Journal of Current Science and Technology, January-April 2023 Vol. 13 No. 1, pp. 136-161

ISSN 2630-0656 (Online)

Cite this article: Arockia Panimalar, S., Krishnakumar, A. (2023, January). A review of churn prediction models using different machine learning and deep learning approaches in cloud environment. *Journal of Current Science and Technology, 13*(1), 136-161. DOI: 10.14456/jcst.2023.12

Journal of Current Science and Technology

Journal homepage: https://jcst.rsu.ac.th

A review of churn prediction models using different machine learning and deep learning approaches in cloud environment

S. Arockia Panimalar*and A. Krishnakumar

*Department of Computer Science, Sree Saraswathi Thyagaraja College, Pollachi, Coimbatore – 642107, India

*Corresponding author; E-mail: sapanimalar@gmail.com

Received 11 May 2022; Revised 3 November 2022; Accepted 29 November 2022; Published online 29 January 2023

Abstract

Customer churn is portrayed as the event that occurs when the customer quits the organization's services or products. The reasons for this dissatisfaction are the higher costs, low-level understanding of service plan, bad support, high subscription rate, and service quality, etc., Companies ought to be capable in the prediction of that customer behavior perfectly to retain on-hand customers and minimize the churn rate of customers in before occurrence. The study elucidates a review analysis of various churn prediction models, striving in different sectors through utilizing different machine-learning approaches, Deep-learning algorithms, metaheuristic optimization techniques, feature extraction-based methods, and hybrid approaches. The paper also surveys commonly utilized machine-learning techniques on a cloud computing platform to determine customer churn patterns. The churn prediction model, with better precision results, facilitates spotting firms which are near to getting churn and directing firms' focus to minimize overall churn percentage, shaping retention policies, and boosting the company's revenue.

Keywords: *CCP-customer churn-prediction; deep-learning; DNN-deep neural-networks; feature classification; feature-selection; LSTM-long short-term memory; machine-learning.*

1. Introduction

Customer churn is one of the significant mounting problems in all industries. The objective of the business organization shifted to retaining customers rather than acquiring new customers. Various research showed that gaining a new customer is around 5 to 10 times more costlier than retaining the already existing users. This can be accomplished through customer satisfaction and making them happier at work. In such a scenario, the customer churn-prediction phase stands out as the more significant research area where it aids business firms in retaining or maintaining their on-hand subscription users or general customers. Since there is an increase in the emergence of innovative business models and competitors, customer-acquisition costs seem to increase subsequently. Therefore, it becomes highly necessary for any business organization or service provider to apply a churn prediction process. Because of the direct impacts on the industries' revenue, the organization looks for an efficient model to estimate churn prediction. This churn prediction process could be explored through machine learning models, Data mining techniques, and within a cloud infrastructure. The rate of churn in an organization substantially affects customers' lifetime value since it significantly impacts the company's revenue and the service length that the organization offers. In the case of Telecom industries, churns frequently vary, irrespective of different subscriber categories, such as mobile-based subscribers or fixed-line subscribers. In an approach to rectify the issue of predicting the churner, it becomes a crucial challenge for every organization. An efficient and reliable model needs to be launched for churn prediction and to gain the performance outcomes leading to revenue uplift within the organization. The present study enumerated various review analyses of churn prediction models, relying on a different domain.

Generally, the Churn-prediction phase is defined as a binary classification process that distinguishes churners from that non-churners. The effective churn-prediction model is constructed by considering various factors, including the feature-selection process, customer social network, techniques utilized, customer-behavior data, etc., those factors assist in the development of the prominent churn-prediction framework and aid in enhancing the model performance. Once the effective model is developed (Jain, Khunteta, & Srivastava, 2021), the model classifies the churners from non-churners. The model can be developed through machine-learning approaches and Deep learning algorithms (trained by neural layers) followed by a classification phas

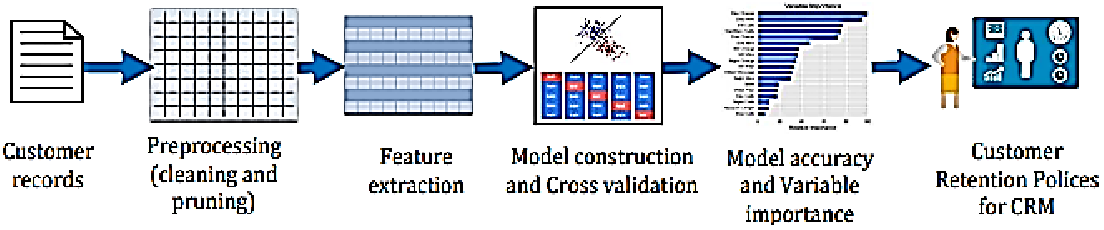

Figure 1 Different churn-prediction model phases

The above Figure 1 illustrates the churn-prediction model phases, such as pre-processing customer records (input data). This is followed through the feature extraction phase, wherein the input data's dominant and necessary features are extracted to frame the churn-prediction model. The next phase is the construction methods, utilizing distinct model classifiers. The cross-model validation is performed in the same phase. The prediction of the churns would be evaluated through different metrics, including variable significance reports and the accuracy of predictions. The last phase in the figure facilitates offering customer retention policies to CRM-customer relationship-management representative (Hartati, & Bijaksana, 2018). The present study depicts the review analysis of the existing churn prediction model using various approaches presented in different domains. The review analysis states that Deep learning and Ml methods can process huge amounts of customer data. Therefore, the technology aids in predicting the churn rate of customers. In a competitive industry, wherein indulging more competitors, the cost of customer acquisition would be higher. Hence the study plays a prominent role and is vital for all businesses to explore customer churn rates precisely in prior stages before they leave the company.

The existing churn prediction model depended on enterprise managers' decisions seems to be simplified in terms of inductive reasoning, and managers could arrange the churn prediction process from a present customer according to churned customers' characteristics. However, the experiences and outcomes from the conventional churn prediction model were also unreliable, specifically of a complicated issue, with nil guidance obtained through experience. The enterprise resources also seem to be limited, and these are invested first in winning the customers back with high churn possibility (Karvana, Yazid, Syalim, & Mursanto, 2019). If enterprises desire to recognize scientific customer churn prediction, they may adopt a few analyses on machines, mathematical algorithms, and tools to determine the association between customer churn and technical indicators. Those methods judge if the customer was churned and offer the customer's churn probability. The logistic regression technique exhibits better effects on prediction based on the level of significance of customer-churn parameters could be seen. In the existing approach, the logistic regression approach utilized in trend prediction within customer churn guides the enterprises to determine earlier customer churn winning signals and to identify the customer churn tendency. However, the review on churn prediction of previous articles does not focus on the churn prediction mechanisms in the different streams with distinct data mining and transfer learning approaches or techniques (Joel, & Srinath, 2021).

Since the customer competition seemed to be increasing aggressively, market saturation turned to elevate higher, and homogenization of services and products competition was intensifying more. Customer needs and development require the force operators to launch greater attractive personalized items or products. However, it could not alleviate the crucial higher churn rate situation. In this new competition face or the patterns, the enterprise competition had shifted gradually from obtaining products as core to taking the customer to be core. As a result, the core industrial competitiveness has turned toward managing the user's scale effects (Vo, Liu, Li, & Xu, 2021). While establishing competitiveness for the customer, the industries in different streams faced current customer churn issues. To decrease customer churn evolved, a focus on industries like the manufacturing sector, banking sector, telecom operators, and business domain.

The primary issue to get rectified in this work was to recognize the significance of the higher value of customer churn prediction method and different techniques in customer churn. Through assessing the characteristics and techniques with algorithms designed to predict the customer churn rate, it is important to perform a review assessment of all churn prediction model that suits the fitted industry. By the big data analysis, in a different sector, the historical customer information prediction was incorporated with different algorithms such as deep learning In accordance to the big data analysis, in a different sector, the historical customer's information prediction was incorporated with different algorithms such deep learning methods, various research was conducted for the prediction of customer churns, and it further digs out capable churned customers within the customer library. This kind of research aids the enterprises in directing the target win-back measures by potential churned customer characteristics.

1.1 Problem identification

Within the business environment, Customer attrition generally represents the customers or the users who leave or switch over from one business organization to another firm anonymously. Subscriber churn is also referred to as attrition. From the perception of machine learning, the prediction of churn seems as a supervised problem (labeled), as stated below. Provided the pre-defined Forecasting horizon, the goal is the prediction of future churners across the horizon. The prediction model needs to determine the possible churners before the customer leaves off the network (Srivastava, & Eachempati, 2021). The model ought to support a CRM stream to prevent subscribers who were likely to churn through formulating subscriber retention policies. Such policies need to do attraction towards likely churners and then retain them permanently. Hence the company's potential losses must be prevented at this stage. For instance, the problem input in a telecom organization consists of information on the previous calls of every mobile subscriber. All business data and personal data are managed by service-provider.

Additionally, in model training of machine-learning approach, labels were provided as churner lists as the outputs. Once the training phase of the model is completed, with precise accuracy and better estimation, the model should be capable of predicting the churners list gained from the real dataset. Therefore, the churn label should not be included in the training phase. The issue will be categorized as predictive modeling or mining in the knowledge discovery phase perception. To overcome the problems, a detailed discussion analysis needs to be organized.

1.2 Objectives

- To propound a detailed survey analysis of existing methods of churn-prediction models using various machine-learning approaches, Deep-learning methods, cloud-based churn-prediction models, feature-selection-based CCP methods, and Hybrid approaches, aiding in the development of a churn-prediction framework.
- To elucidate comparative analysis of different churn-prediction models and present the evaluation metric for each existing study.

1.3 Paper-organization

Section I discusses the introductory concepts of the churn prediction phase. Section II states the research method of the study. Section III discusses the fundamental review analysis, employing different learning methods to facilitate churn prediction. Section IV represents the comparative assessment part of different churn prediction frameworks. Section V provides the challenges inhibited by the existing studies. Section VI enumerates the conclusive analysis of related works. Finally, section VII propounds the conclusive part of the study.

2. Research methodology

The method to elucidate the review analysis of churn prediction was exhibited through method phases illustrated below. Then a detailed analysis of the different churn prediction researchers is presented**.**

2.1 Search keywords

Search terms are specifically given to scholars to build out different literature studies. The used search terms "churn prediction models," "different churn prediction using machine learning and deep learning approaches" prediction of churn in different industries or companies" were entered into the publication databases. The search key terms yielded the results by retrieving many researchers or articles. The sources chosen included literature surveys, research papers, and empirical studies.

2.2 Research questions

The selected papers, from this selection knowledge, are retrieved and evaluated critically based on three kinds of quality evaluation questions, stated as below.

QE1: Do the paper covers out relevant research work, and does it explores research topics comprehensively?

QE2: Does the paper offers clear implication with justifiable outcomes and their conclusions?

QE3: Do the articles, papers, or reviews provide future directions?

