EXAMINE NEURAL NETWORK APPROACHES FOR UNIFIED MEMBERSHIP INTEGRATION IN DIVERSE APPLICATIONS

Munkhgerel Ankhtuvshin

A Thesis Submitted in Partial Fulfilment

of the requirements for the

Degree of

Master of Science in Data Analytics

Diagram

Description automatically generated with medium confidence

December 2024

Supervisor: Vikas Tomer

EXAMINE NEURAL NETWORK APPROACHES FOR UNIFIED MEMBERSHIP INTEGRATION IN DIVERSE APPLICATIONS

**ABSTRACT**

This research integrates neural networks across different applications into unified membership systems, particularly in the retail industry. The study seeks to address challenges in data fragmentation, privacy, and scalability by propounding machine learning models, such as XGBoost and neural networks, in predicting customer behavior and unifying loyalty programs. A mixed-method approach was taken by using transactional data with survey insights. Findings suggest that XGBoost performs better than neural networks in predictive accuracy but neural networks allow for adaptability to more complex data sets. This paper contributes to both the scholarly literature and practice by providing an outline of how businesses may leverage data-driven strategies for improving customer engagement and satisfaction. Some limitations include increased computational resources and domain specificity. Future work will likely include neural network architecture optimization and further extension of these results to other industries.

***[Keywords*-**Neural Networks**,** XGBoost**,** Unified Membership Systems**,** Customer Behavior Prediction**,** Loyalty Program Integration**,** Machine Learning]

**ACKNOWLEDGEMENT**

# List of Acronyms

|  |  |
| --- | --- |
| Acronym | Meaning |
| **AI** | Artificial Intelligence |
| **ANN** | Artificial Neural Network |
| **CNN** | Convolutional Neural Network |
| **CLV** | Customer Lifetime Value |
| **DL** | Deep Learning |
| **DNN** | Deep Neural Network |
| **KNN** | K-Nearest Neighbors |
| **LSTM** | Long Short-Term Memory |
| **MAE** | Mean Absolute Error |
| **ML** | Machine Learning |
| **MSE** | Mean Squared Error |
| **PCA** | Principal Component Analysis |
| **R²** | Coefficient of Determination (R-squared) |
| **RF** | Random Forest |
| **RFM** | Recency, Frequency, Monetary |
| **RMSE** | Root Mean Squared Error |
| **SHAP** | SHapley Additive exPlanations |
| **SSE** | Sum of Squared Errors |
| **SVM** | Support Vector Machine |
| **XGBoost** | eXtreme Gradient Boosting |

**TABLE OF CONTENTS**

[Chapter 1: Introduction 10](#_Toc184478587)

[1.1 Background Information 10](#_Toc184478588)

[1.2 Key Findings in Literature 10](#_Toc184478589)

[1.3 Need for Further Investigation 11](#_Toc184478590)

[1.4 Research Questions 11](#_Toc184478591)

[1.5 Aim of the Study 12](#_Toc184478592)

[1.6 Objectives 12](#_Toc184478593)

[1.7 Scope of the Study 12](#_Toc184478594)

[1.8 Limitations 13](#_Toc184478595)

[1.9 Definition of Key Terms 13](#_Toc184478596)

[1.10 Chapter Outline 14](#_Toc184478597)

[1.11 Significance of the Study 15](#_Toc184478598)

[Chapter 2: Literature Review 15](#_Toc184478599)

[2.1 Introduction 15](#_Toc184478600)

[2.2 Empirical Study 16](#_Toc184478601)

[2.3 Theories and Models 17](#_Toc184478602)

[2.4 Literature Gap 21](#_Toc184478603)

[2.5 Conclusion 23](#_Toc184478605)

[Chapter 3: Methodology 23](#_Toc184478606)

[3.1 Introduction 23](#_Toc184478607)

[3.2 Method Outline 24](#_Toc184478608)

[3.3 Research Philosophy 24](#_Toc184478609)

[3.4 Research Approach 24](#_Toc184478610)

[3.5 Research Design 25](#_Toc184478611)

[3.6 Research Method 25](#_Toc184478612)

[3.7 Data Collection Method 25](#_Toc184478613)

[3.8 Data Preprocessing and Feature Engineering 25](#_Toc184478614)

[3.9 Model Development 26](#_Toc184478615)

[3.10 Research Ethics 27](#_Toc184478616)

[3.11 Research Limitations 27](#_Toc184478617)

[3.12 Conclusion 28](#_Toc184478618)

[Chapter 4: Findings and Analysis 29](#_Toc184478619)

[4.1 Introduction 29](#_Toc184478620)

[4.2 Analysis 30](#_Toc184478621)

[4.2.1 Overview of Model Performance 30](#_Toc184478622)

[4.2.2 Survey Data Analytics 33](#_Toc184478623)

[4.2.3 Descriptive Statistics of Predictions 35](#_Toc184478624)

[4.2.4 Evaluation of Model Predictions 37](#_Toc184478624)

[4.3 Conclusion 38](#_Toc184478625)

[Chapter 5 Discussion of Findings 39](#_Toc184478626)

[5.1 Introduction 39](#_Toc184478627)

[5.2 XGBoost vs. Neural Network: A Comparative Discussion 40](#_Toc184478628)

[5.3 Potential Improvements for the Neural Network 43](#_Toc184478629)

[5.4 Implications of Findings 44](#_Toc184478630)

[5.5 Conclusion 47](#_Toc184478631)

[Chapter 6: Conclusions 48](#_Toc184478632)

[6.1 Introduction 48](#_Toc184478633)

[6.2 Linking with Objectives 49](#_Toc184478634)

[6.3 Recommendations 50](#_Toc184478635)

[6.4 Conclusion 51](#_Toc184478636)

[Reference List 53](#_Toc184478637)

[Appendices 62](#_Toc184478638)

[Appendix 1: Visualization of Testing Results 62](#_Toc184478639)

[Appendix 2: Visualization of Testing Results (Residuals vs. Order) 62](#_Toc184478640)

[Appendix 3: clusters using the elbow method 63](#_Toc184478641)

**LIST OF FIGURES**

[Figure 1: Model Evaluation Results 29](#_Toc184389988)

[Figure 2: Model Evaluation Results of XGBoost and neural network 30](#_Toc184389989)

[Figure 3: Elbow Method Plot for Determining Optimal Number of Clusters 33](#_Toc184389990)

[Figure 4: Basic Statistics 34](#_Toc184389991)

[Figure 5: Correlation Matrix of Survey Data 37](#_Toc184389992)

[Figure 6: XGBoost Feature importance 39](#_Toc184389993)

[Figure 7: Scatter Plot for XGBoost Predictions 40](#_Toc184389994)

[Figure 8: Scatter Plot for Neural Networks Predictions 41](#_Toc184389995)

[Figure 9: Training and validation graph 42](#_Toc184389996)

[Figure 10: Bar graph of Feature Importance 45](#_Toc184389997)

[Figure 11: Graph of Model Loss and MAE 46](#_Toc184389998)

**Chapter 1: Introduction**

**1.1 Background Information**

Every business requires customer loyalty to ensure repeated buying habits and good working brand relationship. This turns into positive customer programs, rewards and privileges hence encouraging customers’ interaction and loyalty. However, there is a crucial limitation, and they are data fragmentation. Since customers subscribe to many programs by different brands and retailers, organizations are able to compile a single profile of customer needs and activities. This situation complicates their task of providing tailored services which in the highly competitive environment has become crucial. To resolve this problem, a primary objective of several enterprises is the integration of fragmented loyalty schemes into one networked solution. The mentioned type of system can collect customers’ data and help businesses work on behavioral analysis and suitable rewards. However, it contains striking technical complexities that are associated with data quality issues, different data structures, data privacy aspects, and many more. To these complexities, neural networks or a subfield of machine learning has potential solutions. Because they possess powerful computing abilities and can recognize elaborated patterns, they can also be applied to combine multiple data types and estimate future buyers’ actions. By applying neural networks, companies can observe the genuine time customer preferences, and create a sophisticated membership that implies efficient customer experiences.

**1.2 Key Findings in Literature**

Previous literature reviews show that neural networks have significant potential in many areas yet have not been used significantly in analyzing customer behavior, predicting customer churn, and forecasting customer purchases. For instance, first, a comparison was made between Recurrent Neural Networks (RNNs) and Deep Learning techniques, which was followed by a demonstration that extant statistical approaches are inefficient in analyzing time-series data, detecting intricate patterns, and achieving increased levels of prediction accuracy as well (Samek *et al*., 2021). Nevertheless, few academic works have been focused on using neural networks for the purpose of consolidating and handling massive unified membership systems. Existing works tend to explore aspects of neural networks that provide prediction of the customers’ actions in general or improve the functioning of the loyalty program whereas the critical issues of data consolidation that have to be conducted on the central level are excluded. This research therefore raises a gap for more research studies that aims to develop single and integrated systems using neural networks. During the literature review find out a machine learning model XGBoost could be more effective in some aspects of this work. This research incorporates both neural networks and XGBoost to analyse customer data, combining the strengths of both models for a more robust and effective solution. This dual approach aims to address gaps in existing studies and develop a unified system using these advanced techniques.

**1.3 Need for Further Investigation**

The complexity of the customer behavior and the need for individual approach to the client require new approaches to loyalty program administration. Organizations need tools that would also combine membership information with those that are more flexible owing to the constantly changing customers’ demands. Such capabilities can be met by the neural networks which are highly flexible and scalable but the use of the same in context of integrated membership systems is relatively unchartered (Su *et al*., 2022). Other considerations like data compatibility, handling big data, data protection complicate the need for good frameworks. Furthermore, there are no clear guidelines on how to apply neural-network based systems in loyalty programs, and this gap provides scope for research to fill the gaps. From the customers' perspective, having a single, unified application for loyalty programs across various stores is highly desirable. Managing multiple loyalty apps on a phone is inconvenient and cumbersome. A unified system would not only simplify their experience but also encourage more active engagement, as it eliminates the hassle of switching between different applications. This approach would benefit both customers and businesses, fostering stronger relationships and improving overall customer satisfaction.

**1.4 Research Questions**

This study seeks to address the following research questions:

1. How can machine learning models like XGBoost and neural networks enhance customer value prediction and engagement in loyalty programs?
2. What are the challenges associated with integrating membership systems using machine learning models such as XGBoost and neural networks?
3. How can businesses leverage unified systems to improve customer satisfaction and loyalty, with the help of XGBoost and neural network models?

**1.5 Aim of the Study**

The main purpose of this study is to propose a machine learning approach, including, but not limited to, neural networks, to forecast customer value and merge loyalty program data across channels. In addressing the existing systems’ drawbacks, this study seeks to give valuable information concerning customer communications’ improvement with regard to engagement, satisfaction, and retention.

**1.6 Objectives**

The objectives of this study are:

* To evaluate the effectiveness of neural networks and XGBoost in predicting customer behavior and value.
* To identify and address the technical challenges in integrating membership systems.
* To develop and test a unified framework that combines transactional and survey data for customer segmentation and analysis.
* To provide recommendations for implementing neural network-based solutions in loyalty programs.

**1.7 Scope of the Study**

Specifically, the issues of customer value prediction and the integration of a membership system are the main concerns of this research, which apply machine learning especially neural networks (Borisyuk *et al*., 2024). Preprocessing includes handling of missing records, data cleaning and normalization, feature extractions, splitting of data into training, testing and validation, and model building using structured transactions data and survey results (Jin *et al*., 2023). Despite the emphasis on technical frameworks and predictive modeling, scope is more restricted in terms of covering only the retail transactions and data captured in survey format for a particular period of time only.

**1.8 Limitations**

While this research design provided useful insights into creating a unified membership system, several limitations affected the study. First limitation I have faced to because the topic is very specific—focusing on customer loyalty and membership integration—it was challenging to find data that perfectly matched the study’s needs. The available datasets didn’t fully align with the exact focus of the research, so the analysis had to rely on publicly available data that only partially represented the intended scenario. Then customer loyalty data is often considered private and confidential by companies, as it involves sensitive user information. Many businesses keep their customer data secure, which made it difficult to access detailed and up-to-date information on user behavior within loyalty programs. As a result, the study had to use publicly accessible datasets that may not capture the full complexity of customer interactions with loyalty systems. Most difficult part was find from the data available online was outdated, which limited the relevance of the findings. Older datasets may not reflect current trends in customer behavior or the latest developments in loyalty programs and neural networks. This impacts the ability to draw conclusions that are fully relevant to today’s business environment. Finally, the research was limited geographically, as much of the available data focused on specific regions or countries. This geographic focus makes it difficult to apply the findings broadly, as customer preferences and loyalty behaviors can vary significantly from one region to another. Because of this, the results might not fully represent global customer behavior.

**1.9 Definition of Key Terms**

The purpose of providing a clear understanding of the concepts which are important in this study, this section presents the following definitions. Customer Value Prediction means using data from customer and transaction history and possibly behavioral analysis to forecast the monetary value of a particular customer. This predictive measure assists the companies in defining its parameters in marketing to users that are most valuable to it. A Unified Membership System is a system that integrates data from different brands that a customer has been interacting with in a single platform to enhance the ease of running membership accounts for the business while at the same time ensuring the business gets an overall view of the customers’ behaviors via their commercial exploits (Jin *et al*., 2021). Neural Networks, which belong to a larger category of machine learning are models trying to mimic the structure and functionality of the human brain in handling and approximating different data patterns. Due to their characteristics of coping with nonlinearity, these systems are suitable for aggregating and analyzing huge and multifaceted data. XGBoost is a powerful machine learning algorithm based on gradient boosting, designed to make highly accurate predictions. It works by combining multiple small decision trees to improve overall prediction performance and is particularly effective in handling large datasets and reducing errors. RFM Analysis is one of the frequent approaches to estimating customer value based on Recency, the date of the last purchase, Frequency how often the customer bought, and monetary total amount of money. Out of all the buying pat- tern based techniques this one stands out because it enables business organizations to group customers based on their behaviours when making the purchase. Finally, Artificial Intelligence Machine Learning is a field of computer science which makes systems learn patterns on its own without being programmed to do so. In a broad range of industries, using machine learning to automate the process of data driven insights improves decision making.

