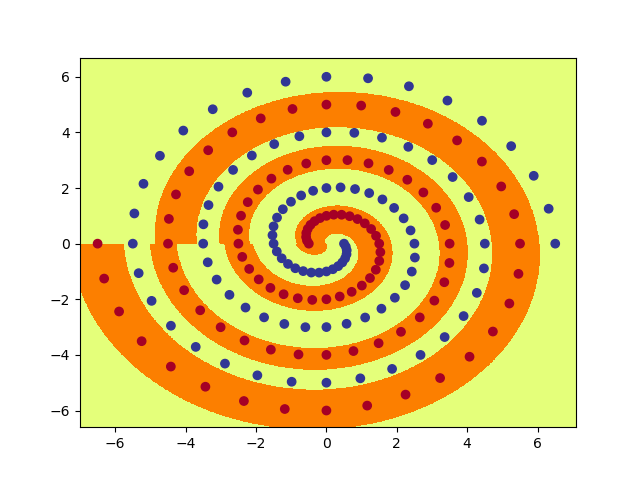
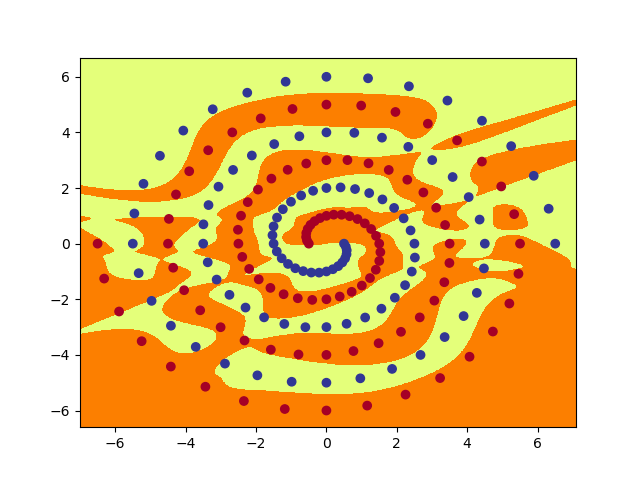
The three models constructed very different hidden layers which vary from shape and usage. The first layer of Polar net has distinct hidden value comparing to the other two because of the unique coordinate system used. It generates spiral shaped masks while others have linear separated hidden layers. Another difference is that PolarNet generates the final answer directly from the first hidden layer by overlapping them while RawNet and ShortNet need to first combine those linear layers to form some curves in second layer to cover the spiral shape. There is no significant difference between the layers of RawNet and ShortNet.

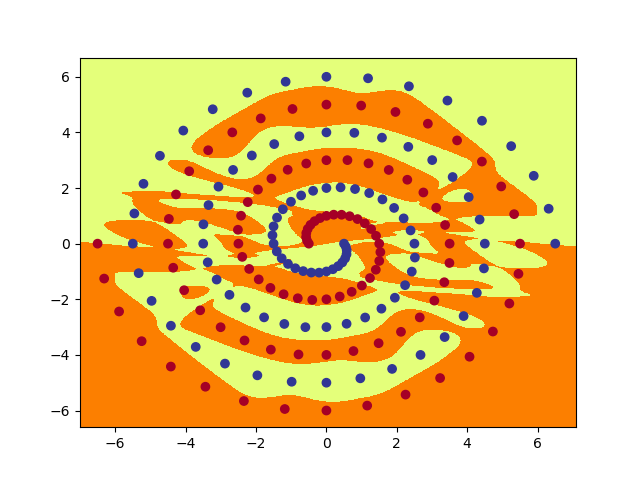
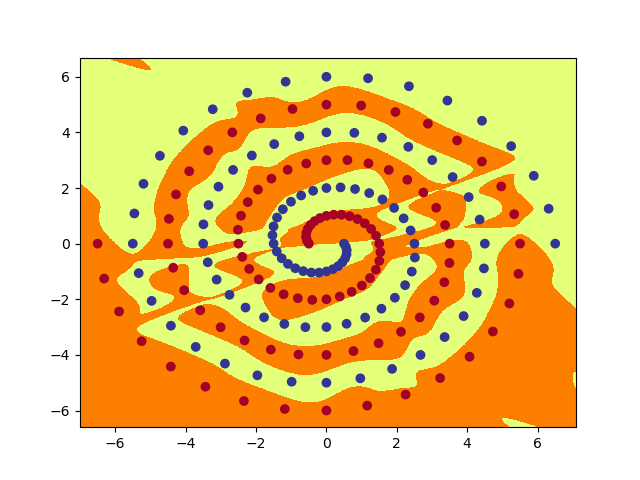
Some experiments were made on initial values for RawNet and ShortNet. It is found that initial value of 0.2 gives quite satisfying performance with 7000 expected epochs. Although sometimes it stuck on some point and epochs needed approach to 20000, it is still much better than those values in both directions. A larger initial weight could save time for approaching to optimal answer however the network may go far away if the initial weight is too large.

Apparently, the output of PolarNet is more natural than the models using Cartesian coordinate system. Meanwhile, the naturalness of the output of ShortNet is relative higher than that of RawNet. The reason for high naturalness of PolarNet is the spiral shape which can be represented approximately in a linear form with polar coordinates. Thus, the regression for this model is more likely to approach the real function of the spiral shape. However, only approximate shape can be imitated by networks using Cartesian coordinates. Representation is regarded as a fixed “layer” to filter data which is designed by human, with an appropriate representation, the model can fit the samples with higher accuracy. The purpose of elaborating representation of data is to cover as much features which are fixed, obvious and do not require machine learning to find out as possible. With a pertinent representation, elevation of both accuracy and speed of a model is possible.

