

MACHINE LEARNING

Dr Ioannis Patras
i.patras@qmul.ac.uk

ECS708 Machine Learning

Slide no: 1

1

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: 80% final exam, 20% coursework

ECS708 Machine Learning

Slide no: 2

2

Feedback

- Feedback on assessments:
 - Written, collective feedback with the common mistakes
 - A Friday lecture session with Tas
- Forum Q/A
- Questions during the lecture (do ask!)
- Feedback to me (lecture/labs/demonstrators/+++)
 - Before/after/during lecture, email, appointment

ECS708 Machine Learning

Slide no: 3

3

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.]

ECS708 Machine Learning

Slide no: 4

4

Machine Learning

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

Applications include:

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

ECS708 Machine Learning

Slide no: 5

5

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)
- “Web rankings today are mostly a matter of machine learning” (Prabhakar Raghavan, former Dir. Research, Yahoo)
- “Machine learning is going to result in a real revolution” (Greg Papadopoulos, former CTO, Sun)

ECS708 Machine Learning

Slide no: 6

6

Machine Learning by examples

Classification

ECS708 Machine Learning

Slide no: 7

7

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.
You can contact the instructors by emailing: 10615-instructors@cs.cmu.edu.

Natural _LoseWeight SuperFood Endorsed by Oprah Winfrey, Free Trial 1 bottle, pay only \$5.95 for shipping mfw rik

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 wieght 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