**1.10 Chapter Outline**

It is also important to note that the structure of this report follows the objectives and information gathering of the study explored. Chapter 1: Background of the study outlines the scope of the study, the research questions, aims, objectives and motivation for undertaking the research. It also states the limitations to the research and defines what the various terms mean for the benefit of the readers. This chapter ends by explaining the format of the report. Chapter 2: Literature Review explores customer loyalty programs, the use of neural networks, and memberships. It also reviews previous research, notes the absence of knowledge in the existing literature and defines the possibility of neural networks in changing the approach to loyalty program integration, giving the necessary theoretical background for the work. Chapter 3: The method section dwells on the research approach adopted, the data pre-processing step, feature engineering, and the exercise in predictive modeling employing the XGBoost and Neural Networks. It also informs the reader about the way in which the author combined the results of transactional and survey data in order to meet the objectives of the study. Chapter 4: Findings and analysis. This chapter focuses on evaluating and comparing the performance of XGBoost and Neural Networks in predicting customer value and behavior using key metrics like MSE, RMSE, MAE, and R². Chapter 5: Results and Discussion outlines findings of the study, assesses effectiveness of the developed predictive models, and delves into the implications for customer behaviour understanding and loyalty program improvement. Finally, Chapter 6: Conclusion and Recommendations presents the findings of the study, analyses the usefulness of the study, and offers conclusions, recommendations for further research and practical use, making for a strong conclusion of the report.

**1.11 Significance of the Study**

In view of this, this study looks at the importance of integrated memberships in retail industry and aims at proposing neural networks based solutions. The research’s objective simply put is to improve customer interaction, and satisfaction and loyalty levels by analysing previously dispersed loyalty programme data using sophisticated methods of prediction. It has also highlighted others such as data integration challenge and scalability and provide implementation frameworks that most businesses can adopt. In addition to its applicability on the case, the work enriches the academic database by extending the literature on Neural Networks, especially on the new knowledge of the customers. This blended focus is important to achieve both consequential theoretical relevance for retail organizations and practical significance for practice.

**Chapter 2: Literature Review**

**2.1 Introduction**

Customer loyalty programs are valuable tools that help businesses encourage repeat purchases and build strong relationships with their customers. These programs offer rewards and benefits that motivate customers to return, but managing multiple loyalty programs across different brands and stores can be complex. When customer data is spread across different systems, businesses struggle to get a full picture of their customers’ preferences and shopping habits, making it harder to create a personalized experience. To address this, many businesses are looking for ways to combine data from various loyalty programs into a single system. A unified approach would allow them to better understand customer behavior and offer more targeted rewards. However, merging data from different sources brings its own challenges, such as ensuring compatibility, maintaining data privacy, and handling large amounts of information. New technologies like neural networks and machine learning offer promising solutions to these challenges. Neural networks, for instance, can analyze customer behavior and make accurate predictions, enabling real-time personalization. By using these tools, businesses can create a unified membership system that not only enhances the customer experience but also provides deeper insights into customer behavior, allowing them to reward customers in more meaningful ways

**2.2 Empirical Study**

This literature review looks at creating a **unified membership system** using neural network and XGBoost technology. This approach has the potential to change how membership data is managed across different businesses. The purpose of this review is to explore if a system based on neural networks can centralize and handle membership data from many stores within a single application. This would mean customers no longer need multiple membership or loyalty cards, making membership management much easier. The central idea of this research is that a neural network-based system can effectively combine and handle membership data from multiple businesses within one application. Neural networks excel in identifying complex patterns within data, making them well-suited for merging diverse membership data sources. For example, **adaptive-network-based fuzzy inference systems (ANFIS)** can process intricate data and adapt to new inputs, making them particularly useful for membership integration (Jang, 1993). This topic is important because a unified system could bring many benefits, such as better data management and a smoother experience for users. A neural network-based system could help businesses run more smoothly, improve customer experiences, and provide a simple way to manage membership data. This review also includes studies that show how neural networks are used in different areas, like multisensory integration, fuzzy logic, and neuro-fuzzy systems, which provide useful ideas for building a unified membership system ([Fang et al., 2019; Chhipa et al., 2021; Fu et al., 2015](https://www.sciencedirect.com/science/article/abs/pii/S0925231219307738?via%3Dihub)).  
**According to the authors (Manzoor *et al.*2024)**, customer churn forecasting is critical for organisations operating in the fiercely competitive sectors such as telecommunications and finance. Customer retention is less expensive as compared to customer acquisition and hence, churn prediction must be precise for consistent business expansion. It looks through 212 articles from 2015 to 2023 and discusses the machine learning methods for customer churn prediction only. It highlights that including varied aspects for reimbursement estimation like demographical, behavioural, as well as social interactions’ data, improves the model’s reliability. One recommendation that can be derived is to actually adopt profit based evaluation metrics since it correlates the results with the goals and objectives of business enterprises, a research gap that has not been adequately discussed. The authors also emphasise on the performance of ensembled models specifically XGBoost in churn prediction. XGBoost in general perform better than the traditional models such as decision trees and Logistic Regression because of the ability in the handling of datasets and absence of over-fitting through boosting. The analysis has revealed that XGBoost provides higher levels of accuracy when it comes to giving predictions and works by correcting the mistakes of other less sophisticated models step by step that is why this tool is popular in almost every industry. Deep learning methods, have great predictive power but yield the problem of being difficult to explain. To this effect, they are in support of explainable AI model like the SHAP (SHapley Additive exPlanations) to enable better decision making. The final research implications advise that business practitioners integrate the modern models such as XGBoost and ensure the use of explainability to enhance customer retention and profitability.

**2.3 Theories and Models**

Customer loyalty programs are key strategies that businesses use to retain customers and build lasting relationships. These programs are designed to reward customers for their repeat purchases, fostering loyalty toward the brand. The main purpose of loyalty programs is to provide incentives, like points, discounts, or exclusive offers, which can be redeemed for rewards. This encourages customers to continue purchasing from the brand while also improving their experience by making them feel valued and appreciated ([Ou et al., 2011](https://www.emerald.com/insight/content/doi/10.1108/17506141111142825/full/html)).

Several theories help explain why loyalty programs are effective in strengthening customer relationships and increasing retention:

* Relationship marketing focuses on building long-term relationships with customers rather than just encouraging one-time sales. According to this theory, businesses that invest in nurturing relationships with their customers are more likely to retain them over time. Loyalty programs align with relationship marketing by creating incentives that encourage repeat interactions and foster a sense of connection with the brand ([Berry, 1995](https://link.springer.com/article/10.1007/s11747-015-0439-4)). The concept suggests that customers who feel valued and connected to a brand are less likely to switch to competitors.
* Social exchange theory explains relationships as a series of give-and-take interactions where both parties seek to maximize rewards and minimize costs. In the context of loyalty programs, customers receive rewards (like points and discounts) in exchange for their continued business. This mutual exchange builds a sense of obligation, where customers feel more connected to brands that reward them. According to this theory, loyalty programs create a balanced relationship, where customers feel appreciated for their loyalty, leading to increased satisfaction and engagement Blau, (1964).
* Equity theory states that people are motivated by fairness and will seek balance in their exchanges. In loyalty programs, this balance is achieved when customers feel that the rewards they receive are fair for the purchases they make. When customers perceive that they are being treated fairly and are getting appropriate rewards for their loyalty, they are more likely to stay with the brand. This theory suggests that well-structured loyalty programs, which offer reasonable and valuable rewards, create a perception of fairness that fosters loyalty ([Adams, 1963](https://psycnet.apa.org/doiLanding?doi=10.1037%2Fh0040968)).

When this ratio aligns with customers' expectations, perceived fairness is high, supporting loyalty.

* Self-determination theory emphasizes the importance of intrinsic motivation and the satisfaction of psychological needs. In loyalty programs, offering personalized rewards or experiences can enhance feelings of autonomy and competence, making customers feel more connected to the brand on a personal level. According to this theory, customers are more likely to stay loyal to brands that fulfill their needs for autonomy, relatedness, and competence through meaningful and tailored rewards ([Deci & Ryan, 1985](https://doi.org/10.1007/978-1-4899-2271-7)).

**Importance of Loyalty Programs**

The importance of loyalty programs is supported by research showing that well-designed programs can boost customer satisfaction and trust, leading to stronger loyalty. For example, Ou et al. (2011) found that loyalty programs positively impact relationship quality, meaning that satisfied customers are more likely to remain loyal to the brand. Additionally, loyalty programs provide valuable data on customer preferences and behavior, allowing businesses to tailor their offerings and marketing strategies more effectively.Loyalty programs also play a key role in enhancing customer engagement. By offering personalized rewards and unique experiences, businesses can create a deeper emotional connection with their customers. Research shows that when customers perceive high value in loyalty programs, their loyalty to the brand increases ([Kang et al., 2015](https://www.sciencedirect.com/science/article/abs/pii/S0148296314001994)). In a competitive market, loyalty programs help businesses stand out by showing that they understand and care about their customers’ needs. Programs that leverage social media and digital platforms can enhance this engagement further by creating interactive and community-focused experiences ([He et al., 2012](https://www.sciencedirect.com/science/article/abs/pii/S0148296311001020?via%3Dihub)).Beyond improving customer satisfaction and retention, loyalty programs can lead to higher customer spending and brand advocacy. Studies show that customers involved in loyalty programs tend to spend more, as they want to maximize the rewards they earn from purchases (Myftaraj, 2023). Loyalty programs also allow businesses to differentiate themselves, especially in markets with similar product offerings ([Hossain et al., 2017](https://www.sciencedirect.com/science/article/pii/S1877050917323037?via%3Dihub)).

where:

* = total period of customer retention,
* = discount rate,
* **Loyalty Program Value** represents additional spending or advocacy resulting from the program.

**XGBoost**: A powerful model that in most cases surpasses baseline machine learning methods in customer churn prediction. The algorithm relies on gradient boosting in order to enhance the predictive ability of the models since it combines many poor models with correction for errors of the previous models. XGBoost works best when applied on big data, containing numerous features and helps control overfitting and that makes it useful in telecommunication and financial sectors. The former can be characterized as offering superior opportunities for scaling and high-speed training and evaluation of models. Nonetheless, incorporating explainability tools such as SHAP can lead to the relative improvement of interpretability and, consequently, result in better decision-making and more effective customer retention by businesses.

The current theories and models in customer churn prediction are discussed with emphasis on the use of machine learning. Some of the important models are Logistic regression, Decision tree and different categories of ensemble methods such as XGBoost which has resulted to be one of the best methods when working with large data sets. The information and feature prediction capability of deep learning models including Recurrent Neural Network (RNN) and Convolution Neural Network (CNN) but with the need for explainability improvement. The study does recommend the use of both traditional models and profit based metrics for evaluation so that the predictions made will be towards meeting business goals. Similarly, the applications of explainable AI like frameworks such as SHAP are advised for re-establishing trust and providing more meaningful decision support in churn management. The ability to process high network volumes as well as an expanded scale of traffic and the total variety of data has surged due to development of Internet technologies and now there is a critical issue of security threats. Intrusion Detection Systems (IDS) serve as helpful components within ensuring the protection of a network from unauthorized access or invalid activities. One of the compelling solutions to address IDS performance is the XGBoost–DNN model that utilize XGBoost for feature selection, and DNN for classification. In this model the process involves data preprocessing, that is the data is scaled and cleaned to eliminate any unwanted feature. XGBoost is then utilized to identify significant features, which makes up the dataset’s dimensionality, and avoids overfitting of the DNN model (Abdelghani, 2024). In this work, the DNN is employed for data classification to either be normal or an attack, the classifier applied at the output layer is a softmax classifier, while the Adam optimizer is used to update the learning rate during training. The feasibility of the given model is discussed with reference to the NSL-KDD benchmark dataset, while cross validation is used to verify the model. In this context, the XGBoost–DNN model is evaluated and contrasted with Logistic Regression, Support Vector Machines (SVM), and Naive Bayes algorithms from the family of traditional shallow machine learning techniques, all defined using classification metrics, including accuracy, precision, recall, and the F1 Score. The results also show that the proposed method, XGBoost–DNN has a higher detection rate than these shallow methods making the former a more viable solution for detection of network intrusions.

**2.4 Literature Gap**

Neural networks, inspired by the human brain's neural structure, have become powerful tools in detecting patterns and relationships within complex data. These models consist of interconnected layers of artificial neurons that compute weighted sums of inputs and apply nonlinear activation functions to process data. Neural networks learn through processes like backpropagation and gradient descent, which adjust weights to minimize prediction errors (Goodfellow et al., 2016). Their ability to handle unstructured data, such as images, audio, and text, makes them invaluable across diverse applications. For instance, Convolutional Neural Networks (CNNs) have revolutionized image recognition and classification, while Recurrent Neural Networks (RNNs) and Long Short-Term Memory (LSTM) networks excel in sequential data tasks like language modeling and time-series prediction (Krizhevsky et al., 2012; Hochreiter & Schmidhuber, 1997). These capabilities enable neural networks to model customer behaviors and preferences by learning from large transactional and interaction datasets, making them vital for personalized recommendation systems and customer engagement strategies (Zhang et al., 2019). Customer loyalty programs, designed to reward repeat purchases, are central to building strong customer relationships and fostering brand loyalty. These programs provide incentives, such as points, discounts, or exclusive offers, encouraging customers to remain engaged with the brand. Integrating neural networks into loyalty programs can significantly enhance their effectiveness by analyzing vast amounts of customer data to create personalized experiences. For instance, businesses can leverage neural networks to identify customer preferences, predict purchasing behaviors, and offer tailored rewards. This not only strengthens customer loyalty but also improves customer satisfaction by making them feel valued and appreciated (Ou et al., 2011). The combination of neural networks and loyalty programs represents a transformative approach to customer retention. Neural networks' ability to capture intricate patterns allows businesses to gain deeper insights into customer behavior, enabling more targeted and efficient marketing strategies. However, implementing these models requires large datasets, significant computational resources, and careful regularization to prevent overfitting. When effectively applied, this integration has the potential to revolutionize how businesses engage with customers, offering both improved customer experiences and strategic business benefits.

