# 5 - Deep Learning Overview

## 5.1 - Deep Learning Summary

Marciel Barros

Setembro, 2020

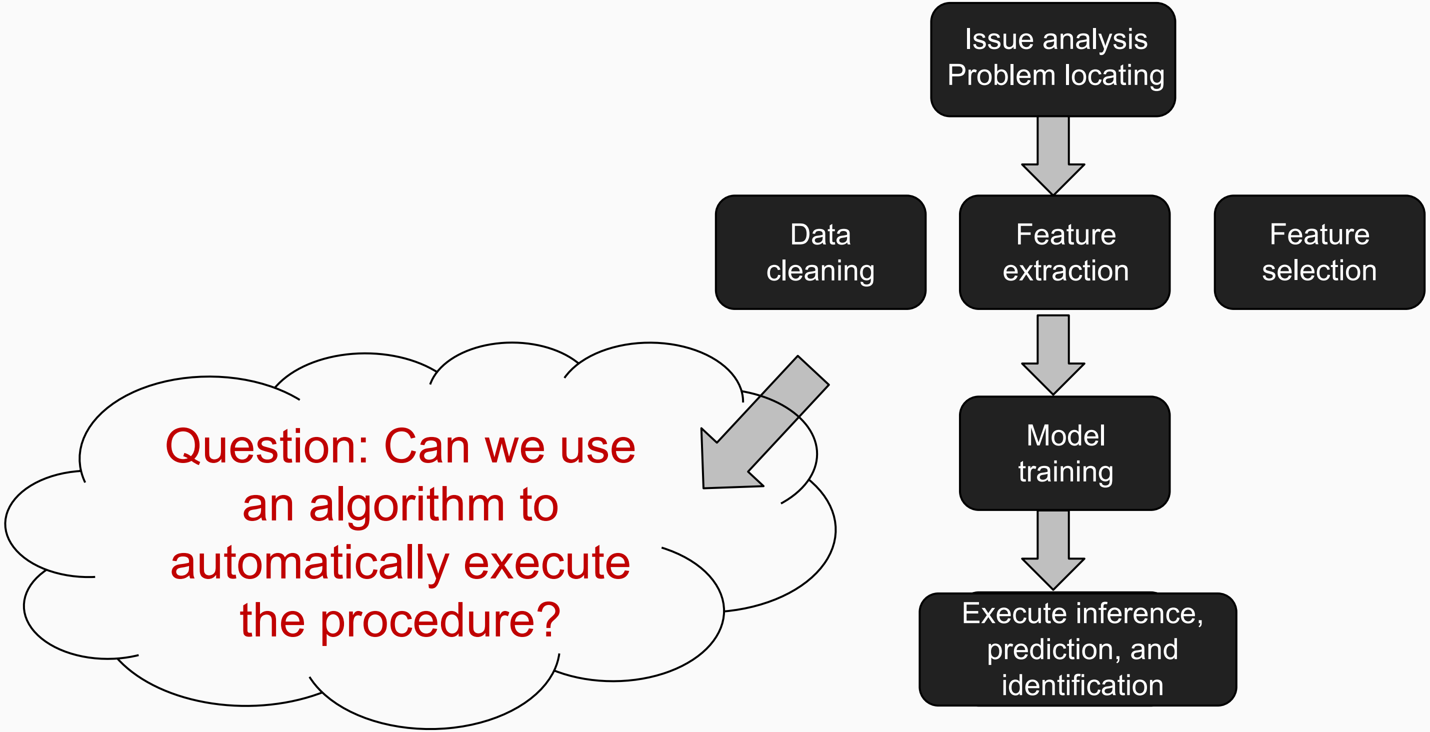
# Index

# Traditional Machine Learning and Deep Learning

As a model based on unsupervised feature learning and feature hierarchy learning, deep learning has great advantages in fields such as computer vision, speech recognition, and natural language processing.

