Machine Learning, Machine Learning (extended)

1 – Introduction

Kashif Rajpoot

k.m.rajpoot@cs.bham.ac.uk

School of Computer Science

University of Birmingham

Outline

- What is machine learning?
- Applications
- Aims
- Learning outcomes
- Assessment
- Relevant texts
- Plagiarism
- Basics of machine learning
- What is the learning problem?
- Classes of learning
- Common terminology
- List of topics

What is machine learning?

- Algorithms that enable computers to learn from examples

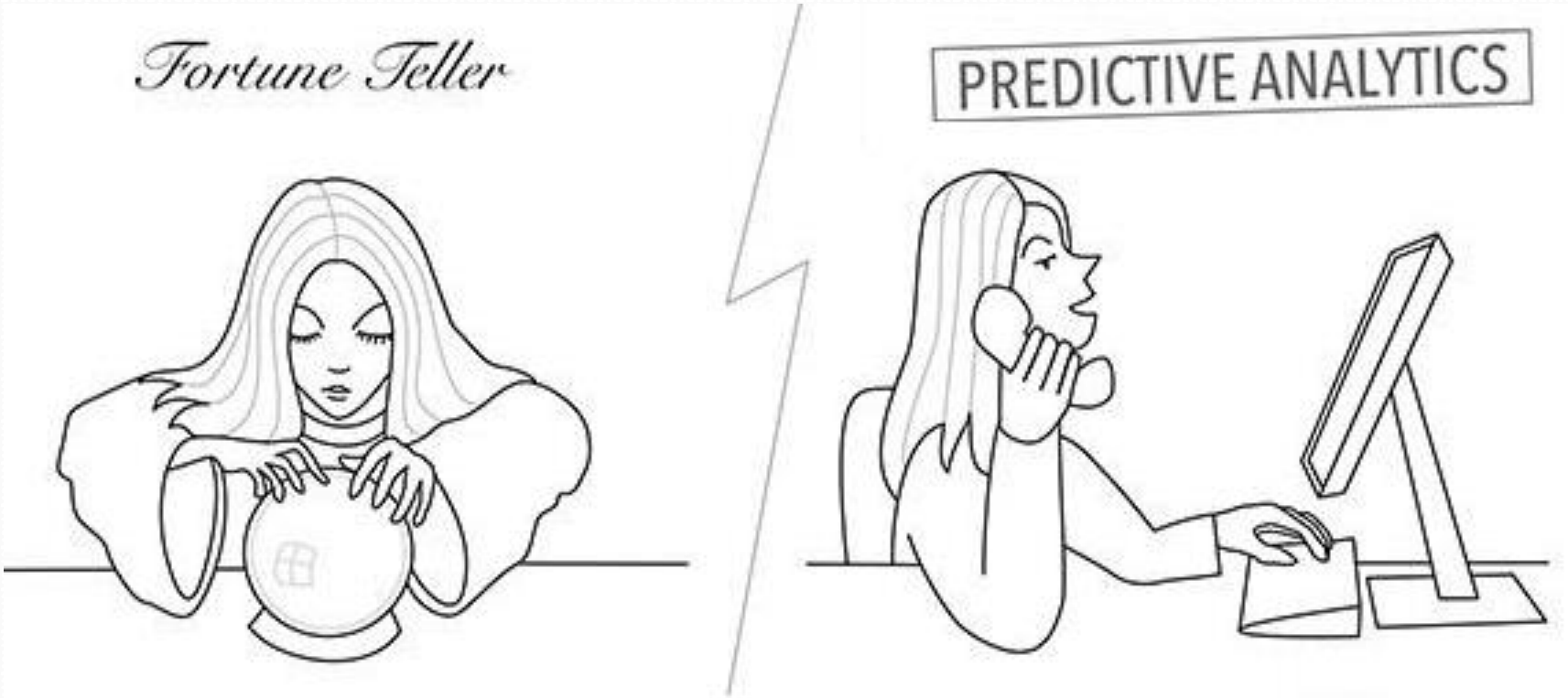
- Algorithms learn from example observations of objects
  - Speech?
  - Image?
  - Health symptoms?
  - Stock price?
  - Personal shopping choice?

What is machine learning?

Given example observations of objects:

- Can we find similar objects?
- Can we make predictions about objects?
- Can we group the objects?

What is machine learning?



"Why the change? Well, I could see where the future was going..."

