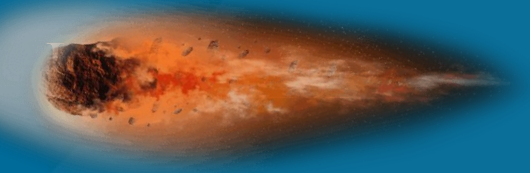
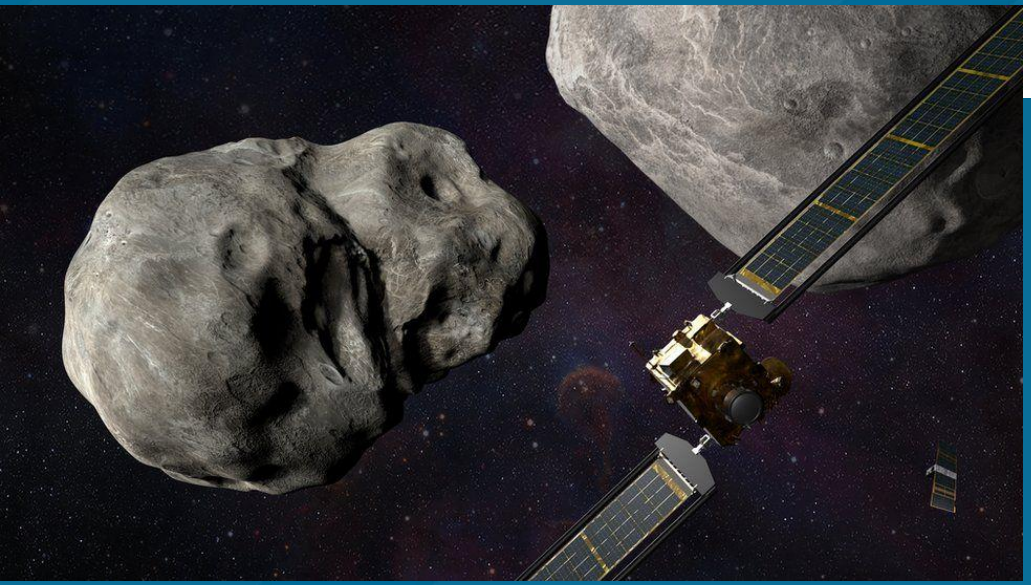


EXPLORATORY DATA ANALYSIS ON ASTEROID FEATURES

Ruining Feng, Andrea Kang, Haozhe Liu, Ameya Mandale, Kunaal Malodhakar



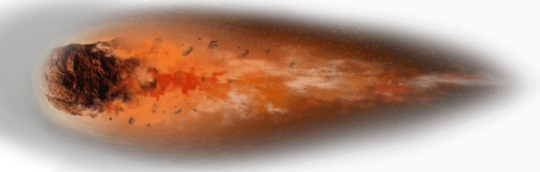
QUESTIONS/OBJECTIVE

- What asteroid features will be critical to inspect in future events?
- To analyze the characteristics of previous asteroids and determine any potential correlations
- Information utilized from this could assist with missions such as DART (what asteroids should we look into for interference?)

DATASETS

ASTEROID	SEMI-MAJOR AXIS	ECCENTRICITY	MAGNITUDE	MEAN MOTION
Phaethon	190197373	0.88989766	14.6	313.938977
Dioretsa	3575929554	0.89950652	13.8	59.8733833
Damocles	1769897305	0.86640318	13.3	250.92822
Jormungandr	219312299	0.85071695	18.6	152.261476

- Open Asteroid Dataset
- NASA Asteroid Classification Dataset
- Variables with little to no data, such as color, were removed



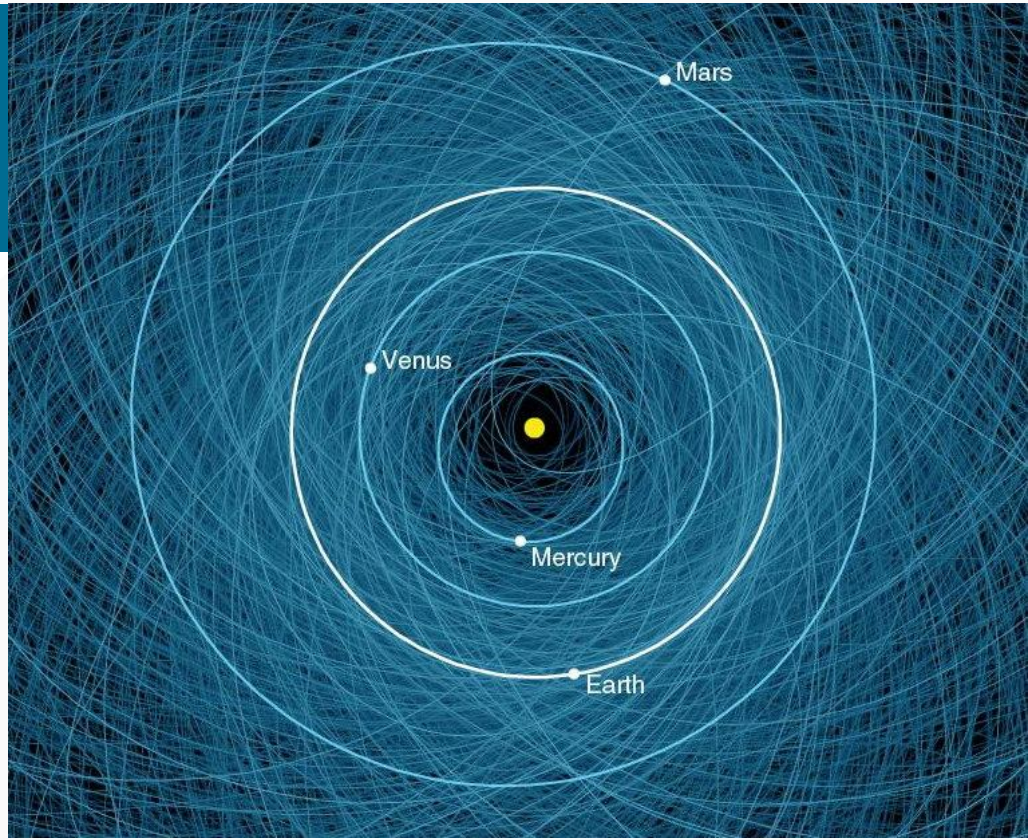
DATASET ANALYSIS

UC San Diego

HAZARDOUS VS NON-HAZARDOUS?

