

MALIGNANT COMMENTS CLASSIFICATION

Submitted by:

Vivek Anand

**ACKNOWLEDGMENT**

One of the pleasant aspects of preparing a project report is the opportunity to thank those who have contributed to make this project possible.

We are extremely thankful to Ms. [**Khushboo Garg**](https://www.flipnwork.com/index.php/team_members/view/17) , whose active interest in the project & insight helped us to formulate, redefine implement our approach to the project.

We are also thankful to our institute & Other seen unseen hands which have given us direct & indirect help in completion of this project.

**INTRODUCTION**

* **Business Problem Framing**

This project is related to the scenario of proliferation of social media enables people to express their opinions widely online. However, at the same time, this has resulted in the emergence of conflict and hate, making online environments uninviting for users. Although researchers have found that hate is a problem across multiple platforms, there is a lack of models for online hate detection.

Online hate, described as abusive language, aggression, cyberbullying, hatefulness, and many others has been identified as a major threat on online social media platforms. Social media platforms are the most prominent grounds for such toxic behaviour.

There has been a remarkable increase in the cases of cyberbullying and trolls on various social media platforms. Many celebrities and influences are facing backlashes from people and must come across hateful and offensive comments. This can take a toll on anyone and affect them mentally leading to depression, mental illness, self-hatred, and suicidal thoughts.

Internet comments are bastions of hatred and vitriol. While online anonymity has provided a new outlet for aggression and hate speech, machine learning can be used to fight it. The problem we sought to solve was the tagging of internet comments that are aggressive towards other users. This means that insults to third parties such as celebrities will be tagged as unoffensive, but “u are an idiot” is clearly offensive.

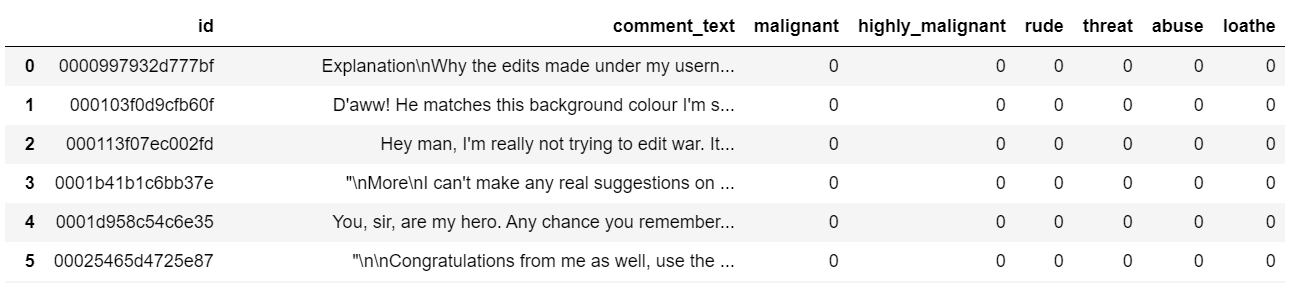
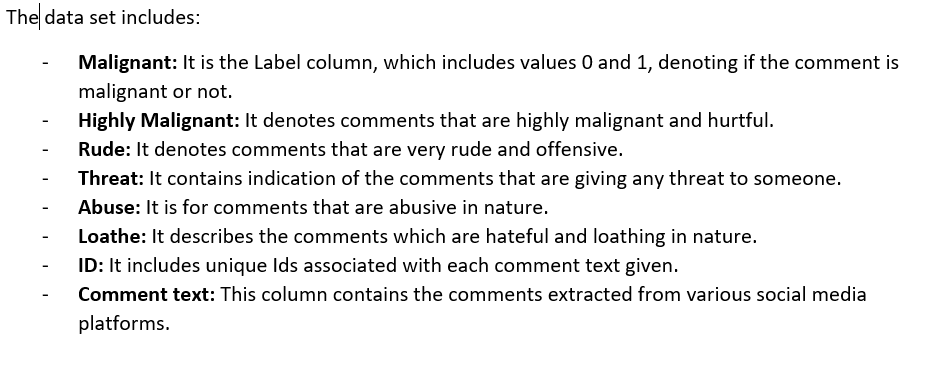
Our goal is to build a prototype of online hate and abuse comment classifier which can used to classify hate and offensive comments so that it can be controlled and restricted from spreading hatred and cyberbullying.

