iJRASET
International Journal For Research in Applied Science and Engineering Technology



INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 8 **Issue:** V **Month of publication:** May 2020

DOI: http://doi.org/10.22214/ijraset.2020.5479

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.429
Volume 8 Issue V May 2020- Available at www.ijraset.com

Telecom Customer Churn Prediction

Morla Manasa[1], Mangunuru Narendra Reddy[2], Maddula Leela Sahithi[3], Pirati Yashwanth Kumar[4], Veerla Sandhya[5]
[1, 2, 3, 4]*Department of Computer Science and Engineering, Gudlavalleru Engineering College, Gudlavlleru, Andhra Pradesh, India*
[5]*Assistant Professor, Department of Computer Science and Engineering, Gudlavalleru Engineering College, Gudlavalleru, Andhra Pradesh, India*

Abstract: ***With the fast advancement of telecommunication industry, the specialist organizations are slanted more towards development of the supporter base. To address the issue of getting by in the serious condition, the maintenance of existing clients has become a huge challenge. In the study done in the Telecom business, it is expressed that the expense of procuring another client is undeniably more that holding the current one. Along these lines, by gathering information from the telecom businesses can help in foreseeing the relationship of the clients as whether they will leave the organization. The necessary activity should be attempted by the telecom businesses so as to start the obtaining of their related clients for making their reasonable worth stale. The relationship of Customer Relationship Management helps in catching purchaser data and the association further utilize this data to fulfill client needs. So as to improve and investigate the client obtaining and maintenance, CRM apparatuses have been created for increment in the productivity and help in the prescient displaying and characterization of different consistent undertakings .Our paper proposes another system for the churn prediction model. The effectiveness and the exhibition of Decision tree and Random forest methods have been looked at.***

Keywords: Telecommunication, Customer Relationship Management, Churn prediction, Decision Tree, Random Forest.

I. INTRODUCTION

The telecommunication area has gotten one of the fundamental businesses in created nations. The specialized advancement and the expanding number of administrators raised the degrees of rivalry. Organizations are striving to make due in this serious market contingent upon different procedures. Three primary systems have been proposed to create more incomes, for example, procure new clients, upsell the current clients, and increment the maintenance time of client. Nonetheless, looking at these systems considering the estimation of income on venture of each into account that the third procedure is the most gainful technique, demonstrates that holding a current client cost a lot of lower than securing another one, in addition to being viewed as a lot simpler than the upselling methodology. To apply the third system, organizations need to diminish the capability of client's churn, known as the client development starting with one supplier then onto the next.

Client's churn is an impressive worry in administration area with high serious administrations. Then again**,** predicting the client who are probably going to leave the organization will speak to possibly enormous extra income source on the off chance that it is done in the early stage. Many research affirmed that machine learning innovation is profoundly productive to anticipate this circumstance. An efficient churn prescient model benefits organizations from numerous points of view.

Early identification of clients prone to leave may assist with building cost effective ways in showcasing procedures. Client maintenance crusades may be restricted to chosen clients yet it should cover the greater part of the client. Inaccurate expectations could bring about an organization losing profits in view of the limits offered to nonstop supporters. Along these lines, the correct predictions of the churn clients have turned out to be exceptionally significant for the organizations. The noticeable job that the telecommunication area has come to possess overall makes it even more imperative to create expectation systems along the lines of churn prediction. Scarcely any insights show the significance of the client holds in this area.

II. LITERATURE REVIEW

Telecom business endeavors from their operational exercises, created, gathered and put away enormous measures of verifiable information and are presently underwriting these information resources now. Such verifiable information include covered inside them designs identifying with the adequacy of their different business forms. Machine learning has been effectively applied in numerous businesses as a reasonable tool for information revelation.Classification is critical errands in a wide assortment of utilization areas. There are countless strategies accessible to play out this undertaking and frequently determination of the most suitable method represents a major test to the professional. Both the procedures have discovered significant use in the field of Customer Relationship Management (CRM) and Churn prediction.

©IJRASET: All Rights are Reserved

A. CRM in Telecom Industry

Globalization and deregulation, joined with progresses in data innovation, have drastically changed the administrative setting of administration businesses. Despite of the fact that the starting of customer relationship the executives was at first in the mechanical setting, the administration business is likewise centered around keeping up and improving client connections. Administrations are created and conveyed by similar foundations. Accomplishment of a specialist co-op is subject to long term connections that create between the supplier and client. Administration industry relies upon consistent pattern of repurchase so maintenance of clients requires hard endeavors by associations Anderson et al.,(1994).

