

Introduction

Physical Address for Prof. Ilias Tagkopoulos

Computer Science:

Office: 3063 Kemper Hall

Phone: (530) 752-4821

Fax: (530) 752-4767

Instructor: Ilias Tagkopoulos
iliast@ucdavis.edu

Genome and Biomedical Sciences Facility:

Office: 5313 GBSF

Phone: (530) 752-7707

Fax: (530) 754-9658

■ 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.

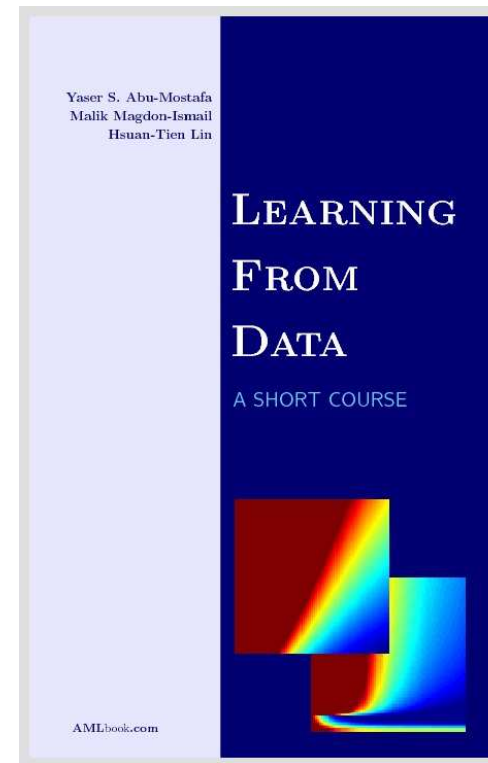
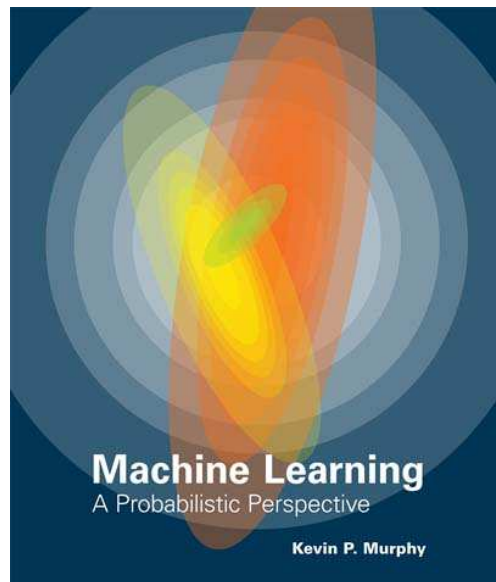
Prerequisites: ECS 60 (Programming); STA 32/ECS 132 (Probability); Mat 22A (Linear Algebra) or **equivalent**

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:

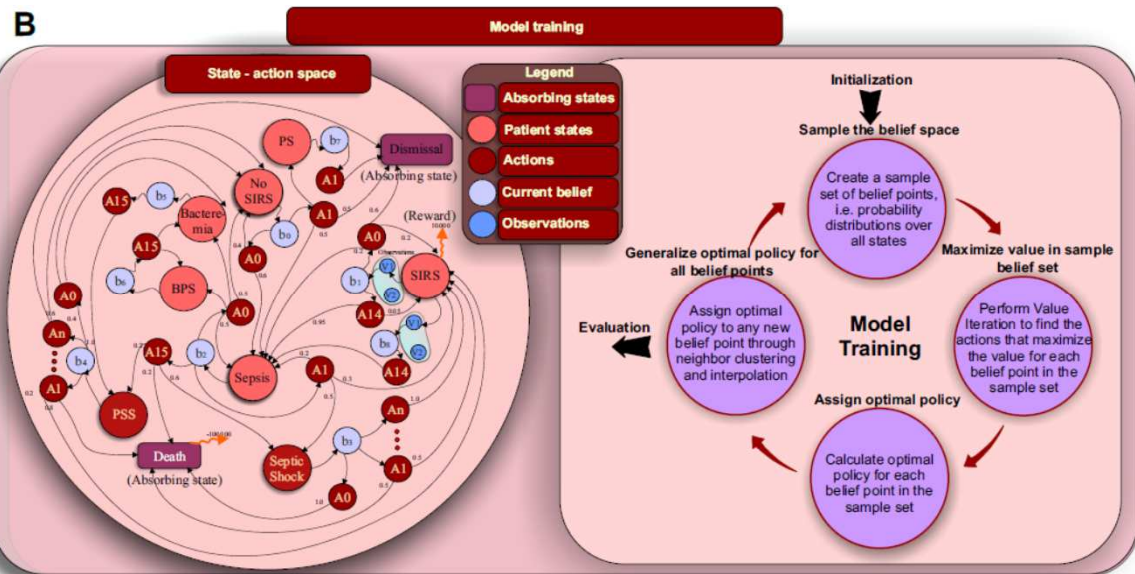
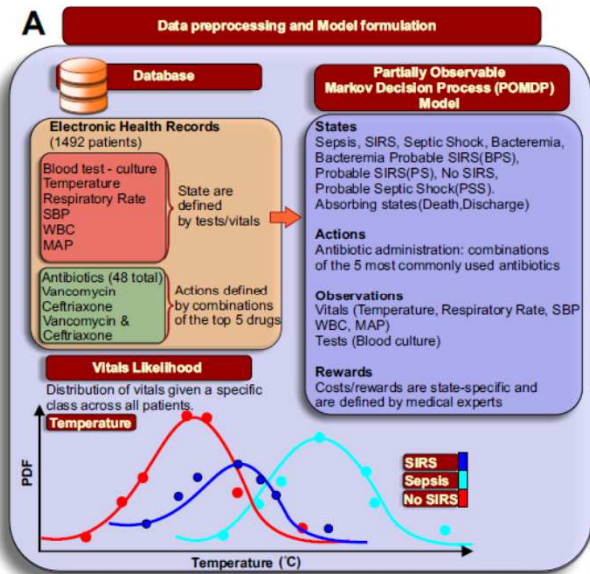
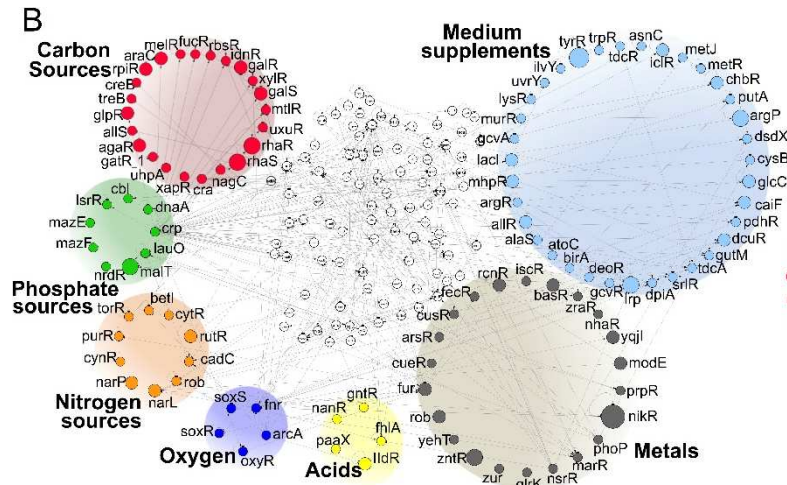
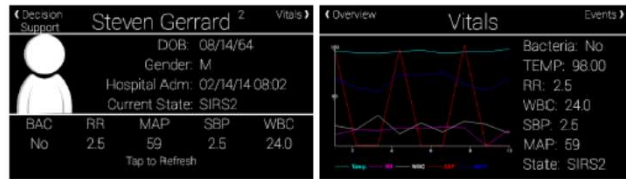
- “Pattern recognition and machine learning” by Christopher Bishop

