NEURAL NETWORKS SUMMARY

Siva Sai Madhumitha Kotala

811257709

**Introduction:**

This study's dataset consists of 50,000 IMDb movie reviews, of which half are classified as "positive" or "good" and the other half as "negative." The objective of the study is to enhance the neural network model's performance through the use of diverse methodologies on the IMDb dataset. The current neural network model will be modified in a number of ways, such as by changing the units, activation function, loss function, number of hidden layers, and regularization techniques like dropout. The results that follow will be thoroughly examined.

**Objective and Approach:**

The basic objective is to iteratively improve the neural network model. This means modifying important parameters including the number of units, activation function, loss function, and hidden layers in addition to adding regularization techniques like dropout. An organized methodology is used in the study to evaluate how these changes affect the prediction power of the model.

**Data Processing Techniques:**

Robust data processing processes are crucial to this study. They encompass data handling, modification, computation, analysis, and organization. These processes are necessary to extract meaningful patterns and insights from the IMDb dataset and to ensure the efficacy of subsequent model training and assessment.

Neural networks could not be used without tensor representations of the integer representations. We shortened the lengthier reviews and padded the shorter ones with zeros to equalize the length of each review. Because of this, every review was represented as a fixed-length vector, where each element represented the index of a dictionary word.

Following the data import, we were able to choose the maximum word count and duration for each review. Next, using only 16 units for the hidden layer, we built a basic neural network model. The techniques we used were binary Cross entropy, Mean Squared Error (MSE) as the loss function, ROI of the hyper-tuned hidden layer parameters and dropout, and the activation functions were the optimization methods Adam, Regularization, and Tanh. Next, we tried to increase the model's utility by looking at the previously suggested methods.

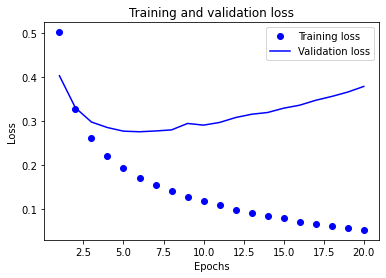
We then created models with one, two, and three hidden layers by changing the quantity of hidden layers. We evaluated, contrasted, and trained the models using the test and training datasets. Compared to using only one hidden layer, we found that adding three hidden layers improved test validity and accuracy.  
The many methods we employed to ensure test accuracy and validity were as follows:

**Approaches:**

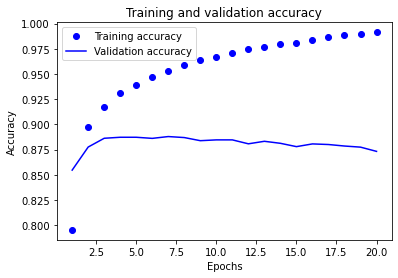
The data was then imported, and the maximum word count and length for each review were established. Next, using a single 16-unit hidden layer, we constructed a simple neural network model. For the hidden layer, we utilized dropout and hypertuned as parameters, Adam, Regularization as the optimizer, and binary Cross entropy, MSE as the loss function. For the activation functions, we used relu and tanh. Next, we examined the previously indicated techniques in an attempt to improve the utility of the model. Next, we varied the number of hidden layers to build models with one, two, and three hidden layers. Using the test and training datasets, we compared, assessed, and trained the models. Comparing the use of three hidden layers to that of one, we discovered that the latter improved test validity and accuracy.

**The many methods we employed for test accuracy and validation are listed below:**

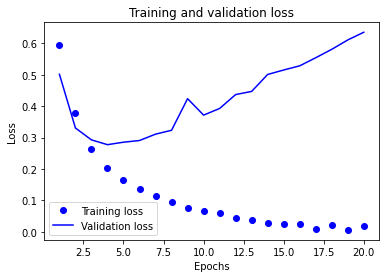
Neural network with – 1-hidden layer,16-units , loss= binary crossentropy,activation=relu



