Deep Learning

COL 865

July 24, 2017

Logistics:-

(I) Who should do this course?

~~this~~ ↳ Anyone who is interested in deep learning

(II) Eligibility:-

A large number of registrations. 54

∴ Need to bring down the numbers:-

(I) At least (A-) in ML
& At least (B) in AI

~~↳ exceptions~~

(II) Exceptions:-

(I) One grade lower (in AI)
but CG 9+.

(II) Mtech/DD Student:- ML (B) is fine. But should be relevant for research

(III) PhD students:- ok if important for research.

Only CSE/SIT students.

Note:- Similar course in EE.

Who should not do:-

(I) No ML background

(II) Has done EE course. (Don't recall the number)

↳ Instructor:- Raghavendra (EE)

Evaluation:-

Three components:-

(I) ~~Exa~~ Assignments:-

Individual

(a) 5 points (Mathematical foundations)
↳ first month

(b) 15 points:- Basic Deep RL coding (CNN) on a known dataset
(Month 2)

(c) 15 points:- More advanced (mini-project) implementation
→ groups of two.
(Month 3)

(II) Paper Presentation:-

Each student will present a paper (groups of 3 or 4)

↳ One hour slot
20 min presentation + Discussion

→ 10 points + 5 Points for writing review.

(2)

Exams:-

One Mid-term (Minor) exam
↳ 20%.

Major Exam → 30%.

Cheating:-

You know the Penalties :→ expect at least a grade down. May be More.

"Deep" learning should be your goal & not "copying" whether shallow or deep :)

Class Timings:-

Move to 2 1/5 hr classes.

This week:-

Monday :- 11 to 12

Thursday:- 12:00 - 1:40

Next week onward:-

Monday:- Afternoon ? 1:00 to 2:30 ?] | 11:00 - 12:30 |?

Thursday:- 12:00 to 1:30 or 1:00 - 2:30 ?] ∈

Any other day? With at least a days gap

Note:- Keep slot H free.

Last 4 weeks:- student presentations
↳ 1 for each.

~~Course content~~

~~Course content~~.

~~Basic~~.

References:-

Deep Learning. Ian Goodfellow, Yoshua Bengio & Aaron Courville.

↳ Available online (Html).

Background:- (I) First 5 chapters should know. (First Assignment based on these).

(II) ML:- Neural Networks / Perceptron / Back Propagation

↳ Read notes online or COL774.

Other References:

(I) Lots of online resources

↳ Youtube Lectures / Blogs

↳ Nando De Freitas (Oxford) Deep Learning Course

(II) Blog :- Machine Learning Mastery by Jason Brownlee

(III) Will post additional references.

(3)

Piazza Page :-

① Useful for discussions

② COL 865: Special Topics in Computer Applications
:- Deep Learning

Code:- col865

③ All announcements on Piazza

Attendance Policy :-
As per Institute norms

~~Course Content~~ :-

Course Objective:- To get basic conceptual understanding as well as hands on experience with emerging area of deep learning

⇒ includes mathematical sophistication

↳ Understand Foundational

Assumed to be known:
- Concepts
  - ↳ Basic Mathematical background (10th step):- Linear algebra, calculus, numerical computation, (matrix), Machine Learning