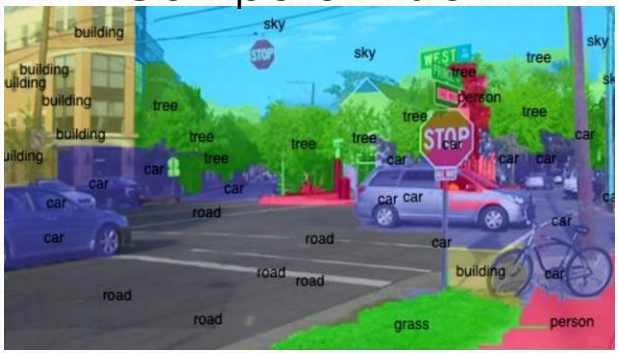
Machine learning

- Variety of algorithms
  - Often the algorithm parameters need to be tuned
  - Each algorithm has its pros and cons
- Important to understand the algorithms
- We will discuss a small selection of algorithms..
  - ..but covering variety

Applications

Speech recognition

Computer vision



Robot control

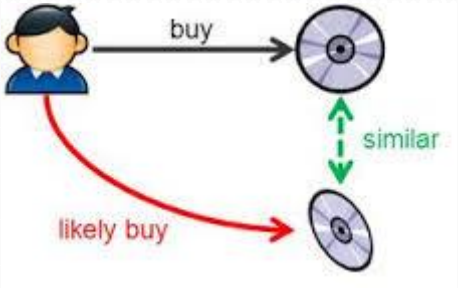
Language translation

Text or document classification (e.g. spam email)

Natural language processing

Applications

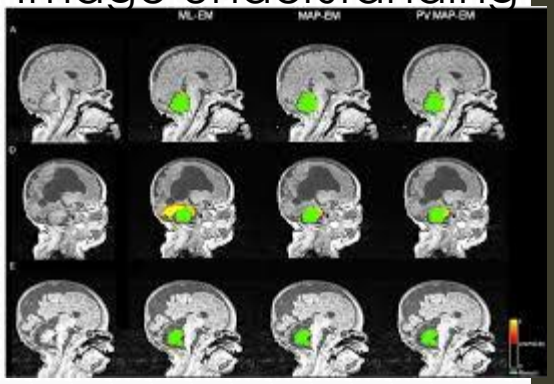
Recommendation system



Personal
software assistant

Google ASSISTANT

Image understanding



Game playing
(e.g. chess, go, backgammon)

Driverless cars

ML in industry

- Companies with lots of data, with a need to 'understand' the data, for which traditional algorithms don't exist

- Google (in almost everything)
- Microsoft (e.g. personal assistant)
- Amazon (e.g. recommendation system)
- Facebook (e.g. friends suggestion, face tagging)
- Uber (e.g. driverless navigation)

Aims

1. Introduce the basic concepts and terminology of machine learning
2. Give an overview of the main approaches to machine learning
3. Show similarities and differences between different approaches
4. Present basic principles for the classification of approaches to machine learning
5. Give practical experience of applying machine learning algorithms to classification and data analysis problems
6. (ML extended only) Develop skills of literature surveying and critical thinking in an area of machine learning

Learning outcomes

On successful completion, the student should be able to:

1. Demonstrate a knowledge and understanding of the main approaches to machine learning
2. Demonstrate the ability to apply the main approaches to unseen examples
3. Demonstrate an understanding of the differences, advantages and problems of the main approaches in machine learning
4. Demonstrate an understanding of the main limitations of current approaches to machine learning, and be able to discuss possible extensions to overcome these limitations
5. Demonstrate a practical understanding of the use of machine learning algorithms
6. (ML extended only) Survey and discuss the research literature in one subfield of machine learning