In the end, some experiments were made on the parameters and architectures. The first finding is that larger batch size does not always gives a better performance. Typically, increase of batch reduces time needed and may increase the efficiency of learning. 1, 97 and 194 were chosen as batch size, and it is found that when batch size is 1, the training process is considerably slow. Even it is faster to train each epoch when 194 is chosen, the efficiency is not satisfying: data can not be fitted within 20000 epochs. So, choosing an appropriate batch size can lead to higher accuracy and speed. Also, different activations were tested and it is found that ReLU is not suitable for this task. Because the spiral shape is smooth and divided equally, and Tanh fits the features well. However, the shape of ReLU is too sharp and it does not give symmetric output, some points would be hard to be fitted as it may be ignored by ReLU.

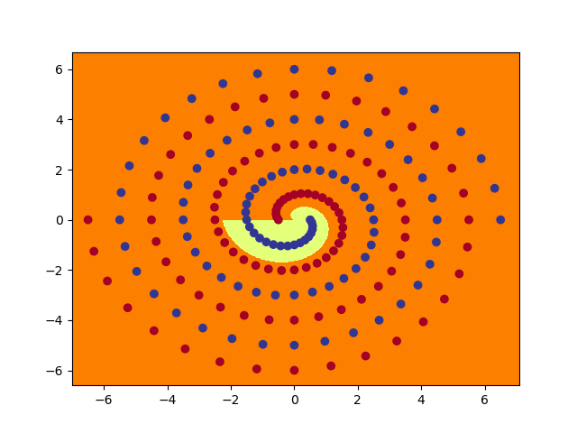
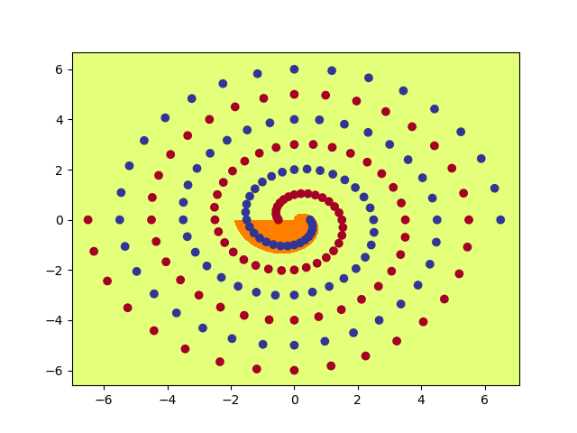
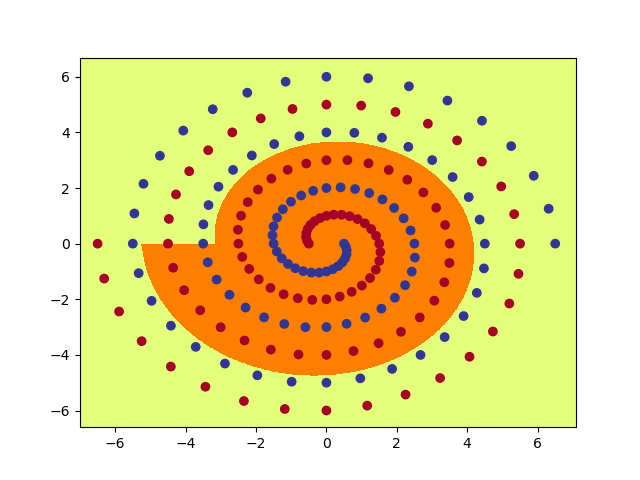
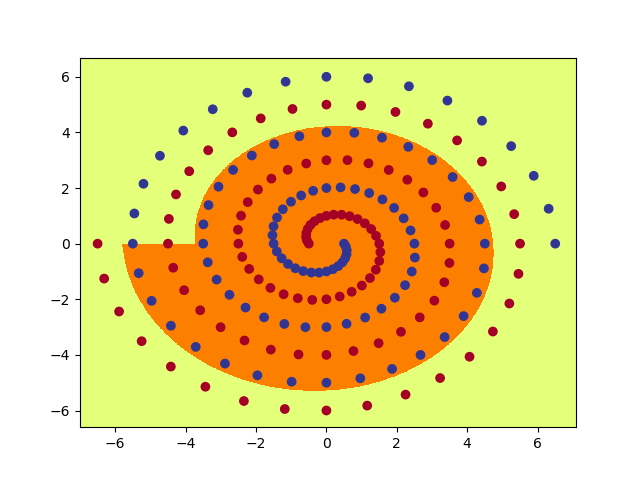
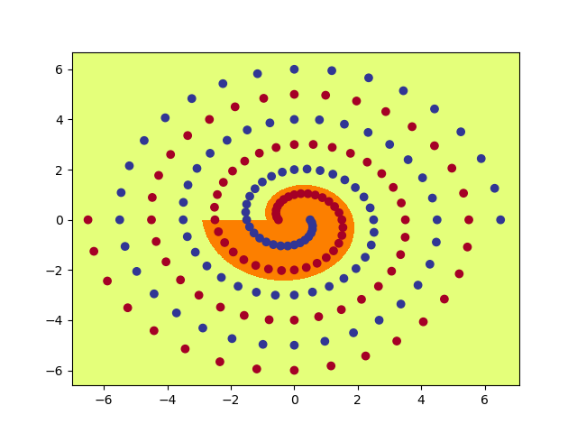
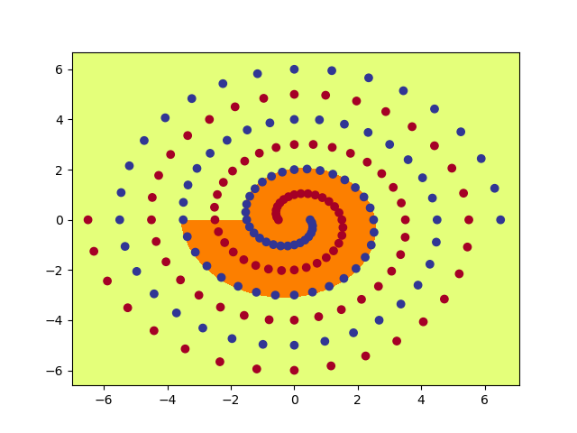
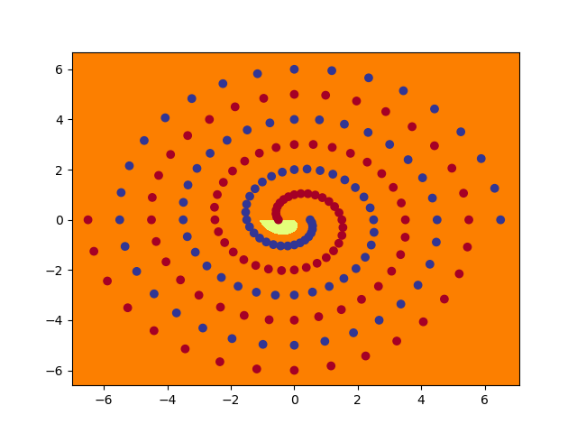
Appendix: Output Functions

Output for PolarNet

