REPORT

Samyuktha Ananthan

**Introduction:**   
  
In natural language processing, sentiment analysis is an essential activity that seeks to ascertain the sentiment conveyed in a text. In this study, we examine sentiment analysis utilizing neural networks on the IMDB dataset, which includes positive and negative movie reviews labeled.  
  
  
**Preprocessing and Data Loading**:  
  
The following is the loading and preprocessing of the IMDB dataset:   
Reviews are shown as sequences of word indices.   
The binary matrices created from these sequences are vectorized, with each word index represented by a 1 in the appropriate location in a 10,000-length binary array.

**Model Architectures:**

Model 1: Comprising three dense layers of 64 units each, employing the hyperbolic tangent activation function (tanh). It also includes dropout layers with a dropout rate of 0.5 after each dense layer, and L2 regularization applied to the first dense layer with a strength of 0.005. The model is compiled using the RMSprop optimizer, mean squared error loss, and accuracy metric.

Model 2: Comprising an extra dropout layer after each dense layer, without explicit regularization applied, is similar to Model 1 and is compiled using the Adam optimizer, mean squared error loss, and accuracy metric.

**Model Training and Evaluation:**

The following is how both models are trained and assessed:  
Twenty epochs and a batch size of 256 are used to train Model 1, while four epochs and a batch size of 512 are used to train Model 2.  
A validation set for evaluation is created using a part of the training data.  
Using the test dataset, both models are assessed and accuracy and loss measures are provided.

**Outcomes and Conclusion:**   
  
Test loss for Model 1 is 0.1616 and test accuracy is 0.8380, while test loss and accuracy for Model 2 are 0.1623 and 0.8723, respectively. Model 2 outperforms Model 1 by a little margin, most likely because of its improved generalization to new data and its more intricate design with extra dropout layers. Using the IMDB dataset, both models successfully analyze sentiment; however, Model 2 exhibits somewhat better accuracy.