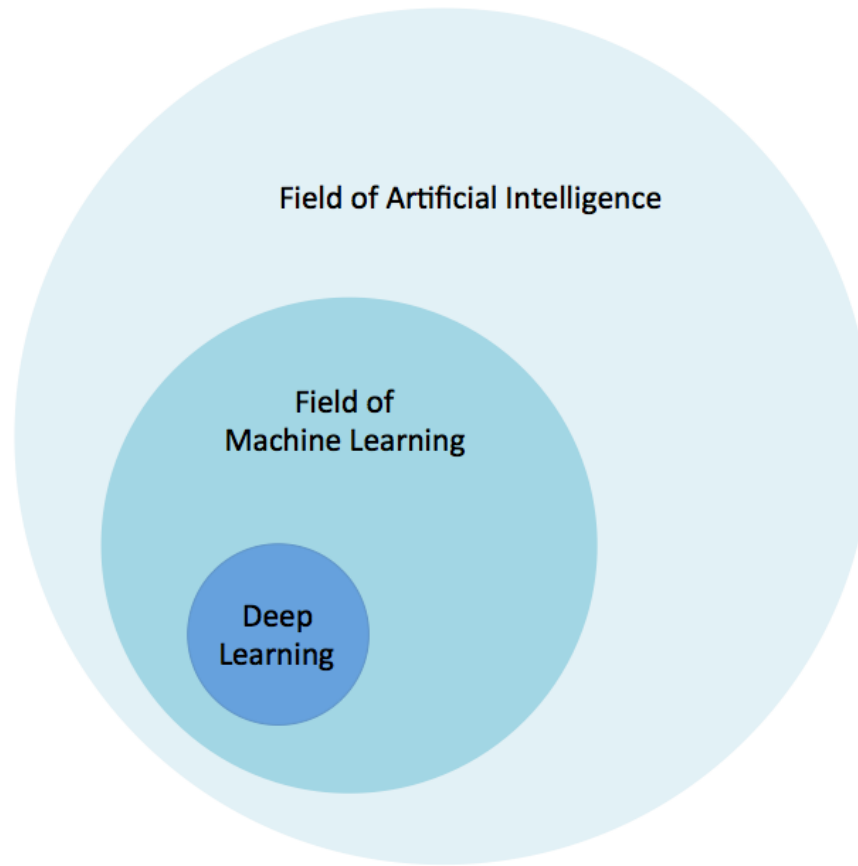


ANN

Artificial Neural Networks. Wrapping the NNs

Where we are? What are we doing?



