (2017) study on missing values imputation emphasizes the importance and recommendation of applying identical transformations to both the training and test datasets. In this study, missing values more than 70% are dropped and lower than 70% are imputed with similar imputation techniques applied to training dataset. Below (figure) shows us the shape of the test data after applying the imputation techniques.

```
Dropped Features due to high missing values more than 70%: ['br_000', 'bq_000', 'bp_000', 'bo_000', 'ab_000', 'cr_000', 'bn_
000', 'cd_000']

Features with Mean Imputation for Missing Values- 5% to 15% : ['ak_000', 'ca_000', 'dm_000', 'df_000', 'dg_000', 'dh_000',
'dl_000', 'dj_000', 'dk_000', 'eb_000', 'di_000', 'ac_000', 'bx_000', 'cc_000']

Features with MICE Imputation for Missing values - 15%-70% : ['class', 'aa_000', 'ad_000', 'ae_000', 'af_000', 'ag_000', 'ag
_001', 'ag_002', 'ag_003', 'ag_004', 'ag_005', 'ag_006', 'ag_007', 'ag_008', 'ag_009', 'ah_000', 'ai_000', 'aj_000', 'al_00
0', 'am_0', 'an_000', 'ao_000', 'ap_000', 'aq_000', 'ar_000', 'as_000', 'at_000', 'au_000', 'av_000', 'ax_000', 'ay_000', 'a
y_001', 'ay_002', 'ay_003', 'ay_004', 'ay_005', 'ay_006', 'ay_007', 'ay_008', 'ay_009', 'az_000', 'az_001', 'az_002', 'az_00
3', 'az_004', 'az_005', 'az_006', 'az_007', 'az_008', 'az_009', 'ba_000', 'ba_001', 'ba_002', 'ba_003', 'ba_004', 'ba_005',
'ba_006', 'ba_007', 'ba_008', 'ba_009', 'bb_000', 'bc_000', 'bd_000', 'be_000', 'bf_000', 'bg_000', 'bh_000', 'bi_000', 'bj_
000', 'bk_000', 'bl_000', 'bm_000', 'bs_000', 'bt_000', 'bu_000', 'bv_000', 'by_000', 'bz_000', 'cb_000', 'ce_000', 'cf_00
0', 'cg_000', 'ch_000', 'ci_000', 'cj_000', 'ck_000', 'cl_000', 'cm_000', 'cn_000', 'cn_001', 'cn_002', 'cn_003', 'cn_004',
'cn_005', 'cn_006', 'cn_007', 'cn_008', 'cn_009', 'co_000', 'cp_000', 'cq_000', 'cs_000', 'cs_001', 'cs_002', 'cs_003', 'cs_
004', 'cs_005', 'cs_006', 'cs_007', 'cs_008', 'cs_009', 'ct_000', 'cu_000', 'cv_000', 'cx_000', 'cy_000', 'cz_000', 'da_00
0', 'db_000', 'dc_000', 'dd_000', 'de_000', 'dn_000', 'do_000', 'dp_000', 'dq_000', 'dr_000', 'ds_000', 'dt_000', 'du_000',
'dv_000', 'dx_000', 'dy_000', 'dz_000', 'ea_000', 'ec_00', 'ed_000', 'ee_000', 'ee_001', 'ee_002', 'ee_003', 'ee_004', 'ee_0
05', 'ee_006', 'ee_007', 'ee_008', 'ee_009', 'ef_000', 'eg_000']

Shape of Test data after missing value removal: (16000, 162)
```

Figure 31 Overview of Features removed and Features Imputed – Test Dataset

Evaluation Metrics:

Loss: Represents the model's error and is a summary of how well the model's predictions match the actual labels. Lower values are better.

Accuracy: The fraction of predictions our model got right. Higher values are better.

Recall (Sensitivity): The ability of the model to find all the relevant cases within the dataset. Higher values are better, especially for safety-critical applications.

AUC-ROC: Reflects the model's capability to discriminate between the positive and negative classes. Higher values indicate better performance.

4.1.6.1 Evaluation - Model 1: Multi-Layer Perceptron - Keras Sequential neural network

```
Test Results:
Loss: 0.2833
500/500 [==============================] - 1s 2ms/step
Accuracy: 93.69%
Recall: 97.87%
AUC-ROC: 98.55%
```

