Project: NASA Near-Earth Object (NEO) Tracking & Insights using Public API
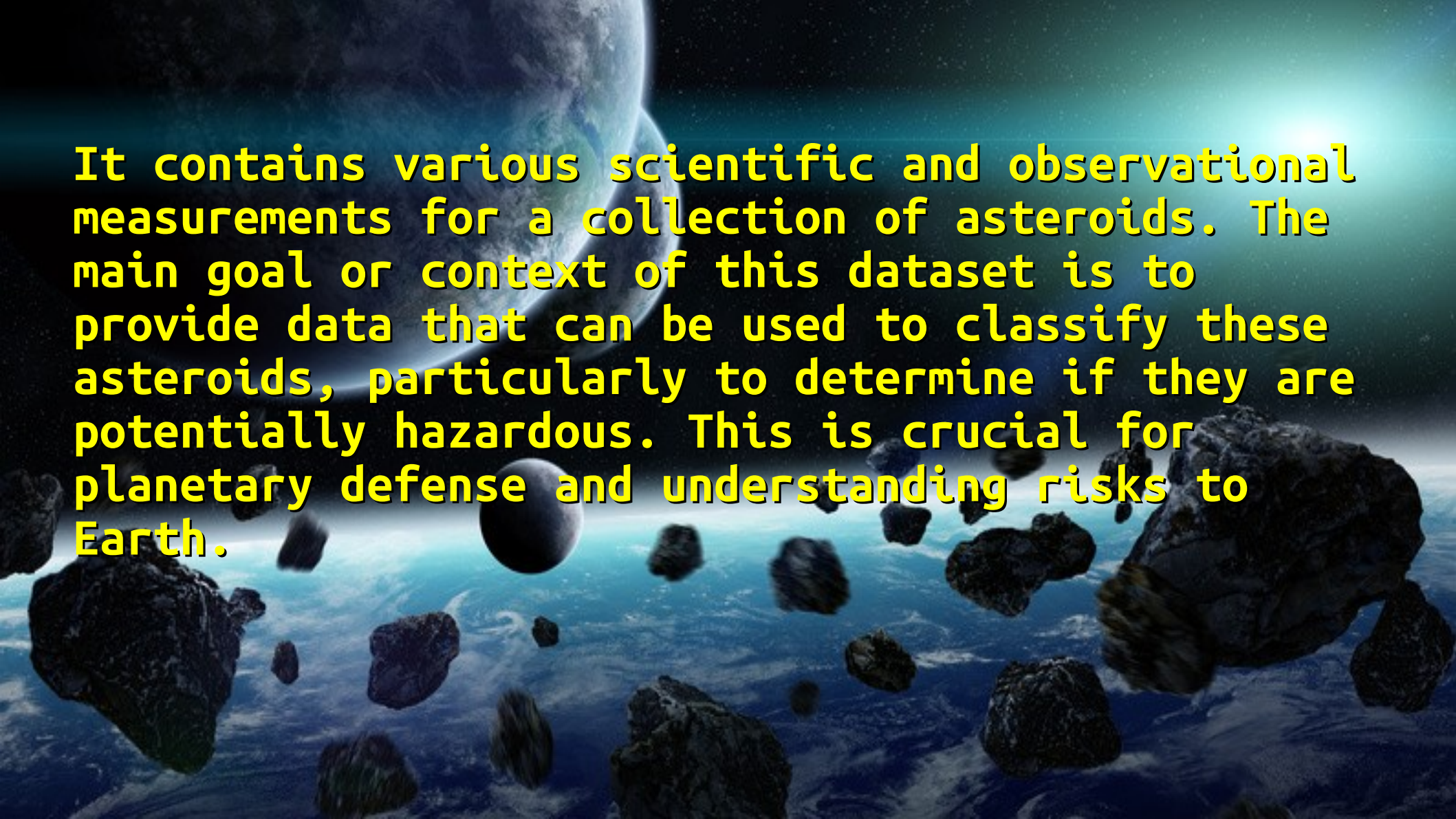
The data is about Asteroids - NeoWs.

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 data-set.