* **Conceptual Background of the Domain Problem**

Conceptual Background in this dataset, i.e.We will use Statistics and machine learning algorithms as well as natural language technic to process the dataset and built the model. Since the target value is classified, we will use the classification model algorithms in it. To visualize the data, we will use matplotlib and seaborn library, which based on python language.

* **Review of Literature**

Review of Literature is basically related to comprehensive summary of dataset as well as descriptions of input variables and output variable.

**** 

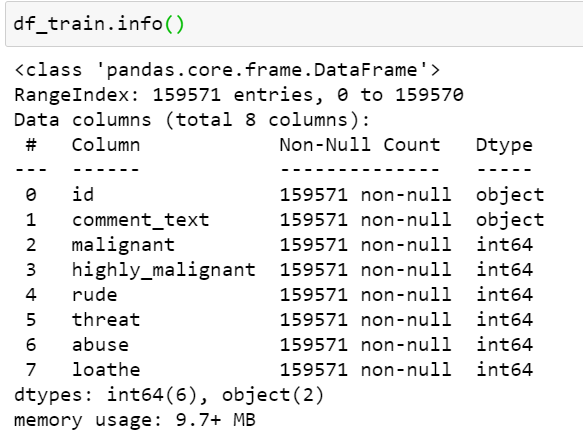
* **Motivation for the Problem Undertaken**

My motivation behind solving this classification problem is that it will help us to classify comments in social media.

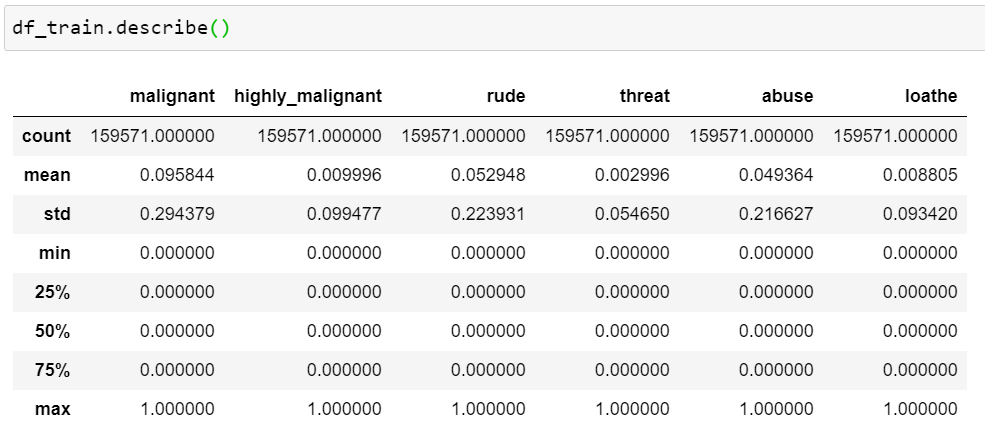
**Analytical Problem Framing**

* **Mathematical/ Analytical Modelling of the Problem**

In this section we understand and describe the mathematical, statistical and analytics modelling done during this project along with the proper justification. First of all for better understanding about dataset and given attributes information we use **Dataframe.info()** command, which tell me, The total number of attributes, what is name of attributes, datatype of attributes and how many Non-null values are present in dataset.



After understanding the dataset values, we take the statistical descriptions of dataset using **Dataframe.describe()** command in python, which tells the following statistical descriptions:



After knowing about statistical description, we move forward in the way of finding co-relation of variables, then move to data-cleaning, then move to visualized the data set and their relationship. After this we move to data pre-processing and modelling as well as testing the accuracy of model.

* **Data Sources and their formats**

Machine learning algorithms as well as Natural Language processing are almost always optimized for raw, detailed source data. Thus, the data environment must provision large quantities of raw data for discovery-oriented analytics practices such as data exploration, data mining, statistics, and machine learning.

Tabular data for machine learning is typically found is .csv files. Csv files are text-based files containing comma separated values (csv). Csv files are popular for ML as they are easy to view/debug and easy to read/write from programs (no compression/indexing).