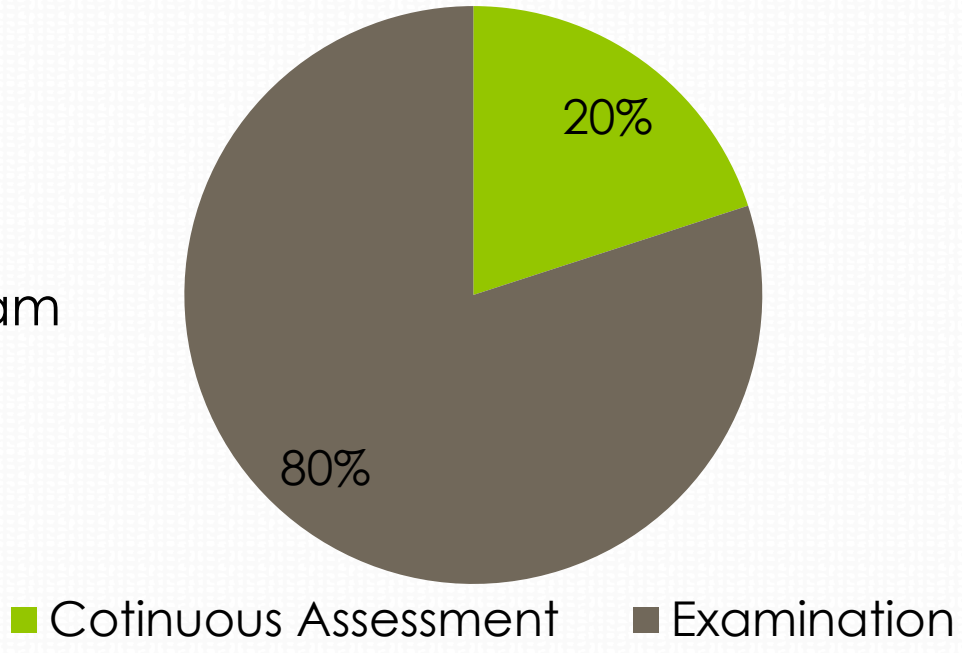
Module focus

- Understanding the fundamental principles
  - Commonly used algorithms
  - Common pitfalls
  - Categories of algorithms

- NOT a module on ML software packages

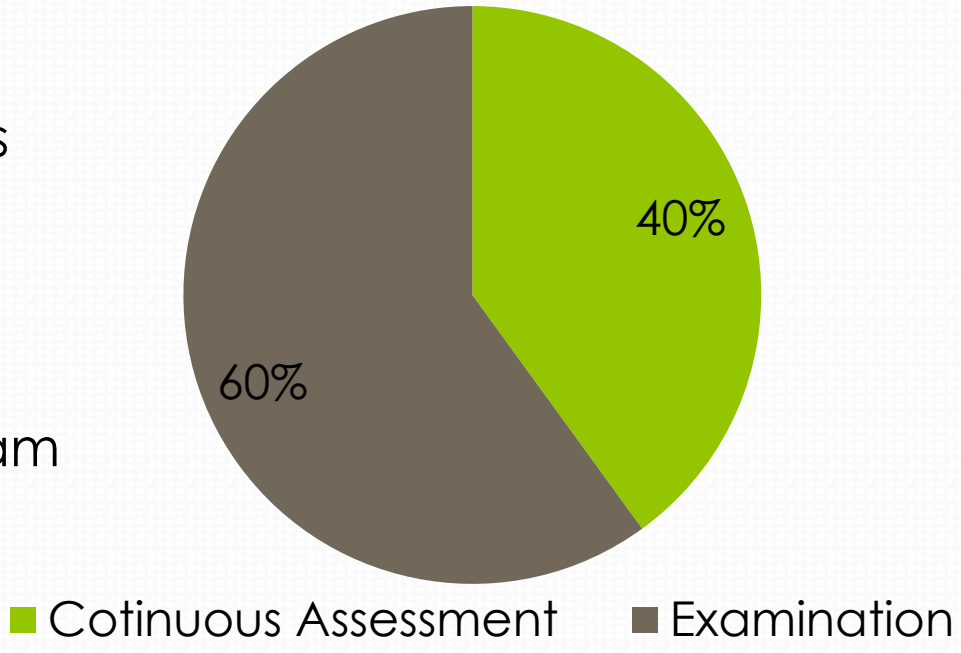
Assessment: machine learning

- Continuous assessment (20%)
  - Class test (20%)
- Examination (80%)
  - 90 minutes written exam
  - Closed-book and closed-notes exam



Assessment: machine learning (extended)

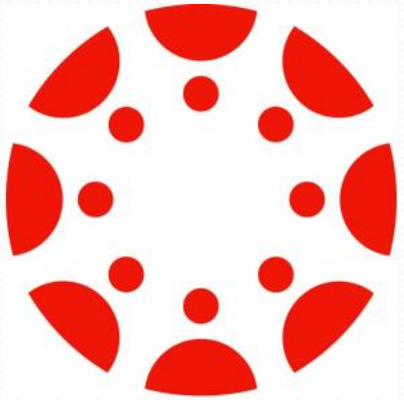
- Continuous assessment (40%)
  - Class test (20%)
  - Computer based tests (2x10%=20%)
- Examination (60%)
  - 90 minutes written exam
  - Closed-book and closed-notes exam



