Amaan Hussain
University of Illinois Chicago
April 2022
Faculty Advisor: Professor Evan McCarty
GitHub Link: https://github.com/AmaanHussain9553/Research_Sentimental_Analysis

Predicting Hate Speech using Recurrent Neural Networks (LSTM)

CS 398 Undergraduate Research Final Report

About the Dataset

The dataset consists of over 30,000 tweets from Twitter and each tweet consists of classification labels of '0' or '1'. '0' denotes that the tweet does not classify as a hatred tweet and '1' denotes that the tweet is classified as a hatred tweet. Hatred is determined by the fact if the tweet is racist or sexist. This data is taken from Kaggle (link) and is provided by Analytics Vidhya.

id	label	tweet
1	0	@user when a father is dysfunctional and is so selfish he drags his kids into his dysfunction. #run
2	0	@user @user thanks for #lyft credit i can't use cause they don't offer wheelchair vans in pdx. #disapointed #getthanked
3	0	bihday your majesty
4	0	#model i love u take with u all the time in urÃ°ÂŸÂ“Â±!!! Ã°ÂŸÂ˜Â™Ã°ÂŸÂ˜ÂŽÃ°ÂŸÂ‘Â„Ã°ÂŸÂ‘Â…Ã°ÂŸÂ‘Â¦Ã°ÂŸÂ‘Â¦Ã°ÂŸÂ‘Â¦
5	0	factsguide: society now #motivation
6	0	[2/2] huge fan fare and big talking before they leave. chaos and pay disputes when they get there. #allshowandnogo
7	0	@user camping tomorrow @user @user @user @user @user @user @user dannyÃCÂ€Â¦
8	0	the next school year is the year for exams.Ã°ÂŸÂ˜Â¯ can't think about that Ã°ÂŸÂ˜Â- #school #exams #hate #imagine #actorslife #revolutionschool #girl
9	0	we won!!! love the land!!! #allin #cavs #champions #cleveland #clevelandcavaliers ÃcÂ€Â¦
10	0	@user @user welcome here ! i'm it's so #gr8 !
11	0	ÃcÂ†Â #ireland consumer price index (mom) climbed from previous 0.2% to 0.5% in may #blog #silver #gold #forex
12	0	we are so selfish. #orlando #standwithorlando #pulseshooting #orlandoshooting #biggerproblems #selfish #heabreaking #values #love #
13	0	i get to see my daddy today!! #80days #gettingfed
14	1	@user #cnn calls #michigan middle school 'build the wall' chant '' #tcot
15	1	no comment! in #australia #opkillingbay #seashepherd #helpcovedolphins #thecove #helpcovedolphins
16	0	ouch...junior is angryÃ°ÂŸÂ˜Â#got7 #junior #yugyoem #omg
17	0	i am thankful for having a paner. #thankful #positive
18	1	retweet if you agree!
19	0	its #friday! Ã°ÂŸÂ˜Â€ smiles all around via ig user: @user #cookies make people
20	0	as we all know, essential oils are not made of chemicals.
21	0	#euro2016 people blaming ha for conceded goal was it fat rooney who gave away free kick knowing bale can hit them from there.
22	0	sad little dude.. #badday #coneofshame #cats #pissed #funny #laughs
23	0	product of the day: happy man #wine tool who's it's the #weekend? time to open up & drink up!
24	1	@user @user lumpy says i am a . prove it lumpy.

1.1 CSV Sample of what the data looks like

Preprocessing Data

The data is read in from a CSV file and stored in a Pandas data frame. The first task involves removing all the punctuation marks and common stop words in English so that the words that are more essential in determining racism/sexism are left. Then the data is filtered by words that occur in the entire dataset 30 or more times and each list of tweets is one-hot encoded to be passed into the LSTM model that we will be developing. Finally, we pad each list of string to be as big as the largest tweet (string) and split the data into train and test (80/20 train/test).