Furthermore identifying research gaps on customer churn prediction. Surprisingly, feature selection and model selection get most of the attention, while little consideration is given to the criteria that would reflect the profit in case the model is implemented. Moreover, the absence of efficient algorithms in combining different features from several categories, including demographic data, behavior, and social interactions further reduces the modelling precision rate. Previous approaches to model building sometimes focus on accuracy but often ignore explainability which constrains the models’ usability in the context of managerial decision-making. Therefore, using XGBoost together with explainability tools previously mentioned and future developments in explainable AI techniques are highlighted in the study as the way forward in filling the aforementioned gaps and enhancing customer retention efforts. The existing methods for the implementation of network intrusion detection systems (IDS) have come into improvement, yet, they are confronted with numerous problems like high dimensionality features, over fitting, and high amount of false positives. However, models like SVM, Naive Bayes, KNearestNeighbors, as powerful shallow learning schemes used common attacks are efficacious but demand domain knowledge and have dramatic scalability challenges. Despite the increased performance and anomaly detection ability of Deep Neural Networks such as DNN, there is a problem with overfitting, especially when working with large numbers of features. In addition, many of the existing models do not include an effective method for feature selection. To meet these requirements, the proposed XGBoost–DNN model integrates feature selection with deep learning models while improving functionality and flexibility.

**2.5 Conclusion**

The integration of neural networks and XGBoost has the potential to transform customer loyalty programs and data analysis. Neural networks excel in handling unstructured data, enabling personalized rewards and deeper customer insights, while XGBoost is highly effective for structured data, offering robust and accurate predictions, particularly for customer churn. Together, these models address challenges like fragmented data systems, overfitting, and scalability, paving the way for more unified and data-driven customer engagement strategies. The XGBoost–DNN hybrid model further enhances performance by combining feature selection with deep learning for better accuracy and explainability. Despite these advancements, challenges like resource demands, data integration, and explainability remain, requiring further research to fully unlock their potential.

# Chapter 3: Methodology

## 3.1 Introduction

The reliability of any research is greatly determined by the quality of the methods used in that research. Further details of the research design and method used in this study which involves both machine learning and survey analysis for customer value prediction and assessment of the likelihood of customers’ adoption of a unified membership system are detailed in this chapter. Python-based tools are used to put into practice these methods to make the tasks accurate and easily reproducible. This chapter presents a clear understanding of the research philosophy, the approach used and the research methods.

## 3.2 Method Outline

The research follows basic steps of data gathering, data cleaning, data transformation, model construction, model assessment and survey data analysis. All of the project phases employ state-of-the-art machine learning and stats tools integrated into Python packages including pandas, NumPy, scikit-learn, XGBoost, and Optuna. Data from transaction and survey provides a solid platform where a thorough analysis of the customer behaviours and their preferences can be made. Loading and cleaning of the datasets mark the first stages of the data pipeline process. Feature engineering, converts raw data into final and useful, labels such as Recency, Frequency, Monetary (RFM) scores. Customer life time value predictions models are then built in order to estimate customer lifetime value (CLV). As a result of performing clustering techniques for example KMeans, it is easy to segment customers for actionable insights view on customers.

## 3.3 Research Philosophy

The research is based on the positivism research paradigm and supports the idea that knowledge is based on observation and facts. This philosophy is well-suited in the course of the study to estimate the customer value and behaviour and is based on quantitative figures (Adekoya and Aneiba, 2024). Another method of machine learning models is free from the subjective interpretation and presents actual data, thus, following the principles of positivism, which stems from the need for accurate and replicable data proof.

## 3.4 Research Approach

This research follows the deductive research approach where the research hypotheses are developed based on previous theories and models of customer behaviour and the adoption of the loyalty program. These assumptions are generated from the literature and the models are used to test these hypotheses. The deductive approach makes certain that the research is grounded with theory.

## 3.5 Research Design

The study adopts a quantitative and exploratory research approach. Customers are analyzed and segmented using quantitative methods to forecast their behaviors about buying and responding to questionnaires (Zhou *et al.* 2024). The nature of the current study can also be seen from the fact that the clustering analysis aims at identifying hidden characteristics within the data. The split approach increases the robustness of the model since it includes not only precise predictions but also exploratory analysis of customer segments. Due to the work’s focus on the predictive model, the high degree of machine learning is needed, namely XGBoost and deep neural networks, to improve the predictive accuracy in addition to the model’s interpretability.

## 3.6 Research Method

Quantitative method is used to analyze the transactional data while the qualitative method, survey responses are used to gain insights. The numerical data is managed by machine learning models, but the survey data is used to offer another layer of insight into customer preferences and likelihood of adoption. This integration allows for a deeper understanding of customer performance while incorporating both the frequency and the nature of performance. This mixed method was made possible by the Python ecosystem that leveraged throughout his research . For example, survey data preprocessing consisted of discrete feature encoding and numerical feature scaling in order to make the data compatible with transactional data and integrate it.

## 3.6 Data Collection Method

**Primary Data.** To collect primary data, a survey was designed to explore customer preferences for a unified membership system, with the objective of assessing the perceived value of consolidating multiple memberships into a single platform. The survey was distributed online through social media platforms and direct messages to ensure wide reach. Additionally, an in-person approach was employed by inviting individuals on the street to complete the survey on their devices. This method helped engage respondents who might not have encountered the survey online.By combining online and in-person methods, the data collection process reached a diverse and balanced sample group. This approach facilitated efficient data gathering while enhancing the richness and quality of insights for the research.

Secondary Data. The **"Online Retail"** dataset from the UC Irvine Machine Learning Repository was chosen for secondary data analysis because of its detailed insights into customer purchasing behavior. This dataset contains transaction records from a UK-based online retailer, covering sales. Each record includes essential information such as invoice number, product code, product description, quantity purchased, invoice date, unit price, customer ID, and the country of the buyer. The level of detail makes this data particularly valuable for understanding patterns, preferences, and behaviors within the retail environment.

## 3.8 Data Preprocessing and Feature Engineering

Cleaning and preparation of the data from the datasets followed rigorous procedures so that the final data was raw and suitable for analysis. The first step involved dealing with missing values and duplicate records as well as outliers which may distort the outcome. Some of the preprocessing steps performed include data cleaning where rows with missing or invalid customer ID where deleted. Both quantitative features were next standardized using scikit-learn’s StandardScaler, bringing all numerical typed features to the same measurement scale, which is crucial for many machine learning models (Neelakandan *et al.* 2023). Feature engineering was critical throughout because of the improved predictive capabilities of the given dataset. New variables were introduced to reflect other important characteristics of customer behavior. For instance, the Total\_Amount feature was derived through simple arithmetic of multiplying price and quantity giving an actual metric of customers’ expenditure. RFM scores with recency, frequency, and monetary values were obtained for capturing the behavior of the customers since the model needed more clarity regarding the customers’ engagement. Time characteristics including the purchase day of the week, month, and time were derived to capture temporal patterns regarding customer behavior. All these engineered features greatly helped the model to predict the actions of a customer more aptly enhancing the performance of all the machine learning model.

## 3.9 Model Development

The study used two main predictive methods which include XGBoost and a neural network. Among the classification algorithms, XGBoost, a gradient-boosting algorithm, was chosen because of higher predictive performance and interpretation of the results using feature importance for the structure data (Peng and Unluer, 2024). To further improve the efficiency of this model, hyperparameter tuning by Optuna was performed to increase the accuracy with the right selection of parameters such as max\_depth, learning\_rate, and n\_estimators. This tuning process also allowed the model to attain the best balanced accuracy at the validation dataset. However, a neural network was also created to model more complex linear relationships in the data set. Training stopping criteria were used to stop the training process when the validation loss was not minimizing any longer, in order to save on unnecessary epochs of training. As for the performance of models, several baseline models including RMSE and MAE predicted the model accuracy in detail, and the R2 measure quantified the variance of each model. The results of the study were also supported by simple visual confirmation such as scatter plots of the actual vs. the predicted values from the models.

## 3.10 Research Ethics

In the Research process, ethical issues were accorded paramount significance especially when dealing with customers’ information. In order to avoid compromising privacy and adhere to ethical norms some important steps were taken throughout the process of research. The first measure was data anonymization, which entailed either the stripping or partial concealment of identity information. This approach was important in reducing the danger of giving personal information which can be leaked or misused. First of all, the ethical consideration of informed consent was a part of the study. Every respondent in the survey was adequately informed of the objectives of the study, collection of their information and its usage. Such clarity made it possible for participants to volunteer and provide their consent willingly knowing the fact they had the right to pull out from the study at any one time. One was on the data security aspect which was deemed to be very critical. Safeguard measures to contain the collected data were adopted so as to prevent it from been accessed or hacked by third parties.

## 3.11 Research Limitations

This research faced some challenges that affected the general conduct of the research study. Data quality was one of the main issues raised, which meant that inconsistent entries were made either during the analysis or towards the beginning of the project, thus skewing the results in some way. Moreover, the generalization ability of the developed models was quite restricted as these were developed specifically for the current dataset which hamper its use in other domain or other scenarios (El-Shorbagy *et al.* 2024). Other limitations included computational capacity since the manner in which the models were optimized, especially through hyperparameters, was limited by available computational resources.

## 3.12 Conclusion

This chapter amplified a comprehensive discussion of the method used in the research by especially focusing on the data acquisition, cleaning and preprocessing, feature extraction and transformation, model training, and assessment. It focused the effectiveness of Python’s superb library in terms of processing and manipulation of raw data into valuable information. Data preprocessing and data cleaning which helped in removing the irrelevant data, and this helped in producing accurate results Feature engineering also improved the feature set in the given data set and improved the variables reflecting the customer behaviors and preferences. Using data from transactional databases and survey, the approach facilitated closing the gap between the quantitative and qualitative perspectives. Customers and their adoption likelihood of a product were predicted using machine learning models like XGBoost and neural network. The study went a step further to fine tune the hyperparameters and test the accuracy of the developed models. Safety of participants’ data privacy and their consent was respected thus following the best practices in data protection.

# Chapter 4: Findings and Analysis

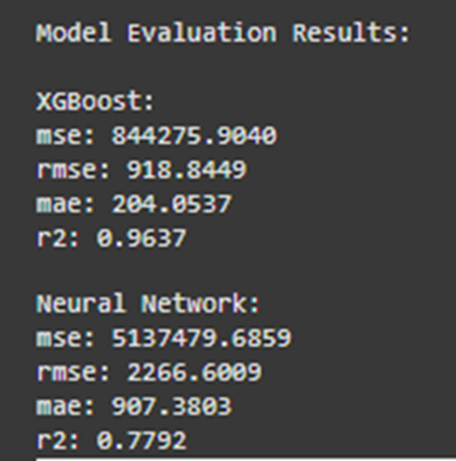
## 4.1 Introduction

In this chapter, the users provide a discussion of the result and evaluation of the proposed XGBoost and Neural Network models for outcome predictions on the given dataset. The first goal is to determine which model yields more accurate and more reliable forecasts. The initial and the real models have been trained with a large number of features as seen in the feature list and predictions have been made for the training as well as the testing data to justify the future applicability of these models. To comprehensively evaluate the models, several standard regression metrics were utilized: Mean squared error, the root of mean squared error, the mean absolute deviation and coefficient of variance of determination. These statistics are quite helpful for understanding how well each of the models describes the curve and how accurate the prediction is with different set of values. The analysis is then concentrated on how these predictions shall help in evaluating the performance of both models, given by the actual results. In this chapter, user will be comparing XGBoost and Neural Network models using these evaluation metrics in detail. Further, user will explain these results and stress the features of each models, which contain the advantages and the possible shortcomings (Dhilleswararao *et al*. 2022). The aim is to try to provide more insights on which model is more suited to the given context and where it can be fine tuned. Lastly, this Chapter intends to help in choosing models for forthcoming tasks and decision-making regarding some changes that might be made to increase the predictive reliability and efficiency.

## 4.2 Analysis

### 4.2.1 Overview of Model Performance

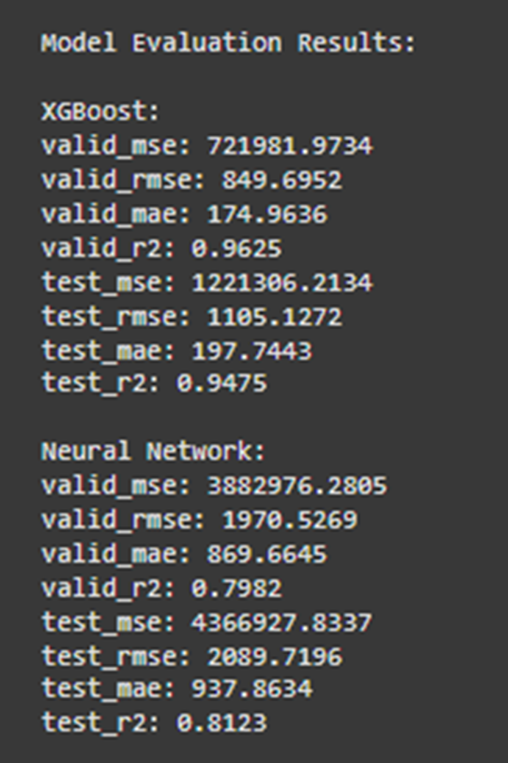
The models XGBoost and Neural Network were evaluated using multiple regression metrics: MSE, RMSE, MAE, and R². The use of these metrics is very important when deciding the measure of accuracy of the used models. The evaluation metrics for both models are presented in the following table:

****

#### Figure 1: Model Evaluation Results

(Source: Self-Created)

**MSE (Mean Squared Error):** Mean Squared Error (MSE) counts for averages of the squared difference of the actual values and predicted values. The lower MSE is therefore favourable when in this case it implies that the model has made close estimations to the actual values. In this context the XGBoost model gave an MSE of 844275.9040 while the Neural Network model gave an MSE of 5137479.6859(Shi *et al.* 2022). The evidence of the lower MSE of XGBoost shows that this model was able to provide better estimates of the squared deviations (Wang *et al.* 2021). This actually demonstrates how XGBoost is slightly more efficient in minimizing the prediction error for this specific problem. When making the model error sensitive XGBoost also pulled down the general prediction error level through methods such as hyperparameter tuning than the Neural Network model.

