Journal of Information Systems Engineering and Management
2025, 10(25s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

Customer Churn Prediction using Machine Learning Approach: A Comprehensive Study

Anurag Bhatnagar[1], Sumit Srivastava[2]
[1,2]Department of Information Technology, Manipal University Jaipur, India
[1]anurag.bhatnagar@jaipur.manipal.edu, 2sumit.srivastava@jaipur.manipal.edu

ARTICLE INFO

Received: 29 Dec 2024
Revised: 13 Feb 2025
Accepted: 22 Feb 2025

ABSTRACT

Churn is a term that combines" Change" and" Turn." The ability to predict customer churn is a signnificant concern for service providers. In today's market, customers are increasingly discerning and seek to access the best services available in their daily lives. This pursuit of superior services often leads to churn or attrition for organizations. Consequently, forecasting churn has emerged as one of the most formidable challenges faced by service providers. The complexity of churn prediction is heightened by the vast amount of customer data, its sparsity, and the imbalanced nature of this data. This paper highlights the research conducted by various scholars on Customer Churn Prediction (CCP) methodologies within the telecommunications sector.

Introduction of Churn, It's Types and Customer Relationship Manager (CRM) :
In the quest of finding better services, customers switch from one company to another, this movement is well-known as "customer churn". The fundamental reason behind customer churn is disappointment with the services, astonishing charges, unfriendly offers, and ferocious consumer provision [1]. The rate of churn is also known as the attrition rate. Identifying churners is a classification problem and whether a customer is availing services or not is fundamental. The world is going through technological advancements every day, and service providers are trying at their level best to provide the best services to their customers by integrating new technical features. Conversely, clients are looking for the greatest deals and high-caliber services at reasonable costs. Sebastiaan Hoppner asserts that maintaining current consumers is less expensive than acquiring new ones [2]. In general, there are two types of churners: those that are voluntary and those that are involuntary or passive [3]. A consumer is considered to have voluntarily switched from one service provider to another when they stop using their services from the former supplier. On the other hand, involuntary churn occurs when a service provider terminates a customer's account due to non-payment or other reasons [4]. Two subcategories of voluntary churn are incidental or cyclical churn and deliberate or active churn [5]. Incidental churn is the result of customers' changing locations or lifestyles. Reasons for purposeful churn include convenience, cost of service, technological improvements, social, psychological, and quality aspects [6]. The basic classification of churners is visualized in Figure-1 [7].

Objectives: The objective of this work is to consolidate existing research into a single document, thereby assisting researchers and scholars in their analysis and future investigations in this field. Additionally, this study also proposes a general model for predicting customer churn in the telecommunications industry, which will serve as a valuable resource for emerging researchers in this area.

Methods: The practice and advancement of computer systems that can operate independently of human intervention and learn from their experiences is known as machine learning, which is one of the most prominent features of the computer industry. To optimize efficacy, various statistical models and algorithms are implemented. Supervised machine learning, a subtype of machine learning and artificial intelligence, is characterized by its reliance on labelled data. The labelled dataset is utilized in supervised machine learning to train the algorithm for accurate data classification and outcome prediction. The second type of machine learning is unsupervised learning approaches. In this machine learning category, the user is not obligated to oversee or train the model. This machine learning enables the model to autonomously uncover previously unnoticed patterns and information. Unsupervised machine learning mostly addresses unlabeled data. The third category of machine learning is reinforcement learning. This category allows the agent to learn from its experiences by utilizing input from previous acts. Regarding the issue of anticipating churn, supervised machine learning offers optimal support to researchers. Supervised machine learning entails categorizing a tagged input dataset into distinct classes. In churn prediction, the dataset is categorized into two classifications: whether the client has departed from the organization or not. It is a binary classification problem.

Results: The findings demonstrate that machine learning approaches, particularly those utilizing advanced algorithms such as ensemble methods and deep learning, offer significant advantages over traditional statistical techniques. These models are more accurate and predictive, enabling businesses to better identify clients who are at risk and carry out targeted retention campaigns. The study underscores the necessity of selecting appropriate features and continuously improving models to adapt to evolving consumer behavior and market conditions.

Conclusions: This comprehensive study has examined the application of machine learning techniques to customer churn prediction, highlighting its potential to transform how businesses perceive and handle client retention. Numerous ML models have been thoroughly explored in this study, providing insightful information about the effectiveness, drawbacks, and strengths of each model in predicting customer attrition.

Keywords: Churn, Attrition, Telecom, Data Sparsity, Imbalanced Data.

INTRODUCTION

In the quest of finding better services, customers switch from one company to another, this movement is well-known as "customer churn". The fundamental reason behind customer churn is disappointment with the services, astonishing charges, unfriendly offers, and ferocious consumer provision [1]. The rate of churn is also known as the attrition rate. Identifying churners is a classification problem and whether a customer is availing services or not is fundamental. The world is going through technological advancements every day, and service providers are trying at their level best to provide the best services to their customers by integrating new technical features. Conversely, clients are looking for the greatest deals and high-caliber services at reasonable costs. Sebastiaan Hoppner asserts that maintaining current consumers is less expensive than acquiring new ones [2]. In general, there are two types of churners: those that are voluntary and those that are involuntary or passive [3]. A consumer is considered to have voluntarily switched from one service provider to another when they stop using their services from the former supplier. On the other hand, involuntary churn occurs when a service provider terminates a customer's account due to non-payment or other reasons [4]. Two subcategories of voluntary churn are incidental or cyclical churn and deliberate or active churn [5]. Incidental churn is the result of customers' changing locations or lifestyles. Reasons for purposeful churn include convenience, cost of service, technological improvements, social, psychological, and quality aspects [6]. The basic classification of churners is visualized in Figure-1 [7].

