Q1 (a): In ipad.

(b): In ipad

(c): This formular is represent a circle in the coordinate system. With the input x and y, it’s output can be larger or smaller than 0. If it’s larger, the data is outside the circle, the opposite is inside the circle. So this perceptron can be useful as a classifier in this case.

(d): Pick random (w1, w2, theta) in (w1, w2, theta) space

Evaluate the residu in this point, R0 = R(w1, w2, theta)

Calculate the gradient in point using the sigmoid function.

Pick a new point that is lies a small step away from point in the opposite direction of the gradient.

If h small enough, then R is guaranteed to be smaller.

(e): One single layer of perceptron can only have one linearly decision boundary. And if we got 2 layers, there are 2 linearly boundary, which makes it non-linera.

Without activation function, the output of perceptron can be either 1 or 0. It’s binary, and can be meanless as the input of the next layer. With non linear activation, it’s possible to define an error sum for the perceptron. Improve and make the non-linear worked.

(f): The iris data has three kinds of flowers, one linearly decision line can only seprate 2 classes. And there is one flowers are in the between of the other two flowers. It’s also means that it’s not linearly classifiable.

(g): With single layer, the output for each node is binary, which means it’s only classify whether the data is this kind of flower or not. For the other node in the same layer is the same result. So we may need to process the output from the first layer to do further classify the unclassify data.

Q2: (a): The possibility of choosing which path is determined by the intensity of pheromone. If more and more ants choose one path, this is because it’s the shorter path and the pheromone is getting stronger. The stygmergy ants can communicate with each other by the pheromone.

It would be much simple, if we see each ant individually. It just follow the pheromone and release pheronmone. However, with stygmergy behaviour, they can easily find the shorter path to their destination.

(b): Make 20 random set of each nodes weights and biases.

Compare the output with labels for each sets of weights and biases. And calculate each sets’ fitness. Record the ratings.

Fitness can be regard with the loss. If the loss is low, the fitness is high.

Take the two that classify the most accuracy in the last round.

Cross them over at a random point to make two new set.

Repeat the random crossover until have another 20 sets.

Occasionally make an arbitrary change, which is mutation.

(c): Boltzmann Machines





