Assignment – 1

NEURAL NETWORKS

Sanjana

Overview:

The IMDb review dataset consists of fifty thousand movie reviews, with twenty-five thousand labeled as "positive" or "good" and the remaining twenty-five thousand as "negative." This study explores various techniques to improve the performance of a neural network model using the IMDb dataset. Modifications can be made to an existing neural network model by altering factors such as the activation function, loss function, number of units, number of hidden layers, and regularization methods like dropout. The results are then analyzed to assess the impact of these changes.

Data Processing:

Several preparatory steps were needed to convert the raw text data from the IMDb reviewer dataset into a format suitable for training a neural network. Since including all the words in the dataset would create a high-dimensional input space, we selected the top 10,000 most frequent words. A dictionary was then used to map the words in the top 10,000 list to their respective indices, transforming the text reviews into integer representations. These integer values were then converted into tensors, a necessary step before feeding the data into neural networks. To ensure consistency in review length, longer reviews were truncated, and shorter ones were padded with zeros. As a result, each review was represented as a fixed-length vector, where each element corresponds to the index of a word.

Approaches:

We set a maximum word count and review duration for each review when the data is merged. Next, using a single 16-unit hidden layer, we constructed a simple neural network model. Adam served as the optimizer, the hidden layer parameters were dropout and hyper tuned, the triggering rates were relu and tanh, the loss parameter was binary Cross entropy, and the optimization was MSE. Next, we investigated the previously recommended strategies to improve the model's efficacy. Next, we changed the total number of hidden layers to produce prototypes with one, two, and three hidden layers that were not visible. We used the test and instruction datasets to compare, assess, and refine the models. Our results show that adding three hidden layers improved test accuracy and validity when compared to the use of only one of them.

Hidden Layers and Accuracy Percentage:

- **1-hidden layer, 16-units Accuracy =** 87.7%
- **3-hidden layer, 16-units Accuracy =** 87.2%
- **3-hidden layer, 32-units Accuracy =** 87%
- **2-hidden layer, 64-units Accuracy** = 86.1%
- **3-hidden layer, 128-units Accuracy** = 87.68 %
- **3-hidden layer, 16-units Accuracy** = 87.4%
- **1-hidden layer, 16-units Accuracy** = 88.72 %
- **3-hidden layer, 16-units Accuracy** = 87.24 %
- **2-hidden layer, 16-units Accuracy** = 86.54 %
- **3-hidden layer, 16-units Accuracy** = 87.40%

Conclusion:

We then implemented dropout regularization to reduce overfitting. By utilizing both training and test datasets, we created a new model with dropout layers. Unlike the baseline model, the results revealed that incorporating dropout regularization led to a decrease in validation accuracy. This suggests that different modifications to the neural network model can produce varying levels of accuracy and loss. Compared to Model Hyper, which yielded the best accuracy and loss results, the three dense layers with a dropout rate of 0.5 seemed to deliver the best performance on the IMDb dataset. The mean squared error (MSE) loss function produced the lowest loss value when compared to binary cross-entropy. However, the tanh activation function showed declining accuracy, likely due to the vanishing gradient problem. It was concluded that the Adam optimizer is effective in building the model. Although Model MSE had the smallest loss value, it was less accurate than Model Hyper. When compared to the other models, Model Regularization showed the weakest performance in terms of accuracy. Based on these findings, it can be concluded that Model Hyper outperforms all the other models tested.