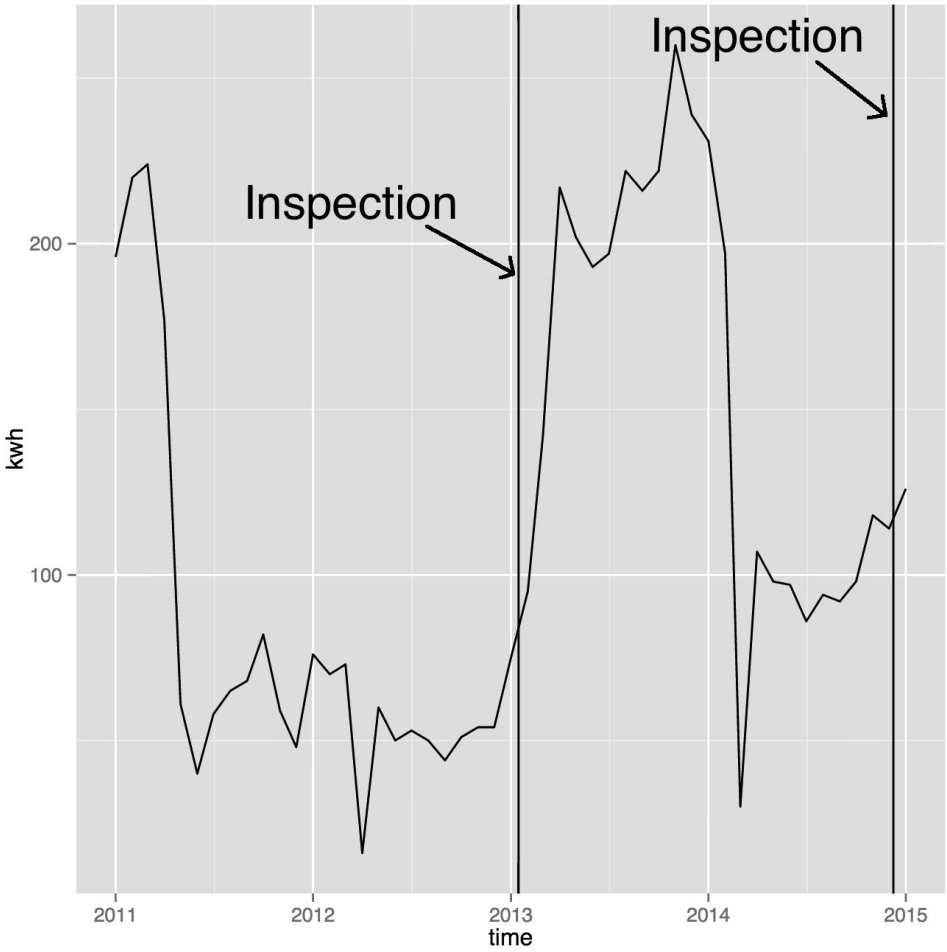
# Machine Learning for Data-Driven Smart Grid Applications

May 22, 2018

**Patrick Glauner**, **Jorge Augusto Meira** and **Radu State**
Interdisciplinary Centre for Security, Reliability and Trust,
University of Luxembourg

# Motivation: non-technical losses (NTL)

Example of NTL: Two assumed occurrences of NTL due to significant consumption drops followed by inspections (visualized by vertical bars).



# About us

- We work on real-world machine learning problems together with industry partners
- Recent research includes detection of electricity theft/non-technical losses, correction of biases in data and augmented reality

SnT

securityandtrust.lu

uni.lu

UNIVERSITÉ DU LUXEMBOURG

# Goals of this tutorial

- Providing an introduction to machine learning
- Understanding the three pillars of machine learning
- Knowing when to use which model

# Contents

- Introduction
- Supervised learning
- Unsupervised learning
- Reinforcement learning
- Deep Learning
- Conclusions

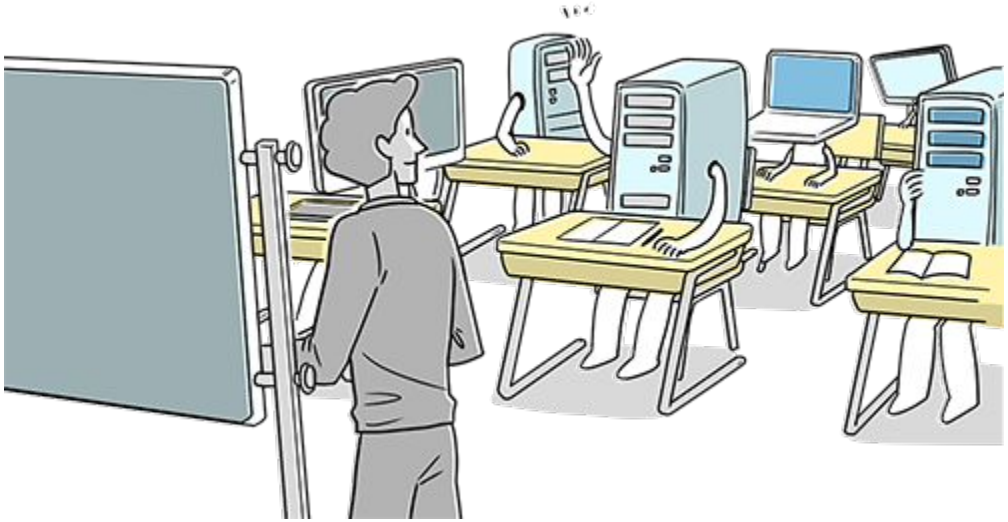
# Introduction

"Artificial intelligence is the science of knowing what to do when you don't know what to do." (Peter Norvig)

https://www.youtube.com/watch?v=rtmQ3xlt-4A

# Introduction

***Arthur Samuel* (1959):** *"Field of study that gives computers the ability to learn without being explicitly programmed"*.



# Introduction

# Introduction

**What do customers buy after viewing this item?**

Best Selling

Lenovo
N22 11.6-Inch HD Chromebook Laptop (Black) - (Intel Celeron N3060, 2 GB RAM, 32 GB EMMC, Chrome OS)
4.5 out of 5 stars 271
£109.99 Prime

Top Rated

Inateck
13-13.3 Inch Macbook Air/ Macbook Pro / Pro Retina Sleeve Case Cover Protective Bag Ultrabook Netbook Carrying Protector
5 out of 5 stars 35
£16.99 Prime

# Introduction

# Introduction

"Machine Learning is a subset of Artificial Intelligence techniques which use statistical models to enable machines to improve with experiences"

