## Question 1. a)

This ANN uses the 194 training points dataset provided to solve the task of using an ANN to separate two spirals. We started with an implementation that used 5 layers which can be found in the 'question1a5layers.ipynb' file. We were able to successfully perform the task to varying levels of accuracy using this implementation.The number of epochs tested with this implementation varied from 5,000 to 200,000. We consistently found the models were more accurate when using epochs closer to the highest part of this range (200,000). We used a combination of error learning curves and visual inspection of the generalisation test applied to all pixels of a section of the x and y plane to evaluate our results.

The second implementation we used 3 layers and had better results than the 5 layer implementation. This ANN can be found in the 'question1a.ipynb' file. The model trained faster and was able to be more successful with less than half of the epochs of the 5 layer model. We found around 100,000 epochs to be the best.

Both implementations seemed to work best using a gradient descent optimizer with a learning rate of about .5. Values both lower and higher were tested and gave less accurate results.

### Three Layer Model

I used a three layer artificial neural network to solve this problem. It uses the 194 data points sample provided in the labs and required for the assignment. I varied the learning rates, epochs and step sizes to find the optimal results. I found the step sizes to work best at around 0.1. Step size of 0.5 generally worked the best and epochs of 50,000 to 200,000 to work the best. More than this the training time is very high and overtraining can occur.

