SNS COLLEGE OF ENGINEERING

Kurumbapalayam(Po), Coimbatore – 641 107
Accredited by NAAC-UGC with 'A' Grade
Approved by AICTE, Recognized by UGC & Affiliated to Anna University, Chennai

Department of AI &DS

Course Name – 19AD602 DEEP LEARNING

III Year / VI Semester

Unit 1-INTRODUCTION
Topic: NEURAL NETWORKS

NEURAL NETWORKS

Case Study

Task: Predict handwritten digits (0-9) from images using a neural network.

Approach: Train a model on the MNIST dataset with one hidden layer and ReLU activation.

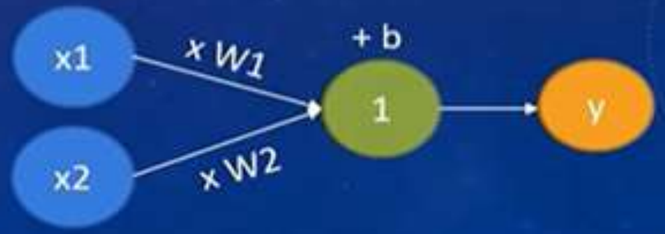
Outcome: The network achieves 95% accuracy in correctly classifying digits.

NEURAL NETWORKS

Neural Networks

- Structure: interconnected web of nodes called neurons and edges that join them together
- Layers
  - First Layer – input layer
  - Hidden Layer(s)
  - Final Layer

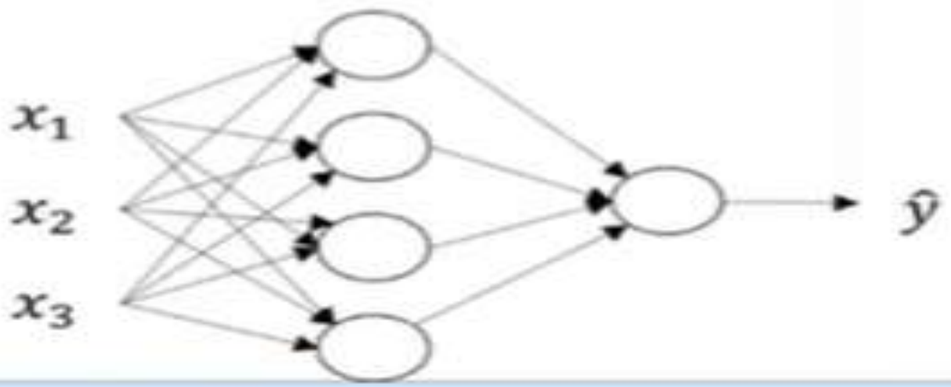
  ] Classifier (SVM, Logistic regression,...)
- Function
  - Receive a set of inputs
  - Perform complex calculations
  - Use the output to solve a problem
- Each edge has unique weight
- Each node has unique bias
- Firing of a classifier / Activation – produces a score
- Process of improving neural network's accuracy - training



NEURAL NETWORKS

Shallow Neural Networks:

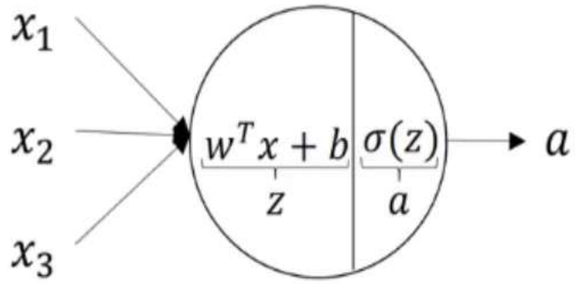
When we hear the name Neural Network, we feel that it consist of many and many hidden layers but there is a type of neural network with a few numbers of hidden layers. Shallow neural networks consist of only 1 or 2 hidden layers. Understanding a shallow neural network gives us an insight into what exactly is going on inside a deep neural network. In this post, let us see what is a shallow neural network and its working in a mathematical context. The figure below shows a shallow neural network with 1 hidden layer, 1 input layer and 1 output layer.