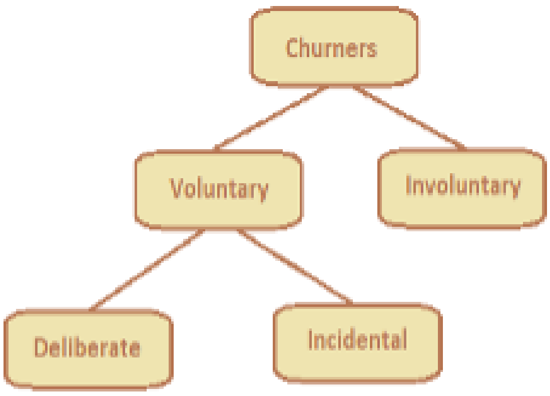

Figure 1: Classification of Churners

Moreover, the activity of churn can also be sub-divided into three sub-parts [8]:

- Total – In case of officially cancelled contract.
- Hidden or concealed – In which the contract is not yet over but the customer has stopped availing services form a long time.
- Partial – In which the contract is not discarded but the customer is not availing services to some extent and using services of Competitor Company.

To focus on customer retention organizations dedicatedly form Customer Relationship Manager (CRM). The primary role of CRM is to pay attention on the services and offers presented to cus- tomers of organization. Despite so many efforts offers and services due to behaviour expectancy acuity and latest offers by competitors provoke customers to churn [9]. Therefore, even though a lot of emphasis has been dedicated to preventing client churn attrition or churn is an ongoing process. Consequently, to prevent or reduce organizational turnover the phrase" churn prediction" must be narrowed down. It is true that new businesses may have frequent or significant attrition but for well-established businesses the churn scenario is the opposite. Not only is it a difficult effort for many telecom companies but attrition and churn prediction are also one of the most difficult occupations in the banking network and wireless sensor fields of networking. These days churn prediction is a routine procedure for both profitable and non-profitable businesses. The amount of the data set also grows with the capture of consumer data [10]. Static churn prediction models that emphasize these data set characteristics have been the subject of research. These models combine certain demographic and behavioral attributes or even integration of social aspects. The sparsity and volume of data are making it increasingly difficult to extract meaningful information. Because of this predicting client attrition grew increasingly difficult necessitating the use of data mining and various machine learning approaches. much while academics are working extremely hard to anticipate churn there are still some areas where their efforts fall short because customer behavior is a very complex attribute in and of itself and it gets much harder when considered monthly.

OBJECTIVES:

MOTIVATION:

While it provides a comprehensive source of information for analysing consumer behaviour it also neces- sitates sophisticated analytical techniques to extract actionable insights. Businesses can develop more effective strategies to mitigate customer churn and gain a more profound comprehension of the factors driving it by utilizing a variety of ML algorithms. The purpose of this paper is to address this lacuna by conducting a comprehensive analysis of customer churn prediction using machine learning techniques. This study endeavours to provide valuable insights into the optimal utilization of these methods for attrition prediction by conducting a comprehensive analysis of a variety of ML models including their strengths and limitations. In essence this research not only enhances the academic comprehension of ML applications in customer attrition prediction but also offers practical advice to practitioners who aspire to leverage the potential of machine learning to achieve superior business results.

OUR CONTRIBUTION:

To explore churn prediction in the telecommunications sector it was essential for us to conduct a thorough examination of the existing research efforts in this area. By analysing studies conducted from 2009 to 2021 as illustrated in Figure-2 we gathered significant insights and methodologies employed by various researchers. These results are systematically presented in a tabular format within this study intended to act as a reference point for emerging researchers entering this field. Through our detailed analysis we have developed a comprehensive architectural framework aimed at predicting customer churn in the telecom industry as clearly depicted in Figure-3. This framework outlines a structured workflow that begins with the crucial task of selecting an appropriate dataset from the telecommunications sector. Following this the data undergoes careful preprocessing which includes data cleaning addressing imbalanced data filtering and feature selection. The processed data is then divided into training and testing sets which are subsequently utilized by different machine learning models as per the selection of researchers to predict customer churn. This carefully constructed architecture is designed to provide essential guidance and support to novice researchers pursuing their academic endeavours in the telecommunications domain with the goal of mastering customer churn prediction.

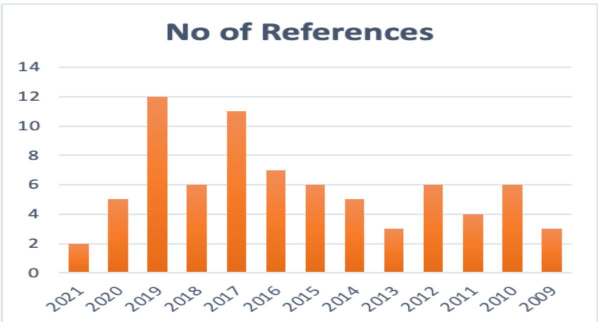

Figure 2: Year Wise representation of Research Papers

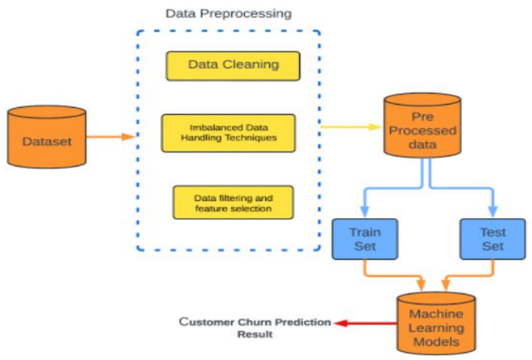

Figure 3: Comprehensive Architectural Framework

PAPER ORGANISATION

The remaining paper is organized as follows. Section-2 highlights the motivation of the paper and our contribution in this field. Section-3 delves into various classification models of supervised machine learning available for classification of churners in the telecom industry. Section-4 covers various techniques for data preprocessing which help classification models to enhance performance and improve accuracy. Section-5 summarizes the state-of-the-artwork done by different researchers. Section-6 discusses the future scope of the work and illustrates lessons learned and concludes the paper. Figure-4 condenses the framework of this paper.

