

# **ITCS 6100 BIG DATA FOR COMPUTATIONAL ADVANTAGE**

## **PROJECT DELIVERABLE-1**

### **PROJECT GROUP 8**

#### **TEAM MEMBERS:**

- Aneela Gannarapu
- Preetham Garre
- Rachana Goli
- Sai Rithwik Reddy Bolla
- Susmitha Dalli

#### **COMMUNICATION PLAN:**

- All the Team members decided to discuss perspectives and insights through slack or Gmail chat and exchange their ideas apparently whenever it's required.
- All the Team members planned to gather or meet via Zoom or Google meet whenever required and will work on finishing the tasks accordingly.
- The project's repository can be accessed on GitHub using URL that's given below:  
<https://github.com/rachanagoli/BigDataGroup8>

#### **DATA SET SELECTION :**

We selected the data set from Kaggle active competitions to analyze. Below is the link to our dataset selected.

#### **Dataset 1 : Asteroid Dataset**

Link:

[Asteroid Dataset | Kaggle](#)

#### **BUSINESS PROBLEM OR OPPORTUNITY, DOMAIN KNOWLEDGE:**

The asteroid dataset offers valuable insights and opportunities for multiple domains, including astronomy and space exploration. It contains information on the materialistic characteristics and the orbital details of more than 17,000 asteroids. Moreover, detecting and monitoring potentially dangerous asteroids can help avert catastrophic incidents, such as an asteroid collision with Earth. The dataset also facilitates the identification of valuable resources, like water and metals, on asteroids, which can pave the way for future asteroid mining expeditions. Consequently, the asteroid dataset provides critical data for diverse scientific and commercial use cases in the field of space exploration.

## **RESEARCH OBJECTIVES AND QUESTIONS:**

During our research work, several questions have come up. The questions listed below are a few of the ones that deserve comprehensive research and should be considered from a wider perspective. We have opted to employ AWS technologies to implement the solution. The crucial task at hand is to determine the most suitable designs and algorithms to analyze the dataset effectively.

We listed out the following questions as an initial attempt towards the project:

- Can we predict the diameter of an asteroid in the solar system?
- Can we predict whether an asteroid is Potentially Hazardous or not ?
- Can we predict the probability of an asteroid colliding with the Earth based on its orbit and eccentricity?
- Can we predict the composition of an asteroid based on its spectral characteristics and other observable properties?