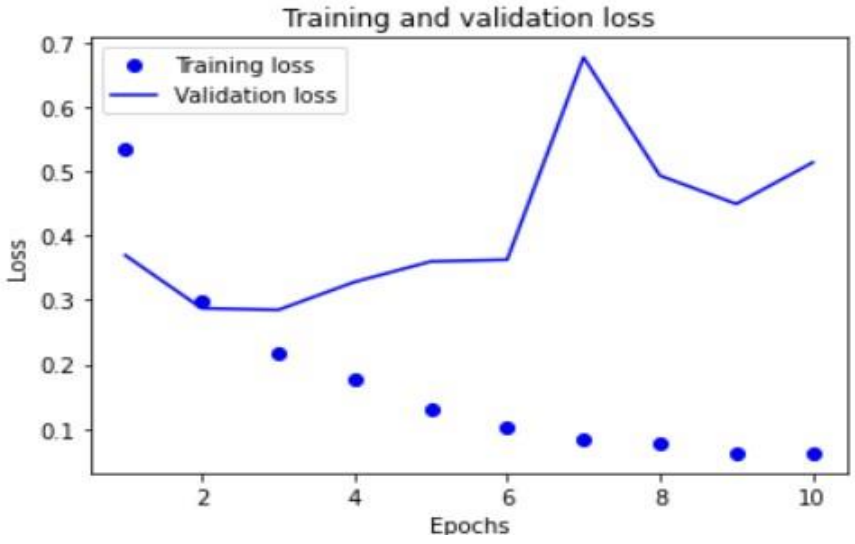
Neural Network Model for IMDB Sentimental Analysis

Machine learning (ML) and deep learning (DL) techniques are used to implement sentiment analysis on the IMDB movie reviews dataset in order to evaluate the model's efficiency.

Activation Functions Used for the Model

- ✓ ReLU (Rectified Linear Unit)
- ✓ Sigmoid
- ✓ Tanh (Hyperbolic Tangent)

```
Epoch 1/10
79/79 [==============================] - 6s 62ms/step - loss: 0.3756 - accuracy: 0.8376 - val_loss: 0.2654 - val_accuracy: 0.8914
Epoch 2/10
79/79 [==============================] - 4s 56ms/step - loss: 0.2044 - accuracy: 0.9220 - val_loss: 0.2689 - val_accuracy: 0.8946
Epoch 3/10
79/79 [==============================] - 5s 66ms/step - loss: 0.1540 - accuracy: 0.9431 - val_loss: 0.2897 - val_accuracy: 0.8875
Epoch 4/10
79/79 [==============================] - 4s 52ms/step - loss: 0.0986 - accuracy: 0.9659 - val_loss: 0.3681 - val_accuracy: 0.8838
Epoch 5/10
79/79 [==============================] - 6s 70ms/step - loss: 0.0515 - accuracy: 0.9842 - val_loss: 0.4519 - val_accuracy: 0.8798
Epoch 6/10
79/79 [==============================] - 4s 52ms/step - loss: 0.0279 - accuracy: 0.9911 - val_loss: 0.5898 - val_accuracy: 0.8780
Epoch 7/10
79/79 [==============================] - 4s 52ms/step - loss: 0.0216 - accuracy: 0.9930 - val_loss: 0.6167 - val_accuracy: 0.8759
Epoch 8/10
79/79 [==============================] - 5s 68ms/step - loss: 0.0160 - accuracy: 0.9951 - val_loss: 0.6612 - val_accuracy: 0.8726
Epoch 9/10
79/79 [==============================] - 4s 53ms/step - loss: 0.0150 - accuracy: 0.9954 - val_loss: 0.6477 - val_accuracy: 0.8762
Epoch 10/10
79/79 [==============================] - 4s 52ms/step - loss: 0.0202 - accuracy: 0.9931 - val_loss: 0.6833 - val_accuracy: 0.8766
```



As we can see accuracy validation loss goes up after 2 epochs due to overfitting.

We will be setting Epochs= 2, Batch size= 512 & RMSProp optimizer to obtain the following results:

OBSERVATIONS

Layers	Activation Fn.	Nodes	Lose Fn.	Accuracy Validation(epoch1)	Accuracy Validation(epoch2)	Accuracy Test
1	tanh	16	mse	86.42	87.28	85.52
2	tanh	16	mse	87.17	87.45	87.57
3	tanh	16	mse	88.02	87.42	88.05

Layers	Activation Fn.	Nodes	Lose Fn.	Accuracy Validation(epoch1)	Accuracy Validation(epoch2)	Accuracy Test
1	tanh	32	mse	87.54	85.6	85.23
2	tanh	32	mse	85.21	86.57	85.78
3	tanh	32	mse	87.79	87.21	88.25

Layers	Activation Fn.	Nodes	Lose Fn.	Accuracy Validation(epoch1)	Accuracy Validation(epoch2)	Accuracy Test
1	tanh	64	mse	86.38	82.76	86.54
2	tanh	64	mse	87.88	85.51	88.45
3	tanh	64	mse	85.28	85.35	85.38