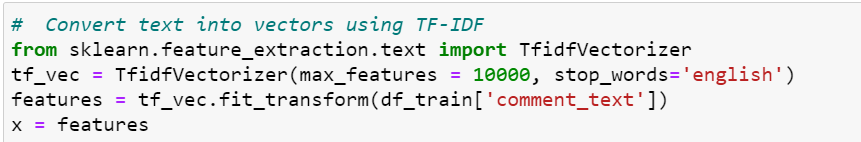
* **Data Pre-processing Done**
  1. Read dataset and make it in proper format.
  2. Encode labels
  3. Convert all cases to lower
  4. Remove punctuations
  5. Remove Stopwords
  6. Check stats of messages
  7. Convert all texts into vectors
  8. Import classifier
  9. Train and test
  10. Check the accuracy/confusion matrix.
* **Hardware and Software Requirements and Tools Used**

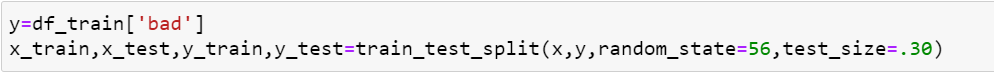
In this project dataset is too large for processing or modelling, that’s why we use good hardware configuration like as above 4GB RAM, above or equal core i3 processor and need good storage HDD. In way of software, we use any operating system which support python language for coding.

**Model/s Development and Evaluation**

* **Identification of possible problem-solving approaches (methods)**





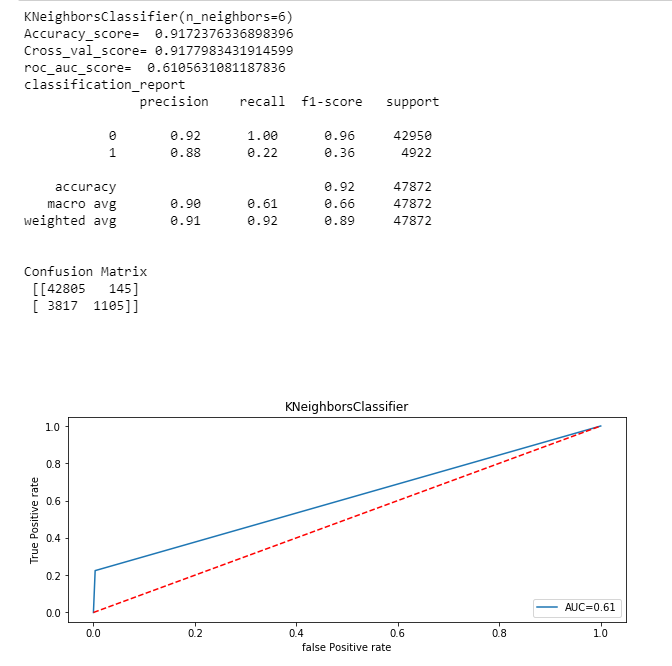


**Testing of Identified Approaches (Algorithms)**

* Convert all texts into vectors
* Import classifier
* ****Train and test
* **Run and Evaluate selected models**