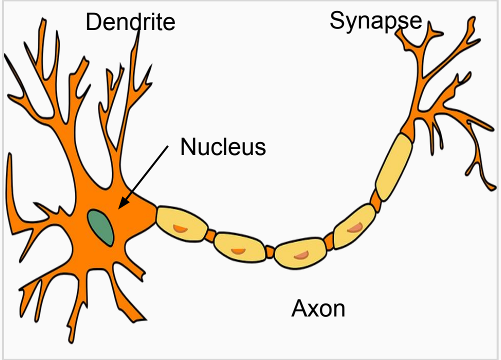
| Traditional Machine Learning | Deep Learning |
| --- | --- |
| Low hardware requirements on the computer: Given the limited computing amount, the computer does not need a GPU for parallel computing generally. | Higher hardware requirements on the computer: To execute matrix operations on massive data, the computer needs a GPU to perform parallel computing. |
| Applicable to training under a small data amount and whose performance cannot be improved continuously as the data amount increases. | Whose performance can be high when high-dimensional weight parameters and massive training data are provided. |
| Level-by-level problem breakdown | E2E learning |
| Manual feature selection | Algorithm-based automatic feature extraction |
| Easy-to-explain features | Hard-to-explain features |

# Traditional Machine Learning

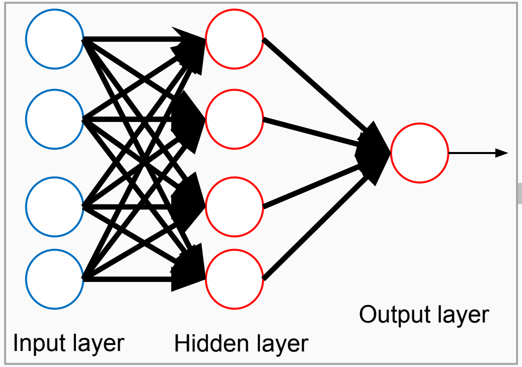


# Deep Learning

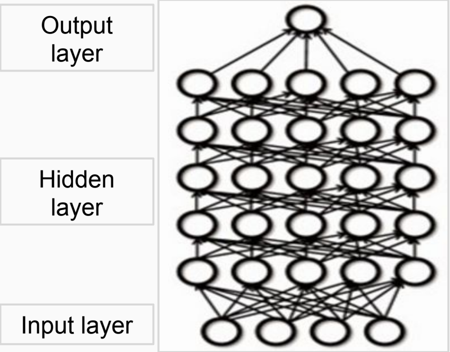
Generally, the deep learning architecture is a deep neural network. "Deep" in "deep learning" refers to the number of layers of the neural network.