ECS708 Machine Learning

Slide no: 8

8

Face detection



Example training images
for each orientation



ECS708 Machine Learning

Slide no: 9

9

Machine Learning by examples

Regression

ECS708 Machine Learning

Slide no: 10

10

Stock market prediction



ECS708 Machine Learning

Slide no: 11

11

Facial landmark localisation



<https://www.youtube.com/watch?v=ONnobi5GBs>

ECS708 Machine Learning

Slide no: 12

12

Machine Learning by examples

Clustering

ECS708 Machine Learning

Slide no: 13

13

Clustering Images

Clustering images

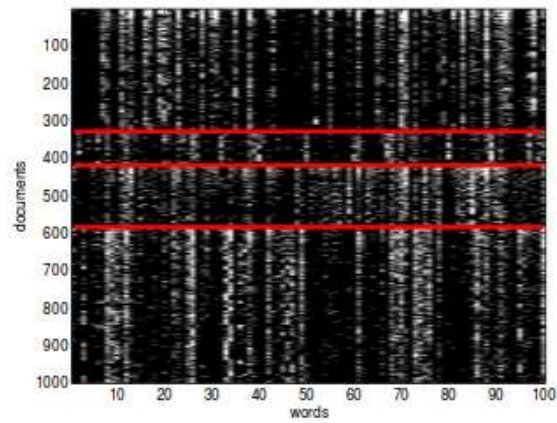


ECS708 Machine Learning

Slide no: 14

14

Clustering documents



ECS708 Machine Learning

Slide no: 15

15

Machine Learning by examples

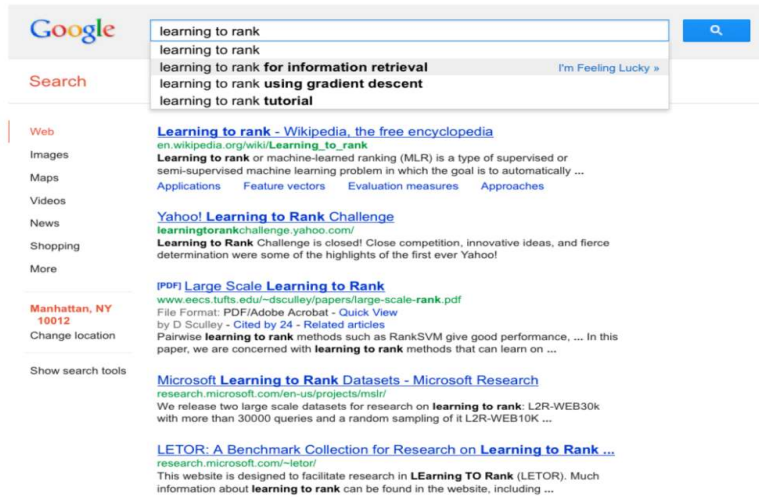
Ranking

ECS708 Machine Learning

Slide no: 16

16

Text-based search

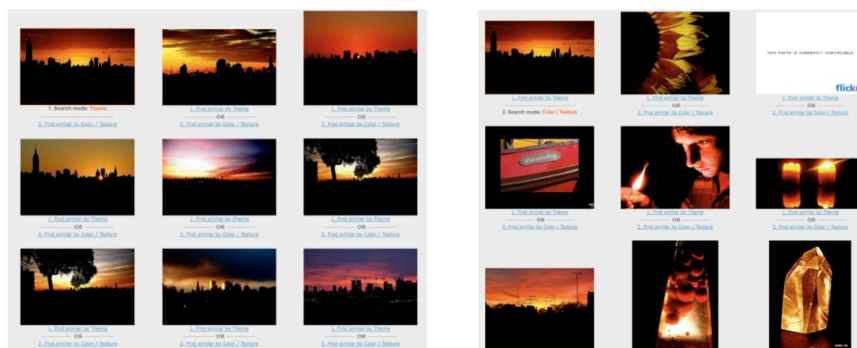


ECS708 Machine Learning

Slide no: 17

17

Visual search – find similar images



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

ECS708 Machine Learning

Slide no: 18

18

Machine Learning by examples

Recommendation

ECS708 Machine Learning

Slide no: 19

19

Recommendation



ECS708 Machine Learning

Slide no: 20

20

Recommendation systems

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

Netflix competition (www.netflixprize.com)
Machine learning competition with \$1m prize

ECS708 Machine Learning

Slide no: 21

21

Types of Learning Problems

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

ECS708 Machine Learning

Slide no: 22

22

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).

23

Supervised Learning (regression)

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

24

Supervised Learning (classification)

Learn to predict the whether a tumour of a given size
(feature) is malignant or not (target)

ECS708 Machine Learning

Slide no: 25

25

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



26

Unsupervised Learning (clustering)

Given the size and hue/colour (features) of several fruits,
group them into clusters

ECS708 Machine Learning

Slide no: 27

27

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.

28

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 Selection: 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 good or bad 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, e.g. the football predicting octopus.

31

Techniques

Techniques covered in ECS708 include:

- Probability and Random Variables
- Neural Networks
- Bayesian Inference

- Clustering
- Hidden Markov Models (HMMs)
- Principal Components Analysis, Independent Component Analysis

32

Probability and Random Variables

- Probability of an event (set of possible outcomes):

$$P(A) \geq 0$$

$$P(\Omega) = 1$$

$$P(A \cup B) = P(A) + P(B) \text{ if } A, B \text{ mutually exclusive}$$

- Random Variables $X: \Omega \rightarrow \mathbb{R}$

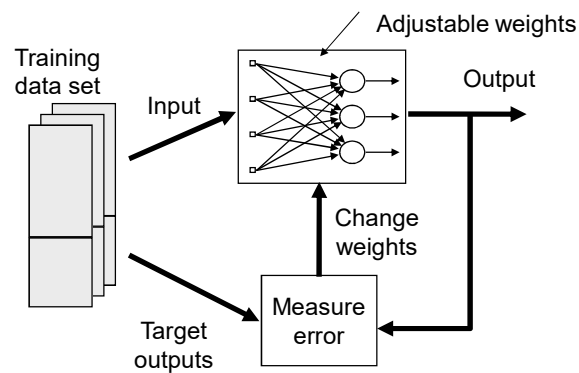
- Distribution function $F_X(x) = P(X \leq x)$

- Density function (pdf) $f_X(x) = \frac{d}{dx} F_X(x)$

(More familiar as $p(x)$)

Neural Networks

Adjust “weight” parameters in a network of simple calculating units to produce desired behaviour.



