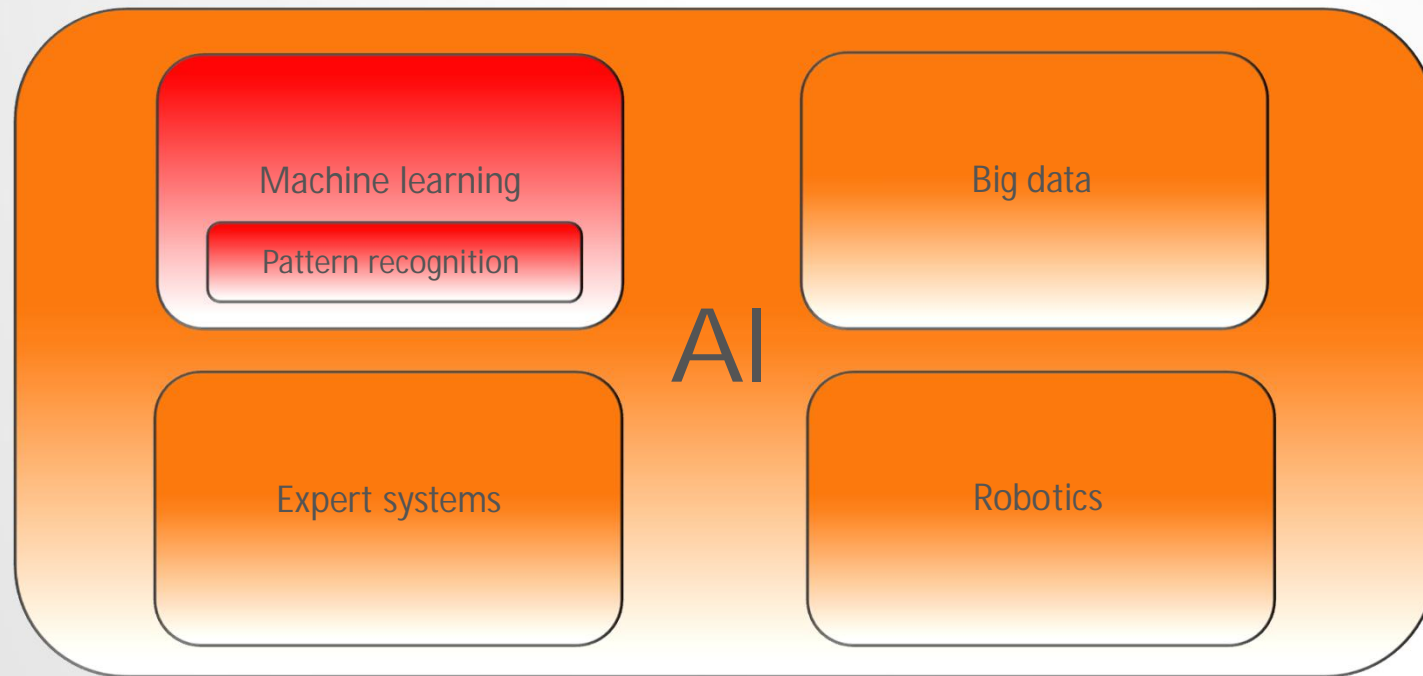


Machine Learning

Dr. Peter Axelberg 2016

Machine Learning is a subfield of AI



What is Machine learning (ML)

Machine learning is a core part of AI and focuses on the development of computer programs with the ability to learn without being explicitly programmed.



Machine learning (ML)

- q Herbert Simon:
"Learning is any process by which a system (computer, robot etc) improves performance from experience".
- q Machine Learning is concerned with computer programs that automatically improve their performance through experience



Herbert Simon
Turing Award 1975
Nobel Prize in Economics 1978

Why Machine learning?

- q Can replace certain manual tedious tasks
- q A well trained ML algorithm can handle large amount of data much more efficiently and accurately than a human
- q Can find new (hidden) knowledge (i.e. pattern or trends) in large datastreams (big data, data mining)
- q Is much faster than a human when doing the same task
- q etc.



Why ML now?

