ML PROJECT:
BANK CHURN PREDICTION

SONAL PRIYA-UCE2022647
SEJAL SONAWANE-UCE2022649

PROBLEM STATEMENT

Predicting Bank Customer Churn: Using machine learning techniques to identify and predict the likelihood of customers leaving a bank, based on historical data and customer attributes.

OBJECTIVE

The objective is to develop a predictive model that can accurately forecast customer churn, allowing the bank to proactively implement retention strategies and improve customer satisfaction and loyalty

INTRODUCTION

No business can thrive without it's customers. On the flip side, customers leaving the business is a nightmare that every business owner dreads.
In fact, one of the key metrics to measure a business' success is by measuring its customer chum rate - the lower the churn, the more loved the company is.
Typically, 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.
This value factors in unique and proprietary predictions of how long a user will remain a customer and is updated every day for all users who have purchased at least one of the products/services. The values assigned are between 1 and 5.

TASK

An up-and-coming startup is keen on reducing its customer churn and has hired you as a Machine Learning engineer for this task. As an expert, we are required to build a sophisticated Machine Learning model that predicts the churn score for a website based on multiple features.

D

DATA SET

The dataset consists of parameters such as the user's demographic and Pil details, membership account details, duration and frequency of their visits to the website, reported grievances and feedback, and the like.

DATA SET

```
df.head()
```

Python

	customer_id	Name	age	gender	security_no	region_category	membership_category	joining_date	joined_through_referral	referr
0	fffe43004900440036003000300038 00	Pattie Morrisey	18	F	XW0DQ7H	Village	Platinum Membership	2017-08-17	No	xxx
1	fffe43004900440032003100300035003700	Traci Peery	32	F	5K0N3X1	City	Premium Membership	2017-08-28	?	CID2
2	fffe43004900440031003900320036 00	Merideth Mcmeen	44	F	1F2TCL3	Town	No Membership	2016-11-11	Yes	CID1
3	fffe43004900440036003000330031003600	Eufemia Cardwell	37	M	VJGJ33N	City	No Membership	2016-10-29	Yes	CID
4	fffe43004900440031003900350030003600	Meghan Kosak	31	F	SVZXCWB	City	No Membership	2017-09-12	No	xxx

