Detection of Potentially Hazardous Asteroids in Near Earth Object

Sahil Gupta
Student at the Department of Computer Science and Engineering
PES University
Bengaluru, India
genius.sahil2001@gmail.com

Sai Mihir J
Student at the Department of Computer Science and Engineering
PES University
Bengaluru, India
saimihir.j@gmail.com

Vishisht Srihari Rao
Student at the Department of Computer Science and Engineering
PES University
Bengaluru, India
vishisht.rao@gmail.com

***Abstract*—Asteroids are irregular shaped rocky bodies that orbit the sun and may or may not be found in the orbital paths of planets. Near Earth Objects are asteroids, comets, or meteors whose trajectories bring it within 1.3 astronomical units of the sun and hence 0.3 astronomical units or 45 million kilometers of the Earth's orbit. Potentially Hazardous Asteroids are asteroids which could potentially make dangerously close approaches to the Earth during its orbit. This paper compares four models which predict whether an asteroid is potentially hazardous or not using the NASA JPL Asteroid Dataset.**

Keywords—PHA, NEO, asteroid, orbit, intersection, magnitude

I. Introduction

Asteroids are formed from the residue of the formation of our solar system about four and a half billion years ago. The formation of Jupiter prevented the further formation of any planetary bodies between Mars and Jupiter. This led to the accumulation of the small objects that were there into the asteroids that we see today. Asteroids can be found throughout the solar system but over 99% of them are found between Jupiter and Mars, this region is known as the main asteroid belt.

Asteroids are too small to be categorized as planets and are sometimes called planetoids or minor planets. Even though there are over 1.1 million asteroids in the solar system, the combined weight of all of them does not exceed the weight of the Earth's moon. The size of asteroids varies greatly, six meters in diameter being one of the smallest which was discovered in 1991 whereas one of the largest to be detected is around 940km in diameter. Asteroids are generally grey in color but owing to space weathering they become redder and darker with time.

Objects such as asteroids, comets, or meteors are classified as Near Earth Objects (NEOs) if their trajectories bring them within 0.3 astronomical units of the Earth. While this does not necessarily mean that these objects will at all times be within 0.3 astronomical units to the Earth, it does indicate that they will at some point in their orbit satisfy this condition. NEOs are of interest as their trajectories allow for simpler observation and analysis through the use of ground based radar. They also provide a representative sample of the materials and history in the solar system. Not all NEOs are potentially hazardous, but all potentially hazardous asteroids fall under NEOs.

Potentially Hazardous Asteroids (PHAs) are asteroids whose trajectories could potentially bring them threateningly close to Earth. The main attributes that contribute toward a NEO being potentially hazardous are MOID (minimum orbit intersection distance) and absolute magnitude (usually denoted by H). If the MOID is less than 0.05 astronomical units and the absolute magnitude is less than 22.0, such asteroids are considered to be potentially hazardous. There are several other mathematical and physical factors that contribute toward an NEO being potentially hazardous. For example, an asteroid satisfying MOID and H conditions may not be potentially hazardous if its size and density are small, making its kinetic energy low.

The article in [1] shows an in-depth study of what the probable consequences would be if an asteroid of diameter 10km, density 2.5 times that of water, impacted at a speed of 20km/s. The immediate impact would cause a large explosion and a crater being formed, in the case that it landed in the ocean, the water displaced would vaporize immediately. If this asteroid landed in the pacific, not only would it cause a megatsunami along the entire pacific rim, it would also result in a complex train of tsunamis. Other severe consequences from such an impact could be a global firestorm or acid rain.

Some noteworthy asteroid impacts on Earth involve the collision 65 million years ago which caused a global firestorm, a global warming that eventually led to the extinction of dinosaurs. In 1490, 10,000 people were killed in the Chinese city of Chiling-yang when an asteroid broke overhead. In 1908, an asteroid exploded 50 meters above the surface of the Earth above Tunguska, Siberia resulting in the loss of 1000 reindeers and trees over the span of 2000 square kilometers.

Considering that such severe outcomes are possible in the case that an asteroid collides with the Earth, it is important to explore the possible preventions and solutions of such an event. To prevent an asteroid from colliding with the Earth altogether we may adopt techniques such as attempting to blow it up using nuclear weapons or projecting electromagnetic energy onto it to throw it off course. The gravitational tractor technique involves sending an unmanned probe on a close orbit around the asteroid to slightly exert some gravitational force on it, over the course of 15 years this could pull the asteroid away from the Earth enough to avoid a collision. Another way to change the course of the asteroid would be to attach a tether to it and a heavy mass on the other end of the tether. While this is not infeasible as space tethers have been used for a long time, it requires time to implement.

As seen in the above paragraph, in order to prepare oneself best in the case of an asteroid impact, it is crucial to accurately detect PHAs as early as possible. Hence, the aim of this paper is to clearly determine which classification model works best in terms of accuracy and performance to predict whether a given asteroid is potentially hazardous or not. The intention is to run multiple classifications models such as RFC, GBC, SVC, LR, ANN, etc. and observe the outcome.