Figure 32 MLP – Evaluation results

The test results indicate a loss of 0.2833, with an accuracy of 93.69%, a recall of 97.87%, and an AUC-ROC of 98.55%. Comparatively, the training results showed a higher accuracy of 96.19% and recall of 95.90%, with a slightly lower AUC-ROC of 98.89%. The validation results were closer to the training ones, with an accuracy of 96.34%, recall of 97.41%, and an AUC-ROC of 99.06%. Generally, the model performed slightly better on the training and validation sets than on the test set, which is common as models tend to perform best on the data they were trained on. However, the differences are not substantial, indicating that the model has generalized well. The recall is consistently high across all sets, which is particularly important in scenarios where the cost of false negatives is high. The high AUC-ROC values across all sets suggest that the model has a good measure of separability and is capable of distinguishing between the positive and negative classes effectively.

```
Test Loss: 0.3080962896347046
Test Accuracy: 94.14%
Test Recall: 96.53%
Test AUC: 98.02%
```

Figure 33 MLP (Tuned model) – Evaluation results

Further optimization is carried out by tuning hyperparameters. However, the baseline model outperforms the tuned model in terms of test recall and AUC-ROC but has a slightly lower test accuracy. The tuned model shows an improvement in test accuracy, but at the cost of a higher test loss, indicating that the tuning may have led to a slight overfit to the training data, as seen by the larger gap between training/validation and test performance metrics.

Detailed overview of the architecture and results of the tuned model can be seen in Appendix section of this report.

4.1.6.2 Evaluation - Model 2: 1D Convolutional Neural Network

```
Test Loss: 9.81102180480957
Test Accuracy: 98.47%
Test Recall: 81.87%
Test AUC: 93.48%
```

Figure 34 1D CNN – Evaluation results

The test results for the Convolutional Neural Network (CNN) indicate a significantly higher test loss of 9.811 compared to the training and validation losses of 0.0195 and 0.0197, respectively. This suggests that the model is overfitting to the training data and not generalizing well to unseen data. Despite the high test loss, the model achieved a test accuracy of 98.47%, which is slightly lower than the training and validation accuracies of 99.35% and 99.39%. However, there is a substantial drop in recall from 99.24% in training and 99.28% in validation to 81.87% in testing, indicating the model is failing to identify a significant portion of the positive class in the test set. The AUC (Area Under the Receiver Operating Characteristic Curve) for the test set is also lower at 93.48%, compared to 99.95% for training and 99.94% for validation. This decrease in test recall and AUC suggests that while the model is accurate, it's less reliable when it comes to identifying true positives in new data, which is critical in many real-world applications. Overall, while the model shows high accuracy, the high-test loss and lower recall and AUC on the test data point towards overfitting, and measures should be taken to improve the model's ability to generalize.

4.1.6.3 Evaluation - Model 3: Recurrent Neural Network (RNN) with a Bidirectional LSTM

```
Test Loss: 0.110493004322052
Test Accuracy: 96.51%
Test Recall: 89.07%
Test AUC: 96.71%
```

Figure 35 RNN LSTM Bidirectional – Evaluation results

The test results from the Recurrent Neural Network (RNN) with LSTM layers indicate a test loss of 0.1149, with an accuracy of 96.51%, a recall of 89.07%, and an AUC of 96.71%. These metrics are slightly lower than the training results, which show an accuracy of 97.36%, a recall of 97.54%, and an AUC-ROC of 99.51%, as well as the validation results, with an accuracy of 97.67%, a recall of 97.75%, and an AUC-ROC of 99.53%. The model's performance on the test data demonstrates a good level of generalization, with only a small drop in performance metrics compared to the training and validation sets. However, the decrease in recall and AUC on the test data indicates that while the model is accurate, it is slightly less effective at correctly identifying all positive cases (recall) and differentiating between classes (AUC) when presented with new data. The results suggest that the model is well-trained but may benefit from further optimization to improve its recall and AUC on unseen data.

Further optimization of the model was systematically approached by identifying critical hyperparameters and employing advanced search strategies to explore the hyperparameter space. Techniques such as cross-validation were utilized to ensure the model's robustness. However, there is no significant difference in evaluation on test data. Hyperparameter-tuned model shows slightly lower accuracy on the test set compared to the baseline (96.22% vs. 96.51%) but has a higher recall (92.08% vs. 89.87%) and AUC (98.44% vs. 96.71%). This suggests that while the tuned model is slightly less accurate overall, it is better at identifying the relevant cases (higher recall) and distinguishing between the classes (higher AUC).

```
Test Loss: 0.10636764764785767
Test Accuracy: 96.22%
Test Recall: 92.00%
Test AUC: 98.44%
```

Figure 36 RNN LSTM Bidirectional (Tuned Hyperparameter) – Evaluation results

Detailed overview of the architecture and results of the tuned model can be seen in Appendix section of this report.

CHAPTER 5 : DISCUSSION

5.1 Comparative Analysis of Neural network model (MLP, 1D CNN, RNN LSTM)

A comparative analysis of three neural network architectures such as 1D Convolutional Neural Network (CNN), a Multilayer Perceptron (MLP), and a Recurrent Neural Network with Long Short-Term Memory (RNN LSTM) is generated and the performance of each model was evaluated based on standard metrics: loss, accuracy, recall, and the area under the Receiver Operating Characteristic curve (AUC-ROC).