Assessment schedule (tentative)

Week	Test (ML)	Test (ML extended)
1		
2	Computer Based Test (ungraded) announced (6th Oct)	Computer Based Test (ungraded) announced (6th Oct)
3		
4	Computer Based Test (ungraded) DUE (20th Oct)	Computer Based Test (ungraded) DUE (20th Oct)
5	Online Test (ungraded) (27th Oct)	Online Test (ungraded) (27th Oct) Computer Based Test 1 (GRADED) announced (27th Oct)
6		
7		Computer Based Test 1 (GRADED) DUE (10th Nov)
8	Class Test (GRADED) (17th Nov)	Class Test (GRADED) (17th Nov) Computer Based Test 2 (GRADED) announced (17th Nov)
9		
10		Computer Based Test 2 (GRADED) DUE (1st Dec)
11		

Module website

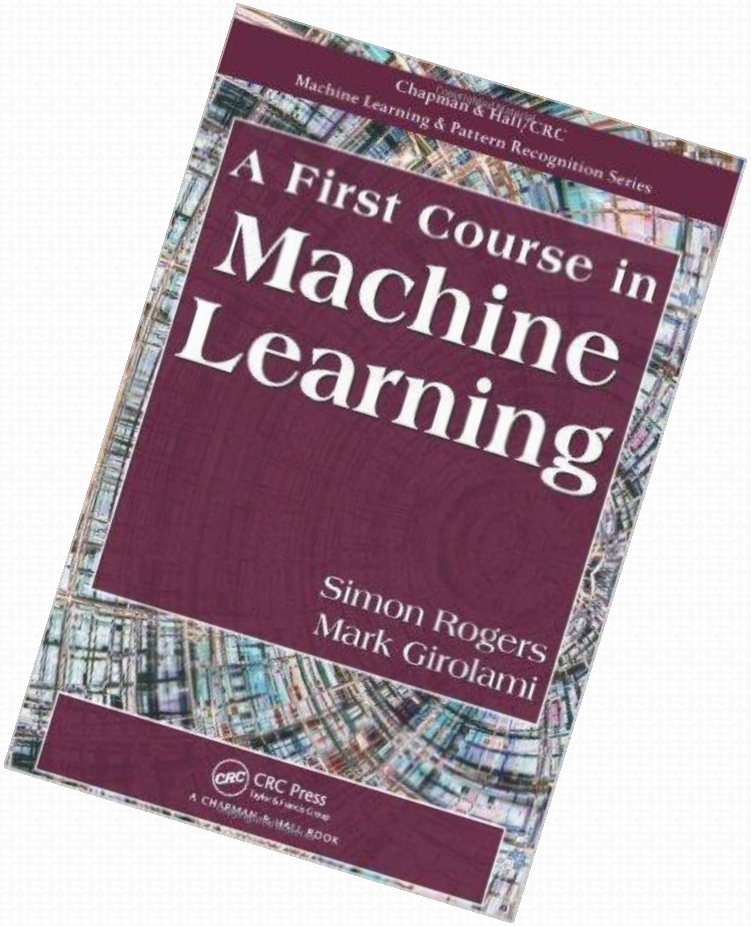


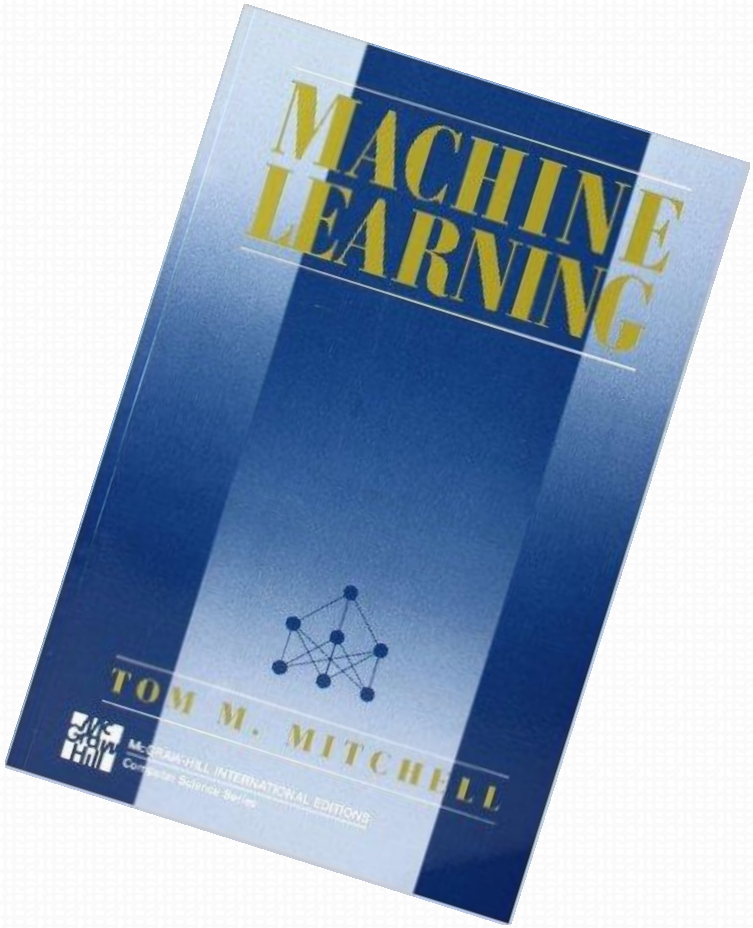
- Module Canvas page
  - https://canvas.bham.ac.uk/courses/27269
- Lecture slides will be uploaded weekly
- Announcements/discussions
- Computer based test submission
- Online class test

Office hours

- Tuesday 9.30am-11am
- Location: LG06d (lower ground floor)

Relevant texts





Plagiarism

- https://intranet.birmingham.ac.uk/as/studentservices/conduct/plagiarism/index.aspx
- https://intranet.birmingham.ac.uk/as/studentservices/conduct/plagiarism/guidance-students.aspx
- http://www.birmingham.ac.uk/Documents/university/legal/plagiarism.pdf

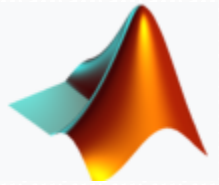
Pre-requisites

- Mathematical techniques for computer science (or equivalent)