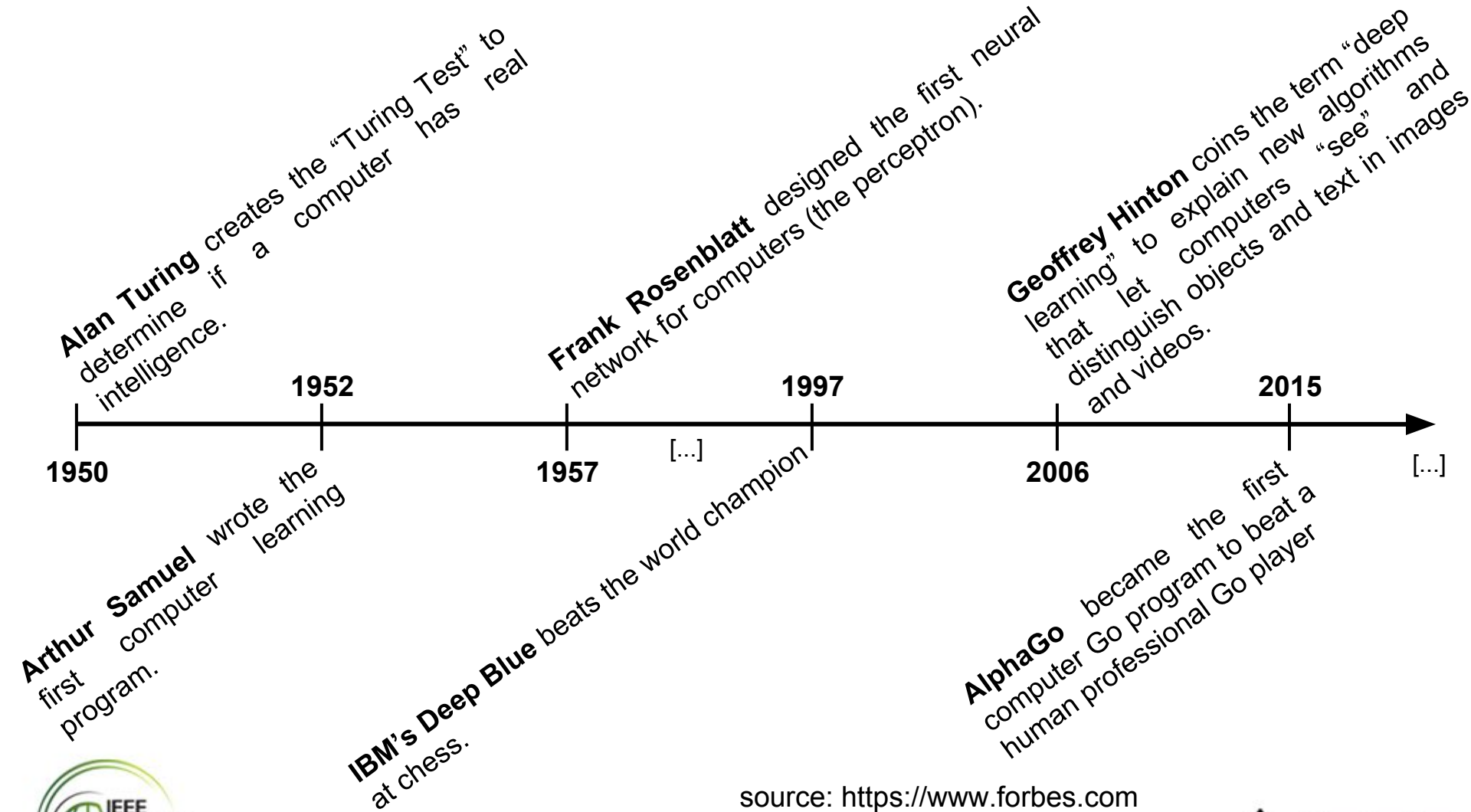
Use cases: *data mining, autonomous cars, recommendation...*

https://rapidminer.com/artificial-intelligence-machine-learning-deep-learning/

# Introduction

**Tom Mitchell (1998):** "*A computer program is said to learn from* ***experience E*** *with respect to some* ***class of tasks T*** *and* ***performance measure P*** *if its* ***performance at tasks in T****, as* ***measured by P****, improves with* ***experience E****."*

# Introduction




source: https://www.forbes.com

# Introduction: three pillars of machine learning

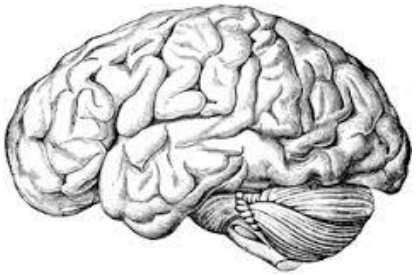
- Supervised learning: induce a function that maps from input to output.
- Unsupervised learning: find hidden structure in data.
- Reinforced learning: reward-based learning.

