Neural Networks (Assignment -1)

The implementation of neural networks for IMDB reviews dataset has been done with several parameters, below are the results for various changes to layers, units, etc.

1. Layers:

To implement the model using different hidden layers, the base setup taken as unit size of 16 and loss function as binary cross entropy and below are the results for it,

Layers	Loss	Accuracy	V_Loss	V_Accuracy	T_Loss	T_Accuracy
1	0.2664	0.914	0.3066	0.8801	0.3165	0.8762
2	**0.2322**	**0.9208**	**0.2799**	**0.8898**	**0.2931**	**0.8842**
3	0.2549	0.9107	0.2886	0.8859	0.3009	0.882

It can be observed that when two hidden layers are used the validation loss and test loss are less when compared to other layers. Using one or three hidden layers did not significantly improve the accuracy of the model compared to using two hidden layers. This suggests that the model's capacity is appropriate for the problem, and adding or removing layers may not help.

2. Hidden Units:

To implement the model using different hidden layers, the base setup taken as hidden layer of 2 and loss function as binary cross entropy and below are the results for it,

Units	Loss	Accuracy	V_Loss	V_Accuracy	T_Loss	T_Accuracy
16	**0.2322**	**0.9208**	**0.2799**	**0.8898**	**0.2931**	**0.8842**
32	0.2284	0.9183	0.3048	0.8755	0.3138	0.8718
64	0.2208	0.9191	0.3009	0.878	0.3117	0.8704

Increasing the number of hidden units did not have a significant effect on the accuracy of the model. But it can be observed that for less unit size of 16 the model is having high accuracy and hence select the model with unit size 16. There isn't any significant increase because the dataset might not be much complex for which higher hidden unit size doesn't have any impact.

3. Loss Function:

To implement the model using different loss functions, the base setup taken as hidden layer of 2, unit size of 16 and below are the results for it,

Loss Function	Loss	Accuracy	V_Loss	V_Accuracy	T_Loss	T_Accuracy
Binary Cross entropy	**0.2322**	**0.9208**	**0.2799**	**0.8898**	**0.2931**	**0.8842**
MSE	0.074	0.9165	0.0868	0.8887	0.0913	0.8832

The MSE loss function doesn't have any impact on the accuracy of the existing model furthermore it reduced the accuracy of the model slightly. Using binary cross entropy would be the optimal choice for the model.

4. Activation Function:

To implement the model using different activation functions, the base setup taken as hidden layer of 2, unit size of 16 and below are the results for it,

Activation	Loss	Accuracy	V_Loss	V_Accuracy	T_Loss	T_Accuracy
Relu	0.2322	0.9208	0.2799	0.8898	0.2931	0.8842
Tanh	0.2138	0.928	0.2696	0.8906	0.2856	0.884

Using Tanh function, the model test accuracy has not been improved. Also, Relu activation is good at capturing non-linear relationship between the data.

5. Regularization and Dropout:

The original model setup with Unit size =16, Activation function = Relu, Layers =2, Loss Function= Binary cross entropy has the validation loss is 0.2799 and accuracy is 0.8898

For this model, regularization L2 is applied with 0.0002 and dropout rate as 0.3 the results obtained are validation loss= 0.2572 and accuracy = 0.9107. Using these fundamentals, it allowed the model to train, and the validation loss and accuracy has improved.

```
Epoch 1/20
49/49 [==============================] - 3s 39ms/step - loss: 0.5324 - accuracy: 0.7527 - val_loss: 0.3335 - val_accuracy: 0.8980
Epoch 2/20
49/49 [==============================] - 3s 71ms/step - loss: 0.3572 - accuracy: 0.8723 - val_loss: 0.2572 - val_accuracy: 0.9107
Epoch 2: early stopping
```

Summary:

From the results obtained it can be observed that a neural network model for IMDB dataset have,

- Right number of hidden layers, if the layers are less the model capability will be reduced and if the layers are more even then the capability will be reduced. In this model 2 layers are considered.
- Right unit size based on the dataset, If the dataset is not too complex or dataset is less the unit size should be less. Here in this model unit size of 16 is optimal.
- Proper selection of loss function. It can be observed that loss function binary cross entropy has higher accuracy rate than mean square error loss function.
- Relu activation function has high capability of finding the non-linear relationship between datapoints rather than tanh activation function
- To improve the original, regularization and dropout techniques are used. With the technique used, validation loss has been decreased slightly and accuracy is increased. These additions have improved the model slightly by 2%.

Overall, the best performance was achieved by using two hidden layers with 16 units in each layer, binary cross-entropy loss function, relu activation function, and adding L2 regularization and dropout regularization with a dropout rate of 0.3 to the model. The validation accuracy of this model was around 0.91.