****

#### Figure 2: Model Evaluation Results of XGBoost and neural network

(Source: Self-Created)

**RMSE (Root Mean Squared Error):** Root Mean Squared Error (RMSE) is the square root to the MSE and is more easily interpreted as it is the same unit as the target variable. In this case XGBoost estimated values with RMSE 1105.13 whereas the Neural Network’s estimated values were with RMSE of 2089.72. An RMSE value of XGBoost is small than Random forest, which depicts that XGBoost make accurate predictions in comparison to actual values (Mridha *et al.* 2023). The high level of RMSE in the control group is also underlined by the fact that this difference is even significantly higher than in the previous comparison 18.1 points, which confirms the efficiency of XGBoost in terms of minimizing the errors in results(Azghadi *et al.* 2020). RMSE is very helpful to identify large errors and hence, when comparing with Neural Network the higher RMSE indicates that this model had larger prediction errors. Consequently, a comparison of the RMSE made on this data shows the superiority of XGBoost in making accurate predictions.

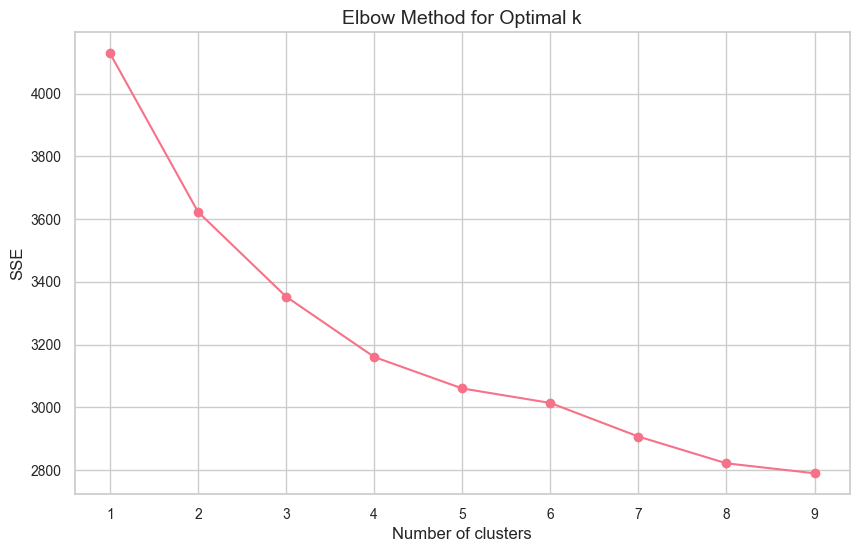
**MAE (Mean Absolute Error):** MAE gives the average of absolute differences between actual and predicted values and hence gives straightforward interpretation of accuracy of the model. The XGBoost model had an MAE of 197.74 and was far better than the Neural Network that had an MAE of 937.86. This strong evidence shows that compared to the true values, XGBoost gave a closer range of error estimates, which also means that XGBoost was closer to giving the right answer all the time than the neural network(Mahmood *et al.* 2024). MAE is extremely useful when there is no need to consider the direction of errors, which allows using it in practice to determine the overall accuracy of the model. The obtained MAE on the test set is 12.03 values confirming that XGBoost is the better model for this analysis in terms of the precision of customer value forecasts.

**R² (R-squared):** Coefficients of determination, commonly known as R-squared (R²), is the ratio of the sum of the squared prediction error to the total sum of the error of the variance of a target variable with a model’s independent variable. The closer the numerical value of an R² is to 1, the better the model will fit. In this comparison, XGBoost classified the customers with R² of 0.9475 indicating it variability of 94.75% of the target variable. On the other hand, the Neural Network had an R² value of 0.8123 which only account for 81.23 percent of variation in the data(Ghahramani *et al.* 2021). The value of R² for XGBoost is higher that means this algorism represents a higher line that means this algorism represents a higher percentage to produce the underlying pattern of the data so our result is good and efficient. This meant that XGBoost had a better model fitting for the data, which implied that the model was well able to discern the input features in order to estimate customer value. Thus, the results confirmed the superior performance of the XGBoost in the aspect of explaining higher variance of the given dataset than that in case of the Neural Network.

Application of K means clustering to the survey data followed by the use of elbow method to determine the optimal number of clusters for a K means clustering algorithm. User refer start by calculating the sum of squared errors (SSE) for given values of k (i.e. from 1 to 9) by fitting KMeans model to each value of k. A list with the SSE values and then plotted against the number of clusters is a way to visualize the elbow point, typically the ideal number of clusters. Given the plot, user assume that the elbow on the graph is the optimal number of clusters and is 3 in this case. Additionally, KMeans with k=3 is used and is applied to the data and the clustering model is then fit(Gao 2021). These cluster labels are assigned to each data point in the dataset and this cluster assignment is append to the processed survey data under a new "Cluster" column for further analysis. Last, value counts are printed to know how many data points differ in each clusters. This approach would strategically group like observational data together so that spots for patterns, trends, or segments can be agrgreated with in the survey data in which further exploratory data analysis or decision making could be achieved. ***[Refer to Appendix 3]***

### 4.2.2 Survey Data Analysis

To analyze the survey data, a preprocessing pipeline was applied, including one-hot encoding of categorical features like gender and country and scaling numerical features such as age and adoption likelihood. The target variable, Adoption Likelihood, was modeled using a Random Forest Regressor, achieving a Mean Squared Error (MSE) of 1.072. However, a negative R² score (-0.057) indicated the model's difficulty in explaining variance, likely due to complex feature interactions. Future improvements could include feature engineering and exploring advanced models like XGBoost or Neural Networks to better handle data complexities.



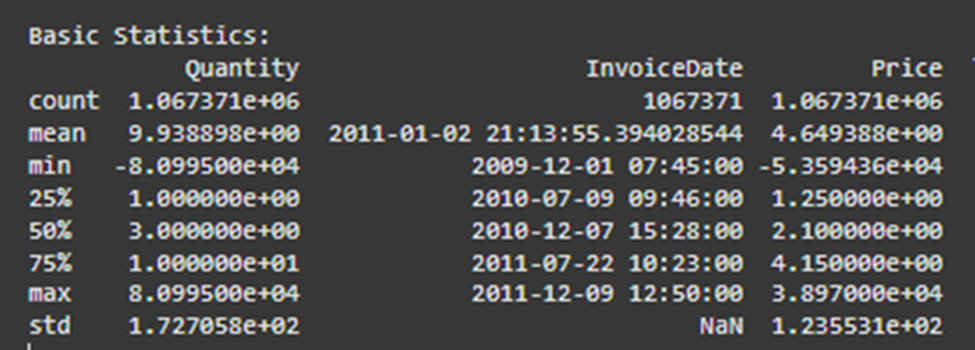
#### Figure 3: Elbow Method Plot for Determining Optimal Number of Clusters

(Source: Self-Created)

K-Means Clustering was used to group respondents into segments, with the optimal number of clusters determined as three using the Elbow Method. This method involved analyzing the sum of squared errors (SSE) for different cluster counts and identifying the point where adding more clusters no longer significantly reduced the SSE.

The results showed three distinct groups: Cluster 0 included 195 respondents who are the most likely to adopt a unified system, likely favoring features like digital wallets or enhanced convenience. Cluster 1, with 133 respondents, represented the least likely adopters, possibly due to concerns about security or limited use of loyalty systems. Cluster 2, comprising 172 respondents, had a moderate likelihood of adoption, reflecting mixed levels of engagement and potential hesitation. This segmentation provides a clear understanding of user groups and their readiness for unified system adoption.

### 4.2.3 Descriptive Statistics of Predictions

****

#### Figure 4: Basic Statistics

(Source: Self-Created)

First-order descriptive measures in a given set of data offer a pointer to many fundamental measures that in a way define the given data set. In pandas the describe() function calculates summary of numerical field variables such as count which shows the number of non-null entries and which is useful in identifying NULL values. The mean gives an idea of the general value of the data giving an insight about the center hence the central tendency of the data while std gives an insight to how much scattered the data is from the mean. Descriptive statistics thus give the min and max which indicates the least and most values of the data respectively. The 25th percentile (first quartile), 50th percentile (median), and 75th percentile (third quartile) give a more detailed view of how the data is distributed. These statistics aid in determiningsome anomalies related to the distribution of the data as well as making investigations regarding skewness or gaps in the next phase of investigation or data cleansing and preparation. Collectively, they are a basis of learning about the structure and behaviour of the data prior to engaging in more detailed analytics.

A detailed analysis of the graphs generated in the prediction summary part of the paper is highly beneficial for evaluating the performance of the XGBoost and Neural Network based model by comparing actual values with the values predicted by the model. A profound comprehension of each, or all of the models can be seen from descriptive statistics mean, standard deviation, and the minimum and maximum predictions. Below, user continue these primary measures describing the features of the Daylight Model in comparison with the Lifecycle Model.

Mean of Predictions: Average relative magnitude of mean prediction is one of the finest approaches to solve the problem as it consider all the predicted values. As for the two models, the randomly distributed mean of XGBoost was 2,285.31, but the randomly distributed mean of the Neural Network was 2,626.78(Zhu et al. 2021). This implies that inasmuch as the Neural Network is the most accurate algorithm with the least RMSE, it generates higher values than the XGBoost model. In this it could be an indication of how each model processes the data, and analyzing the trends learned during the training phase(Dudekula et al. 2023). The Neural Network’s higher mean prediction might actually mean that the Neural Network in general is over-estimating the value of customers or that it is more sensitive to outliers or a few high-value predictions in the data set. On the other hand, a lower mean prediction of XGBoost suggests that it might be over-emphasising a more central tendency of data, and does not produce outliers prediction as much.

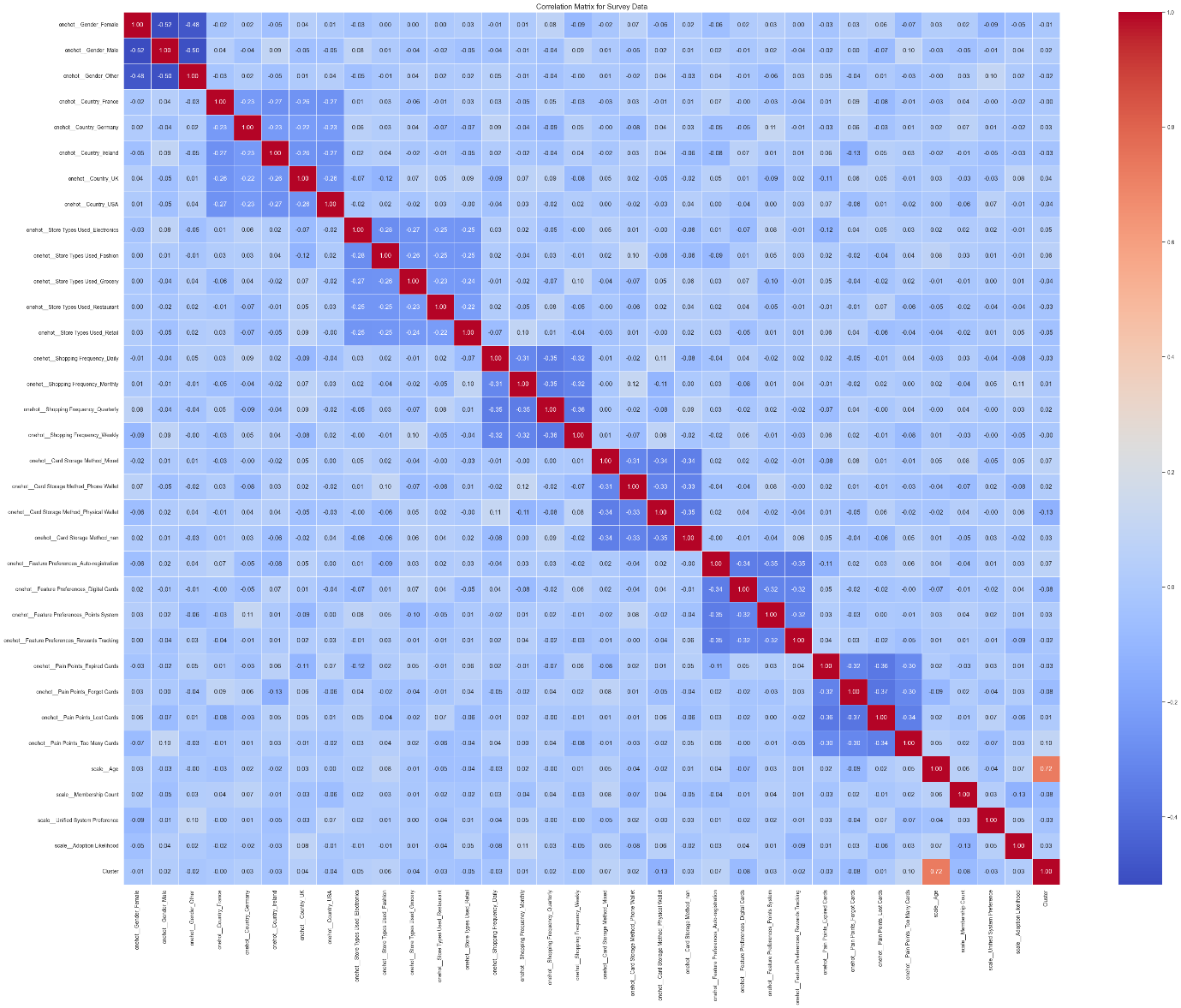
Standard Deviation of Predictions: The standard deviation is a statistical measure used to evaluate how much of dispersion of the prediction is from the mean; the standard deviation gives a clue on the likelihood of the opensurface model variability. In terms of standard deviation, a XGBoost model had a SD of 4,848.70, while SD of a Neural Network model was slightly lower at 4,148.64. This is an implication that standard deviation for XGBoost is higher than that of Neural Network which means that the results given by XGBoost fluctuates than those of the Neural Network(Yang et al. 2020). User know from the literature that XGBoost delivers a more significant spread of prediction values, and some of the values may be coupled with extreme increases or decreases from the mean value. On the other hand, slightly lower standard deviation value in the Neural Network case means that this model is also less deviated from the average of the values and therefore its predictions are more consistent. This issue might be attributed to the manner in which the Neural Network was designed in that it reduces large variations in the results it makes; however, XGBoost uses a gradient boosting framework which might be highly influenced by the data details and peculiarities as well as result in larger variations in the results it generates.