The field of CRM start in 1990 (Ling and Yen, 2001; Xu et al., 2002) and was characterized as an mechanism for creating relationship with one's customer and increment client faithfulness and benefit (Ngai, 2005). After this second, CRM was characterized in different habits by different creators. Swift (2001) characterized CRM as a venture way to deal with under and impact client behavious through significant correspondence so as to improve client procurement, maintenance, reliability and productivity.

Then again, Kotler and Keller (2006), characterized CRM as the way toward overseeing itemized data about individual clients and cautiously dealing with all client "contact focuses" to expand client dedication. CRM, as per Kincaid (2003), was the key utilization of data, procedures, innovation and individuals to oversee client relationship with an organization.

As an elective view, Richards and Jones (2008) characterized CRM as a lot of business exercises bolstered by both innovation and procedures that is coordinated by system and is intended to improve business execution in a region of client/customer the board. Inspite of different definition, it tends to be comprehended that the main issue of CRM is client and is the investigation of consumer loyalty and client relationship.

Olafsson et al. (2009) consumer loyalty and important clients in an organization is dynamic and the relationship develops and changes after some time. As per Poel and Larivie're (2004), the principle objective of CRM is to forestall client churn, which is the most significant danger to telecommunication organizations. In like manner, Ryals (2005) assembled CRM into two classes. The principal classification of assignments engaged with pulling in new clients and the second class of errands concentrated on keeping the current clients. In continuation with this examination, Reinartz et al. (2005) revealed that deficient distribution to client maintenance endeavors has extraordinary effect on long term client gainfulness when contrasted with lacking portion to client procurement endeavors. Comparative discoveries were additionally announced by Chu et al. (2007).

The rise of electronic business has expanded the measure of accessible data and has enormously helped organizations to take care of client's desires. Simultaneously, this improvement additionally made client mindful of other market openings, which thus, made them all the more requesting. Inability to fulfill these needs made the clients change their past specialist co-op. This procedure whenever called "Churning" (Lejeune, 2001). This changed over the market from business-driven to client driven (Bose,2002).

III. AN OVERVIEW OF PROPOSED SYSTEM

In this paper, we audit the current takes a shot at churn prediction in three alternate points of view: datasets and models**.** Firstly, we present the insights concerning the accessibility of open datasets and what sorts of client details are accessible in each dataset for predicting client churn. Besides, we thoroughly analyze the different prescient demonstrating models that have been utilized in the writing for anticipating the churners utilizing various classes of client records, and afterward quantitatively look at their exhibitions. At long last, we sum up what sorts of execution measurements have been utilized to assess the current churn prediction techniques**.** Analyzing all these three points of view is critical for building up an increasingly proficient churn prediction framework for telecom businesses.

While there are other churn prediction studies accessible in the writing, they fundamentally centered around various demonstrating strategies. As far as we could possibly know, none of those studies assessed the datasets and measurements for assessing the beat expectation models. Henceforth, we accept that this study can give a guide to the two analysts and client relationship administrators to all the more likely comprehend the area and difficulties in detail.

IV. METHODOLOGY

The procedure we utilized in our model is classification algorithms. The possibility of classification algorithms is quite basic. You anticipate the objective class by breaking down the preparation dataset. This is one of the most, if not the most fundamental idea you study when you learn Data Science. We utilize the preparation dataset to show signs of improvement limit conditions which could be utilized to decide each target class. When the limit conditions are resolved, the following undertaking is to anticipate the objective class. The entire procedure is known as Classification.

©IJRASET: All Rights are Reserved

ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.429
Volume 8 Issue V May 2020- Available at www.ijraset.com

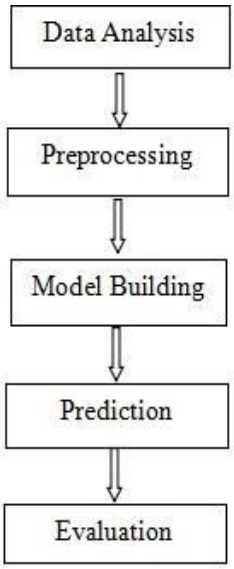


Fig. 1 Architecture

A. Decision Tree

A Decision Tree is a straightforward representation for grouping models. It is a Supervised Machine Learning where the information is constantly part as indicated by a specific parameter.

