ieeta instituto de engenharia electrónica e telemática de aveiro

universidade de aveiro

Departamento de Eletrónica, Telecomunicações e Informática

Machine Learning

LECTURE 1 : INTRODUCTION

Petia Georgieva
(petia@ua.pt)

Course Structure & Evaluation

Lectures/labs: 3 hours per week

Student Attendance: IEETA, office 1.13 :

Tuesday - 14h30-15h30
Thursday - 14h30-15h30

1.Practical component – 50% of the final grade
(evaluation in a group of two students)

2. Theoretical component - 50% of the final grade
(individual evaluation)

Evaluation

Practical component - 50% of the final grade

The practical component consists of two projects, developed in a group of two students.

The first project is evaluated based on a submitted report (IEEE format) and a short (10-15 min.) oral presentation.

The second project is evaluated based on a submitted paper (IEEE format).

The students are encouraged to use Latex text editor.
Overleaf is a convenient platform for collaborative writing and publishing using Latex (https://www.overleaf.com/) .

Theoretical Component – 50% of the final grade

Final individual exam during the last class.

PROGRAM

Supervised learning
Linear (univariate/ multivariate) regression
Logistic regression. Regularization
Artificial Neural Networks (ANN)
Support Vector Machines (SVM)
Naive Bayes. k-NN

Unsupervised learning
K-means clustering
Data dimensionality reduction
Principal components analysis (PCA)

Reinforcement learning
Markov Decision Process (MDP), Bellman equations
Value iteration and policy iteration
Learning a model for MDP
Fitted Value iteration

Deep Learning
Deep Neural Networks
Softmax Regression

Overview of ML

Grew out of work in Artificial Intelligence (AI) and the increasing computational power of computers.

Main reasons for ML algorithms:

- exponential growth of data – need for data mining

E.g., IoT, medical records, biology, engineering

- Applications can't be programed by hand.

E.g., Autonomous driving, handwriting recognition, Natural Language Processing (NLP), Computer Vision.

Computational Neuroscience - brain imaging (BCI)

A bit of history

- **1950,** Alan Turing: "Computing Machinery and Intelligence" define the question "Can machines think?“ =>Turing test.

- **1956** –The field of Artificial Inteligente (AI) formally established at the conference in Dartmouth College.

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

- **1998,** Tom M. Mitchell: “ Can the computer program learn from experience ? “.

Formal Definition of ML

„A computer program is said to learn from experience E with respect to some task T and some performance measure P, if its performance on T, as measured by P, improves with experience E.“ (T. Mitchell 1998)

- Given:
  - a task T
  - a performance measure P
  - some experience E with the task

- Goal:
  - generalize the experience in a way that allows to improve your performance on the task

Learning to recognize Spam-Mail

- Task:
  - sort E-mails into categories (e.g., Regular / Spam)
- Performance Measure:
  - Weighted Sum of Mistakes (letting spam through is not so bad as misclassifying regular E-mail as spam)
- Experience:
  - Handsorted E-mail messages in your folder

Computer vision
(e.g. Learning to detect & recognize faces)

Example training images
for each orientation

Learning to classify documents

Company home page

vs

Personal home page

vs

University home page

vs

…

Machine Learning Approaches

Supervised Learning

Given examples with “correct answer” (labeled examples)

(e.g. given dataset with spam/not-spam labeled emails)

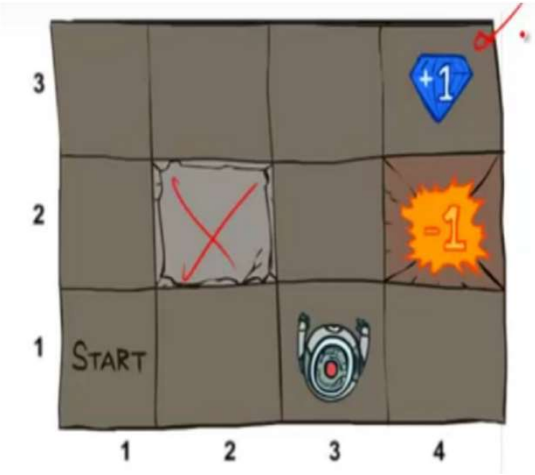
Unsupervised Learning

Given examples without answers (no labels).

Reinforcement Learning

On-line learning by taking actions and getting rewards/penalties.

