*HATE SPEECH DETECTION- BY NEHA POKHARIYA*

**Acknowledgement**

First of all, I would like to thank our institution “Graphic Era deemed to be university” for giving us an opportunity to work on such great projects. I am also thankful to our mentor Garima ma’am for guiding me through the project.

**Motivation**

I chose this project because I feel like this project will help me understand more about languages like python, and its different user friendly libraries. The project also helps me understand a little more about how to work with a dataset and use it to analyse and classify messages. As I really like to work with data, this project was an amazing chance to learn more about data and how they work, here, specifically for hate speech classifiers.

This project also helps me to understand and deploy a website through stream lit and heroku.

**What is hate speech?**

Any speech that disparages a group of people because of their race, religion, nationality, national origin, sexual orientation, or gender identity is considered hate speech. Hate speech is frequently used to promote bigotry and hatred. Additionally, it can be applied to threaten and intimidate others. People may experience loneliness, anxiety, and fear as a result. Hate crimes may also result from it. Additionally, hate speech can harm ties between various racial and ethnic groups. It is crucial to identify hate speech since doing so can assist in stopping these negative impacts.

🟧 Numpy:

NumPy is a Python library used for working with arrays. It also has functions for working in the domain of linear algebra and matrices. It is an open-source project, and you can use it freely. NumPy stands for Numerical Python.

🟧 Pandas:

Pandas is a software library written for the Python programming language for data manipulation and analysis. In particular, it offers data structures and operations for manipulating numerical tables and time series.

🟧 Countvectorizer

A fantastic utility offered by the Python scikit-learn module is CountVectorizer. It is used to convert a given text into a vector based on the number of times (count) that each word appears across the full text.

🟧 Train test split

Our data is separated into train and test sets using the train test split() technique. We must first separate our data into features (X) and labels (y). Divided into X train, X test, Y train, and Y test, the dataframe is. The model is trained and fitted using the X train and y train sets.

🟧 Decision tree

A supervised machine learning algorithm called Decision Tree uses a set of principles to make judgments, much like how people do.

**Content:**

1. Data cleaning

2. Training the Hate Speech Detection Model

3. Streamlit and pycharm-website and deployment

**Data Cleaning**

Data cleaning is the process of preparing incorrectly formated data for analysis by deleting or modifying the incorrectly formatted data which is generally not necessary or useful for data analysis, as it can hinder the process or provide inaccurate results. I will perform the process of data cleaning by using the re and nltk library in Python.

• Lower case

• Tokenization

• Removing stop words and punctuation

• Removing special characters

• Stemming

## Training the Hate Speech Detection Model

Now, before training the model, let’s split the data into a training set and a test set. After that we’ll be using decision tree to predict if the given string is or input is offensive or not.

We can also test and train using f1 score.

## Streamlit and pycharm-website and deployment

* Preprocess(transform)
* Vectorize
* Predict(using multinominal)
* Display(streamlit)

We will be using streamlit and Heroku to deploy a host website to check for any input for offensiveness.