Neural Networks Assignment

Chandrashekar Reddy Boreddy

Intoduction: The aim of this endeavor is to delve into diverse methodologies to enhance the efficiency of a neural network framework on the IMDb dataset. We will adjust an existing neural network model and juxtapose the outcomes of various strategies such as altering the quantity of concealed layers, units, loss function, activation function, and regularization methodologies like dropout.

Data Set: We utilized the IMDb dataset, encompassing movie critiques categorized as affirmative or negative. The dataset comprises 25,000 movie critiques for training and an equivalent count for testing.

Methodology: We initiated by loading the data and stipulating the utmost count of words to be contemplated in each critique and the maximal length of each critique. Subsequently, we erected a foundational neural network model with one concealed layer containing 16 units. We utilized binary\_crossentropy as the loss function and relu as the activation function for the concealed layer.

We then scrutinized diverse approaches to augment the efficacy of the model. Initially, we experimented with the count of concealed layers by erecting models with one and three concealed layers. We trained and assessed the models on the training and test datasets and compared the outcomes. We ascertained that utilizing three concealed layers culminated in marginally elevated validation and test accuracy in comparison to utilizing one concealed layer.

Next, we endeavored to utilize layers with more concealed units or fewer concealed units, specifically 32, 64, and 128 units. We trained and evaluated the models with diverse counts of concealed units and charted the validation accuracy for each model. We observed that amplifying the count of concealed units generally resulted in superior validation and test accuracy, but amplifying the count of concealed units excessively can precipitate overfitting.

Subsequently, we tried utilizing the mse loss function instead of binary\_crossentropy. We trained and evaluated the model with mse loss and juxtaposed the outcomes with the foundational model. We found that utilizing mse loss did not significantly influence the performance of the model.

Conclusion: Ultimately, we attempted utilizing dropout regularization to forestall overfitting. We formulated a novel model with dropout layers and trained and evaluated the model on the training and test datasets. We ascertained that employing dropout regularization led to a higher validation accuracy in comparison to the foundational model. It can be inferred that varied permutations of the neural network models exhibit differing levels of accuracy and loss. The Model\_Hyper attained the supreme accuracy and loss, intimating that employing three dense layers with a dropout rate of 0.5 can yield optimal performance for the IMDB dataset. Utilizing the MSE loss function yielded the minimum loss value, in contrast to binary cross-entropy. The tanh activation function manifested diminished accuracy owing to the vanishing gradient quandary. The Adam optimizer function was discerned to be adept at computing the model. Regularization mitigated overfitting and yielded diminutive losses, with the L-2 model evincing marginally superior accuracy. Lastly, the dropout technique abated the loss function, albeit negligibly impacting accuracy. Grounded on the graph, it is observable that the Model\_Hyper boasts the highest accuracy with a reasonably modest loss. The Model\_MSE records the minimum loss value but lags in accuracy compared to the Model\_Hyper. The Model\_tanh portrays substandard accuracy relative to other models, while the model\_regularization exhibits an elevated loss and diminished accuracy vis-à-vis other models. Consequently, it can be surmised that the Model\_Hyper stands as the preeminent performing model among those evaluated.