

Project Report

Project: Analysis of SpaceX Falcon 9 Dataset

Dhiraj H. Gawhare

Roll Number: 22M0062

M. Tech (Aerodynamics)

IIT Bombay.

Project Report: Analysis of SpaceX Falcon 9 Dataset

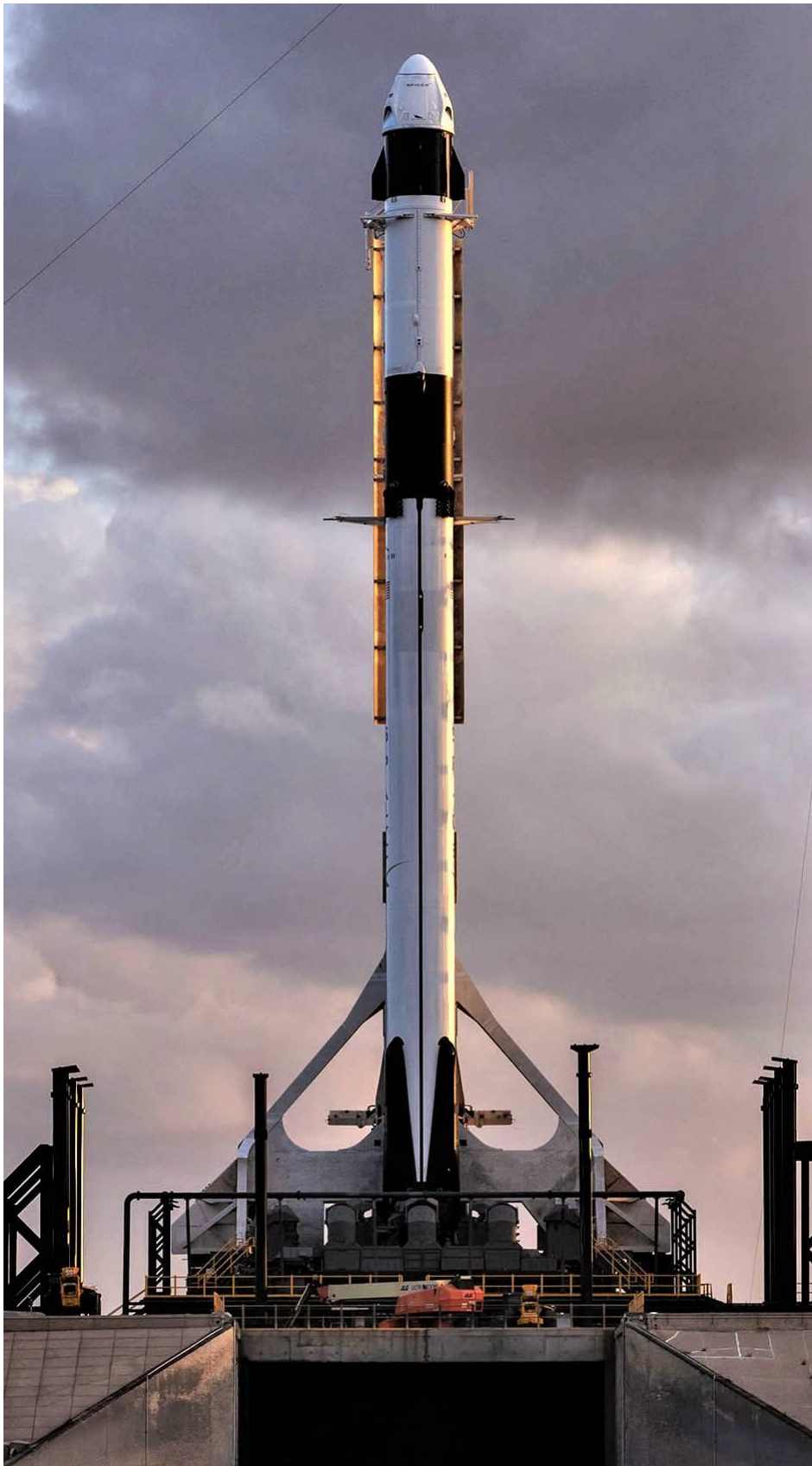


1. Introduction

The aim of this project was to perform a comprehensive analysis of the SpaceX Falcon 9 dataset. The dataset contains information about various attributes related to Falcon 9 rocket launches. The project involved several stages, including Exploratory Data Analysis (EDA), feature engineering, interactive visual analysis, and the implementation of machine learning models.

2. Data Overview

Briefly describe the dataset, including the number of instances, features, and the nature of the data.



3. Methodology

3.1 Exploratory Data Analysis (EDA)

Provide an overview of the EDA process, including the following points:

- Summary statistics of key features.
- Distribution of important variables.
- Identification of missing values and data cleaning strategies.
- Insights gained from initial visualization.
-

3.2 Feature Engineering

Explain the feature engineering techniques applied to enhance the quality of the dataset:

- Creation of new relevant features.
- Transformation of existing features.

3.3 Interactive Visual Analysis

Detail the interactive visualizations created using Folium and Plotly:

- Purpose of visualizations (e.g., geographical distribution, launch success over time).
- Interactivity and user engagement.
- Insights derived from the visualizations.

3.4 Machine Learning Model Implementation

Describe the machine learning algorithms implemented:

- SVM, Logistic Regression, Decision Trees, and KNN.
- Explain the rationale behind choosing these algorithms.

3.5 Model Performance Evaluation

Discuss the accuracy metrics and evaluation process used to assess model performance:

- Accuracy scores of each algorithm.
- Justification for selecting accuracy as the evaluation metric.

3.6 Hyperparameter Tuning

Explain the process of hyperparameter tuning using GridsearchCV:

- Parameters tuned for each algorithm.
- Impact of hyperparameter tuning on model performance.

4. Results

Present the results obtained from the analysis:

- Accuracies achieved by each algorithm.
- Maximum accuracy achieved after hyperparameter tuning.
- Comparison of the models' performance.

5. Conclusion

Summarize the key findings and insights from the project:

- Highlight the importance of hyperparameter tuning in improving model performance.
- Discuss the most effective algorithm based on results and tuning efforts.
- Emphasize the role of interactive visualizations in gaining insights.

6. Future Work

Suggest potential areas for further exploration and improvement:

- Incorporating more advanced machine learning techniques.
- Collecting additional features to enhance prediction accuracy.
- Exploring other evaluation metrics beyond accuracy.

7. Acknowledgments

Acknowledge any external resources, datasets, or libraries used in the project.