Journal of Computing & Biomedical Informatics
ISSN: 2710 - 1606
Volume 07 Issue 01
2024
Research Article
https://doi.org/10.56979/701/2024

An Analysis of Supervised Machine Learning Techniques for Churn Forecasting and Component Identification in the Telecom Sector

Hira Farman[1*], Abdul Wahab Khan[2], Saad Ahmed[3], Dodo Khan[4], Muhammad Imran[5], and Priyanka Bajaj[6]

[1]Department of Computer Science, Iqra University, Karachi, Pakistan.
[2]Department of Computer Science, Mohammad Ali Jinnah University, Karachi, Pakistan.
[3]Iqra University, Karachi, Pakistan.
[4]Department of CS &IT, TIEST, NED University of Engineering & Technology, Karachi, Pakistan.
[5]Department of Metallurgical Engineering, NED University of Engineering & Technology, Karachi, Pakistan.
[6]Department of Business administration, Salim Habib University, Karachi, Pakistan.
[*]Corresponding Author: Hira Farman. Email: hira.farman@iqra.edu.pk

Received: April 01, 2024 **Accepted:** May 29, 2024 **Published:** June 01, 2024

Abstract: The term "business intelligence" (BI) refers to a broad range of tools and software intended to collect, process, and evaluate data so that business users may decide on the best course of action. Getting the right information to the right decision-makers at the right time is the main goal of business intelligence (BI). The telecom business creates massive amounts of data every day because of its big clientele. Decision-makers and business professionals emphasized that maintaining existing clientele is less expensive than recruiting new ones. In addition to detecting patterns of behavior from the data on existing attrition clients, business analysts and customer relationship management (CRM) analysts need to understand the reasons for customer attrition. This paper focuses on customer churn, a critical metric that represents the percentage of customers ending their relationship with a company over a specific period. Using detailed datasets and advanced data analysis and machine learning techniques, the key churn predictors were found in this study, along with practical recommendations on how to improve customer retention. In order to create predictive customer churn models, several important machine learning algorithms have been surveyed and compared in this study. This study looks at more than just churn and non-churn classification; it also evaluates the accuracy of different data mining techniques. Uses a variety of performance indicators and confusion matrices to assess the effectiveness of three classification models: Random Forest (RF), Decision Tree (DT), and Logistic Regression (LR). With the best AUC (0.985), F1 score (0.934), Precision (0.935), and MCC (0.830), the Random Forest model outperformed the others, demonstrating a strong balance between recall and precision. The Decision Tree model also performed well, with notable accuracy). Logistic Regression, while effective, showed comparatively lower metrics, with an AUC of 0.848 and an F1 score of 0.800. The confusion matrices further validated these results, highlighting the Random Forest model's robustness and superior classification capabilities. The findings show that with the RF algorithm, our suggested churn prediction model generated superior churn categorization. Furthermore, this research delves into the fundamentals of BI and presents optimization strategies crucial for making dynamic, optimal decisions in today's corporate landscape.

Keywords: Business Intelligence (BI); Churn; Data Mining; Telecom.

1. Introduction

In the modern world, telecom businesses are producing enormous amounts of data at a very rapid pace. Numerous telecom service providers are in competition with one another to gain share of the market.

Consumers can choose from a variety of alternatives, including more affordable and superior services. Maximizing profits and surviving in a cutthroat market is telecom firms' ultimate goal. When a large portion of customers are dissatisfied with any telecom company's services, this is known as customer churn. Customers begin to migrate to different suppliers of services as a result of it.

Business Intelligence (BI) is essential for modern businesses aiming to stay competitive in a rapidly evolving market. BI tools and applications facilitate the collection, accessibility, and analysis of data, supporting informed decision-making processes. A key application of BI is in understanding and mitigating customer churn. Customer churn is a critical business metric indicating the percentage of customers who terminate their relationship with a company within a given timeframe. High churn rates can significantly impact a company's profitability and growth. This paper explores the factors influencing churn and retention through comprehensive data analysis and machine learning techniques, offering actionable insights to improve customer retention strategies. A range of software programs are utilized to evaluate an organization's raw data in order to make wise decisions that will lead to commercial success. These programs are collectively referred to as business intelligence, or BI. Data mining, online analytical processing, querying, and reporting are some of the associated tasks that make up business intelligence (BI) as a field. Multidimensional studies, mathematical projection, modeling, ad hoc inquiries, and "canned" reports are some of the techniques used. Business Intelligence is the capacity of an organization to use its data in a useful way. Competitive, consumer, industry, commodity, and operational intelligence are just a few of the numerous fields that business intelligence (BI) encompasses. According to earlier research, the term "business intelligence" (BI) in this work refers to a collection of ideas and techniques based on fact-based support systems for enhancing decision making [1-5].Over the past 15 to 20 years, the way businesses have utilized technology and data has changed considerably. This is all thanks to the advancements in data analysis that has introduced the term business analytics. Business analytics is a broad term used to describe the various ways in which business data can be extracted, stored, and analyzed in order to uncover incredible consumer insights in such depth that was never possible before. One such jewel under the umbrella of business analytics is Business Intelligence (BI). BI refers to the tools and techniques used to gather, store, access, and analyze data about a company in order to make better informed business decisions. It works to reveal patterns, trends, or potential issues. The field of business BI has undergone a remarkable evolution over the years, shaping the way organizations make use of data. This drive through the historical perspective of business Intelligence provides valuable insights into its revolutionizing path. The concept of BI was born in the 1960s, focusing primarily on using technology to support and execute transactional processing. According to the University of Georgia, Michael Scott Morton's work in 1967 can be attributed with the generation and advancement of management decision systems in the management processes to aid the managers in making informed decisions based on facts. To deal with it, the purpose of this study is to present an elaborated synthesis of the investigated field, including the discussed and the underdeveloped research domains, as well as to outline the interaction between the BI process's stages and its organizational environment.

The term Business Intelligence itself was first coined and promoted by the former Gartner analyst, Howard Dresner in 1989, who described BI as 'a broad category of applications, technologies, and processes for gathering, storing, accessing, and analyzing data to help business users make better decisions.BI doesn't just refer to the tools and applications used to get the required data out, it is used to get the data in (to a mart or warehouse such as modern day electronic inventory management systems).In the early of BI, decision makers used to have a hard time utilizing BI as it wasn't so mainstream and user friendly as now. Managers had to contact BI vendors and a BI trained IT team to get what they required. These were quite expensive and time consuming. Hence the usage of BI was comparatively low. The following decades saw a shift in the organization's view and behavior towards BI as the technology continued to improve. The organizations recognized the importance of extracting actionable insights from their growing volume of data. Within the last 5 years, the use of business Intelligence and the tools available have immensely changed. There has been a move towards creating business intelligence software that provides simplicity and ease of use for staff members across a company. 'Data Discovery', a fairly recent innovation, is a term used to describe the act of discovering opportunities that are hidden inside the company data. Data discovery not only uncovers hidden data, but also allows data to be presented through various grids, visualization tools and dashboards to actually see your data other than in just number form.

Other significant benefits of implementing BI are the cost reductions, as the software is much cheaper now. Improved decision making, focus and time efficiency are all due to BI. Business Intelligence has grown into a multidisciplinary subject that encompasses data warehousing, data mining, reporting, and analytics.

Churn prediction is vital in the telecom industry because managing the customer's lifetime value association is vital to telecom operators, and they must retain their valued clients [24, 25]. According to the literature CRM, the toughest job is to retain the existing customers over the new ones [23-26]. They demarcate the factors that may lead to a client or a group of customers to leave. In literary context, the people who left are referred to as dissatisfied customers albeit without regard of the factors and conditions that led to their decision. Since churners are not one and same in their behaviors, there is no need to manage all of them in a similar manner. It is also evident that would certain clients are more likely to leave a given business than others are. Such a prediction model is needed to help ascertain consumers who are likely to churn and the retention strategies that can be used which specifically includes promotional strategies that are customized to meet definite categories among the churn variables. Churn prediction model is presented in this work, which incorporates numerous types of machine learning methods. It means that the quality of the classifier depends on the dataset that is available at the given instance of time. Subsequently, we examine the proposed churn prediction model using information retrieval indices. From the TP rate, FP rate, precision, recall, F-measure, as well as the ROC area, the validity of a churn prediction model is measured.