Human neural network




Perceptron



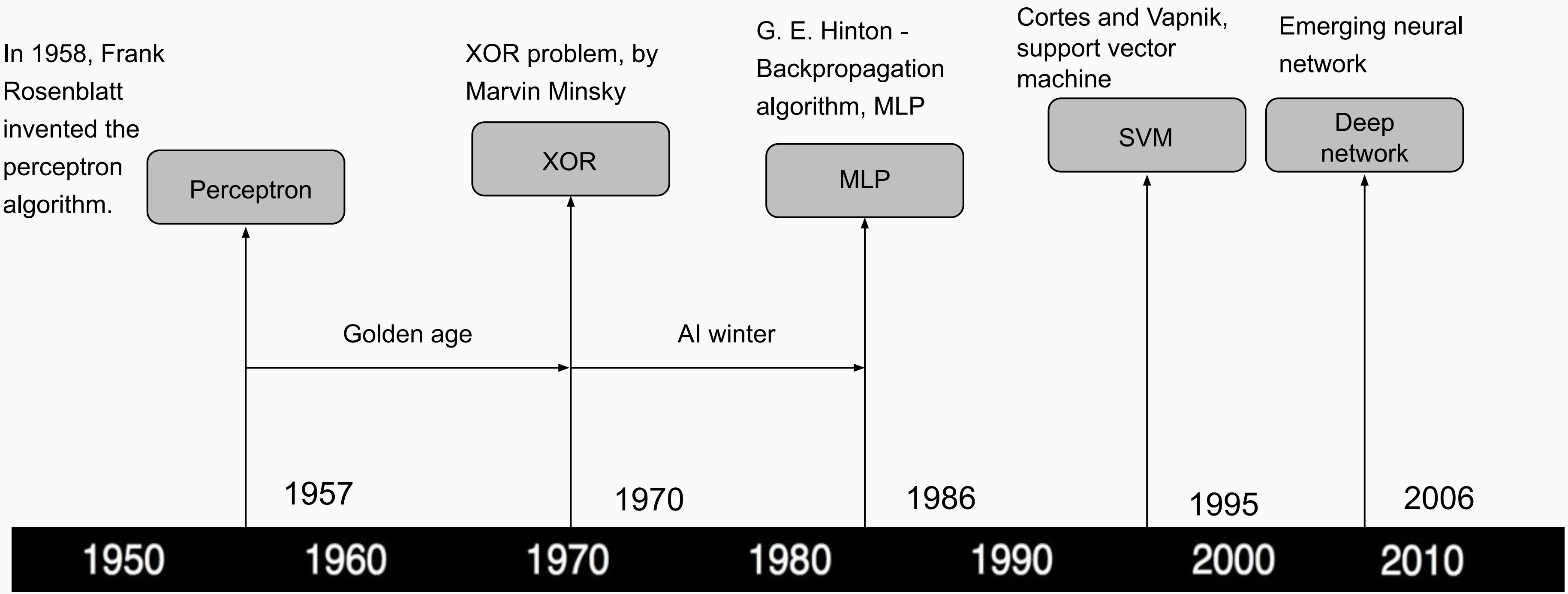

Deep neural network

# Neural Network

- According Hecht Nielsen, a neural network researcher in the U.S., it is a **computer system composed of simple and highly interconnected processing elements**, which process information by dynamic response to external inputs.
- It can be simply expressed as **an information processing system designed to imitate the human brain structure and functions** based on its source, features, and explanations.
- **Artificial neural network (neural network)**: artificial neurons connected to each other, which extracts and simplifies the human brain's microstructure and functions.

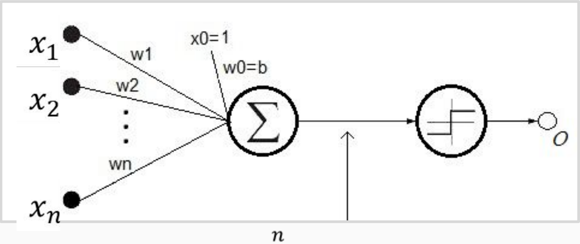
HUAWEI

# Development History of Neural Networks



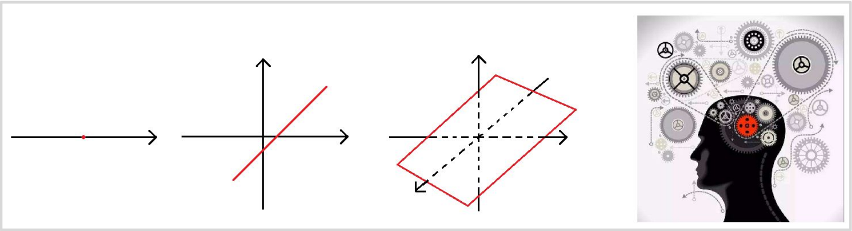
# Single-Layer Perceptron

- Input vector: $X = [x_0, x_1, \dots, x_n]^T$
- Weight: $W = [\omega_0, \omega_1, \dots, \omega_n]^T$, in which $\omega_0$ is the offset.
- Activation function: $O = sign(net) = \begin{cases} 1, net > 0, \\ -1, otherwise. \end{cases}$




