Assignment 1 Report
Neural Network for IMDB Sentiment Classification

Introduction

The purpose of this assignment was to develop and evaluate a neural network model to classify movie reviews from the IMDB dataset as positive or negative. Sentiment classification is a common processing task that helps understand customer feedback and opinions.

Neural networks can be created with different hidden layers, hidden units, activation functions, and loss functions. These design choices affect how well the model learns patterns in the data and how accurately it can perform on unseen data.

The goal of this assignment is to see how different neural network design changes can affect the validation accuracy and overall model performance.

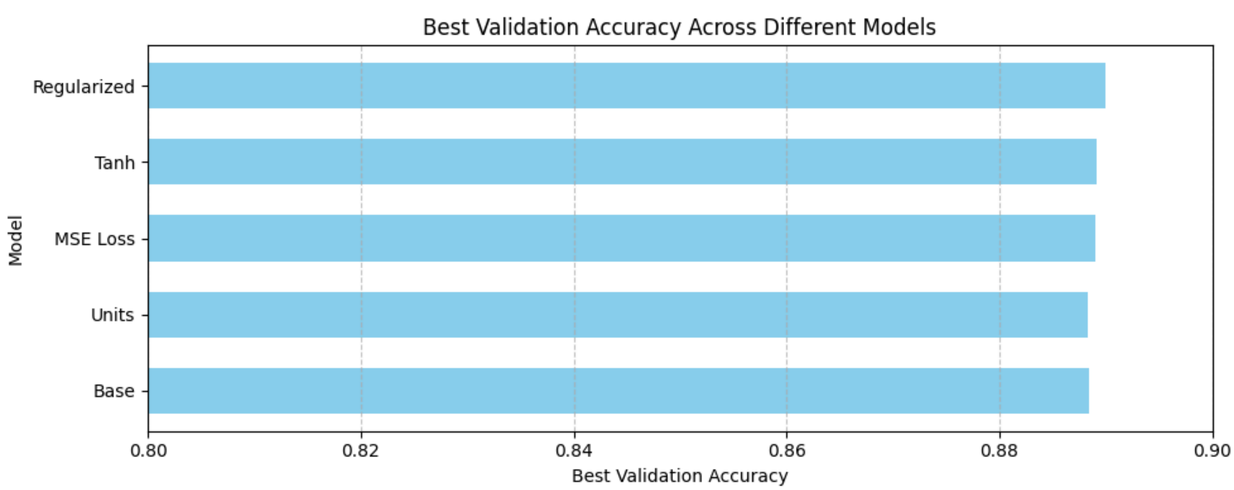
Methodology

The IMDB dataset consists of over 50,000 movie reviews labeled as positive or negative. The dataset was split into three sets. A training set (used to train the model), a validation set (used to evaluate model performance during training), and a test set (used to evaluate final performance). I tested different neural network models such as testing the number of hidden layers 1, 2, and 3 layers. I tested the hidden units 16, 32, and 64 units. I tested the changing the loss function from binary-crossentropy to mean squared error. I also tested activation functions ReLu and tanh. Finally I tried using regularization techniques such as L2 and dropout to help validation accuracy. Each of these models were trained for multiple epochs and validation accuracy was used to compare performance.