Exploring the existing machine learning and data mining methodologies, detailing a model for customer churn predictions, finding out factors that cause customer attrition, and presenting measures of retaining customers are the objectives of the study. The results of the experiments have shown that the proposed model is superior to the other models in the aspect of correctly identifying churners. In this work, the churn prediction model that has been proposed and found to be of good accuracy is a contribution. Feature selection methods include information gain and correlation attribute ranking filter to select the most important characteristics. Used several machine learning approaches for both churn and non-churn analysis on big datasets of telecom industry and they reported that Random forest algorithm is superior to other methods in terms of accuracy.

1.1. Business Intelligence

Business Intelligence is an umbrella term referring to a set of concepts that involve using technology to gather, process, and report information to business executives and other end-users in a organization with a view to facilitating optimum decision making. BI can be characterized as a set of tools, applications, and methodologies allowing organizations to gather data from internal and external information systems, prepare it for analytical processing, create and run queries on the data, and present the results of analysis to corporate executives and managers in the form of reports, dashboards, and data visualizations.

1.2. Customer Churn

Customer churn, or customer attrition, is the loss of clients or customers. In the context of contractual services, churn occurs when a subscriber stops paying for the service. Churn rate is an important indicator of customer satisfaction and loyalty, and it is used as a key performance indicator in many businesses, especially those in the service sector.

1.3. Predictive Analytics in Churn Prediction

Predictive analysis is the process of identifying, measuring, forecasting and comparing patterns that are thought to represent future trends based on past and present data. As for customer churn, many factors that cause churn can be discovered by predictive analytics to guide enterprises to maintain potential lost customers.

1.4. Objective of Study

1. To take a deeper look at the process of data analysis, by utilizing Business Intelligence tools and techniques in order to enhance business decision making.

This objective focuses on exploring the methodologies and processes involved in data analysis, emphasizing the use of Business Intelligence (BI) tools. By investigating these tools and techniques, the study aims to illustrate how they can be applied to improve the quality and effectiveness of business decision-making processes.

2. To understand how BI tools such as the Orange software are used to uncover informative insights from a raw dataset collected from the website Kaggle.com.

This objective aims to demonstrate the practical application of specific BI tools, particularly the Orange software, in analyzing raw datasets. By working with a dataset sourced from Kaggle.com, the study seeks to showcase the capabilities of Orange in transforming raw data into valuable insights that can inform business strategies and operations.

3. How do businesses leverage machine learning for market analysis and predicting future trends?

This objective is to explore the role of machine learning in business contexts, particularly in market analysis and trend prediction.

The study aims to examine how businesses apply machine learning algorithms to analyze market data, identify patterns, and forecast future trends, thereby gaining a competitive edge and making more informed strategic decisions.

Together, these objectives encapsulate the study's aim to delve into data analysis using BI tools and machine learning, highlighting their significance in modern business practices.

1.5. Research Questions

- To understand the impact of balanced and Imbalanced dataset on the performance of Machine learning algorithms.
- How does the implementation of business intelligence tools enhance data-driven decision-making in call center operations?
- How effective are various machine learning models in predicting customer churn?
- How do businesses use machine learning algorithms to analyze consumer sentiment from social media, reviews, and other online platforms, and how does this analysis contribute to predicting shifts in market preferences and trends?

The remaining of this paper is organized as follows: Section two is the related work as stated in the previous research papers and articles. Section 3 is devoted to the gap analysis and comparison Section 4 introduces the proposed customer churn prediction model. Section 5 contains a brief on the experimental results achieved and analysis. Section 6 focused on customer profiling and retention. Finally, in Section 7, the conclusion of the discussion is provided along with the directions for future work.

2. Materials and Methods

Churn prediction has been done in the previous studies using various techniques like hybrid techniques, data mining, and the use of machine learning techniques. These techniques help in decision and customer relationship management while helping business organizations to either recognize, predict or manage churning customers. As for the techniques that are most known to predict the matters connected with client attrition, these are the decision trees. Various works also emphasize the idea that BI may enhance call center in a positive manner. They emphasize the ways in which workforce management, enhanced call routing, and effective resource allocation are made possible by BI tools and analytics. These results imply that improving operational performance is achieved through integrating BI into contact center initiatives [1]. This article [2] emphasizes how BI raises the bar for customer service. Call centers may analyze consumer data, predict requirements, and personalize interactions with the help of BI technologies. Research suggests that using BI analytics to generate a personalized approach increases customer happiness and loyalty. This article [3] explores the influence of business intelligence (BI) on contact center operations and customer service is described in the literature using a variety of theoretical frameworks. In order to comprehend user acceptability of BI technologies in contact centers, the Technology Acceptability Model (TAM) is frequently cited. Also, frameworks such as the Service Quality Model facilitate the evaluation of the ways in which BI affects various aspects of customer service. Even while previous research [4] has pointed out the advantages of BI, there are still significant gaps and restrictions. Some research ignores the particular difficulties encountered by smaller contact centers in favor of concentrating mostly on large-scale businesses. In addition, the temporal dimension of BI impact is not sufficiently investigated, raising concerns regarding the long-term durability of beneficial outcomes. Many research efforts [5] largely ignore the rich qualitative components of consumer and employee experiences in favor of quantitative techniques. Comprehensive insights need an understanding of the human perspective in the context of BI deployment. A comprehensive investigation of the ethical issues

[6] surrounding BI in contact centers is lacking in the literature. Concerns about privacy, permission, and the appropriate use of consumer information must be addressed when BI technologies collect and handle massive amounts of personal data. Research on BI [7] over the last 35 years has been immense however there is a limited understanding of how BI processes are connected to organizational contexts. This gap has resulted in incomplete literature that doesn't allow informed growth in the field. This study aims to understand this issue and come up with a solution using a combined framework that. The aim here is to find well researched areas as well as those that need more exploration within the BI field. This paper uses a process framework to evaluate the eight key aspects related to BI wish are; environmental influence, organizational factors, decision making, individual and managerial elements, the BI process itself, strategic results, firm performance, and organizational intelligence. The review highlights the structures and contradictions within these areas in order to provide directions for future research to expand upon. This study [8] identifies and analyzes the various trends and opportunities in the field of business created by integrating machine learning, and Artificial Intelligence into Business Intelligence. These trends include predictive analytics used to identify patterns in data to predict future outcomes in different business situations like optimizing operations, foreseeing customer needs. Moreover, AI-powered tools such as catboats improve customer interaction and assist in various tasks. These technologies allow firms to reveal insights, smooth out processes, and discover many more opportunities that provide competitive advantages in the market. Along with opportunities, the study [9] takes into account certain important considerations that can be better taken care of through BI. The development of BI has been aided by advancements such as the creation of data warehouses, improvements in data purification, hardware and software capabilities, and the introduction of web architecture. This research aims to create a framework for developing a business intelligence system. The paper further talks about the implementation and integration of AI technologies and AI robots into business processes, improving analysis, decision making, and security measures. This paper [10] brilliantly elaborates the history of BI from its birth to its development till date. It aims to identify the main aspects of BI, analyze the impact of technological innovation, and consider industries or factors that have majorly affected its evolution [11]. This research takes into account the impact of the pandemic on increasing digitalization, especially in the healthcare sector, focusing on the role of BI in making use of generated data for decision making and improved management. The study targets the therapeutic path of oncological patients, especially women with breast cancer. A decision support system (DSS) is developed with two versions; an experience-driven version and a data-driven version. The research investigates which of the two versions is more accurate in estimating costs associated with different treatment plans. [12]. This article promotes and encourages the implementation of industry 4.0, a set of technologies aimed at increasing innovation ensuring quick responses in dynamic markets. The core components of industry 4.0 are business intelligence, machine learning, predictive analysis and digital technology all of which help in revolutionizing organizational operations and developmental strategies. This article primarily focuses on the state of BI technology, emphasizing on its positive impacts on economic and business decision making across various sectors. It further predicts the future of industry 4.0, especially in the arena of BI. [13] The notion of business intelligence (BI) as seen through the prism of management. The notion and depth of business intelligence (BI) is explored, focusing on a method and later growing to cover items and tools. According to the writers, business intelligence (BI) is the act of finding, storing, and turning data into data which may be used in making decisions. They argue that utilizing data for better making choices is BI's final goal. The review notes lack of research, particular in respect to the value of decision-making within BI and a lack of technical support. To fill the knowledge gaps and provide a deeper understanding of BI's role in choices, researchers urge more research. [14] "Competition and Business Strategy in Historical Perspective " offers an in-depth study of the advancement of business strategy and competition perspectives during the fifty years preceding the present. It charts the development of early research and commercial work in determining the financial viability and competitive position of the industry. The repercussions of a market for strategic ideas and the integration of a historical component to the investigation of business methods are also covered in this piece of writing. This paper enriches our awareness of competitive edge and business strategies by integrating past patterns [15]. The paper congruencies with the 50th milestone of the Hawaii International Conference on Systems Sciences (HICSS), the paper covers the 28-year history of the Business Intelligence and Analytics (BI&A) minitrack. From Executive Information Systems (EIS) to Data