$$net = \sum_{i=0}^{n} \omega_i x_i = \boldsymbol{W}^T \mathbf{X}$$

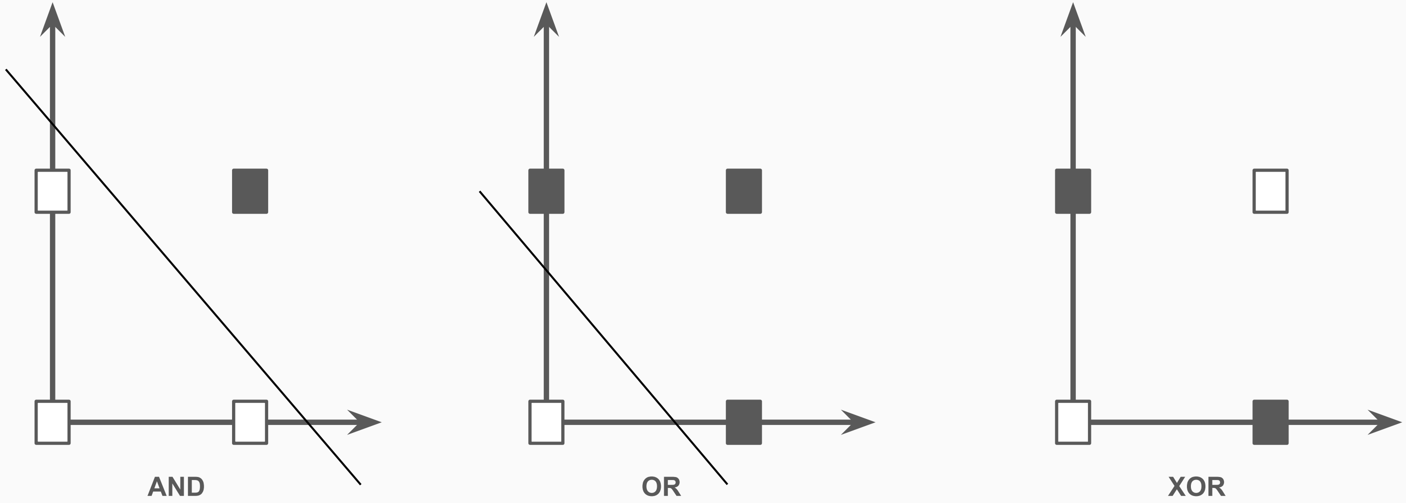
- The preceding perceptron is equivalent to a classifier. It uses the high-dimensional $X$ vector as the input and performs binary classification on input samples in the high-dimensional space. When $\boldsymbol{W}^T\mathbf{X} > 0$, $O = 1$. In this case, the samples are classified into a type. Otherwise, $O = -1$. In this case, the samples are classified into the other type. The boundary of these two types is $\boldsymbol{W}^T\mathbf{X} = 0$, which is a high-dimensional hyperplane.



Classification point $Ax + B = 0$ | Classification line $Ax + By + C = 0$ | Classification plane $Ax + By + Cz + D = 0$ | Classification hyperplane $W^T X + b = 0$

# XOR Problem

- In 1969, Minsky, an American mathematician and AI pioneer, proved that a perceptron is essentially a linear model that can only deal with linear classification problems, but cannot process non-linear data.




https://blog.csdn.net/qq_18515405/article/details/42123697

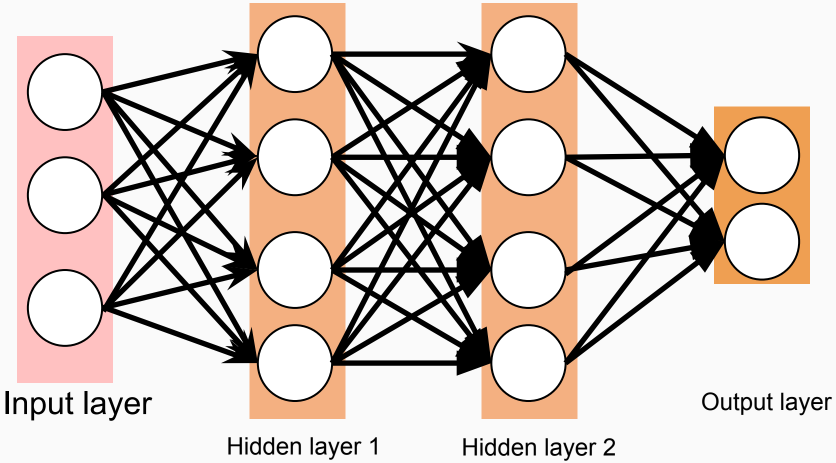
# Feedforward Neural Network

**The multilayer perceptron is a feedforward neural network:** the simplest neural network with neurons arranged in layers.