1) The principle goal of utilizing Decision Tree in this exploration work is the forecast of target class utilizing choice standard taken from earlier information.

2) It utilizes nodes and interior nodes for the forecast and arrangement. Root nodes group the occurrences with various highlights.

3) Root nodes can have at least two branches while the leaf nodes speak to characterization.

In each stage, Decision tree picks every node by assessing the most elevated data gain among all the characteristics.

B. Random Forest

Random forest is a supervised learning algorithm. It tends to be utilized both for classification and regression. It is likewise the most adaptable and simple to utilize algorithm. A forest is involved trees. It is said that the more trees it has, the more powerful a forest is. Random forest makes decision trees on arbitrarily chose information tests, gets prediction from each tree and chooses the best arrangement by methods for casting a ballot. It additionally gives a quite decent pointer of the element significance.

Random forest has an assortment of utilizations, for example, proposal motors, picture arrangement and highlight determination. It tends to be utilized to characterize faithful credit candidates, recognize fake movement and predict diseases. It lies at the base of the Boruta algorithm, which chooses significant highlights in a dataset

The Random Forest is a model comprised of numerous decision trees.Rather than basically averaging the expectation of trees , this model uses two key ideas that gives it the name irregular:

1) Random inspecting of preparing information focuses whenever building the trees

2) Random subsets of highlights thought about while parting nodes

a) Random Testing of Preparing Perceptions :

i) The tests are taken with substitution, known as bootstrapping, implies that a few examples will be utilized on numerous occasions in a solitary tree.

ii) The thought is that via preparing every tree on various examples, although every tree may have high fluctuation as for a specific arrangement of the preparation information, by and large, the whole forest will have small difference yet not at expense of expanding the predisposition.

b) Random Subsets of Highlights for Parting Nodes:

i) The other primary idea in this algorithm is solitary a subgroup of the considerable number of highlights are considered for parting every hub in every decision tree

ii) This algorithm can likewise be prepared thinking about all the highlights at each hub as normal in relapse.

©IJRASET: All Rights are Reserved

C. How Random Forest Algorithm Works

This algorithm works on two phases, first one is random forest creation,and the second one is to make an expectation from the random forest classifier made in the principal stage. The entire procedure is demonstrated as follows

Here the creator starts the Random Forest creation pseudocode:

1) Randomly select "K" highlights from absolute "m" highlights where k << m
2) Among the "K" highlights, figure the node "d" utilizing the best part point
3) Split the hub into girl hubs using the best split
4) Redo the a to c ventures until "l" number of hubs has been reached
5) Build the forest by rehashing stages a to d for "n" number occasions to make "n" number of tree

D. KNN

K-Nearest Neighbor is one of the simplest Machine Learning algorithms dependent on Supervised Learning method. It expect the likeness between the new case/information and accessible cases and put the new case into the class that is generally like the accessible classifications. It stores all the accessible information and characterizes another information point dependent on the likeness. This implies when new information shows up then it very well may be effectively characterized into a well suite classification by utilizing K-NN algorithm.It can be utilized for Regression just as for Classification however for the most part it is utilized for the Classification problems.K-NN is a non-parametric calculation, which implies it doesn't make any supposition on hidden data.It is likewise called a lazy learner algorithm since it doesn't gain from the preparation set quickly rather it stores the dataset and at the hour of arrangement, it plays out an activity on the dataset. At the preparation stage just stores the dataset and when it gets new information, at that point it arranges that information into a classification that is a lot of like the new information.

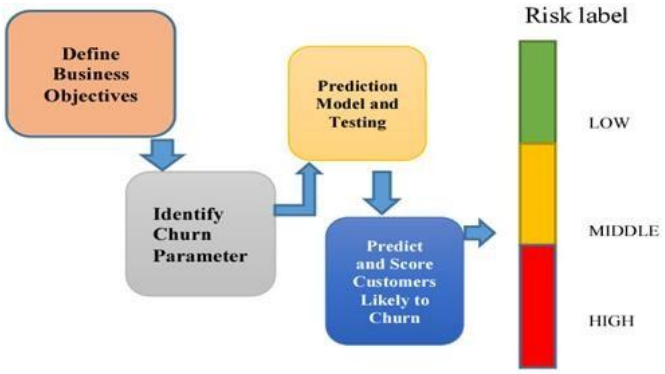


Fig. 2 Schematic of churn prediction model

V. RESULTS AND DISCUSSIONS

Accuracy values of classification algorithms we used like KNN, Decision Tree and Random Forest are displayed in below table. Out of all these algorithms Random Forest algorithm got the highest Accuracy. The below figure shows the accuracy values of classification algorithms we used.

