

Abhishek Kumar Srivastava

Student Id: 861307778

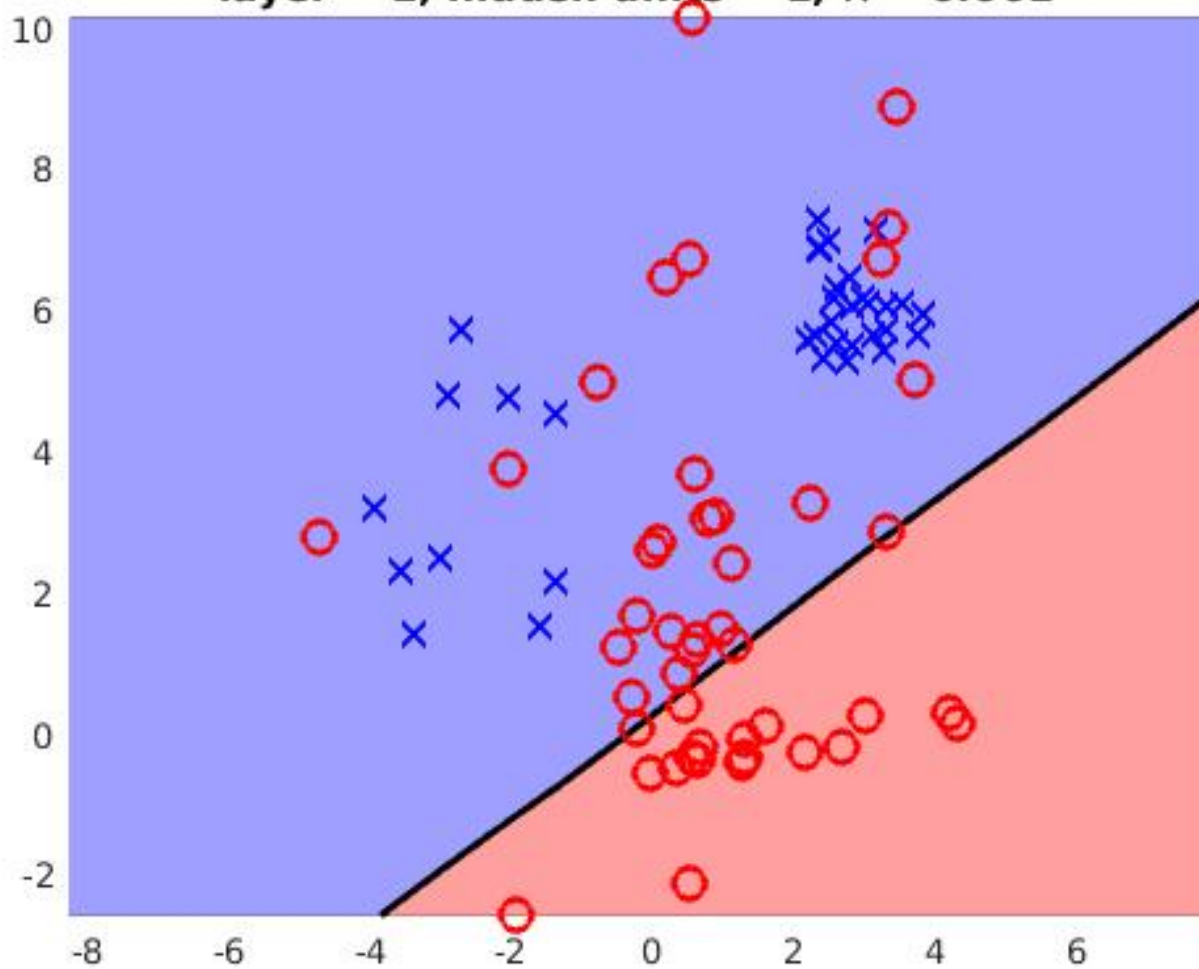
November 8, 2016

CS 229 Machine Learning

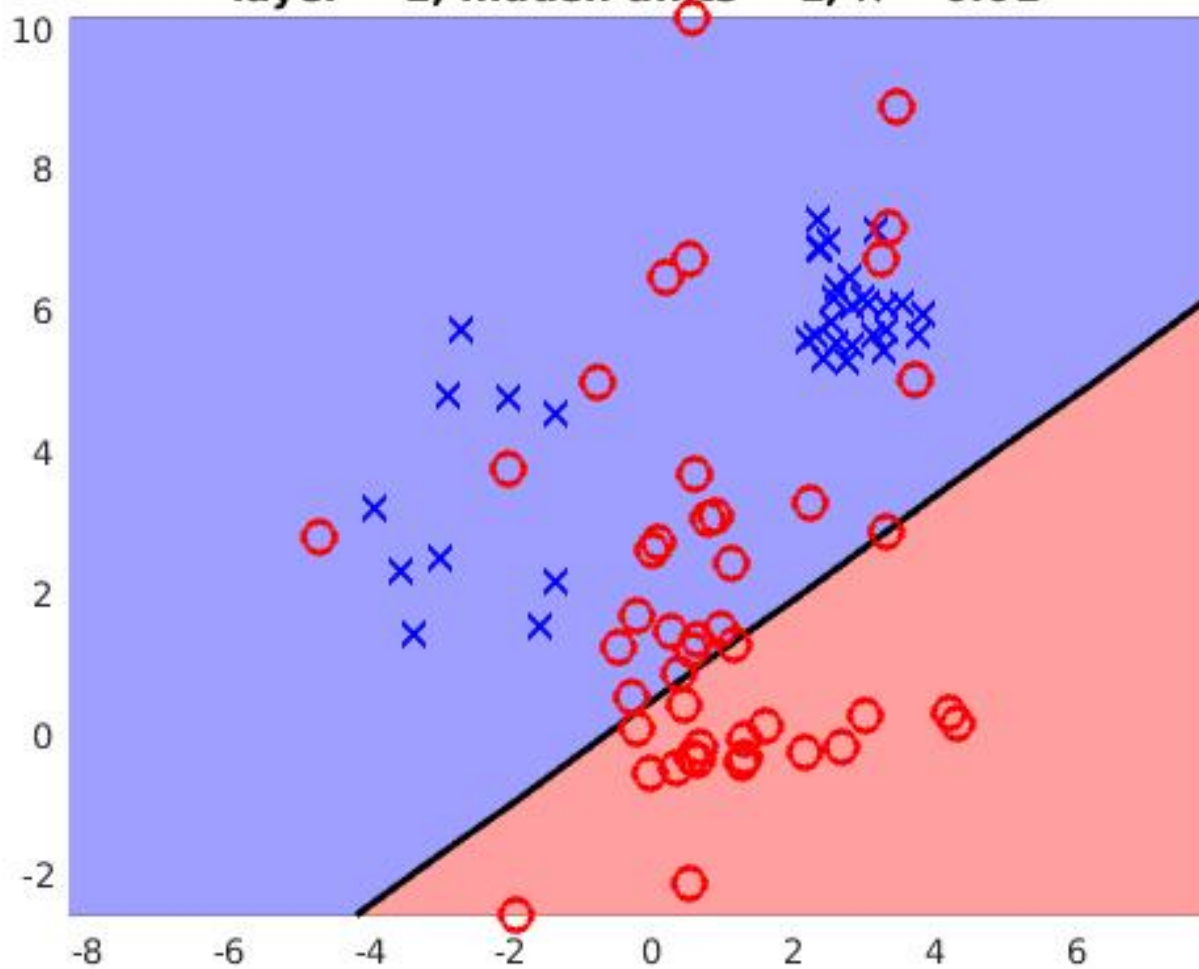
PS5

Question 1

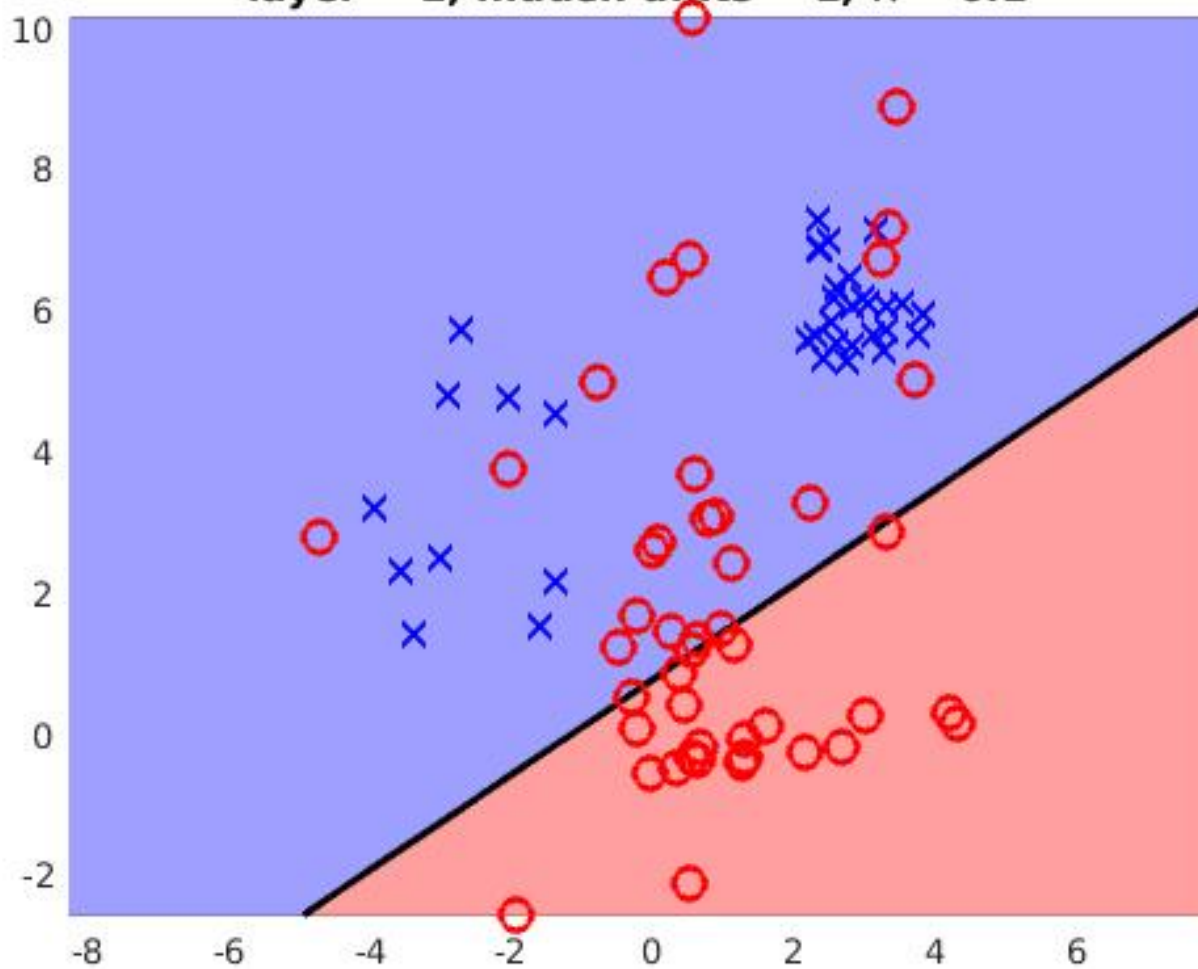
**layer = 2, hidden units = 1,  $\lambda = 0.001$**



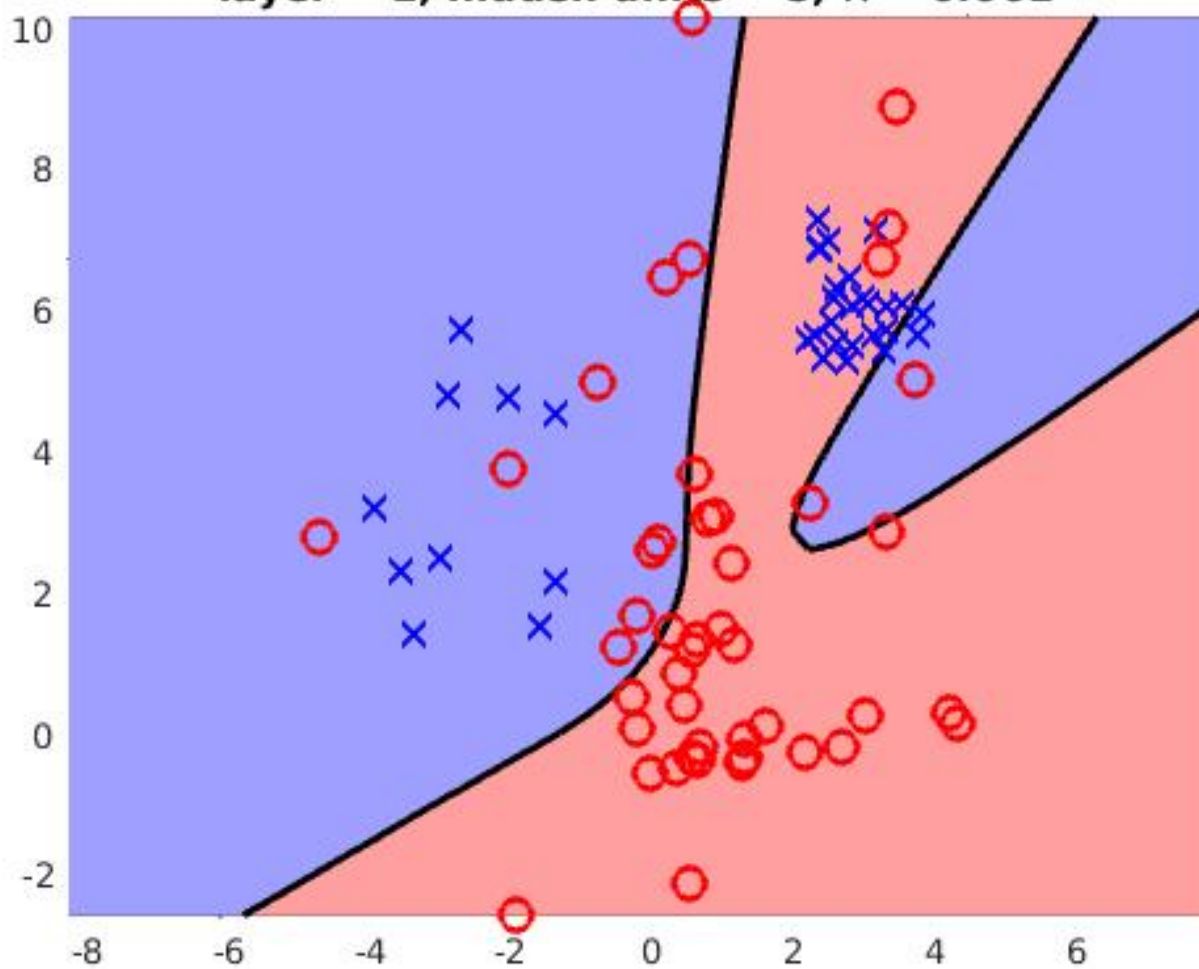
**layer = 2, hidden units = 1,  $\lambda = 0.01$**



