

MALIGNANT COMMENT CLASSIFIER

Submitted by:

SAI CHANDER REDDY I

**INTRODUCTION**

* Business Problem Framing
* 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.
* 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 thought
* Conceptual Background of the Domain Problem
* 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.

.

* Motivation for the Problem Undertaken
* 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.

**Analytical Problem Framing**

* Mathematical/ Analytical Modeling of the Problem

The Data consisted of more normal comments and less comments which are malignant ,rude,threat,loathe,abuse. So we have to find the number of data which is malignant and normal data and we have made a label column which showed the sample has malignant comments or not with 0 and 1 logistic representation.

* Data Preprocessing Done

The Comments\_text is the column which has been preprocessed using nlp techniques to remove all the extra spaces, punctuations, stopwords,html tags and so much noisy data and it has been preprocessed.

* Data Inputs- Logic- Output Relationships

The comments\_text is the input data we provided to the model as the comments are preprocessed and made into vector representation using TFIDF vectorizer. The data is presented in vector form and output is predicted using 0 and 1 logic , where 0 represents normal comment and 1 represents malignant comment

* Hardware and Software Requirements and Tools Used

Hardware Requirements:

Ram : 4GigaBytes

Rom:500GigaBytes

Processor: Intel core i3 or more

Software Requirements:

Windows 10

Python Programming Language is used .

The packages used are:

Numpy,Pandas,Seaborn,Scipy,Scikitlearn,MatplotLib,imblearn,NLTK

**Model/s Development and Evaluation**

* Identification of possible problem-solving approaches (methods)

The Problem is Classification Type of Problem. We have used Logistic regression to solve the problem. The Testing metrics we have used are accuracy\_score,classification report,confusion matrix, auc\_roc-score..

* Training of Model:

We have used models like Decision trees and Logistic Regression

**CONCLUSION**

This project Malignant comment classifier is modelled to perform predictions to malignant texts. This model predicts 0 as the text is normal and 1 if the text is malignant and abuse.