# Supervised learning

**Raw Data**

**Features**

**Models**



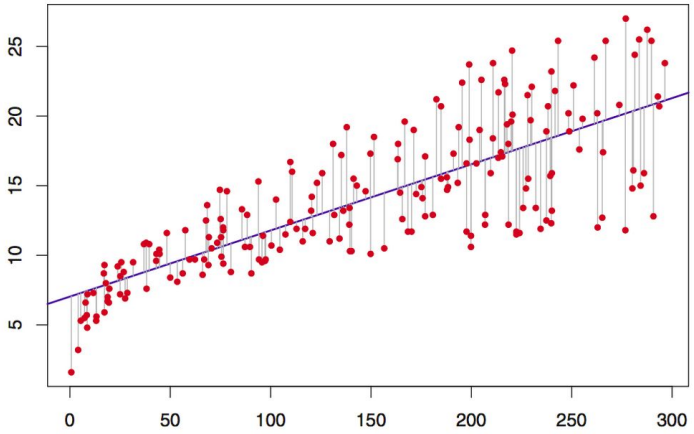
# Supervised learning

Regression

Classification

# Supervised learning

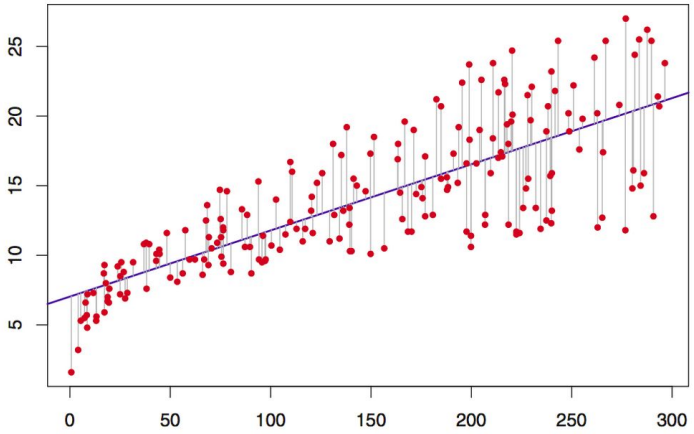
Regression



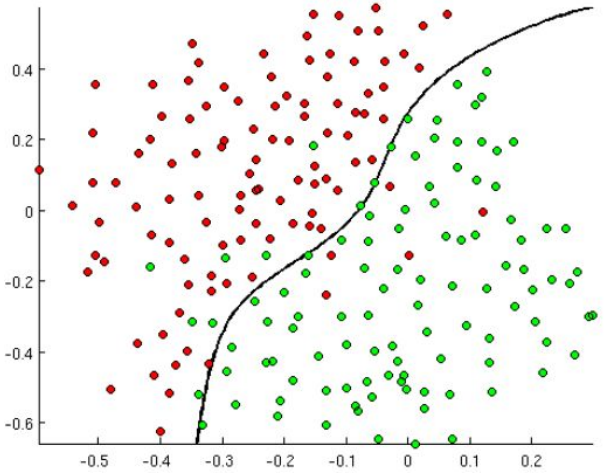

Classification

# Supervised learning

Regression

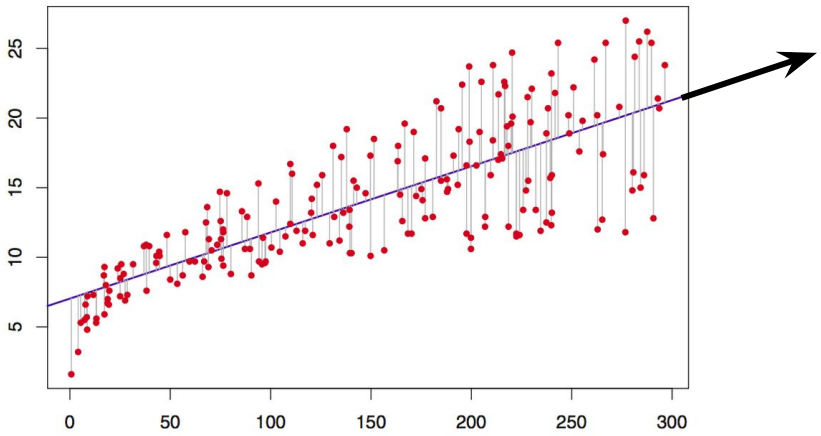



Classification



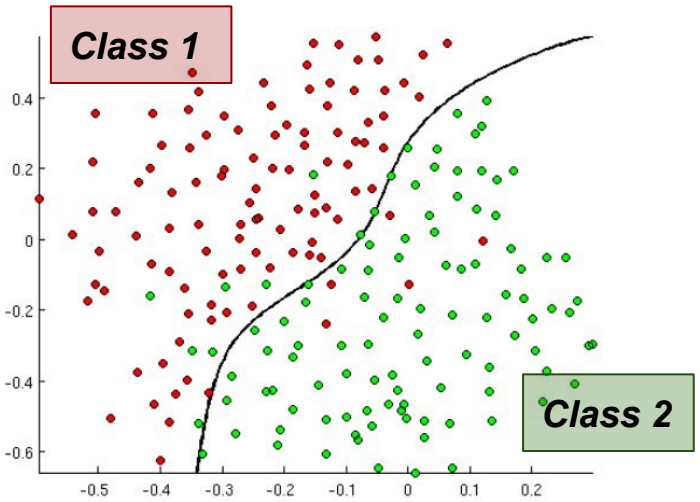
# Supervised learning

## Regression




***Continuous values***

## Classification




***Discrete values (categories)***

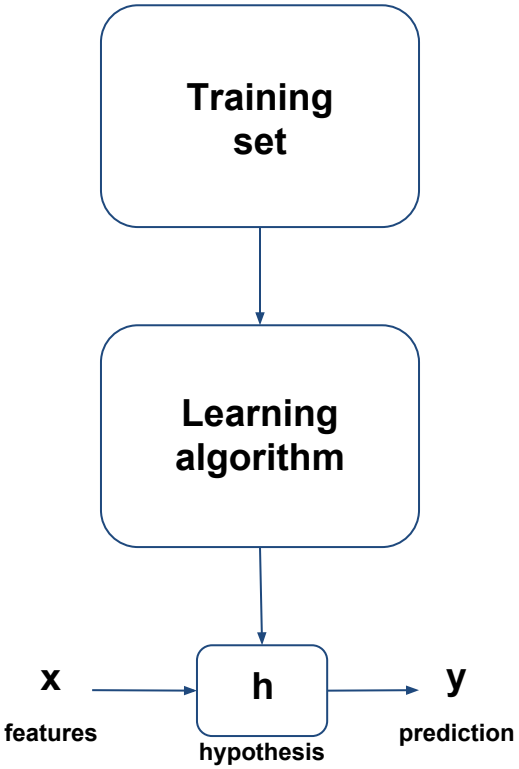
# Supervised learning: use cases

- Detection of anomalies
- Forecasting
- Medical diagnosis
- …

# Supervised learning: models

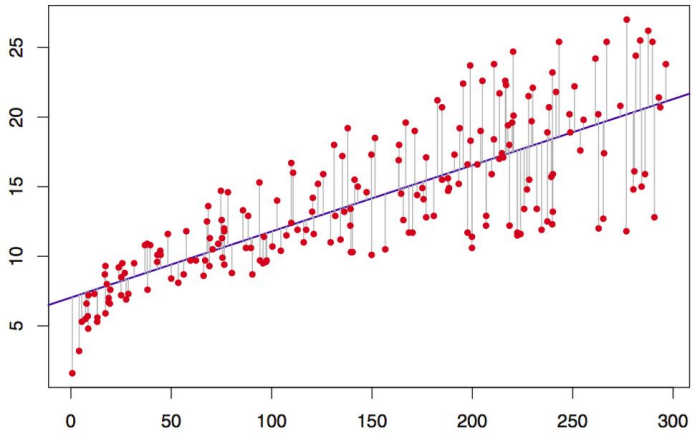
- Linear/logistic regression
- Decision tree, random forest
- Support vector machine
- …
- Neural networks, Deep Learning

