The Ciência & Engenharia - Science & Engineering Journal
ISSN: 0103-944X
Volume 11 Issue 1, 2023
pp: 1158 - 1169

mSVM: An Approach for Customer Churn Prediction using modified Support Vector Machine and various Machine Learning Techniques

Dr.S Brinthakumari
Associate Professor, New Horizon Institute of Technology and Management, Maharashtra, India
brinthakumaris@nhitm.ac.in
Mrs.Priyanka Dhiraj Sananse and Mrs.Punam Bagul
Assistant Professor, K.C College of Engineering and Management Studies and Research, Maharashtra, India
priyanka.sonawane@kccemsr.edu.in, bagulpunam07@gmail.com

Abstract
Consumer attrition is a major issue across the world companies and that is one of their primary worries. Organizations, specifically in the telecommunication sector, are working to build technologies to forecast future customer churn due to the obvious direct impact on profits. It's essential to define the variables that lead to customer churn before making the required efforts to minimize churn. An important impact was the creation of an attrition estimation method that aids telecommunication corporations in predicting which consumers are willing to churn. The methodology proposed in this study employs mathematical methodologies on a big data framework to provide a unique strategy to feature design and evaluation. This research work has proposed a customer churn prediction using a modified Support Vector Machine Learning (mSVM) classifier. The significant contribution of this research is that we have changed the distance function of SVM in both training and testing. Similar machine learning algorithms are also validated on similar datasets such as Naïve Bayes (NB), K-nearest Neighbor (KNN) and Decision Tree (DT) with J48 classifiers. The BigML dataset has been used for detecting churn in the telecom industry in a real-time scenario. In an extensive experimental analysis, we demonstrate mSVM which produces higher accuracy over the traditional machine learning algorithm for different cross-validation methods.
Keywords: Deep learning, deep convolutional neural network, image recognition, machine learning, image sentiment classification

1 Introduction

Nowadays Consumers go through a long decision-making process while connecting to the many optional Telecommunication service accessible today. Telecommunications companies offer a wide range of services, and phone porting is common and also customer retention is an issue. As a consequence, it becomes more and more critical than ever for telecoms providers to detect

The Ciência & Engenharia - Science & Engineering Journal
ISSN: 0103-944X
Volume 11 Issue 1, 2023
pp: 1158 - 1169

events that cause clients to terminate their services and take proactive steps to maintain them. Begin by determining how many users churned throughout that particular month. The total monthly user days is further divided by the number of continues to pump each user day. The monthly churn rate is calculated by multiplying the amount by the number of monthly days. Data mining methodologies are more effective in forecasting client loss, according to research conducted over the previous few years. Developing a successful churn forecasting models is a time-consuming procedure that encompasses it all from identifying relevant predictor characteristics from a big amount of available consumer information to selecting the most appropriate forecast data mining methods for the extracted features. Aside from network data, telecom companies collect a plethora of client-related information, like client profiling, calling habits as well as democratic information. Based on their history of calling conduct and general behavior, it is possible to categorize a consumer's attitude toward leaving or staying. Data mining techniques are more effective in projecting turnover, according to research undertaken over the previous decade. Predictive modeling-based churn forecast solutions are more effective. Churn forecast systems and sentiment analysis use categorization and clustering techniques to evaluate churn consumers and the factors for their exit. Since, we acquire enormous volumes of information every day in the telecommunication industry, mining it using specific data mining techniques takes duration, and interpretation forecasts utilizing standard methods is tough. Various scholars have described static and dynamic strategies for reducing churn in large data collections. However, true churn detection remains a serious difficulty for such systems. Because telecommunication data might occasionally include churn, it's vital to recognize search flaws. To properly detect churn from extensive data, strong administration of client relationships is required.

Utilizing Natural Language Processing (NLP) and ML approaches, we suggested churn identification and forecasts from huge telecommunications data sets in this research. The first sector is associated with both the process of tactical Natural Language Processing, which involves information that has been standardized, feature extraction, and feature engineering. As feature extraction techniques, term frequency-inverse document frequency, Stanford Natural Language Processing, and connection between occurrences has been suggested. Statistical and machine learning approaches were used to train and assess the whole curriculum.

