Enhancing Customer Churn Prediction in the Telecommunications Sector: Benchmarking Neural Networks Against Traditional Machine Learning Models

MSc Research Project
MSc in AI for Business

Ebubekir Ayhan
Student ID: x23141727

School of Computing
National College of Ireland

Supervisor: Victor Del Rosal

National College of Ireland
Project Submission Sheet
School of Computing

Student Name:	Ebubekir Ayhan
Student ID:	x23141727
Programme:	MSc in AI for Business
Year:	2024
Module:	MSc Research Project
Supervisor:	Victor Del Rosal
Submission Due Date:	16/09/2024
Project Title:	Enhancing Customer Churn Prediction in the Telecommunications Sector: Benchmarking Neural Networks Against Traditional Machine Learning Models
Word Count:	5186
Page Count:	16

I hereby certify that the information contained in this (my submission) is information pertaining to research I conducted for this project. All information other than my own contribution will be fully referenced and listed in the relevant bibliography section at the rear of the project.

ALL internet material must be referenced in the bibliography section. Students are required to use the Referencing Standard specified in the report template. To use other author's written or electronic work is illegal (plagiarism) and may result in disciplinary action.

Signature:	Ebubekir Ayhan
Date:	14th September 2024

PLEASE READ THE FOLLOWING INSTRUCTIONS AND CHECKLIST:

Attach a completed copy of this sheet to each project (including multiple copies).	☐
Attach a Moodle submission receipt of the online project submission, to each project (including multiple copies).	☐
You must ensure that you retain a HARD COPY of the project, both for your own reference and in case a project is lost or mislaid. It is not sufficient to keep a copy on computer.	☐

Assignments that are submitted to the Programme Coordinator office must be placed into the assignment box located outside the office.

Office Use Only	
Signature:	
Date:	
Penalty Applied (if applicable):	

Enhancing Customer Churn Prediction in the Telecommunications Sector: Benchmarking Neural Networks Against Traditional Machine Learning Models

Ebubekir Ayhan
x23141727

Abstract

Customer churn prediction is very crucial in the telecommunications industry due to the high cost of gaining new customers and the current competitive environment. This study aims to investigate how accurately newly developing neural network technologies can make predictions compared to old traditional methods and to examine the results. The study uses the CRISP-DM methodology to examine a Kaggle dataset of telecom customers. The dataset contains details about service usage and demographics that are essential for churn prediction. Using RapidMiner, extensive cleaning and transformation procedures are part of the data preprocessing process. Then, six machine learning models are implemented and benchmarked: Support Vector Machine (SVM), Naïve Bayes, Random Forest, k-nearest Neighbors (k-NN), Random Networks, and Deep Learning. The results show that SVMs and neural networks perform better than other models, with accuracies of 78.76% and 80.02%, respectively. These models demonstrate strong recall and precision metrics for both retained and churned customers, highlighting their efficacy in churn prediction tasks. Conversely, Naïve Bayes exhibits imbalanced recall rates and lower accuracy (44.31%), underscoring its limitations in this particular context. The results highlight how crucial it is to choose models that are suitable for the subtleties of telecom churn prediction. The industry will be affected by the use of precise churn forecasts to create proactive retention plans that will raise customer satisfaction and lower churn rates. Subsequent investigations may examine supplementary customer metrics and sophisticated modeling methodologies to enhance prognostic precision and practicality in actual telecom environments.

1 Introduction

1.1 Background and Motivation

The telecommunications industry has become a huge market with many services such as networks, internet infrastructure, film, TV, and music. With the internet and the developing network infrastructure, almost all adults and even children use mobile devices to make calls or access the internet in their daily lives. Competition in this growing market is also fierce. Since there are so many active, rival service providers, the telecommunications industry faces difficult obstacles recentlyUmayaparvathi and Iyakutti (2016). Many companies are trying to win customers using different strategies with various campaigns and packages. The reason that it is far more expensive to acquire new customers than to retain current ones, is that churn prediction is also crucial for companies. With billions of telecom users, even a tiny percentage of churn results in a significant loss of revenueDalvi et al. (2016). According to a study, it is six times more expensive to acquire a new customer than it is to keep a churn-prone existing oneAhmad et al. (2019). Therefore, it has become very important for telecommunication companies to retain their existing customers. Customer Churn can be defined as the process of consumers moving anonymously from one service provider to anotherUmayaparvathi and Iyakutti (2016). In this context, the telecommunications sector is one of the sectors most affected by customer churn due to its repetitive market. Customer dissatisfaction, increased costs, poor quality, a lack of features, and privacy concerns are some of the reasons why customers leave businessesSharma and Rajan (2017). In order to prevent customer churn and regain existing customers, companies implement various strategies and conduct various campaigns for the continuation of their subscriptions by contacting their customers before their contract expires. But even more important is to identify which customers are closer to churn and contact the customer before the customer churns and win the customer back with various strategies. In this way, customer loyalty will increase and the number of customers will be increased at a lower cost. Therefore, customer churn prediction is of great importance for telecommunication companies.

Before artificial intelligence and machine learning, companies tried to identify these customers with various data analyses and various rule-based systems. While customer churn was predicted using data mining techniques, the model did not perform well. On the other hand, because of the difficulty in handling them, the massively large data sources were disregarded Ahmad et al. (2019). With machine learning methods, it has become possible to detect them with higher accuracy rates. With the development of machine learning methods, various traditional machine learning methods have been developed. Random forests, decision trees, and regressions are just a few of them. However, these classical learning methods were incomplete in understanding some complex data and the relationship between that data and transferring it to modeling. Therefore, more advanced model machine learning techniques were needed.