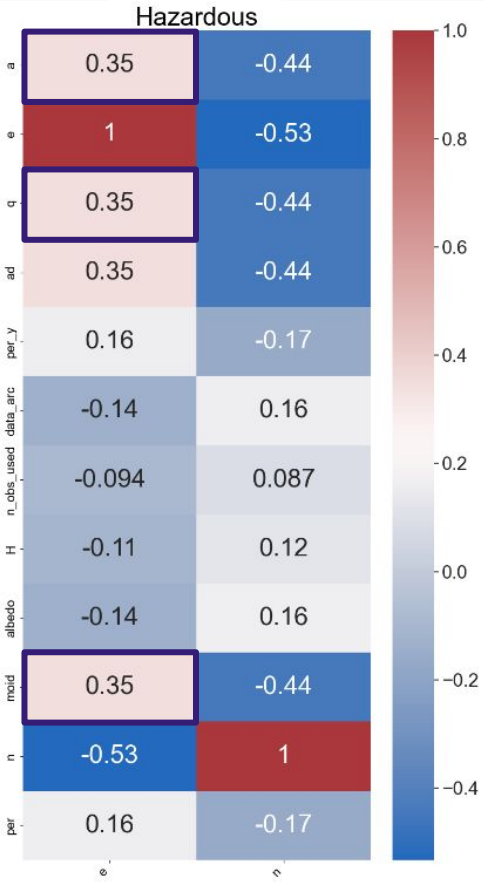
- **Original Condition for Physically Hazardous Asteroids:**
  - Minimum Orbit Intersection Distance (moid) < 7.5 million km
  - Absolute magnitude (H) >= 22

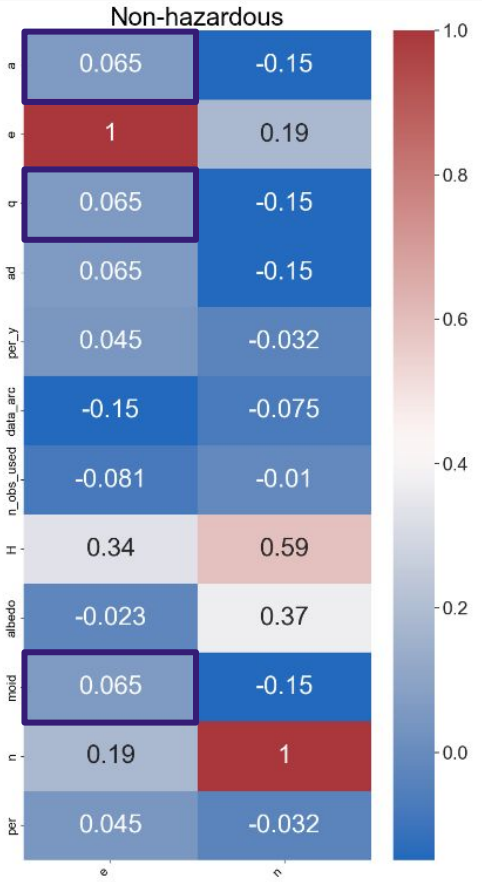
- Are there more correlations across hazardous asteroids beyond this guideline? (For orbit shape and angular orbit speed, potentially)



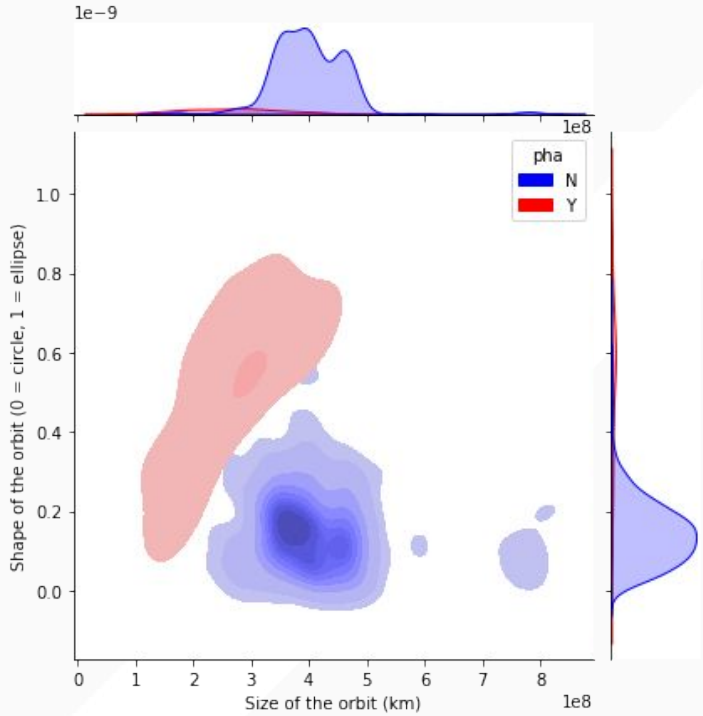
CORRELATIONS

a	Size of the orbit (km)
e	Shape of the orbit (0 = circle, 1 = ellipse)
q	Closest distance from sun (perihelion) in orbit (km)
ad	Farthest distance from sun (aphelion) in orbit (km)
per_y	Orbital period (years)
data_arc	Data arc-span (d)
n_obs_used	# of observations used
H	Asteroid's absolute magnitude
albedo	Geometric albedo (brightness ratio)
moid	Earth Minimum Orbit Intersection Distance (km) To classify hazardous objects
n	Mean motion (deg/day) Angular speed necessary for one complete orbit
per	Orbital period (days)

