Deep Learning Concepts

Adapted from: https://towardsdatascience.com/the-beginners-glossary-of-neural-network-terms

Artificial Neural Network (NN):

A stack of simple learning algorithms (called layers) that sequentially process the input, producing an output. This embodies the idea of deep learning by design, as each layer learns a more refined understanding of the input.

Network Layers:

All neural networks are made up of layers. A layer defines an operation that takes some inputs, some parameters, and produces a set of outputs.

Dense Layer:

Dense Layer is a simple layer of neurons in which each neuron receives input from all the neurons of previous layer

Activation Function:

The activation function defines the output of a neuron / node given an input or set of input (output of multiple neurons). It’s the mimic of the stimulation of a biological neuron.

ReLU:

Short of Rectified Linear Unity, defined as ReLU(x) = max(0, x). This is one of the simplest (and most efficient) ways of making something non-linear.

Dense Network / Fully-Connected Network / Multi-Layer Perceptron:

A collection of dense layers interleaved with activation functions.

Perceptron Model:

In 1958, Frank Rosenblatt created a computational model of a neuron: the perceptron, which is the basis for most of what came afterward. Simply put, the perceptron is a weighted sum followed by an activation function.

Weights:

A weight represent the strength of the connection between units. In other terms they are real values that tell the importance of a feature in predicting the final value. The DL algorithms learn by updating/improving weights

Training:

The actual learning is performed by the training loop. In simplified terms, training means to feed inputs to a model, collect its outputs, compare them with the expected outputs, and change the weights to make to correct the outputs.

Loss Function:

The loss, or error, is a function that measures the wrongness of the model. In practice, all we want is to reduce the loss of our models. In other words, we want them to be as correct as possible.

Backpropagation:

An algorithm used in artificial intelligence (AI) to fine-tune mathematical weight functions and improve the accuracy of an artificial neural network's outputs.

Gradient Descent:

Gradient Descent is an optimization algorithm for finding a local minimum of a differentiable function.

Epochs:

An epoch means training the neural network with all the training data for one cycle. In an epoch, we use all of the data exactly once.

A forward pass and a backward pass together are counted as one pass

An epoch is made up of one or more batches, where we use a part of the dataset to train the neural network.

Batch:

A group of training samples. We often have more data than we can fit in memory or in GPU. Thus, we cannot compute the gradient for all examples we have. Instead, we can compute it for a small subset (a batch) and apply it

SGD / RMSprop / Adam / Nadam / Radam / Ranger / etc:

Optimizers: algorithms that enhance backpropagation to make it faster and better.

Learning Rate:

The learning rate is a hyperparameter that controls how much to change the model in response to the estimated error each time the model weights are updated.