ASTEROID DIAMETER PREDICTION

USING DEEP LEARNING AND MACHINE LEARNING

SHARON R
B.TECH(ADS)

Agenda

ABSTRACT

This project aims to estimate the size of asteroids in our solar system based on various observable characteristics such as their absolute magnitude, albedo, distance from the Sun, and their orbital parameters. Accurately predicting the size of asteroids is crucial for understanding their potential impact on Earth and for planning future space missions to study these objects.The project involves analyzing the set of observational data published by JPL's Solar System Dynamics (SSD) group, to create models that can predict the diameter of asteroids based on their observable characteristics. Ultimately, the goal of the project is to improve our understanding of the asteroid population and its potential impact on our planet. As someone who is truly fascinated by astronomy and data mining, I am looking forward to working on the project that will help me learn captivating facts about the field I take interest in, not only from the textbook, but from the real astronomical data, and also enhance my skills both through careful attention to the details of data and machine learning, when it comes down to the actual use of the knowledge gained during the analysis.

OUTLINE OF THE PROJECT LOOKS LIKE

To begin with, I will be exploring the dataset in search of interesting relationships between the features and explaining the relationships if possible. That is what I will be done in the *Exploring the Data* section. Then I will be preparing the data for training the diameter prediction models further in the project. This step comprises feature engineering, encoding the categorical features, scaling the data. After having done that, the primary focus will be on constructing and fine-tuning three models: MLP neural network, CatBoost, and Light GBM. By exploring and optimizing these models, I aim to identify the most suitable one for our specific task - prediction of asteroid diameter. Through a comparative analysis of their performance, you will see me make an informed decision about which model yields the best results.

INTRODUCTION

- Asteroids are small, rocky objects that revolve around the Sun in elliptical orbits. They are sometimes called minor planets or planetoids and are remnants left over from the early formation of our solar system about 4.6 billion years ago.

- Asteroids range in size from tiny specks to hundreds of kilometers across, and they are located primarily in the asteroid belt between Mars and Jupiter, though, they can also be observed in other regions of the solar system. Most asteroids are irregularly shaped, only some of them are close to spherical, and they are often pitted or cratered. But wait, what do we mean by the diameter of the asteroid if it is of irregular shape?

- Basically, the diameter of an asteroid of irregular shape is usually represented by its **equivalent diameter**, which is the diameter of a sphere that has the same volume as the asteroid. This equivalent diameter is a handy way to compare the size of different asteroids, but it should be noted that it does not necessarily represent the actual physical size or shape of the asteroid. In some cases, the equivalent diameter may be an overestimate or an underestimate of the true size of the asteroid, depending on its shape and density.

- There are **two classes of asteroids** that are of particular interest to astronomers and planetary scientists: **NEAs and PHAs**. NEAs stands for **"Near-Earth Asteroids"**, and, as the name suggests, these are asteroids that come relatively close to Earth's orbit, meaning they have the potential to collide with Earth. On the other hand, PHA stands for **"Potentially Hazardous Asteroids"** and these are a subset of NEAs that are particularly worrisome because they have the potential to come even closer to Earth's orbit, and their size is big enough that they could cause significant damage if they were to collide with Earth.

- Apart from other reasons, diameter prediction is extremely important for identifying NEAs and PHAs, since the size of an asteroid can have a big impact on how it interacts with Earth. Smaller asteroids are more likely to burn up in Earth's atmosphere before reaching the surface, while larger asteroids can cause more damage upon impact. By predicting the diameter of NEAs and PHAs, scientists can better assess the potential threat these asteroids pose to Earth and develop strategies to diminish any potential harm. In addition, the diameter prediction can also provide valuable information about the composition, structure, and evolution of asteroids, which can lead to a deeper understanding of the origins of the Solar System we live in.