(intelligent robotics =>



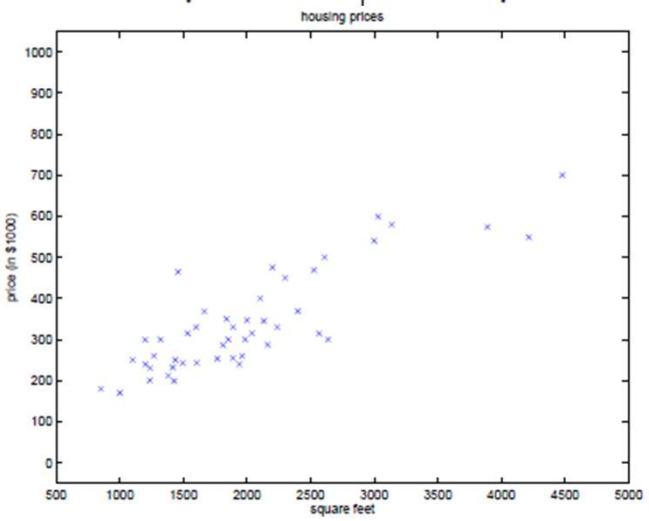
Deep Learning

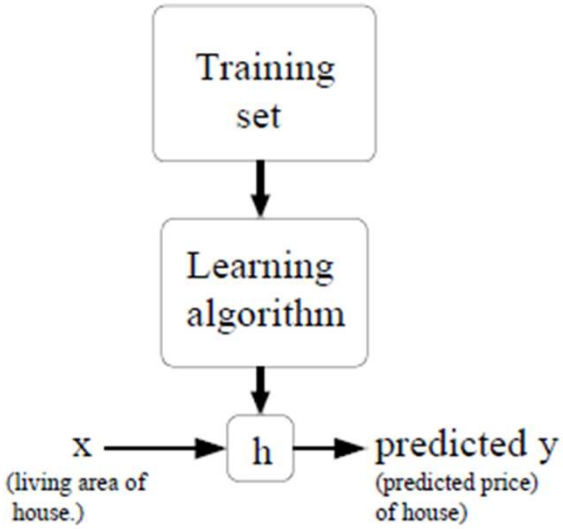
Automatically extract hidden features (in contrast to hand-crafted features). Need a lot of data (Big data) . Need for very high computational resources (GPUs). .

Supervised Learning – univariate regression

Problem: Learn a model to predict the house price as a function of the house area

Living area (feet^2)	Price (1000$s)
2104	400
1600	330
2400	369
1416	232
3000	540
⋮	⋮





$$h_\theta(x) = \theta^T x = \theta_0 + \theta_1 x_1$$

Supervised Learning – multivariate regression

Problem: Learn a model to predict the house price as a function of the house area and number of bedrooms.

Living area (feet^2)	#bedrooms	Price (1000$s)
2104	3	400
1600	3	330
2400	3	369
1416	2	232
3000	4	540
⋮	⋮	⋮

$$h_\theta(x) = \theta_0 + \theta_1 x_1 + \theta_2 x_2$$

Supervised Learning – classification

Problem: Learn to predict if the property is a house (H) or an apartment (A), as a function of its living area, # of bedrooms and the price.

Living area ($feet^2$)	#bedrooms	Price (1000$s)
2104	3	400
1600	3	330
2400	3	369
1416	2	232
3000	4	540
⋮	⋮	⋮

House/App.
H
A
H
A
H
….

Supervised Learning

- Given a collection of examples (*data set*). Each example contains:
  - set of features (attributes)
  - output - a label (continuous value or discrete values(class))
- Divide data into training (60-80%) and testing (40-20%) sub sets.
- **Training stage** : build classification/regression model with the training set
- **Model validation**: Apply the learned model to examples in the testing set.
- **Goal:** maximize the accuracy of the model on the testing set

Learning/training mode

Batch Learning

The model is updated only after all training examples are processed.

Stochastic/Incremental/On-line Learning

The model is updated after every single training example is processed.

Mini-batch learning

Batch learning is not practical for deep learning with big data. The model is updated with mini-batches of training examples.

Types of Features

1. **Nominal (Qualitative) features**
   - Ex: days of the week, seasons, job type
   - Ex: good, better, best
   - Color (red, green, blue)

   Integer encoding (1,2,3) / one-hot encoding(1,0)

