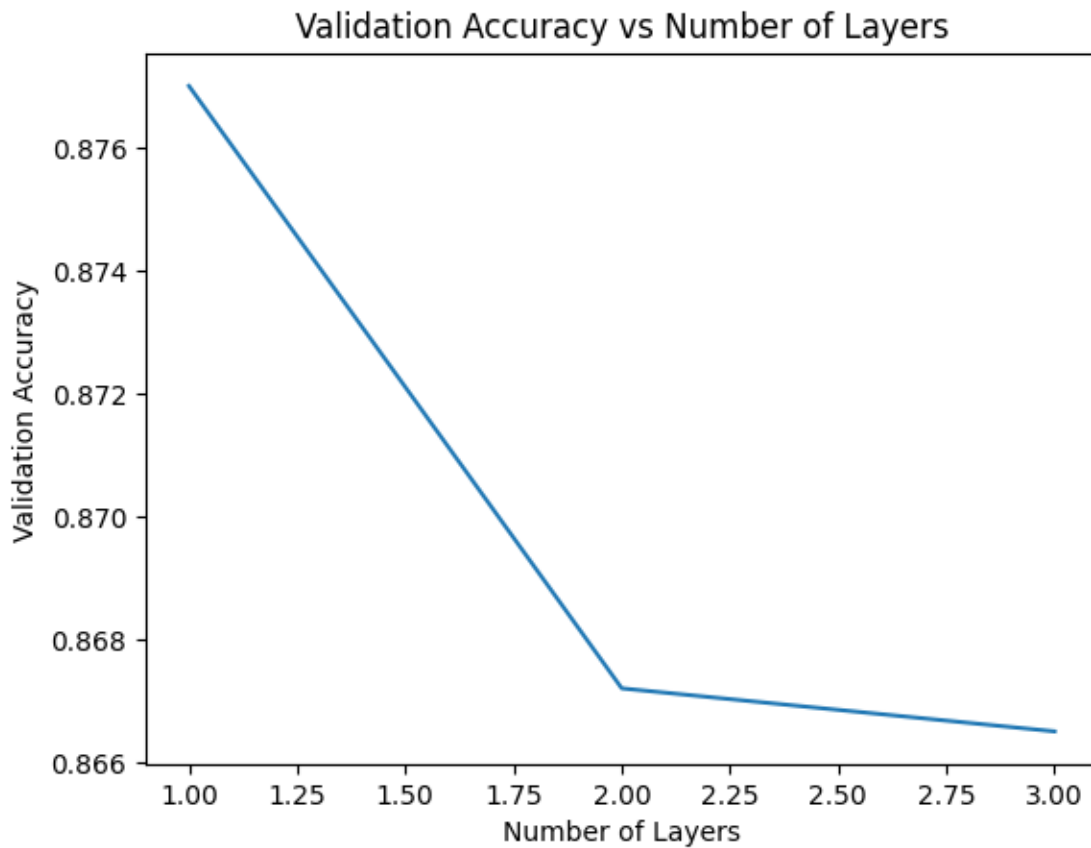# Assignment 1

Overview

In this assignment, we will conduct a series of experiments to tune the parameters of a neural network mode, for the IMDB movie review binary classification problem. The main goal was to understand how different architectural and training choices may affect the accuracy of the model.

Results

Experiment	Validation Accuracy	Test Accuracy
Number of Layers (1)	0.8770	0.86764
Number of Layers (2)	0.8672	0.85484

Experiment	Validation Accuracy	Test Accuracy
Number of Layers (3)	0.8665	0.85328
Number of Units (32)	0.8677	0.85468
Number of Units (64)	0.8763	0.86468
MSE Loss	0.8738	0.86452
Tanh Activation	0.8678	0.85288
Dropout Regularization	0.8814	0.87364

Findings

Based on the results, we find that as the number of hidden layers increased from 1 to 3, both the validation and test accuracies decreased. This suggests a simpler architecture with fewer hidden layers may be more effective. Besides, increasing the number of hidden units from 32 to 64 led to an improvement in both validation and test accuracies. Using the Mean Squared Error loss function instead of binary cross-entropy also resulted in a relatively good performance. If we replace the ReLU activation function with the Tanh activation function, it did not lead to significant improvements in accuracy. After applying dropout regularization with a rate of 0.5, the model achieved the highest validation and test accuracies among all the experiments.

Conclusion

In conclusion, the model with dropout regularization performed the best in terms of both validation and test accuracy. This implies that regularization is a crucial technique for improving the generalization ability of the model on unseen data. For the architecture, a relatively simple model with an appropriate number of hidden units may be sufficient for this problem.