COMP1804

Applied Machine Learning

Lecture 1: Course Introduction

COMP1804
Learning Outcomes

- **On completing this module successfully, you will be able to:**

A. Rationalise appropriate scenarios for machine learning applications and evaluate the choice of machine learning methods for given application requirements.

B. Demonstrate competency in using appropriate libraries/toolkits to solve given real-world machine learning problems and develop and evaluate suitable application.

C. Understand and apply the relevant input data preparation and processing required for the machine learning models used, and quantitatively evaluate and qualitatively interpret the learning outcome.

D. Recognise and critically address the ethical, legal, social and professional issues that can arise when applying machine learning technologies.

Course Contacts

Dimitrios Kollias
D.Kollias@greenwich.ac.uk
Office Hours: Contact by email
Personal Website

A K M Mahtab Hossain
a.k.m.hossain@greenwich.ac.uk
Office Hours: Contact by email
Personal Website

- Use your Greenwich email account to email
  otherwise your emails my be caught by our spam-filter
- Put COMP1804 in the subject line
- We will aim to reply within 48 hours during term-time.
  If you don't get a reply by then, send again!

COMP1804
structure

- **Structure**
  - One hour weekly lecture: pre-recorded, 09:00-10:00 Fridays
  - Two hour weekly lab and Q&A: online, 10:00-12:00 Fridays

- **Course Website**

  Moodle: COMP-1804-M01-2020-21 Applied Machine Learning

Essential Resources and Further Reading

- Further reading (ebooks/guide):
  - Real World Machine Learning (eBook)
    Brink, H., Richards, J., and Fetherolf, M. - www.allitebooks.in/real-world-machine-learning/
  - Hands-On Machine Learning with Scikit-Learn and TensorFlow: Concepts, Tools, and Techniques to Build Intelligent Systems
    Géron, A., O'Reilly Media
- Useful resources:
  - UCI Machine Learning Repository
  - UCI KDD Archive
  - Statlib
  - Delve
  - Kaggle for datasets, project example and online competitions
  - SeedBank (a collection of ML examples by Google)

General Resources: Journals

- IEEE Transactions on Pattern Analysis and Machine Intelligence (TPAMI)
- International Journal of Computer Vision (IJCV)
- Journal of Machine Learning Research
- Machine Learning
- Neural Computation
- Neurocomputing
- Neural Networks
- IEEE Transactions on Neural Networks and Learning Systems (TNNLS)
- Pattern Recognition
- Annals of Statistics
- Journal of the American Statistical Association

etc.

General Resources: Conferences

- IEEE Conference on Computer Vision and Pattern Recognition (CVPR)
- Neural Information Processing Systems (NeurIPS)
- International Conference on Machine Learning (ICML)
- International Conference on Learning Representations (ICLR)
- Conference on Artificial Intelligence (AAAI)
- European Conference on Computer Vision (ECCV)
- International Conference on Computer Vision (ICCV)
- International Joint Conferences on Artificial Intelligence (IJCAI)
- European Conference on Machine Learning (ECML)
- Uncertainty in Artificial Intelligence (UAI)
- International Conference on AI & Statistics (AISTATS)
- Computational Learning Theory (COLT)
- International Conference on Artificial Neural Networks (ICANN)

etc.

COMP1804
assessment

- Logbook (20%)
  - individual work
  - four selected exercises throughout the course
  - all exercises should be completed and uploaded by 19/03/2021
    (advice: complete each exercise after each lab or within a week)

- Practical Coursework (80%)
  - could be work in groups but students should upload individual assignments
  - involves practical development and report in IEEE-style paper
  - is due by 09/04/2021

Avoiding Plagiarism

- Discussing the course, helping each other with tutorials and coursework is great! (the goal should be to understand better, not copy work)

BUT!!!

- Copying from another student, submitting same coursework report, allowing someone to copy your work, or getting someone else to do your work
- Copying from Web, book etc. without clear and explicit acknowledgement
- Not acknowledging who you got help from/did help in your tutorial/coursework

IS PLAGIARISM

- We do check for it and penalties are severe :(
- Please don't put yourself or your friends at risk
- If in doubt - ask the staff

But...