# Supervised learning: workflow



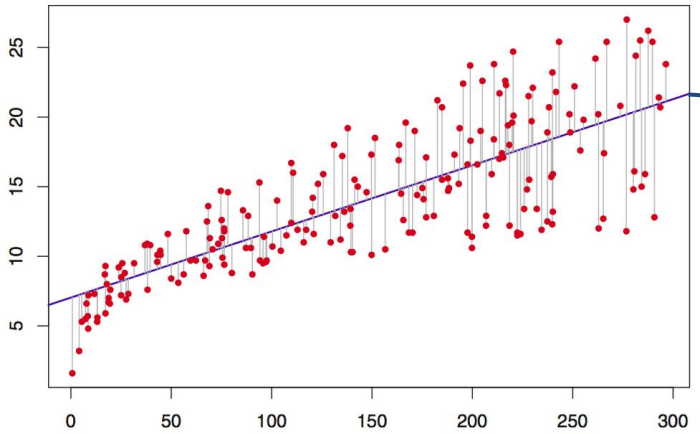
# Supervised learning: regression

## How to represent ‘h’ (hypothesis)



# Supervised learning: regression

## How to represent ‘h’ (hypothesis)

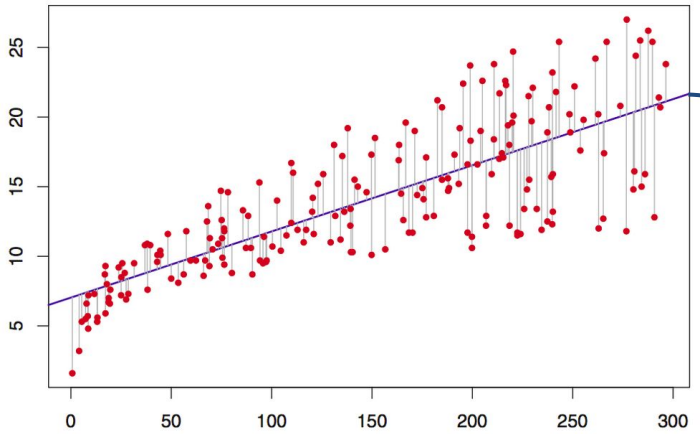



$h_\Theta(x) = \Theta_0 + \Theta_1 x$

$y = ax + b$

# Supervised learning: regression

## How to represent ‘h’ (hypothesis)




$h_\Theta(x) = \Theta_0 + \Theta_1 x$

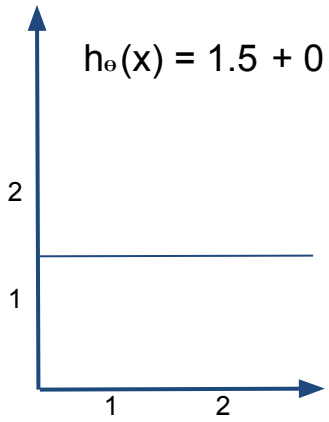
$y = ax + b$

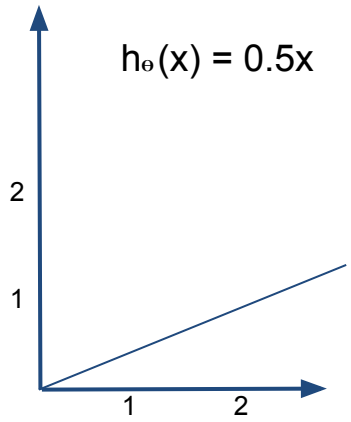
For example, housing price.

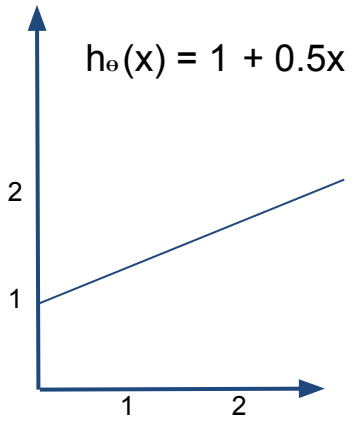
| 2014 | 2015 | 2016 | 2017 | 2018 |
|---|---|---|---|---|
| ~180k | ~182k | ~184k | ~186k | ??? |

# Supervised learning: regression

## How to represent ‘h’ (hypothesis)

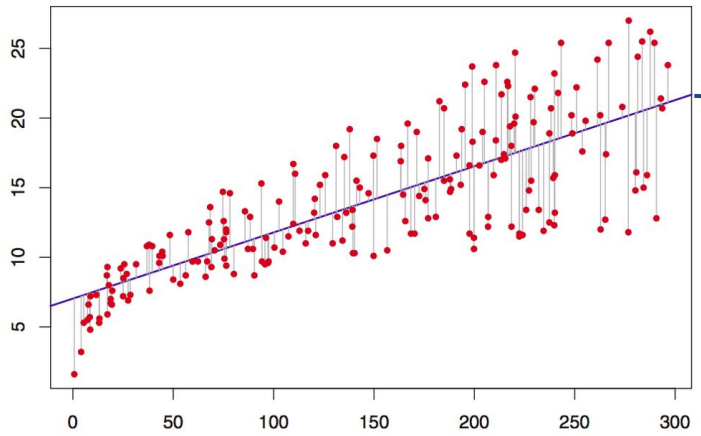


# Supervised learning: regression

## How to represent ‘h’ (hypothesis)

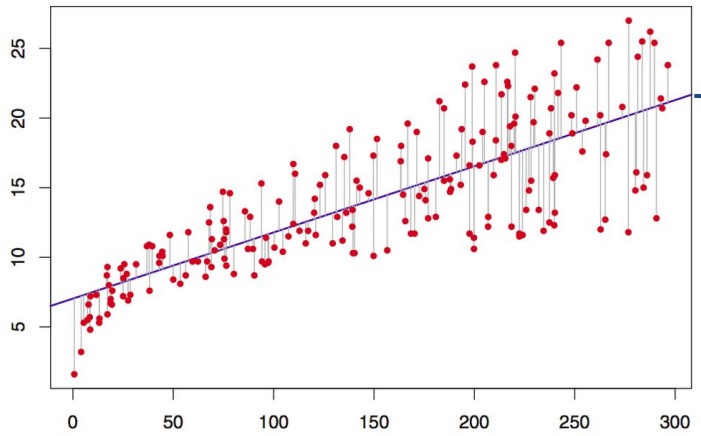



$h_\theta(x) = \Theta_0 + \Theta_1 x$

Choose $\Theta_0\Theta_1$ so that $\mathbf{h_\theta(x)}$ is close to **y** for our training set

# Supervised learning: regression

## How to represent ‘h’ (hypothesis)




$h_\theta(x) = \Theta_0 + \Theta_1 x$

Choose $\Theta_0\Theta_1$ so that $\mathbf{h_\theta(x)}$ is close to **y** for our training set

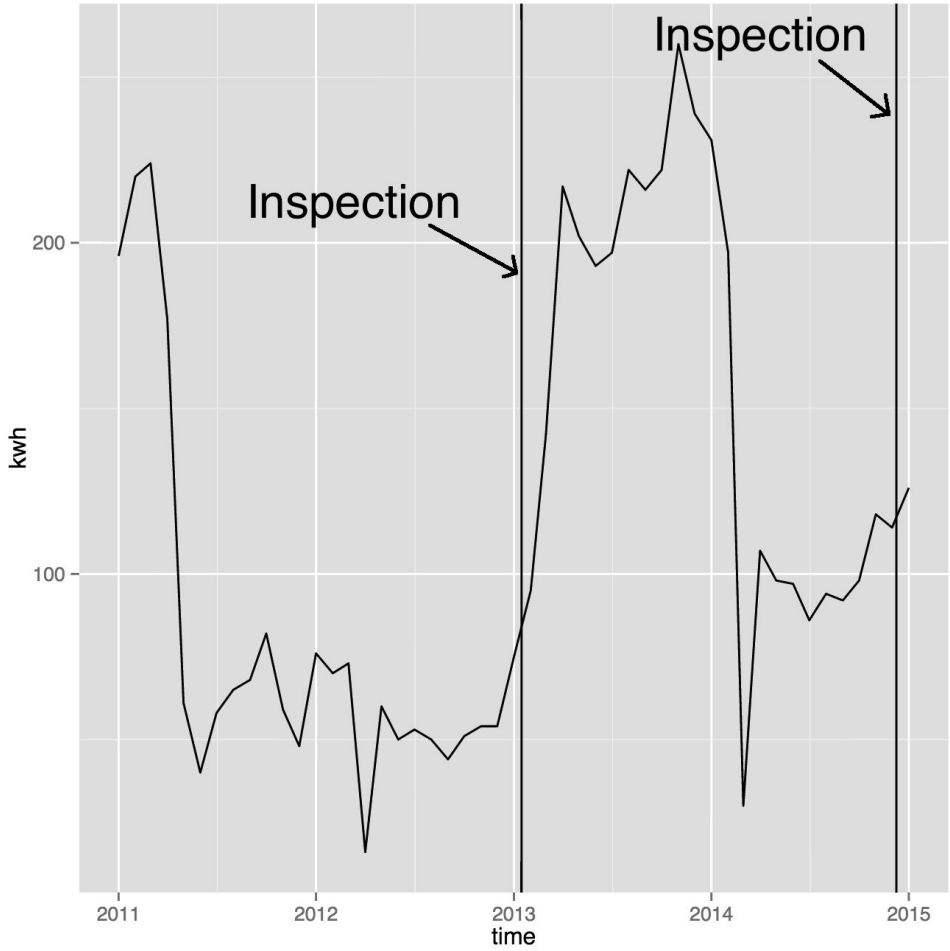
The idea is to minimize $\Theta_0\Theta_1$, so that $\mathbf{h_\theta(x)-y}$ tends to decrease.