2.3 Article selection

The selection of the articles for this review analysis undergoes certain criteria to enter into the selection list. The researcher or the article that satisfies the below article selection criteria was finally taken for review analysis study.

Inclusion criteria

- It must consist of Meta-analyses.
- The paper must present research papers, a literature review, and a review paper with a defined search process, data extraction, and research question. The papers were included regardless of whether the review or paper was part of the main component or any articles.
- The entire research work must be associated with the main concept or theme of the paper, like churn prediction.

Exclusion criteria

The papers can be excluded if the articles, papers, or reviews if the records follow the below condition

- If the articles or papers were duplicate reports or papers of similar research studies.
- The informal literature reviews were excluded, with no defined research question, no defined search process, and nil defined data extraction phase.
- If the articles or records were not written in the English language.

3. Review analysis of various customer churn-prediction methods

The following section explicates the detailed review analysis of the existing churn-prediction model employed through different learning approaches.

3.1 Pre-processing, imbalance problem, and sampling-based churn prediction

Various Churn prediction models have been proposed in the recent decade depending on different machine-learning techniques and deep-learning algorithms. In the overall prediction phase, the machine-learning prediction model comprises of pre-processing phase, feature selection, Feature classification, and finally, the prediction stage. Pre-processing phase indulges in pruning and cleaning of data. Determining capable churning users or customers is achieved based on prior behaviors and past customer information. One such study, performing Churn-prediction through data-mining algorithms involving the K-means algorithm, is used for the pre-processing phase of the dataset. The attributes were extracted from the pre-processed input data using the mRMR-Maximum Redundancy and maximum-Relevance technique. The attributes are chosen with higher co-relation with the output class and less correlation among the attributes. Based on extracted features, customer churn prediction or separation was examined through the SVM-Support-vector machine technique, and PSO-Particle Swarm-optimization was utilized. The hyperparameters optimization of the SVM approach is achieved through PSO. However, in such feature selection techniques, local optimal solution issues need to be addressed with vast data feature extraction with high accuracy in feature classification (Ahmad, Jafar, & Aljoumaa, 2019). Generally, churn prediction has skewed dataset distribution, such that a single class consists of many instances counted compared to other classes. The class having fewer samples is referred to as the minority class, and the class with different relative instances is defined as the majority class. The imbalance ratio among classes depicts imbalance instances distribution within the dataset. In the churn-prediction model, the count of non-churners was higher than the churners count (Dwiyanti, & Ardiyanti, 2016).

However, the predictive model's performance was highly impacted when the real-world data set was imbalanced. The dataset is considered imbalanced if the sample size from a single class seems smaller or larger than the other remaining classes. The most generally utilized method for this is under-sampling to handle the imbalance issue in different domains. Therefore, the well-known sampling methods were reviewed, and a comparison of the different sampling methods' performance is performed, like the adaptive synthetic-sampling approach, (MTDF)mega trend diffusion-function, majority weighted minority-sampling method, immune centroids over-sampling method and couples top N-reverse k-nearest neighbor method (Amin et al., 2016).

Machine-learning implementation upon Telco-customer churn data, which consists of four consecutive phases, ought to proceed. The stages include Data-preparation, Pre-processing of data, Feature extraction step by using feature selection optimization approaches, feature classification phases using model classifiers, and finally, the prediction phases based on the classification results. Customer churn is a prominent and challenging factor for the telecommunication stream. Assessing if the churned customer is a big issue is highly required. Then, the organization's management would take essential strategies to reduce the churn rate and sustain the customer. The customer's churn data, categorized as churn APS-Atas-Permintaan Sendiri is considered imbalanced data; hence, this problem is another challenging activity within the machine-learning technique. The imbalance class could be handled in the churn-prediction model by integrating the RUS-Random User-Sampling method and the SMOTE-Synthetic Minority Over-sampling technique. The integrated method combines the bagging technique to gain better churn-prediction outputs. The research initially utilized this pre-processing data phase to clear out imbalanced data. Then, the imbalanced dataset is balanced out through pre-processing phase, and the features are extracted through sampling methods RUS and SMOTE techniques (Gui, 2017). After proceeding with feature selection and pre-processing techniques, the churn prediction model is built out. C4.5 method and bagging methods are applied in the research. The generally utilized techniques in dealing with imbalanced data within the churn prediction model are comprised of two types: resampling approaches and cost-sensitive learning techniques. The cost-sensitive learning techniques adjust the relative error costs in the model-training phase. The Resampling data perform well in handling imbalanced data and sorting out with balanced information before the training phases of the model (Nguyen, & Duong, 2021).

Reducing the attributes in imbalanced data may lead to a higher chance of information loss. Hence to avoid this kind of information loss, one approach is to assign the weights of the attributes through a domain expert. It is not alone expensive but also necessitates the human expert in such a field. Therefore, an efficient technique is required to assign suitable weights automatically with the involvement of a domain expert. The study, as an effort, presented a novel feature-weighting method. This technique applies the (GA) genetic algorithm for automatically assigning weights values to the attributes based on the Naïve-Bayes classification (Amin et al., 2019b).

Deciding the appropriate usage of resources to perform user requests within the cloud environment was not trivial. Hence effective resource prediction could play a significant role within the resource management sector. The model estimates the required resources properly. In such research, the ensemble CPU load-prediction method through Bayesian information selects the best constituent technique in every time slot based on the usage history of cloud resources. Further, two kinds of smooth filters were applied to reduce the negative effects of outliers in noticed data points. The framework contributing to churn prediction for resource management in cloud computing was presented, including a module in the prediction of resource usage high precisely. The research proposes a new prediction churn prediction algorithm using cloud resources based on resource usage history evaluation. The prediction algorithm was based on this history assessment. First, the usage history of cloud resources (Tofighy, Rahmanian, & Ghobaei-Arani, 2018) was analyzed for a specified time window with appropriate finite sizes. After this, outlier points in this window were discovered. In order to perform this, fitting algorithms were used in smoothening out the outliers and BIC in assessing the outlier's effects. After the data wrangling process, the resource usage history point sets out in the same approach. It aids in dealing with fluctuating impacts of the usage of cloud resources.

Another model in heterogeneous workloads was proposed in the execution of cloud-based applications using a workload clustering-based resource-provisioning mechanism. The mechanism used (BBO) Biogeography based-optimization method utilizing K-means clustering method to classify cloud workloads according to customer QoS-quality of service necessities. Besides this, the Bayesian learning method specifies the appropriate resource provisioning action. The Bayesian learning technique specified the resource provisioning task to satisfy those cloud-based applications' QoS demands. The outcomes of simulation provided by simulation. The simulation outcomes demonstrated that the solution reduced the delay of action, cost, the energy of consumption, and SLA-violation ratio in comparison with clustering-based resource provisioning workload mechanisms (Ghobaei-Arani, 2021).

Another research by Amin and Khan contributed to the formalization of customer-churn prediction, wherein the rough-set theory was utilized as one of the classifiers. The multi-class classifier is employed for examining the trade-off to choose the efficient classification model for customer-churn prediction. The experimental analysis explores the various rule-generation algorithm performance like the genetic and exhaustive. From outcomes, it is determined that this rough set is a multi-class classifier and a one-class classifier depending on GA, yields appropriate performance in comparison with other different rule-generation algorithms. Further to this, through employing the proposed method, like the rough-set multi-class classification, upon the public dataset, the outcomes depicted that this multi-class classifier offers highly accurate outcomes for multi-binary classification issues(Amin, Khan, Ali, & Anwar, 2014).

3.2 Feature selection-based churn prediction methods

Feature-selection techniques are the phase to eliminate the irrelevant attributes or features from the input dataset. The feature selection in the churn prediction model extracts relevant features while managing acceptable accurate classification outcomes. The extracted features play a significant role in impacting the resulting classification accuracy directly. One such research proposed a method comprising two categories of phases, selection of features and model classification based on selected features (Sivasankar, & Vijaya, 2019). The first phase utilizes the filter and wrapper methods for feature selection with the ROS-Random Oversampling technique. Through this method, attribute size could be reduced, and misclassification errors could be minimized. In the second model phase, extracted attributes are obtained as data inputs through classification methods such as K-Nearest Neighbors methods, DT-Decision-Tree technique, SVM technique, ANN-Artificial neural network, and NB-Naïve-Bayes method. The classification accuracy of the churn prediction relies on classifier optimizations and performance along with other phase potentials. However, the Imbalance of the class stands out as a significant characteristic of the customer churn-prediction method. Various standardized model classifier frequently tends to favor the majority class. The scenario led to poor performance classification of the minority class.

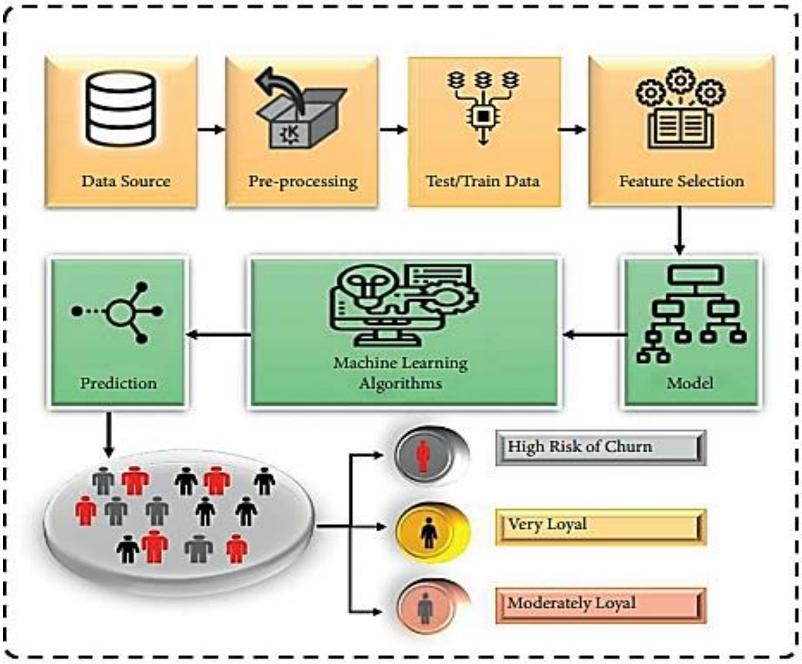

Figure 2 General CCP framework based on feature selection phase (Faritha Banu et al., 2022).