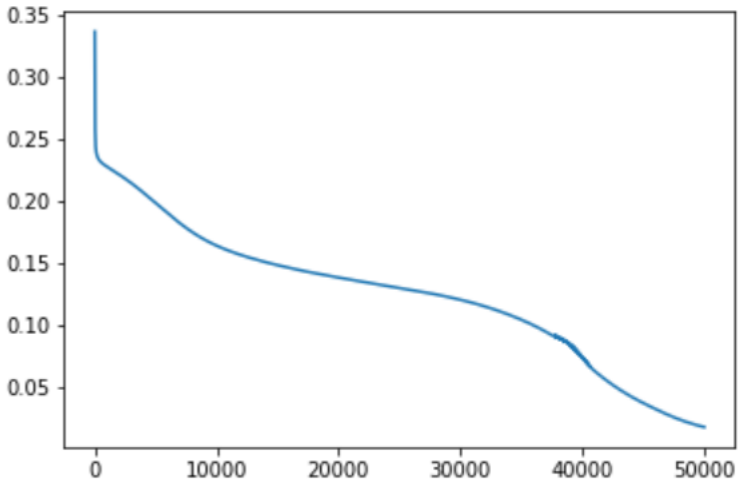
---

Learning rate - .5

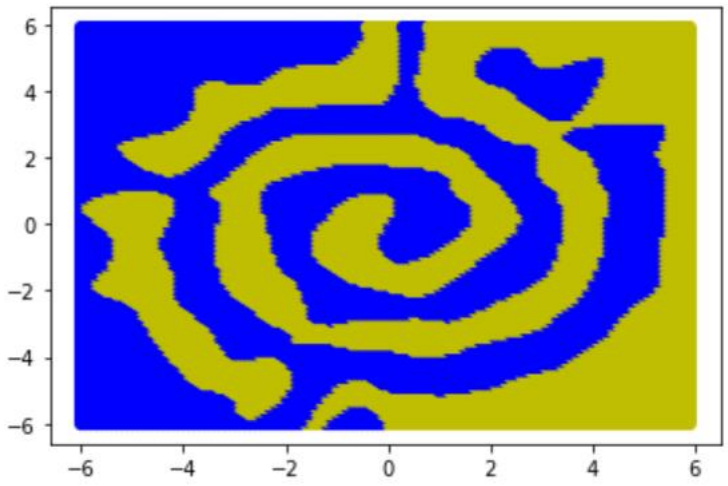
Epochs – 50000

Step size 0.1

Error Learning Curve



Spirals



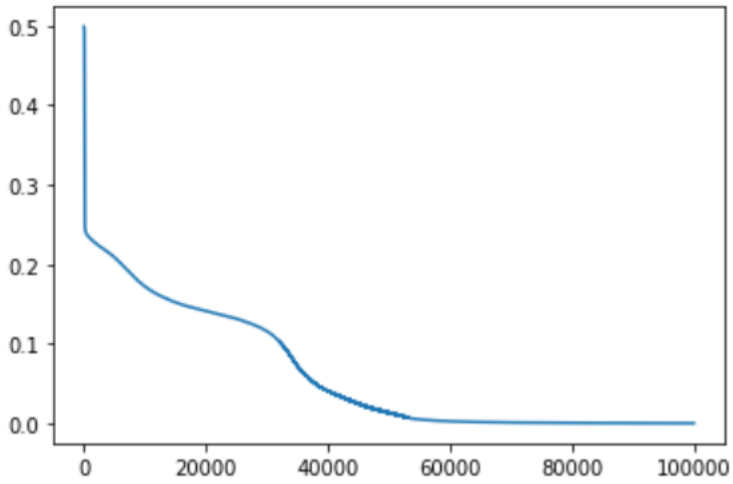

---

Learning rate - .5

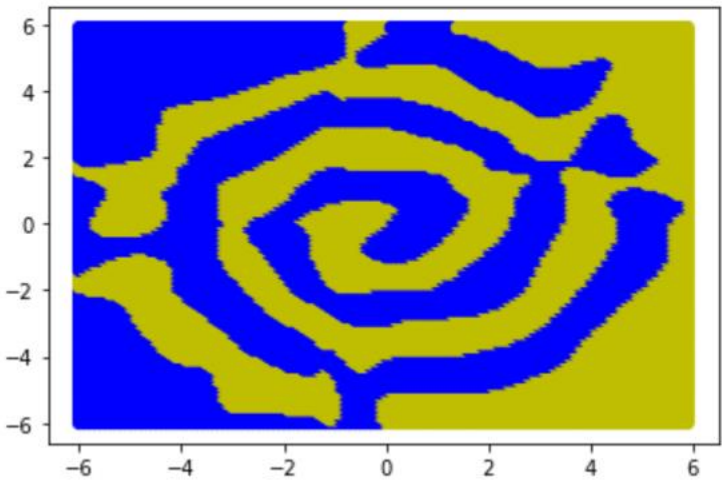
Epochs – 50000

Step size 0.1

Error Learning Curve




Spiral




**Five layer Model**

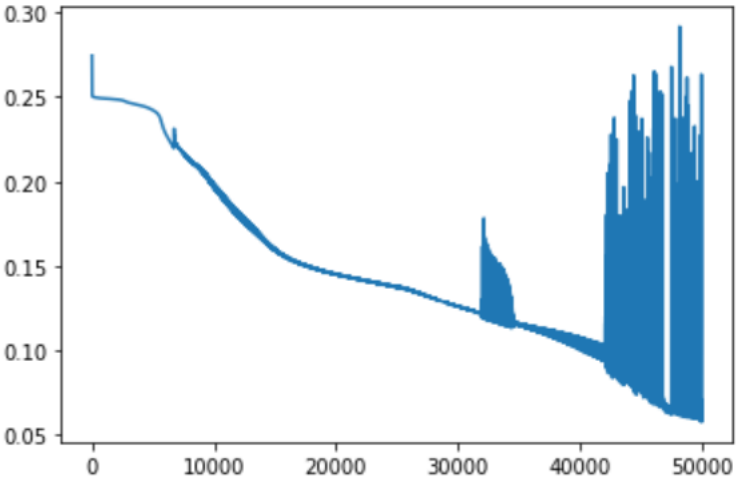
A neural network similar to the three layer model except it uses five layers. Requires higher epochs to get working. This causes longer training times. Doesn't perform as well as the three layer model.