■ 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



■ Definition: Machine Learning

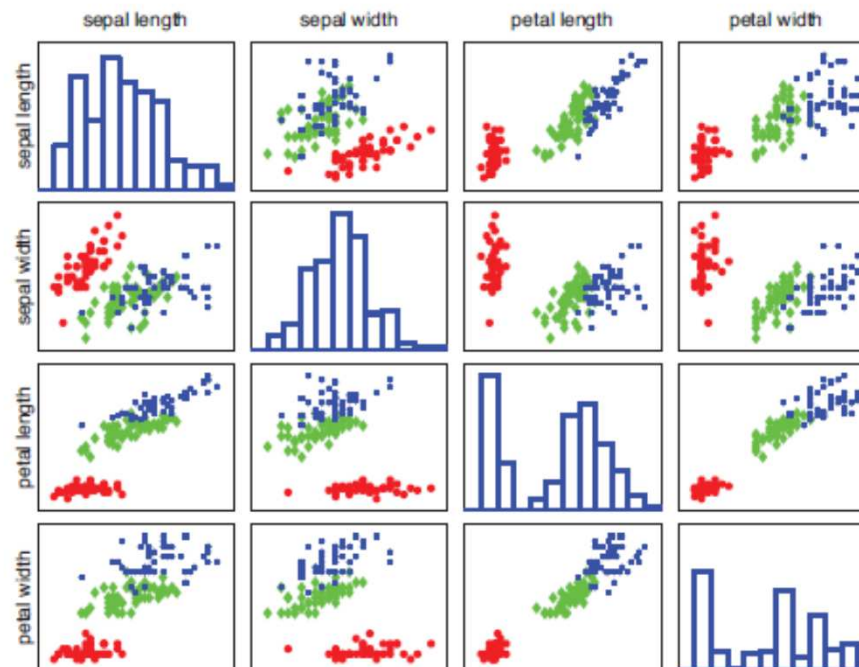
Machine learning: Algorithms and methods for building systems that learn how to perform tasks from data.

- “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

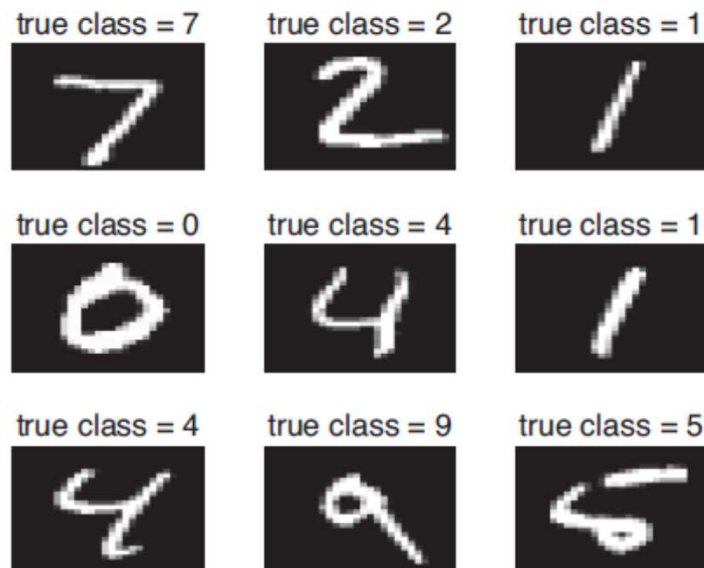
- **Supervised learning**: learn a mapping from input x to output y , given a labeled set of input-output pairs (training set $\mathcal{D} = \{(\mathbf{x}_i, y_i)\}_{i=1}^N$).
 - x_i can be a D-dimensional vector of **features** (also called attributes or covariates) for sample i
 - y_i is the **response 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