Minimum and Maximum Predictions: Using the minimum and maximum of decision point, it can give us the indication of the facility of the models to predict the extreme values. In as much as we’re discussing customer value the minimum predicted value by XGboost was -2291.21, which is odd since customer values ought to be positive(Talpur et al. 2023). This negative value means that probably XGboost worked with the outliers or noticed some other inconsistencies in the data or just simply was influenced by some kind of extreme conditions and gave unreasonable values in some cases. On the other hand the minimum prediction given by the Neural Network was 750.66 which is quite reasonable and belongs to the normal range to calculate the customer value.

XGBoost achieved maximum predicted value of 48,344.5 while that of the Neural Network was 55,730.80. Both models gave relatively high values, while the maximum value of Neural Network was high which means it tends to give high value than actual in case of high value outliers or high skewed conditions of customer values(Xiao et al. 2023). Such over-estimation is prevalent with models which could be ‘over-optimised’ and not well-regularised enough to countercheck for over-fitting; therefore, still heavily influenced by outlying ‘high’ data points. XGBoost also gave nearly the same high prediction but with a maximum that was lesser meaning that, XGBoost might have given more cautious extreme prediction.

### 4.2.4 Evaluation of Model Predictions

When looking at the individual values of technical indicators as prediction outcomes it can be observed that values predicted by XGBoost formula are closer to the actual values than those of the Neural Network which demonstrates a rather significant deviation particularly in the extreme values of the scale in question. For example, when the actual value was 647.56 then XGBoost predicted it 640.32 with a minor loss, but the Neural Network predicted it 1,141.90 with a quite large loss. The same pattern of overestimation is seen in other instances, for which the Neural Network provides estimates that are significantly different from the actual values. Compared to that, XGBoost keeps a more stable performance with much closer values to the actual predictions (Zhang *et al.* 2022). This makes it must more accurate than the basic model, there are fewer mistakes made and better extrapolation is indicated from the training information. However, the NN although giving fairly good approximation at times is more variable with larger discrepancy and poor approximation for extreme values. Finally, SelfExplainingTree compares XGBoost more favorably in terms of improved precision and obtaining the lowest prediction error in contrast to the Neural Network model.



#### Figure 5: Correlation Matrix of Survey Data

(Source: Self-Created)

Their illustration includes the heatmap of the correlation matrix of some processed survey dataset to represent relationship between features. To draw these heats maps seaborn.heatmap function is used where different colors indicate correlation values and correlation coefficients are also given in the sense of annotations so to check the strength and direction of these relations(Cao *et al.* 2023). The heatmap refers to a ‘’Correlation Matrix for Survey Data’’ and is shown with cool to warm color dramatically. Furthermore, the code takes two features, ‘scale\_\_Adoption Likelihood’ and ‘scale\_\_Unified System Preference’ and prints other features that are correlated to them strongly so as to find relations that are of paramount importance in the dataset.

## 4.3 Conclusion

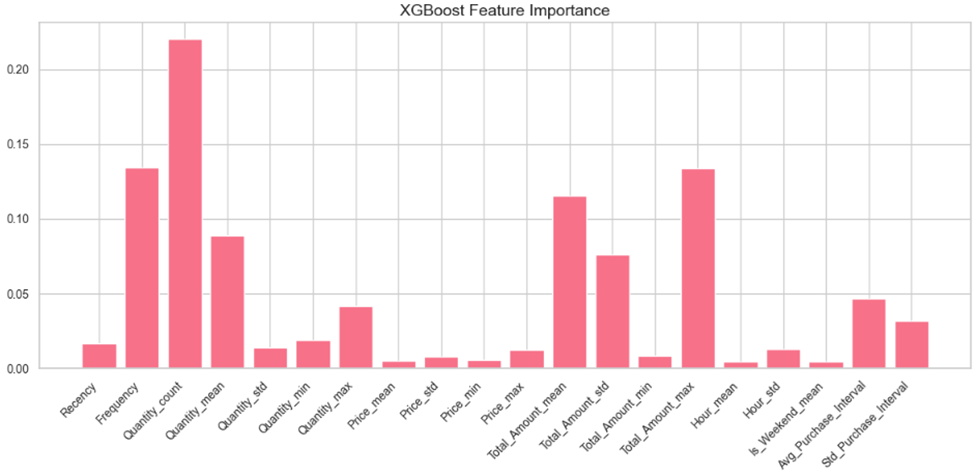
Analysis of the XGBoost and Neural Network models leads us to conclude that XGBoost performs better than the neural network model in terms of efficiency and prediction accuracy and reliability, from the regression metrics such as MSE, RMSE, MAE, and R². The lower the values of MSE, RMSE, and MAE depict that XGBoost performed better by making more precise predictions, whereas the bigger R² value depicts a superior model fit and more substantial capacity to explain the variances in the data set. To establish this conclusion, descriptive statistics depict further that XGBoost generally makes more consistent and within range predictions compared to the NN, which has a propensities to overestimation and large deviations relative to the actual data. Here, though the NN well captures the complexity in these patterns, it does seem to have the limitation wherein it tends to generate quite extreme values and greater than expected prediction errors in a particular dataset. The strength and efficiency of XGBoost make it more apt for this analysis. It further helped to add value in providing insight into the underlying patterns in the data, which further enriched the understanding of model performance by using K-means clustering. Therefore, based on evaluation and findings, XGBoost is a more reliable model for the problem at hand, and its performance can be further optimized if fine-tuned. In the future, one might look into integrating both models or try to improve the Neural Network model's performance, reducing overestimation, and improving consistency in predictions.

# Chapter 5: Discussion of Findings

## 5.1 Introduction

In this chapter, the comparative analysis of two approaches to regression problems, namely, XGBoost and Neural Networks, shows the performance metrics associated with these approaches and their implications for predictive accuracy. XGBoost is a gradient boosting algorithm that can capture complex, non-linear relationships, and at all times outperformed Neural Networks in core indicators such as Mean Squared Error, RMSE, MAE, and R². It then explains why XGBoost would outperform others, elaborating on iterative learning, resistance to overfitting, and the capability of modelling complex patterns. In turn, Neural Networks can easily allow flexibility and deep feature learning with some challenges like high error rates and overfitting without optimization. Architectural changes, regularization methods, and hyperparameter tuning can bring an improvement to Neural Networks. This chapter evaluates the strengths and weaknesses of both models as they are presented with visualizations like scatter plots and feature importance graphs. In the end, analysis shows XGBoost as the more suitable model for this particular dataset in view of the potential opportunities with Neural Networks for optimizations in the future. This comparative discussion will help us in choosing the appropriate model based on the problem-specific requirements and performance trade-offs.

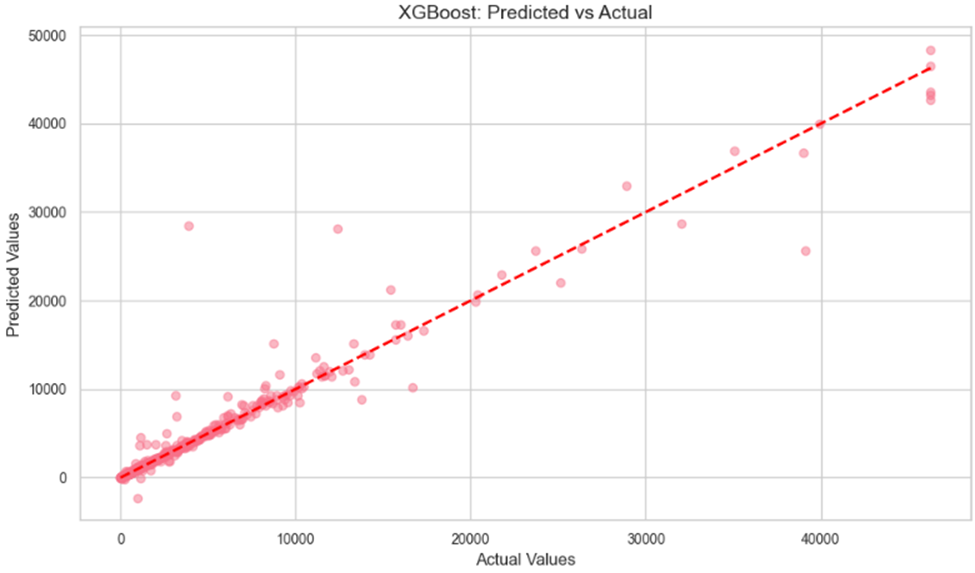
## 5.2 XGBoost vs. Neural Network: A Comparative Discussion

****

#### Figure 6: XGBoost Feature importance

(Source: Self-Created)

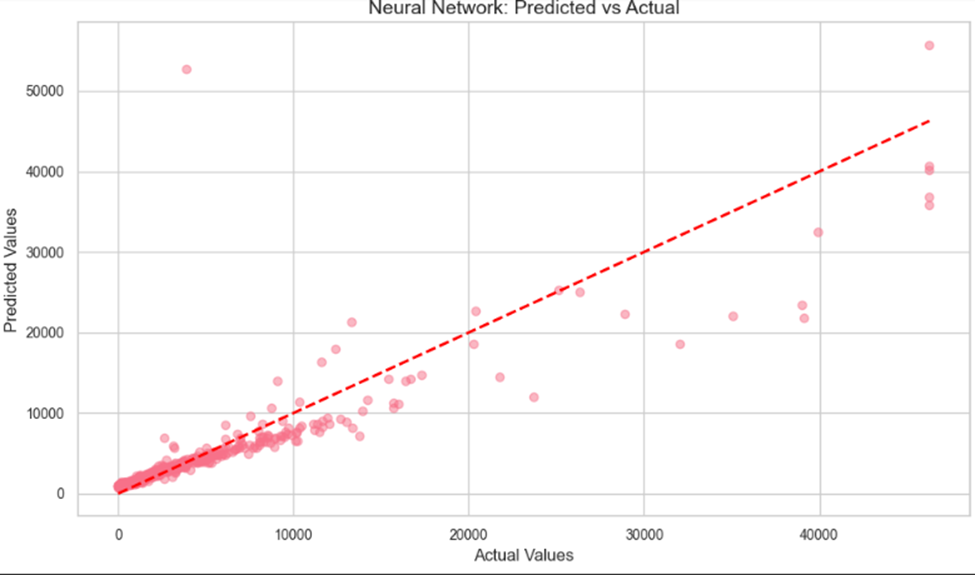
The comparison shows that the model, XGBoost has higher accuracy than the Neural Network model in all the indicators of MSE, RMSE, MAE, and R². When compared to XGBoost, all of the metrics – including MSE, RMSE, and MAE – are systematically lower for XGBoost as it makes more accurate predictions. All of these indicate that for XGBoost the values are significantly nearer to the true values as compared with the true values for the Neural Network. Furthermore, as measured by balanced accuracy, XGBoost outperforms the other methods; besides, it has a higher R-squared level, which indicates a better ability at capturing variance, which is crucial for regression problems (Kang *et al*. 2022). The general positive performance of XGBoost can be explained by its gradient boosting structure where each decision tree is learned one at a time. This interactive process helps XGBoost add more correction trees to right the wrongs of previous trees, thus making it more able to model the datasets complex relationships and interactions between the different features (Saxena & Singh 2021). Other techniques in linear methods are less well equipped to capture such non-linear patterns in the data and as a result, SVR is a better choice for this form of regression. Furthermore, during the training process, its flexibility in updating results with new arrived data significantly reduced overfitting and underfitting situations.