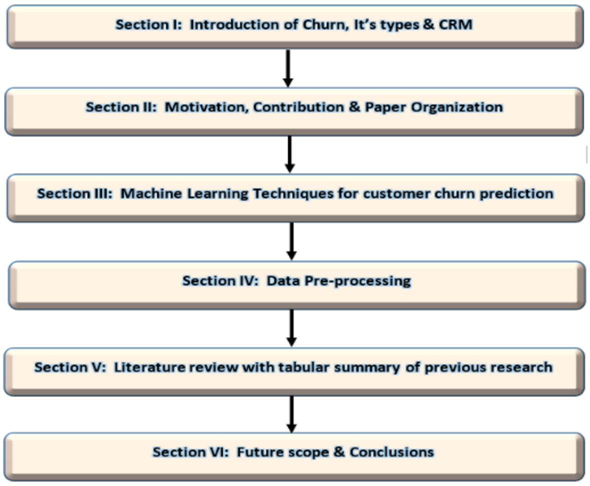

Figure 4: Paper Structure

METHODS: MACHINE LEARNING TECHNIQUES USED IN CUSTOMER CHURN PREDICTION

The practice and advancement of computer systems that can operate independently of human intervention and learn from their experiences is known as machine learning, which is one of the most prominent features of the computer industry. To optimize efficacy, various statistical models and algorithms are implemented.

Further, the machine learning algorithms are divided into three basic categories as shown in the figure-3 given below.

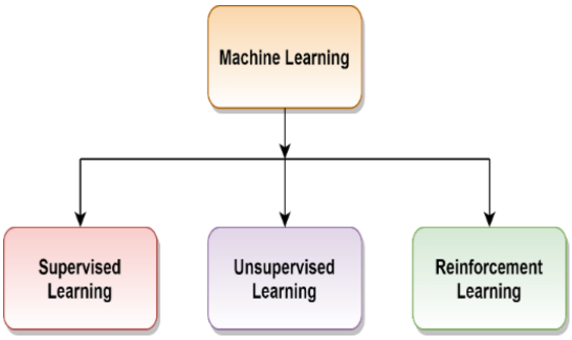

Figure 5: Types of machine learning

Supervised machine learning, a subtype of machine learning and artificial intelligence, is characterized by its reliance on labelled data. The labelled dataset is utilized in supervised machine learning to train the algorithm for accurate data classification and outcome prediction. The second type of machine learning is unsupervised learning approaches. In this machine learning category, the user is not obligated to oversee or train the model. This machine learning enables the model to autonomously uncover previously unnoticed patterns and information [54].

Unsupervised machine learning mostly addresses unlabeled data. The third category of machine learning is reinforcement learning. This category allows the agent to learn from its experiences by utilizing input from previous acts. Regarding the issue of anticipating churn, supervised machine learning offers optimal support to researchers. Supervised machine learning entails categorizing a tagged input dataset into distinct classes. In churn prediction, the dataset is categorized into two classifications: whether the client has departed from the organization or not.

It is a binary classification problem.

Some of the classification models which are frequently used to predict churn are as follows:

LOGISTIC REGRESSION

Logistic regression is a statistical analysis method employed to predict data values based on previous observations of a dataset [55]. Logistic regression predicts the value of a dependent variable based on the relationship between one or more independent data factors. Logistic regression is a widely utilized classification algorithm due to its simplicity and efficiency in addressing binary and linear classification issues [32]. This is applicable to both categorical and continuous variables. This algorithm analyses the independent variable to forecast the value of a singular dichotomous dependent variable.

SUPPORT VECTOR MACHINE (SVM)

The Support Vector Machine (SVM) is a very dependable machine learning model. Support Vector Machine (SVM) is employed for binary classification tasks. After the model is trained and provided with labelled training data, the SVM can categorize new text for each category.

K-NEAREST NEIGHBOURS (KNN)

This is one of the simplest machine learning algorithms. KNN is helpful in both the classification and regression problems [21, 22]. The implementation and understanding of KNN is easier, but the major bottleneck of KNN is its slow speed. As the size of data increases, KNN becomes slower. In cases where high accuracy is prioritized instead of human-readability, KNN is preferred.

RANDOM FOREST

This algorithm is useful in solving both classification and regression problems. The fundamental principle for classification is constructing a multitude of decision trees during training time. For regression tasks, the mean or average prediction of the individual trees is returned. The collection of decision trees is collectively known as the random forest. Bagging and the feature randomness techniques are two important methods used in random forest during the creation of individual trees whose prediction by committee is more accurate than that of any individual tree [30, 42]. In this algorithm, multiple decision trees are made and they are then merged to get a more accurate and stable prediction result.

NAIVE BAYES

It is a probabilistic classifier based on the implementation of Bayes' Theorem. The basic principle of this classifier is that each pair of features classified is independent of each other [22]. On the other side, naive Bayes is also known as a bad estimator because probability-based outputs are not to be taken too seriously.

DECISION TREE

This is another probability-based classifier that enables researchers to decide on the classification process. A decision tree can be described as a map or chart that helps determine a course of action or exhibit a statistical probability [17]. It is a graphical presentation of all possible solutions for a given condition. A decision tree uses all the features/attributes of a dataset to provide the result, whereas random forest selects specific features and dataset to build multiple decision trees and then provide the result by averaging the computed result [26].

An assortment of Data Preprocessing

PRE-PROCESSING OF DATASET

The" dataset" is a fundamental component in every research endeavor. Data is often unclean and poorly handled, necessitating data pre-processing. If the dataset is not clean and does not meet the specified quality standards the results will likewise be subpar. Proper data processing is essential prior to the use of any machine algorithm to prevent the occurrence of anonymous outcomes. Dimension reduction, data cleansing, and data transformation are all strategies for data pre-processing [10].