NEURAL NETWORKS

THE NEURON:

- The neuron is the atomic unit of a neural network.
- Given an input, it provides the output and passes that output as an input to the subsequent layer.
- A neuron can be thought of as a combination of 2 parts:
- The first part computes the output ***Z***, using the inputs and the weights.
- The second part performs the activation on **Z** to give out the final output **A** of the neuron.



NEURAL NETWORKS

THE HIDDEN LAYER:

- The hidden layer comprises of various neurons, each of which performs the above 2 calculations.
- The 4 neurons present in the hidden layer of our shallow neural network compute the following:

$$z_1^{[1]} = w_1^{[1]T} x + b_1^{[1]}, a_1^{[1]} = \sigma\left(z_1^{[1]}\right)$$
$$z_2^{[1]} = w_2^{[1]T} x + b_2^{[1]}, a_2^{[1]} = \sigma\left(z_2^{[1]}\right)$$
$$z_3^{[1]} = w_3^{[1]T} x + b_3^{[1]}, a_3^{[1]} = \sigma\left(z_3^{[1]}\right)$$
$$z_4^{[1]} = w_4^{[1]T} x + b_4^{[1]}, a_4^{[1]} = \sigma\left(z_4^{[1]}\right)$$

NEURAL NETWORKS

In the above equations,

- The superscript number ***[i]*** denotes the layer number and the subscript number ***j*** denotes the neuron number in a particular layer.
- ***X*** is the input vector consisting of 3 features.
- ***W[i]j*** is the weight associated with neuron ***j*** present in the layer ***i***.
- ***b[i]j*** is the bias associated with neuron ***j*** present in the layer ***i***.
- ***Z[i]j*** is the intermediate output associated with neuron ***j*** present in the layer ***i***.
- ***A[i]j*** is the final output associated with neuron ***j*** present in the layer ***i***.
- ***Sigma*** is the sigmoid activation function. Mathematically it is defined as:

$$\sigma(x) = \frac{1}{1 + e^{-x}}$$

NEURAL NETWORKS

- As we can see, the above 4 equations seem redundant. Therefore we will vectorize them as:

$$Z^{[1]} = X^{[1]T} X + b^{[1]}$$

$$A^{[1]} = \sigma\left(Z^{[1]}\right)$$

1.The first equation computes all the intermediate outputs ***Z*** in single matrix multiplication.
2.The second equation computes all the activations **A** in single matrix multiplication.

NEURAL NETWORKS

SHALLOW NEURAL NETWORKS:

- A neural network is built using various hidden layers.
- Now that we know the computations that occur in a particular layer, let us understand how the whole neural network computes the output for a given input ***X***.
- These can also be called the ***forward-propagation*** equations.

