ABES ENGINEERING COLLEGE, GHAZIABAD (U.P), INDIA :-
- A P J ABDUL KALAM TECHNICAL UNIVERSITY, LUCKNOW

-:***RESEARCH PAPER:-***

Asst. Professor of Department Of Computer Science Data Science
ABES Engineering College , Ghaziabad
----------@abes.ac.in

CUSTOMER CHURN PREDICTION

Raj Pratap Singh
Department Of Computer Science Data science
ABES Engineering College , Ghaziabad
raj.20b1541038@abes.ac.in

Ms. Dimple, Pandey **Mr. Ravi Shankar Pal, Mr. Amit Kumar**

Abstract-> Customer churn analysis and prediction in telecommunication sector is an issue now a days because it's very important for telecommunication industries to analyze behaviors of various customer to predict which customers are about to leave the subscription from telecommunication company. So machine learning techniques and algorithm plays an important role for companies because its help to retail in buisness in today's commercial conditions because gaining a new customer's cost is more than retaining the existing ones or to gaining to one we can loose more customer . This project focuses on many various machine learning techniques for predicting customer churn through which we can build the classification models such as Logistic Regression, Random Forest and lazy learning and also compare the performance of these models.

This is based on data mining techniques and algorithm plays an important role for companies in commercial conditions because gaining a new customer's cost is more than retaining the existing ones. Mainly the data mining is a part of big data where the data is stored in large form . So, in this prediction we use meta-heuristic algorithm so, this algo predict or mathematically optimize the good solution of complex and difficult problems

The primary objective of churn is to accurate and analysis of customer survival in companies and custromer hazard

function . the gaining of customer churn is give us the data to tell us the time for which each customer was active, and also predict percentage of customer is increase or decrease . the lower percentage of churn , the better will be.

INTRODUCTION-> The customer churn is tell us the percentage of the product of the customer of company who stop buying product at a certain period of time , the customer who cease or stop purchasing a product or service for a given period is referred as **churner**.so, to predict the percentage or analysis we use customer churn prediction. It is mainly use in a telecommunication company, the individual who has opted service from a firm is referred to as Churn .It is also called customer attrition . Many industries build a model like a churn as a common application or prediction for data mining technique. Mobile telephone organizations present across the globe are almost on the verge of building their own churn model. Furthermore, to retain the customers, churn results can be efficiently utilized for various other goals. While performing the analysis using machine learning customer experience tends to provide valuable insights. Some people will change their service providers from time to time because companies can not full fill the service of customer as they like . Increase or decrease in the calling rate will also depend on different job responsibilities. Based on the availability of the data various situations may reflect**.** mainly the churn prediction use the start up companies because the major company have already spread their services and fame like amazon flipkart . they don't need like this prediction but also uses because to use they can analysis the product rate and predict if it is low selling then provide more sale and discount to increase the production .

But in start up company , they have not fame so they do more hard work and always use this production to full fill the needs of customer to increase the selling and production .

1. METHODOLOGY-

Problem Statement-> No business can thrive without it's customers. Means "to run a business we need customer "On the flip side, "customers leaving the business is a nightmare that every business owner dreads! " it is big problem if customer is running away and company will be lost that is big effect to company

