Customer Churn Analyzer

Overview - A machine learning-based analyzer that predicts customer churn using service usage, billing, and interaction data to support proactive retention strategies.

Name-Nandni Jaiswal

Roll no – 2312res965

Email ID- nandni_2312res965@iitp.ac.in

Problem Statement & Business Impact

- Problem: Identify customers likely to cancel subscription.

- Business Impact:
- Preventing churn is more cost-effective than acquiring new users.
- Retention strategies improve revenue & customer loyalty.

Dataset Overview

- 1002 rows,10 columns
- Key Features:
  - Contract, SupportCalls, MonthlyBill, PaymentMethod
  - Billing Issues, DataUsageGB, TenureMonths, AutoPay
  - Target: Churn (Yes/No)

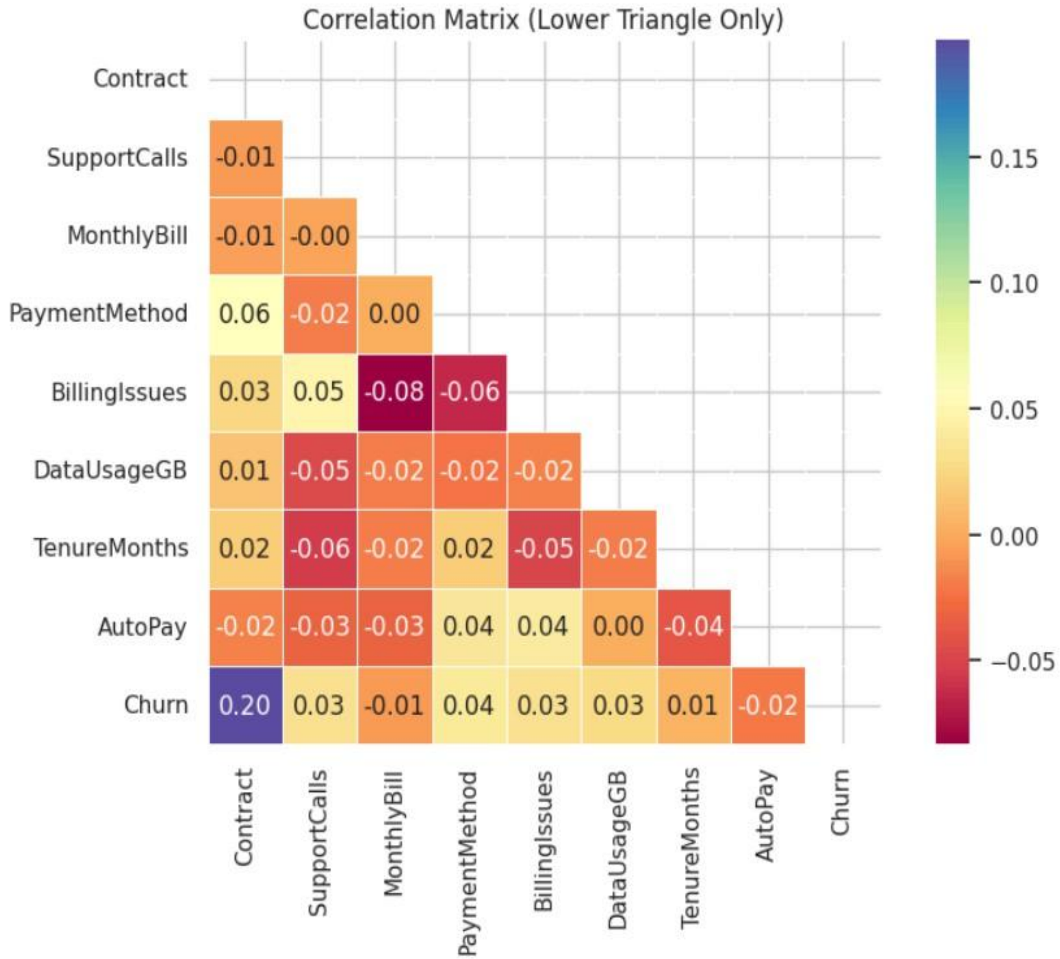
```
  CustomerID Contract  SupportCalls  MonthlyBill PaymentMethod  BillingIssues  \
0       C001  Monthly             5       120.00    CreditCard              1
1       C002   Annual             1        80.00           UPI              0
2       C003  Monthly             5        74.12    CreditCard              1
3       C004   Annual             1        46.44           UPI              0
4       C005  Monthly             4       105.61     DebitCard              0

   DataUsageGB  TenureMonths  AutoPay Churn
0        60.00             6        0   Yes
1        95.00            24        1    No
2       105.81            36        1   Yes
3        74.44            21        0    No
4       105.19            16        1   Yes
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 1002 entries, 0 to 1001
Data columns (total 10 columns):
 #   Column         Non-Null Count  Dtype
---  ------         --------------  -----
 0   CustomerID     1002 non-null   object
 1   Contract       1002 non-null   object
 2   SupportCalls   1002 non-null   int64
 3   MonthlyBill    1002 non-null   float64
 4   PaymentMethod  1002 non-null   object
 5   BillingIssues  1002 non-null   int64
 6   DataUsageGB    1002 non-null   float64
 7   TenureMonths   1002 non-null   int64
 8   AutoPay        1002 non-null   int64
 9   Churn          1002 non-null   object
dtypes: float64(2), int64(4), object(4)
memory usage: 78.4+ KB
None
Churn
No     528
Yes    474
Name: count, dtype: int64
```