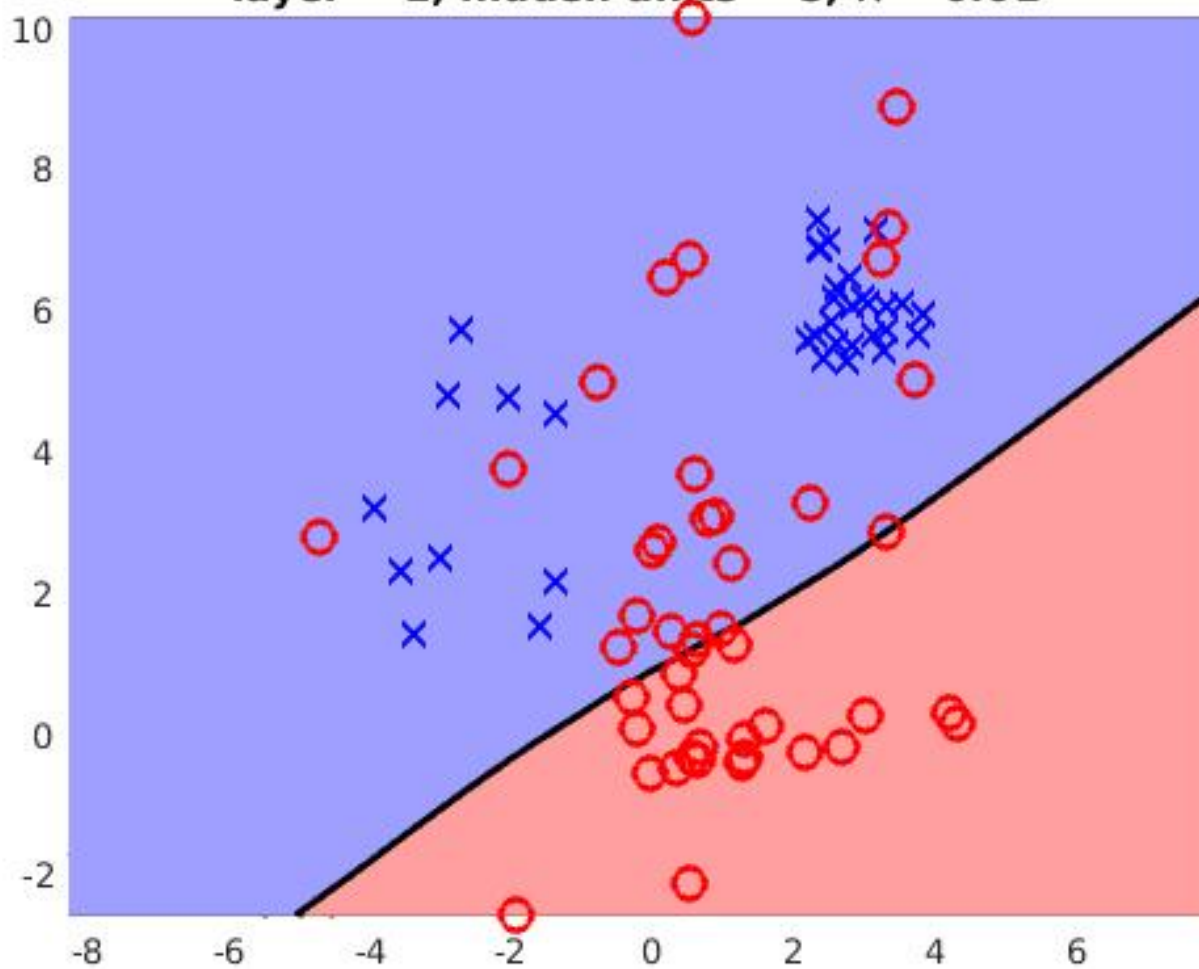
**layer = 2, hidden units = 1,  $\lambda = 0.1$**



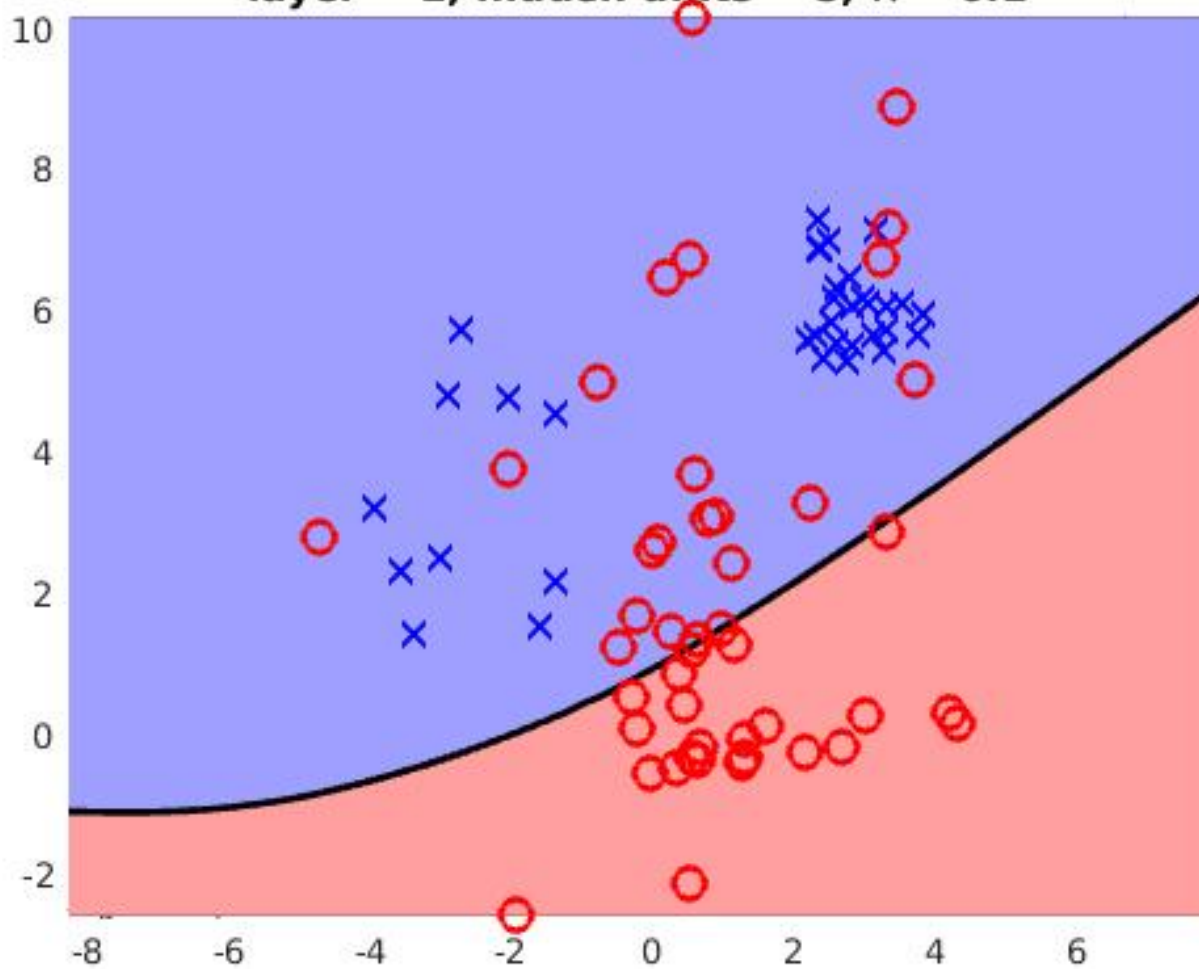
**layer = 2, hidden units = 5,  $\lambda = 0.001$**