Overview of Machine learning

Text classification or prediction is a broader framework that gives a certain output to a given input text for categorization and aggregation. Series marking gives a category to each component of a sequence of numbers (PoS tagging provides a form of a word to each term in an input text); parsing assigns a distance matrix to a feature vector, establishing the linguistic meaning of the sentence; and so on. Estimation, which allocates a productive capacity to each input, is another example; stochastic classification is a subset of classification. This kind of algorithm use statistical analysis to get the optimal categorization for a specific scenario. Unlike conventional

The Ciência & Engenharia - Science & Engineering Journal
ISSN: 0103-944X
Volume 11 Issue 1, 2023
pp: 1158 - 1169

equations, which provide a 'best' class, probabilistic formulae produce a probability that the example is a participant of one of the training photos. Typically, the best class is picked as the one with the greatest probability. However, compared to – anti-classification models, such an approach offers significant benefits.

The whole paper is divided as follows. Part 2 describes a thorough state of art system analysis done by previous researchers. Part 3 proposed system has been demonstrated with system architecture and detailed explanation. While part 4 indicates an algorithm design and experimental set-up with result and discussion are evaluated in part 5. In part 6 conclusion and future work are considered for developed system.

2 Literature Survey

Telecommunications companies aren't usually the most well-liked by their clients. In the telecommunication sector, customer satisfaction is essential to succeed. Be it confusing invoicing, unsolicited promotional mails, lousy customer care, data rate, connectivity, or pricey packages, individuals routinely grumble about internet providers' effectiveness. As a reason, it's no wonder that telecom businesses have a higher proportion of client churn. Client churn (attrition) is debilitating in this industry since telecommunication providers retain huge fixed facilities that can only be compensated by income. Client retention is usually prioritized, followed by customer retention. Going to get a new client, on either extreme, can cost up to five times more than keeping an old one. Including a study conducted by Bain & Company, increasing client retention rates by 5percent can improve earnings by 25 percent to 95 percent. Customer attrition, often known as churn, is a term that measures the lot of consumers that discontinue performing business with the organization or service. Several companies might use this data to identify out what's generating their excessive attrition and develop proactive actions to address those problems. But then if you knew in advance that a particular consumer was just about to quit and might initiate appropriate steps to keep them from doing so?

Consumers may quit for a range of factors, such as poor service quality, prolonged customer service longer waits, pricing changes; new firms enter the industry, and so forth. Typically, there is no one explanation for customer discontent, but rather a series of events that lead to it. If your company fails to identify these warning signals and respond before the cancel button is clicked, your customer will have already departed. You do, however, have something valuable in the shape of data. Your customer gave you lots of cues as to where you were falling short. It might aid in the gathering of crucial data and the training of consumer churn systems. ML is all about gaining information from the past and using it to improve future experiences. The telecommunications industry has a lot of room for expansion. Carriers' large volumes of customer data might aid them in shifting from a reactive to a proactive mode of operation. The advancement of advanced Artificial Intelligence (AI) and information analytics tools has made it even easier using this rich information, to combat churn.

The Ciência & Engenharia - Science & Engineering Journal
ISSN: 0103-944X
Volume 11 Issue 1, 2023
pp: 1158 - 1169

Churn forecast employs two techniques: ML and DM. The decision-tree approach [1] is a good way to anticipate churn. Researchers also employ a neural network technique [7], information certainty [8], as well as Particle Swarm Optimization (PSO) for churn prediction. As per a system [2,], a modern group of technologies is being developed to raise the bar for recognizing potential churners. The responsibilities are classed as dealing, demand pattern, and call pattern modifications summary tasks, and are gathered from demand data and customer accounts. The properties are analyzed using different probabilistic data mining methods: Naive Bayes (NB) and Bayesian Network. The outcomes are presented to those produced with the C4.5 Decision Tree (DT), a popular categorization and predictive approach. Customers may readily move to competitors as a result of these factors, among other considerations. One way for doing this is to enhance churn forecast from large volumes of information utilising mining. According to [3] a literature review and formalization of the gathering process's time frame. Second, this study explores the increase in churn system accuracy by expanding the period of consumer activities from 1 to 17 years, using Logistic Regression (LR), classification trees as well as bagging combined with classification trees.

