IBM Developer
SKILLS NETWORK

Winning Space Race with Data Science

Manas Gupta
28thJuly,2023

Outline

- Executive Summary
- Introduction
- Methodology
- Results
- Conclusion
- Appendix

Executive Summary

The Winning Space Race with Data Science project aims to leverage advanced data science techniques to enhance space exploration and propel our nation to the forefront of space missions. Through a comprehensive methodology that integrates machine learning, big data analytics, and artificial intelligence, we seek to optimize spacecraft design, mission planning, and resource allocation.

Introduction

The introduction provides an overview of the current state of the space race, highlighting the competitive landscape and the importance of technology and innovation. The role of data science is introduced as a crucial element that can revolutionize space exploration, making missions more efficient, cost-effective, and safe. The objective of this report is to explore how data science can be utilized to gain a competitive edge in the space race and ultimately achieve groundbreaking discoveries.

Section 1

Methodology

Methodology

1. This section delves into the methodologies employed to harness the power of data science in the space race. It begins with a discussion on data collection techniques, including satellite sensors, space telescopes, and other data-gathering instruments. Moreover, the report highlights the challenges in handling massive volumes of space-related data and how big data analytics can help process, manage, and extract valuable insights.
2. The subsequent part focuses on the application of machine learning algorithms to space data. Machine learning can analyze historical space mission data to predict potential risks and optimize mission planning. It can also aid in autonomous decision-making during space missions, reducing the reliance on ground control and enabling quicker responses to dynamic situations.
3. Additionally, the report delves into the significance of interdisciplinary collaboration between space scientists, engineers, and data scientists. It emphasizes the need for knowledge-sharing and integration of expertise to make the most of data science applications in space exploration.

Data Collection

- Data collection is a fundamental process in the field of data science and plays a crucial role in various applications, including space exploration. It involves gathering relevant and accurate information from different sources to analyze and draw insights that can inform decision-making and drive advancements in various domains. In the context of winning the space race with data science, data collection encompasses obtaining data from a wide range of sources related to space missions, space objects, astronomical observations, astronaut health, and other relevant aspects.
- Satellite Sensors and Space Telescopes: Satellite sensors and space telescopes are primary sources of data in space exploration. Satellites equipped with various sensors capture valuable information, such as imagery, atmospheric data, and gravitational measurements. Space telescopes, like the Hubble Space Telescope, observe celestial objects, galaxies, and exoplanets, providing essential data for astronomical research.
- Robotic Probes and Rovers: Robotic probes and rovers deployed on celestial bodies like the Moon, Mars, and other planets gather on-site data about the surface, geology, and environmental conditions. These instruments transmit data back to Earth, enabling scientists to study these extraterrestrial environments.
- Astronaut Data: Astronauts on space missions generate extensive data related to their physiological health, cognitive performance, and interactions with the spacecraft. Monitoring these parameters helps ensure astronaut well-being and safety during extended space missions.
- Spacecraft Telemetry: Spacecraft send telemetry data to Earth, providing vital information about their status, performance, and health. This telemetry includes data on propulsion systems, power usage, and onboard equipment, helping ground control monitor and manage the spacecraft effectively.
- Space Debris Tracking: Collecting data on space debris is essential to prevent collisions and protect spacecraft in orbit. Ground-based and space-based sensors track debris objects, enabling the calculation of their orbits and potential collision risks.
- Ground Observatories: Ground-based observatories around the world gather astronomical data, such as radio waves, X-rays, and gamma rays, from distant celestial objects and events. This data complements space-based observations, enhancing our understanding of the universe.
- Space Weather Monitoring: Space weather, including solar flares and cosmic radiation, can affect space missions and satellite operations. Monitoring space weather data helps mitigate potential risks and safeguard space assets.
- Historical Mission Data: Data from past space missions provides valuable insights for future endeavors. Analyzing historical mission data helps identify patterns, risks, and opportunities, aiding in optimized mission planning and decision-making.

Data Collection – SpaceX API

- Present your data collection with SpaceX REST calls using key phrases and flowcharts

- Add the GitHub URL of the completed SpaceX API calls notebook (must include completed code cell and outcome cell), as an external reference and peer-review purpose

Place your flowchart of SpaceX API calls here

Data Collection - Scraping

- Present your web scraping process using key phrases and flowcharts
- Add the GitHub URL of the completed web scraping notebook, as an external reference and peer-review purpose

Place your flowchart of web scraping here

Data Wrangling

- Describe how data were processed
- You need to present your data wrangling process using key phrases and flowcharts
- Add the GitHub URL of your completed data wrangling related notebooks, as an external reference and peer-review purpose

EDA with Data Visualization

- Summarize what charts were plotted and why you used those charts
- Add the GitHub URL of your completed EDA with data visualization notebook, as an external reference and peer-review purpose

EDA with SQL

- Using bullet point format, summarize the SQL queries you performed
- Add the GitHub URL of your completed EDA with SQL notebook, as an external reference and peer-review purpose

Build an Interactive Map with Folium

- Summarize what map objects such as markers, circles, lines, etc. you created and added to a folium map
- Explain why you added those objects
- Add the GitHub URL of your completed interactive map with Folium map, as an external reference and peer-review purpose

