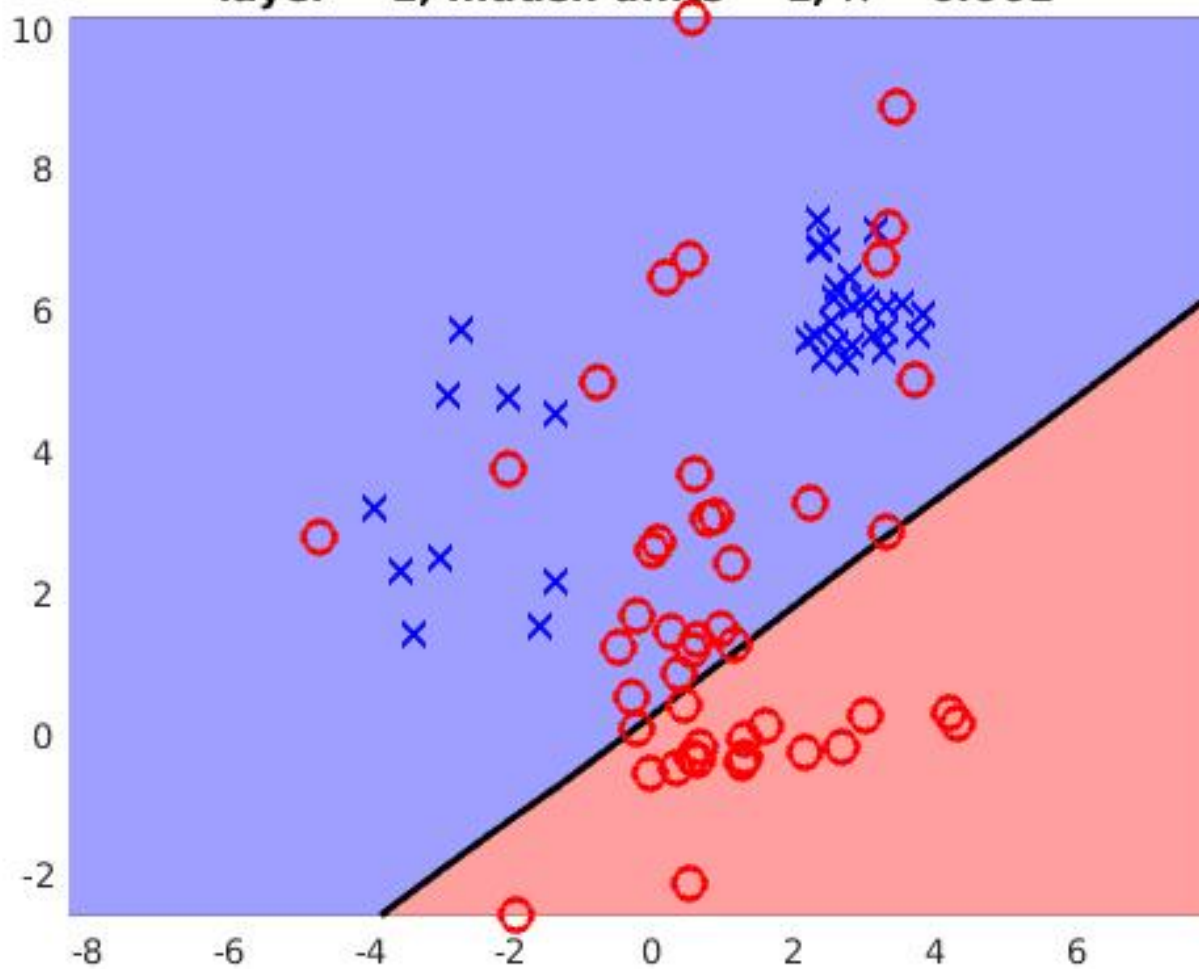
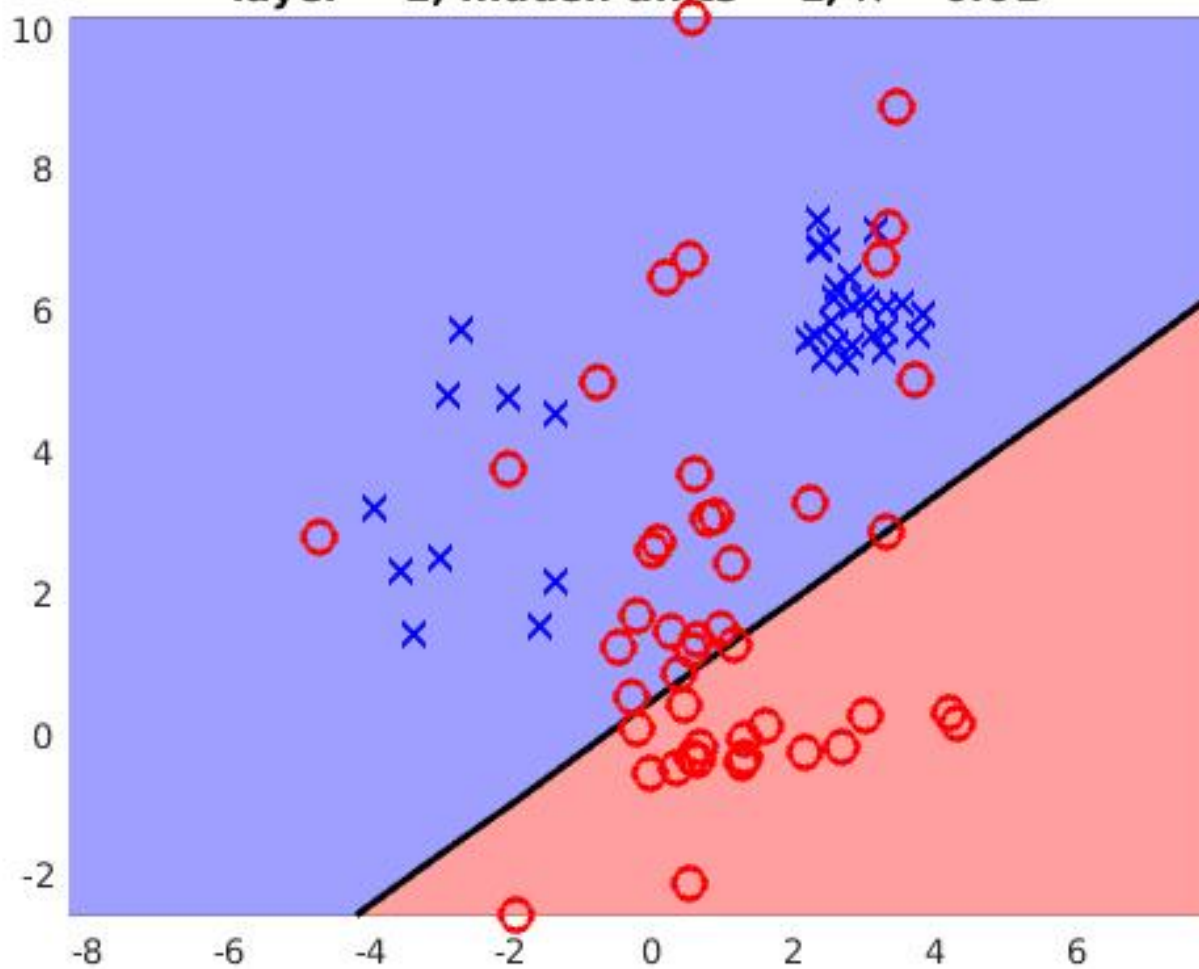