FEATURE SELECTION

In machine learning, using nearly all aspects of a dataset may diminish the generalization ability to address some real-world problems. In a dataset, it is uncommon for all attributes to be beneficial for model implementation. Utilizing features of a dataset in a machine learning model presents two primary challenges: incorporating all attributes may lead to an overfitted model, while employing too few attributes can result in an underfitted model. Consequently, feature selection is among the most critical responsibilities. Filter, Wrapper, and Embedded approaches are the most utilized feature selection techniques [11, 12].

ENSEMBLE METHODS

Ensemble methods use multiple learning algorithms to obtain better predictive performance of classifi- cation models. Common types of ensembles are:

BAGGING

It is also known as Boot Strap Aggregation. Accuracy and stability of a machine learning model may be improved by using this ensemble technique. If there is a dataset D with tuples – d in it, then this technique will work as follows – if 'i' is the iteration variable then for (i = 1, 2...n) a training set

Di of d tuples will be sampled from the original dataset. A classifier model Sample Di is used to train classification model CMi. Each classifier model CMi assigns a class label for an unknown tuple X that counts as one vote. The tuple X is classified by a classifier CMi by taking a majority vote among the predictions made by each base classifier [13].

BOOSTING

Robert and Yoav proposed the boosting ensemble technique. In the first proposal, the boosting algorithm was not adaptive. Schapire and Freund proposed an adaptive boosting algorithm known as AdaBoost. This algorithm manipulates the training dataset [13]. In boosting, each tuple is given some weightage. K classifiers series is iteratively constructed. The basic principle of boosting algorithm provides weight to each tuple and then feeds these tuples in classification model Ci. The weight of the misclassified tuples is increased in next step and weight of correctly classified tuples is decreased [14].

STACKING

This ensemble technique combines various classifier models built by different algorithms [15]. The following are two of the steps of Stacking: On the original dataset various models are applied. Stacking ensemble technique uses the meta-learning algorithms which help it to learn which classification model is the best suitable model for prediction out of two or more base models. According to Wolpert's the original data and the classification is known as level-0 dataset and level 0 classification model [30]. Similarly in the second stage Meta models are used to derive a classifier from level 0 training data. This is one of the most rarely used ensemble techniques as compared to bagging and boosting [14].

VOTING

The concept of a voting ensemble technique revolves around the idea of merging predictions from various individual classification models. This methodology is employed to enhance the overall performance and accuracy of the model through a collective decision-making process. By leveraging the diverse perspectives and strengths of multiple models the voting ensemble approach aims to achieve superior predictive capabilities compared to any single model working in isolation. One of the other ensemble techniques is voting. Each of the classifiers assigned a class label of each instance is determined [12].

LITERATURE REVIEW AND TABULAR PROJECT OF PREVIOUS RESEARCH WORK:

This section presents an overview of research work done in the prediction of customer churn in the telecom industry.

Nadia Alboukaey et al. stated that customer churn not only has its impact on revenue but also on the customer base [16]. According to him, in churn prediction both accuracy and the time when churn prediction is done is equally important. He proposed that whether a statistic or dynamic dataset is referred to as churn prediction may be delayed because nobody can predict on which day customer will decide to churn. So, he discussed MTN operator of Syrian Arab Republic customers' daily behaviour as multivariate time series over a period of 150 days and four models were proposed – Featured-based models → RFM-based model, Statistics-based model and deep learning-based models → LSTM-based model and CNN-based models. This research stated that daily models performed more accurately than monthly models in churn prediction. However, LSTM-based models are equal to CNN-based model and RFM-based models. Moreover, these three performed better than the statistics-based model. The first outcome of this paper is extracting meaningful features from time series-based dataset using RFM-based model. Random forest and decision tree models are also used on monthly datasets to compare results with time series-based models. The second outcome is based on deep learning-based models which suggest LSTM and CNN models.

Another impact of this work was an increase in the efficiency of customer retention as it was implemented daily so the churners can be identified at an early stage.

Sebastiaan Hoppner et al. stated that customer churn prediction is a binary classification problem which is now a day has become one of the biggest challenges for service providers in various domains [17]. Traditional churn prediction models are not aligned with the core of business profit maximization as these models were focused on not only the misclassification cost but also the advantages of correct classification. So, the proposed work was more focused on the core of profit maximization in business. An expected maximum profit measure for customer churn (EMPC) has been proposed in this work which is directed towards the selection of the most profitable churn model. ProfTree was proposed in the paper which is derived from an evolutionary algorithm for learning profit-driven decision trees. The outcome of this paper stated that the ProTree achieves significant profit improvements compared to traditional tree-based models.9 real life churn dataset - Belg1, Belg2, Belg3, Chile, Duke1, Duke2, Korean1, Korean2, and UCI were taken into consideration. Some of the three based models such as ProfTree, C4.5, CTree, C4.5 (pruned), EvTree, CART, CART (pruned) was applied on the mentioned dataset, and it was observed that ProfTree aims to be the most profitable model among all applied models.

ProfTree is an effective model not only because of its high hit rate and recall rate but also because of its effectiveness in accurately identifying churners and indicating future churners. Thus, ProfTree model is proven the best among all aligned models with the business requirement of profit maximization.

Jaehyun Ahn et al. proposed churn prediction in internet services, games, insurance, and management [18]. They differentiated between churn definitions by collecting it from various business fields like administration, marketing, IT, telecom, newspapers, insurance, and psychology. Based on this definition, churn loss, feature engineering, and prediction models have been classified. The significance of this work is directly giving benefits to researchers as they can select the definition of churn and its associated model they are interested in. This paper also suggested some feature modification techniques which are mentioned below:

1. **One-Hot encoding:** To convert categorical values into numerical values this technique is used.

