

Binary Classification of Asteroids

Background

There are massive numbers of asteroids and comets orbiting the Sun. A tiny fraction of them can follow a path which passes by Earth. These Near-Earth Objects (NEOs) which have high traveling speeds can be the size of a pebble or a mountain! NEOs have collided with Earth since its formation, and some made the life as we know it possible. They have also resulted in some extreme changes on Earth surface which caused an occasional extinction of some living organisms such as dinosaurs! ^[1]

As of November 2021, there are 2,223 known Potentially Hazardous Asteroids (PHAs) (about 8% of the total near-Earth population), of which 160 are estimated to be larger than one kilometer in diameter! ^[2] NASA has three main tools in its planetary-defense arsenal. The first is to detonate an explosive device near an oncoming asteroid to break it up into smaller, less dangerous chunks. The second is to fire lasers that could heat up and vaporize the space rock enough to change its orbital path. The third is to send a spacecraft to slam into the asteroid, knocking it off its trajectory.^[3]

Problem Statement

If there's a need for intervention, an enormous budget and planning is needed to stop an asteroid from hitting Earth. The resources needed depend on the characteristics of the asteroid. The objective of this project is to explore the available data and build a classification model to be able to identify if an asteroid is hazardous or not! NASA and government agencies can use this information as part of their planetary-defense strategies.

Data

The data about asteroids for this project has been collected from the NASA open API and available on [Kaggle](#). NEOWs ([Near-Earth Object Web Service](#)) is a RESTful web service for near earth Asteroid information. With NEOWs a user can search for Asteroids based on their closest approach date to Earth, lookup a specific Asteroid with its NASA JPL small body id, as well as browse the overall dataset. The dataset has 4,687 observations and 40 columns. The target is whether an asteroid is hazardous or not (True or False). The feature columns include asteroid's speed, some dimensions, and other orbital information.

Tools

Different Python libraries will be used for performing Exploratory Data Analysis, visualization, building and selecting models. In addition, Tableau will be used for additional insights provided through interactive visualizations.

MPV Goal

The MPV will include the preliminary EDA and the base classification model. The suitable metrics will be selected based on these results for further evaluation of future models.

References

[1] [Reference 1](#)

[2] [Reference 2](#)

[3] [Reference 3](#)