In today's world, technology is crucial to meeting customer demands and reaching a certain level of satisfaction. It is imperative for the business to create a productive and effective model for handling customer churn with modern technologyFujo et al. (2022). Neural networks are more prominent than other models as a machine learning method inspired by neurons in the human brain. With this method, it has become possible to detect more complex flows and relationships. Customer churn prediction is one of the areas where neural networks can be used. With this method, it is possible to better analyze customer behavior and movements and make more accurate predictions. This study aims to benchmark using a neural network model and other models in customer churn prediction. With this study, customer churn prediction modeling with both classical machine learning methods and neural network method and comparative analysis will be performed.

1.2 Research Question and Objective

By comparing neural network performance to conventional machine learning models, this study seeks to improve customer churn prediction in the telecom industry. The following are the main goals of this study:

-Analyze the performance of several conventional machine learning models in customer churn prediction. This covers models like naive bayes, random forests, and support vector machines.

-In the context of churn prediction, evaluate the effectiveness of neural networks, encompassing both shallow and more intricate architectures like deep and recurrent neural networks.

-The best strategy for churn prediction in the telecom industry can be found by contrasting and comparing these models' predictive abilities.

1.3 Research Question

How can customer churn prediction in the telecommunications industry be improved by benchmarking neural networks against traditional machine learning models, using a dataset that contains comprehensive details about monthly fees, additional fees, subscription services, data usage, contract type, and customer status (joined, stayed, or churned) from a Californian telecommunication company in Q2 2022?

1.4 Structure of the Paper

This is how this report is organized:

- Section 2: This section offers a critical analysis of earlier research and its conclusions, reviewing relevant work in the area of churn prediction.
- Section 3: This section describes the research methodology used in this study, including how the features were chosen, how the data was collected, and the evaluation metrics that were used to compare the models.
- Section 4: The architecture of the neural networks and the parameters of the conventional machine learning models are among the design specifications of the models that are covered in this section.
- Section 5: The implementation process, along with the tools and programming languages utilized, is covered in this section.
- Section 6: The models' evaluation is presented in this section, with the results displayed through a variety of visual aids, including charts and tables and a detailed discussion of the findings has been made.
- Section 7: The conclusion and prospects for additional study on the subject are covered in this section.

2 Related Work

In the telecommunications industry, the task of predicting customer churn has received a lot of attention because it has a significant impact on customer retention strategies and overall profitability. Over the years, many strategies—from cutting-edge machine learning techniques to traditional statistical methods—have been put forth and put into practice.

2.1 Traditional Machine Learning Models

To improve the predictive accuracy of churn models, traditional machine learning techniques have been the focus of efforts in customer churn prediction. The study conducted by Dalvi et al. uses two popular, established techniques —logistic regression and decision trees— to forecast customer churn in the telecom sector. These techniques are prized for being easy to understand and straightforward, which makes them useful instruments for spotting trends in customer churn and supporting well-thought-out customer retention campaigns. By predicting their propensity to switch to competitors, shows how logistic regression and decision trees can be used to analyze customer data and derive insights that help retain customers. Lastly, they recommend that users use our comparative analysis to determine which algorithm best fits their data in order to save time when searching for an accurate modelDalvi et al. (2016).

Due to their similar characteristics, traditional machine learning techniques for customer churn prediction frequently have trouble differentiating between churn and non-churn customers. A novel method incorporating a certainty estimation mechanism based on the distance factor within the dataset is proposed by Amin et al. With this method, churn predictions are made more accurately by dividing the data into zones of high and low certainty. The study shows that classifier certainty, especially in high-certainty zones, can greatly improve prediction accuracy. Utilizing conventional assessment metrics like accuracy, F-measure, precision, and recall, this approach offers a strong foundation for enhancing churn prediction in the telecom sectorAmin et al. (2019).

Traditional machine learning techniques continue to be essential in the effort to forecast customer attrition and keep hold of current clients. A new framework for churn prediction is proposed by Dahiya and Bhatia and is implemented with WEKA Data Mining software. The use of logistic regression and decision tree techniques is the main focus of their research. The authors hope to accurately predict customer churn by utilizing these conventional techniques, and they also hope to offer telecom companies useful insights to improve customer retention. Their study highlights the interpretability and ease of implementation of decision trees and logistic regression as major benefits, underscoring their ongoing relevance in the field of customer churn predictionDahiya and Bhatia (2015).

Conventional regression-based models have demonstrated considerable promise in the endeavor to forecast customer attrition in the telecommunications industry. A machine learning framework based on regression was created by Ele et al. to predict customer churn. Nine different regression models were examined in their study: support vector regression, random forest regression, ridge regression, stochastic gradient descent regression, polynomial regression, Lasso regression, elastic net regression, and robust regression. Lasso Regression outperformed the others in terms of performance metrics, proving the usefulness of conventional regression methods for churn prediction. Regression models are a good choice for telecom businesses looking to anticipate and reduce customer attrition because of their effectiveness and interpretability, as demonstrated by this studyEle et al. (2023).

2.2 Neural Networks

Thanks to deep learning techniques, especially neural networks, machine learning models can handle complex data patterns, and neural networks have become more and more popular in the field of customer churn prediction. A Deep Backpropagation Artificial Neural Network (Deep-BP-ANN) was deployed by Fujo et al. to forecast customer attrition in the telecom sector. The model included an early stopping strategy to avoid overfitting in addition to feature selection techniques like Variance Thresholding and Lasso Regression. The study illustrated the superior performance of the neural network model in reducing overfitting and improving predictive accuracy by contrasting dropout and activity regularization strategies. The outcomes demonstrated that the Deep-BP-ANN model significantly outperformed conventional machine learning methods, particularly when enhanced with Lasso Regression for feature selection and activity regularizationFujo et al. (2022).

