# Space Mission Failure Prediction

Ruizi Wang, Nathan Anye, Tianning Li, Maryam Kameli

## 1 Introduction

Space exploration is a risky task that depends on various factors. It was always a mystery if the lunch ahead would be a success or not. By analyzing the factors that contribute to successful missions, we can gain valuable insights. This project leverages historical space mission data, data analysis techniques, and machine learning algorithms to uncover patterns that could potentially predict the failure of future launches.

## 2 Motivation

Current estimates suggest that failed rocket launches can cost the industry up to $5000 millions annually as shown in Figure 2. This highlights the importance of improving mission success rates, which can happen using data analysis and machine learning. Our project aims to identify attributes that contribute to successful launches which can potentially lead to significantly more efficient lunches. The success of space missions is important because the rockets that launch satellites play a crucial role in our world. The satellites they carry provide a wide range of essential services, including research, communication, weather forecasting, and GPS navigation. These services have become deeply integrated into our daily lives, which makes the reliability of rocket launches a critical concern. For instance, satellites help to:

- Set up networking capabilities: Communication satellites enable phone calls, internet access, and other forms of data transmission across vast distances [4].
- Forecast weather: Weather satellites provide vital data used in weather forecasting, allowing us to plan our activities and prepare for potential hazards [3].
- GPS: GPS satellites provide the foundation for global positioning systems, enabling navigation in cars, smartphones, and various other applications [7].

Given the importance of these services by satellites, improving the success rate of rocket launches becomes even more important, especially considering the high cost involved. As can be seen in Figure 1, the average rocket launch was very expensive in 2020, costing around 50 million dollars. Therefore, increasing mission success rates can lead to significant cost savings in the long term.

## 3 Related Work

Our project compares the results of various methods to predict the failure of rocket launches. Failure prediction is more common in the field as we can see similar goal have been followed in previous work [6, 2, 1] with different approaches. M.W. Ahmad et al. [1] proposed a new machine learning framework to improve the safety of autonomous aerial vehicles (UAVs) by predicting, detecting, and classifying control surface failures. The framework utilizes a Long Short-Term Memory to analyze sensor data collected during flight. This allows

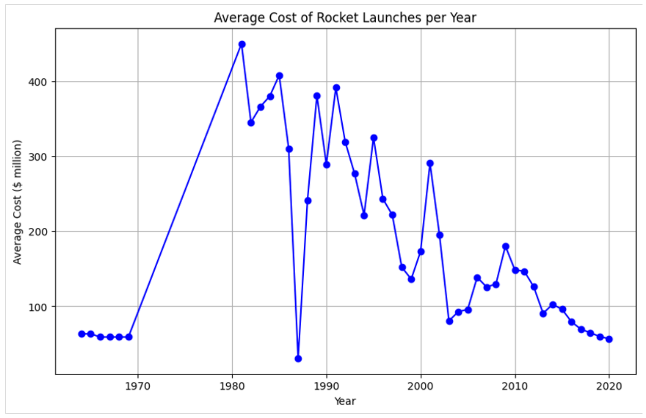


Figure 1: Average Cost of Rocket Lunches per Year

the system to predict failures an average of 19 seconds before they happen. The framework can also detect failures almost instantly (with 0.74 seconds delay) and precisely identify the type of failure. Another related work is proposed by Y. Shao et al. [5], which uses a time series model to analyze spacecraft component temperature data collected via remote sensing. They utilized an ARIMA (auto regressive integrated moving-average) model for temperature prediction. Deviations from the predicted temperature values signal potential component failures. Our work expands this concept by considering a broader range of variables for more robust prediction and comparing the performance of various machine learning algorithms to identify the optimal model for this purpose. Our contribution is implementing efficient yet precise algorithms which do not require high computational power.