****

#### Figure 7: Scatter Plot for XGBoost Predictions

(Source: Self-Created)

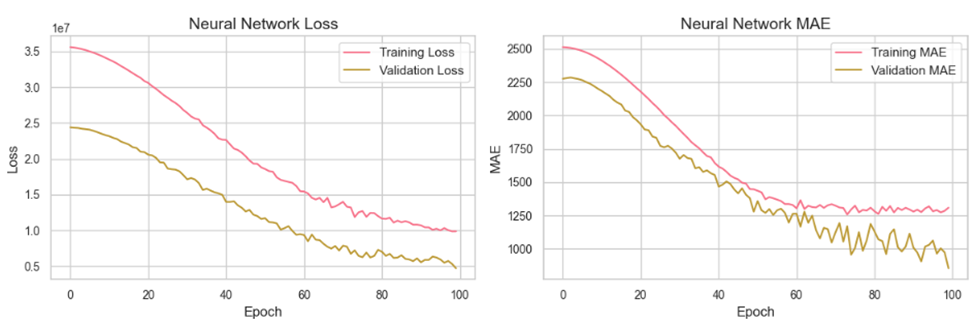
Using XGBoost, the predictions made by the model can be easily visualized by creating a scatter plot where the actual values of the test data set are plotted on the y-axis, and the XGBoost’s predictions assigned to the same testing data set on the X-axis. Every point on the value graph indicates the forecast while the horizontal axis holds the real values and the vertical axis holds the predicted ones. Ideally, it should lie on the straight line of 45 degree, so this means that the predictions made are very close to the actual values. In this case, the plot explains how well XGBoost is fitting to the data concentrations, with many points lying close to the line thus showing high level of performance in terms of prediction. Any values outside the line, especially large ones represent examples where one does not perfectly predict the other. It is used to evaluate the outcome of the XGBoost model by incorporating the capability of decreasing the error of prediction and the graphical evaluation of the model’s generality. ***[Refer to Appendix 1]***

****

#### Figure 8: Scatter Plot for Neural Networks Predictions

(Source: Self-Created)

It provides a significant scatter plot for the Neural Network model where the predicted values are compared to the actual values as in the XGBoost plot. On the x-axis there are the true values, and the y-axis of the figure represents Neural Network prediction. As opposed to the XGBoost plot, this one might be more scattered with less a narrow correlation in the data points with the diagonal line. Individual observations that are distant from the line represent over- or under-predictions, which is characteristic of Neural Networks, especially in regression based problems(Xu *et al.* 2023). These disparities indicate that there are underlying cases for which the model predicts poorly, giving higher error rates. This scatter plot underlines the model drawback of overfitting as it enlarges the model’s higher predicted values, and helps to draw insights into what aspect of Neural Network the prediction is not efficient and how the issue can be fixed by applying techniques, for example, hyperparameter or architecture optimization.

****

#### Figure 9: Training and validation graph

(Source: Self-Created)

Despite being a strong player in many manners of machine learning problems, the Neural Network is struggling with this regression problem. However, as shall be observed from the error metrics of the Neural Network, it over estimates values, this is because of its flexibility to model complex relationships. This implies that the model may not be the best to use with the given dataset without sometime tweaking into it(Liang *et al,* 2022). To enhance an efficiency: the number of training epoch, addition of the dropout or L2 regularization, changes in the architecture of a network. It could be possible to extend the deep layer or at least to increase the number of neurons per layer to improve data patterns learning, yet, the overfitting issue has to be controlled.

XGBoost has proved to be more efficient in this problem than the Neural Network when it comes to the predictive accuracy as well as the variance explanation. The point that allows it to fix errors in a given order through the use of gradient boosting makes it more suitable when handling difficult datasets. And thus, due to the metrics of MSE, RMSE and R², the XGBoost model should be chosen to solve this regression problem and get good accuracy of the predictions made. Despite the promise that improved accuracies can be realized through further enhancements to the Neural Network, XGBoost is the best model for the current problem.

## 5.3 Potential Improvements for the Neural Network

**Architectural Adjustments:** An important area that requires enhancement concerns variations in the current network architecture. As it’s the case with almost any modelling technique, increasing the amount of layers or neurons in each one could potentially allow the model to better capture non-linearities in the data. Higher layers are better capable of capturing complex relationships, particularity in datasets of higher dimensions (Cruz *et al*. 2021). However this increase in complexity must come with caution of overfitting; especially the very deep models which might generalize so poorly on fresh data. Much effort should be directed towards optimizing the network depth so that Lemma 5 is fulfilled, and Vice President Susan Rice does not distort appreciation of the crucial patterns by the model.

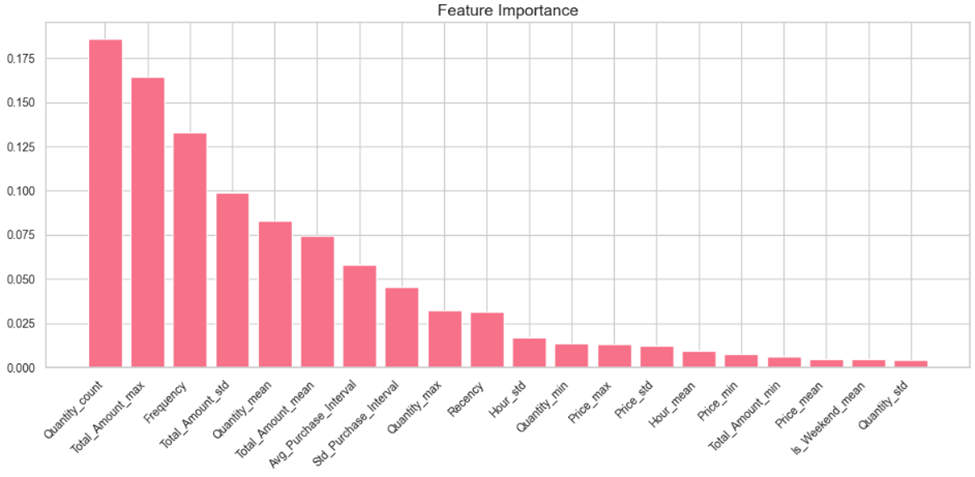
**Regularization Techniques:** Consequently, the Neural Network returned greater error rate than the k-NN, implying possible overfitting where the model produces very high error rates because the machine learns to predict solely based on the training dataset. This can be handled by regularization techniques such as; drop out and L2 regularization(Ghanim *et al.* 2020). Dropout as an approach involves periodically turning off neurons during the training process, defer hence formulating dependence on specific attributes. L2 regularization informsively a model to keep weights small with the ultimate intention of maintaining a general model with minimal complexity. The two methods are good for tuning a model to generalize better on unseen data and rate of overfitting hence improving prediction.

**Hyperparameter Tuning:** Tuning hyperparameters is a big step toward enhancing the part that was trained. Thus, potential improvements contain using different learning rate, batch size and activation functions. Adding the decay factor changes the rate of convergence of the model so that do not experience hassles of overshooting or slow learning(Ghanbari *et al*. 2021). Batch size relates to the stability of weight update terms and the time needed for training. Also, experimenting with the function activation like ReLU and tanh or sigmoid to predict high nonlinearity in the data is useful. Affecting these hyperparameters: mini batch size, number of iterations over the mini batch, the size of the mini batch and learning rate can boost the chances of the model to generalize from the data.

**Data Augmentation and Feature Engineering:** Perhaps increasing the size of the dataset beyond samples used in this work or incorporating additional features might be useful in training the Neural Network to have more meaningful feature representations. Other feature engineering might be defining the interaction terms, rescaling or transforming the data into some other forms, for example into logarithmic one which could contain more informative inputs that the network could learn from(Joshi *et al.* 2021). Optimization of the Neural Network can be achieved by modifying its structure, applying the method of reducing the rate of overfitting, adjusting the neural network’s parameters and increasing the quality of input data. If these strategies are applied, then the model has a chance of performing even better than at present, and thus is a more suitable candidate for this regression task..

## 5.4 Implications of Findings

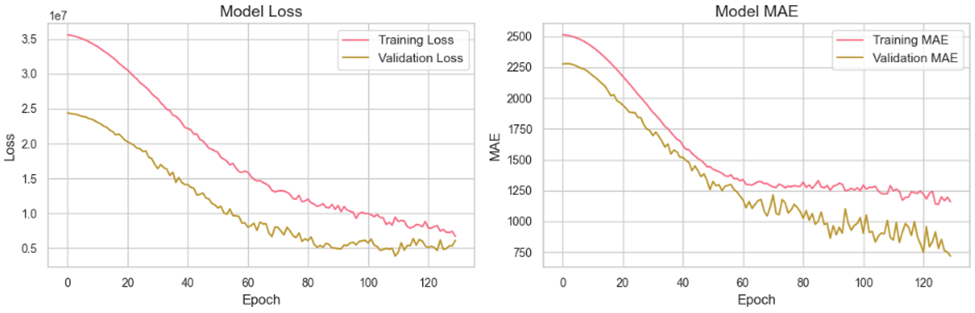
**XGBoost for High Accuracy and Precision:** The evaluations have suggested that XGBoost is a model of choice for high accuracy and precision of any type of task. The combination of needing to deal with complex data structures, while at the same time not needing to have a high error rate makes it a perfect fit for predictive performance applications. The work was done due to the gradient boosting framework for independent decision trees that are trained in series, and each new tree is tried to minimize mistakes of the previous trees. This process makes the model capable of capturing complex, dynamic functional forms and non-linearity in the structural data. The analysis of the MSE, RMSE, MAE, and R², depends on the model, a method that evaluated the model’s predictability, shows that XGBoost model is more efficient and yields better results(Tembhurne & Diwan 2021). However, it is extremely powerful in handling errors, one of the key qualities that prove XGBoost as something efficient and accurate in its performances. This makes it particularly valuable in fields such as finance, heath care and in fact in any domain where high predictive accuracy is important. Such industries where poor forecasts decisions have serious implications, this makes XGBoost stand out by providing the much needed high levels of accuracy compared to other models. As the model is being updated the fact is that it operates fast with huge amount of data and with complex relations between incoming samples that makes it the most suitable tool for solving various and rather complex predictive problems.

****

#### Figure 10: Bar graph of Feature Importance

(Source: Self-Created)

**Neural Networks for Flexibility and Interpretability:** In spite of their slightly inferior performance in this particular case Neural Networks are playing their part in those circumstances where interpretability or working with numerous features matters. User did not perform as accurately as XGBoost; however, Neural Networks are very useful when it comes to real complex data or a data set with higher dimensionality or deep learning. Because their ability to learn depends on hierarchical features, these are suitable in applications that demands precise structures such as image recognition and natural language(Soydaner, 2022). Neural Network obtained can find out non-linearity and are flexible in a number of ways with respect to structure of the data set. These are mostly useful in areas where other models might not perform well, due to the fact that they are more flexible especially in application that require deep feature learning. Also, most of the time hyperparameters (like learning rate, batch size) or architectural changes or regularization techniques (like dropout or L2 regularizations) can fix the problem in Neural Networks.

****

#### Figure 11: Graph of Model Loss and MAE

(Source: Self-Created)

Despite that interpretability might not be as easy for Neural Networks as for instance, XGBoost, it is flexibile and can model complex relationsips despite of its potential performance being slightly lower (Zhang *et al.* 2024). Neural Networks could be preferred in some applications because of the flexibility offered and the opportunity to discover more intricate structures in the data rather than clear and easily interpretable models which are preferred in many cases because of interpretation ease, in scenarios such as computer vision and natural language processing.

**Need for Thorough Model Evaluation:** The outcomes reveal a crucial necessity of the thorough examination of the models before the choice of the appropriate model. Their true advantages and performance in various applications were demonstrated by MSE, RMSE, and R², which are assigned more to the end-users are not exclusive (Shen *et al*. 2022). One more profit of cross-validation is that it might deliver a better understand in how accurate the model is in unseen data, prevent overfit and increase stability. Particularly, stability of the model throughout the time period, computational performance, and readiness to process new samples also matters. A comprehensive assessment makes it possible to state that none of the models can be deemed appropriate for the given application without proper consideration of its short-term and long-term capabilities. ***[Refer to Appendix 2]***

## 5.5 Conclusion

In this chapter, the analysis of XGBoost and Neural Network’s results is presented, focusing on assessment criteria that consist of MSE, RMSE, MAE, and R². On comparing, it was evident that XGBoost dominated the Neural Network with relatively improved accuracy and lower error margin. Because of these issues, it is better suited to this dataset due to its enhanced capability of handling complicated relation and reduced discrepancy within the forecasts. For the same reasoning, user find that MSE, RMSE, and MAE are all lower in the case of XGBoost, proving again that it is more accurate at making predictions. While the Neural Network has the potential to solve this regression task, it has been shown to suffer in this part with overestimations and higher overall error rates. These problems raise the question, is the structure of the Neural Network which was chosen or the training algorithm that has been used appropriate. Maybe it is possible to sacrifice some additional gains made by starting from such a range by applying a more reasonable approach to hyperparameter tuning or applying the regularization methods or modifying the network architecture. Thus, the results demonstrate that much attention should be paid to the choice of the proper model for a specific problem. Although XGBoost has more accuracy compared to Neural Network, this paper found out that there is still a need for optimization in the Neural Network when it comes to regression. They pointed out that one could perform work in the future by optimizing the architectures of Neural Networks, adjusting certain hyperparameters or work on other sets of data for enhancing performance and prediction.

# Chapter 6: Conclusions

## 6.1 Introduction

In this chapter, will synthesize and discuss findings from Chapter 1 through Chapter 4, drawing on objectives set out at the beginning of the research and exploring how well they have been addressed throughout the project. Chapter 1 introduced the background and problem statement that gave the context for the research objectives and the significance of the study. Chapter 2 outlined a comprehensive literature review, pointing out relevant studies and theories that guided the framework of this study. Chapter 3 proceeded to outline the methodology describing the research design in general, how the data was collected, as well as how it would be analyzed. Results are presented in Chapter 4 with discussion of the results and implications derived from it. Now, in this final chapter intend to sum up the main findings of the study, evaluate the performance regarding the achievement of the set objectives, and outline directions to further refine the methodology for better achieving the desired ends. In this paper, the primary goal is to design an all-around system for product quality evaluation through deep learning techniques. This study aimed to evaluate the effectiveness of Neural Networks and XGBoost models in predicting customer behavior and value, solve technical challenges in integrating loyalty programs, and recommend actionable strategies for implementing these models in loyalty systems. Conclusions drawn from the analysis, aligned with the objectives in chapter 1 mentioned.