Warehousing (DW), Business Intelligence (BI), Business Analytics (BA), and, lately, big data, the authors chronicle the rise of BI&A research through several stages. It notes the way the minitrack's growth is in step with market patterns as well as global it is, with efforts from the USA, Europe, Australia, Asia/Pacific, and Africa [43] [44]. The article illustrates the fluid nature of the BI&A field and offers novel topics such as IoT data, artificial intelligence, and cognitive computing. It promotes HICSS as the biggest global forum for BI&A research. For a greater grasp of the collective effect of BI&A researchers, the authors pledge to maintain their research in future years, which includes social network studies and referrals. The imperative for regular collaboration within the BI&A community to address fresh problems and mold the field's future comes at the conclusion of the study. [16] the article is an overview of the creation of business continuity planning (BCM), defining major advances linked to the emergence of novel technologies and laws and emphasizing its lack of an alone, recognizable event or specific policy. The terrorist acts of 9/11 mark a major turning point with a bearing on organization and managerial advances in an array of various sectors. The study highlights how the need for BCM has contributed to the convergence of certain company procedures, particularly in the monetary, assistance, usefulness, non-profit, and government sector. Propagation of norms leads to the description of competing standards, depicting the growth of BCM from an Anglo centric, information technologically driven effort to a global approach. By stressing the legalization of BCM practices in companies prior to their acceptance by meta-institutions, the study adds to the debate on company history and organizational theory. In the long run, BCM changed from a context-specific need to a method that must be done all over, acting as an emblem of globally best practices for handling crises. [17] In a bid to serve as a coherent structure, the paper "Business Models: Origin, Development and Future Research Perspectives" discusses the idea of business models. A unified picture arises from a study of the conceptual history and origins of business models. The study groups 681 peer-reviewed publications into four key research zones: trend, efficiency & governance, innovation, transition & evolution, and transformation. Responding to a survey of global experts, layout, creativity, and development & change are considered major areas of future study. The article improves the area of business model research by highlighting how business models impact competitive performance as well as leadership studies worldwide [39]. A single structure that generates terms, opinions, and key elements offers a single envisioning of company models. In overall, it tries to bring up various viewpoints on company structures, showing theoretic growth, and identifies key areas for further study. [18] It includes that the Share market is one of the most unpredictable and places of high interest in the world. There are no significant methods to predict the share price. Mainly people use three ways such as fundamental analysis, statistical analysis and machine learning to predict the share price of share market but none of these methods are proved as a consistently acceptable prediction tool. So developing a prediction tool is one of the challenging tasks as share price depends on many influential factors and features [19] [40].

This study proposed a stock analysis model based on analyzing the stock market from both quantitative and qualitative perspectives, the whole reasoning process of the model is relatively transparent, and the decision results can be interpreted. The experimental results show that the proposed model is suitable for analyzing this kind of investment decision making problems under an uncertain and risky environment. [20] Forecasting the stock market has become very important in planning business activities. The prediction of stock price has driven many researches in a variety of disciplines, including computer science, statistics, economics, finance, and operations research. Recent studies have shown that the enormous amount of online information that is available in the public domain, such as Wikipedia, the social forums, news from media, have a significant impact on the investor's opinion towards the financial markets. From the Machine learning predictions, we are able to predict the stock market movements which helps the investors to invest money at the correct place and in a timely manner. The sentiment scores obtained from the analysis of the news articles is a powerful indicator of stock movements and can be used to effectively leverage the prediction of short-term trends. [21] The ability to predict stock prices is essential for informing investment decisions in the stock market. However, the complexity of various factors influencing stock prices has been widely studied. Stock relationship information is mainly obtained through industry classification data from third-party platforms, but these data are often approximate and subject to time lag. Experimental results have indicated that the K-means algorithm is indeed capable of extracting correlations between stocks from daily trading data, and these relationships can facilitate improved stock price predictions. This contribution fills a gap in the current literature and provides a more comprehensive perspective on stock prices and relationships, which is an important guide to decision-making and trading strategies in financial markets. This study has demonstrated its effectiveness for improving stock price prediction by leveraging both temporal and relational information. Stock market prediction [22] has always wedged the attention of many analysts and researchers. Popular theories suggest that stock markets are essentially a random walk and it is a fool's game to try and predict them. Predicting stock prices is a challenging problem in itself because of the number of variables which are involved. In the short term, the market behaves like a voting machine but in the longer term, it acts like a weighing machine and hence there is scope for predicting the market movements for a longer timeframe. Financial markets provide a unique platform for trading and investing, where trades can be executed from any device that can connect to the Internet. With the advent of stock markets, people have the opportunity to have multiple avenues to make their investment grow. Research can play an important role in paving the way how stock markets will be analyzed and made more robust in the future. Online trading services have already revolutionized the way people buy and sell stocks. The financial markets have evolved rapidly into a strong and interconnected global marketplace.

This research [23] describes a simulation-based machine learning methodology for assessing the performance of call centers with diverse demand kinds and heterogeneous server sets. In order to sample quality of service (QoS) outcomes as measured by service level (SL), a model for simulation for a call center with multi-skill agents and multi-class customers was first developed. A machine learning algorithm was then trained on just a few of modeling samples in order to quickly generate a look-up table of QoS for all applicant schedules [42]. One of the key performance metrics employed in the study [27] to offer insight considerations from a cost-effective perspective is the expected maximum profit standard. ProfLogit and EPMC techniques are also used to tackle the prediction problem [28]. Using a convolutional neural network (CNN), deep learning is used to forecast churn [29]. Pareto multi-criteria optimization is used to identify the best features in order to increase operational efficiency through the acquisition of new features and reusable procedures [30].In this study, decision rules are used to segment customers in the first phase, and each leaf of the tree has a model constructed for it in the second step. For churn prediction, this hybrid method is contrasted with logistic regression, decision trees, random forests, and logistic model trees [31]. This study [34] put up a theory regarding the viability of employing clever techniques to forecast client attrition utilizing deep learning neural networks. By emphasizing the usefulness of ensemble algorithms and the crucial role that data balancing strategies play in optimizing churn prediction models, this work [35] closes gaps in the body of literature. Using machine learning techniques [36] like KNN, Random Forest, and XG Boost, the suggested prediction model finds the characteristics that have a significant impact on customer attrition. The churn has been predicted using analysis of the IBM Watson dataset. EMPC (anticipated maximum profit measure for customer churn) metrics is directly integrated into the model formulation in this study's [37]. In this study [39], a metaheuristic-based churn prediction method based on large scale telecom data churn prediction is presented. As the classifier, a modified version of the Firefly algorithm is employed. This study presents [40] [43] a hybrid recommendation technique for targeted retaining activities and describes the process of predicting E-Commerce customer attrition based on support vector machines.

Previous studies have utilized a range of techniques to predict customer churn, including logistic regression, decision trees, and neural networks. Factors such as customer demographics, service usage patterns, and interaction history have been identified as significant predictors. However, the complexity and diversity of datasets require continuous refinement of models and methods to improve accuracy and actionable insights.

3. The telecommunications Sector, Churn Modeling Comparison

Churn modeling in the telecommunications sector involves the use of various machine learning algorithms to predict whether a customer will churn (leave the service) or not. Several studies have employed different datasets and algorithms to identify the factors that contribute to customer churn, allowing telecom companies to implement strategies to retain customers. Table 1 describes the churn comparison table 1.

