Lecture 1

Introduction

Physical Address for Prof. Ilias Tagkopoulos
Computer Science:
Office: 3063 Kemper Hall
Phone: (530) 752-4821
Fax: (530) 752-4767

Genome and Biomedical Sciences Facility:
Office: 5313 GBSF
Phone: (530) 752-7707
Fax: (530) 754-9658

Instructor: Ilias Tagkopoulos
iliast@ucdavis.edu

Course Technical Announcements

- Syllabus, lecture slides, reading materials, homework and lab assignments will be uploaded to **SMARTSITE**

- **Lectures:** TTh, 4:40-6:00 pm, 126 Wellman

- **Discussion:** 9-9:50am 166 Chemistry
  - TA-led discussion on homework material

- **Office hours:** TTh, 10:30 – 11:30pm, 3063 Kemper Hall

- **TA:** Ameen Eetemadi (eetemadi@ucdavis.edu)
  Minseung Kim (msgkim@ucdavis.edu)

- Please contact the TAs directly for **anything** that has to do with the homework.

Course Technical Announcements

Grading:
4 sets (30%), midterm (20%), project (20%), final (30%)

Homework:
Due before class according to schedule.

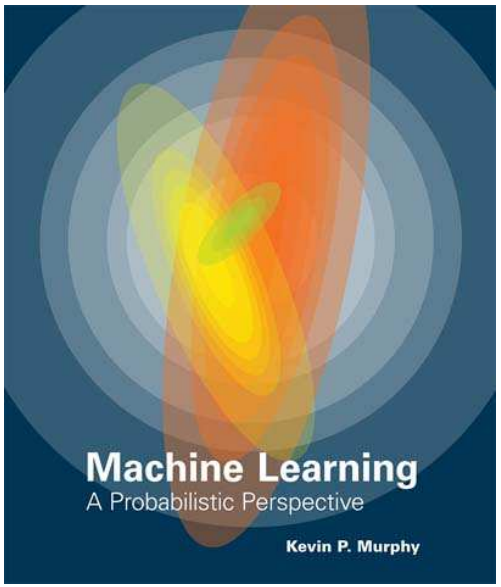
Prerequisites: ECS 60 (Programming); STA 32/ECS 132 (Probability); Mat 22A (Linear Algebra) or **equivalent**

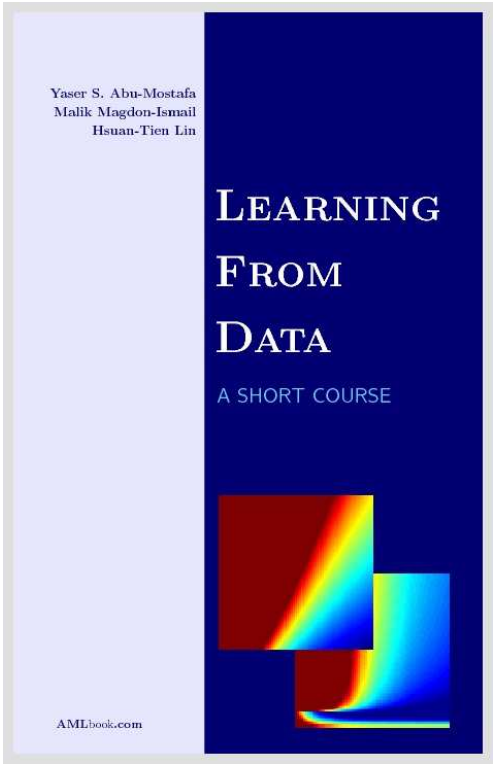
Important: If you would like to contact me, please start your email subject by "ECS 171 " to ensure proper handling as your email may be discarded otherwise (filters)

Textbook

No textbook required.

However, we will quote some sections from these books:





The Murphy book is on reserve and available (VPN, campus) here:
http://site.ebrary.com/lib/ucdavis/detail.action?docID=10597102

Other supplementary books:

- "Pattern recognition and machine learning" by Christopher Bishop

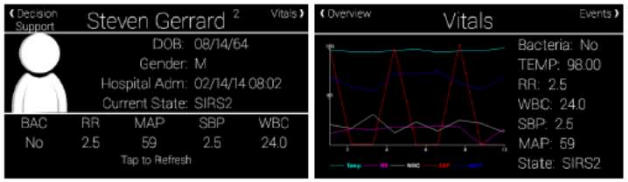
Syllabus

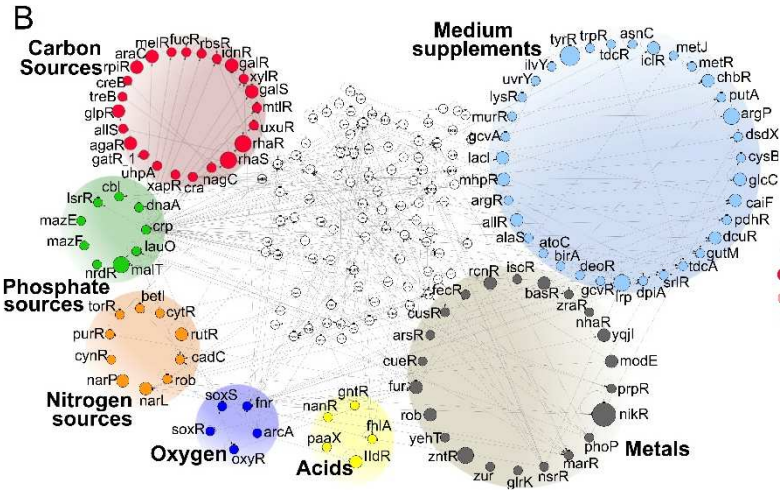
Lecture	Date	Topic	Comments
1	9/24/2015	**Introduction**	HW1 posted
2	9/29/2015	Linear Regression	
3	10/1/2015	Other Regression methods	
4	10/6/2015	**Classification**	HW1 due - HW2 posted
5	10/8/2015	Artificial Neural Networks	**Project Topics**
6	10/13/2015	Artificial Neural Networks	
7	10/15/2015	Support Vector Machines	**Projects Assigned**
8	10/20/2015	Support Vector Machines	
9	10/22/2015	Support Vector Machines	HW2 due - HW3 posted
10	10/27/2015	**Midterm**	
11	10/29/2015	Classification issues: Kernels, Overfitting, Regularization	
12	11/3/2015	Dimensionality Reduction	
13	11/5/2015	Reinforcement Learning	
14	11/10/2015	Decision support: Markov Decision Processes	
15	11/12/2015	Graphical Models - Naïve Bayes	
16	11/17/2015	**Clustering**: K-means - Hierarchical	HW3 due - HW4 posted
17	11/19/2015	Special topics: Deep Learning	
18	11/24/2015	Project Presentation I	**Project Reports Due**
19	11/26/2015	**NO CLASS** (Thanksgiving)	HW4 due
20	12/1/2015	Project Presentation II	
21	12/3/2015	Project Presentation III - Overview	

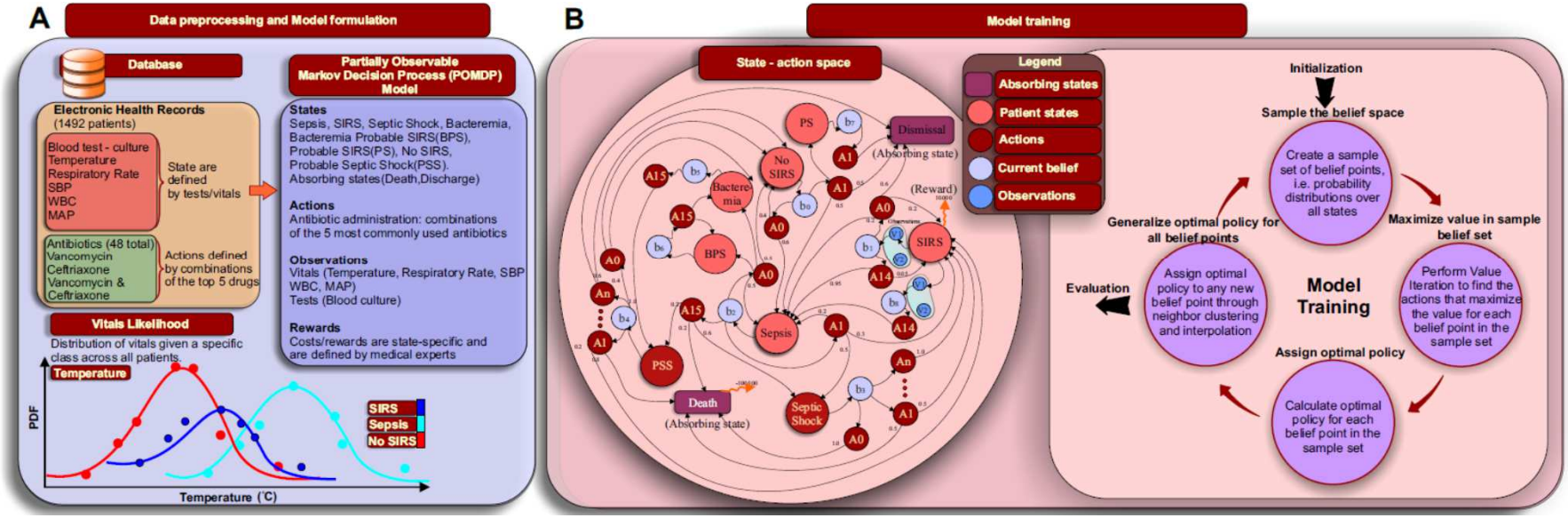
Notes:

1. The syllabus is tentative - we will adapt our schedule accordingly during the quarter
2. The TAs will lead the Wednesday discussion (no discussion session on Monday)
3. Final is on 12/10 3:30-5:30, 126 Wellman

Project?


Decision Support
Steven Gerrard 2
Vitals
DOB: 08/14/64
Gender: M
Hospital Adm: 02/14/14 08:02
Current State: SIRS2
BAC RR MAP SBP WBC
No 2.5 59 2.5 24.0
Tap to Refresh
Overview
Vitals
Events
Bacteria: No
TEMP: 98.00
RR: 2.5
WBC: 24.0
SBP: 2.5
MAP: 59
State: SIRS2


B
Carbon Sources
Medium supplements
Phosphate sources
Nitrogen sources
Oxygen
Acids
Metals


A
Data preprocessing and Model formulation
Database
Electronic Health Records
(1492 patients)
Blood test - culture
Temperature
Respiratory Rate
SBP
WBC
MAP
State are defined by tests/vitals
Antibiotics (48 total)
Vancomycin
Ceftriaxone
Vancomycin & Ceftriaxone
Actions defined by combinations of the top 5 drugs
Partially Observable Markov Decision Process (POMDP) Model
States
Sepsis, SIRS, Septic Shock, Bacteremia, Bacteremia Probable SIRS(BPS), Probable SIRS(PS), No SIRS, Probable Septic Shock(PSS). Absorbing states(Death,Discharge)
Actions
Antibiotic administration: combinations of the 5 most commonly used antibiotics
Observations
Vitals (Temperature, Respiratory Rate, SBP WBC, MAP)
Tests (Blood culture)
Rewards
Costs/rewards are state-specific and are defined by medical experts
Vitals Likelihood
Distribution of vitals given a specific class across all patients.
Temperature
PDF
Temperature (°C)
SIRS
Sepsis
No SIRS
B
Model training
State - action space
Legend
Absorbing states
Patient states
Actions
Current belief
Observations
Dismissal
(Absorbing state)
(Reward)
Death
Initialization
Sample the belief space
Create a sample set of belief points, i.e. probability distributions over all states
Maximize value in sample belief set
Perform Value Iteration to find the actions that maximize the value for each belief point in the sample set
Assign optimal policy
Calculate optimal policy for each belief point in the sample set
Evaluation
Assign optimal policy to any new belief point through neighbor clustering and interpolation
Generalize optimal policy for all belief points
Model Training

Definition: Machine Learning

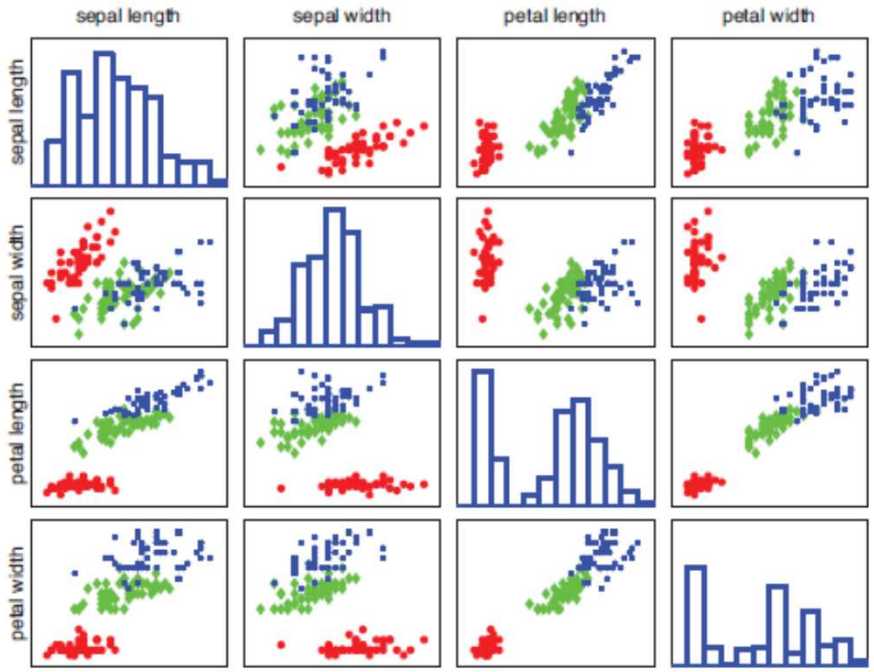
Machine learning: Algorithms and methods for building systems that learn how to perform tasks from data.