About the Model

The machine learning model used to train and eventually test our preprocessed data is the LSTM (long-short term memory) model which is a recurrent neural network. We create a sequential model where the first layer is meant to establish the word embeddings based on our vocabulary size (unique words occurring 30 or more times – 1304 words) which is our input dimension, we decide an output dimension which is generally decided as: *output_dimension = input_dimension ** 0.25*, and we set the *input_length* which is the padded size of each tweet. Then our second layer is the LSTM layer where we created a 1024 hidden node model which returns the sequences each time and is flattened and processes one output for us after passing it through a sigmoid activation function. To compile the model, I am using *'adam'* as my optimizer since it is meant for training deep learning models and *'binary cross entropy'* as my loss function since it is a binary classification problem.

LSTM Model Tuning Hyperparameters

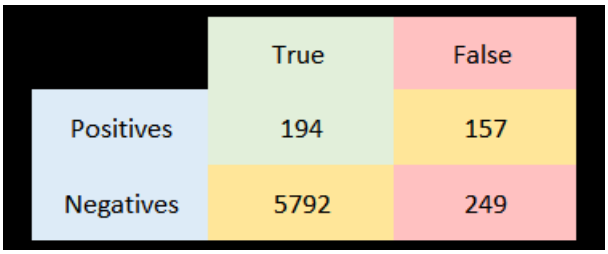
Embedding Size	Batch Size	Accuracy
8	8	90.76%
	16	95.58%
	32	94.23%
	64	95.39%
	128	95.06%
32	8	96.63%
	16	96.40%
	32	96.16%
	64	96.13%
	128	95.64%
64	8	97.29%
	16	96.40%
	32	96.48%
	64	96.35%
	128	95.94%
128	8	97.66%
	16	97.19%
	32	96.88%
	64	96.72%
	128	96.41%
256	8	98.44%
	16	97.85%
	32	97.25%
	64	96.94%
	128	96.00%
512	8	98.69%
	16	98.47%
	32	98.01%
	64	97.63%
	128	97.05%

1.2 Data Recorded of various model parameters tested to achieve maximum accuracy

Model Results

To check the accuracy of my model against the test data which was untouched until after the completion of my model, I used a confusion matrix which would allow me to see true positives, true negatives, false positives, and false negatives.

We can see from the diagram below that the model was significantly better at correctly predicting the tweets that were not racist/sexist than accurately predicting the tweets that were.



	True	False
Positives	194	157
Negatives	5792	249

1.3 Confusion Matrix results from test data.

Conclusion

To summarize the research experience, I got the chance to use real-life data from social media and apply key concepts and knowledge towards it. Under the guidance of Professor Evan McCarty, I was able to learn new techniques when dealing with machine and deep learning models while being exposed to software used by companies at the highest level. It was very intuitive to be able to go from a simple CSV file to being able to preprocess data and pass it through a complex neural network that can help predict hatred in tweets sent out by people. I was able to apply the knowledge learnt in many of my classes while having to take extra steps in researching and reaching my end goal.

References:

Brownlee, J. (2021, February 1). *How to use word embedding layers for deep learning with keras*. Machine Learning Mastery. Retrieved April 23, 2022, from https://machinelearningmastery.com/use-word-embedding-layers-deep-learning-keras/

Brownlee, J. (2020, September 2). *Sequence Classification with LSTM recurrent neural networks in python with keras*. Machine Learning Mastery. Retrieved April 23, 2022, from https://machinelearningmastery.com/sequence-classification-lstm-recurrent-neural-networks-python-keras/

Brownlee, J. (2021, October 12). *Your first deep learning project in python with keras step-by-step*. Machine Learning Mastery. Retrieved April 23, 2022, from https://machinelearningmastery.com/tutorial-first-neural-network-python-keras/

Dataset:

Toosi, A. (2019, January 6). *Twitter sentiment analysis*. Kaggle. Retrieved January 20, 2022, from https://www.kaggle.com/datasets/arkhoshghalb/twitter-sentiment-analysis-hatred-speech