BRAC University

CSE440: Natural Language Processing (NLP II)

Final Project

[Explanatory Detection of Online Sexism (EDOS Classification)]

Submitted To:

Dr. Farig Yousuf Sadeque

Associate Professor, BRAC University

Submitted By:

Group 1 - CSE440 Section 1

Group Member's Information:

Name: Ashok Lamichhane ID: 21201785 Department: CSE

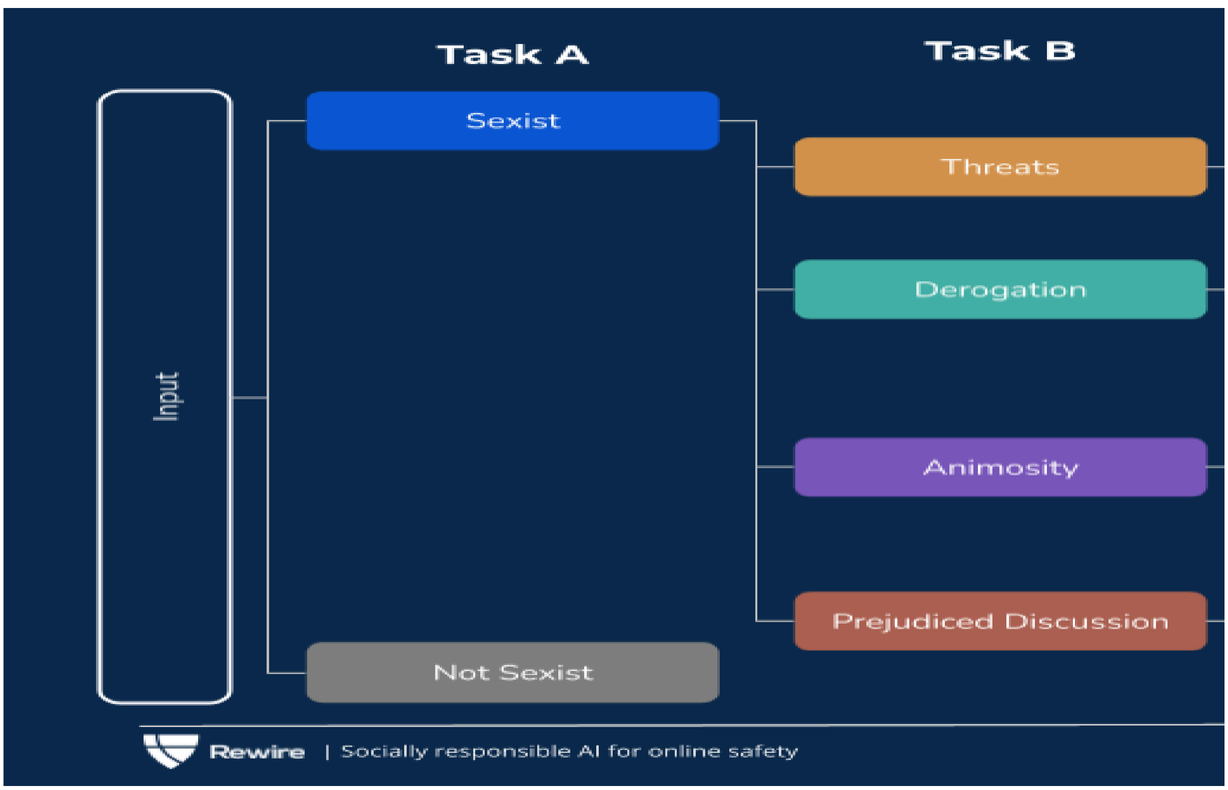
Name: Mukund Prasad Singh ID: 20201202 Department: CSE

Explanatory Detection of Online Sexism (EDOS Classification)

Introduction

Sexism is a form of discrimination based on a person's gender or sex. Sexism can lead to violence and create an oppressive environment that prevents most women from fully participating in public life. Creating a model that can identify sexist texts could lead to several positive outcomes for society. Such a model could be used to prevent and counter sexist language in different situations, such as in public and private communications, media, and social networks. This could help promote gender equality and create a more inclusive and respectful environment.

Using the given dataset, we are going to perform a text classification task using deep learning techniques. For task a and task b proposed by SemEval(2023), we used the bi-LSTM model.

