11-785 - Deep Learning 21/04/2020

Lecture 1 - Introduction

- supplement the slides with your notes ; motivation, observations
- course always covers implementation (use history to assist for the future)
- A taxonomy between AI, ML introduced
- ML as statistical paradigm vs symbolic AI
- neural networks beating previous benchmarks, humans
- ASR (2016) - Microsoft
- machine translation - Big data, GPUs
- neural networks 'able to generalise', but relies on statistical emul.? of expert knowledge
- some very grand claims about neural networks (they're hyped atm)
  - be careful

Q: What are neural networks?

- I/O device
- Brain posited as a neural network → conflation of model and phenomenon
- understanding brain is still final frontier
- move from what is thinking perhaps to what does it do
- there may be undiscovered/unmined insights from cybernetics that can be of use.

H: History + philosophical context → spare time.

- Associationism
- Connectionism
- Afferent / efferent connections
- Alexander Bain (1873) - Information is in connections
- Bain (1873) - Neural groups → foundations of connectionism.
  - A learning mechanism, interpreted as a prediction / foreshadowing of Hebbian learning
- Russell quote - remember this
- Alexander Bain (1873) - did back of the envelope calculations; but RECANTED (he was spot on about no. of neurons) due to self-doubt

- connectionism:-
  Alexander Bain (1873) - Mind and Body
- connectionist machines → knowledge stored in connections between elements

- Infant apoptosis
- connectionist machines are counterpoint to von Neumann machines
- we use von Neumann machines to emulate neural networks.
- neural networks are fall into category of connectionist machines.
- Biological model of NN
  - information based interpretation of neurons
- crits for Neural Basis of Cognition
- McCulloch and Pitts (1943)
- mathematical model of a neuron
  (11): Excitatory + inhibitory synapses -
    - can be seen as weights

(18) McCulloch + Pitts
(18) - Read (11) -
- Relation between structure of synaptic model and Boolean logic gates
- computer science → logic can be emulated with arbitrary components of gates

criticism: - overstatements
  - no learning mechanism.

Hebb (1949) :- organisation of behaviour
  - neurons that fire together wire together.

- look at chemical basis of Hebbian learning in neurons if you are interested in developing field

$w_i = w_i + \eta x_i y$

- Good initial idea
- Modifications of Hebbian learning → no negative feedback

  - see slides → tendency for everything to be more connected.

- Ad-hoc mod. trajectory:-
  - Oja's rule
  - Generalised Hebbian learning (Sanger)
  - Renormalisation can yield stability

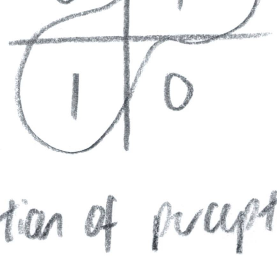
- Rosenblatt (1958) - Perceptron
- You've done this already - refresh

(A2) read the slides in spare time

- Minsky & Papert (1968)
- One neuron cannot simulate XOR
- But multiple neurons can e.g. MLP rather than perceptrons
- universal approximators
- MLPs - lecture 2
- Summary is excellent → spare time, understand context, history
- Sigmoid unit vs step function

(real-valued) vs (discrete)

- Allows continuous real-valued rep.
- Sigmoid as an approximation to Boolean
- Perceptron as a linear classifier → restricted to hyperplanes
- XOR:



→ cannot be done with 1 hyperplane

- composition of perceptrons for complicated decision boundaries

- ML spaces
  ↳ 784 dimensional for 28x28
- what region of ~~dim~~ high dimensions can you construct a hypersurface
- slide is excellent pedagogically
- (61): MLPs — connectionist
  — Boolean functions
  — Boolean machines

Continuous-valued outputs

- MLP as continuous-valued regression (square pulse)
(8) - (A3) - see relation between this and regression
- see Kolmogorov results + in context of tons of data

- (71): summary

(78) A4 - see skipped slides

- (71): neural nets as complex concatenation of functions
  I/O device
- (L2) - neural networks as universal approximators