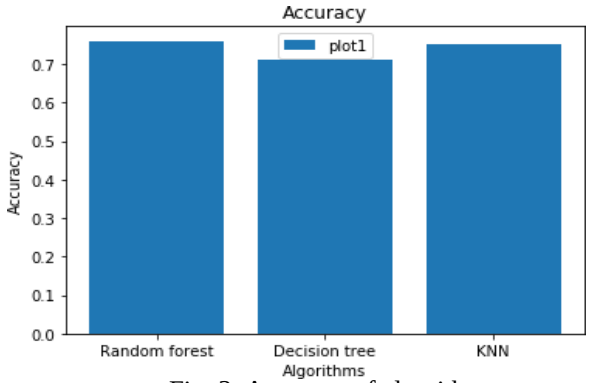


Fig. 3 Accuracy of algorithms

©IJRASET: All Rights are Reserved

A. Development of GUI

The model that gave the maximum accuracy for the test data was Random Forest. So, Random Forest is used for creating the GUI. The GUI is created using Tkinter package in python . Two GUIs are created, one for predicting and the other for training new data. TheGUI contains input fields for all attributes in the dataset. The system will predict whether the Customer will leave the company or not based on the trained model. The GUI will be a useful tool for telecommunication industry.The below figures are user interfaces. It will predicts that the customers who are likely to churn.

The below figure predicts that the customer will not leave the company.

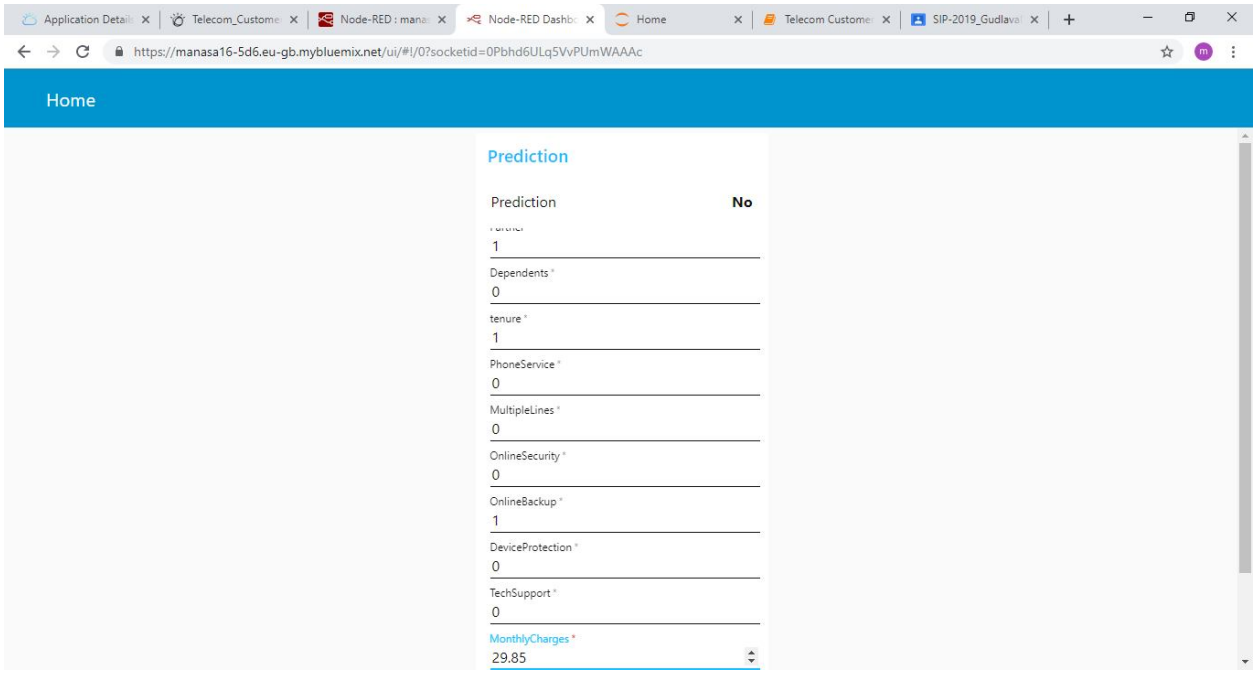


Fig. 4 Predicting the false output

The below figure predicts that the customer will leave the company.