# 4 Data Analysis

We used datasets that provide information about space missions, and mission status, to analyze other attributes' effect on mission status. We prepossessed data including cleaning the data to handle missing values, standardizing formats across datasets, and linking relevant attributes to facilitate analysis. The scope of the project aligns with the original plan outlined in the project proposal. After analyzing the three proposed datasets, we decided to use two of them, as the third dataset is more about astronauts' information that is not included in the other two, and this may not be an important factor like the physical factors or company. The datasets that are being used will be discussed more in the following sections.

## 4.1 Space Flights' Mission Status (SM)

This dataset [1] contains the following attributes: Company, Launch Date, Launch Time, Launch Site, Temperature (°F), Wind Speed (MPH), Humidity (%), Vehicle Type, Liftoff Thrust (kN), Payload to Orbit (kg), Rocket Height (m), Fairing Diameter (m), Payload Name, Payload Type, Payload Mass (kg), Payload Orbit, Mission Status, and Failure Reason. This dataset has 120 successful missions and 30 failed missions, which makes it imbalanced. There were some missing values in many columns, especially Failure Reason and some weather information. This dataset does not have the geolocations of the launch sites, which we need for getting missing weather data. To solve this, we retrieved the geolocations as Longitude/Latitude coordinates using the Google Maps Geolocation API. Next, we used the Open Meteo API to retrieve the missing weather data at a given launch time. Four rows did not have a launch time so we could not retrieve weather data for them. We opted to remove these rows instead. Because the Open Meteo API provides many more weather variables, we added these to the dataset. This increased our variables from 18 to 49. One issue we encountered with the Open

[1] https://www.kaggle.com/datasets/rosetabares/spacemissionsflightstatus

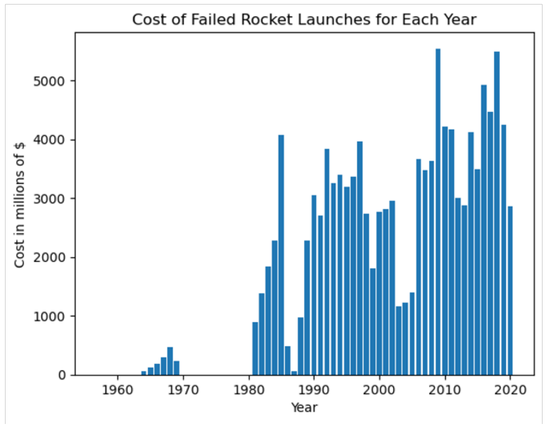


Figure 2: Cost of Failed Rocket Launches for Each Year

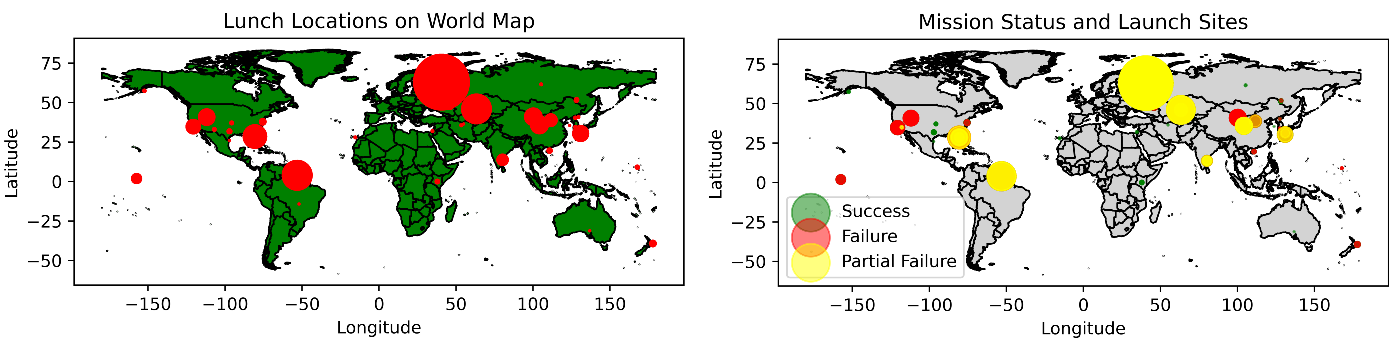


Figure 3: Map Visualization using Geopandas

Meteo data is it did not provide weather information for the hour 23:00. We accounted for this by picking the hour at 22:00 where 23:00 was not available. Finally, we need to compare the flights between this dataset and the other space flight dataset.