↳ Feed Forward Networks

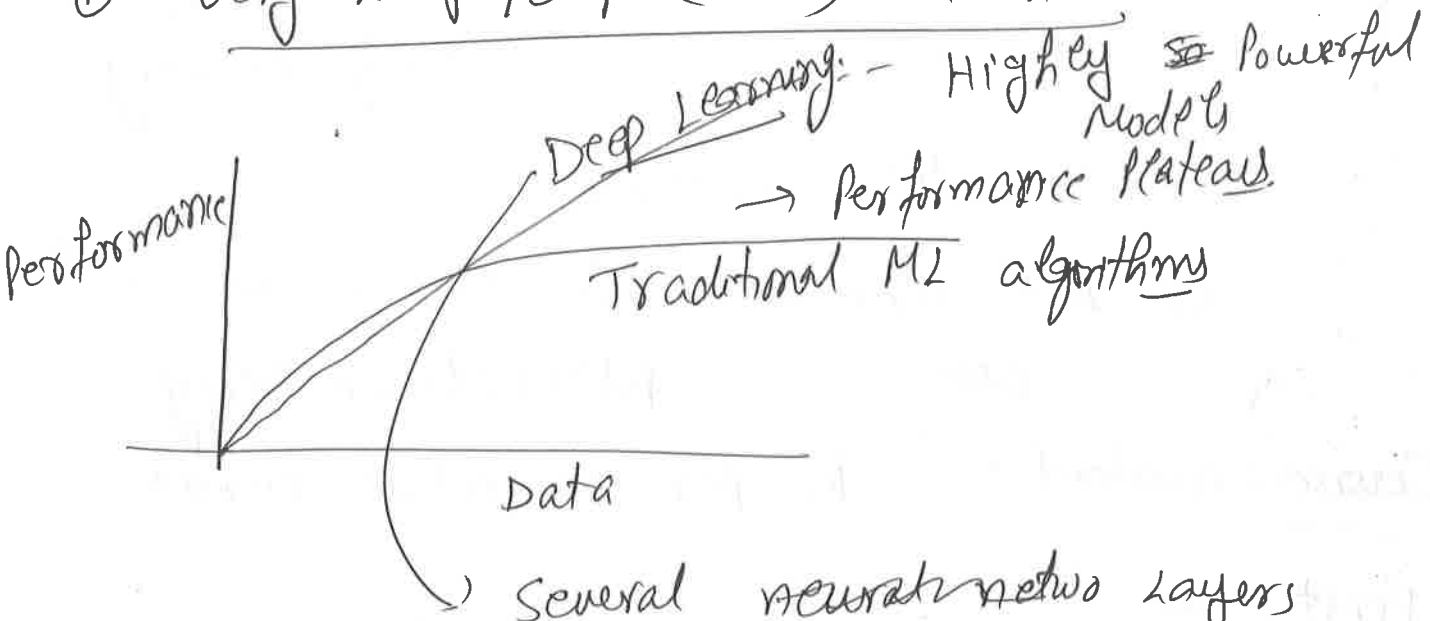
(Deep) → Neural Networks with multiple layers

↳

What is Deep Learning? Ask

Multiple Different Perspectives :-

① Very Large / Deep (Neural) Networks :-



↳ Several neural network layers on the architecture.

Highly complicated models :-

An extremely large number of parameters (prone to overfitting).

How to make them work?

↳ Availability of GPUs (Computations)

↳ Large Amounts of Data (Millions of examples)

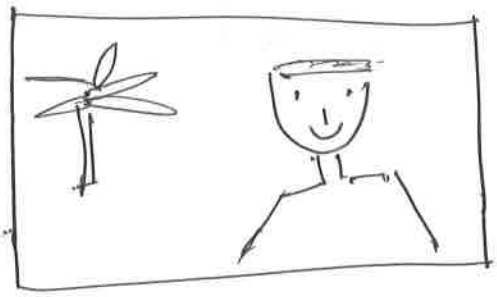
↳ Improved Models:
- ↳ Regularization
- ↳ ~~Better Conceptual understanding~~
- ↳ Optimization (improved) is escaping local minima
- ↳ Initialization
- ↳ Activation Function (Non-Linearity)

Feed Forward Neural Networks ← Chapters 6, 7, 8 in book

④ Second Perspetive:-

Deep Learning Models as Feature Extractors:-

①



(spatial structure)

⇒ Recognize objects:-
generate a hierarchy of features (automatically)
lowest level (Pixels) to highest level (object)

CNN:- Convolution Neural Networks
(Vision)

② & Handling sequential Data:- (Temporal structures)

In 2009, he was still a student.

He was still a student in 2009.

RNN:- Recurrent Neural Network

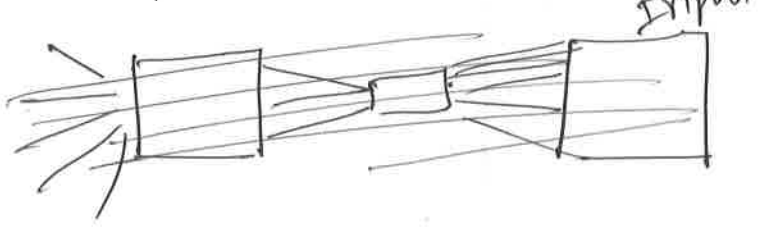
LSTM:- Long Short-Term Memory

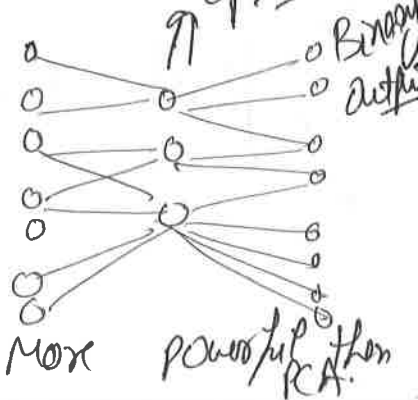
③ Generative Models:- (Autoencoders)