- "Field of study that gives computers the ability to learn without being explicitly programmed". A. Samuel (1959)
- **Related** to:
  - Data Mining
  - Pattern Recognition
  - Artificial Intelligence
  - Statistical Inference

Learning categories and types

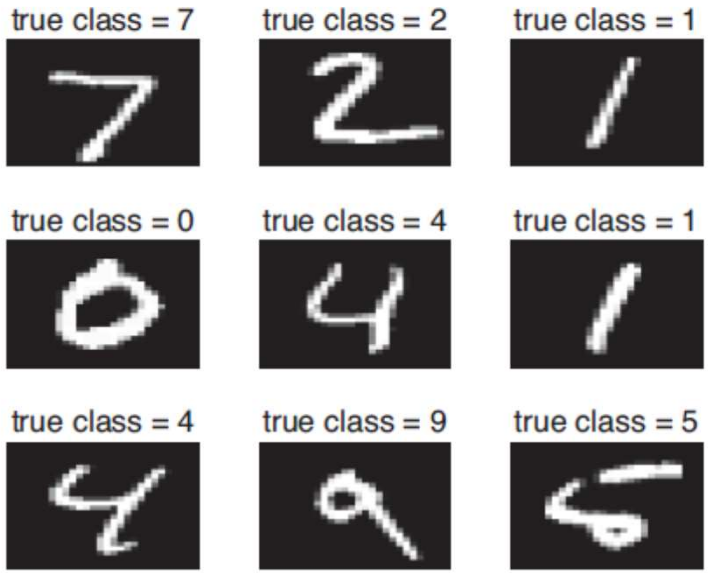
- **Supervised learning**: learn a mapping from input x to output y, given a labeled set of input-output pairs (training set $\mathcal{D} = \{(\mathbf{x}_i, y_i)\}_{i=1}^{N}$).
  - x_i can be a D-dimensional vector of **features** (also called attributes or covariates) for sample i
  - y_i is the **response variable**
    - When y_i is **categorical** (i.e. label of a class), then the task is to assign a class to a new sample and it is called **classification.**
    - When y_i is real-valued, then the task is called **regression**.
  - **Examples**: SVM, ANN, Linear Regression

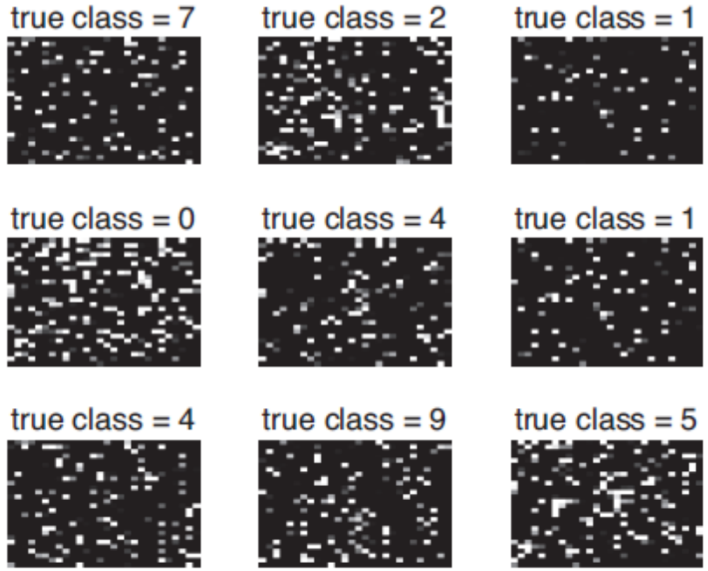
Learning categories and types: Classification



