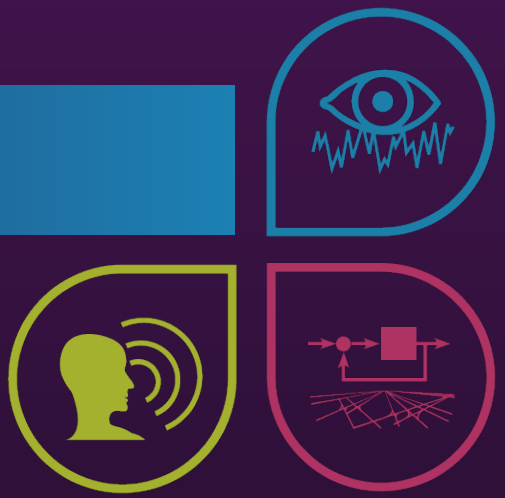
gipsa-lab

Grenoble | images | parole | signal | automatique | laboratoire

Machine learning overview



MLST presentation

Julien MUZEAU

GIPSA-lab

16/01/2018

cnrs Grenoble INP Université Grenoble Alpes UMR 5216

Outline

History

According to Wikipedia : « Machine learning is a field of computer science that gives computers the ability to learn without being explicitly programmed »

1959 **A. Samuel** coined "Machine Learning"

History

According to Wikipedia : « Machine learning is a field of computer science that gives computers the ability to learn without being explicitly programmed »

1959 **A. Samuel** coined "Machine Learning"

1981 **P. Werbos** MLP + BackPropagation

History

According to Wikipedia : « Machine learning is a field of computer science that gives computers the ability to learn without being explicitly programmed »

- 1959 **A. Samuel** coined "Machine Learning"
- 1981 **P. Werbos** MLP + BackPropagation
- 1986 **J.R. Quinlan** Decision Tree

History

According to Wikipedia : « Machine learning is a field of computer science that gives computers the ability to learn without being explicitly programmed »

1959 **A. Samuel** coined "Machine Learning"

1981 **P. Werbos** MLP + BackPropagation

1986 **J.R. Quinlan** Decision Tree

1995 **V. Vapnik & C. Cortes** Support Vector Machine

History

According to Wikipedia : « Machine learning is a field of computer science that gives computers the ability to learn without being explicitly programmed »

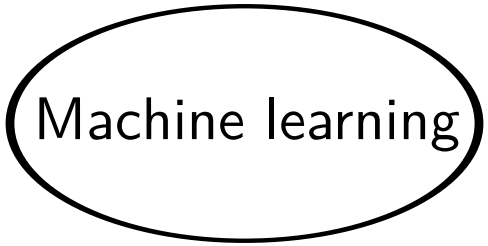
1959	**A. Samuel** coined "Machine Learning"
1981	**P. Werbos** MLP + BackPropagation
1986	**J.R. Quinlan** Decision Tree
1995	**V. Vapnik & C. Cortes** Support Vector Machine
2001	**L. Breiman** Random Forest

History

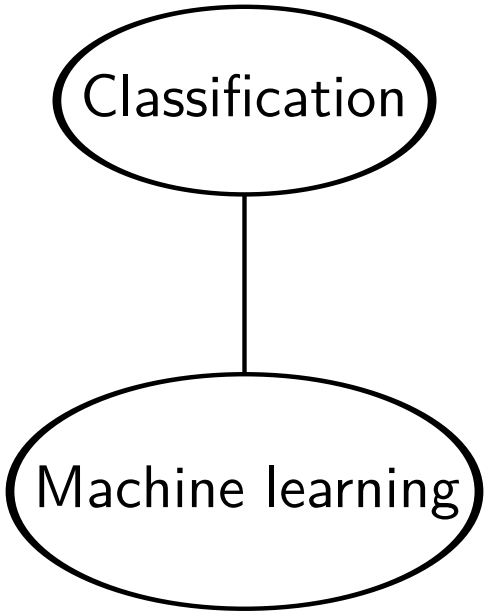
According to Wikipedia : « Machine learning is a field of computer science that gives computers the ability to learn without being explicitly programmed »

1959	**A. Samuel** coined "Machine Learning"
1981	**P. Werbos** MLP + BackPropagation
1986	**J.R. Quinlan** Decision Tree
1995	**V. Vapnik & C. Cortes** Support Vector Machine
2001	**L. Breiman** Random Forest
2005	New era of Deep Learning : CNN, AE, GD, ...