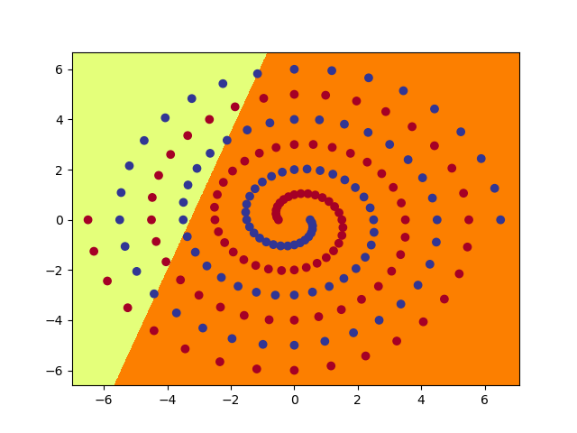
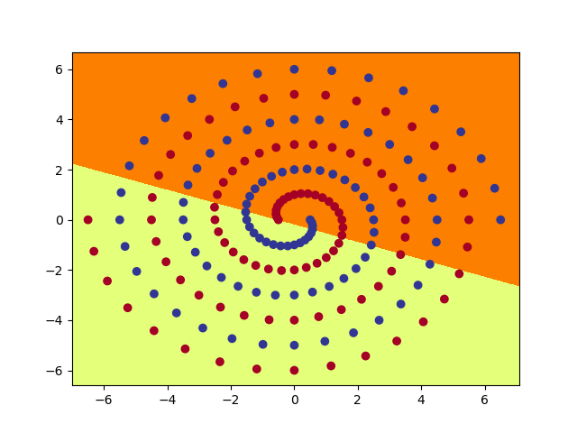
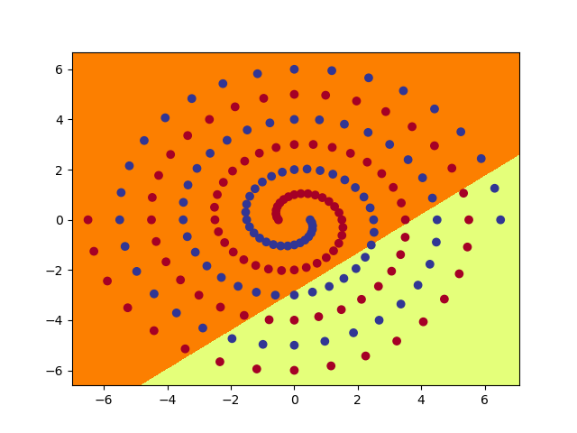
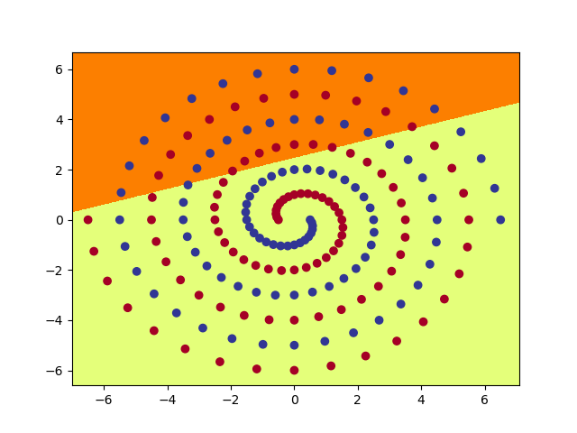
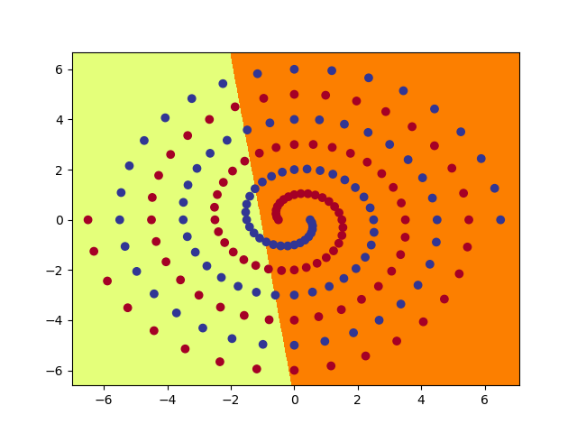
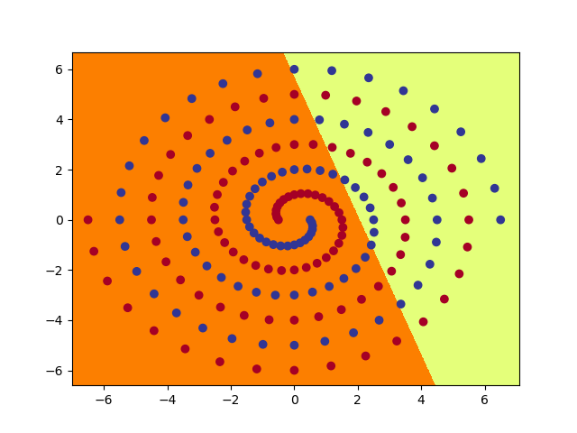
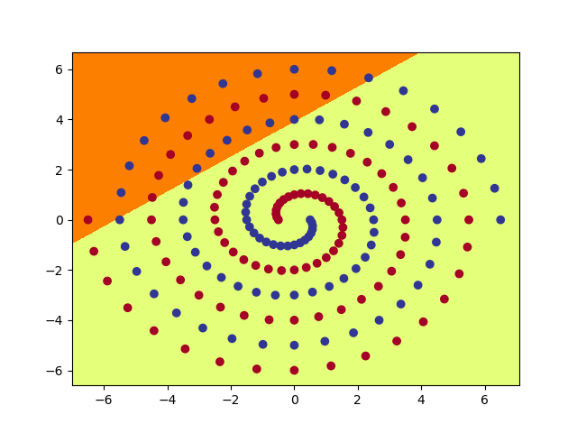
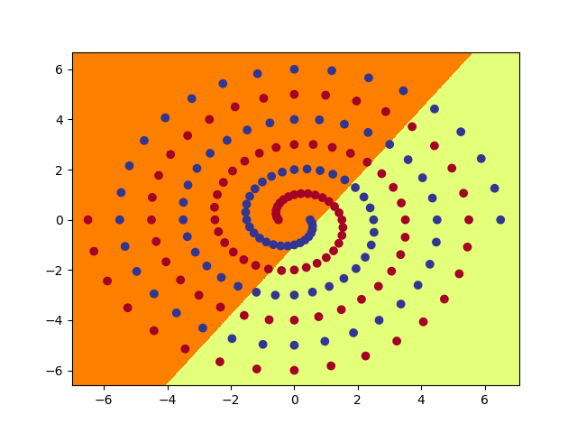
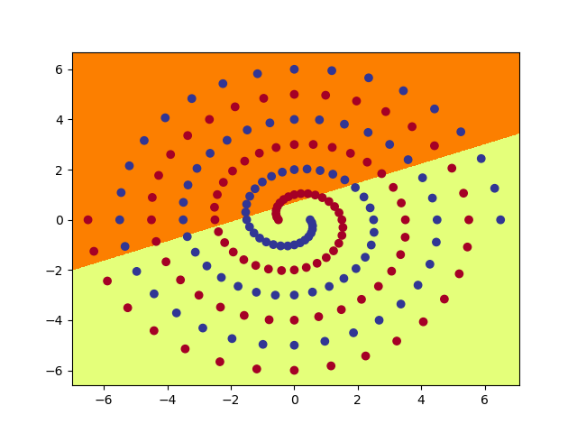
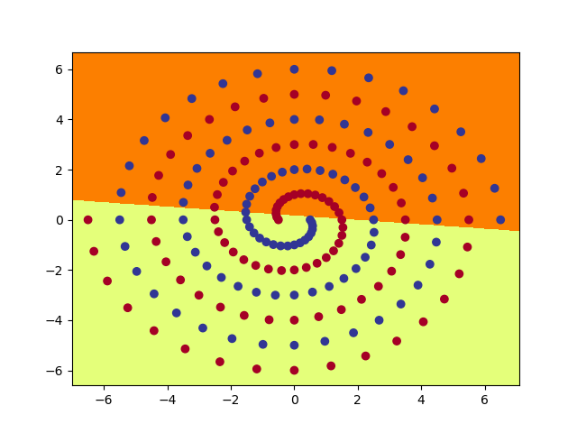
Output for RawNet

Output for ShortNet