LITERATURE SURVEY

The study by Victor Basu titled “Prediction of Asteroid Diameter with the Help of Multi-Layer Perceptron Regressor“ explores the application of machine learning techniques, specifically a Multi-Layer Perceptron (MLP) regressor, to predict the diameter of asteroids. The research addresses the important task of estimating asteroid sizes, which has significant implications for planetary defense and our understanding of celestial bodies. This literature survey aims to provide context and examine related work in the field of asteroid diameter prediction and the use of neural networks for this purpose.

Estimating the diameter of asteroids is a crucial aspect of planetary science and space exploration. Accurate size predictions are essential for assessing the potential impact threat of near-Earth objects and for planning future space missions, such as asteroid mining or exploration. Traditional methods of size estimation involve astronomical observations, radar imaging, and light curve analysis. However, these methods may have limitations in terms of accuracy and coverage.

Victor Basu's research paper introduces the use of an MLP regressor for predicting asteroid diameters. The study likely includes data preprocessing steps, feature selection, and model training details. It is important to explore the specific methodologies and results presented in Basu's paper to understand the contribution in depth.

In conclusion, Basu's work on predicting asteroid diameter using an MLP regressor represents a significant advancement in the field of asteroid characterization. By harnessing the power of machine learning, this research may offer more accurate and efficient methods for estimating asteroid sizes, thereby contributing to our understanding of near-Earth objects and planetary defense efforts.Further investigation into the paper's methodologies and findings is necessary to evaluate its significance and potential applications in asteroid science and space exploration.

EXISTING SYSTEM

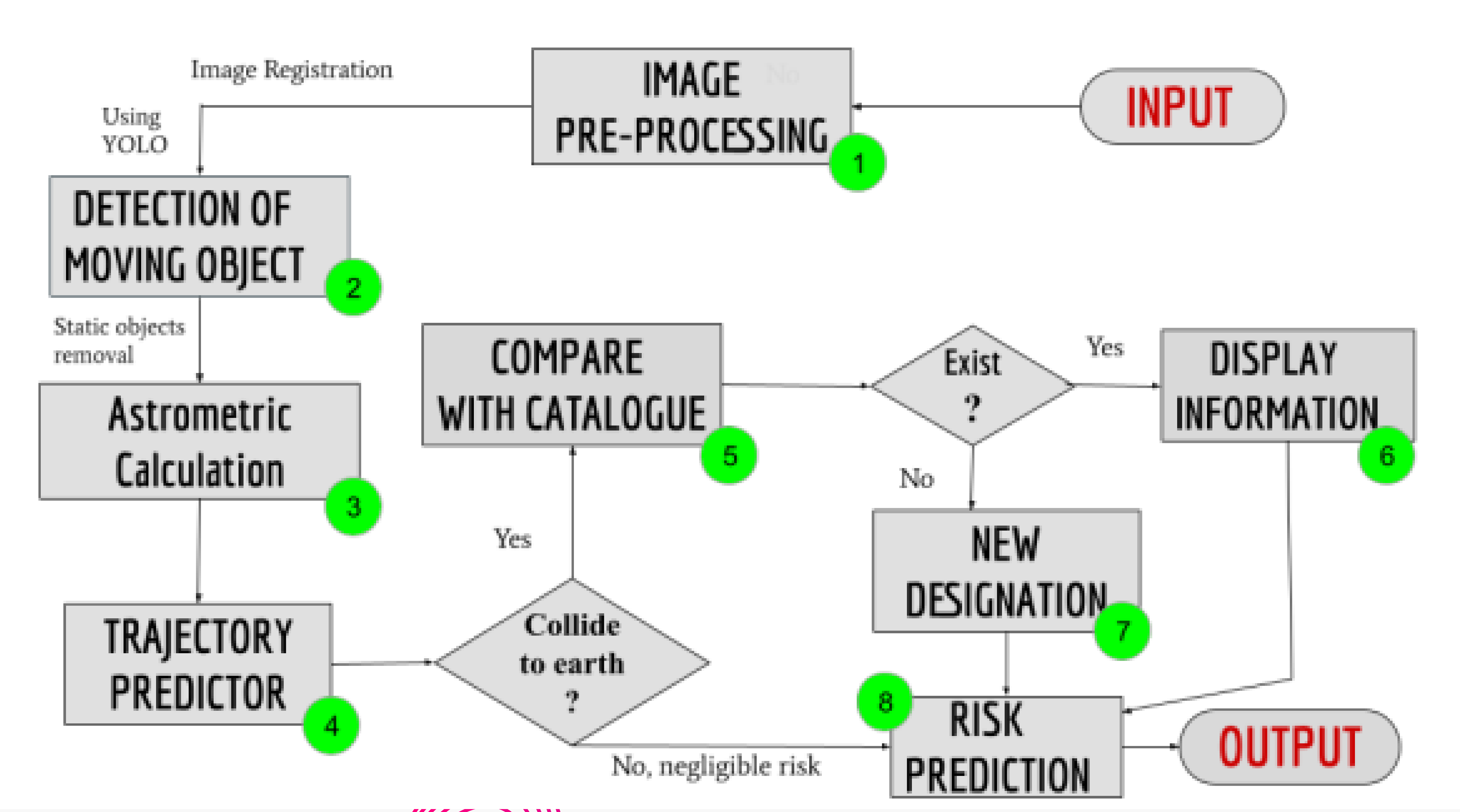
- The existing system for predicting asteroid diameters, as derived from Victor Basu's research paper titled "Prediction of Asteroid Diameter with the Help of Multi-Layer Perceptron Regressor," operates through a well-defined set of processes and components.
- It commences with the collection of diverse data sources, including astronomical databases, ground-based telescopes, and space mission observations, which together form a comprehensive dataset for asteroid analysis.
- This dataset then undergoes meticulous data preprocessing, involving tasks such as cleaning, feature engineering, and normalization, to ensure the quality and consistency of the input data.
- At the core of this system lies a Multi-Layer Perceptron (MLP) regressor, a neural network architecture designed for regression tasks.
- The architecture of this model is optimized through experimentation, and key hyperparameters are tuned during training.
- This training process is complemented by rigorous validation, often employing techniques like cross-validation to assess the model's generalization capabilities effectively.

