CSE440: NATURAL LANGUAGE PROCESSING II

Project Title: Explainable Detection of Online Sexism

Team: 15

Section: 02

Spring 2024

ID	NAME
21301470	Syeda Ifroza Ahmed
21201124	Mim Raihan
21201209	Radia Tasnim Arony
21201247	Subrina Nowrin Shupti

INTRODUCTION

The objective of our study is to distinguish between sexist and non-sexist comments using a shallow recurrent neural network. First, we import the dataset and employ pre-trained GloVe vectors for review representation, with a 100-dimensional vector representing each word. For the categorization in our project, we employ the Python programming language and popular libraries such as Numpy, NLTK, Scikit-learn, Keras, and Pandas. Additionally, the training set has undergone 20 model runs and the dataset has been preprocessed. Moreover, a shallow model with a bidirectional LSTM layer and a unidirectional embedding layer is provided. Finally, the performance of several models is evaluated along with an analysis of the LSTM layers' efficacy.

DATASET DESCRIPTION

In our data set we have rewire_id, text, label_sexist, label_category, label_vector features. Among these, we don't need the id column to do our classifications.Here we have 14001 rows and 5 columns.We have so many null values in the dataset as well.To clean the dataset,we have performed data pre-processing.

DATA PRE-PROCESSING

From the dataset, unnecessary columns were dropped, such as- rewire_id, label_vector. Further pre-processing was done by converting given sentences to lowercase, removing multiple spaces, single character, html tags, punctuation, numbers and stopwords. Along with those, tokenization and imputation were also carried out on the given data.The dataset was divided into 80% and 20% for training and testing respectively.

DATA VISUALIZATION

The given dataset was visualized using bar charts and word clouds, before training and testing the models. After the implementation of models, individual graphs of model accuracy and model loss, confusion matrix, and two final bar charts of accuracy comparison of models, for both binary and multiclass classification were shown.

MODEL DETAILS

For the binary classification task, four models have been trained and then tested- Shallow model, Unidirectional LSTM, bidirectional LSTM, GRU. An overview of some hyperparameters used were- activation functions ‘ReLU’ and ‘sigmoid’, along with an Adam optimizer, epoch of 20, and a batch size of 25. Moreover, empirical regularization techniques like dropout and early stopping were utilized with values set to 0.2 and patience= 3, respectively. Lastly, the loss function, binary cross entropy loss was applied as well.

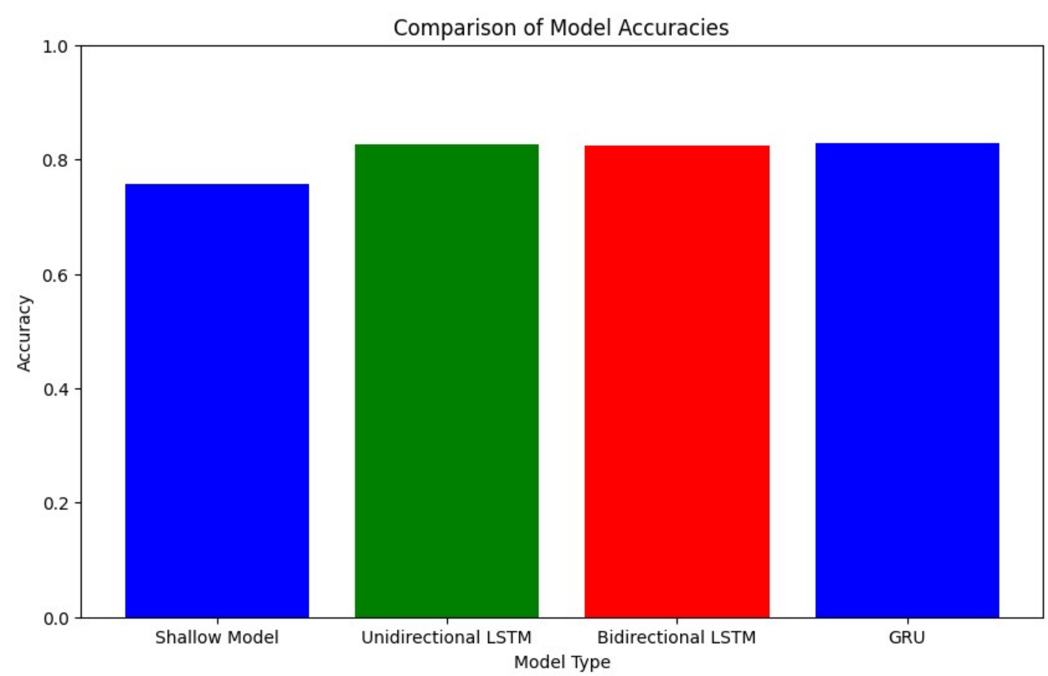
For the multiclass classification task, three models have been trained and then tested- Dense model, Unidirectional LSTM, bidirectional LSTM. An overview of some hyperparameters used were- activation functions ‘ReLU’ and ‘softmax’, along with an Adam optimizer, epoch of 20, and vocabulary size of 500. Also, empirical regularization techniques like dropout and early stopping were utilized with values set to 0.2 and patience= 3, respectively. Finally, the loss function, sparse categorical cross entropy was applied as well.

