MALIGNANT COMMENTS CLASSIFICATION

MALIGNANT PROJECT:

MALIGNANT PROJECT PROJECT STATEMENT:

The 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 have to 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.

Data Analysis:

First of all we are importing the dictionaries that we are needed for this project and then importing the csv file into the jupiter note book and displaying maximum columns in it,then checking the shape of the data base given and Next checking the datatype given in the data base by using df.info command,there are float, integer and object type,further we are dropping the null values present if any in case! further we are checking the column we see that the entire column is same value so we have no use with that and In this project all the variables are in the numeric variables so need to to convert to

EDA Concluding Remarks:

we come to the next part of the model which is Exploratory Data Analysis ,In that we are visualizing the label using count plot and finding the number of comments per class,and finding the % of comments in various categories and created word cloud for better understanding to visualize the model by most used word in the comments will be represented big and less used words will be view small

Pre -Processing Pipeline:

After the EDA part,Not all categories have 5000 rows, So we should count them first and make them balanced,balancing the data is very much important for accurate results

Building Machine Learning Models:

In this project we used six models to se the prediction in it and they are log regression, KNN, BernoulliNB, MultinomialNB, SVM and Random Forest and their out put for each model is 0.8 for log ression and 0.17 for KNN and 0.77 for bernoulli and 0.87 for Multi nomial and 0.87 for SVM and 0.83 for random forest and all the output are plotted in the line graph on F1 Score of ML models (TF-IDF)

Concluding Remarks:

By comparing these algorithms we choose Random Forest Classifier as our final Model because the accuracy is high and less the difference between the accuracy, Finally we get an accuracy of 92% for this model,hence it is a good model.And it will be very helpful to classify and predict it is a toxic comment are not by the client, thank you.