As a consequence, data-related activities such as information gathering, processing, and evaluation will be greatly reduced for researchers. The length and promotional nature of a subscription affect the cost of the subscription. They will get a letter from the newspaper business alerting them that their subscription will be stopped. Lastly, if they want to retain their subscription, walk them through the process. Consumers are unable to terminate their subscriptions, but after their membership has ended, they get a four-week grace period.

According to [4] the most successful customer contact strategies may be used to effectively boost customer satisfaction. The researchers employed a Multilayer Perceptron (MLP) neural network method to calculate the number of customers in one of Malaysia's largest telecoms businesses. Multiple Regression Analysis and Analyzing Logistic Regression were used to compare the data to the most commonly utilized churn prediction techniques. The most important neural network architecture, using the learning method Levenberg Marquardt (LM), comprises of 14 inputs, 1 hidden and 1 output node. A Multilayer Perceptron (MLP) neural network technique was utilized to calculate customer attrition at one of Malaysia's top telecommunications enterprises, in comparison to the most used churn forecasting models, including Multiple Regression Analysis as well as LR evaluation. The method [5] creates an effective and informative statistical churn system by focusing on strongly connected intervals in data sets using a Partial Least Square (PLS) approach. Early findings show that the proposed method generates highly precise outcomes than existing prediction systems and better recognizes crucial properties to describe churning patterns.

Furthermore, techniques for network, overage, and problem management are provided and studied in the context of a number of important marketing campaigns. In churn prediction models, Bures and Van den Poel [6] the Random sampling efficiency, Improved Under-Sampling, Gradient Boosting Technique as well as Weighted Random Forest (RF) are some of

The Ciência & Engenharia - Science & Engineering Journal
ISSN: 0103-944X
Volume 11 Issue 1, 2023
pp: 1158 - 1169

the techniques examined. To evaluate the concept, metrics were employed (AUC, Lift). The sampling technique is shown to be superior to the other strategies investigated in the study. Gavril et al. [7] deliver a novel data mining method for detecting client turnover across a huge range of corpora types. Around 3500 customer records are evaluated based on incoming and departing input calls and messages. For development and research, conventional machine learning approaches were applied. For the whole dataset, the methods computed mean efficiency around 90 percent. Asma Baccouche et al. [8] used a Neural Network technique to build a prediction model to handle the issue of customer attrition for a big Chinese telecommunications corporation with 5.23 million associates. The mean exactness was 91.1 %, suggesting that the results were very predictable.

Idris [9] suggested a genetic engineering method to simulate AdaBoost-churning telecom issues. The two known Data Sets were used to evaluate the series, having 89 percent and 63 percent efficiency, one from Orange Telecommunications and another from cell2cell. Huang et al. [10] examined client's attrition using only a big data framework. Big data, depending on the quantity, variation, and volume of information, considerably improves the churn prediction loop, according to the study. Info from China's largest telecoms company's Project Management as well as Assistance Departments was supposed to be stored in a gigantic data warehouse for fracture engineering. The forest tool was used to identify AUC at randomly. K-means as well as fuzzy c-means clustering methodologies are used to divide subscribers into different groups depending on clustered input characteristics, as according to [11]. Several classifications are used to create the Adaptive Neuro-Fuzzy Inference System (ANFIS), a forecast system for active churn management. Neuro-fuzzy parallel categorization is the early phase in the prediction process. The Neuro-fuzzy classifier's results are used mostly by FIS to determine churner operations. Performance metrics could be used to identify ineffective behaviors. Client service infrastructure and services, operations, and performance are all linked to churn administration metrics. The flexibility of GSM numbers is a critical consideration when choosing a churner. To improve the identification of possible churners, System [12] has been updated with a new set of apps. The attributes are gathered from telephone records and consumer data and are classed as agreement, call behavior, and call pattern alterations descriptive characteristics. The items are analyzed using two probabilistic data mining methods: Naive Bayes (NB) and Bayesian Network. The outcomes are compared to those from a C4.5 decision tree, which is commonly used in categorization and forecast uses. Consumers may easily shift to competitors as a result of these factors, among other reasons. One way for doing this is to enhance churn forecast from huge volumes of information mining. According to [13], the technique for selecting time windows has been formalized, and a literature study has been conducted. Second, utilizing Logistic Regression (LR), classification trees, as well as bagging in conjunction with classification trees, this article investigated the rise in churn model accuracy by stretching the background of client events between 1 and 17 years. As a result, data-related challenges in data storage, planning, and research may be considerably reduced. Customers must pay a certain

The Ciência & Engenharia - Science & Engineering Journal
ISSN: 0103-944X
Volume 11 Issue 1, 2023
pp: 1158 - 1169

amount based on the duration of their membership and the promotional value. They get a letter from the newspaper company informing them that their membership is about to run out. Then, if they want to extend their subscription, walk them through the process. Clients are unable to terminate their subscriptions, but after their membership has ended, they get a four-week grace period.

