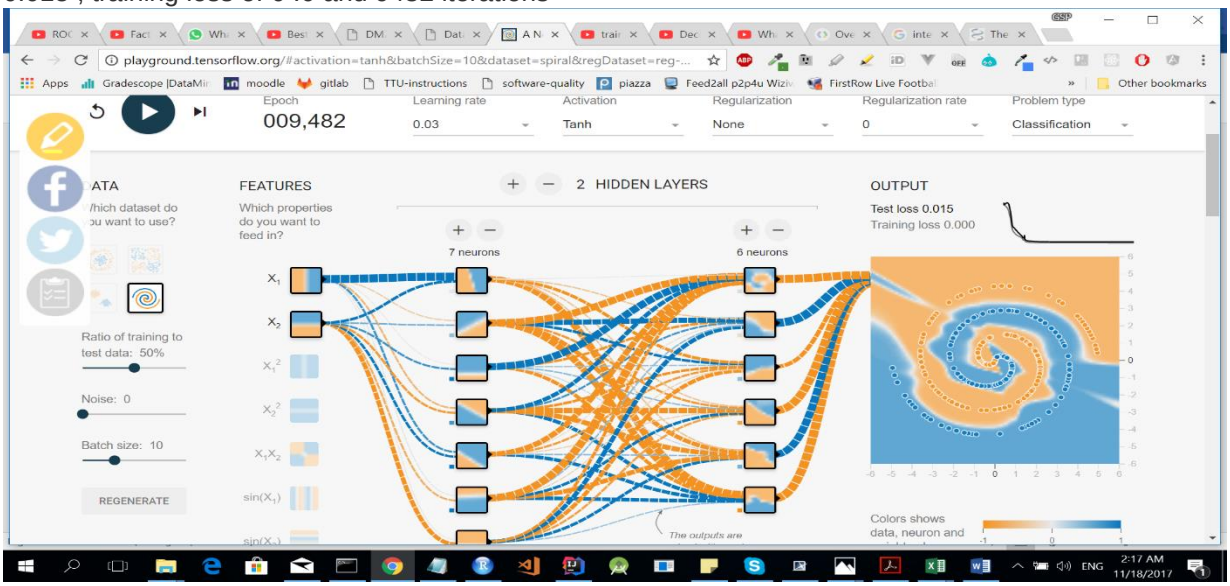
Exe1

(a) ***Describe your search process in finding the architecture (combination of neurons and layers) which***

I realized that while increasing the number of neurons in the hidden layer there was an increase in performance i.e. the test loss and training loss kept going and low .

b

To solve the task , I had 2 hidden layers with 7 neurons on the left and 6 on the right with a test loss of 0.015 , training loss of 0 .0 and 9482 iterations



#c. ***What do the colors orange and blue stand for in: (c1) connections between neurons; (c2) inside neurons; (c3) inside the output image (the one with a spiral).***

In c1 connections between neurons, blue shows a positive weight, which means the network is using that output of the neuron as given. An orange line shows that the network is assigning a negative weight

Inside neurons orange shows negative values while blue shows positive values.

Inside the output image the blue and orange color indicate which network is predicting a particular area

d# Which input features have you chosen for your architecture? Why are they important for your network to solve the problem*?*

I have chosen to increase the number of ***neurons in the hidden layer.*** Increasing the number of neurons is important because when many neurons are created and connected together they strengthen the connections that lead to success and diminishing the failure rate.

Exe2

- Each classifier consists of three components i.e. **representation, evaluation and optimization.** Representation means classifier must be represented in a formal language that a computer

understands.
Evaluation is used to distinguish between good classifiers from bad ones. (page 79)

- If building a classifier, it's better to set aside some of the data at the beginning and only use it to test your classifier at the very end, followed by learning your classifier on the whole data set.

- The success or failure of machine learning projects largely depend on the features used. Learning is easy if you have many independent features that each correlate well with the class. On the other hand, if a class is a very complex functions of the features, it may be hard to learn it. (page 84)

- If you have constructed the best set of features, but the classifiers you receive are not accurate enough. Then there are two main choices you can do either to design a better learning algorithm or gather more data. (page 84)

- As a rule of thumb a dumb algorithm with lots and lots of data beats a clever one with modest amount of it. (page 84)

- As a rule of thumb a dumb a logarithm with lots and lots of data beats a clever one with modest amount of it.

- Learn many models not just one. Researchers found that instead of learning many models and selecting the best one, its better if we combine the variations, the results are much better. (page 85)

- Simplicity does not mean accuracy. Simpler hypothesis should be preferred because simplicity is a virtue in its own right, not because of hypothetical connection with accuracy. (page 86)

- Representation does not imply learnable. Just because a function can be represented does not mean it can be learned. For example, decision tree learners cannot learn a tree with more leaves. (page 87)

Exe3