We will use data mining techniques like clustering, classification, and data visualization in our study.

## 4.2 All Space Missions from 1957 (SC)

The dataset [2] includes data on all space missions since the beginning of the Space Race (1957). The data includes features like Company Name, Location, Datum, Detail, Status Rocket, Rocket, and Status Mission Where 'Rocket' attribute has a few missing values, but other attributes do not have any null values. So we do not have any concern about null values in this dataset. This dataset has 3879 successful missions, 339 failures, 102 partial failures, and 4 pre-launch failures, which makes the classes imbalanced. For this dataset, we used the provided launch locations and times to add weather data as we did with the first dataset. Therefore, we had 31 more weather variables to use in data exploration.

## 4.3 Map Visualization

We extracted location information for each launch. We want to use this information to extract any possible rules about the relationship between location and mission status. We used GeoPandas to plot longitude and latitude data on a map, which provides a comprehensive spatial visualization of our information. Figure 3 visualizes the distribution of lunches and missions across various locations. The left side shows the number

[2] https://www.kaggle.com/datasets/agirlcoding/all-space-missions-from-1957

of lunches at each site, with marker sizes proportional to frequency. This reveals a surprising diversity of locations beyond the more well-known ones. The right figure use the same logic to show the frequency of mission statuses at each site. It was interesting to see in Figure 3 that the number of launches a site has is irrelevant to the success rate; if they don't analyze previous missions to identify reasons for failure, they're likely to experience continued failures. This demonstrates the importance of our work. Even a site with few launches could achieve greater success through careful analysis of past missions.

# 5 Result

All the results and data of this project are available on the GitHub: `https://github.com/maryamkameli/Space-Mission-Failure-Prediction`.

## 5.1 Database1-based Analysis (SM)

This dataset has 50 attributes which we will explore more in the following sections. We want to find factors of interest through visualization and apply them to model building. We use a k-means classifier and correlation matrix to find the relationship between features. Based on this analysis, we choose a decision tree as the base categorization because our datasets are imbalanced. Furthermore, we explore alternative methods such as random forest classifier and gradient boosting classifier, which are similar to a decision tree, to compare the results.

### 5.1.1 Visualization

We explored different aspects of the given attributes in this dataset. To understand investment patterns, we analyzed the distribution of payload costs across categories. As shown in Figure 4 (right), communication satellites represent the category with the highest investment. This aligns with our earlier observation about the skewed balance within the dataset, further illustrated in Figure 4 (left). Finally, Figure 5 examines the number of failed launches per month.

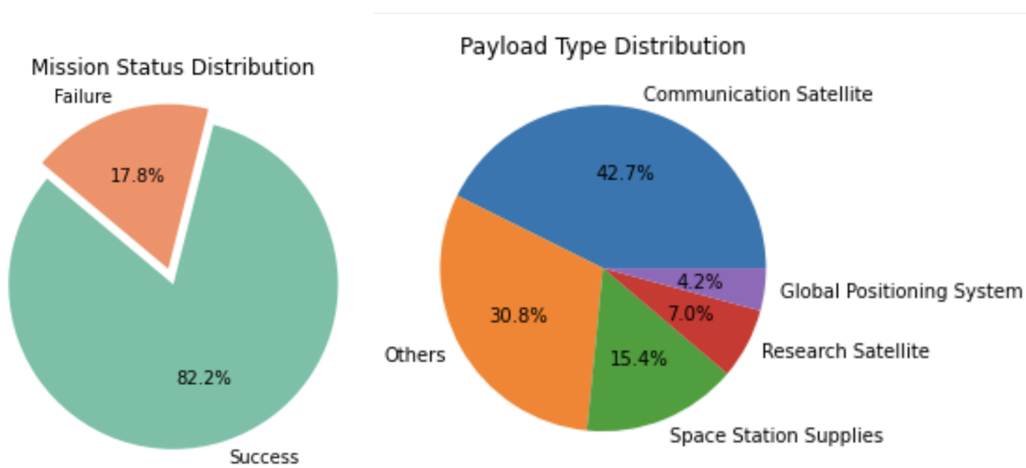


Figure 4: Data Distribution in Dataset1 (SM)

