

WORKING OF ARTIFICIAL NEURAL NETWORK

Artificial Neural Network (ANN) is a computational model inspired by the biological neural network of the human brain. It is widely used in Artificial Intelligence and Machine Learning for solving complex problems such as classification, prediction, speech recognition, and image processing. ANN consists of interconnected neurons organized into layers. These neurons process input data, apply mathematical transformations, and produce meaningful outputs.

1. Structure of ANN

ANN consists of three main layers: Input Layer, Hidden Layer(s), and Output Layer. The input layer receives raw data. Hidden layers perform computations and feature extraction. The output layer provides final predictions. Each neuron is connected with weights and bias values.

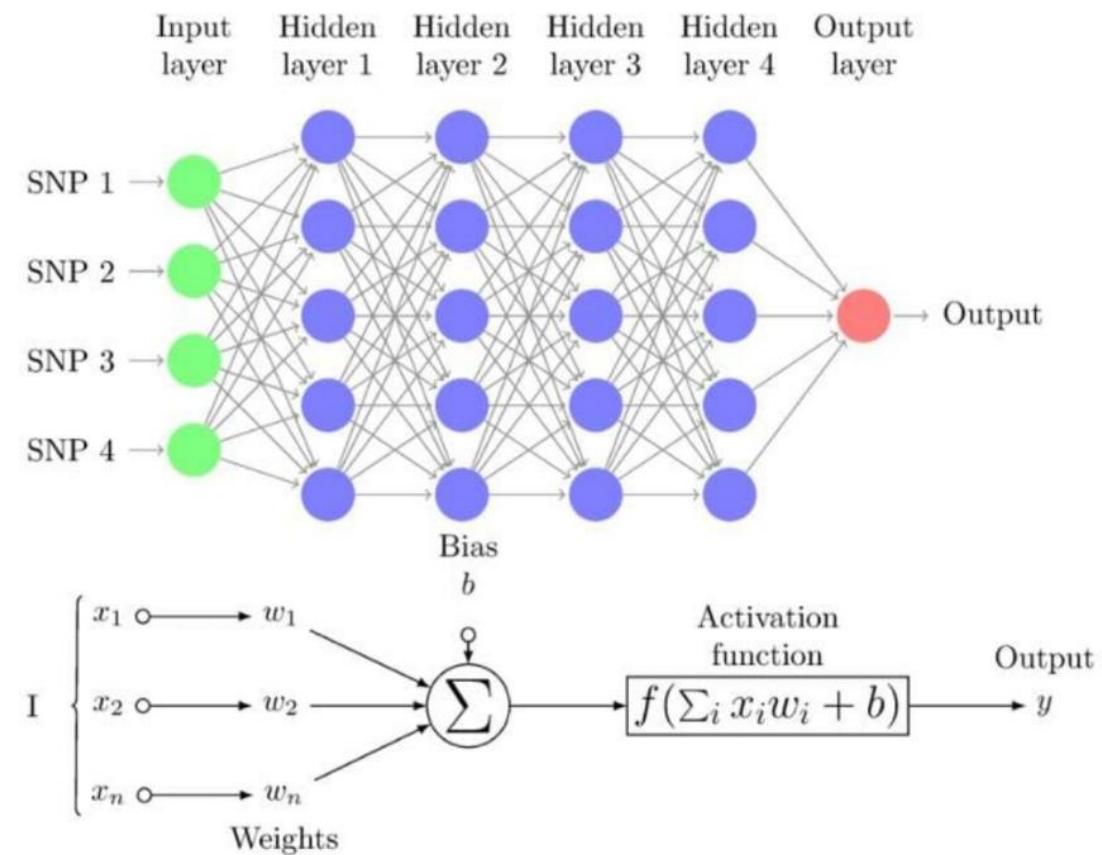


Figure 1: Multilayer Perceptron Architecture

2. Working Mechanism

The working mechanism of an Artificial Neural Network (ANN) begins when the input layer receives raw data in the form of numerical values. Each input is multiplied by a corresponding weight, and all the weighted inputs are summed together along with a bias value. This result is then passed through an activation function such as Sigmoid, ReLU, or Tanh to introduce non-linearity and produce the output of the neuron. The output from one layer becomes the input to the next layer, and this process continues until the final output layer generates the prediction or classification result. During training, the network performs forward propagation to compute the output, calculates the error using a loss function by comparing the predicted output with the actual value, and then applies backpropagation to adjust the weights using a gradient descent algorithm. This process is repeated multiple times until the error is minimized, enabling the network to make accurate predictions on new data.

3. Activation Functions

An activation function is a mathematical function used in an Artificial Neural Network (ANN) to determine the output of a neuron. After calculating the weighted sum of inputs and adding the bias, the result is passed through the activation function. Its main purpose is to introduce non-linearity into the network so that it can learn and solve complex problems. Without an activation function, the neural network would behave like a simple linear model and would not be able to handle complicated patterns such as image recognition or speech processing.

Some common activation functions are Sigmoid, which produces output between 0 and 1 and is mainly used for binary classification; ReLU (Rectified Linear Unit), which outputs zero for negative values and the same ...

4. CNN Architecture

Convolutional Neural Networks (CNN) are mainly used for image processing tasks. They consist of convolution layers, pooling layers, and fully connected layers. Convolution layers extract spatial features, pooling reduces dimensionality, and fully connected layers perform classification.