What is machine learning?

“Instead of trying to produce a programme to reproduce the adult mind, why not rather try to produce one which simulates the child’s?”

Alan Turing, 1950

So… Machine Learning
The COOLEST TOPIC IN SCIENCE

- “A breakthrough in machine learning would be worth ten Microsofts” (Bill Gates, Chairman, Microsoft)
- “Machine learning is the next Internet” (Tony Tether, 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, Dir. Research, Yahoo)
- “Machine learning is going to result in a real revolution” (Greg Papadopoulos, CTO, Sun)
- “Machine learning is today’s discontinuity” (Jerry Yang, CEO, Yahoo)

How would you teach someone to distinguish...

between a puppy and a muffin?

Puppy or Muffin?

Machine Learning is learning the rules

plus experience.

Some definitions of Machine Learning

- A definition (by Tom Mitchell):
  “How can we build computer systems that automatically improve with experience, and what are the fundamental laws that govern all learning processes?”

- More technically:
  "A computer program is said to learn from experience E with respect to some class of tasks T and performance measure P, if its performance at tasks in T, as measured by P, improves with experience E"

Why shall the Machine “Learn”? (I)

In general, the machine learns to address the case where the human has not found or is not able to find the law to explain the observed data, such as:

- the human is not able to explicitly encode his/her expertise in computer algorithms
  → speech/face recognition
- the human has no expertise
  → e.g., automatic navigation on Mars
- problems of which the solution needs to be adapted to particular cases
  → user biometrics

etc.

Why shall the Machine “Learn”? (II)

Machine Learning provides smart ways to analyse vast volumes of data. Nowadays there is an abundance of data to learn from, in almost all application domains.

By developing fast and efficient algorithms and data-driven models for real-time processing of data, Machine Learning can produce fast and accurate results and analysis.

What's happening in the industry & a timeline... (I)

1959	The term 'machine learning' was coined by computer scientist Arthur Samuel
1964	Eliza, the first chatbot, is developed at MIT
1997	IBM's Deep Blue defeated Garry Kasparov, the world's chest champion
2009	Chris Bishop: "Even the most sophisticated computers can't tell a cat from a dog!"
2011	IMB's Watson Q&A machine beat humans on Jeopardy! game; Google Brain used Deep Neural Networks to identify and categorise objects; Apple integrates Siri, a personal voice assistant, into the iPhone
2012	How many computers do you need to identify a cat? 1600
2014	Facebook developed DeepFace, a deep-learning facial recognition system; Youtube recognizes cats from videos

What's happening in the industry & a timeline... (II)

2015 Microsoft & Google beat humans in image recognition

2016 Machine Learning algorithms outperform inexperienced radiologists; Google's AplhaGo defeated a Go world champion

2018 Google AI can spot advanced breast cancer more effectively than humans. It's up 99% accurate under the right conditions.

Today: Revolution & Technologies

- Moore's Law + GPUs
- Cloud computing
- More and richer data
- Github + Open source
- New and refined techniques

APIs and Open Source Tools
Google

Machine Learning as an API

Cloud Vision API

Cloud Speech API

Cloud Natural Language API

Cloud Translation API

Cloud Video Intelligence API

Use your own data to train models

TensorFlow

Cloud Machine Learning Engine

APIs and Open Source Tools
Amazon

Recommendations

AMAZON PERSONALIZE »

Forecasting

AMAZON FORECAST »

Image and Video Analysis

AMAZON REKOGNITION »

Advanced Text Analytics

AMAZON COMPREHEND »

Document Analysis

AMAZON TEXTRACT »

Voice

AMAZON POLLY »

