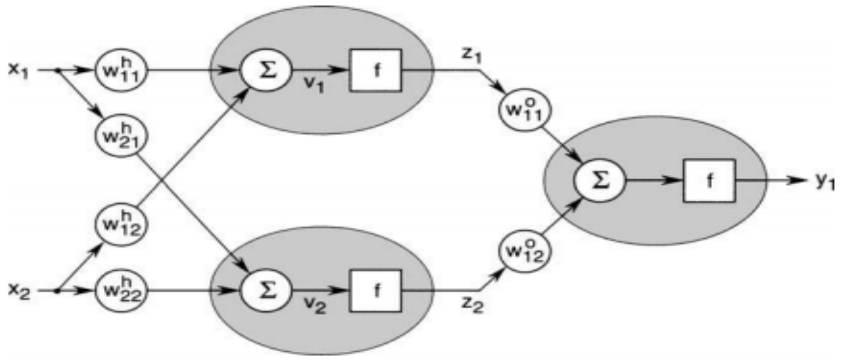
Project #2

Submit all work including computer codes

1. (40 points) Consider the neural network below with bias:

Fig. 1



to classify two classes of linearly non-separable data on the plane.

Fig. 1 with bias included is expressed as:

$$
\begin{aligned}
v_1 &= w_{11}^h x_{d1} + w_{12}^h x_{d2} - \theta_1, \\
v_2 &= w_{21}^h x_{d1} + w_{22}^h x_{d2} - \theta_2, \\
z_1 &= f(v_1), \\
z_2 &= f(v_2), \\
y_1 &= f\left(w_{11}^o z_1 + w_{12}^o z_2 - \theta_3\right),
\end{aligned}
$$

where θ_1, θ_2, θ_3 are bias parameters; (x_{d1}, x_{d2}) is input data.

Download training_data.txt and test_data.txt from class website. The 1st column contains labels 0 and 1, the 2nd and 3rd columns store coordinates of the data points. The data points are plotted in Fig. 2. Left: training_data. Right: test_data. Colors for labels: 0 (red), 1 (blue).

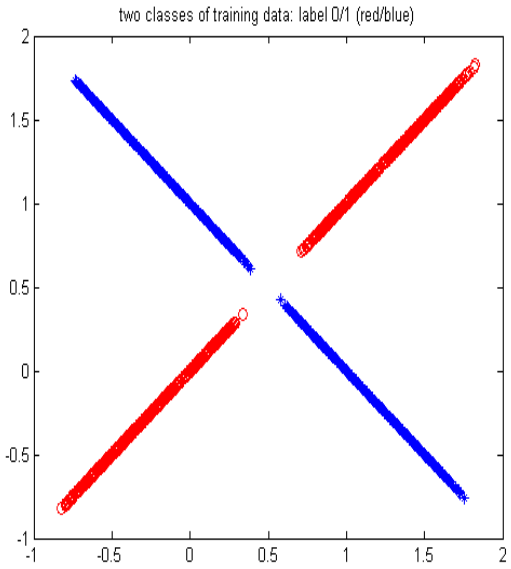
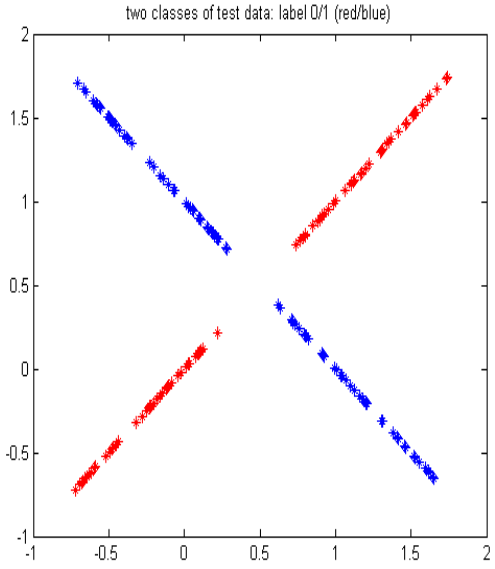


Fig.2



1) Program gradient descent with exact formulas or use tensorflow (e.g. tf.train.GradientDescentOptimizer)
to train the network in Fig. 1 (f = sigmoid) on training_data.txt.

Training loss can be either square loss or log loss. Select proper learning rate and number of epochs to descend and reach a small loss value (as much as possible).

2) Print and plot training loss vs. epoch numbers.

3) Round the network output to 0 or 1, then compute and report test accuracy (percentage of correct label predictions) on test_data.txt.