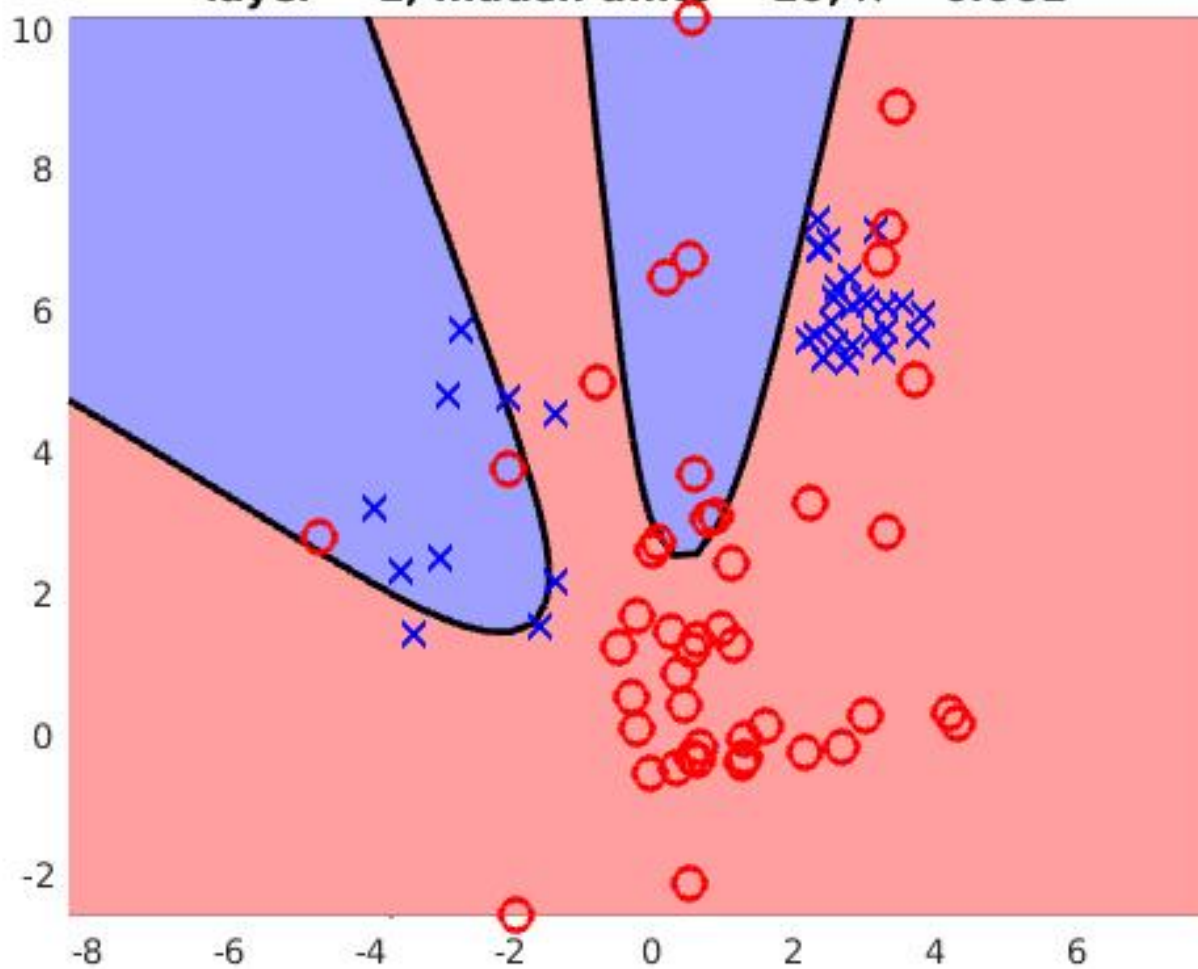
**layer = 2, hidden units = 5,  $\lambda = 0.01$**