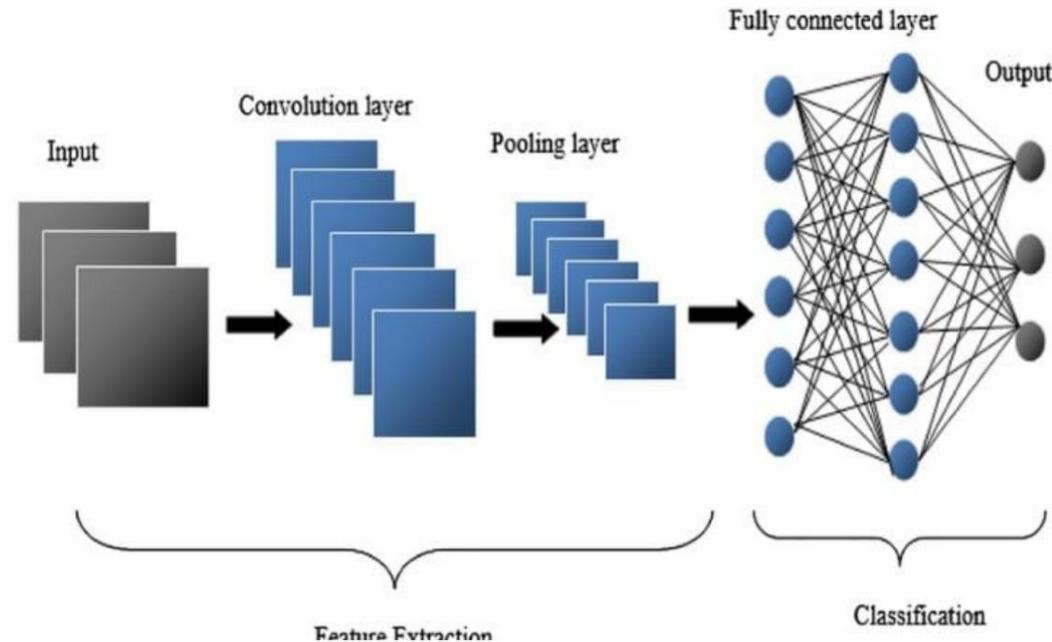


Figure 2: CNN Architecture Flow

5. RNN Architecture

Recurrent Neural Networks (RNN) are used for sequential data such as text and speech. Unlike traditional networks, RNN has memory that stores previous information. LSTM is an advanced form of RNN designed to overcome long-term dependency problems.

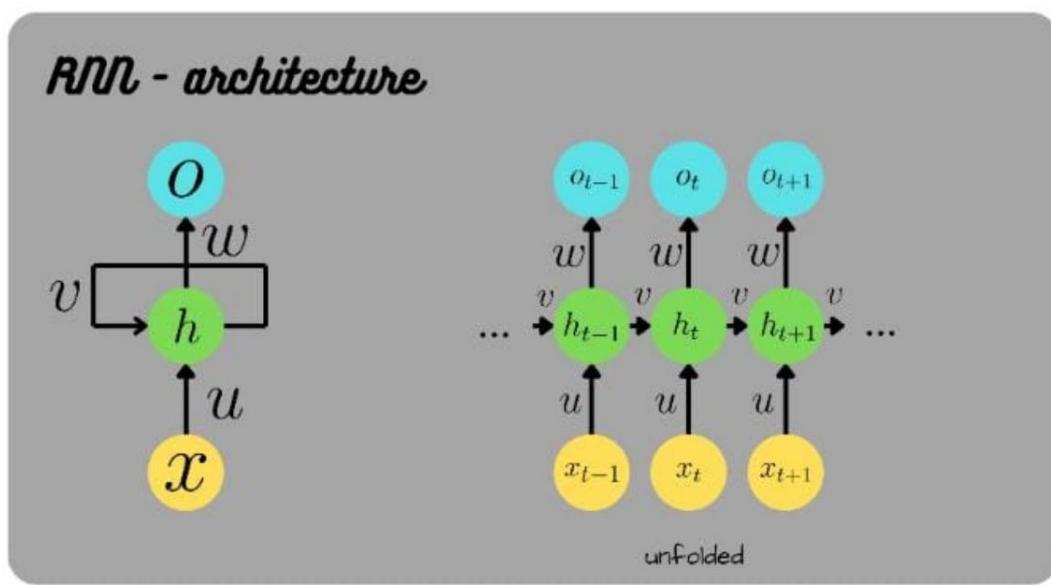


Figure 3: RNN Structure

6. Advantages and Disadvantages

Advantages of Artificial Neural Network (ANN)

1. Learns automatically from data.
2. Can solve complex and non-linear problems.
3. High accuracy in prediction and classification tasks.
4. Capable of handling large amounts of data.
5. Supports parallel processing.

Disadvantages of Artificial Neural Network (ANN)

1. Requires a large amount of training data.
2. Training process can be time-consuming.
3. High computational cost.
4. Difficult to interpret results (black box model).
5. Selection of proper architecture is complex.

7. Conclusion

Artificial Neural Networks have revolutionized modern Artificial Intelligence. With architectures like MLP, CNN, and RNN, ANN systems can solve complex real-world problems. Despite high computational requirements, their performance and adaptability make them highly valuable.