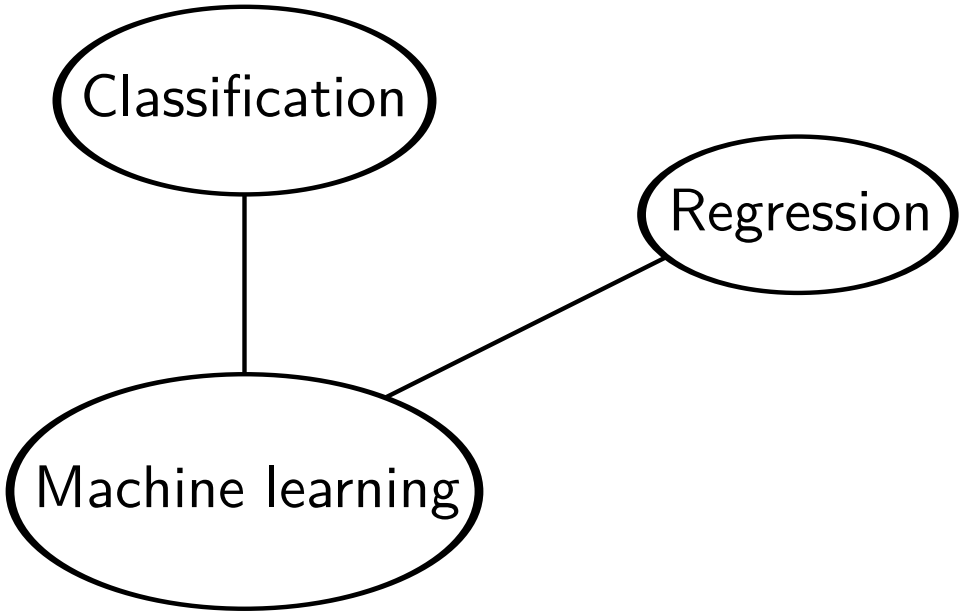
Topics



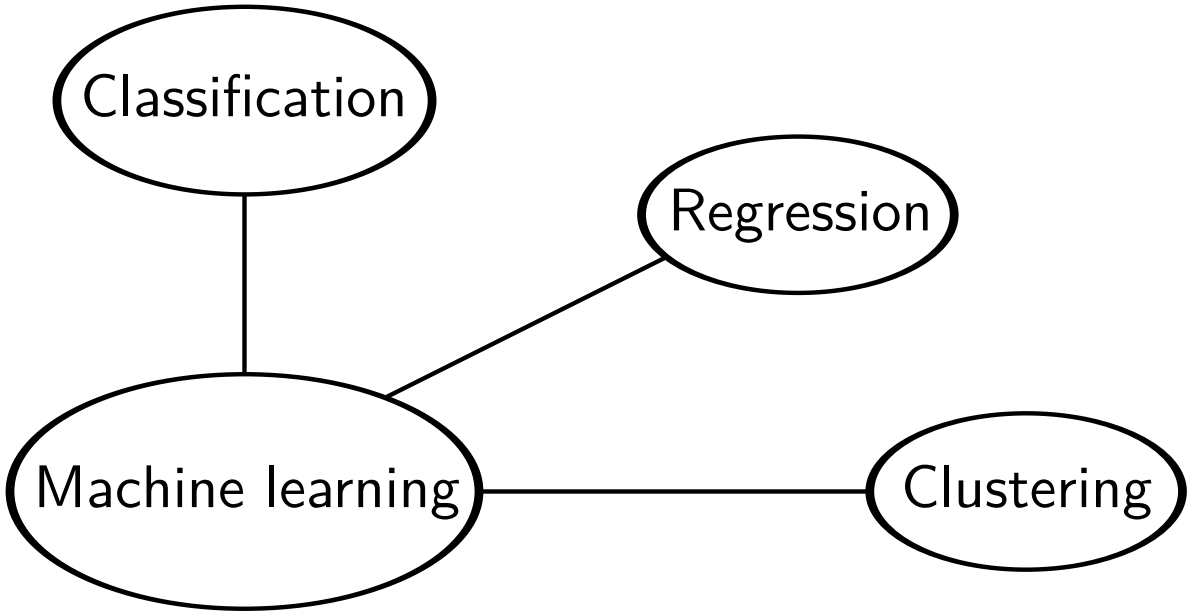
Topics



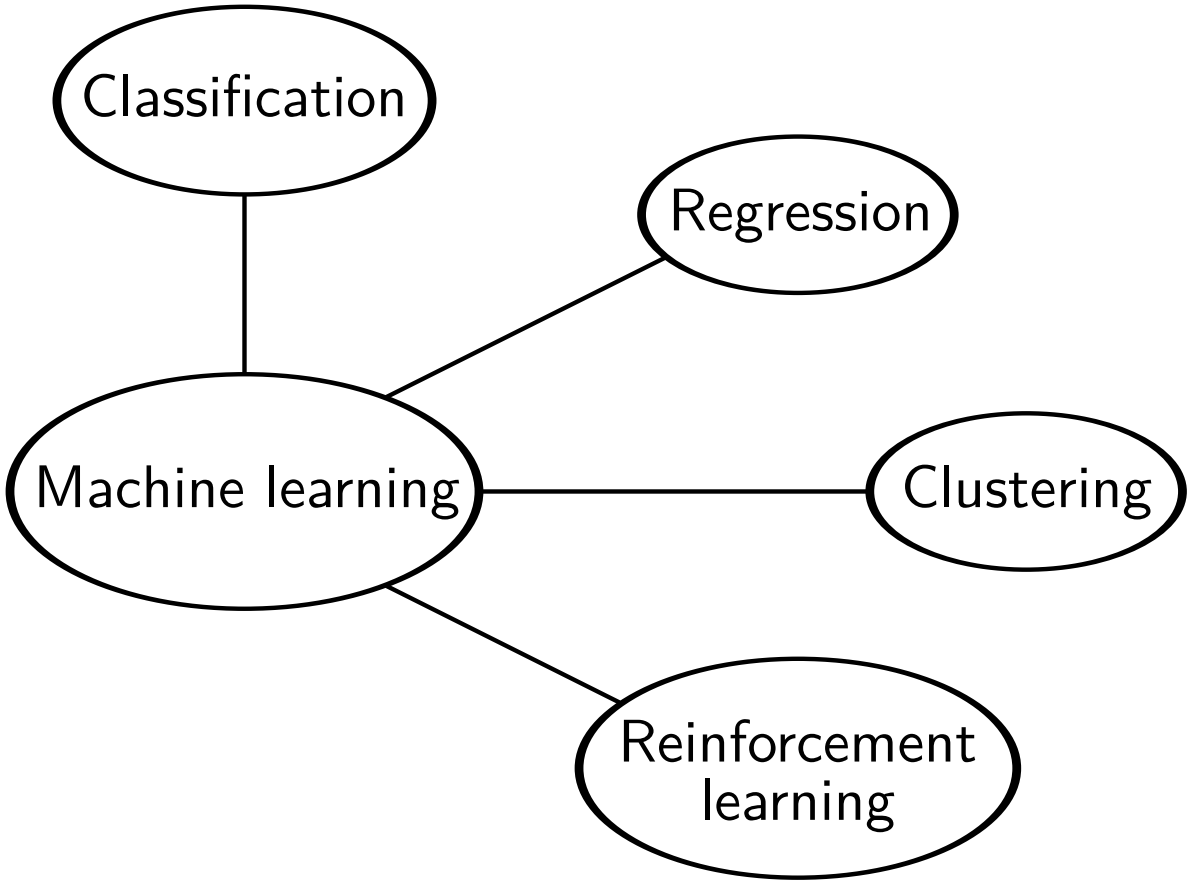
Topics