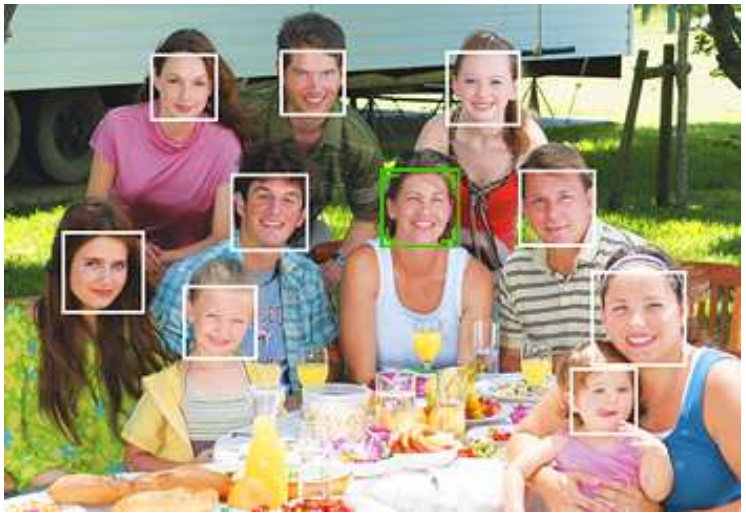
Learning categories and types: Classification

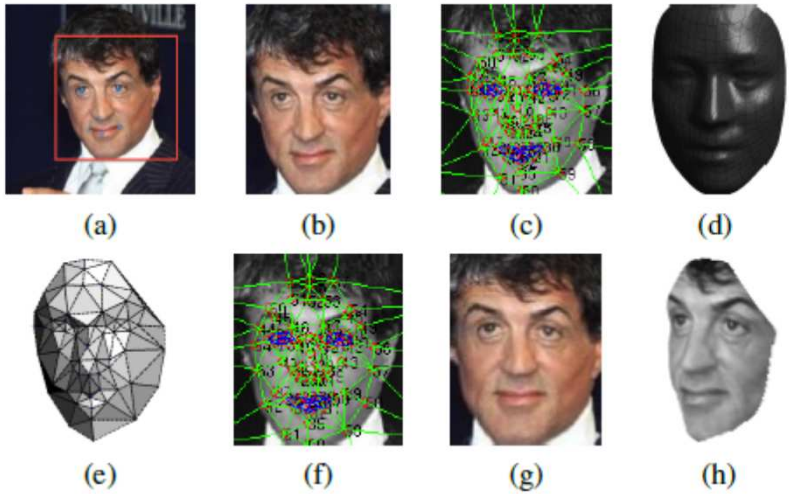
In which case will a classifier perform **best**?





Learning categories and types: Object detection - recognition





- Object detection must be invariant in subtle changes
- Object recognition needs more features and data
  - Quite accurate: DeepFace (Facebook) almost as good as humans (97.25% vs. 97.53%)

DeepFace: Closing the Gap to Human-Level Performance in Face Verification

Yaniv Taigman Ming Yang Marc'Aurelio Ranzato Lior Wolf

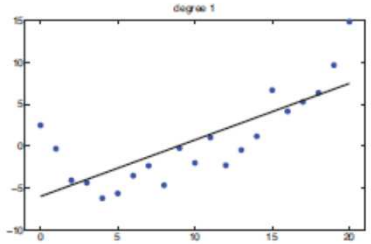
Facebook AI Research
Menlo Park, CA, USA
{yaniv, mingyang, ranzato}@fb.com

Tel Aviv University
Tel Aviv, Israel
wolf@cs.tau.ac.il

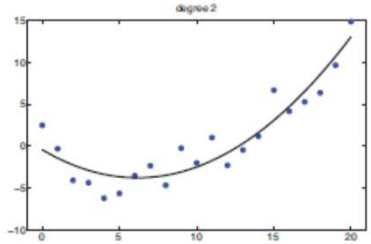
Learning categories and types: Regression

Linear Regression Trendlines and Volatility-Based Price Probability Ranges

A = (852.34) trendline of recent tops
B = (834.04) upper bound 96% probability based on 3-mo volatility
C = (820.32) upper bound 96% probability based on 1-yr volatility
D = (767) street consensus 1-year target (56 analysts)
E = (744.08) regression trendline from January 2009 low
F = (647.95) regression trendline from date iPhone introduced
G = (526.63) trendline from recent bottoms
H = (477.24) regression trendline for last 10-years
I = (407.96) lower bound 96% probability based on 1-yr volatility
J = (399.47) lower bound 96% probability based on 3-mo volatility
Mkt = 576.80



(a)

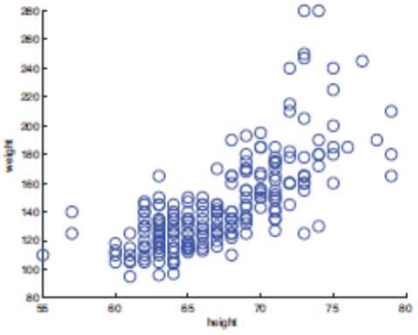


(b)