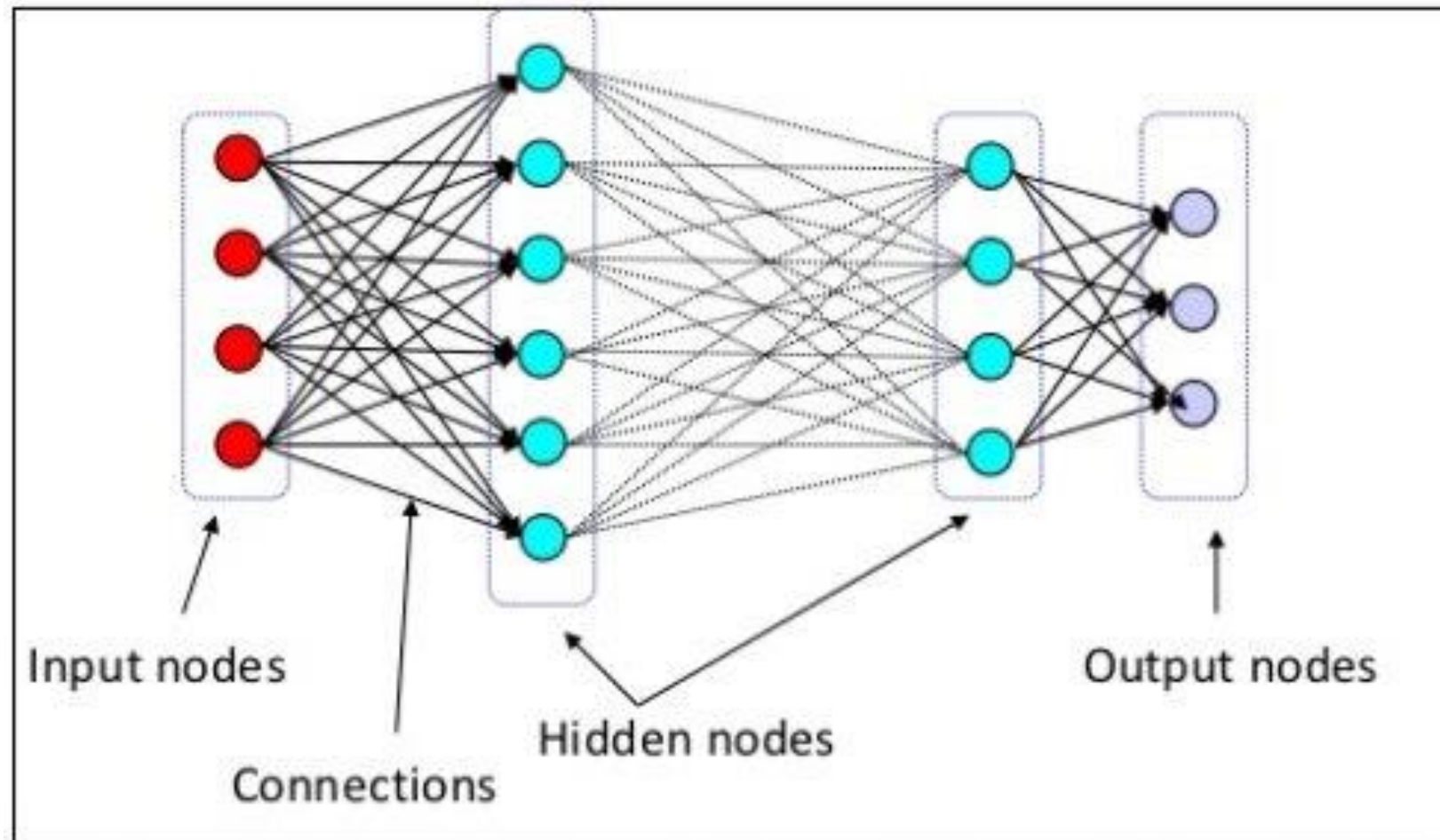
There are various types of NNs

- ANN
- CNN
- RNN
- LSTM
- GFNN - GirlFriendsNN
- EGFNN - ExGirlFriendsNN

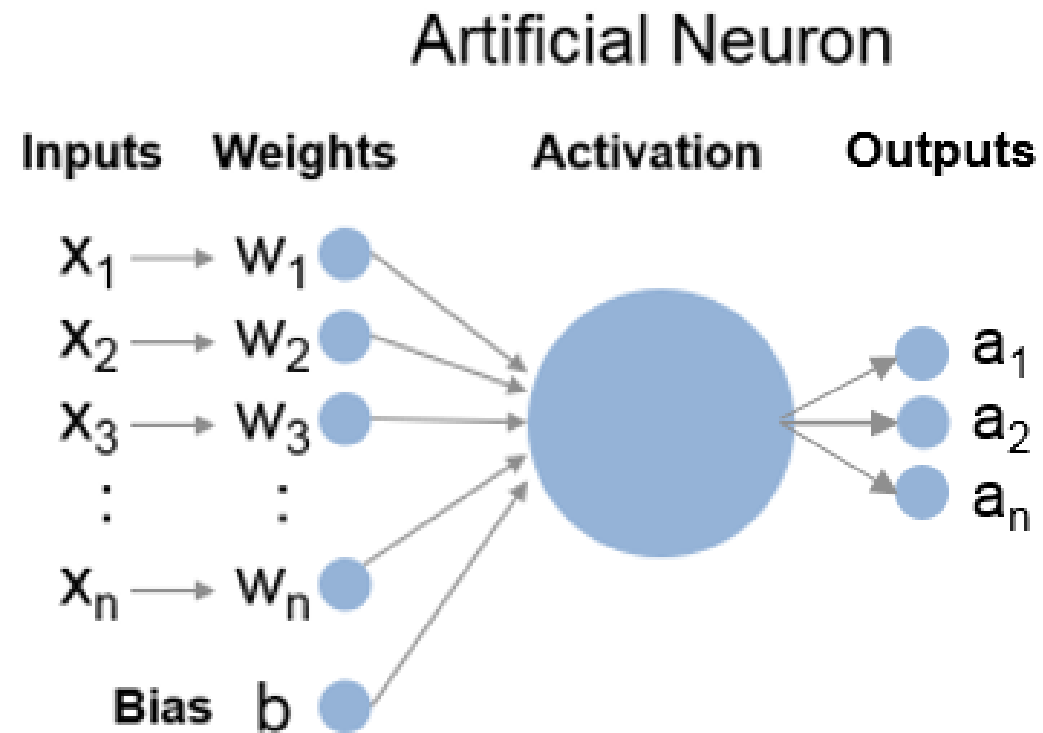
ANN

- ANN is basically a simple Neural Network
- CNN, RNN and LSTM are extended versions of NNs
- They all have a set of same principles
- and almost the same building algorithm

NN structure. Intuition



Neuron structure. Intuition



Okay, this all good...

- But what about practical stuff?
- There are some steps which are repetitive: ...

Data preprocessing

- Data importing
- Data encoding
- Data splitting
- Feature scaling

Building the network (1/2)

- Choosing the proper NN
- Choosing the:
 - Right # of layers
 - Right number of neurons (average of output and input parameters)
 - Relevant loss function
 - Proper activation function

Building the network (2/2)

- Compilation of NN
- Fitting the NN to your training set

Test data preparation

- Usually the same steps as for training data
- Take care about !DIMENSIONS!
- *NB: if we work in TF, Theano and Keras keep in mind that they work w/ numpy arrays*

Train your model

- Loops ... epochs ... batches

Prediction

- The prediction is easy in most of the times. One line
- One more line for fitting... Some times

Output

- Plot
- Print
- etc.

In depth. Data labelling

- It is necessary when it comes to Categorical Data. Usually String vars.
- We can use pandas and/or SciKitLearn
- Convert Strings into numbers
 - France, Spain, Germany – 0,1,2 / 001,010,100 (dummy variable) etc.
 - Male, Female – 1,2