According to [14], to lessen client retention rate, client retention tactics should be implemented. The study proposes a neural network method for assessing turnover rates of client's in one of Malaysia's main telecommunications carriers using Multilayer Perceptron (MLP). Multiple Regression Analysis and Logistic Regression Analysis were used to compare the findings to the most prevalent churn prediction approaches. The optimum design of the neural network using the Levenberg Marquardt (LM) learning algorithm comprises 14 input, one hidden and one output node. When comparing the far more general churn prediction methodologies, like Multiple Regression as well as LR Analysis, a Multilayer Perceptron neural network technology was used to forecast customer churn at one of Malaysia's leading telecoms firms. The approach [15] explains about using a Partial Least Square (PLS) methodology to develop an accurate and thorough predictive churn approach depending on strongly correlated huge datasets among parameters. Initial results suggest that the suggested model surpasses existing prediction systems and reveals critical factors for interpreting churning behavior. The preliminary churn marketing practices described and analyzed are system, overage and problem administration.

Burez and Van den Poel [16] evaluated the validity of Random Sampling,Gradient Boosting Prototype, Improved Under-Sampling, , as well as Weighted RF in churn estimation techniques using imbalanced data. The system was evaluated using (AUC and Lift) metrics. The under-sampling strategy outperformed the other tactics studied, as per the test outcomes.

Gavril et al. [17] developed a novel data mining model for detecting prepaid churn rate using a data of 3333 call details containing 21 factors and a conditional churn variable with two parameters: Yes/No. Data on the amount of inbound and outbound texts, and also mails for each client, were amongst the characteristics. The researcher used Principle Component Analysis (PCA) to decrease the size of the dataset. Three machine learning algorithms were used to estimate the churn aspect: Neural Networks (NN), Bayesian network, as well as Support Vector Machine (SVM). The study developed AUC to assess the programs' effectiveness. AUC rates of 99.10 %, 99.55 %, & 99.70 % are found in Bayesian Networks, NN, and SVM, correspondingly. There were no missing values in the dataset included in this investigation. To fix the challenges of customer defection in a huge Chinese telecommunication firm with roughly 5.23 million consumers, He et al. [18] developed a predictive model depending on the NN method. The overall efficiency rate was utilized as a baseline for forecast accuracy, and it was 91.1 percent. Idris [19] demonstrated a genetic programming approach that used AdaBoost to mimic telecom churn. Two main data sets were used to validate the hypotheses. One data was produced by Orange Telecom, while another was supplied by cell2cell, with the cell2cell data possessing 89 percent efficiency and the other possessing 63 percent. The problem of customer churn in the big

The Ciência & Engenharia - Science & Engineering Journal
ISSN: 0103-944X
Volume 11 Issue 1, 2023
pp: 1158 - 1169

data framework was investigated by Huang et al. [20]. The investigators wanted to show that big data might assist anticipate churn significantly based on the quantity, diversity, and speed of the information. A big data framework was required to produce the fractures in information from China's most known telecom firm's Operational and Business Support divisions. AUC was used to assess the RF approach.

3 Proposed work

To identify churn, the planned research effort will employ text analysis, natural language processing, and machine learning. Observe the client's changing habit pattern during forecast. To discover which variables have the most influence on the accuracy of churn forecasting. Month after month and day after day, examine and calculate the churn rate, which aids in enhancing the model's quality of service. To enhance model effectiveness, the suggested research would develop and implement a solution for churn prediction based on Natural Language Processing and ML. Later, throughout the forecast, we see the client's changing attitude pattern. We have also looked at the elements which have the greatest impact on churn forecast accuracy. Every month and day, we examine and calculate churn rates, which helps us to improve the system's customer experience. In this research, we proposed leveraging big data to forecast churn. The process begins with an artificial information gathering from telecom that contains certain imbalance information. Information preprocessing, standardization, extraction and selection of characteristics should all be done as needed. All through the run, optimization strategies were utilized to reduce identical characteristics that could lead to a high failure rate. The training and validation of the proposed system is later done. After all of the procedures have been accomplished, the plan specifies the data set's classification accuracy.