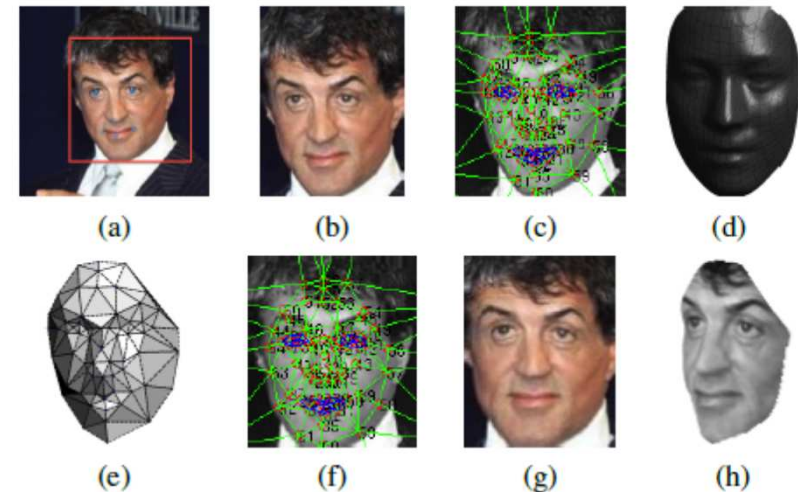
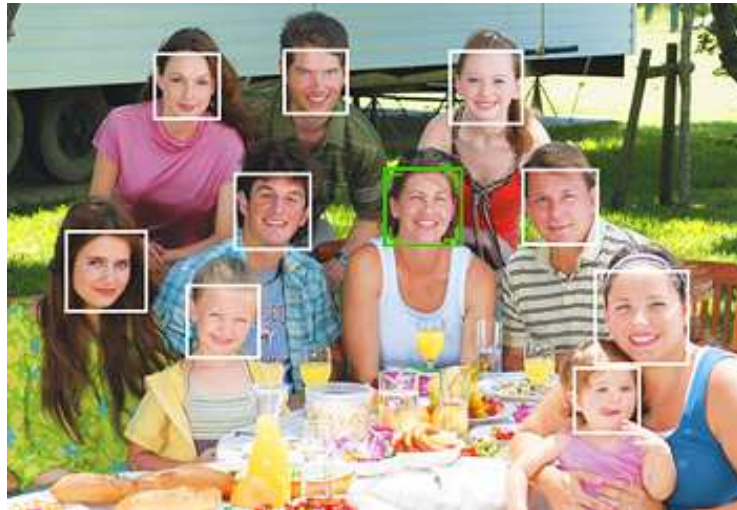


■ Learning categories and types: Classification

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

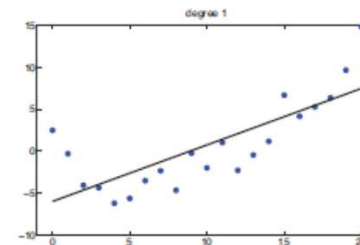
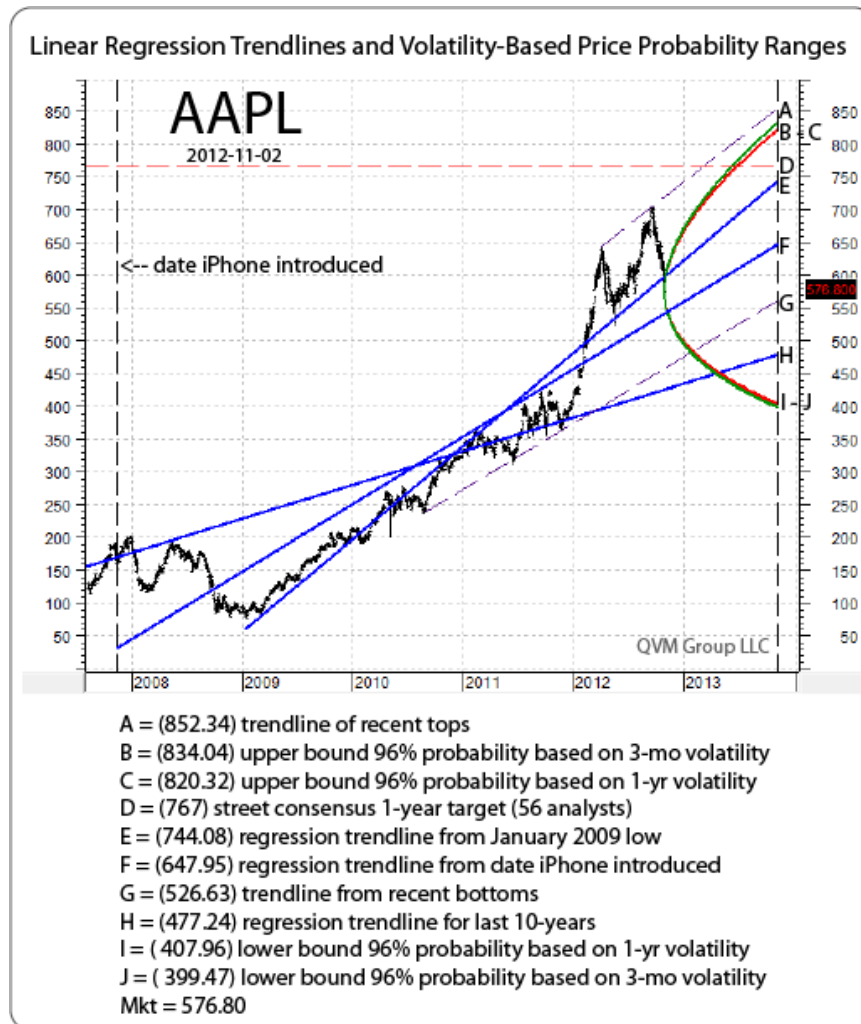
Facebook AI Research
Menlo Park, CA, USA

