Basics of Neural Network

Sayeda Suaiba Anwar
Lecturer, Dept. of CSE
East Delta University

Neuron

A **neuron** is the fundamental cellular unit of the nervous system.

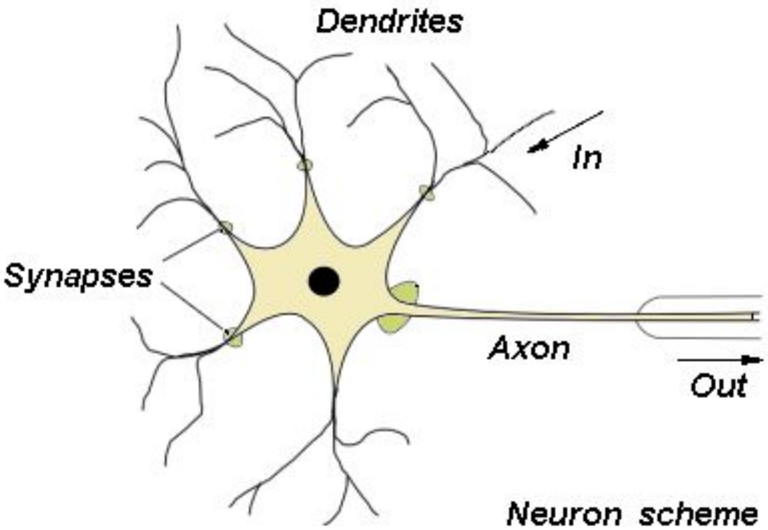
Deep Learning studies the basic unit of a brain called a brain cell or a neuron.

Parts of neuron is given below:

- **Dendrite**: Receives signals from other neurons
- **Cell Body (Soma)**: Sums all the inputs
- **Axon**: It is used to transmit signals to the other cells

As we know that our brain consists of multiple connected neurons forming a neural network, we can also have a network of artificial neurons called **perceptrons** to form a Deep neural network.

Neuron



Biological	Artificial
Dendrites	Inputs
Cell Nucleus	Nodes
Synapses	Weights
Axon	Outputs

Perceptron

An **Artificial Neuron** or a **Perceptron** is a simple model of a biological neuron in an artificial neural network.

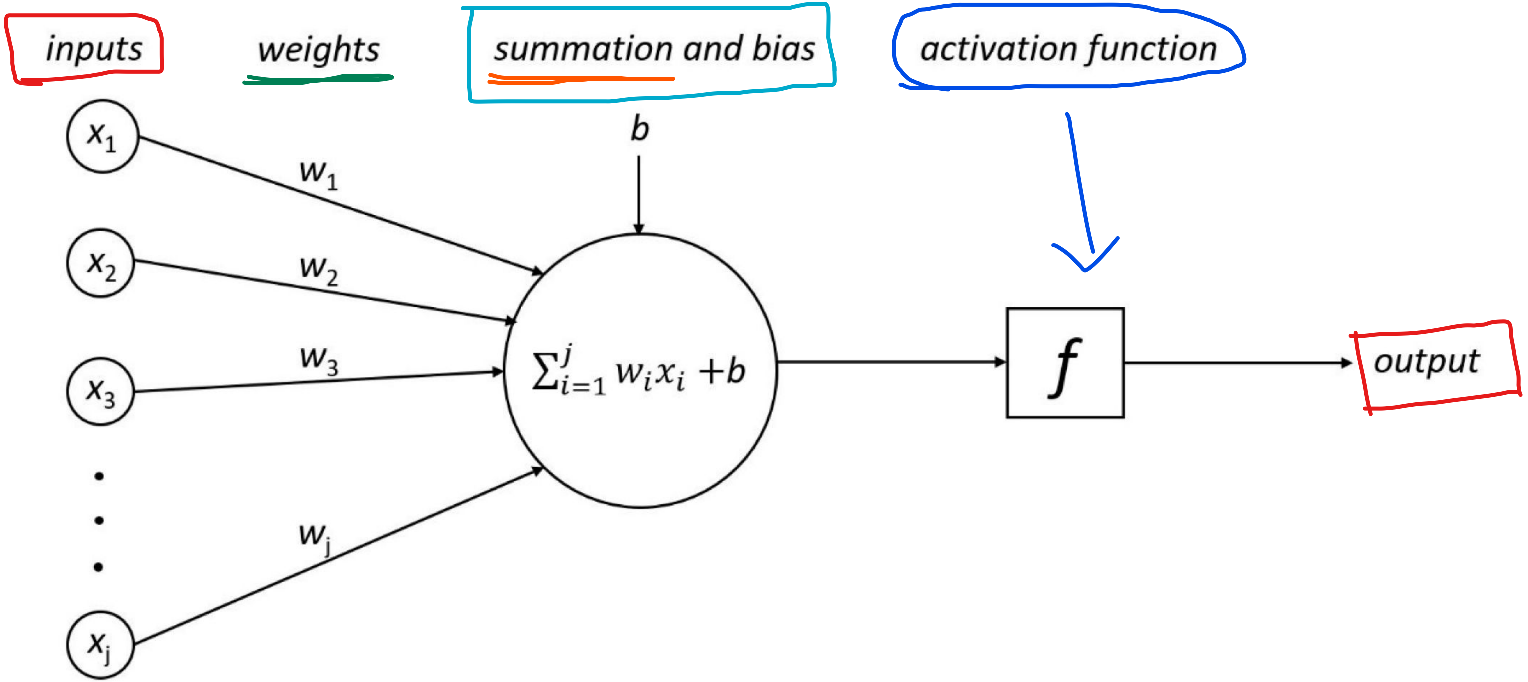
It models a neuron which has a set of inputs, each of which is given a specific weight.

