

**Group 02 CCNLP Mini project Report**

* **Group members**

| **Name** | **Roll no** | **PRN** |
| --- | --- | --- |
| Ajinkya Karnik | PD29 | 1032180678 |
| Samruddhi Agarwal | PD26 | 1031280507 |
| Shivam Pawar | PD35 | 1032181169 |
| Abhishek Patil | PE15 | 1032180111 |

* **Introduction**

This project predicts the emotion in a particular text that is given as input by the user.

Furthermore, an arbitrary dataset including texts from various conversations on the internet is picked and analysed with the help of a machine learning model to train and predict emotions in the input given by a new user.

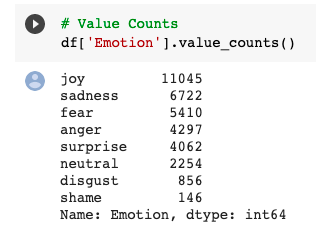
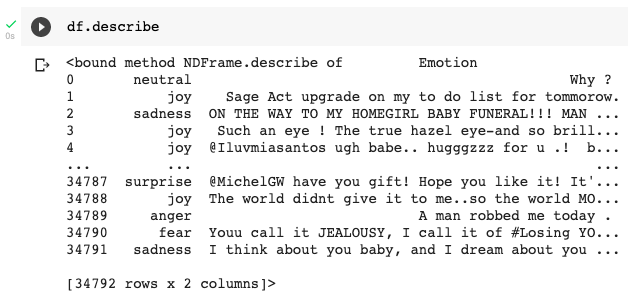
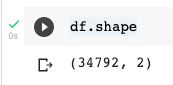
* **Problem definition**

Process, analyse and classify an input string of characters and predict the emotion associated with it.

* **Dataset description and data pre-processing**

Dataset format : Comma Separated Values (*.csv*)

Dataset - **emotion\_dataset\_raw.csv**

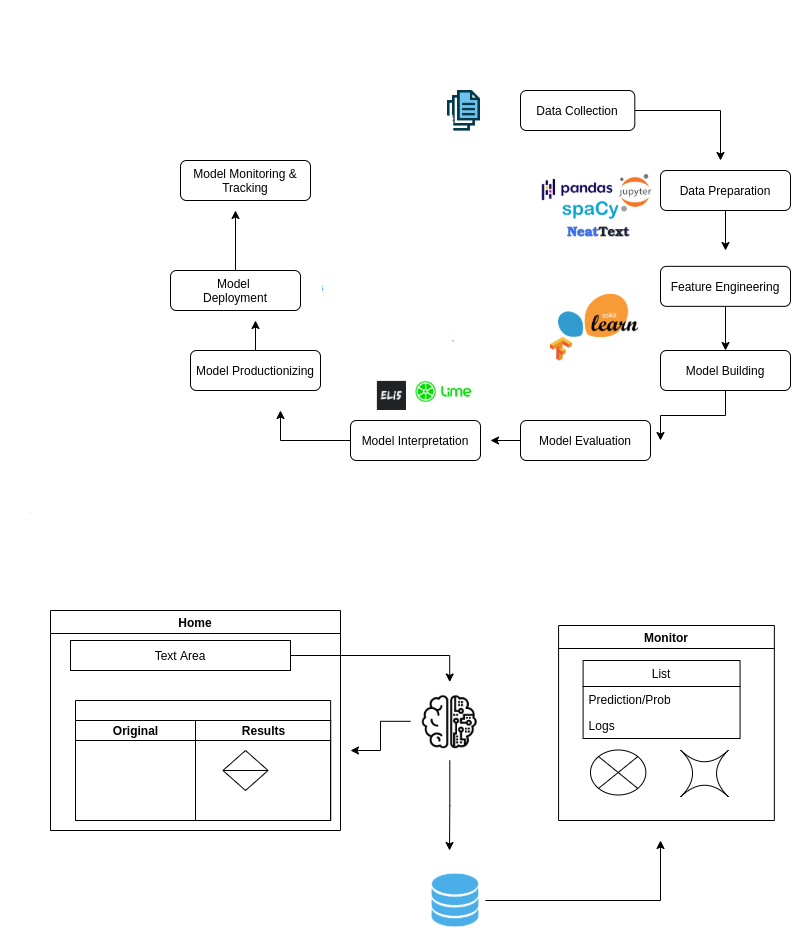
****

**Preprocessing involves**

* Removing User handles. *( as the data set contains online conversations )*

**

* Removing Stop words.
* **System architecture and system design**

****

* **Tasks performed**

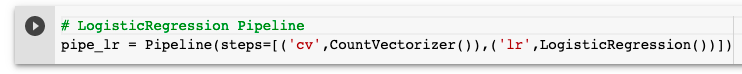
1. Making an interactive UI with streamlit
2. Processing and analysing of emotions in a given text
3. Showing the probability chart amongst different emotions
4. Showing the confidence level of the result
5. Keeping record of the input texts
6. Visualising the result

* **Methodologies**

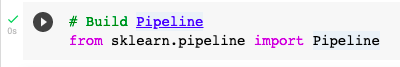
Algorithms used :

**Logistic Regression**

We chose this algorithm because this project majorly depended upon the probability of a certain emotion being felt and observed in a text snippet.

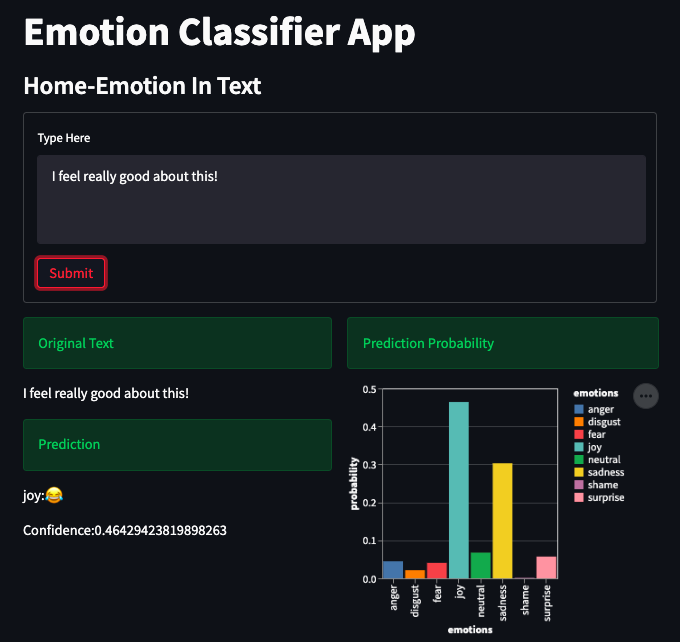


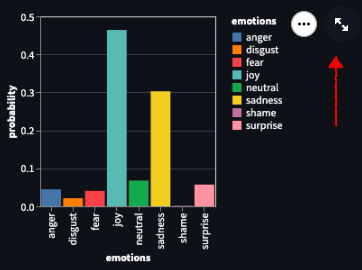
**Developed pipelines** with **sklearn.pipeline**

****

They helped us to develop a sequential flow of data from one estimato the next till it reaches the final prediction algorithm. It ensures there is no data leakage between train, test, and validation sets. The pipeline also makes a program more automated to be used as a functional code.

* **Output and visualizations**

****



Furthermore the visualization can be magnified for more detail :