Table 1. Churn Modeling Comparison

Ref	year	Algorithms	Dataset	Determined the Factors Important for Churn
[25]	2013	Naïve Bayes,Decision Trees,Support Vector Machine	BM Watson dataset release	21 attributes and 7043 instances make up the data set. The remaining qualities indicate whether or not they churn, with 1869 churners and 5174 non-churners.
[36]	2019	KNN, RF,XGBOOST	2015 marked the publication of the IBM Watson dataset, which includes 21 attributes and 7043 cases.	Internet service type and monthly fees
[37]	2020	A novel classifier that incorporates the expected maximum profit measure for customer churn EMPC metric into the framework building process directly	statistics from actual use from several telecom service providers	NA
[38]	2019	Pearson_correlation(OC), k Nearest neighbor (KNN) algorithm	On the Kaggle platform, a public dataset titled "Telco Customer Churn" comprises 7042 rows and 21 attributes.	NA
[39]	2017	Hybrid classification based on fireflies	Orange dataset	NA

4. Methodology

4.1. Dataset Description

The dataset Telecom churn insights: Unraveling customer Behavior in telecommunication industry" has been carefully selected from kaggle.com to show how Business intelligence is used to collect, store, and analyze customer data for the telecommunication industry in order to uncover useful insights regarding the telecom customer churns. It contains a systematic collection of customer interactions, patterns, and churn behaviors thoroughly organized for in-depth analysis. This dataset offers a unique path to understanding the relationship of customers within the telecom industry. The telecom churn insight provides a huge and diverse dataset showing a wide variety of customer data such as; attributes, usage patterns, and service preferences. This dataset is ideal for data analysts, scientists, machine learning practitioners, business intelligence professionals, marketing and customer engagement teams, and many more data enthusiasts within the telecom industry as well as from outside, allowing them to discover thoughtful insights into customer behavior. Data set contain Features; 20 and Instances; 7043 Key feature description of datasets are

A. customer ID (Alpha-numeric)

B. customer demographics: include features such as;

- Gender (categorical-nominal): male or female
- Partners(categorical-nominal): yes, or no
- Dependents (categorical-nominal): yes, or no
- senior citizens (categorical): 0(no) or 1(yes)
- tenure month (numerical-discrete): free range starting from one

C. Usage patterns; features that provide understanding of customer engagement such as; multiple lines (categorical-nominal): No, yes, or No phone service

- Internet service type (categorical-nominal): No, DSL, or Fiber Optic

- Online security (categorical-nominal): No, yes, or No internet service
- Online backup (categorical-nominal), No, yes, or No internet service
- Device protection (categorical-nominal), No, yes, or No internet service
- Tech support (categorical-nominal): No, yes, or No internet service
- streaming movies (categorical-nominal): No, yes, or No internet service
- streaming tv (categorical-nominal), No, yes, or No internet service
- Contract type (Alphanumeric), Month to month, one year, or two years
- Monthly charges (numerical-continuous):
- Total charges (numerical-continuous).

D. Customer preference; on the specific services customers prefer such as;

- Paperless billing (categorical-nominal): No or Yes
- Payment method type (categorical-nominal): Electronic check, Mailed check, Bank transfer or credit card (automatic)

E. Churn prediction (categorical-nominal) target variable; No or Yes. churn label explores predictive analytics, identifying major patterns and signals that point towards potential customer churn.

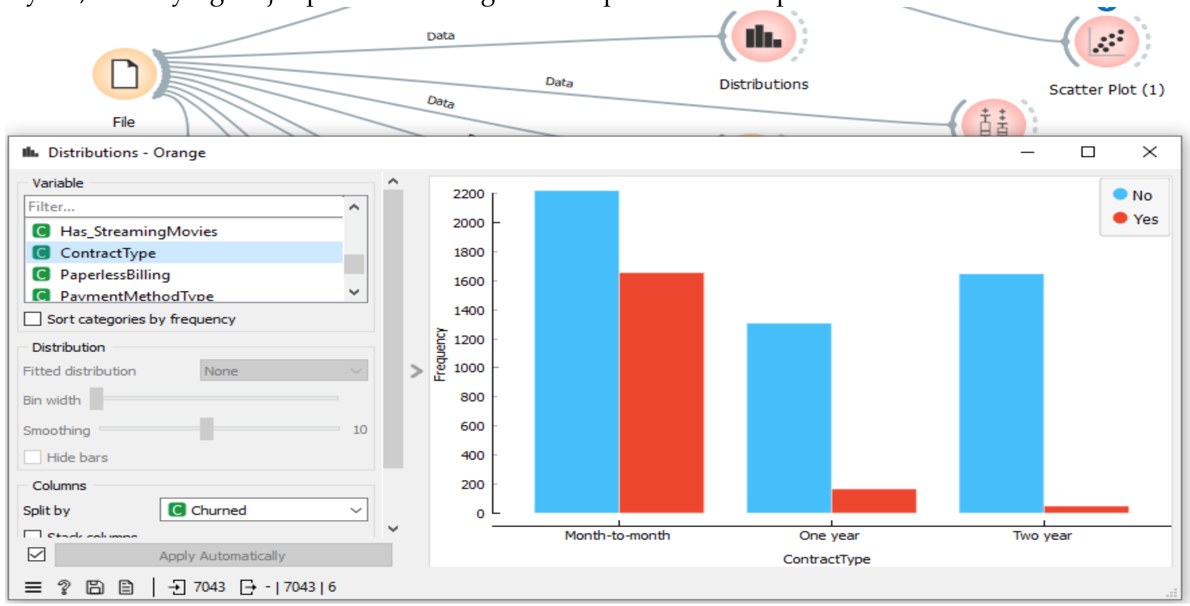

Figure 1. Dataset Feature contract type

The bar chart figure 1 generated using Orange provides a visual analysis of customer churn across different contract types. The x-axis categorizes contracts into "Month-to-month," "One year," and "Two year," while the y-axis represents the frequency of customers. The chart uses blue bars to indicate customers who did not churn and red bars for those who did. It is evident from the chart that the "Month-to-month" contract type has the highest number of customers, with a notable portion of them having churned, as indicated by the substantial red bar. In contrast, customers with "One year" and "Two year" contracts show significantly lower churn rates, highlighted by the smaller red bars in these categories. This visualization suggests a strong correlation between contract length and churn rates, indicating that customers with shorter-term contracts are more likely to leave the service compared to those with longer-term commitments.

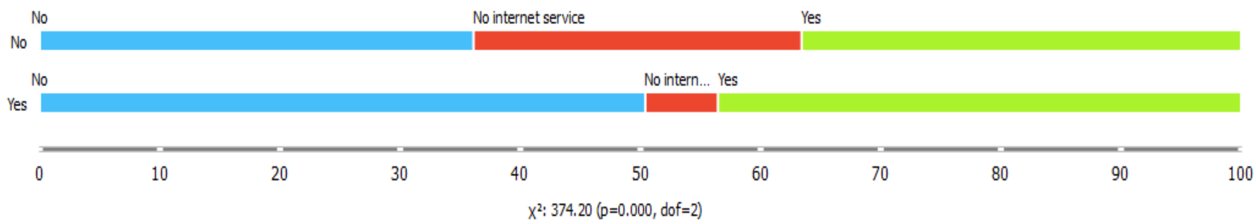

Figure 2. Box Plot analysis

The bar chart in figure 2 visualizes the distribution of internet service usage among customers, divided by their churn status. The horizontal axis shows the percentage of customers, from 0 to 100%. The chart categorizes customers into three groups: those with no internet service, those with internet service, and those who churned. The top bar represents customers who did not churn, with most of them, shown in blue, having internet service. A smaller segment, depicted in red, represents customers with no internet service, and a final segment in green represents customers with internet service. The bottom bar shows customers who churned, with a notable portion in blue indicating no internet service, a small red segment for no internet service, and a green segment for those with internet service. The chi-square statistic (χ^2) of 374.20, with a p-value of 0.000 and 2 degrees of freedom, indicates a significant association between internet service usage and customer churn, suggesting that internet service status is significantly related to the likelihood of customer churn.

4.2. Proposed methodology

The methodology diagram outlines a process for predicting customer churn in the telecom industry using supervised machine learning techniques. The process begins with the identification of the target variable, which is whether a customer will churn or not churn. This data is fed into a supervised machine learning model. The telecom churn data, which includes various features related to customer behavior and demographics, is used to train the predictive model. The predictive model employs different machine learning algorithms such as regression, decision trees, and random forests to analyze the data and make predictions. The outcome of the model is a classification of customers into two categories: churn or not churn. Based on this classification, appropriate actions can be taken to address customer retention. The process concludes with the final outcome, which determines whether a customer is likely to churn or not, thereby helping telecom companies to strategize their retention efforts effectively.

