**Summary Report**

**Assignment -2**

**Introduction:**

In this study, I explored the methods to improve the performance of an existing neural network model using the iris dataset and made the modifications. These consist of changing the loss function, exploring different activation functions, adjusting the number of hidden layers, and varying the number of hidden units. To further improve the model's performance, I will also use dropout and regularization methods. The results of these changes on the model's validation and test accuracy will be presented in this report, providing insights into how different methods impact neural network performance.

1. **You used two hidden layers. Try using one or three hidden layers and see how doing so affects validation and test accuracy.**

**A graph with blue dots

Description automatically generatedA graph with blue lines

Description automatically generated**

**A graph with blue lines

Description automatically generatedA graph with blue dots

Description automatically generated**

This same dataset was used for both model training and evaluation. An accuracy of 96.93% was attained by the One Hidden Layer Model and 98.05% by the Three Hidden Layers Model.

Compared to the simpler One Hidden Layer Model, the performance of the models indicates that the Three Hidden Layers Model's increased complexity did not significantly improve accuracy.

1. **Try using layers with more hidden units or fewer hidden units: 32 units, 64 units, and so on.**

**A graph with blue lines

Description automatically generatedA graph with blue lines and white text

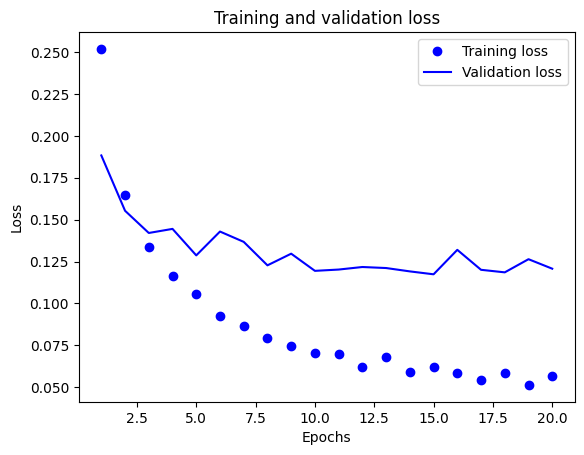
Description automatically generatedA graph with blue lines

Description automatically generatedA graph with blue dots

Description automatically generated**

The accuracy of the Fewer Hidden Units Model was 98.25%, whereas the More Hidden Units Model was 97.65%. Comparing the models with more hidden units to those with fewer, there was no noticeable improvement in accuracy. This implies that there was little to no benefit in this case from the More Hidden Units Model's increased complexity.

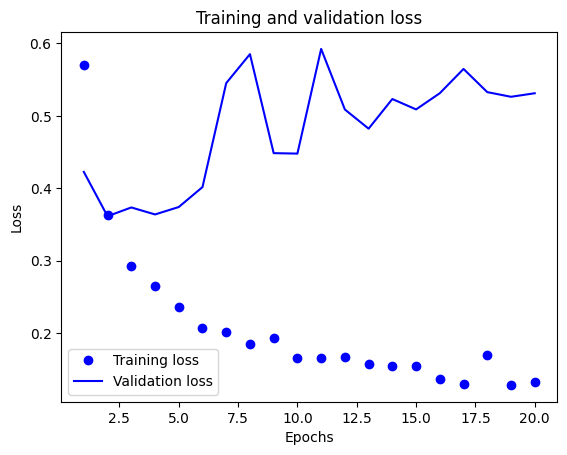
1. **Try using the mse loss function instead of binary\_crossentropy.**

**A graph with blue dots

Description automatically generated**

The model was trained and evaluated using MSE loss. The model achieved an accuracy of 87.3%.

1. **Try using the tanh activation (an activation that was popular in the early days of neural networks) instead of relu.**

**A graph with blue dots

Description automatically generated**

L2 regularization with a coefficient of 0.001 applied to the kernel of each hidden layer. There is a slight fluctuation in both training and validation accuracy across epochs, indicating some degree of overfitting. The training and validation losses also show fluctuations, suggesting that the model may benefit from further regularization or dropout layers to improve generalization.

The accuracy for the model with tanh activation function is approximately 86.95%.

1. **Use any technique we studied in class, and these include regularization, dropout, etc., to get your model to perform better on validation.**

A graph with blue dots

Description automatically generatedA graph with blue dots

Description automatically generated

The model includes dropout layers with a 50% dropout rate after each hidden layer to reduce overfitting. The accuracy of the model on the validation data after 20 epochs is approximately 87.23%.

**Conclusion:**

Using the iris dataset, a neural network model's performance was enhanced by a few adjustments. Accuracy was not significantly increased by adding more hidden layers or units. Neither the tanh activation function nor the mean squared error (MSE) loss function showed significant improvement. Using regularization and dropout methods improved generalization and decreased overfitting, achieving an accuracy of approximately 87.23% on the validation data. The model that performed the best was binary\_crossentropy with a loss function of less hidden units. Reducing dropout rates or exploring into different regularization methods may provide further optimizations.

**Accuracy table:**

|  |  |  |
| --- | --- | --- |
| Approach | Training Accuracy | Validation Accuracy |
| 1. One hidden layer 2. Three hidden layers | 1. 98.05% 2. 96.93% | 1. 85.73% 2. 86.58% |
| 1. Fewer hidden units 2. More hidden units | 1. 98.25% 2. 97.65% | 1. 86.06% 2. 86.83% |
| mse loss function instead of binary\_crossentropy | 96.07% | 87.30% |
| tanh activation | 97.74% | 86.95% |
| Regularization | 96.31% | 87.23% |