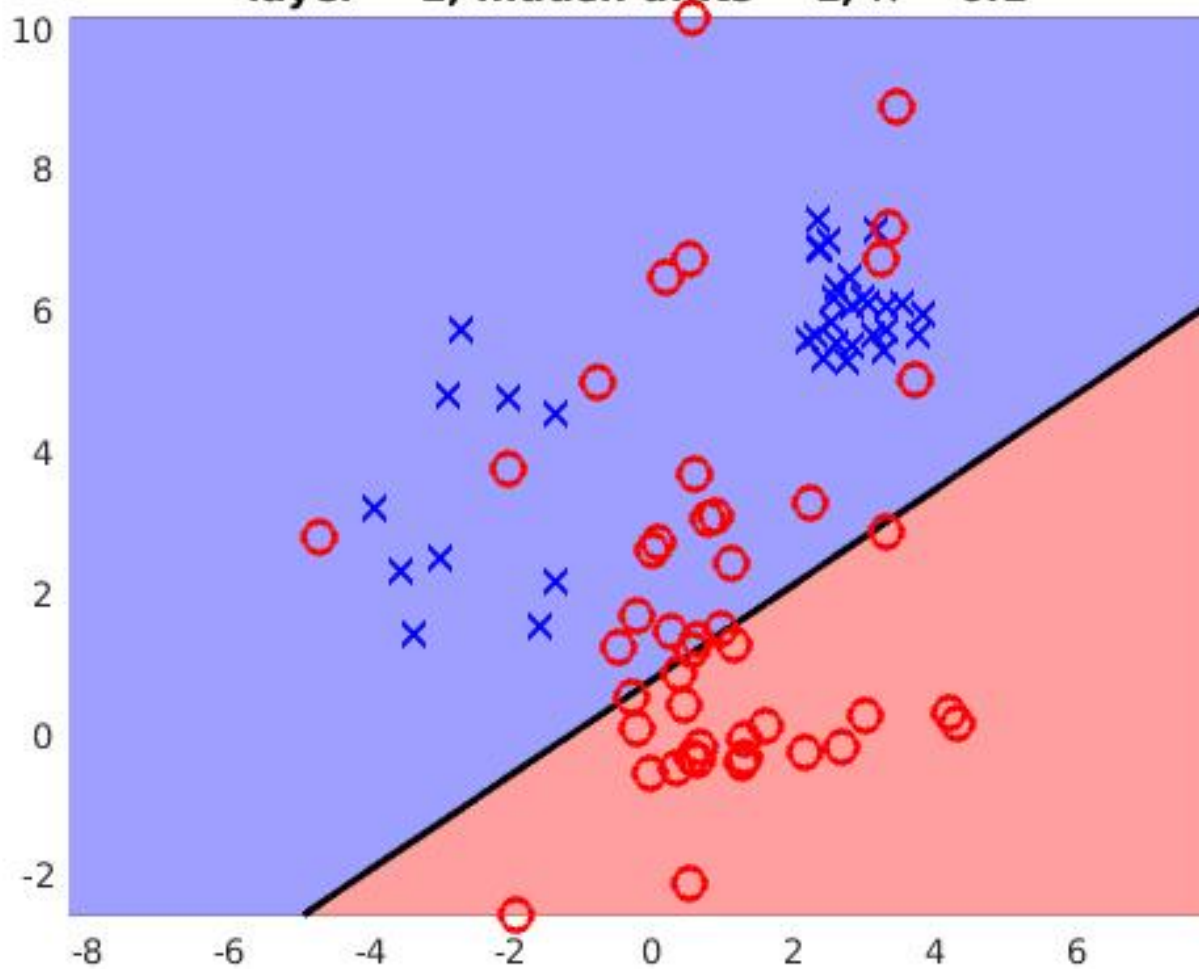
layer = 2, hidden units = 1,  $\lambda = 0.001$



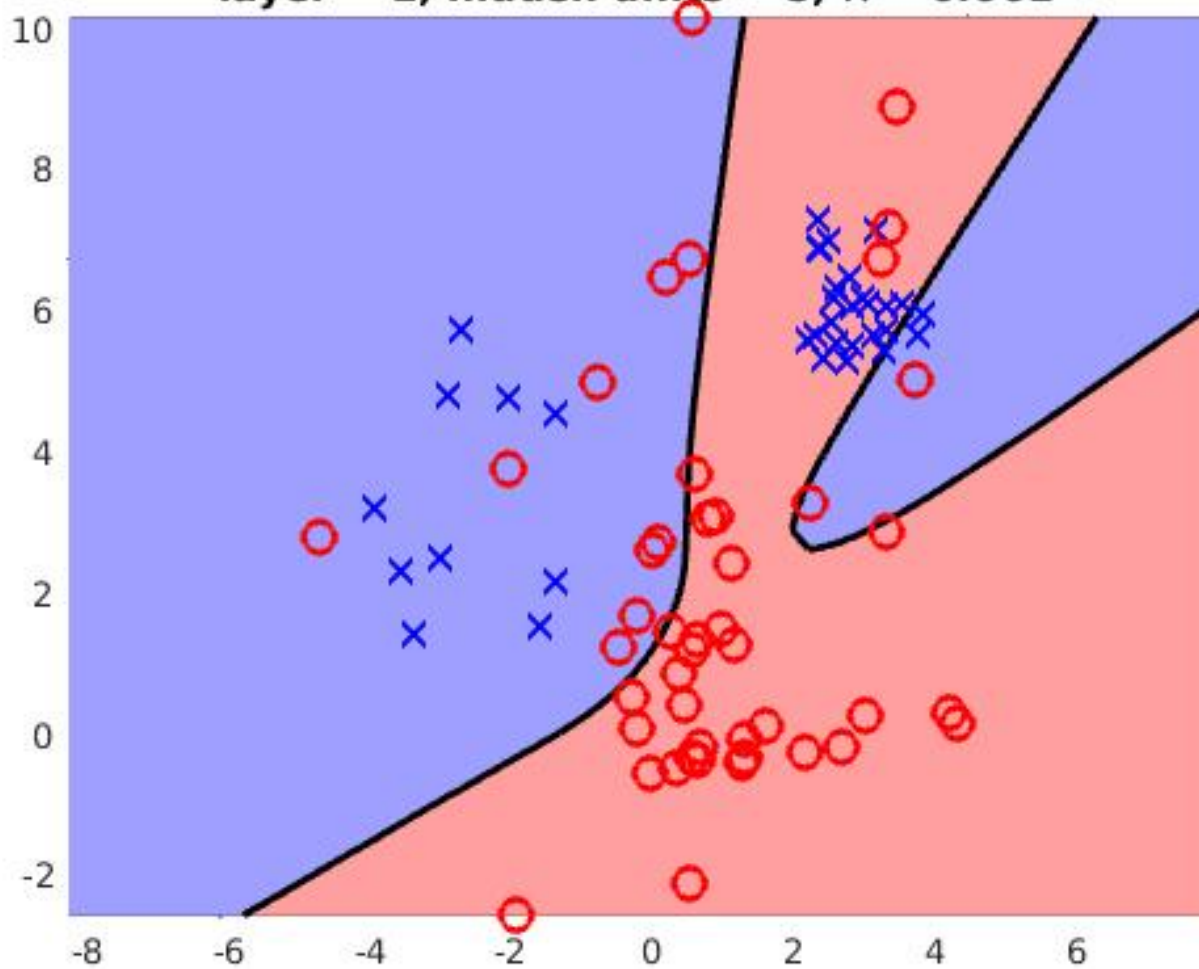
**layer = 2, hidden units = 1,  $\lambda = 0.01$**



