

Comparative Analysis of Machine Learning Algorithms for Predicting Telecom Customer Churn: An Experimental Study

G. Yaswanth

PG Scholar, Dept. of Computer Science Sri Venkateswara University, Tirupati

Abstract— Client churn forecast should be the significant part of any organization. This aids in the discovery of clients who are probably going to drop a membership to a help. As of late, the versatile telecom market has transformed from a quickly developing business sector into a condition of immersion. The focal point of media transmission organizations is to move from developing of enormous client into keeping clients in house. Consequently, it is important to know which clients are probably going to change to a rival in future. This review presents a relative examination of two AI calculations specifically Decision Tree and Naïve Bayes for the grouping of client churn expectation. The trial results assess the exhibition of these calculations regarding exactness, accuracy, and review. The discoveries add to understanding the viability of various choice tree variations with regards to client churn expectation.

I. INTRODUCTION

With the parts of improvement and the headways all over the planet correspondence is one of the really required substances in the turn of events. Correspondence requirements to quick and dependable for individuals for their works and necessities to be finished. In created nations telecom businesses post a significant job and have turned into a piece of necessities expected for individuals to live [1]. The turns of events and mechanical advancement and the consistent expansion in the working organization the opposition between them on crested. Organizations need to flourish to make due in the serious market by executing new procedures and strategies for getting the client base all over the planet [3]. The fundamental subject behind these techniques to create and produce more pay to the organizations. There are various techniques followed by the showcasing group in the company to draw in new clients, making the current clients to purchase or move up to new administrations inside a similar organization lastly to keep the client base for longer period.

Clients who are probably going to churn is the significant issue in help areas with the high contest on the lookout. On the opposite side, breaking down the clients who are probably going to leave the organization will address possibly huge extra income source assuming it is finished in the beginning stage [6]. Numerous scientists affirmed that AI procedures is profoundly valuable and effective to foresee the stirring and non-agitating occasions by gaining from the past organization information. The information utilized in this contains all client data all through a timeframe. In this examination, it fundamentally centers around AI techniques and calculations for expectation of churn in telecom businesses; we basically focussed on naïve bayes and decision tree calculation for productive expectation model for client churn. There are various organizations to fabricate their own client churn expectation models since obtaining another client costs multiple times higher than the expense of holding the client liable to churn.

II. METHODOLOGY

Request models are a strategy for high importance used in various fields. In class confirmation, gathering models are used to sort out which class the data has a spot with. The request model is a model that works by making assumptions. In our survey, models were made including Naive Bayes and Decision Tree systems to fabricate a model which will foresee regardless of whether a specific client will churn.

2.1 Naive Bayes

Naive Bayes characterization is a famous AI calculation that depends on Bayes' hypothesis with a supposition of freedom between the highlights. It is a basic yet powerful probabilistic model utilized for grouping undertakings [2].

The calculation is classified "guileless" in light of the fact that it expects that the presence or nonappearance of a specific element is irrelevant to the presence or nonattendance of different highlights. All in all, it expects that all elements are autonomous of one another, which isn't generally evident in true situations. Notwithstanding this working on suspicion, Naive Bayes frequently performs well by and by and can give solid outcomes [4].

The Naive Bayes calculation works by computing the probabilities of an example having a place with every conceivable class in light of the noticed element values. It then doles out the example to the class with the most noteworthy likelihood. The

computation of these probabilities includes assessing the probability of each component given each class and the earlier likelihood of each class [5].

The calculation is especially valuable while working with high-layered datasets and when the presumption of component autonomy is sensible. It is known for its computational proficiency and is in many cases utilized in message order, spam sifting, opinion examination, and other comparable assignments.

2.2 Decision Tree

A decision tree is an overseen learning computation that is obviously appropriate for portrayal issues, as mentioning classes on a definite level is proficient [4]. Decision Tree estimations are used for the two assumptions as well as portrayal in computer based intelligence. Using the decision tree with a given course of action of information sources, one can design the various outcomes that are a result of the results or decisions [5]. It works like a stream chart, confining snippets of data into two relative characterizations at the same time from the "tree trunk" to "branches," to "leaves," where the classes become even more limitedly near. This makes classes inside groupings, taking into account regular game plan with limited human oversight. This decision tree is an outcome of various different evened out progresses that will help you with showing up at explicit decisions [5][7]. To develop this tree, there are two phases - Selection and Pruning. In enrollment, we build a tree however, in pruning, we kill the couple of complexities of the tree.

III. EXPERIMENTAL RESULTS

The examinations were led utilizing the Python programming language, using the strong Scikit-learn library for information portrayal, control, and investigation. For this review, the Telecom Agitate expectation dataset from the Kaggle was utilized [8]. This dataset comprises of 7043 examples, each containing 21 highlights, alongside an objective variable demonstrating the Stir (1869 occasions) or No Beat (5174 cases).

