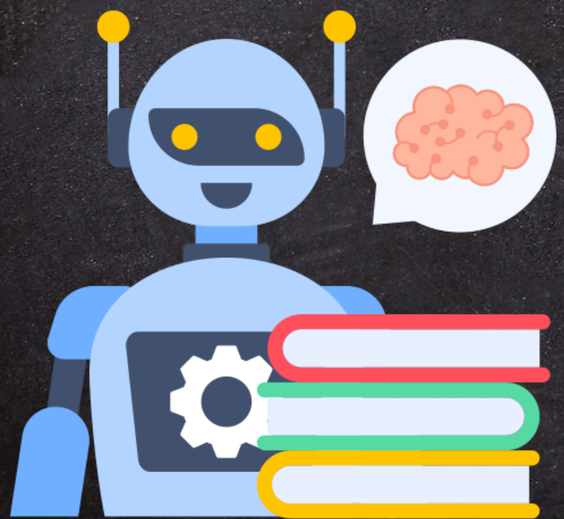
ASTRONOMIX

TEAM:INFINITY8
(PUNJAB ENGINEERING COLLEGE)

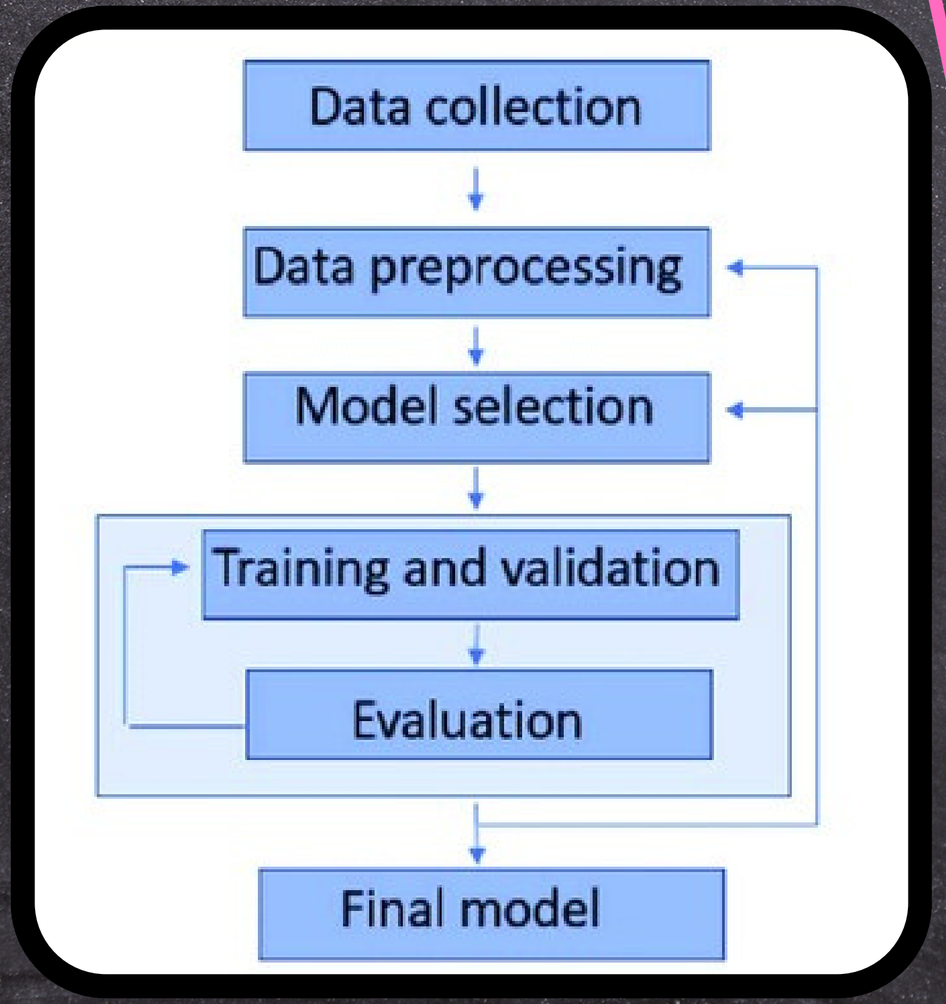
CHALLENGE OVERVIEW: BUILD PREDICTIVE MODELS THAT CAN DETERMINE WHETHER AN ASTEROID IS ON A COLLISION COURSE WITH EARTH OR NOT.