Because of its capacity to represent intricate patterns in data, neural networks have gained popularity in the field of predicting customer churn. In their comparative analysis of prediction techniques, Huang et al. included multilayer perceptron neural networks. According to the study, churn prediction performance can be greatly enhanced by combining neural networks with a strong feature set. This demonstrates neural networks' capacity to manage the complex relationships found in customer data, providing telecom companies with an effective tool to predict customer churn and pursue focused retention campaignsHuang et al. (2012).

Given their propensity to represent intricate and non-linear relationships in data, neural networks are being used more and more to predict customer churn. The use of Back Propagation Neural Network (BPNN) and other models for early warning of customer churn in the telecommunications industry was investigated by Zhou et al. Data cleaning, oversampling, and standardization were among the preprocessing techniques used in their study on a sizable dataset consisting of 900,000 telecom customer records. Neural networks are robust when processing large and diverse datasets, as demonstrated by the evaluation of the BPNN model and other models for churn predictionZhou et al. (2023).

2.3 Comparative Studies

The effectiveness of various machine learning models for predicting customer attrition has been assessed in a number of comparison studies. A thorough comparison between the adaptive Naïve Bayes model by Amin et al. and a number of conventional and sophisticated machine learning models, such as Deep-BP-ANN, CNN, neural networks, linear regression, XGBoost, KNN, Logit Boost, SVM, and PCA-LB techniques, was carried out. The results of their study demonstrated that the adaptive learning model demonstrated notable enhancements in precision, recall, F1-score, MCC, and overall accuracy, in addition to outperforming numerous conventional models. The aforementioned comparison highlights the significance of adaptive methodologies in augmenting the accuracy of churn prediction and exhibits the possibility of amalgamating conventional models with evolutionary computation for superior outcomesAmin et al. (2023).

Chong et al. evaluated various machine learning algorithms to forecast customer churn within a telecommunications company. The six supervised learning algorithms that were the subject of the study were AdaBoost, XGBoost, Random Forest, SVM, K-Nearest Neighbors, and Logistic Regression. They found that XGBoost was the most successful classifier by comparing the performance of these models; it achieved 79.67% accuracy, 64.67% precision, 51.87% recall, and 57.57% F1-score. Along with highlighting the potential of various machine learning techniques in churn prediction, the study offers telecommunication companies practical insights that they can use to retain customers. These insights include personalized customer experiences, loyalty programs, and the integration of AI into customer relationship managementChong et al. (2023).

Umayaparvathi and Iyakutti conducted a thorough survey that offers a detailed analysis of customer churn prediction in the telecom sector with an emphasis on the datasets, techniques, and metrics employed in the literature. This study emphasizes the value of building successful churn prediction models and the relevance of data mining techniques in predicting customer attrition. Customer care service details, credit scores, billing and payment details, usage patterns, and value-added services are the six primary categories into which the authors divide customer data. Additionally, they examine a number of publicly available datasets, highlighting the variety of customer attributes found in each, such as the PAKDD 2006 Data Mining Competition Dataset and the ACM KDD Cup 2009 Orange Labs DatasetUmayaparvathi and Iyakutti (2016).

Sikri et al. address the problem of customer attrition in the telecom sector, stressing the value of keeping current clients over attracting new ones. To address data skewness in churn prediction, they suggest a novel method utilizing ratio-based data balancing. When compared to conventional over- and under-sampling techniques, this method improves predictive accuracy. Perceptron, Multi-Layer Perceptron, Naive Bayes, Logistic Regression, K-Nearest Neighbor, Decision Tree, and ensemble techniques like Gradient Boosting and XGBoost are just a few of the machine learning algorithms that are evaluated in this study. The findings suggest that hybrid approaches, especially when utilizing XGBoost at a 75:25 ratio, produce the most encouraging outcomes. The study emphasizes the significance of data balancing strategies for churn prediction model optimizationSikri et al. (2024).

2.4 Gaps and Justification for Current Study

Despite the extensive body of research, there remain gaps in the comprehensive evaluation of neural networks versus traditional machine learning models specifically tailored to the telecommunications sector. Previous studies often focus on a limited set of models or do not fully leverage the rich, detailed datasets available today. This research aims to address these gaps by conducting a thorough benchmarking of various models using a dataset that includes monthly fees, additional fees, subscription services, data usage, contract type, and customer status (joined, stayed, or churned) from a Californian telecommunication company in Q2 2022.

By providing a detailed comparative analysis, this study seeks to offer valuable insights into the most effective techniques for customer churn prediction, thereby aiding telecom operators in implementing more robust and accurate retention strategies.

3 Methodology

In this study, CRISP-DM methodology is applied which means a Cross-Industry Standard Process for Data Mining and developed in 1999. It is stated that the technologies used in data mining are all somewhat interchangeableShafique and Qaiser (2014). There are 6 steps of the CRISP-DM. They are Business Understanding, Data Understanding, Data preparation, Modelling, Evaluation, and Deployment. A well designed CRISP-DM steps are shown in Figure 1. After understanding the business, data collection and understanding the data is required.