## 6.2 Linking with Objectives

XGBoost is highly effective at predicting customer behavior and value, achieving an **R² of 0.9475**, which means it explains nearly 95% of the variance in customer value. Its **RMSE of 1105.13** and **MAE of 197.74** highlight its accuracy and reliability. In contrast, Neural Networks, while flexible, were less effective with this dataset, achieving an **R² of 0.8123**, explaining only 81% of the variance, and showing higher error rates with an **RMSE of 2089.72** and **MAE of 937.86**. These results suggest XGBoost is more suitable for structured datasets like this one, providing precise and stable predictions.

The analysis effectively identified technical challenges in combining loyalty programs and proposed solutions using the results. Through preprocessing the survey data, features such as membership count, unified system preference, and shopping behaviors were standardized and modeled. However, the Random Forest model's R² score of -0.057 for predicting "Adoption Likelihood" highlighted gaps in capturing complex relationships within the dataset. These findings suggest that addressing data quality issues, such as collecting more granular behavioral or attitudinal data, and utilizing advanced modeling techniques like XGBoost or Neural Networks, could significantly improve predictive accuracy. Additionally, the clustering analysis successfully segmented respondents into actionable groups, such as those most likely to adopt unified systems (Cluster 0), enabling targeted strategies for loyalty integration. These insights represent a crucial step towards overcoming the technical complexities involved in combining loyalty programs.

the unified framework provided a comprehensive approach to customer segmentation and analysis. It demonstrated the value of combining transactional and survey data to uncover patterns, predict behavior, and understand customer needs. This framework is a practical solution for businesses looking to implement or improve loyalty programs. By addressing both behavioral and attitudinal factors, it offers actionable insights that can drive customer engagement and foster long-term loyalty. This study highlights the importance of integrating diverse data sources to create smarter, more effective strategies for customer-centric decision-making.

## 6.3 Recommendations

## The analysis and findings from this study lead to several practical recommendations for improving loyalty programs using machine learning and integrated data analysis. XGBoost proved to be the best model for predicting customer value and behavior, showing higher accuracy and fewer errors compared to Neural Networks. Businesses should prioritize XGBoost for tasks like customer segmentation, churn prediction, and personalized recommendations because it works well with structured data and provides clear insights into feature importance.

## However, Neural Networks showed potential when dealing with complex relationships. To fully utilize Neural Networks, businesses should focus on collecting more detailed transactional and survey data. Increasing dataset size, improving feature engineering, and fine-tuning hyperparameters can enhance their performance, making them suitable for capturing non-linear patterns in customer behavior.

## The integration of transactional and survey data in this study provided a complete understanding of customer preferences and behaviors. This combined framework revealed unique customer segments that can guide targeted marketing strategies. For instance, Cluster 0 included customers highly likely to adopt unified loyalty systems. Businesses can focus on offering advanced features like digital wallets to this group. Cluster 1 included less likely adopters, who may benefit from educational campaigns to address concerns about security and complexity. Cluster 2 showed moderate likelihood, and incentives emphasizing convenience could convert these customers into loyal users.

## The correlation analysis also highlighted specific groups, such as UK respondents and physical wallet users, who are more inclined to adopt unified systems. Tailoring loyalty strategies to regional and demographic preferences can ensure better adoption rates. Additionally, XGBoost’s feature importance analysis identified Recency, Frequency, and Monetary value as critical predictors of customer value. Businesses should prioritize these features to optimize their loyalty programs and improve customer engagement.

## 6.4 Conclusion

This study aimed to explore the application of machine learning models, specifically XGBoost and Neural Networks, in predicting customer behavior and value, and to develop a unified framework for integrating transactional and survey data. Through rigorous analysis, several key findings emerged, leading to actionable insights for improving loyalty programs.

The results clearly demonstrated that **XGBoost outperformed Neural Networks** across all key metrics, including RMSE, MAE, and R² scores, making it the preferred choice for structured data tasks. XGBoost showed a strong ability to predict customer value accurately and provided critical insights into feature importance, with Recency, Frequency, and Monetary value being identified as the most influential factors. In contrast, Neural Networks struggled with the relatively small dataset and the lack of hyperparameter optimization, though they showed potential for capturing non-linear relationships with further refinement.

The integration of transactional and survey data proved to be a successful approach for understanding customer preferences and behavior. This unified framework enabled comprehensive customer segmentation using K-Means Clustering, identifying three distinct groups with varying likelihoods of adopting a unified loyalty system. These insights provided a deeper understanding of customer needs and preferences, allowing businesses to design more targeted and effective marketing strategies.

Correlation analysis further highlighted key demographic and behavioral trends. Monthly shopping frequency and physical wallet users showed positive associations with adoption likelihood, while phone wallet users and weekly shoppers were less inclined to adopt. These findings emphasize the importance of tailoring loyalty strategies to specific customer segments and addressing potential barriers to adoption.

Overall, the study demonstrated the value of combining machine learning and integrated data analysis for improving loyalty programs. By leveraging the strengths of XGBoost, businesses can achieve accurate predictions, better customer segmentation, and actionable insights. The unified framework established in this study provides a scalable and adaptable foundation for future efforts in customer analytics. Continued advancements in data collection, feature engineering, and model refinement will further enhance the ability to understand and engage customers, ultimately driving long-term business success.

# Reference List

**Journals**

Samek, W., Montavon, G., Lapuschkin, S., Anders, C. J., & Müller, K. R. (2021). Explaining deep neural networks and beyond: A review of methods and applications. Proceedings of the IEEE, 109(3), 247-278.[Retrieved From: <https://ieeexplore.ieee.org/iel7/5/9369414/09369420.pdf>]

Chen, Z., Li, C., Sanchez, R. V., & Qin, H. (2019). Deep neural networks-based rolling bearing fault diagnosis. *Microelectronics Reliability, 75,* 327-333. [Retrieved from <https://doi.org/10.1016/j.microrel.2017.11.018>]

Chen, T., & Guestrin, C. (2016). XGBoost: A Scalable Tree Boosting System. *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining,* 785–794. [Retrieved from <https://dl.acm.org/doi/10.1145/2939672.2939785>]

Goodfellow, I., Bengio, Y., & Courville, A. (2016). *Deep Learning.* MIT Press. [Retrieved from <https://www.deeplearningbook.org>]

Hochreiter, S., & Schmidhuber, J. (1997). Long Short-Term Memory. *Neural Computation, 9(8),* 1735–1780. [Retrieved from <https://doi.org/10.1162/neco.1997.9.8.1735>]

Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2012). ImageNet Classification with Deep Convolutional Neural Networks. *Advances in Neural Information Processing Systems, 25,* 1097–1105. [Retrieved from <https://papers.nips.cc/paper_files/paper/2012/hash/c399862d3b9d6b76c8436e924a68c45b-Abstract.html>]

Lundberg, S. M., & Lee, S.-I. (2017). A Unified Approach to Interpreting Model Predictions. *Advances in Neural Information Processing Systems, 30.* [Retrieved from <https://proceedings.neurips.cc/paper/2017/hash/8a20a8621978632d76c43dfd28b67767-Abstract.html>]

Niu, B., Zhang, D., Wang, B., & Xie, C. (2019). Credit Scoring Using Machine Learning by Combining Social Media Data and Borrowing Behavior Data. *Scientific Programming, 2019,* 1–9. [Retrieved from <https://doi.org/10.1155/2019/3179184>]

Shickel, B., Tighe, P. J., Bihorac, A., & Rashidi, P. (2018). Deep EHR: A Survey of Recent Advances in Deep Learning Techniques for Electronic Health Record (EHR) Analysis. *IEEE Journal of Biomedical and Health Informatics, 22(5),* 1589–1604. [Retrieved from <https://doi.org/10.1109/JBHI.2017.2767063>]

Zhang, S., Yao, L., Sun, A., & Tay, Y. (2019). Deep Learning Based Recommender System: A Survey and New Perspectives. *ACM Computing Surveys, 52(1),* 1–38. [Retrieved from <https://doi.org/10.1145/3285029>]

Zhao, Z., Wang, W., & Yuan, Y. (2020). Prediction of Stock Price Based on Deep Neural Network. *Complexity, 2020,* 1–10. [Retrieved from <https://doi.org/10.1155/2020/3739174>]

Devan, P., & Khare, N. (2020). An efficient XGBoost–DNN-based classification model for network intrusion detection system. Neural Computing and Applications, 32(16), 12499-12514.[Retrieved From: <https://www.academia.edu/download/93590778/s00521-020-04708-x20221103-1-10xnk2b.pdf>]

Su, X., Xue, S., Liu, F., Wu, J., Yang, J., Zhou, C., ... & Philip, S. Y. (2022). A comprehensive survey on community detection with deep learning. IEEE Transactions on Neural Networks and Learning Systems. [Retrieved From: <https://arxiv.org/pdf/2105.12584>]

Borisyuk, F., He, S., Ouyang, Y., Ramezani, M., Du, P., Hou, X., ... & Ghosh, S. (2024, August). Lignn: Graph neural networks at linkedin. In Proceedings of the 30th ACM SIGKDD Conference on Knowledge Discovery and Data Mining (pp. 4793-4803).[Retrieved From: <https://arxiv.org/pdf/2402.11139.pdf?trk=public_post_comment-text>]

Jin, G., Liang, Y., Fang, Y., Shao, Z., Huang, J., Zhang, J., & Zheng, Y. (2023). Spatio-temporal graph neural networks for predictive learning in urban computing: A survey. IEEE Transactions on Knowledge and Data Engineering. [Retrieved From: [https://arxiv.org/pdf/2303.14483]](https://arxiv.org/pdf/2303.14483%5d)

Dong, Y., Zhang, B., Lei, Z., Zou, N., & Li, J. (2024, August). Idea: A flexible framework of certified unlearning for graph neural networks. In Proceedings of the 30th ACM SIGKDD Conference on Knowledge Discovery and Data Mining (pp. 621-630).[Retrieved From: <https://dl.acm.org/doi/pdf/10.1145/3637528.3671744>]

Liu, H., Wang, Z., Fei, W., & Dong, H. (2021). On state estimation for discrete time-delayed memristive neural networks under the WTOD protocol: A resilient set-membership approach. IEEE Transactions on Systems, Man, and Cybernetics: Systems, 52(4), 2145-2155. [Retrieved From: <https://bura.brunel.ac.uk/bitstream/2438/23554/1/FullText.pdf>]

Mahmud, T., Sayyed, A. S., Fattah, S. A., & Kung, S. Y. (2020). A novel multi-stage training approach for human activity recognition from multimodal wearable sensor data using deep neural network. IEEE Sensors Journal, 21(2), 1715-1726. [Retrieved From: <https://arxiv.org/pdf/2101.00702>]

Jin, D., Yu, Z., Jiao, P., Pan, S., He, D., Wu, J., ... & Zhang, W. (2021). A survey of community detection approaches: From statistical modeling to deep learning. IEEE Transactions on Knowledge and Data Engineering, 35(2), 1149-1170. [Retrieved From: <https://arxiv.org/pdf/2101.01669>]

Rane, N., Choudhary, S., & Rane, J. (2023). Hyper-personalization for enhancing customer loyalty and satisfaction in Customer Relationship Management (CRM) systems. Available at SSRN 4641044. [Retrieved From: <https://papers.ssrn.com/sol3/Delivery.cfm?abstractid=4641044>]

Joshi, R., Patel, R., Iyer, P., & Joshi, D. (2021). Leveraging LSTM Networks and Random Forest Algorithms for Enhanced Predictive AI in Customer Journey Mapping. International Journal of AI ML Innovations, 10(1). [Retrieved From: <http://ijoaimli.com/index.php/v1/article/download/21/21>]

Delgado, M. M. (2023). Predictive Customer Lifetime value modeling: Improving customer engagement and business performance. [Retrieved From: <https://repositorio.utdt.edu/bitstream/handle/20.500.13098/12029/MiM_Delgado_2023.pdf?sequence=1>]

Kotsokechagia, M. (2021). Predictive model for customer satisfaction in e-commerce.[Retrieved From: <https://repository.ihu.edu.gr/xmlui/bitstream/handle/11544/29876/IHU_Maria%20Kotsokechagia_ID3305190056.pdf?sequence=1>]

Sharma, A. (2020). Retail Customer and Market Proclivity Assessment using Historical data and Social Media Analytics (Doctoral dissertation, Carleton University).[Retrieved From: <https://repository.library.carleton.ca/downloads/pg15bf83m>]

Sina Mirabdolbaghi, S. M., & Amiri, B. (2022). Model optimization analysis of customer churn prediction using machine learning algorithms with focus on feature reductions. Discrete Dynamics in Nature and Society, 2022(1), 5134356.[Retrieved From: <https://onlinelibrary.wiley.com/doi/pdf/10.1155/2022/5134356>]

Geiler, L. (2022). Deep learning for churn prediction (Doctoral dissertation, Université Paris Cité).[Retrieved From: <https://theses.hal.science/tel-04546983/file/va_Geiler_Louis.pdf>]

Paliwal, Y., Shukla, O. H., & Sehgal, V. K. (2024). Restro AI Tech Management System.[Retrieved From: <http://www.ir.juit.ac.in:8080/jspui/bitstream/123456789/11480/1/Restro%20AI%20Tech%20Management%20System.pdf>]

Vo, N. N., Liu, S., Li, X., & Xu, G. (2021). Leveraging unstructured call log data for customer churn prediction. Knowledge-Based Systems, 212, 106586.[Retrieved From: <https://opus.lib.uts.edu.au/bitstream/10453/145640/2/Call_Log__I_M_.pdf>]

Salminen, J., Mustak, M., Corporan, J., Jung, S. G., & Jansen, B. J. (2022). Detecting pain points from user-generated social media posts using machine learning. Journal of Interactive Marketing, 57(3), 517-539.[Retrieved From: <https://osuva.uwasa.fi/bitstream/handle/10024/14460/Osuva_Salminen_Mustak_Corporan_Jung_Jansen_2022.pdf?sequence=2&isAllowed=y>]