Tel Aviv University
Tel Aviv, Israel

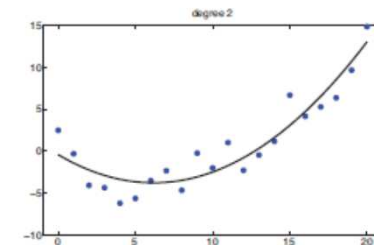
{yaniv, mingyang, ranzato}@fb.com

wolf@cs.tau.ac.il

Learning categories and types: Regression



(a)

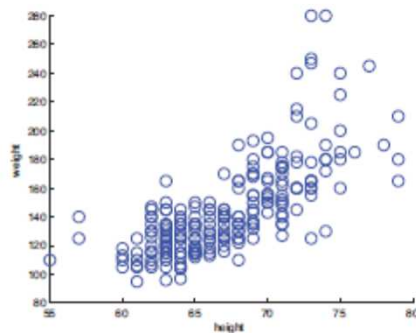


(b)

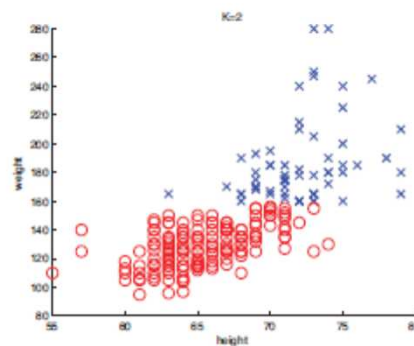
■ Learning types

- **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

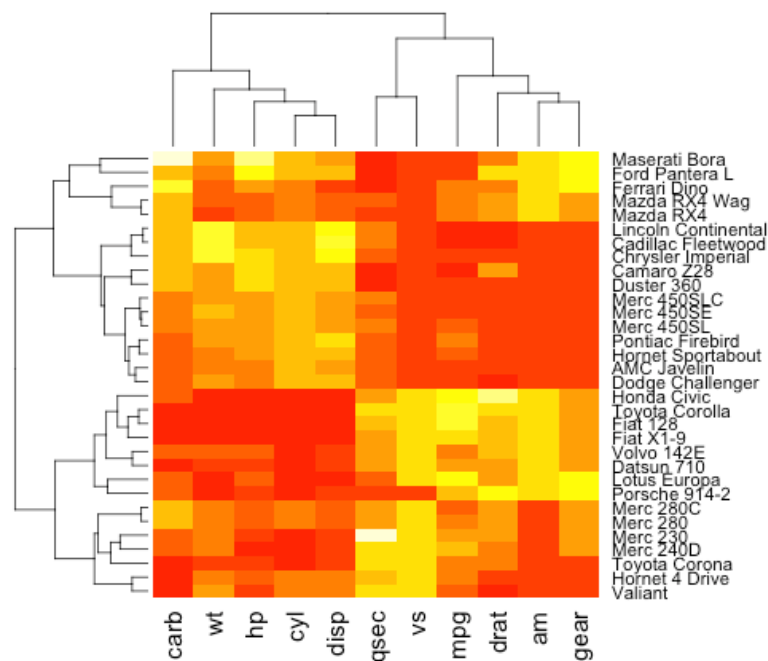
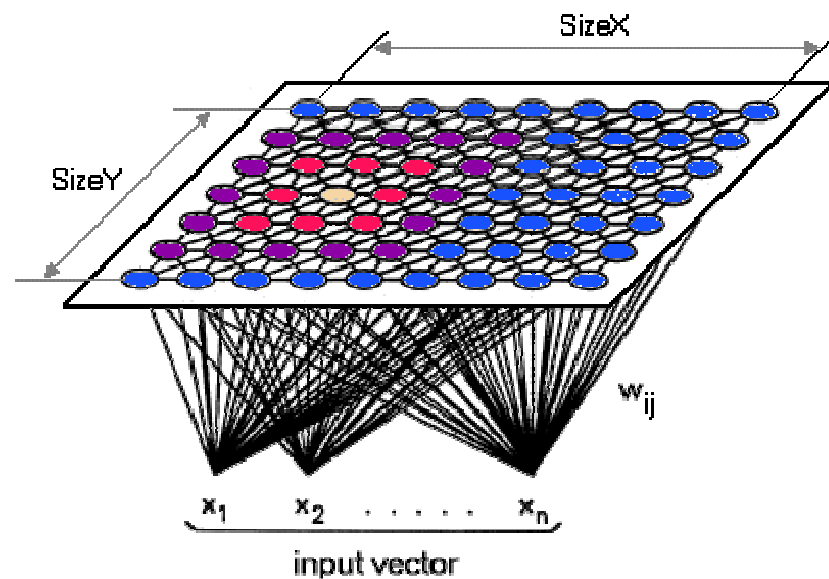
Learning categories and types: Unsupervised learning