2. **Numeric (Quantitative) features**

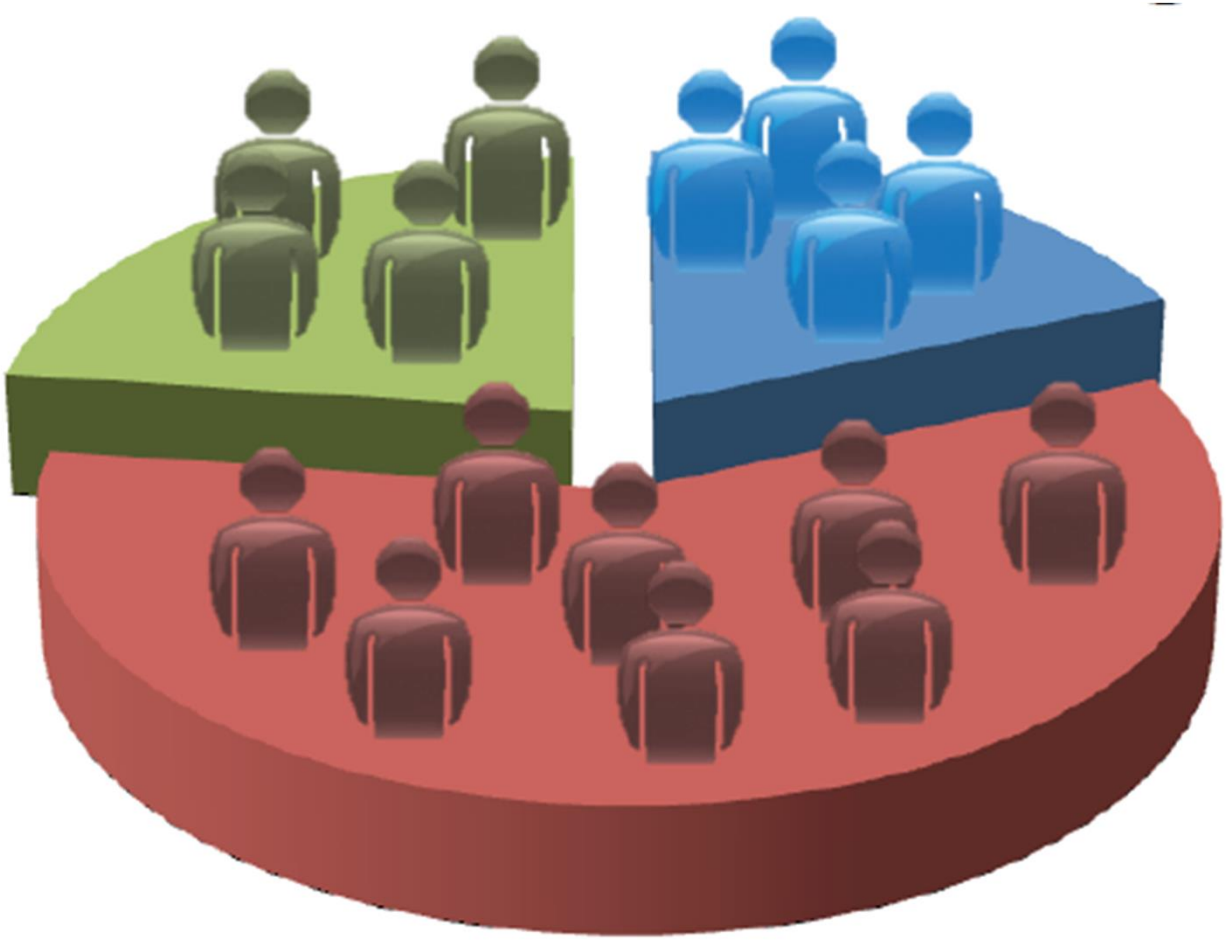
   2.1 Discrete features (integer numbers)
   - Ex: age, years, counts
   - Binary features (0,1)

   2.2 Continuous features (real numbers)

   Ex: temperature, height, weight, humidity, el. current, etc.