Learning rate - .5

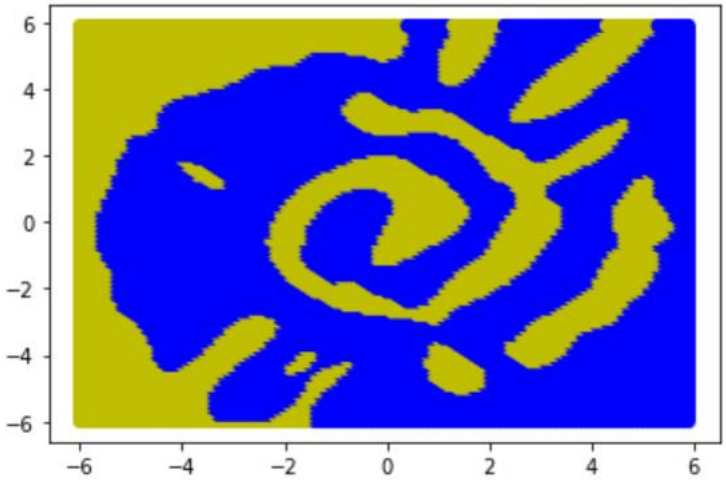
Epochs – 50000

Step size 0.1

Error Learning Curve



Spiral

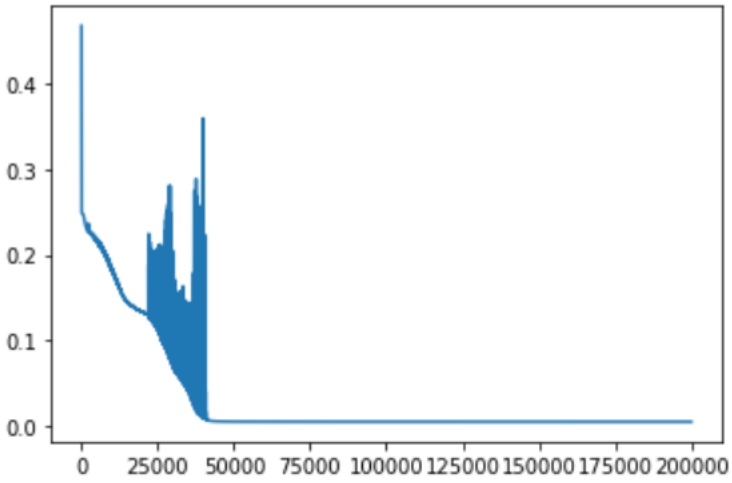



Learning rate - .5

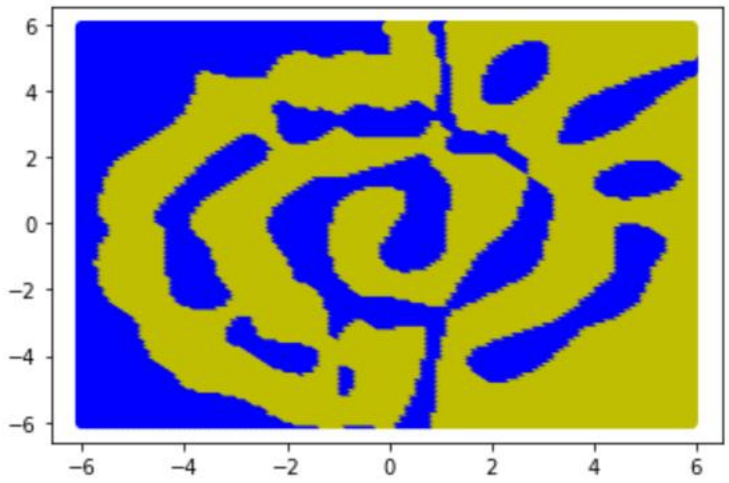
Epochs – 100000

Step size 0.1

Error Learning Curve




Spiral

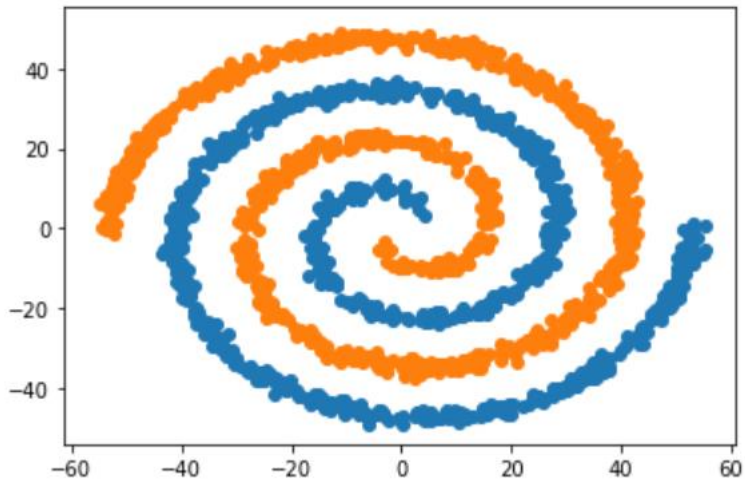



**Question 1. b)**

The file for the implementation of this model is called ‘question1b.ipynb’. To generate the variation of the two spiral task, we used a modified version of the spiral generator given in the assignment specification. During the course of creating this model, we found that the quality of the data provided has a large impact on the success of the model. To minimise problems, the spirals we generated contained 800 data points (less wouldn’t be enough and more would increase training time). We also found that the lower the range of randomness for the data points of each spiral, the more easily we were able to create a successful model.