Artificial intelligence (AI) applications were utilized widely for business purposes. Research wherein the customer churn data is predicted through a feature selection-based model, as illustrated in Figure 2, is exposed. Feature Selection facilitates better data understanding ability. The feature selection-based technique also minimizes computation time and minimizes storage requirements. However, the constraint of feature-selection approaches is the different performance metrics in the case of a large volume of features. Some feature-selection models with model classification techniques for churn prediction may be limited to performing classification depending on features present in the corresponding data set only. One example is the multi-filter feature-selection-based classification model in determining software defect prediction. The framework utilized the ANN method. Oversampling techniques utilized inside the framework to assess the class imbalance impact problems on the performance of classification. However, several Machine learning classifiers should also be included in developing the prediction model to enhance the performance. The outcomes of the study imply that the feature-selection technique could choose specific features for training out the churn-prediction model and generate effective performance. The research also confirmed that the feature-selection process must be utilized for extracting relevant features for training the prediction model (Iqbal, & Aftab, 2020). This feature-selection method prediction model category also evidences that not all data features need to be relevant for prediction. Only the appropriate features could be identified and obtained from the feature-selection technique. The performance outcomes proved the necessity to build a higher-quality churn-prediction model by employing a feature-selection process. The research proposed the implementation of feature selection to retrieve relevant attributes and generates enhanced customer-churn prediction framework performance (Iqbal, & Aftab, 2020). Further, a classification algorithm is utilized for feature classification, using the Naïve-Bayes classifier. The framework provided the optimized better performance in churn prediction, employing SBFS-Sequential Backward-floating selection, feature selection method (with feature number), and SBS-Sequential backward selection (Iqbal, & Aftab, 2020).

Another method is the power-aware technique for deciding the allocation of (VM) virtual machines within physical resource data centers. Virtualization was used to be superior technology for power-aware VM allocation techniques. Since this allocation of VM is more tedious, the evolutional algorithms usages may solve the issue. The efficient micro-genetic algorithm was presented for selecting appropriate destinations among the VM machine's physical hosts. This sort of evaluation within the simulation environment depicted that this micro-genetic approach offers invaluable model improvement by the power-consumption rate compared with the conventional technique (Tarahomi, Izadi, & Ghobaei-Arani, 2021).

The cloud platform might face issues in under-provisioning or under-provisioning issues because of the late user arrival rate to those cloud applications differing from time to time if the feature of elasticity is not managed correctly. Hence it will necessitate the elasticity management issue for resources standing as a major problem in a cloud-computing environment. For such purpose, an elastic controller based on colored Petri-nets is proposed for the management of cloud infrastructures manually. The proposed algorithm efficacy is analyzed under three kinds of workloads. The simulation results implied that the elastic controller diminishes the response time of the resources from 4.80%, maximizes the resources utilization, and increases the elasticity range to 9.30% and 6.70% compared with conventional approaches. For this purpose, live migration of virtual machines and dynamic consolidation were applied (Shahidinejad, Ghobaei-Arani, & Esmaeili, 2020). Hence this optimization paves the way to determine churn prediction rate within cloud infrastructure management.

3.3 Machine learning methods for churn-prediction

Due to direct impacts on organization profits, specifically in the telecommunication sector, firms focus on different approaches for developing a way for customer churn prediction. To concentrate on this scenario, a predictive approach is employed such that larger telecommunication input data is utilized for model development, potential in prediction, explaining, and classifying customer churn issues. Machine learning techniques are utilized for feature classification, including the Random-forest method, XGBoost model classifier, logistic-regression model classifier, and SVM classifier. Additionally, the Ensemble technique is applied to enhance the classifier's performance. The results depicted which model features to select for handling customer churn (Chabumba, Jadhav, & Ajoodha, 2021). The research outcomes gained an accuracy rate of around 80.18 percent for the RF method, a 78.99 % prediction accuracy rate implemented with the XGBoost classifier, and 79.83 % in SVM methods. In addition, they retrieved 73.8 percent with a logistic-regression method. Among all the performance outcomes of machine-learning approaches, Random-forest techniques outperform other churn-prediction –models. However, also the research could be tested on massive data volume. One scenario that has a larger data volume is the telecom sector. Because of the high client base, larger data volumes were generated in the telecom sector daily (Ullah et al., 2019).

Customer Relationship-management and Business-analytics analyzers require the exact reasons for churn users and their behavioral patterns against the line of existing customer information (churn users). One of the research shows the design of a churn prediction model, utilizing clustering techniques and classification approaches to determine churn users (customers). The factors standing behind customer churn rate were all analyzed within the telecom sector. In such research, feature selection was made using a correlation attribute-ranking filter and information-gain methods. The proposed framework initially classifies churn customer information through classification algorithms in machine learning. From the experimental analysis, the Random-forest-algorithm revealed better performance with 88.7 percent precise classifying of data instances (features). Another important action of CRM in avoiding churners is the creation of efficient retention organizational policies. After model classification, the proposed framework segmented churning consumer data by categorizing churn customers in separate groups utilizing cosine-similarity to offer group-based retention formulations. The research also determines churn factors, which is essential to identify churn root causes. Through the knowledge of predominant churn factors obtained from customer information, CRM could enhance the organization's productivity range, formulate recommendations of appropriate promotions towards likely churn customer groups depending on behavioral patterns and enhance the company's marketing campaigns excessively. Further, the prediction model of machine-learning techniques could be expanded to explore customers' changing churn customer's behavioral-patters through Artificial-Intelligence for trend analysis and Prediction phase.

Similarly, an optimized churn-prediction model is not sufficient to handle with proactive-churners. This prediction model, which emerged as a better-integrated model having an effective retention strategy, is beneficial for the prediction of pro-active churners in day to day competitive market (Mishachandar, & Kumar, 2018). Hence, building out and designing such models, with opted prediction accuracy and a good precision rate, has become a tough constraint for many researchers. Hence to overcome such an aspect, the research assumed two machine-learning algorithms based on classification for churn prediction. Specifically, the Decision-tree algorithm and Naïve-Bayes classification are taken into consideration. A machine-learning algorithm with a big-data analytics software tool for churn prediction is implemented to generate an opted environment for prediction phases. Generally, businesses prefer churn prediction to bring out an effective retention strategy, standing out prevention approach, compared to prediction design. The prediction model could be evaluated with ML algorithms for performance assessment in the testing phase.

Further model efficiency is increased through this evaluation. However, the outcome accuracy rate of various customer-retention approaches could be attempted in the future. However, various conventional studies are tried out for performing churn prediction, utilizing machine-learning approaches. However, a multi-factor hybrid technique creates reason validation for efficient talent management and employee retention. The prediction model was not tuned to provide recommendations for talent management and employee retention. In such a scenario, a robust hybrid churn prediction model is built for any organization. The factors identifying churn rate are inspired by employee data-set from the source website Kaggle. The research method determines crucial factors which govern the churn rate of employees and recommends the organization HR manager the way to retain disgruntled and valuable workers. The comparison of churn-prediction performance, such as the Gradient-boost method, DNN-Deep neural-networks-based models, and random-forest models, was attempted, and their significance in transforming the turnover status within an organization is formulated. From experimental analysis, it is determined that DNN outperformed RF and gradient-boost model by accuracy parameter for prediction on FMCG-companies, aided for validation phase (Srivastava, & Eachempati, 2021).

In such a prediction model, among those behavioral variables, RFM-Recency, Frequency & Monetary variable stands out as the best effective predictors in distinguishing potential-churner(non-loyal) and loyal users. Along with those RFM parameters, distribution-time prizes count, returned items count, and purchased items count are added to RFM, generating RFMITSDP (Augmented RFM model) design. Hence the research proposed a new model for predicting and determining churners in commercial business (Khodabandehlou, & Rahman, 2017). In the model, three categories of supervised ML methods as Decision-tree algorithms, Random-forest method, and ANN methods, were used extensively for customer churn prediction. The capability of the ML methods in churn prediction was also evaluated. From the research analysis, it is observed that the RFMITSDP framework possesses a higher accuracy rate in churn prediction and performance efficiency in comparison with RFM. The highest variation among the two categories of model observed is 24%, implemented using the ANN-Multi-Layer perceptron technique. However, extending the current RFM framework could further maximize prediction accuracy-percentage of churn.

Similar to this, another primary motivation is the dire necessity of the business to retain the existing customers, coupled with higher costs related to acquiring new ones. The review analysis of that field explicated the effective, rule-based CCP methods based on (RST) Rough set theory for extracting the significant decision rules associated with non-churn and churn data. The proposed technique efficiently performs the churn classification from the non-churn customer data. The prediction is also the performance of that customer who would possibly churn or directly churn in the future. Extensive experiments of simulations were carried out to assess the proposed RST-based CPP method performance through four rule-generation mechanisms as (CA) covering algorithm, LEM2 algorithm, GA, and (EA) exhaustive algorithms (Amin et al., 2017).

Likewise, the comparative analysis outcomes are bought out through another study that compares MLB packages and Apache Spark-ML Techniques arranged by the training phase and accuracy parameter. The comparison of the performance of ML methods (decision tree) upon transaction dataset (bank customers) for churn prediction is assessed (Sayed, Abdel-Fattah, & Kholief, 2018). From the inferences of the study, it is clearly defined that MLiB-package with RDD-based API technique shows effective results in their model training time, which is caused by internal package transformation. The outcomes imply that the machine-learning churn-prediction framework could outperform other existing methods in terms of churn-prediction accuracy and more rapid and qualified outputs. The outcomes facilitated to get implemented in bank firms and other businesses handling numerous records and clients in churn probability prediction. However, also the research can be expanded for more comparative detailed studies, for more platforms and packages to check out the weighted, accurate churn-prediction model in various circumstances.

3.4 Deep learning algorithm for churn prediction

In many scenarios, the churn prediction framework development may be perceived as tedious, like dealing with a large volume of data, imbalanced data distribution, and higher dimensional features, specifically in telecom firms. The solution is presented in one research for such inherent churn-prediction issues through Ensemble-based Meta-classification and transfer-learning methods (deep-learning approaches). Initially, the first step applies Transfer-learning through fine-tuned Multiple Pre-trained Deep-CNN (convolutional-neural networks). Telecom datasets were generally defined in vector representation, wherein vector forms are transformed into 2-dimensional images since Deep CNN methods possess higher learning potentials upon those images (Ahmed et al., 2019). In the second phase, Churn prediction is accomplished by appending Deep CNN data with the original feature vector. In the final stage, final feature-vector forms are built for the higher-level Adaboost Based ensemble-classifier and GP-Genetic-Programming model classifier. Then, those classifiers engage in the prediction of churns. The performance of the proposed framework Transfer learning (TL) DeepE model was compared with existing techniques through tenfold cross-validation. Due to their challenging approach to churn-prediction activity, exploitation of GP-Adaboost meta-classification, Transfer-learning, and Deep-learning techniques, and the choice of chosen features seem to be highly effective in developing the churn-prediction model.

