ASPIRENEX

Data Science
Internship Presentation

Tanveer Ahmad

Final Year, B.Tech - CS

ABES Engineering College

Task 1 : Credit Card Fraud detection

Objective: Develop a machine learning model to detect fraudulent credit card transactions.

Goal: Ensure accurate detection of fraudulent transactions to prevent customer charges for unauthorized purchases.

Dataset Link :

GitHub Link :

Dataset: European transactions, Sept 2013; 492 frauds out of 284,807 (0.172% fraud rate); PCA features, 'Time', 'Amount'.

Data Preparation: Cleaned data, under-sampled for balance, used PCA components, 'Time', and 'Amount'.

Model Building: Logistic Regression; 80/20 train-test split with stratification.

Model Evaluation: Training accuracy: **92.86%**; Test accuracy: **90.0%**; Evaluated with accuracy score, confusion matrix, classification report.

Task 2 : Movie Rating Prediction with Python

Objective: Develop a model to predict movie ratings using features like genre, director, and actors.

Goal: Analyze historical movie data to create an accurate rating prediction model.

Dataset Link :

GitHub Link :

Data Preparation: Used Pandas, NumPy, Matplotlib, Seaborn, Scikit-Learn; cleaned IMDb Movies India dataset.

EDA: Analyzed ratings per year, top directors, popular genres, and duration vs. ratings.

Feature Engineering: Created features based on average ratings for genre, director, and actors.

Model Building: Linear Regression; predictors: Year, Duration, Votes, Genre Mean Rating, Director Rating, Actor Ratings; 70/30 train-test split.

Model Evaluation: MSE: 0.3529, MAE: 0.4431, R2: 0.7964; model explains 79.64% of rating variance; future work to refine model and explore advanced algorithms.

Thank you!