### 5.1.2 Analysis on Different Models

Some data is of non-numeric type. We will drop them when analyzing the dataset using some models. The feature "Fairing Diameter (m)" is accompanied by missing data. We will drop this feature to preserve the

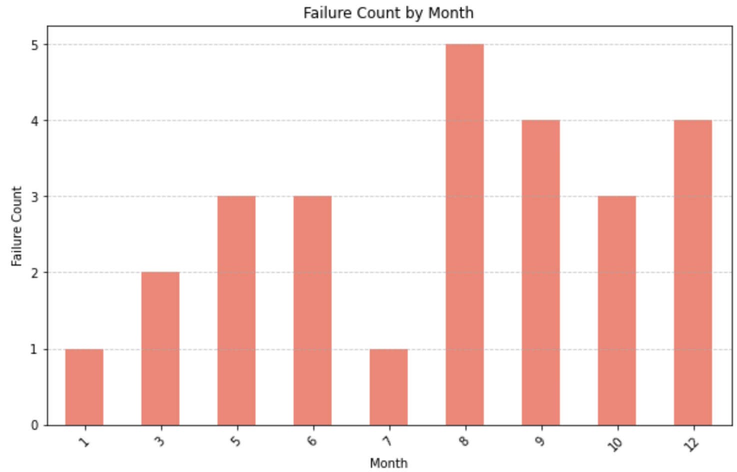


Figure 5: Number of Failed Lunches per Month

mission status data when analyzing the importance of features based on different models.

When using the k-means classifier, we extract 29 features from 50 attributes to analyze the relationship between those features. Based on the correlation matrix shown in Figure 6 and the similarities between features, we divide 29 attributes into 6 categories. They are atmospheric conditions, pressure and cloud cover, wind conditions, radiation and evapotranspiration, soil moisture, and soil temperature.

We chose to check the distribution of the values of the features in the case of launch failures. We find that some of the values of features are clustered together, while the distribution of some does not seem to have any pattern. For the features with concentrated distributions, we can speculate that they may be related to launch failures.

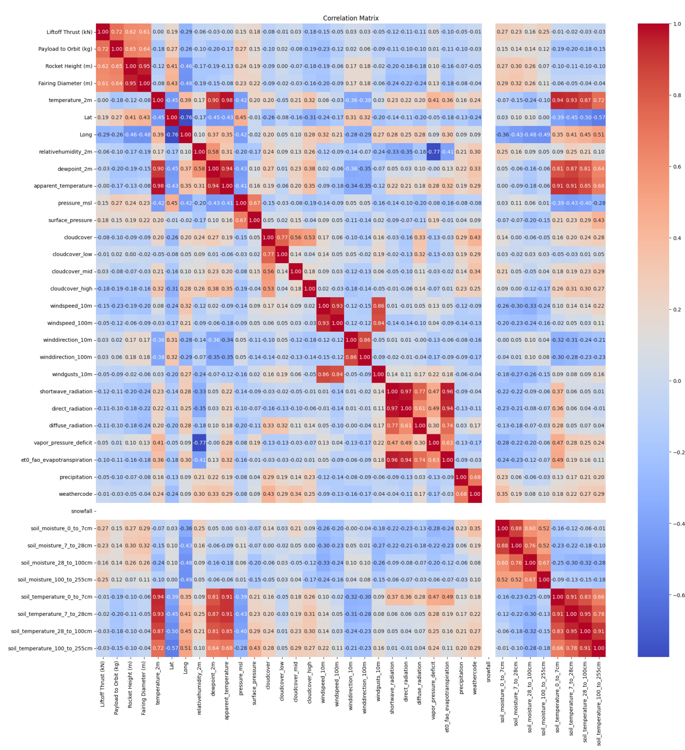


Figure 6: Correlation Matrix of Some Features in Dataset1 (SM)