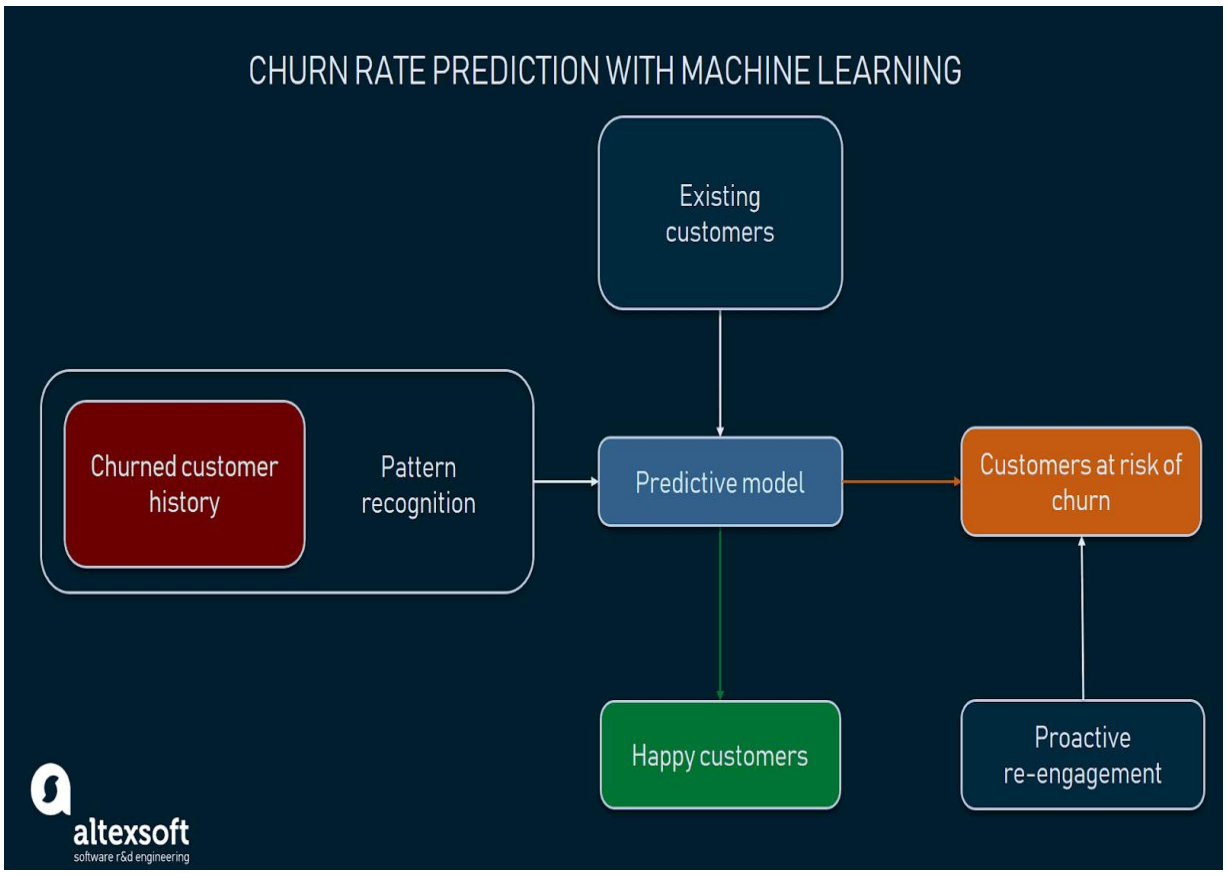
In fact, one of the key metrics to measure a buisness success is by measuring its customer chum rate -the lower the churn, the more customer loved the company . , every user of a product or a service is assigned a prediction value that estimates their state of chum at any given time. This value may be based on multiple factors such as the user's demographic, their browsing behavior and historical purchase data, among other details.

- Now to solve the problem statement we use **dataset ,** by dataset we take data of user of product or service to predict churn

The formula to calculate the customer churn rate:-

$$\text{CCR} = \frac{\text{(customer beginning of month-customer ending of month)}}{\text{Customer beginning of month}}$$

- By using machine learning we can solve the churn rate .So , there are some following method and models which is uses to predict the churn rate .



1. SK learn model selection :->

- In this model, we are splitting the data , first we need to split the dataframe into x, y or any two variable.
- then use train-test-spilt function(data) to split the data
- use random state to have same data each time you run the program
- use stratify to cut the data with the same portion
- Perform the cleaning by calling the cleaning function
- Display all numerical in the data

