



# **EECS708: Machine Learning**

## **Introduction to Machine Learning Module**

Dr. Ioannis Patras  
School of EECS

## **Outline**

### [Course Logistics](#)

Examples of Machine Learning problems

A rough categorization of Machine Learning problems

What will cover in this module

## Course Format

- 12 weeks, 2 (+1 occasionally) hours of lectures per week
- 2 coursework assignments organised in 6 lab sessions of 2 hrs each as follows:
  - 2hrs lab sessions for part 1 of assignment X
  - 2hrs lab sessions for part 2 of assignment X
- Assessment: 60% final exam, 40% coursework

## Feedback

- Feedback on assessments:
  - Written, collective feedback with the common mistakes
  - Online lecture session with Tas
- Forum Q/A
- Questions during the interactive lecture session
  - Questions to be posed on the chat channel
- Feedback to me (lecture/labs/demonstrators/+++)
  - By email, msteams appointment

## Recommended texts

- Kevin Murphy. Machine Learning. A probabilistic perspective. MIT Press. <http://www.cs.ubc.ca/~murphyk/MLbook/>  
[Main book for the module]
- D. Barber: Bayesian Reasoning and Machine Learning  
[Comprehensive, a bit advanced, free online pdf]
- Duda, Hart and Stork: Pattern Classification (2001)  
[Good, comprehensive]
- Bishop: Pattern Recognition and Machine Learning (2006)  
[Good and reasonably affordable.]
- Barber (2002). [Thorough coverage of Probability, RVs, etc.]

## Machine Learning

- Use of “intelligent” techniques for analysis and processing of signals and data.

Applications include:

- Autonomous vehicles
- Speech recognition
- Medical diagnosis
- Search engines (google search, google image search)
- News grouping, Ad placement
- Netflix recommendations
- Spam Filtering
- Stock market prediction

## A Few Quotes

- “A breakthrough in machine learning would be worth ten Microsofts” (Bill Gates, Chairman, Microsoft)
- “Machine learning is the next Internet” (Tony Tether, former director, DARPA)
- “Machine learning is the hot new thing” (John Hennessy, President, Stanford)
- “Machine learning is going to result in a real revolution” (Greg Papadopoulos, former CTO, Sun)

## Outline

Course Logistics

[Examples of Machine Learning problems](#)

A rough categorization of Machine Learning problems

What will cover in this module

# Machine Learning by examples

## Classification

## Spam Filtering

Osman Khan to Carlos [show details](#) Jan 7 (6 days ago) [Reply](#)

sounds good  
+ok

Carlos Guestrin wrote:  
Let's try to chat on Friday a little to coordinate and more on Sunday in person?  
Carlos

**Welcome to New Media Installation: Art that Learns**

Hi everyone,

Welcome to New Media Installation: Art that Learns

The class will start tomorrow.  
\*\*Make sure you attend the first class, even if you are on the Wait List.\*\*  
The classes are held in Doherty Hall C316, and will be Tue, Thu 01:30-4:20 PM.

By now, you should be subscribed to our course mailing list: [10615-announce@cs.cmu.edu](mailto:10615-announce@cs.cmu.edu).  
You can contact the instructors by emailing: [10615-instructors@cs.cmu.edu](mailto:10615-instructors@cs.cmu.edu).

**Natural\_LoseWeight SuperFood Endorsed by Oprah Winfrey, Free Trial 1 bottle, pay only \$5.95 for shipping mfw rik** [open link](#)

Jaquelyn Halley to rhenlein, bcc: thehorney, bcc: ang [show details](#) 9:52 PM (1 hour ago) [Reply](#)

=== Natural WeightLOSS Solution ===

Vital Acai is a natural WeightLOSS product that Enables people to lose weight and cleansing their bodies faster than most other products on the market.

Here are some of the benefits of Vital Acai that You might not be aware of. These benefits have helped people who have been using Vital Acai daily to Achieve goals and reach new heights in there dieting that they never thought they could.

- Rapid WeightLOSS
- Increased metabolism - BurnFat & calories easily!
- Better Mood and Attitude
- More Self Confidence
- Cleanse and Detoxify Your Body
- Much More Energy
- BetterSexLife
- A Natural Colon Cleanse

➡ Spam/no spam

## Face detection



Example training images  
for each orientation



Machine Learning by examples

# Regression

## Stock market prediction



## Facial landmark localisation



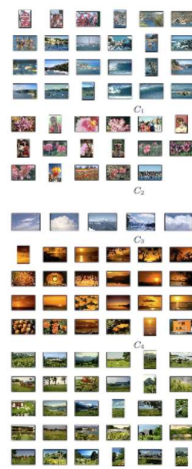
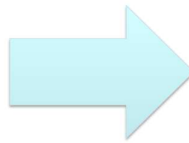
<https://www.youtube.com/watch?v=ONnobin5GBs>

Machine Learning by examples

# Clustering

## Clustering Images

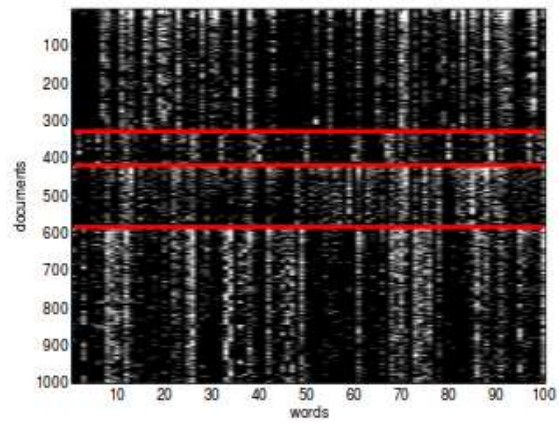
Clustering images



[Goldberger et al.]