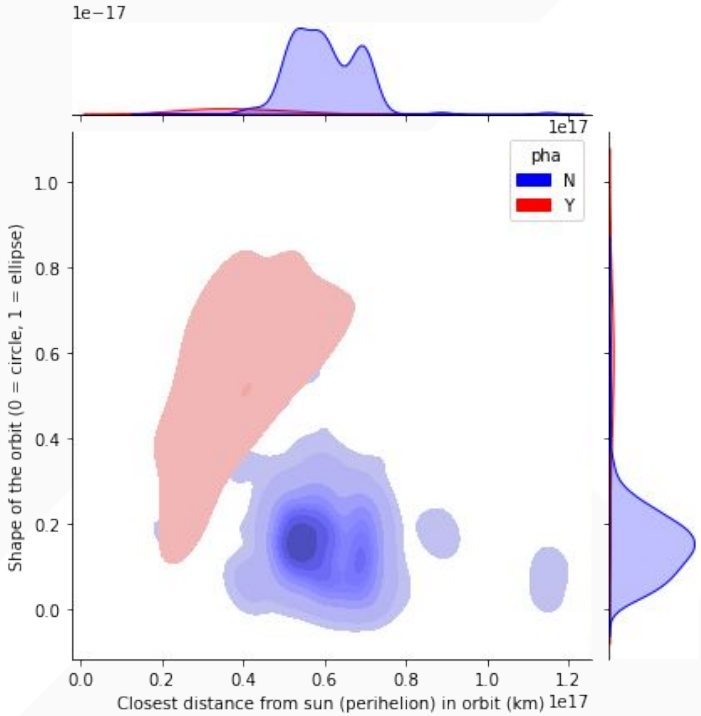




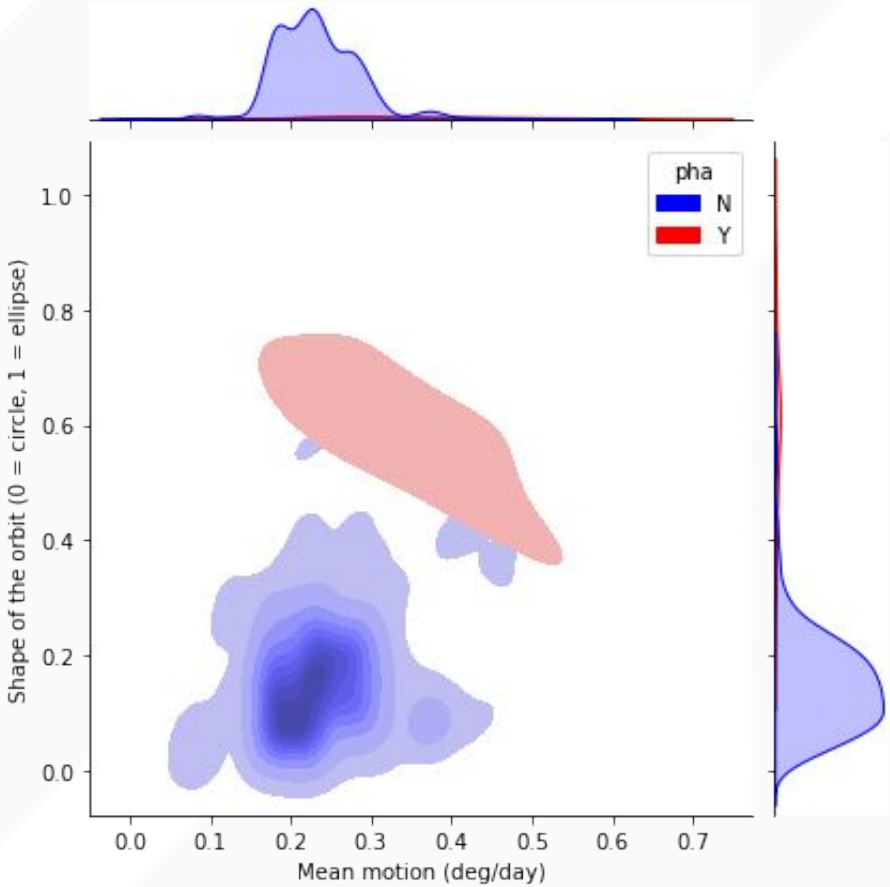
ORBIT SHAPE/ECCENTRICITY



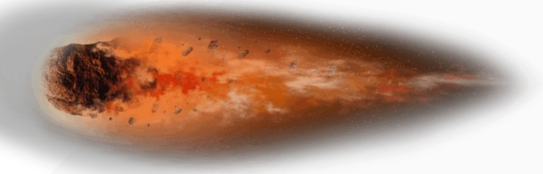
- **Hazardous:**
  - ↑ Eccentricity, ↑ Semi-major axis (shorter than non-hazardous), ↑ Closest sun distance
- **Non-Hazardous:**
  - All variables concentrated in a certain area
  - More circular orbits



ORBIT SHAPE + SPEED



- Eccentricity + Mean angular velocity = Distinct clusters
- **Hazardous:**
  - ↑ Mean angular orbit speed, ↓ Eccentricity
- **Non-hazardous:**
  - Mostly circular orbits compared to hazardous counterparts
  - Angular speed slower on average



CLASSIFICATION MODELS

- Correlated features acquired; how to classify?

UC San Diego

CLASSIFIERS

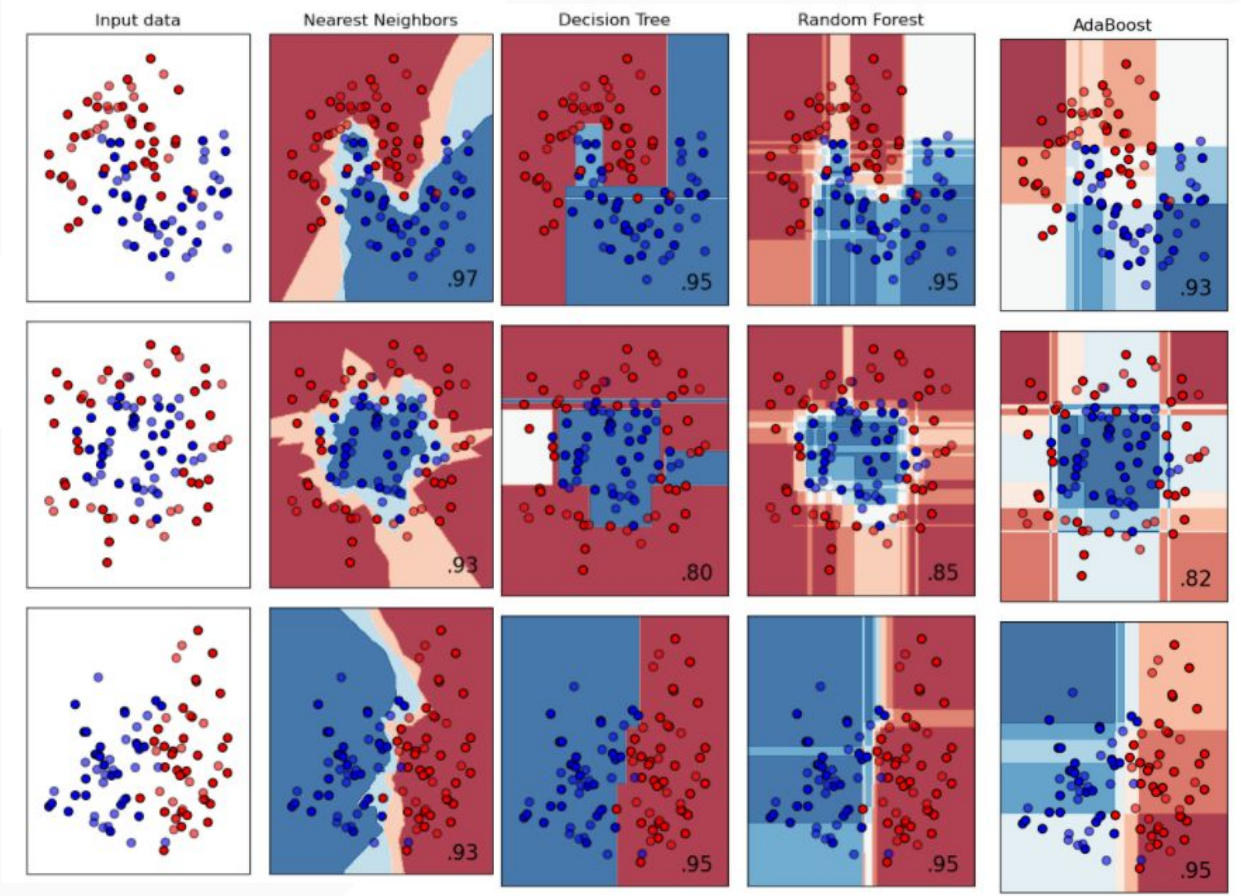
Six classifiers were used for further EDA:

Baseline Classifiers:

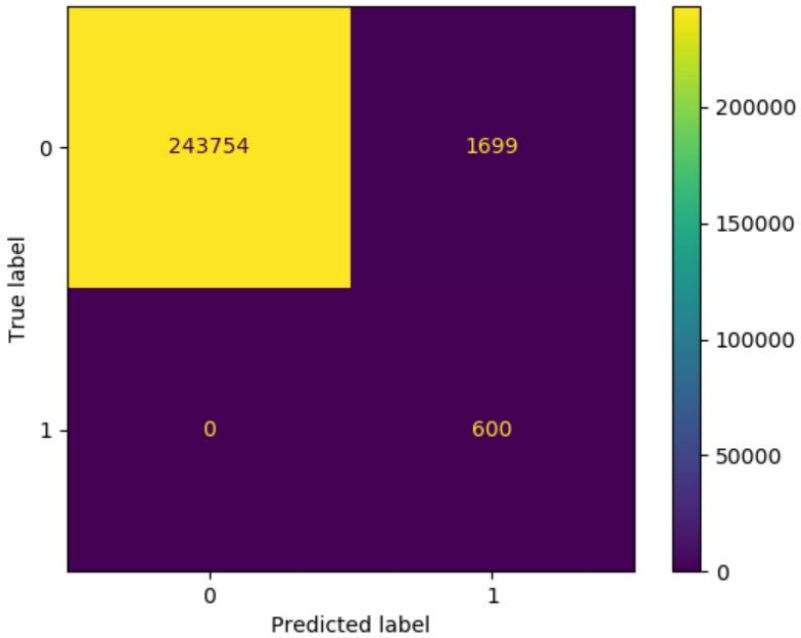
- Logistic Regression
- K-Nearest Neighbor

Target Classifiers:

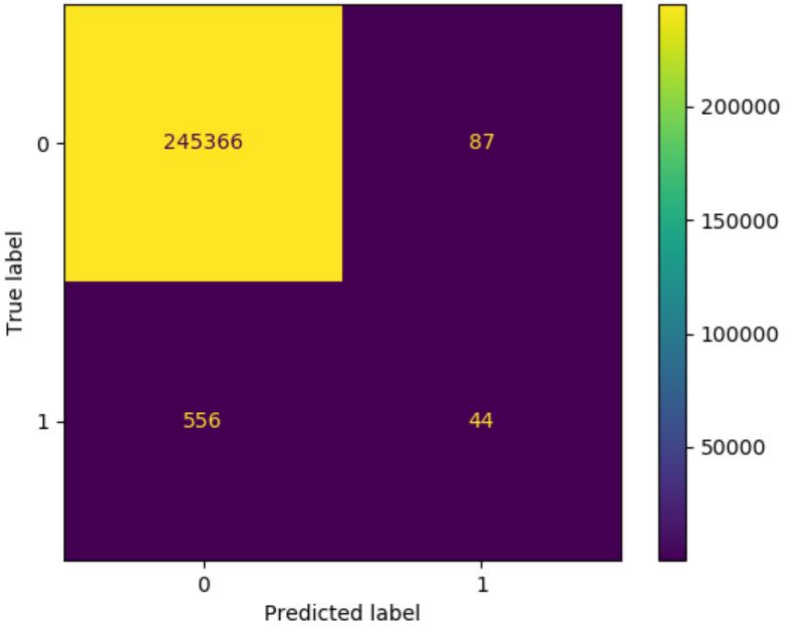
- Random Forest
- Decision Trees
- AdaBoost
- XgBoost



CONFUSION MATRIX (BASELINE)

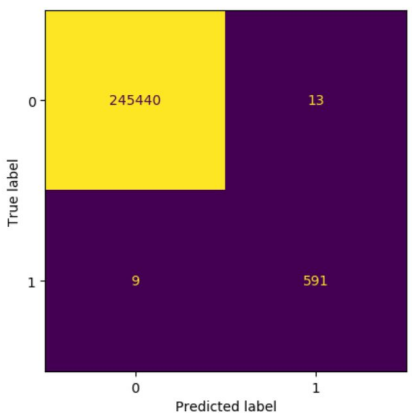


Logistic Regression

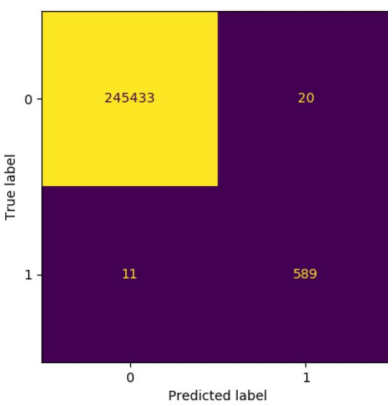


KNN

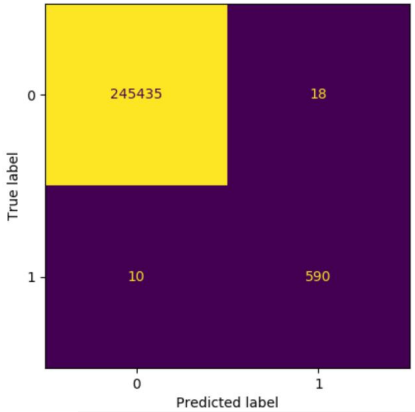
CONFUSION MATRIX (TARGET)