Thus, we can define the cost function $\mathbf{J(\Theta_0\Theta_1)}$ aiming to minimize $\Theta_0\Theta_1$**:**

$$\mathbf{J(\Theta) = \tfrac{1}{2}\sum (h_\theta(x) - y)^2}$$

# Supervised learning: classification

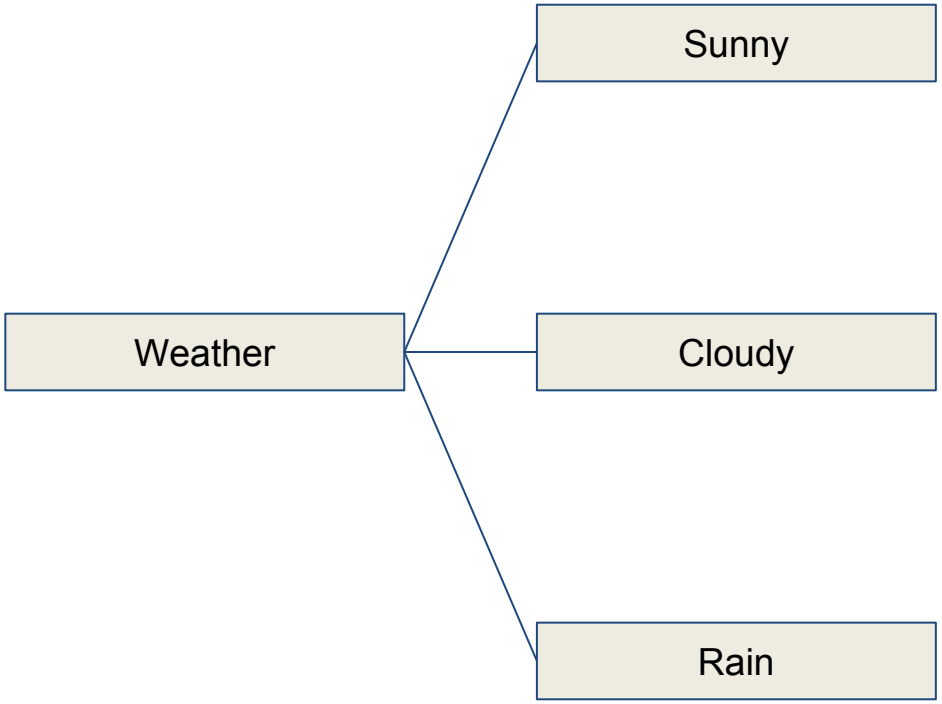
Example of non-technical losses (NTL): Two assumed occurrences of NTL due to significant consumption drops followed by inspections (visualized by vertical bars).