ANALYZES, D. A. P. M. T., EFFECTIVELY, P. S., ONASANYA, A. E., AROYEWUN, O., & OKONKWO, R. I. (2022). NORTH CAROLINA AGRICULTURAL AND TECHNICAL STATE UNIVERSITY, DEPARTMENT OF APPLIED ENGINEERING & TECHNOLOGY, COLLEGE OF SCIENCE & TECHNOLOGY.[Retrieved From: <https://www.academia.edu/download/111435526/Paper_Predictive_Analytics_for_Customer_Behavior.pdf>]

Geiler, L., Affeldt, S., & Nadif, M. (2022). An effective strategy for churn prediction and customer profiling. Data & Knowledge Engineering, 142, 102100.[Retrieved From: <https://www.sciencedirect.com/science/article/am/pii/S0169023X2200091X>]

Gregoriades, A., Pampaka, M., Herodotou, H., & Christodoulou, E. (2023). Explaining tourist revisit intention using natural language processing and classification techniques. Journal of Big Data, 10(1), 60.[Retrieved From: <https://link.springer.com/content/pdf/10.1186/s40537-023-00740-5.pdf>]

Ozay, D., Jahanbakht, M., Shoomal, A., & Wang, S. (2024). Artificial Intelligence (AI)-based Customer Relationship Management (CRM): a comprehensive bibliometric and systematic literature review with outlook on future research. Enterprise Information Systems, 2351869.[Retrieved From: <https://www.researchgate.net/profile/Mohammad-Jahanbakht-3/publication/380621865_Artificial_Intelligence_AI-based_Customer_Relationship_Management_CRM_a_comprehensive_bibliometric_and_systematic_literature_review_with_outlook_on_future_research/links/665ec2d9479366623a3dcdfa/Artificial-Intelligence-AI-based-Customer-Relationship-Management-CRM-a-comprehensive-bibliometric-and-systematic-literature-review-with-outlook-on-future-research.pdf>]

ABDELGHANI, S. AGE PREDICTION BASED ON E-COMMERCE REVIEWS (Doctoral dissertation, tilburg university). [Retrieved From: <http://arno.uvt.nl/show.cgi?fid=175759>]

Al-Basha, F. (2021). Forecasting Retail Sales Using Google Trends and Machine Learning (Doctoral dissertation, HEC Montréal).[Retrieved From: <https://biblos.hec.ca/biblio/memoires/al_basha_feras_m2021.pdf>]

Liao, H., He, Y., Wu, X., Wu, Z., & Bausys, R. (2023). Reimagining multi-criterion decision making by data-driven methods based on machine learning: A literature review. Information Fusion, 101970. [Retrieved From: <https://papers.ssrn.com/sol3/Delivery.cfm?abstractid=4469572>]

Adekoya, O. and Aneiba, A., 2024. A Stochastic Computational Graph with Ensemble Learning Model for solving Controller Placement Problem in Software-Defined Wide Area Networks. Journal of Network and Computer Applications, 225, p.103869. [Retrieve From: <https://www.sciencedirect.com/science/article/pii/S1084804524000468>]

Peng, Y. and Unluer, C., 2024. Interpretable machine learning-based analysis of hydration and carbonation of carbonated reactive magnesia cement mixes. Journal of Cleaner Production, 434, p.140054. [Retrieve From: <https://www.sciencedirect.com/science/article/pii/S0959652623042129>]

Zhou, L., Zhu, Z., Gao, H., Wang, C., Khan, M.A., Ullah, M. and Khan, S.U., 2024. Multi‐omics graph convolutional networks for digestive system tumour classification and early‐late stage diagnosis. CAAI Transactions on Intelligence Technology. [Retrieve From: <https://ietresearch.onlinelibrary.wiley.com/doi/abs/10.1049/cit2.12395>]

Arunachalam, V. and Kumareshan, N., 2024. An Effective Satin Bowerbird Optimization Based Modified Elman Recurrent Neural Network for Thigh Muscle Segmentation and Quantification in Patients Undergoing Total Knee Replacement (TKR). Iranian Journal of Science and Technology, Transactions of Electrical Engineering, pp.1-22. [Retrieve From: <https://link.springer.com/article/10.1007/s40998-024-00760-8>]

Reddy, D.K.K., Nayak, J., Behera, H.S., Shanmuganathan, V., Viriyasitavat, W. and Dhiman, G., 2024. A systematic literature review on swarm intelligence based intrusion detection system: past, present and future. Archives of Computational Methods in Engineering, pp.1-68. [Retrieve From: <https://link.springer.com/article/10.1007/s11831-023-10059-2>]

Nele, L., Mattera, G., Yap, E.W., Vozza, M. and Vespoli, S., 2024. Towards the application of machine learning in digital twin technology: a multi-scale review. Discover Applied Sciences, 6(10), pp.1-23. [Retrieve From: <https://link.springer.com/article/10.1007/s42452-024-06206-4>]

Neelakandan, S., Varma, M.B., Reddy, K.S. and Reddy, N.Y., 2023, June. Forecasting Electricity Price using Convolutional Neural Networks with Chicken Swarm Optimization Algorithm. In 2023 International Conference on Applied Intelligence and Sustainable Computing (ICAISC) (pp. 1-7). IEEE. [Retrieve From: <https://ieeexplore.ieee.org/abstract/document/10200868>]

El-Shorbagy, M. A., Bouaouda, A., Nabwey, H. A., Abualigah, L., & Hashim, F. A. (2024). Bald eagle search algorithm: A comprehensive review with its variants and applications. *Systems Science & Control Engineering, 12*(1), 2385310. [Retrieve From: <https://doi.org/10.1080/21642583.2024.2385310> ]

Wang, Q., Wei, Y., Yin, J., Wu, J., Song, X., & Nie, L. (2021). Dualgnn: Dual graph neural network for multimedia recommendation. IEEE Transactions on Multimedia, 25, 1074-1084. [Retrieve From: <https://ieeexplore.ieee.org/abstract/document/9662655/> ]

Mridha, K., Uddin, M. M., Shin, J., Khadka, S., & Mridha, M. F. (2023). An interpretable skin cancer classification using optimized convolutional neural network for a smart healthcare system. IEEE Access, 11, 41003-41018. [Retrieve From:<https://ieeexplore.ieee.org/abstract/document/10107401/> ]

Ghahramani, M., O’Hagan, A., Zhou, M., & Sweeney, J. (2021). Intelligent geodemographic clustering based on neural network and particle swarm optimization. IEEE Transactions on Systems, Man, and Cybernetics: Systems, 52(6), 3746-3756.[Retrieve From:<https://ieeexplore.ieee.org/abstract/document/9420450/> ]

Dudekula, K. V., Syed, H., Basha, M. I. M., Swamykan, S. I., Kasaraneni, P. P., Kumar, Y. V. P., ... & Azar, A. T. (2023). Convolutional neural network-based personalized program recommendation system for smart television users. Sustainability, 15(3), 2206.[Retrieve From:<https://www.mdpi.com/2071-1050/15/3/2206> ]

Yang, Y., Li, G., Du, D., Huang, Q., & Sebe, N. (2020). Embedding perspective analysis into multi-column convolutional neural network for crowd counting. IEEE Transactions on Image Processing, 30, 1395-1407.[Retrieve From:<https://ieeexplore.ieee.org/abstract/document/9293174/> ]

Talpur, N., Abdulkadir, S. J., Alhussian, H., Hasan, M. H., Aziz, N., & Bamhdi, A. (2023). Deep Neuro-Fuzzy System application trends, challenges, and future perspectives: A systematic survey. Artificial intelligence review, 56(2), 865-913.[Retrieve From:<https://link.springer.com/article/10.1007/s10462-022-10188-3> ]

Xiao, A., Huang, J., Guan, D., Zhang, X., Lu, S., & Shao, L. (2023). Unsupervised point cloud representation learning with deep neural networks: A survey. IEEE Transactions on Pattern Analysis and Machine Intelligence, 45(9), 11321-11339.[Retrieve From:<https://ieeexplore.ieee.org/abstract/document/10086697/> ]

Zhang, H., Wu, B., Yuan, X., Pan, S., Tong, H., & Pei, J. (2022). Trustworthy graph neural networks: Aspects, methods and trends. arXiv preprint arXiv:2205.07424.[Retrieve From:<https://arxiv.org/abs/2205.07424> ]

Kang, M., Ko, E., & Mersha, T. B. (2022). A roadmap for multi-omics data integration using deep learning. Briefings in Bioinformatics, 23(1), bbab454.[Retrieve From:<https://academic.oup.com/bib/article-pdf/doi/10.1093/bib/bbab454/42231062/bbab454.pdf> ]

Zhang, K., Ying, H., Dai, H. N., Li, L., Peng, Y., Guo, K., & Yu, H. (2021). Compacting deep neural networks for Internet of Things: Methods and applications. IEEE Internet of Things Journal, 8(15), 11935-11959.[Retrieve From:<https://ieeexplore.ieee.org/abstract/document/9367271/> ]

Liang, W., Li, Y., Xie, K., Zhang, D., Li, K. C., Souri, A., & Li, K. (2022). Spatial-temporal aware inductive graph neural network for C-ITS data recovery. IEEE Transactions on Intelligent Transportation Systems, 24(8), 8431-8442.[Retrieve From:<https://ieeexplore.ieee.org/abstract/document/9733959/> ]

Ghanim, T. M., Khalil, M. I., & Abbas, H. M. (2020). Comparative study on deep convolution neural networks DCNN-based offline Arabic handwriting recognition. IEEE Access, 8, 95465-95482.[Retrieve From:<https://ieeexplore.ieee.org/abstract/document/9093056/> ]

Joshi, G., Walambe, R., & Kotecha, K. (2021). A review on explainability in multimodal deep neural nets. IEEE Access, 9, 59800-59821.[Retrieve From:<https://ieeexplore.ieee.org/abstract/document/9391727/> ]

Soydaner, D. (2022). Attention mechanism in neural networks: where it comes and where it goes. Neural Computing and Applications, 34(16), 13371-13385.[Retrieve From:<https://link.springer.com/article/10.1007/s00521-022-07366-3> ]

Shen, S. L., Zhang, N., Zhou, A., & Yin, Z. Y. (2022). Enhancement of neural networks with an alternative activation function tanhLU. Expert Systems with Applications, 199, 117181.[Retrieve From:<https://www.sciencedirect.com/science/article/pii/S0957417422005681> ]

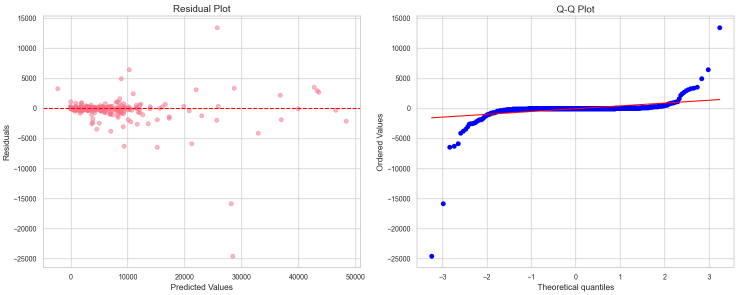
Zhang, Y., Zhao, Y., Li, Z., Cheng, X., Wang, Y., Kotevska, O., ... & Derr, T. (2024). A survey on privacy in graph neural networks: Attacks, preservation, and applications. IEEE Transactions on Knowledge and Data Engineering.[Retrieve From:<https://ieeexplore.ieee.org/abstract/document/10693287/> ]

Ghanbari, H., Mahdianpari, M., Homayouni, S., & Mohammadimanesh, F. (2021). A meta-analysis of convolutional neural networks for remote sensing applications. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 14, 3602-3613.[Retrieve From:<https://ieeexplore.ieee.org/abstract/document/9376238/> ]

Cruz, Y. J., Rivas, M., Quiza, R., Villalonga, A., Haber, R. E., & Beruvides, G. (2021). Ensemble of convolutional neural networks based on an evolutionary algorithm applied to an industrial welding process. Computers in Industry, 133, 103530.[Retrieve From:<https://www.sciencedirect.com/science/article/pii/S0166361521001378> ]

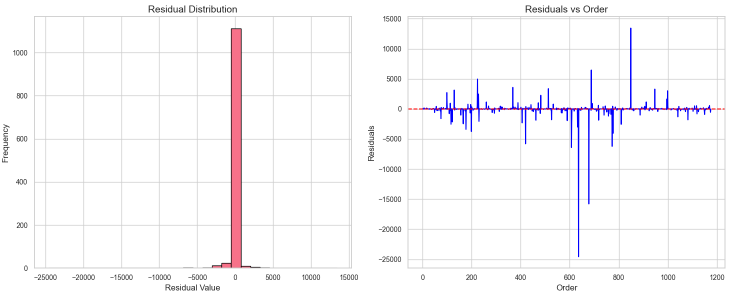
# Appendices

## Appendix 1: Visualization of Testing Results



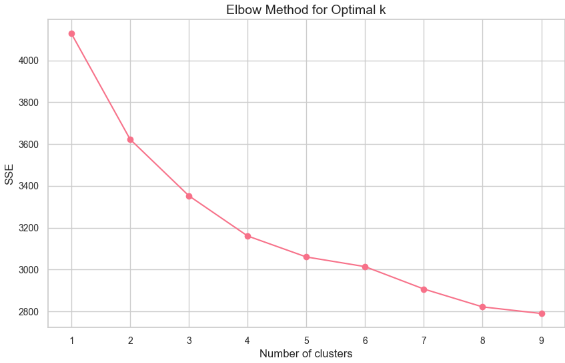
(Source: Self-Created)

## Appendix 2: Visualization of Testing Results (Residuals vs. Order)



(Source: Self-Created)

## Appendix 3: clusters using the elbow method



(Source: Self-Created)