# Machine Learning and Data Mining

26. May 2021.

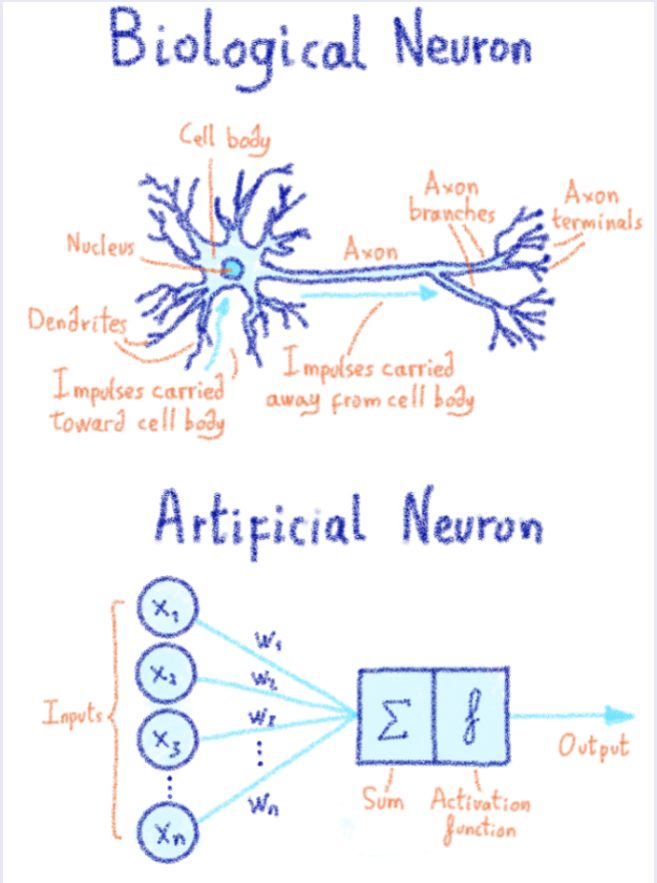


Figure 1: This Picture is from the https://www.ovh.com/blog/what-does-training-neural-networks-mean/

# Artificial neural network

## neural network (NN)

- a learning algorithm that is inspired by the structure and functional aspects of biological neural networks
- used when the exact nature of the relationship between inputs and output is not known.
- non-linear statistical data modeling tools
- used to model complex relationships between inputs and outputs
- used to find patterns in data
- used to capture the statistical structure in an unknown joint probability distribution between observed variables
- an example of soft computing

NNs learn the relationship between inputs and output through training!