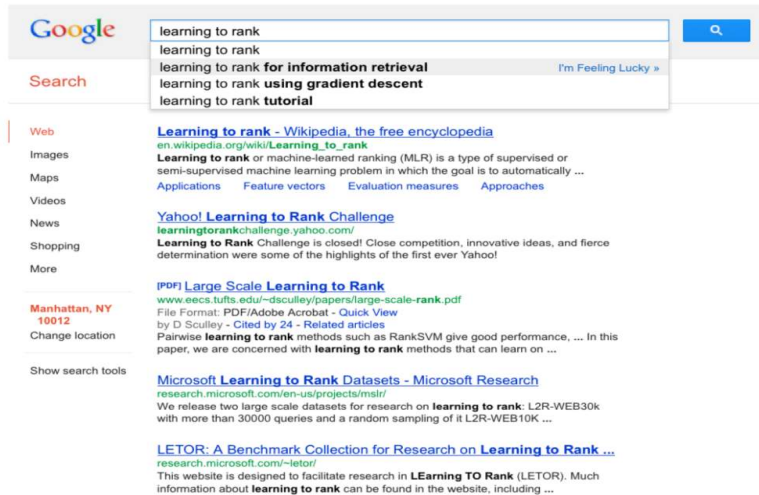
## Clustering documents



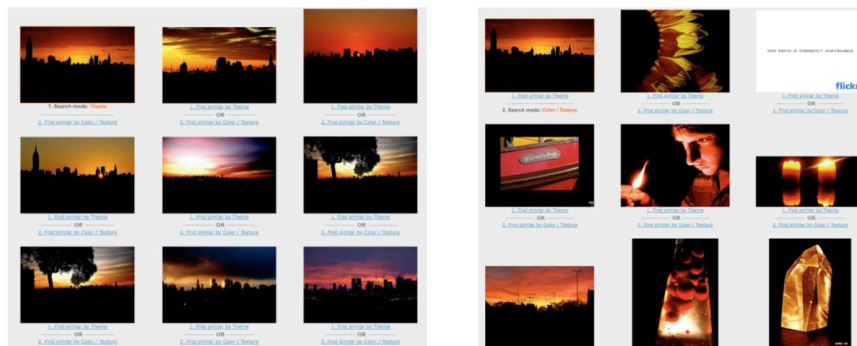
## Machine Learning by examples

# Ranking

# Text-based search



# Visual search – find similar images



<http://www.tiltomo.com/>

# Machine Learning by examples

## Recommendation

## Recommendation

The screenshot displays the Amazon website's recommendation engine. At the top, the navigation bar includes the Amazon logo, user account links, and a search bar. Below the navigation bar, the page is titled "Recommended for You" and lists four book recommendations. Each recommendation includes the book's title, author, average customer review, and price. The books are:

- Causality: Models, Reasoning and Inference** by Judea Pearl (September 14, 2009). Average Customer Review: 4.6 out of 5 (120). List Price: \$40.00, Price: \$32.49. 61 used & only from \$28.00.
- The Lady Tasting Tea: How Statistics Revolutionized Science in the Twentieth Century** by David Salsburg (May 1, 2002). Average Customer Review: 4.6 out of 5 (120). List Price: \$18.99, Price: \$13.88. 61 used & only from \$9.00.
- The Eighth Day of Creation: Makers of the Revolution in Biology, 25th Anniversary Edition** by Horace Freeland Judson (November 1, 1996). Average Customer Review: 4.6 out of 5 (120). List Price: \$40.00, Price: \$38.00. 39 used & only from \$28.00.
- The Machinery of Life** by David S. Goodell (April 28, 2009). Average Customer Review: 4.6 out of 5 (161). List Price: \$35.00, Price: \$17.49. 52 used & only from \$12.00.

Each book entry includes a small thumbnail image of the book cover and buttons for "Add to Cart" and "Add to Wish List".

## Recommendation systems

	1		?	3	5	?
	?	1				2
		4		4	5	?

Netflix competition ([www.netflixprize.com](http://www.netflixprize.com))

Machine learning competition with \$1m prize

## Outline

Course Logistics

Examples of Machine Learning problems

[A rough categorization of Machine Learning problems](#)

What will cover in this module

## Types of Learning Problems

- Supervised Learning  
Regression, Classification
- Unsupervised Learning  
Clustering
- Reinforcement Learning  
Policy learning

## Supervised Learning: learn a prediction function

Learning a function  $f$  when the target is known for the training data.