Statistical Inference

Use of probability theory to estimate the “best” answer

Uses estimate of prob. of observation x given classes ω_i

$$p(x|\omega_i)$$

1) Maximum Likelihood – Choose $\max_i p(x|\omega_i)$

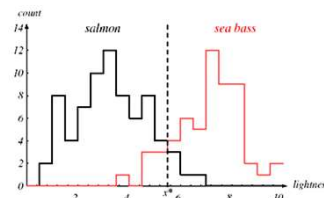
E.g. Classification of Fish:

2) Use Bayes Theorem:

$$p(\omega_i|x) = \frac{p(x|\omega_i)p(\omega_i)}{p(x)}$$

Choose $\max_i p(\omega_i|x)$

[After all, we have x and want to know ω_i]



ECS708 Machine Learning

Slide no: 35

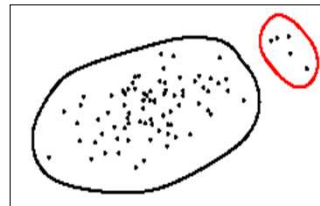
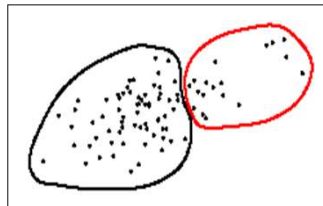
35

Clustering

- Collecting together “similar” observations or signals.
- E.g. cluster similar documents, music, bacteria,...

Issues:

- Similarity measures to use?
- Same / different sized clusters?
- How many clusters?



ECS708 Machine Learning

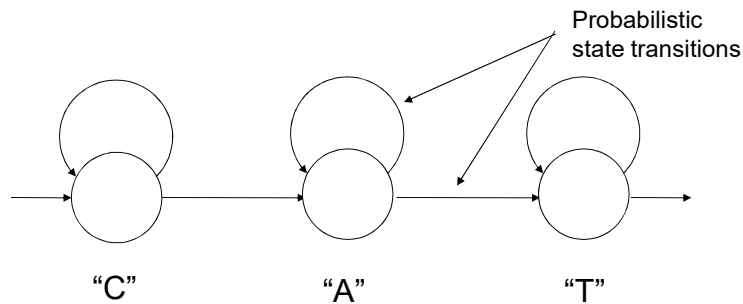
Slide no: 36

36

Hidden Markov Models (HMMs)

Hidden Markov Models (HMMs)

- Statistical model of speech utterances
- We don't "see" the states ("hidden"), only their output
- Choose the most probable word given the utterance



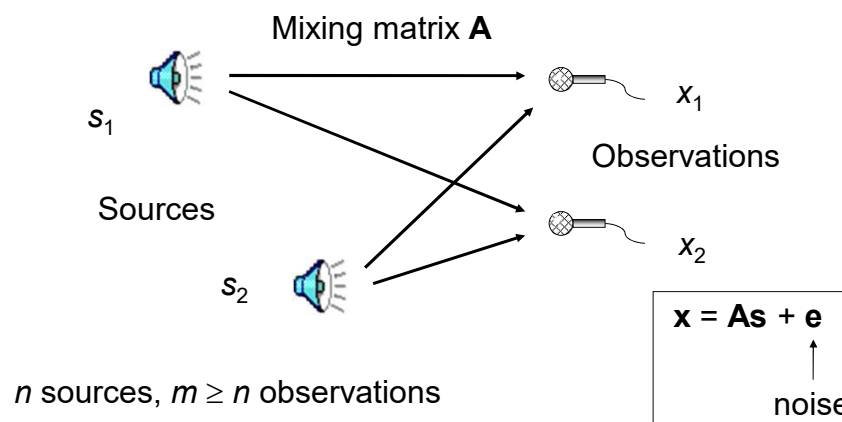
ECS708 Machine Learning

Slide no: 37

37

Independent Component Analysis

Blind Sources Separation - want to "unmix" observed signals that contain more than one source signal – the "Cocktail Party Problem"



ECS708 Machine Learning

Slide no: 38

38