APIs and Open Source Tools
Microsoft

Machine Learning service

AI + MACHINE LEARNING

Translator Text API

AI + MACHINE LEARNING

Machine Learning Studio

ANALYTICS

Face API

AI + MACHINE LEARNING

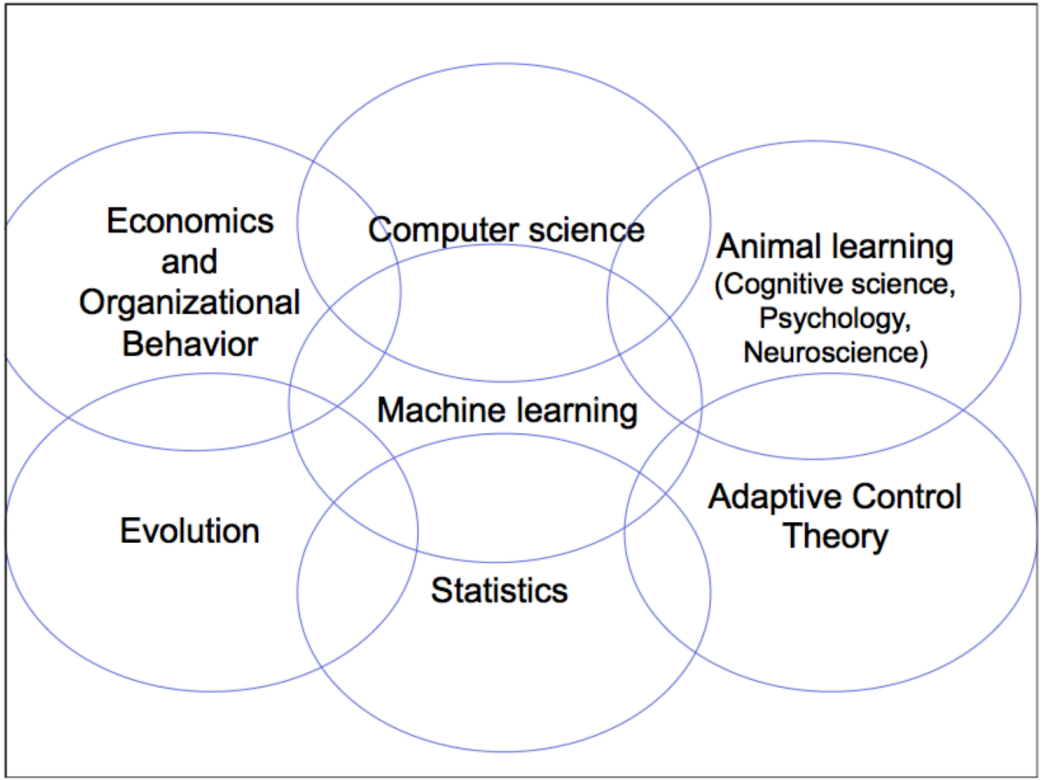
Batch

COMPUTE

Data Factory

DATABASES