II. REVIEW OF LITERATURE

A. Understanding the distribution of Near-Earth Asteroids (NEAs)

The main aim of the research paper [2], was to understand the distribution of Near-Earth Asteroids. The paper tries to deduce the orbital and size distributions of the near-Earth asteroids (NEAs) via 3 methods: (i) numerically integrating NEAs from their source regions to their observed orbits, (ii) estimating the observational biases and size distribution associated with asteroids on those orbits, and (iii) creating a model population that can be fit to the known NEAs.

The authors in the paper predict that there are ~900 NEAs with absolute magnitude less than 18 (that is, kilometer-sized), of which 29, 65, and 6% reside on Amor, Apollo, and Aten orbits, respectively.

In conclusion, based on the result, this paper proves that roughly 40% of the kilometer-sized NEAs have been found. The remainder, on highly eccentric and inclined orbits, are more difficult to detect.

B. Identifying Potentially Hazardous Asteroids (PHAs) among NEAs

This paper [3], tries to find a new model to identify the hazardous asteroids. As far as asteroids are concerned, there are many asteroids called near-earth asteroids, but all are not hazardous.

The target of this paper was to identify those hazardous asteroids and classify them with non-hazardous types. This was achieved by training various machine learning models with the data features and later comparing those results to find the most accurate model which gives the most accurate classification. The paper's analysis suggests that random forest and xgbclassifier give the most accurate prediction.

C. Lincoln Near-Earth Asteroid Research (LINEAR)

The Lincoln Near-Earth Asteroid Research (LINEAR) program has applied electro-optical technology developed for Air Force Space Surveillance applications to the problem of discovering near-Earth asteroids (NEAs) and comets. This application is natural due to the commonality between the surveillance of the sky for man-made satellites and the search for near-Earth objects (NEOs). Both require the efficient search of broad swaths of sky to detect faint, moving objects.

Currently, the Air Force Ground-based Electro-Optic Deep Space Surveillance (GEODSS) systems, which operate as part of the worldwide U.S. space surveillance network, are being upgraded to state-of-the-art charge-coupled device (CCD) detectors. These detectors are based on recent advances made by MIT Lincoln Laboratory in the fabrication of large format, highly sensitive CCDs.

In addition, state-of-the-art data processing algorithms have been developed to employ the new detectors for search operations. In order to address stressing space surveillance requirements, the Lincoln CCDs have a unique combination of features, including large format, high quantum efficiency, frame transfer, high readout rate, and low noise, not found on any commercially available CCD. Systems development for the GEODSS upgrades has been accomplished at the Lincoln Laboratory Experimental Test Site (ETS) located near Socorro, New Mexico, over the past several years. Starting in 1996, the Air Force funded a small effort to demonstrate the effectiveness of the CCD and broad area search technology when applied to the problem of finding asteroids and comets.

D. Impact Monitoring of Near-Earth Objects (NEOs)

The Impact Monitoring (IM) of Near-Earth Objects (NEOs) is a young field of research, considering that 22 years ago precise algorithms to compute an impact probability with the Earth did not exist.

In the last five years three systems for the detection of imminent impactors (small asteroidal objects detected a few days before the possible impact with the Earth) have been developed: SCOUT (at JPL/NASA), NEORANGER (at University of Helsinki) and NEOScan (at University of Pisa/SpaceDyS). The IM science, in addition to being useful for planetary protection, is a very fascinating field of research because it involves astronomy, physics, mathematics and computer science.

This paper [5] reviews the mathematical tools and algorithms of the IM science, highlighting the historical evolution and the challenges to be faced in the future.

III. DATASET

JPL is a federally funded research and development center managed by Caltech for NASA. It is responsible for exploration of space, other planets, the sun, etc. and aims to understand our place in this universe and to search for the possibility of life beyond earth.

The dataset that has been selected is NASA's Jet Propulsion Laboratory of California Institute of Technology. The entire dataset has been extracted directly from the JPL website and all JPL-authored documents are sponsored by NASA under Contract NAS7-030010.

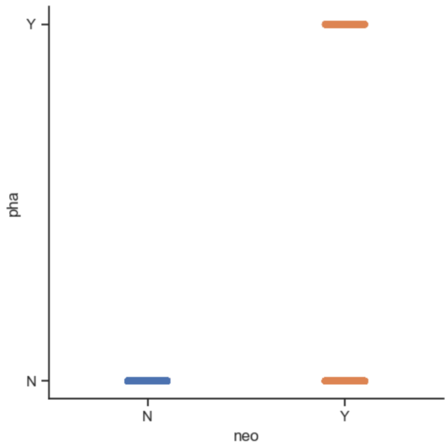
It contains 958,524 unique values which correspond to characteristics of as many asteroids. It has 45 columns of which some important ones are NEO flag, PHA flag, absolute magnitude (H), diameter, geometric albedo, eccentricity, perihelion distance, inclination angle with respect to x-y ecliptic plane, minimum orbit earth intersection distance, etc.