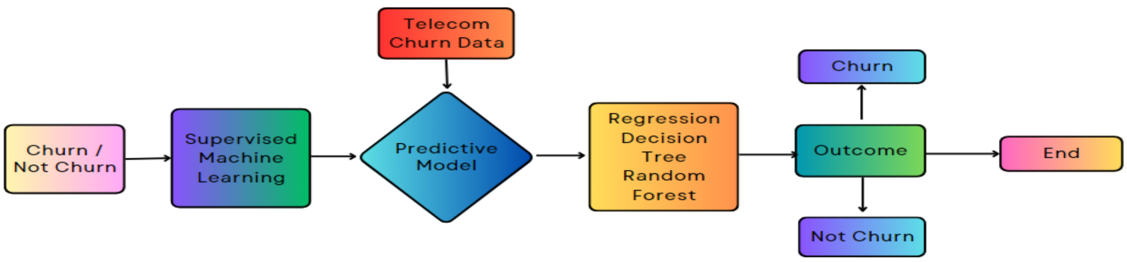

Figure 3. Proposed Methodology

The Orange model result diagram presents a workflow for predicting customer churn using various machine learning techniques. It begins with a data file that is processed through a data sampler to ensure a manageable and representative subset of the dataset. Selected columns of the dataset are then chosen for analysis. This refined data is fed into three different machine learning models: Random Forest, Decision Tree, and Logistic Regression. Each model processes the data to generate predictions. The predictions are then evaluated using multiple methods. A scatter plot visualizes the relationships in the data, while distributions offer insights into the spread and central tendencies of the predictions. Additionally, a confusion matrix is used to assess the performance of the models by comparing the predicted values against actual outcomes, allowing for a comprehensive evaluation of the models' accuracy and reliability in predicting customer churn.

4.3. Machine Learning Model Used in Study

In the telecom industry, customer churn is an important variable they have to make sure of in order to retain their customers and have steady customers that will give them constant flow of income. Many machine learning methods have been used in this problem to capture features as well as signs that may lead to customer churn. Here, outline three typical machine learning models for churn prediction in the telecommunication sector.

4.3.1. Random Forest

The Random Forest ensemble learning approach performs their process by creating many of these decision trees during training and deciding on the overall average voting for regression tasks or majority vote for classification tasks as shown in figure 5. Random Forest works well on a range of datasets and is resistant to over fitting. Equation 1 illustrates how it manages missing values and keeps accuracy even when dealing with a large number of features

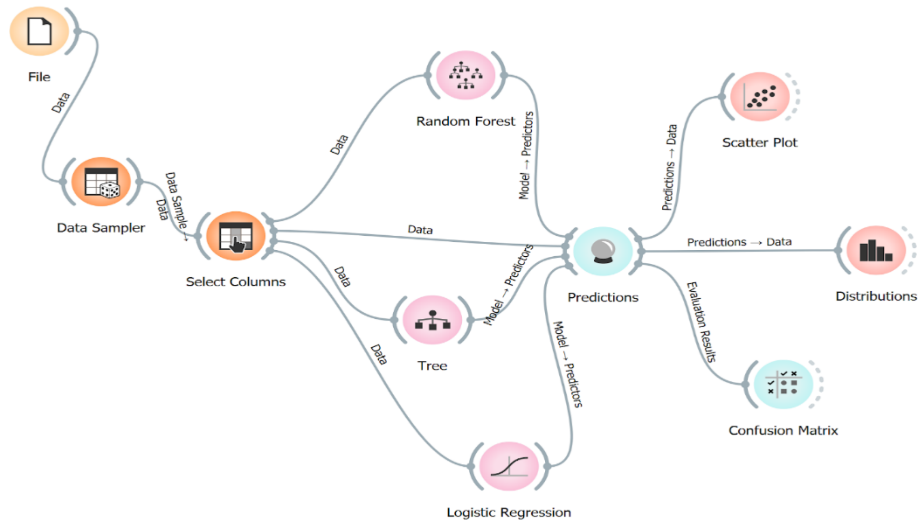

Figure 4. Complete Flow Diagram

.

Classification: Let Cˆ_b(x) be the bOth random-forest tree's class prediction.

Then $\hat{C}_{rf}^{B}$ (x) = majority vote $\{\hat{C}_b$ (x)$\}B_1$ … (1)

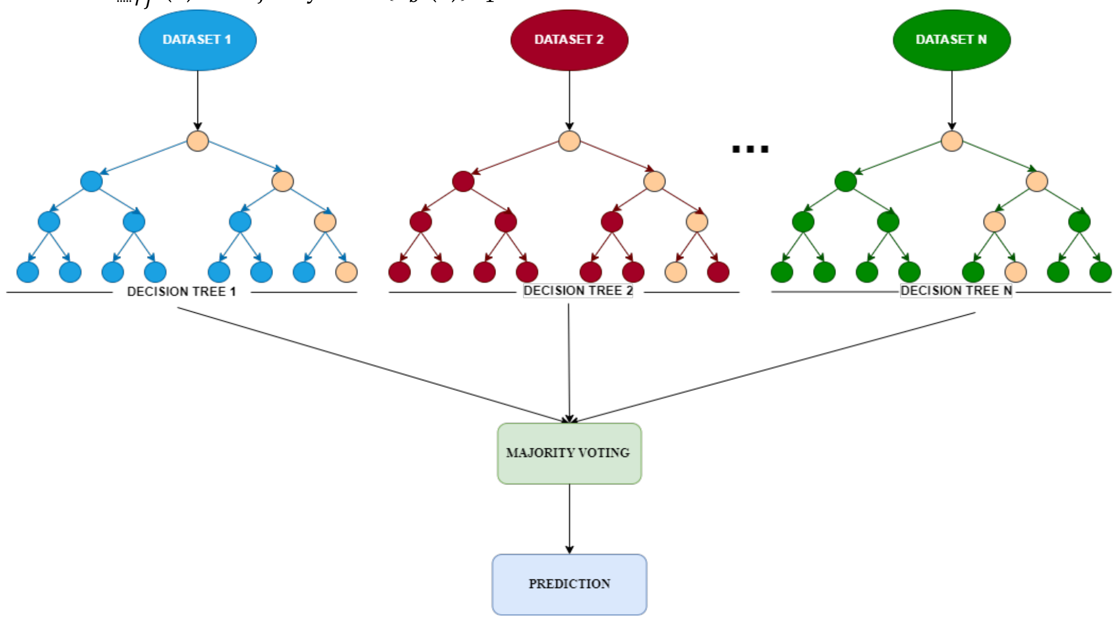

Figure 5. Representation of Random Forest

4.3.2. Logistic Regression

In case of binary classification scenarios, an instance's probability of belonging to a certain class is predicted through a statistical model which is referred to as the logistic regression. In order to make sure that the best outcome is availed and also accurate forecast of rainfall is made, the work exploits the support vector machine (SVM) classifier and logistic regression. Though its name says the opposite, logistic regression is not a regression algorithm. It is an algorithm for classification. It uses the logistic function that Figure 6 shows to derive the probability of the positive class. The logistic function that would transform every real-valued input into a number between 0 and 1. The equation for the logistic function is shown in Equation 2:

Sigmoid (z) $=\frac{1}{1+e^{-z}}$ (2)

Where:

z is a linear combination of the features and their respective coefficients, represented as $z=\beta0+\beta1x_1+\beta2x_2+\ldots+\beta n\,.\,x_n$, where β0 is the intercept, βi are the coefficients, and xi are the feature values.