**layer = 2, hidden units = 5,  $\lambda = 0.1$**

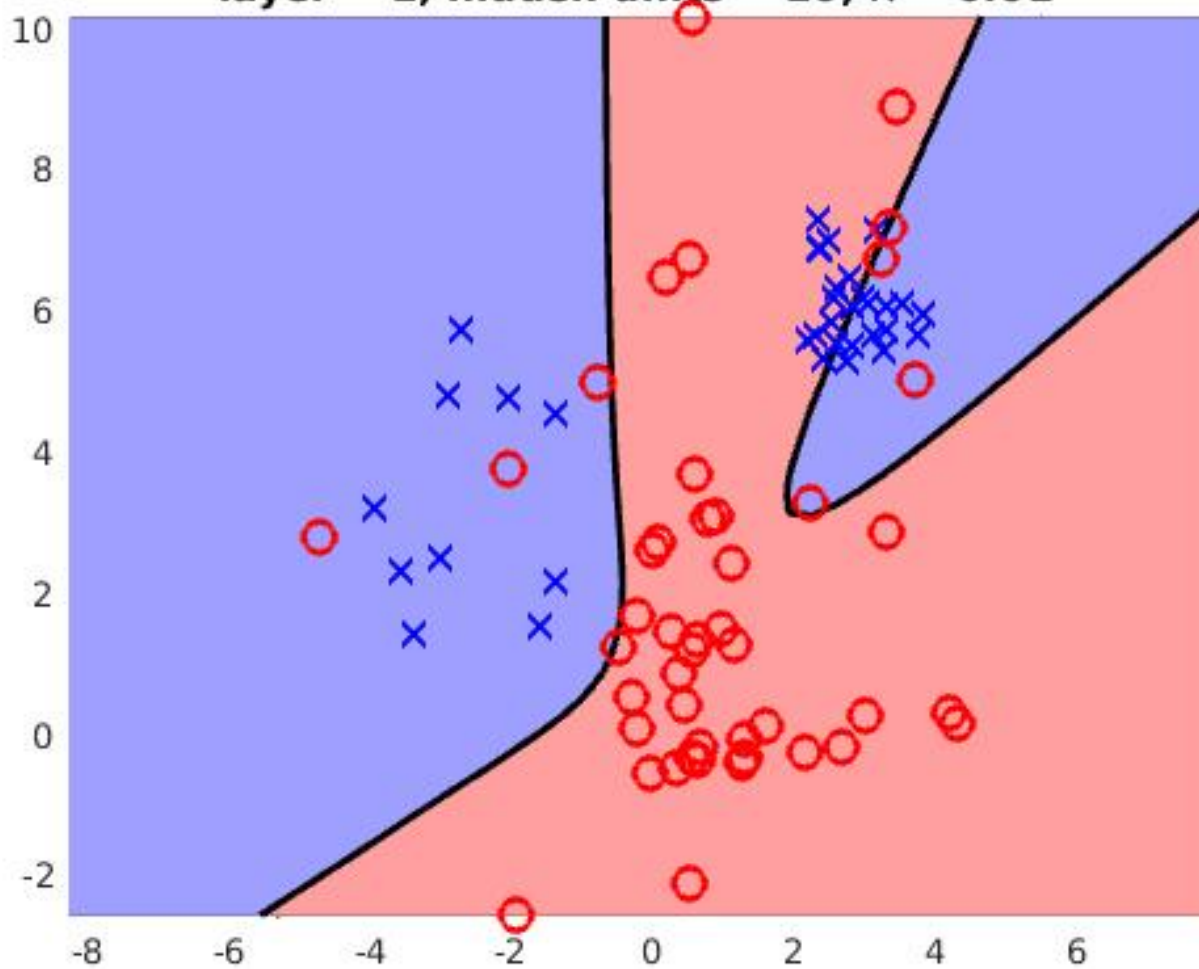


**layer = 2, hidden units = 20,  $\lambda = 0.001$**

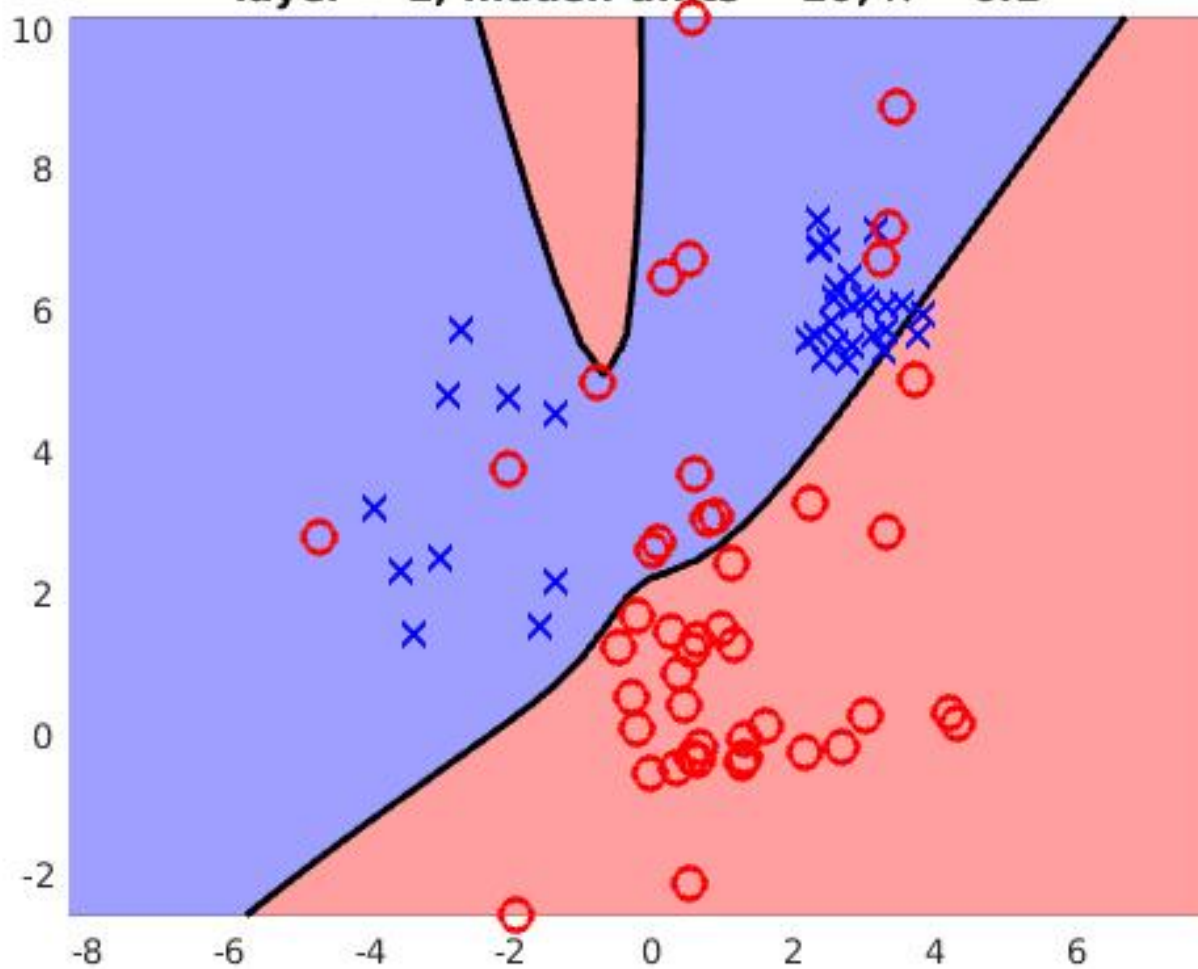