	id	spkid	full_name	pdes	name	prefix	neo	pha	H	diameter	...	sigma_i	sigma_om	sigma_w	sigma_ma	sigma_ad	sigma_n	sigma_tp	sigma_per	class
0	a0000001	2000001	1 Ceres	1	Ceres	NaN	N	N	3.40	939.400	...	4.608900e-09	6.168800e-08	6.624800e-08	7.820700e-09	1.111300e-11	1.196500e-12	3.782900e-08	9.415900e-09	MBA
1	a0000002	2000002	2 Pallas	2	Pallas	NaN	N	N	4.20	545.000	...	3.469400e-06	6.272400e-06	9.128200e-06	8.859100e-06	4.961300e-09	4.653600e-10	4.078700e-05	3.680700e-06	MBA
2	a0000003	2000003	3 Juno	3	Juno	NaN	N	N	5.33	246.596	...	3.223100e-06	1.664600e-05	1.772100e-05	8.110400e-06	4.363900e-09	4.413400e-10	3.528800e-05	3.107200e-06	MBA
3	a0000004	2000004	4 Vesta	4	Vesta	NaN	N	N	3.00	525.400	...	2.170600e-07	3.880800e-07	1.789300e-07	1.206800e-06	1.648600e-09	2.612500e-10	4.103700e-06	1.274900e-06	MBA
4	a0000005	2000005	5 Astraea	5	Astraea	NaN	N	N	6.90	106.699	...	2.740800e-06	2.894900e-05	2.984200e-05	8.303800e-06	4.729000e-09	5.522700e-10	3.474300e-05	3.490500e-06	MBA

5 rows × 45 columns

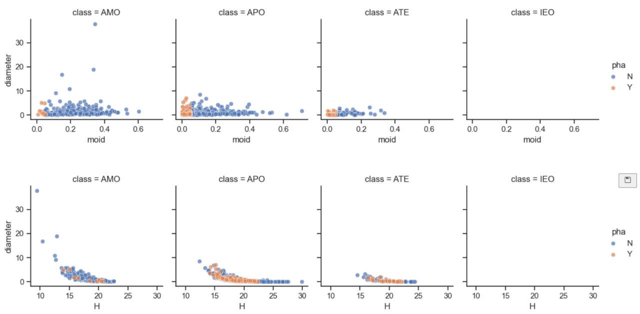
IV. INITIAL INSIGHTS

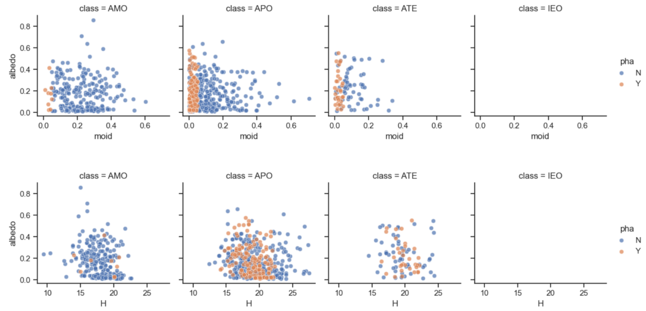
Initially it was noticed that 937,000 asteroids were not potentially hazardous whereas only 2066 were potentially hazardous. Therefore, we attempted to reduce the dataset by plotting NEO vs PHA.

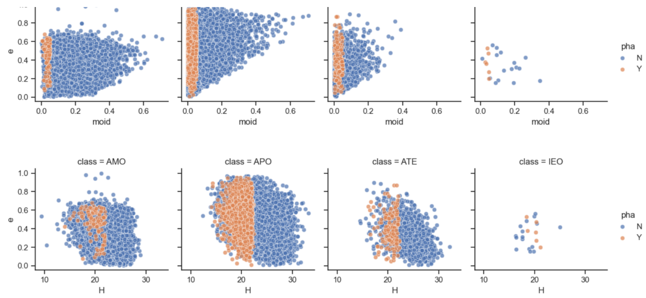


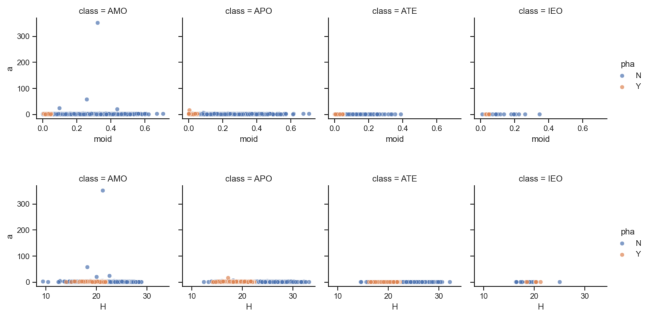
It can be clearly seen from this graph that there is no potentially hazardous asteroid which is not a near earth object.