Model Performance:

Table 3 Overview of Evaluation Results on Test data for all Model

S. No	Model Description	Model Evaluation Metrics			
		Loss	Accuracy	Recall	AUC
1	Multilayer Perceptron (MLP) (Baseline Model)	0.2833	93.69%	97.87%	98.55%
2	1D Convolutional Neural Network (CNN)	9.811	98.47%	81.87%	93.48%
3	Recurrent Neural Network (RNN) LSTM (Baseline Model)	0.11049	96.51%	89.87%	96.71%

Multilayer Perceptron (MLP):

Loss: 0.2833 (lowest, indicating the best fit to the test data)

Accuracy: 93.69% (good, but lower than CNN)

Recall: 97.87% (highest, crucial for identifying most true positives)

AUC: 98.55% (highest, excellent discriminative ability)

1D Convolutional Neural Network (CNN):

Loss: 9.811 (significantly higher, indicating potential overfitting or reporting error)

Accuracy: 98.47% (highest, suggesting excellent overall predictive performance)

Recall: 81.87% (lowest, could miss critical true positive predictions)

AUC: 93.48% (moderate, indicating good discriminative ability)

Recurrent Neural Network (RNN) LSTM (Baseline):

Loss: 0.11049 (low, suggesting a good fit)

Accuracy: 96.51% (second-best, indicating high reliability)

Recall: 89.87% (moderate, better than CNN but lower than MLP)

AUC: 96.71% (second highest, showing strong discriminative ability)

DISCUSSION:

The MLP model, despite not having the highest accuracy, shows a strong balance across all metrics, particularly in recall and AUC, which are crucial for the application of predicting brake failure where missing a true failure case could be catastrophic. The high recall rate suggests that the MLP is the most reliable model for ensuring that few failure cases are missed, even though at the potential cost of more false positives.

The CNN, while showing the highest accuracy, has a suspiciously high loss value which could be indicative of an anomaly in the model's evaluation or potential overfitting. Its lower recall rate also suggests it might not be as reliable for detecting all true brake failure cases, which is a significant consideration for this application.

The LSTM model demonstrates a good compromise between the MLP and CNN, with well-rounded performance in all metrics. It has a lower loss compared to the CNN and a better recall rate than the CNN, but it does not outperform the MLP in any metric.

Considering the critical nature of brake failure prediction, where the cost of false negatives (undetected failures) could be extremely high, the MLP model stands out as the most suitable option for this application. Its superior recall and AUC values suggest that it will be the most effective at identifying potential brake failures without missing many actual failure events. However, each model's performance should also be assessed in the context of operational costs and the acceptable threshold for false positives in the real-world deployment of such a system.

5.2 Assessing Predictive Efficacy of Multi-Layer Perceptron for Brake Failure Detection (Chosen Model)

Multi-Layer Perceptron demonstrated substantial potential in improving the anticipation and prevention of brake failures in heavy vehicles.

High Accuracy (93.69%): The model correctly predicts brake failures in the majority of cases. This level of accuracy suggests that the model can be a reliable tool for maintenance teams, allowing them to trust the system's predictions and take preemptive actions to prevent brake failures.

Exceptional Recall (97.87%): High recall indicates that the model is capable of identifying almost all true instances of potential brake failures. This is crucial in a predictive maintenance context because missing a true failure could lead to a breakdown or an accident. A model with high recall ensures that very few failure events go unnoticed, thus significantly improving safety.

Excellent AUC-ROC (98.55%): The Area Under the Receiver Operating Characteristic curve is a measure of a model's ability to distinguish between classes. An AUC-ROC score close to 100% means that the model has excellent capability in differentiating between the classes, in this case, likely brake failure versus non-failure. This suggests that the model can confidently discriminate between the conditions that lead to brake failures and those that don't, which is essential for accurate early warnings.

The combination of these performance metrics indicates a model that is both precise and robust, making it an effective tool for the early detection of potential brake failures. By incorporating this model into the regular maintenance schedule of heavy vehicles, it is possible to conduct repairs or replacements of brake components before a failure occurs, thereby preventing vehicle downtime and potential accidents. Furthermore, the high predictive performance may allow for more efficient use of resources, as maintenance can be scheduled based on actual condition data rather than predefined intervals, leading to a more effective and cost-efficient maintenance process.

However, it's important to note that the practical application of the model also depends on the ability to integrate it into existing systems and workflows, the quality of the incoming data during real-world operations, and the acceptance by the end-users who will interact with the model's predictions.