RESULT COMPARISON

Task A

Model	**Accuracy**
Dense Model	75.70%
Unidirectional LSTM	82.18%
Bidirectional LSTM	82.67%
GRU	85.32%

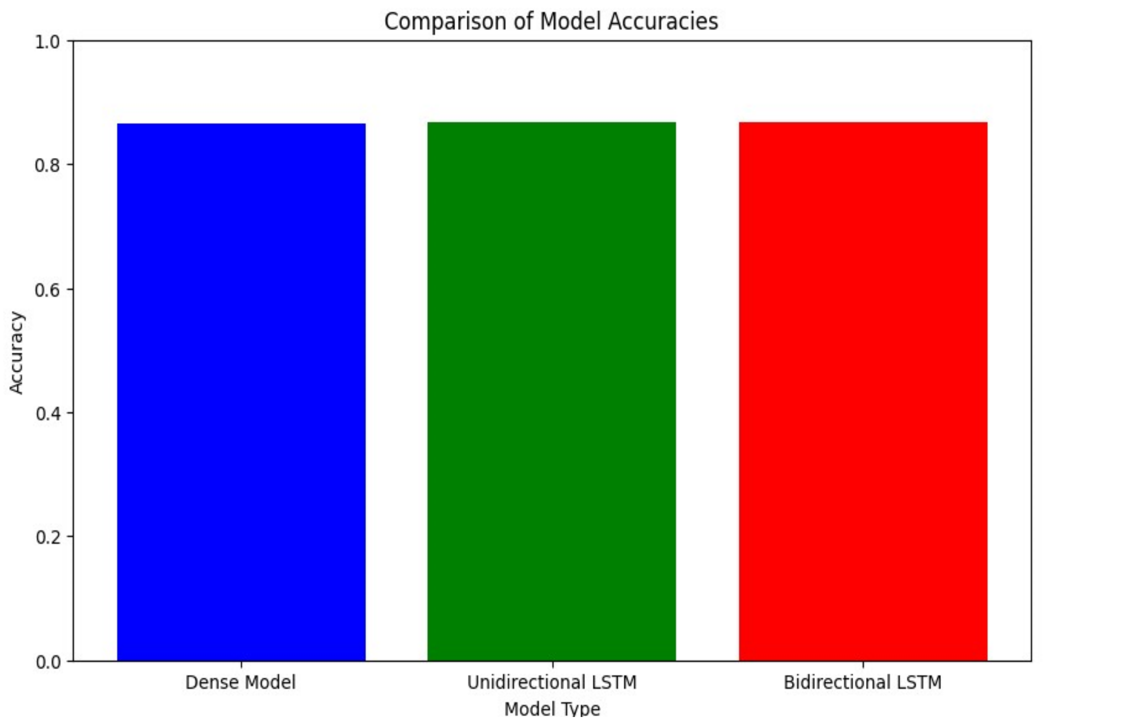
The final bar chart comparison of accuracies is provided below-



Task B

Model	**Accuracy**
Dense Model	86.57%
Unidirectional LSTM	86.71%
Bidirectional LSTM	86.68%

The final bar chart comparison of accuracies is provided below-



HOW WE IMPROVED OUR ACCURACY:

To improve the model, the combined use of regularisers, optimizers, and GRUs should be considered. Because GRU has fewer parameters and a simpler architecture, GRUs train faster. Furthermore, when working with small datasets, a simpler model like an LSTM may be more prone to overfitting, whereas the simpler structure of GRUs may promote greater generalization. Furthermore, eliminating certain neurons from the dataset simplifies the neural network and tackles overfitting which is known as dropout.