CUSTOMER CHURN PREDICTION

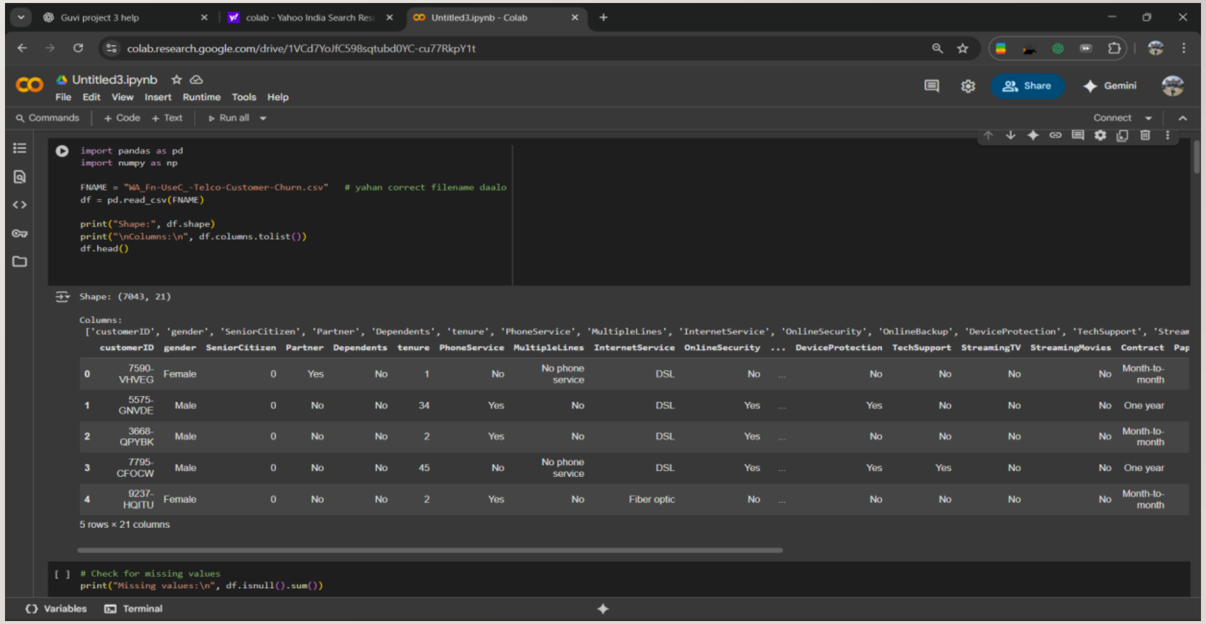
PROBLEM DEFINITION

- The objective of this project is to predict customer churn in a telecom company. Customer churn refers to customers leaving the company. By predicting churn, businesses can take preventive actions to retain customers.

Problem

DATASET OVERVIEW

- \- Dataset: Telco Customer Churn (7043 rows, 21 columns)
- \- Features: Demographics, services subscribed, billing information
- \- Target: Churn (Yes/No)



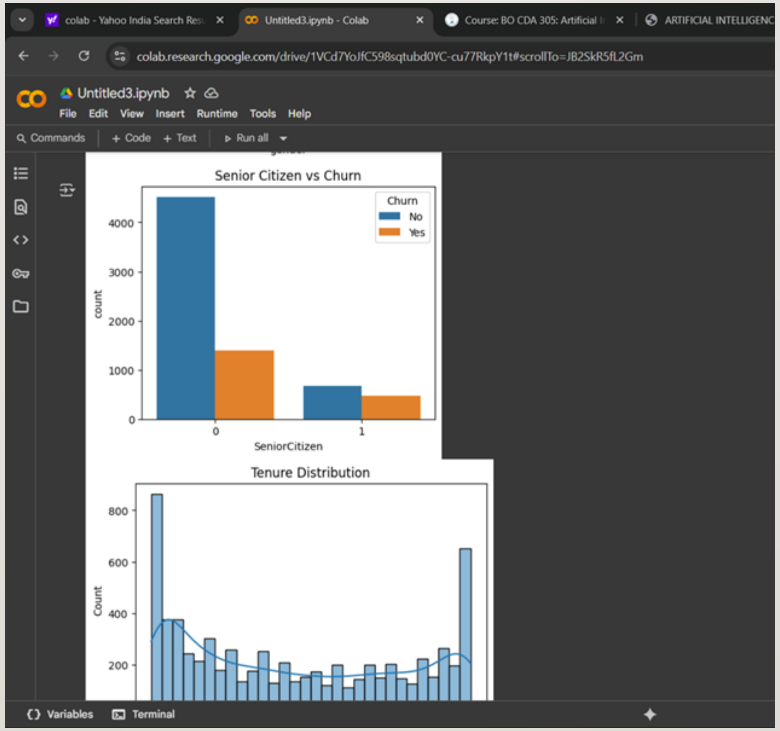
DATA PREPROCESSING

- - Handled missing values
- - Converted categorical data to numerical (Label Encoding)
- - Train-test split applied (80-20)
- Standardized numerical features