CHAPTER 6 : RESEARCH CONCLUSION

This research project has achieved its goals, resulting in the development of an innovative and practical neural network-based system for predicting brake failure in Scania trucks. To answer the research question of how neural network technology can enhance the prediction and prevention of brake failure, we diligently adhered to our outlined research objectives:

Objective 1: Comprehensive Literature review: This study initiated by conducting a thorough review of the literature, gaining an in-depth understanding of brake failure mechanisms and exploring the current landscape of neural network applications in predictive maintenance.

Objective 2: Data Collection: A comprehensive dataset was collected, encompassing a wide range of brake parameters and historical failure data from Scania trucks, which included both normal operation scenarios and failure instances to ensure a well-rounded analysis.

Objective 3: Data Preprocessing: Through careful data preparation, we cleaned and normalized the collected data, conducted a comprehensive exploration and analysis. This allowed us to uncover crucial patterns and relationships within the data creating an optimal foundation for the application of machine learning techniques.

Objective 4: Design and implementation of Three Neural network Model: Three distinct neural network models were designed and implemented, each undergoing rigorous training and validation processes to ascertain their accuracy and reliability with results including performance metrics such as loss, accuracy, recall, and AUC-score.

Objective 5: Evaluation of Three neural Network with test dataset: The models were then evaluated against a separate, real-world operational dataset to determine their predictive capabilities, providing valuable insights into their practical utility.

Objective 6: Comparison of Three Neural Network Model Post external testing: An in-depth comparative analysis of the three models was conducted post-testing, leading to the identification of the model with superior predictive accuracy and operational viability.

Objective 7: Conclusion and Strategic Recommendations for Neural Network-Based Predictive Maintenance in Heavy Vehicles: Finally, informed by the analysis, we selected the most effective model, drew conclusions about its practical effectiveness in operational environments, and formulated strategic recommendations for its deployment and for future research directions.

By fulfilling these research objectives, we have conclusively addressed the research aim "To develop a cutting-edge neural network model capable of accurately classifying and predicting brake failures in Scania trucks before they lead to failure, thereby enhancing vehicular safety, reducing downtime, and potentially saving lives by preempting on-road breakdowns and accidents" and the question "How can neural network-based predictive model improve the anticipation and prevention of brake failures in heavy vehicles?". The deliverable is a state-of-the-art deep learning algorithm focused on predictive maintenance for Scania trucks, designed to advance the prevention of brake failures.

The deployed Multi-Layer Perceptron model, trained on a comprehensive dataset, has exhibited exceptional performance, achieving notable metrics in predicting brake failure incidents with a Accuracy of 93.69%, Recall of 97.87%, and an AUC-ROC score of 98.55%. These results underscore the model's capability in effectively predicting brake failures, thereby significantly contributing to the field of predictive maintenance. Consequently, this research has substantially bridged the knowledge gap, providing a valuable and reliable tool for the heavy vehicle industry, enhancing safety, and optimizing maintenance operations.

6.1 Recommendation for Future Works

Based on the results obtained from the research, few recommendations are presented for future scope of works as follows

Integration into Maintenance Workflows: It is recommended that Scania integrates the predictive maintenance system into their regular maintenance workflows to ensure continuous monitoring and early detection of potential brake failures.

Ongoing Model Training and Updating: To maintain the accuracy and relevance of the predictive model, it is suggested that the model be regularly updated with new data, to adapt to changing patterns in brake wear and failure modes.

Expansion to Other Vehicle Systems: Considering the success of the model in predicting brake failures, a similar approach could be used to predict other critical component failures within the trucks, such as engine or transmission systems, to provide a comprehensive predictive maintenance suite.

Training for Technicians and Operators: Develop training programs for maintenance technicians and operators to effectively interpret the predictive model’s outputs and take appropriate preventive actions.

Feedback Mechanisms: Implement feedback mechanisms to continuously collect data on the system’s performance, which can be used for further improvement and validation of the predictive model.

Further Research: Further research is recommended to explore the long-term impacts of using such a predictive maintenance system on the operational costs and safety records of trucking operations.

These recommendations aims to maximize the impact of the research findings and to ensure that the developed predictive maintenance system can be effectively used to improve safety, reliability, and efficiency in the heavy vehicle industry.

6.2 Research Limitations

Time Restrictions: The research timeline limited the ability to explore a broader range of methodologies and to extend the model training and validation phases to achieve potentially better results.

Data Limitations: The accuracy and robustness of the neural network are highly dependent on the quality and quantity of the data used for training. Any limitations in the dataset, such as biases, missing values, or lack of representation of all possible failure scenarios, could affect the model's performance

Dependency on Historical Data: Predictive accuracy of the model is contingent on the historical data it has been trained on. The higher reliance upon historical data might not capture the emerging trends in braking system technology. If the model has not been exposed to certain types of brake failure scenarios during training, it might fail to predict them accurately.