2. **Bucketing:** This is also known as discrete binning or Data binning. This is useful to convert wide and sparse features which have large variance into categorical data type.

3. **Data Imputation:** In the dataset, missing values may lead to wrong prediction thus it is important to remove them. This technique is useful for replacing missing values with any means.

4. **Normalization:** This technique scales individual samples into unit norm.

In this work, the researcher compared churn prediction analysis techniques using log data. The work has its focus on internet services and games, insurance, and management. High sparse data into low dimensional dense data is achieved using feature embedding. The time window is used to select customer churn than selecting the complete churner dataset. CAC (Customer Acquisition Cost) or CLV (Customer Lifetime Value) is used to calculate loss cost of churners. Deep learning is outperforming conventional customer churn prediction techniques as it can capture minute changes. Whereas the future scope of the paper is to compare the performance of individual models as stated by the researcher. Praveen et al. proposed six phases in their research work in [19]. In the first phase, the most suitable dataset was selected by considering variance analysis and correlation matrix. In the second phase, data cleaning and filtering was done. In the third phase, feature selection using gravitational search algorithm was taken into consideration. In the fourth phase, predictive models mentioned below were applied by the researchers. In phase five, cross-validation was done using K-folds, and in the final phase six, confusion matrix and AUC curve were used to state the result. They have divided their dataset into 80-20 for training and testing respectively. Naive Bayes, Logistic Regression, Random Forest, Decision Tree, and support vector machine, as well as boosting and ensemble techniques, have been applied to predict customer churn in the next phase. K-fold cross-validation is also applied to observe the accuracy of prediction models and to avoid overfitting. At the end, the confusion matrix and AUC curve had been taken into consideration for the results. The research work stated that Adaboost and XGboost classifiers are better than others in terms of accuracy.

J. Vijaya et al. stated that due to huge competition in the telecom industry, it has become challenging to avoid churn and maintain the losses in revenue [20]. Conversely, forecasting churners is challenging due to the sparse characteristics and limited amount of information. This research elucidated particle swarm optimization (PSO) and proposed three variations of PSO for forecasting customer turnover. Three types of PSO are employed: PSO using feature selection as a pre-processing method, PSO integrated with simulated annealing, and PSO with a combination of both feature selection and simulated annealing. These three variations have been evaluated against decision trees, naive Bayes, KNN, SVM, and random forest. Accuracy, true positive rate, true negative rate, false positive rate, precision, recall, and F- measure are parameters that have garnered interest for comparative analysis. Research indicates that metaheuristics demonstrated superior efficiency and enhanced predictability.

Praveen et al. identified that in the data set which is imbalanced the classification model's performance decreases [21]. The proposed approach states that the KNN and centroid is useful to remove class imbalance from a dataset so that performance of the classification model can be improved. To improve the performance of the model, Principal Component Analysis (PCA) is also used by researchers to implement dimensional reduction. Thus, the approach based on centroid oversampling technique to generate synthetic samples by using KNN algorithm to identify the nearest neighbors and compute centroid of these nearest data points in the space was proposed. In this paper, the accuracy to predict churn in the telecom industry is achieved 94% using KNN by using the value of k as 3. Table-1 shows a summary of the research done by various researchers to identify customer churn in telecom industry.

Table 1: A Summarized Research Work for Customer Churn Prediction in Telecom

S. No.	Title	Method	Source
1	Analysis of Churn Prediction: A Case Study on Telecommunication Services in Macedonia	LR (Logistic Regression), NB (Naive Bayes), KNN (K Nearest Neighbours)	Aleksandar J. Petkovski, Biljana L. Risteska Sto-jkoska [22]
2	Churn Analysis in Telecommunication using Logistic Regression	Logistic Regression	Helen Treasa Sebastia, Rupali Wagh [23]
3	Application of data mining techniques in customer relationship management: A literature review and classification	Association rule, Decision tree, Genetic algorithm, Neural networks, K-Nearest Neighbour, Linear/Logistic regression	E.W.T. Ngai, Li Xiu [24]
4	On the operational efficiency of different feature types for telco Churn prediction	Pareto multicriteria optimization	Sandra Mitrovi´c, Bart Bae-sens [25]
5	A new hybrid classification algorithm for customer churn prediction based on logistic regression and decision trees	Logit leaf model	Arno De Caigny [26]
6	Churn prediction on huge telecom data using hybrid firefly-based classification	Firefly algorithm, Simulated annealing	Ammar A. Q. Ahmed, Maheswari D. [27]

7	Prediction of customer attrition of commercial banks based on SVM Model	SVM	Benlan He, Yong Shi [28]
8	K Nearest Sequence Method and Its Application to Churn Prediction	KNS (K-nearest sequence)	Dymitr Ruta, Detlef Nauck [29]
9	A comparative analysis of data preparation algorithms for customer churn prediction: A case study in telecommunication	Logistic regression, Bagging, Bayesian network, Naive Bayes, Decision tree, Neural network, Random forests, SVM	Kristof Coussement, Stefan Lessmann, Geert Verstraeten [30]
10	Attributes in predictive models: A case study in churn prediction in the energy sector	Decision tree, Logistic regression, SVM	Julie Moeyersoms, David Martens [31]
11	Churn Prediction in Telecommunication Using Data Mining Technology	Data Mining Function	Rahul J. Jadhav [32]
12	A new hybrid classification algorithm for customer churn prediction based on logistic regression and decision trees	Hybrid approach using Decision Tree and Logistic Regression	Arno De Caigny [33]
13	A Survey on Customer Churn Prediction using Machine Learning Techniques	Neural networks, Ensemble classifier, Boosting, Genetic Algorithm	Saran Kumar A, Chandrakala D. [34]
14	Genetic algorithm based neural network approaches for predicting churn in cellular wireless network services	Genetic Algorithm (GA) based neural network (NN)	P.C. Pendharkar [35]
15	Customer churn prediction in telecommunications Simulation Modelling: Practice and Theory	SVM-POLY using AdaBoost	T. Vafeiadis, K.I. Diamantaras, G. Sarigiannidis, K. Chatzisavvas [36]
16	A customer churn prediction model in telecom industry using boosting	Boosting9Logistic Regression	L. Ning, L. Hua, L. Jie, Z. Guangquan [37]
17	A Comparative Analysis of Data Preparation Algorithms for Customer Churn Predic- tion: A Case Study in the Telecommunication Industry	Predictive analytics, Data preparation techniques	Kristof Coussement, Stefan Lessmann [38]
18	Effectual Predicting Telecom Customer Churn using Deep Neural Network	Neural Network	BhawnaNigam,Himanshu Dugar [39]
19	Bagging and boosting classification trees to predict churn	Bagging, Stochastic gradient boosting, Binary logit model	Aur´elie Lemmens, Christophe Croux [40]
20	Churn predicitionn in subscription services: An ap- plication of support vector machines while comparing two parameter selection techniques	Support vector machines, Logistic regression, Random forests	Kristof Coussement, Dirk Van den Poel [41]
21	Application of data mining techniques for customer churn in banking	Decision tree, Neural networks	T. Huang, C. Ho [42]