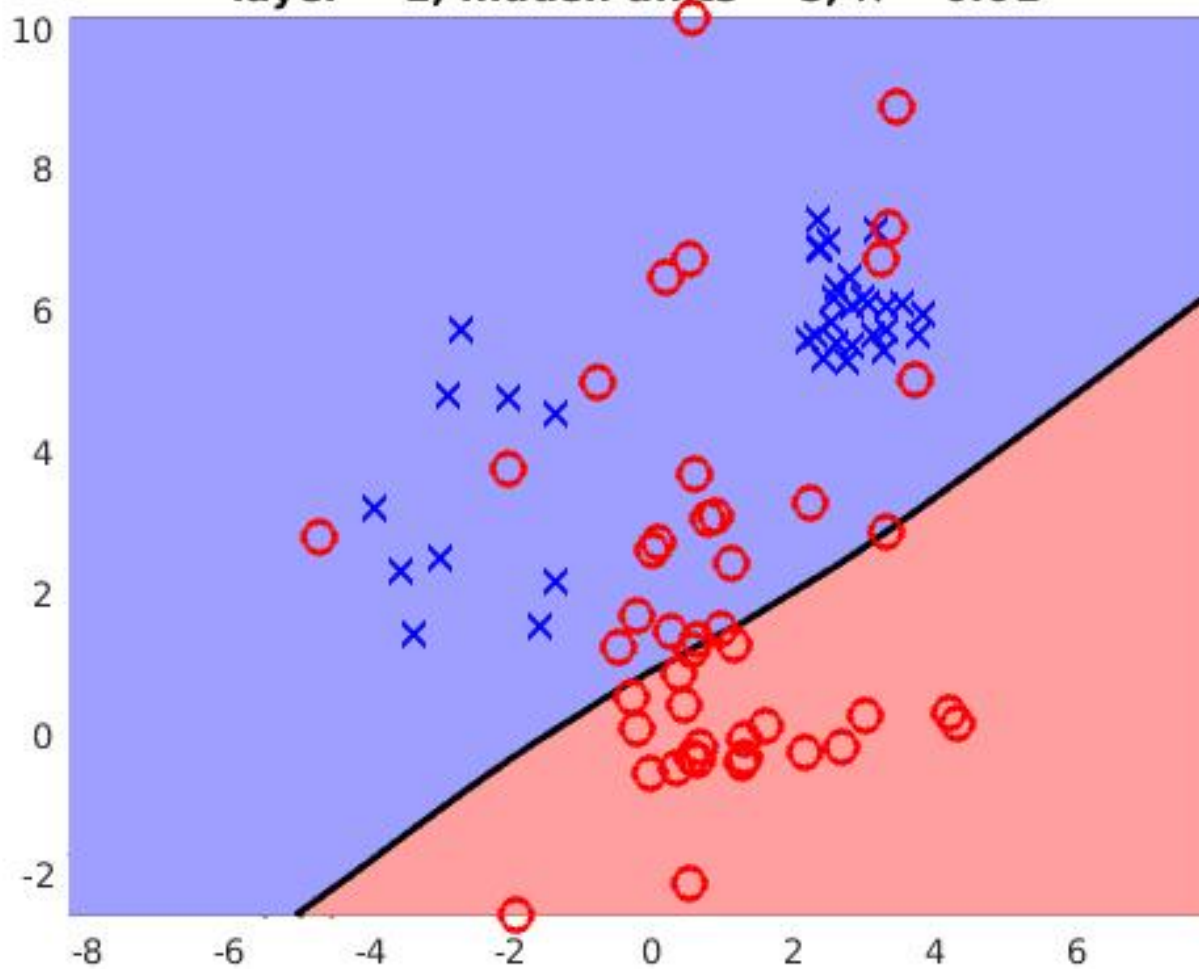
**layer = 2, hidden units = 1,  $\lambda = 0.1$**



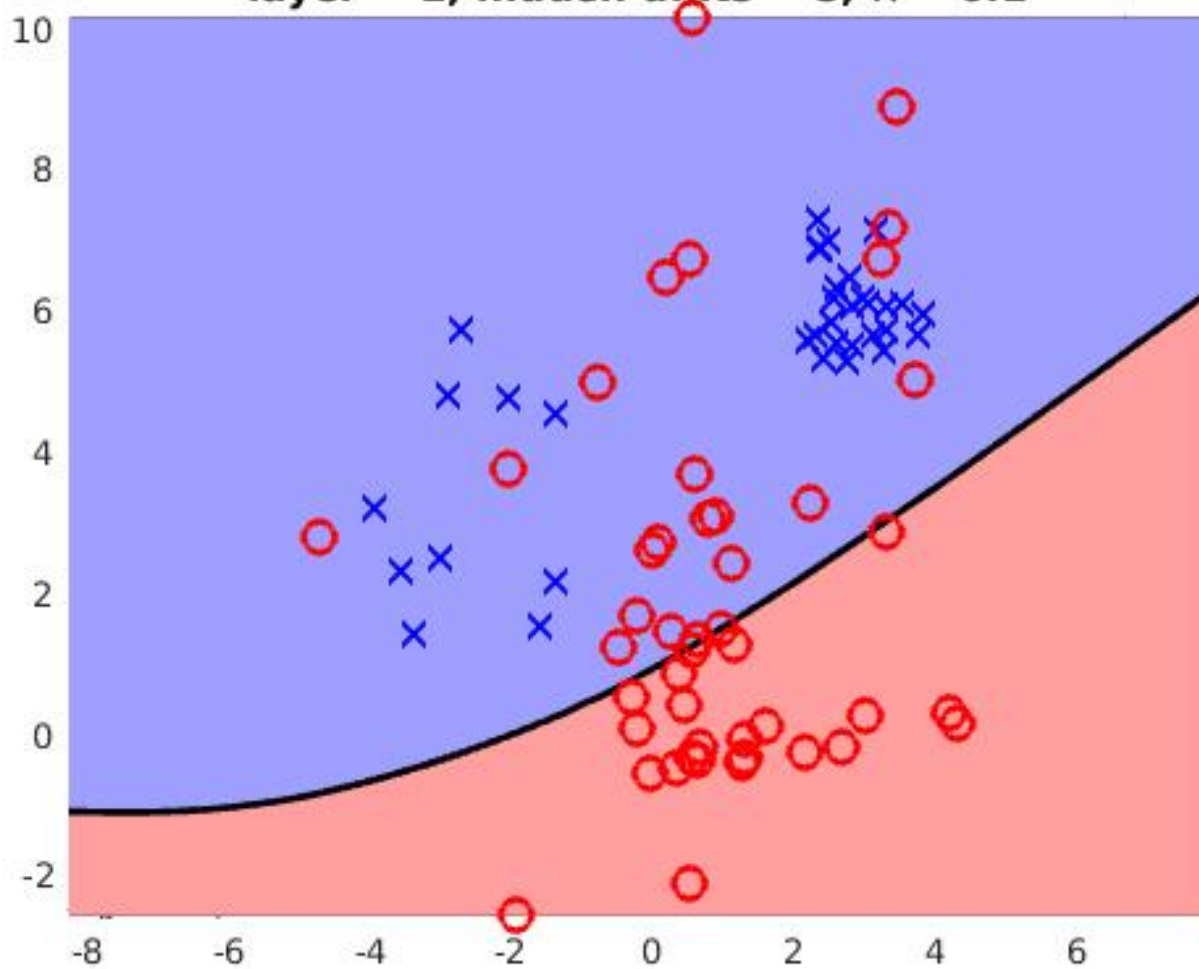
**layer = 2, hidden units = 5,  $\lambda = 0.001$**



