DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING

PROJECT REPORT ON –"Customer churn prediction"

Submitted to:

Mr Saumitra chattopadhyay
(class co-ordinator)

Submitted by:

Rithak
University Roll. No.: 2018645
Class Roll No: 43
CSE-J-VI-Sem
Session: 2022-23

CERTIFICATE

This is certified to be "**Rithak**" Roll no-**2018645** has developed mini project on "Customer churn prediction" for the CSE VI semester Mini project in Graphic Era Hill University, Dehradun during the academic year 2022-23.The project carried out by student is their own work as best of my knowledge.

..... **.....**

MR Saumitra chattopadhyay

Examiner's signature

Class Co-oordinator

CSE Department

GEHU Dehradun

GEHU Dehradun

ACKNOWLEDGMENT

I would like to express our gratitude to the Almighty, the most Beneficent and the most Merciful, for Successful completion of project work .

I wish to thank our parents for their continuing support and encouragement. We also wish to thank them for providing us with the opportunity to reach this far in our studies.

I would like to thank particularly our class co-ordinator Mr saumitra chattopadhyay for their patience, support and encouragement throughout the completion of practical work.

At last but not the least I greatly indebted to all other persons who directly or indirectly helped me during this task.

Rithak
Roll no -2018645
CSE-J-VI-SEM
SESSION-2022-23
GEHU, Dehradun

Table of content

INTRODUCTION

Customer churn prediction refers to the process of using historical customer data and predictive modeling techniques to identify customers who are at risk of ending their relationship with a business or discontinuing their use of its products or services. The objective is to predict which customers are likely to churn in the future, allowing the business to take proactive measures to retain those customers and minimize revenue loss.

Customer churn, also known as customer attrition, can have a significant impact on a business's profitability and growth. By accurately predicting which customers are likely to churn, companies can implement targeted retention strategies, such as personalized offers, improved customer service, or loyalty programs, to increase customer satisfaction and reduce customer churn rates.

To perform customer churn prediction, historical customer data is typically analyzed, including information such as demographics, transaction history, service usage patterns, customer interactions, and feedback. Machine learning algorithms, such as logistic regression, decision trees, random forests, gradient boosting, or neural networks, are applied to build predictive models based on these data.

The predictive models learn patterns and relationships between customer attributes and churn behavior from the historical data. These models are then used to make predictions on new or unseen customer data, indicating the likelihood of churn for each customer. The predictions can be used to prioritize customer retention efforts, allocate resources efficiently, and develop proactive strategies to retain valuable customers.

Customer churn prediction models can provide businesses with valuable insights into customer behavior and enable them to take timely actions to reduce churn rates, improve customer satisfaction, and drive long-term business success.

Problem Statement

- The objective of this project is to build a predictive model that can accurately identify customers who are likely to churn.
- By identifying potential churners in advance, businesses can implement targeted retention strategies and reduce customer attrition.

Data collection

- For this purpose we selected data set from kaggle which is in form of csv file

Exploratory Data Analysis

- Descriptive statistics and visualizations were used to gain insights into the data.
- The distribution of the target variable (churn) was analyzed, and patterns or trends in customer behavior were identified.
- Correlations between variables were explored to uncover potential predictors of churn.
- Univariate Analysis:Univariate analysis focuses on examining individual variables in isolation. It helps to understand the distribution and characteristics of each variable independently.
- Bivariate Analysis:Bivariate analysis involves exploring the relationship between two variables, usually the predictors and the target variable (churn). It helps identify patterns, associations, and dependencies.

Model Building

- Decision Tree Classification:

Decision tree classification is a popular machine learning algorithm used for both binary and multi-class classification problems, including customer churn prediction. It builds a tree-like model of decisions and their possible consequences.

- Random Forest Classification:

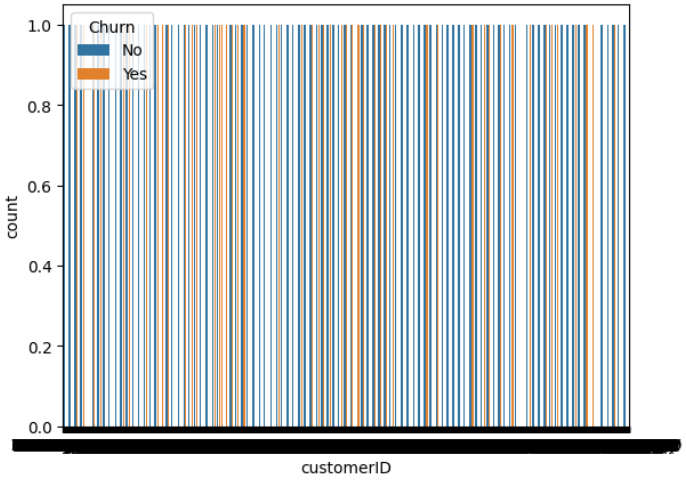
Random Forest is an ensemble learning method that combines multiple decision trees to improve predictive performance and mitigate overfitting.