Owing to the feature selection or parameter selection, a non-sufficient empirical basement occurs for a better understanding of various hyperparameters' performance in the DNN model during their usage in the churn-prediction phase. Due to the above research gap, a derived empirically heuristic knowledge, aiding in hyper-parameters selection in churn-prediction model while in DNN utilization for modeling churn-prediction design, is still lacking. To rectify the issue, an experimental assessment of the effects of distinct hyper-parameters is evolved through a study (Domingos, Ojeme, & Daramola, 2021). In such research, the Feed-forward DNN method was utilized for Customer churn prediction in bank organizations from input data. It concentrates on DNN training using supervised learning methods to evaluate prediction accuracy before the hyper-parameters tuning phase. The tuning process of hyper-parameters is achieved through experimentation with multi-classifier parameters. However, in the future, more longitudinal studies are required to test experiment reproducibility with numerous collected data samples over a longer time obtained from several banks. The results should be combined with the generalization of inference to that bank sector for churn prediction. Another study limitation is the unbalanced dataset in their distribution. Even though Churn prediction obtained critical significance for existing users, it also facilitates future customers' trends predictions. Hence one such churn prediction implemented for Telco-dataset is employed through the Deep-learning technique. A Multi-layered neural network is developed for building out non-linear model classification. This churn-prediction technique works on support, context, usage, and customer features (Agrawal, Das, Gaikwad, & Dhage, 2018). The churn possibility and the determined churn factors were estimated. The model trained was then employed on features' final weight values and does the estimation of churn possibility for the concerned customer. Through this Deep-learning churn-prediction model, 80.10 % of prediction accuracy is attained. Since the churn-prediction model provided churn factors, it could be utilized by other organizations to assess the reason for those churn factors. It aids the organization in formulating appropriate actions or steps for eliminating churn factors.

The churn factors are used to maintain customer retention within the organization. This retention sustainability in the marketing field seems crucial for gaining a reduced cost rate through retaining high profits from existing customers (long-term) and retaining temporary customers. The conventional churn prediction was achieved through classification methods, including discriminant analysis, decision-tree methods, and neural networks with exclusive engineered features throughout the customer's lifetime. However, the behavior of customers differs in different lifecycle phases. The validity and reliability of churn-prediction model outcomes depend on simple consolidation feature values at all lifetime levels. Hence, it is crucial to incorporate time-correlation attributes or features in simultaneous classification model implementation (Sung, Higgins, Zhang, & Choe, 2017). Breaking every time-correlation feature into multiple phases could depict the customer behavior changes in a detailed phase. To focus on this concept, Deep-learning algorithms are leveraged for churn prediction with Time-correlated features in one structured approach. The raw input-data usage from the model could result in robust performance because ConvNets discovers novel features in higher-dimensional inputs. The work studies implementing Deep-learning algorithms for churn prediction with the concept of time-correlation attributes or features within a cloud-computing platform.

Similarly, the Deep Backpropogation-ANN method, employed by another researcher Wael, Ahmed, and Subramanian, utilizes two methods for feature selection, such as the Lasso regression and variance thresholding method. Additionally, the model gets strengthened through an earlier stopping approach to stop training to prevent overfitting issues. Then study deliberated comparative analysis in its efficiency to reduce the overfitting issue between activity-regularization strategies and drop-out strategies for two other datasets, such as Cell2Cell and IBM Telco. The outcomes delineated that implemented model outperforms well for feature selection with lasso-regression and prior stopping method for picking out epochs and larger count of neurons to input-hidden layers for reducing overfitting issues in both the datasets (Wael Fujo, Subramanian, & Ahmad Khder, 2022).

In another research, Elnasir and Ebrah assessed two datasets, Cell2Cell and IBM Watson, which comprise nearly 57 attributes and 71047 observations visualized through orange software. For such churn prediction, from these attributes, three different predictive-model are used decision tree, SVM, and Naïve bayes method, through Matlab. The study focused on determining the best accurate-churn prediction model, specifically in the telecom sector, and determining the most significant that makes the customers churn data. The model's performance was measured through the area under the curve(AOC), wherein the best values of AUC are 0.82 for the IBM dataset and a score value of 0.99 for the Cell2Cell dataset (Ebrah, & Elnasir, 2019). Another research by Ahmed uses ensemble-based transfer learning meta-classification is applied. The Telecom dataset was converted to two-dimensional images as this Deep CNN possesses high learning bulk on the pictures. The predictions obtained from Deep CNN were enriched to initial features of the telecom dataset and then utilized for setting the final features for Adaboost and high-level Genetic-programming utilizing an ensemble classifier. The resulting assessment was performed on the Cell2cell and orange datasets. The inference brought the accuracy and AUC score value. The method suggested could be potentially efficient in addressing the complications and concerns of the telecom sector (Ahmed et al., 2019). Owing to this, another technique that utilizes DL approaches removed manual feature engineering. For this perspective, three DNN architecture has been developed, and the corresponding CCP method is designed using cell2cell and crowd analytics dataset. The comparison is performed among the dataset with other models as well. The efficiency of the transfer-learning (TL) Deep extraction method made compared with other conventional methods for Cell2cell and orange dataset through the usage of 10-fold cross-validation (Umayaparvathi, & Iyakutti, 2017). the accuracy of the prediction gained is 68.20% and 75.40 % through conducting experiments on cell2cell dataset and then orange data-set, while it also obtained AUC score of 0.74 and 0.83 (Ahmed et al., 2019).

3.5 Cloud environment churn prediction methods

CCP-customer churn prediction stands out as a tedious task for many decision-makers and the ML community due to one fact, where non-churn users and churn users have similar features. Similarly, the widespread usage of the cloud-computing platform and IoT Internet of-things enable it to be feasible in the customer data collection phase in performing CCP. Therefore, the CCP model is presented through ML algorithms within a cloud-computing platform. The proposed framework of the CCP model was implemented in three different phases: data-collection, P-AGBPNN pre-processing technique, and AGBPNN-Adaptive Gain with Back-propagation Neural-network. In the initial stages, customer data collection is adopted using different IoT devices such as wearables, smartphones, laptops, etc. Then, the collected data were sent to the targeted CDS Cloud-data server through IoT gadgets. After the process, pre-processing stage occurs such that missing dataset values has been inputted efficiently. Finally, the p-AGBPNN framework was examined utilizing a benchmark dataset. Hence this AGBPNN model could be a suitable CCP tool within a cloud environment (Jeyakarthic, & Venkatesh, 2020). The P-AGBPNN model outperformed other models, with outstanding outcomes having an accuracy rate of churn prediction (91.8%), F1-score parameter percentage (95.2%), and specificity value (70.50%). However, the presented model performance could be improved further by using clustering techniques and optimized feature-selection methods.

Like machine-learning techniques, the Deep-learning model was applied for effective CCP. The study presented the Adagrad Optimizer method with the EHO-Elephant herding-optimization-based Bi-LSTM model for the customer churn-prediction process within IoT enabled cloud-computing platform. The framework initially acquires customer data using smart-watch, smartphones, laptops, etc. Then the collected data were then classified through the Bi-LSTM model, determining customers as non-churners or churners (Venkatesh & Jeyakarthic, 2020b). The Bi-LSTM model efficiently was maximized by using hyperparameter tuning approaches such as the Adagrad optimizer and EHO method for selecting the parameter values optimally. The parameter factors in the model evaluation phase are a count of hidden layers, epochs, and the learning rate. The simulation outcomes depicted the model supremacy wherein the AG-EHO Bi-LSTM method exhibits superior performance compared to other techniques (having an accuracy rate of 98.36%). According to this model, managing churns within the organization seems like a crucial challenge for VAS providers and mobile operators. For a better understanding of churn patterns and user behavior, and effective action or have to formulate retentions for user's retentions (churns conversion). The churn-prediction model could assist the aspect in estimating the churn users periodically. Various machine-learning methods are utilized for user churn estimation.

Additionally, Decision-tree algorithms, Random-forest methods, Gradient boosting methods, and Logistic-regression techniques were utilized for comparative assessment of customer data with their parameter values such as calls-length, service calls of customers, day's call minutes, and voice-call messages. The outcomes of the predictive-model comparative evaluation explicated that the Gradient boosting technique and Random-forest machine learning method exhibited better prediction capability in churn detection (Ibrahim, Aborizka, & Maghraby, 2018). Machine-learning and analytical techniques were employed for analyzing and managing the data or the logs of users. Such information may be utilized as information, aiding in cloud-resources management for reaching out effective resource utilization and generating more revenue from those organization service-users. Machine-learning techniques may utilize the information to predict changing user behavior based on past charging history. The historical data will be clubbed into a cluster set, depending on the charging logs' similarity within the self-adaptive prediction model. The model learns to form current and old charging-transaction also. Hence this user behavior (represented as feature vectors) aids in getting fed into a machine-learning classifier for the model's training. The trained model further provides the churn prediction outcomes to the organization. The varying user behavior is handled in the study and how it is utilized for the churn-prediction phase within the cloud environment.

According to this aspect, the organization should focus on a highly customer-oriented approach, rather than a product-oriented one. The organization employs CRM to achieve customer-oriented outputs since customers are valuable assets. Therefore, the organization should have to achieve retaining profitable consumers. For necessity, data mining methods were applied to determine the deliberate churn of users. The new churn-prediction model is implemented to enhance the existing boosted-tree method's performance (Kaur, & Vashisht, 2015). The technique was employed on a cloud platform wherein many cloud-computing features, including used resources, pay, security, virtualized computer dynamic-provisioning technique, and low cost, etc., were presented for effectively mining out the vast volume of data. The framework was utilized for churn attrition of clients. Hence, as a result, proactive measures would be considered by the organization to prevent churn.

The outputs exposed that the proposed framework performs well with a high churn prediction accuracy rate compared to other conventional churn-prediction techniques. However, the model needs to be developed in such an approach that it will work parallel in the cloud environment for multiple-machine to minimize processing time. Besides the increase in computational and operational cost, the higher power of consumption decreases the lifetime and reliability of hardware resources. Furthermore, environmental challenges and gases emission were taken into consideration. Hence it is mandatory to develop and use power-effective mechanisms in the environmental sector (Ghobaei-Arani, Shamsi, & Rahmanian, 2017). The network management and storage devices' power consumption were maintained effectively through previous researchers. However, also it has been claimed that between resources of servers, including RAM, network instruments, storage resources, and CPU, this will consume a higher power portion, which turns out the churn customers. Hence a model must rectify this issue for resources provisioning solution, accomplished through the meta-heuristic method in assessing cloud workloads.

The clustering method utilized the integration of fuzzy C-means and genetic algorithm in determining similar network clusters by the QoS parameters of users. For these benefits, the grey-wolf optimizer is utilized in making proper scaling-decision for offering the cloud resources to serve cloud workloads(Ghobaei-Arani & Shahidinejad, 2021). The comprehensive framework showed interaction among cloud providers, resource provisioning brokers, and users within the workload clustering process. The cloud workloads submitted to cloud providers were heterogeneous to cloud work-loads analysis, quality parameters, and cloud workload management for satisfying Qos requirements (Shahidinejad, Ghobaei-Arani, & Masdari, 2021). A hybrid approach to resource provisioning problems within the cloud environment was presented. For such purposes, K-means and (ICA) imperialist competition – algorithms are used to cluster the submitted workload by those end-users. The scaling decisions using a decision algorithm for effective resource-provisioning are proposed in the research.