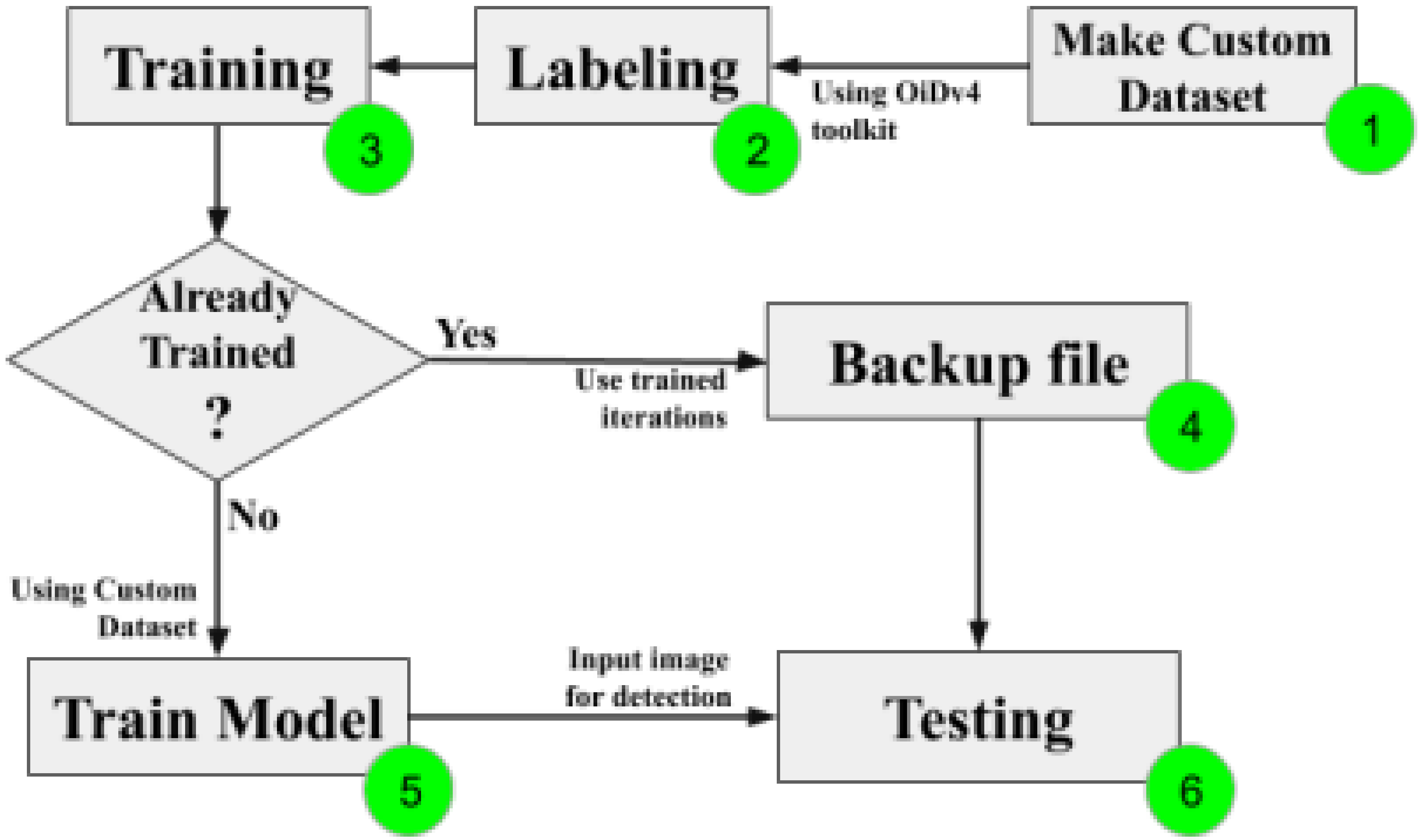
PROPOSED SYSTEM

- The proposed system for predicting asteroid diameters incorporates a systematic approach to ensure accurate and reliable results. It begins with data collection from diverse sources, including astronomical databases and observations, followed by preliminary analysis to assess data quality and structure. Subsequently, Exploratory Data Analysis (EDA) techniques are employed to visualize data distributions and establish correlations between predictor variables such as classes (a, e, ad, q, H, albedo, data_arc, no. of observations) and the target variable, asteroid diameter. Feature engineering is employed to extract relevant attributes, and data processing includes encoding categorical features, splitting the dataset into training and testing subsets, and scaling numerical features to achieve consistent scales.

- The core of the system is a Multi-Layer Perceptron (MLP) regressor, a robust machine learning model known for its predictive capabilities. Model fine-tuning involves optimizing hyperparameters using techniques like grid search or random search. Model