SENTIMENT ANALYSIS PROJECT REPORT

This is the report for the project “Sentiment Analysis using API”
The Framework used for the project is **FLASK.**
The language used for the project is **Python.**

Different Classification models were used for the project. Some of them are mentioned below with the metrics.

- Naïve Bayes Classifier with an accuracy of 59.34%
- Catboost Classifier with an accuracy of 89.4%
- RandomForest Classifier with an accuracy of 90.57%

Since RandomForest Classifier had the largest accuracy, the model was implemented with this algorithm.

Methods used to tune the model:

- First stopwords were downloaded using the NLTK library. After that, the texts were filtered and all the special characters including “@”,”#” etc. were removed.
- All the UpperCase characters were converted to lowercase.
- Each individual word from the texts were separated and were applied the process of “Stemming”, but since words were reduced without meaningless representation, “Lemmatization” was used instead because it provided better accuracy.
- After that , the process of CountVectorization was applied to convert the filtered sentences to arrays,to train the model.

Conclusion:

Hence, Using FLASK Framework, stopwords, Lemmatizer and RandomForest Classifier, the sentiment analysis project was implemented and the model managed to achieve a accuracy of 90.57% on an average basis.

Work Done by:

Vamsi Krishna S

Vamsiknight0@gmail.com

+91 7550106194