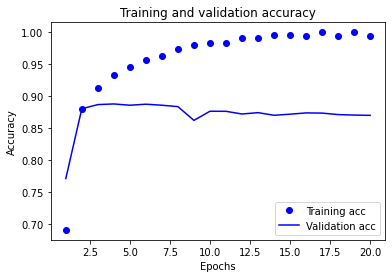
**Accuracy is 88.78%**



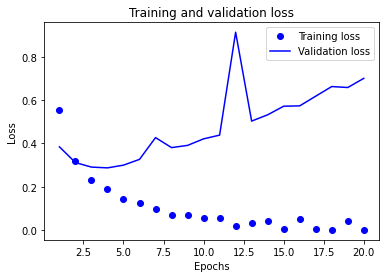
**Neural network with – 3-hidden layer,16-units , loss= binary crossentropy,activation=relu**

****

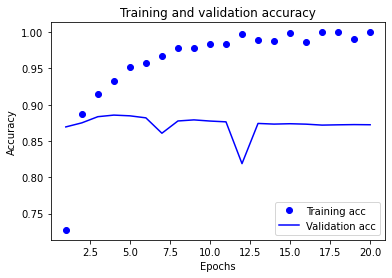
**Accuracy=87.45%**

**** •

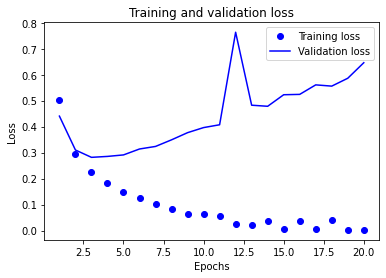
**Neural network with – 3-hidden layer,32-units , loss= binary crossentropy,activation=relu**

****

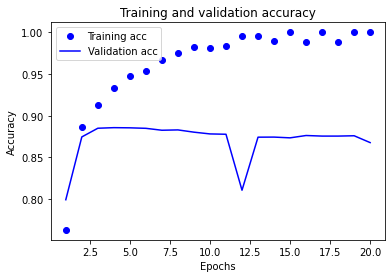
**Accuracy=86.75%**

****

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

****

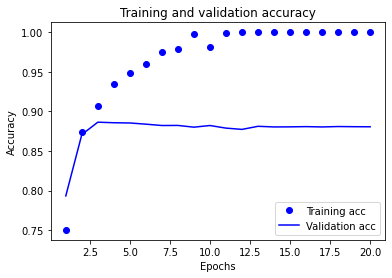
**Accuracy=85.74%**

****

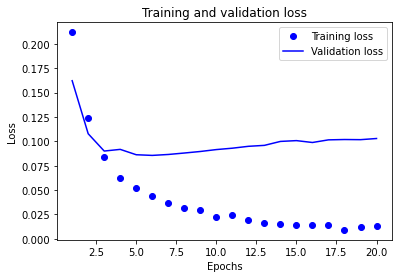
**Neural network with – 3-hidden layer,128-units ,loss=binarcrossentropy,activation=relu**

****

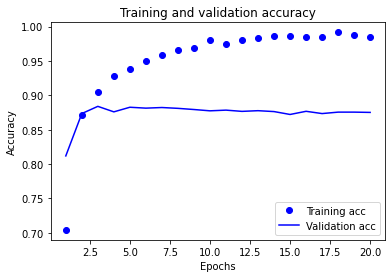
**Accuracy=87.07%**

****

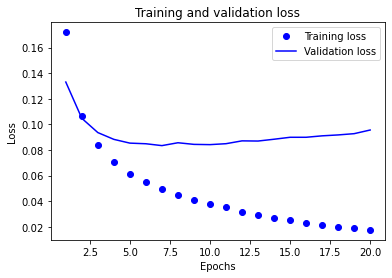
**Neural network with – 3-hidden layer,16-units ,loss=MSE , activation=relu**

