ASSIGNMENT 1: NEURAL NETWORKS

RISHITHA RAPURI

811283595

Introduction:

The dataset used for this study involves 50,000 movie reviews from IMDb, with half labelled as "positive" or "good" and the other half as "negative." The research aims to optimize the performance of a neural network model by implementing various strategies on the IMDb dataset. The existing neural network model will undergo multiple adjustments, including alterations to the number of hidden layers, units, activation function, loss function, and the application of regularization methods like dropout. The subsequent outcomes will be subjected to detailed analysis.

Objective and Approach:

Improving the neural network model iteratively is the main goal. This entails adjusting crucial variables such the quantity of hidden layers, units, activation function, and loss function as well as incorporating regularization strategies like dropout. The research follows a structured approach to assess the impact of these modifications on the model's predictive capabilities.

Data Processing Techniques:

Integral to this study are Robust data processing procedures, which include data handling, manipulation, computing, analysis, and organizing, are essential to this study. In order to guarantee the effectiveness of later model training and assessment, these procedures are essential for obtaining significant patterns and insights from the IMDb dataset.

Tensor representations of the integer representations were required in order to use neural networks. To make all of the reviews the same length, we padded the shorter reviews with zeros and reduced the longer ones. For this reason, each review was represented as a fixed-length vector with each element denoting a dictionary word's index.

The maximum word count and time for each review were then specified once we imported the data. Next, we constructed a simple neural network model where the hidden layer consists of just 16 units. We employed binary Cross entropy, Mean Squared Error (MSE) as the loss function, return on investment (ROI) of dropout and hyper-tuned hidden layer parameters, and optimization algorithms Adam, Regularization, and Tanh as the activation functions. We then examined the previously indicated approaches in an attempt to improve the model's usefulness. By varying the number of hidden layers, we then produced models with one, two, and three hidden layers. With the test and training datasets, we assessed, compared, and trained the models. We discovered that, in comparison to, adding three hidden layers improved test validity and accuracy as opposed to employing only one hidden layer.

The different approaches that we used for validation and test accuracy:

Neural network with – 1-hidden layer,16-units , loss= binary crossentropy, activation=relu

A graph with blue dots

Description automatically generatedA graph with blue dots

Description automatically generated

Neural network with – 3-hidden layer,16-units , loss= binary crossentropy,activation=relu

A graph with red lines

Description automatically generatedA graph with red dots

Description automatically generated

Neural Network with 32 Hidden units & 3 layers.

A graph with red lines

Description automatically generatedA graph with red dots

Description automatically generated

Neural network with – 3-hidden layer,32-units, loss=binary Cross entropy, activation=relu, optimizer=rmsprop(regularization), droupout=0.5, Hyper tuned parameters (kernel\_regularizer=regularizers. l2(0.0001))

A graph with red dots

Description automatically generatedA graph with red dots

Description automatically generated

Accuracy: 87.1%

Neural network with – 3-hidden layer,16-units ,loss=binary Cross entropy , activation=relu, optimizer=rmsprop(regularization),dropout=0.5

A graph with red dots

Description automatically generatedA graph with red dots

Description automatically generated

Accuracy: 87.5%

Neural network with – 2-hidden layer,16-units ,loss=binary Cross entropy , activation=relu, optimizer=rmsprop(regularization)

A graph with green lines and dots

Description automatically generatedA graph with green lines and dots

Description automatically generated

Accuracy: 86.95%

Neural network with – 3-hidden layer,16-units ,loss=binary crossentropy , activation=relu,optimizer=adam

A graph with red dots

Description automatically generatedA graph with red dots

Description automatically generated

Accuracy: 85.83%

Neural network with – 1-hidden layer,16-units ,loss=MSE , activation=tanh

A graph of training and validation

Description automatically generatedA graph with green lines and numbers

Description automatically generated

Accuracy: 86.55%

Neural network with – 3-hidden layer,16-units ,loss=MSE , activation=relu

A graph with red lines and numbers

Description automatically generatedA graph with red lines

Description automatically generated

Accuracy : 86.5%

Neural network with – 3-hidden layer,128-units ,loss=binarcrossentropy,activation=relu

A graph with green lines and dots

Description automatically generatedA graph with green lines and dots

Description automatically generated

Accuracy: 87.37%

Neural network with – 2-hidden layer,64-units , loss= binarcrossentropy,activation=relu

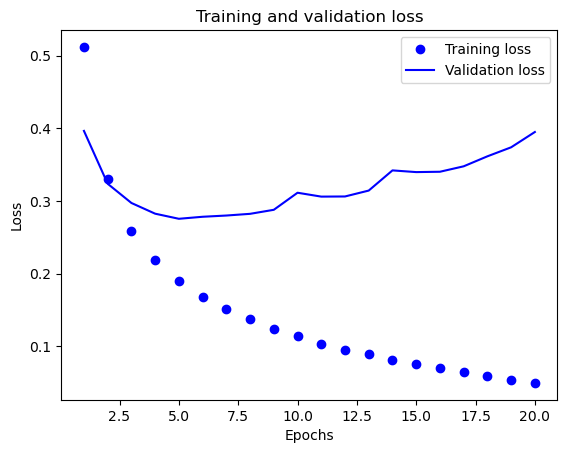
A graph with green lines and white text

Description automatically generatedA graph with green lines and dots

Description automatically generated

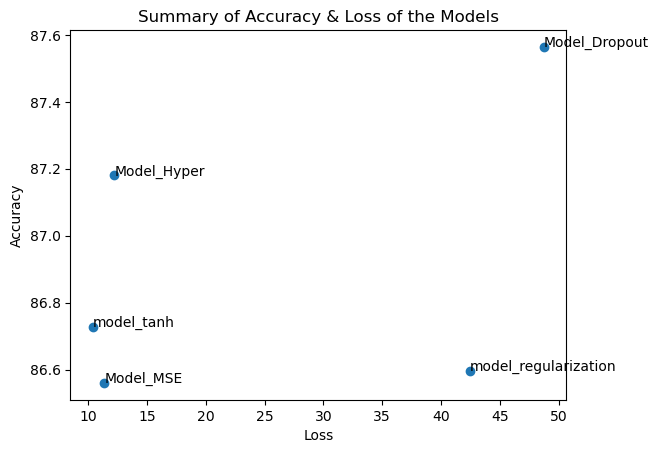
Accuracy: 86.95%

Neural network with – 1-hidden layer,16-units , loss= binary crossentropy,activation=relu

 A graph with blue dots

Description automatically generated

The picture below illustrates the many models that are employed together with their accuracy and validation loss performance, making it easier for us to understand each model.



Conclusion:

In order to avoid overfitting, we finally attempted dropout regularization. Using dropout layers, we created a new model containing training and test datasets. Compared to the baseline model, we discovered that using dropout regularization improved the validation accuracy. Different neural network model modifications are therefore expected to have different accuracy and loss functions. The best accuracy and loss were achieved by the Model Hyper, suggesting that the IMDB dataset would benefit from the use of three thick layers with a dropout rate of 0.5. The loss value for the MSE loss function was lower than that of binary cross-entropy. The issue of the vanishing gradient reduces the precision of the tanh activation function. It was demonstrated that the model could be calculated effectively.