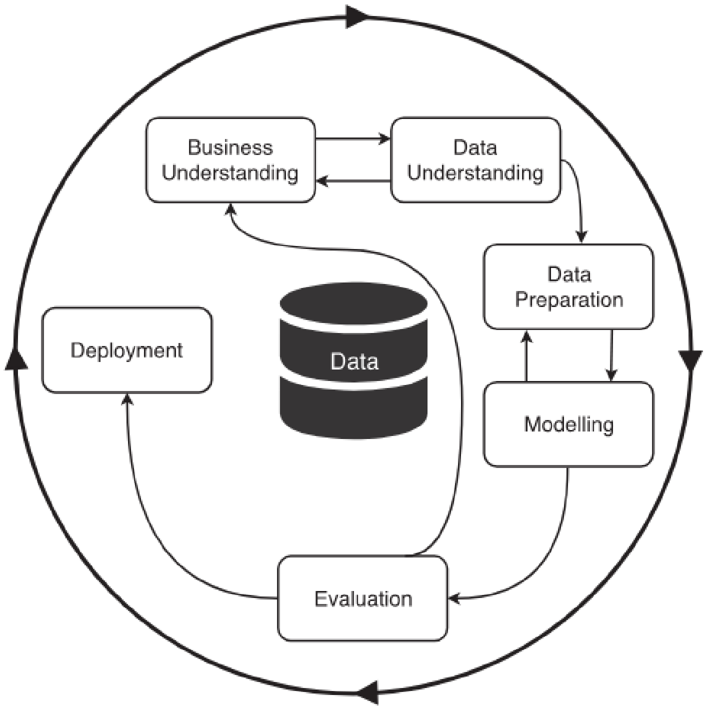

Figure 1: CRISP-DM Steps (Martínez-Plumed et al. (2019))

3.1 Data Collection

Since telecommunication sector customer churn data contains highly private information, it is the type of data that companies do not want to share. Therefore, it is not possible to request this data directly from the companies. The dataset used in this study is publicly shared on Kaggle and is completely available[1]. The current dataset contains demographic information such as the customer's gender, age, marriage status, city of residence, as well as various data that will help with customer churn prediction, such as the contract type, monthly payment amount, and extra services used. The data set consists of 38 columns and 7043 rows of customer data. At the end, there is information about whether the customer churned or stayed. When exploratory data analysis (EDA) was made, as seen in Figure 2, the customers' gender was almost equally distributed. Nearly all existing customers also use the phone service. When we look at the graphs in Figure 3, it is

[1] `https://www.kaggle.com/datasets/shilongzhuang/telecom-customer-churn-by-maven-analytics/data`

possible to say that almost all customers use the internet service. While nearly half of the customers have made a subscription in the form of monthly payments, it is observed that other customers have an annual (1-year or 2-year) contract. Finally, Figure 4 shows the churn status of the customers. Among these customers, those who have joined will not be included in the model as they have just joined the company, and this will negatively affect the accuracy of the model.

3.2 Data Preparation

After understanding business and data, data preparation is the other step that needs to be followed. Cleaning is one of the parts of the preparation. First of all, the columns that don't affect the model like customer ID, latitude, and zip code need to be removed. Then customer status needs to be filtered for the joined customers which they shouldn't be in the model. After doing this, replacing missing values is another step of the cleaning. All the missing values need to be handled. Adding an average answer or average number to missing values is an option to handle those. After finishing cleaning, the data transformation part is the other part of the data preparation. To apply a machine learning model, all the variables except the variable that needs to be predicted (customer status) need to turn into numerical values. To see more details of these cleaning steps, please check the configuration manual document that is attached to this study.

3.3 Modelling

Modeling is the next step after finishing data preparation. Before applying the model, a 70-30 split was used to divide the dataset into training and testing sets. This guarantees that the performance of the model can be assessed on unobserved data. Then model can be applied and the model that is built can be run. In this study, the Rapidminer tool is used to apply data cleaning and modeling. Since this study is benchmarking different machine learning models, the models need to be changed and run again every time for neural networks and other traditional machine learning models. Again, there are more details regarding this step in the configuration manual document which is attached. The models that are applied to the dataset are as below.

- Neural Network
- Deep Learning
- K Nearest Neighbour (k-NN)
- Naive Bayes
- Random Forest
- Support Vector Machine (SVM)

3.4 Evaluation

After completing the modeling, it is necessary to make an evaluation. After running the model, Rapidminer measures the performance of the model with various metrics. The evaluation metrics listed below were used to compare the models' performances. These metrics are accuracy, recall, and precision.