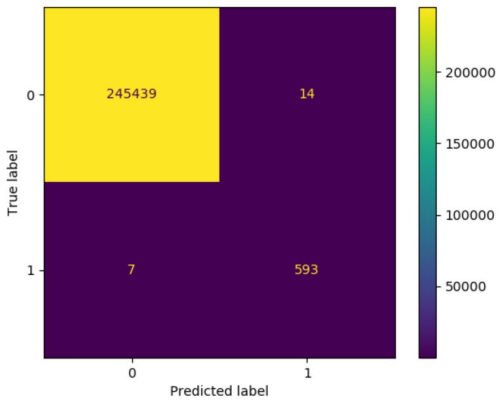
Decision Tree



Random Forest



AdaBoost

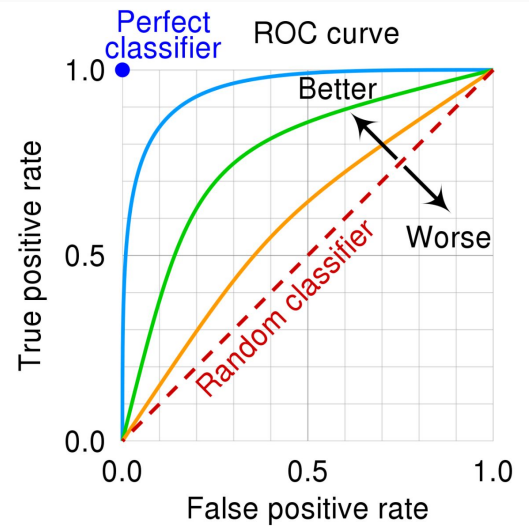


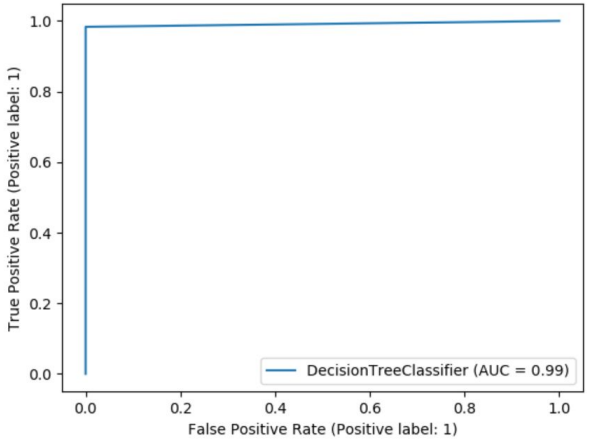
Xgboost

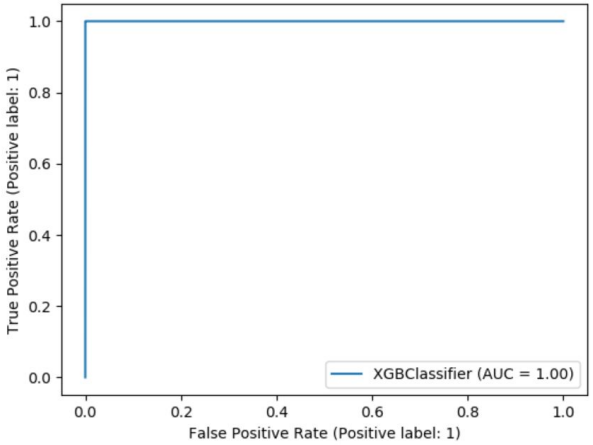
- Many non-hazardous asteroids were classified correctly - this is likely due to the unbalanced ratio of non-hazardous to hazardous asteroids

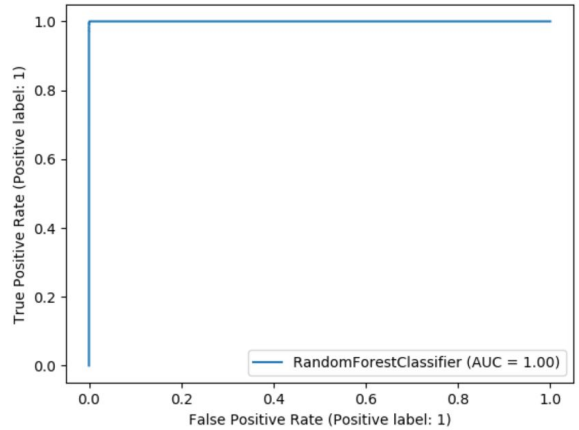
ROC

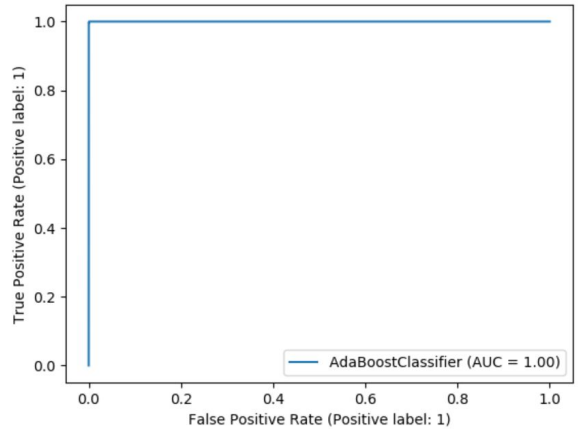
- All target classifiers performed near perfectly based on the ROC