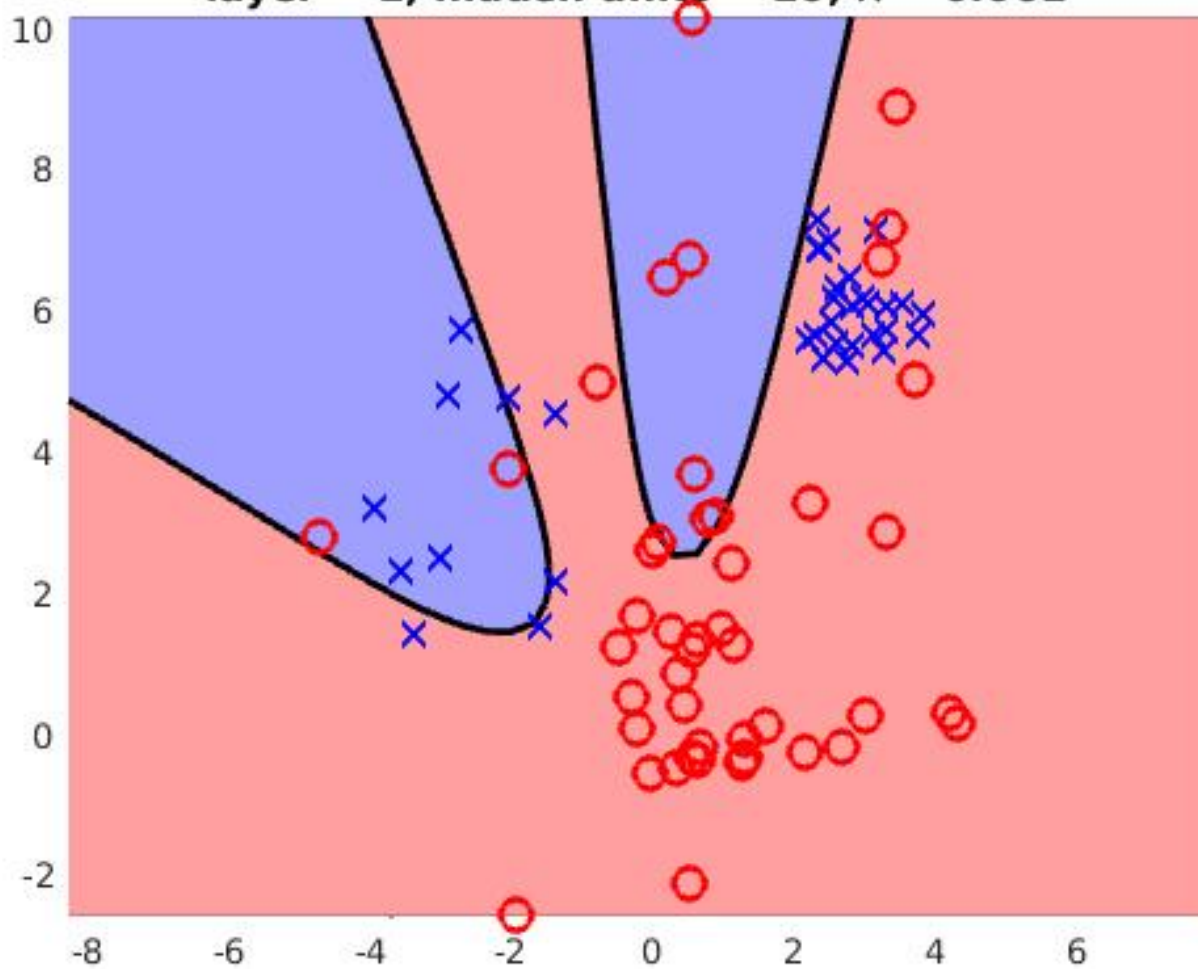
**layer = 2, hidden units = 5,  $\lambda = 0.01$**



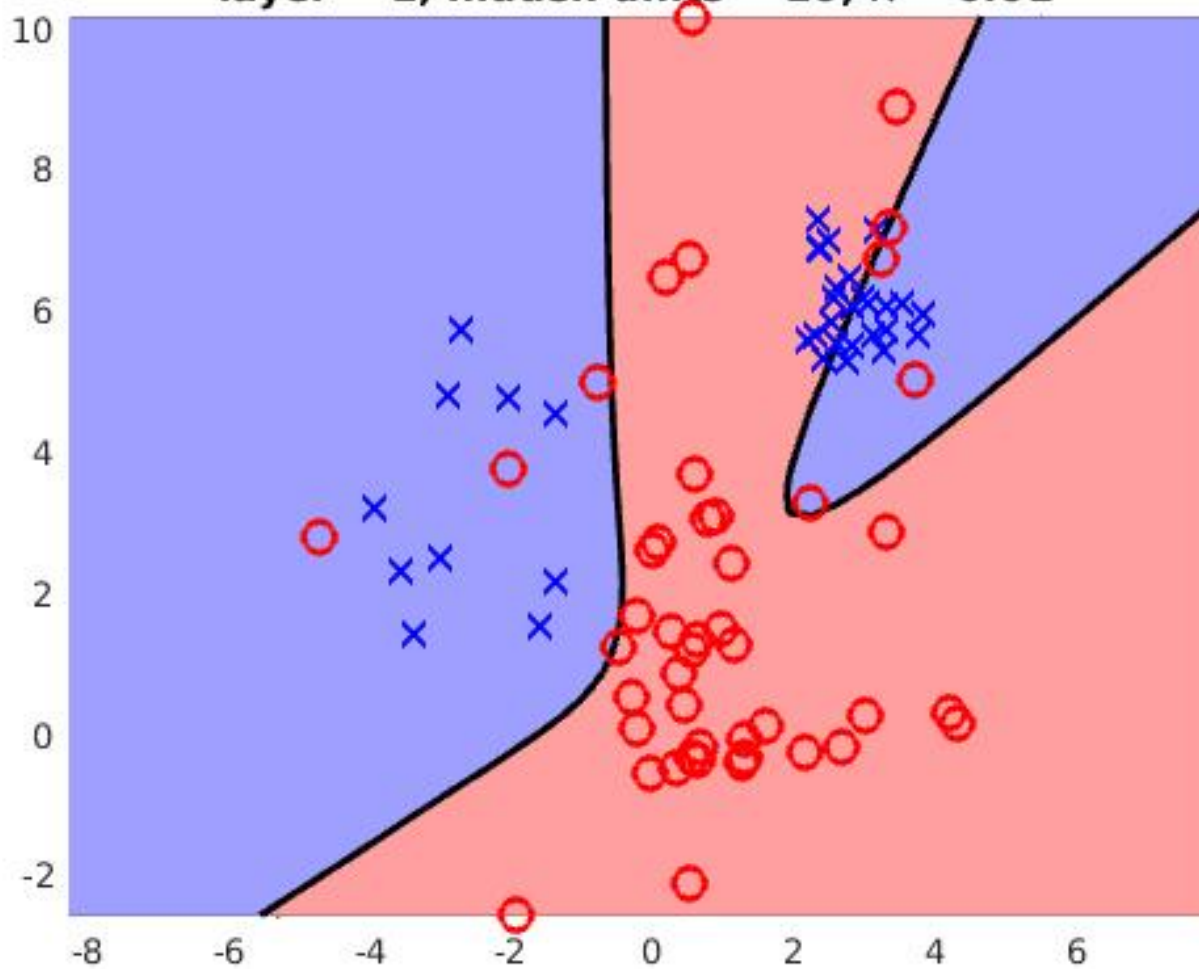
**layer = 2, hidden units = 5,  $\lambda = 0.1$**



