Name of faculty and U.ID:

Nahita Pathania 19372

1. Project Name:

Generic analysis of a dataset and development of a basic ML model based on user query in high-level language.

1. Problem Statement:

Developing insights for a given dataset play a crucial role in addressing the real-world data problems. The process of gaining insights is usually time-consuming as it involves writing code to plot various graphs for different data points, understanding the statistics like outliers, range, distribution type, interquartile range and many more. Our aim is to streamline this process to the highest possible extent. What if the developer/user can understand all of this by typing-in the query in a high-level language (English) and gain all those insights? Let’s not stop there, what if we could predict label values based on the features, apply ML, just by processing the given high-level language statement? And this applies for almost every kind of dataset.

1. Technologies to be used:
   1. Machine Learning
   2. Deep Learning
   3. Natural Language Processing
   4. Data Processing tools
   5. Feature extraction tools
   6. Data Visualization tools
   7. Image Processing
   8. React js
   9. Node js
   10. Flask
   11. Firebase
2. Project Outcomes:

A conventional data scientist follows seven steps to develop a solution in-terms of Machine learning: Preprocess data, understand data, gain insights, develop a basic model, fine-tune parameters and features, scale the model and integrate. With this project, first ***FOUR*** steps in the conventional process would be optimized to the highest extent. A data scientist can preprocess, understand data, gain insights and even develop a basic model with performance evaluation by just typing-in the queries in English language which streamlines the process of developing a basic model.

Team Details :

1. Prasanth G. 11601526
2. Anil Raj T. 11601794
3. Raja Rahul M. 11601793
4. Nikhil JSK 11601789