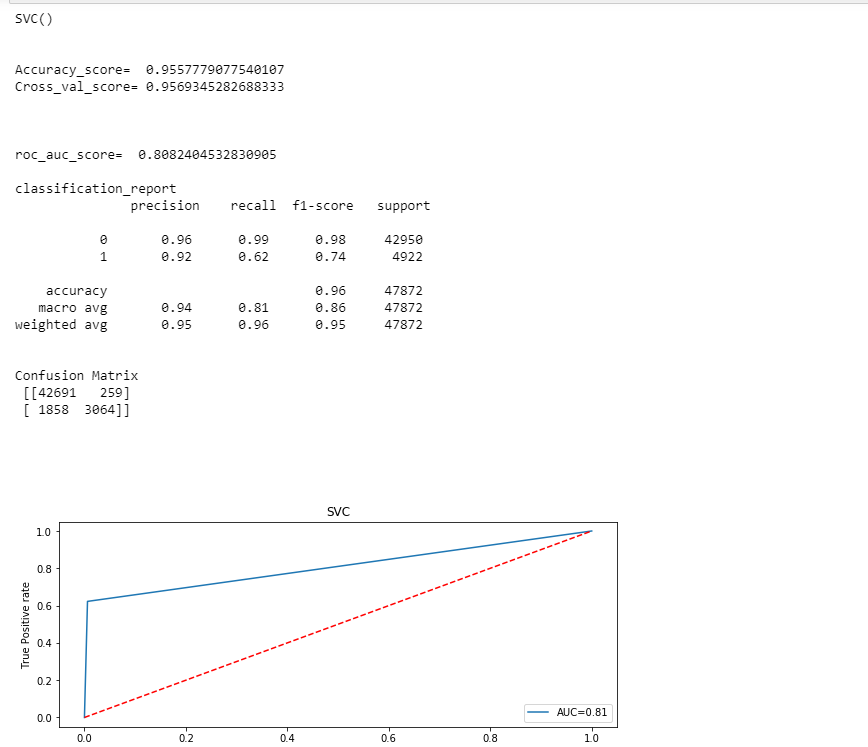
1. **kneighborsclassifier algorithm**





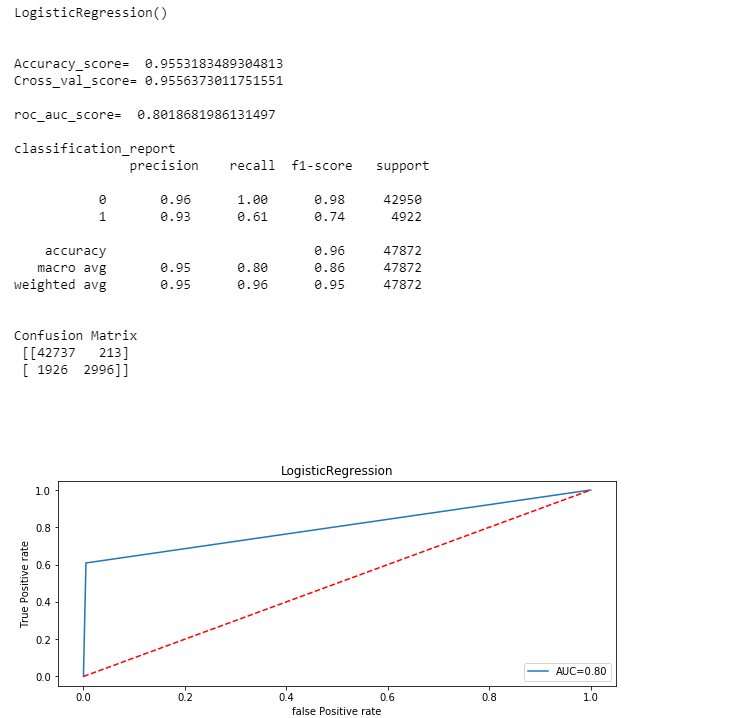
1. **svc algorithm**

****

****

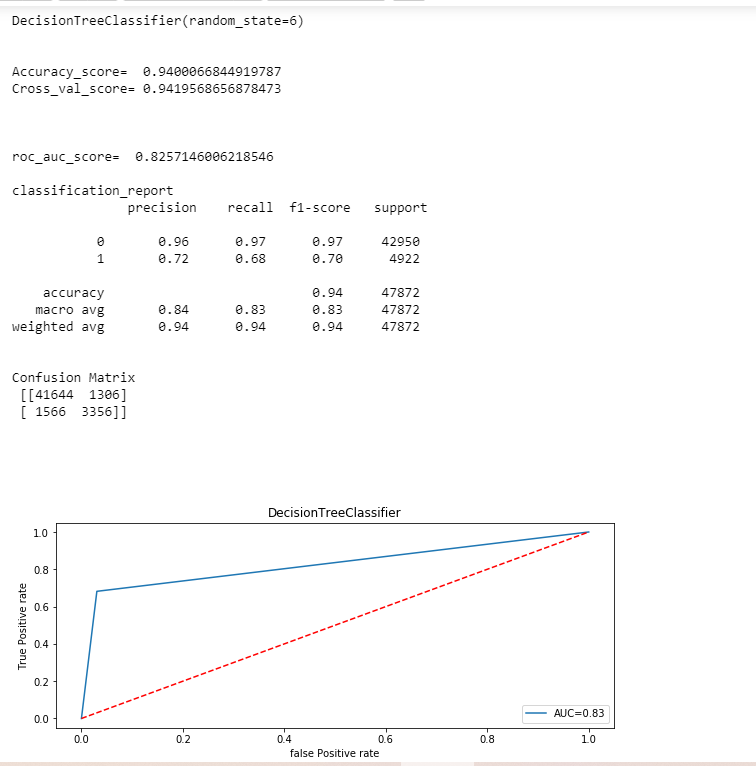
1. **Logistic regression**

****

****

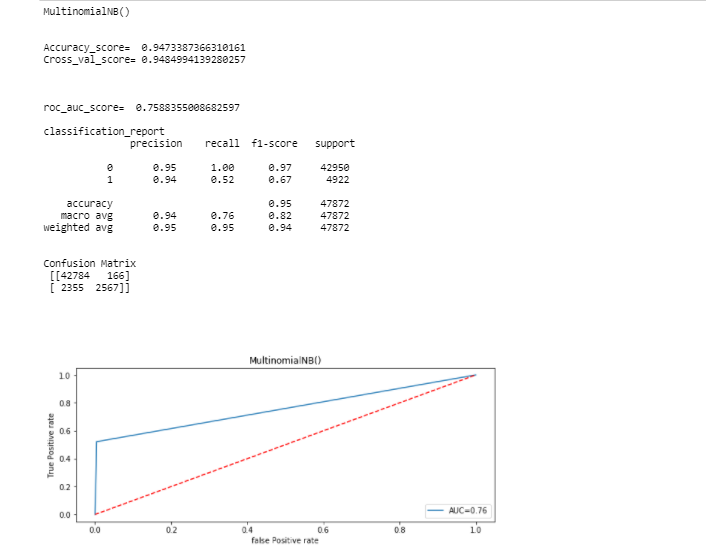
1. **Decision tree classifier**