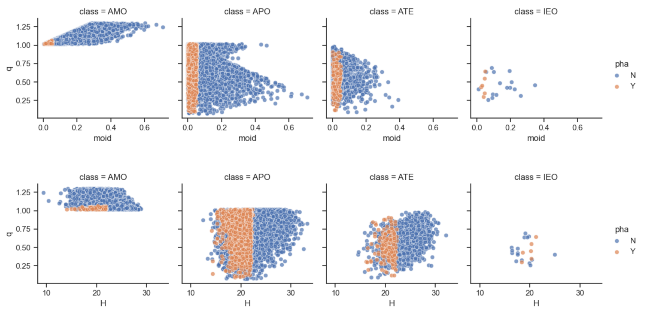
Initially we plot for various classes of asteroids the MOID and H against several parameters such as diameter, albedo, eccentricity (e), semi-major axis (a), perihelion distance (q), inclination angle with respect to x-y ecliptic plane (i), time of perihelion passage (tp).

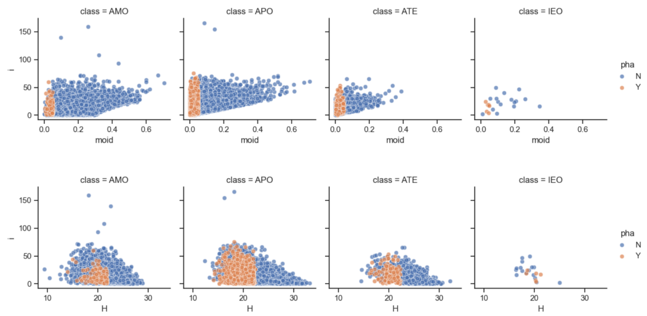


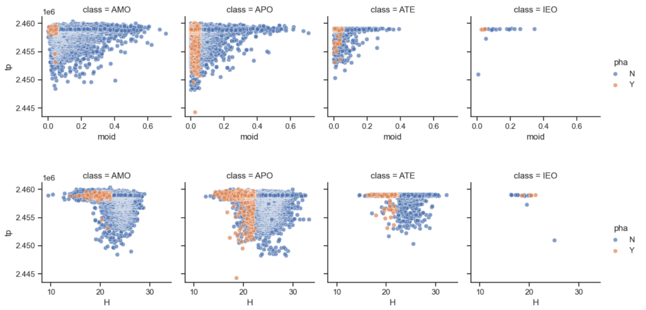










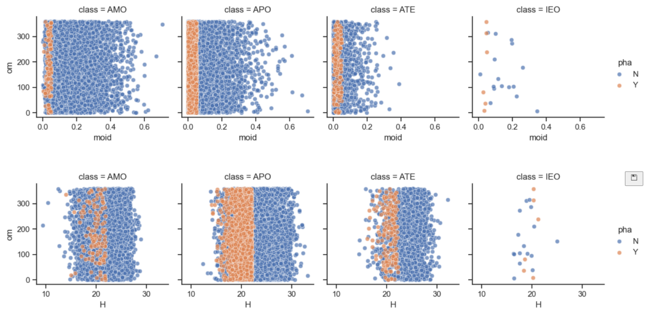


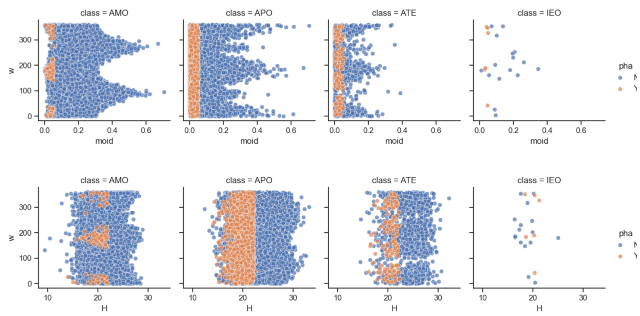
We clearly observe that an asteroid is potentially hazardous only if MOID is less than 0.05 au and H is less than 22.0 holds true. Further it can be seen that for albedo, eccentricity and tp the plots are more scattered as compared to the other parameters. This would mean that it is more difficult to perform the classification based on albedo, eccentricity and tp as compared to the other parameters.

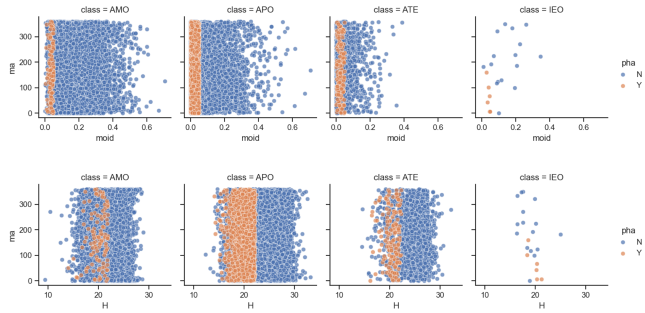
We also plot the spread of these chosen parameters to observe its mean, standard deviation and range.

