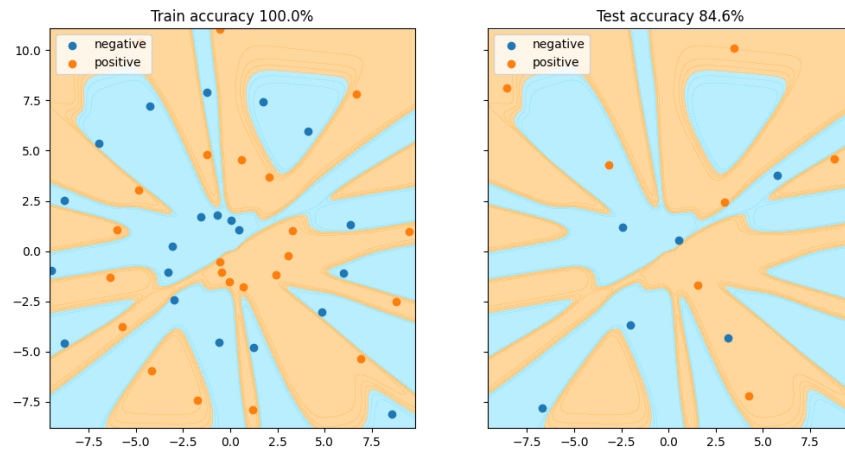2.   a.   Could not get it working, but I would convert the Cartesian coordinates to polar coordinates. From there I would then I would plug the modified data and compare the distance of the data to points from into spiral equation to make a classification.

b.

Figure 1: N_iters = 10000, learning_rate = 1, alpha = 1, penalty = "l2", layer dimension of 32
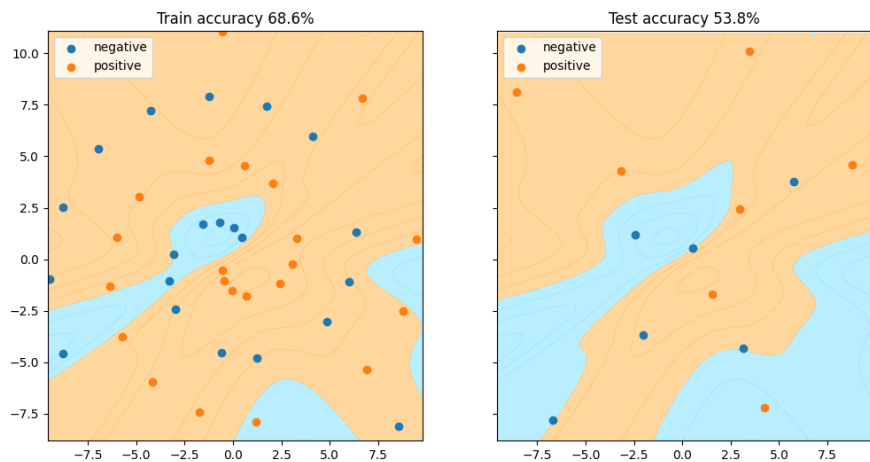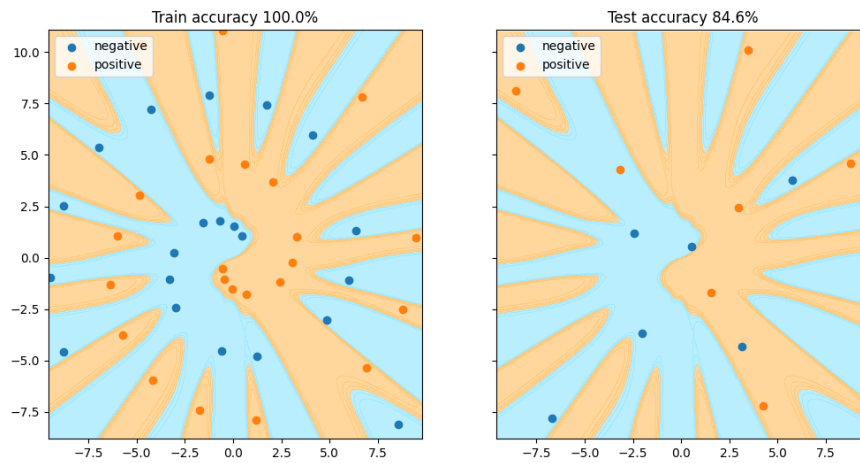

Figure 2: N_iters = 200, learning_rate = 1, alpha = 1, penalty = "l2", layer dimension of 32

Number of iteration, learning rate, had no effect. Penalty L2 was more accurate than L1. Alpha also effected the fit of the model with 1 being the most accurate fit.

Train accuracy 100.0% • Test accuracy 84.6%

c.

Figure 3: most accurate plot

N_iters = 10000, learning_rate = 1, alpha = 1, penalty = "l2", layer dimension of 128 were the hyperparameters of the most accurate plot that I could find. Alpha had the most effect on the accuracy of the test.