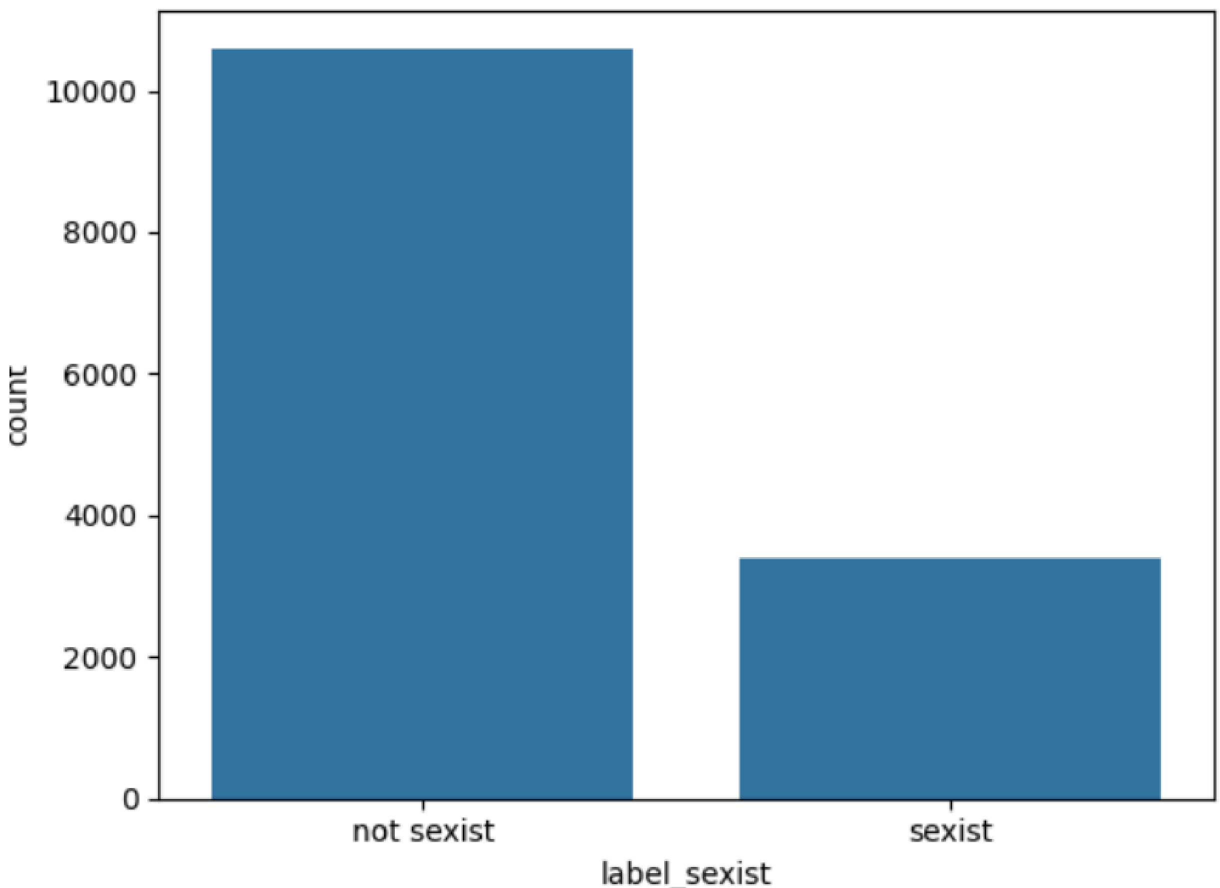


Dataset

From the given dataset, we just used train_all_tasks.csv, and removed all other data files from there. Then we split the data file in train dev and test. There are two different tasks for each classification:

A. Binary Sexism Detection: In this task, systems are required to predict whether a post is sexist or not, by performing a two-class (or binary) classification. The results are divided into two classes, where the Sexiest is more than 11,000 and the not sexiest is around 3,000 instances.

B. Category of Sexism: This task involves classifications that are identified as sexiest into one of four categories such as Threats, Derogation, Animosity, or Prejudiced discussion, by performing a four class classification.