$$Z^{[1]} = W^{[1]T} X + b^{[1]}$$

$$A^{[1]} = \sigma\left(Z^{[1]}\right)$$

$$Z^{[2]} = W^{[2]T} A^{[1]} + b^{[2]}$$

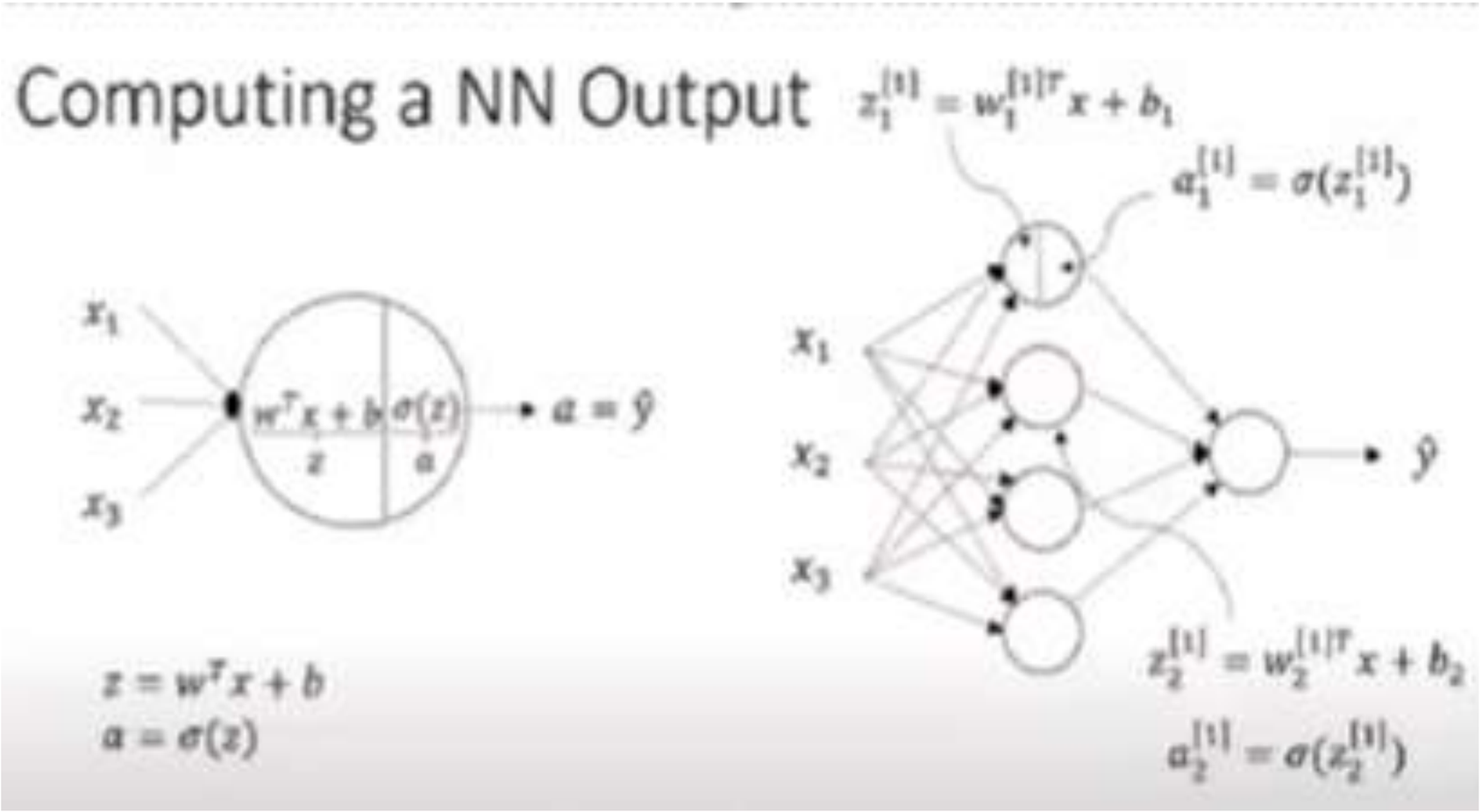
$$\hat{y} = A^{[2]} = \sigma\left(Z^{[2]}\right)$$

NEURAL NETWORKS

- The first equation calculates the intermediate output ***Z[1]*** of the first hidden layer.
- The second equation calculates the final output ***A[1]*** of the first hidden layer.
- The third equation calculates the intermediate output ***Z[2]*** of the output layer.
- The fourth equation calculates the final output ***A[2]*** of the output layer which is also the final output of the whole neural network.

NEURAL NETWORKS

Computing a NN Output



NEURAL NETWORKS

Activity

Objective: Show how inputs pass through layers of a neural network.

1. Write down three features (e.g., age, income, education) as inputs on sticky notes.
2. Draw a network with input, hidden, and output nodes; calculate outputs manually using simple weights.
3. Discuss how activations (like ReLU) and weights influence the prediction.

THANK YOU