Assignment 02: Artificial Neural Networks

Tazmeen Afroz
Student ID: 22P-9252
Course: AI3003

May 26, 2025

1 Task 0

1.1 Dataset Analysis Questions

Q#	Question	Answer
1	How many features are in the inputs variable, and how much is its length?	Input Features :2 Length : 1000
2	What many output features are in the outputs variable, and how much is its length?	Output Features :1 Length : 1000
3	How many points belong to cluster 0, and how many points belong to cluster 1?	Cluster0 : 500 Cluster1 :500

Table 1: Task 0 Questions and Answers

2 Task 1: 0 Hidden Layer ANN Architecture

2.1 Network Architecture

The network consists of:

- Input layer: 2 features (x, y coordinates)
- Output layer: 1 feature (classification label)
- Weights matrix: (2×1)
- Bias vector: (1×1)

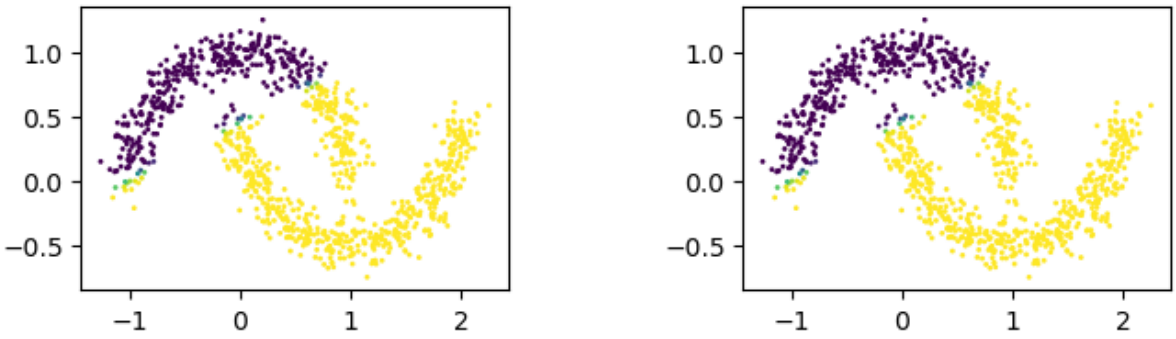
2.2 Task 1 Questions and Results

Q	Question	Answer
1	Change from the sigmoid activation to tanh. What is the effect on the scatter plot?	Tanh causes inconsistent scatter plots - sometimes shows proper class separation with distinct colors, other times shows all points as one color (complete misclassification).
2	Change from the tanh activation to relu. What is the effect on the scatter plot?	ReLU scatter plot shows all points classified as one class (single color) because ReLU outputs only non-negative values, failing to separate the two classes.
3	Change from the relu activation to softmax. What is the effect on the scatter plot?	Softmax scatter plot shows all points as one class because softmax is designed for multi-class problems with multiple outputs, not single-output binary classification.

Table 2: Task 1 Questions and Answers

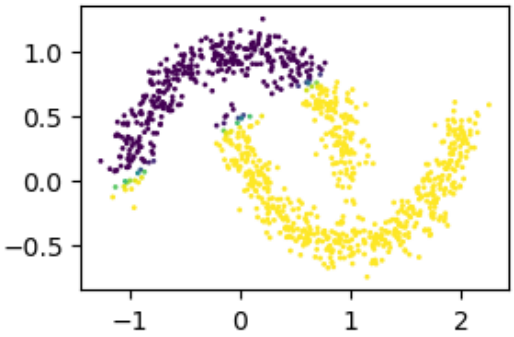
2.3 Learning Rate Analysis

What is the effect of changing the learning rate to 0.1, 0.01, 0.001, 0.0001? for the dataset,Lower learning rates (0.01, 0.001, 0.0001) result in similar minimum error losses around 0.116-0.121, while a higher learning rate (0.1) slightly reduces the error to 0.112 but may risk instability

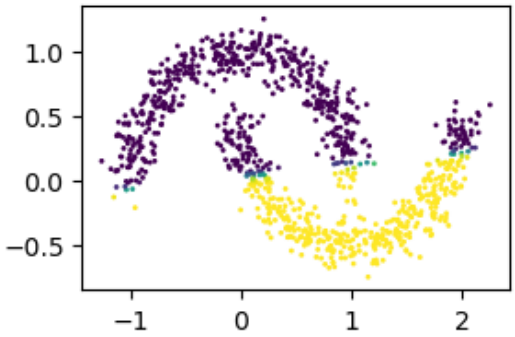


(a) Learning Rate = 0.1

(b) Learning Rate = 0.01



(c) Learning Rate = 0.001



(d) Learning Rate = 0.0001

Figure 1: Scatter plots for different learning rates

Learning Rate	Minimum Error Loss
0.1	0.11224506484000543
0.01	0.11579961983645566
0.001	0.12054599312544417
0.0001	0.11599062751930556

Table 3: Minimum error loss for different learning rates

3 Task 2: 1 Hidden Layer ANN Architecture

3.1 Network Architecture

The network consists of:

- Input layer: 2 features
- Hidden layer: Variable size (2, 4, 8, 16, 20, 24 neurons)
- Output layer: 1 feature

