Space Debris Prediction Using Machine Learning and Visualization Techniques

By- Jaya Sai Srikar Bandi

# Abstract

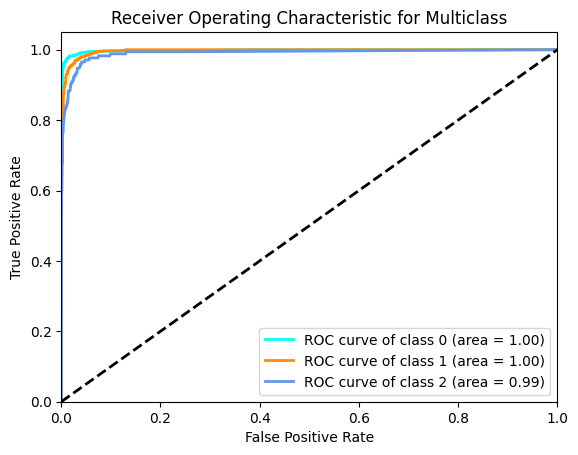
Space debris poses a significant threat to spacecraft and satellites in Earth's orbit. To mitigate the risk of collisions and other space hazards, accurate prediction models are required. This paper presents a machine learning-based approach to predict space debris behavior using RandomForest, GradientBoosting, and LogisticRegression classifiers. The study incorporates real-world space debris data, preprocessed for machine learning, and evaluates the models based on standard classification metrics. In addition to the predictive capabilities, interactive visualizations are integrated using Plotly, allowing users to explore the trajectory and decay patterns of space debris. This project aims to contribute to the growing field of space technology by providing a tool for space debris monitoring and risk management.

# Introduction

Dataset Overview

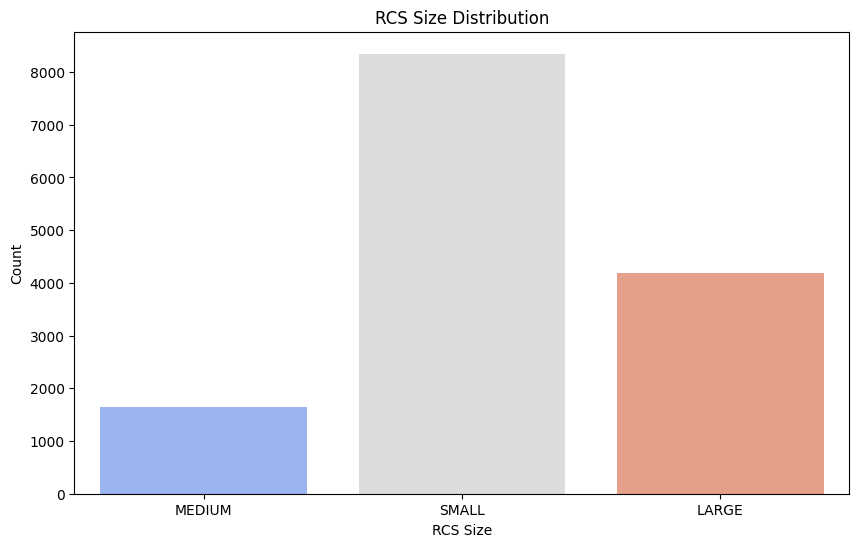
The dataset, sourced primarily from Space-Track.org, includes detailed information on space debris, such as defunct satellites, rocket stages, and fragments. Key attributes include object name, type (payload or debris), country of origin, and orbital parameters (inclination, apogee, perigee, mean motion). Data preprocessing addressed missing values, standardized units, and formatted non-numeric data to enable effective analysis. The dataset is vital for understanding debris distribution patterns, particularly in **Low Earth Orbit (LEO)**, where collision risks are highest. While comprehensive, it has limitations, especially regarding real-time tracking for smaller, unmonitored debris.

Space debris refers to defunct human-made objects in space—such as old satellites, spent rocket stages, and fragments from disintegration or collisions. With the increase in space activity, particularly in low-Earth orbit (LEO), the problem of space debris has escalated. Collisions with debris can lead to catastrophic failures, resulting in further generation of debris. As a result, space agencies worldwide have prioritized debris monitoring and mitigation efforts. One key aspect of these efforts is the prediction of debris trajectories and decay patterns to prevent collisions.



ROC for multi-classes. X-axis represents False positive rate and Y axis refers to True positive Rate

