

Complete Overview

1

Artificial Neural Network (ANN)

2

Deep Learning

3

Weight

4

Bias

5

Activation Function

6

Multilayer Perceptron (MLP)

7

Convolutional Neural Network (CNN)

8

Recurrent Neural Network (RNN)

9

Long Short-Term Memory (LSTM)

10

Transformer Models

11

Generative Adversarial Network (GAN)

12

Backpropagation

13

Gradient Descent

14

Optimizer

15

Learning Rate

16

Loss Function

17

Epoch

18

Batch Size

19

Overfitting

20

Underfitting

21

Regularization

22

Dropout

23

Early Stopping

24

Batch Normalization

25

Transfer Learning

26

Fine-tuning

27

Attention Mechanism

28

Word Embeddings

29

Precision

30

F1 Score

1 Artificial Neural Network (ANN)

ANN stands for Artificial Neural Network. In data science, it refers to a computational model loosely inspired by the structure and function of the human brain.

2 Deep Learning

Deep Learning is associated with a machine learning algorithm (Artificial Neural Network, ANN), which uses the human brain concept to facilitate modeling arbitrary functions. ANN requires a vast amount of data, and this algorithm is highly flexible when it comes to modeling multiple outputs simultaneously.

3 Weight

A parameter within a neural network that transforms input data within the network's layers. It is adjusted during training so that the network predicts the correct output.

4 Bias

A term added to a model's output that allows the model to represent patterns that do not pass through the origin.

5 Activation Function

The activation function calculates a weighted total and then adds bias to decide whether a neuron should be activated. It aims to introduce non-linearity into a neuron's output. Examples include sigmoid, ReLU (Rectified Unit), and tanh.

6 Multilayer Perceptron (MLP)

A type of neural network consists of at least three layers of nodes: an input layer, one or more hidden layers, and an output layer. Unlike CNNs or RNNs, MLPs are fully connected, meaning each neuron in one layer connects to every neuron in the following layer.

7 Convolutional Neural Network (CNN)

Convolutional Neural Networks (CNNs) are a powerful type of deep learning model that excels at processing data with a grid-like structure, primarily images. They are inspired by how the human visual cortex functions and are particularly adept at tasks like image recognition, object detection, and image segmentation.

8 Recurrent Neural Network (RNN)

RNN stands for Recurrent Neural Network. Unlike traditional ANNs that process data point by point, RNNs are specifically designed to handle sequential data, where the order of information matters.

9 Long Short-Term Memory (LSTM)

A special kind of RNN is capable of learning long-term dependencies, including gates that regulate the flow of information.

10 Transformer Models

A type of neural network architecture that eschews recurrence and instead relies entirely on self-attention mechanisms to draw global dependencies between input and output, which has been revolutionary in tasks like translation and text generation.

11 Generative Adversarial Network (GAN)

Generative Adversarial Network (GAN): Ian Goodfellow and his colleagues designed a class of machine learning frameworks where two neural networks compete in a game.

12 Backpropagation

In neural networks, if the estimated output is far from the actual output (high error), we update the biases and weights based on the error. This weight and bias updating process is known as Back Propagation. Back-propagation (BP) algorithms determine the loss (or error) at the output and then propagate it back into the network. The weights are updated to minimize the error resulting from each neuron.

13 Gradient Descent

Gradient descent is a first-order iterative optimization algorithm used to find the minimum of a function. We use a gradient descent algorithm in machine learning algorithms to minimize the cost function. It finds out the best set of parameters for our algorithm.

14 Optimizer

An algorithm or method is used to change the neural network's attributes, such as weights and learning rate, to reduce the losses. Common optimizers include SGD (Stochastic Gradient Descent), Adam, and RMSprop.

15 Learning Rate

The size of the training algorithm's step on the loss surface. A smaller learning rate might make the training more reliable but also make it slower to converge.

16 Loss Function

A function that measures how well the neural network models the expected outcome.

17 Epoch

This deep learning term – epoch, refers to a single complete pass of the training dataset through a machine learning model. Imagine a loop where you train the model on all your data points once. Each completion of that loop is considered an epoch.

18 Batch Size

The number of training examples utilized in one iteration (a single batch) of the model training.

19 Overfitting

A model is said to overfit when it performs well on the training dataset but fails on the test set. This happens when the model is too sensitive and captures random patterns that are present only in the training dataset.

20 Underfitting

Underfitting occurs when a statistical model or machine learning algorithm cannot capture the underlying trend of the data. It refers to a model that neither models on the training data nor generalizes to new data. An underfit model is unsuitable as it will perform poorly on the training data.

21 Regularization

Regularization is a technique used to solve the overfitting problem in statistical models. In machine learning, regularization penalizes the coefficients so that the model can be generalized better.

22 Dropout

A regularization technique for neural networks that prevents overfitting by randomly setting a fraction of input units to zero at each update during training.

23 Early Stopping

Regularization is used to avoid overfitting when training learners with an iterative method, such as gradient descent. Training is stopped as soon as the performance on a validation dataset starts to degrade.

24 Batch Normalization

A technique to improve the training of deep neural networks that normalizes the inputs to a layer for each mini-batch.

25 Transfer Learning

Transfer learning is applying a pre-trained model to a completely new dataset. A pre-trained model is a model created by someone to solve a problem. This model can be applied to solve a similar problem with similar data.

26 Fine-tuning

The process of taking a pre-trained deep learning model (the network has already been trained on a related task) and continuing the training on a new dataset specific to a second task, which can be smaller in size, leveraging the learned features.

27 Attention Mechanism

A component in complex neural networks, particularly in sequence-to-sequence models, allows the network to focus on different parts of the input sequentially rather than considering the whole input simultaneously, improving the performance in tasks like machine translation.

28 Word Embeddings

A type of word representation that allows words with similar meanings to have a similar representation. They are a set of language modeling and feature learning techniques in NLP where words or phrases from the vocabulary are mapped to vectors of real numbers.

29 Precision

A metric that quantifies the number of correct positive predictions made. It is defined as the number of true positives divided by the number of true positives plus the number of false positives.

30 F1 Score

A measure of a test's accuracy and considers both the precision and the recall of the test to compute the score: $2 * (\text{precision} * \text{recall}) / (\text{precision} + \text{recall})$. It is particularly useful when the class distribution is uneven.