# Machine Learning CS 401 18/09/17

Updates have been added to the course.

1st: partly based on St John’s University (They do everything from original material/studies)

Barak will do both original & more recent.

2nd: Extra points gained for taking notes in class, reviewing them with Barak, and submitting via github.

3rd: Project: Optionally, students may take on a project. This would look great on our CV, but it’s not mandatory.

## What is Machine Learning?

Machine Learning grows out of different disaplines.

It does stuff by looking at data & figuring out what that data tells us.

Examples: Processing cheques in a bank, reading the information from these cheques. or… A Post Office reading numbers off of envelopes.

Observe:

Digits can be written many different ways in handwriting. For example, an ‘o’, might loop around on itself, or have a gap and not join up. A hand writing digit recogniser would need to ‘learn’ how to recognise various handwriting versions of letters.

We usually think of Machine Learning as taking great big data sets.

Google has lots of data on us. It uses ‘big fancy’ algorithms to analyse this data to learn about our browsing habits and therefore interests etc

Best minds in CS are dedicated to collecting such data, looking for example at the correlation between location, credit card transactions, types of purchases etc.

In this module we’re looking at data that’s easier to characterise.

1st Machine Learning systems were like those by Arthur Samuel in IBM, to program a computer to run chequers. They wanted to improve its performance with use.

Remember: this was pre-internet. So, not too many games were played for the computer to learn from them. Arthur Samuels idea was to have the computer play itself.

These early systems developed into systems like those we use today. ie: Alpha go & Deep Blue.

These are an example of where main source of data is internally generated.

“It’s the field of study that gives computers the ability to learn without being explicitly programmed” as per Arthur Samuel.

Inputs 🡪 🡨 outputs

binary switches

knobs

Goal in Machine Learning is to adjust these knobs.

Statisticians do this.

Statistics Math Physics Neuroscience Cognitive Science Information theory

esp (thermos) dynamics

Control Theory

ML

Statisticians were very dubious about Machine Learning. Stats were defined by making guarantees and putting in error bars. They saw ML as being a ‘bunch of cowboys’ who threw stuff together.

Question: How to formulate this?

Need to have some kind of inputs & need to have some kind of quality metrics.

For example, you could freeze a version of chequers, and let another version run. Then after a time compare the version that had longer to learn with the one that had less to see how many wins they each have to evaluate.

In thermo dynamics there’s a similar concept. ie: moving to a low energy state. Maths for this is the same as maths for ML.

In Neuroscience - Nervous system is made up of voltages, where neurons take in inputs from other neurons.

In modelling neuroscience, Machine Learning started with one layer per-ceptrons then multi layered per-ceptrons. Example of one model layer per-ceptrons… mobile adapts to radio frequency in a room.

Over time, figured out how to get really deep structures -> led to deep learning.

Brains -> slow stuff

Computers -> fast stuff

Cognitive Science -> questions such as does the subject go before or after the verb?

Observed that animals & children learned without thinking about things at depth.

People try modelling the behaviour of rabbits for example. By blowing some air into their eye causing the eye to close. Then by ringing a bell before the air would be blown into the eye. Then by ringing a different sounding bell that didn’t have a puff of air blown into the eye. Further tests might have included presenting a flash of light or different stimuli and provision of food.

Information theory. Example: a bit is a unit of information. Lots of maths comes from information theory. EG: given loads of inputs, check is a load of text spelled correctly? Can the next word be predicted? etc. Inside the brain, information theory seems to be very important

Control theory: Usually think of simple examples. -> pressure point on a boiler, -> cruise control in a car.

### REMAINDER OF LESSON - REFER TO SLIDES

When viewing the sheep farming image…

See a biped, 2 quadpeds. And lots of other detailed information. You might be able to predict direction or next likely action. However if an alien were to view the image they would not be able to make sense of it having no experience of our world.

Supervised Learning: -> figure out structure of a big amount of data. eg: if input is from a Sumerian alphabet. It doesn’t know the Sumerian Language but recognises the patterns and therefore can identify the language without understanding it.

Used in predicting things

Unsupervised Learning: -> reasoning. A baby doesn’t do much, they get fed and looked after. They pretty much just lie there or are carried. Yet they learn a lot about the world just by touching things, seeing and hearing.

input -> ie sensory data

output -> a reward or a delayed reward

Example, learning to ski.

Reinforcement Learning: Is the Black Sheep of Machine Learning. Until relatively recently it was difficult to achieve good success rates. This is no longer the case nowadays.

Re: The x & o graph…

The dotted line refers to the hypothesis.

It represents a 2D set of data.

x’s & o’s represent two different classes of data.

Example: fraudulent credit card use & non fraudulent credit card use.

Re: The Regression Graph…

Output isn’t binary

Y is the hypothesis

Re: The cluster graph…

Might be used to predict when a motor might fail.