- q Floods of available data need to be analyzed or processed
 - Internet and big data information
- q Increased computational power
 - things that were not possible to achieve in the past is possible to achieve today thanks to more powerful computers
- q New useful and powerful ML algorithm has been developed in recent years
- q Demand driven: The Industry asks for ML solutions to be implemented in a variety of applications

ML applications



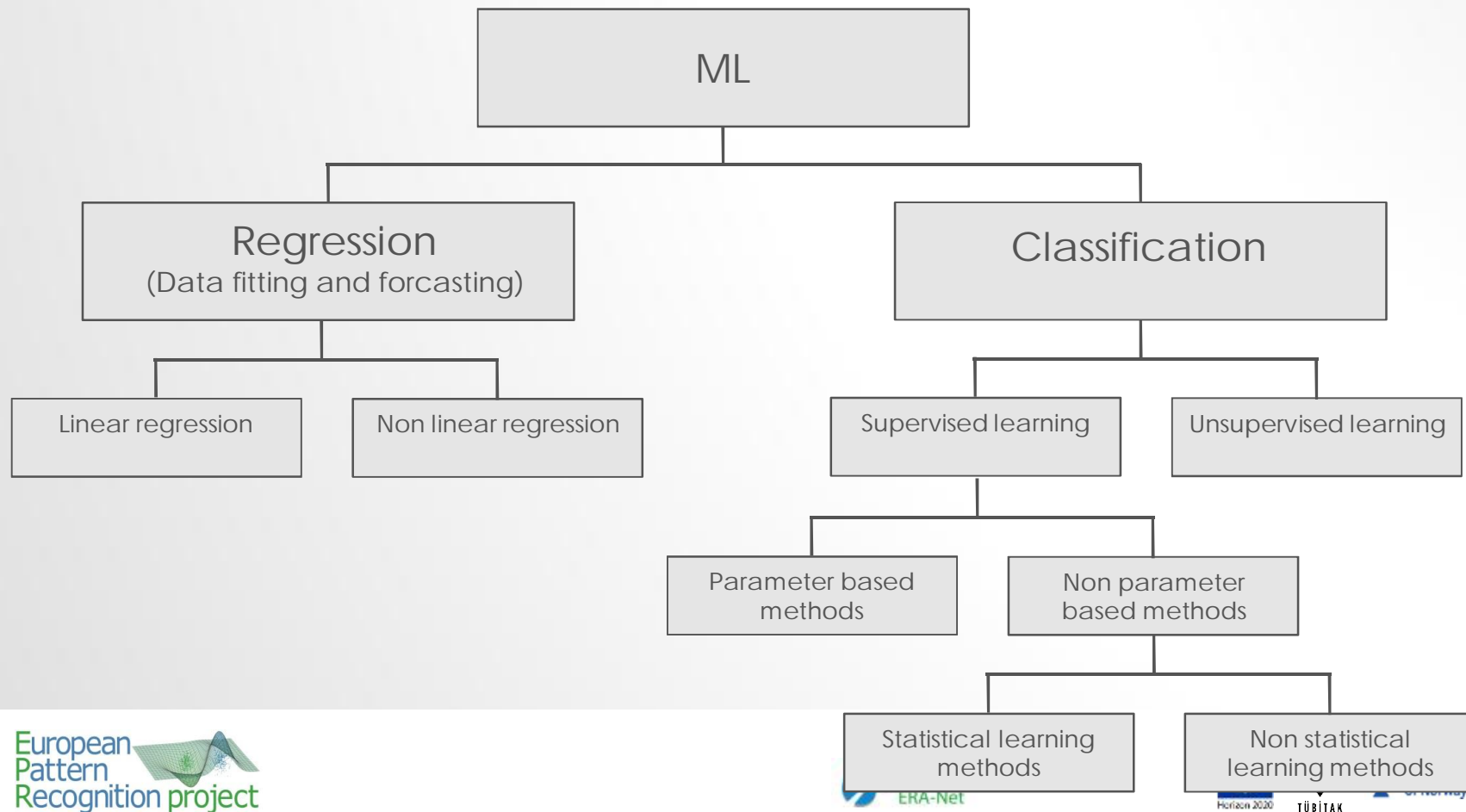
More ML applications

Optical Character Recognition (OCR)	<ul style="list-style-type: none">• <u>Handwritten</u>: sorting letters by postal code.• <u>Printed texts</u>: reading machines for blind people, digitalization of text documents.
Biometrics	<ul style="list-style-type: none">• <u>Face recognition, verification, retrieval.</u>• <u>Finger prints recognition.</u>• <u>Speech recognition.</u>
Diagnostic systems	<ul style="list-style-type: none">• <u>Medical diagnosis</u>: X-Ray, EKG (ElectroCardioGraph) analysis.