Inspiration

1.Finding potential hazardous and non-hazardous asteroids

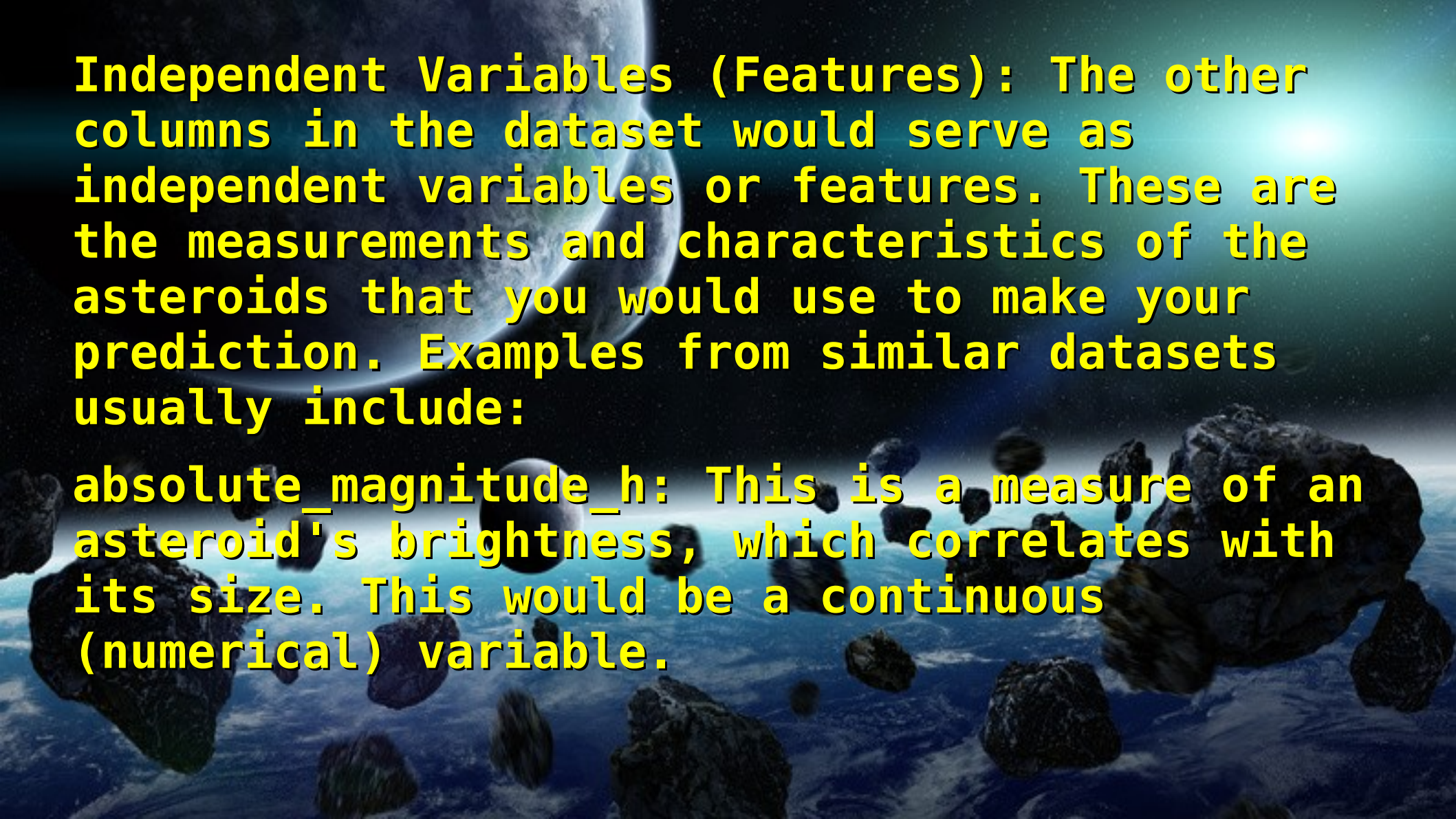2.Features responsible for claiming an asteroid to be hazardous

