Assignment 3

Training Models Across Different Sample Sizes:

The values are set to:

Cutoff reviews set to 150 words.

training samples = 100

Validate samples = 10,000

words= 10,000

The models' test accuracy and loss were documented in the table below. The models were trained with varying training sample sizes ranging from 100 to 10,000.

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:

• First Configuration:

1. The assignment's import of the IMDB review dataset has been completed.

2. The model's first setup involved gathering 100 training samples, each review lasting no more than 150 words, for a total of 10,000 words that were used as input.

3. In addition, 10,000 validation samples of both good and negative reviews are used to validate this model.

4. Since the classification model included an optimizer named Adam, the loss function known as "binary cross-entropy" was employed.

• Trained Models:

Using accuracy as a performance indicator, four models were trained, verified, and evaluated using the original configuration.

2. A hot-encoded sequence model yielded a test loss of 0.4380 and an accuracy of 0.801.

3. The test accuracy and loss for the embedded model without masking were 0.4455 and 0.798, respectively.

4. A masking-embedded model produced a test accuracy of 0.4349 and a test loss of 0.811.

5. Global Vectors for Word Representation (GloVe), a pre-trained model, provided an assessment

The investigation's findings demonstrated that, when it came to sentiment analysis, RNNs with embedded layers outperformed alternative word embedding strategies, such one-hot encoded sequences. Test accuracy and test loss were consistently higher with the embedded layer-based models than with other methods. Additionally, a comparison of several embedded layer kinds, such as masked and conventional embedded layers, is conducted. When compared to masked embedded layers, the normal embedded layer-based models performed marginally better in terms of test accuracy. This model implementation shows that masking has no effect on the provided IMDb dataset, despite the fact that it enables the model to ignore padding tokens and concentrate only on the actual word embeddings, producing more meaningful representations and better performance.

CONCLUSION:

• For all training sample sizes and cutoff reviews, the validation accuracy for the embedding layer model is greater than the test accuracy. This implies that there's a chance the model is overfitting the training set.

• For some cutoff reviews and training sample sizes, the validation accuracy of the pre-trained model is higher than the test accuracy, but lower for other situations. This implies that the model's performance is less consistent than that of the embedding layer model.

• Contrary to popular belief, which holds that pre-trained embeddings improve model performance, the results showed that the simple embedding layer model outperformed the pre-trained model. Generally speaking, it's critical 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. In essence, improving the embeddings might result in improved performance.Finally, because these results are based on a small number of training samples and a constrained set of hyperparameters, we should exercise caution when extrapolating inferences from them. alternative results could be obtained with alternative hyperparameters or additional training data.