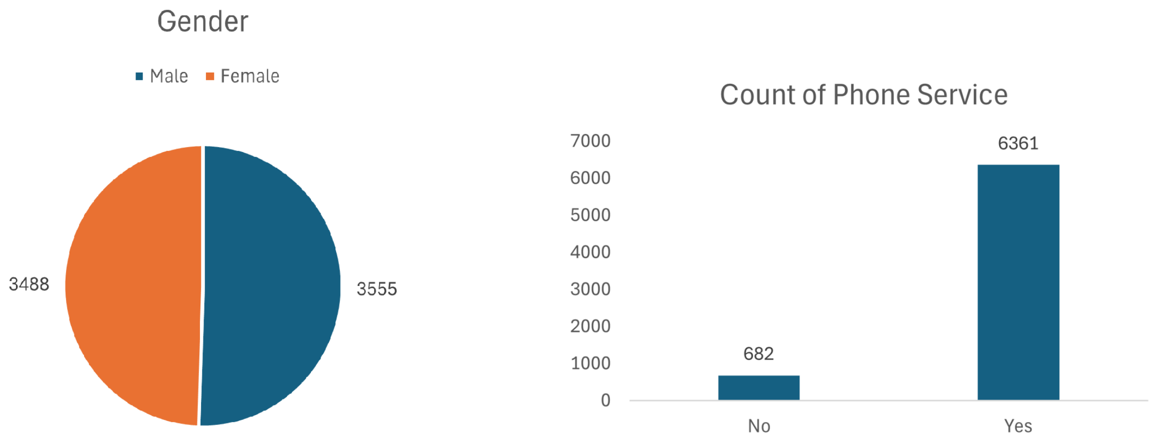

Figure 2: EDA 1

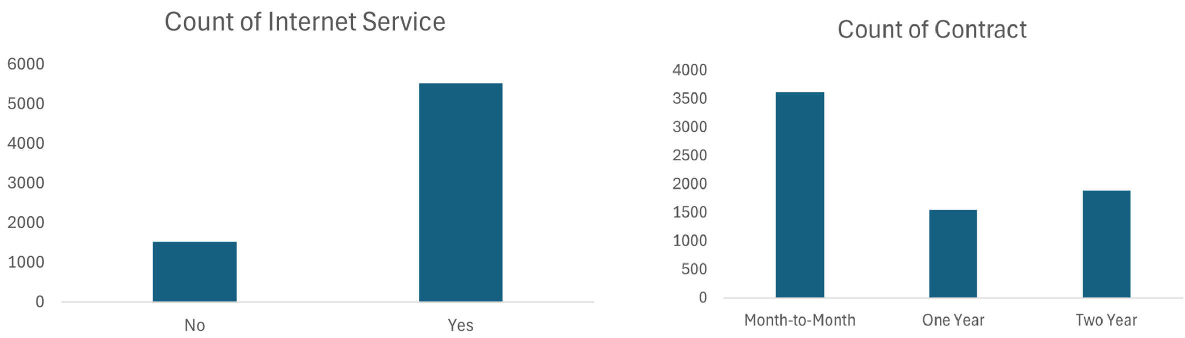

Figure 3: EDA 2

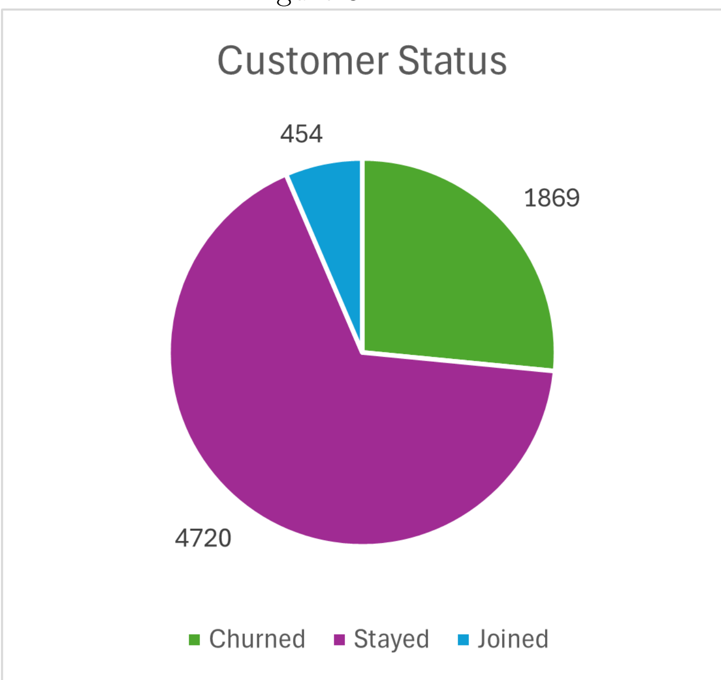

Figure 4: Customer Status

- **Accuracy:** calculates the percentage of correctly predicted cases out of all the instances.
- **Precision:** shows the percentage of all positive predictions that are actually positive.
- **Recall:** evaluates how well the model can represent every pertinent instance of the positive class.

4 Design Specification

The system analyzes customer data and finds possible churners by using machine learning models—both conventional models and neural networks—that are integrated into RapidMiner. Preparing the data, training the models, evaluating the system, and deploying it are its main parts. Figure 5 shows the data flow diagram of the study. This project complies with data science and machine learning best practices by adhering to this structured data flow, guaranteeing that every stage advances the main objective of precisely forecasting customer churn and offering insightful business information.