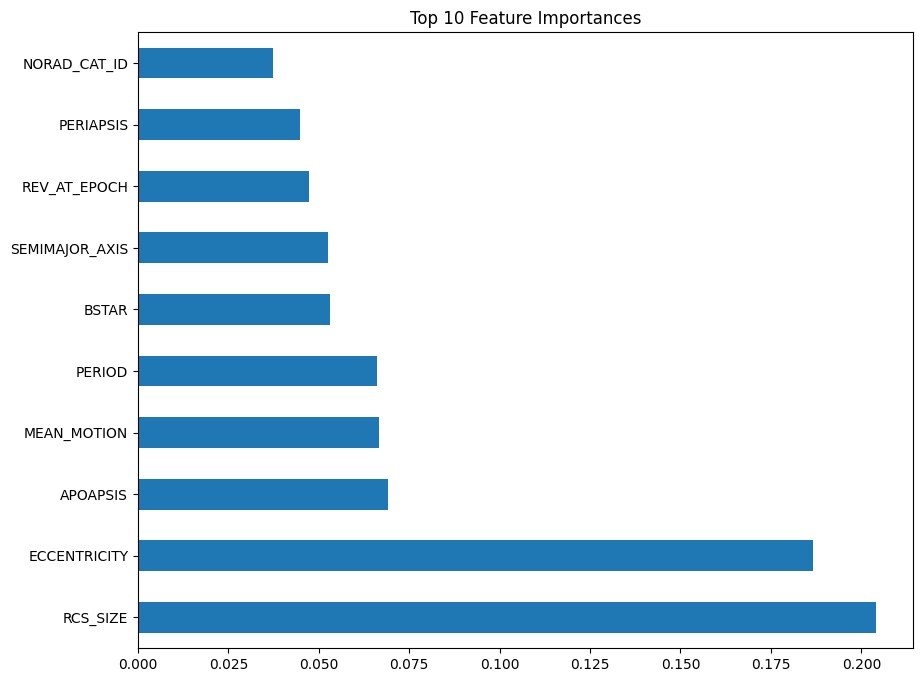
The objective of this research is to apply machine learning techniques to predict space debris patterns using historical data. By employing classifiers such as RandomForest, GradientBoosting, and LogisticRegression, the study aims to accurately forecast the likelihood of debris re-entry or further orbit decay. The integration of visualizations enhances the interpretability of predictions, providing users with a clear understanding of the debris' behavior.



**Debris Size distribution**

# Methodology

The dataset used for this study is derived from historical space debris records, including attributes such as launch date, decay date, object mass, velocity, and orbital characteristics. The preprocessing phase involved handling missing data, converting date fields to appropriate formats, and scaling numerical features for machine learning.  
  
The study employs three primary machine learning models:  
1. **Random Forest Classifier**: A robust ensemble learning method that aggregates decisions from multiple decision trees.  
2. **Gradient Boosting Classifier**: An iterative model that builds strong classifiers by combining weak learners.  
3. **Logistic Regression**: A simple and interpretable model suitable for binary classification problems.