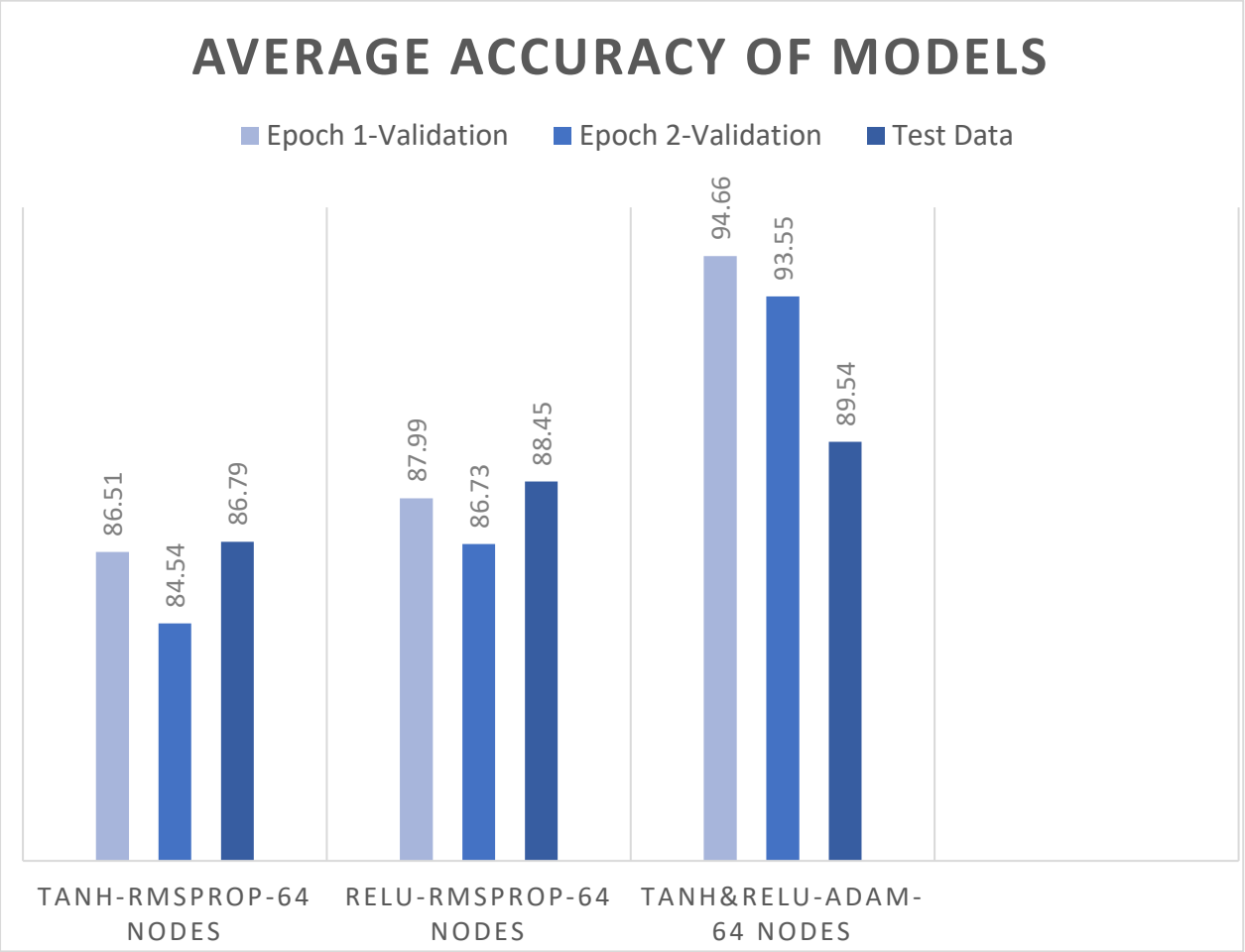
Layers	Activation Fn.	Nodes	Lose Fn.	Accuracy Validation(epoch1)	Accuracy Validation(epoch2)	Accuracy Test
1	binary cross entropy	64	mse	87.49	86.42	88.11
2	binary cross entropy	64	mse	88.23	87.43	88.46
3	binary cross entropy	64	mse	88.26	86.35	88.31

Improving Accuracy using drop-out method and hyper-parameter tuning

✓ Number of Hidden layers	3
✓ Number of Nodes	64
✓ Drop-out Rate	0.1
✓ Activation Function	ReLU & tanh
✓ Optimizer	Adam
✓ Hyper Parameter- Learning Rate	0.0001

Layers	Activation Fn.	Nodes	Lose Fn.	Accuracy(epoch1) Validation	Accuracy(epoch2) Validation	Accuracy Test
1	tanh & ReLU	64	mse	94.62	93.28	89.41
2	tanh & ReLU	64	mse	95.37	95.68	89.63
3	tanh & ReLU	64	mse	94.01	91.71	89.59

RESULTS



- ✓ As we can see from the above results the accuracy of the model has significantly improved by 1%.
- ✓ The model was reformed using 64 nodes at input level and 1 node at output layer with sigmoid activation function.
- ✓ The shortcoming of tanh activation function of vanishing gradient is overcome by combining it with ReLU function.
- ✓ Dropout is used at 20-30% to randomly drop out (set to zero) some units in a layer during training and this helps in capturing the important features by nodes in each level so as to prevent overfitting.
- ✓ Adam is preferred over RMSProp because it combines the benefits of Adagrad and RMSProp.This allows Adam to adaptively change the learning rate for each parameter, which can result in faster convergence and better performance than other optimization algorithms.