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

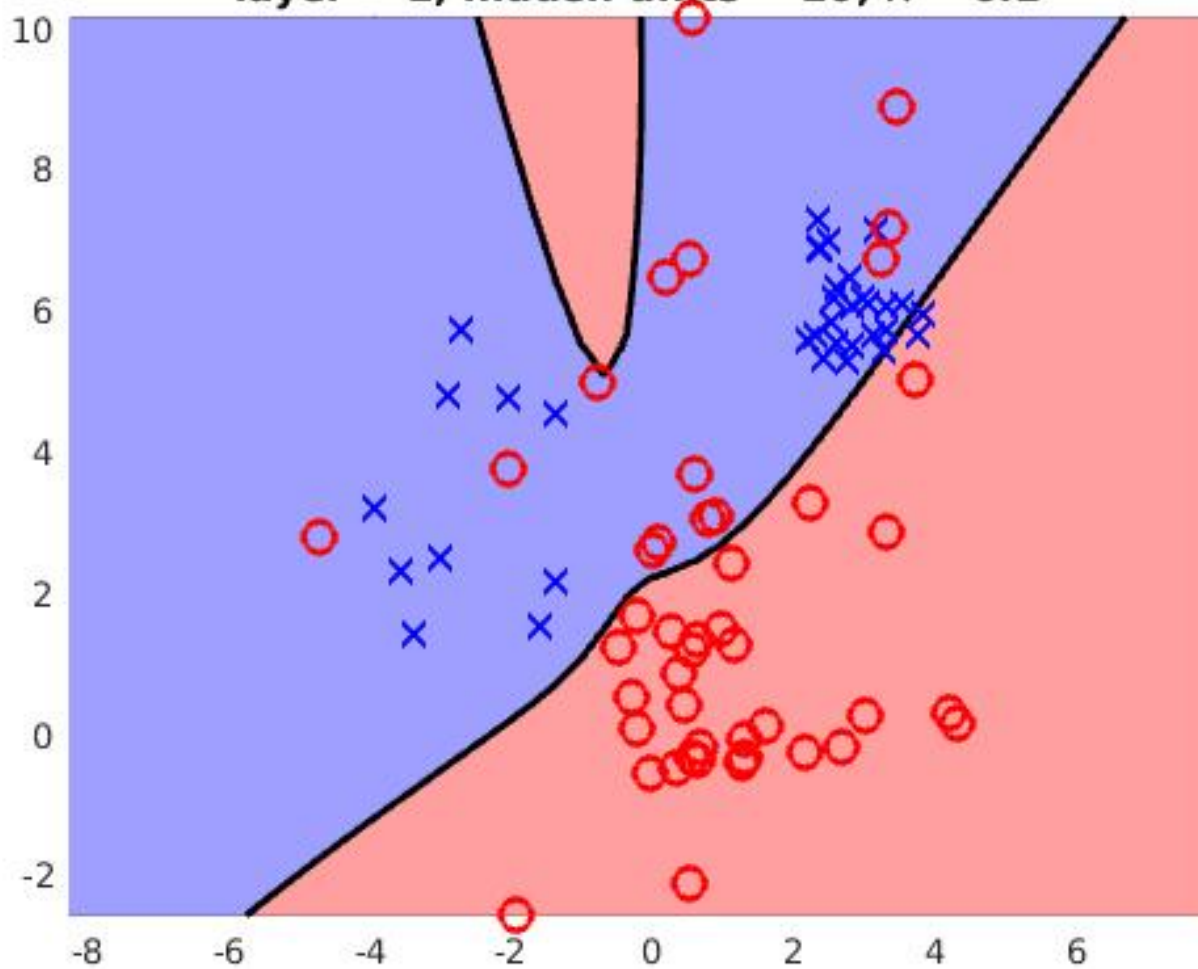


**layer = 2, hidden units = 20,  $\lambda = 0.01$**

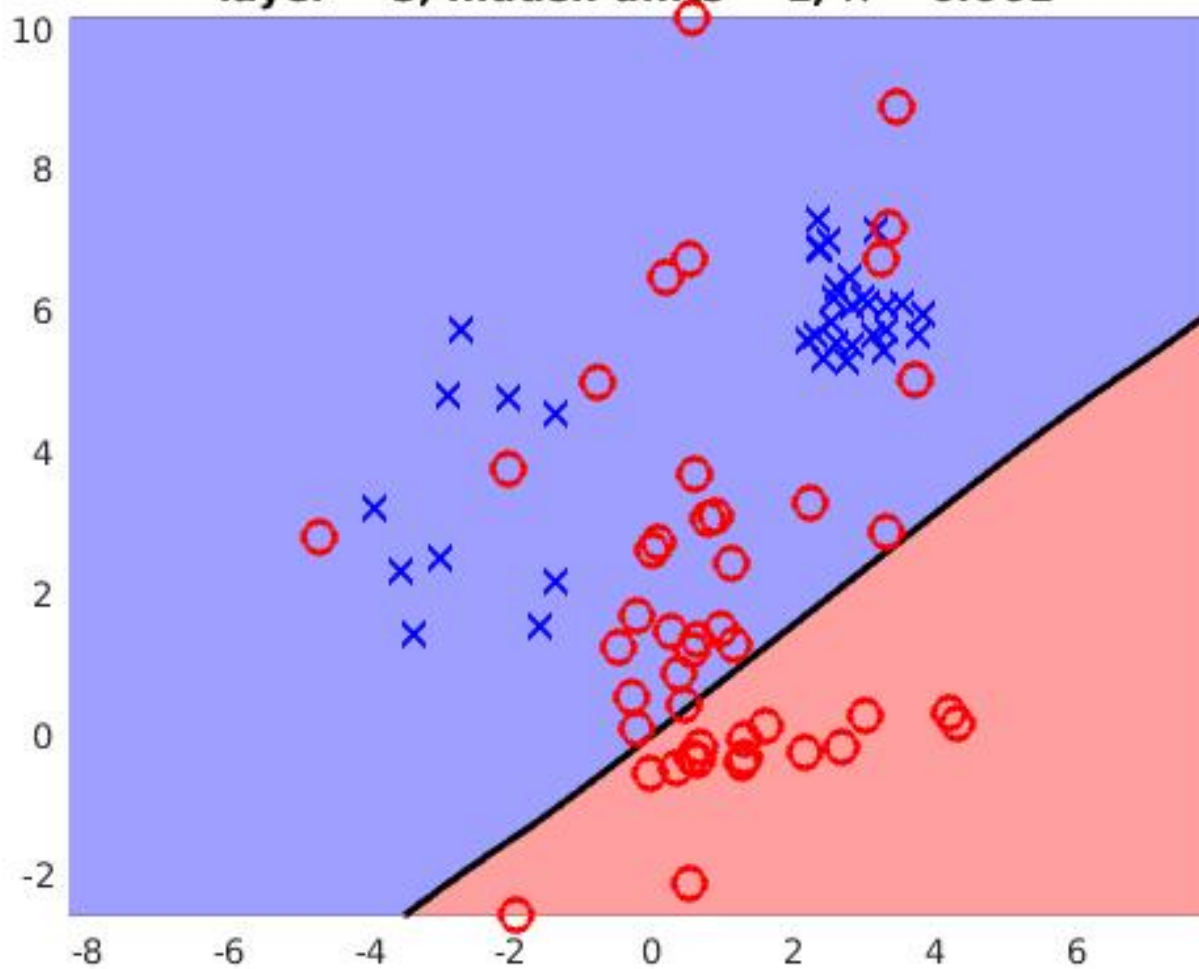




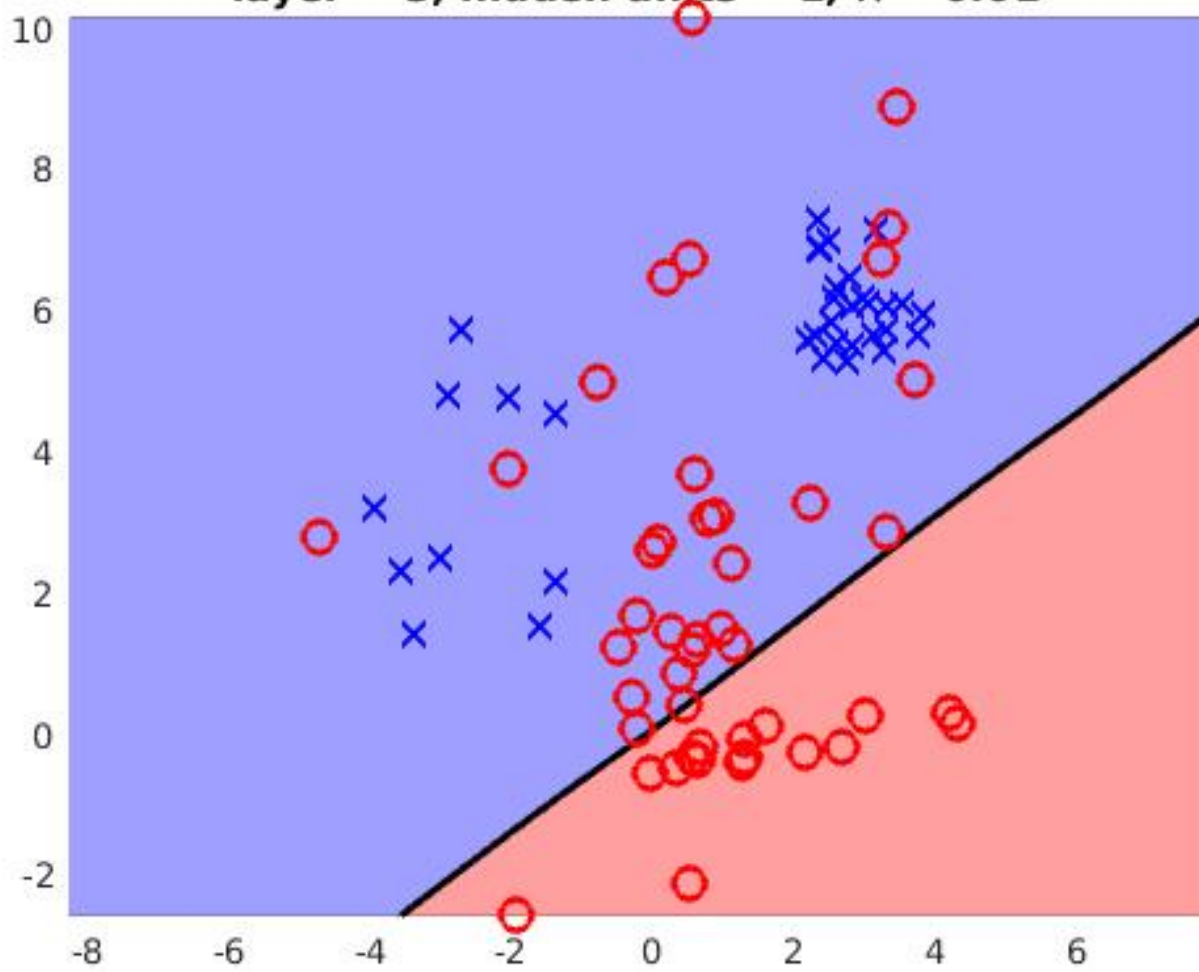
**layer = 2, hidden units = 20,  $\lambda = 0.1$**