In this review, two AI calculations, specifically Naive Bayes, Decision Tree were applied to the telecom client stir dataset. The presentation of every calculation was assessed utilizing exactness, accuracy, and review as assessment measurements. The exploratory outcomes are summed up in the table-1 and same displayed in the figure-1:

**TABLE 1
PERFORMANCE OF CLASSIFIERS**

Algorithm	Accuracy	Precision	Recall
Decision Tree	89.78	89.7	90
Naive Bayes	86.53	87	87

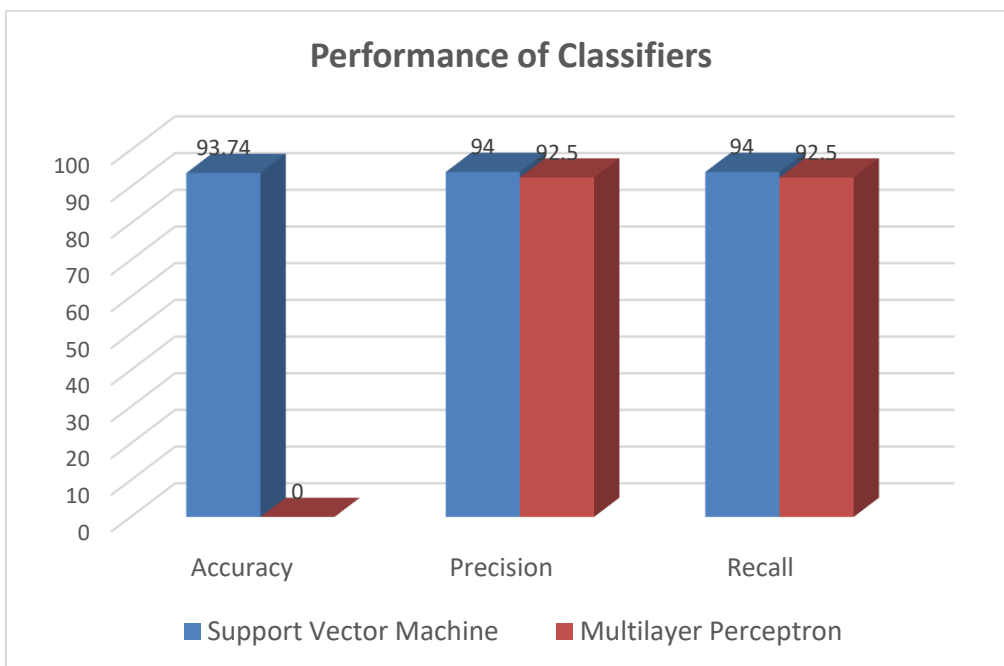


Figure-1: Performance of algorithms

From the figure-1, the outcomes show that the two calculations accomplished sensibly high exactness rates, demonstrating their true capacity for client agitate expectation. Decision Tree displayed somewhat higher exactness (89.78%) contrasted with Naive Bayes (86.53%).

Taking a gander at accuracy, Decision Tree accomplished an accuracy pace of 89.7%, demonstrating a low bogus positive rate in the characterization of stir forecast. Then again, Naive Bayes accomplished an accuracy pace of 87%. Albeit somewhat lower than decision tree, it actually exhibits a decent capacity to precisely distinguish client beat.

As far as review, decision tree accomplished a review pace of 90%, which connotes the calculation's capacity to accurately recognize positive occurrences of client stir. Naive Bayes, be that as it may, showed a fundamentally higher review pace of 87%, demonstrating its capacity to catch a bigger extent of positive occasions.

IV. CONCLUSION

The principal objective of the venture is to foresee the clients that are probably going to beat in a telecom industry, examination of the organizations with their past information for a while which assists them with understanding that day's circumstance of the organization and assists them with surveying stir variable of the organization so they can execute new designs to bring new clients and the clients currently in their organization. This framework is utilized to decrease the time and weight of organization by going through the authentic information and insight about the organization to evaluate and respond to the circumstances.

In this review features the similar examination of Decision Tree and Naive Bayes calculations for client agitate grouping, giving bits of knowledge into their exhibition in view of exactness, accuracy, and review measurements. These discoveries can act as an important reference for scientists and specialists working in the field of client beat order and related applications.

REFERENCES

- [1] Customer Churn Prediction in Telecommunication with Rotation Forest Method.
https://www.researchgate.net/publication/282981765_Customer_churn_prediction_in_telecommunication
- [2] D. Hand, H. Mannila, P. Smyth.: Principles of Data Mining. The MIT Press. (2001)
- [3] G. Ravi Kumar, K. Tirupathaiah and B. Krishna Reddy, "Client Churn Prediction of Banking and fund industry utilizing Machine Learning Techniques", International Journal of Computer Sciences and Engineering, Volume-7, Issue-6, e-ISSN: 2347 — 2693, PP: 871-875, June 2019
- [4] J.Han and M.Kamber,"Data Mining concepts and Techniques", the Morgan Kaufmann series in Data Management Systems, 2nded.San Mateo, CA; Morgan Kaufmann, 2006.
- [5] N.Michael, "Artificial Intelligence – A Guide to Intelligent Systems", 2nd Edition, Addison Wesley 2005
- [6] S. A. Qureshi, A. S. Rehman, A. M. Qamar, A. Kamal, and A. Rehman, "Telecommunication subscribers' churn prediction model using machine learning," in Proc. 8th Int. Conf. Digit. Inf. Manage., Sep. 2013, pp. 131–136
- [7] Surya Bhupal Rao, S.Rahamat Basha, G. Ravi Kumar, "A Comparative approach of Text Mining: Classification, Clustering and Extraction Techniques", Journal of Mechanics of Continua and Mathematical Sciences, Special Issue, No.-5, January (2020) pp 120-131, ISSN: 0973-8975.
- [8] <https://www.kaggle.com/code/ashydv/telecom-churn-prediction>.