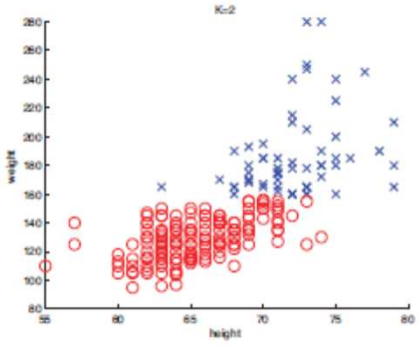
Learning types

- **Unsupervised learning**: find the **structure** or **patterns** in the data (knowledge discovery).
  - Make assumptions about the sample/feature distributions or relationships in the data
  - Examples: K-means, SOM, Hierarchical clustering
- **Reinforcement learning:** Given a certain environment, find the policy (i.e. actions) that maximizes the cumulative reward.
  - **Observability** can be full or partial
  - **States, actions, rewards, observations**
  - **Examples**: (Partially Observable) Markov Decision Processes, Q-learning

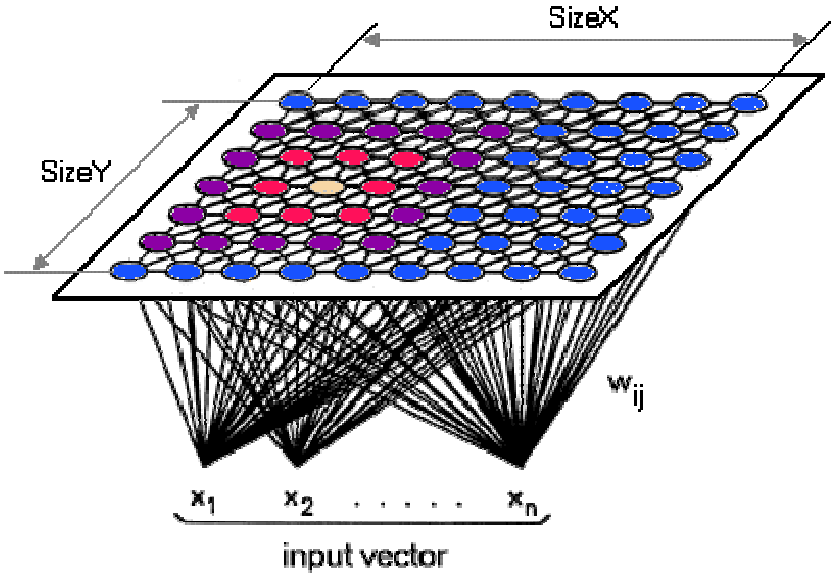
Learning categories and types: Unsupervised learning

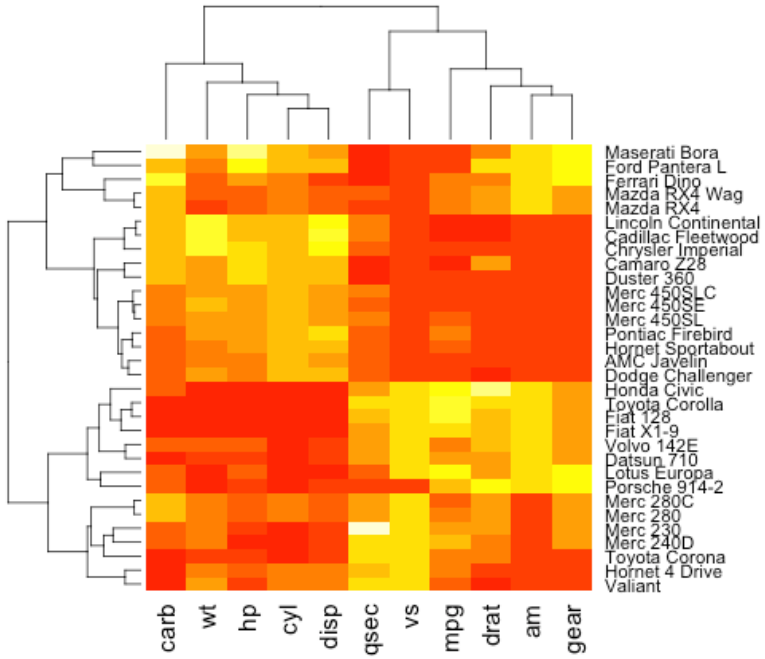


(a)