The Ciência & Engenharia - Science & Engineering Journal
ISSN: 0103-944X
Volume 11 Issue 1, 2023
pp: 1158 - 1169

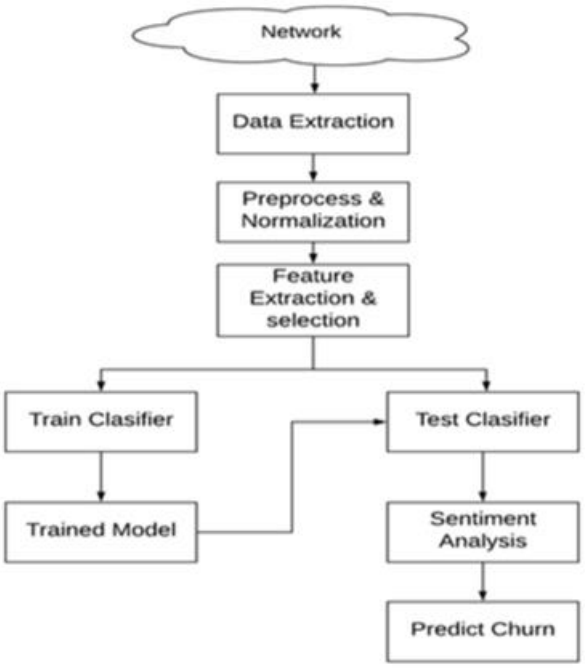

Fig.2 Proposed system overview

The goal of this type of study in the telecom sector is to assist businesses in increasing their profits. Forecasting turnover is already one of the biggest major sources of revenue for telecommunication providers. As a consequence, the article's goal was to develop a methodology for the Telecom Company to forecast consumer attrition. These forecasting models will have significant AUC rates. The sample group was split into two parts for evaluation and prototype creation: 30 percent and 70 percent for testing and training respectively. We used 10-fold cross-validation to analyze and improve hyper parameters. Engineering tools were used, as well as effective function translation and a selection method. They're making machine learning algorithms more user-friendly. There was another problem: the information was imbalanced. The churn of clients contributes for only about 5% of overall entries. To solve a problem, under sampling or tree approaches that are unaffected by this issue were applied. Our multiple classifiers can be more successful at detecting churn and making correct predictions in large datasets. This research makes a contribution to the creation of a supervised method for recognizing dimensional classes, selecting acceptable characteristics, and reducing repetition by evaluating their association. The findings show that using the correlation methodology leading to a higher f-score in the weighted frequency of the term. The use of weighted word frequency to select features is critical in this scenario. To avoid overlap, the connection between qualities in a category of aspects is assessed.

1. Data gathering: First, certain features are used to obtain information for distinct Telecom Sector consumers.

The Ciência & Engenharia - Science & Engineering Journal
ISSN: 0103-944X
Volume 11 Issue 1, 2023
pp: 1158 - 1169

2. Preprocessing: Next, to verify that our data is comprehensive, we'll perform textual analysis, elimination of stop word, stem (using Porter's method), index term extraction, and data filtering.
3. Lexical analysis: The input alphabet is divided into two types via lexical analysis: 1) alphabetic characters (a-z) and 2) separators between words (e.g. space,tab).
4. Stop word removal: It is the technique of deleting recurring phrases from publications.
5. Stemming: It is the process of transforming all versions of a phrase with a single stem phrase. Auxiliary verbs, parenthetical phrase, third-person omniscient endings, past tense suffixes, and other variations exist.
6. Data Training: We aggregate fake and real-time news data from the internet and train any machine learning classifier.
7. **Machine learning testing:** We use any computational modeling classifier or weight analyzer for actual or artificial data input to forecast online news.
8. Evaluation: We showed the developed system's effectiveness and compared it to other existing systems.