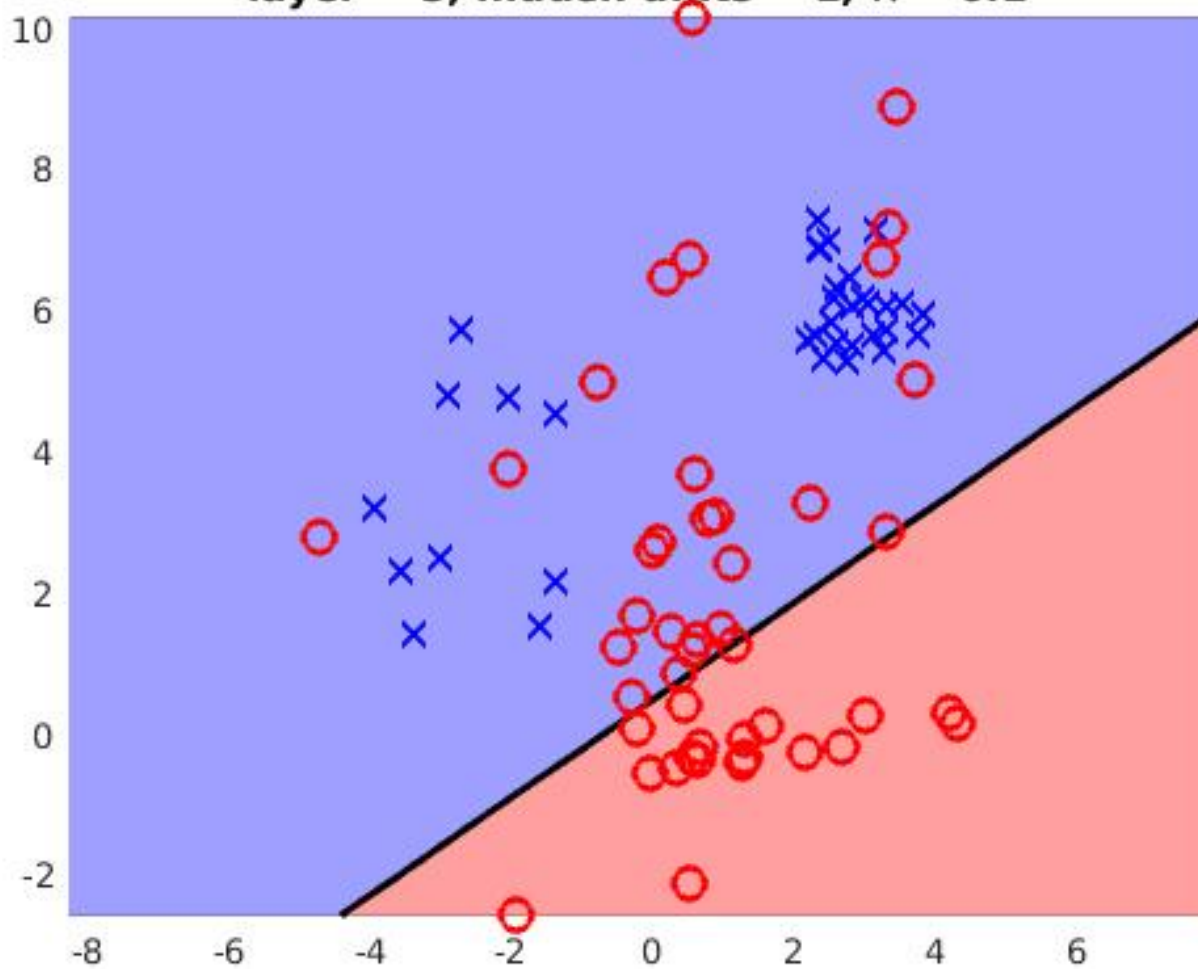
**layer = 3, hidden units = 1,  $\lambda = 0.001$**



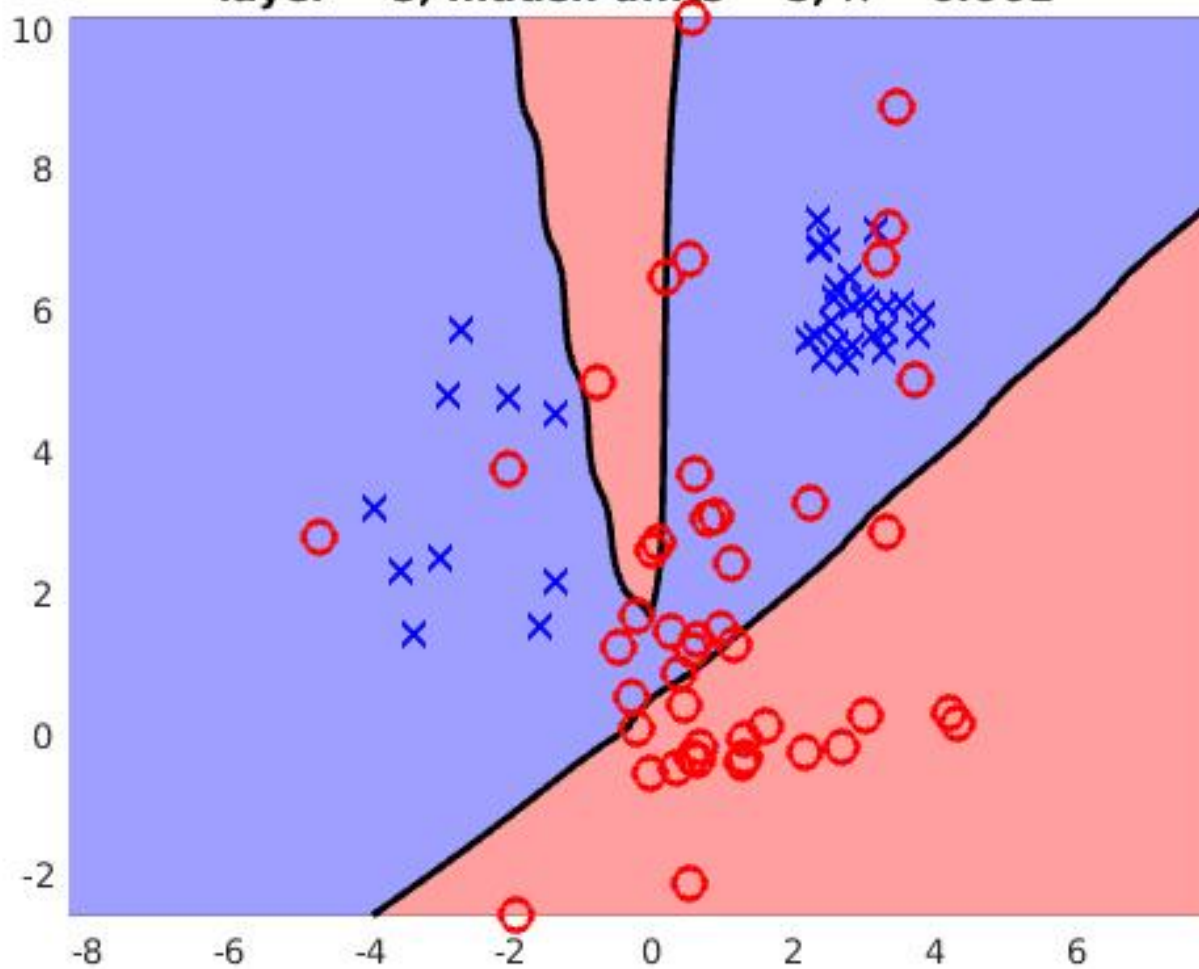
A scatter plot illustrating a linear classification problem. The x-axis ranges from -8 to 6, and the y-axis ranges from -2 to 10. The plot is divided into two regions by a black linear decision boundary. The region above the boundary is shaded blue, and the region below is shaded red. Data points are represented by blue 'x' marks and red 'o' marks. The blue 'x' marks are primarily located in the blue region, while the red 'o' marks are primarily located in the red region. There is some overlap between the two classes, particularly in the central area around the decision boundary.