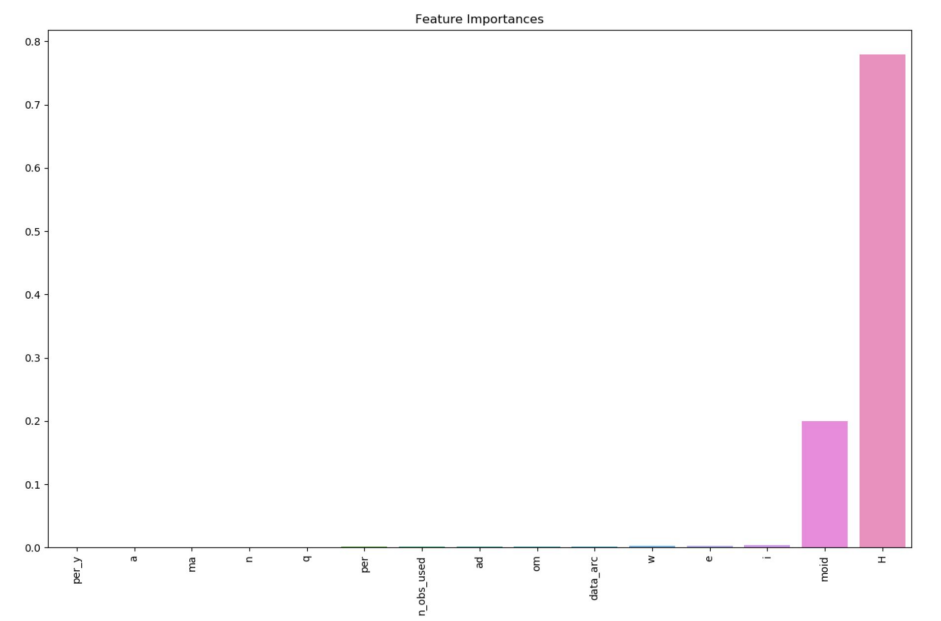




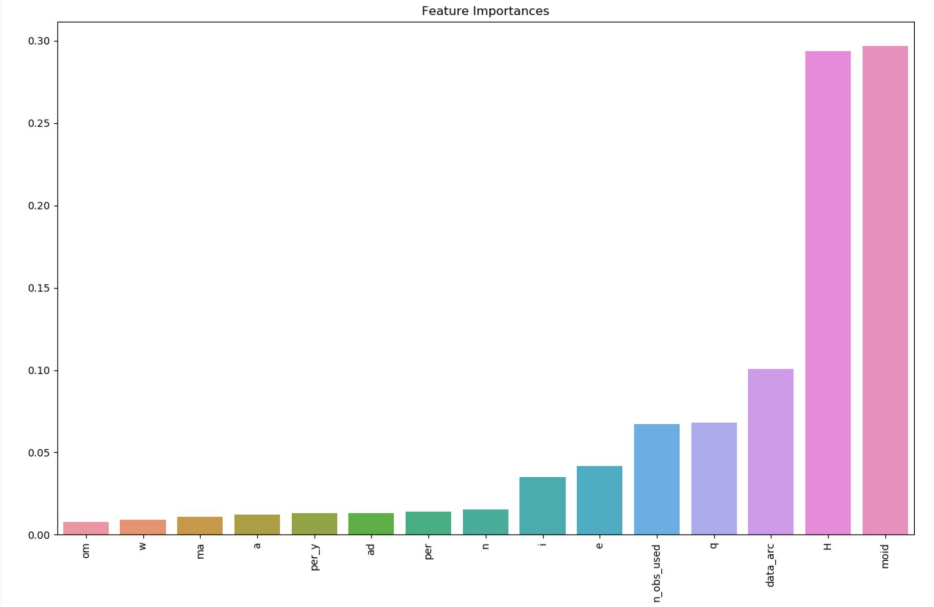




FEATURE IMPORTANCE

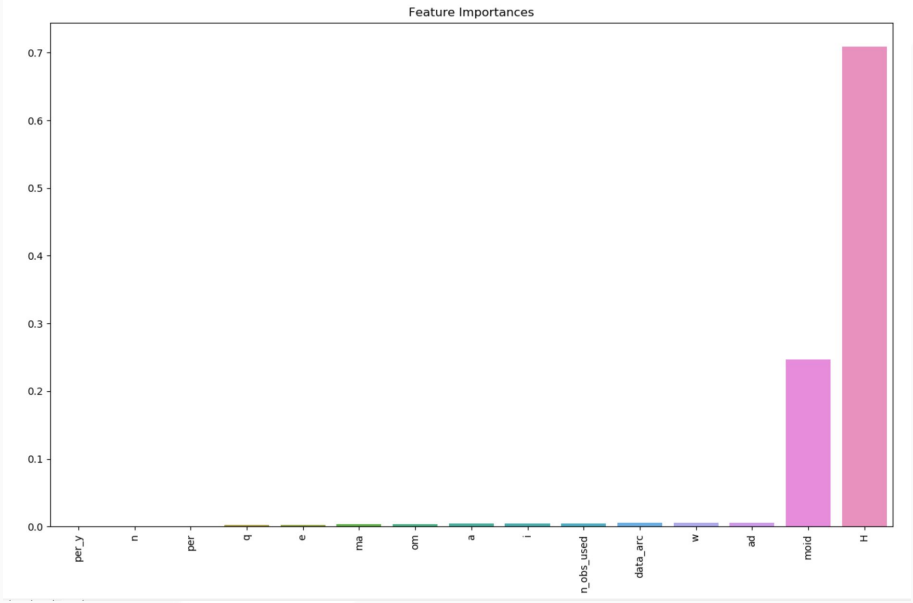


Decision Tree

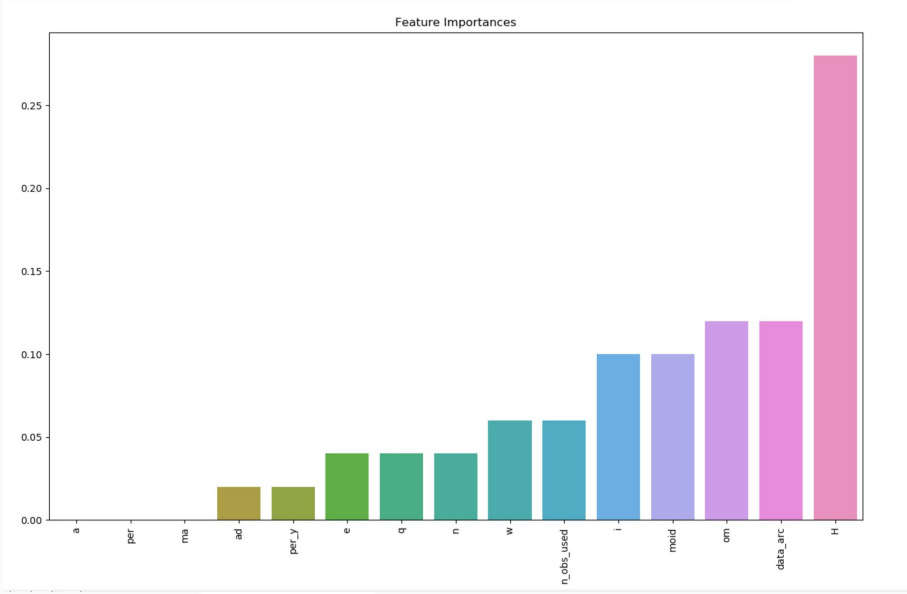


Random Forest

FEATURE IMPORTANCE



AdaBoost



Xgboost

PRECISION/F1

- Precision was high likely due to the large number of non-hazardous asteroids
- However, F1 was as expected for Baseline vs Target Classifiers
  - Baseline: <= 70%
  - Target: > 99%

	Precision	Recall	F1-score
Logistic Regression	0.9930	0.4139	0.7052
Decision Tree	0.9988	0.9768	0.9987
Random Forest	0.9991	0.9817	0.9995
KNN	0.9973	0.3203	0.6432
AdaBoost	0.9988	0.9768	0.9994
XgBoost	0.9991	0.9826	0.9996

