# How to Run

1. The dataset file should be present in the same directory as the jupyter notebook
2. The entire code is present in the form of a jupyter notebook and each cell of it can be executed by Shift+Enter
3. Has included the HTML output file which has the Code+generated Output in it.

# Approach

1. Extrapolatory data analysis to understand the distribution of data and number of classes and type of problem statement as **Multi Label Text Classification**
2. Encoding the prediction class according to Multi label classification using Multilabel binarizer
3. Using TF IDF Vectorizer for the dataset and splitting the data in training and test set
4. Define model metrics like
    a. Exact Match Ratio
    b. Hamming Loss
    c. Recall
    d. Precision
    e. F1 Score
5. Using SVM,SGD and Decision Tree Classifers for model training and prediction and find their scores
6. It is observed that Decision Tree outpeforms the other classifiers so it's hyperparameters is further tuned to get better performance
7. Finally classification report for tuned decision tree model is generated and classwise report for analyzing the further improvements and strong and weak classifying classes
8. Model is tested on some sample sentences as in the document