Sentiment Analysis Project Report

ID - 20241003 Name - Ibney Tousif Rupom
ID - 20101013 Name - Rahmat Ullah Mahin
ID - 20201045 Name - Emon hossen
ID - 21301053 Name - Md Foysal Kabir

Overview

The goal of the project is to perform sentiment analysis on the IMDB dataset using various recurrent neural network architectures, including SimpleRNN, Unidirectional LSTM, and Bidirectional LSTM. The primary objectives were to preprocess the data, build and train different models,leverage pre-trained GloVe word embeddings, build and train different models, and evaluate their performance.

Data Preparation

The IMDB dataset was loaded and preprocessed using the Keras Tokenizer. The reviews were converted into sequences, and GloVe word embeddings were used to represent words as vectors. The dataset was split into training and testing sets, and the sentiment labels were converted to binary format.

Model Architectures

Three recurrent neural network architectures were explored:

SimpleRNN Model:

- Embedding layer with GloVe weights
- SimpleRNN layer with 128 units
- Dense layer with sigmoid activation

LSTM Model:

- Embedding layer with GloVe weights
- LSTM layer with 10 units
- Dense layer with sigmoid activation

Bidirectional LSTM Model:

- Embedding layer with GloVe weights
- Bidirectional LSTM layer with 10 units
- Dense layer with sigmoid activation

Model Training and Evaluation

Models were compiled with binary cross entropy loss and Adam optimizer,training was conducted for 20 epochs with a batch size of 32. Model performance was evaluated using accuracy and confusion matrices. Additionally, a user-defined function **(predict_sentiment1)** was created to make predictions on user input.

Results

```
Improved Model with Bidirectional LSTM Accuracy: 0.8585
Classification Report for Improved Model with Bidirectional LSTM:
              precision    recall  f1-score   support

           0       0.87      0.84      0.86      4981
           1       0.85      0.87      0.86      5019

    accuracy                           0.86     10000
   macro avg       0.86      0.86      0.86     10000
weighted avg       0.86      0.86      0.86     10000
```

```
Original Model Accuracy: 0.7267
Classification Report for Original Model:
              precision    recall  f1-score   support

           0       0.75      0.67      0.71      4981
           1       0.70      0.78      0.74      5019

    accuracy                           0.73     10000
   macro avg       0.73      0.73      0.73     10000
weighted avg       0.73      0.73      0.73     10000
```

```
Improved Model with LSTM Accuracy: 0.8556
Classification Report for Improved Model with LSTM:
              precision    recall  f1-score   support

           0       0.86      0.85      0.85      4981
           1       0.86      0.86      0.86      5019

    accuracy                           0.86     10000
   macro avg       0.86      0.86      0.86     10000
weighted avg       0.86      0.86      0.86     10000
```

We can see that the Bidirectional LSTM has the highest accuracy among all these 3 so bidirectional LSTM will be the most effective for the sentiment analysis task on this particular dataset. Here's a brief comparison of the models:

Bidirectional LSTM:

- Accuracy: 0.8585
- Advantages: The Bidirectional LSTM leverages information from both past and future context, providing a more comprehensive understanding of the review text. This bidirectional information flow helps capture nuanced patterns and relationships, contributing to higher accuracy.

Original Model (SimpleRNN):

- Accuracy: 0.7267
- Observations: The SimpleRNN model might struggle with capturing long-term dependencies in sequential data. Its lower accuracy suggests limitations in understanding intricate patterns and context, especially relevant in sentiment analysis.

LSTM:

- Accuracy: 0.8556
- Strengths: The LSTM model performs better than the SimpleRNN, benefiting from its ability to capture long-term dependencies. However, it falls slightly behind the Bidirectional LSTM, suggesting that incorporating information from both directions further enhances performance.

The Bidirectional LSTM is the best model for sentiment analysis on this dataset, by these accuracy metricsSince the SimpleRNN model lacks a memory mill, to use it would lead not only is less accurate than either standalone LSTM or original (SimpleRNN) models but also more expensive; and bidirectional information flow provides additional advantages other than capturing long-term dependencies.

Improvements

Model Hyperparameters: Experiment with different hyperparameters such as the number of epochs, batch size, and the number of units in the recurrent layers to optimize model performance.
Architecture Exploration: Try different architectures, including variations of LSTM and Bidirectional LSTM, to identify the most suitable model for sentiment analysis.
Fine-tuning Pre-trained Embeddings: Fine-tune the weights of the pre-trained GloVe embeddings during training to adapt them to the specific sentiment analysis task.
Data Augmentation: Explore techniques like data augmentation to increase the diversity of the training data and potentially improve model generalization.

Conclusion

Finally, using the IMDB dataset this project implemented different recurrent neural network architectures for sentiment analysis and obtained satisfactory results.The results offer important information about the models 'ability to capture subtleties of sentiment in movie reviews.The SimpleRNN model was able to provide [insights], which shows that it has, among other things, the following strengths and limitations.With its long-term memory capability, the LSTM model offers [observations], particularly useful for other applications.The Bidirectional LSTM model's ability to gather information from both past and future context is shown by its findings of [give data], in which it displayed proficiency at dealing with several aspects.Although the models were able to attain high accuracies, there are many ways in which they need further improvement.Further hyperparameter tuning steps and testing with different architectures may reveal even better model performance.With training, fine-tuning pre-trained GloVe embeddings could make the word representations fit closer to the sentiment analysis task.Moreover, the investigation of data augmentation methods could broaden training sets by increasing their variability.Adding transformed copies of existing reviews to the data set is likely required so that models can learn how to accommodate more varied input situations.In a nutshell, this project not only compares the effectiveness of various recurrent neural network architectures for sentiment analysis. It also forms a springboard on which to build future advances.These areas identified for improvement give exciting insights into how model accuracy and robustness of the sentiment prediction can be further refined in real-world tasks.The project results further enhance current sentiment analysis methodologies, also providing some possibilities for applications across a wide spectrum of natural language processing areas.