****

****

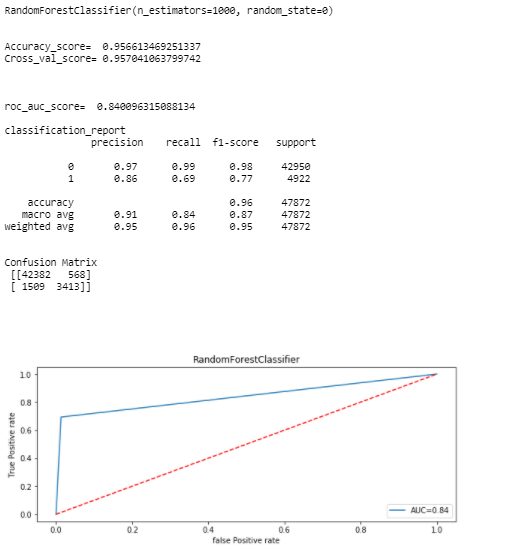
1. **MultinomialNB**

****

****

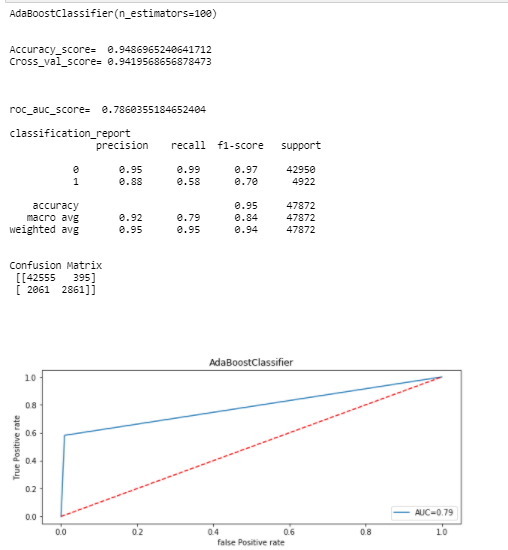
1. **RandomForestClassifier**

****

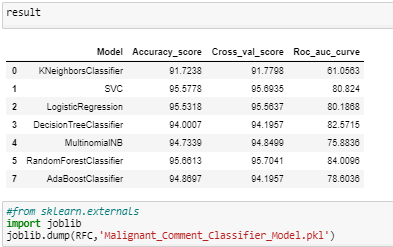
****

**7.AdaBoostClassifier**

****

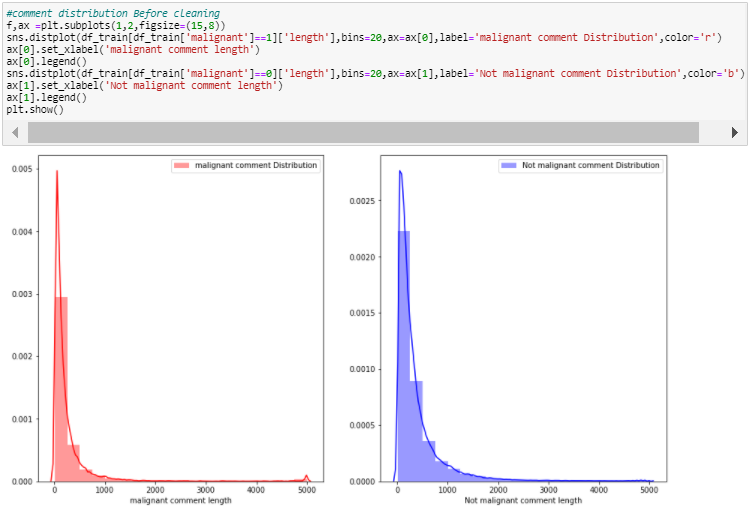
****

* **Final Result of models**

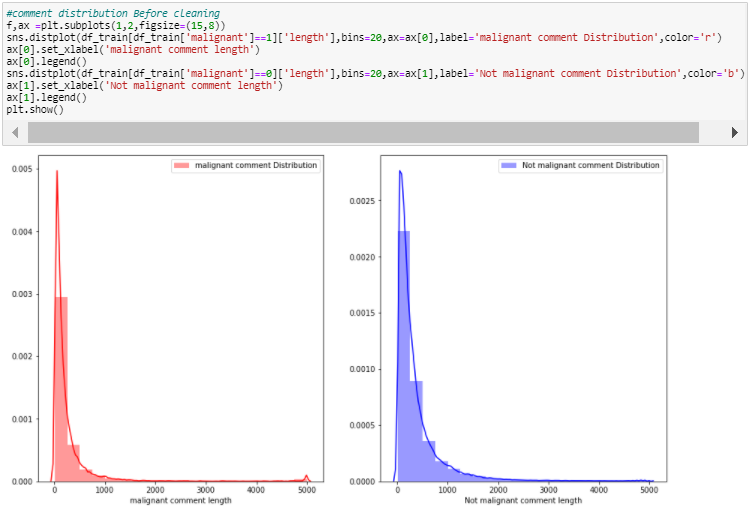