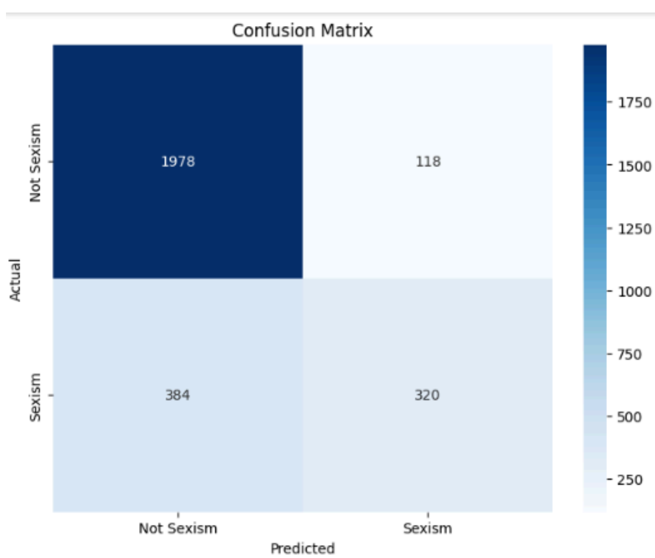
Model Description (Bi-LSTM & GloVe):

Here we use a model bi-LSTM with a2_glove.6B Glove for embedding. To correctly apply the Bi-LSTM model, we have imported various libraries for data manipulation, visualization, text-processing and deep learning(tensorflow.keras).It defines a Bidirectional LSTM model using Keras. This model architecture is commonly used for sequential data like text. The bidirectional aspect allows the model to learn from both past and future context of a word in a sequence. The LSTM layers help in capturing long-range dependencies in the data, which is particularly useful in text processing tasks. In data preprocessing, we perform various steps like tokenization, padding, and splitting the data into training and testing sets. We had 1 embedding layer, two Bidirectional LSTM layers and 1 output layer. All the hidden units had ’Relu’ and the output layer had a ’Softmax’ activation function. We downloaded the GloVe from the verified online sources and uploaded to google drive, made public and accessed from it.

Results & Analysis (F1 Score & Accuracy)

After constructing the models we ran the dataset over the model. Of the 80% training data, we utilized a validation split of 0.2. The models were run for 25 epochs each as we set up patience. We used a high number of epochs and early stopping for better performance with dropouts of 0.25.

For Task A: The Bi-Directional LSTM has accuracy with test data, 82.07% accuracy score, 82.07% F1 score and 0.41 loss for binary EDOS classification.



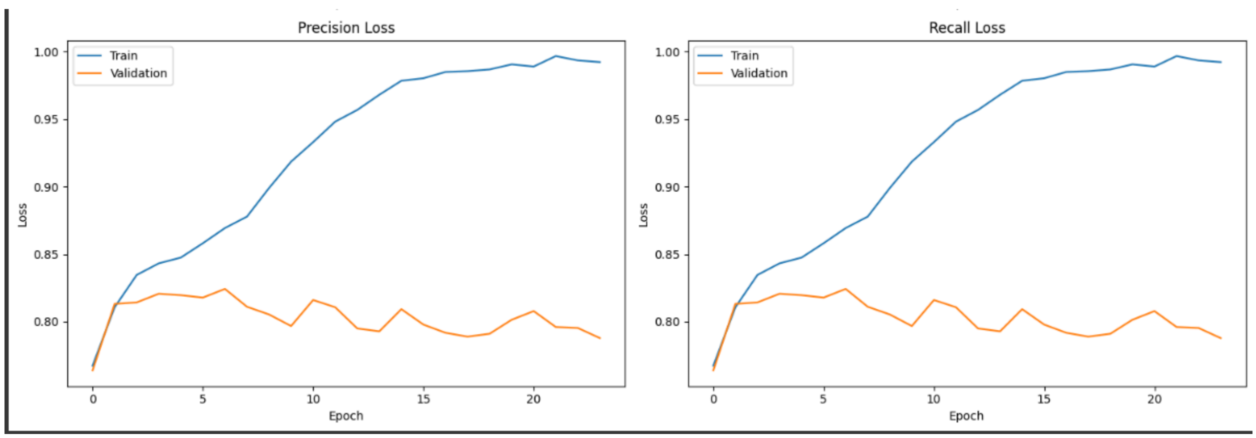
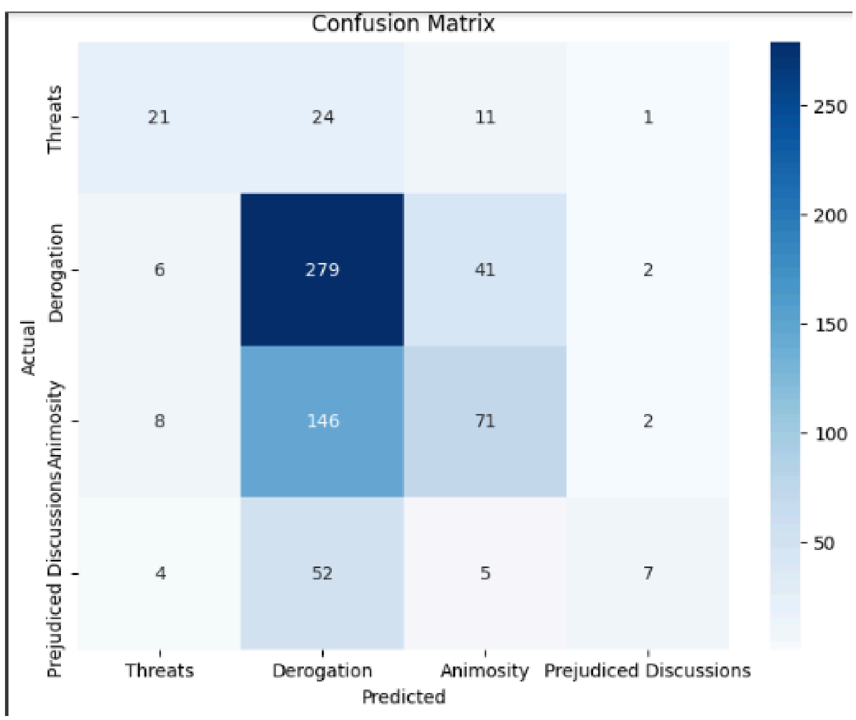


