**CORONAVIRUS TWEET SENTIMENT ANALYSIS**

**Instructions:**

i) Please fill in all the required information.

ii) Avoid grammatical errors.

| **Team Member’s Name, Email and Contribution:** |
| --- |
| Sudhanshu Chouhan  Email - [Sudhanshurocks1998@gmail.com](mailto:Sudhanshurocks1998@gmail.com)  Contribution :  1.Data Wrangling   * Data Analysis * Null values of removal.   2.Country wise analysis of tweets  3.Removal of unnecessary things from original tweet.  4.. Analysis of cleaned data set.  5.. Employing various classification model and analysing the result.  6. Technical Documentation.  7.Video presentation.  Kapil Narayan Singh  Email – [kapilalmabetter@gmail.com](mailto:kapilalmabetter@gmail.com)  Contribution :  1.Data Wrangling   * Analysis of Data set. * Visualization of data through wordplay.   2.Cleaning the data .  3.Power point presentation  4.Employing the data set and analyzing the result.  5.Usage of confusion matrix.  6.Video presentation. |
| **Please paste the GitHub Repo link.** |
| Github Link:- https://github.com/itsmesudhanshu/Coronavirussentiment |
| **Please write a short summary of your Capstone project and its components. Describe the problem statement, your approaches and your conclusions. (200-400 words)** |
| **Problem Statement**    **Insights**   * In this Project we have given the data set of tweets done by the people during the coronavirus pandemic from various countries. * We have to see in this data what type of comment that has been done during the pandemic. * We get to know from our data that there are five types of sentiment that are extremely positive, extremely negative , positive, negative and neutral. * Then we have to perform data visualization. From there we get to know which countries' sentiments have been made. * After that we have to clean our tweets by removing unnecessary things like urls, hashtags and many more. * We have imported various types of libraries in which word play is the best library that we have used for data visualization.From this library we have displayed types of tweet sentiment and their frequency. * Then we have deployed various classification models and come to a conclusion that stochastic gradient descent performs the best in the case of binary classification and logistic classification in multi class classification. |