(a)



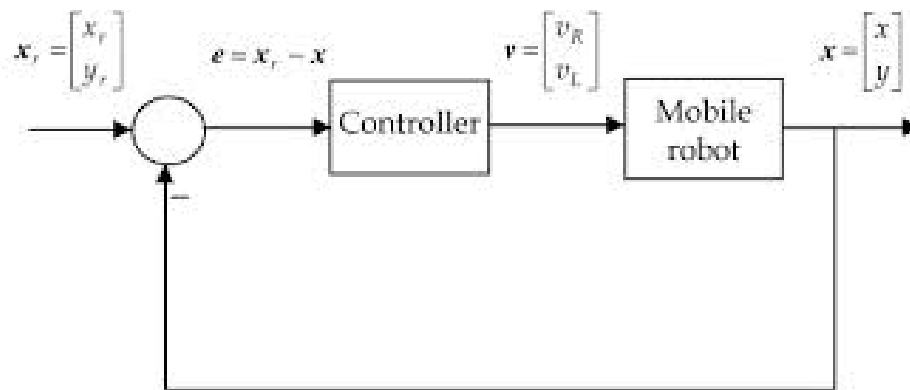
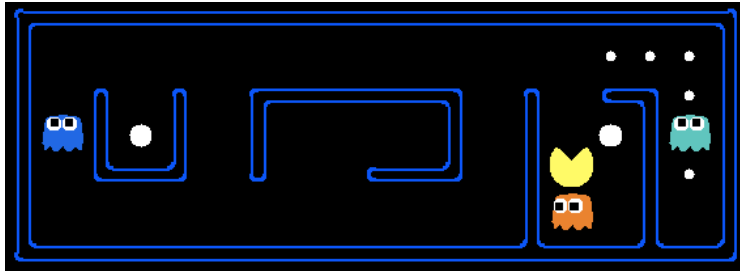
(b)



Maserati Bora
Ford Pantera L
Ferrari Dino
Mazda RX4 Wag
Mazda RX4
Lincoln Continental
Cadillac Fleetwood
Chrysler Imperial
Camaro Z28
Duster 360
Merc 450SLC
Merc 450SE
Merc 450SL
Pontiac Firebird
Hornet Sportabout
AMC Javelin
Dodge Challenger
Honda Civic
Toyota Corolla
Fiat 128
Fiat X1-9
Volvo 142E
Datsun 710
Lotus Europa
Porsche 914-2
Merc 280C
Merc 280
Merc 230
Merc 240D
Toyota Corona
Hornet 4 Drive
Valiant

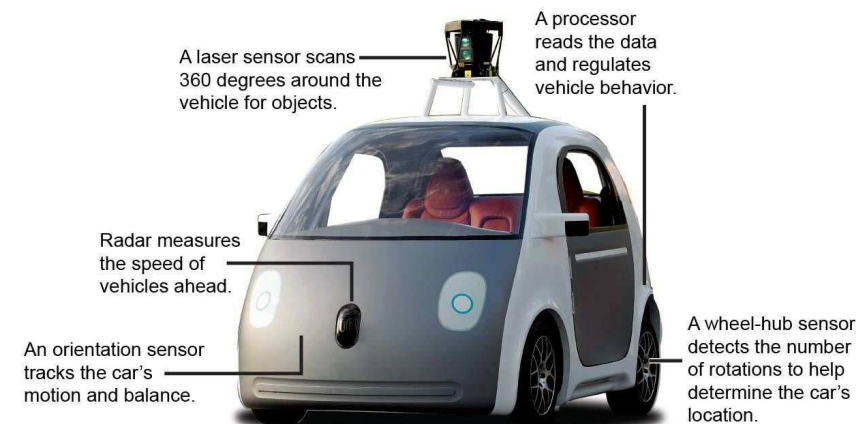
carb wt hp cyl disp qsec vs mpg drat am gear