Various enterprises might provide different web services having different potentials. Wherein that web service could be integrated to offer the complete software application functionality that meets the request of users. Hence this approach may aid in reducing the churn rate of customers. Hence the composition of services to be NP-hard optimization issues in combining the heterogeneous and distributed web services were introduced. The linear programming towards the web-service-composition issue, referred to as LP-WSC, was selected to choose the effective service in each request in distributed geographical cloud environment to enhance QoS criteria (Ghobaei-Arani & Souri, 2019). The elasticity resource analytical design could play a prominent role in managing cloud resources in buffer management. The (ControCity) Controlling elasticity model is proposed to control the resource's elasticity by elasticity management and buffer management. The components are called elasticity manager and buffer manager in middleware, an application layer. The buffer management controls the user's request input queue impacting the churn rate. The elasticity management controlled the cloud platform's elasticity and placed it within buffer control. The control mechanism was implemented through learning automata, offering the solution to control the cloud platform's elasticity. The experimental outcomes imply that this ControCity diminishes the response time through 3.70% and it maximizes the elasticity and utilization of resources by 5.40% and 8.40%, respectively (Ghobaei-Arani, Souri, Baker, & Hussien, 2019).

3.6 Meta-heuristic methods for CCP

Customer churn prediction is a crucial process in business decision-making that properly determines churn users. Non-churn users and churn users do have resembling attributes or features. Such features are extracted and trained in a neural network to predict churn users (machine learning approaches). Metaheuristic methods are used for feature extraction and classification in the prediction model. A dynamic customer churn-prediction strategy is designed through text analytics with a meta-heuristic approach within business intelligence (CCPBI-TAMO). In addition, the chaotic-pigeon-inspired optimization method (CPIO-FS) is utilized in the feature-selection process and minimizes the complexity of computation. The LSTM model with the SAE-Stacked Auto-encoder technique is employed for feature classification. In the LSTM-SAE method, the SAE contribution is identifying compact data features combined with LSTM classification capability. The hyperparameter tuning process occurs through the SFO-Sunflower optimization method, enhancing the performance of CCP (Pustokhina et al., 2021). The experimental outcomes reveal the proposed framework's superior performance compared with other conventional techniques exhibiting 95.5 % of prediction accuracy. The framework of churn prediction could be recognized in a cloud-based environment and IoT platform for churn prediction in real-time applications, including travel planning, E-commerce, telecommunication, etc.

The metaheuristic methods were implemented to find the optimal solution from the set of functions. An optimal metaheuristic-based FS model is applied in churn prediction model development with one Gradient-Boosting involved in the feature-classification process in CCP. The model consists of four major phases: data acquisition, Pre-processing technique, feature-extraction process, and classification method (Venkatesh, & Jeyakarthic, 2020a). The data gathered were accomplished through IoT gadgets (wearable, Smartphones, laptops, etc.), and such sensed data were transmitted to the CDS cloud data server through IoT gadgets. Then the data gathered were then undergone pre-processing, and missing values were inputted after the phase. Metaheuristic approach, the ACO-Ant-Colony optimization algorithm is placed as a feature-selector for choosing out the optimal-features set. Then GBT algorithm was integrated into the feature-classification technique for classifying data into non-churn and churn users. However, clustering methods can be added to improvise the classification outcomes for the churn prediction process.

Similarly to this ACO algorithm, PSO-Particle swarm-optimization algorithm is utilized for carrying out analysis for the churn-predictor. The pre-processing phase was represented as a variant incorporating the feature of the extraction process. The local search algorithm is the Simulated-annealing method, performing better in imbalanced data scenarios. Further, scalability related to the search algorithms (Vijaya, & Sivasankar, 2019) brings out the best user(candidate) within the churn-prediction process. But also, in the future, the algorithm can be enhanced by extending the research for more effective and rapid churn prediction performance.

3.7 Hybrid churn prediction methods

Generally, Churn-datasets suffer from imbalanced data distribution, and if the model classifiers undergo a training phase with such imbalanced data sets, it might create bias in the majority class. As s result, the performance of classification also gets degraded. To deal with imbalanced data distribution, researchers focused on pattern recognition. Hybrid techniques (SOS-BUS) integrate under-sampling and over-sampling techniques (Salunkhe, & Mali, 2018). The over-sampling method is performed by utilizing SMOTE method for under-sampling, rectifying the issue of random-under sampling methods. The hybrid method concentrates on major class board line instances. Those instances were assumed as a significant data class. These instances or significant data features were crucial in finalizing our decision boundary among classes.

Similarly, another research utilizes machine-learning techniques in a hybrid model for churn prediction purposes. In such a model, the PSO technique, a nature-inspired optimization algorithm, is employed as integrated with hidden feed-forward neural networks in this hybrid design. The PSO optimization algorithm, implemented for weighing the customer's features and neural-network structure optimization, is carried out simultaneously. The hybrid model facilitates handling imbalanced data distribution by employing an advanced over-sampling technique in the training stage (Faris, 2018). The Random-weight neural network was chosen as a feature classifier for churn prediction since it exhibits various benefits compared to other categories of gradient-descent-based classical neural networks. The training phase produces an extremely rapid learning process with less human intervention in tuning out initial parameters. However, the proposed hybrid approach is efficient in its weighting approach, and classification could be investigated more based on other categories of under-sampling and over-sampling techniques.

Another hybrid approach, the LLM-Logit leaf model, is formulated for data classification. The idea beyond the LLM model is that those distinct models were constructed upon data segments instead of the entire data set, leading to prediction performance, and it has to maintain comprehensibility. The LLM comprises two phases, the segmentation process and the prediction phase (De Caigny, Coussement, & De Bock, 2018). In this segmentation phase, the customer segments were determined through decision rules. Then, in the second phase, a model is created for each leaf inside the tree. The hybrid techniques were benchmarked against logistic regression, logistic-model trees, decision trees, and random forests regarding comprehensibility and predictive performance. The hybrid approach aids analysts who face data having heterogeneity among the users or customers. Then it also benchmarked against effective algorithms for customer churn prediction. However, in the churn-prediction model, real-time churn prediction was required, and model training ought to be employed on the fly as well model training-time stands as a significant metric for consideration in performance improvement.

3.8. Comparative analysis of various customer churn prediction methods using different techniques

The article explains its significance in the business and industrial sectors for organizational performance and quality improvement. As shown in Table 1, the information gained from the review assessment of the various predictive model could be utilized for decision-making in customer retention management systems and the commercial enterprise sector. The churn prediction model seems important to high-education communities in predicting customer churn rates to impact revenues positively. The machine learning and DL methods, with high-accuracy prediction models, play a prominent role in revenue enhancement and keeping resource utilization by assessing the behavior of users based on charging logs.

Table 1 Comparative analysis of various customer churn prediction methods using different techniques

S. No.	Author	Description	Techniques used	Advantages/Disadvantages	Evaluation metrics	Prediction Accuracy
1	Amuda and Adeyemo (2019)	A predictive model for churn prediction in financial organization	MLP-Multi-layer perceptron of ANN	Beneficial in customer retention management in the decision-making phase. Other ANN architectures were not fitted in the model and employed.	Accuracy ROC Precision, Recall, F-measure	97.52% 89% 97.7 % 99.8 % 98.8 %
2	Dalmia, Nikil, and Kumar (2020)	Prominent Churn prediction model applied in the Bank sector.	Supervised machine learning, K-nearest neighbor-algorithm and XGBoost algorithm	The accurate prediction performance forecasts provide insights into formulating company strategies against churn reduction. Churn predictions can be performed only on an annual or quarterly basis	Accuracy Error rate Sensitivity Specificity	86.85% 13.15% 87.96% 79.03%
3	Sam, Suresh, Kanya, Tamilselvi and Tejasria (2021))	The churn prediction model performs prediction and does churn classification	Machine-learning technique-Random-forest, Neural networks, and SVM.	Random-forest, aiding loan prediction in banks, and churn prediction design beneficial for bank officials' operations.	Accuracy Recall precision	85% 84% 83%

S. No.	Author	Description	Techniques used	Advantages/Disadvantages	Evaluation metrics	Prediction Accuracy
		depending on criteria. Feature classification through Random-forest and training within Neural-network for dataset analysis and prediction summarisation		Less prediction accuracy and higher computation time complexity.		
4	Amin, et al. (2019a)	The effectiveness of churn-impacting factors, such as Upper-distance and lower distance among data samples, was determined through the proposed framework for CCP.	Machine learning	Various uncertain churn factors were not addressed for CCP in TCI(Telecommunication industry) Research should yield empirical outcomes on balanced dataset distribution with multi-base model classifiers.	Accuracy Recall F-Measure Precision	90% 87.50 % 83% 81 %
5	Pustokhina, Pustokhin, Nguyen, Elhoseny and Shankar (2021)	Presents enhanced SMOTE-Synthetic minority-over sampling method with OWELM-Optimal weighted-Extreme machine-learning for CCP	synthetic minority The oversampling technique (SMOTE) for unbalanced data and optimal weighted-extreme machine-learning technique.	SMOTE technique rectifies imbalanced data distribution issues and ROA characteristics in identifying the optimal sampling rate. Does not apply in real-time commercial regions.	Sensitivity Specificity Accuracy F1-Score	95% 93% 94% 92%
6	Britto & Gobinath (2021)	The Enhanced Churn-prediction model for bank sector evaluation addresses the comparative assessment with several Deep-learning techniques.	Hybrid Attention-based GRU Bi-LSTM Model	The performance efficiency of Bi-LSTM techniques was maximized through hyper-parameters tuning techniques to yield better performance. The learning rate of the proposed model is constrained.	Precision Recall F1-Score PR-AUC	83.74 % 99.04% 99.08% 23.21%
7	He, Xiong and Tsai (2020)	The study explores multiple	(ML techniques)	The standard Data pipeline for churn	AUC	score- 68%

S. No.	Author	Description	Techniques used	Advantages/Disadvantages	Evaluation metrics	Prediction Accuracy
		techniques for churn user classification using optimal models, including Gradient Boost and Extra-Tree model classifiers.	Extremely Randomized-Trees Classifier and Gradient Boosting Model	prediction could save resources, time, and human labor in better churn prediction. Even though the under-sampling method minimized data dimensions and boosted training, it lost information from most of the churn class.		
8	Alboukaey, Joukhadar and Ghneim (2020)	The study explicates two major churn prediction approaches, (RFM-based model) and then a Deep-learning model in churn prediction daily through multivariate time series.	Deep-learning techniques LSTM-based and CNN-based DL model for the automatic-feature selection phase.	Deep-learning models outperformed other statistical-based churn prediction-model with good performance. The study was not extended by adding out dummy-variable in RFM-based & LSTM models, denoting the user's response in learning out respective churn uplift models.	AUC F1-Score Log Loss Lift EMPC(expected maximum profit measure for churn)	90% 80% 80% 76.70% 80%

