

# Neural Networks and Deep Learning (NNDL) FAQ

## 1) What is the place of NNDL in Machine Learning (ML)?

ML has many algorithms each solving a specific problem. NNDL algorithms can solve all of these problems with the same accuracy. Often the same NNDL algorithm can be used to solve many problems.

More importantly there many difficult problems that cannot be solved by other ML algorithms. They can only be solved by NNDL algorithms.

## 2) What are the fundamental NNDL algorithms?

Artificial Neural Network (ANN), Convolutional Neural Network (CNN), and Recurrent Neural Network (RNN).

## 3) What type of Learning do these neural networks use?

All of them use Supervised Learning.

## 4) What is supervised learning?

The network uses labeled data. The input variables and the corresponding outputs are collected from past experience. This collection is used to train the network.

## 5) What happens during the training?

The output of a Neural Network depends on the input and certain variables inside the program called the parameters. During training, the parameters are slowly adjusted.

## 6) Is there a strict definition for the parameters?

All the weights and biases of all the neurons. For CNNs include the filter coefficients.

## 7) What adjusts the parameters?

The Backpropagation Algorithm.

## 8) Where did the backpropagation algorithm come from?

Gradient Descent Algorithm from Convex Optimization.

## 9) What does Gradient Descent Algorithm do?

Find the location of the global minimum point of a Convex Multivariable Function.

## 10) What does backpropagation algorithm do?

Consider the loss or error at the output of a neural network as a function of the parameters and find the parameter values that minimize loss or error.

11) What are the most useful loss or error functions?

Some of Squared Error (for regression), Binary Cross-entropy (for binary classification), and Categorical Cross entropy for multi-class classification).

12) What is training?

The available labeled-data is roughly divided equally into two parts. One part is used for training. During training, each set of input variables are fed to the input of the network, the output of the network is calculated and compared to the label, the error or loss is calculated, and back propagation algorithm is used to update the parameters. Now this is repeated with the next set of data.

13) What is testing?

After training is over, the parameters are fixed and the overall error or loss is calculated using the second part of the labeled-data.

14) What is an Artificial Neural Network (ANN)?

ANN is the simplest neural network that can be used to perform classification or regression. Studying ANN is crucial in understanding CNN and RNN.

An ANN consists of fully connected layers of neurons (followed by activating functions) that takes in a 1-dimensional array and produces a class label (classification) or a rational number (regression or function approximation).

15) Is regression very different from classification, mathematically speaking?

No. Classification is a subset of regression. Regression approximates a continuous function. Classification approximates a discontinuous function.

16) How many layers are necessary in an ANN?

Two layers are sufficient for any classification or regression problem.

17) What does fully connected mean?

The output of each neuron in a particular layer is connected as inputs to all the neurons in the next layer.

18) What are in the layers of an ANN?

Neurons followed by activating functions. Each neuron's output passes through an activation function before being connected to the inputs of the neurons of the next layer.

19) What is the most important activating function?

ReLU. This is the best for ANNs and CNNs. RNN use Sigmoid and Tanh as well.

20) Why do we use activating functions?

Without the activating functions, the output of a neural network would be a linear function of the inputs (not very useful).

21) Is that all what activating functions do?

No. The ReLu activating function has the effect of dividing the domain in to sub-regions and selecting particular neurons in the first layer to be responsible for the output of the network. In each sub-region, some neurons' outputs will be zero, while others will be non-zero.

22) Do neural networks use polynomials to approximate functions?

No. They use piece-wise-linear functions (when we use ReLu). Recall that ReLu is piece-wise linear. Therefore their composition is piece-wise-linear too.