performance is assessed using the R-squared (R2) and Mean Squared Error (MSE) metrics, providing insights into the accuracy of the predictions. Importantly, the system conducts a thorough evaluation with and without the inclusion of the approximate diameter feature to understand its impact on predictive accuracy. In conclusion, this proposed system offers a comprehensive framework for advancing our understanding of asteroid characteristics and is particularly noteworthy for achieving high prediction accuracy using the MLP regressor, which has implications for planetary science and space exploration.

ARCHITETURE DIAGRAM





Training
3
Labeling
2
Using OiDv4
toolkit
Make Custom
Dataset
1
Already
Trained
?
Yes
Use trained
iterations
Backup file
4
No
Using Custom
Dataset
Train Model
5
Input image
for detection
Testing
6

NEO AND PHA

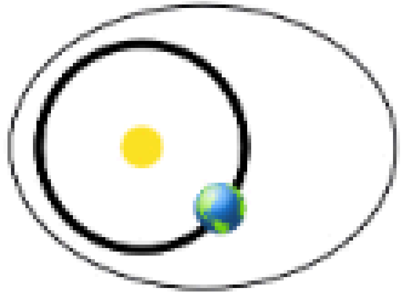
The two most noteworthy classifications for asteroids are “Near Earth Asteroid“ (NEA) or “Near Earth Object“ (NEO) and “Potentially Hazardous Asteroid“ (PHA). Despite the names, neither classification indicates any direct danger to Earth.