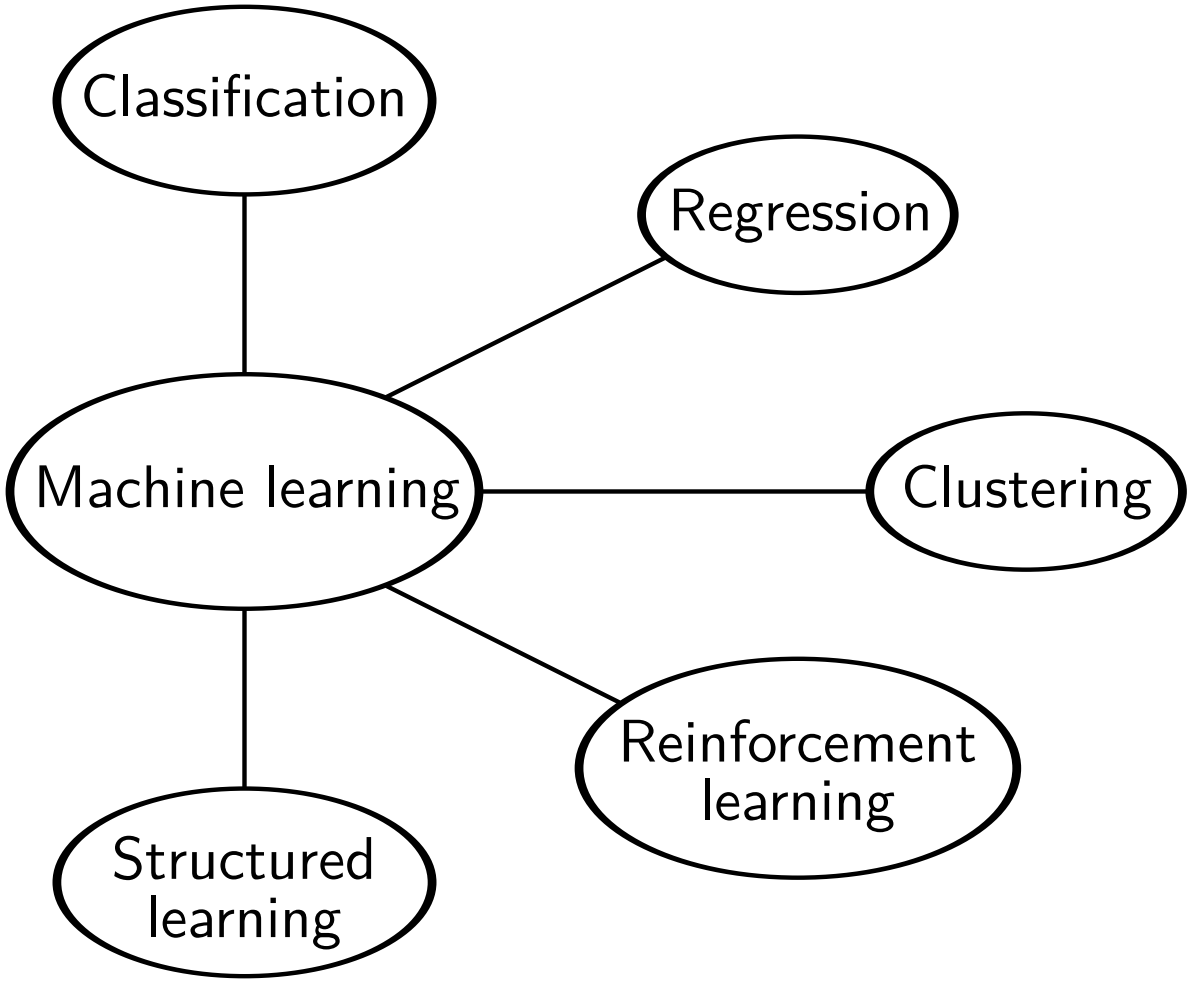
Topics



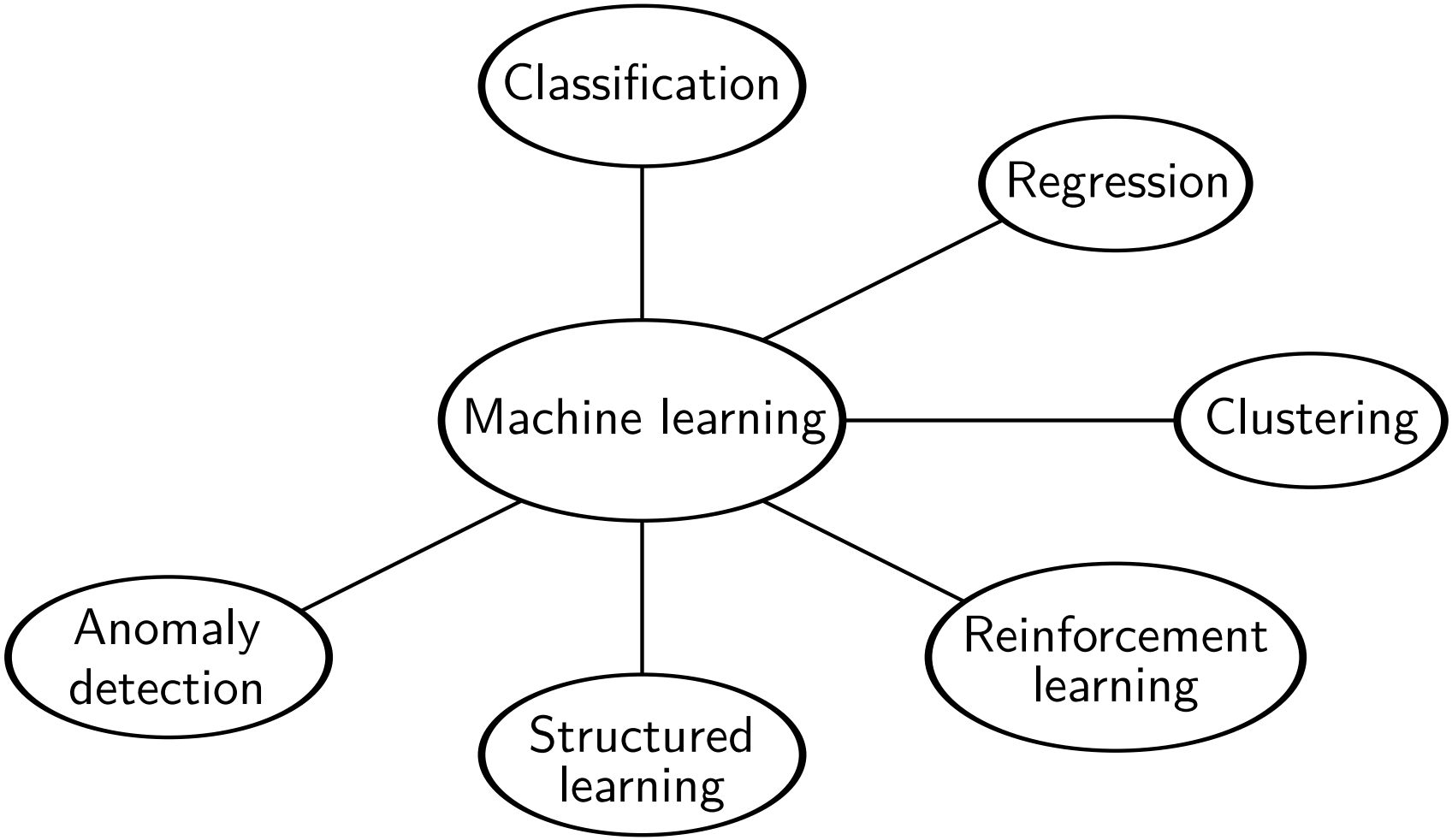
Topics


Classification
Regression
Machine learning
Clustering