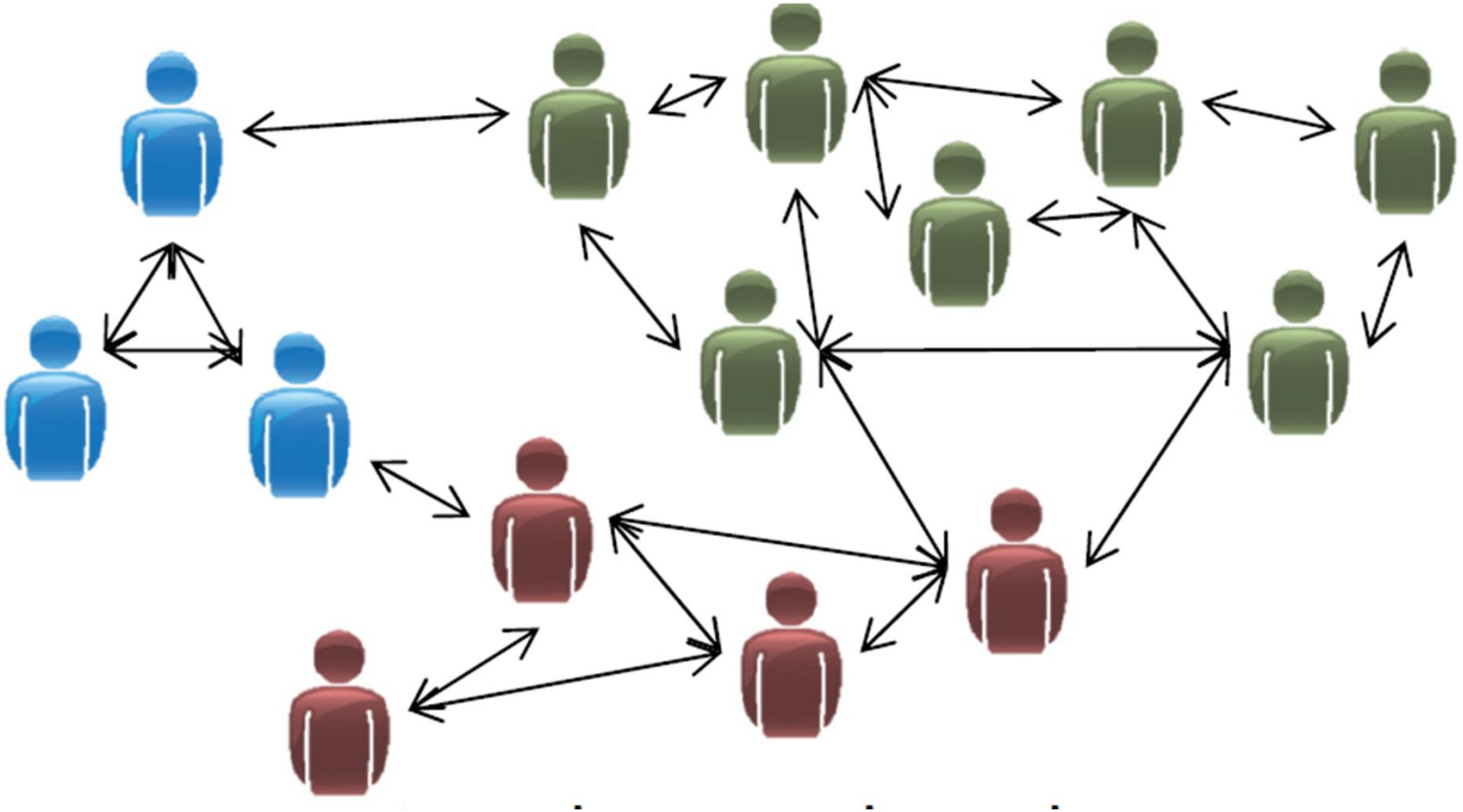
Unsupervised learning -

Market segmentation (data base of customers)