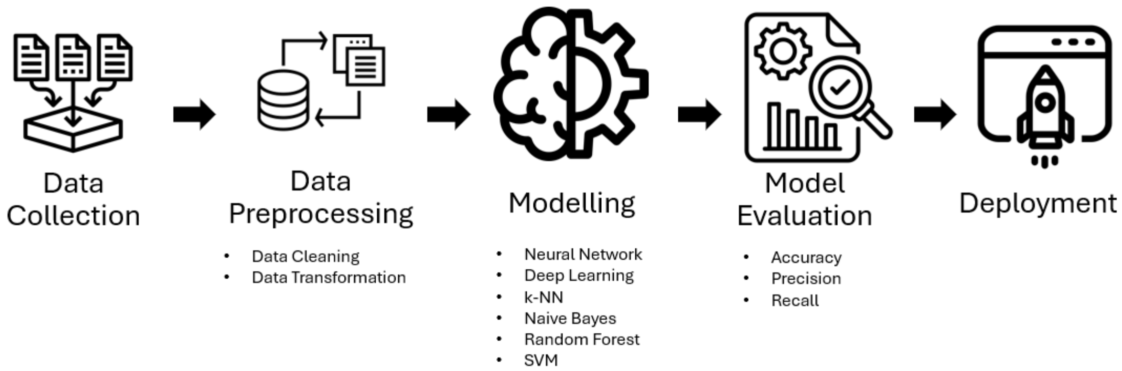

Figure 5: Data Flow Diagram

5 Implementation

In this study, the Rapidminer machine-learning software tool is used for modeling. First, the customer churn data downloaded from Kaggle was uploaded to the software. Then we moved on to data preprocessing steps. As part of the first step, data cleaning, variables such as churn category, churn reason, customer ID, latitude, longitude, and zip code, which did not affect the model or would negatively affect the model's accuracy, were removed. Later, customer status was changed from a polynomial value to a binomial because the main data consisted of customer status churned, stayed, and joined values, and since joined customers were still new, it was not appropriate to include them in churn modeling. Therefore, joined customers were filtered and cleaned from the data set.

Data cleaning continued by filling in missing values. We can eliminate missing values by filling in the missing numerical values such as average monthly GB download and average monthly long-distance charges with the average of the current values. Data cleaning continued by filling in missing values. We can eliminate missing values by filling in the missing numerical values such as average monthly GB download and average monthly long-distance charges with the average of the current values. Apart from these, missing values that only have a yes/no option can also be filled in (no), assuming that the feature does not exist. For example, the Device Protection Plan is one of them. While there are 2390 yes (customers with this plan) there are 2855 no. This data is missing for 1344 customers. Therefore, it can be assumed that these customers did not purchase this plan either. With this method, all missing values of variables such as device protection plan, multiple lines, online backup, online security, premium tech support, streaming movies, streaming music, streaming TV, and unlimited data were filled in as no. Finally, the missing values of the variable that has cable, fiber optic, and DSL internet type options must be filled in. For this purpose, missing values were filled as "none". Figure 6 shows the steps of the Rapidminer Modelling. The first step is uploading the dataset. The 4 steps after the first step consist of the data cleaning steps so far. After that, you need to decide what you are predicting with the set role process. In the dataset, "customer status" is the variable that we want to predict (churned/stayed).

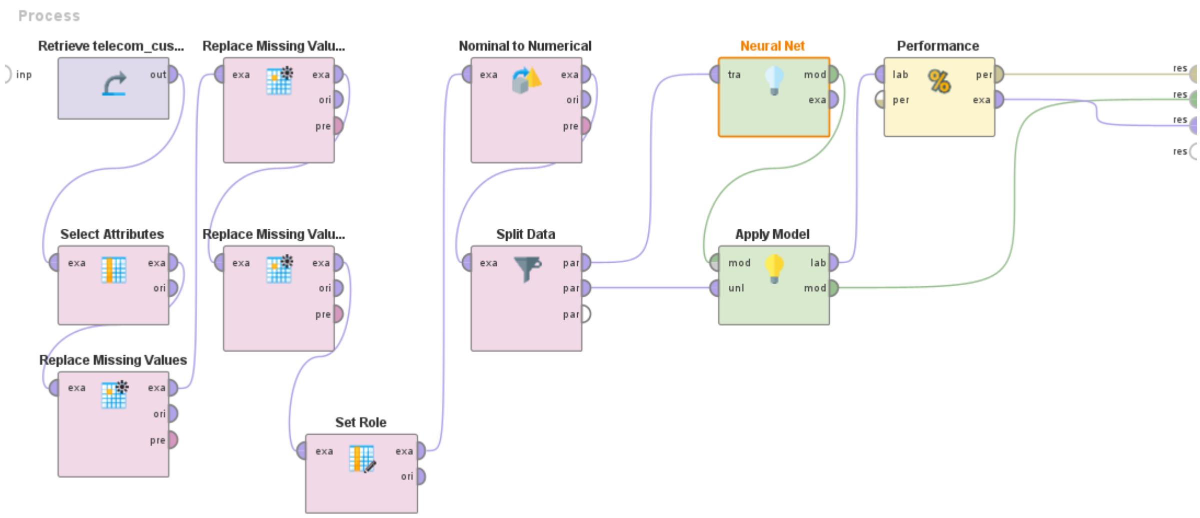

Figure 6: Rapidminer Modelling

After data cleaning is completed, all of the data (except the target variable) must be converted to numerical values, otherwise machine learning models will not understand anything from the texts. Doing this is quite simple with Rapidminer. All data is digitized with the "nominal to numerical" process.After all the data is numeric, the dataset is ready for modeling. 70% of the data can be reserved for training and 30% for testing. Split data processing is used for this.

After all this is completed, some of the split data is transferred to the relevant machine learning model, and the remaining 30% is transferred to the applied model process for testing. The neural network is used in this example. Implementation of other models works the same way. Only the machine learning model process needs to change. Finally, the completed model can be run.

6 Evaluation

The performance of different machine learning models in predicting customer churn in the telecom industry is critically examined in this section. In this analysis, classic machine learning algorithms like K-Nearest Neighbors (KNN), Naive Bayes, Random Forest, and Support Vector Machines (SVM) are benchmarked against neural networks and deep learning.

6.1 Neural Networks

The first applied model is the neural network which has **80.02%** accuracy.This outcome is consistent with the research conducted by Sikri et al., who also observed that neural networks are useful for managing intricate, non-linear relationships in churn prediction data.

Table 1: Performance Table of Neural Network

	True Stayed	**True Churned**	**Class Precision**
Predicted Stayed	1411	390	78.35%
Predicted Churned	5	171	97.16%
Class Recall	99.65%	30.48%	

6.2 Deep Learning

The second applied model to the dataset is deep learning which has **77.04%** accuracy. This is consistent with research by Huang et al., which emphasizes deep learning's ability to represent complex data structures and dependencies.