The logistic regression model calculates the odds of a binary result, such as one of zero or one. The logistic function predicts the likelihood that an occurrence is in the positive class (class 1):

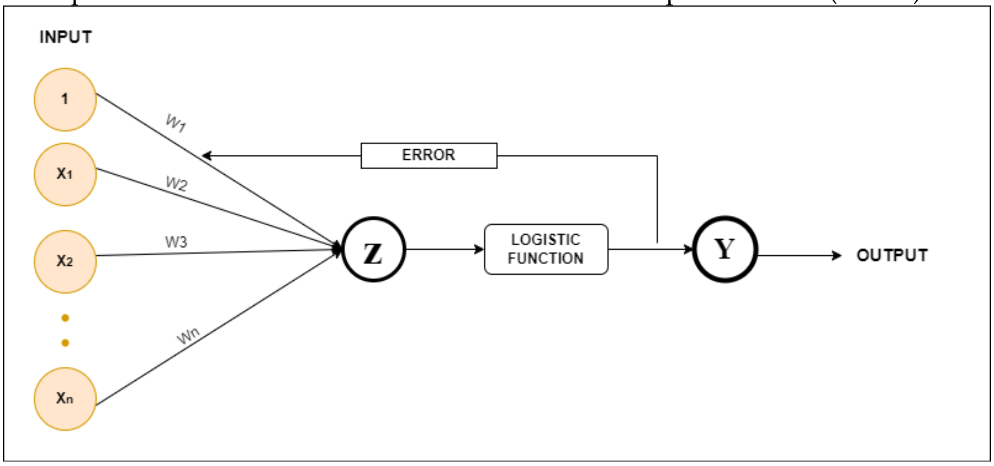

Figure 6. Representation of Logistic Regression

4.3.3. Decision Tree

Using a decision tree machine learning approach, problems with regression and classification are addressed. Its organization is comparable to a flowchart, where a decision rule is represented by each branch, an attribute or characteristic by each internal node, and a result or class label by each leaf node.

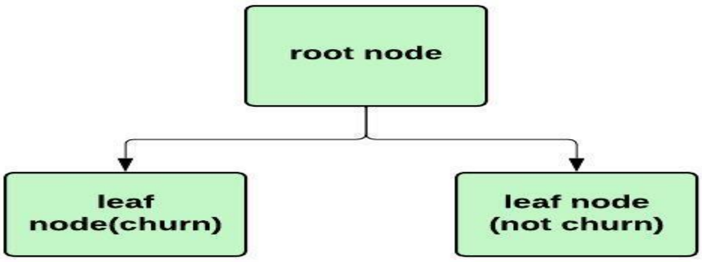

Figure 7a. Representation of Tree

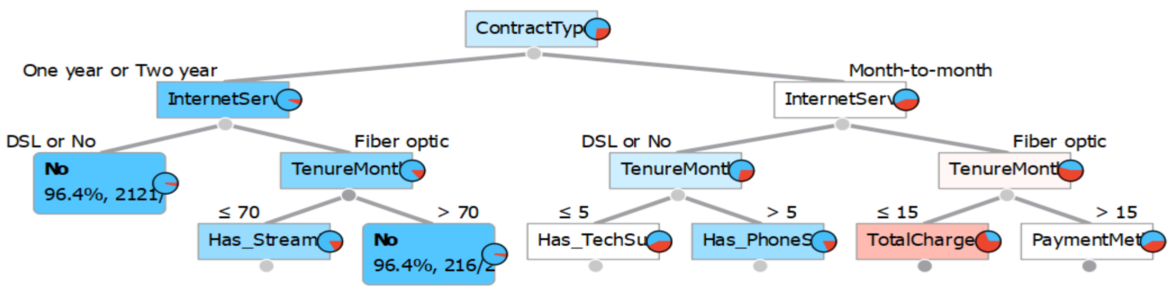

Figure 7b. Representation of Tree

The decision tree figure 7b illustrates the classification process based on various customer attributes, with "Contract Type" serving as the root node. The tree branches into two primary paths: "One year or Two year" contracts and "Month-to-month" contracts. For "One year or Two year" contracts, the next significant feature is "Internet Service," which further divides into "DSL or No," with a predominant classification of

"No" at 96.4% accuracy for 2,127 instances, and "Fiber optic," which splits based on "Tenure Month." For customers with "Fiber optic" service and a tenure of more than 70 months, the model predicts "No" with 96.4% accuracy for 216 instances. On the "Month-to-month" contract path, the tree again considers "Internet Service," dividing into "DSL or No" and "Fiber optic." For "DSL or No," further splits are based on "Tenure Month," leading to additional decision points such as "Has_TechSu" and "Has Phones." For "Fiber optic," the tenure is split into two categories: ≤ 15 months and > 15 months, leading to decisions based on "Total Charge" and "Payment Met," respectively. The decision tree highlights the importance of contract type, internet service type, and tenure in determining customer classification, with notable accuracy at various decision nodes.

5. Results

The results indicate the performance metrics of three different models: Logistic Regression, Random Forest, and Decision Tree. The Random Forest model achieved the highest AUC (0.985), indicating superior discrimination ability. It also had the highest F1 score (0.934), Precision (0.935), and MCC (0.830), reflecting a strong balance between precision and recall, and an overall high-quality prediction. The Decision Tree model performed nearly as well, with an AUC of 0.974 and the highest CA (0.942), Recall (0.942), and F1 score (0.940), indicating excellent accuracy and recall. The Logistic Regression model, while still robust, had comparatively lower scores across all metrics, with an AUC of 0.848, CA of 0.805, and an F1 score of 0.800. Overall, the Random Forest model demonstrated the best performance across most metrics, followed closely by the Decision Tree model as shown in figure 8.

Model	AUC	CA	F1	Prec	Recall	MCC
Logistic Regression	0.848	0.805	0.800	0.797	0.805	0.477
Random Forest	0.985	0.935	0.934	0.935	0.935	0.830
Tree	0.974	0.942	0.940	0.942	0.942	0.848

Figure 8. Results of Machine learning models

5.1. Criterion for Evaluating Model Confusion Matrix

In machine learning, performance metrics are measurements that indicate how successfully a model is operating on a particular job. The type of task—classification, regression, clustering, or other particular goals—determines which metrics apply. A confusion matrix is a table that displays the percentages of true positives, true negatives, false positives, and false negatives in order to evaluate the effectiveness of a categorization system.

-True Positives (TP) are cases that were correctly predicted to be positive.

-True Negatives (TN) are cases that were correctly predicted to be negative.

-False Positive (FP): Situations in which a Type I error resulted in a false positive prediction.

-False Negative (FN): Instances in which the negative prediction (Type II mistake) was made incorrectly.

5.1.1 Accuracy

All it measures is the frequency with which the classifier makes accurate predictions. The ratio of the number of accurate forecasts to the total number of predictions (see equation 3) can be used to determine accuracy.

$$\text{Accuracy}=\frac{number\ of\ correct\ prediction}{total\ number\ of\ predictions} \quad (3)$$

The "Number of Correctly Classified Instances" indicates the number of data records that the algorithm correctly classified. The "Total Number of Instances" parameter indicates the total number of data records in the dataset.

5.1.2. Precision (Positive Predictive Value)

It makes the understanding of why a significant number of instances that were accurately classified as positive occurred. Precision is beneficial in situations depicted in equation 4 to show cases in which False Positives are worse than False Negatives.

$$\text{Precision: } \frac{TP}{TP+FP} \quad (4)$$

5.1.3. Specificity (True Negative Rate)

Equation 5 defines specificity as the ratio of accurate negative predictions to the total number of actual negatives observed.

Specificity: $\frac{TN}{TN+FP}$ (5)

5.1.4. Recall (Sensitivity, True Positive Rate)

It makes clear how many actual positive cases our model correctly anticipated. Recall is a useful metric when False Positive is more worrying than False Negative (see equation 6).

Sensitivity: $\frac{TP}{FN+FP}$ (6)

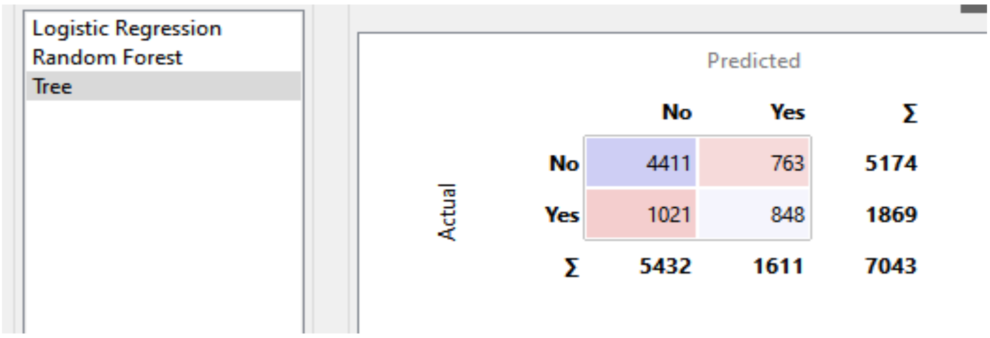

Figure 9. Confusion Matrix Decision tree

The Decision Tree model's confusion matrix shows in figure 9 that it predicted 4,411 true negatives and 848 true positives. However, it also had 1,021 false negatives and 763 false positives. This indicates that while the model is relatively accurate in predicting true positives and negatives, it has a considerable number of misclassifications, particularly false negatives, where the model incorrectly predicted "No" when the actual value was "Yes."