Unsupervised learning

Social network analysis

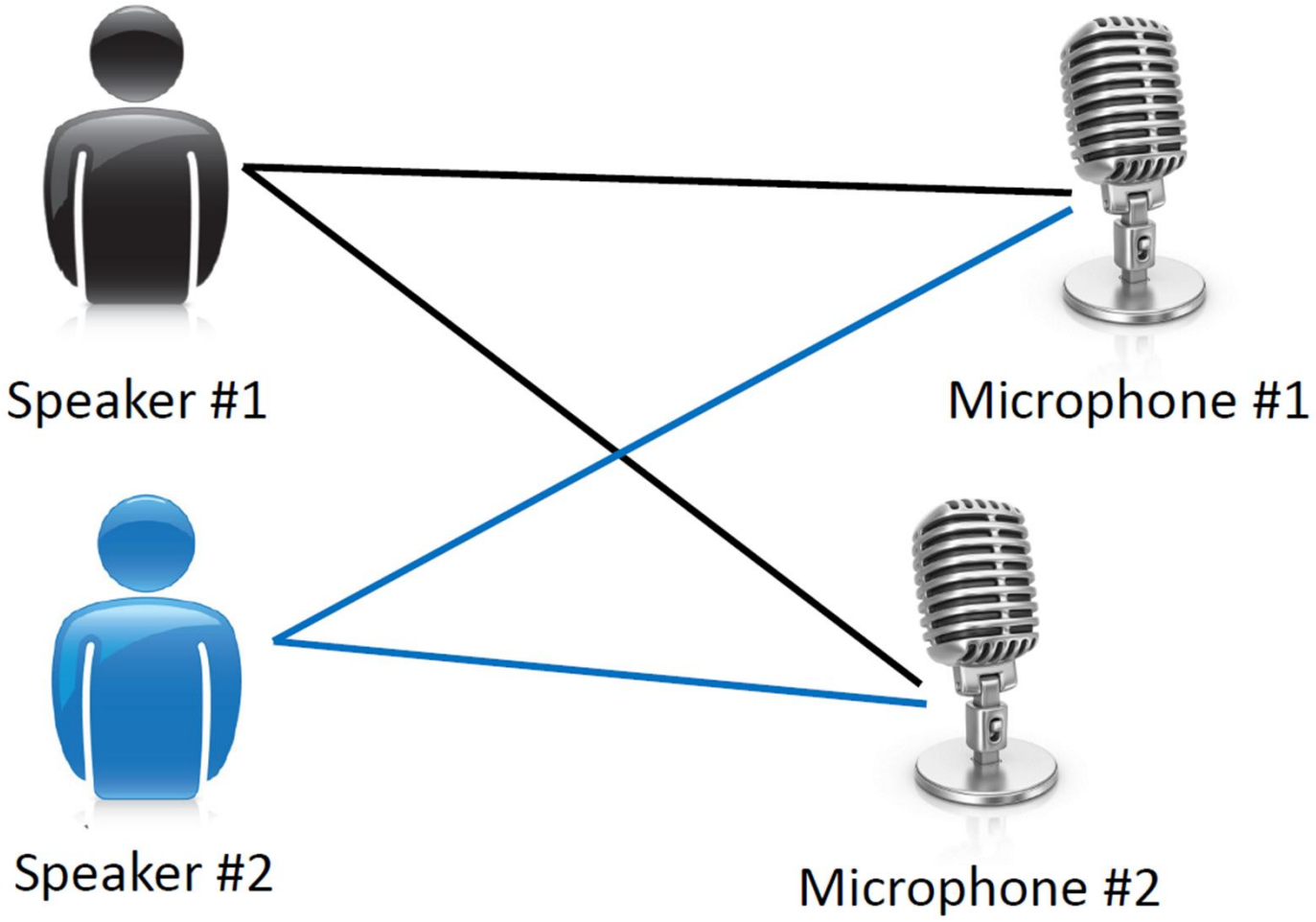


Unsupervised learning

Understand galaxies - astronomical data analysis

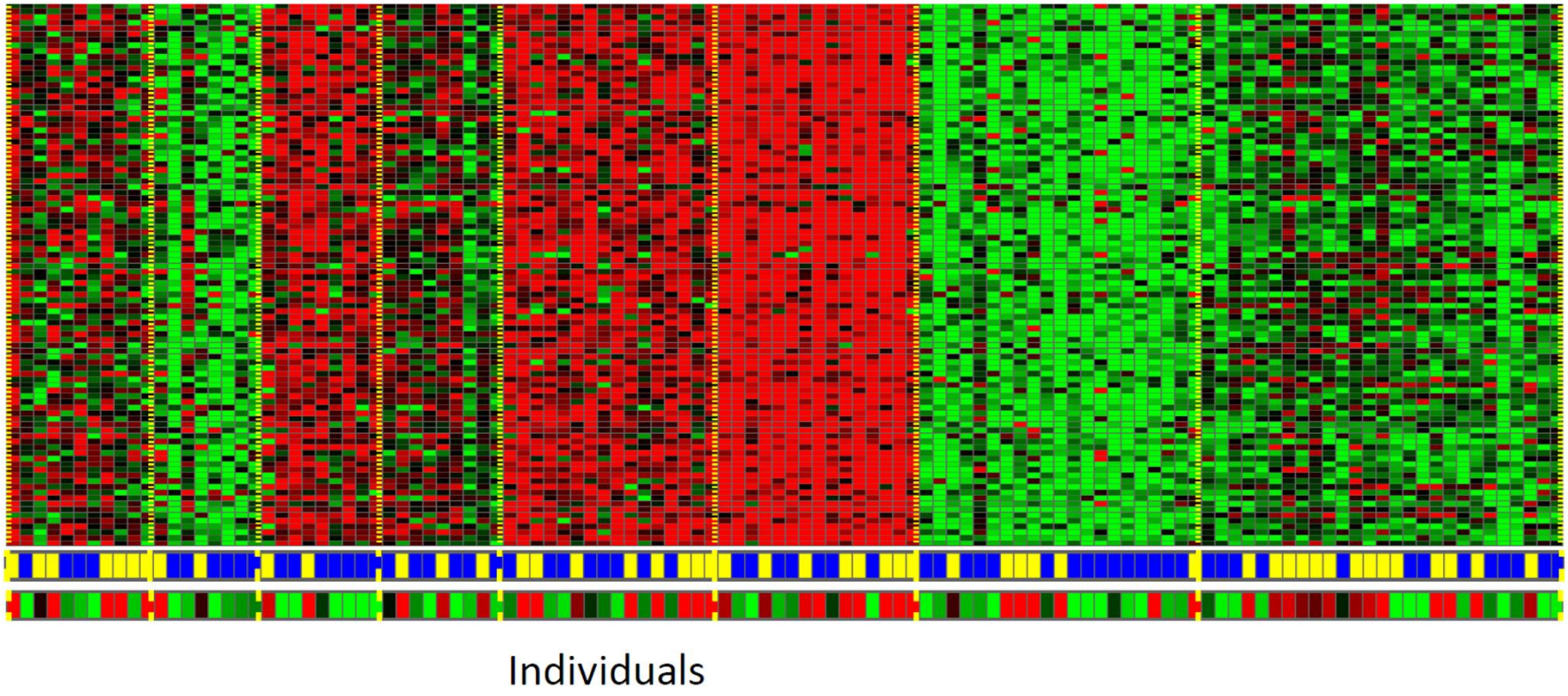
Unsupervised learning

Cocktail party problem (blind source separation)