- Introduction to AI (or equivalent)
- Math refresher material is available on Canvas
  - Linear algebra
  - Probability theory

Math refreshers

- Linear Algebra
  - Canvas (https://canvas.bham.ac.uk/files/4348230/download?download_frd=1)
  - A First Course in Machine Learning (section 1.3)
- Probability theory
  - Canvas (https://canvas.bham.ac.uk/files/4348231/download?download_frd=1)
  - A First Course in Machine Learning (sections 2.2 to 2.6)

MATLAB



- MATLAB is a very popular numerical computing environment
  - For computer bases tests, solution is required to be developed in MATLAB
  - Available for free through University's campus-wide license (https://mysoftware.bham.ac.uk)
- MATLAB basics (vectors, matrices, loops, plotting, etc)
  - http://www.cyclismo.org/tutorial/matlab/
  - http://users.rowan.edu/~shreek/networks1/matlabintro.html
- MATLAB primer (by Mathworks)
  - http://au.mathworks.com/help/pdf_doc/matlab/getstart.pdf

Basics of machine learning

What is the learning problem?

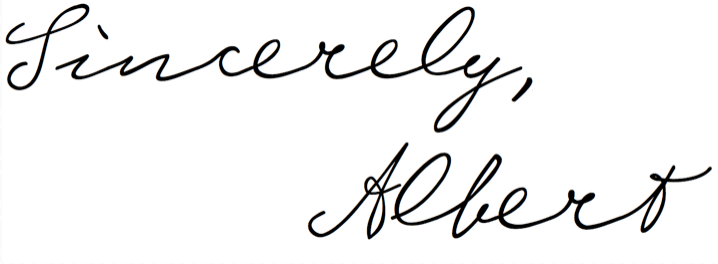
- Ability to improve *performance* (or to make accurate predictions) through *experience* to perform a *task*
  - Improve at task T, with respect to performance measure P, based on experience E
- Task?
- Performance measure?
- Experience?

What is the learning problem?

- Learning to play checkers
- Task T?
- Performance measure P?
- Experience E?