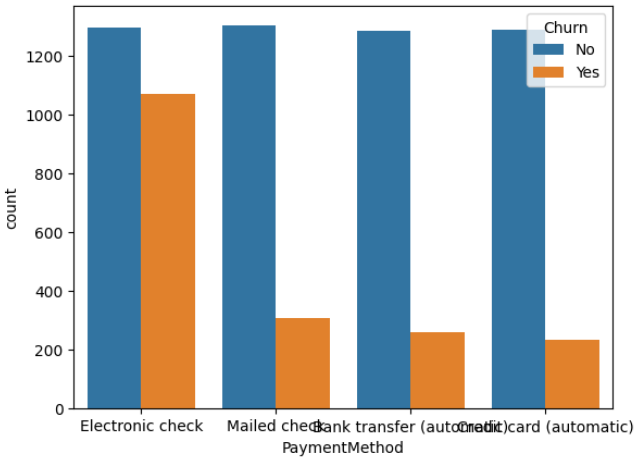
- For building customer churn prediction we use Random forest classification because of its High Predictive Accuracy, Robustness and Feature Importance.

Results and Insights

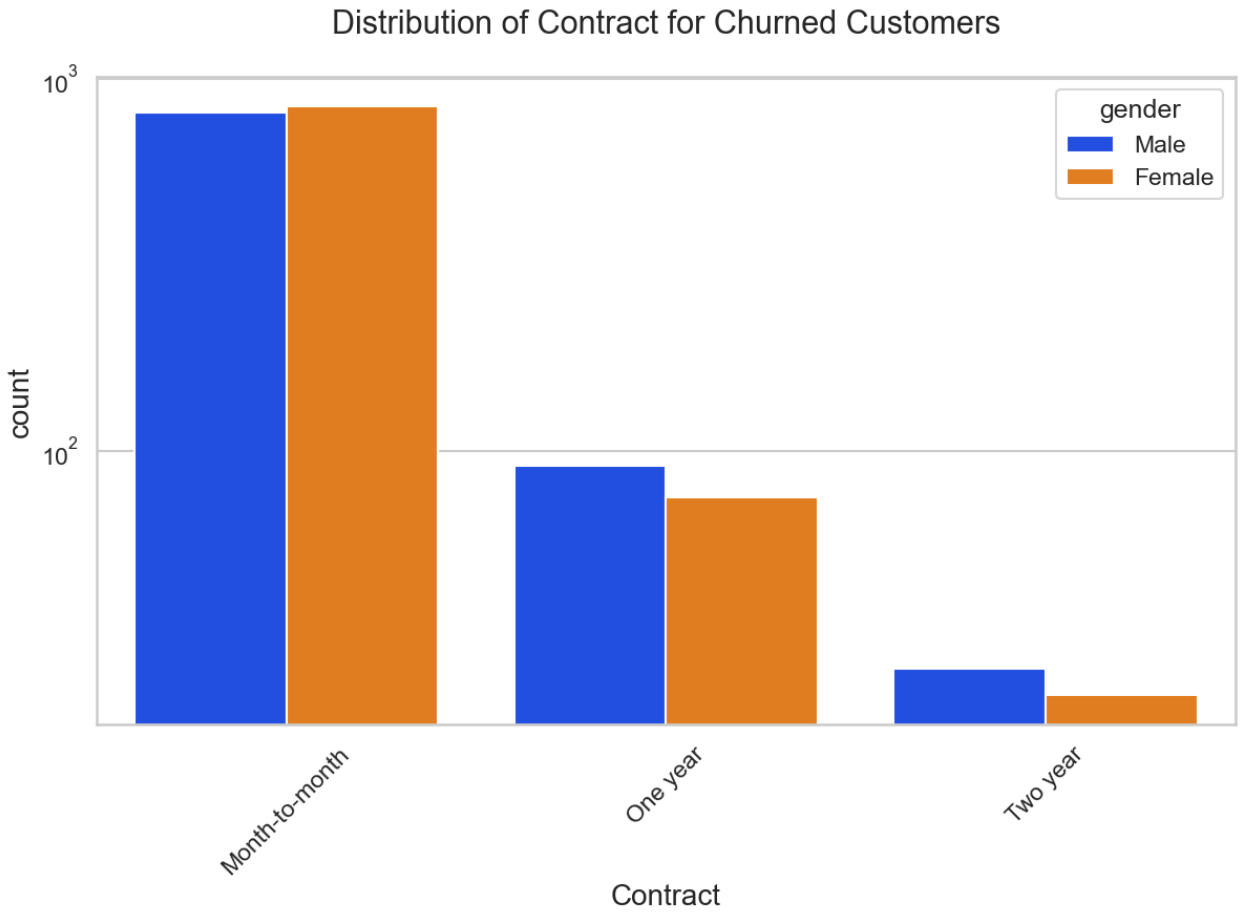
Expolatory data ansaylsis

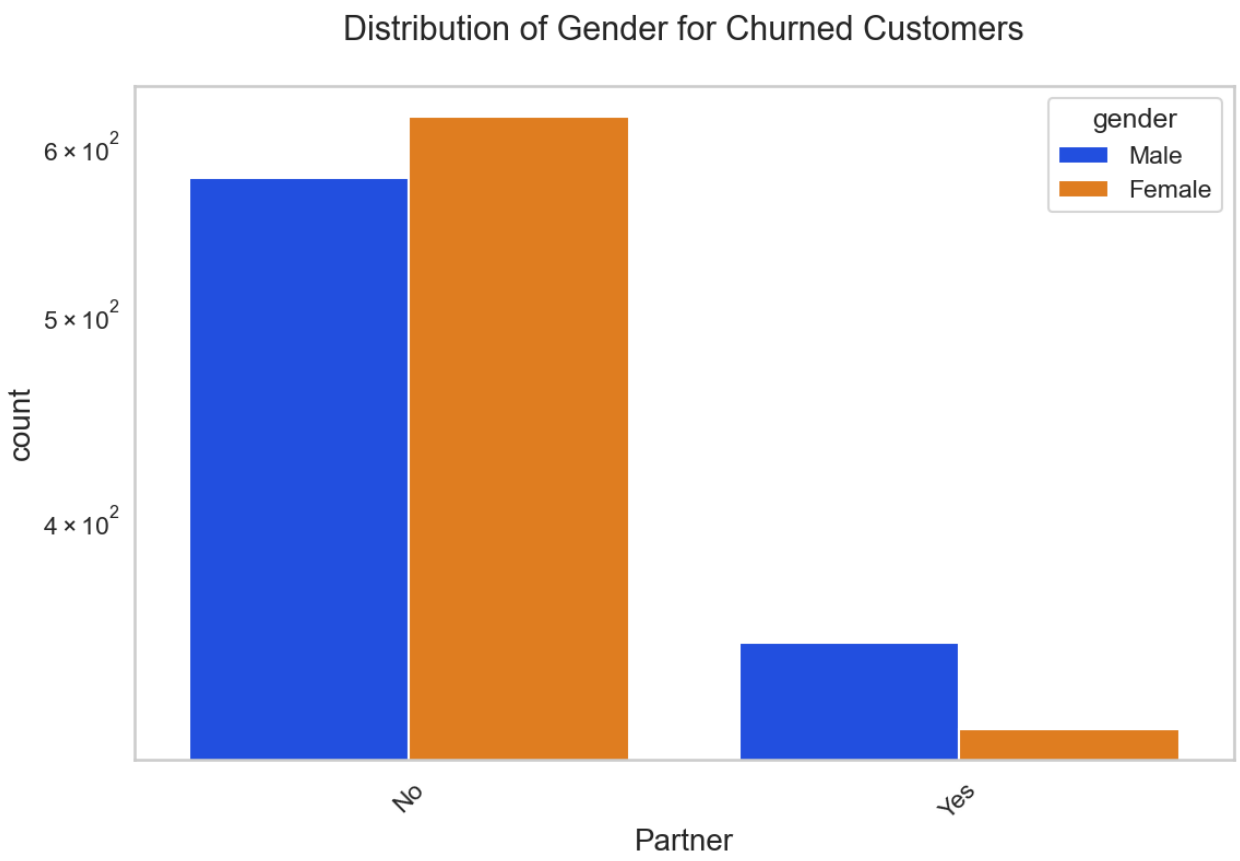
Univariate





Bivarite


Distribution of Contract for Churned Customers
gender
Male
Female
count
10^3
10^2
Month-to-month
One year
Two year
Contract



Models building

Decision tree classification

Classification report

0.934412265758092

	precision	recall	f1-score	support
0	0.97	0.88	0.93	540
1	0.91	0.98	0.94	634
accuracy			0.93	1174
macro avg	0.94	0.93	0.93	1174
weighted avg	0.94	0.93	0.93	1174

confusion matrix

[[477 63]
[14 620]

Random forest classification

Classification report

0.9427350427350427

	precision	recall	f1-score	support
0	0.95	0.92	0.93	518
1	0.94	0.96	0.95	652

accuracy			0.94	1170
macro avg	0.94	0.94	0.94	1170
weighted avg	0.94	0.94	0.94	1170

confusion matrix

[[478 40]
[27 625]]

- with RF Classifier, also we are able to get quite good results, infact better than Decision Tree.

Implemenation and deployement

- We save our model using pickle in python
- And deploy it using python API flask
- Flask is a popular web framework for building APIs in Python. It provides a lightweight and flexible approach to developing web applications and APIs. Here's an example of how to create a basic Flask API:

Conclusion

- The customer churn prediction project successfully developed a model that accurately identifies customers at risk of churn.
- The insights provided by the model have enabled the business to take proactive measures to retain valuable customers and improve overall customer retention rates.
- Ongoing monitoring and continuous improvement of the model will be crucial for maintaining its effectiveness and adapting to changing customer behaviors.