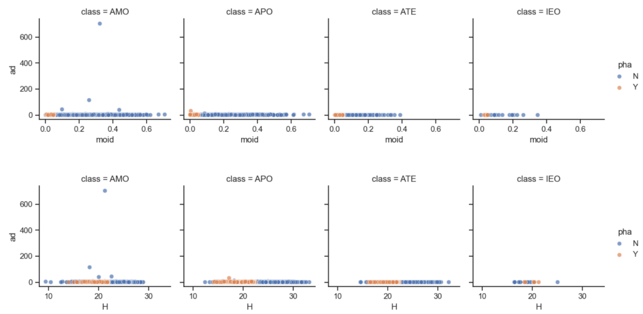
	diameter	albedo	e	a	q	i	tp
count	849.000000	802.000000	22895.000000	22895.000000	22895.000000	22895.000000	2.289500e+04
mean	1.168998	0.180098	0.443916	1.782582	0.913972	12.519644	2.458346e+06
std	1.918393	0.139816	0.176964	2.426534	0.232710	11.125158	1.457585e+03
min	0.002500	0.009000	0.002846	0.555418	0.070511	0.013518	2.444268e+06
25%	0.351000	0.057250	0.313677	1.310277	0.783904	4.580640	2.458453e+06
50%	0.700000	0.153500	0.457464	1.713459	0.963881	8.843493	2.458868e+06
75%	1.400000	0.268750	0.569426	2.194203	1.068407	17.741780	2.459087e+06
max	37.675000	0.856000	0.996476	352.628376	1.299988	165.541000	2.460373e+06

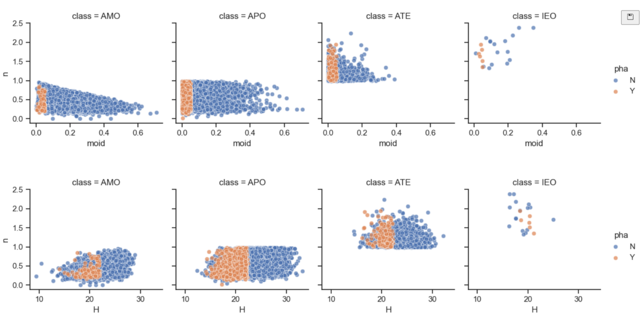
We now plot for various classes of asteroids the MOID and H against some other parameters, namely, omega (om), w, median anomaly (ma), ad, n, rms value of signal (rms).

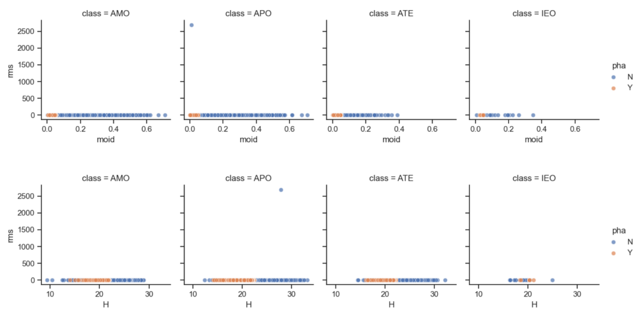












Once again we observe that an asteroid is potentially hazardous only if MOID is less than 0.05 and H is less than 22.0, further, all these plots are more scattered compared to the previous set of plots. We also observe the spread of these features in the following table.

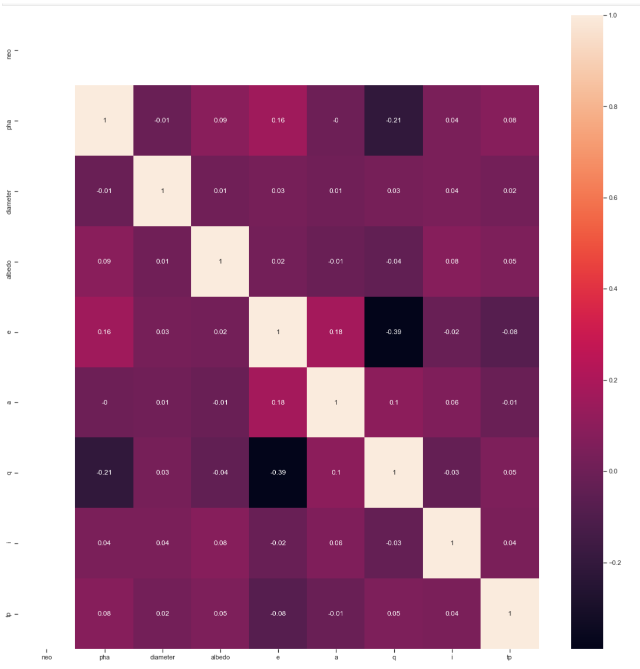
	om	w	ma	ad	n	rms
count	22895.000000	22895.000000	22895.000000	22895.000000	22895.000000	22895.000000
mean	172.560956	181.982241	171.727375	2.651191	0.518787	0.599343
std	103.490498	104.201613	121.187638	4.836224	0.280655	17.752762
min	0.025992	0.007924	0.005210	0.653773	0.000149	0.037189
25%	82.401768	92.478178	52.424668	1.707376	0.303242	0.396545
50%	172.457893	184.088932	163.400370	2.478915	0.439434	0.472730
75%	253.722575	271.584225	290.269795	3.394601	0.657142	0.555490
max	359.977940	359.982032	359.998040	704.014132	2.381082	2686.600000

