**Sentimental Analysis**:-

-By Yash Rajnikant Tawde

The aim of the project was to analyze the feelings and sentimental value of the reviews for a given dataset. The dataset consisted of well-labeled data, which was imported from Kaggle Datasets as a ‘tsv file’. We had to achieve the task of understanding and classifying the tone of the reviews, with a high level of accuracy. We analyzed and computed the results for classification using 4 different methods. Those methods were as given below:-

1. Support Vector Classifier Algorithm. (72%)
2. Pipeline Creation (Support Vector Classifier and Count Vectorization). (79.2%)
3. MultinomialNB Classifier Algorithm. (74.4%)
4. Pipeline Creation (MultinomialNB Classifier and Count Vectorization). (78.4%)

Each of the above methods present to us a different level of accuracy score. The highest level of accuracy was found out using the Pipeline Creation (Support Vector Classifier and Count Vectorization). The accuracy obtained was 79.2%. The imported dataset was converted to a Dataframe. On further analysis, we found out that there were no null values present in the dataset. Further, we used the ‘iloc[]’ command in order to convert the columns of the dataframe into numpy arrays. For SVC algorithm and the MultinomialNB algorithm it is very important to have vectorized the x variables. The advantage of using the pipeline creation method is that we don’t need to vectorize and transform the x variables manually, thus saving a lot of time. As the number of entries in the dataframe were 1000, we sought to use the train\_test\_split method in order to create separate datasets for training and testing purposes. Train\_test\_split offers an actual insight on the viability of the model that has been trained. Further, we serialized the pipeline model using the ‘JOBLIB’ library. This model was further integrated into a web app, to review the sentiments of reviews by customers. We could use this for multiple other applications, such as Tweet’s sentiment analysis. We could use web-scraping to read the tweet data, and then perform a classification, by using the Pipeline methods.