22	SVM model for customer attrition in the retail industry	SVM	B. He, Y. Shi [43]
23	Boosting Decision Trees for Telecom Churn Prediction	Trees Decision, Boosting	C. Kuang, D. Wang [44]
24	Genetic algorithms and random forest hybrid method for predicting churn	Genetic algorithm, Random Forest	P. Wang, Y. Wang [45]
25	Customer retention in the mobile telecommunications market using hybrid classifiers	Hybrid classifiers	R. Kumar, A. Sinha [46]
26	Neural network models for customer churn prediction in telecom	Neural networks	Y. Zheng, X. Li [47]
27	Costsensitive churn prediction models for telecom industry	Cost-sensitive algorithms	B. Liu, X. Li [48]
28	Machine learning models for predicting customer churn in subscription services	Random forest, Gradient boosting	J. Davis, A. Smith [49]
29	A novel deep learning approach to predict churn	Deep learning, Neural networks	K. Lee, H. Kim [50]
30	A robust ensemble learning model for predicting churn in telecom	Ensemble learning	T. Chen, X. Gu [51]
31	Survival analysis for churn prediction in subscription services	Survival analysis	E. Saas, S. Barnett [52]
32	Customer churn prediction using extreme gradient boosting (XGBoost)	XGBoost	P. Brown, G. Taylor [53]
33	Predictive modeling of customer churn in telecom using logistic regression	Logistic regression	C. Evans, M. Martin [54]

FUTURE SCOPE AND CONCLUSION

This article offers a comprehensive analysis of machine learning methodologies for predicting customer churn, although there are other potential research directions that could augment the comprehension and implementation of these strategies.

FUTURE SCOPE

- **Exploration of Advanced Machine Learning Techniques:** Future research may investigate the incorporation of sophisticated ML methodologies like deep learning and reinforcement learning, which could enhance accuracy and predictive capabilities. Examining the comparison between these sophisticated models and traditional methodologies may provide significant insights into their practical effectiveness for churn prediction.

- **Incorporation of Real-Time Data:** The present study predominantly utilizes historical data for churn prediction, lacking the integration of real-time data. Future study may concentrate on the incorporation of real-time data streams, including customer interactions and behavioral data, to improve the promptness and precision of churn predictions. This entails creating systems proficient in real-time data processing and analysis to deliver dynamic churn predictions.

- **Cross-Industry Analysis:** Although this study concentrates on certain businesses, expanding the research to include a cross-industry analysis could yield a more comprehensive understanding of the application of various ML models. Analyzing churn prediction models across diverse sec- tors, including retail, telecommunications, and financial services, may elucidate industry-specific determinants and necessary model modifications.

- **Customer Segmentation:** Analyzing the responses of various customer segments to churn pre- dictors might provide more detailed data. Future research may concentrate on creating customized models that consider the distinct attributes of different client categories, resulting in more person- alized and effective retention methods.

- **Influence of External Factors:** External factors, including economic conditions, competitive dynamics, and technology innovations, might affect customer behavior. Subsequent study may integrate these characteristics into the machine learning models to evaluate their influence on churn prediction and enhance the models' robustness.

CONCLUSION

This comprehensive study has examined the application of machine learning techniques to customer churn prediction, highlighting its potential to transform how businesses perceive and handle client retention. Numerous ML models have been thoroughly explored in this study, providing insightful information about the effectiveness, drawbacks, and strengths of each model in predicting customer attrition.

The findings demonstrate that machine learning approaches, particularly those utilizing advanced algorithms such as ensemble methods and deep learning, offer significant advantages over traditional statistical techniques. These models are more accurate and predictive, enabling businesses to better identify clients who are at risk and carry out targeted retention campaigns. The study underscores the necessity of selecting appropriate features and continuously improving models to adapt to evolving consumer behavior and market conditions.