****

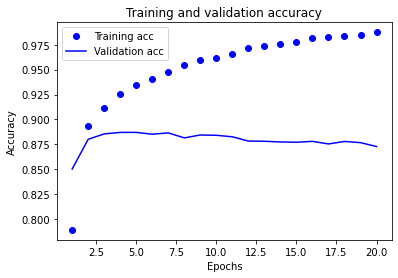
**Accuracy=86.36%**

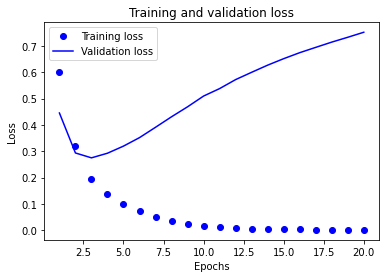
****

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

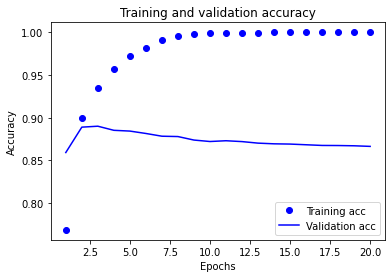
****

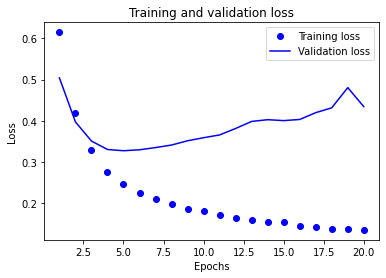
**Accuracy=86.65%**

****

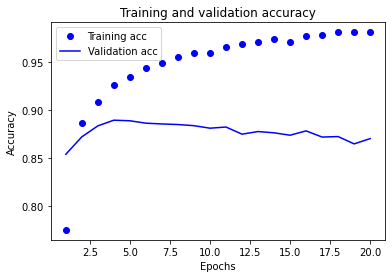
****

**Accuracy=85.74%**

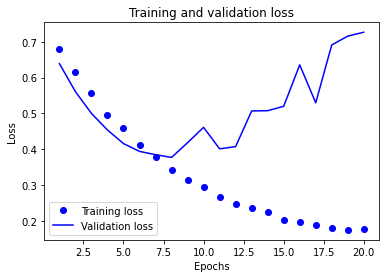
****

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

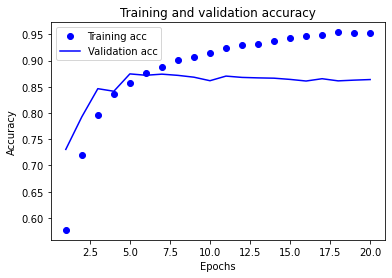
**Accuracy-85.60%**

****

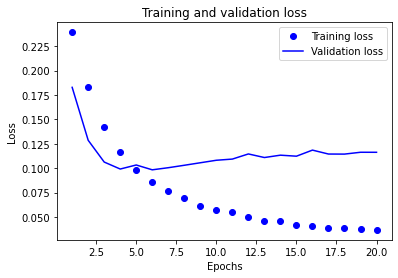
**Neural network with – 3-hidden layer,16-units ,loss=binary Cross entropy , activation=relu, optimizer=rmsprop(regularization),dropout=0.5**

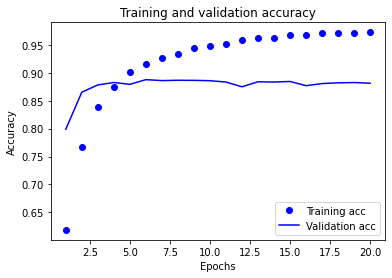
****

**Accuracy-86.28%**

****

**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))**

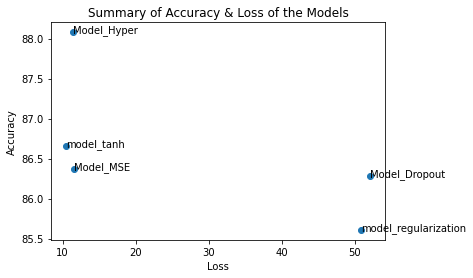
****

**Accuracy-88.08%**

**Conclusion:**

Finally, we tried dropout regularization to prevent overfitting. We developed a new model with training and test datasets using dropout layers. We found that using dropout regularization increased the validation accuracy when compared to the baseline model. Thus, it is expected that different alterations to neural network models will have varied loss functions and accuracy. The Model Hyper produced the best accuracy and loss, indicating that three thick layers with a dropout rate of 0.5 would be advantageous for the IMDB dataset. Compared to binary cross-entropy, the MSE loss function has a smaller loss value. The vanishing gradient problem lowers the tanh activation function's precision. Effective calculation of the model was shown to be possible.

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

****

With the lowest loss value, Model MSE is less accurate than Model Hyper. The Model Regularization shows poor accuracy when compared to other models.

**Thus, we may draw the conclusion that, among all the models examined, the Model Hyper is the most effective.**