TRAJECTORY VISUALIZATION

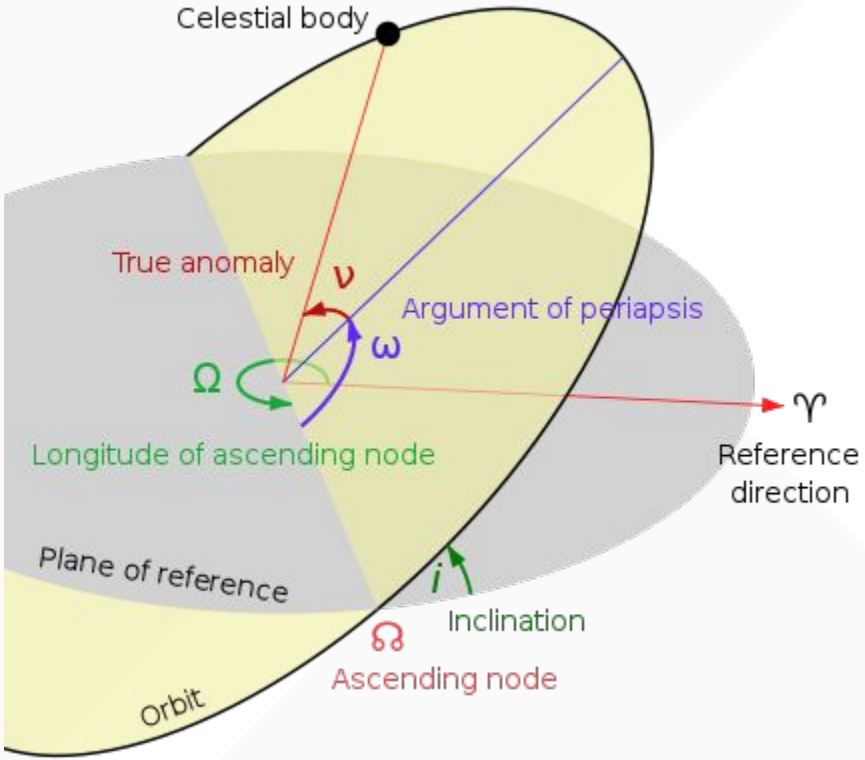
- Lets gaze the Skies

UC San Diego

ORBITAL ELEMENTS

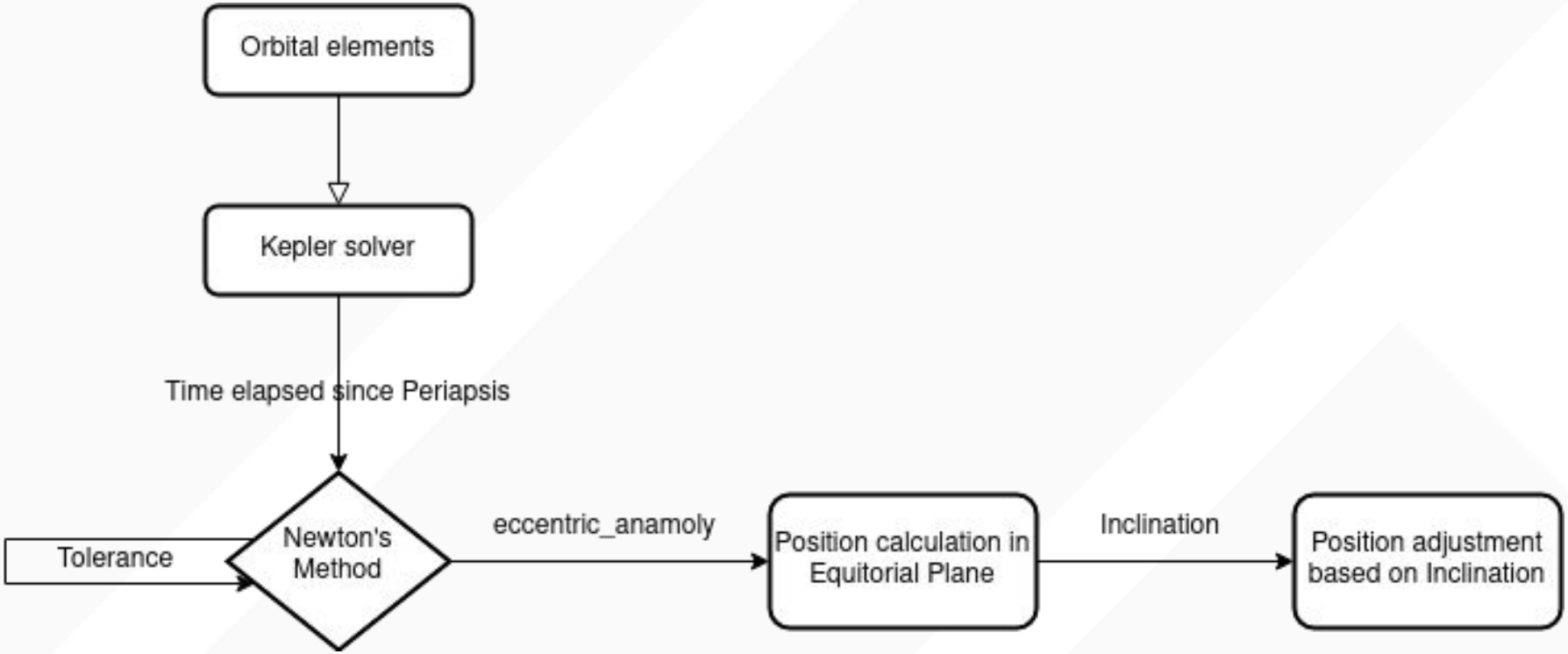
Plane of reference : Equatorial plane

- Semi-major axis
- Eccentricity
- Mean anomaly
- Inclination: *i*
- Argument of Periapsis: ω
- Longitude of Ascending Node: Ω