The initial dataset contained a large number of values (almost one million) of which only about 2000 were potentially hazardous. From our exploratory data analysis we observe that an asteroid is not potentially hazardous if it is not a near earth object. Hence we can bring our dataset down from a million values to approximately 22,000 values of which 2066 are potentially hazardous.

We can further drop values where MOID is greater than 0.05 or H is greater than 22.0 as there is no asteroid which is potentially hazardous for these values. This drastically reduces the size of our dataset such that our models can be trained more efficiently.

We then observed that the majority of the null values in our dataset occurred in the attributes “albedo” and “diameter”. Since these are numerical columns, we replaced these null values with the mean of the attribute.

To ensure that none of our attributes have too much covariance with one another, we plotted a heat map, which verified that the exploratory data analysis on our dataset has been done thoroughly. The results of the heat map are shown below.



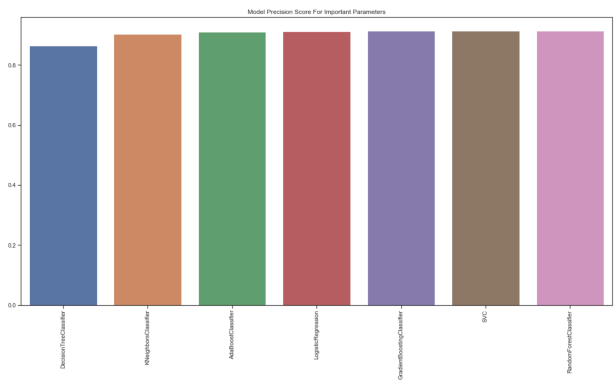
After modifying our dataset as per the results of our exploratory data analysis we may proceed to train it on a number of classification models to determine the optimal model.

V. Methodology

To identify which models should be further explored for this task, a set of models were created which included LR, KNN, SVC, Decision Tree, RF, Adaboost and GBC. The accuracies of these models were obtained on our dataset to provide an insight as to which ones may be more effective. The results are shown below.

	Score
RandomForestClassifier	0.914173
SVC	0.913955
GradientBoostingClassifier	0.913737
LogisticRegression	0.911990
AdaBoostClassifier	0.910898
KNeighborsClassifier	0.902817
DecisionTreeClassifier	0.865036

As seen from the results, all models apart from the Decision Tree Classifier provide similar accuracies when run on our dataset. We decided that we would further explore the Random Forest Classifier, the Gradient Boost Classifier, SVC and an Artificial Neural Network to determine which one of them could give us optimum results. The precision scores for all the models can be seen below and further support our selection.



We begin our analysis on the Random Forest Classifier model. As can be seen from the results below, we obtain an accuracy of 0.9135 along with an F1 score of 0.95.

```
Accuracy: 0.9135182354225814

Classification Report:
              precision    recall  f1-score   support

         0.0       0.92      0.99      0.95      4182
         1.0       0.51      0.07      0.12       397

    accuracy                           0.91      4579
   macro avg       0.71      0.53      0.54      4579
weighted avg       0.88      0.91      0.88      4579
```

We have used only the important parameters that have been identified earlier in the above model. We also captured a classification report when the Random Forest Classifier Model is run with all parameters which can be seen below.

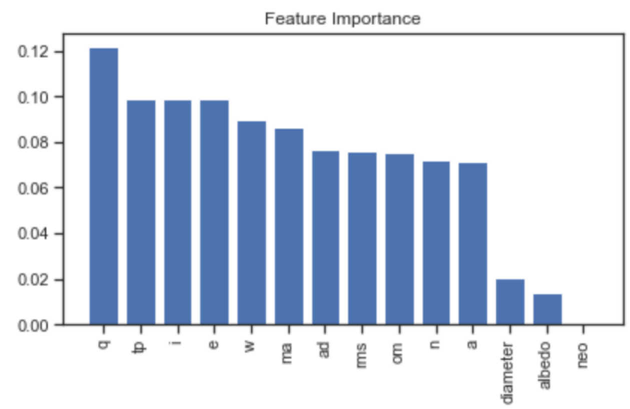
```
Accuracy: 0.9148285651889059

Classification Report:
              precision    recall  f1-score   support

         0.0       0.92      1.00      0.96      4182
         1.0       0.59      0.06      0.10       397

    accuracy                           0.91      4579
   macro avg       0.76      0.53      0.53      4579
weighted avg       0.89      0.91      0.88      4579
```

Interestingly, if we include all the parameters, we obtain a slightly better accuracy of 0.9148 and F1 score of 0.96. To try and ascertain why this may be the case, we plotted the importance of each feature towards classification of the model.



