PROJECT REPORT

Executive Summary

This project aimed to predict the successful landing of SpaceX's Falcon 9 first stages using machine learning models. Data was collected from web scraping, followed by preprocessing to prepare it for analysis. Various classification algorithms, including Logistic Regression, Support Vector Machines, Decision Trees, and K-Nearest Neighbors, were applied. The models were fine-tuned using GridSearchCV, and performance was evaluated with a confusion matrix.

Results showed that one of the models outperformed the others in predicting successful landings, helping to understand key factors influencing outcomes. The project highlights the potential of machine learning in optimizing rocket launches, with an interactive Dash application providing insights through data visualization.

Methodology

- The data was collected by web scraping APIs from SpaceX website.
- Data was standardized for Machine Learning Algorithms and the target variable was converted into a binary variable was success and failure as part of data pre-processing.
- Categorical variables such as launch site and rocket type were converted into numerical formats using techniques like one-hot encoding, enabling the machine learning algorithms to process these features.
- Exploratory data analysis (EDA) was performed using visualization libraries in python to determine important features and relationship among the features.
- Interactive visual analytics was done using Folium and Plotly Dash.
- Predictive analysis using classification models was done to determine best model for prediction.

Results and Conclusion

Exploratory data Analysis and feature engineering resulted in choosing the important features. The Orbit type, Payload mass, launch site were deemed to be detrimental in launch outcome. The Map with Folium revealed the close proximity of water body with launch pad. This makes sense as launch failure will result in catastrophic downfall of rocket and water bodies are a safe space to dispose a rocket.

Classification algorithms revealed that SVM classifier performs best on the data set. It achieved an accuracy score of 1.0 on test data with best parameters.

Other algorithms also performed very well with all their accuracy being over 0.9. Overfitting was not evident in the models as each models improved accuracy on the test set.

PROJECT REPORT

Future Applications

- **Improved Mission Planning**: SpaceX launch prediction models can optimize mission planning by predicting potential launch success and identifying the best launch windows based on weather, technical conditions, and past performance.
- **Cost Optimization**: These models can help SpaceX minimize costs by predicting the likelihood of launch delays or failures, enabling better resource allocation and more efficient logistics.
- **Enhanced Safety Measures**: By identifying patterns in failed launches, prediction models can enhance safety protocols, reducing the risk of accidents and improving the reliability of rocket landings.