We use decision tree, random forest classifier, and gradient boosting classifiers to analyze the importance of different features based on the corresponding model. We can see that the importance distribution of the three

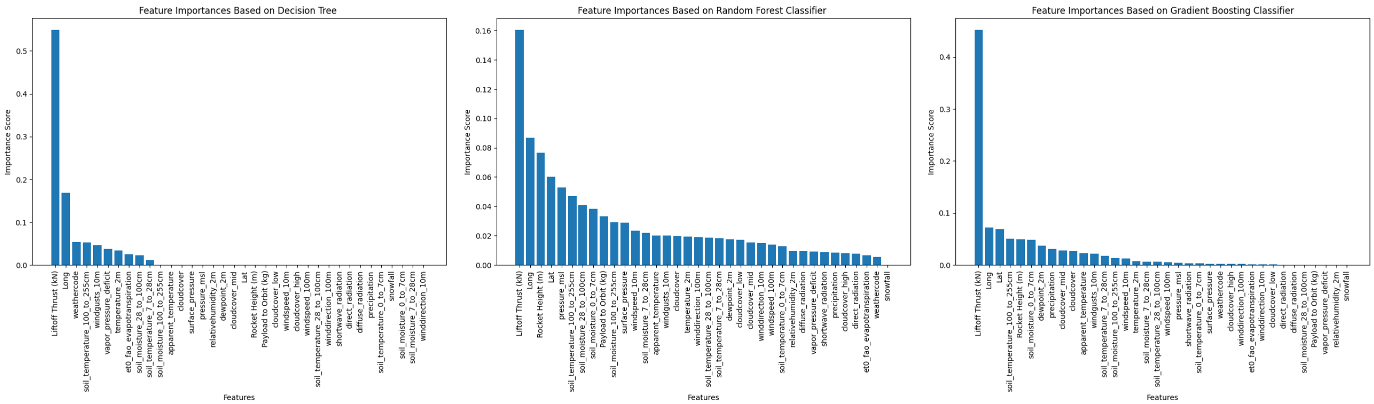


Figure 7: Importance of Features Based on Three Models

models is slightly different in Figure 7. However, no matter for which model, ”liftoff thrust” has the greatest influence on the mission status.

#### 5.1.3 Accuracy of Prediction

From the table 1, three models performed the same. We can see the data in database1(SM) is simple and small. In such cases, the ensemble nature of the random forest classifier and the gradient boosting classifier may not provide significant benefits over a single decision tree. Moreover, the accuracy of the Support Vector Machine (SVM) is lower than other methods on the test set, which suggests that the performance of SVM on imbalanced data is unstable and needs to be optimized by parameter tuning.

| | Decision Tree | Random Forest Classifier | Gradient Boosting Classifier | SVM |
|---|---|---|---|---|
| Training accuracy | 0.993 | 0.993 | 0.993 | 0.891 |
| Testing accuracy | 0.875 | 0.875 | 0.875 | 0.75 |

Table 1: Accuracy of Dataset1 (SM)

### 5.2 Database2-based Analysis (SC)

We have explored different aspects of attributes in this dataset. We also use methods such as SVM and the gradient boosting classifier to predict the mission status.

#### 5.2.1 Visualization

- **Summary**
  In Figure 8, the right and left figures are the proportion of the first and second datasets about the number of successful launches and the number of failures, respectively. The first dataset that we use has fewer data but the proportion of launch failures is close to 20%, which is more helpful for us to study the causes of launch failures.

- **Successes and Failures Over Time**
  We try to look for factors that are affecting the launch success rate from the visualization. As shown in Figure 9(right), time is an important factor because the success rate increases and the failure rate decreases as time changes. In Figure 9(left), we find an interesting factor. The failure rate may also be related to the day of the week since Wednesday has the lowest failure rate.