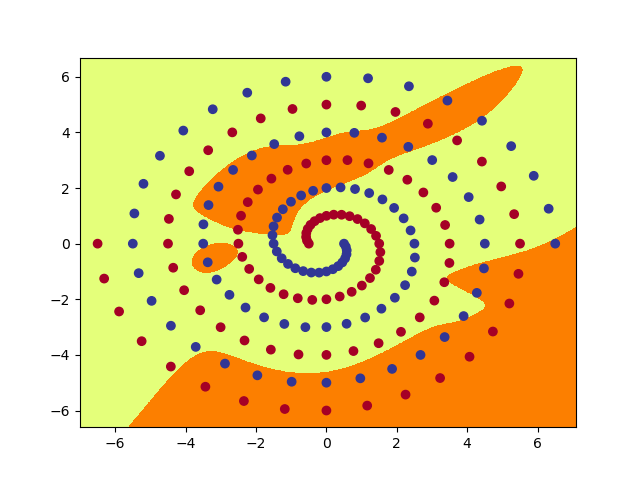
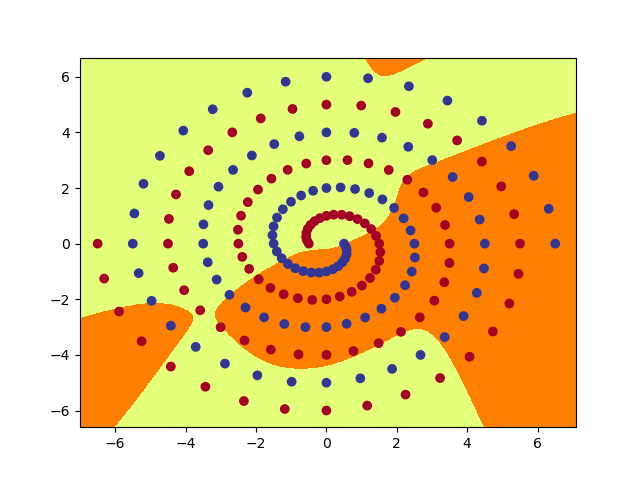
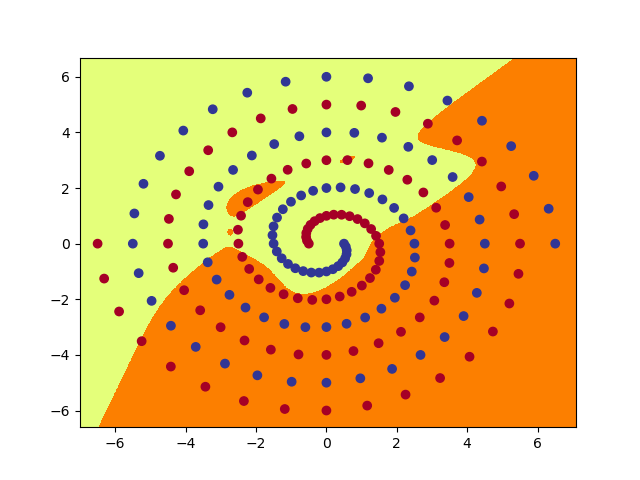
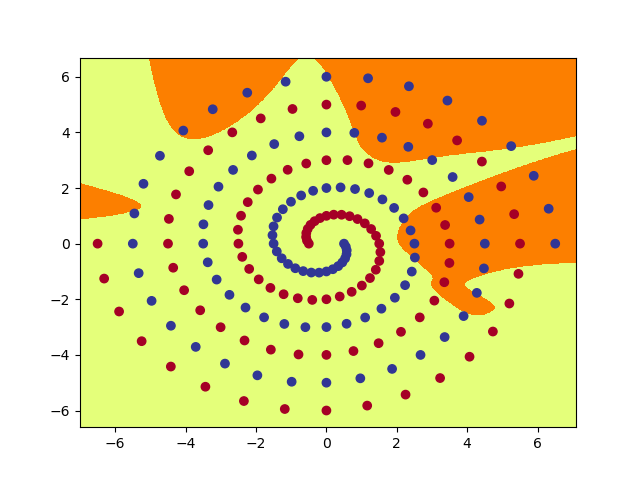
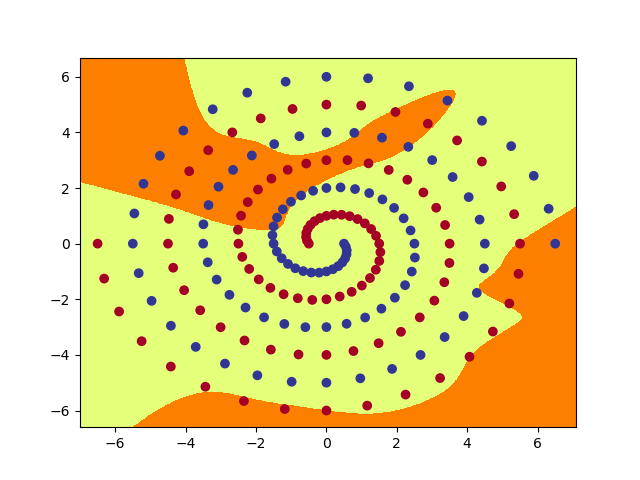
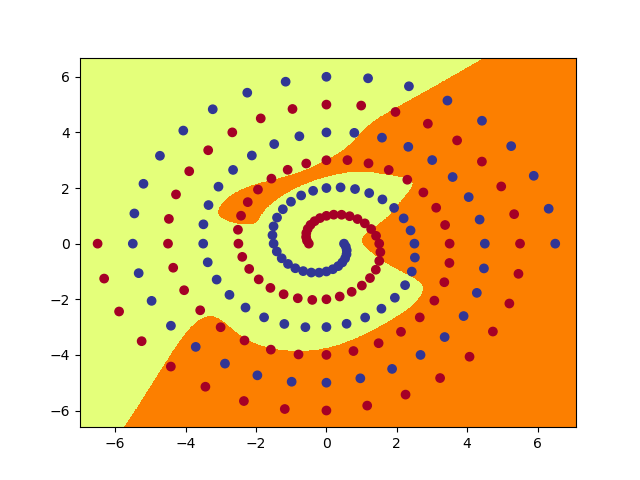
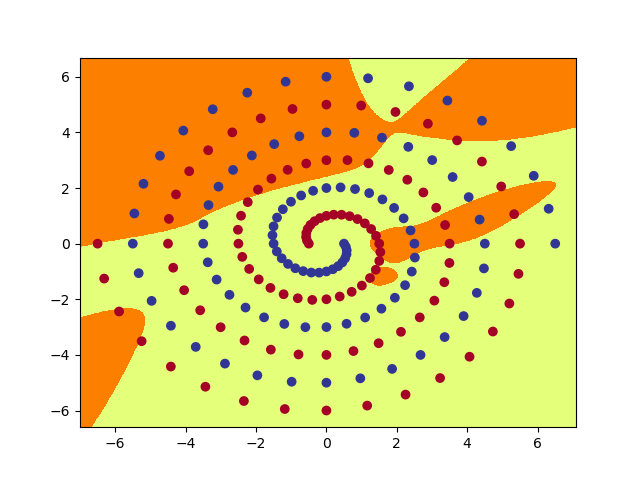
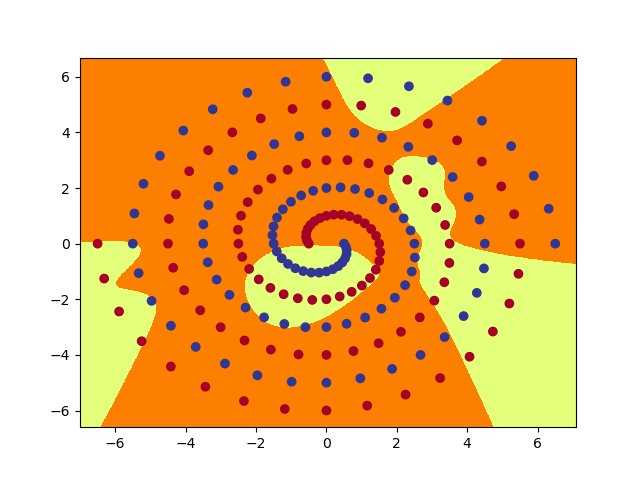
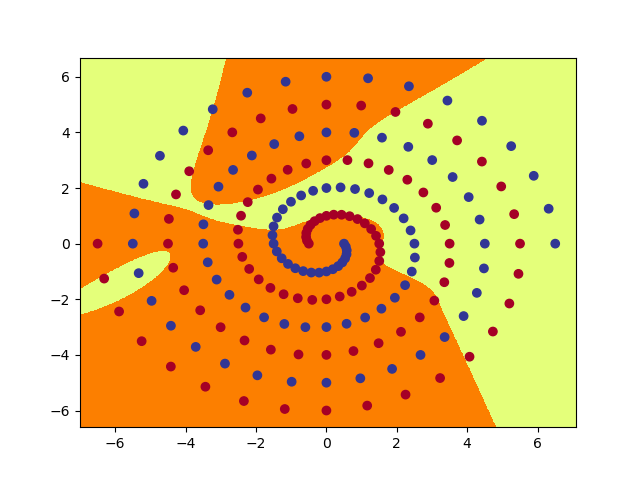
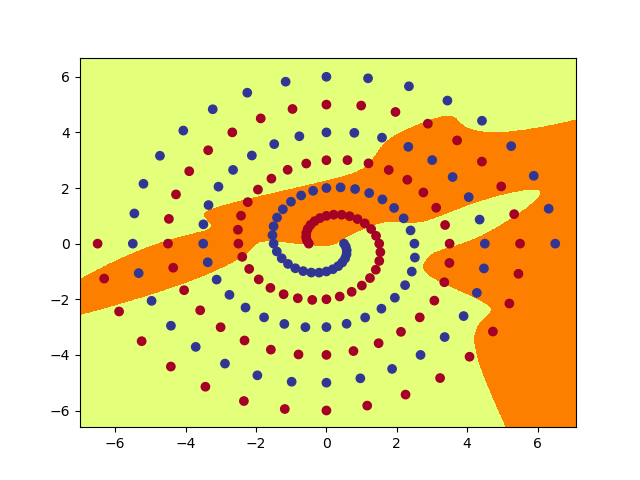
Appendix: Hidden Layers