Many Communities relate to Machine Learning

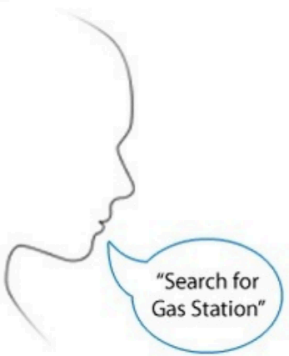


Machine Learning is Everywhere

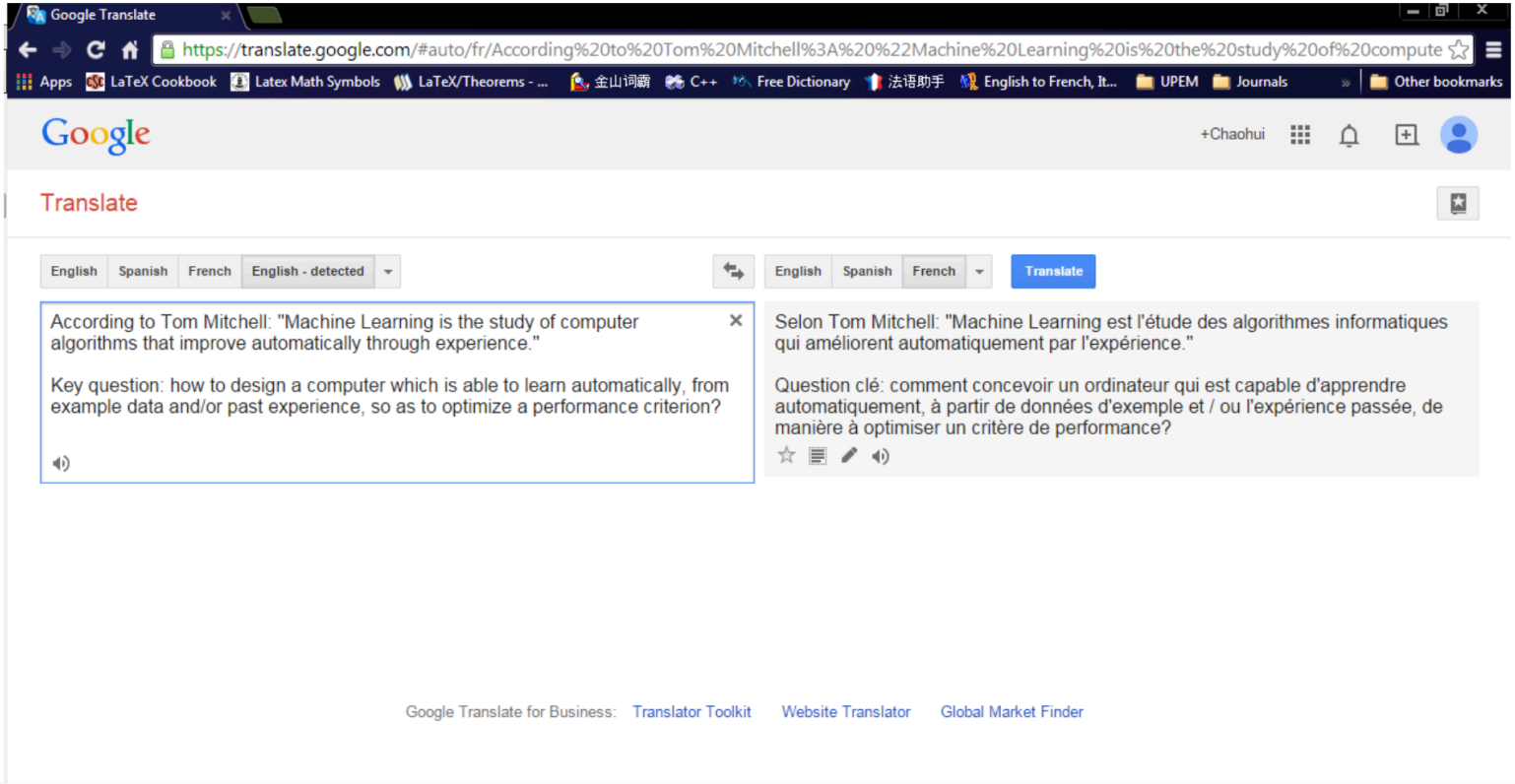
- Speech Recognition, Natural Language Processing
- Computer Vision
- Robotics
- Medical Outcomes' Analysis
- Computational Biology
- Sensor Networks

etc.