4. Critical analysis of the study

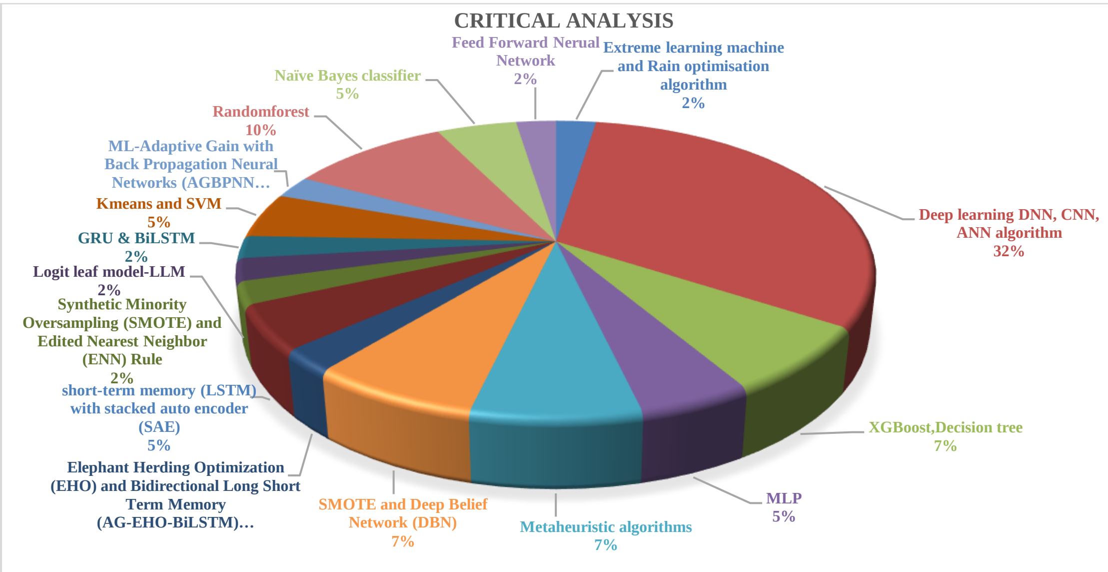

Figure 3 Critical analysis of the study in terms of deep analysis of algorithms

The above Figure 3 illustrates the critical analysis of the different algorithms utilized for churn prediction models. Different algorithms of deep learning, machine learning, and metaheuristics algorithms are employed in churn prediction. The figure clearly defines the major churn prediction models designed using deep learning methods such as ANN, CNN, GRU model, and Bi-LSTM contributing to 32% of articles. The accuracy of those DL approaches attains above 90% in its performance and exhibits higher precision, AUC, recall, and F1-score rate. On the other hand, the least number of algorithms were implemented from Decision tree algorithm, meta-heuristic approaches, (SMOTE)-Synthetic Minority Oversampling and (ENN)-Edited Nearest Neighbour Rule, LLM, ML-Adaptive Gain with Back Propagation Neural Networks (AGBPNN and powerful learning algorithm. Implementing a deep learning algorithm for churn prediction rate exhibited more efficiency to yield the benefits in retrieving the churned data forecasting. The critical analysis of the model implementing different approaches of ML, DL, metaheuristic approaches, Cloud computing churn prediction design, hybrid design churn prediction model, and feature selection approach utilized not alone in one sector but in other different business verticals, that makes the implementer go off with the effective churn prediction algorithm giving out best precise churned results, where there is higher customer participation. The critical analysis yields out to take off better plans in algorithmic selection for the churn prediction model for the researchers.

5. Challenges

- The adherence to organizational norms, particularly in the COVID pandemic situation, has specific impacts on the decision-making of employees to sustain or leave off from the organization. This is because that workplace transforms from a remote workstation to a work-from-home option, with limited working operations. Therefore, if social distancing norms are not followed strictly, the employees may leave the organization (which affects the churn rate). Hence, research must incorporate this retention decision factor with the study to investigate their impacts on employee churn rate.
- The effects of sinusoidal activation functions, such as spline function and sine function, can be assessed in the study. The churn-prediction model has to consider the impacts of various variants of ReLU upon DNN performance for the churning system modeling phase. The experimental analysis of the sinusoidal activation-function impacts upon deep neural architecture stands still as an active research area and demands more investigation (Domingos et al., 2021).
- A cognitive churn-analysis method was not addressed in any research. Hence the design of cognitive churn-prediction techniques needs to be propounded through Watson technology. Through this design, more diverse customer behavioral attributes (unmeasurable parameter values), including customer feedback, can be monitored. Similarly, the challenge for better prediction accuracy and performance capability depends on such churn analysis.
- The customer behavioral attributes have to be scored with values relying on the weight parameters of those churn impacts. Therefore, the score evaluation of customer behavior attributes needs to be planned in future research. In the research, the metric ought to be adhered to for evaluating which category of attributes possesses more impact on churn rate in the prediction model (Sung et al., 2017)

6. Conclusive representation and discussion of related works

Table 2 Advantages and disadvantages of various techniques

Author	Utilized Technique	Evaluation Tools	Performance metrics	Datasets	Advantages	Disadvantages
Amin et al. (2019c)	The model was devised for CCCP through data transformation methods.	Underlying classifiers performance through Naïve Bayes classifier, DNN, Gradient boosting techniques, SRI-Single rule induction, DNN,	POD-probability of detection - 0.92 POF-0.12% probability of false-alarm GM-Gmean - 0.29.	WCCP stands for Within Company churn-prediction and Cross-Company Churn Prediction	The Z-score-based transformation attains the best results in enhancing the prediction performance f of the classifier by reducing the error rate compared to	Does not have prominent outcomes in terms of general evaluation metrics, including POD, POF, Gmean, and AUC

Author	Utilized Technique	Evaluation Tools	Performance metrics	Datasets	Advantages	Disadvantages
		and KNN as baseline model classifiers through data transformation techniques.	AUC- area under the curve -0.52	CCCP Subject datasets	conventional techniques.	
Xiahou and Harada (2022)	An efficient loss prediction model based on K-means and SVM algorithmic technique through clustering segmentation loss	SVM and K-means algorithm for churn prediction	Accuracy rate-92% Recall rate - 97% Precision rate-86% AUC rate-92.70%	Balanced dataset.	The study provides high significance to establishing profitability for CRM-customer relationship management for ecommerce business in Business to customer approach.	Further study has limitations like the comparison and collection of behavioral data of customer need to be a performance for enhancing the generalization of the model.
Pustokhina et al. (2021)	To identify the Optimal Sampling – rate of SMOTE- synthetic minority over-sampling technique and tuning of parameters through WELM- weighted extreme ML.	(MOROA)- multi-objective rain optimization algorithm	Accuracy-0.94	Unbalanced dataset	Customer data indulges high-class labeling and data normalization.	Performance rate could be improvised further through different feature-selection methods.
Wu et al. (2022)	MBST- Multivariate Behavior Sequence Transformer has two attention mechanisms, exploring. Behavioral and temporal information.	Tree-based model classifier with attention mechanisms	F1-Score-82.72% AUC-93.75%	Tencent QQ browser-dataset	From the customer's behavioral data, the temporal characteristics were explored within browsers.	The model can be employed, and churn prediction can be performed from multiple datasets.
Banday and Khan (2021)	Deep learning ANN	Data-analytics tools – ANN training model	Precision-0.77 Recall- 0.51 F1-Score-0.61 Accuracy-87% ROC- 0.73	Kaggle dataset	The usage of the churn prediction model was credited to the CRM system.	The dataset size seems limited with churned customers that form smaller data fractions. In this scenario, bigger datasets could be utilized in the churn prediction model.
Zhang, Moro and Ramos (2022)	Logistic regression and Fisher discriminant equations-analysis	Statistical tool	Accuracy-93.94%	Telecom dataset	The outcomes of churn predicted data applicable for telecom industries to reduce the local host and monthly-fixed costs, increasing the	The model must inhibit the data from different operators, which might increase the model's reliability.

Author	Utilized Technique	Evaluation Tools	Performance metrics	Datasets	Advantages	Disadvantages
					customer retaining possibility in telecom customers.	
Xu, Ma and Kim (2021)	ensemble-learning technique	Stacking model inclusive of Decision-tree, Naïve-bayes ML, soft voting algorithms l	Accuracy-98.09% Precision-95% Recall-95.48% F1-Score-95.54%	open-source dataset	Feature construction expands the feature space and discovers the latent data acquired from implicit features.	The phase of feature construction is neglected in the model's development process. Hence better prediction system is required for effectiveness in retrieving churners.

As shown and summarized in Table 1 and Table 2, the study could be utilized to review studies of different churn prediction models appropriate for service fields. The researchers can acquire a high interest through incorporating fragmented churn researchers in academic and industrial fields such as IT, marketing, newspapers, psychology, business administration, and marketing and enumerate the difference in churn rate. Due to the distribution of fragmented past studies, there are certain difficulties for the researcher in launching new churn researchers and developing a prediction model. Hence to point out those issues, the survey paper elucidates the variations in the churn prediction definition, models, and different algorithms in its implementation in business fields, including insurance, telecommunication, marketing, newspaper publishing, and marketing, and makes a comparison of the difference in feature engineering and churn loss in different verticals. The study propounds the classification information based on taxonomy with highly detailed technologies upon churn in a broader range compared to past survey papers. The research could diminish the confusion regarding the churn criteria, which have been fragmented and used over multiple business or industrial fields. This paper could aid researchers in applying the churn ideas and concepts in prediction frameworks, and those different domains of churn prediction models vary in terms of frameworks and performance. The review propounded to contribute to the knowledge in the non-contractual business-to-business customer prediction domain and can be employed in the telecommunication sector. Specifically, the availability of common business data sources like invoice data if it is possible in devising the predictive models of churn rate through different historical data in various domains. This historical data in the complete review analysis is utilized to devise the features on model performance if the alternative churn definition would yield out the models that perform perfectly. Those approaches will be the foundation for discussing employee churn rate retention activities. The use cases of the churn prediction model were effective in B2B domains such as music and streaming services, Telecom organizations, media stream, and software as service providers in terms of improvising customer outreach, the value of customers, and customer services.