It contains various scientific and observational measurements for a collection of asteroids. The main goal or context of this dataset is to provide data that can be used to classify these asteroids, particularly to determine if they are potentially hazardous. This is crucial for planetary defense and understanding risks to Earth.
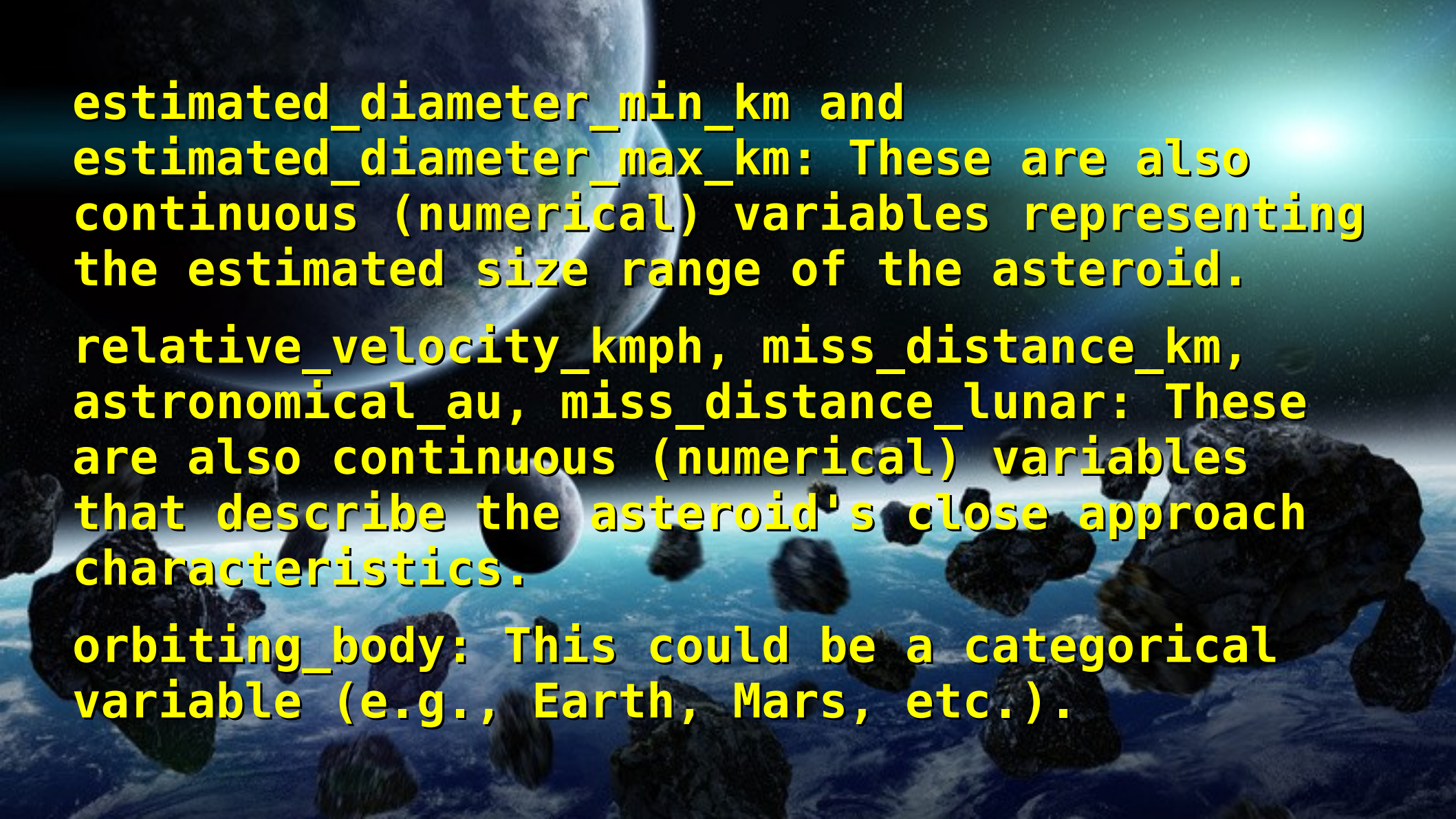
From a machine learning perspective, this dataset is ideally suited for supervised learning, specifically classification.

Dependent Variable: The key dependent variable (what you're trying to predict) would most likely be a column indicating whether an asteroid is is_potentially_hazardous_asteroid or similar. This is a categorical variable, typically represented as True/False or 1/0.