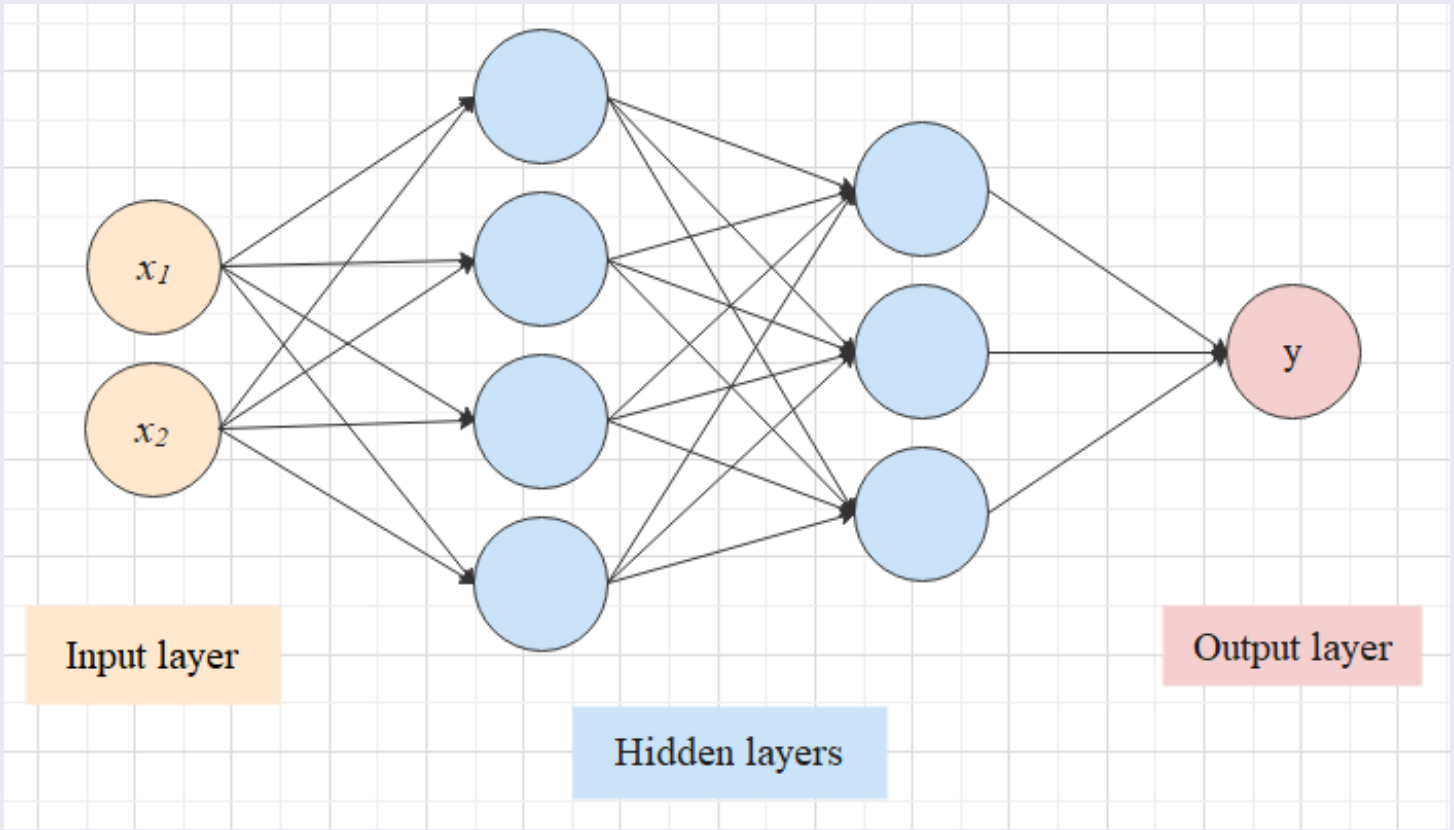


Figure 2: A simple artificial neural network with two hidden layers. This Figure is from the dissertation entitled "Stochastic modeling and statistical properties of the Limit Order Book", Dragana Radojicic

# Activation functions

## Activation functions

- a step function (e.g. a sign function)
- a sigmoid function ($\sigma(x) = \frac{1}{1+e^{-x}}$)
- a hyperbolic tangent function ($tanh(x) = \frac{e^{x}-e^{-x}}{e^{x}+e^{-x}}$)

- input $x$ and produces output $y$
- $y = f(wx + b)$, where $f$ is an activation function, an $w$ and $b$ are both the neuron's scalar parameters

# The training process of a neural network

## The training process of a neural network

- The input is defined
- The weights of the network are assigned some starting value.
- An output is obtained by passing the input through neurons in each layer.
- The generated output is compared with the expected output or label.
- The error between the prediction and label is then used to update the weights of each node.
- The error is then propagated in backwards direction through every layer, to update the weights in each layer such that they minimize the error.

This is from the dissertation entitled "Stochastic modeling and statistical properties of the Limit Order Book", Dragana Radojicic

# The training phase

## The training phase based on the backpropagation

- Online (stochastic) method
- Batch method

## stochastic gradient descent (SGD) method

- most commonly used to update the weights
- The samples are randomly taken from the whole set of data → stochastic

## Batch learning

- the dataset is divided into several batches (parts)
- all batches are passed into the neural network
- Gradient descent optimize the learning rate
- One epoch is when an entire dataset is passed forward and backward