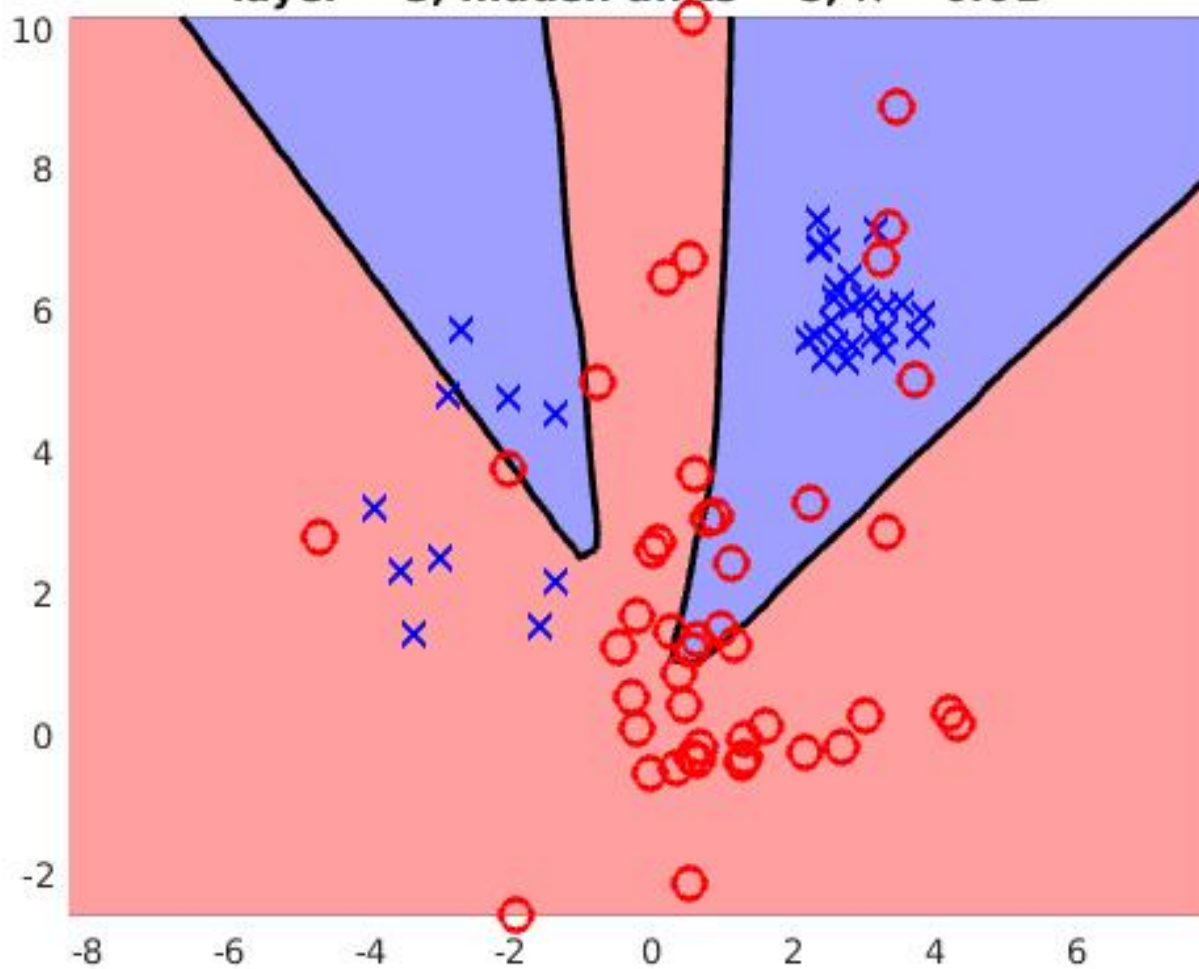
**layer = 3, hidden units = 1,  $\lambda = 0.1$**



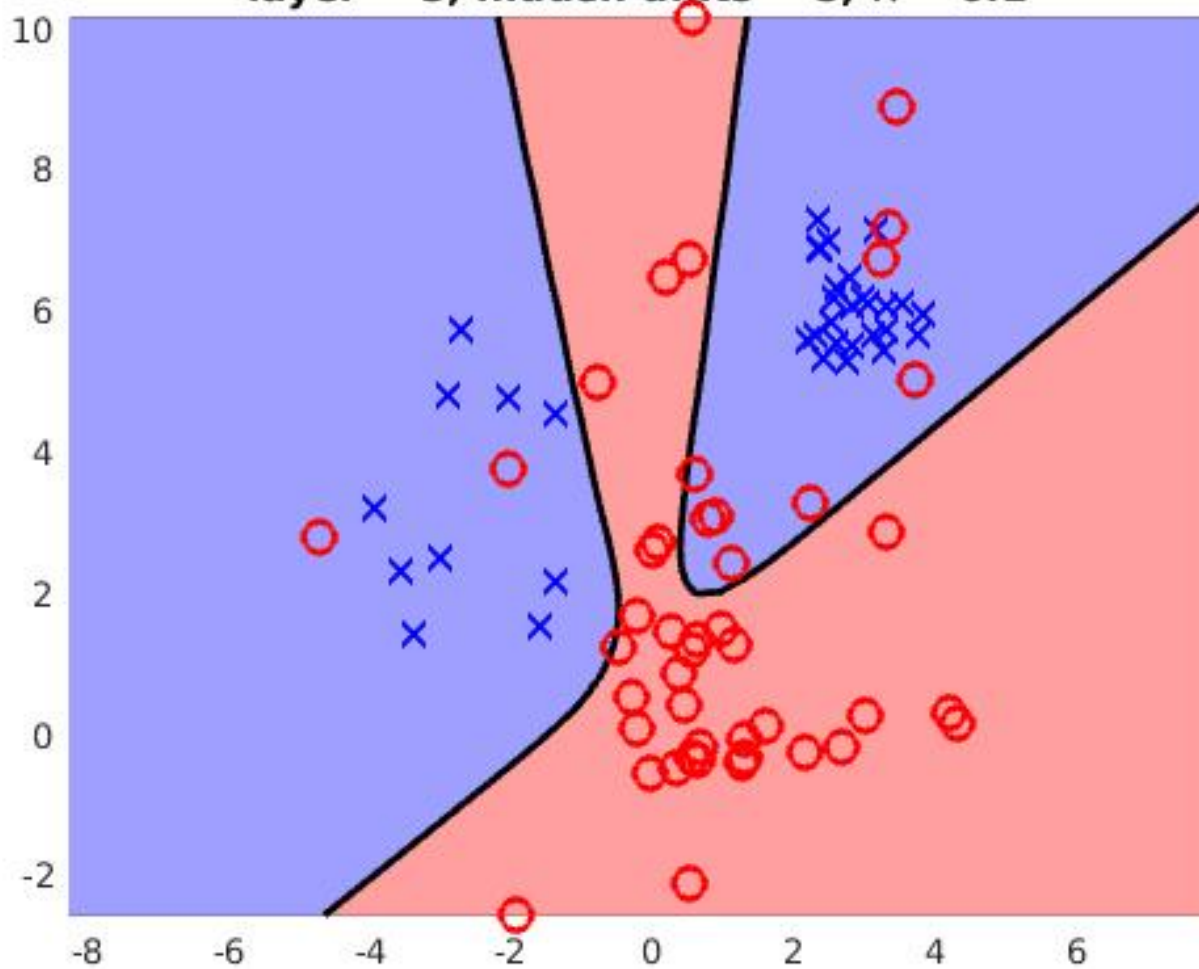
**layer = 3, hidden units = 5,  $\lambda = 0.001$**