Build a Dashboard with Plotly Dash

- Summarize what plots/graphs and interactions you have added to a dashboard
- Explain why you added those plots and interactions
- Add the GitHub URL of your completed Plotly Dash lab, as an external reference and peer-review purpose

Predictive Analysis (Classification)

- Summarize how you built, evaluated, improved, and found the best performing classification model
- You need present your model development process using key phrases and flowchart
- Add the GitHub URL of your completed predictive analysis lab, as an external reference and peer-review purpose

Results

- Exploratory data analysis results
- Interactive analytics demo in screenshots
- Predictive analysis results

Section 2

Insights drawn from EDA

Flight Number vs. Launch Site

- Show a scatter plot of Flight Number vs. Launch Site
- Show the screenshot of the scatter plot with explanations

Payload vs. Launch Site

- Show a scatter plot
  of Payload vs. Launch Site

- Show the screenshot of the
  scatter plot with explanations

Success Rate vs. Orbit Type

- Show a bar chart for the success rate of each orbit type

- Show the screenshot of the scatter plot with explanations

Flight Number vs. Orbit Type

- Show a scatter point of Flight number vs. Orbit type

- Show the screenshot of the scatter plot with explanations

Payload vs. Orbit Type

- Show a scatter point of payload vs. orbit type

- Show the screenshot of the scatter plot with explanations

Launch Success Yearly Trend

- Show a line chart of yearly average success rate

- Show the screenshot of the scatter plot with explanations

All Launch Site Names

- Find the names of the unique launch sites
- Present your query result with a short explanation here

Launch Site Names Begin with 'CCA'

- Find 5 records where launch sites begin with `CCA`
- Present your query result with a short explanation here

Total Payload Mass

- Calculate the total payload carried by boosters from NASA
- Present your query result with a short explanation here

Average Payload Mass by F9 v1.1

- Calculate the average payload mass carried by booster version F9 v1.1
- Present your query result with a short explanation here

First Successful Ground Landing Date

- Find the dates of the first successful landing outcome on ground pad
- Present your query result with a short explanation here

Successful Drone Ship Landing with Payload between 4000 and 6000

- List the names of boosters which have successfully landed on drone ship and had payload mass greater than 4000 but less than 6000

- Present your query result with a short explanation here

Total Number of Successful and Failure Mission Outcomes

- Calculate the total number of successful and failure mission outcomes
- Present your query result with a short explanation here

Boosters Carried Maximum Payload

- List the names of the booster which have carried the maximum payload mass
- Present your query result with a short explanation here

2015 Launch Records

- List the failed landing_outcomes in drone ship, their booster versions, and launch site names for in year 2015

- Present your query result with a short explanation here

Rank Landing Outcomes Between 2010-06-04 and 2017-03-20

- Rank the count of landing outcomes (such as Failure (drone ship) or Success (ground pad)) between the date 2010-06-04 and 2017-03-20, in descending order

- Present your query result with a short explanation here

Section 3

Launch Sites Proximities Analysis

<Folium Map Screenshot 1>

- Replace <Folium map screenshot 1> title with an appropriate title
- Explore the generated folium map and make a proper screenshot to include all launch sites’ location markers on a global map
- Explain the important elements and findings on the screenshot

<Folium Map Screenshot 2>

- Replace <Folium map screenshot 2> title with an appropriate title
- Explore the folium map and make a proper screenshot to show the color-labeled launch outcomes on the map
- Explain the important elements and findings on the screenshot

<Folium Map Screenshot 3>

- Replace <Folium map screenshot 3> title with an appropriate title

- Explore the generated folium map and show the screenshot of a selected launch site to its proximities such as railway, highway, coastline, with distance calculated and displayed

- Explain the important elements and findings on the screenshot

Section 4

Build a Dashboard with Plotly Dash

<Dashboard Screenshot 1>

- Replace <Dashboard screenshot 1> title with an appropriate title
- Show the screenshot of launch success count for all sites, in a piechart
- Explain the important elements and findings on the screenshot

<Dashboard Screenshot 2>

- Replace <Dashboard screenshot 2> title with an appropriate title
- Show the screenshot of the piechart for the launch site with highest launch success ratio
- Explain the important elements and findings on the screenshot

<Dashboard Screenshot 3>

- Replace <Dashboard screenshot 3> title with an appropriate title

- Show screenshots of Payload vs. Launch Outcome scatter plot for all sites, with different payload selected in the range slider

- Explain the important elements and findings on the screenshot, such as which payload range or booster version have the largest success rate, etc.

Section 5

Predictive Analysis (Classification)

Classification Accuracy

- Visualize the built model accuracy for all built classification models, in a bar chart
- Find which model has the highest classification accuracy

Confusion Matrix

- Show the confusion matrix of the best performing model with an explanation

Conclusions

- Point 1
- Point 2
- Point 3
- Point 4
- ...

Appendix

- Include any relevant assets like Python code snippets, SQL queries, charts, Notebook outputs, or data sets that you may have created during this project

Thank you!