Fig: Confusion matrix and Precision & Recall performance of Task A

For Task B: The Bi-Directional LSTM has accuracy with test data, 55.26% accuracy score, 49.43% F1 score and 0.99 loss for 4 classes EDOS classification.



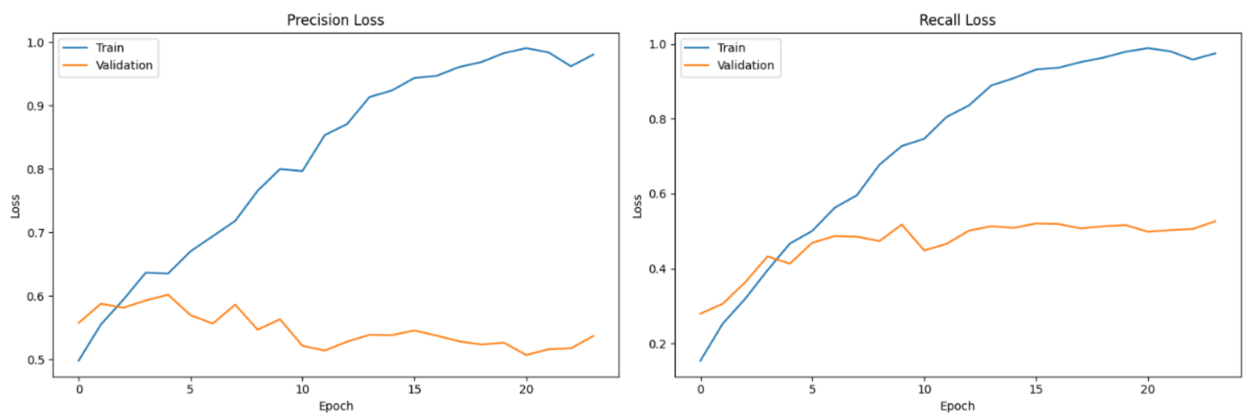


Fig: Confusion matrix and Precision & Recall performance of Task B

Conclusion:

Overall, our study highlights the importance of selecting appropriate models and regularization techniques for classification tasks of varying complexity. Future research in this area could explore the use of other regularization techniques or investigate the impact of different hyperparameters on model performance and use specific data preprocessing techniques. In today's society social media and its use is increasing rapidly so evaluating this type of problem and suggesting some solutions for decreasing it to some extent is very crucial for everyone, especially for the national government of every nation. So using this type of Deep Learning & Classification approach of Natural Languages using Different NLP hyperparameter techniques can be helpful for identifying the issue. In this project we successfully classified for both tasks with a good enough f1 score and accuracy.

Model Name	Loss	Accuracy	F1 Score
Task A (Bi-LSTM)	0.41	0.82	0.82
Task B (Bi-LSTM)	0.99	0.55	0.49

Fig: All results in one frame

Thank You