Group Name: AO
Name: Alaa Eddine Osta
Email: ostaalaaeddine@gmail.com
Country: Lebanon
College: Üsküdar University
Specialization: NLP
Project Deadline: 11/25/2021

Problem Description

The term hate speech is any form of expression (verbal, written, or behavioral communication) that attacks or uses derogatory or discriminatory language against a person or group. This hate could be towards religion, ethnicity, nationality, race, color, ancestry, sex, or other identity factors. In this problem, we need to build a machine learning model that predicts which tweets have hate speech in them.

Hate Speech Detection is generally a task of sentiment classification. To classify hate speech from a piece of text, we need to train the model on data used to classify sentiments. So for the task of the hate speech detection model, we will use Twitter tweets to identify tweets containing hate speech.

Data Cleansing and Transformation

```
# Remove symbols from tweets and lower case all characters
def process_tweet(tweet):
    return " ".join(re.sub("(@[A-Za-z0-9]+)|([^0-9A-Za-z \t])", " ",tweet.lower()).split())
```

Figure 1: The regex function used to remove the noise from data

	label	tweet	processed_tweets
12227	0	@user â□□my mom says my smile is captivatingâ...	my mom says my smile is captivating i says hap...
14709	0	in 3 days i will be meeting my sis-n-law, coun...	in 3 days i will be meeting my sis n law coune...
19319	0	hating the conservative homophobes using this ...	hating the conservative homophobes using this ...
4308	0	awee if this doesn't #scream #friday #acewe...	awee if this doesn t scream friday acewellstuc...
24055	0	fathersday #fatherÃ¢Â□Â□s #day #god! #Ã□Â« #...	fathersday father s day god tony a smith buy t...

Figure 2: Result after shuffling and applying the function shown in Figure 4 on the data

- The featurization techniques used are TfidfVectorizer and Universal-Sentence-Encoder.