What is the learning problem?

- Learning to recognize handwritten words
- Task T?



- Performance measure P?
- Experience E?

What is the learning problem?

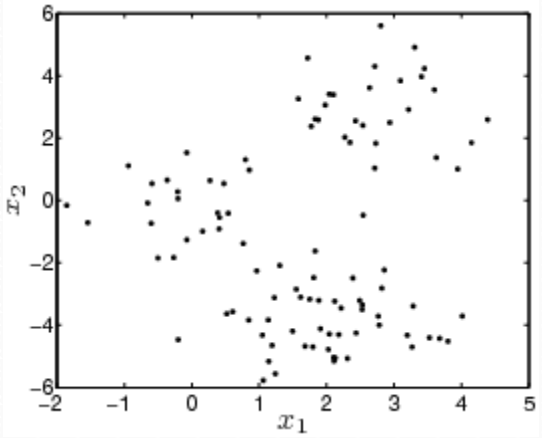
- Learning to recognize faces
- Task T?
- Performance measure P?
- Experience E?

What is the learning problem?

- Learning to drive autonomously
- Task T?
- Performance measure P?
- Experience E?

What is the learning problem?

- Learning to find clusters in data
- Task T?
- Performance measure P?
- Experience E?



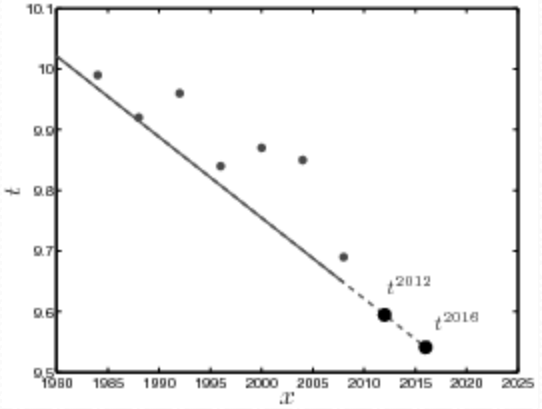
What is the learning problem?

- Learning to interpret image scene
- Task T?

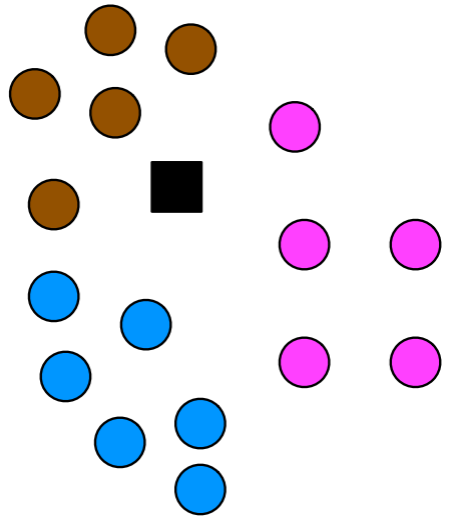
- Performance measure P?
- Experience E?

Classes of learning

- **Regression**: learning a continuous function from a set of past examples
  - Predict a real value target for a future example
  - e.g. predict winning time in Olympic race

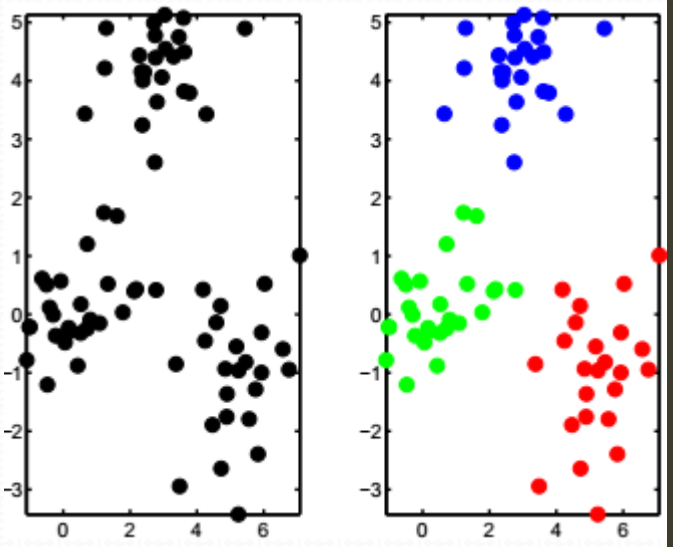


- **Classification**: Learning a function that can separate past examples of different types from one another
  - Assign a discrete target label/type for a future example
  - e.g. document classification

