



Asteroid Size Predictor Web Application

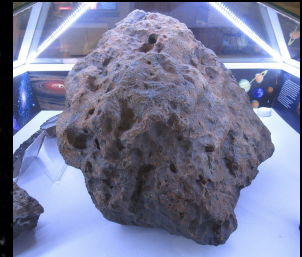
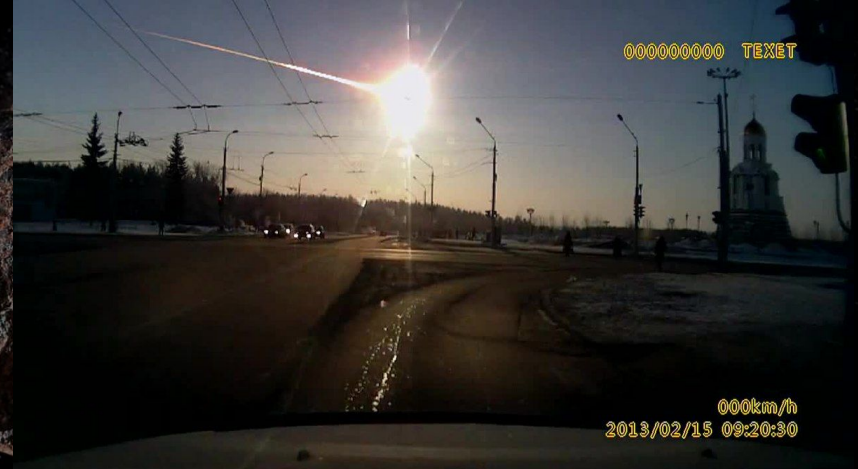
Nate DiRenzo

A composite image of Earth and a large asteroid in space. The Earth is on the left, showing blue oceans and white clouds. The asteroid is on the right, a large, dark, irregularly shaped rock with many small holes and a rough, textured surface. The background is a dark, starry space with a faint, glowing nebula or galaxy structure in the upper left.

Introduction

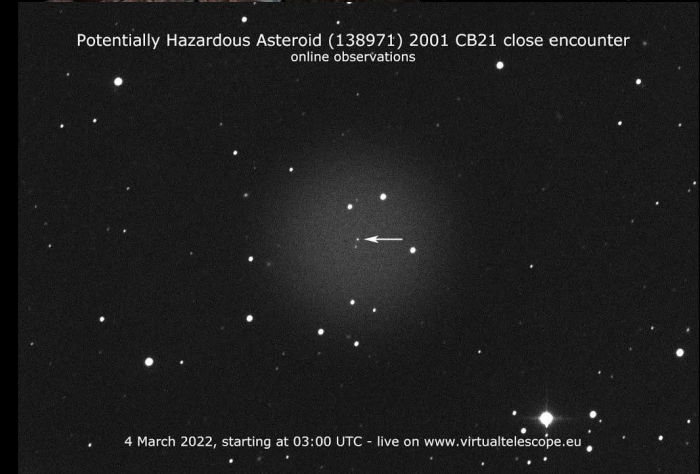
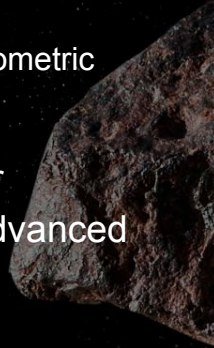
Why Asteroids?

- There are over 1,000,000 known asteroids in our solar system
- A relatively small (140-300m) impact would have continent-wide effects
- Lots of public data available



Current State

- Asteroid sizes are typically estimated from thermal radiometry
 - Function of Absolute Magnitude, Geometric Albedo
- Precise measurements require stellar occultation and direct imaging with advanced instruments on Earth (not many)



Project Goal

Build a machine learning model to predict asteroid diameter.

Determine efficacy of model.

Make model publicly available for others to use.