Table 2: Performance Table of Deep Learning

	True Stayed	**True Churned**	**Class Precision**
Predicted Stayed	1153	191	85.79%
Predicted Churned	263	370	58.45%
Class Recall	81.43%	65.95%	

6.3 k-NN

The third applied model is k-NN (k nearest neighbor) which has **78.81%** accuracy. This model's trade-off between recall and precision is consistent with findings in the literature, indicating that while k-NN is capable of efficiently classifying instances, further tuning may be necessary to increase recall rates.

Table 3: Performance Table of k-NN

	True Stayed	**True Churned**	**Class Precision**
Predicted Stayed	1294	297	81.33%
Predicted Churned	122	264	68.39%
Class Recall	91.38%	47.06%	

6.4 Naïve Bayes

The fourth applied model is Naïve Bayes which has **44.31%** accuracy.

Table 4: Performance Table of Naive Bayes

	True Stayed	**True Churned**	**Class Precision**
Predicted Stayed	468	153	75.36%
Predicted Churned	948	408	30.09%
Class Recall	33.05%	72.73%	

6.5 Random Forest

The fifth applied model is Random Forest which has **71.83%** accuracy.

Table 5: Performance Table of Random Forest

	True Stayed	**True Churned**	**Class Precision**
Predicted Stayed	1416	557	71.77%
Predicted Churned	0	4	100.00%
Class Recall	100.00%	0.71%	

6.6 Support Vector Machine (SVM)

The sixth applied model is Support Vector Machine (SVM) which has **78.76%** accuracy. SVM can function as a solid baseline model for churn prediction thanks to its balanced precision and recall, which guarantees consistent performance on a variety of datasets.

Table 6: Performance Table of SVM

	True Stayed	**True Churned**	**Class Precision**
Predicted Stayed	1117	121	90.23%
Predicted Churned	299	440	59.54%
Class Recall	78.88%	78.43%	

6.7 Discussion

Different approaches yield different results when evaluating six machine learning models for telecom industry customer churn prediction. Table 7 shows the comparison of the results of those models. With an accuracy of 80.02%, neural networks outperformed the other models. The neural network model showed strong performance in accurately identifying churn cases, even with its high accuracy. Its precision for predicting customers who actually churned was 97.16%. Its recall rate for churned customers was, however, lower at 30.48%, indicating possible difficulties in recording every instance of churn. Our findings about neural networks' superior performance align with those of Fujo et al. (2022), who also showed how well neural networks handled complex data patterns for predicting customer churn. This supports the idea that complex relationships found in customer data can be captured by neural networks in an efficient manner.

Deep learning, on the other hand, showed competitive performance with a precision of 85.79% for stayed customers, despite achieving a slightly lower accuracy of 77.04%. But when it came to churn prediction, its accuracy fell to 58.45%, indicating a disparity in correctly identifying customers who are likely to leave. The model's recall rates, which were 81.43% for customers who stayed and 65.95% for customers who left, showed that it was reasonably balanced in capturing examples of both classes, though there was still an opportunity for refinement.

With a 78.81% accuracy rate and balanced precision metrics—81.33% for remaining customers and 68.39% for churned customers—the k-Nearest Neighbors (k-NN) approach produced good results. Its recall rate for customers who had churned, however, was significantly lower at 47.06%, which may suggest that it is not possible to reliably identify every instance of churn based only on the approach of the closest neighbors.

Conversely, Naïve Bayes demonstrated the least accuracy of all the models at 44.31%, with a precision of 75.36% for customers who remained and a noticeably lower precision of 30.09% for customers who churned. Furthermore, there were significant difficulties in accurately predicting customer churn with this model, as evidenced by the unbalanced recall rates, which were 33.05% for customers who stayed and 72.73% for those who left.

In terms of precision, Random Forest performed perfectly (100.00%) for customers who remained, but only 71.77% for those who left. Its accuracy was 71.83%. Its recall rate for customers who left, however, was a pitiful 0.71%, indicating that there may have been an overfitting bias in favor of predicted retention rates rather than churn rates.

Finally, the Support Vector Machine (SVM) model demonstrated a 78.76% accuracy rate, exhibiting robust precision metrics of 90.23% for remaining customers and 59.54% for customers who churned. With balanced recall rates of 78.88% for remaining customers and 78.43% for churning customers, the model demonstrated a strong ability to distinguish between the two classes. Our study's observed accuracy rates for SVM models closely match those of Huang et al. (2012), demonstrating the validity of SVMs for categorizing customer churn data. This validation demonstrates the validity of SVMs as a workable method for telecom industry churn prediction.

Table 7: Comparison of Machine Learning Models for Customer Churn Prediction

Model	**Accuracy (%)**	**Precision**	**Recall**	**Strengths**	**Weaknesses**
Neural Networks	80.02	High	High	Captures non-linear patterns well	Complex to train and tune
Deep Learning	77.04	Moderate	High	Models complex structures	Requires large datasets
k-NN	78.81	Moderate	Moderate	Effective with simple implementation	May need tuning for recall
Naïve Bayes	44.31	Low	High	High recall for churned customers	Assumes feature independence
Random Forest	71.83	High	Moderate	Handles high-dimensional data	Complexity increases with features
SVM	78.76	High	High	Strong generalization capability	Computationally intensive

To sum up, certain models, like SVMs and neural networks, performed well in terms of accuracy and balanced precision-recall metrics, but other models, like Random Forest and Naïve Bayes, had a difficult time accurately predicting churn, especially when recall rates were uneven. The significance of selecting an appropriate model according to particular business goals and dataset attributes is highlighted by these results. Additionally, they highlight the possibility of enhancing the model further via feature engineering, model tuning, and ensemble approaches in subsequent studies and applications.