I used a spiral made up of data points that I generated based on a variation of the provided spiral generation algorithm. The tighter the range of values (tighter lines in spiral) and more data points increases the success of the model. This variant requires larger epochs and training times. Learning rate of 0.1 worked the best.

Generated spiral

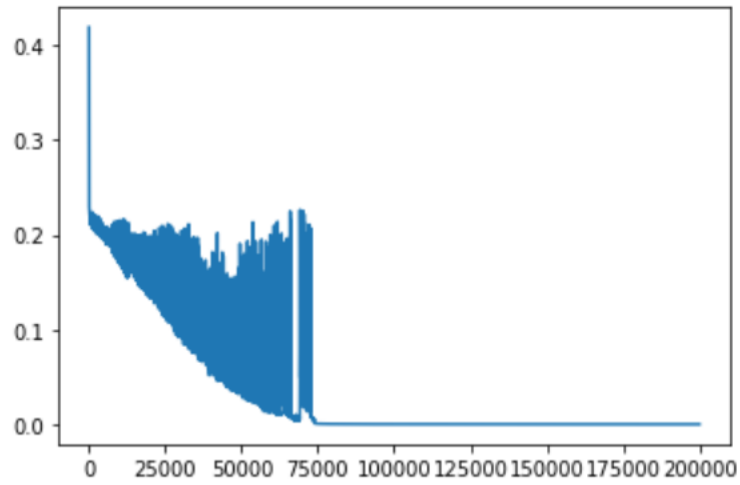



Learning rate - .5

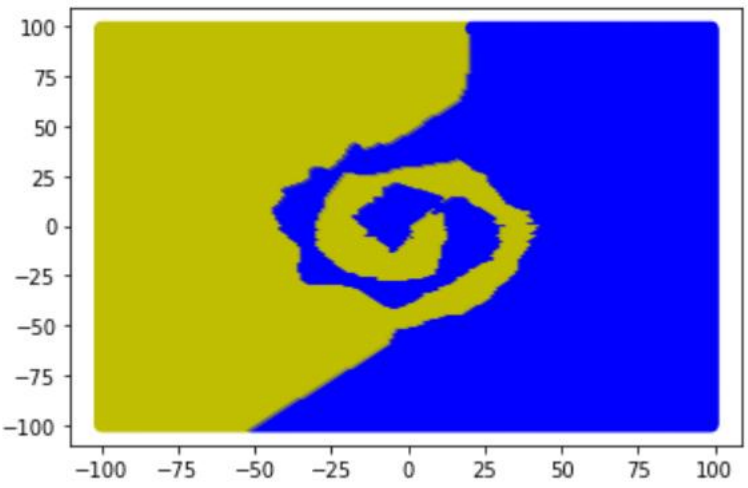
Epochs – 200000

Step size 0.1

Error Learning Curve




Spiral

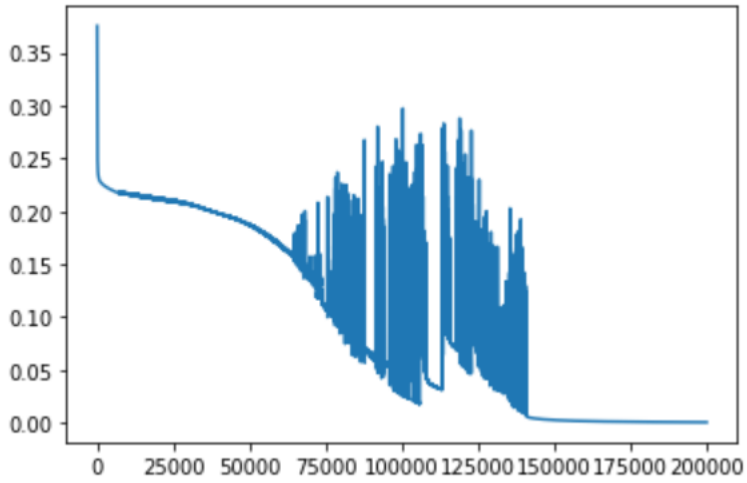


Learning rate - .1

Epochs – 200000

Step size 0.1

Error Learning Curve




Spiral