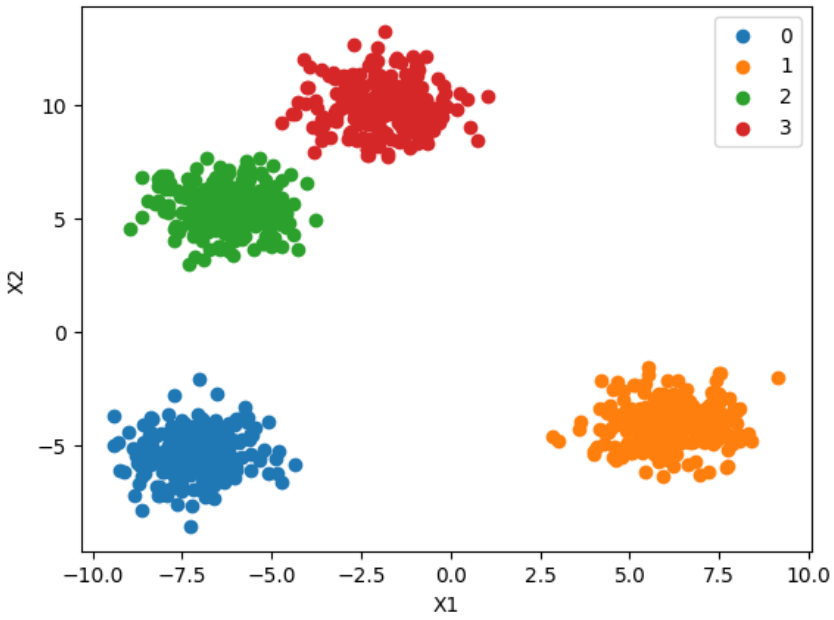
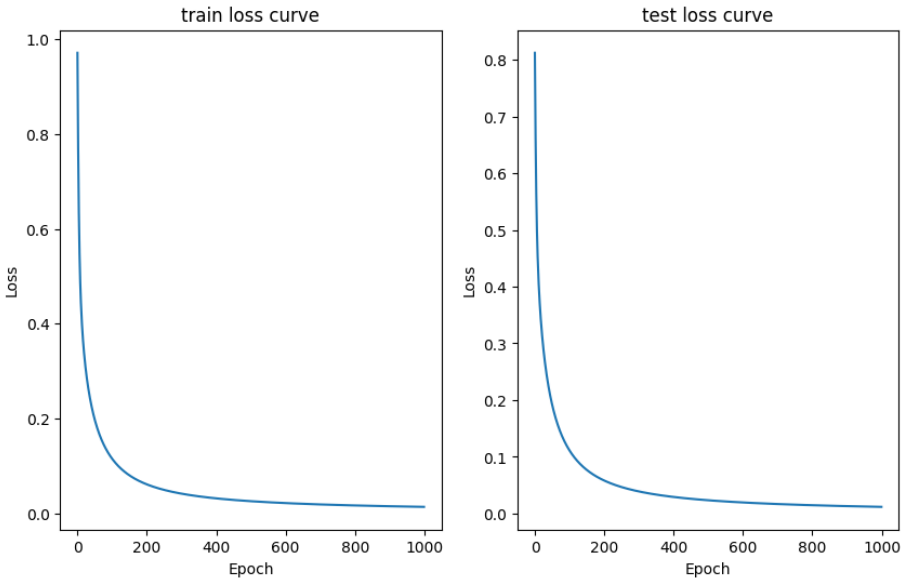
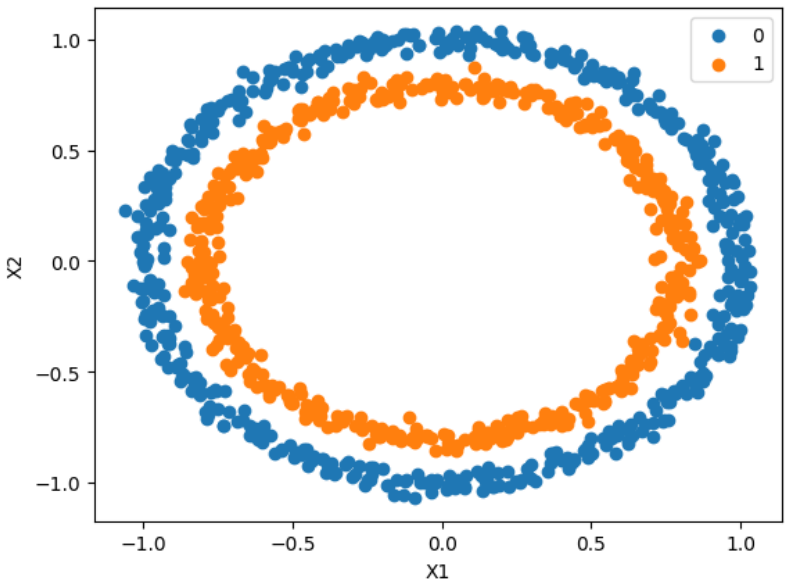
Report on the completion of the fifth task.