It receives n inputs and then sums those inputs, applies a transformation and produces an output.

It has two functions:

- Summation
- Transformation(Activation)

Schematic Representation of a Perceptron



Weight and Bias

Weight:

The weight shows the effectiveness of a particular input. More the weight of input, more it will have an impact on the neural network.

Bias:

Bias is an additional parameter in the Perceptron which is used to adjust the output along with the weighted sum of the inputs to the neuron which helps the model in a way that it can fit best for the given data.

Bias

Bias is the error that is introduced by the model's prediction and the actual data.

Bias = Predicted – Actual

High Bias means the model has created a function that fails to understand the relationship between input and output data.

Low Bias means the model has created a function that has understood the relationship between input and output data.

There is only bias which is **unique** to every layer of the neural network.

Variance

Variance of a deep learning model is the amount by which its performance varies with different portions of training data set.

Low variance means the performance of the model does not vary much with the different data set.

High variance means the performance of the model varies considerably with different data set.

- Models with high bias will have low variance.
- Models with high variance will have a low bias.

Variance

A model that does not match a data set with a high bias will create an inflexible model with a low variance that results in a suboptimal machine learning model.

Characteristics of a **high variance model** include:

- Noise in the data set
- Potential towards overfitting
- Complex models
- Trying to put all data points as close as possible

	under	over
Tain	✗	✓
Test	✗	✗

Underfitting & Overfitting

The terms **underfitting** and **overfitting** refer to how the model fails to match the data. The fitting of a model directly correlates to whether it will return accurate predictions from a given data set.

Underfitting:

- If a model fails to show satisfactory accuracy during the training phase itself it means the model is underfitting.
- occurs when the model is unable to match the input data to the target data.
- This happens when the model is not complex enough to match all the available data.

Underfitting & Overfitting

Overfitting:

- If a model is showing high accuracy during the training phase but fails to show similar accuracy during the testing phase it indicates overfitting.
- This occurs when dealing with highly complex models where the model will match almost all the given data points and perform well in training datasets.
- However, the model would not be able to generalize the data point in the test data set to predict the outcome accurately.

Bias vs. Variance : A Trade-Off

- Bias and variance are inversely connected.

Low Bias, High Variance : The model will fit with the data set while increasing the chances of inaccurate predictions.

High Bias, Low Variance : It will reduce the risk of inaccurate predictions, the model will not properly match the data set.

- Having a higher variance does not indicate a bad ML algorithm. Machine learning algorithms should be able to handle some variance.

Bias vs. Variance : A Trade-Off

We can tackle the trade-off in multiple ways:

Increasing the complexity of the model to count for bias and variance, thus decreasing the overall bias while increasing the variance to an acceptable level.

Increasing the training data set can also help to balance this trade-off, to some extent. This is the preferred method when dealing with overfitting models.
A large data set offers more data points for the algorithm to generalize data easily. However, the major issue with increasing the trading data set is that underfitting or low bias models are not that sensitive to the training data set.

Activation Function

An **activation function** is a function that is added into an artificial neural network in order to help the network learn complex patterns in the data.

Activation functions add **non-linearity** to the model.

There are many functions that are used as Activation Functions, like:

- Step Function
- Linear Function
- Sigmoid Function
- Tanh Function
- ReLU Function (Rectifier Linear Unit)
- Leaky ReLU Function

Necessity of Activation Function

- An artificial neuron without an activation function will just produce the sum of dot products between all inputs and their weights.

- So, without activation function the network will produce a linear output and there is not any necessary of hidden layer.

- But, in the real world hardly any data is so simple that it can be represented with just a linear mapping function. The activation function added this non-linearity property especially in the neurons of hidden layers to the network.

Necessity of Activation Function

- Thus we can say that Neural networks are used to implement complex functions, and non-linear activation functions enable them to approximate arbitrarily complex functions.

- Without the nonlinearity introduced by the activation function, multiple layers of a neural network are equivalent to a single layer neural network.

Think and Answer

How can you solve the problem of overfitting and underfitting?