Near Earth Asteroids (NEO):

- Near-Earth Objects (NEOs) are comets and asteroids that have been nudged by the gravitational attraction of nearby planets into orbits that allow them to enter the Earth’s neighborhood.
- asteroids with perihelion distance (q) less than 1.3 au. NEAs are divided into groups (Atira, Aten, Apollo, Amor) according to their perihelion distance (q), aphelion distance (ad) and their semi-major axes (a).

Amors

Earth-approaching NEAs with orbits exterior to Earth's but interior to Mars' (named after asteroid (1221) Amor)



$$a > 1.0 \text{ AU}$$
$$1.017 \text{ AU} < q < 1.3 \text{ AU}$$

Apollos

Earth-crossing NEAs with semi-major axes larger than Earth's (named after asteroid (1862) Apollo)

$$a > 1.0 \text{ AU}$$
$$q < 1.017 \text{ AU}$$

Atens

Earth-crossing NEAs with semi-major axes smaller than Earth's (named after asteroid (2062) Aten)

$$a < 1.0 \text{ AU}$$
$$Q > 0.983 \text{ AU}$$

Atiras

NEAs whose orbits are contained entirely within the orbit of the Earth (named after asteroid (163693) Atira)

$$a < 1.0 \text{ AU}$$
$$Q < 0.983 \text{ AU}$$

(q = perihelion distance, Q = aphelion distance, a = semi-major axis)

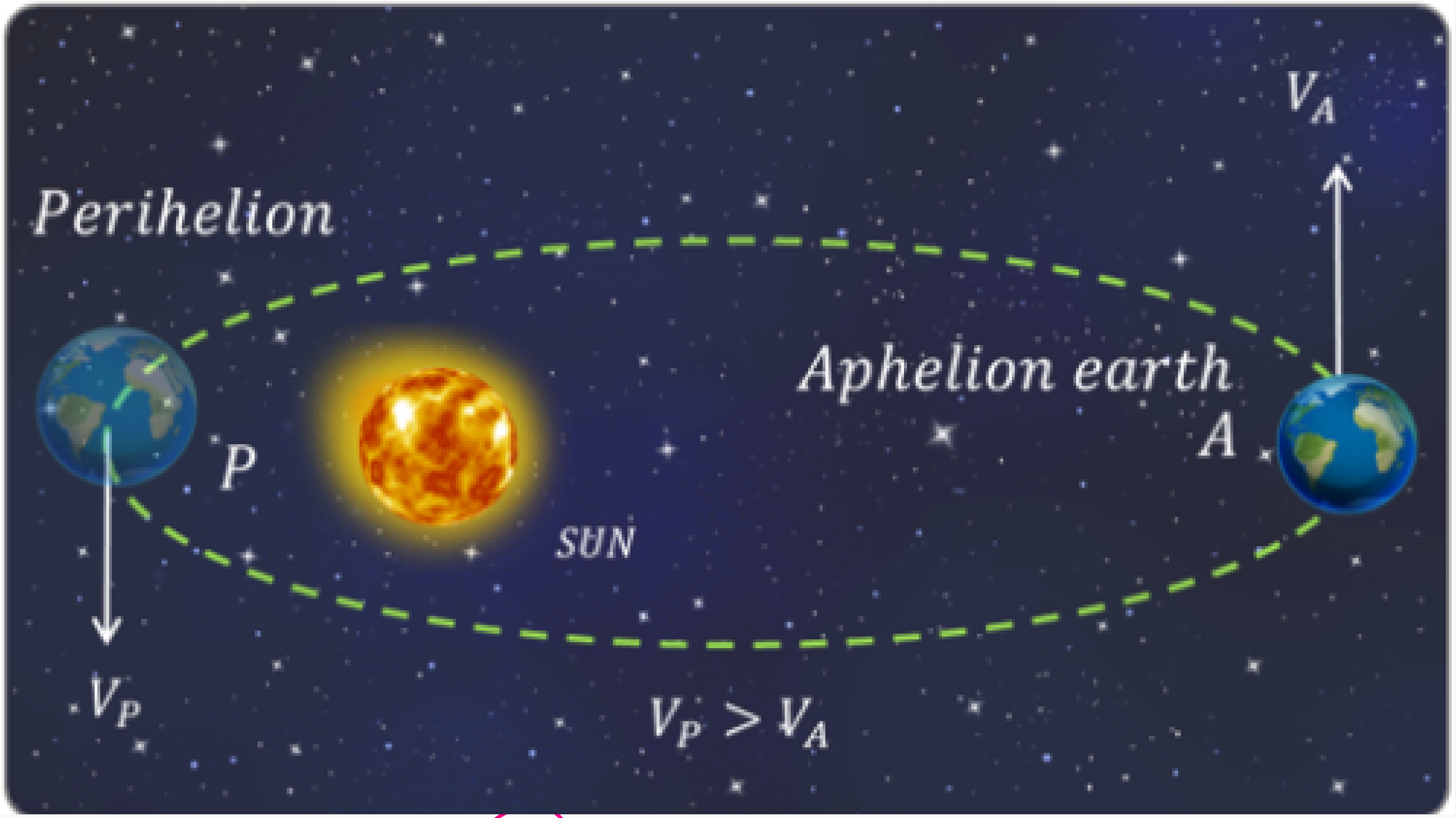
Potentially Hazardous Asteroids (PHA): NEAs whose orbits passes within 0.05 au of Earth's (Earth Minimum Orbit Intersection Distance moid < 0.05 au) and whose absolute magnitude H is 22.0 or brighter (less). PHAs are defined based on parameters that measure asteroid's potential to make threatening close approaches to the Earth. Specifically, asteroids that can't get any closer to the Earth (i.e. moid) than 0.05 au or are smaller than 140 m in diameter (i.e. H=22.0 with assumed albedo of 14%) are not considered PHAs.