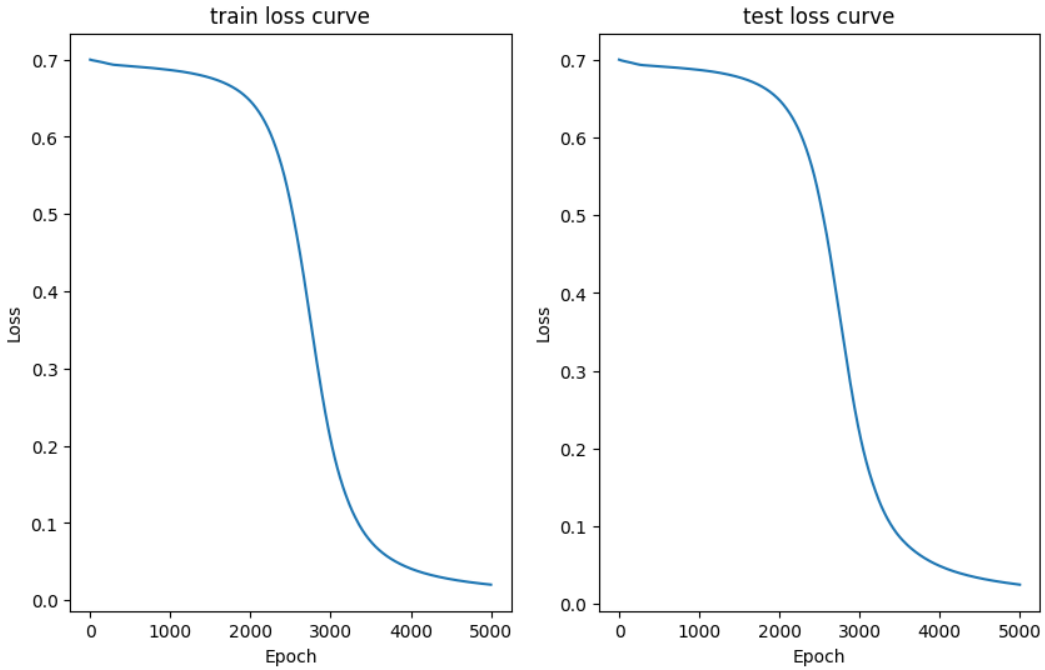
1. **The purpose of the task.**
   The main goal of this task is to investigate how does the different neural network parameters such as count of hidden layers and hidden units, learning rates and activation function affect the performance on linear and non-linear data. Try to reach the 100% accuracy on test sets for every model.
2. **Experiments.**
   a) **Experiment A – 4 blobs classifier (hidden layers=2, hidden units per layer=6, lerning rate=0.01, activation function=softmax, loss=CrossEntropy, epochs=1000) – max accuracy is 100%.**



