IJCRT.ORG ISSN : 2320-2882

INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)

An International Open Access, Peer-reviewed, Refereed Journal

An Experimental Study Comparing Machine Learning Algorithms For Telecom Customer Churn Prediction

Dr P Devaraju
Assistant Professor, Dept. Of Computer Science & Technology,
Sri Krishnadevaraya University, Anantapuramu

ABSTRACT

A key component of any organization should be the client beat forecast. This makes it easier to identify customers who are most likely to discontinue their membership. Recently, the diverse telecom market has changed from a rapidly expanding industry to one that is immersed Media transmission companies are focusing on retaining their current clientele rather than expanding their clientele. Therefore, it's critical to identify which customers are most likely to switch to a competitor in the future. In order to group client beat expectations, this review provides a comparative analysis of two AI computations, namely Decision Tree and Naïve Bayes. The trial results evaluate how well these computations are displayed in terms of precision, accuracy, and review. The findings advance our knowledge of the feasibility of different choice tree variations in relation to client stir expectations.

1. Introduction

With the advancements and improvements occurring all over the world, one of the things that is truly needed in the current situation is correspondence. People need correspondence to be reliable and fast so they can complete their tasks and needs. In developed countries, telecom companies provide a large number of jobs and have become an essential part of people's daily lives [1]. The conflict between them increased as a result of the events that transpired the development of machinery, and the steady growth of the working organization. Businesses must thrive in order to survive in the competitive market by implementing innovative practices and strategies to expand their clientele globally [3]. The basic idea behind these methods is to generate and generate additional compensation for the organizations. The company's showcasing group uses a variety of strategies to attract new customers, persuade existing customers to upgrade to new services within the same company, and, finally, maintain the clientele for an extended length of time. The major problem in areas where there is a high level of competition for services is clients who are likely to stir. On the other hand, if it is completed in the early stages, identifying the clients who are most likely to leave the company will address a potentially significant additional revenue source [6]. Many scientists have confirmed that AI algorithms are extremely useful and efficient in anticipating both exciting and uninteresting events by

using historical organization data. All client data over a period of time is included in the data used in this. This analysis primarily focuses on AI methods and computations for telecom companies' expectation of disruption; we primarily examined naïve bayes and decision tree computation for a client beat productive expectation model. Since acquiring a new client is far more expensive than holding the current client accountable for agitation, there are several organizations that create their own client beat expectation models.

2. Methodology

Request models are an important strategy that is used in many different fields. To determine which class the data belongs to, class confirmation uses collection models. One model that operates by making assumptions is the request model. In order to create a model that can predict whether a particular client will stir or not, models such as Naive Bayes and Decision Tree systems were created for our survey.

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, 8].

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

Given that mentioning classes on a specific level is proficient, a decision tree is an overseen learning computation that is clearly suitable for portrayal problems [4, 8]. The two assumptions and representation in computer-based intelligence are estimated using decision trees. One can design the different outcomes that result from the decisions or results by using the decision tree with a specific course of action of information sources [5]. Similar to a stream chart, it simultaneously confines data fragments into two relative characterizations, from the "tree trunk" to "branches," to "leaves," where the classes get even closer. This creates classes within groups while considering a regular game plan with minimal human supervision. This decision tree is the result of several balanced steps that will assist you in making clear decisions. [5] [7]. There are two stages in the development of this tree: selection and pruning. We construct a tree during enrollment, but we eliminate a few of its complexities during pruning.

