

Report on Neural Network Model Building

Kiran Kour

The objective of the Neural Network Model:

To build a NN model which gives the highest accuracy when tested over the test set.

Model building:

As seen in the summary table below, we built ten models with different layers, nodes, and hyper tuning parameters

Model	Layers	Nodes	Activation	Optimizer	Loss	Regularization	Dropout
Model 1	2	16	relu	rmsprop	binary_crossentropy	–	–
Model 2	1	32	relu	adam	binary_crossentropy	–	–
Model 3	3	32	relu	rmsprop	binary_crossentropy	–	0.5
Model 4	1	64	tanh	rmsprop	binary_crossentropy	–	–
Model 5	2	64	relu	adam	mse	–	–
Model 6	2	16	tanh	rmsprop	mse	–	–
Model 7	3	64	relu	rmsprop	binary_crossentropy	L2 (0.001)	–
Model 8	2	16	relu	rmsprop	binary_crossentropy	L2 (0.001)	–
Model 9	2	16	relu	rmsprop	binary_crossentropy	–	0.5
Model 10	2	16	relu	rmsprop	binary_crossentropy	–	0.3

Model	Epochs	Batch Size	Optimal Epochs	Test Accuracy	Test Loss
Model 1	25	512	4	88.48%	28.93%
Model 2	30	512	4	88.38%	29.53%
Model 3	40	520	5	88.42%	30.37%
Model 4	40	512	5	87.98%	30.26%
Model 5	25	512	2	87.70%	9.01%
Model 6	25	512	4	88.26%	8.72%
Model 7	25	512	2	87.58%	43.45%
Model 8	25	512	3	88.44%	33.73%
Model 9	25	512	8	88.55%	30.13%
Model 10	25	512	7	88.77%	27.78%

Final Observations:

- ❖ The maximum number of models was built using the **relu activation function**. Relu or “**Rectified Linear Unit**” does not activate all the neurons at the same time and overcomes the vanishing gradient problem, allowing models to learn faster and perform better. Likewise, the **binary crossentropy loss function** was used for most models because they work well with classification models.
- ❖ **Model 9** and **Model 10**, with the highest accuracy, were built using two layers, 16 neurons, relu being the activation function. Model 9 was built using a **dropout rate of 0.5**, resulting in 88.55% accuracy. When the dropout rate was **reduced to 0.3** during building Model 10, it resulted in maximum accuracy of 88.77%. Dropout is a technique where randomly selected neurons are ignored during training, thereby increasing the efficiency and accuracy of the model.
- ❖ Talking about **regularization**, Model 7 and Model 8 was built using L2(0.001) regularization but showed no significant increase in accuracy. These two models have the greatest loss values.
- ❖ The models were built using **different epochs** but increasing, or decreasing epochs showed no significant increase in accuracy. The least loss value and highest accuracy were achieved mostly in the initial epochs.
- ❖ If we talk about **minimum loss value** and **considerable accuracy**, **Model 6** did a good job. It had a loss value of 8.72%, the least among all the other models, and an accuracy of 88.26%. It was built using **tanh** as the activation function. But, notice the loss function used was **mse**, which is not an ideal loss function for the classification model, and we are dealing with the IMDB dataset, which is about classification.
- ❖ Changing the number of **hidden layers** also did not show any considerable change in the performance of the models. The models with layer two and layer one showed some increase in the accuracy of models.

Conclusion:

To conclude, the **best model observed was Model 10**, which had an **accuracy of 88.77%** and a **loss value of 27.78%**, built with two hidden layers, 16 neurons, relu as the activation function, rmsprop as the optimizer, binary crossentropy as the loss function and dropout rate of 0.3