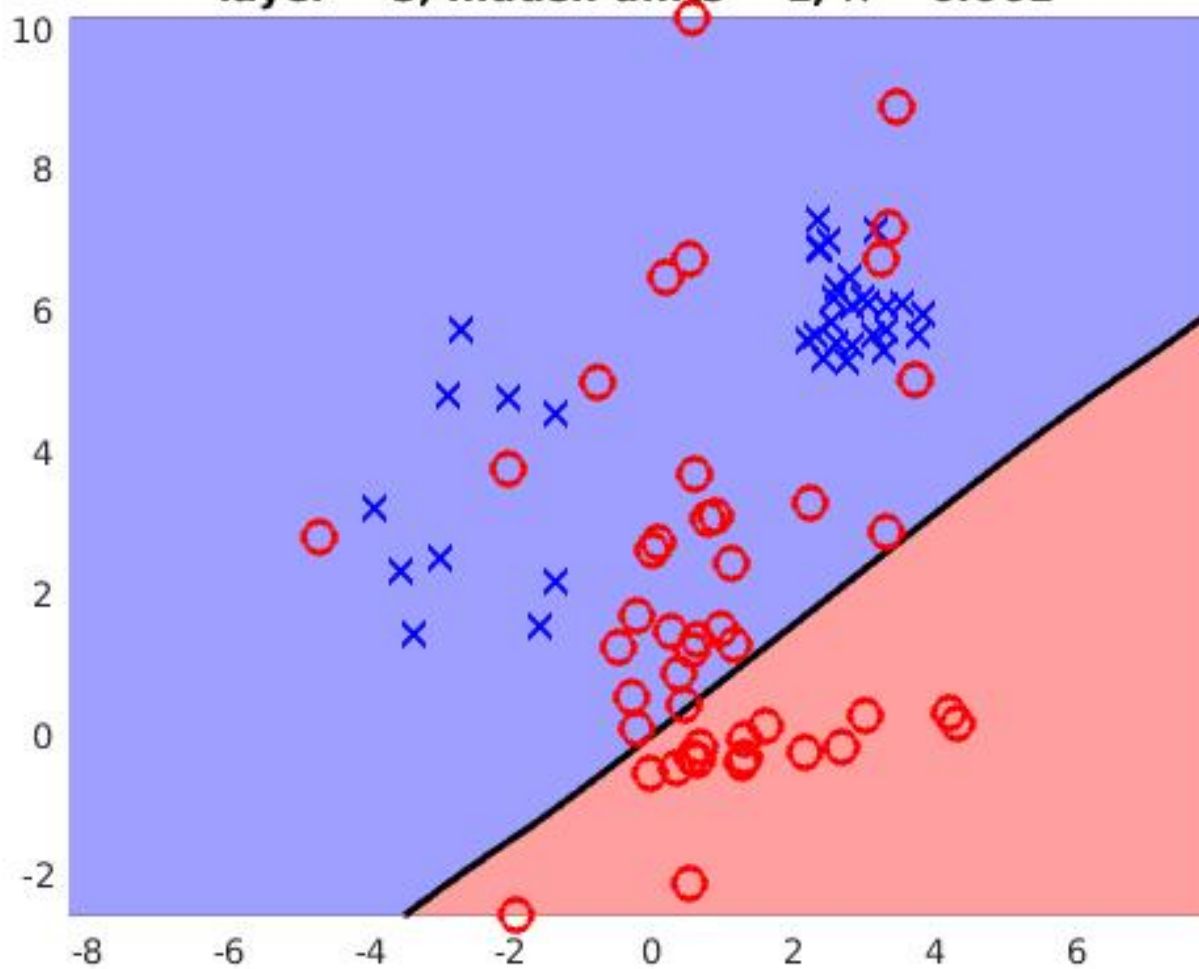
**layer = 2, hidden units = 20,  $\lambda = 0.01$**



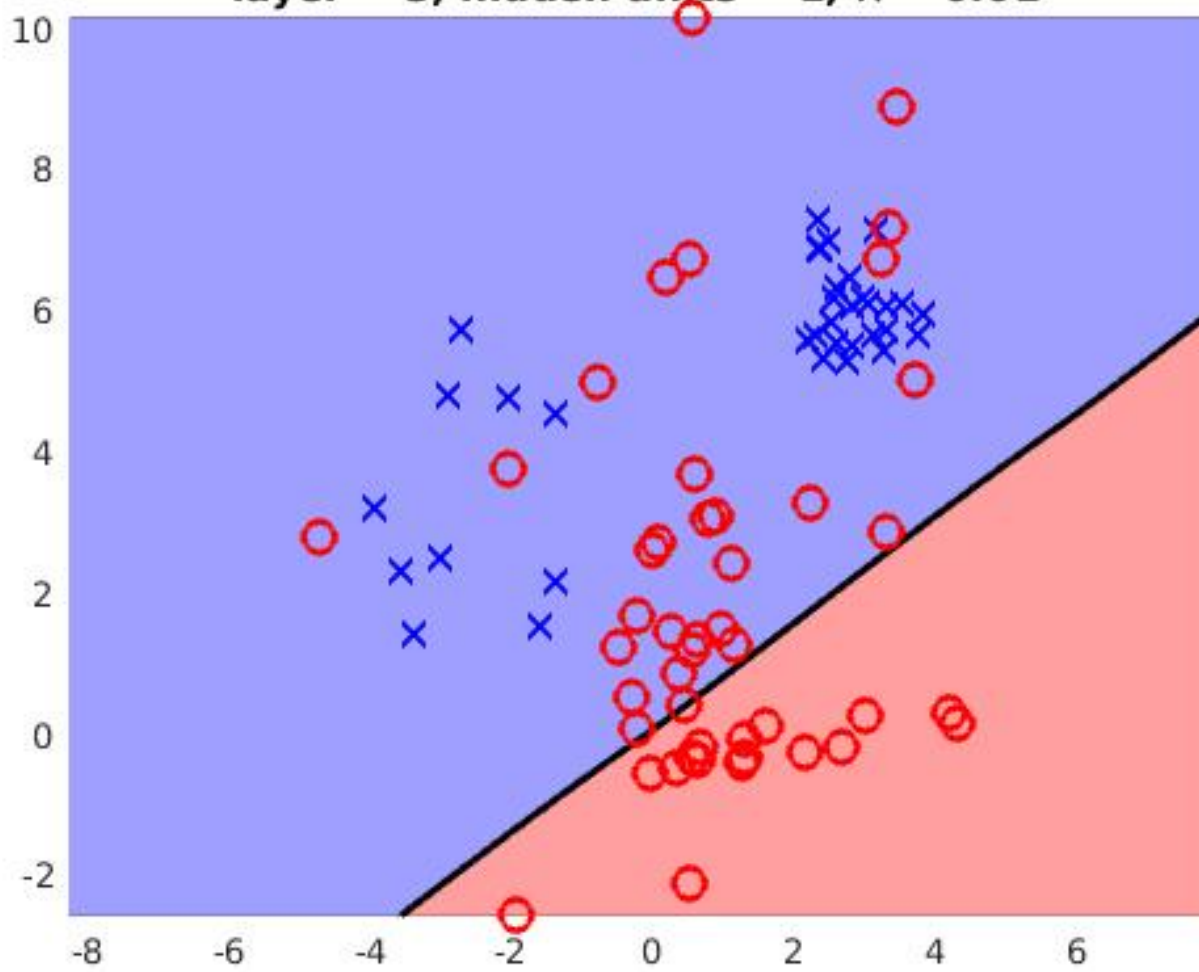
**layer = 2, hidden units = 20,  $\lambda = 0.1$**