3.1 System overview

Input: inp 1…..n all input parameters, Desired Threshold Th.
Output: Executed for output as lable.
Step 1 : Read all records from database (R into DB)
Step 2: Parts [] ← Split(R)
Step 3: $\mathrm{CVal} = \sum_{k=0}^{n} \mathbf{Parts[k]}$
Step 4: check (Cval with Respective threshold)
Step 5: T← get current state with timestamp
Step 6 : if(T.time > Defined Time)
Read all measure of for TP and FN
Else continue. Tot++
Step 7: calculate score = (TP *100 / Tot)
Step 8 : if (score >= Th)
Generate event
end for

The Ciência & Engenharia - Science & Engineering Journal
ISSN: 0103-944X
Volume 11 Issue 1, 2023
pp: 1158 - 1169

Decision Tree Classifier

Input : Selected feature of all test instances D i….n , Training database policies {T 1 ……T n }
Output: No. of probable classified trees with weight and label.
Step 1: Read (D into D[i])
V←Extract features (D)
Step 2: N←Count_Features(D)
Step 3: for each(c into TrainDB)
Step 4: Nc[i] ==Ext_Features(c)
Step 5: select relevant features of w= {Nc[i], N}
Step 6: Statement (w>t)
Step 7: Return Tree Insatnce { Nc[i], N, w, label}

KNN Classifier

Input: Train_DatasetF TrF[], Test_DatesetF TsF[], Threshold T.
Output: Classified label
Step 1: Read R {All attributes} from current parameters.
Step 2: Map with train features with each sample.
Step 3: Calculate distance of train DB with same evidences

$$distance = \sum_{k=0}^{n} (TrF, TsF)$$

Step 4 : evaluate distance> threshold
Step 5: Return the predicted label

Updated Random Forest Classifier (URF)

Input: Training Rules Tr[], Test Instances Ts[], Threshold T.
Output: Weight w{0-1}
Step 1 : Read each test instance from (TsInstnace from Ts)
Step 2 : Read each test instance from (TsInstnace from Ts)
Step 3 : TsIns $= \sum_{k=0}^{n} \{Ak \ldots An\}$
Step 4 : Read each train instance from (TrInstnace from Tr)
Step 5 : $TrIns = \sum_{j=0}^{n} \{Aj \ldots \ldots Am\}$
Step 6 : $w = WeightCalc(TsIns, TrIns)$
Step 7 : if (w >= T)
Step 8 : Forward feed layer to input layer for output
OutLayer[] ← {Tsf,w}
Step 9 : optimized feed layer weight, Cweigt ← OutLayer [0]
Step 10 : Return Cweight

3.2 Datasets used

The system categorization graph is shown below. The graphs show how the system categorizes the aggregate inputs into separate cases. The suggested system uses a modified SVM

The Ciência & Engenharia - Science & Engineering Journal
ISSN: 0103-944X
Volume 11 Issue 1, 2023
pp: 1158 - 1169

combination that produces good results in all areas. 5000 instances were supplied for training while 1500 reviews were given for assessment with various classification models and performance assessment. The projected outcomes are compared to two three similar machine learning techniques. The dataset has been collected from the below url.
https://bigml.com/user/francisco/gallery/dataset/5163ad540c0b5e5b22000383
The comparative study of multiple classification methods for the recommended time series forecasting module is shown in Table 1. KNN has the lowest accuracy, but mSVM classification has the best accuracy with a 95% from different cross validation rate. Figure 2 below shows a similar set of data.

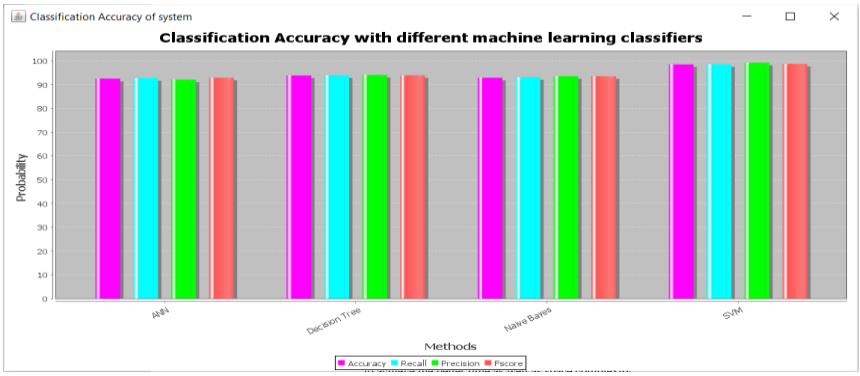

Fig.2 Accuracy of proposed model using various machine learning algorithms for churn prediction