In some researchers, there may be a disproportionate structure of the dataset that has derogatory impacts on the typical ML and DL method's effectiveness in CCP. Hence it becomes crucial to use well-structured and clean datasets in CCP, as this performance relies on the dataset's characteristics. The potential in generalizing the experimental study seems significant to their validity. The count and the category of datasets used in the experimental process might impact the research findings' generalizability in various ways. As a result, the two frequently and major used datasets in CPP with various features were UCI and kaggle. Those datasets were accessible for free to the public and used widely for model training and to assess CCP models. Most of the researchers conducted in the review study handled many datasets such as Imbalanced datasets, public datasets, transaction datasets, telecom datasets, telco datasets, IBM telco datasets, Cell2cell, IBM Watson datasets, orange datasets, churn, CCCP datasets, Subject datasets, Tencent QQ browser-dataset, Kaggle, open-source dataset, and WCCP, relying on different ML and DL methods for churn prediction. Some researchers choose Cell2cell and telecom datasets while working with DL methods since this dataset consists of many record features compared to other datasets. Before depending on a large data volume, researchers must know that DL methods work efficiently with those datasets containing significant features toward the target. Because of this, certain researchers applied and recommended feature-powerful techniques, such as the XgBoost method, which does data visualization depending on the independent and dependent variables. This phase decides if the dataset comprises more relevant and significant features.

Moreover, DL methods and ML techniques also work well with balanced datasets. It seems tough to obtain a balanced dataset while handling the customer-churn issues as most customers were not churning. The inferences also recommended that the telecom industry acquire churned-customer data from archives or history for balancing datasets utilized for model training. Another approach, through utilizing a balanced technique like the ROS method. Some authors point his/her research to feature-based ML algorithms, showing efficient performance in churn prediction, including the SVM method, Fuzzy clusters, Multi-layer perceptrons, and KNN measures. XGBoost method also enhances the model's prediction accuracy, thus tuning the hyperparameters and boosting the existing performance of CCP.

Additionally, the DL CNN approach seems to be the add-on to ML techniques and could be efficient for future researchers. The Telecom dataset was generally large, hence those DLs must be effective in handling those big datasets with a large count of features. In the previous researchers, the minimum count of features obtained was 10, and the maximum level seems to be 722. Among the attributes, the most general attributes used by the researchers are the recharge attributes, duration attributes, demographic attributes, relational attributes, customer-care attributes, and service attributes. Hence for the model to be more effective, the CCP must be capable of dealing with small and large-feature sets.

The better feature extraction seems significant for minimizing the feature set, like in reduction in dimensionality. In this perspective, PCA is explicated to be efficient in dimensionality reduction, and the correlation matrix is another good example of reducing dimensionality. CNN is also exposed as a better DL method; the generally used methods are the decision tree and logistic regression methods. The accuracy metric is a primary performance measure, however, CCP cannot completely on the accuracy metric for performance analysis. Some researchers also added a few more performance-measure, such as confusion-matrix, recall, precision, and F1-score. CNN possesses inbuilt feature extraction that gives better outcomes and saves time. The confusion matrix is a significant performance measure utilized to measure performance. Some researchers attempted to employ all CCP to achieve current-state of art methods with more modification to develop effective hybrid models in the above-discussed dataset for prediction. The best effective technique could be clustered into a single group. It would also be presented as a single comparison factor for those industries in implementation in the targeted domain prominently in the communication sector.

7. Conclusion

The paper enumerates a details discussion of churn prediction techniques and their inferences, which are beneficial in various industries. The review analysis clearly illustrates the high significance of the churn-predictive model, specifically in the communication sector, for holding out customer retention and turning out high essentials for all categories of service industries. The existing studies of churn prediction models relied on machine learning approaches, Deep-learning algorithms, meta-heuristic optimization techniques, cloud-based churn-prediction models, hybrid approaches integrating two learning techniques, and feature-based churn-prediction models elucidated in the study. The existing research assessment guides the formulation of customer retention policies within an organization by implementing an enhanced churn-predictive model. Further, Hybrid churn-prediction approaches integrating two or more methods for feature extraction and feature-classification phase were also discussed in the study. The comparative analysis of different machine-learning and DL methods contributing to the churn prediction model were propounded in the study as well. The prediction accuracies corresponding to each model classifier of the churn-prediction model were defined in the assessment. The research gaps determined in the existing churn prediction model were enumerated. The study outlines the dimensions utilized for churn prediction, and present trends were explored. However, in the future, an enhanced multi-dimensional customer data set can be employed in churn prediction techniques, and decision-based analysis can be formulated to attain high prediction accuracy.

The enhancement of the original dataset and large dataset with tie features and centrality seems to be the key to improvising the churn prediction performance. The review assessment also recommends utilizing PCA for measuring the feature contribution in the churn model seems to attain success and validity. Similarly, a better dataset with more features with deep learning of features shall be utilized for the DL method for accurate prediction instead of using the same data sets category.

8. Declaration

- **Conflict of Interest**: The author reports that there is no conflict of Interest
- **Funding**: None
- **Acknowledgment**: None

9. References

Agrawal, S., Das, A., Gaikwad, A., & Dhage, S. (2018, July). Customer churn prediction modelling based on behavioural patterns analysis using deep learning. In *2018 International conference on smart computing and electronic enterprise (ICSCEE)* (pp. 1-6). IEEE. https://doi.org/10.1109/ICSCEE.2018.8538420

Ahmad, A. K., Jafar, A., & Aljoumaa, K. (2019). Customer churn prediction in telecom using machine learning in big data platform. *Journal of Big Data*, *6*(1), 1-24. DOI: https://doi.org/10.1186/s40537-019-0191-6

Ahmed, U., Khan, A., Khan, S. H., Basit, A., Haq, I. U., & Lee, Y. S. (2019). Transfer learning and meta classification based deep churn prediction system for telecom industry. *arXiv preprint arXiv:1901.06091*. https://doi.org/10.48550/arXiv.1901.06091

Alboukaey, N., Joukhadar, A., & Ghneim, N. (2020). Dynamic behavior based churn prediction in mobile telecom. *Expert Systems with Applications*, *162*, 113779. https://doi.org/10.1016/j.eswa.2020.113779

Amin, A., Al-Obeidat, F., Shah, B., Adnan, A., Loo, J., & Anwar, S. (2019a). Customer churn prediction in telecommunication industry using data certainty. *Journal of Business Research*, *94*, 290-301. DOI: https://doi.org/10.1016/j.jbusres.2018.03.003

Amin, A., Anwar, S., Adnan, A., Nawaz, M., Alawfi, K., Hussain, A., & Huang, K. (2017). Customer churn prediction in the telecommunication sector using a rough set approach. *Neurocomputing*, *237*, 242-254. https://doi.org/10.1016/j.neucom.2016.12.009

Amin, A., Anwar, S., Adnan, A., Nawaz, M., Howard, N., Qadir, J., ... & Hussain, A. (2016). Comparing oversampling techniques to handle the class imbalance problem: A customer churn prediction case study. *IEEE Access*, *4*, 7940-7957. https://doi.org/10.1109/ACCESS.2016.2619719

Amin, A., Khan, C., Ali, I., & Anwar, S. (2014). Customer churn prediction in telecommunication industry: with and without counter-example. In *Mexican international conference on artificial intelligence* (pp. 206-218). Springer, Cham. https://doi.org/10.1007/978-3-319-13650-9_19

Amin, A., Shah, B., Abbas, A., Anwar, S., Alfandi, O., & Moreira, F. (2019b). Features weight estimation using a genetic algorithm for customer churn prediction in the telecom sector. In *World conference on information systems and technologies* (pp. 483-491). Springer, Cham. https://doi.org/10.1007/978-3-030-16184-2_46

Amin, A., Shah, B., Khattak, A. M., Moreira, F. J. L., Ali, G., Rocha, A., & Anwar, S. (2019c). Cross-company customer churn prediction in telecommunication: A comparison of data transformation methods. *International Journal of Information Management*, *46*, 304-319. DOI: https://doi.org/10.1016/j.ijinfomgt.2018.08.015

Amuda, K. A., & Adeyemo, A. B. (2019). Customers churn prediction in financial institution using artificial neural network. *arXiv preprint arXiv:1912.11346*.

Banday, S. A., & Khan, S. (2021). Evaluation Study of Churn Prediction Models for Business Intelligence. *Big Data Analytics* (pp. 201-213). Auerbach Publications.

Britto, M. M. J., & Gobinath, D. R. (2021). Improved Churn Prediction Model In Banking Industry And Comparison Of Deep Learning Algorithms. *International Journal of Aquatic Science, 12*(2), 2521-2529.

Chabumba, D. R., Jadhav, A., & Ajoodha, R. (2021). Predicting telecommunication customer churn using machine learning techniques. In *Interdisciplinary Research in Technology and Management* (pp. 625-636): CRC Press. https://doi.org/10.1201/9781003202240-98

Dalmia, H., Nikil, C. V., & Kumar, S. (2020). Churning of Bank Customers Using Supervised Learning *Innovations in Electronics and Communication Engineering* (pp. 681-691): Springer, Singapore. https://doi.org/10.1007/978-981-15-3172-9_64

De Caigny, A., Coussement, K., & De Bock, K. W. (2018). A new hybrid classification algorithm for customer churn prediction based on logistic regression and decision trees. *European Journal of Operational Research, 269*(2), 760-772. https://doi.org/10.1016/j.ejor.2018.02.009

Domingos, E., Ojeme, B., & Daramola, O. (2021). Experimental Analysis of Hyperparameters for Deep Learning-Based Churn Prediction in the Banking Sector. *Computation, 9*(3), 34. https://doi.org/10.3390/computation9030034

Dwiyanti, E., & Ardiyanti, A. (2017). Handling imbalanced data in churn prediction using rusboost and feature selection (case study: Pt. telekomunikasi indonesia regional 7). In *Recent Advances on Soft Computing and Data Mining: The Second International Conference on Soft Computing and Data Mining (SCDM-2016), Bandung, Indonesia, August 18-20, 2016 Proceedings Second* (pp. 376-385). Springer International Publishing. https://doi.org/10.1007/978-3-319-51281-5_38

Ebrah, K., & Elnasir, S. (2019). Churn prediction using machine learning and recommendations plans for telecoms. *Journal of Computer and Communications*, *7*(11), 33-53. https://doi.org/10.4236/jcc.2019.711003

Faris, H. (2018). A hybrid swarm intelligent neural network model for customer churn prediction and identifying the influencing factors. *Information*, *9*(11), 288. https://doi.org/10.3390/info9110288

Faritha Banu, J., Neelakandan, S., Geetha, B., Selvalakshmi, V., Umadevi, A., & Martinson, E. O. (2022). Artificial Intelligence Based Customer Churn Prediction Model for Business Markets. *Computational Intelligence and Neuroscience*, *2022*. https://doi.org/10.1155/2022/1703696

Ghobaei-Arani, M. (2021). A workload clustering based resource provisioning mechanism using Biogeography based optimization technique in the cloud based systems. *Soft Computing*, *25*(5), 3813-3830. DOI: https://doi.org/10.1007/s00500-020-05409-2

Ghobaei-Arani, M., & Shahidinejad, A. (2021). An efficient resource provisioning approach for analyzing cloud workloads: a metaheuristic-based clustering approach. *The Journal of Supercomputing*, *77*(1), 711-750. DOI: https://doi.org/10.1007/s11227-020-03296-w