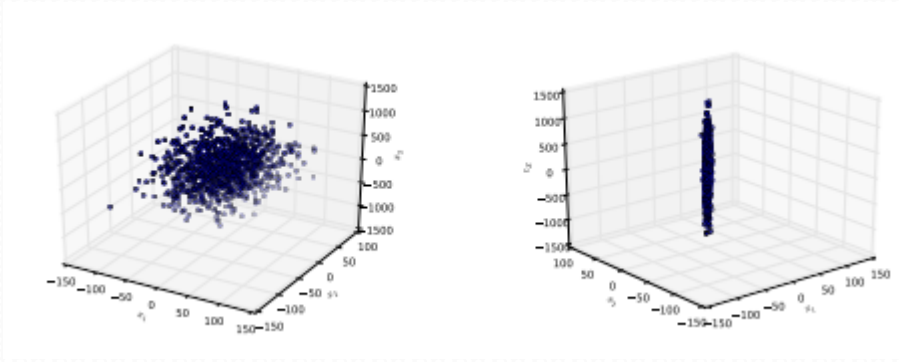


Classes of learning

- **Clustering**: partition examples into groups, each group having similar examples
  - e.g. brain regions with similar activation



- **Dimensionality reduction**: transform high-dimensional data into a lower-dimensional preserving representation
  - e.g. reducing unnecessary attributes



Training experience

- Direct or indirect feedback may be available
  - Chess game move
  - Digit recognition
  - Face recognition
- With or without a teacher
  - Examples (i.e. experience) with or without target labels
  - e.g. face recognition, data grouping
- Is the training experience representative of the performance goal?
  - e.g. digit recognition
  - How well the training examples distribution represent the true examples distribution?

Forms of machine learning

- **Supervised learning**: learner receives set of labelled examples (i.e. direct feedback) in order to learn to classify unseen examples
  - Classification, regression
- **Unsupervised learning**: learner receives set of unlabelled examples (i.e. no teacher) in order to learn to categorize unseen examples
  - Clustering
- **Dimensionality reduction**: transform high-dimensional data into a lower-dimensional preserving representation

ML: important questions

- How much training data is sufficient?
- What algorithms exist for learning general target functions from specific training examples?
- Can we transfer what is learned from one task to improve learning in other related tasks?
- What is the relationship between different learning algorithms, and which should be used when?

ML: important questions

- Can we build never ending learners?
- Can machine learning theories and algorithms help explain human learning?
- Can we design programming language containing machine learning primitives?

Common terminology

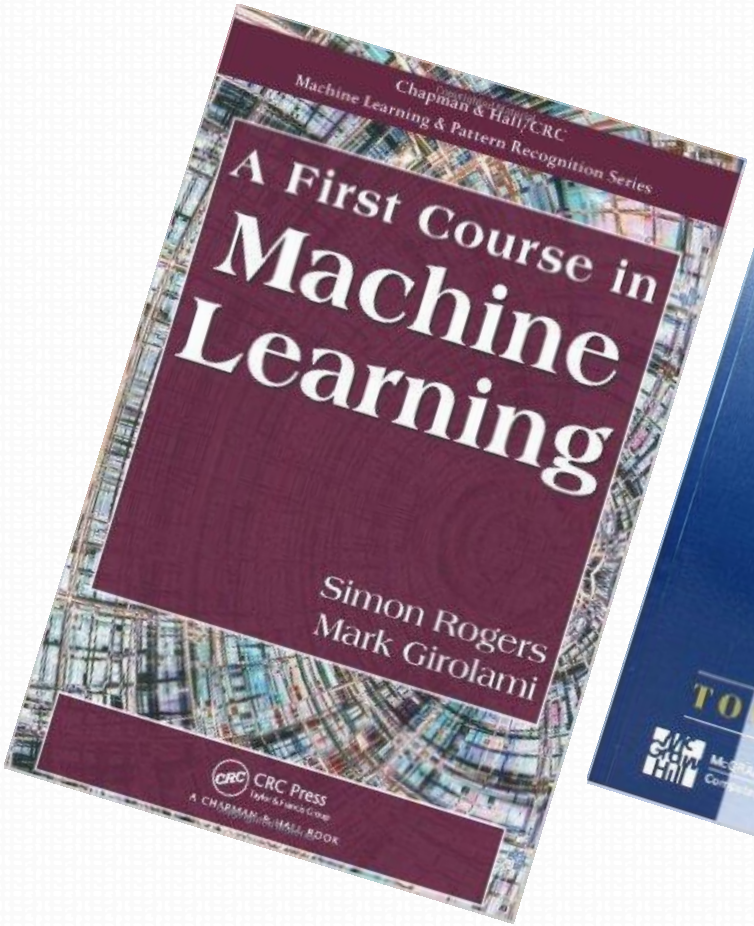
- *Examples*: items of data used for learning or evaluation
- *Features*: attributes that characterize an example
- *Labels*: values or categories assigned to examples
- *Task*: a prediction activity that the algorithm needs to learn
- *Performance*: measure of prediction accuracy of an algorithm
- *Experience*: past examples which can be used in learning
- *Training*: learning to predict from examples
- *Testing*: predicting previously unseen examples