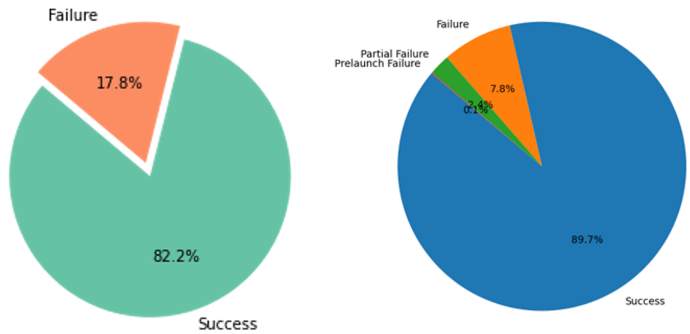


Figure 8: Class Imbalance: Success and Failure Distribution in Datasets

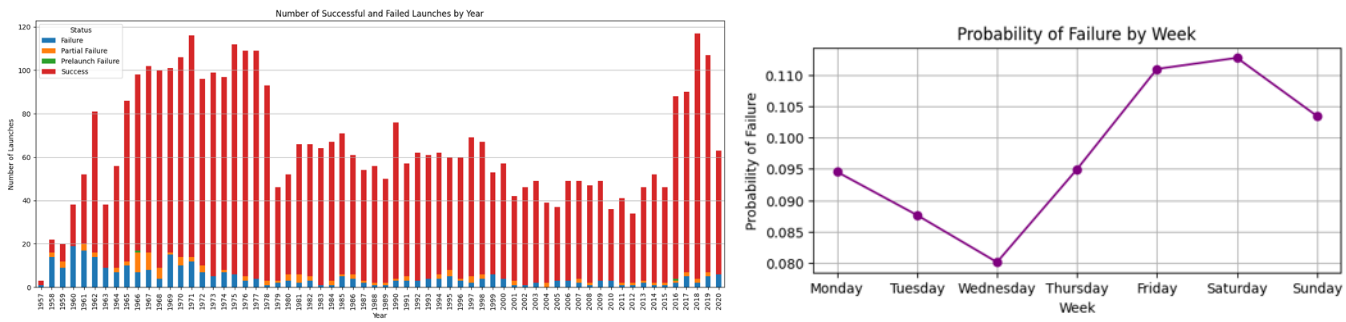


Figure 9: Number of successful and failed launches by years (Right image), Probability of failure by week (Left image)

- **Cost**
  Figure 10 shows the trend in the average spend per mission, and it can be seen that while some failed missions cost similar money to the successful spend, most of the failures seem to be due to under-spending.

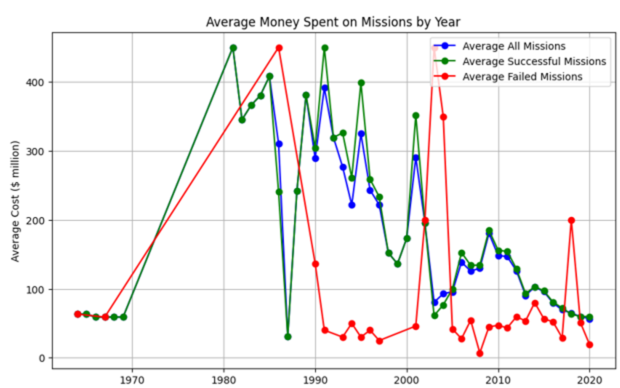


Figure 10: Average money spend on missions by year

### 5.2.2 Analysis on Different Models

We build decision tree, random forest, gradient boosting classifiers, and SVM model to analyze 33 features from 40 features of Dataset2(SC). Those 33 features contain numerical data without missing values and the id feature. From Figure 11, we can see the relationship between those features.