Reinforcement learning

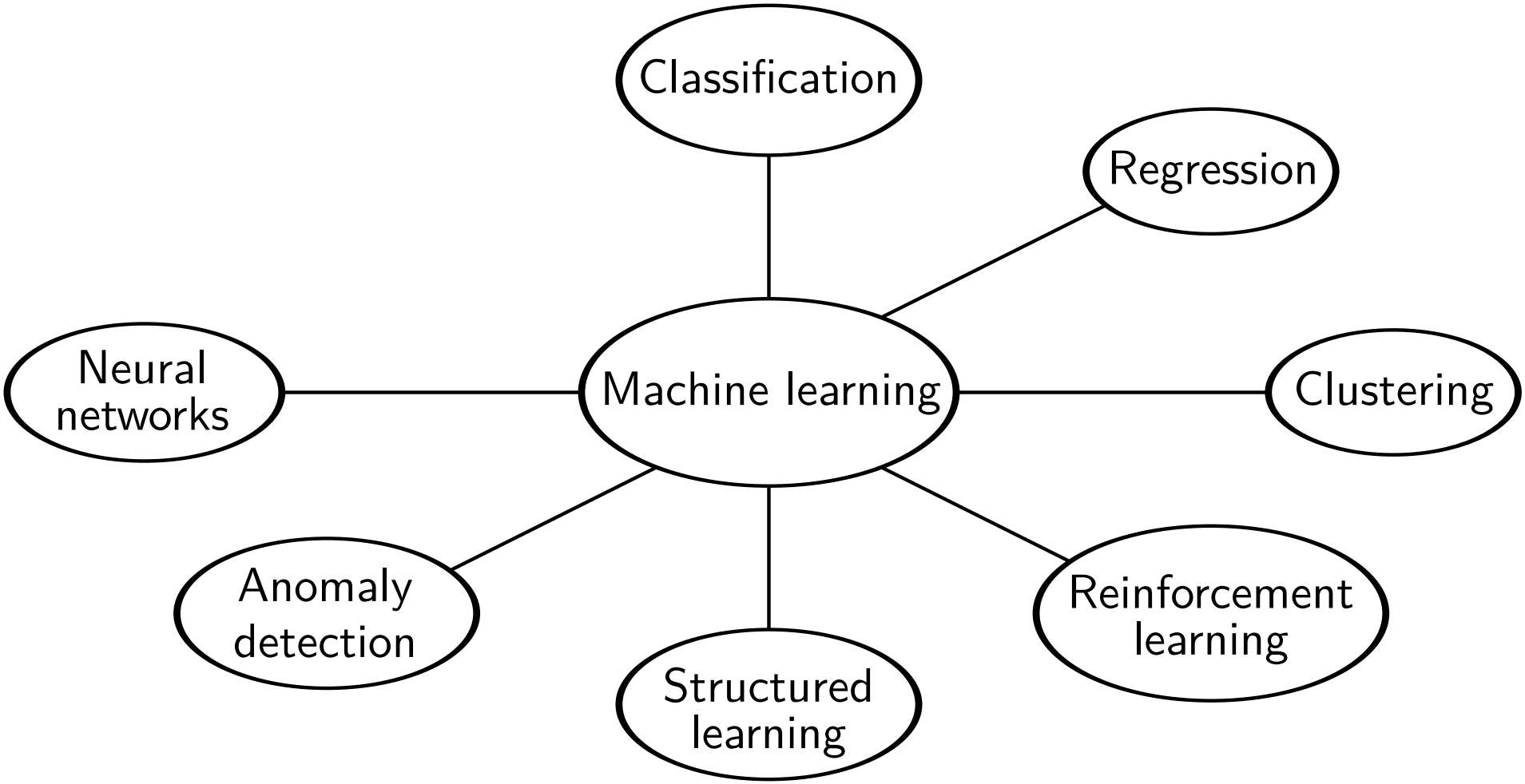
Topics


Classification
Regression
Machine learning
Clustering
Reinforcement learning
Structured learning