3. 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 Telecon Agitate expectation dataset from the Kaggle was utilized [9]. 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 in the figure-1:

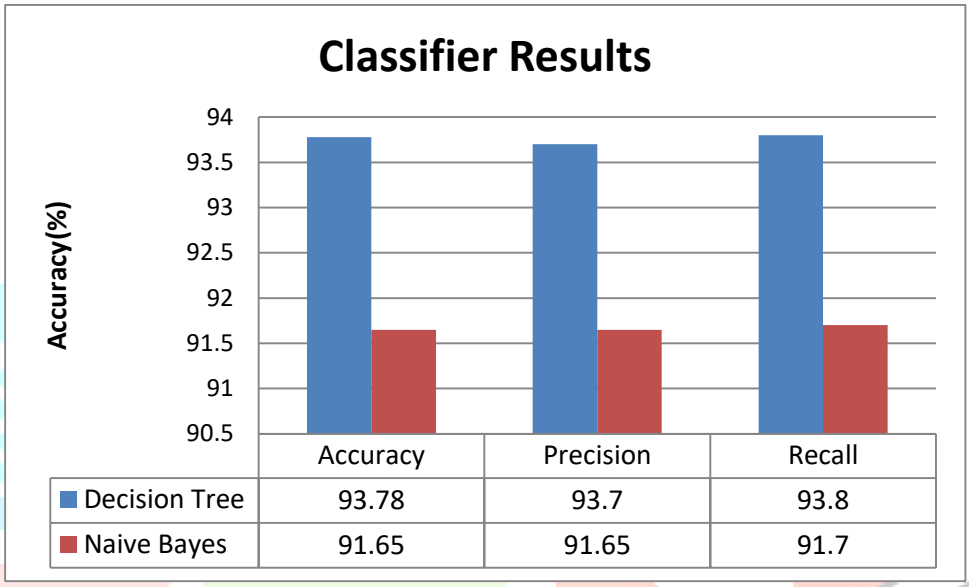

Figure-1: Performance of algorithms

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 (93.78%) contrasted with Naive Bayes (91.65%).

Taking a gander at accuracy, Decision Tree accomplished an accuracy pace of 93.7%, demonstrating a low bogus positive rate in the characterization of stir forecast. Then again, Naive Bayes accomplished an accuracy pace of 91.65%. 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 93.8%, 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 91.7%, demonstrating its capacity to catch a bigger extent of positive occasions.

4. Conclusion

The venture's main goal is to predict which clients are likely to succeed in the telecom sector. They do this by analyzing the companies' historical data, which helps them comprehend the current state of the company and survey its key variables so they can implement new strategies to attract both new and existing clients. By reviewing the organization's real information and insights, this framework is used to reduce the time and burden of the organization in order to assess and react to the situation.

In this review, Decision Tree and Naive Bayes calculations for client agitation grouping are similarly examined, providing tidbits of information about their display in terms of precision, accuracy, and review metrics. For scientists and experts working in the field of client beat order and related applications, these findings can serve as a valuable resource.

References

1. Customer Churn Prediction in Telecommunication with Rotation Forest Methodhttps://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. G. Ravi Kumar, K. Venkata Sheshanna, S. Rahamat Basha, and P. Kiran Kumar Reddy, "An Improved Decision Tree Classification Approach for Expectation of Cardiotocogram", Proceedings of International Conference on Computational Intelligence, Data Science and Cloud Computing, Lecture Notes on Data Engineering and Communications Technologies 62, Springer Nature Singapore Pte Ltd. 2021, PP:327-333, ISBN 978-981-33-4967-4
8. Dr. M. V. Lakshmaiah, Dr. G. Ravi Kumar and Dr. G. Pakardin, "Frame work for Finding Association Rules in Bid Data by using Hadoop Map/Reduce Tool", International Journal of Advance and Innovative Research, Volume 2, Issue 1(I), PP:6-9, ISSN 2394 -7780, Impact Factor : 1.437,Publisher: Indian Academicians and Researchers Association, UGC Journal No: 63571, January-March 2015
9. T. Aditya Sai Srinivas, A. David Donald , G. Thippanna "Sentimental Roadmap: Labeling a Dataset for Targeted Analysis", Journal of Network Security and Data Mining Volume 6 Issue 3.
10. https://www.kaggle.com/code/ashydv/telecom-churn-prediction