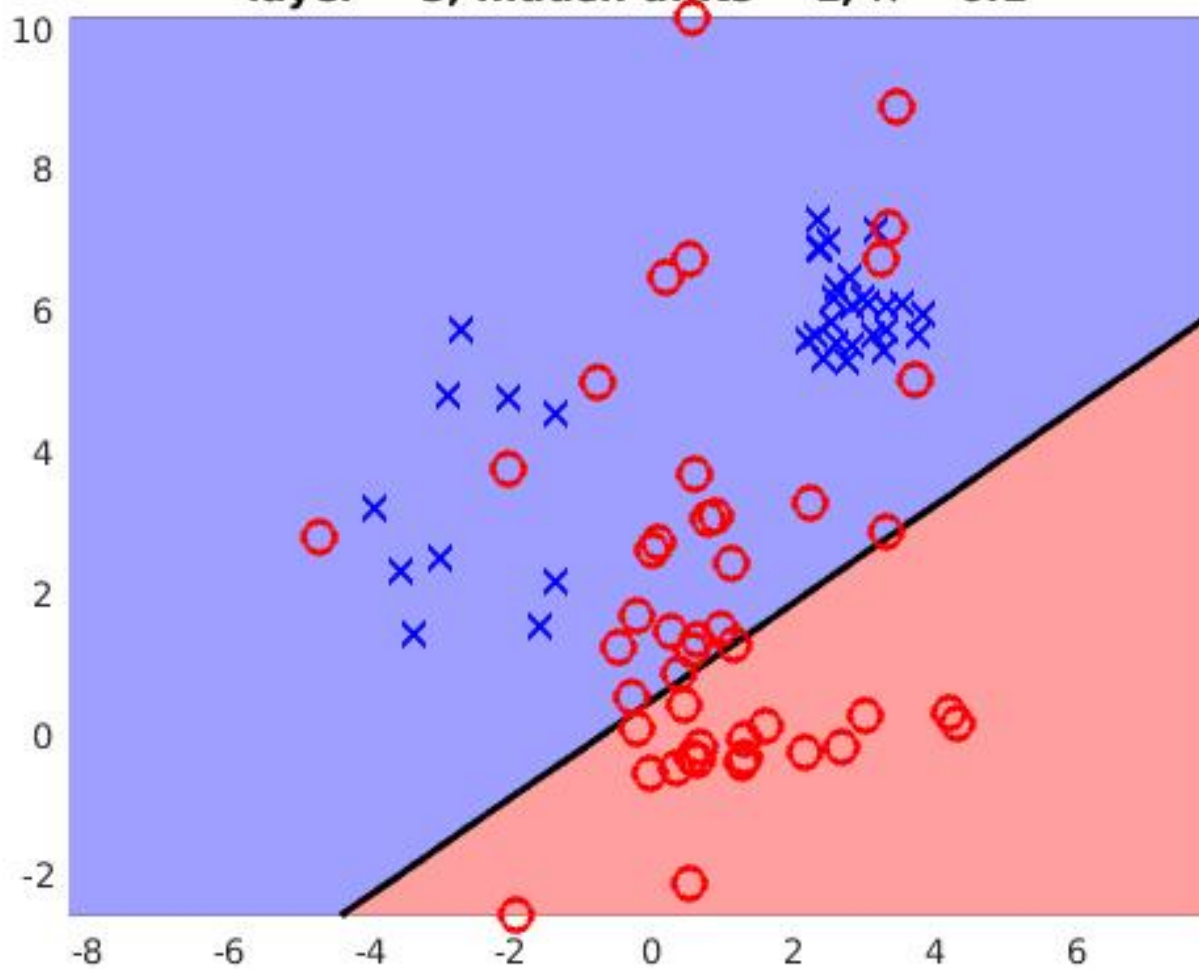
**layer = 3, hidden units = 1,  $\lambda = 0.001$**



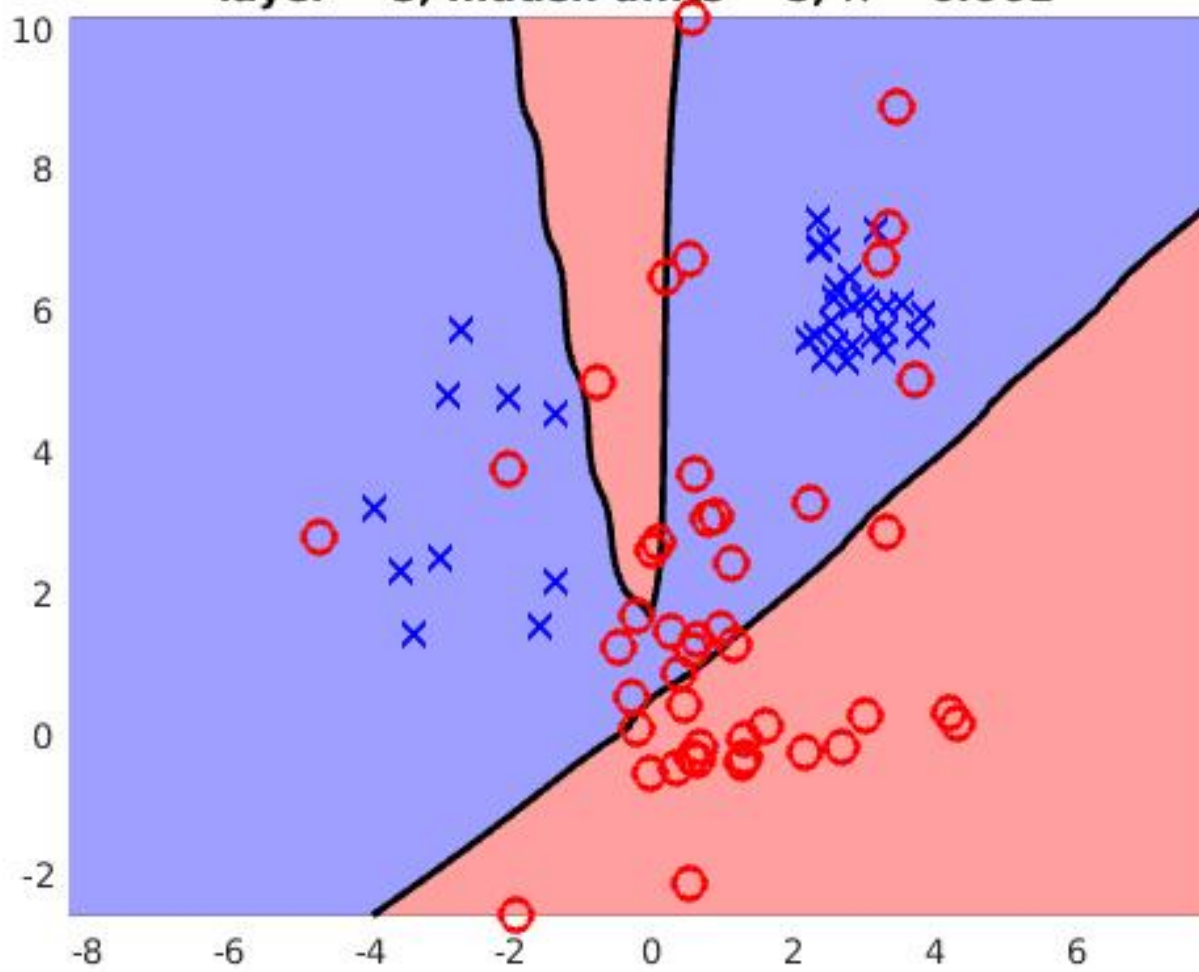
The figure is a scatter plot illustrating a linear classification problem. The x-axis ranges from -8 to 6, and the y-axis ranges from -2 to 10. A solid black line represents the linear decision boundary, which is approximately defined by the equation  $y = x + 3$ . The region above this line is shaded light blue, and the region below is shaded light red. There are two classes of data points: blue 'x' marks and red 'o' marks. The blue 'x' marks are primarily located in the blue region, with a small cluster around (2.5, 6) and another around (-3, 4). The red 'o' marks are primarily located in the red region, with a large cluster around (0, 0) and another around (3, 5). There is some overlap between the two classes, particularly in the region around (0, 2) and (3, 6).