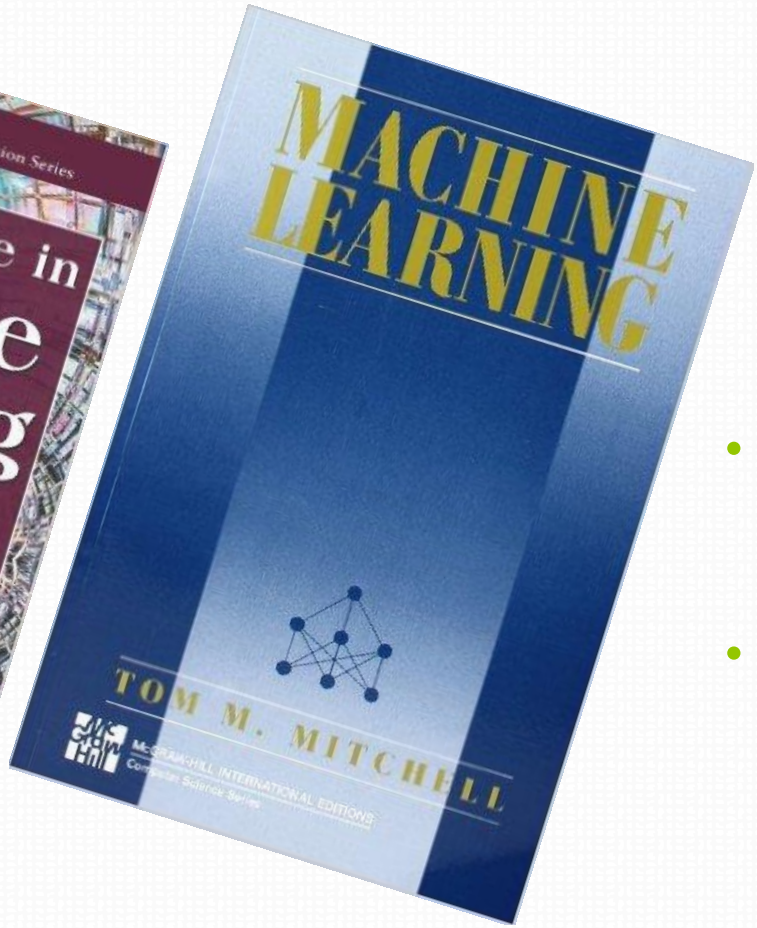
Common terminology

- *Cross validation*: distribute data into k-folds to train and evaluate algorithm performance
- *Training samples*: examples used to train algorithm
- *Validation samples*: examples used to tune algorithm parameters
- *Test samples*: examples used to evaluate algorithm
- *Loss function*: performance (loss) measure function
- *Learner function/model*: a function or model that is learnt to predict labels from features
- *Hypothesis set*: set of functions mapping features to labels

List of topics (tentative)

- Basics
- Supervised learning
  - Regression: linear modelling by least squares
  - Regression: linear modelling by maximum likelihood
  - Classification: Bayesian classification
  - Classification: instance-based classification
  - Classification: discriminative classification
- Unsupervised learning
  - Clustering: k-means clustering
  - Clustering: hierarchical clustering
  - Dimensionality reduction: principal component analysis
- Ensemble methods
  - Boosting
  - Random forests





Author's material (Simon Rogers)

- Ata Kaban's material from previous years
- Various other sources for graphical illustration

Thank You