PERIHELION AND APHELION

Perihelion Distance : The perihelion distance is an orbital parameter that describes the closest distance between an asteroid and the Sun during its orbit around the Sun. Specifically, perihelion distance is the point in the asteroid's orbit where it is closest to the Sun.

At perihelion, the asteroid is closest to the Sun and therefore receives the highest amount of solar radiation, which can cause its surface to heat up and possibly undergo physical changes such as sublimation of volatile materials.

Aphelion Distance: The aphelion distance of an asteroid is the point in its orbit where it is farthest from the Sun. It is the opposite point of the orbit from the perihelion distance, which is the point where the asteroid is closest to the Sun.



Perihelion
P
SUN
Aphelion earth
A
V_A
V_P
$V_P > V_A$

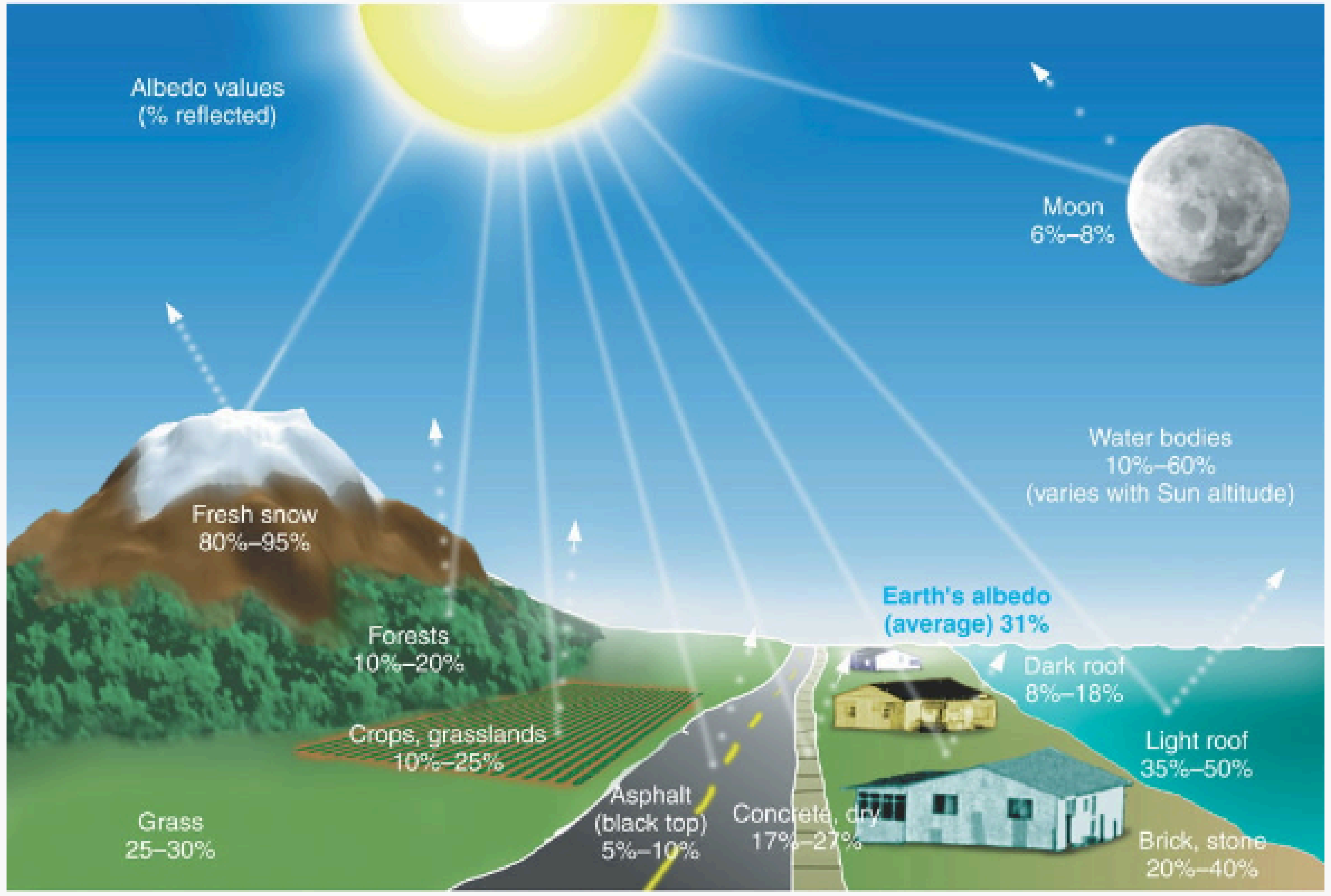
ALBEDO

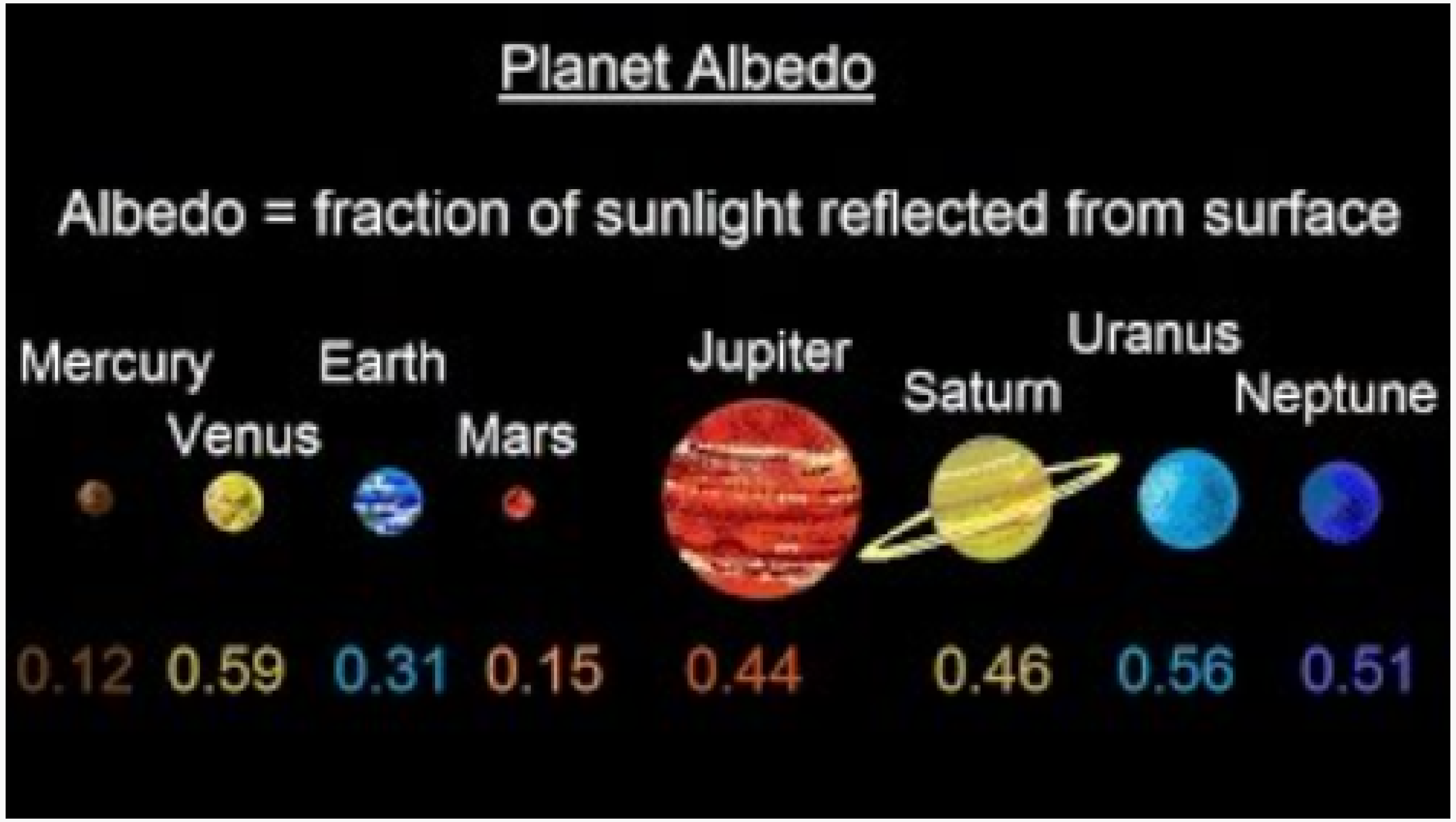
- Albedo in the context of asteroids refers to the measure of the reflectivity of an asteroid’s surface.

- It is defined as the fraction of solar radiation that is reflected by the asteroid’s surface.

- Albedo is usually measured on a scale of 0 to 1, where 0 represents an object that absorbs all incoming radiation and 1 represents an object that reflects all incoming radiation.

- Asteroids with higher albedo appear brighter in telescopes and have a more reflective surface, while asteroids with lower albedo appear darker and have a more absorptive surface. To compare, our Moon has a very low albedo (0.07), while Venus has a high albedo (0.60).





The albedo combined with the absolute magnitude can help determine the size of an asteroid.

CONCLUSION

- The project aimed to assess the performance of three models for predicting asteroid diameters: MLP
- The evaluation was based on two metrics: mean squared error (MSE) and R2 score.
- The MLP model demonstrated the best performance, achieving the lowest MSE of 1.893 and the highest R2 score of 0.978. This indicates that the MLP model provides accurate predictions and effectively explains the variance in asteroid diameters.
- In conclusion, the MLP model proved to be the most suitable and accurate model for predicting asteroid diameters in this project. Although the MLP model showed the best performance, all three models achieved satisfactory results in predicting asteroid diameters, with the MLP model being the most accurate among them, having achieved an MSE of 1.893 and the highest R2 score of 0.978.