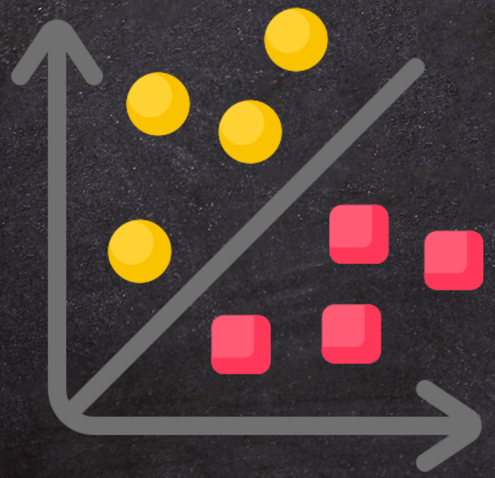
ASTEROID COLLISION PREDICTOR

HERE'S A STEP-BY-STEP GUIDE ON HOW TO APPROACH THIS PROBLEM:



objectives


Data collection
Data preprocessing
Model selection
Training and validation
Evaluation
Final model



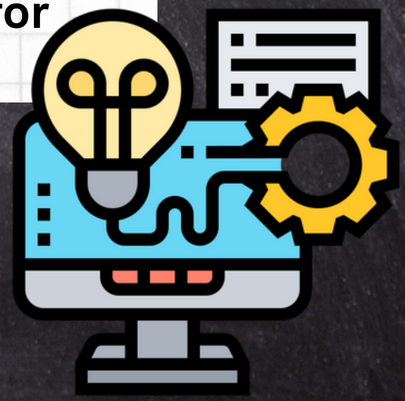
Step 1: Importing the dependencies

PANDAS :pandas is a fast,powerful, flexible and easy to use open source data analysis and manipulation tool, built on top of the Python programming language.

MATPLOTLIB:Matplotlib is a comprehensive library for creating static, animated, and interactive visualizations in Python.

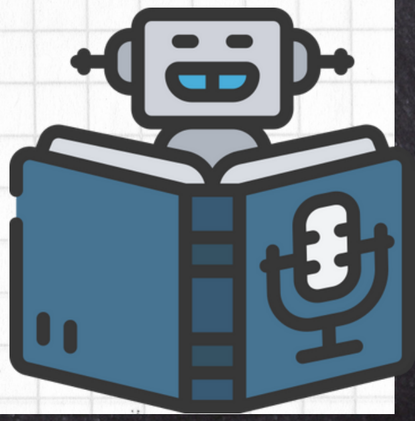
SKLEARN:Scikit-learn (Sklearn) is the most useful and robust library for machine learning in Python.

NUMPY:NumPy is a Python library used for working with arrays.It also has functions for working in domain of linear algebra, fourier transform, and matrices.



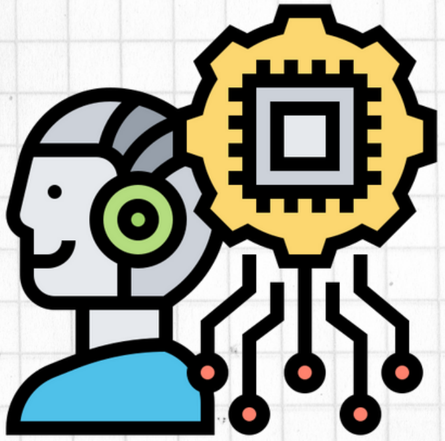
Step 2: Data Collection and Processing

- Loading the dataset to a pandas Dataframe.
- Checking number of rows and columns
- Analyze the distribution of Data.
- We found a null element in Dataset(Asteroid Magnitude).
- We filled up the null space by Imputation(by Median)!!
- Determined the Feature and Target Datatype to be achieved



Step 3: Splitting the features and target

- So here we found that our target data type is ’Hazardous’.
- Then we cleaned and Dataprocessed our target. We made two variables X and Y, having numerical dataset and target dataset respectively.



Step 4: Splitting the data into training and test data.

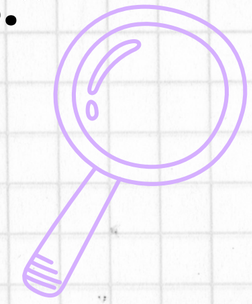
- Creating train and test data variables in X and Y.
- Verifying shape of data variables in X and Y.
- Verifying standard deviation of X and Y.

Step 5: Hyper Parameter Training

- Applying logistic regression to the dataset.
- Training the logistic regression model with training data.
- Verifying accuracy score on training data.
- Accuracy score on training data = 88.25516 %
- Verifying accuracy score on testing data.
- Accuracy score on testing data = 88.39142 %
- Therefore no condition of overfitting.

Step 6: Building a predictive system

- Changing the input data to a numpy array.
- It helps us to reshape the input data.
- Reshape the numpy array as we are predicting only one data point.
- So if prediction is false, our model prints “Collision doesn’t occurs.” , otherwise our model prints “ Collision occurs.”



Conclusion

Through the Model, based on the Orbital Data provided , we can estimate the collision probability of Asteroids with high accuracy upto 88-90%.

Contributed By:-

Priyanshu

Sahil Kumar

Varun Pratap Singh

PUNJAB ENGINNERING COLLEGE