Model Generalization: While the model demonstrates high accuracy on the test dataset, its ability to generalize to other truck models, brands, or different environmental conditions may be limited. Generalization is a common challenge in machine learning and requires extensive validation.

Real-World Validation: The model has been tested in a controlled environment, and its performance in real-world scenarios may vary. Real-world factors such as sensor noise, unexpected operating conditions, and data transmission errors can affect the model's predictive capabilities.

Field Testing and Validation: Conducting extensive field tests to validate the model's predictions can be time-consuming and expensive, requiring collaboration with truck operators and possibly interrupting regular operation.

References

Abarro, C.C., Caliwag, A.C., Valverde, E.C., Lim, W. and Maier, M., 2022. Implementation of IoT-Based Low-Delay Smart Streetlight Monitoring System. *IEEE Internet of Things Journal*, *9*(19), pp.18461-18472.

Agarwal H. & et al., 2016. Artificial neural network (ANN)-based prediction of depth filter loading capacity for filter sizing. *Biotechnology Prog*ress, 32, pp. 1436–1443.

Ak, R. & et. al., 2013. A genetic algorithm and neural network technique for predicting wind power under uncertainty. *Chem Eng Trans*action, Vol. 33, pp. 925–930.

Arena, F., Collotta, M., Luca, L., Ruggieri, M. and Termine, F.G. (2022). Predictive Maintenance in the Automotive Sector: A Literature Review. *Mathematical and Computational Applications*, [online] 27(1), p.2. doi:https://doi.org/10.3390/mca27010002.

Basham, J.D., Gardner, J.E. and Smith, S.J., 2020. Measuring the implementation of UDL in classrooms and schools: Initial field test results. *Remedial and Special Education*, *41*(4), pp.231-243.

Bhatikar, S. & et. al., 2018. Artificial neural network-based energy storage system modeling for hybrid electric vehicles. SAE Technical paper 2000-01-1564.

Bowlin, C. L., Subramanian, S., Darbha, S. & Rajagopal, K. R. (2006). Pressure control scheme for air brakes in commercial vehicles. *IEE Proceedings - Intelligent Transport Systems 153*(1):21- 32.

Braun, V. & Clarke, V., 2019. Reflecting on reflexive thematic analysis. *Qualitative Research in Sport, Exercise, and Health, Vol. 11*, Bo, 4, pp. 589-597.

Carta, M.G., Ghacem, R., Milka, M., Moula, O., Staali, N., Uali, U., Bouakhari, G., Mannu, M., Refrafi, R., Yaakoubi, S. and Moro, M.F., 2020. Implementing WHO-quality rights project in Tunisia: results of an intervention at Razi hospital. *Clinical practice and epidemiology in mental health: CP & EMH*, *16*(Suppl-1), p.125.

Chen, L., & Liu, Y. (2019). Mitigation strategies for brake failures in heavy-duty trucks: A review. *Journal of Transportation Engineering, Part C: Civil Engineering,* 27(4), p.04019023.

Creswell, J. W., 2013. *Research Design: Qualitative, Quantitative, and Mixed Methods Approaches.* 4th Edition, New Jersey: Sage Publications Inc.

Crotty, M., 2013. *The Foundations of Social Research: Meaning and Perspectives in the Research Process*. 3rd ed. London: Sage, p.10.

Dammak, A., 2015. Research Paradigms: Methodologies and Compatible Methods. *The Academic Journal of St Clement Education Group*, 6(2).

Damschroder, L.J., Reardon, C.M., Widerquist, M.A.O. and Lowery, J., 2022. The updated Consolidated Framework for Implementation Research based on user feedback. *Implementation science*, *17*(1), p.75.

Davis, K.A., Coleman, J.R., Adams, M., Allen, N., Breen, G., Cullen, B., Dickens, C., Fox, E., Graham, N., Holliday, J. and Howard, L.M., 2020. Mental health in UK Biobank–development, implementation and results from an online questionnaire completed by 157 366 participants: a reanalysis. *BJPsych open*, *6*(2), p.e18.

de Miranda Costa, M.M., Santana, H.T., Hernandez, P.S., Carvalho, A.A. and da Silva Gama, Z.A., 2020. Results of a national system-wide quality improvement initiative for the implementation of evidence-based infection prevention practices in Brazilian hospitals. *Journal of Hospital Infection*, *105*(1), pp.24-34.

Dhar, S., 2010. *Development of a diagnostic algorithm for air brakes in trucks*. PhD Thesis, Texas A & M University, College Station, TX, USA.