In depth. Feature scaling

- Normalization

$$X_{\text{norm}} = (x - x_{\text{min}}) / (x_{\text{max}} - x_{\text{min}})$$

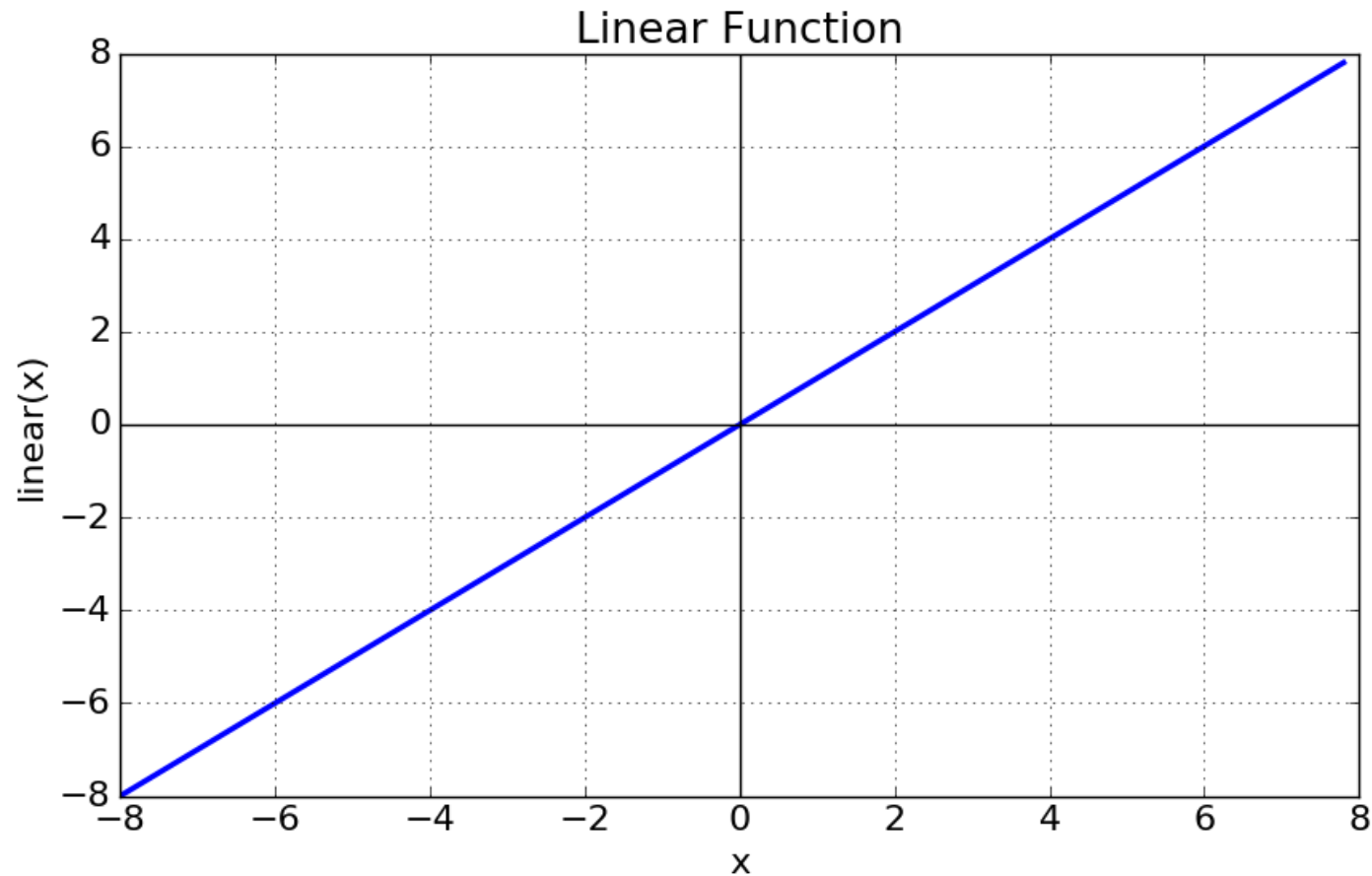
- Standardization

$$X_{\text{new}} = (x - \mu) / \sigma$$

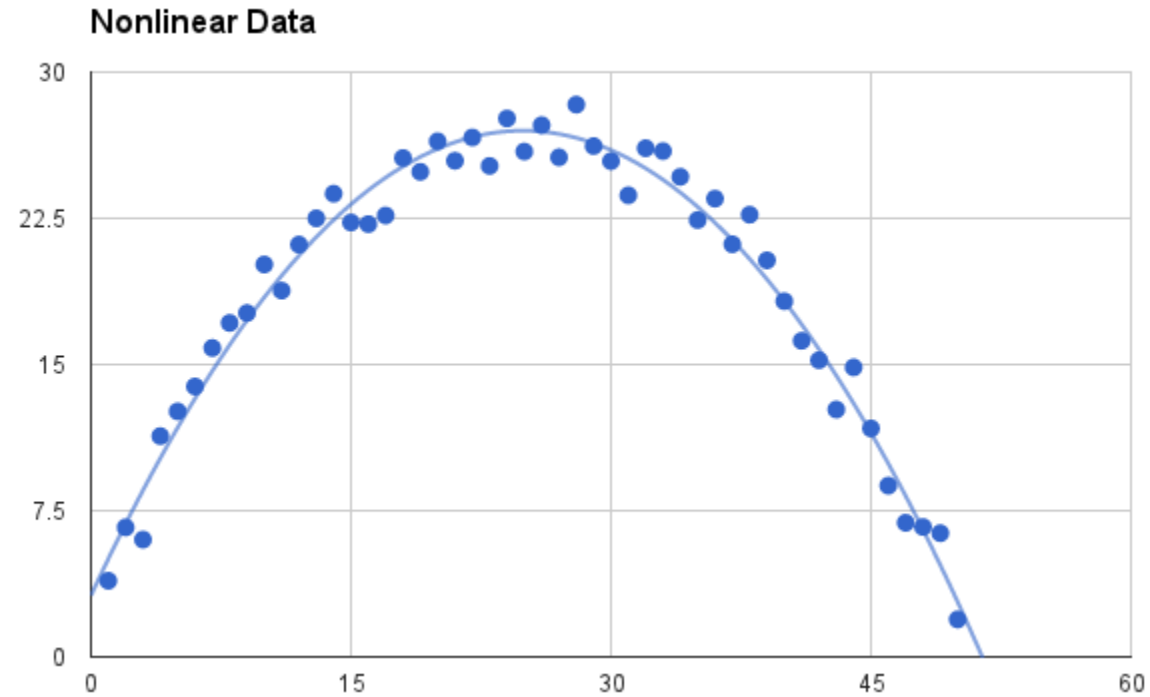
Activation functions