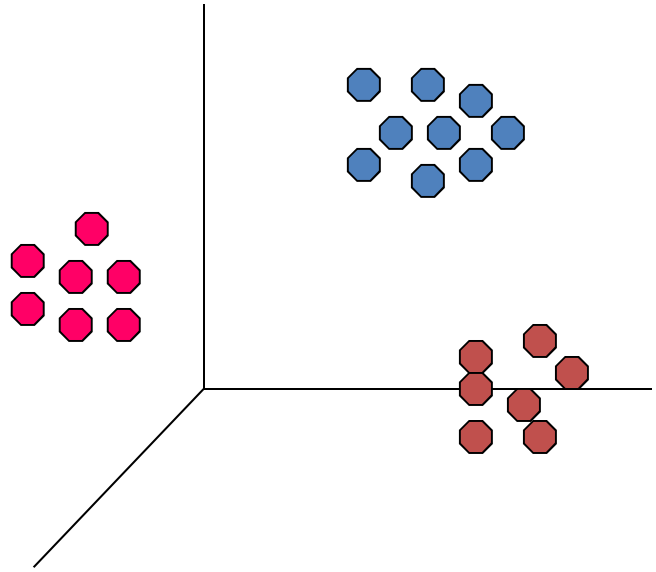
Unsupervised learning

Genes grouping

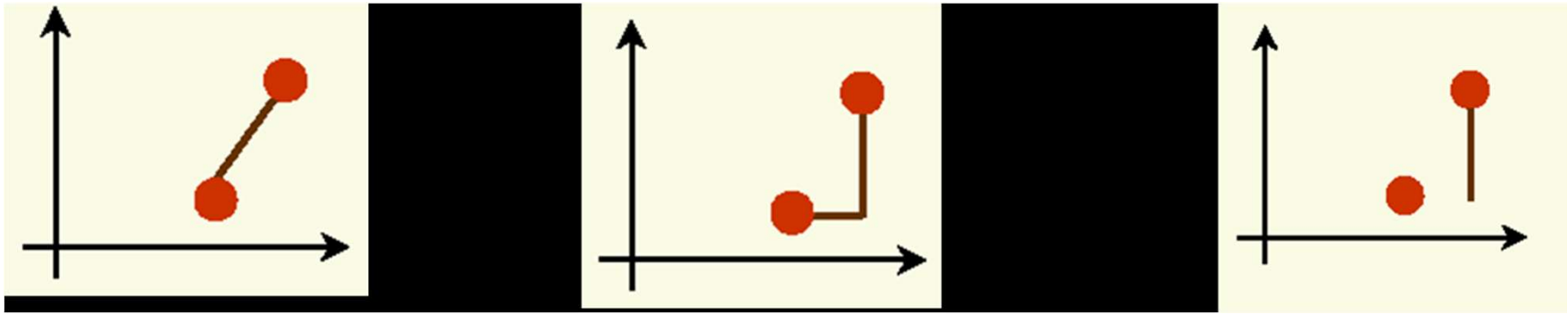


Clustering (unsupervised learning)

Given a collection of examples. Each example is a point in the multidimensional space of features. Find a similarity measure that separates the points into clusters.



Distance (similarity) measures



Euclidian Distance (L2 norm)

$$d(p,q)=\sqrt{(x_p-x_q)^2+(y_p-y_q)^2}$$

Manhattan Distance (L1 norm)

$$d(p,q)=|x_p-x_q|+|y_p-y_q|$$

Chebyshev distance

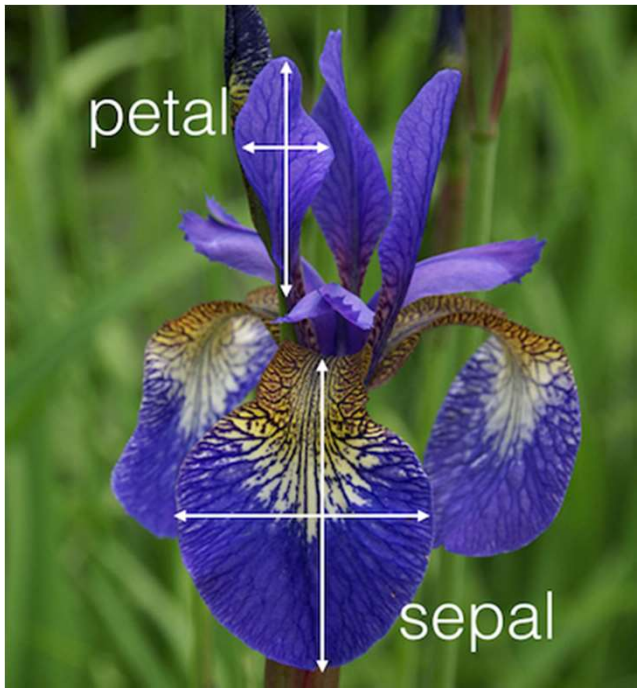
$$d(p,q)=\max(|x_p-x_q|,|y_p-y_q|)$$

Iris Plant Data Set

- Iris Plant data – benchmark dataset for illustration of ML methods.
  - UCI Machine Learning Repository
    http://www.ics.uci.edu/~mlearn/MLRepository.html