Dunn, AL., Guenther, D. & Radlinski, R., 2019. Application of air brake performance relationships in accident reconstruction and their correlation to real vehicle performance. *SAE Int J Commer Veh*icles, 5, pp. 251–259.

Dzinamarira, T., Kamanzi, C. and Mashamba-Thompson, T.P., 2020. Key stakeholders' perspectives on implementation and scale up of HIV self-testing in Rwanda. *Diagnostics*, *10*(4), p.194.

Easterby-Smith, M. & et al., 2018. *Management Research*. 3rd ed. London: SAGE Publications Ltd.

Eriksson, L. & Nielsen, L., 2014. Modeling and control of engines and drivelines. Hoboken: Wiley, 2014.

Fei J, Zhao N, Shi Y, et al., 2016. Compressor performance prediction using a novel feed-forward neural network based on Gaussian kernel function. *Adv Mech Eng*, 8, pp.1–14.

Flick, U., 2012. *An Introduction to Qualitative Research*. Thousand Oaks, CA: SAGE.

Gondek, C., Halfner, D. & Sampson, O. (2016). Prediction of Failures in the Air Pressure System of Scania Trucks Using a Random Forest and Feature Engineering. *Adv. Intell. Data Anal. 9897,* 398–402 DOI:10.1007/978-3-319-46349-0_36

Hellstrom, E., Ivarsson, E., Nielsen, L. & Aslund, J., 2019. "Look-ahead ¨ control for heavy trucks to minimize trip time and fuel consumption". *Control Engineering Practice, vol. 17,* pp. 245–254.

Holt, M., Lee, E., Lea, T., Bavinton, B., Broady, T., Mao, L., MacGibbon, J., Keen, P., Murphy, D., Bear, B. and Crawford, D., 2020. HIV preexposure prophylaxis cascades to assess implementation in Australia: results from repeated, national behavioral surveillance of gay and bisexual men, 2014–2018. *JAIDS Journal of Acquired Immune Deficiency Syndromes*, *83*(3), pp.e16-e22.

Jarrett, RP. & Clark, N., 2012. *Weighting of parameters in artificial neural network prediction of heavy-duty diesel engine emissions.* SAE Technical paper 2002-02-2878.

Kafunah, J., Ali, M. I. & Breslin, J. G. (2021). Handling Imbalanced Datasets for Robust Deep Neural Network-Based Fault Detection in Manufacturing Systems. *Appl. Sci. 11*, 9783. https://doi.org/10.3390/app11219783

Kalogirou, S., Chondros, T. & Dimarogonas, A., 2020. Development of an artificial neural network-based fault diagnostic system of an electric car. SAE Technical paper 2000-01-1055.

Kirches, C., Bock, H., Schloder, J. & Sager, S., 2013. "Mixed-integer ¨ nmpc for predictive cruise control of heavy-duty trucks" in 2013 European Control Conference, July 2013, pp. 4118–4123.

Klaiman, T., Silvestri, J.A., Srinivasan, T., Szymanski, S., Tran, T., Oredeko, F., Sjoding, M.W., Fuchs, B.D., Maillie, S., Jablonski, J. and Lane-Fall, M.B., 2021. Improving prone positioning for severe acute respiratory distress syndrome during the COVID-19 pandemic. An implementation-mapping approach. *Annals of the American Thoracic Society*, *18*(2), pp.300-307.

Kotlicka-Antczak, M., Podgórski, M., Oliver, D., Maric, N.P., Valmaggia, L. and Fusar-Poli, P., 2020. Worldwide implementation of clinical services for the prevention of psychosis: the IEPA early intervention in mental health survey. *Early intervention in psychiatry*, *14*(6), pp.741-750.

Lammert, M., Duran, A., Diez, J., Burton, J. & Nicholson, A., 2014. "Effect of platooning on fuel consumption of class 8 vehicles over a range of speeds, following distances, and mass". *SAE International Journal of Commercial Vehicles, Vol. 7,* pp. 626–639.

Li, X., Zhang, Y., & Chen, W. (2019). Predicting brake pad wear using machine learning techniques. *International Journal of Vehicle Systems Modelling and Testing,* 34(2), pp.156-168.

Limpert, R., 2019. *Brake design and safety*. Warrendale, PA: Society of Automotive Engineers.

Ljung, L. & Glad, T., 2016. Modeling and Identification of Dynamic Systems. Lund: Studentlitteratur.

Maher, C. & et al., 2018. Ensuring Rigor in Qualitative Data Analysis. *International Journal of Qualitative Methods, Vol. 17*, No. 1, pp. 1-13.

Martinez, C., & Williams, R. (2018). Investigating the impact of brake failure on heavy-duty truck accidents. *Transportation Research Record,* 2679(1), pp.15-24.

