NATURAL LANGUAGE PROCESSING
Implementation Details

NAME : SURYA S   **NAME :** NILAVAN I
REG.NO : 20BCE1071   **REG.NO :** 20BCE1080
NAME : SUMEGH S GONUGADE   **NAME :** NITHISH KANNA G
REG.NO : 20BCE1100   **REG.NO :** 20BCE1304

SUBMITTED TO:

Dr. Ilakiyaselvan N

AIM:

To develop a club recruitment tool using NLP and Machine Learning.

STEPS INVOLVED IN THE IMPLEMENTATION OF THE PROJECT:

1. Extract dataset from Kaggle.
2. Perform exploratory data analysis – Clean the data (remove NA values) and Visualize the data.
3. Organize each post (for bag_of_words model) – Partition the posts into 16 personality traits.
4. Tokenize the posts (using nltk – NLP toolkit) using bag_of_words model.
5. Partition the cleansed data into train and test data.
6. Fit the machine learning models on the train data.
7. Predict using test data and check the accuracy for train and test data.
8. After using many ML models, choose the appropriate model and make it as the base model.
9. The model should return the personality trait using the social media post as input (using MBTI() and tellmemyMBTI() functions).
10. Classify the 16 personality traits into the department that suits (using the dept() function which takes trait as input).
11. Feed real life test data (Quora social media posts) and check the department that the candidate belongs to.