7 Conclusion and Future Work

Assessing how well different machine learning models performed in predicting customer churn in the context of the telecom sector was the main goal of this thesis. In this particular customer churn prediction task, our specific goal was to evaluate the performance of neural networks compared to other conventional machine learning models. All things considered, the investigation has effectively assessed and contrasted six different machine-learning techniques. With accuracy rates of 80.02% and 78.76%, respectively, neural networks and SVMs were the best at predicting both stayed and churned customers. They also showed strong precision and recall metrics. Conversely, models such as Naïve Bayes showed difficulties in accurately predicting customer churn, with lower accuracy (44.31%) and imbalanced recall rates.

The results highlight how crucial it is to choose machine learning models that are suitable for the particular predictive task of customer churn. Because of their balanced performance metrics, neural networks and SVMs have important implications for telecom companies looking to improve their customer retention strategies. These models offer insightful information about customer behavior that can guide proactive retention campaigns, which may lower churn and raise customer satisfaction.

In the future, further customer features and behavioral metrics, like usage patterns, customer service interactions, and satisfaction surveys, can be investigated to improve predictive accuracy and potentially better capture churn patterns. Additionally, to increase overall predictive performance and robustness, experiments can be made with ensemble methods or hybrid models, which combine the advantages of multiple algorithms. Creating more complex predictive models by examining the effects of external factors (such as the competitive environment and the state of the economy) on churn behavior will also help to improve accuracy.

In summary, this thesis has improved our knowledge of machine learning applications for telecom industry customer churn prediction. We have determined the advantages, disadvantages, and potential areas for advancement in predictive accuracy and useful implementation by assessing and contrasting six distinct models. The knowledge gathered from this study paves the way for additional research and useful applications that will lower customer churn and improve business outcomes in the telecom industry and elsewhere.

References

Ahmad, A. K., Jafar, A. and Aljoumaa, K. (2019). Customer churn prediction in telecom using machine learning in big data platform, *Journal of Big Data* **6**(1): 1–24.

Amin, A., Adnan, A. and Anwar, S. (2023). An adaptive learning approach for customer churn prediction in the telecommunication industry using evolutionary computation and naïve bayes, *Applied Soft Computing* **137**: 110103.

Amin, A., Al-Obeidat, F., Shah, B., Adnan, A., Loo, J. and Anwar, S. (2019). Customer churn prediction in telecommunication industry using data certainty, *Journal of Business Research* **94**: 290–301.

Chong, A. Y. W., Khaw, K. W., Yeong, W. C. and Chuah, W. X. (2023). Customer churn prediction of telecom company using machine learning algorithms, *Journal of Soft Computing and Data Mining* **4**(2): 1–22.

Dahiya, K. and Bhatia, S. (2015). Customer churn analysis in telecom industry, *2015 4th International Conference on Reliability, Infocom Technologies and Optimization (ICRITO)(Trends and Future Directions)*, IEEE, pp. 1–6.

Dalvi, P. K., Khandge, S. K., Deomore, A., Bankar, A. and Kanade, V. (2016). Analysis of customer churn prediction in telecom industry using decision trees and logistic regression, *2016 symposium on colossal data analysis and networking (CDAN)*, IEEE, pp. 1–4.

Ele, S. I., Alo, U. R., Nweke, H. F. and Ofem, O. A. (2023). Regression-based machine learning framework for customer churn prediction in telecommunication industry, *Journal of Advances in Information Technology* **14**(5): 1046–1055.

Fujo, S. W., Subramanian, S., Khder, M. A. et al. (2022). Customer churn prediction in telecommunication industry using deep learning, *Information Sciences Letters* **11**(1): 24.

Huang, B., Kechadi, M. T. and Buckley, B. (2012). Customer churn prediction in telecommunications, *Expert Systems with Applications* **39**(1): 1414–1425.

Martínez-Plumed, F., Contreras-Ochando, L., Ferri, C., Hernández-Orallo, J., Kull, M., Lachiche, N., Ramírez-Quintana, M. J. and Flach, P. (2019). Crisp-dm twenty years later: From data mining processes to data science trajectories, *IEEE transactions on knowledge and data engineering* **33**(8): 3048–3061.

Shafique, U. and Qaiser, H. (2014). A comparative study of data mining process models (kdd, crisp-dm and semma), *International Journal of Innovation and Scientific Research* **12**(1): 217–222.

Sharma, R. and Rajan, S. (2017). Evaluating prediction of customer churn behavior based on artificial bee colony algorithm, *International Journal of Engineering and Computer Science* **6**(1): 20017–20021.

Sikri, A., Jameel, R., Idrees, S. M. and Kaur, H. (2024). Enhancing customer retention in telecom industry with machine learning driven churn prediction, *Scientific Reports* **14**(1): 13097.

Umayaparvathi, V. and Iyakutti, K. (2016). A survey on customer churn prediction in telecom industry: Datasets, methods and metrics, *International Research Journal of Engineering and Technology (IRJET)* **3**(04).

Zhou, Y., Chen, W., Sun, X. and Yang, D. (2023). Early warning of telecom enterprise customer churn based on ensemble learning, *Plos one* **18**(10): e0292466.