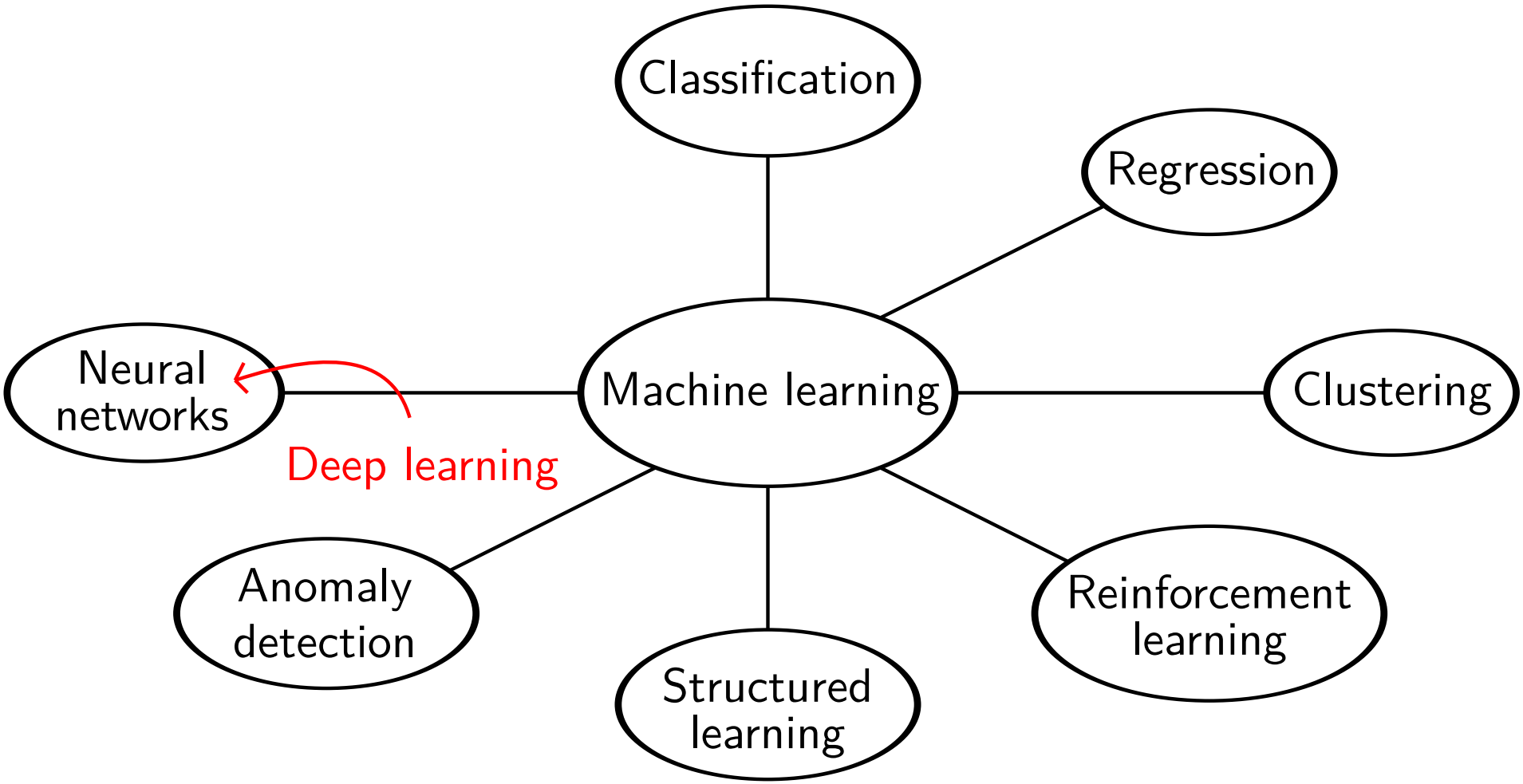
Topics


Classification
Regression
Machine learning
Clustering
Anomaly detection
Structured learning
Reinforcement learning

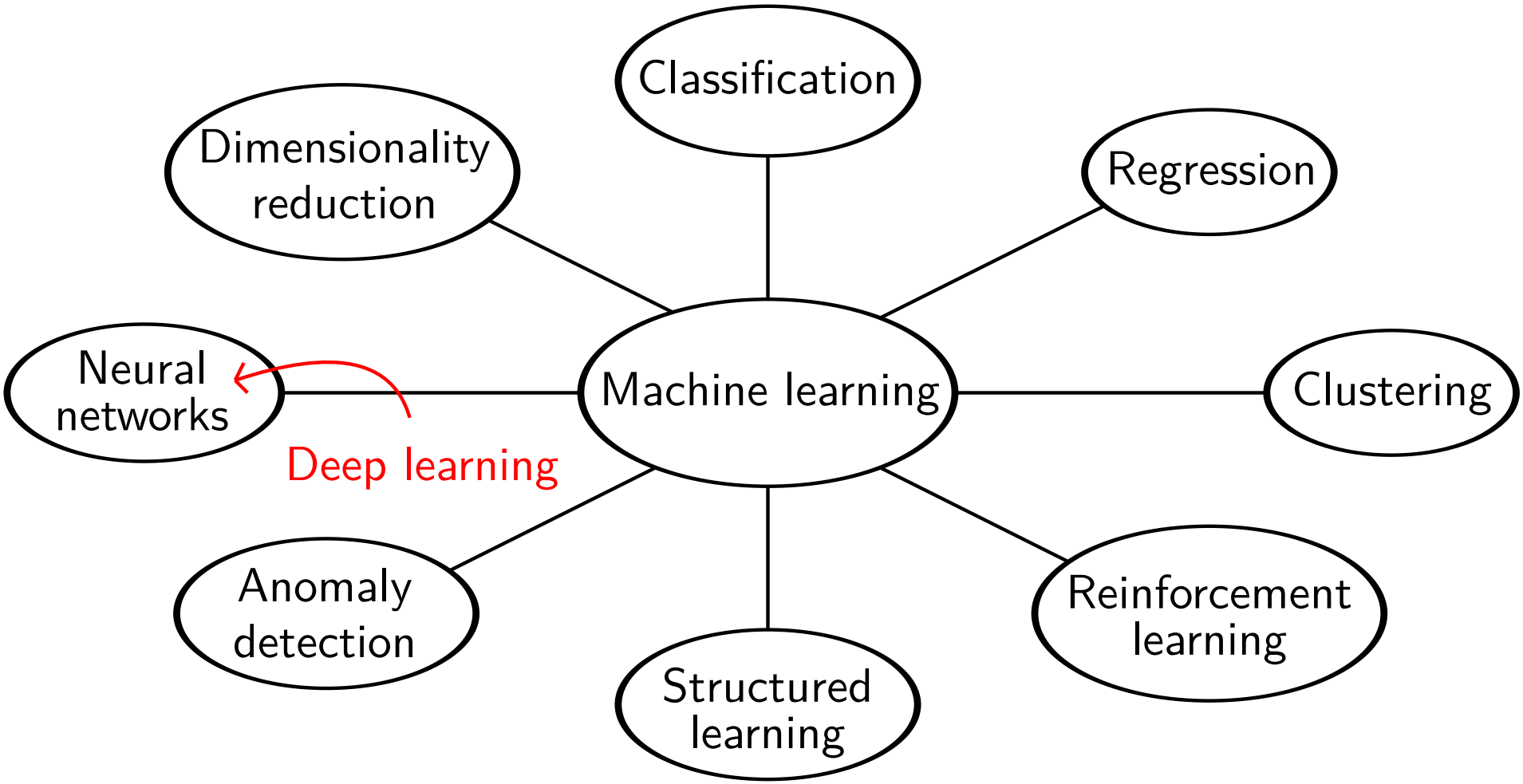
Topics


Classification
Regression
Neural networks
Machine learning
Clustering
Anomaly detection
Structured learning
Reinforcement learning

Topics


Classification
Regression
Neural networks
Machine learning
Clustering
Deep learning
Anomaly detection
Structured learning
Reinforcement learning

Topics


Classification
Dimensionality reduction
Regression
Neural networks
Machine learning
Clustering
Deep learning
Anomaly detection
Structured learning
Reinforcement learning



Methods (1)

Classification - Regression (Supervised learning)

- Decision trees
- Ensemble learning
- k-Nearest Neighbors
- Support/Relevance Vector Machine

Clustering (Unsupervised learning)

- k-means
- Expectation-Maximization
- Mean-shift

Methods (2)

Dimensionality reduction

- Principal Component Analysis
- Independent Component Analysis
- Linear Discriminant Analysis
- t-distributed Stochastic Neighbor Embedding