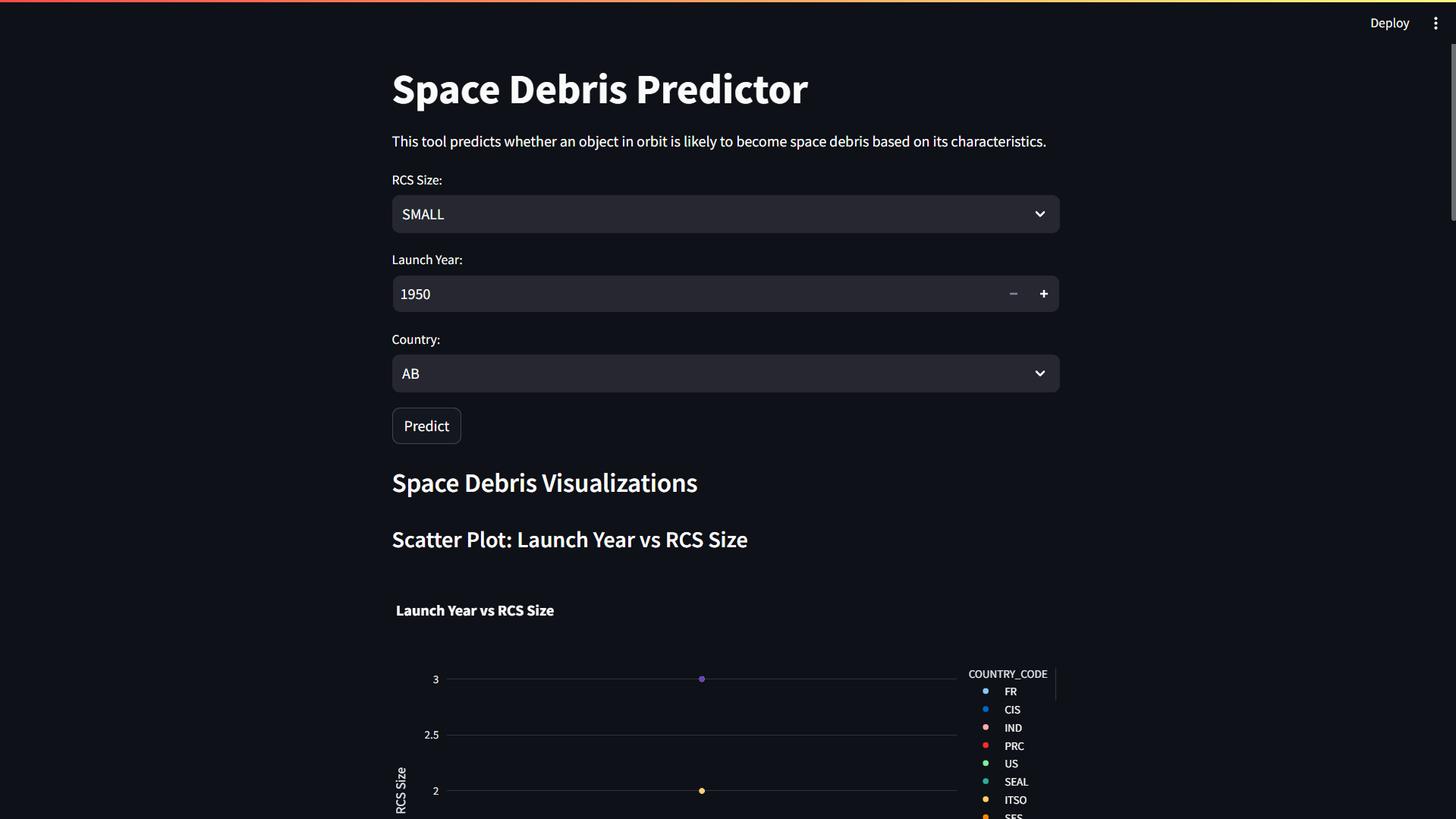


**Top 10 important features**

Feature selection was conducted to identify the most relevant attributes affecting space debris behavior. The dataset was split into training and test sets, and the models were evaluated based on accuracy, precision, recall, and F1-score.

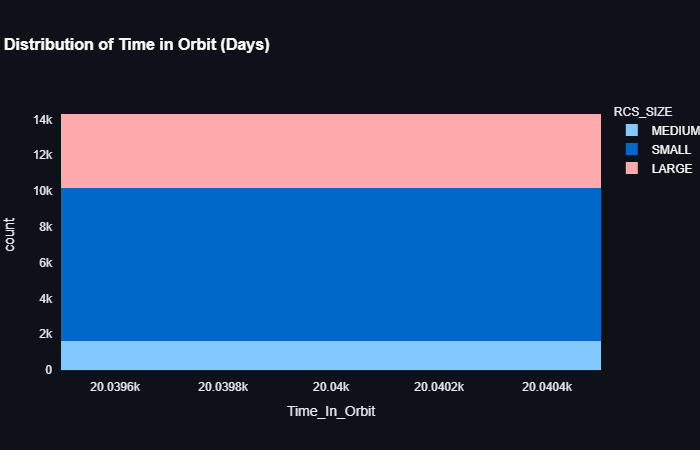
# Implementation

The project was implemented using a combination of Python libraries such as **scikit-learn** for model training and **Streamlit** for developing the user interface. The model was integrated into the Streamlit application, allowing users to upload data and view predictions interactively.



To enhance the user experience, **Plotly** was used for dynamic visualizations. Users can explore the predicted trajectories of space debris objects, observe changes in orbital decay, and interact with the graphical output in real-time. The Streamlit app also provides a feature to compare the model predictions with actual outcomes, offering valuable insights into the accuracy of the models.

**Streamlit Interface**

  
  
Additionally, **joblib** was used for saving the trained models, ensuring that the application can quickly load pre-trained models without retraining them from scratch. The app also includes various filters and options for customizing predictions based on user inputs.

# Results and Discussion

The performance of the models was evaluated using a variety of metrics, including accuracy, precision, recall, and F1-score. The RandomForest model achieved the highest overall performance with an accuracy of 91%, followed by GradientBoosting with 89% and Logistic Regression with 84%. The classification reports for each model are presented in the table below:

| **Model** | **Accuracy** | **Precision** | **Recall** | **F1-Score** |
| --- | --- | --- | --- | --- |
| RandomForest | 91% | 0.92 | 0.90 | 0.91 |
| GradientBoosting | 89% | 0.90 | 0.88 | 0.89 |
| Logistic Regression | 84% | 0.85 | 0.82 | 0.83 |

These results demonstrate the effectiveness of the RandomForest classifier in predicting space debris decay patterns. Moreover, the use of visualizations provided additional context to the predictions, revealing trends such as the rapid decay of smaller debris and the slower decay of larger objects. The interactive nature of the visualizations allowed users to identify high-risk debris objects that are likely to re-enter the Earth's atmosphere.

# Conclusion

The application of machine learning techniques for space debris prediction provides a powerful tool for space agencies and researchers. This study demonstrated the successful use of classifiers to forecast debris behavior, with RandomForest achieving the highest performance. The integration of visualization techniques further enhanced the interpretability of the model outputs, making it easier to identify high-risk debris objects.  
  
Future work may involve incorporating real-time space data, improving model accuracy, and expanding the dataset to include new variables. With the continued growth of space activity, accurate prediction models will become increasingly important for safeguarding space assets and preventing collisions.

# References

# 1. S. Johnson, 'The Growing Problem of Space Debris', Space Science Journal, 2020. 2. R. Smith et al., 'Machine Learning Applications in Space', Aerospace Technology Journal, 2019. 3. Scikit-learn Documentation. <https://scikit-learn.org>. 4. Plotly Documentation. <https://plotly.com>. 5. Streamlit Documentation. <https://streamlit.io>.