HW to Chapters 2 “The Perceptron”

Name: Jing Cao

NUID: 002899660

Non-programming Assignment

1. Describe the Perceptron and how it works:

The Perceptron is the simplest type of artificial neural network used for supervised learning of binary classifiers. It is a type of linear classifier, meaning it separates input data into two classes using a linear decision boundary. A perceptron works by aggregating input signals, multiplying them by corresponding weights, and applying an activation function (usually a step function). If the weighted sum exceeds a certain threshold, the perceptron produces a binary output (1 or 0). The perceptron learns by adjusting the weights based on the error between the predicted output and the actual target output during the training process.

1. What is forward and backpropagation propagation for the Perceptron?

Forward Propagation: This is the process where the perceptron computes the output based on the given input data. The input vector is multiplied by the corresponding weights, and the sum is passed through an activation function to generate the output. This process involves no weight updates and simply moves information through the network.

Backpropagation: This is the process used to update the weights of the perceptron after each training example. It compares the perceptron’s calculated output with the target output and calculates the error. This error is then used to adjust the weights in such a way that the perceptron will improve its prediction for the next training iteration. In the basic perceptron, this involves adjusting weights using a simple rule based on the error between the actual and expected output​.

1. What is the history of the Perceptron?

The perceptron was invented in 1943 by Warren McCulloch and Walter Pitts as the first artificial neuron. The first hardware implementation, the Mark I Perceptron machine, was developed in 1957 by Frank Rosenblatt at the Cornell Aeronautical Laboratory. Rosenblatt's perceptron machine was a significant milestone in neural networks and was funded by the U.S. Office of Naval Research. Despite its initial promise, the perceptron faced criticism in the late 1960s when Marvin Minsky and Seymour Papert highlighted its limitations, specifically its inability to solve non-linearly separable problems like the XOR function. This caused research in neural networks to stagnate for several years, only to be revived in the 1980s when multi-layer perceptrons were introduced​​.

1. What is Supervised Training?

Supervised training refers to the process where a neural network (like the perceptron) is trained using a labeled dataset. Each training example consists of an input vector and a known target output (label). The perceptron uses this information to adjust its weights to minimize the error between the calculated output and the target output. The goal of supervised training is to find the optimal weight configuration that best separates the input data into the correct classes​.

1. Why is Perceptron referred to as a binary linear classifier?

The perceptron is referred to as a binary linear classifier because:

Binary: It classifies the input into one of two possible categories (0 or 1, positive or negative).

Linear: The decision boundary that the perceptron uses to classify inputs is a straight line (or a hyperplane in higher dimensions), which makes it a linear model. The perceptron separates the input data into two classes by adjusting the weights to fit a linear boundary between them​​.

1. What are the disadvantages of binary linear classification?

The primary disadvantage of binary linear classification using the perceptron is that it can only solve linearly separable problems. If the data cannot be separated by a straight line, as is the case with problems like XOR, the perceptron will fail to classify the data correctly.

Other disadvantages include:

It cannot model nonlinear relationships between inputs and outputs.

It is sensitive to the initial weights and the chosen learning rate.

The perceptron may not converge (find a solution) if the dataset is not linearly separable​​.