Project Delivery 3

Group Name: The Survivor
Names: Abdelrahman Atef Abouzid
Email: Hemanysat@gmail.com
Country: Egypt
College: Ain Shams University
Specialization: NLP

Problem Statement

The term hate speech is understood as any type of verbal, written or behavioural communication that attacks or uses derogatory or discriminatory language against a person or group based on what they are, in other words, based on their religion, ethnicity, nationality, race, colour, ancestry, sex or another identity factor. In this problem, We will take you through a hate speech detection model with Machine Learning and Python.

Hate Speech Detection is generally a task of sentiment classification. So for training, a model that can classify hate speech from a certain piece of text can be achieved by training it on a data that is generally used to classify sentiments. So for the task of hate speech detection model, We will use the Twitter tweets to identify tweets containing Hate speech.

Data Understanding

The train set consists of 31962 rows with 3 columns each. One for index, one for the tweet and the final one is the label. The main feature is the tweet which contains the contents of extracted tweets. The Dataset does not contain any missing values and not outliers hence it is in natural language. Therefore, no problems in the data to overcome.

Data Cleaning and Transformation

In this section, we applied various data cleaning and transformation tasks on a Twitter NLP dataset. Firstly, we downloaded the necessary packages,

including NLTK and spaCy, and defined functions for removing punctuation, checking if a string contains numbers, stemming, and tokenizing the text data. Next, we define a custom class called DataFrameDataset that utilizes the torchtext fields and dataset classes to prepare the dataset for the PyTorch model. The class converts a Pandas dataframe into a torchtext dataset and splits it into training and validation datasets. It then builds the vocabulary for the text and label fields and prints out information on the size of the vocabulary and the most commonly used words in the text data.

Github Repository Link

hemany1/Hate-speech-classification (github.com)