McHugh, S., Presseau, J., Luecking, C.T. and Powell, B.J., 2022. Examining the complementarity between the ERIC compilation of implementation strategies and the behaviour change technique taxonomy: a qualitative analysis. *Implementation Science*, *17*(1), pp.1-23.

Mohammadpour, J., Franchek, M. & Grigoriadis, K., 2012. A survey on diagnostic methods for automotive engines. *Int J Engine Resources*, 13, 41–64.

Oh, E. & Lee, H. (2020). An Imbalanced Data Handling Framework for Industrial Big Data Using a Gaussian Process Regression-Based Generative Adversarial Network. *Symmetry 12*(669), 1-19. doi:10.3390/sym12040669

Patrick, DTO. & Andre, K., 2020. *Practical reliability engineering*. 5th ed. Hoboken, NJ: John Wiley & Sons.

Pellitteri, F., Campagna, N., Castiglia, V., Damiano, A. and Miceli, R., 2020. Design, implementation and experimental results of an inductive power transfer system for electric bicycle wireless charging. *IET Renewable Power Generation*, *14*(15), pp.2908-2915.

Prakash, A., Patil, K. & Kalyani, U., 2013. Artificial neural network-based driver modelling for vehicle systems. SAE Technical paper 2013-01-2860.

Rafsunjani, S., Safa, R. S., Imran, A. A., Rahim, S. & Nandi, D. (2019). An Empirical Comparison of Missing Value Imputation Techniques on APS Failure Prediction. *I.J. Information Technology and Computer Science, 2*, 21-29

Rahi, S., 2017. Research design and methods: A systematic review of research paradigms, sampling issues, and instruments development. *International Journal of Economics & Management Sciences, Vol. 6,* pp. 1-5.

Rajaram, V. & Subramanian, S., 2015. A model-based rear-end collision avoidance algorithm for heavy commercial road vehicles. *Proc IMechE, Part D: J Automobile Engineering*, 229, pp. 550–562.

Raveendran, R., Suresh, A., Rajaram, V. and Subramanian, S.C. (2018). Artificial neural network approach for air brake pushrod stroke prediction in heavy commercial road vehicles. *Proceedings of the Institution of Mechanical Engineers, Part D: Journal of Automobile Engineering*, 233(10), pp.2467–2478. doi:https://doi.org/10.1177/0954407018794594.

Rawat , S.D. (2020). Predict component failure related with Air Pressure System at Scania Trucks using various machine learning methods. *Faculty of Engineering Environment and Computing* .

Rawat, S. D. (2020). Predict component failure related with Air Pressure System at Scania Trucks using various machine learning methods. DOI:10.13140/RG.2.2.29587.20003

Reddy, S., Rogers, W., Makinen, V.P., Coiera, E., Brown, P., Wenzel, M., Weicken, E., Ansari, S., Mathur, P., Casey, A. and Kelly, B., 2021. Evaluation framework to guide implementation of AI systems into healthcare settings. *BMJ health & care informatics*, *28*(1).

Roelfsema, M., van Soest, H.L., Harmsen, M., van Vuuren, D.P., Bertram, C., den Elzen, M., Höhne, N., Iacobuta, G., Krey, V., Kriegler, E. and Luderer, G., 2020. Taking stock of national climate policies to evaluate implementation of the Paris Agreement. *Nature communications*, *11*(1), p.2096.

Rogers, L., De Brún, A. and McAuliffe, E., 2020. Defining and assessing context in healthcare implementation studies: a systematic review. *BMC Health Services Research*, *20*(1), pp.1-24.

Russell, S., 2016. Artificial Intelligence: a Modern Approach. Pearson Education Limited.

Scania. (2016). *Scania Safari.* [Online] Available from https://www.scania.com/content/dam/scanianoe/market/ke/experience-scania/publications/Scania-News-3.pdf [Accessed November 04, 2023].

Semerikov, S., Mintii, M. and Mintii, I., 2021. Review of the course "Development of Virtual and Augmented Reality Software" for STEM teachers: implementation results and improvement potentials. CEUR Workshop Proceedings.

SERRADILLA, O., ZUGASTI, E. and ZURUTUZA, U. (2020). *Deep learning models for predictive maintenance: a survey, comparison, challenges and prospect*. [online] Available at: https://arxiv.org/pdf/2010.03207.pdf [Accessed 26 Dec. 2023].

Silva, W. and Capretz, M. (2019). Assets Predictive Maintenance Using Convolutional Neural Networks. *IEEE*. doi:https://doi.org/10.1109/snpd.2019.8935752.

Smith, J. D., & Johnson, A. B. (2020). Brake failure in heavy-duty trucks: Causes and prevention. *Journal of Commercial Vehicle Safety,* 25(2), pp.45-61.