5 rows × 25 columns

FEATURE SELECTION

customer_id

age

security_no

membership_category

joined_through_referral

preferred_offer_types

internet_option

days_since_last_login

avg_transaction_value

points_in_wallet

offer_application_preference

complaint_status, feedback

Name

gender

region_category

joining_date

referral_id

medium_of_operation

last_visit_time

avg_time_spent

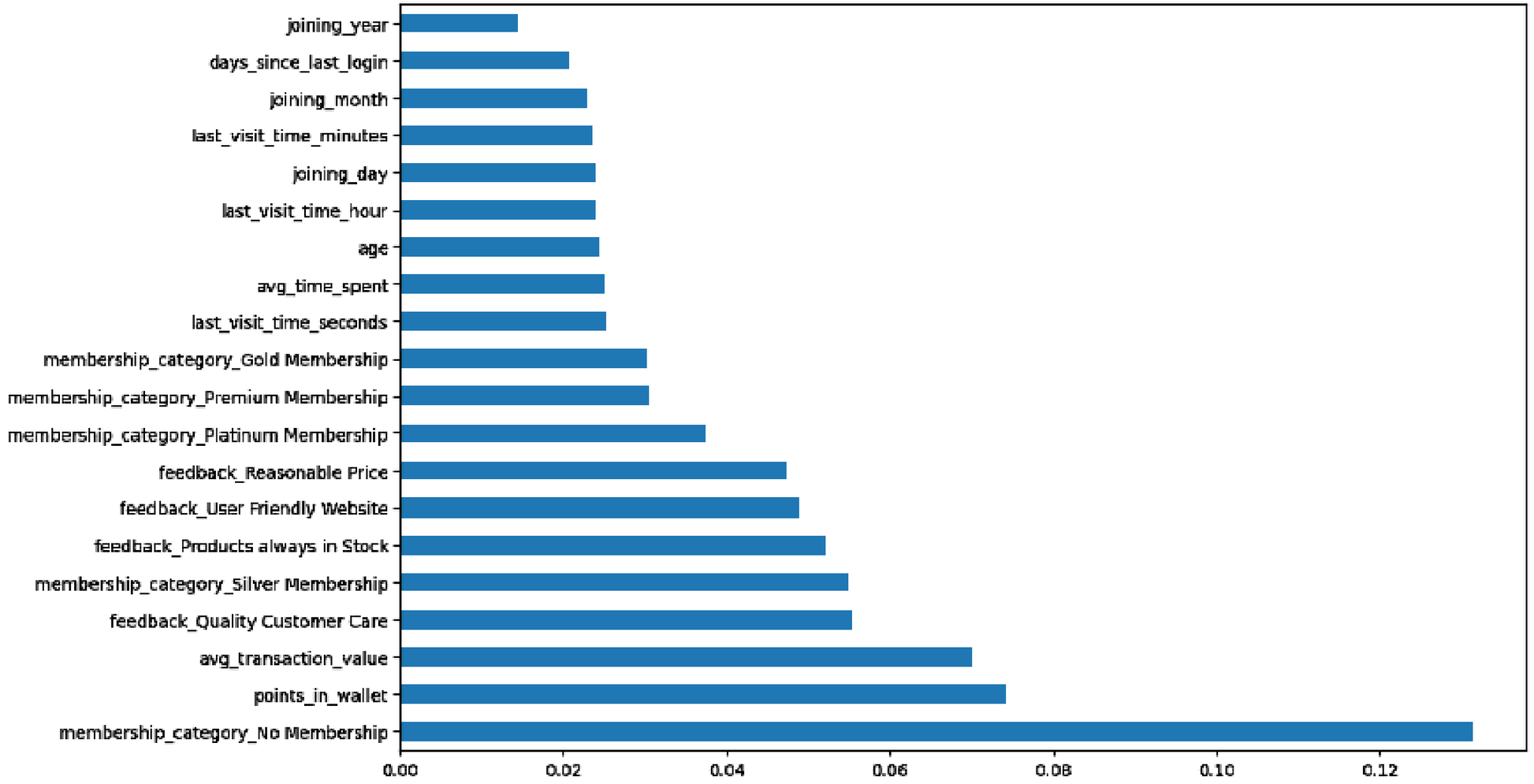
avg_frequency_login_days

used_special_discount

past_complaint

churn_risk_score

IMPORTANT FEATURE



USED MODELS

- Random Forest
- Decision Tree
- XG Boost(maximum accuracy is obtained in XG Boost)

Welcome to State Bank of India

Customer Churn Prediction Service

we understand the critical importance of retaining customers. Every lost customer represents not just a transactional loss, but a potential long-term relationship that could impact your bottom line. That's why we've developed a cutting-edge Customer Churn Prediction Service powered by machine learning algorithms.

Prediction

FLASK AND HTML

app.py - Flask API

- loads models
- receives customer details through GUI
- computes and returns predictions
- returns customer info table

HTML/CSS/Bootstrap

- HTML template and CSS styling to enter customer info.
- jinja - loads data passed by Flask:
  1. predictions
  2.customer info table

DEPLOYMENT

- The API interfacing for the deplyment on Localhost is done using Flask.
- The server is run on Local system during the staging of the project.
- Older deployment was done on Google Cloud Platform

PREDICTIONS

- The final prediction of the model is the range of 1-5 churn for that customer.
- The prediction signifies the chances of the customer to leave the services of the bank which makes the bank to focus more on such such customers and try to retain them using Sales and Marketing strategies about which I have worked in this GitHub module.

CONCLUSION

The implementation of machine learning algorithms like XGBoost, Decision Trees, and Random Forest has significantly improved our ability to predict and manage customer churn. Among these algorithms, XGBoost stands out for its exceptional accuracy and robustness. Its advanced boosting technique enables precise predictions, empowering proactive retention strategies and driving business success.

REFERENCES

- https://scikit-learn.org/stable/modules/generated/sklearn.tree.DecisionTreeClassifier.html
- https://xgboost.readthedocs.io/en/stable/
- https://www.geeksforgeeks.org/random-forest-algorithm-in-machine-learning/

THANK YOU!