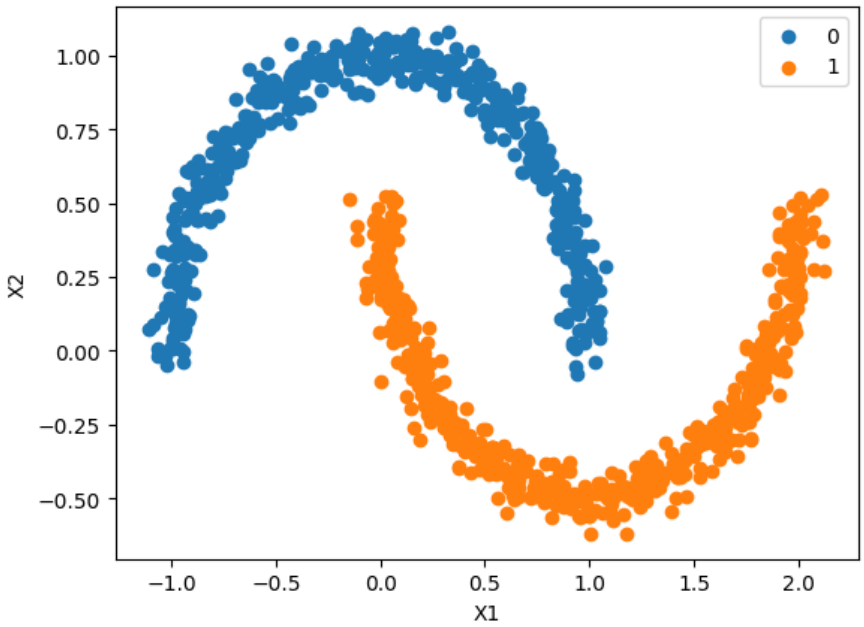


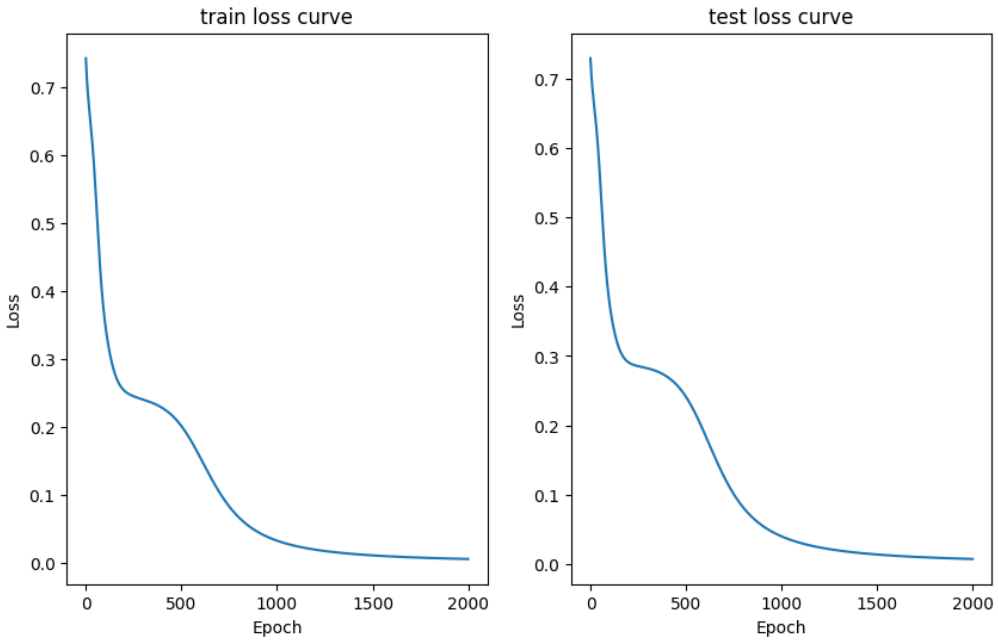
b) **Experiment B – 2 circles classifier (hidden layers=3, hidden units per layer=10, lerning rate=0.05, activation function=ReLU and Sigmoid, loss=BinaryClassEntropy (BCE), epochs=5000) – max accuracy is 100%.**





c) **Experiment C – 2 moons classifier (hidden layers=3, hidden units per layer=5, lerning rate=0.1, activation function=Tanh and sigmoid, loss=BinaryCrossEntropy, epochs=2000) – max accuracy is 100%.**





3. **Conclusion.**

-From these graphs we can conclude that non-linear data basically needs more epochs to gain a little loss and a high accuracy. This can be seen in the softness of the loss curves.

-ReLU and Tanh activation functions give the same performance for experiment B, but ReLU fails more on the experiment C and gives an accuracy of about 88%, while Tanh performs well with 100% accuracy and a loss less than 0.01.

-Experiment B required a specific learning rate of 0.05 because default rates of 0.01 and 0.1 were not able to come up with an ideal result and failed.

-Increasing the number of hidden units didn’t give significant changes in models’ performances. On the other hand, increasing the number of hidden layers did, but that required more time for training.