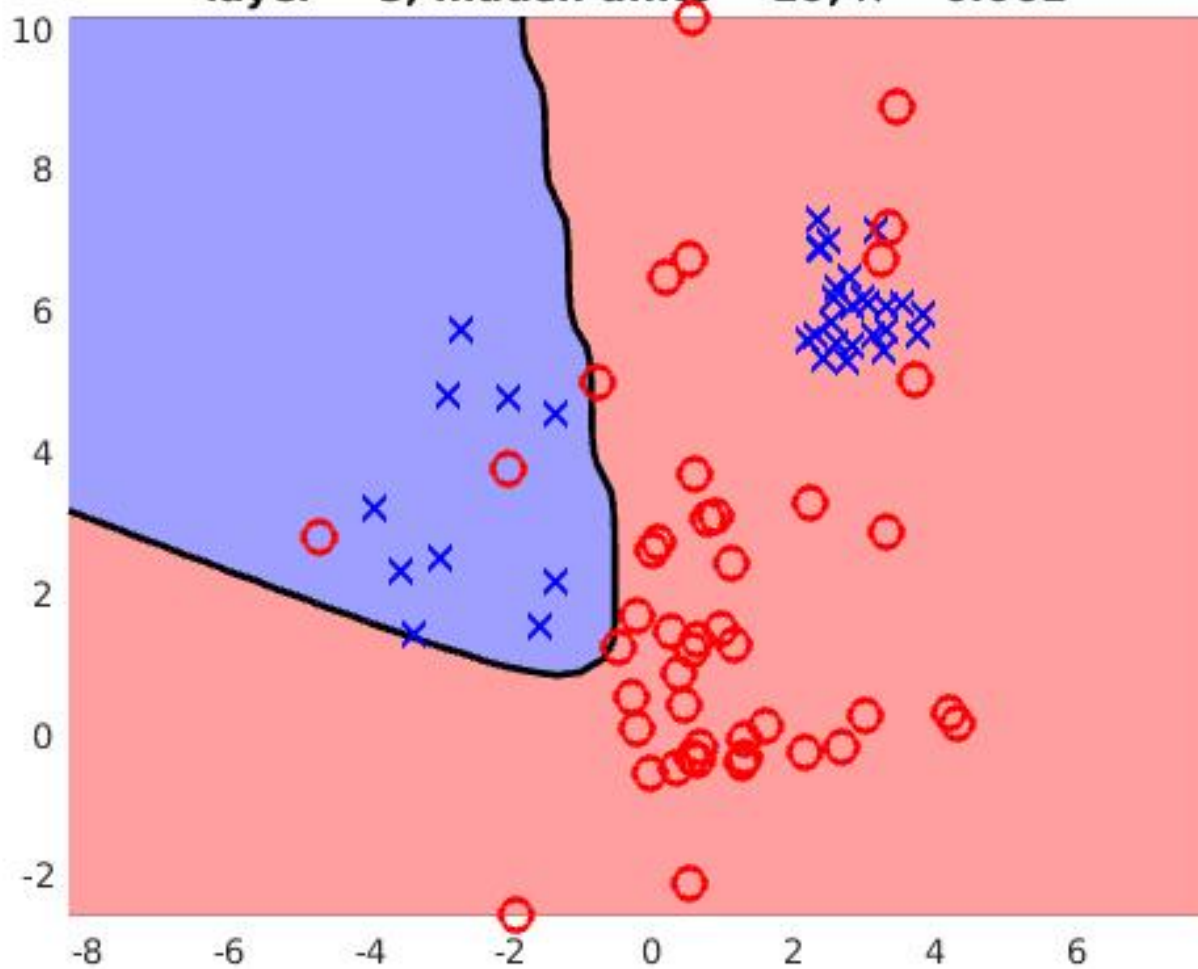
**layer = 3, hidden units = 5,  $\lambda = 0.01$**



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

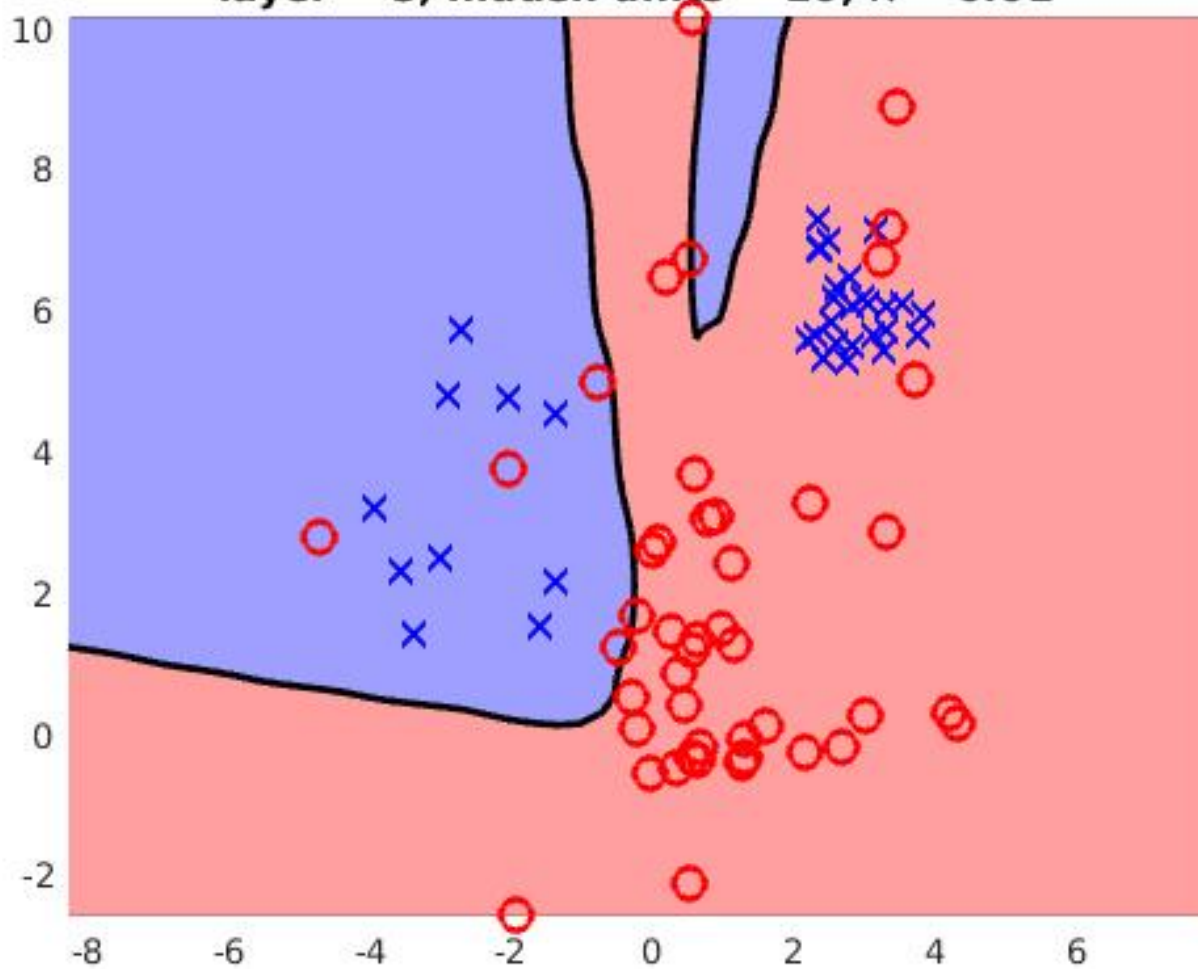


**layer = 3, hidden units = 20,  $\lambda = 0.001$**





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



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