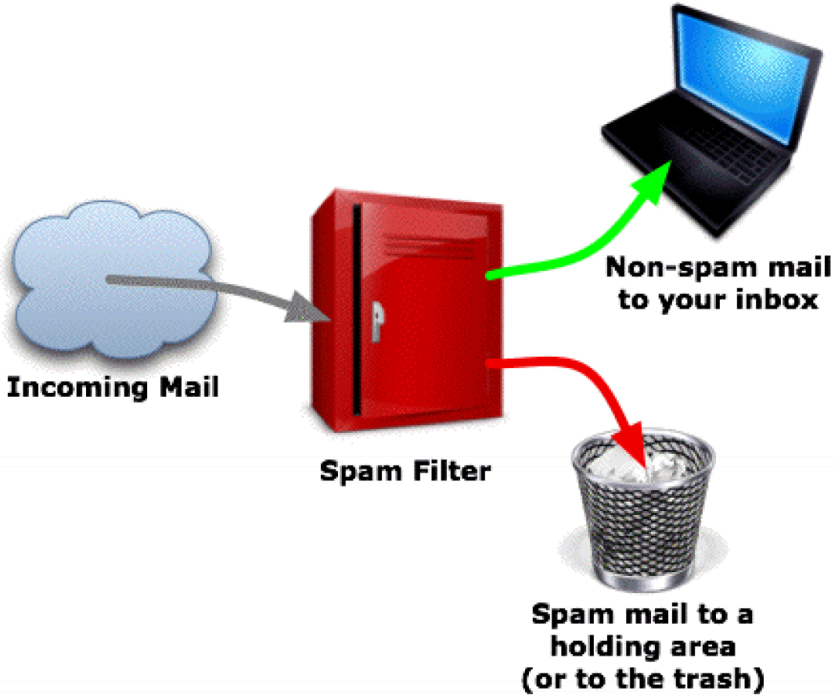
Case Study 1
Speech Recognition



Case Study 2
Machine Translation



Case Study 3
Spam Detection



Dear Sir.

First, I must solicit your confidence in this transaction, this is by virture of its nature as being utterly confidencial and top secret. …

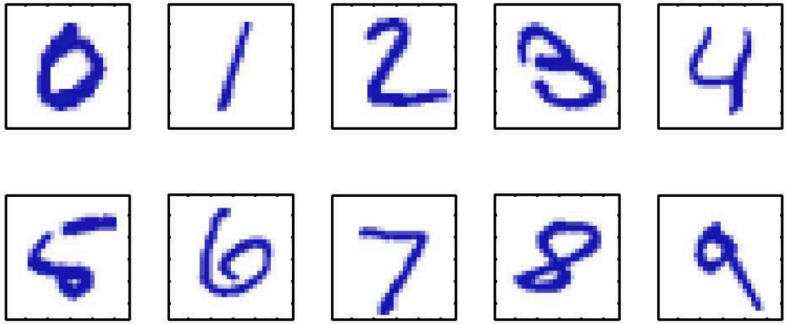
TO BE REMOVED FROM FUTURE MAILINGS, SIMPLY REPLY TO THIS MESSAGE AND PUT "REMOVE" IN THE SUBJECT.

99 MILLION EMAIL ADDRESSES
FOR ONLY $99

Ok, Iknow this is blatantly OT but I'm beginning to go insane. Had an old Dell Dimension XPS sitting in the corner and decided to put it to use, I know it was working pre being stuck in the corner, but when I plugged it in, hit the power nothing happened.

Case Study 4
Computer Vision: Optical Character Recognition (OCR)

Handwritten digit recognition: >99% accuracy



Applications: recognising addresses, checks, books, pen input, etc.

Main difficulty: wide variability of same number/character

Case Study 5
Computer Vision: Face Recognition

Training examples of one person

Testing images

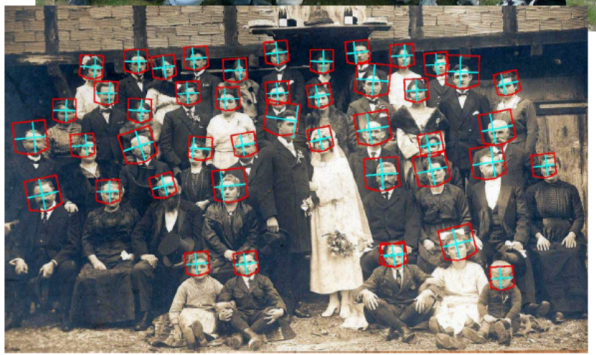
Main difficulty: large intra-class variations, due to various factors:

- viewpoint change
- illumination change
- hair-style change
- age change

etc.