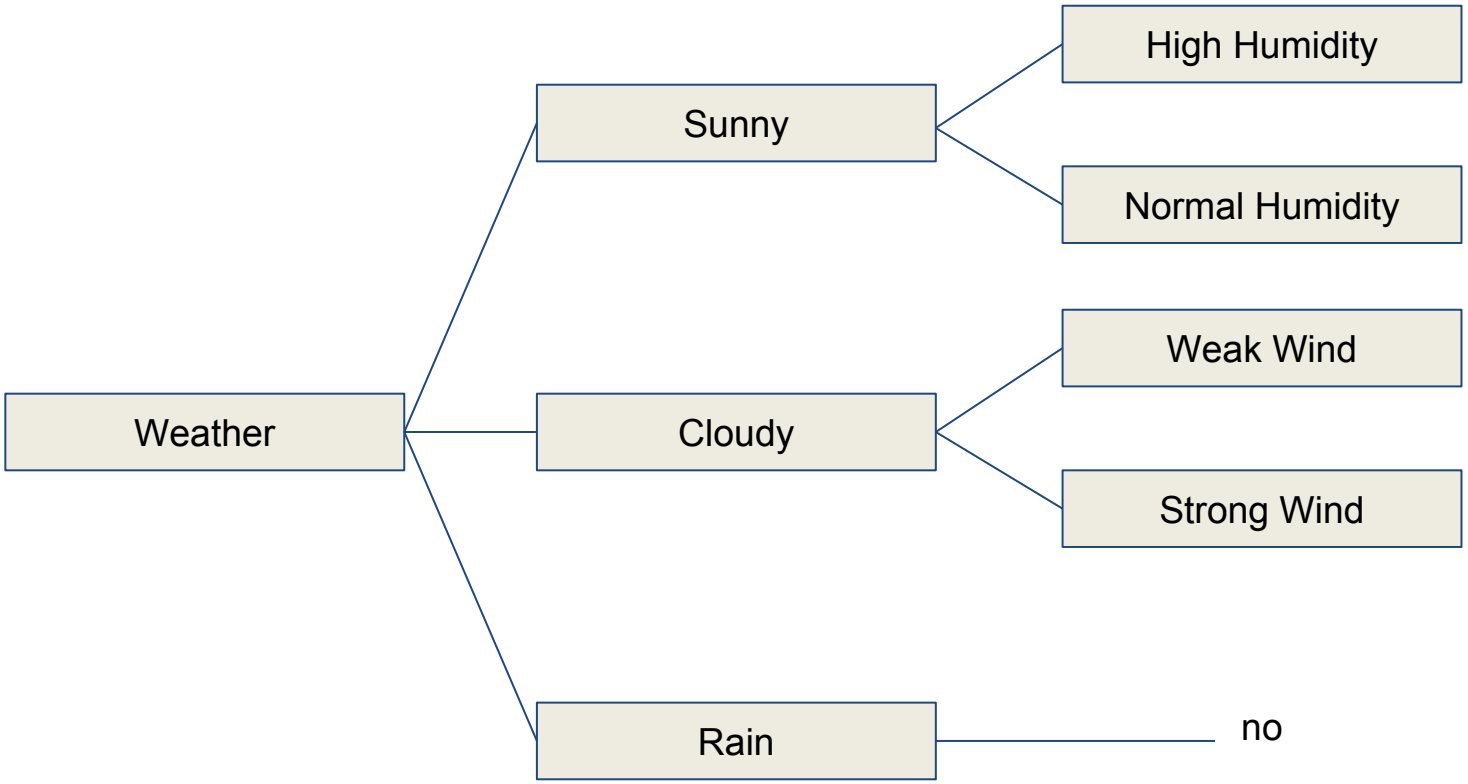
# Supervised learning: decision tree

Weather

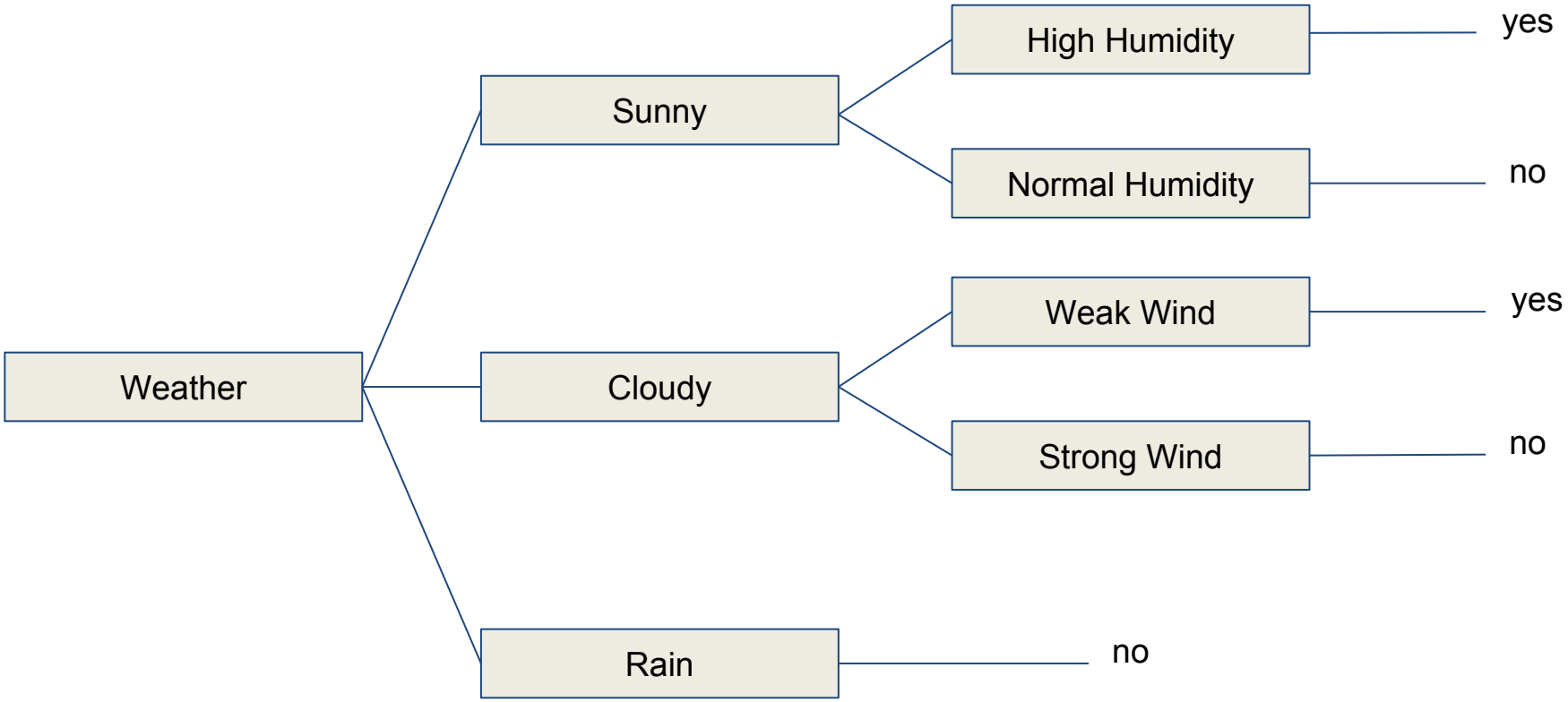
# Supervised learning: decision tree



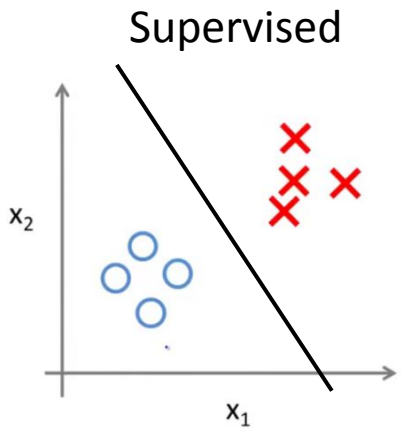
# Supervised learning: decision tree



# Supervised learning: decision tree



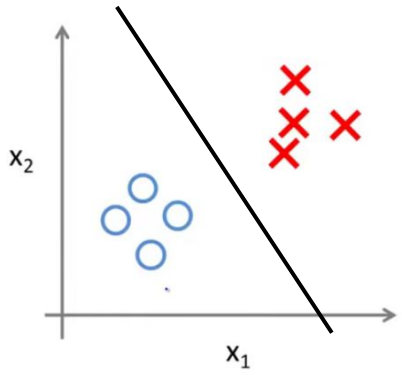
# Unsupervised learning




Known labels

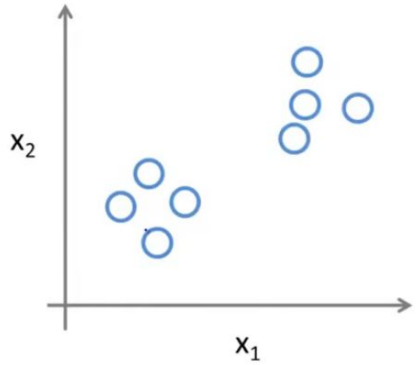
# Unsupervised learning

Supervised




Known labels

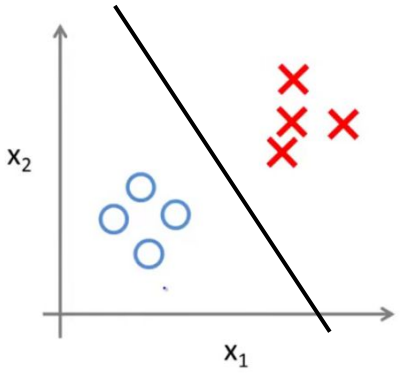
Unsupervised




Unknown labels

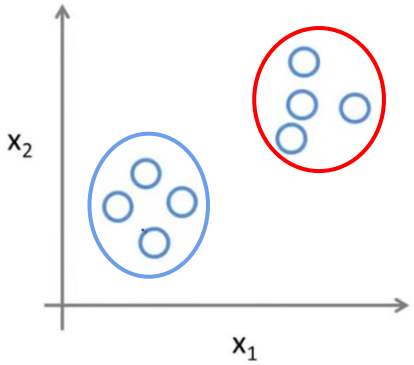
# Unsupervised learning

Supervised




Known labels

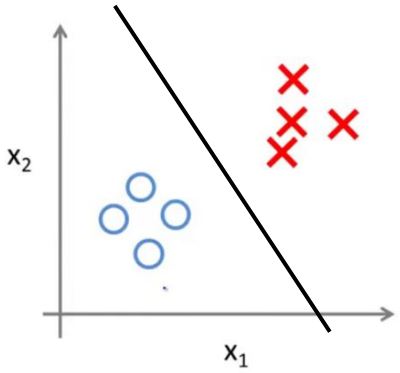
Unsupervised




Unknown labels

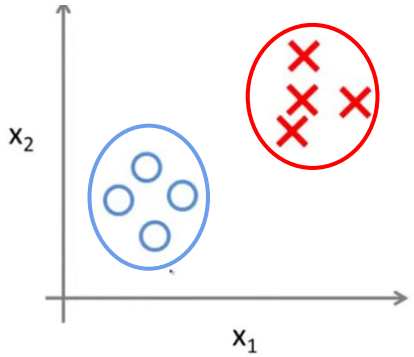
# Unsupervised learning

Supervised




Known labels

Unsupervised

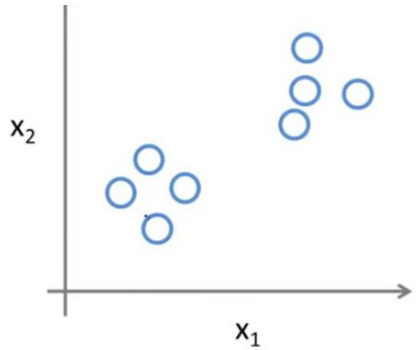



Unknown labels

# Unsupervised learning: clustering

**K-means algorithm**

1: Define **K** centroids randomly.