Basic Terminology

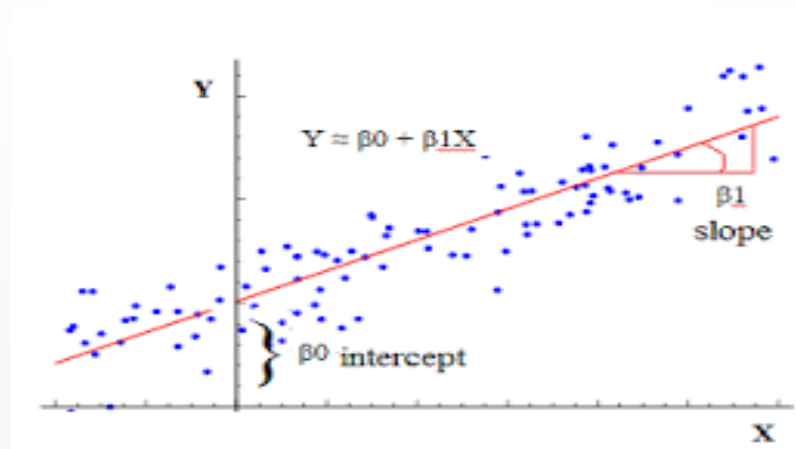
- q Object
 - something to be recognized by a ML algorithm
- q Feature (attribute)
 - anything that characterizes a certain object (color, weight, physical dimensions etc.)
- q Feature extraction
 - preparation of a feature vector
- q Feature vector
 - an n-dimensional vector of numerical features that represent an object
- q Training samples (training data)
 - a set of objects stored in a feature matrix that characterize an object
- q Test sample
 - an unknown object to be classified by the ML algorithm
- q Classifier
 - the nomenclature of a ML algorithm

Categories of ML



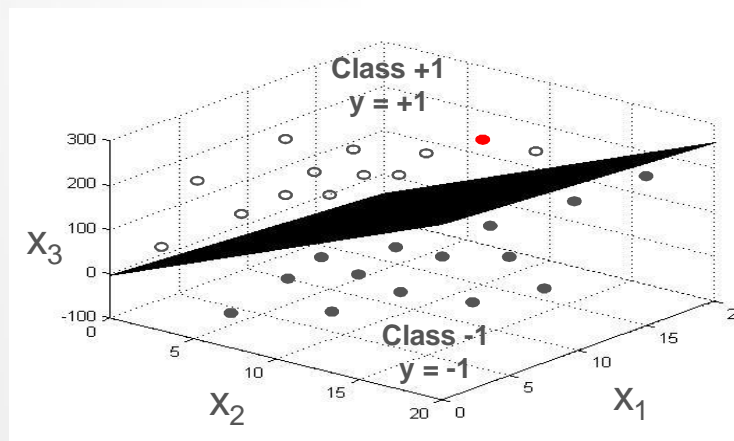
Regression

Regression is the ability of a ML algorithm to discover a trend in a datastream (data fitting)



Classification

Classification is the ability of a ML algorithm to decide to which class a particular test object belongs to



Let's take a look of the basic idea behind ML classification

The fundamental question in
ML classification:

How can I distinguish between different objects
that I want to classify?