****

* **Visualizations**

**1.**

****

**2.**

****

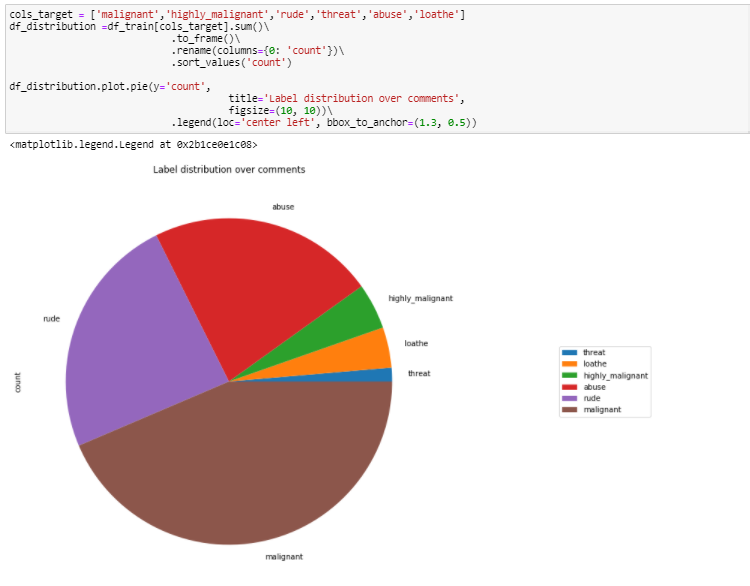
**3.**

****

**4.**

****

**4.**

****

**CONCLUSION**

After analysing data, visualization, and modelling, we conclude that using the Random Forest Classifier is suitable for modelling of comment's category prediction.