

1. In your scikit-learn implementation, wet hidden_layer_sizes to the following values (5,5), (10,10), (30,10) and keep other parameters the same or as default. Report their training and testing ac- curacy using the same train-test split in the test_Part1().

TRAINING REPORT

Table 1:

Error vs Epochs: [30, 10]

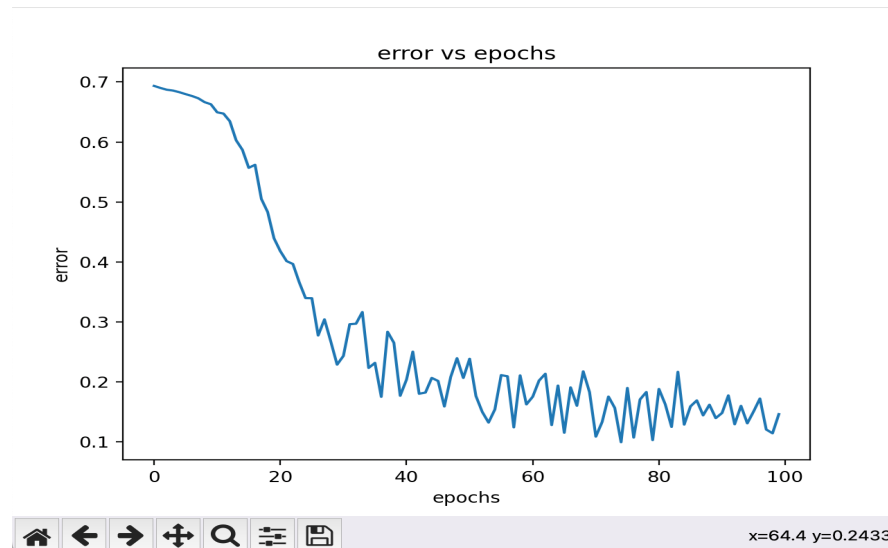


Table 2:

Error vs Epochs: [5, 5]

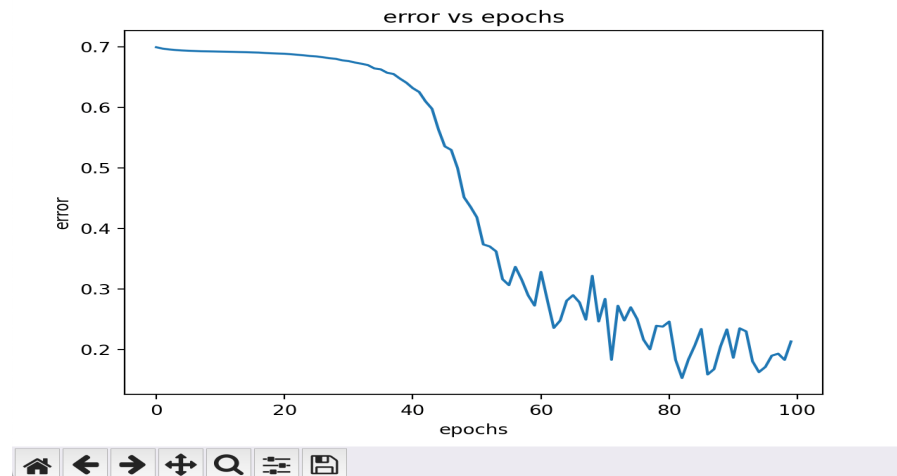
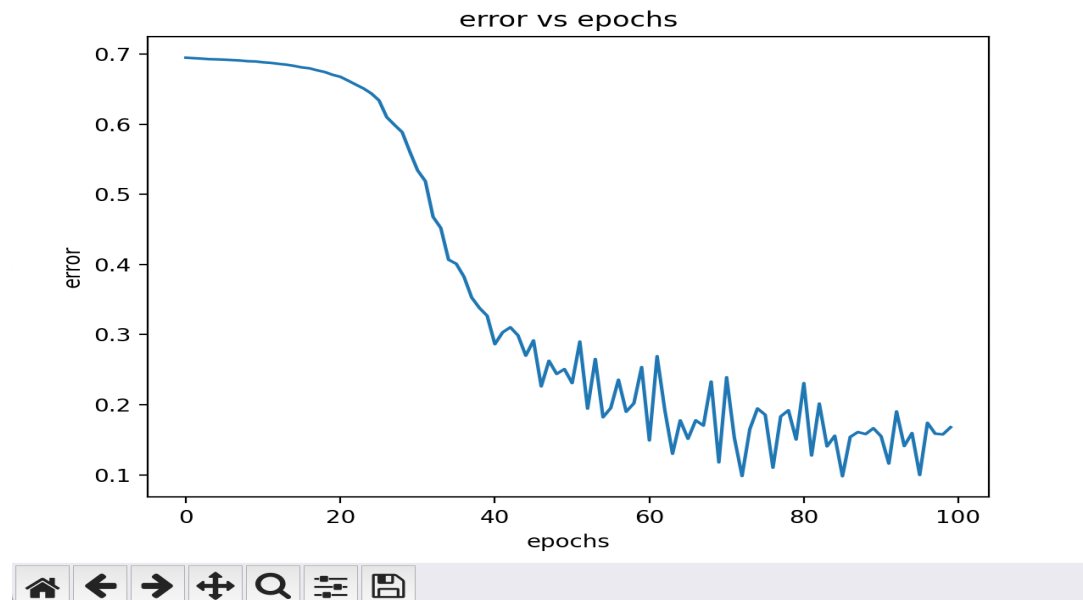


Table 3

Error vs Epochs: [10, 10]



Results from Terminal:

For hidden layer [30, 10]

Train accuracy 0.975

Test accuracy 0.95

Confusion Matrix is from Part 1a for test is:

```
[[ 7  1]
 [ 0 12]]
```

Confusion Matrix from Part 1b is:

```
[[ 7  1]
 [ 1 11]]
```

For hidden layer [5, 5]

Train accuracy 0.925

Test accuracy 0.95

Confusion Matrix is from Part 1a for test is:

```
[[ 7  1]
 [ 0 12]]
```

Confusion Matrix from Part 1b is:

```
[[ 7  1]
```

[1 11]]

For hidden layer [10, 10]

Train accuracy 0.875

Test accuracy 0.95

Confusion Matrix is from Part 1a for test is:

[[7 1]

[0 12]]

Confusion Matrix from Part 1b is:

[[7 1]

[1 11]]

Reflection on Hidden Layers and their impact:

Different hidden layer sizes can significantly impact the performance of a neural network. Here are a few considerations:

- Overfitting vs Underfitting: increasing the hidden layer size might result in a more complex model that could better fit the training data, but it could also lead to overfitting, where the model learns to memorize the training data rather than generalize to new, unseen data.
- Training time: larger hidden layer sizes can also increase the training time since more parameters need to be learned.
- Complexity (computation): larger hidden layer sizes lead to higher computational costs during both training and inference.
- Generalization: smaller hidden layer sizes might lead to underfitting, where the model is too simple to capture the underlying patterns in the data, leading to poor performance on both training and test data.
- Interpretability and model complexity: smaller hidden layer sizes usually result in simpler models, which are easier to understand and interpret. This can be valuable, especially in fields where model interpretability is crucial.

Therefore, the choice of the hidden layer size is often a trade-off between these factors, and it might require some experimentation and validation on a separate validation dataset to find the optimal size for a given problem.