It can be seen that diameter and albedo do not contribute much here towards the classification of an asteroid, whether it is potentially hazardous or not.

We then move on to the next model, the Gradient Boost Classifier model. We ran the model with several different learning rates, all of them returning similar accuracies.

```
Learning rate:  0.05
Accuracy score (training): 0.909
Accuracy score (validation): 0.913
Learning rate:  0.075
Accuracy score (training): 0.909
Accuracy score (validation): 0.913
Learning rate:  0.1
Accuracy score (training): 0.909
Accuracy score (validation): 0.913
Learning rate:  0.25
Accuracy score (training): 0.909
Accuracy score (validation): 0.913
Learning rate:  0.5
Accuracy score (training): 0.909
Accuracy score (validation): 0.913
Learning rate:  0.75
Accuracy score (training): 0.910
Accuracy score (validation): 0.912
Learning rate:  1
Accuracy score (training): 0.910
Accuracy score (validation): 0.910
```

The classification report for the important parameters is shown below.

```
Accuracy: 0.9135182354225814

Classification Report:
              precision    recall  f1-score   support

         0.0       0.91      1.00      0.95      4182
         1.0       0.53      0.02      0.04       397

    accuracy                           0.91      4579
   macro avg       0.72      0.51      0.50      4579
weighted avg       0.88      0.91      0.88      4579
```

We explored the classification report when using all parameters as well and observed the following.

```
Accuracy: 0.9126446822450317

Classification Report:
              precision    recall  f1-score   support

         0.0       0.91      1.00      0.95      4182
         1.0       0.38      0.01      0.02       397

    accuracy                           0.91      4579
   macro avg       0.65      0.51      0.49      4579
weighted avg       0.87      0.91      0.87      4579
```

The accuracy obtained while using the important parameters turned out to be 0.9135 whereas the accuracy while using all the parameters was 0.9126. Hence for this model we can conclude that the parameters that we selected in our exploratory data analysis gives a better result than if we use all the parameters.

The third model that we explored was the Support Vector Classifier. The results for the linear SVC are shown below followed by the results for the SVC using an RBF kernel.

```
Accuracy: 0.911771129067482

Classification Report:
              precision    recall  f1-score   support

         0.0       0.91      1.00      0.95      4182
         1.0       0.11      0.00      0.00       397

    accuracy                           0.91      4579
   macro avg       0.51      0.50      0.48      4579
weighted avg       0.84      0.91      0.87      4579
```

```
Accuracy: 0.9135182354225814

Classification Report:
              precision    recall  f1-score   support

         0.0       0.91      1.00      0.95      4182
         1.0       1.00      0.00      0.01       397

    accuracy                           0.91      4579
   macro avg       0.96      0.50      0.48      4579
weighted avg       0.92      0.91      0.87      4579
```

As seen above, the SVC with the RBF kernel outperforms the linear SVC with an accuracy of 0.9135 against an accuracy of 0.9117.

We then explore a Logistic Regression model and see that it has been outperformed by all the above models with an accuracy of 0.9108.

```
Accuracy: 0.9108975758899323

Classification Report:
              precision    recall  f1-score   support

         0.0       0.91      1.00      0.95      4171
         1.0       0.00      0.00      0.00       408

    accuracy                           0.91      4579
   macro avg       0.46      0.50      0.48      4579
weighted avg       0.83      0.91      0.87      4579
```