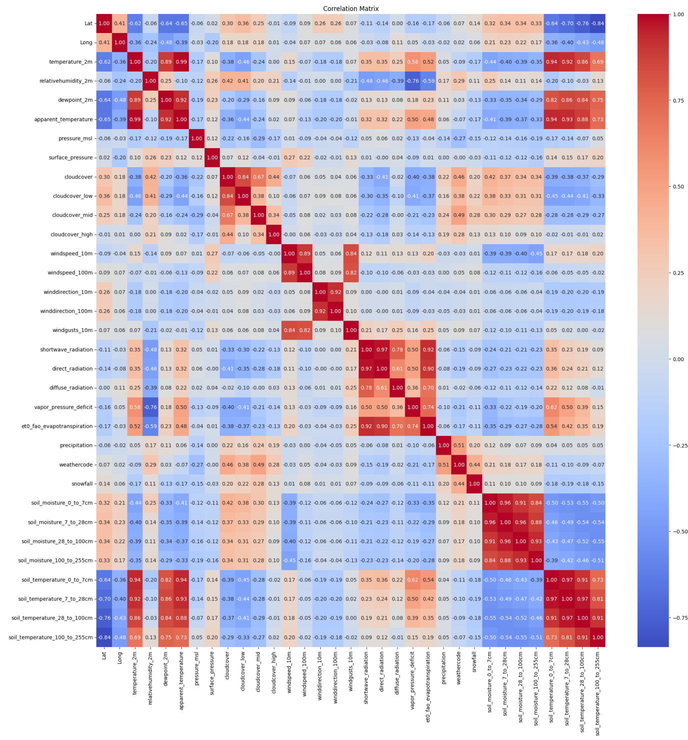


Figure 11: Correlation Matrix of Some Features in Dataset2 (SC)

We use decision tree, random forest classifier, and gradient boosting classifier to analyze the importance of different features based on the corresponding model. We can see that the importance distribution of the three models is slightly different in Figure 12.

For the decision tree classifier, “apparent temperature”, ”surface pressure”, and ”soil temperature from 28 cm to 100 cm” have more influence on the model. For the random forest classifier, ”surface pressure”, ”wind speed at 100m”, and “apparent temperature” have more influence on the model. For the gradient boosting classifier, ”surface pressure”, ”pressure”, soil moisture from 28 cm to 100 cm”, and ”soil temperature from 100 cm to 255cm” have a stronger impact on the mission status.

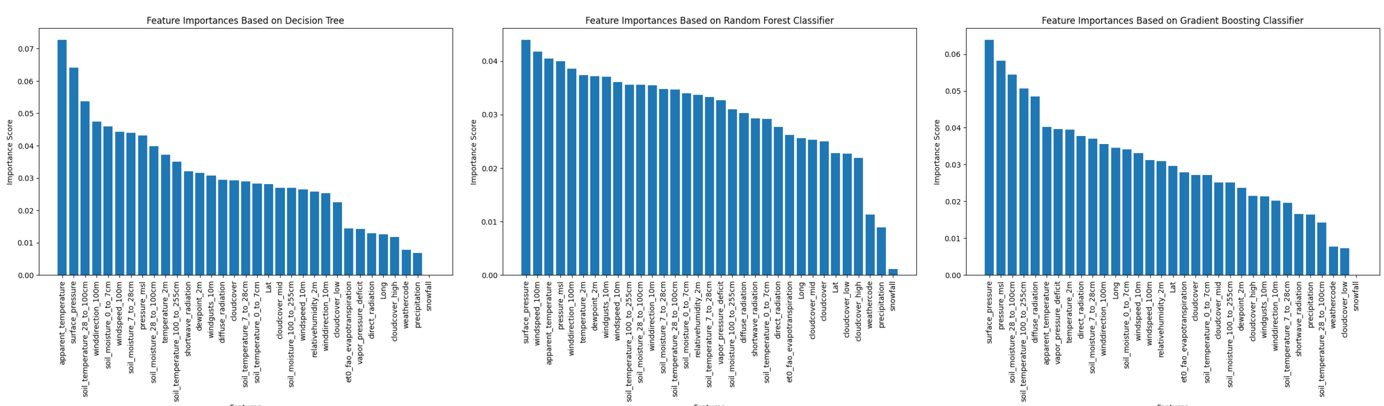


Figure 12: Importance of Features Based on Three Models

### 5.2.3 Accuracy of Prediction

From the table 2, we can see the four models perform well. Compared to individual decision trees, ensemble methods like Random Forest and Gradient Boosting tend to perform better.