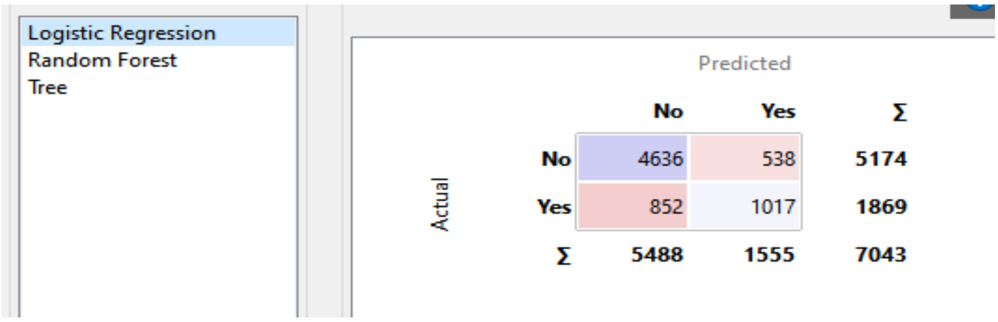

Figure 10. Confusion Matrix logistic regression

The Logistic Regression model's confusion matrix in figure 10 reveals it predicted 4,636 true negatives and 1,017 true positives. It had 852 false negatives and 538 false positives. This model shows an improvement over the Decision Tree in terms of true positive and true negative predictions. However, it still has a significant number of false negatives, though fewer than the Decision Tree, indicating a better balance between sensitivity and specificity.

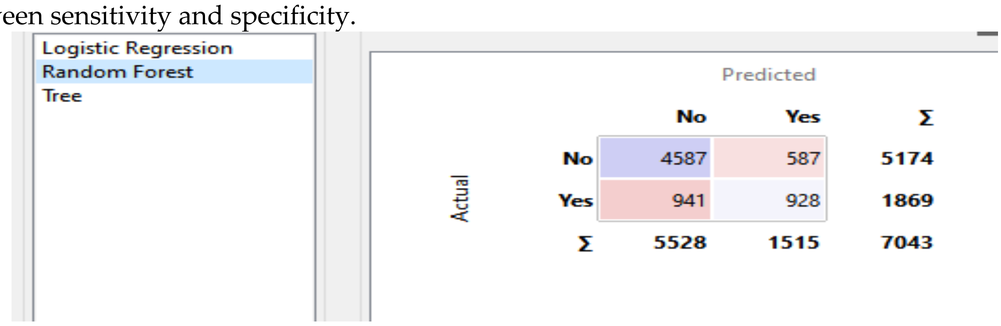

Figure 11. Confusion Matrix Random Forest

The Random Forest model's confusion matrix in figure 11 indicates it predicted 4,587 true negatives and 928 true positives. It had 941 false negatives and 587 false positives. Although this model had slightly fewer true negatives compared to Logistic Regression, it demonstrated a strong ability to correctly classify both positive and negative cases. The number of false negatives and false positives is also lower than the Decision Tree model, showcasing its overall robustness and superior performance in handling the classification task.

6. Conclusion

Telecom churn data and overall procedure for predicting customer churn: Employ supervised machine learning strategy to predict churn. Consequently, the first step involves the specification of the target variable, which in this analysis is churn or not churn, with the second aspect being data pre-processing and feature extraction. Three different approaches to machine learning are employed, and they are Decision Tree, Logistic Regression, and Random Forest. The specific process of the Orange model reflects the processes from data selection and sampling to the choice of appropriate columns to use and the use of these three models. The obtained predictions are then analyzed using the scattered plot, probability distribution, and a confusion matrix for better analysis. Based on the different metrics for the three interesting models, both the Logistic Regression and the Random Forest arrive to perfect scores for all the metrics tested well AUC, CA, F1 Score, precision, recall, MCC for customer churn predicting Performa. The Decision Tree model also yielded satisfactory performance but had slightly lower evaluation metrics suggesting potential for enhanced effectiveness in comparison to the other two models.
In conclusion, it can be stated that the key results of the presented paper demonstrate how machine learning models, accompanied by quantitative measurements and graphical visualizations, can be used as an efficient approach to address the customer churn problem in the context of the telecom industry. The outperforming of Logistic Regression and the Random Forest models illustrates the enormous potential of using these models for customer retention strategies.

7. Future Work

For future work, it would be possible to focus on expanding the decision tree model's possibilities and accuracy with the help of the following tactical approaches. One of these is the feature engineering method where one is able to create more features that help in capturing harder patterns into models such as interactions features and temporal features. Another useful strategy is using aggregation methods, for example, Random Forest or Gradient Boosting to increase the accuracy of models and minimize over fitting. To get slightly more accuracy, fine-tuning of the parameters could also be accomplished with the help of grid search or random search. Further, it is possible to balance the target data by using methods such as SMOTE or via attempting to balance the class weights with equal emphasis on both the larger and the smaller class. Applying cross-validation check methods will help in achieving a more accurate estimate of the degree of the model's generalizability. Applying some techniques revealed in this paper, such as SHAPE values, contributes to the enhancement of interpretability to allow stakeholders to interpret feature contributions to predictions. The idea of constructing a real-time prediction and monitoring would help in getting the necessary data and information which can be useful for the necessary interventions. The inclusion of data from diverse sources, including callback logs and social media interactions will provide a broader perspective and additionally, enhance the forecasting precision. Another approach, which could be useful in the long term, is to set the feedback loop where the results of the analysis in response to user interactions are taken into account and then reintroduced back to the model to help adjust for changes in the patterns over time. Lastly, comparing the results with other machine learning algorithms will assist in discovering the best performing algorithm on the given problem. Thus, it is possible to improve the decision tree model, make it more stable, more accurate, and more applicable to real-world tasks, based on these aspects.