3.2 Task 2 Questions and Results

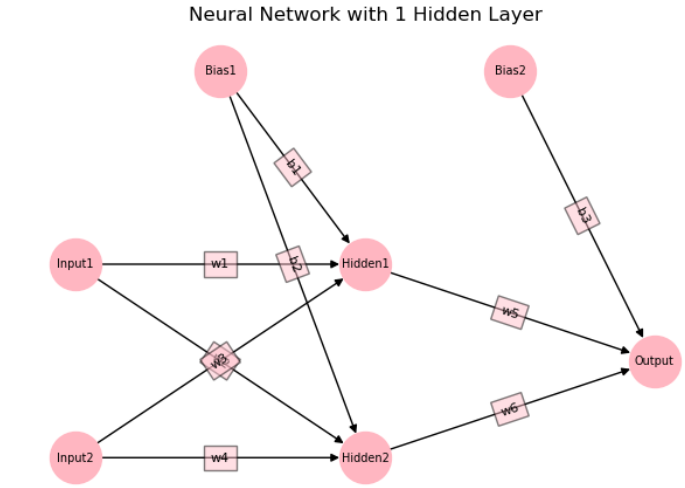
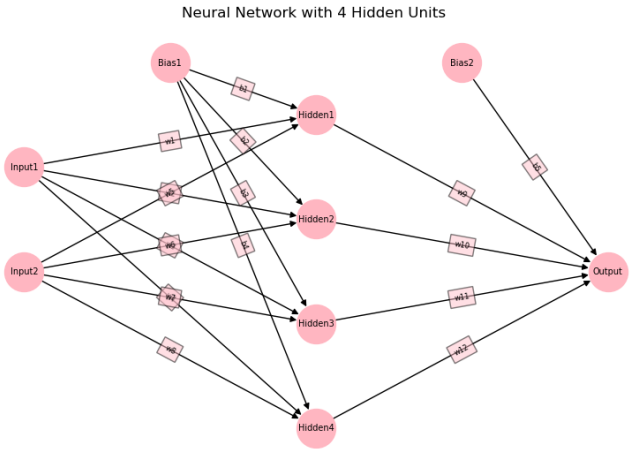
Q	Question	Answer
1	Explain why we have the combination of 2, 2 and 2, 1 for the weight matrices, and 1, 2, and 1, 1 for the bias?	(2, 2) for weights from 2 input neurons to 2 hidden neurons, (2, 1) for weights from 2 hidden neurons to 1 output neuron, (1, 2) for biases added to the 2 hidden neurons, and (1, 1) for the bias added to the 1 output neuron, with the first number indicating rows and the second indicating columns or terms.
2	Draw the network diagram of the ANN architecture, inclusive of bias.	
3	Is the current result of Task 2 better than Task 1? Answer on the basis of minimum error (cost_graph).	The previous error was 0.11, and it remained 0.11 with no major change.
4	Report whether Task 2 with larger ANN network (4 hidden units) is better than Task 1?	Yes, Task 2 with larger ANN network (minimum error 0.09) is better than Task 1 (minimum error 0.11)
5	Draw the network diagram of ANN architecture for the larger size (4 hidden units).	

Table 4: Table with ANN Details and Network Diagrams

3.3 Comprehensive Analysis Table

Learning Rate	Hidden Layer Size				
	2	**8**	**16**	**20**	**24**
1	0.1	0.09	0.09	0.09	0.08
0.1	0.1	0.1	0.1	0.1	0.09
0.01	0.08	0.08	0.0017	0.0006	0.003
0.001	0.09	0.08	0.08	0.08	0.08

Table 5: Minimum error (cost_graph) for different hidden layer sizes and learning rates

Minimum error reported from the table above: 0.0006

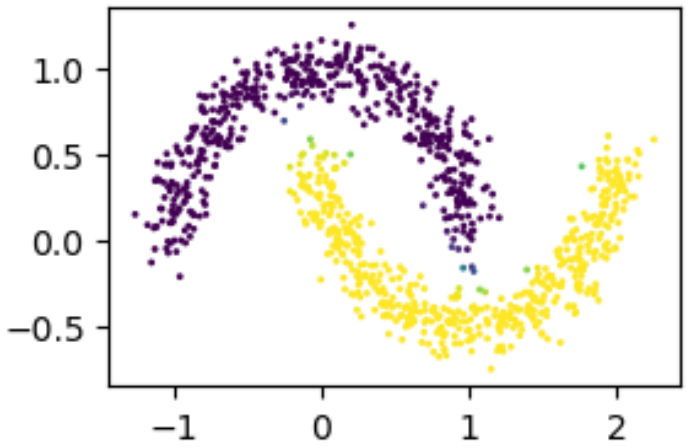


Figure 2: Scatter plot for the configuration with minimum error

4 Task 3: 2 Hidden Layers ANN Architecture

4.1 Network Architecture

The network consists of:

- Input layer: 2 features
- Hidden layer 1: 2 neurons
- Hidden layer 2: 2 neurons
- Output layer: 1 feature

4.2 Task 3 Questions and Results

Q	Question	Answer
1	Choose a learning rate that is larger than the one you found in Task 2. Then for that particular learning rate, find a hidden layer size that gives you same or better results than Task 2.	Learning rate 0.1 Hidden layer 1 size: 20,Hidden layer 2 size: 8 Min error from all combinations up to for h1-n-size in [2, 4, 6, 8, 16, 20, 24]: for h2-n-size in [2, 4, 6, 8, 16, 20, 24, 28]: got 0.07 which is not as better as 0.0006

5 Task 4

5.1 Task 4 Questions and Results

Q	Question	Answer
1	Which ANN architecture is best suited for this dataset?	A multi-class classification ANN with 2 input neurons, a hidden layer, and 3 output neurons using Softmax activation. (error : 1.66)
2	What if we think of the labels as the z-axis for the scatter plot? Should it work now?	Yes. Treating labels as z-values makes it a regression problem. Using 1 output neuron with linear activation reduced error to 0.4.

Table 6: Task 4 Questions and Answers