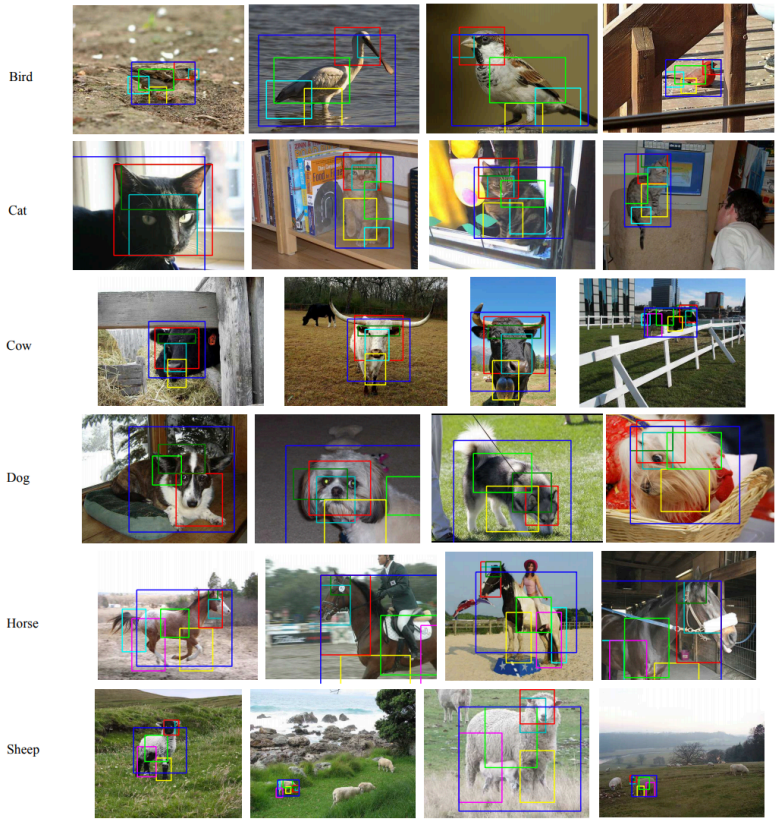
Case Study 6

Computer Vision: Face Detection



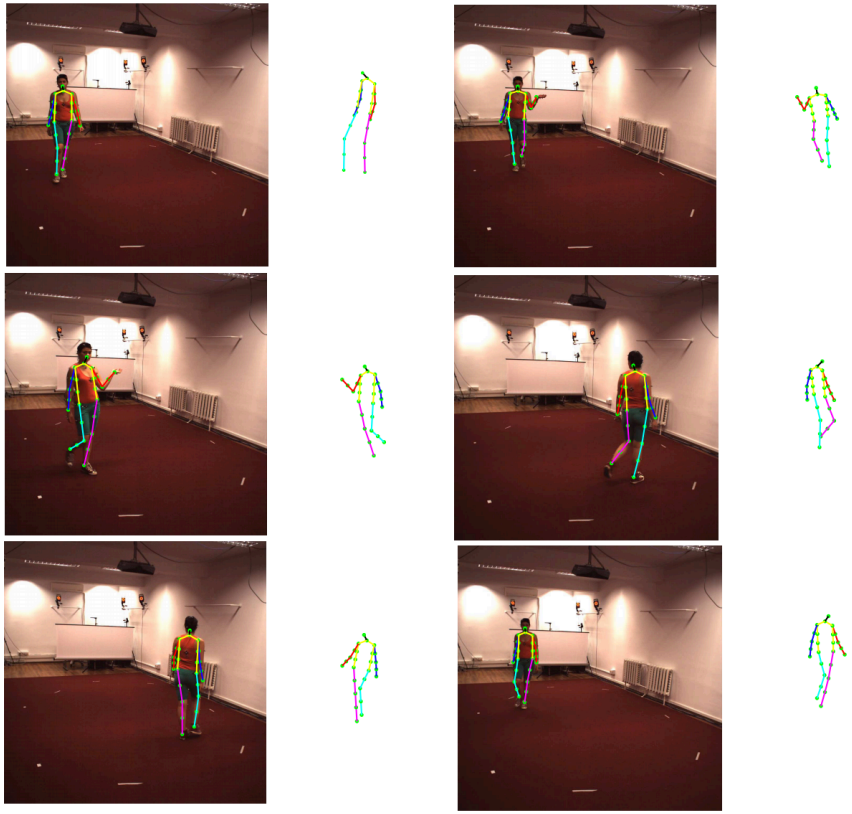
Case Study 7

Computer Vision: Object Detection



Case Study 8

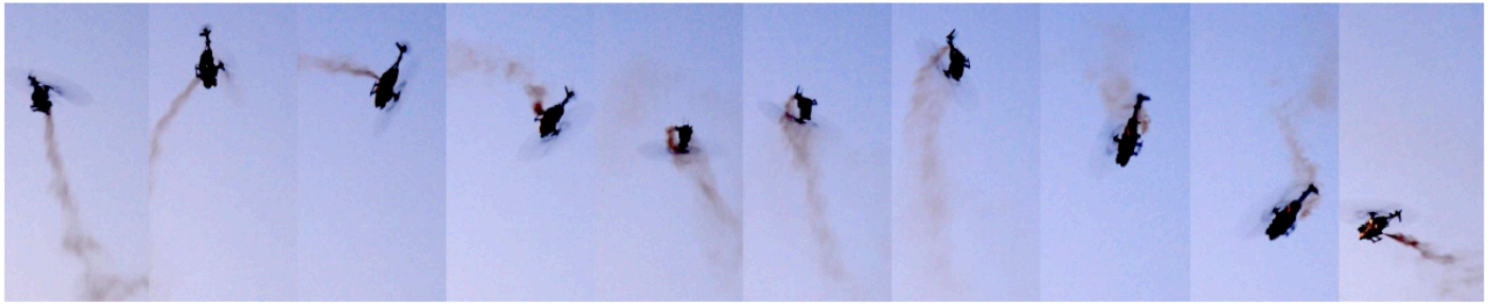
Computer Vision: Pose Estimation



Case Study 9
Computer Vision & Robotics

- Self Driving Cars: e.g., Google Prototype:

- Autonomous Helicopter Control: e.g., http://heli.stanford.edu/

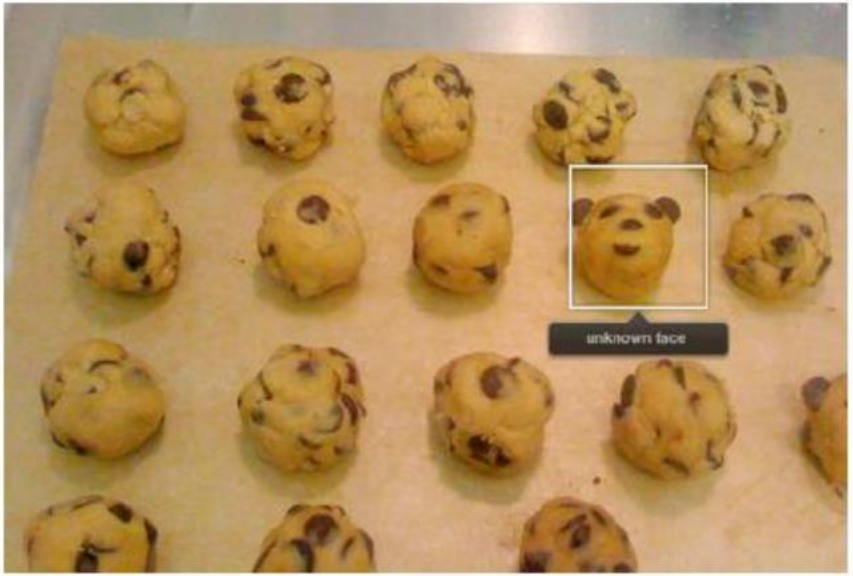


- Navigation of Mars Lander, . . .

However there is still work to be done (I)

Face recognition

Not always perfect!



However there is still work to be done (II)

"A refrigerator filled with lots of food and drinks"

"A yellow school bus parked in a parking lot"

Back to the definition

"computer systems that automatically improve with experience"

"A computer program that … improves it performance at tasks in T, as measured by performance measure P, improves with experience E"

- So WHAT is "experience"?