A large, dark, cratered asteroid or meteorite is positioned on the right side of the frame, appearing to float in space. To its left is a small, blue and white Earth, showing swirling cloud patterns. The background is a deep black space filled with numerous small, distant stars and a faint, wispy nebula or galaxy structure. The text "Solution Path" is written in a clean, white, sans-serif font, centered horizontally between the Earth and the asteroid.

Solution Path

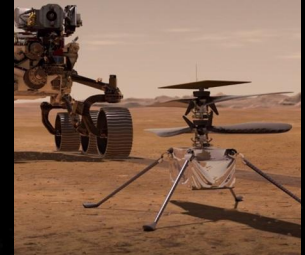
The Data

- [Small-Body Database](#)
 - Courtesy of NASA Jet Propulsion Laboratory, CalTech
- Contains 1.2 million + distinct asteroids, with 30+ features

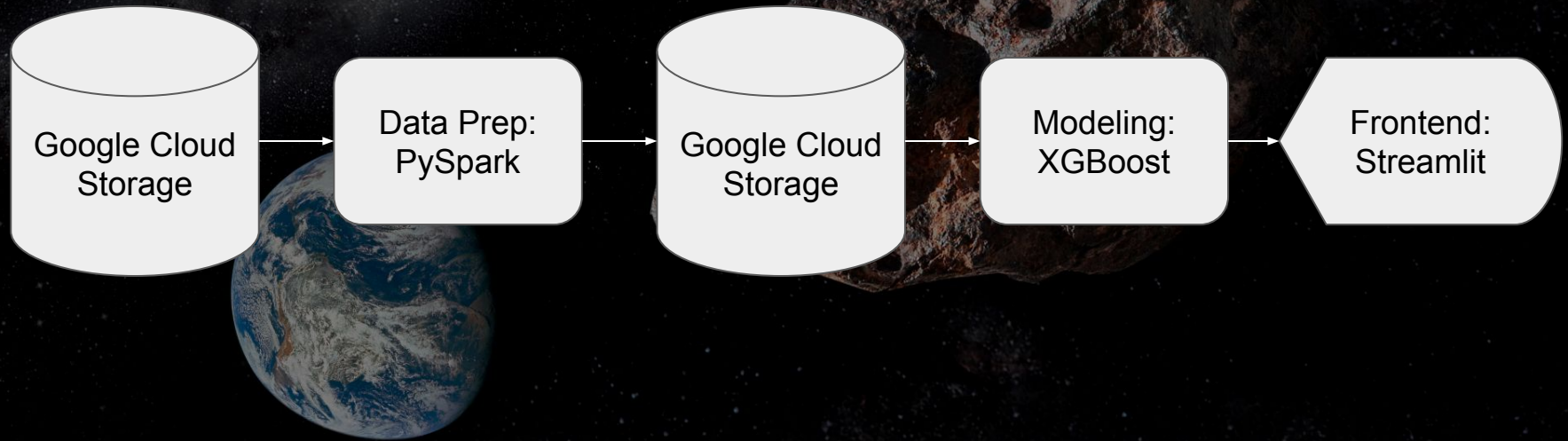
JPL

Jet Propulsion Laboratory
California Institute of Technology

Caltech



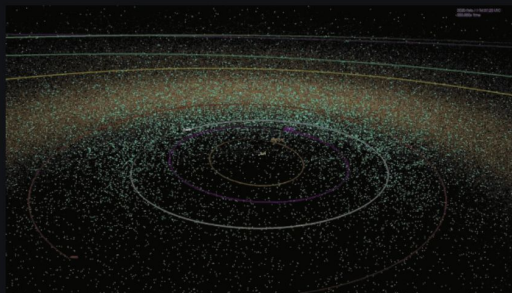
Data Pipeline



Streamlit Application

Asteroid Diameter Predictor

There are over 1,000,000 asteroids in our solar system.



