NEURAL NETWORKS

SUMMARY

AKHILA CHINTA 811308674

Overview

The IMDb review dataset consists of fifty thousand movie reviews, of which twenty-five thousand are categorized as "positive" or "good," and the remaining twenty-five thousand as "negative." This study looks at various strategies to improve the performance of a neural network model using the IMDb dataset. Activation function, loss function, units, number of hidden layers, and regularization methods like dropout are just a few of the numerous changes that may be made to an existing neural network model. The results are then analyzed.

DataProcessing

To convert the raw text data from the IMDb reviewer dataset into a format that might be utilized for neural network training, a few preprocessing processes needed to be carried out. Taking onto account all of the words in the dataset would result in an incredibly high-dimensional input space, thus we selected just the top 10,000 phrases. Then, in order to convert the text reviews through integer representations, we transferred the definitions in the top 10,000-word list to the appropriate indices using a dictionary. Applying neural networks requires us to translate the integer approximations into tensors. We decreased longer reviews and added zeros to shorter reviews to ensure that all assessments was the same length. Consequently, every review was portrayed as a fixed-size vector, whereby each element indicated the index of a dictionary word.

Approaches

After merging the data, we establish a maximum word count and review duration for each review. Then, we built a basic neural network model with a single 16-unit hidden layer. The triggering rates were relu and tanh; the hidden layer parameters were dropout and hyper tuned; the loss parameter was binary Cross entropy; Adam was the optimizer; and the optimization was MSE. Next, we looked into the previously suggested methods to increase effectiveness of the model.

Then, in order to create prototypes with one, two, and three layers hidden from view, we varied the total number of hidden layers. We compared, evaluated, and improved the models applying the test and instruction datasets. Our findings indicate that the addition of three hidden layers increased test validity and accuracy when compared to the use of only one of them.

1-hidden layer, 16-units Accuracy = 88.6%

3-hidden layer, 16-units Accuracy=88.4%

3-hidden layer, 32-units Accuracy=86.3%

2-hidden layer, 64-units Accuracy=86.1%

3-hidden layer, 128-units Accuracy=82.9%

3-hidden layer, 16-units Accuracy = 86.5%

1-hidden layer, 16-units Accuracy=86.9%

3-hidden layer, 16-units Accuracy = 85.9%

2-hidden layer, 16-units Accuracy = 86.1%

3-hidden layer, 16-units Accuracy=86.6%

3-hidden layer, 32-units Accuracy=88.2%

Conclusion

We subsequently tried dropout regularization in an attempt to prevent overfitting. By using training and test datasets, we established a novel model with dropout layers. As contrasted to the baseline model, our results demonstrated that the use of regularized dropouts reduced validation precision. As a result, it is believed that variable neural network modeling modifications have varied accuracy values and loss functions.

In contrast to the Model Hyper, which had the best accuracy and loss, the three thick layers with a rate of drop-out of 0.5 may be used to provide the best results for the IMDB information set. The mean square error (MSE) loss function showed a smallest loss value when binary cross-entropy was taken into account. Because of the declining accuracy of the tanh activator function, disregarding the vanishing gradient problem. It has been established that building the model effectively utilizes the Adam optimized function.

Because Model MSE has a tiny loss value, it is somewhat inaccurate than Model Hyper. Whenever contrasted to other models, the Model Regularization shows poor accuracy.

Thus, it may be extrapolated that the Model Hyper executes the best over all the models investigated.