For our final model, we move on to Artificial Neural Network. For this network we have added two hidden layers, the first with twelve neurons and the second with eight neurons. We have used the ReLU activation function for these. The output layer consists of a single neuron as this is a binary classification task and uses the sigmoid activation function. The loss function that has been used is the binary cross entropy function and an adam optimizer has been applied as well. The batch size has been set to sixteen. The following are the results of the training.

```
Epoch 1/25
1145/1145 [==============================] - 3s 2ms/step - loss: 574.9518 - accuracy: 0.8331
Epoch 2/25
1145/1145 [==============================] - 2s 2ms/step - loss: 503.8534 - accuracy: 0.8360
Epoch 3/25
1145/1145 [==============================] - 2s 2ms/step - loss: 494.6414 - accuracy: 0.8366
Epoch 4/25
1145/1145 [==============================] - 2s 2ms/step - loss: 376.5508 - accuracy: 0.8359
Epoch 5/25
1145/1145 [==============================] - 2s 2ms/step - loss: 437.3798 - accuracy: 0.8311
Epoch 6/25
1145/1145 [==============================] - 2s 2ms/step - loss: 404.0928 - accuracy: 0.8346
Epoch 7/25
1145/1145 [==============================] - 2s 2ms/step - loss: 363.3628 - accuracy: 0.8361
Epoch 8/25
1145/1145 [==============================] - 2s 2ms/step - loss: 373.8309 - accuracy: 0.8354
Epoch 9/25
1145/1145 [==============================] - 2s 2ms/step - loss: 364.7731 - accuracy: 0.8346
Epoch 10/25
1145/1145 [==============================] - 2s 2ms/step - loss: 318.0490 - accuracy: 0.8341
Epoch 11/25
1145/1145 [==============================] - 2s 2ms/step - loss: 360.6249 - accuracy: 0.8350
Epoch 12/25
1145/1145 [==============================] - 2s 2ms/step - loss: 329.1669 - accuracy: 0.8338
Epoch 13/25
1145/1145 [==============================] - 2s 2ms/step - loss: 317.5667 - accuracy: 0.8354
Epoch 14/25
1145/1145 [==============================] - 2s 2ms/step - loss: 231.6043 - accuracy: 0.8344
Epoch 15/25
1145/1145 [==============================] - 2s 2ms/step - loss: 291.8924 - accuracy: 0.8348
Epoch 16/25
1145/1145 [==============================] - 2s 2ms/step - loss: 262.4603 - accuracy: 0.8333
Epoch 17/25
1145/1145 [==============================] - 2s 2ms/step - loss: 230.9279 - accuracy: 0.8353
Epoch 18/25
1145/1145 [==============================] - 2s 2ms/step - loss: 196.5343 - accuracy: 0.8335
Epoch 19/25
1145/1145 [==============================] - 2s 2ms/step - loss: 199.2587 - accuracy: 0.8369
Epoch 20/25
1145/1145 [==============================] - 2s 2ms/step - loss: 202.9566 - accuracy: 0.8348
Epoch 21/25
1145/1145 [==============================] - 2s 2ms/step - loss: 146.8848 - accuracy: 0.8335
Epoch 22/25
1145/1145 [==============================] - 2s 2ms/step - loss: 151.8328 - accuracy: 0.8322
Epoch 23/25
1145/1145 [==============================] - 2s 2ms/step - loss: 155.8151 - accuracy: 0.8335
Epoch 24/25
1145/1145 [==============================] - 2s 2ms/step - loss: 140.4905 - accuracy: 0.8348
Epoch 25/25
1145/1145 [==============================] - 2s 2ms/step - loss: 136.9465 - accuracy: 0.8362
144/144 [==============================] - 0s 708us/step - loss: 47.4861 - accuracy: 0.9133

Accuracy of test: 91.33
```

VI. Conclusion and Future Works

We conclude that among the models chosen, Random Forest Classifier gave us the highest precision score. We also conclude that our assumptions of choosing certain parameters as important turned out to be right as it had an equal influence for both the datasets and overall impact on the influence of near-earth asteroids.

In the future, we can work on improving each individual model by fine tuning hyper parameters. We can also test each model on different datasets collected to see how they would perform on new data as well as real time data.

Acknowledgment

We would like to express our sincere gratitude to Dr. Gowri Srinivasa and all members of the Data Analytics team for providing us with this incredible opportunity to work on this project and enabling us to gain practical experience in this field. We would also like to thank them for their support and guidance along the way. We would like to extend our gratitude towards the department of Computer Science and Engineering at PES University for providing us with the resources to work on this project and for inspiring us throughout.

References

[1] 1989PASP..101..500V Page 500 (harvard.edu) J. Clerk Maxwell, A Treatise on Electricity and Magnetism, 3rd ed., vol. 2. Oxford: Clarendon, 1892, pp.68–73.

[2] William F. Bottke and Robert Jedicke and Alessandro Morbidelli and Jean-Marc Petit and Brett Gladman. "Understanding the Distribution of Near-Earth Asteroids." DOI: https://doi.org/10.1126/science.288.5474.2190. URL: https://www.science.org/doi/abs/10.1126/science.288.5474.2190

[3] "Hazardous Asteroid Classification through Various Machine Learning Techniques" URL: https://www.irjet.net/archives/V7/i3/IRJET-V7I31084.pdf.

[4] Grant H. Stokes, Jenifer B. Evans, Herbert E.M. Viggh, Frank C. Shelly, Eric C. Pearce. "Lincoln Near-Earth Asteroid Program (LINEAR)", Icarus, Volume 148, Issue 1, 2000. ISSN: 0019-1035. DOI: https://doi.org/10.1006/icar.2000.6493. URL: https://www.sciencedirect.com/science/article/pii/S0019103500964935.

[5] Tommei, G. On the Impact Monitoring of Near-Earth Objects: Mathematical Tools, Algorithms, and Challenges for the Future. Universe 2021, 7, 103. DOI: https://doi.org/10.3390/universe7040103. URL: https://www.mdpi.com/journal/universe.