Ghobaei-Arani, M., Shamsi, M., & Rahmanian, A. A. (2017). An efficient approach for improving virtual machine placement in cloud computing environment. *Journal of Experimental & Theoretical Artificial Intelligence*, *29*(6), 1149-1171. DOI: https://doi.org/10.1080/0952813X.2017.1310308

Ghobaei-Arani, M., & Souri, A. (2019). LP-WSC: a linear programming approach for web service composition in geographically distributed cloud environments. *The Journal of Supercomputing*, *75*(5), 2603-2628. DOI: https://doi.org/10.1007/s11227-018-2656-3

Ghobaei-Arani, M., Souri, A., Baker, T., & Hussien, A. (2019). ControCity: an autonomous approach for controlling elasticity using buffer Management in

Cloud Computing Environment. *IEEE Access*, *7*, 106912-106924. https://doi.org/10.1109/ACCESS.2019.2932462

Gui, C. (2017). Analysis of imbalanced data set problem: The case of churn prediction for telecommunication. *Artificial Intelligence Research*, *6*(2), 93-. https://doi.org/10.5430/air.v6n2p93

Hartati, E. P., & Bijaksana, M. A. (2018). Handling imbalance data in churn prediction using combined SMOTE and RUS with bagging method. In *Journal of Physics: Conference Series* (Vol. 971, No. 1, p. 012007). IOP Publishing. https://doi.org/10.1088/1742-6596/971/1/012007

He, Y., Xiong, Y., & Tsai, Y. (2020). Machine learning based approaches to predict customer churn for an insurance company. In *2020 Systems and Information Engineering Design Symposium (SIEDS)* (pp. 1-6). IEEE. https://doi.org/10.1109/SIEDS49339.2020.9106691

Ibrahim, K., Aborizka, M., & Maghraby, F. (2018). Prediction of users charging time in cloud environment using machine learning. *International Journal of Intelligent Computing and Information Sciences*, *18*(2), 39-57. https://doi.org/10.21608/ijicis.2018.30121

Iqbal, A., & Aftab, S. (2020). A Classification Framework for Software Defect Prediction Using Multi-filter Feature Selection Technique and MLP. *International Journal of Modern Education & Computer Science, 12*(1), 18-25. https://doi.org/10.5815/ijmecs.2020.01.03

Jain, H., Khunteta, A., & Srivastava, S. (2021). Telecom churn prediction and used techniques, datasets and performance measures: a review. *Telecommunication Systems*, *76*(4), 613-630. https://doi.org/10.1007/s11235-020-00727-0

Jeyakarthic, M., & Venkatesh, S. (2020). An Effective Customer Churn Prediction Model using Adaptive Gain with Back Propagation Neural Network in Cloud Computing Environment. *Journal of Research on the Lepidoptera*, *51*(1), 386-399. https://doi.org/10.36872/LEPI/V51I1/301034

Joel, M. R., & Srinath, M. (2021). Optimizing profit by retaining customers using machine learning techniques. *J. Sci. Trans. Environ. Technov*, *14*(4), 193-198. https://doi.org/10.56343/STET.116.014.004.007

Karvana, K. G. M., Yazid, S., Syalim, A., & Mursanto, P. (2019, October). Customer churn analysis and prediction using data mining models in banking industry. In *2019 International Workshop on Big Data and Information Security (IWBIS)* (pp. 33-38). IEEE. DOI: 10.1109/IWBIS.2019.8935884

Kaur, K., & Vashisht, S. (2015). A novel approach for providing the customer churn prediction model using enhanced boosted trees technique in cloud computing. *International Journal of Computer Applications, 114*(7), 1-7. https://doi.org/10.5120/19987-6449

Khodabandehlou, S., & Rahman, M. Z. (2017). Comparison of supervised machine learning techniques for customer churn prediction based on analysis of customer behavior. *Journal of Systems and Information Technology*, *19*(1/2), 65-93. https://doi.org/10.1108/JSIT-10-2016-0061

Mishachandar, B., & Kumar, K. A. (2018). Predicting customer churn using targeted proactive retention. *International Journal of Engineering & Technology, 7*(2.27), 69. https://doi.org/10.14419/ijet.v7i2.27.10180

Nguyen, N. N., & Duong, A. T. (2021). Comparison of Two Main Approaches for Handling Imbalanced Data in Churn Prediction Problem. *Journal of advances in information technology*, *12*(1), 1-7. https://doi.org/10.12720/jait.12.1.29-35

Pustokhina, I. V., Pustokhin, D. A., Aswathy, R. H., Jayasankar, T., Jeyalakshmi, C., Díaz, V. G., & Shankar, K. (2021a). Dynamic customer churn prediction strategy for business intelligence using text analytics with evolutionary optimization algorithms. *Information Processing & Management*, *58*(6), 102706. https://doi.org/10.1016/j.ipm.2021.102706

Pustokhina, I. V., Pustokhin, D. A., Nguyen, P. T., Elhoseny, M., & Shankar, K. (2021). Multi-objective rain optimization algorithm with WELM model for customer churn prediction in telecommunication sector. *Complex & Intelligent Systems*, 1-13. DOI: https://doi.org/10.1007/s40747-021-00353-6

Salunkhe, U. R., & Mali, S. N. (2018). A hybrid approach for class imbalance problem in customer churn prediction: A novel extension to under-sampling. *International Journal of Intelligent Systems and Applications, 10*(5), 71. DOI: 10.5815/ijisa.2018.05.08

Sam, D., Suresh, K. C., Kanya, N., Tamilselvi, C., & Tejasria, M. V. S. L. (2021). An Improved Bank Customer Churn and Loan Prediction Model using Supervised Machine Learning Approach. *Turkish Journal of Physiotherapy and Rehabilitation, 32*, 3.

Sayed, H., Abdel-Fattah, M. A., & Kholief, S. (2018). Predicting potential banking customer churn using apache spark ML and MLlib packages: a comparative study. *IJACSA) International Journal of Advanced Computer Science and Applications*, *9*(11), 674-677. https://doi.org/10.14569/IJACSA.2018.091196

Shahidinejad, A., Ghobaei-Arani, M., & Esmaeili, L. (2020). An elastic controller using Colored Petri Nets in cloud computing environment. *Cluster Computing, 23*(2), 1045-1071. https://doi.org/10.1007/s10586-019-02972-8

Shahidinejad, A., Ghobaei-Arani, M., & Masdari, M. (2021). Resource provisioning using workload clustering in cloud computing environment: a hybrid approach. *Cluster Computing, 24*(1), 319-342. https://doi.org/10.1007/s10586-020-03107-0

Sivasankar, E., & Vijaya, J. (2019). A study of feature selection techniques for predicting customer retention in telecommunication sector. *International Journal of Business Information Systems*, *31*(1), 1-26. https://doi.org/10.1504/IJBIS.2019.099524

Srivastava, P. R., & Eachempati, P. (2021). Intelligent employee retention system for attrition rate analysis and churn prediction: an ensemble machine learning and multi-criteria decision-making approach. *Journal of Global Information Management (JGIM), 29*(6), 1-29. https://doi.org/10.4018/JGIM.20211101.oa23

Sung, C., Higgins, C. Y., Zhang, B., & Choe, Y. (2017). Evaluating deep learning in chum prediction for everything-as-a-service in the cloud. In *2017 International Joint Conference on Neural Networks (IJCNN)* (pp. 3664-3669). IEEE. https://doi.org/10.1109/IJCNN.2017.7966317

Tarahomi, M., Izadi, M., & Ghobaei-Arani, M. (2021). An efficient power-aware VM allocation mechanism in cloud data centers: a micro genetic-based approach. *Cluster Computing, 24*(2), 919-934. https://doi.org/10.1007/s10586-020-03152-9

Tofighy, S., Rahmanian, A. A., & Ghobaei-Arani, M. (2018). An ensemble CPU load prediction algorithm using a Bayesian information criterion and smooth filters in a cloud computing environment. *Software: Practice and Experience, 48*(12), 2257-2277. https://doi.org/10.1002/spe.2641

Ullah, I., Raza, B., Malik, A. K., Imran, M., Islam, S. U., & Kim, S. W. (2019). A churn prediction model using random forest: analysis of machine learning techniques for churn prediction and factor identification in telecom sector. *IEEE Access, 7*, 60134-60149. DOI: https://doi.org/10.1109/ACCESS.2019.2914999

Umayaparvathi, V., & Iyakutti, K. (2017). Automated feature selection and churn prediction using deep learning models. *International Research Journal of Engineering and Technology (IRJET), 4*(3), 1846-1854.

Venkatesh, S., & Jeyakarthic, M. (2020a). Metaheuristic based Optimal Feature Subset Selection with Gradient Boosting Tree Model for IoT Assisted Customer Churn Prediction. *Journal of Seybold Report ISSN NO, 1533*, 9211.

Venkatesh, S., & Jeyakarthic, M. (2020b). Adagrad Optimizer with Elephant Herding Optimization based Hyper Parameter Tuned Bidirectional LSTM for Customer Churn Prediction in IoT Enabled Cloud Environment. *Webology, 17*(2), 631-651. DOI: 10.14704/WEB/V17I2/WEB17057

Vijaya, J., & Sivasankar, E. (2019). An efficient system for customer churn prediction through particle swarm optimization based feature selection model with simulated annealing. *Cluster Computing, 22*(5), 10757-10768. https://doi.org/10.1007/s10586-017-1172-1

Vo, N. N., Liu, S., Li, X., & Xu, G. (2021). Leveraging unstructured call log data for customer churn prediction. *Knowledge-Based Systems, 212*, 106586. https://doi.org/10.1016/j.knosys.2020.106586

Wael Fujo, S., Subramanian, S., & Ahmad Khder, M. (2022). Customer Churn Prediction in Telecommunication Industry Using Deep Learning. *Information Sciences Letters, 11*(1), 24. http://dx.doi.org/10.18576/isl/110120

Wu, X., Li, P., Zhao, M., Liu, Y., Crespo, R. G., & Herrera-Viedma, E. (2022). Customer churn prediction for web browsers. *Expert Systems with Applications*, 118177. https://doi.org/10.1016/j.eswa.2022.118177

Xiahou, X., & Harada, Y. (2022). B2C E-Commerce Customer Churn Prediction Based on K-Means and SVM. *Journal of Theoretical and Applied Electronic Commerce Research, 17*(2), 458-475. DOI: https://doi.org/10.3390/jtaer17020024

Xu, T., Ma, Y., & Kim, K. (2021). Telecom churn prediction system based on ensemble learning using feature grouping. *Applied Sciences, 11*(11), 4742. https://doi.org/10.3390/app11114742

Zhang, T., Moro, S., & Ramos, R. F. (2022). A Data-Driven Approach to Improve Customer Churn Prediction Based on Telecom Customer Segmentation. *Future Internet, 14*(3), 94. https://doi.org/10.3390/fi14030094