| | Decision Tree | Random Forest Classifier | Gradient Boosting Classifier | SVM |
|---|---|---|---|---|
| Training accuracy | 1 | 1 | 0.963 | 0.935 |
| Testing accuracy | 0.833 | 0.926 | 0.901 | 0.926 |

Table 2: Accuracy of Dataset2 (SC)

# 6 Conclusion

Our goal is to find features that predict rocket failures. Our data analysis and various model exploration have exposed many such variables. We learned that failures are geographically correlated, though that on its own is not useful since rocket funding and quality probably vary regionally. Our analysis included the extraction of launch location data. Using GeoPandas, we visualized this spatial information, which revealed a surprising diversity of launch sites in Figure 3. Interestingly, we found no direct correlation between the frequency of launches at a site and its mission success rate. This highlights the importance of our project – even sites with infrequent launches could significantly improve their success by thoroughly analyzing past mission data and identifying the root causes of failures. We also learned that increasing rocket cost tends to increase the rocket's success rate. Most importantly, from our models, we learned that the liftoff thrust of the rocket is likely the most important observed feature in the success rate of rocket launches. The surface pressure is also an important variable; it is likely that with increased surface pressure more thrust is needed. Furthermore, the temperature variables are quite relevant, but that is likely due to the relationship between temperature and pressure. It is hard to say for sure if there is a direct relationship between temperature and success rate as a result. We found some interesting daily patterns in failures, with more failures happening on and before weekends. The lowest failure rates were on Wednesdays followed by Tuesdays as Figure 9 shows. Why this is the case is unclear, but it is likely because the people working on the launches get more tired and restless as the week progresses. Performing most launches in the middle of the week can reduce failure chances by a significant amount, hence ultimately reducing cost.

# References


[1] Muhammad Waqas Ahmad, Muhammad Usman Akram, Rashid Ahmad, Khurram Hameed, and Ali Hassan. Intelligent framework for automated failure prediction, detection, and classification of mission critical autonomous flights. *ISA transactions*, 129:355–371, 2022.

[2] Siddhartha R Dalal, Edward B Fowlkes, and Bruce Hoadley. Risk analysis of the space shuttle: Pre-challenger prediction of failure. *Journal of the American Statistical Association*, 84(408):945–957, 1989.

[3] Robert R Hoffman, Daphne S LaDue, H Michael Mogil, Paul J Roebber, and J Gregory Trafton. *Minding the weather: How expert forecasters think*. MIT Press, 2023.

[4] Marko Höyhtyä, Antti Anttonen, Mikko Majanen, Anastasia Yastrebova-Castillo, Mihaly Varga, Luca Lodigiani, Marius Corici, and Hemant Zope. Multi-layered satellite communications systems for ultra-high availability and resilience. *Electronics*, 13(7):1269, 2024.

[5] Yinghua Shao and Yuanhang Zhang. A failure prediction method for spacecraft loads based on time series model. In *2018 12th International Conference on Reliability, Maintainability, and Safety (ICRMS)*, pages 277–281, 2018.

[6] Scott J Uder, Robert B Stone, and Irem Y Tumer. Function based risk assessment and failure prediction for unmanned space missions. In *ASME International Mechanical Engineering Congress and Exposition*, volume 47055, pages 271–287, 2004.

[7] Richa Verma, Bipin Kumar Singh, and Farah Zahidi. Management of gps tracking systems in transportation. In *Intelligent Transportation System and Advanced Technology*, pages 251–263. Springer, 2024.