→ unsupervised

(Dimensionality Reduction)

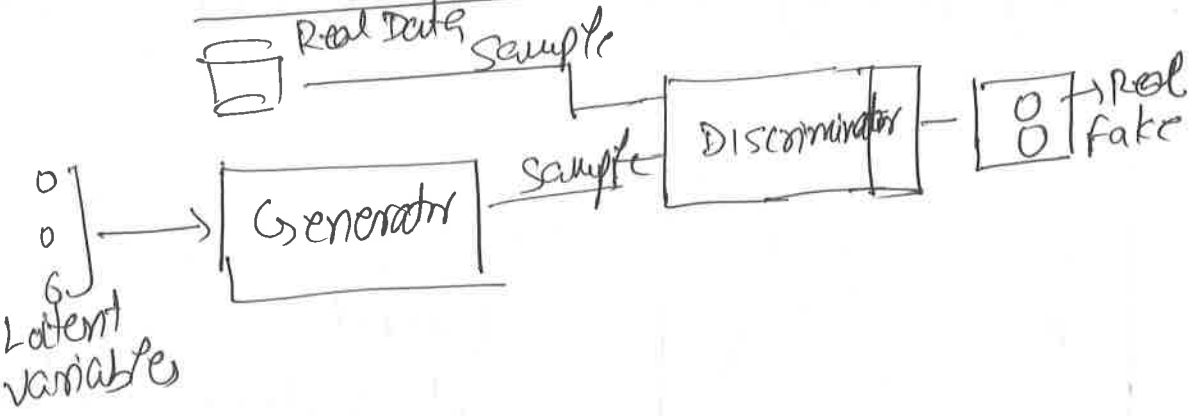
Encoding of Input





④ ③ GANs:-

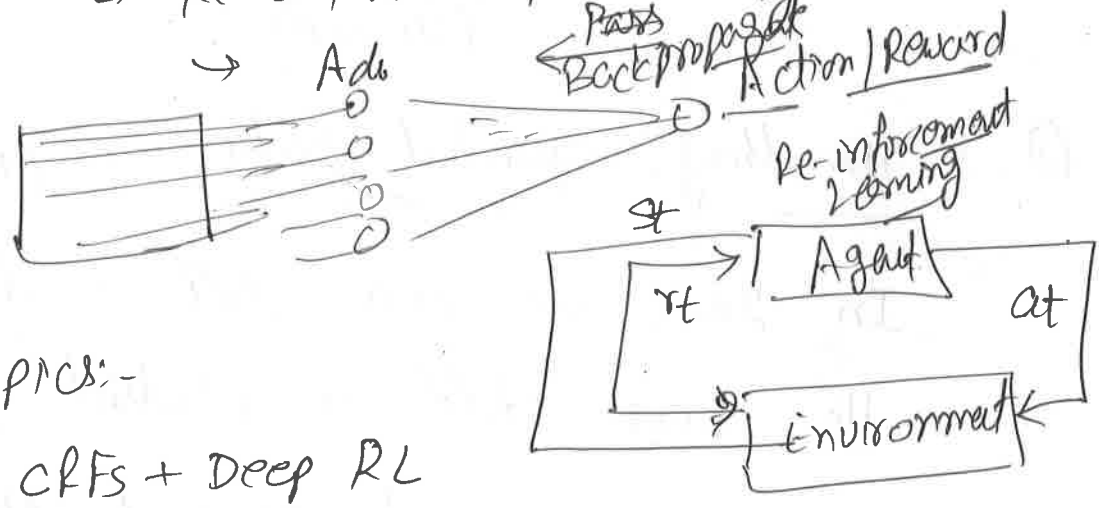
Generative Adversarial Networks (GANs):-



⑤ ~~②~~

⑤ Deep Re-inforcement Learning :-

~~PE~~ ↳ Re-inforcement Learning



5 Weeks - 7 Weeks

Advanced Topics:-

① CRFs + Deep RL
② Sum Product Networks
③ Learning Sparse Networks
④ Variational Autoencoders
⑤ Discriminative Adversarial Networks

Last 4 weeks

Implementation:-

Introduction

TensorFlow	Google Brain, 2015 (rewritten DistBelief)
Theano	University of Montréal, 2009
Keras	François Chollet, 2015 (now at Google)
Torch	Facebook AI Research, Twitter, Google DeepMind
Caffe	Berkeley Vision and Learning Center (BVLC), 2013