REFRENCES

[1] Saran Kumar, A and Chandrakala, D. "A survey on customer churn prediction using machine learning techniques." International Journal of Computer Applications, vol. 975, pp. 8887, 2016.
[2] Aleksandar J. Petkovski, Biljana L. Risteska Stojkoska, Kire V. Trivodaliev, and Slobodan A. Kala- jdziski. "Analysis of churn prediction: a case study on telecommunication services in Macedonia." In 2016 24th Telecommunications Forum (TELFOR), pp. 1–4, IEEE, 2016.
[3] Goran Kraljevi´c and Sven Gotovac. "Modeling data mining applications for prediction of prepaid churn in telecommunication services." Automatika, vol. 51, no. 3, pp. 275–283, 2010.
[4] Hee-Su Kim and Choong-Han Yoon. "Determinants of subscriber churn and customer loyalty in the Korean mobile telephony market." Telecommunications Policy, vol. 28, no. 9-10, pp. 751–765, 2004.
[5] Goran Kraljevi´c and Sven Gotovac. "Modeling data mining applications for prediction of prepaid churn in telecommunication services." Automatika, vol. 51, no. 3, pp. 275–283, 2010.
[6] Rob Mattison. The Telco Churn Management Handbook. Lulu.com, 2006.
[7] Anurag Bhatnagar and Sumit Srivastava. "Performance Analysis of Hoeffding and Logistic Algo- rithm for Churn Prediction in Telecom Sector." In 2020 International Conference on Computation Automation and Knowledge Management (ICCAKM), pp. 377–380, IEEE, 2020.
[8] Vladislav Lazarov and Marius Capota. "Churn Prediction." Expert Systems with Applications, 2007.
[9] Anurag Bhatnagar and Sumit Srivastava. "Performance Analysis of Hoeffding and Logistic Algo- rithm for Churn Prediction in Telecom Sector." In 2020 International Conference on Computation Automation and Knowledge Management (ICCAKM), pp. 377–380, IEEE, 2020.
[10] Jiawei Han and Micheline Kamber. Data Mining: Concepts and Techniques. Morgan Kaufmann, 2006.
[11] Ian H. Witten and Eibe Frank. "Data mining: practical machine learning tools and techniques with Java implementations." Acm Sigmod Record, vol. 31, no. 1, pp. 76–77, 2002.
[12] Helmut Berger, Dieter Merkl, and Michael Dittenbach. "Exploiting partial decision trees for feature subset selection in e-mail categorization." In Proceedings of the 2006 ACM symposium on Applied computing, pp. 1105–1109, 2006.
[13] What Is Data Mining. Introduction to Data Mining. Springer, 2006.
[14] Yoav Freund, Robert E. Schapire, et al. "Experiments with a new boosting algorithm." In ICML, vol. 96, pp. 148–156, Citeseer, 1996.
[15] David H. Wolpert. "Stacked generalization." Neural Networks, vol. 5, no. 2, pp. 241–259, 1992.
[16] Nadia Alboukaey, Ammar Joukhadar, and Nada Ghneim. "Dynamic behavior-based churn prediction in mobile telecom." Expert Systems with Applications, vol. 162, pp. 113779, 2020.

[17] Sebastiaan H¨oppner, Eugen Stripling, Bart Baesens, Seppe vanden Broucke, and Tim Verdonck. "Profit driven decision trees for churn prediction." European Journal of Operational Research, vol. 284, no. 3, pp. 920–933, 2020.
[18] Jaehyun Ahn, Junsik Hwang, Doyoung Kim, Hyukgeun Choi, and Shinjin Kang. "A survey on churn analysis in various business domains." IEEE Access, vol. 8, pp. 220816–220839, 2020.
[19] Praveen Lalwani, Manas Kumar Mishra, Jasroop Singh Chadha, and Pratyush Sethi. "Customer churn prediction system: a machine learning approach." Computing, vol. 104, no. 2, pp. 271–294, 2022.
[20] J. Vijaya and E. Sivasankar. "An efficient system for customer churn prediction through particle swarm optimization-based feature selection model with simulated annealing." Cluster Computing, vol. 22, pp. 10757–10768, 2019.
[21] Praveen Lalwani, Manas Kumar Mishra, Jasroop Singh Chadha, and Pratyush Sethi. "Customer churn prediction system: a machine learning approach." Computing, vol. 104, no. 2, pp. 271–294, 2022.
[22] Aleksandar J. Petkovski, Biljana L. Risteska Stojkoska, Kire V. Trivodaliev, and Slobodan A. Kala- jdziski. "Analysis of churn prediction: a case study on telecommunication services in Macedonia." In 2016 24th Telecommunications Forum (TELFOR), pp. 1–4, IEEE, 2016.
[23] Sebastian H. T and Wagh R. "Churn Analysis in Telecommunication Using Logistic Regression." Orient Journal of Computer Science and Technology, vol. 10, no. 1.
[24] Eric W. T. Ngai, Li Xiu, and Dorothy C. K. Chau. "Application of data mining techniques in customer relationship management: A literature review and classification." Expert Systems with Applications, vol. 36, no. 2, pp. 2592–2602, 2009.
[25] Sandra Mitrovi´c, Bart Baesens, Wilfried Lemahieu, and Jochen De Weerdt. "On the operational efficiency of different feature types for telco churn prediction." European Journal of Operational Research, vol. 267, no. 3, pp. 1141–1155, 2018.
[26] Arno De Caigny, Kristof Coussement, and Koen W. De Bock. "A new hybrid classification algorithm for customer churn prediction based on logistic regression and decision trees." European Journal of Operational Research, vol. 269, no. 2, pp. 760–772, 2018.
[27] Ammar A. Q. Ahmed and D. Maheswari. "Churn prediction on huge telecom data using hybrid firefly-based classification." Egyptian Informatics Journal, vol. 18, no. 3, pp. 215–220, 2017.
[28] Benlan He, Yong Shi, Qian Wan, and Xi Zhao. "Prediction of customer attrition of commercial banks based on SVM model." Procedia Computer Science, vol. 31, pp. 423–430, 2014.
[29] Dymitr Ruta, Detlef Nauck, and Ben Azvine. "K nearest sequence method and its application to churn prediction." In International Conference on Intelligent Data Engineering and Automated Learning, pp. 207–215, Springer, 2006.
[30] Kristof Coussement, Stefan Lessmann, and Geert Verstraeten. "A comparative analysis of data preparation algorithms for customer churn prediction: A case study in the telecommunication in- dustry." Decision Support Systems, vol. 95, pp. 27–36, 2017.
[31] Julie Moeyersoms and David Martens. "Including high-cardinality attributes in predictive models: A case study in churn prediction in the energy sector." Decision Support Systems, vol. 72, pp. 72–81, 2015.
[32] Rahul J. Jadhav and Usharani T. Pawar. "Churn prediction in telecommunication using data mining technology." International Journal of Advanced Computer Science and Applications, vol. 2, no. 2, 2011.
[33] Arno De Caigny, Kristof Coussement, and Koen W. De Bock. "A new hybrid classification algorithm for customer churn prediction based on logistic regression and decision trees." European Journal of Operational Research, vol. 269, no. 2, pp. 760–772, 2018.
[34] Akshara Santharam and Siva Bala Krishnan. "Survey on customer churn prediction techniques." International Research Journal of Engineering and Technology, vol. 5, no. 11, pp. 3, 2018.
[35] Parag C. Pendharkar. "Genetic algorithm-based neural network approaches for predicting churn in cellular wireless network services." Expert Systems with Applications, vol. 36, no. 3, pp. 6714–6720, 2009.
[36] Thanasis Vafeiadis, Konstantinos I. Diamantaras, George Sarigiannidis, and K. Chatzisavvas. "A comparison of machine learning techniques for customer churn prediction." Simulation Modelling Practice and Theory, vol. 55, pp. 1–9, 2015.
[37] Ning Lu, Hua Lin, Jie Lu, and Guangquan Zhang. "A customer churn prediction model in telecom industry using boosting." IEEE Transactions on Industrial Informatics, vol. 10, no. 2, pp. 1659– 1665, 2012.
[38] Kristof Coussement, Stefan Lessmann, and Geert Verstraeten. "A comparative analysis of data preparation

