Assignment – 4

Training Models with Varying Sample Sizes:

The values are set to:

Cutoff reviews set to 150 words.

Training samples = 100

Validate samples = 10,000

Words= 10,000

The training sample sizes for the models ranged from 100 to 10,000, and the test loss and accuracy are listed in the table below :

Sample size	One hot encoded sequence		Embedded		Embedded masked		Pre trained	
	Test Loss	Test Accuracy	Test Loss	Test Accuracy	Test Loss	Test Accuracy	Test Loss	Test Accuracy
100	0.6218	0.6629	0.6708	0.5857	0.6586	0.608	0.6787	0.6134
500	0.697	0.565	0.7138	0.6067	0.7351	0.6192	0.6193	0.6696
2000	0.6596	0.5992	0.7226	0.7108	0.8353	0.7032	0.5391	0.7248
5000	0.4891	0.7961	0.5375	0.7924	0.765	0.7645	0.5137	0.7836
10000	0.4380	0.801	0.4455	0.798	0.4349	0.811	0.4573	0.783

Train sample 100, Validation 10000:

- Initial Setup:

1. For this assignment, the IMDB review dataset has been imported.
2. A total of 10,000 words were used as input for the model in its initial setup, which involved collecting 100 training samples, each review having a maximum length of 150 words.
3. Additionally, 10,000 validation samples of both positive as well as negative reviews are used to validate this model.
4. The classification model with optimizer "Adam" was utilized, and the loss function "binary cross-entropy" was applied.

- Models Trained:

1. Using performance metrics such as accuracy, four models were trained, validated, and tested using the original setup.
2. Test accuracy and test loss for one hot-encoded sequence model are 0.801 and 0.4380, respectively.

3. The test accuracy and loss for the embedded model without masking were 0.4455 and 0.798, respectively.
4. An embedded model produced a test accuracy of 0.4349 and a test loss of 0.811.
5. A pre-trained model called Global Vectors for Word Representation (GloVe) produced test accuracy and loss values of 0.783 and 0.4573, respectively.

The findings of the investigation revealed that RNNs with embedded layers outperformed alternative word embedding approaches, such as one-hot encoded sequences, in sentiment analysis. The embedded layer-based models regularly beat other strategies in terms of test loss and accuracy.

Moreover, several embedded layer types—such as normal embedded and masked embedded layers—are compared. In comparison to masked embedded layers, normal embedded layer-based models performed marginally better in terms of test accuracy. It is observed that masking has no effect on the provided IMDb dataset in this model implementation, despite the fact that the masking technique enables the model to ignore padding tokens and focus only on the actual word embeddings, leading to more meaningful representations and improved performance.

CONCLUSION:

- For all cutoff reviews and training sample sizes, the validation accuracy for the embedding layer model is greater than the test accuracy. This implies that the training data may be overfitting the model.
- For certain cutoff reviews and training sample sizes, the validation accuracy of the pre-trained model is higher than the test accuracy, but it is lower for other situations. This implies that the model's performance is less consistent than that of the embedding layer model.
- The results defied the widely held belief that pre-trained embeddings improve model performance, showing that the simple embedding layer model outperformed the pre-trained model. It is often crucial to keep in mind that the pre-trained model in this case is not ideal for the given job and did not refine the embeddings throughout training. Primarily, optimizing the embeddings could potentially improve performance.
- In conclusion, given that these findings are derived from a restricted number of training samples and a narrow set of hyperparameters, caution should be exercised when extrapolating inferences from them. Other results might arise by adjusting the hyperparameters or using additional training data.

loss of 0.4573 and test accuracy of 0.783.

The results of the analysis showed that RNNs with embedded layers performed significantly better than other word embedding techniques, such as one-hot encoded sequences, in the task of sentimentanalysis. The embedded layer-based models consistently outperformed other techniques in terms ofboth test loss and test accuracy.

Furthermore, with comparison different types of embedded layers, including standard embedded and masked embedded layers. The standard embedded layer-based models showed slightly better performance in terms of test accuracy as compared to masked embedded layers. Although the masking technique allows the model to ignore padding tokens and focus only on the actual word embeddings, leading to more meaningful representations and improved performance here in this model implementation it can be observed that there is no impact of masking on the given IMDb dataset.

CONCLUSION:

- For the embedding layer model, the validation accuracy is higher than the test accuracy for all cutoff reviews and training sample sizes. This suggests that the model might be overfitting to the training data.
- For the pre-trained model, the validation accuracy is higher than the test accuracy for some cutoff reviews and training sample sizes, but lower for others. This suggests that the performance of the modelis more variable than the embedding layer model.
- The findings indicated that the straightforward embedding layer model performed better than the pre- trained model, which goes against the common assumption that pre-trained embeddings enhance modelperformance. Generally, it’s important to consider that the pre-trained model here is not optimized for the task specified and didn’t fine-tune the embeddings during training. Essentially, Fine-tuning the embeddings might lead to better performance.
- Finally, we should be attentive in drawing conclusions from these results as they are based on a smallnumber of training samples and a limited set of hyperparameters. It's possible that different hyperparameters or more training data could lead to different conclusions.