

## Proposal

### Problem Statement Formation

The business problem under investigation is to explore the dependencies of the "Status Mission" feature, serving as the target variable, on other influential variables in a dataset related to global space launches. I will predict rocket launch success based on features such as launch location, company name, rocket specifications, and mission details, including the rocket cost. The aim is to identify specific factors influencing mission statuses and their impact on successful mission outcomes.

### Context

Exploring mission statuses in global space launches is essential for understanding the dynamics and challenges faced by space agencies and private companies in achieving successful missions. By analyzing mission statuses, stakeholders can gain insights into the factors contributing to mission success or failure, thereby informing strategic decision-making and resource allocation for future space endeavors. Descriptive statistics of the dataset provide sufficient quality and detail for effective feature engineering and predictive modeling.

### Criteria for Success

The key criteria for success in this project include developing predictive models with high accuracy in forecasting mission statuses based on relevant variables. Additionally, uncovering actionable insights and patterns in the data will contribute to a deeper understanding of mission outcomes. I will provide clear and interpretable findings that space agencies, aerospace companies, and government entities can use to optimize mission planning, execution, and risk mitigation strategies.

### Scope of Solution Space

This business initiative will exclusively investigate the relationship between mission statuses and various factors, such as launch location, company name, rocket specifications, and mission details. However, the rocket cost feature has missing values in approximately 78% of the rows, posing challenges in analysis. Addressing this issue may require extensive efforts to search multiple sources and append the cost via the mission's name, potentially involving more complex algorithms to scrape the web from diverse sources.

During initial exploratory data analysis, I observed a class imbalance in the dataset, with a majority-to-minority class split of approximately 9:1. Consequently, I will employ techniques such as oversampling the minority class to address this imbalance.

### Constraints within Solution Space

Several constraints may impact the success of this business initiative, starting with the availability and quality of data, which may hinder the accuracy and reliability of predictive models. Secondly, external events such as weather conditions, geopolitical tensions, and

technical failures may influence mission outcomes, adding complexity to the analysis. Lastly, limited time, budget, and expertise may restrict the depth and breadth of the analysis and modeling techniques employed.

#### Stakeholders to Provide Key Insight

Key stakeholders involved in this project include space agencies, aerospace companies, government entities, and research institutions. They are concerned with optimizing mission success rates while minimizing risk and resource allocation. Data will be sourced from reputable sources such as space mission databases, government agencies, and industry reports. The findings and recommendations will be presented to senior management, mission planners, and data science teams responsible for space mission analytics and decision-making.

#### Key Data Sources

The primary data sources for this project include space mission databases providing information on launch dates, mission objectives, and outcomes, as well as government and industry reports offering insights into rocket specifications, launch locations, and company profiles. Additionally, natural language text data from mission details will provide qualitative information on mission parameters, challenges, and achievements.

#### Summary

In summary, this project aims to leverage data analytics and predictive modeling techniques to uncover patterns and insights in global space launch data. The goal is to enhance mission planning, execution, and success rates in the space exploration industry.