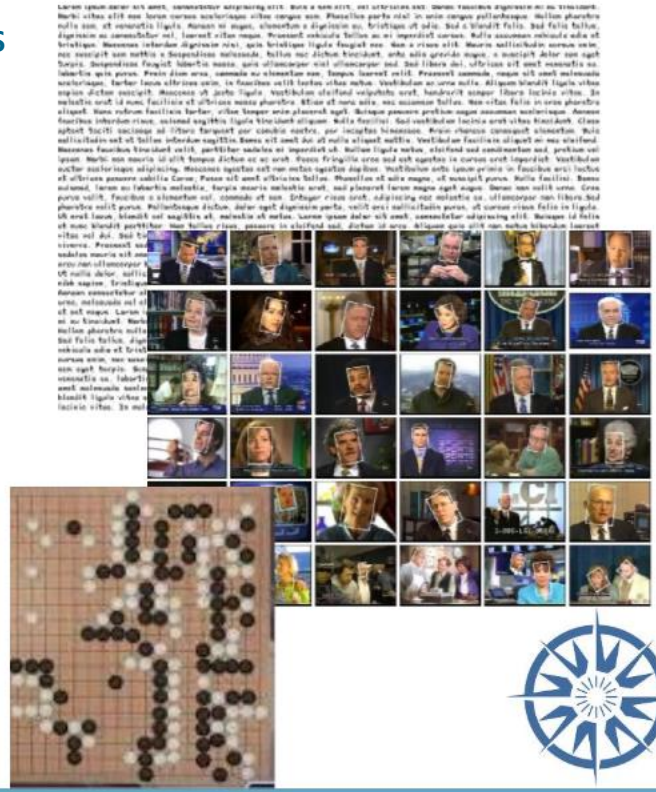
lines of text

labeled images

games played

Representing data

- Machine learning algorithms (typically) only see numbers
- Typically, we create one vector representing each experience
  - Either with raw data values (pixel, characters) or preprocessed data (words, colours, shapes)
  - Vectors organised in a table



Machine Learning Problems
terminology

- Data is often presented in tables
- Columns are called input variables or features or attributes.
- The columns we are trying to predict (outcome and time) are called output variables/targets.
- A row in the table is called a training example or instance.
- The whole table is called a data set.

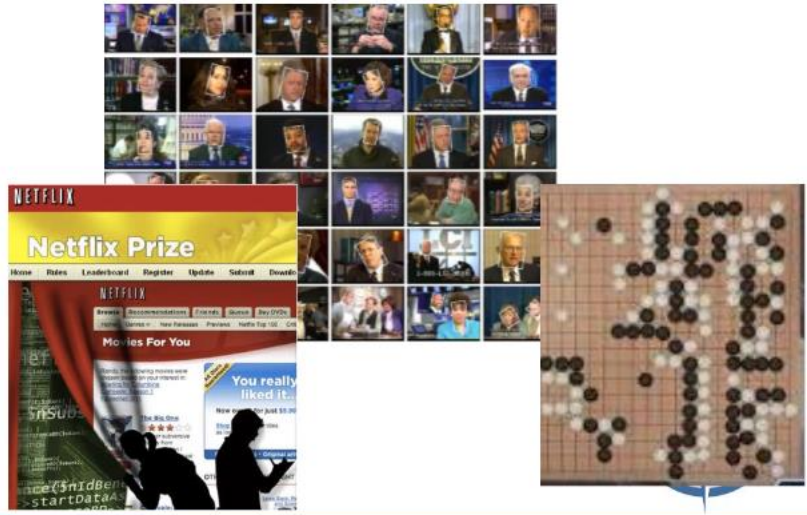
tumor size	texture	perimeter	...	outcome	time
18.02	27.6	117.5		N	31
17.99	10.38	122.8		N	61
20.29	14.34	135.1		R	27
...					

Back to the definition

"computer systems that automatically improve with experience"

"A computer program that ... improves it performance at tasks in T, as measured by performance measure P, improves with experience E"

- So WHAT is "tasks"?
- We've seen some examples
- Can we categorise them?



Main types of machine learning

- Supervised learning
  - Classification
  - Regression
- Unsupervised learning
- Semi-Supervised learning
- Reinforcement Learning

Thank you for the attention

Special Mention on Plagiarism

Please check and think what is the problem in the following slide

So… Machine Learning
The COOLEST TOPIC IN SCIENCE

- A breakthrough in machine learning would be worth ten Microsofts
- Machine learning is the next Internet
- Machine learning is the hot new thing
- Web rankings today are mostly a matter of machine learning
- Machine learning is going to result in a real revolution
- Machine learning is today’s discontinuity