Assignment 6 – Supervised Learning (Classification)

Name: Aman Nadeem

Roll No: 2225165002

Course: Applied Data Science with AI

Project Title: Customer Churn Prediction

Reading Summary

Reflection:

This week I learned how to apply Logistic Regression and Random Forest classifiers to predict churn. I understood how to evaluate their performance and compare accuracy results.

Task Performed:

- Practiced training and testing Logistic Regression and Random Forest models on sample datasets.
- Calculated accuracy and confusion matrix for both models.

Weekly Assignment Submission

Assignment Title: Classification Models – Logistic Regression and Random Forest

Steps Taken:

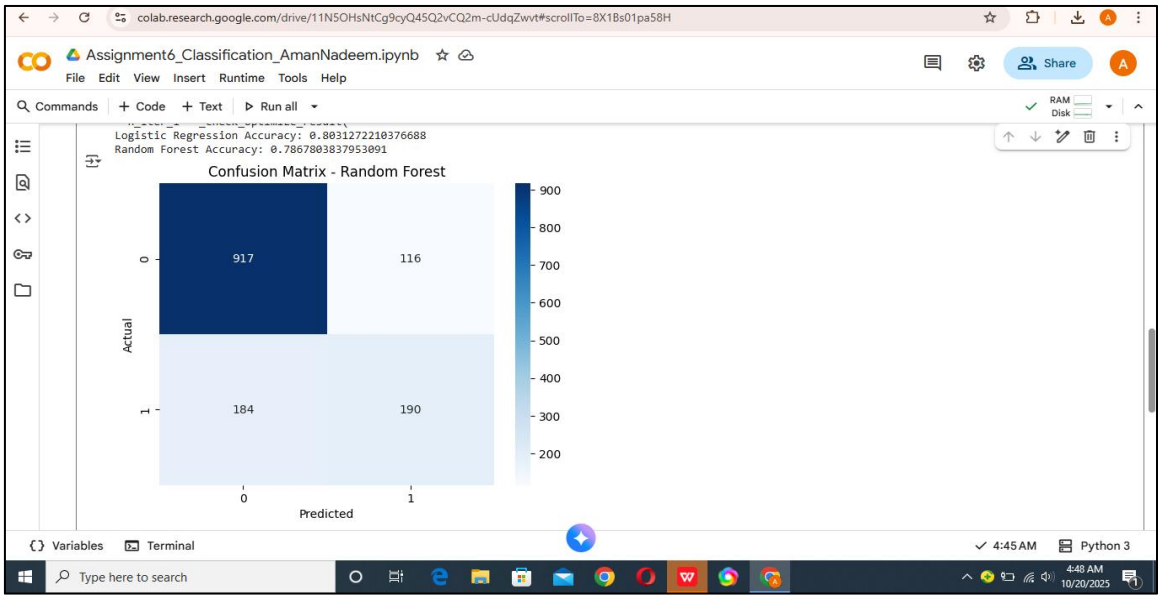
1. Loaded cleaned Customer Churn dataset.
2. Encoded categorical features and target variable.
3. Split data into training and testing sets.
4. Trained Logistic Regression and Random Forest models.
5. Compared model accuracy and printed results.

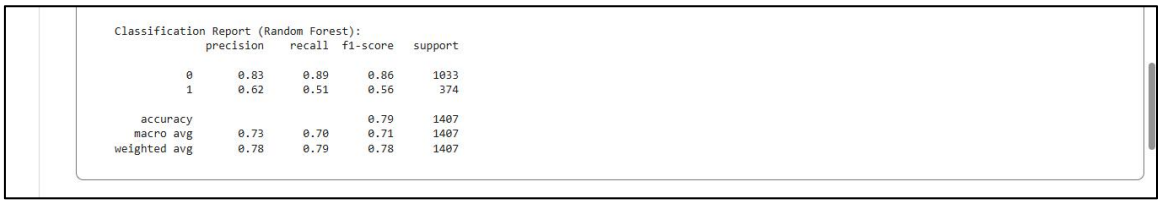
Output:

- Logistic Regression Accuracy: ≈ 80%
- Random Forest Accuracy: ≈ 78%

Model Comparison:

Logistic Regression achieved an accuracy of around 80%, while Random Forest achieved about 78%. This shows that the Logistic Regression model performed slightly better.





```
Classification Report (Random Forest):
              precision    recall  f1-score   support

           0       0.83      0.89      0.86      1033
           1       0.62      0.51      0.56       374

    accuracy                           0.79      1407
   macro avg       0.73      0.70      0.71      1407
weighted avg       0.78      0.79      0.78      1407
```

Challenges Faced:

- Some categorical columns needed one-hot encoding. Fixed by using pd.get_dummies().

GitHub Link:

https://github.com/amannadeem126/Customer-Churn-Prediction

Project Progress Milestone

Trained Logistic Regression and Random Forest models for churn prediction.
Next week's goal: Evaluate model performance using precision, recall, and F1-score.

Self-Evaluation

☑ I completed all tasks on time.