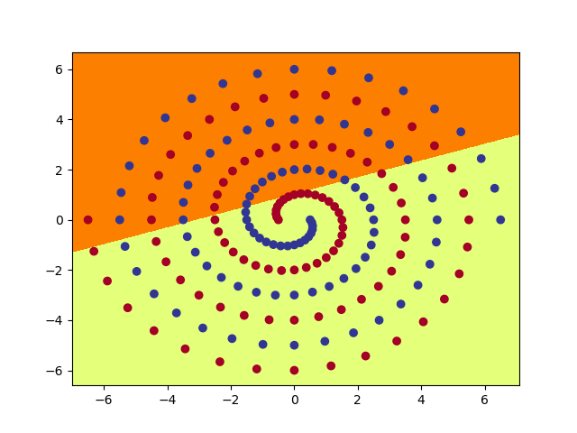
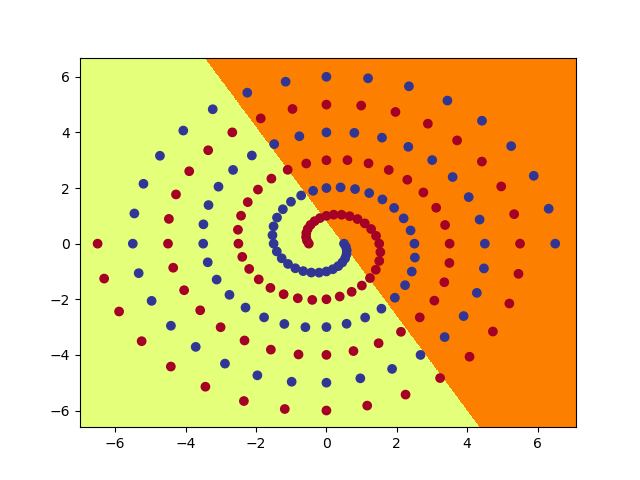
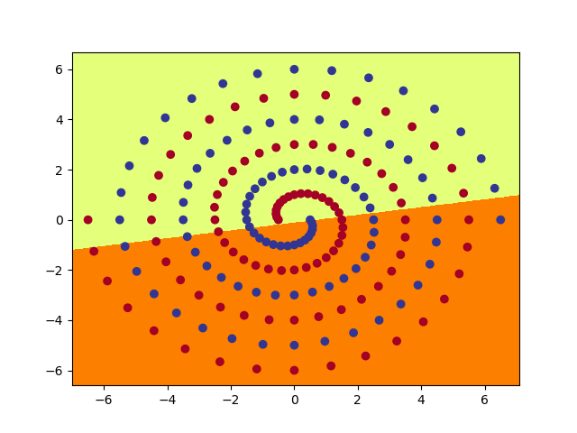
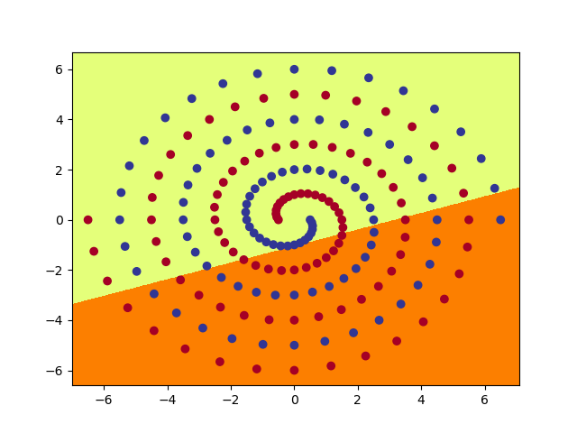
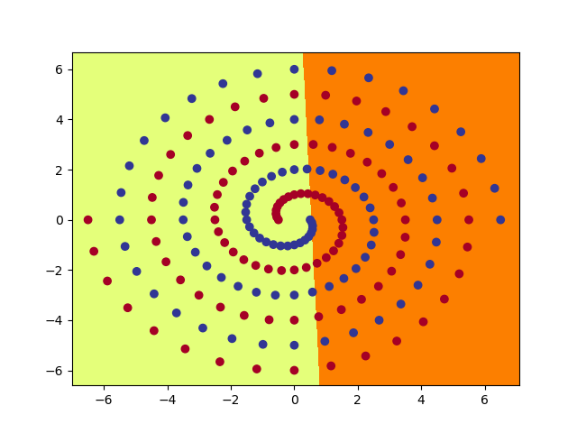
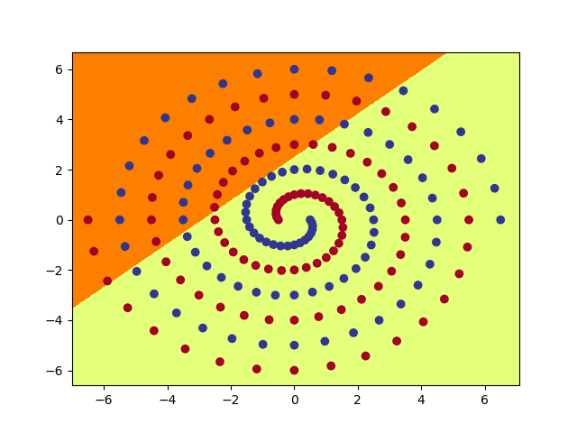
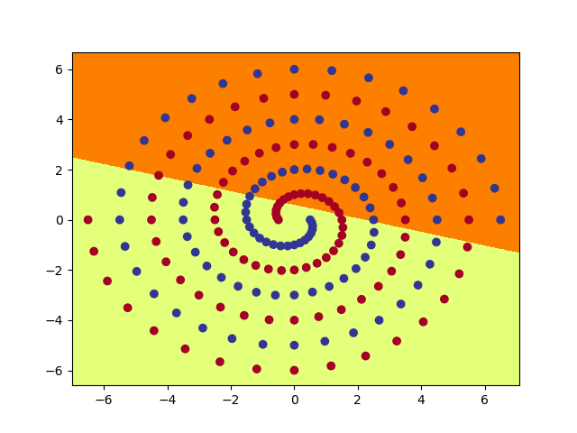
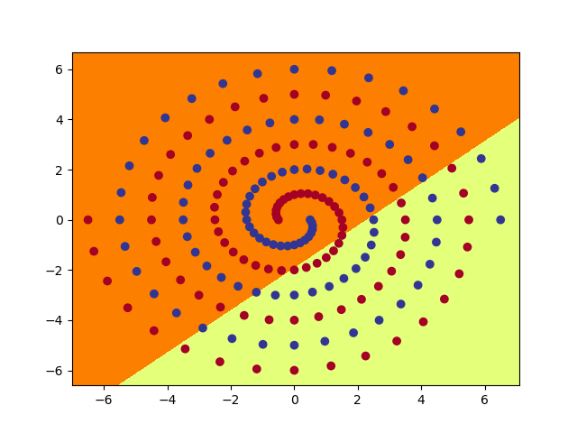
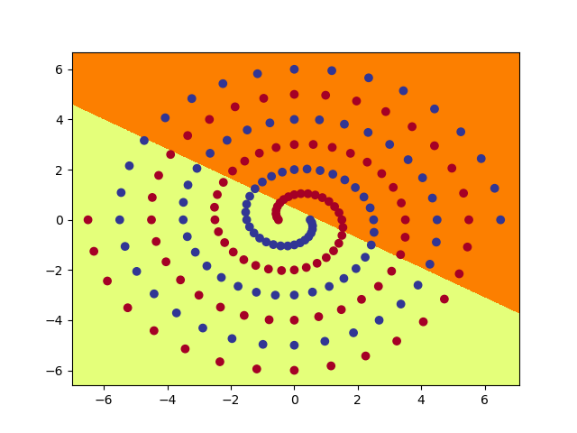
Hidden Layer for PolarNet



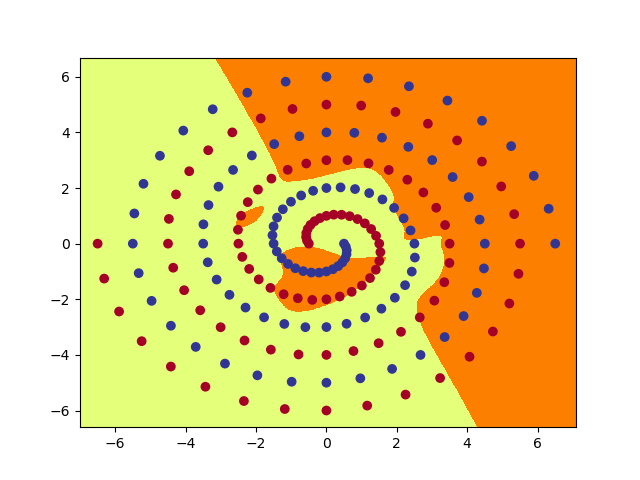