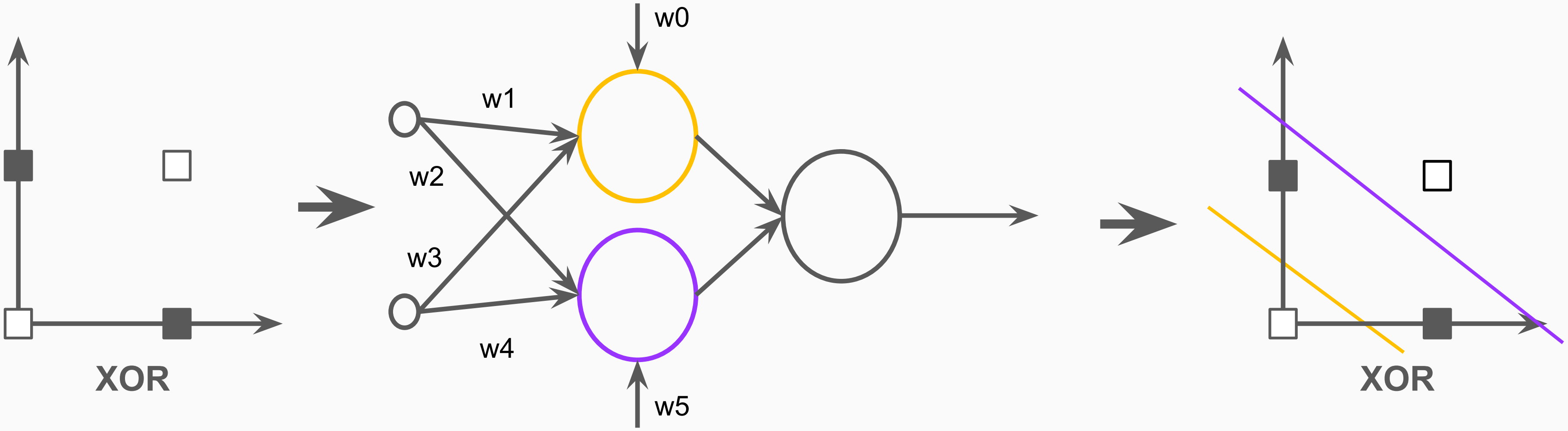
Input nodes **do not have the calculation capability**, only used to **represent element values** of the input vector.

The neuron having the computing capability at each layer is referred to as the **computing unit.**

A unidirectional multi-layer structure is used to receive the output of the previous layer and send the output to the next layer.



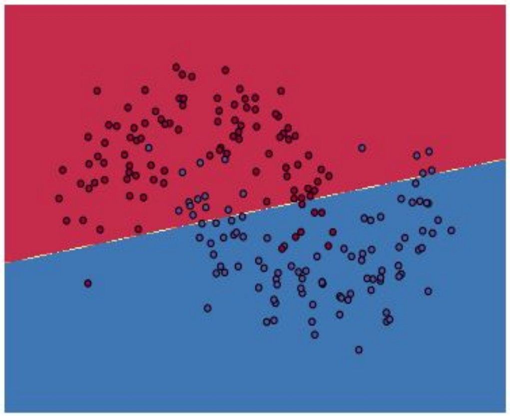
# Solution of XOR



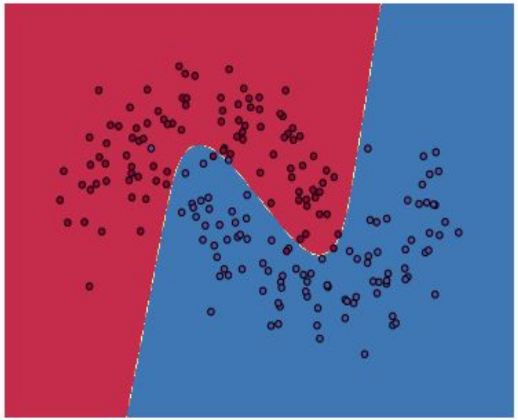

https://playground.tensorflow.org/#activation=tanh&batchSize=10&dataset=xor&regDataset=reg-plane&learningRate=0.03&regularizationRate=0&noise=0&networkShape=2&seed=0.82805&showTestData=false&discretize=true&percTrainData=50&x=true&y=true&xTimesY=false&xSquared=false&ySquared=false&cosX=false&sinX=false&cosY=false&sinY=false&collectStats=false&problem=classification&initZero=false&hideText=false

# Impacts of Hidden Layers on A Neural Network

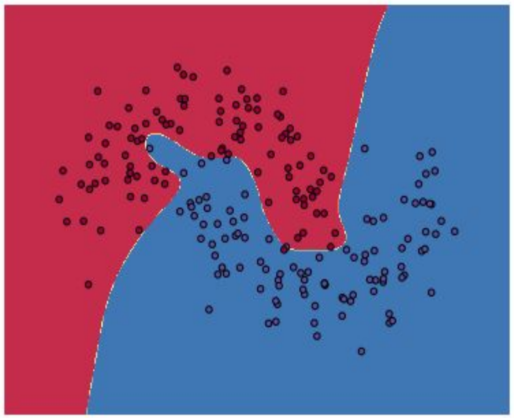
More hidden layers indicate the stronger identification capability of the neural network.



0 hidden layers



3 hidden layers



20 hidden layers

# 5 - Deep Learning Overview

## Next: 5.2 - Training Rules

Marciel Barros

Setembro, 2020