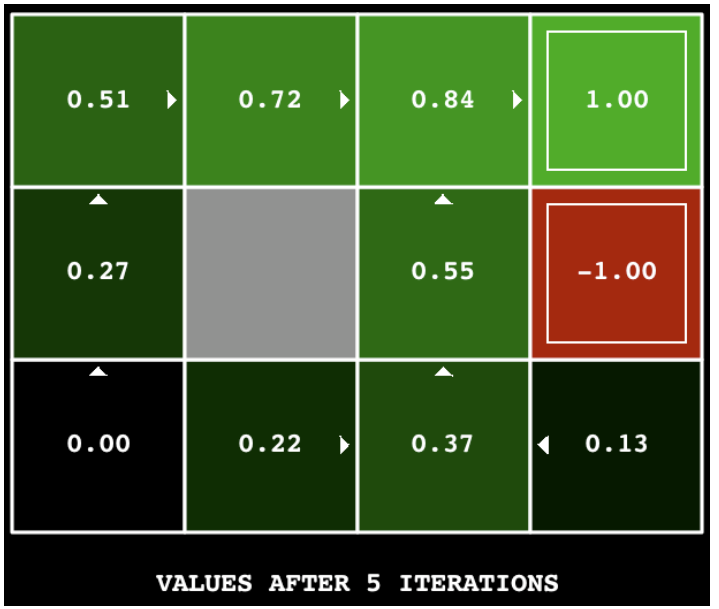


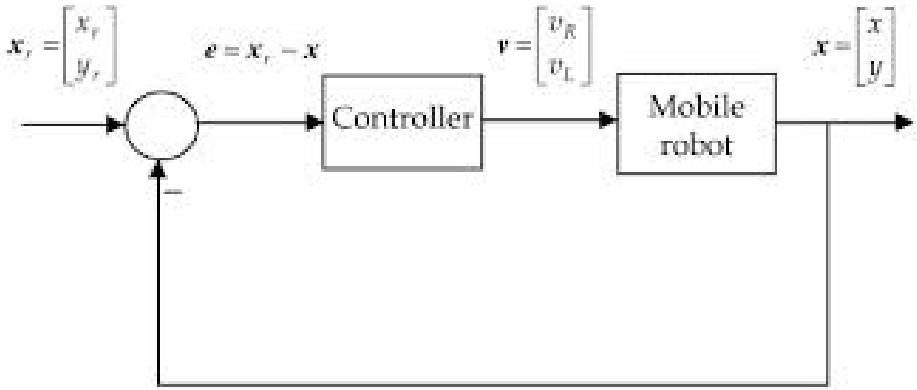
(b)

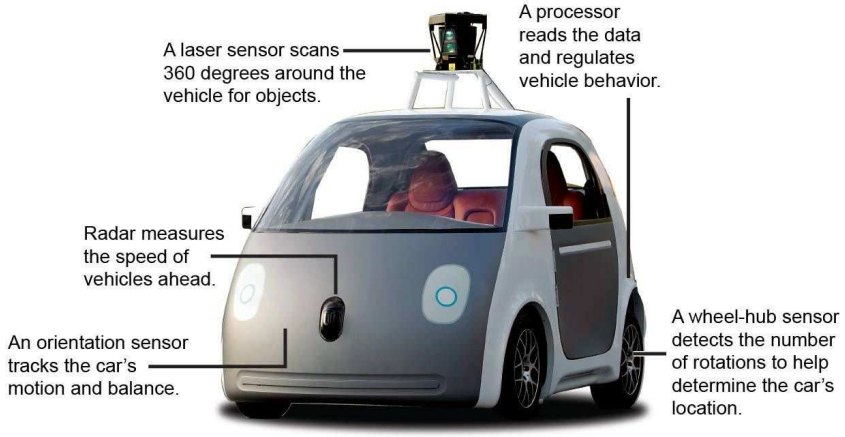




Learning categories and types: Reinforcement learning







Source: Google

Raoul Rañoa / @latimesgraphics

General overview of a machine learning workflow

Step 1. Get enough data!

Dataset

Step 2. Do all of the data samples have labels?

$$\begin{bmatrix} x_{11} & \cdots & x_{1m} \\ \vdots & \ddots & \vdots \\ x_{n1} & \cdots & x_{nm} \end{bmatrix} \begin{matrix} y_1 \\ \dots \\ y_n \end{matrix}$$

Yes

No

$$\begin{bmatrix} x_{11} & \cdots & x_{1m} \\ \vdots & \ddots & \vdots \\ x_{n1} & \cdots & x_{nm} \end{bmatrix}$$

SUPERVISED LEARNING

Step 3: The task is to predict a continuous variable, assign a new sample to a class, or perform an optimal action?