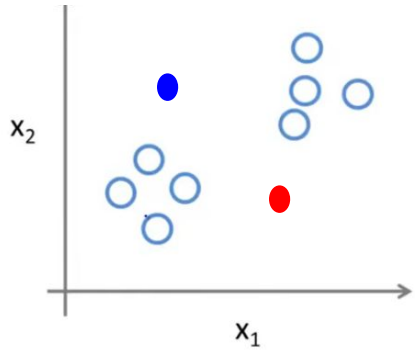


# Unsupervised learning: clustering

**K-means algorithm**

1: Define **K** centroids randomly.

2: Associate every observation according to the nearest centroid.

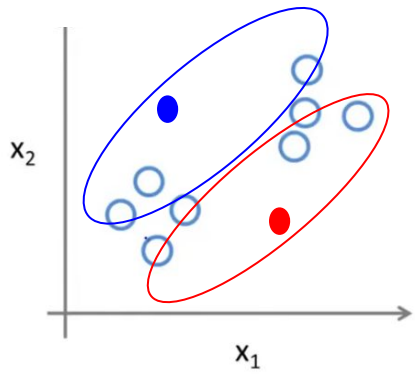


# Unsupervised learning: clustering

**K-means algorithm**

1: Define **K** centroids randomly.

2: Associate every observation according to the nearest centroid.



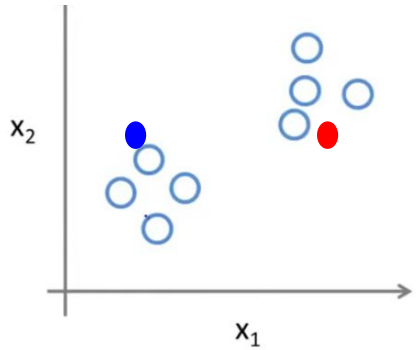
# Unsupervised learning: clustering

**K-means algorithm**

1: Define **K** centroids randomly.

2: Associate every observation according to the nearest centroid.

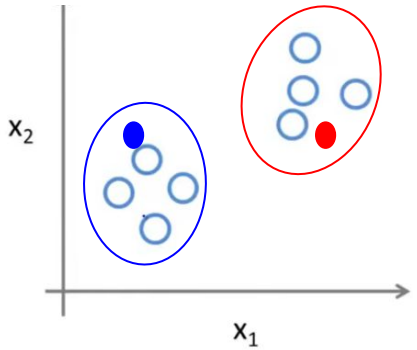
3: Define new centroids according to the mean of the clusters.



# Unsupervised learning: clustering

**K-means algorithm**

1: Define **K** centroids randomly.
2: Associate every observation according to the nearest centroid.
3: Define new centroids according to the mean of the clusters.
4: Repeat step 2 and 3 to converge.



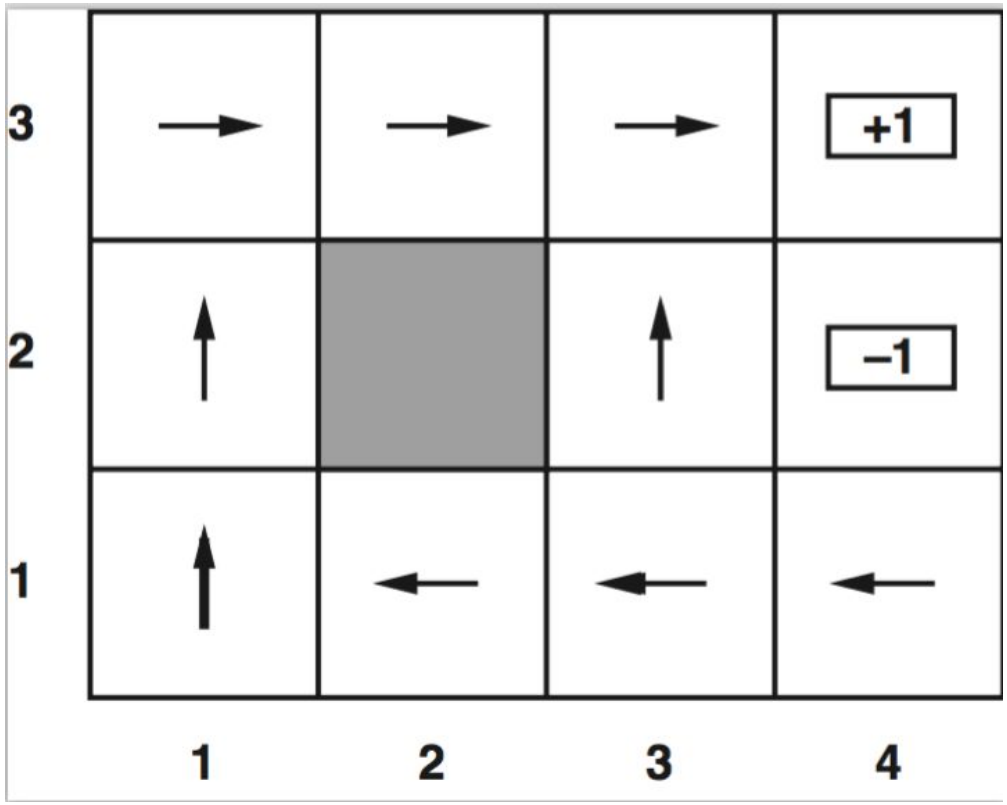
# Unsupervised learning: use cases

- Market segmentation
- Clustering of customers, news, etc.
- Dimensionality reduction of data

# Unsupervised learning: models

- k-means clustering
- Expectation-maximization clustering
- Principal component analysis
- …