EXPLORATORY DATA ANALYSIS (EDA)

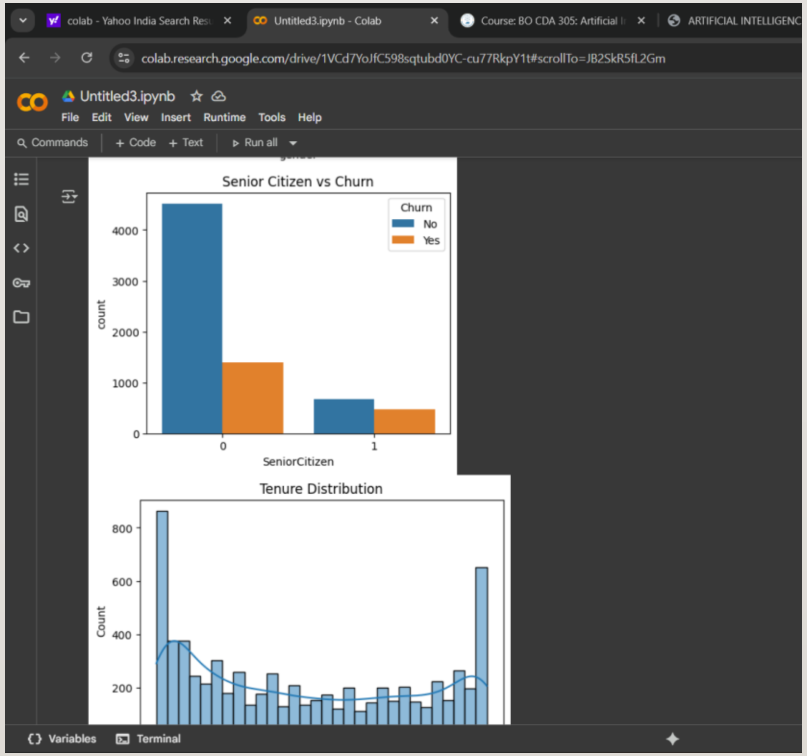
- Following visualizations were created:
- 1. Churn Distribution
- 2. Gender vs Churn
- 3. Senior Citizen vs Churn
- 4. Tenure Distribution
- 5. Monthly Charges vs Churn
- 6. Correlation Heatmap

EDA - SENIOR CITIZEN VS CHURN



Senior citizens have a higher churn rate compared to younger customers.

EDA - TENURE DISTRIBUTION



Customers with shorter tenure are more likely to churn.

MODEL BUILDING

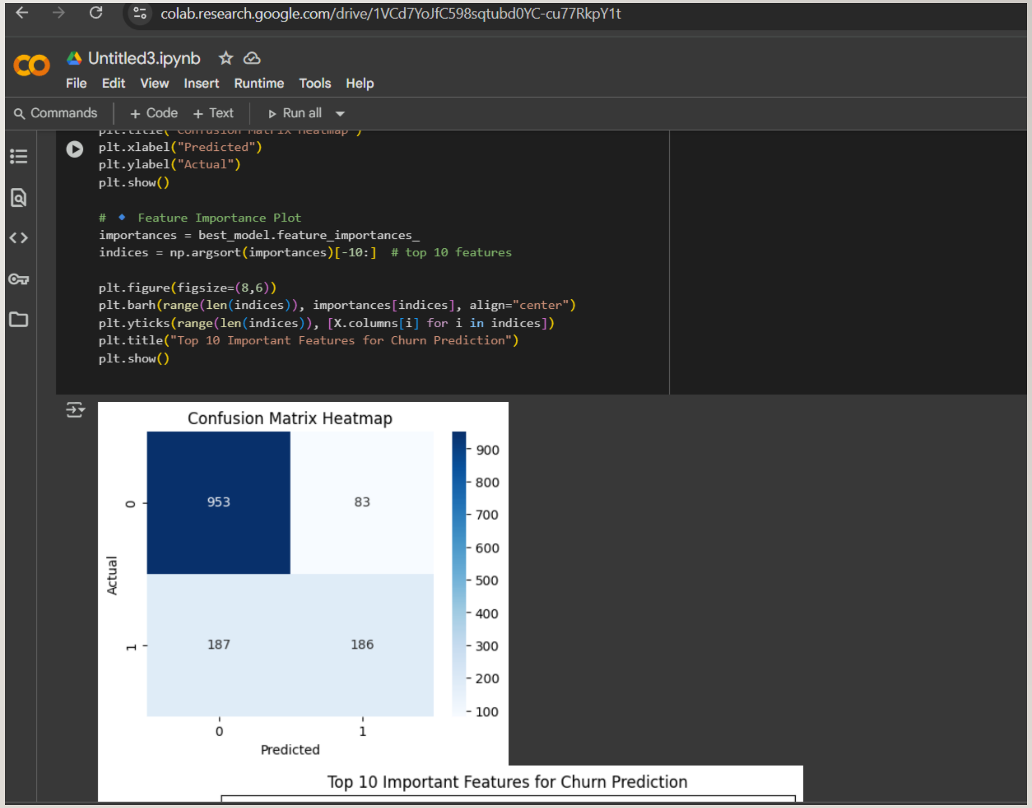
- - Algorithm Used: Random Forest Classifier
- - Hyperparameter Tuning: GridSearchCV
- - Best model selected with optimal parameters
- Model performs well for predicting non-churn customers, but recall for churn class is moderate (~50%).