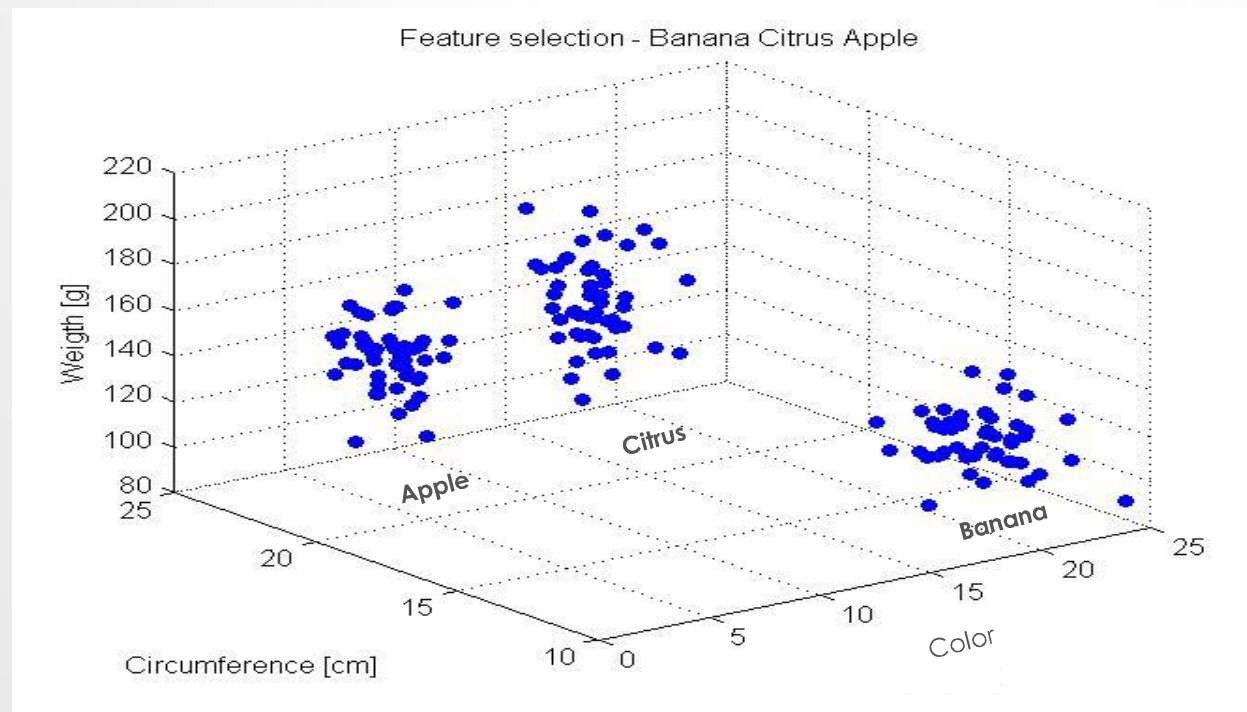
Feature extraction



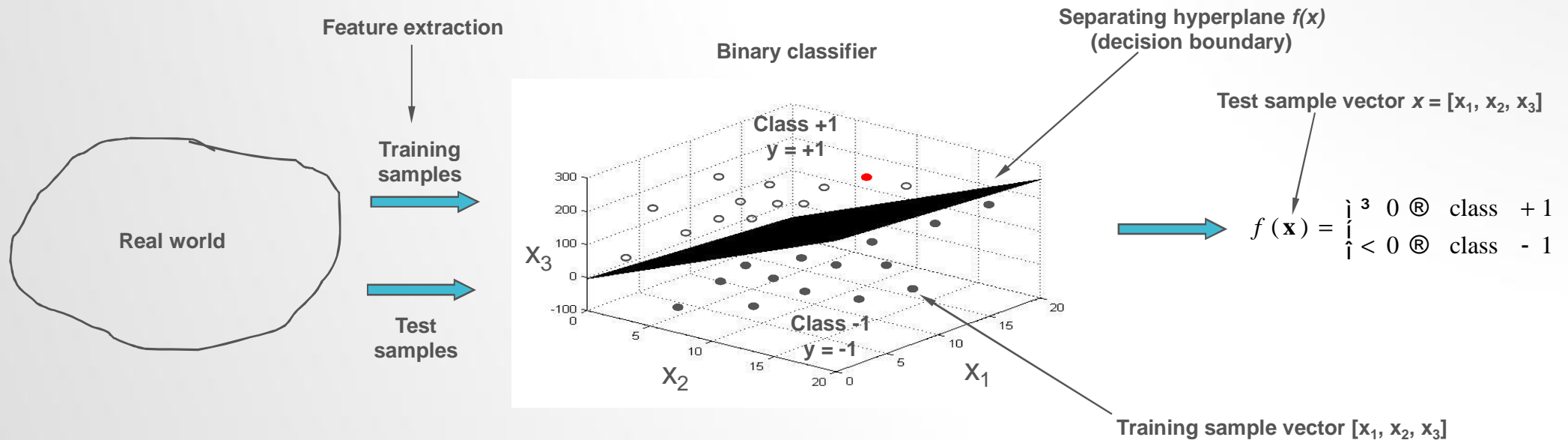
<u>Features</u>	<u>Banana</u>	<u>Citrus</u>	<u>Apple</u>
Length:	20 cm	15 cm	10 cm
Circumference:	12 cm	18 cm	22 cm
Weight:	120 g	180 g	140

The most important task: To define powerful features (attributes) that will distinguish one object from another

Results of feature extractions shown in a diagram



The Learning Process



Supervised learning

Uses a (large) number of *training samples* to train (learn) a *classification algorithm* (*classifier*) to be able to decide if *test data* belong to class +1 or class -1 .

Workflow