# Reinforcement learning



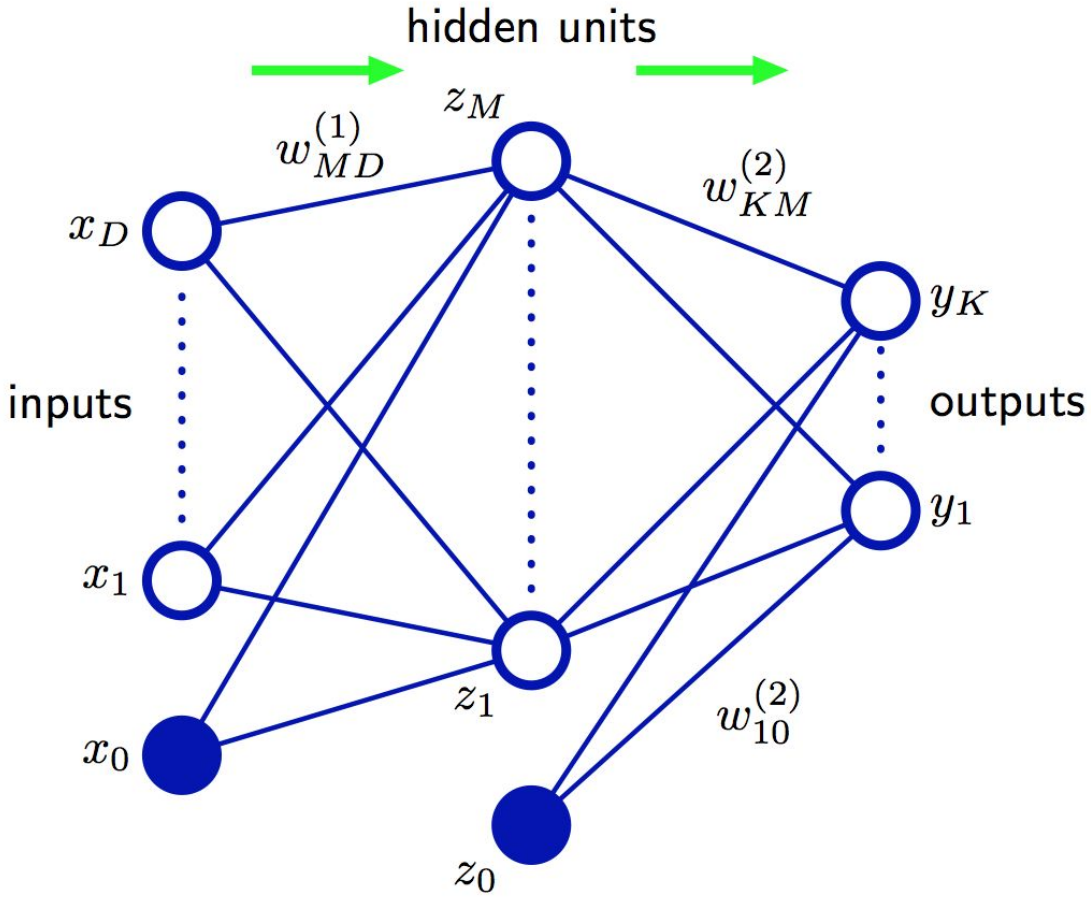
# Reinforcement learning: use cases

- Planning
- Playing games, e.g. the game of Go
- …

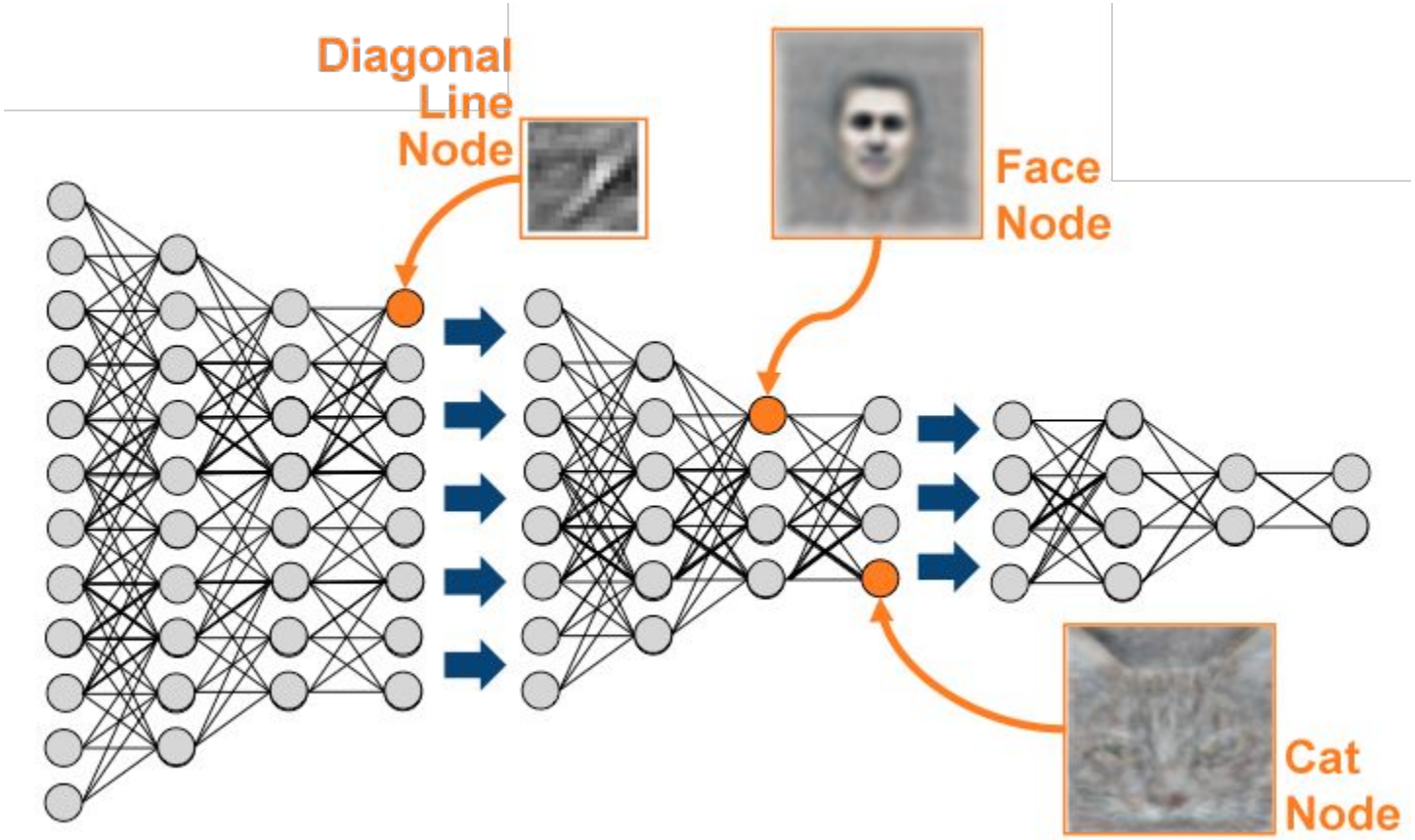
# Reinforcement learning: models

- Value/policy iteration
- Q-learning
- Deep reinforcement learning
- …

# Deep Learning: neural network



# Deep Learning



# Conclusions

- Machine Learning allows to learn complex statistical patterns from data
- Not much domain knowledge required
- Many applications in daily life
- Tell us more about your workflows so that we can figure out how Machine Learning can help you!

# References

[1] T. Mitchell, "**Machine learning**", McGraw Hill, 1997.
[2] C. Bishop, "**Pattern recognition and machine learning**", Springer, 2006.
[3] P. Glauner, et al., "**The Top 10 Topics in Machine Learning Revisited: A Quantitative Meta-Study**", Proceedings of the 25th European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning (ESANN 2017), Bruges, Belgium, 2017.
[4] P. Glauner, et al., "**Impact of Biases in Big Data**", Proceedings of the 26th European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning (ESANN 2018), Bruges, Belgium, 2018.