Name	Diameter (km)
Ceres	939.4000
Pallas	545.0000
Juno	246.5960
Vesta	525.4000
Astraea	106.6990
Hebe	185.1800
Iris	199.8300
Flora	147.4910
Metis	190.0000
Hygiea	407.1200

A composite image of Earth and a large, cratered asteroid in space. The Earth is on the left, showing blue oceans and white clouds. The asteroid is on the right, dark and heavily pitted with numerous impact craters. The background is a deep black space filled with a faint, glowing spiral galaxy and scattered stars.

Conclusions



MAE: .979

R2: .79

Model Results

Areas for Future Work

- More nuanced data preparation & feature engineering
- Deploy PySpark model
- Improve Streamlit Application UI & Interactivity
- Provide functionality for predicting on new inputs



A composite image set against a black background filled with stars and a faint, glowing nebula. On the left, a small, realistic Earth is shown with blue oceans and white clouds. To its right is a large, dark, irregularly shaped asteroid with a heavily cratered and textured surface. The text "Thank you!" is centered in white, sans-serif font between the two celestial bodies.

Thank you!

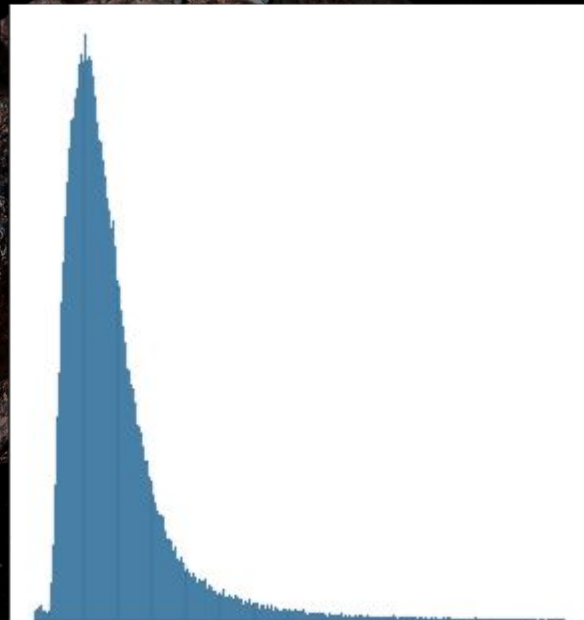
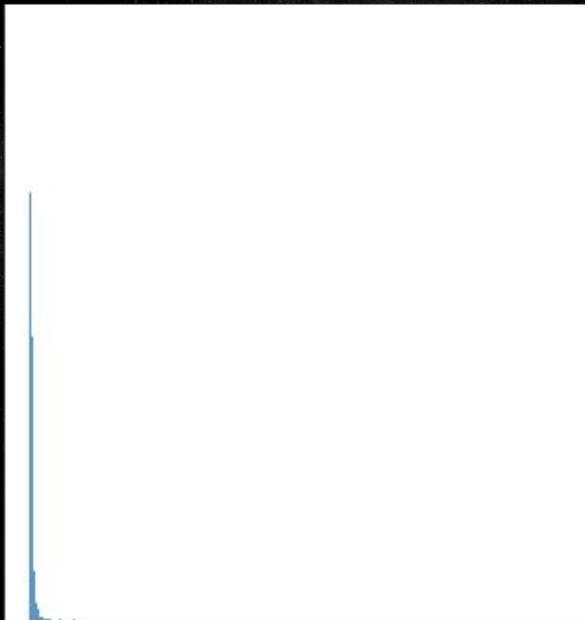
A composite image set against a black background filled with stars. On the left is a small, realistic Earth showing blue oceans and white clouds. On the right is a large, irregularly shaped, brownish-grey rock with numerous dark, circular craters, resembling an asteroid. The word "Appendix" is written in white, sans-serif font across the center of the image, overlapping the asteroid.

Appendix

Correlation Heatmap



Asteroid Diameter Distributions - Pre/Post Quantile Trim



XGBoost Model Feature Importance