Learning categories and types: Reinforcement learning



0.51 ▶	0.72 ▶	0.84 ▶	1.00
0.27 ▲		0.55 ▲	-1.00
0.00 ▲	0.22 ▶	0.37 ▲	0.13 ◀

VALUES AFTER 5 ITERATIONS



Source: Google

Raoul Raïoa / @latimesgraphics

General overview of a machine learning workflow

Step 1. Get enough data!



Dataset

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

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

Yes

No

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

SUPERVISED LEARNING

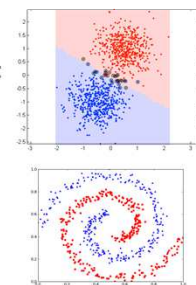
Step 3: The task is to predict a continuous variable, assign a new sample to a class, or perform an optimal action?

Assign to a class

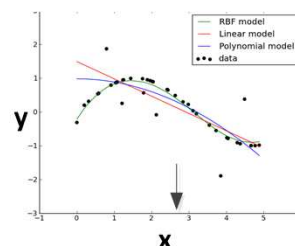
Predict continuous variable

Perform optimal actions

CLASSIFICATION



REGRESSION



Linear, polynomial, logistic, ...

REINFORCEMENT LEARNING (*)

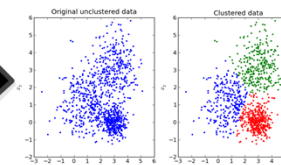


Markov Decision Process (MDP), POMDP, Q-learning, ...

UNSUPERVISED LEARNING

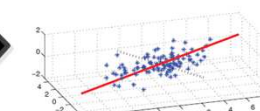
Step 3: The task is to cluster data together, find latent factors or complete missing data?

Clustering



- K-means
- Hierarchical clustering
- SOM

Dimensionality Reduction



- PCA
- ICA

Missing Data

	users				
movies	1	?	3	5	?
	?	1			2
			4	4	5
					?

- Collaborative filtering
- Market Basket analysis
- ...

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

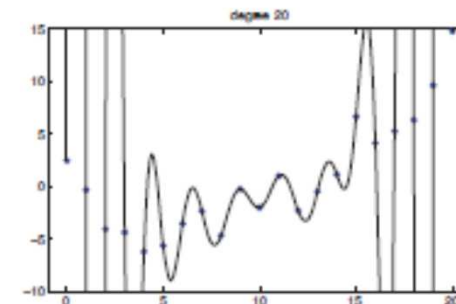
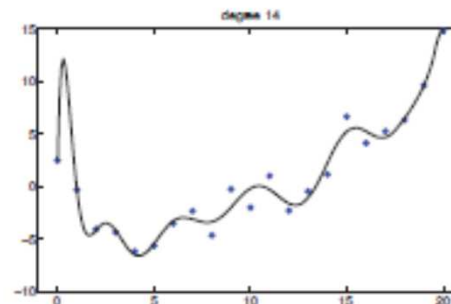
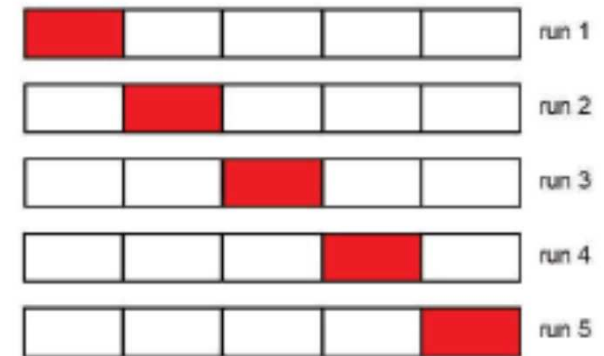
Parametric vs. Non-parametric: Models that have a fixed number of parameters are called parametric. Models where the number of parameters increases with data size are called non-parameteric.

- 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.

■ Basic concepts

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