Results

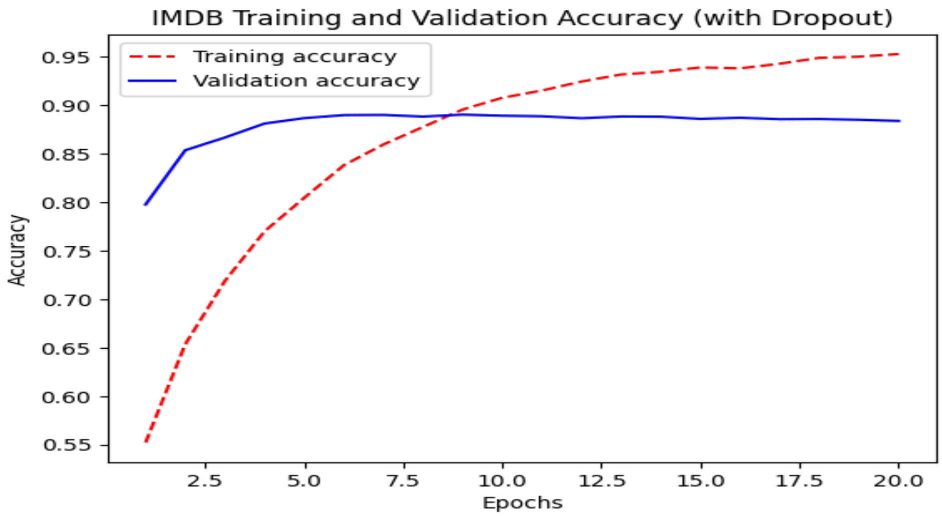
Model performance accuracy

Model	Best Validation Accuracy	Test Accuracy
Baseline (2 layers, 16 units, relu, BCE)	0.8862	0.8574
1 Hidden Layer	0.8846	0.8592
3 Hidden Layers	0.8883	0.8552
32 Units	0.8890	0.8590
64 Units	0.8891	0.8528
MSE Loss	0.8890	0.8590
tanh Activation	0.8891	0.8528
Dropout Model	0.8900	0.8768



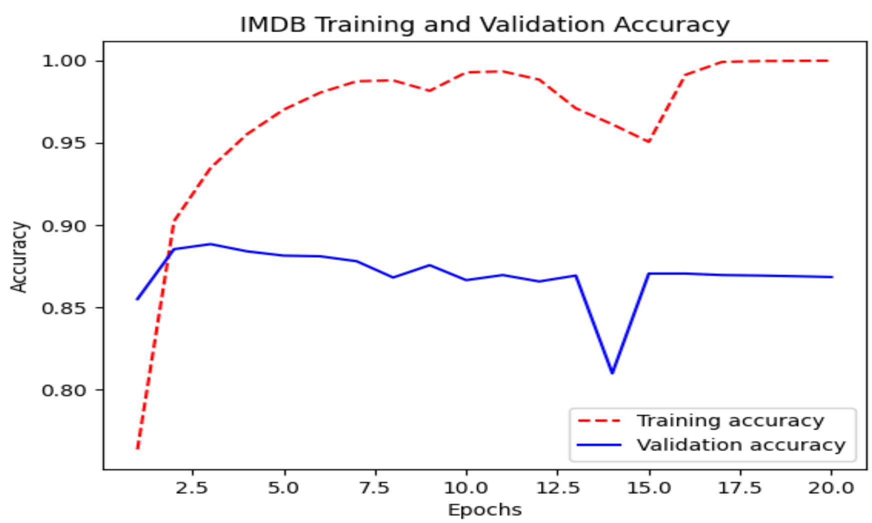
In this figure shows the best validation accuracy across all models tested. This shows that the dropout and regularized model had the highest validation accuracy.

Dropout Model



In this figure shows the validation accuracy over 20 epochs with the training accuracy. You can see that with adding these changes to the model it helps improve the overfitting of the model.

Baseline Model



In this figure you can see the baseline model after 20 epochs. The training accuracy hits around 100% while you can see the validation accuracy go down because of overfitting starting to take over.

Discussion

The results show that neural network design does affect model performance. Increasing the number of hidden layers from 2 to 3 produced a slightly higher validation score but showed more unstable training behavior and showed signs of overfitting.

Increasing the number of hidden units improved performance. The model with 64 hidden units had a validation accuracy of 0.8891. This was higher than the base model. This shows that in some cases model capacity can help a model learn more complex patterns.

The binary-crossentropy loss function and mean squared loss function produced similar validation accuracy results. Binary-crossentropy is more appropriate for this type of classification problem and provides more stable training.

The tanh activation function performed similarly to relu but did not improve performance. Relu remains more efficient as it was designed for more modern neural networks.

The dropout and regularization model had the best performance with the highest validation accuracy of .8900 and the highest test accuracy of 0.8768. This shows that dropout and L2 regularization can reduce overfitting a model and improve generalization.

The training graphs also showed validation accuracy increased during the early epochs but became less stable later showing that overfitting was happening.

Conclusion

The best performing model was the dropout and regularization model. The results show that neural networks performance depends on these design choices and regularization. These results show that neural network performance depends on model design choice and regularization techniques to get optimal performance. Based on these experiments, multiple hidden layers, relu activation, binary-crossentropy, and dropout and regularization help create a good balance between generalization and not overfitting the model.