Structured prediction

- Bayesian network
- Conditional Random Field
- Hidden Markov Model

Methods (3)

Anomaly detection

- k-Nearest Neighbors
- Local outlier factor

Neural network learning

- Auto-encoder
- Multi-Layer Perceptron
- Recurrent Neural Network
- Convolutional Neural Network

Methods (4)

Reinforcement learning

- Q-learning
- Temporal Difference



K-means (1)

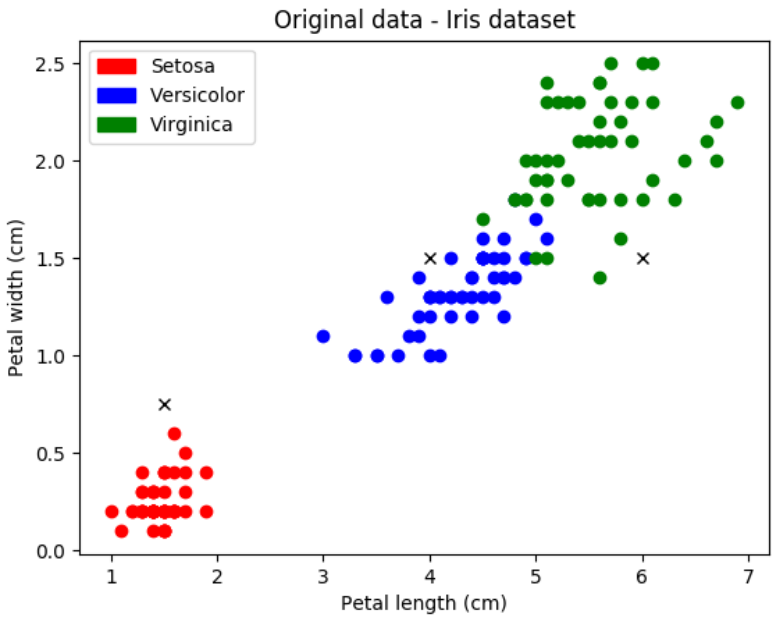
Goal

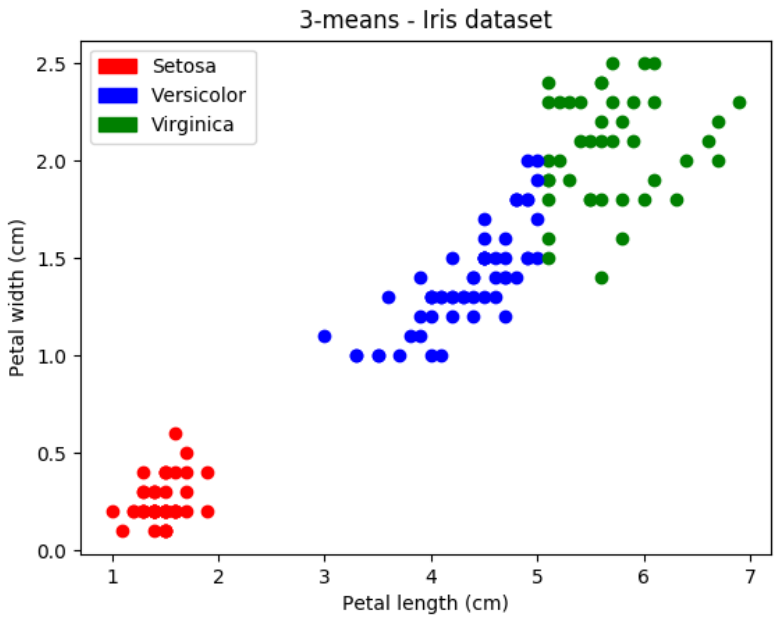
Iteratively divide data into K *clusters* so that the distance between each point within each cluster is minimal

Two steps

- Assign each data point to an only cluster
- Update the centroid of each cluster

K-means (2)


Original data - Iris dataset
Setosa
Versicolor
Virginica
Petal width (cm)
Petal length (cm)
0.0
0.5
1.0
1.5
2.0
2.5
1
2
3
4
5
6
7


3-means - Iris dataset
Setosa
Versicolor
Virginica
Petal width (cm)
Petal length (cm)
0.0
0.5
1.0
1.5
2.0
2.5
1
2
3
4
5
6
7

Gaussian mixture (1)