EDA - CHURN DISTRIBUTION

- Insert Graph Here
- Insight: Around 26% customers churn, showing imbalance in dataset.

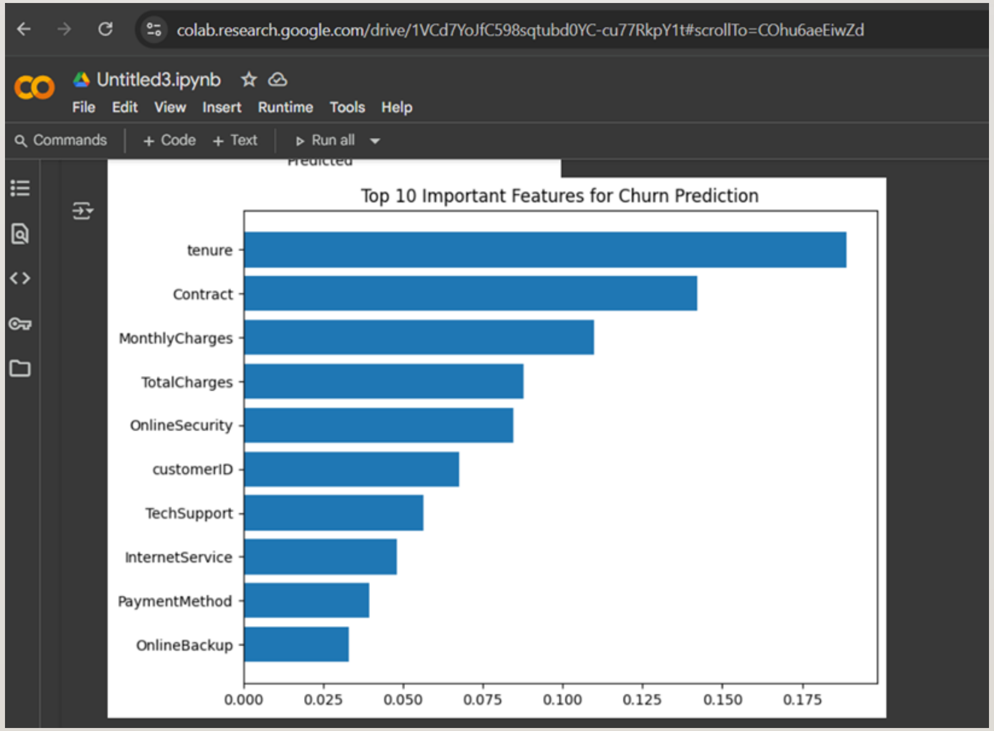
MODEL EVALUATION

- \- Metrics: Accuracy, Precision, Recall, F1-score
- \- Confusion Matrix Heatmap



Model performs well for predicting non-churn customers, but recall for churn class is moderate (~50%).

CONFUSION MATRIX



Contract type, tenure, and monthly charges are the most important drivers of churn.

FEATURE IMPORTANCE

- feature importance analysis was performed using the Random Forest model to identify the most significant factors that influence customer churn. The results highlight that **tenure, monthly charges, contract type, payment method, internet service, and total charges** play a key role in determining whether a customer is likely to leave the telecom company. For instance, customers with shorter tenure and higher monthly charges showed a higher probability of churn. Similarly, customers on **month-to-month contracts** or those using **electronic check payments** were more likely to discontinue services compared to long-term contract holders. Identifying these critical features provides actionable business insights, allowing the company to design targeted retention strategies such as offering discounts, promoting long-term contracts, or improving service quality for high-risk groups. This step ensures that the churn prediction model is not only accurate but also valuable for real-world business decision-making.

RESULTS & INSIGHTS

- - Accuracy: ~80–82%
- - Model predicts non-churn customers better
- - Senior Citizens and customers with high monthly charges churn more
- - Insights help design loyalty programs and offers

CONCLUSION

- ✔️ Successfully predicted customer churn
- ✔️ Identified key factors influencing churn
- ✔️ Provides actionable insights for business retention strategies

- The churn prediction model provides actionable insights to reduce customer loss. With 82% accuracy, it helps identify at-risk customers. Future improvements include testing deep learning models and using real-time data.