  - Three flower types (classes):
    - Setosa
    - Virginica
    - Versicolour

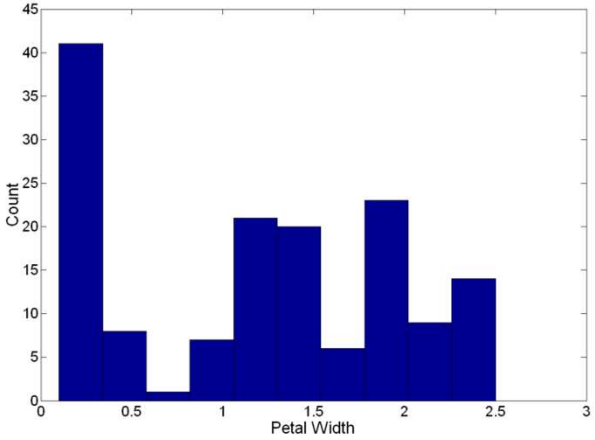
  - Four attributes
    - Sepal width and length
    - Petal width and length

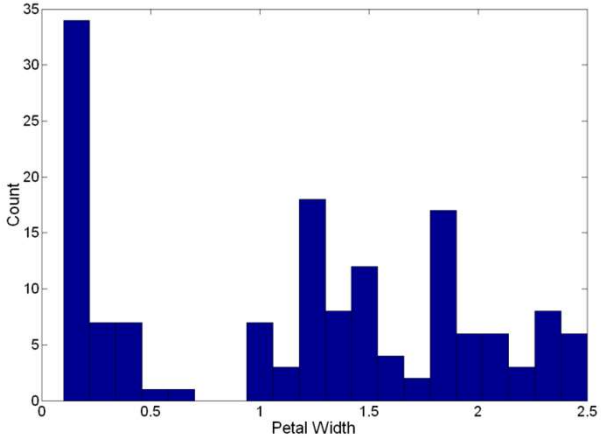


Data Visualization (1)

- **Histograms**
  - Show the distribution of values of a single feature
  - Divide the range of values of a single feature into bins and show bar plots of the number of examples in each bin.
  - Histogram shape depends on the number of bins

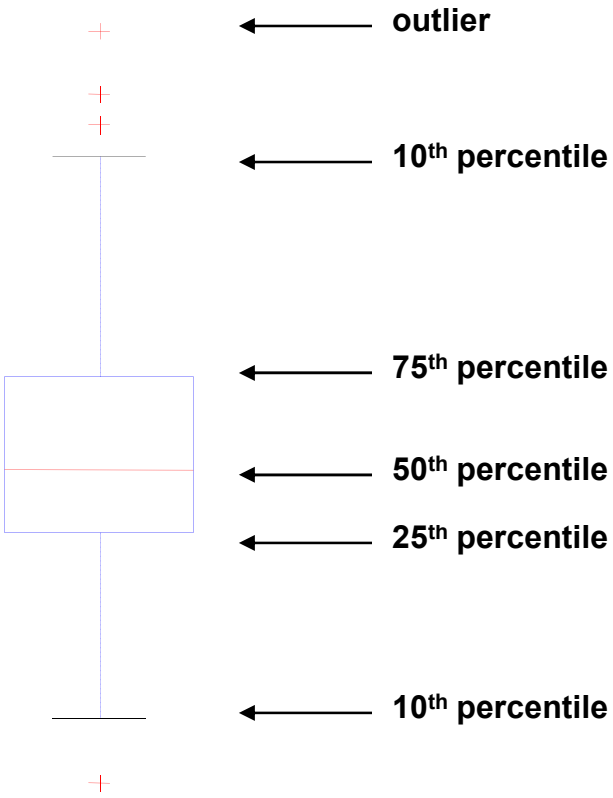
- Example: Petal Width (10 and 20 bins, respectively)