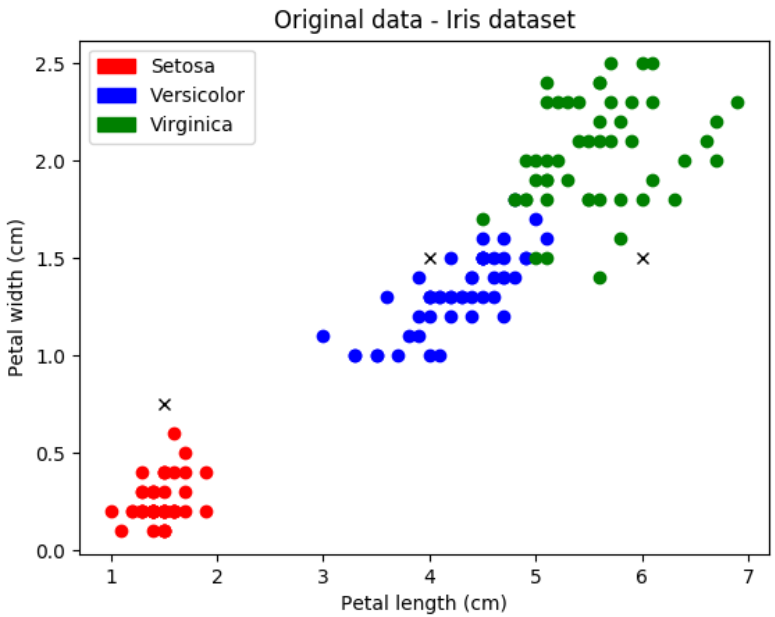
Goal

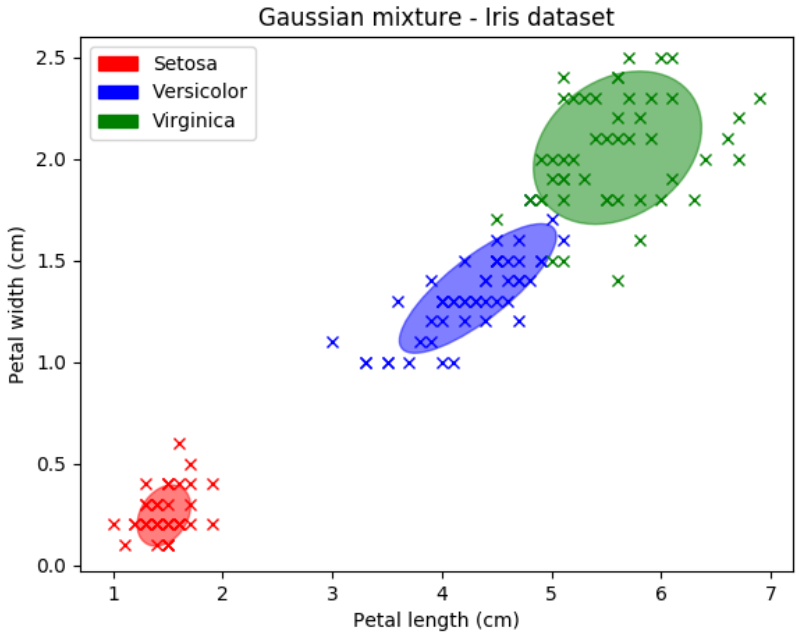
Estimate the mean and covariance of the multivariate Gaussians representing the data

Expectation-Maximization algorithm

- Expectation
  Compute expected value of likelihood function with respect to current parameters estimate
- Maximization
  Estimate the parameters which maximize the expected value

Gaussian mixture (2)





PhD-related example

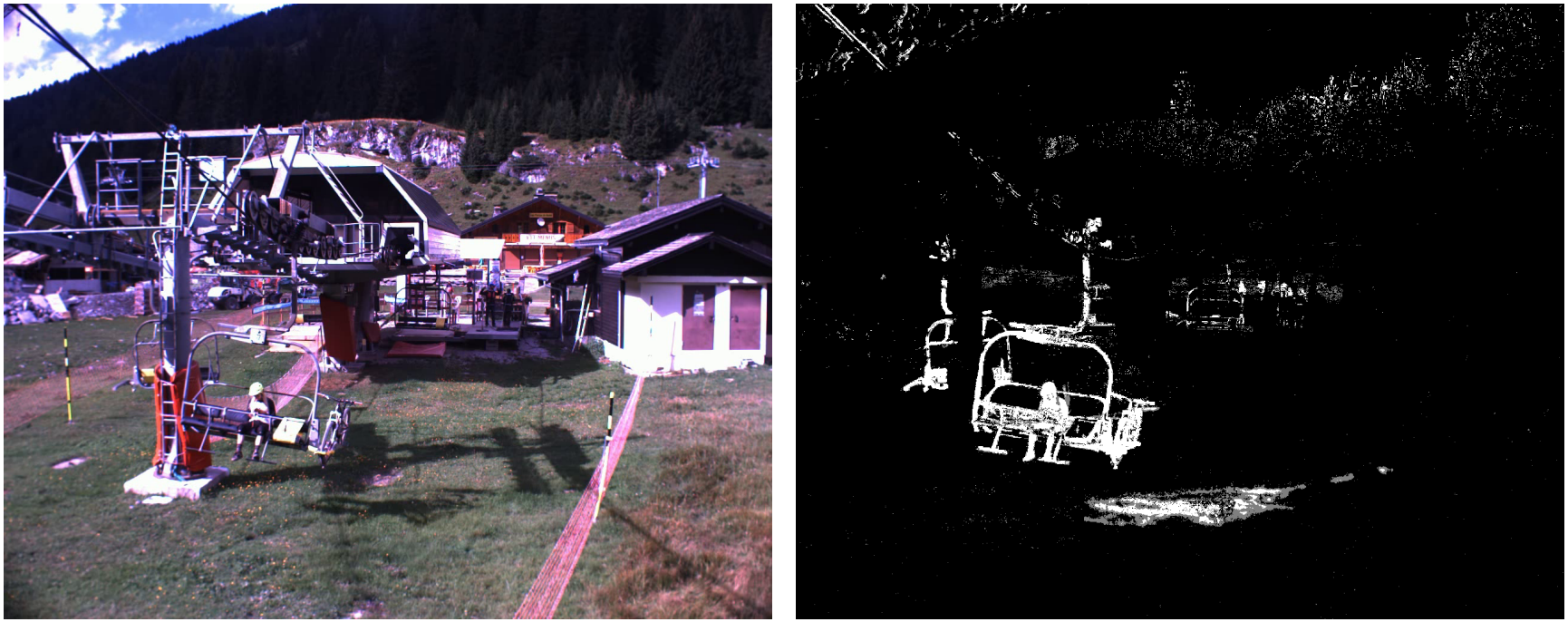
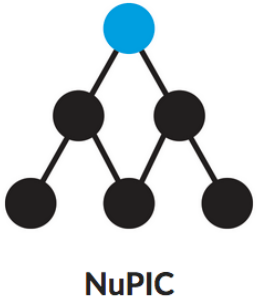
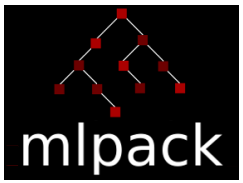
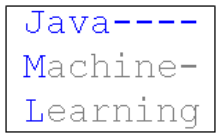


FIGURE – Foreground/Background/Shadow separation (Avoriaz - Lindarets)