References

1. Prem, M. J., & Karnan, M. (2013). Business intelligence: optimization techniques for decision making. International Journal of Engineering, 2(8), 1081-1092.
2. Niu, Y., Ying, L., Yang, J., Bao, M., & Sivaparthipan, C. B. (2021). Organizational business intelligence and decision making using big data analytics. Information Processing & Management, 58(6), 102725.
3. Trkman, P., McCormack, K., De Oliveira, M. P. V., & Ladeira, M. B. (2010). The impact of business analytics on supply chain performance. Decision Support Systems, 49(3), 318-327.
4. Subramaniam, L. V., Faruquie, T. A., Ikbal, S., Godbole, S., & Mohania, M. K. (2009, March). Business intelligence from voice of customer. In 2009 IEEE 25th International Conference on Data Engineering (pp. 1391-1402). IEEE.
5. Chen, H., Chiang, R. H., & Storey, V. C. (2012). Business intelligence and analytics: From big data to big impact. MIS quarterly, 1165-1188.
6. Grigori, D., Casati, F., Castellanos, M., Dayal, U., Sayal, M., & Shan, M. C. (2004). Business process intelligence. Computers in industry, 53(3), 321-343.
7. Talhaoui, Y., & Kohtamäki, M.35 years of research on business intelligence process: a synthesis of a fragmented literature. Management Research Review, 44(5), 677-717. (20
8. Bharadiya, J. P. Machine learning and AI in business intelligence: Trends and opportunities. International Journal of Computer (IJC), 48(1), 123-134. (2023)
9. Bharadiya, J. P. A comparative study of business intelligence and artificial intelligence with big data analytics. American Journal of Artificial Intelligence, 7(1), 24. (2023)
10. Marjamäki, P. Evolution and trends of business intelligence systems: a systematic mapping study. (2017)
11. Basile, L. J., Carbonara, N., Pellegrino, R., & Panniello, U. Business intelligence in the healthcare industry: The utilization of a data-driven approach to support clinical decision making. Technovation, 120, 102482. (2023)
12. Tavera Romero, C. A., Ortiz, J. H., Khalaf, O. I., & Ríos Prado, A. Business intelligence: business evolution after industry 4.0. Sustainability, 13(18), 10026. (2021)
13. Shollo, A., & Galliers, R. D. (2015). Towards an understanding of the role of business intelligence systems in organizational knowledge. Information Systems Journal, 26(4), 339–367. https://doi.org/10.1111/isj.12071.
14. Ghemawat, P. (2002). Competition and Business Strategy in Historical Perspective. Business History Review, 76(1), 37–74. https://doi.org/10.2307/4127751
15. Marjanovic, O., & Dinter, B. (2018). Learning from the history of Business Intelligence and Analytics Research at HICSS – A Semantic Text Mining Approach. Communications of the Association for Information Systems, 43(43), 775–791. https://doi.org/10.17705/1cais.04340
16. Herbane, B. (2010). The evolution of business continuity management: A historical review of practices and drivers. Business History, 52(6), 978–1002. https://doi.org/10.1080/00076791.2010.511185
17. Wirtz, B. W., Pistoia, A., Ullrich, S., & Göttel, V. (2016). Business Models: Origin, Development and Future Research Perspectives. Long Range Planning,49(1), 36–54. https://doi.org/10.1016/j.lrp.2015.04.001
18. Cheng Zhao, Ping Hu, Xiaohui Liu, Xuefeng Lan, Haiming Zhang. (Published: 24 February 2023). "Stock Market Analysis Using Time Series Relational Models for Stock Price Prediction."
19. Yu-Wang Chen, Long-Hao Yang. (Published: 09 February 2023). "Stock market analysis method." The University of Manchester, United Kingdom; Fuzhou University, China
20. Dev Shah, Haruna Isah, Farhana Zulkernine, Mi Zhou. (Published: 27 May 2019). "Stock Market Analysis: A Review and Taxonomy of Prediction Techniques." Hefei University of Technology, China.
21. P Kavya, S Saagarika, R T Subavarsshini, C G Nivethini, Marimuthu Muthuvel, Palaniswamy Velvadivu. (Published: July 2021). "Stock Market Analysis."
22. Tapashi Gosswami, Sanjit Kumar Saha. (September 2018). "Stock Market Data Analysis and Future Stock Prediction using Neural Network." International Journal of Computer Science and Information Security, 16(9):6. Georg-August-Universität Göttingen; Jahangirnagar University; Comilla University.
23. Li, S., Wang, Q., & Koole, G. (2019). Predicting call center performance with machine learning. In Advances in Service Science: Proceedings of the 2018 INFORMS International Conference on Service Science (pp. 193-199). Springer International Publishing.

A. Idris and A. Khan, ‘‘Customer churn prediction for telecommunication: Employing various features selection techniques and tree based ensemble classifiers,’’ in Proc. 15th Int. Multitopic Conf., Dec. 2012, pp. 23–27.

24. M. Kaur, K. Singh, and N. Sharma, ‘‘Data mining as a tool to predict the churn behavior among Indian bank customers,’’ Int. J. Recent Innov. Trends Comput. Commun., vol. 1, no. 9, pp. 720–725, Sep. 2013.

25. V. L. Miguéis, D. van den Poel, A. S. Camanho, and J. F. e Cunha,''Modeling partial customer churn: On the value of first product-category purchase sequences,'' Expert Syst. Appl., vol. 12, no. 12, pp. 11250–11256, Sep. 2012.
26. B. Zhu, B. Baesens, and S. K. van den Broucke, ''An empirical comparison of techniques for the class imbalance problem in churn prediction,'' Inf. Sci., vol. 408, pp. 84–99, Oct. 2017.
27. E. Stripling, S. van den Broucke, K. Antonio, B. Baesens, and M. Snoeck, ''Profit maximizing logistic model for customer churn prediction using genetic algorithms,'' Swarm Evol. Comput., vol. 40, pp. 116–130,June. 2018.
28. Mishra and U. S. Reddy, ''A novel approach for churn prediction using deep learning,'' in Proc. IEEE Int. Conf. Comput. Intell. Comput. Res., Dec. 2017, pp. 1–4.
29. S. Mitrović, B. Baesens, W. Lemahieu, and J. D. Weerdt, ''On the opera- tional efficiency of different feature types for telco churn prediction,'' Eur. J. Oper. Res., vol. 267, no. 3, pp. 1141–1155, Jun. 2018.
30. D. Caigny, K. Coussement, and K. W. D. Bock, ''A new hybrid classifi-cation algorithm for customer churn prediction based on logistic regression and decision trees,'' Eur. J. Oper. Res., vol. 269, no. 2, pp. 760–772,Sep. 2018.
31. Manzoor, A., Qureshi, M. A., Kidney, E., & Longo, L. (2024). A Review on Machine Learning Methods for Customer Churn Prediction and Recommendations for Business Practitioners. IEEE Access.
32. Imani, M. (2024). Customer Churn Prediction in Telecommunication Industry: A Literature Review.
33. Morozov, V., Mezentseva, O., Kolomiiets, A., & Proskurin, M. (2022). Predicting customer churn using machine learning in IT startups. In Lecture Notes in Computational Intelligence and Decision Making: 2021 International Scientific Conference" Intellectual Systems of Decision-making and Problems of Computational Intelligence", Proceedings (pp. 645-664). Springer International Publishing.
34. Sikri, A., Jameel, R., Idrees, S. M., & Kaur, H. (2024). Enhancing customer retention in telecom industry with machine learning driven churn prediction. Scientific Reports, 14
35. Pamina, J., Raja, B., SathyaBama, S., Sruthi, M. S., & VJ, A. (2019). An effective classifier for predicting churn in telecommunication. Jour of Adv Research in Dynamical & Control Systems, 11.
36. Höppner, S., Stripling, E., Baesens, B., vanden Broucke, S., & Verdonck, T. (2020). Profit driven decision trees for churn prediction. European journal of operational research, 284(3), 920-933.
37. Sjarif, N., Rusydi, M., Yusof, M., Hooi, D., Wong, T., Yaakob, S., ... & Osman, M. (2019). A customer Churn prediction using Pearson correlation function and K nearest neighbor algorithm for telecommunication industry. Int. J. Advance Soft Compu. Appl, 11(2).
38. Ahmed, A. A., & Maheswari, D. (2017). Churn prediction on huge telecom data using hybrid firefly based classification. Egyptian Informatics Journal, 18(3), 215-220.
39. Naz, S., Aslam, M., & Sayed, A. (2023). Prevalence of Anemia and its Determinants among the Rural Women of Khyber Pakhtunkhwa-Pakistan. Annals of Human and Social Sciences, 4(4), 42-50.
40. Shobana, J., Gangadhar, C., Arora, R. K., Renjith, P. N., Bamini, J., & devidas Chincholkar, Y. (2023). E-commerce customer churn prevention using machine learning-based business intelligence strategy. Measurement: Sensors, 27, 100728.
41. Zaidi, A., Karim, A. A., Mohiuddin, S., Khan, A., Syed, A., Jehangir, M., & Afzal, I. (2018). Dental Sensitivity Associated With Consumption Of Fizzy Drinks: A Cross Sectional Study. Pakistan Journal of Medicine and Dentistry, 7(4), 5-5.
42. Zhong, X. J., Liu, S. R., Zhang, C. W., Zhao, Y. S., Sayed, A., Rajoka, M. S. R., ... & Song, X. (2024). Natural Alkaloid Coptisine, Isolated from Coptis chinensis, Inhibits Fungal Growth by Disrupting Membranes and Triggering Apoptosis. Pharmacological Research-Modern Chinese Medicine, 100383.
43. Abbas, M., Arslan, M., Bhatty, R. A., Yousaf, F., Khan, A. A., & Rafay, A. (2024). Enhanced Skin Disease Diagnosis through Convolutional Neural Networks and Data Augmentation Techniques. Journal of Computing & Biomedical Informatics, 7(01).
44. Hussain, S.K., Ramay, S.A., Shaheer, H., Abbas T., Mushtaq M.A., Paracha, S., & Saeed, N. (2024). Automated Classification of Ophthalmic Disorders Using Color Fundus Images, Volume: 12, No: 4, pp. 1344-1348 DOI:10.53555/ks.v12i4.3153