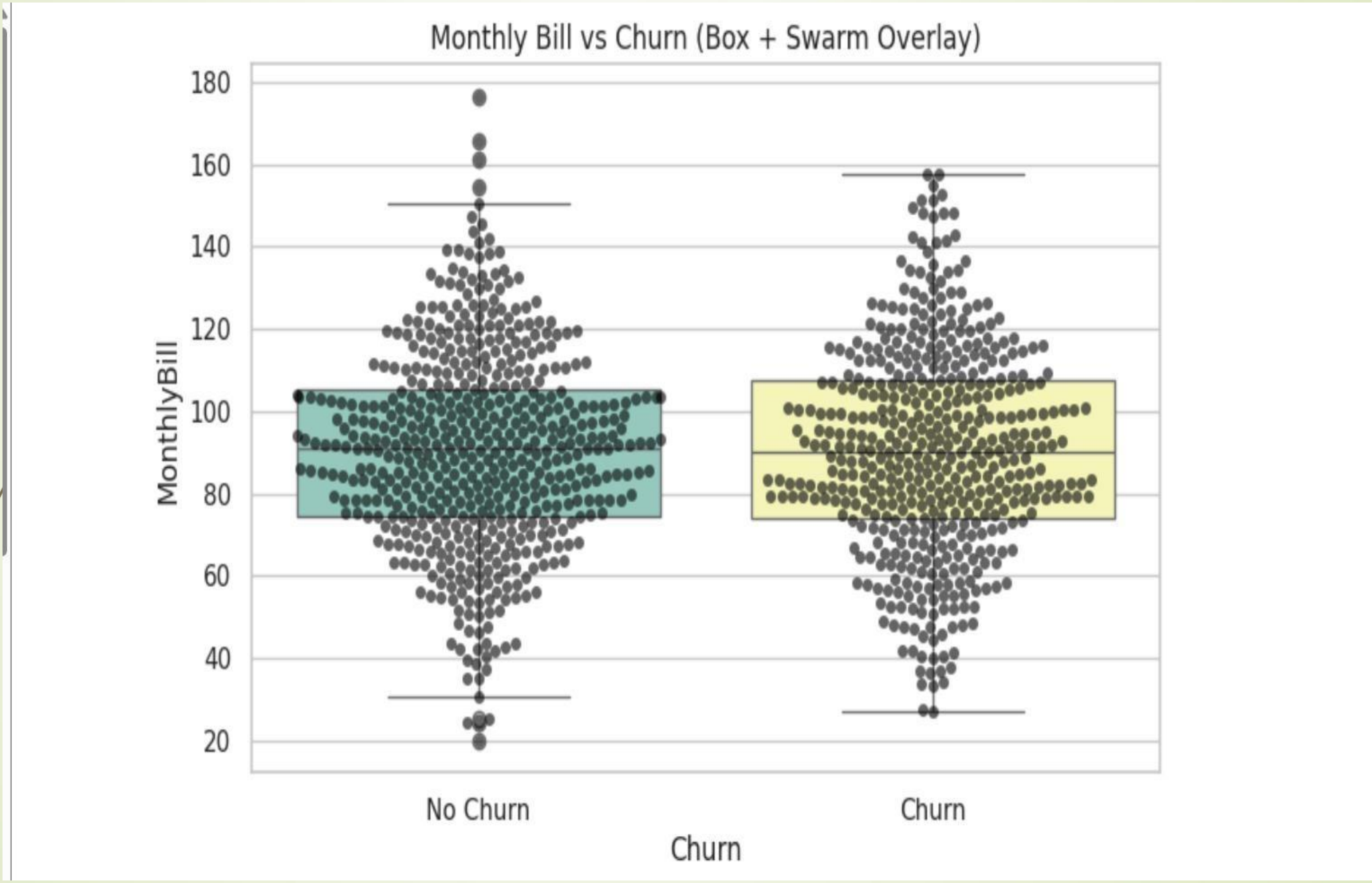
EDA - Correlation Heatmap

9/8/25, 12:26 PM

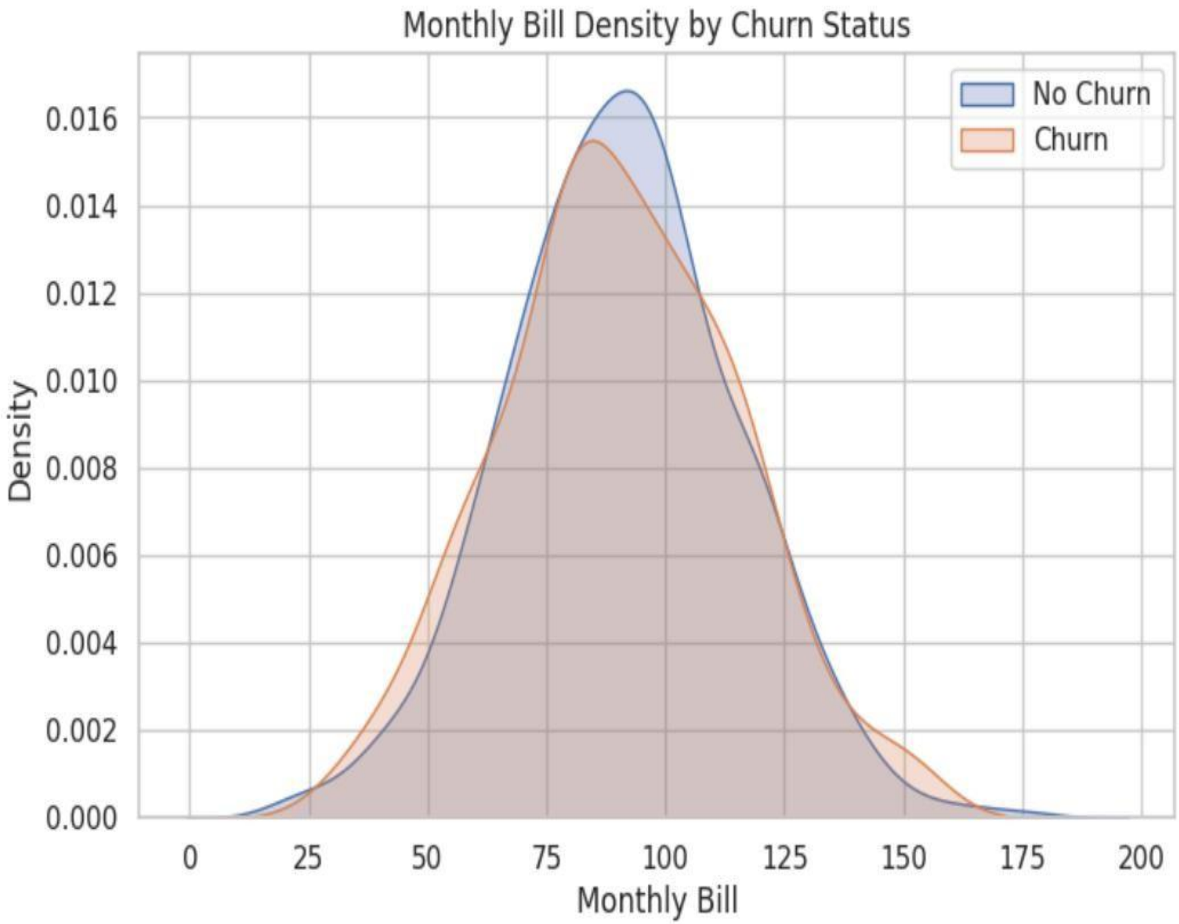
Customer Churn Analyzer.ipynb - Colab



EDA - Monthly Bill vs Churn

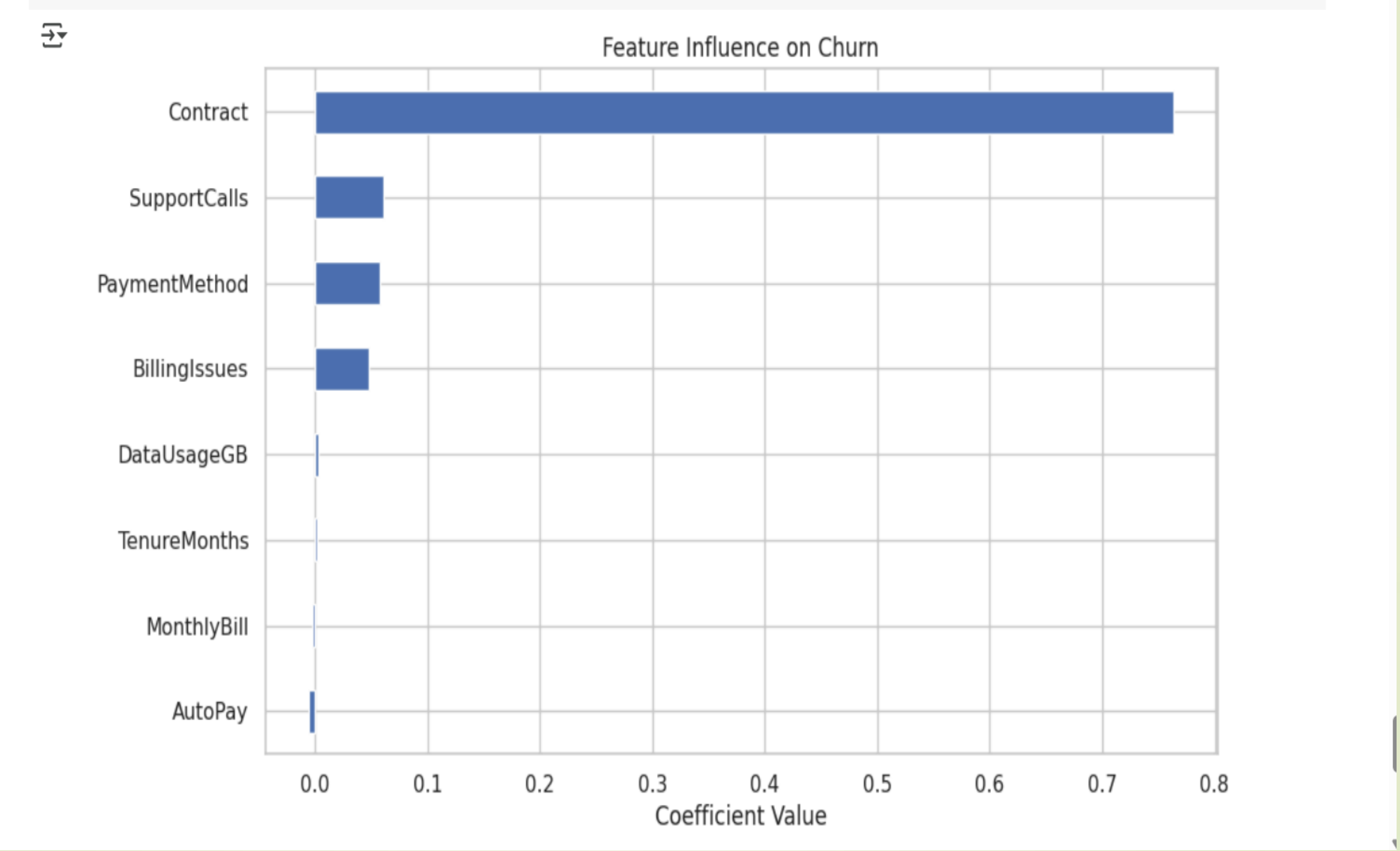


EDA - Monthly Bill Density by Churn



Avg Monthly Bill (Churned): ₹90.37
Avg Monthly Bill (Retained): ₹90.71

Feature Importance



Model Evaluation

- Confusion Matrix:
- [[89 69]
- [69 74]]

Performance:

- Accuracy: 54%
- F1 Score: 0.52
- Precision/Recall ~0.52-0.56

```
Confusion Matrix:
 [[89 69]
 [69 74]]

Classification Report:
              precision    recall  f1-score   support

           0       0.56      0.56      0.56       158
           1       0.52      0.52      0.52       143

    accuracy                           0.54       301
   macro avg       0.54      0.54      0.54       301
weighted avg       0.54      0.54      0.54       301

F1 Score: 0.5174825174825175
```

Prediction Example

❖ **Input:**

- Contract: Monthly, Support Calls: 4, Monthly Bill: 110, Payment Method: Credit Card,
- Billing Issues: 0, Data Usage GB: 85, Tenure Months: 12, Auto Pay: Yes
- Prediction: Churn = **Yes**

Conclusion & Future Work

- Logistic Regression baseline achieved ~54% accuracy (F1 = 0.52)

- Future Improvements:
  - Try Decision Trees, Random Forest, XGBoost
  - Hyperparameter tuning
  - Feature engineering
  - Handle class imbalance


THANK YOU!