Lesson 5: Perceptron

Table of Contents

5.1 Introduction

The perceptron is a type of artificial neural network that was first introduced in the 1950s. It is a simple model that is used to classify inputs into one of two classes. The perceptron consists of a single layer of artificial neurons, each of which takes in a set of inputs and produces an output. The output of the perceptron is determined by the weights and biases of the neurons, as well as the input values.
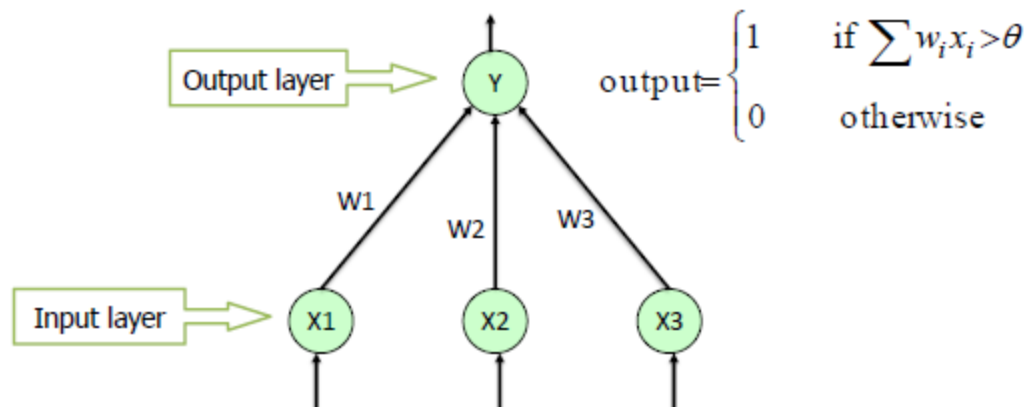
The perceptron is trained using a supervised learning algorithm, which adjusts the weights and biases of the neurons in order to minimize the error between the predicted output and the actual output. The training process continues until the perceptron is able to accurately classify the inputs into one of the two classes. One of the main advantages of the perceptron is its simplicity and ease of implementation. However, it is limited in its ability to solve problems that are not linearly separable, meaning that the two classes cannot be separated by a single straight line. Despite this limitation, the perceptron has played an important role in the development of more complex neural network models, such as multi-layer perceptron and convolutional neural networks.

In summary, the perceptron is a simple model of artificial neural network used for binary classification. It consists of a single layer of artificial neurons and it is trained using supervised learning algorithm. Despite its simplicity and easy implementation, it is limited in its ability to solve problems that are not linearly separable.

5.2 Single layer perceptron

A single layer perceptron is a type of artificial neural network that consists of only one layer of artificial neurons. The inputs are fed directly into the neurons, and the output is determined by the weights and biases of the neurons. The main advantage of a single layer perceptron is its simplicity, as it is easy to implement and train. However, it is limited in its ability to solve problems that are not linearly separable, meaning that the two classes of data cannot be separated by a single straight line. This is because it can only make linear decision boundaries, which is not sufficient for more complex data. Additionally, it can only solve binary classification problems and cannot handle more than two classes. Despite its limitations, the single layer perceptron has been an important step in the development of more advanced neural network models.

Single Layer Perceptron

$$output = \begin{cases} 1 & \text{if } \sum w_i x_i > \theta \\ 0 & \text{otherwise} \end{cases}$$

5.3 Learning algorithm

A learning algorithm is a method used to train an artificial neural network, such as a perceptron, to perform a specific task. The most commonly used learning algorithm for a perceptron is the Perceptron Learning Algorithm (PLA).The PLA is a supervised learning algorithm that adjusts the weights and biases of the neurons in the perceptron in order to minimize the error between the predicted output and the actual output. The algorithm works by iteratively updating the weights and biases based on the training data.

The basic steps of the PLA are:

1. Initialize the weights and biases of the perceptron to small random values.

2. For each training example, compute the predicted output of the perceptron and compare it to the actual output.

3. If the predicted output is different from the actual output, update the weights and biases of the perceptron in order to reduce the error.

4. Repeat steps 2 and 3 until the perceptron is able to accurately classify the training examples or a stopping criterion is met.

The PLA is a simple and effective algorithm for training a perceptron, but it can become stuck in a local minimum if the data is not linearly separable. The algorithm also requires that the training data is labeled, and it can only be used for binary classification.

5.4 Perceptron as maximum likelihood classifier

A perceptron can be used as a maximum likelihood classifier, which is a method of classifying data based on the probability that it belongs to a particular class. The goal of maximum likelihood classification is to assign an input to the class that has the highest probability of producing that input.

In the case of a perceptron, the maximum likelihood classifier is achieved by adjusting the weights and biases of the perceptron such that the probability of the perceptron producing the correct output for a

given input is maximized. This is done by minimizing the negative log-likelihood of the training data, which is equivalent to maximizing the likelihood of the training data.

To calculate the likelihood of an input, the perceptron uses a sigmoid function, which maps the input to a probability value between 0 and 1. The sigmoid function takes the dot product of the inputs and weights as input, and produces a probability value that represents the likelihood that the input belongs to a particular class.

The perceptron as a maximum likelihood classifier is a powerful method for classification as it allows the perceptron to model the probability of different classes, and thus it can handle multiple classes and more complex data. Additionally, it gives the possibility to use it as a probabilistic model that can make predictions with a degree of uncertainty.

5.5 Lesson 5 Questions
1. What is an artificial neural network?
2. What is the purpose of a perceptron?
3. How does a perceptron work?
4. What is a supervised learning algorithm?
5. What are the advantages and limitations of a single layer perceptron?
6. What is the difference between a perceptron and a multi-layer perceptron?
7. What are the most commonly used activation functions in neural networks?
8. How does backpropagation work in a neural network?
9. What are some popular applications of neural networks?
10. What are some challenges and limitations of training neural networks?