algorithms for customer churn prediction: A case study in the telecommunication in- dustry." Decision Support Systems, vol. 95, pp. 27–36, 2017.

[39] Bhawna Nigam, Himanshu Dugar, and M. Niranjanamurthy. "Effectual predicting telecom customer churn using deep neural network." International Journal of Engineering and Advanced Technology, vol. 8, no. 5, 2019.

[40] Aurelie Lemmens and Christophe Croux. "Bagging and boosting classification trees to predict churn." Journal of Marketing Research, vol. 43, no. 2, pp. 276–286, 2006.

[41] Kristof Coussement and Dirk Van den Poel. "Churn prediction in subscription services: An ap- plication of support vector machines while comparing two parameter-selection techniques." Expert Systems with Applications, vol. 34, no. 1, pp. 313–327, 2008.

[42] Jonathan Burez and Dirk Van den Poel. "Separating financial from commercial customer churn: A modeling step towards resolving the conflict between the sales and credit department." Expert Systems with Applications, vol. 35, no. 1, pp. 497–514, 2008.

[43] Kristof Coussement and Dirk Van den Poel. "Integrating the voice of customers through call center emails into a decision support system for churn prediction." Information & Management, vol. 45, no. 3, pp. 164–174, 2008.

[44] Parag C. Pendharkar. "Genetic algorithm based neural network approaches for predicting churn in cellular wireless network services." Expert Systems with Applications, vol. 36, no. 3, pp. 6714–6720, 2009.

[45] Yaya Xie, Xiu Li, E.W.T. Ngai, and Weiyun Ying. "Customer churn prediction using improved balanced random forests." Expert Systems with Applications, vol. 36, no. 3, pp. 5445–5449, 2009.

[46] Chih-Fong Tsai and Yu-Hsin Lu. "Customer churn prediction by hybrid neural networks." Expert Systems with Applications, vol. 36, no. 10, pp. 12547–12553, 2009.

[47] Nicolas Glady, Bart Baesens, and Christophe Croux. "Modeling churn using customer lifetime value." European Journal of Operational Research, vol. 197, no. 1, pp. 402–411, 2009.

[48] Kristof Coussement, Dries F. Benoit, and Dirk Van den Poel. "Improved marketing decision making in a customer churn prediction context using generalized additive models." Expert Systems with Applications, vol. 37, no. 3, pp. 2132–2143, 2010.

[49] Bingquan Huang, Brian Buckley, and T.-M. Kechadi. "Multi-objective feature selection by using NSGA-II for customer churn prediction in telecommunications." Expert Systems with Applications, vol. 37, no. 5, pp. 3638–3646, 2010.

[50] Adem Karahoca and Dilek Karahoca. "GSM churn management by using fuzzy c-means clustering and adaptive neuro fuzzy inference system." Expert Systems with Applications, vol. 38, no. 3, pp. 1814–1822, 2011.

[51] Xiaobing Yu, Shunsheng Guo, Jun Guo, and Xiaorong Huang. "An extended support vector machine forecasting framework for customer churn in e-commerce." Expert Systems with Applications, vol. 38, no. 3, pp. 1425–1430, 2011.

[52] Wouter Verbeke, David Martens, Christophe Mues, and Bart Baesens. "Building comprehensible customer churn prediction models with advanced rule induction techniques." Expert Systems with Applications, vol. 38, no. 3, pp. 2354–2364, 2011.

[53] Hossein Abbasimehr, Mostafa Setak, and Javad Soroor. "A framework for identification of high- value customers by including social network-based variables for churn prediction using neuro-fuzzy techniques." International Journal of Production Research, vol. 51, no. 4, pp. 1279–1294, 2013.

[54] Thanasis Vafeiadis, Konstantinos I. Diamantaras, George Sarigiannidis, and K. Chatzisavvas. "A comparison of machine learning techniques for customer churn prediction." Simulation Modelling Practice and Theory, vol. 55, pp. 1–9, 2015.

[55] Sahar F. Sabbeh. "Machine-learning techniques for customer retention: A comparative study." International Journal of Advanced Computer Science and Applications, vol. 9, no. 2, 2018.