- Linear Activation Function
- Non-linear Activation Functions

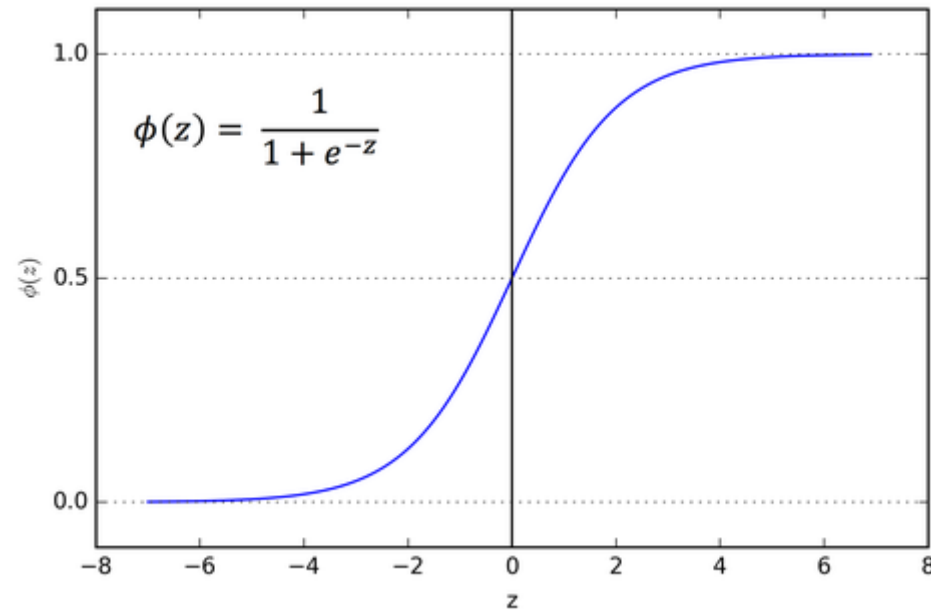
Linear or Identity Activation Function $f(x) = x$



Non-linear Activation Function (are the most used)



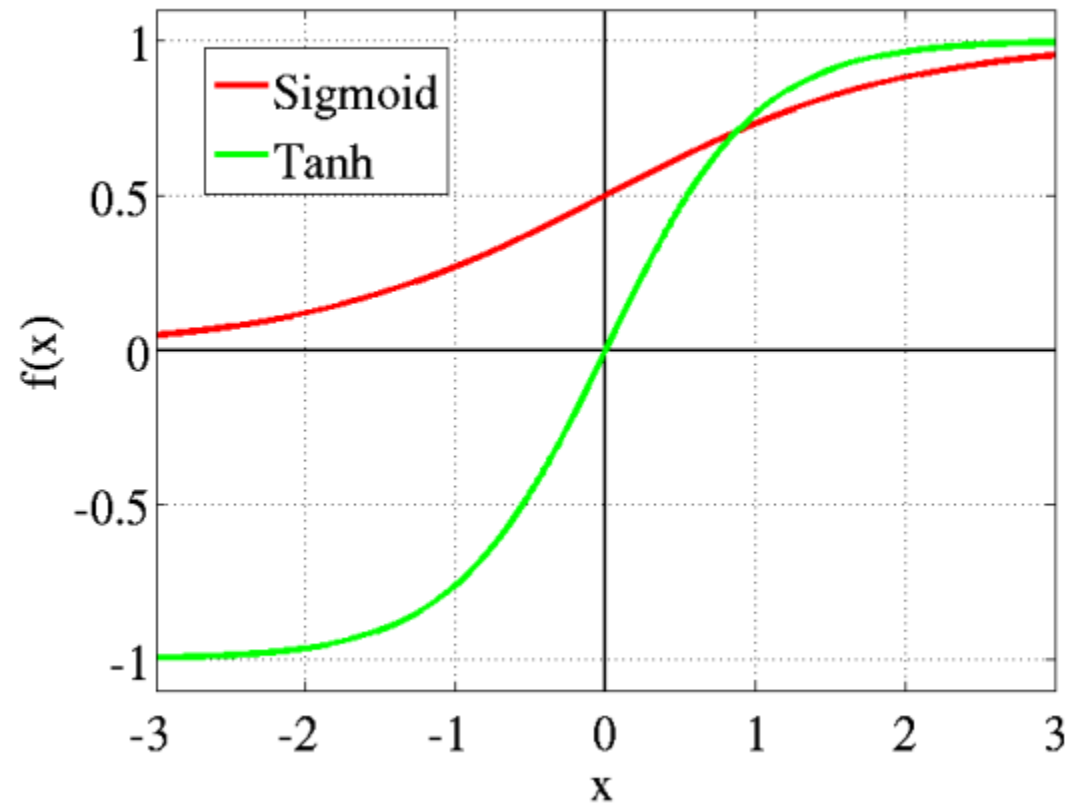
Sigmoid or Logistic Activation Function. Probability