Smith, J., Johnson, A., & Brown, M. (2018). Predictive modeling of braking distances in real-world scenarios. *Journal of Automotive Engineering,* 42(3), pp.123-136.

Smith, J.D., Li, D.H. and Rafferty, M.R., 2020. The implementation research logic model: a method for planning, executing, reporting, and synthesizing implementation projects. *Implementation Science*, *15*, pp.1-12.

Subramanian, S., 2016. *A diagnostic system for air brakes in commercial vehicles*. PhD Thesis, Texas A & M University, College Station, TX, USA.

Tercan, H. and Meisen, T. (2022). Machine learning and deep learning based predictive quality in manufacturing: a systematic review. *Journal of Intelligent Manufacturing*, 33, pp.1879–1905. doi:https://doi.org/10.1007/s10845-022-01963-8.

Turri, V., 2015. “Fuel-efficient and safe heavy-duty vehicle platooning through look-ahead control”. Licentiate Thesis, KTH Royal Institute of Technology.

Uska, M., Wirasasmita, H., Usuluddin, U. & Arianti, B., 2020. “Evaluation of rapidminer-aplication in data mining learning using Persia model”. *Edumatic: Journal Pendidikan Informatika, Vol. 4,* no. 2, pp. 164–171.

Valueva, M.V., Nagornov, N.N., Lyakhov, P.A., Valuev, G.V. and Chervyakov, N.I., 2020. Application of the residue number system to reduce hardware costs of the convolutional neural network implementation. *Mathematics and computers in simulation*, *177*, pp.232-243.

Westerhof, B. & Kalakos, D., 2017. *Heavy vehicle braking using friction estimation for controller optimization*. Gothenburg: Chalmers University of Technology.

Wu, J., Hu, K., Cheng, Y., Wang, J., Deng, C. and Wang, Y. (2019). Ensemble Recurrent Neural Network-Based Residual Useful Life Prognostics of Aircraft Engines. *Structural Durability & Health Monitoring*, 13(3), pp.317–329. doi:https://doi.org/10.32604/sdhm.2019.05571.

Yazdinejad, A., & et.al., 2023. “Secure intelligent fuzzy blockchain framework: Effective threat detection in IoT networks”. *Computers in Industry, Vol. 144*, pp. 103801.

Yazdinejad, R., Parizi, A., Dehghantanha, G., Srivastava, S., Mohan, K. & Rababah, A., 2020. “Cost optimization of secure routing with untrusted devices in software defined networking”. Journal of Parallel and Distributed Computing, vol. 143, pp. 36–46, 2020.

Yazdinejad, R., Parizi, A., Dehghantanha, Q. & Choo, R., 2020. “An energy-efficient controller architecture forgot networks with blockchain-based security”. *IEEE Transactions on Services Computing, Vol. 13,* no. 4, pp. 625–638.

Zhang, B., Zhang, S. and Li, W. (2019). Bearing performance degradation assessment using long short-term memory recurrent network. *Computers in Industry*, 106, pp.14–29. doi:https://doi.org/10.1016/j.compind.2018.12.016.

Zhang, R., Peng, J., Li, H., Chen, B., Liu, W., Huang, Z. and Wang, J. (2021). A predictive control method to improve pressure tracking precision and reduce valve switching for pneumatic brake systems. *Iet Control Theory and Applications*, 15(10), pp.1389–1403. doi:https://doi.org/10.1049/cth2.12130.

APPENDIX A – Research Project Plan

Project Title: Prediction of Brake Failure in Scania trucks using Neural Networks

Madhankumar Duraisamy

2070486

Project Start: 17/10/23

Display Week: 1

TASK	START	END
Phase 1 : Literature review & Research Planning		
Research on Brake Failure and Neural Networks	17/10/23	23/10/23
Define Research Problems, question and objectives	24/10/23	28/10/23
Research feature extraction / engineering methodologies.	29/10/23	4/11/23
Phase 2 Data Collection and Analysis		
Collect data from data sources.	5/11/23	9/11/23
Conduct missing data analysis, correlation analysis	10/11/23	15/11/23
Preprocesses the extracted data and Impute data.	16/11/23	22/11/23
Interpret and discuss results	23/11/23	30/11/23
Phase 3 : Training and Validation, Evaluation		
Build Neural Netowrk models	1/12/23	9/12/23
Train the model and Validate it	10/12/23	15/12/23
Compare models and choose best model	16/12/23	20/12/23
Evaluate the model with test data	21/12/23	25/12/23
Phase 4 : Summarize the findings		
Conclude findings based on the results obtained.	26/12/23	1/1/24
Document the findings and conclusions, recommendation	2/1/24	12/1/24
Dissertation		
Write and submit the dissertation for review	5/11/23	22/1/24