Julia Thacker

Advanced Machine Learning – Assignment 3

Original example:

I started by running the original model from the IMDB example in chapter 6.2. As the visual below shows, this model peaked at about 55% Validation loss and 84% validation accuracy.

Chart, line chart

Description automatically generated

IMDB example with modifications:

Next, I created a model based on the criteria provided in the assignment.

This included:

1. Reviews cutoff after 150 words.
2. A training sample set of 100.
3. Validation data of 10,000 samples (40% of the total data).
4. Only the top 10,000 words were considered.

This model did not perform as well as the original example, with 70% validation loss, and 50% validation accuracy, as seen below.

Chart, line chart

Description automatically generated

Embedding layer:

In my next model, I used the same criteria as above, and also used the embedding layer approach. This also returned about 70% validation loss and about 57% validation accuracy.

Chart, line chart

Description automatically generated

Pretrained word embedding:

My next model also featured the same criteria but used a pretrained word embedding instead of an embedding layer. This model showed slightly higher validation loss and slightly lower validation accuracy. Based on my third and fourth models and their results, I decided to use the embedding layer in my final model, because it performed the best of the two.

Chart, line chart

Description automatically generated

Embedding Layer, training sample size adjustment:

In my final attempt, I increased the training data from 100 to 15,000 training samples. This resulted in 86% validation accuracy and only 33% validation loss, the best performance of any model I have built.

Chart, line chart

Description automatically generated