Beginner's Hypothesis

Data Science Group IITR

February-March 2023

1 Introduction.

It is the year 3472. Humanity is at its pinnacle, and its only ark **Prospectus** has reached its destination- **Zion**. As humanity prepares to inhabit their new home, they find a remnant of past extraterrestrial encounters of Zion. Among the remains is an artefact containing astronomical sensor data and their findings. The data warns them about the existence of black holes in their local cluster. This will prove pivotal in deciding the future of humanity henceforth.

2 Problem Statement.

The data given to you contains the information decoded by astronomers aboard. Each sample(11-tuple) corresponds to one of the black holes viz. **Zeta Draconis**, **Alpha Boreas**, **Beta Eioneous**, **Alpha Hemithea**, and **Gamma Saggita**. Your task is to build a **Classifier** from this data which when given a new sample would predict the black hole from which this sample is most likely to have been generated from. The given data contains **2000 samples**. Your performance will be evaluated on the basis of your classifier's accuracy on the hold-out data of **500 samples**. Any further exploratory analysis and findings from the data would be duly rewarded.

3 Resources

You can start building your data science skills during the brief vacation before the beginner's hypothesis starts. It is recommended that you go through the following link where step-by-step articles are provided for you to get started with Machine Learning. You can cover the **Introduction to Machine Learning**, **Data and its Preprocessing** and **Supervised Learning** sections. You can also reach out to your fellow DSG seniors for any additional resources.
Link: `https://www.geeksforgeeks.org/machine-learning/?ref=ghm`

4 How to submit?

This problem statement will be hosted on **Kaggle**. You will submit your predictions through a **.ipynb** file. Give a brief description of your code through **Markdown** cells in the notebook. You can view the leaderboard and make changes to your code all throughout the competition. The winners of this competition will qualify directly to the interview stage.
Link: **COMING SOON**

In case of any queries, you can contact:
Ved Umrajkar-8446887816
Abhay Kumar-8867392453