credits : planetary.org

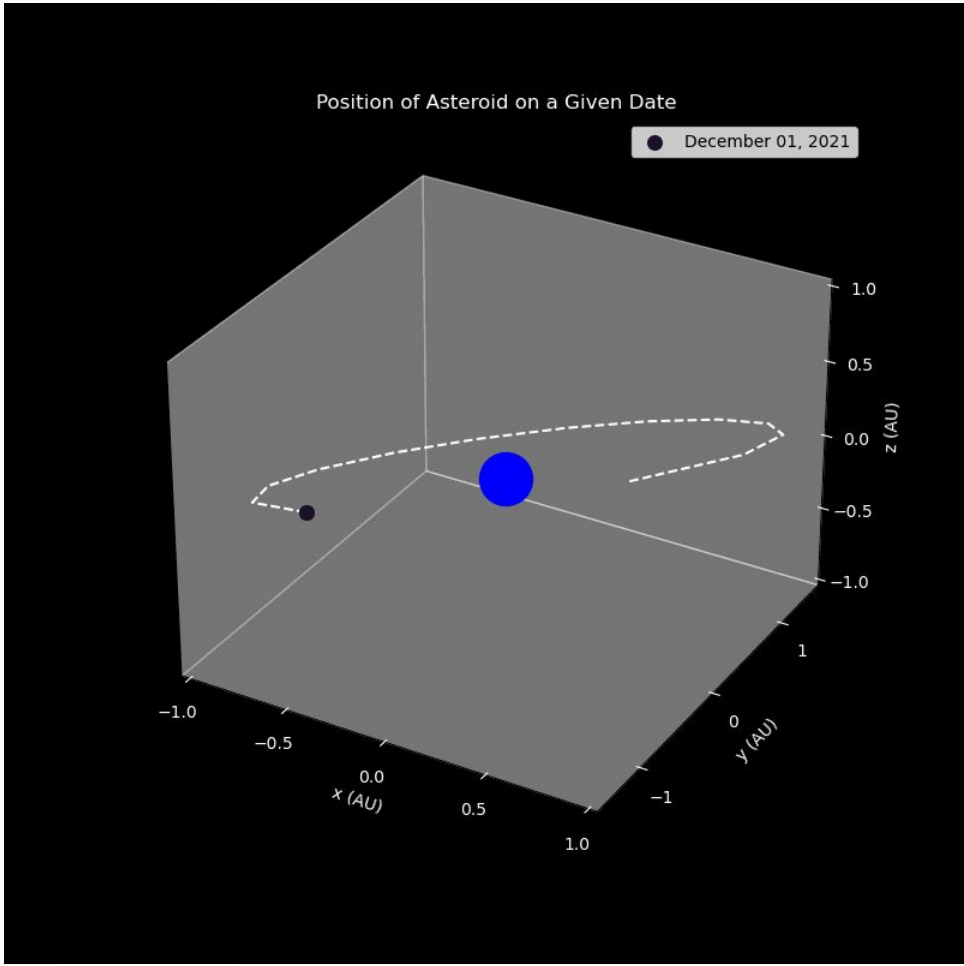
METHODOLOGY



NON-HAZARDOUS

Mapping trajectory of Asteroid ID: 3723955

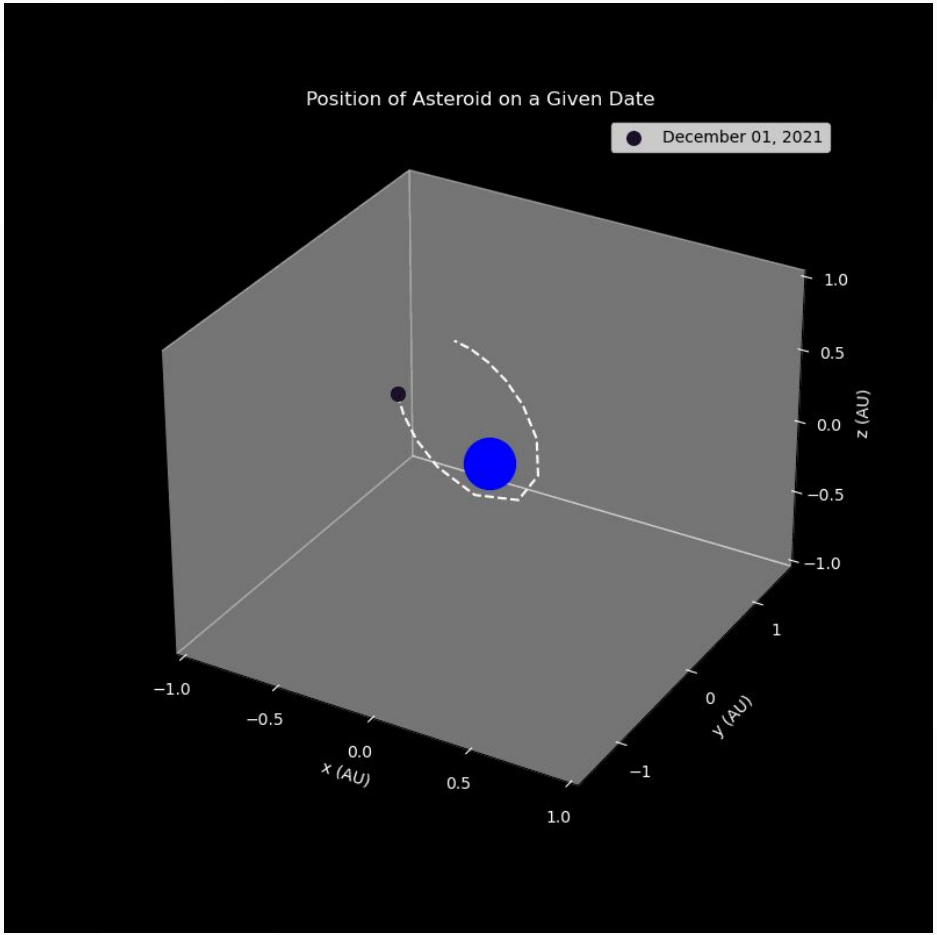
- Mapping from December 2021 to December 2022.
- Orbit shape is fairly elliptical ⟶ e = 0.35
- Orbital Period is estimated at 426 days
- Throughout the asteroid's path, it always maintains a safe distance from earth
- It is marginally closer to Earth on June 1, 2022.



HAZARDOUS

Mapping trajectory of Asteroid ID: 2306383

- Mapping from December 2021 to June 2022.
- Orbit shape is highly elliptical ⟶ e = 0.55
- Orbital Period is estimated at 300 days.
- The asteroid is closest to Earth at around February 1, 2022.
- The trajectory visualization merely confirms result obtained from regression analysis.



CONCLUSION

- For hazardous asteroids:
    - As eccentricity increases → The asteroids tend to travel closer to earth due to orbital shape
    - Orbital period is lesser for asteroids with more elliptical orbits
- Determining orbit shape and speed along with original conditions → Prediction of location in order to interfere with potentially hazardous asteroids.
- Visualization of the asteroid will assist in determining its risk, then trajectory prediction can be used to prevent said asteroid from causing any danger on Earth.
- Prior knowledge to estimate further correlations amongst asteroids to be discovered in the future!