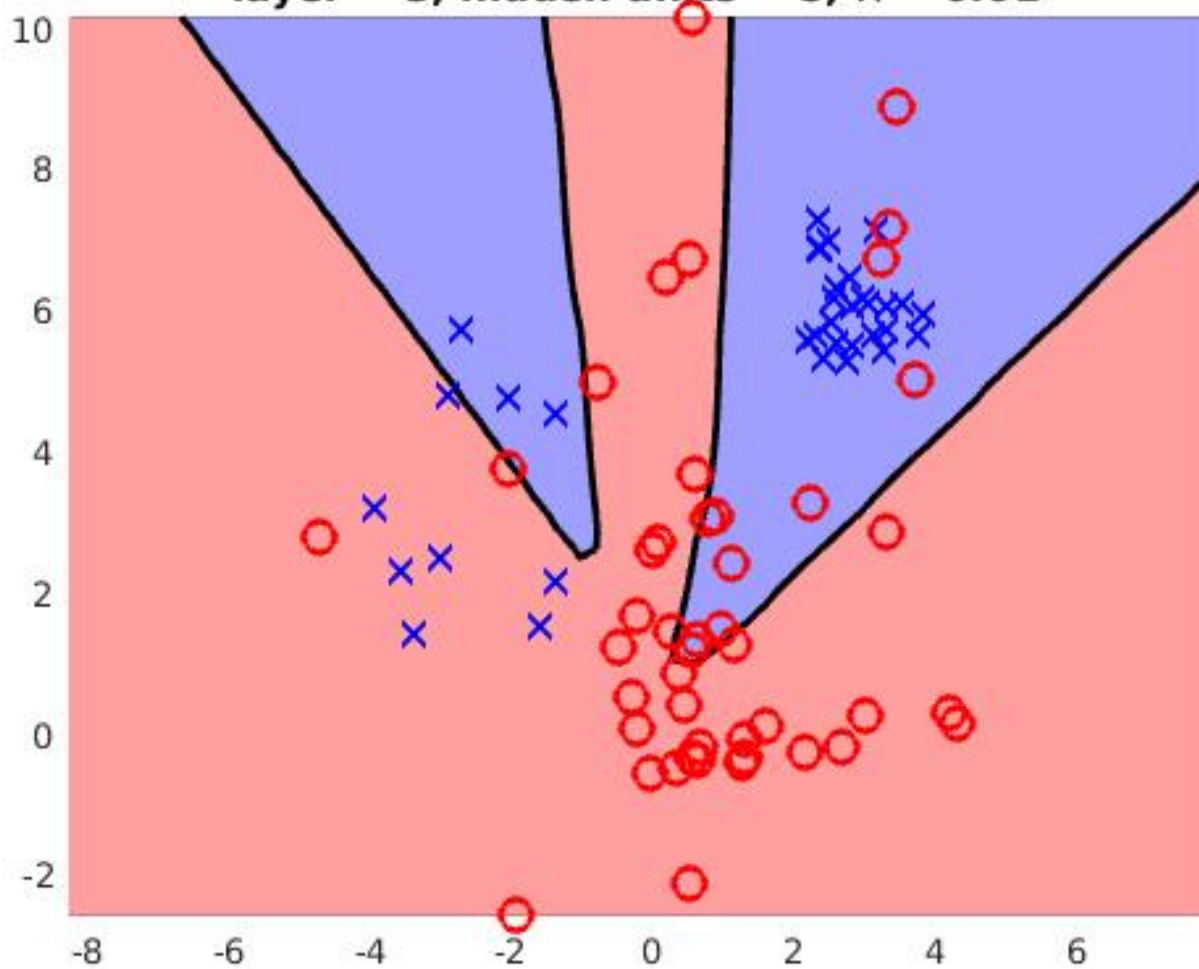
**layer = 3, hidden units = 1,  $\lambda = 0.1$**



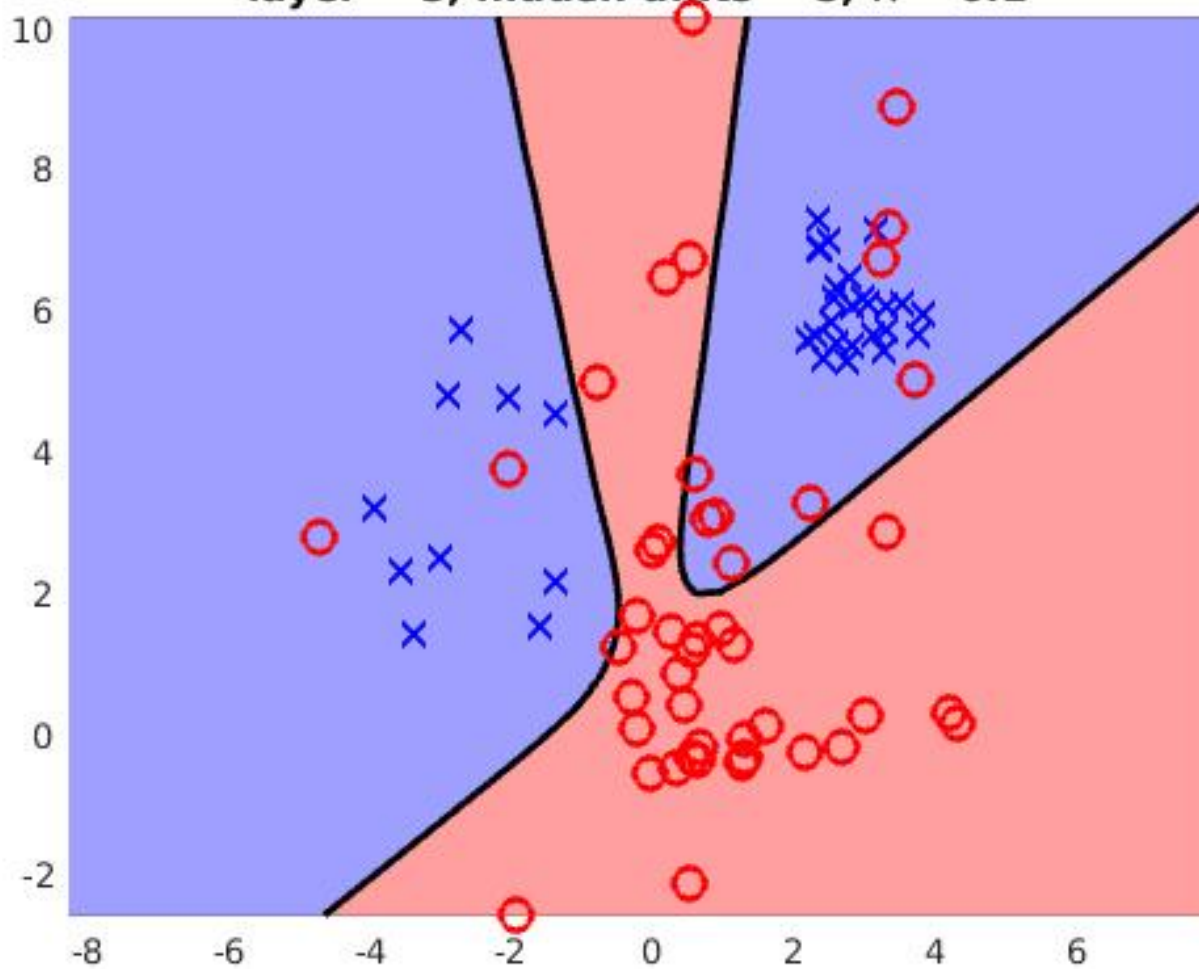
**layer = 3, hidden units = 5,  $\lambda = 0.001$**