Independent Variables (Features): The other columns in the dataset would serve as independent variables or features. These are the measurements and characteristics of the asteroids that you would use to make your prediction. Examples from similar datasets usually include:
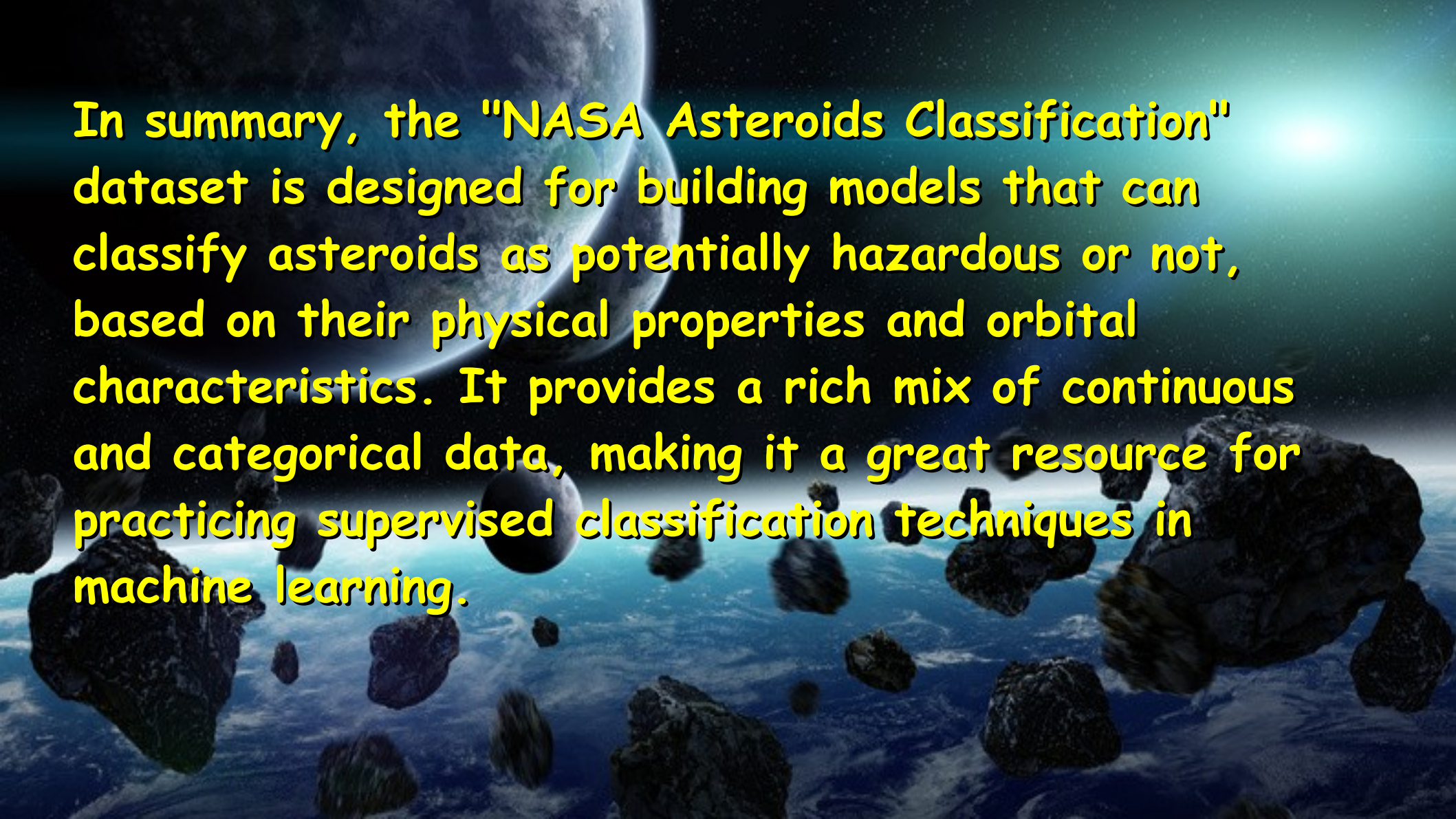
absolute_magnitude_h: This is a measure of an asteroid's brightness, which correlates with its size. This would be a continuous (numerical) variable.

estimated_diameter_min_km and estimated_diameter_max_km: These are also continuous (numerical) variables representing the estimated size range of the asteroid.

relative_velocity_kmph, miss_distance_km, astronomical_au, miss_distance_lunar: These are also continuous (numerical) variables that describe the asteroid's close approach characteristics.

orbiting_body: This could be a categorical variable (e.g., Earth, Mars, etc.).

In summary, the "NASA Asteroids Classification" dataset is designed for building models that can classify asteroids as potentially hazardous or not, based on their physical properties and orbital characteristics. It provides a rich mix of continuous and categorical data, making it a great resource for practicing supervised classification techniques in machine learning.