MODULE 10 PROJECT

OANKAR SANTOSH GHORPADE

DATE:25/08/2025

EXECUTIVE SUMMARY

- This project demonstrates the complete workflow from data collection and wrangling to exploratory data analysis (EDA), interactive visualizations, and predictive modeling.

INTRODUCTION

- The purpose of this project is to apply data science techniques to analyze a dataset, gain insights through EDA, and build predictive models for classification tasks.

DATA COLLECTION & WRANGLING METHODOLOGY

- Data was collected from provided CSV datasets and underwent preprocessing including cleaning, handling missing values, encoding categorical variables, and standardization.

EDA & INTERACTIVE VISUAL ANALYTICS METHODOLOGY

- Exploratory Data Analysis was performed using statistical summaries, SQL queries, and visualization tools. Interactive analytics were developed using Plotly and Folium.

PREDICTIVE ANALYSIS METHODOLOGY

- We applied machine learning models including Logistic Regression, SVM, Decision Trees, and KNN. Hyperparameters were tuned using GridSearchCV with cross-validation.

EDA RESULTS WITH VISUALIZATIONS

- Various plots such as histograms, scatterplots, bar charts, and correlation heatmaps were generated to uncover trends and distributions.

EDA RESULTS WITH SQL

- SQL queries were used to extract insights such as launch success rates, mission outcomes, and relationships between categorical variables.

INTERACTIVE MAP WITH FOLIUM RESULTS

- A Folium map was created to visualize launch site locations and outcomes interactively.

PLOTLY DASH DASHBOARD RESULTS

- A Plotly Dash interactive dashboard was developed to integrate dynamic charts and filters for real-time exploration of the dataset.

PREDICTIVE ANALYSIS RESULTS

- Each classification model was evaluated using cross-validation and test sets. Logistic Regression, SVM, Decision Trees, and KNN achieved varying accuracy scores, with the best method selected based on performance.

CONCLUSION

- The project successfully demonstrated end-to-end data analysis. EDA revealed key insights, and predictive analysis achieved strong accuracy. Interactive dashboards enhanced user understanding.

CREATIVITY & INSIGHTS

- Innovations included interactive dashboards, Folium maps, and advanced hyperparameter tuning. Additional insights were derived from combining EDA with predictive models.