**layer = 3, hidden units = 5,  $\lambda = 0.01$**

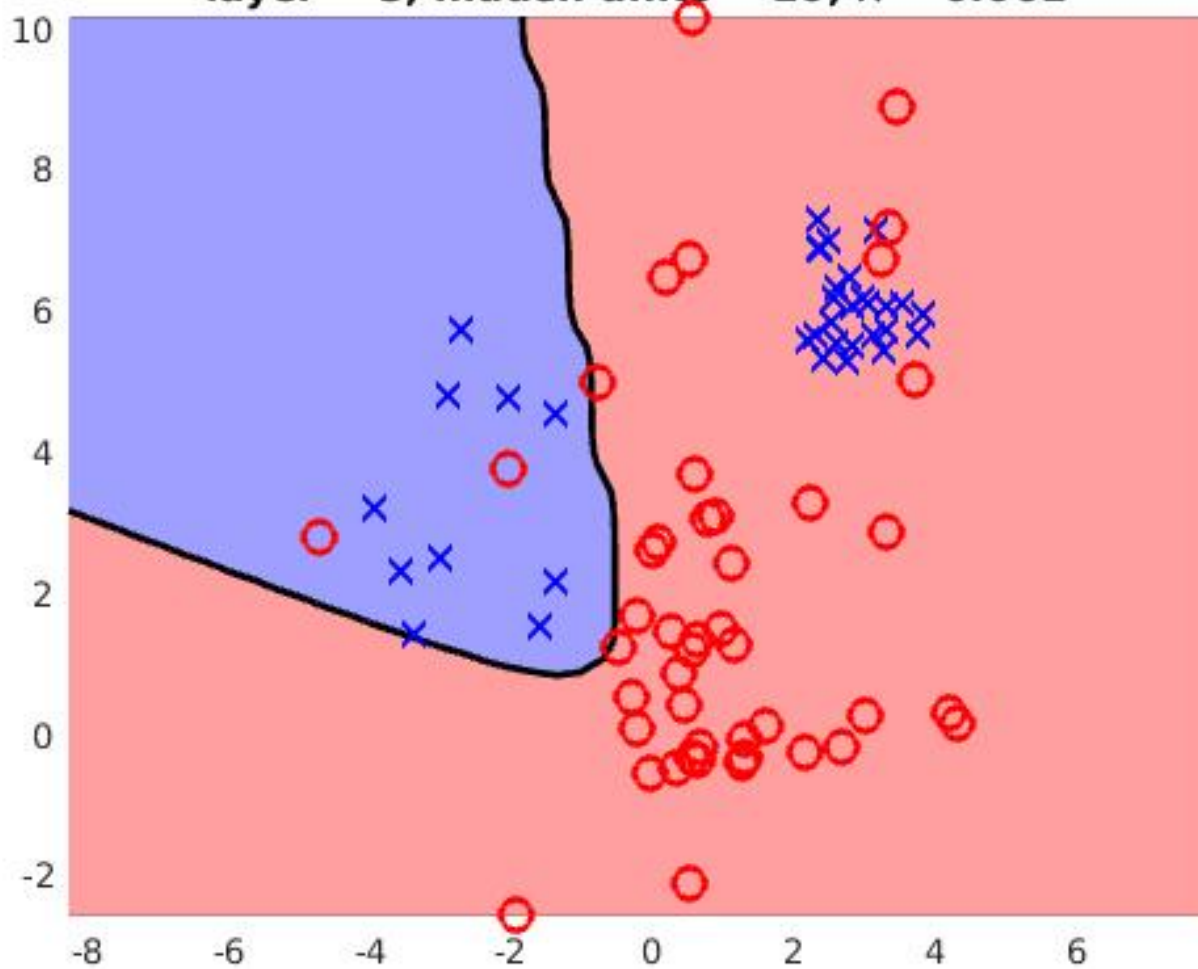


**layer = 3, hidden units = 5,  $\lambda = 0.1$**

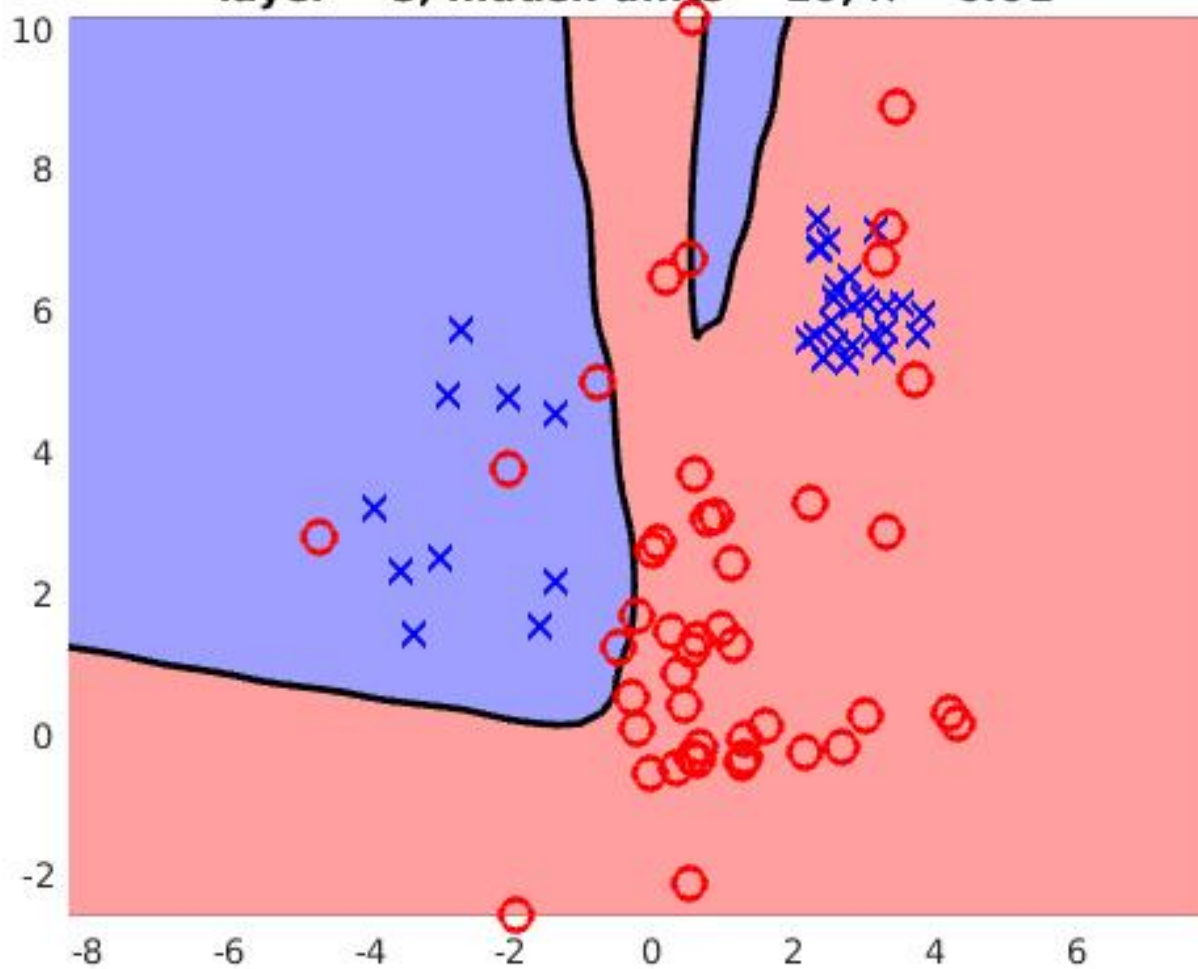




**layer = 3, hidden units = 20,  $\lambda = 0.001$**



**layer = 3, hidden units = 20,  $\lambda = 0.01$**



**layer = 3, hidden units = 20,  $\lambda = 0.1$**

