Introduction to the Perceptron

Generated by LectureMate

February 21, 2025

1 Introduction

The lecture, led by Jillian Kraft, focused on the perceptron, a fundamental neural network architecture. The perceptron serves as a model for a single neuron and is pivotal in understanding more complex neural networks. The session aimed to provide insights into how perceptrons function as classifiers and how they learn from data.

2 Overview of Neural Networks

Neural networks are often viewed as computational simulations of human behavior. The perceptron is the simplest form of a neural network, representing a single neuron. The lecture outlined the structure of a biological neuron, including its components: the cell body, dendrites, and axon.

2.1 Neuronal Structure

- **Dendrites:** Inputs from other neurons.
- **Axon:** Outputs to other neurons.
- **Synapse:** Connection between neurons that regulates firing based on input.

The perceptron simplifies this biological model into a computational device that processes numerical inputs.

3 The Perceptron Model

The perceptron takes multiple inputs, each associated with a weight. The output is determined by a weighted sum of the inputs, followed by an activation function that decides whether the neuron "fires."

3.1 Mathematical Representation

Let the inputs be denoted as $x_1, x_2, \ldots, x_n$ and their corresponding weights as $w_1, w_2, \ldots, w_n$. The output function can be expressed as:

$$y = f\left(\sum_{i=1}^{n} w_i x_i - \theta\right)$$

where θ is the threshold and f is the activation function.

3.2 Activation Function

The simplest activation function is a step function:

$$f(x) = \begin{cases} 1 & \text{if } x \geq 0 \\ 0 & \text{otherwise} \end{cases}$$

4 Logical Functions and the Perceptron

The perceptron can model simple logical functions such as AND and OR. For example, with two binary inputs and weights of 0.5 and a threshold of 1, the perceptron can compute the AND function.

4.1 AND Function Example

- Inputs: $(0, 0) \to 0$
- Inputs: $(1, 1) \to 1$

However, the perceptron cannot model the XOR function, which requires a more complex decision boundary.

5 Limitations of the Perceptron

The perceptron can only learn linearly separable functions. The XOR function is not linearly separable, which means it cannot be modeled by a single-layer perceptron.

5.1 Linearly Separable Functions

A function is linearly separable if a line (or hyperplane in higher dimensions) can be drawn to separate the different classes of outputs. The perceptron is limited to this class of functions.

6 Improving the Perceptron

To overcome the limitations of the perceptron, multiple layers of neurons can be used, leading to a multi-layer perceptron (MLP). This architecture allows for the modeling of more complex functions.

7 Learning in the Perceptron

The lecture also covered the perceptron learning rule, which adjusts weights based on training examples. The learning process involves:

- Initializing weights.

- Computing the output.
- Comparing the output to the target.
- Updating the weights based on the error.

The weight update rule is given by:

$$w_i \leftarrow w_i + \eta(t - y)x_i$$

where η is the learning rate, t is the target output, and y is the actual output.

8 Conclusion

The perceptron is a foundational concept in neural networks, illustrating how simple computational units can learn from data. While limited in its capabilities, understanding the perceptron sets the stage for exploring more complex architectures like multi-layer perceptrons.