4 Observations

- Many systems dealings with mSVM have adopted various types of boosting strategies to obtain superior time and space complexity, according to the above literature study.
- To detect the churn from heterogeneous dataset, it is applied with better classification accuracy.
- The proposed model obtains the higher accuracy over the traditional machine learning classifiers such as J48, NB and KNN etc.
- The BigML dataset, Twitter testing dataset, and customer review dataset are the most often utilized datasets.
- By combining modified SVM, the best accuracy may be achieved with the least amount of temporal complexity.

5 Conclusion

The study aimed at identifying and detecting churn customers from large telecoms corpora and the state-of-the-art analyses of churn prediction models developed by various studies. Several methods still have problems converting linguistic information, which can contribute to high rate of failures throughout implementation. To get superior data organization findings, most scholars have recommended merging Natural Language Processing techniques with various ML methods. Suppose a training algorithm interacts with that approach. In that case, the complete data set

must be tested or confirmed using even sampling strategies that eliminate data imbalance issues and give a trustworthy predicted flow of data. If we cope with the developed models with the HDFS structure and concurrent machine learning methods, which can provide improved outcomes at a lower cost of estimation, in the future path to incorporate a proposed methodology with different methods to obtain higher accurateness, especially when the input signal is big in terms of size and volume.

References

1. S. Mohan, C. Thirumalai and G. Srivastava, "Effective Heart Disease Prediction Using Hybrid Machine Learning Techniques," in IEEE Access, vol. 7, pp. 81542-81554, 2019,
2. Li Yang, Haibin Wu, Xiaoqing Jin, and Pinpin Zheng4 et al., "Study of cardiovascular disease prediction model based on random forest in eastern China ",Scientific Reports (2020) 10:5245
3. Youness Khourdifi and Mohamed Bahajm," Heart Disease Prediction and Classification Using Machine Learning Algorithms Optimized by Particle Swarm Optimization and Ant Colony Optimization", International Journal of Intelligent Engineering and Systems, Vol.12, No.1, 2019
4. Fahd Saleh Alotaibi," Implementation of Machine Learning Model to Predict Heart Failure Disease", (IJACSA) International Journal of Advanced Computer Science and Applications,Vol. 10, No. 6, 2019.
5. Lewlyn L. R. Rodrigues, Dasharathraj K Shetty, Nithesh Naik , Chetana Balakrishna Maddodi and et al., "Machine learning in coronary heart disease prediction: Structural equation modelling approach ", Rodrigues et al., Cogent Engineering (2020), 7: 1723198.
6. Mohd Ashraf, M. A. Rizvi and Himanshu Sharma, "Improved Heart Disease Prediction Using Deep Neural Network", Asian Journal of Computer Science and Technology , ISSN: 2249-0701 Vol.8 No.2, 2019, pp. 49-54 © The Research Publication.
7. Sumit Sharma, Mahesh Parmar , "Heart Diseases Prediction using Deep Learning Neural Network Model ", International Journal of Innovative Technology and Exploring Engineering (IJITEE) ISSN: 2278-3075, Volume-9 Issue-3, January 2020.
8. Asma Baccouche, Begonya Garcia-Zapirain , Cristian Castillo Olea and Adel Elmaghraby, "Ensemble Deep Learning Models for Heart Disease Classification: A Case Study from Mexico", Information 2020, 11, 207;
9. N. Sowri Raja Pillai , K.Kamurunnissa Bee, J.Kiruthika , "Prediction of heart disease using RNN algorithm", International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395-0056 Volume: 06 Issue: 03 | Mar 2019 www.irjet.net p-ISSN: 2395-0072.
10. M.Ganesan and Dr.N.Sivakumar, "IoT based heart disease prediction and diagnosis model for healthcare using machine learning models ", Proceedings of International Conference on Systems computation automation and networking 2019, IEEE- 978-17281-1524