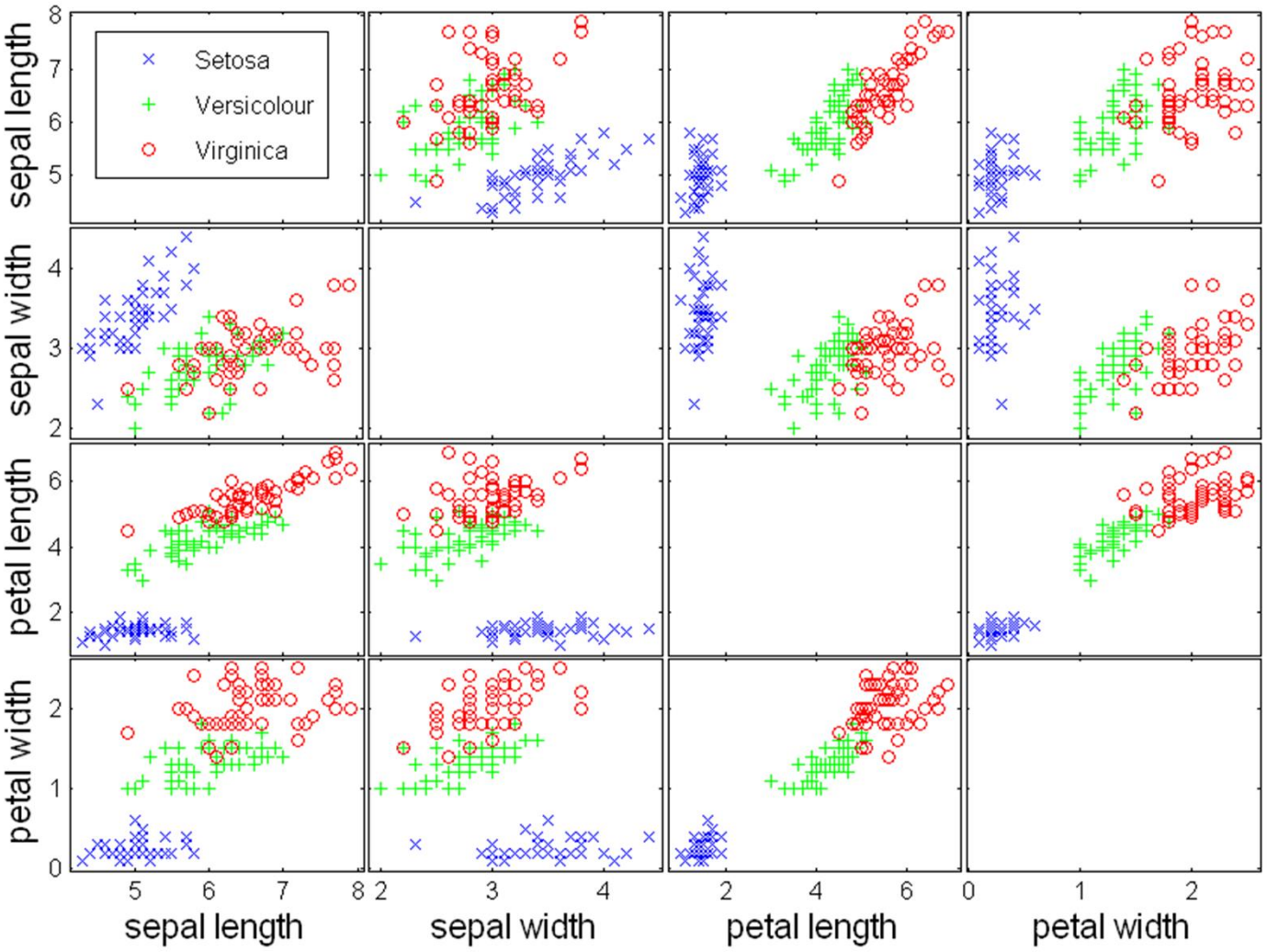
Data Visualization (2)

- **Box Plots**
  - Another way of displaying the distribution of data



Data Visualization (3)

Scatter Plot Array



RECOMMENDED BIBLIOGRAPHY

- Tom Mitchell, Machine Learning. McGraw-Hill, 1997.
- Christopher Bishop, Pattern Recognition and Machine Learning. Springer, 2006.
- David Barber, Bayesian Reasoning and Machine Learning, Cambridge University Press , 2012, (available on-line *http://web4.cs.ucl.ac.uk/staff/D.Barber/textbook/091117.pdf*).
- http://cs229.stanford.edu/
- MOOC (Massive Open Online Courses)
  e.g. https://www.coursera.org/

Food for Thought

Don't forget the ethical issues that the advances of Artificial Intelligence and Machine Learning raise !!!

How ethics and human values could be embedded into the ML algorithms used in AI ?

One hundred year study of AI, Stanford University, August, 2016
ttps://ai100.stanford.edu/

Google, Facebook, Amazon, IBM and Microsoft created Partnership on Artificial Intelligence to Benefit People and Society

https://www.partnershiponai.org/ => **Socially responsible AI**

ML as part of AI

Presence of AI in mainstream technologies:

- robot motion planning and navigation
- computer vision (e.g. object recognition)
- natural language processing and speech recognition

Recent future of AI:

- autonomous vehicles (e.g. drones, self-driving cars)
- medical diagnosis and treatment
- physical assistance for elderly

AI challenges for economy & society:

- Potential threat to humankind (?)
- AI experts have different opinions (?)
- Jobs are missing due to AI (!)
- Militarized AI is a commonly shared concern (!)

ANACONDA 3

1) Install Anaconda 3 for Python 3:

https://www.anaconda.com/distribution/

2) Learn how to use Jupyter Notebook (part of Anaconda)

https://www.dataquest.io/blog/jupyter-notebook-tutorial/