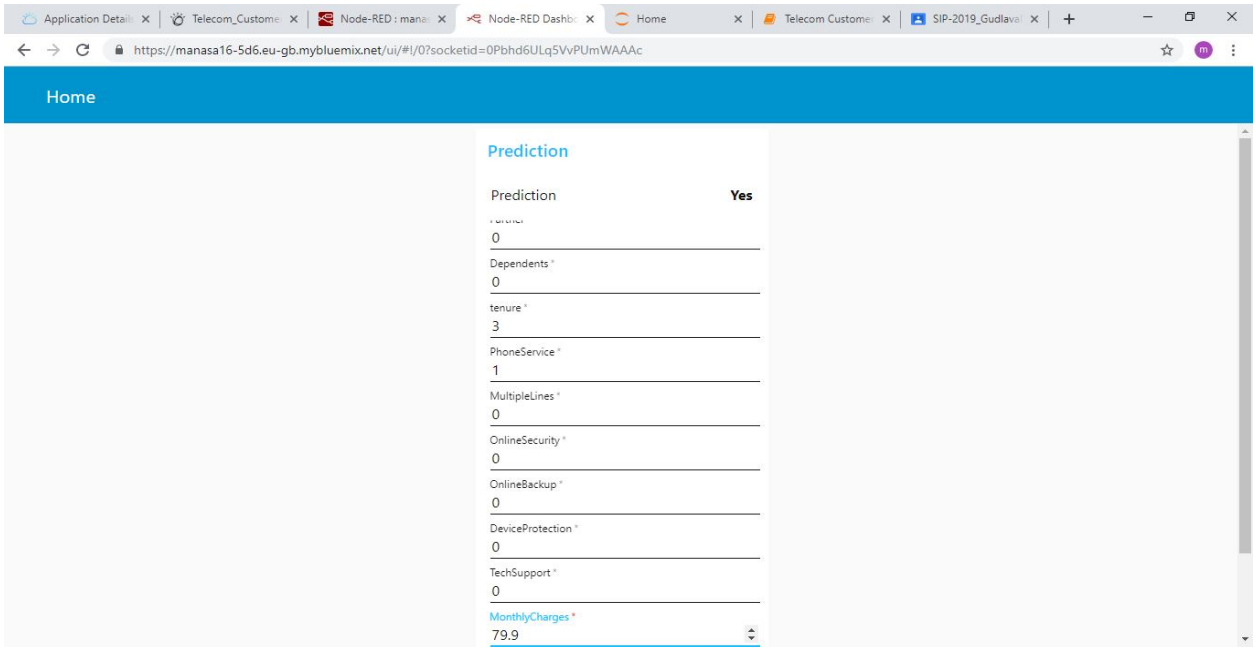


Fig. 5 Predicting the true output

©IJRASET: All Rights are Reserved

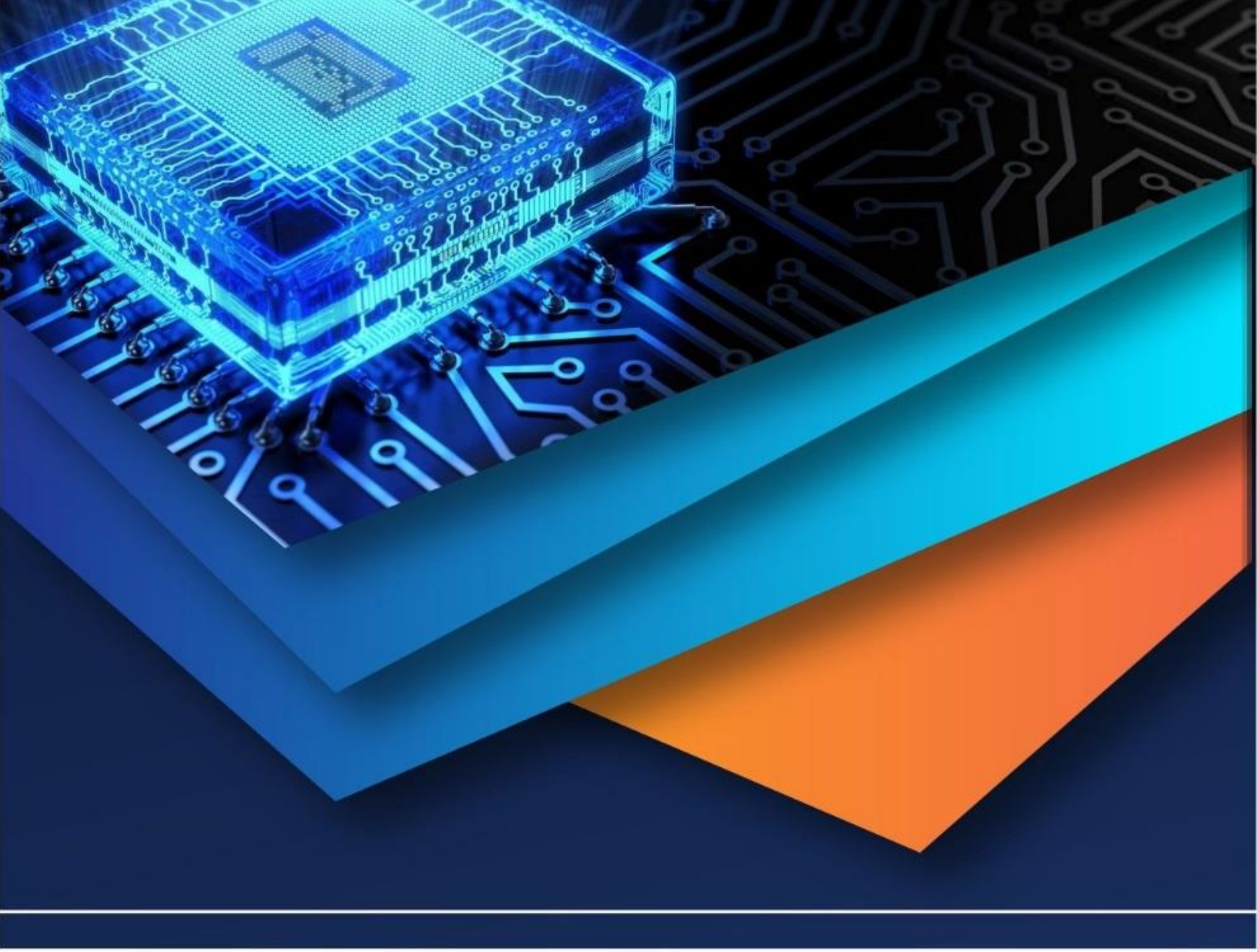
VI. CONCLUSION

Telecommunication industry has experienced high churn rates and tremendous churning misfortune. Despite the fact that the business misfortune is unavoidable, yet churn can be overseen and kept in an adequate level. Great techniques should be created and existing strategies must be improved to prevent the telecommunication industry to confront difficulties. In this paper we talked about the different prediction models and furthermore thought about the quality proportions of prediction models like irregular random forest, decision trees. We found that the accuracy achieved with random forest is far much higher than the decision tree technique which clearly states that random forest is an efficient technique.

REFERENCES

[1] Huang, B., Kechadi, M.T., Buckley, B.: Customer churn prediction in telecommunications. Expert Syst. Appl. 39, 1414–1425,2012
[2] Yen-Hsien Lee, Chih-Ping Wei, Tsang-Hsiang Cheng, Ching-Ting Yang, "Nearest-neighbor-based approach to time-series classification", Decision Support Systems, Volume 53, Issue 1, 2013, pp. 207–217.
[3] Wolniewicz, R.H., Dodier, R.: Predicting customer behavior in telecommunications. IEEE Intell. Syst. 19, 50–58 (2004)
[4] https://patents.google.com/patent/US6009199A/en
[5] https://www.kaggle.com/blastchar/telco-customer-churn
[6] Saradhi, V.V., Palshikar, G.K.: Employee churn prediction. Expert Syst. Appl. 38, 1999–2006,2011
[7] Saran Kumar, A. and Chandrakala, D., 2016. “A Survey on Customer Churn Prediction using Machine Learning Techniques”. International Journal of Computer Applications, 975, p.8887.
[8] J. Hadden, A. Tiwari, R. Roy, D. Ruta, Computer assisted customer churn management: state-of-the-art and future trends,Comput. Oper. Res., 34 (2007) 2902-2917

©IJRASET: All Rights are Reserved



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Call : 08813907089 (24*7 Support on Whatsapp)