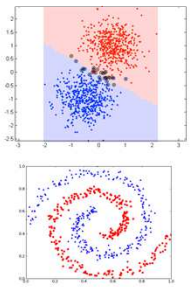
Assign to a class

Predict continuous variable

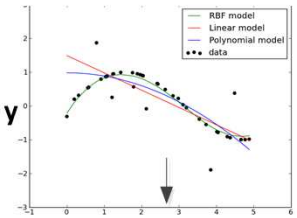
Perform optimal actions

CLASSIFICATION

- Bayesian Classification (Naïve Bayes)
- Linear Discriminant Analysis
- Artificial Neural Networks
- Decision Trees
- Support Vector Machines

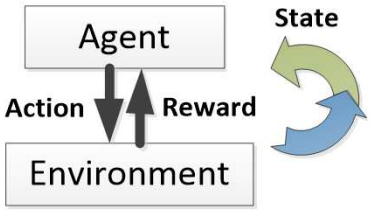


REGRESSION



Linear, polynomial, logistic, ...

REINFORCEMENT LEARNING (*)

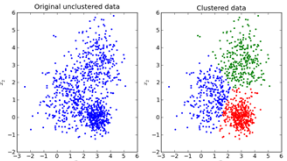


Markov Decision Process (MDP), POMDP, Q-learning, ...

UNSUPERVISED LEARNING

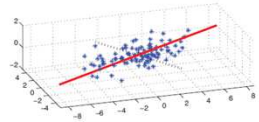
Step 3: The task is to cluster data together, find latent factors or complete missing data?

Clustering



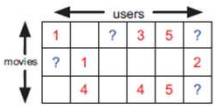
- K-means
- Hierarchical clustering
- SOM

Dimensionality Reduction



- PCA
- ICA

Missing Data



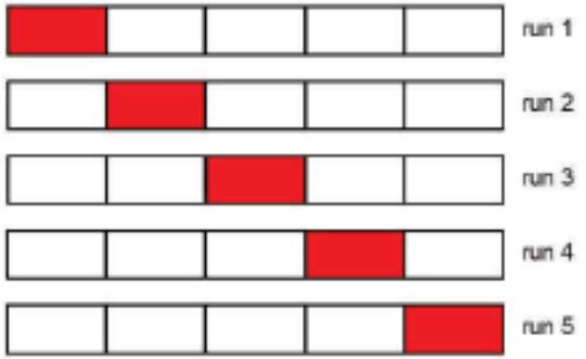
- Collaborative filtering
- Market Basket analysis
- ...

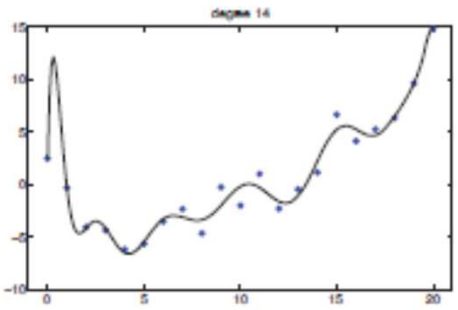
Parametric vs. Non-parametric: Models that have a fixed number of parameters are called parametric. Models where the number of parameters increases with data size are called non-parameteric.

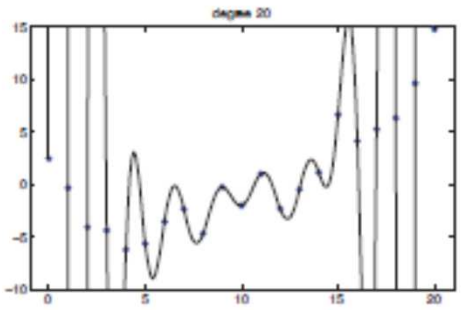
- Parametric approaches **assume** a particular shape/distribution of the data. Non-parametric have no or minimal assumptions (same in statistics: Pearson vs. Spearman's rank correlation).
- Non-parametric approaches are better when we have **no information about the underlying distribution** but need more data.

Other issues:

- Linearity
- Dataset issues: size and structure
- Independence
- Curse of Dimensionality
- Feature Selection
- Overfitting and Noise
- Bias – variance trade-off
- Evaluation







End of Lecture 1

Note: Please install and run either **Matlab** (if you have a license) or **Octave** (open source /GNU) as it will be needed for the homework assignments