Given  $\{(x_i, y_i)\}_{i=1}^N$ ,  $x_i \in X, y_i \in Y$

Learn  $f: X \rightarrow Y$

House price prediction [R]

Stock market prediction [R]

Categorizing (Classification).

## Supervised Learning (regression)

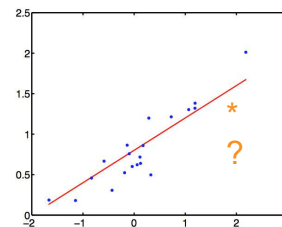
Learn to predict the price of a house (target  $y$ ) given the size of the house (features  $x$ )

Assume an unknown continuous function  $y=f(x)$

Given examples  $(x_i, y_i)$ , which may be noisy

Learn  $f(x)$ , to enable prediction of  $y^*$  given new point  $x^*$ .

It should generalise well to new  $x^*$



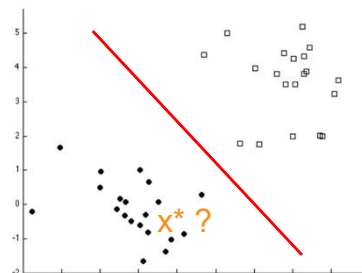
## Supervised Learning (classification)

Learn to predict the whether a tumour of a given size and “sharpness” (features  $x$ ) is malignant or not (target  $y$ , squares/circles)

Assume an partition of the feature space:  $y=f(x)$

Given examples  $(x_i, y_i)$ , which may be noisy

Learn  $f(x)$ , to enable prediction of  $y^*$  given new point  $x^*$ . It should generalise well to new  $x^*$



# Unsupervised Learning

**Categorise** these LEGO bricks into groups

Can you write an algorithm to explain how you grouped them?

No explicit target

was given.

Reduce dimensions

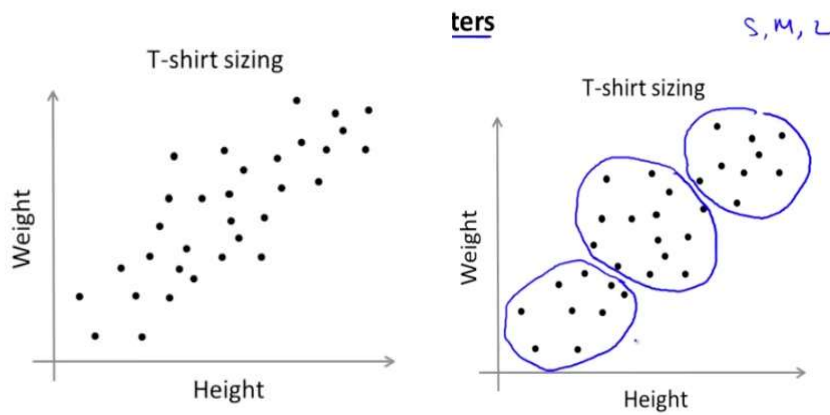
Compress data

Visual Hierarchies



## Unsupervised Learning (clustering)

Given the height and circumference (features) of the general population, group them into three (3) clusters so as to design S, M, L size T-shirts.



## Reinforcement Learning



- LEGO example. Choose 2 blocks one after the other, and I'll tell you how well you've done.
- A kind of denuded supervised learning where you just have a hotter/colder signal, not the complete right answer. Its as if I give you an exam and a mark, but never the model answers. At least there is more guidance than with unsupervised learning.

## Reinforcement Learning

Given the outcome (reward) of previous checkers games,  
learn the move (action) you should make given a  
checkerboard configuration (state).



## Other ways for Classifying Machine Learning Methods

**Parametric Methods:** Learn a low dimensional set of parameters, e.g. weights in a neural network, throwing away the training data points.

**Non-parametric:** Keeps the training data points throughout, e.g. k-nearest neighbor methods.

## Think of the whole of ML like this

**Feature Engineering/Design:** Pre-process the data based on domain specific/expert knowledge.

**Model:** Choose a machine to make the prediction using these features.

**Cost Function:** Write an equation that describes how well or badly your model is doing.

**Minimize cost function:** Write an algorithm to minimize/maximize your cost function, preferably provably so.

**Generalization:** Check that you have not overfitted or underfitted the data.

## Outline

Course Logistics

Examples of Machine Learning problems

A rough categorization of Machine Learning problems

[What will cover in this module](#)

## Techniques

- Probability and Random Variables
- Regression (Linear, non Linear, Multivariate)
- Classification I (Linear, non Linear, regularisation)
- Classification II (Decision Trees, Naïve bayes, metrics)
- Neural Networks
- Clustering (k-means, hierarchical)
- Density Estimation (parametric distributions)
- Dimensionality reduction
- Deep Learning, convolutional NN
- Ensembles

# Summary

Course Logistics

Examples of Machine Learning problems

A rough categorization of Machine Learning problems

What will cover in this module