NB: Sigmoid

- The logistic sigmoid function can cause a neural network to get stuck at the training time.
- The **softmax function** is a more generalized logistic activation function which is used for multiclass classification.

Tanh or hyperbolic tangent Activation Function



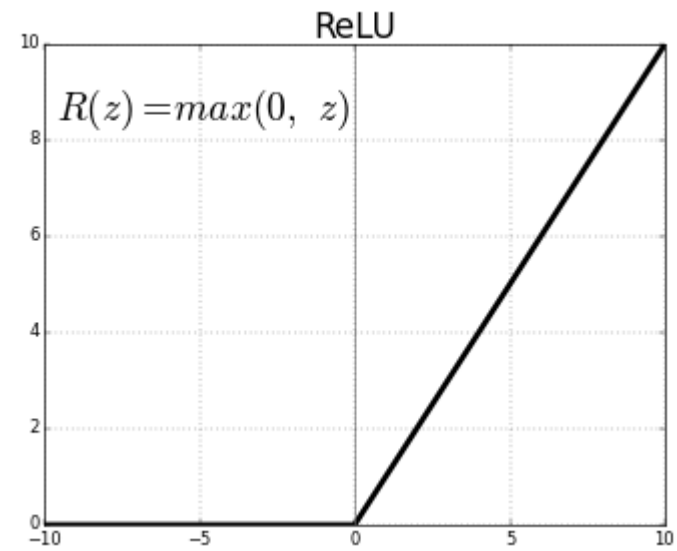
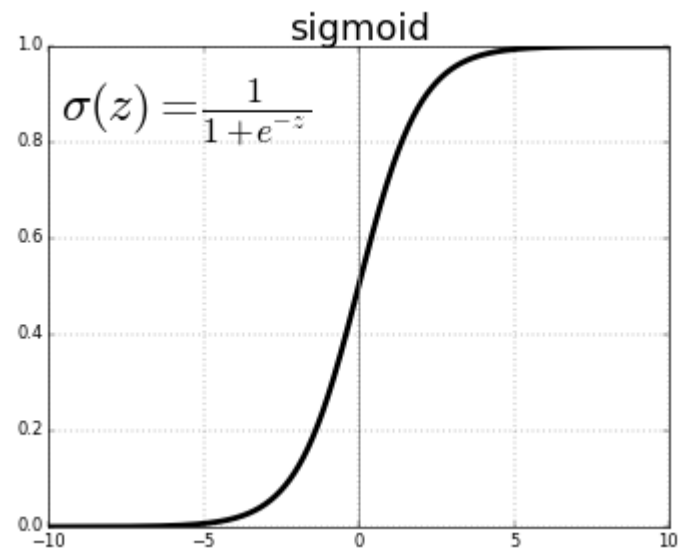
ND: Tanh

- The advantage is that the negative inputs will be mapped strongly negative and the zero inputs will be mapped near zero in the tanh graph.

ReLU (Rectified Linear Unit) Activation Function

- The ReLU is the most used activation function in the world right now. Since, it is used in almost all the convolutional neural networks or deep learning.

ReLU



NB: ReLU

- The issue is that all the negative values become zero immediately which decreases the ability of the model to fit or train from the data properly.

FIN.