# Recurrent neural networks (RNNs)

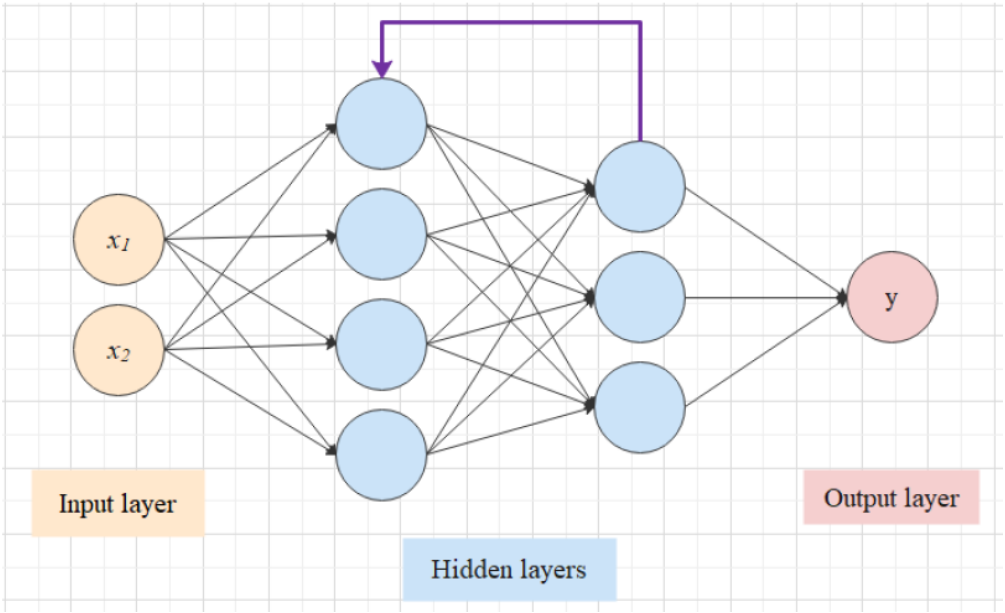


Figure 3: A simple recurrent neural network. This Figure is Figure 3.4. from the dissertation entitled "Stochastic modeling and statistical properties of the Limit Order Book", Dragana Radojicic

# Recurrent neural networks (RNNs)

## Recurrent neural networks (RNNs)

- a class of neural networks that remember the past
- employed when dealing with sequential data types

## Recurrent neural networks (RNNs)

- $x = (x_1, x_2, ..., x_N)$ given sequence
- RNN updates its hidden state $h_t = \phi(h_{t-1}, x_t)$, nonlinear function $\phi$, $h_0$ is initially set to 0.
- $y = (y_1, y_2, ..., y_N) = (h_1, h_2, ..., h_N)$ the hidden state output

# Long short-term memory (LSTM)

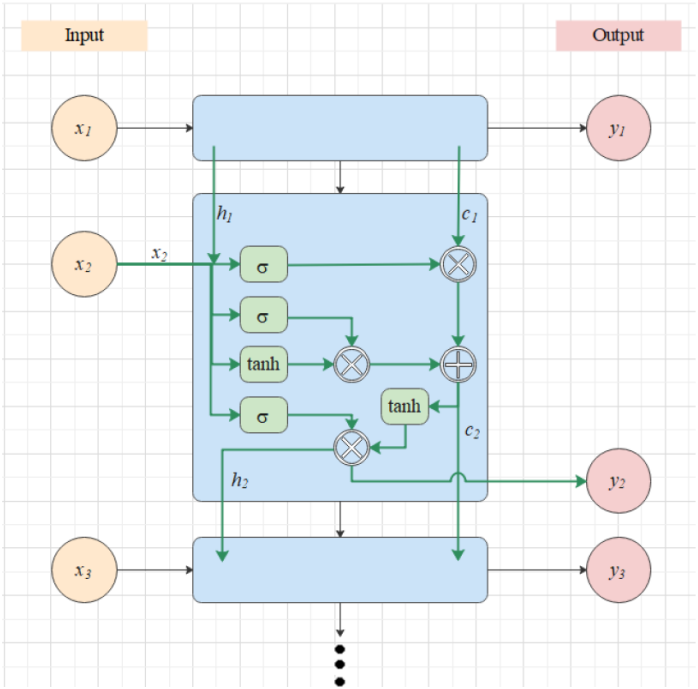


Figure 4: A simple representation of Long short-term memory architecture. This Figure is Figure 3.4. from the dissertation entitled "Stochastic modeling and statistical properties of the Limit Order Book", Dragana Radojicic

# LSTM, GRU

## Long short-term memory (LSTM)

- LSTM unit is composed of a cell, an input gate, an output gate and a forget gate.
- Sepp Hochreiter; Jürgen Schmidhuber (1997). "Long short-term memory". Neural Computation. 9 (8)

## Gated Recurrent Unit (GRU)

- Cho, Kyunghyun; van Merrienboer, Bart; Gulcehre, Caglar; Bahdanau, Dzmitry; Bougares, Fethi; Schwenk, Holger; Bengio, Yoshua (2014). "Learning Phrase Representations using RNN Encoder-Decoder for Statistical Machine Translation".
- GRU cell contains two gates: update gate and reset gate.
- gating mechanisms are used to manage the information flow between cells