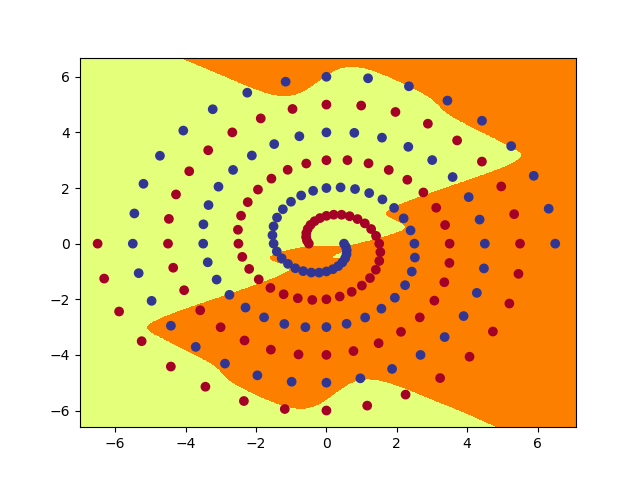
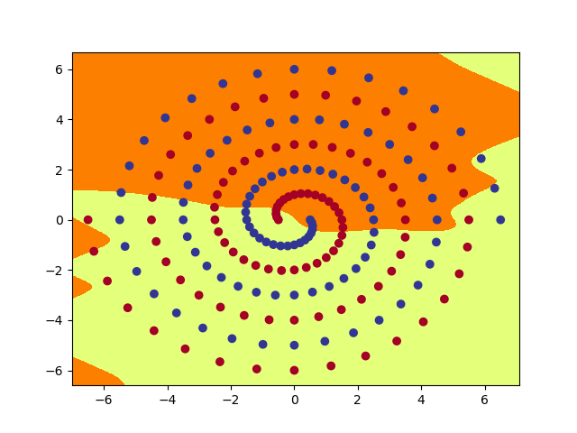
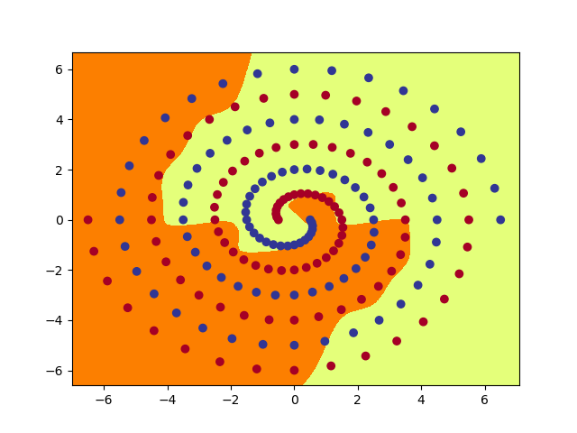
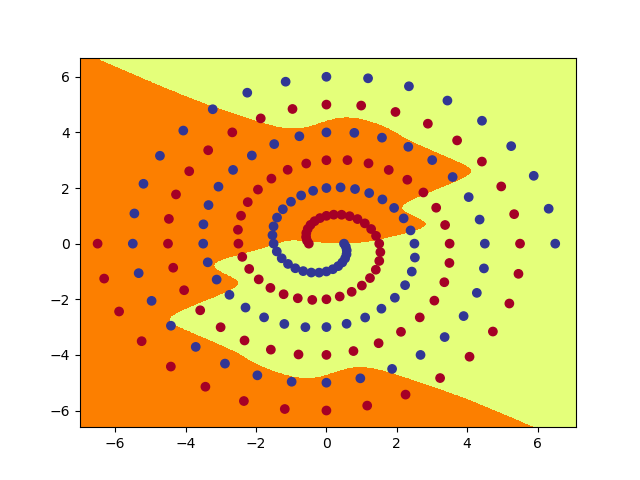
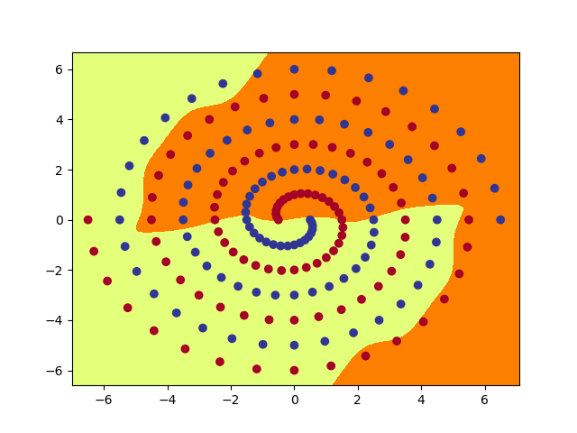
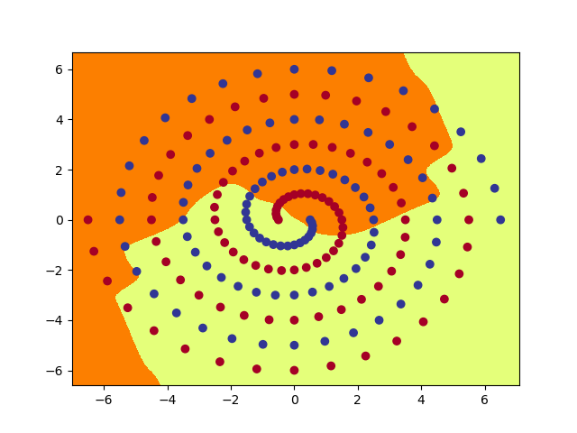
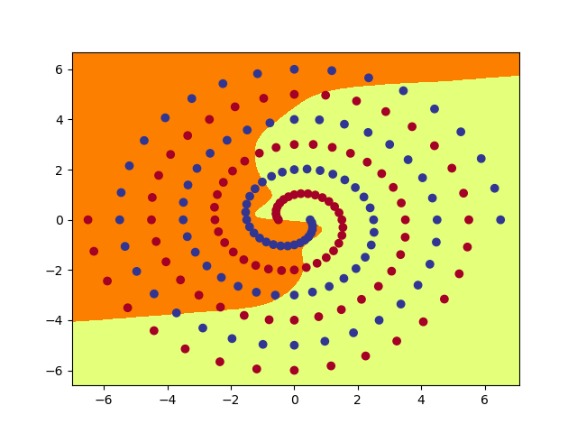
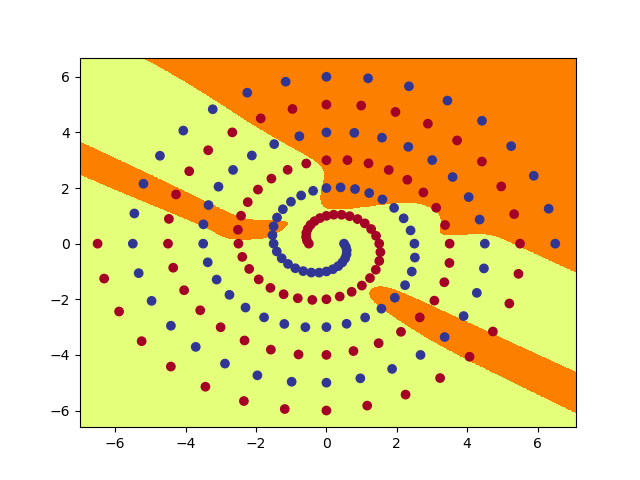
Layer 1 for RawNet



Layer 2 for RawNet



Layer 1 for ShortNet



Layer 2 for ShortNet