Machine learning tools

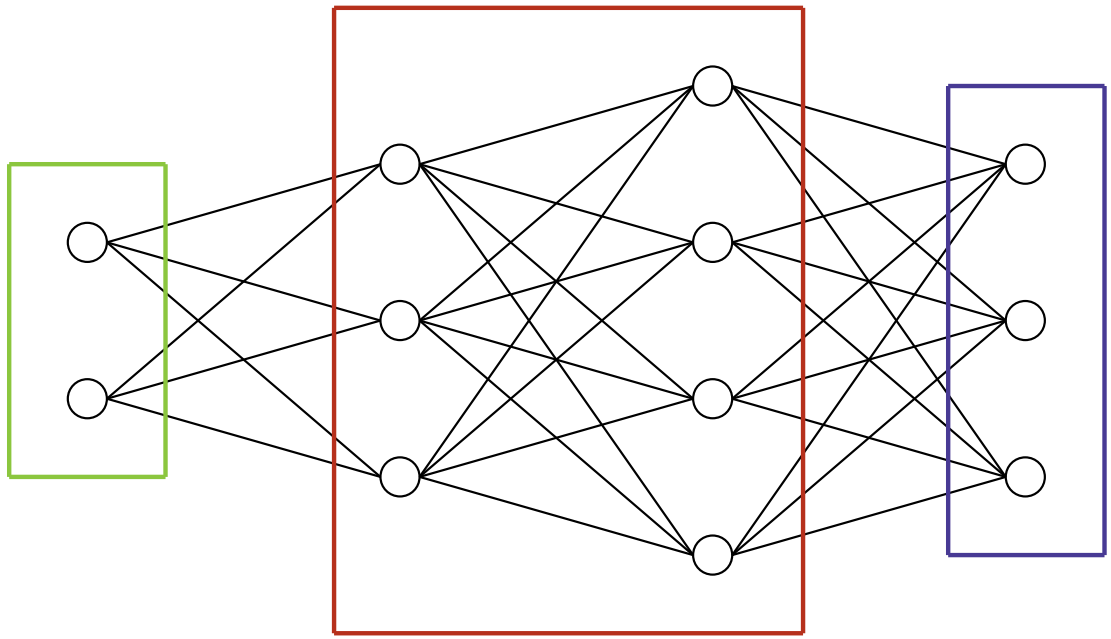






Basic architecture

- Inspired from the human brain
- Structure made up of layers of neurons
- Each neuron contain a value called **activation**



Weights and biases

A **weight** is associated to each connection between neurons of two consecutive layers. This represents the influence of one neuron on another neuron.

Weights and biases

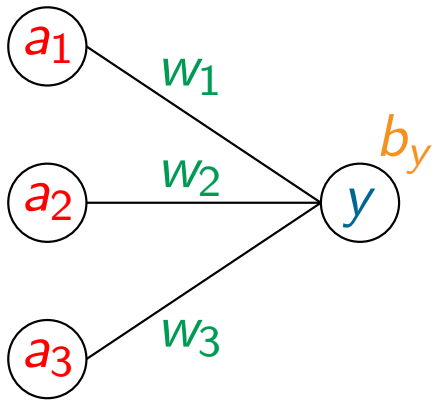
A **weight** is associated to each connection between neurons of two consecutive layers. This represents the influence of one neuron on another neuron.

A **bias** is associated to each neuron. It is a threshold beyond which the neuron is activated and means "something".

Weights and biases

A **weight** is associated to each connection between neurons of two consecutive layers. This represents the influence of one neuron on another neuron.

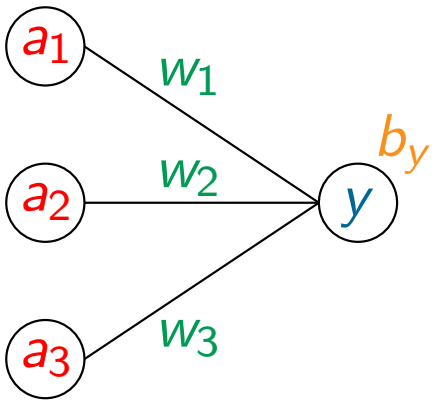
A **bias** is associated to each neuron. It is a threshold beyond which the neuron is activated and means "something".



Weights and biases

A **weight** is associated to each connection between neurons of two consecutive layers. This represents the influence of one neuron on another neuron.

A **bias** is associated to each neuron. It is a threshold beyond which the neuron is activated and means "something".



$$\Longrightarrow \quad y = w_1 a_1 + w_2 a_2 + w_3 a_3 + b_y$$
$$= \sum_i \left[w_i a_i \right] + b_y$$

Activation function

However, it is often desired to obtain an activation value within a specified interval.

Activation function

However, it is often desired to obtain an activation value within a specified interval.
$\Longrightarrow$ solution : include an **activation function** σ

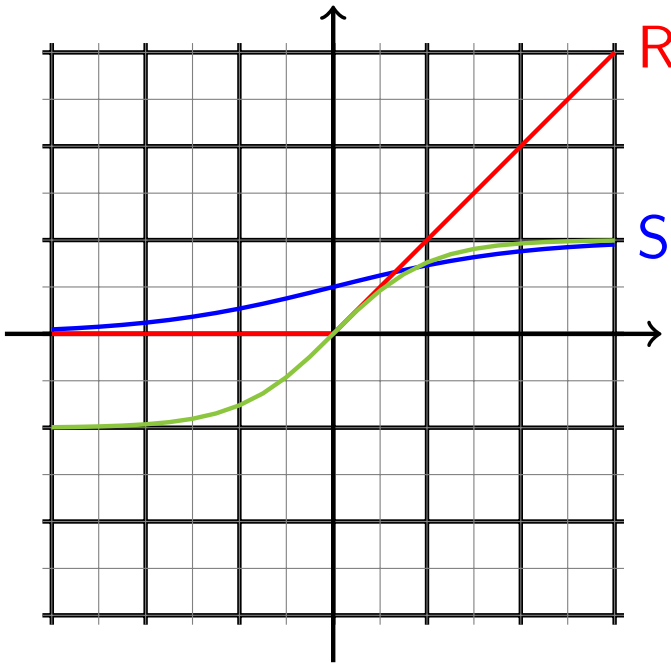
$$y = \sigma\left(\sum_i \left[w_i a_i\right] + b_y\right)$$

Activation function

However, it is often desired to obtain an activation value within a specified interval.

$\Longrightarrow$ solution : include an **activation function** σ

$$y = \sigma\left(\sum_i \left[w_i a_i\right] + b_y\right)$$



ReLU : $\sigma_1(x) = max(0, x)$

Sigmoid : $\sigma_3(x) = \frac{1}{1+e^{-x}}$

$\sigma_2(x) = tanh(x)$

Gradient descent and back propagation (1)

Three steps when working with neural networks : (a) **training**, (b) **validation** and (c) **testing**. That means one set for each step.

(a) The network is trained in a supervised way, *i.e.* with data for which groundtruth is available. The optimal weights are subsequently determined.

(b) The network parameters are precisely computed : number of layers, size of each layer, ...

(c) The network is tested with "real" data and its power is confirmed (or invalidated).

Gradient descent and back propagation (2)

Question

How to train a neural network ?

Answer

Gradient descent algorithm alongside backpropagation

Workflow

I **Propagation**
  i Forward propagation (input ➔ output)
  ii Error calculation according to known output by means of a *loss function*
  iii Backward error propagation in order to compute gradients (output ➔ input)

II **Weights update**
  i Weights modification according to gradients

Different types of neural networks

- ➜ Convolutional Neural Network [CNN]
- ➜ Recurrent Neural Network [RNN]
- ➜ Long Short-Term Memory [LSTM]
- ➜ Auto-Encoder [AE]
- ➜ Multi-Layer Perceptron [MLP]

Thank you all for your attention !