2. Logistic Regression:-

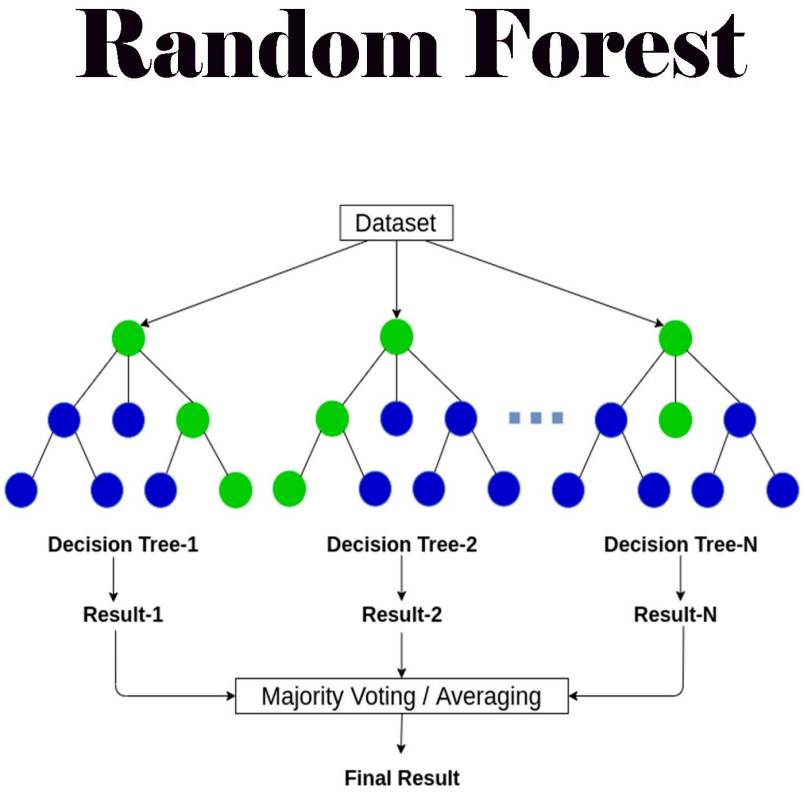
logistic regression is a statistical model and also known as logit model. It is often used to classification and predictive analytic. It is a supervised learning technique which is dependent variable using a given set of independent variables. Logistic regression gives outcome in boolen or binary form like 'yes or no' , 1 or 0, true or false in caterogical or discrete value. It is a linear classification model

For example :-

(a) . To predict whether an mobile no. is spam (1) or (0)
(b) Whether the brain tumor is large (1) or not (0)

3. Random Forest Classifier:-

Random forest classifier is a supervised learning technique which is a popular machine learning . It is a process of combination of multiple classifiers to solve a complex or difficult problem and to improve the performance of the model.Random Forest is a classifier that contains a number of decision nodes on various subsets of the given datasets and takes the average to improve or predict analysis accurately of that dataset. It takes the prediction from each node and based on the majority points of predictions, then it gives the final output. We use random forest because , it takes or make less training time to compare to other algorithm. And it predict the high accuracy in giving output and also it can maintain the accuracy when a delete or large records with the missing value.



3. Gradient Boosting Classifier:-

Gradient boosting algo is a machine learning which is used for regression . so , before using this algorithm we can familiar with adaboost and trade off between bias and varience. Now we have a high measurement data Gradient Boost is used to predict or analyse a continuous value, like Weight. It is Using for Regression is different from doing Linear Regression, so while the two methods are related, don't get confused with each other.

- Consider a data in dataset having different data points and analyse it.
- Now, give equal weight or bias to each of the data points.
- Assume this weight as an input for the model.
- Identify the data points that are in wrong way classified.
- Increase the weight for data points in step 4.
- If you get appropriate output then remove this process else follow steps 2 and 3 again.

It helps us to get a predictive model in form of an whole of weak time to prediction models such as decision nodess. Whenever a decision node performs as a weak learner or understanding then the resulting algorithm is called gradient-boosted trees.

Uses:-

Customer churn prediction is different based on the company's line of business (LoB), operation workflow, and data architecture.

Some use cases for churn prediction are in:

Telecommunication (cable or wireless network segment),
Software as a service provider (SaaS), Retail market,

Subscription-based businesses (media, music and video streaming services, etc.),

Financial institutions (banking, insurance companies, Mortgage Companies, etc.),

Marketing, Human Resource Management (Employee turnover).

REFERENCES:-

https://www.youtube.com/watch%20v=4L2N1AzLDIk&t=750s%20%20%20%20%20%20%20%20:-%20 :-Codejaye

https://ieeexplore.ieee.org/

Abhishek and Ratnesh ,“Predicting Customer Churn Prediction in Telecom Sector"

Anuj and Prabin ,“A Neural Network based Approach for Predicting Customer Churn in Cellular Network Services”

Customer churn prediction in telecom using machine learning in big data platform Abdelrahim Kasem Ahmad, Assef Jafar & Kadan Aljoumaa

Raj Pratap Singh
CSE(DS)
Rollno, :-2000321540043

Raj Pratap Singh
CSE(DS)
Rollno, :-2000321540043