AI5100: Deep Learning /
CS5370: Deep Learning for Vision
Jan-Apr 2021

Instructor:
- Vineeth N Balasubramanian (vineethnb@iith.ac.in)

TAs:
- TBA

CONTENT:
Deep learning is a sub-area of machine learning, considered to be the reincarnation of neural networks. However, with the increased availability of vast amounts of data and compute capability, it has evolved to a field of its own in the last few years with numerous applications in computer vision, speech understanding and natural language processing. While the fundamental methods to train deep neural networks has remained the same for three decades now, a wide variety of architectures for different application settings, and a large group of methods for training them more effectively and efficiently have emerged over these years. This course will describe all these various facets of deep learning, including its applications to different domains. Like machine learning, deep learning has mathematical foundations, and this course will cover these foundations as and when possible/required. This course will also have a focus on deep learning for computer vision, although the methods themselves can be used for other kinds of data such as speech or text too. Students can expect to achieve the following objectives at the end of the course:
- Be familiar with the fundamentals of deep learning methods
- Apply deep learning to image and video data.
- Be familiar with the programming frameworks commonly used for deep learning today
- Be able to apply the learned concepts and methods to a real-world problem
- Learn to appreciate the mathematical rigor behind these methods where possible

PRE-REQUISITES:
- Prior completion of Machine Learning course at the level of CS6510/CS5590 (or equivalent)
- Familiarity with basic Probability Theory, Linear Algebra, Calculus (especially Matrix Calculus)
- Programming proficiency in Python

.
COURSE PORTAL:
We will use Google Classroom for sharing materials and discussions in the course. The Classroom link will be shared with the students registered for the course.

COURSE MODE AND SCHEDULE (TENTATIVE):

Considering the online nature of the semester, this course will be taught through video recordings shared at the beginning of each week, followed by a live online discussion every Saturday morning from 10:00 - 11:30 AM. Students are welcome to join these live sessions to clarify any doubts, questions or have discussions. The buffer weeks are exam weeks in each segment and can be used for catching up on previously missed content.

Week	Topics -- Lecture Videos (V) -- Slides (S)		
Jan 11 - Jan 15	Course Introduction -- V1 -- S1 History -- V2 -- S2 Image Formation -- V3 -- S3 Image Representation -- V4 -- S4 Linear Filtering, Correlation, Convolution -- V5 -- S5 Image in Frequency Domain -- V6 -- S6 Image Sampling -- V7 -- S7	A0	
Jan 18 - Jan 22	Edge Detection -- V8 -- S8 From Edges to Blobs and Corners -- V9 -- S9 Scale Space, Image Pyramids and Filter Banks -- V10 -- S10 SIFT and Variants -- V11 -- S11 Image Segmentation -- V12 -- S12 Other Feature Spaces -- V13 -- S13 Human Visual System [OPTIONAL] -- V14 -- S14	A1 assigned on Jan 23	
Jan 25 - Jan 29	Feature Matching -- V15 -- S15 Hough Transform -- V16 -- S16 From Points to Images: Bag-of-Words and VLAD Representations -- V17 -- S17 Image Descriptor Matching -- V18 -- S18 Pyramid Matching -- V19 -- S19 From Traditional Vision to Deep Learning -- V20 -- S20		Q1 on Jan 30
Feb 1 - Feb 5	PyTorch Tutorial Week	A1 due on Feb 4; A2 assigned on Feb 5	
Feb 8 - Feb 12	Neural Networks: A Review -- V21 (part 1) -- V22 (part 2) -- Slides Feedforward Neural Networks and Backpropagation -- V23 (part 1) -- V24 (part 2) -- Slides Gradient Descent and Variants -- V25 (part 1) -- V26 (part 2) -- Slides		Q2 on Feb 13

	Regularization in Neural Networks -- V27 (part 1) -- V28 (part 2) -- Slides Improving Training of Neural Networks -- V29 (part 1) -- V30 (part 2) -- Slides		
Feb 15 - Feb 19	Buffer week		
Feb 22 - Feb 26	Convolutional Neural Networks: An Introduction -- V31 -- S32 -- Slides Backpropagation in CNNs -- V33 -- Slides Evolution of CNN Architectures for Image Classification -- V34 -- V35 -- Slides Recent CNN Architectures -- V36 -- Slides Finetuning in CNNs -- V37 -- Slides	A2 due on Feb 24; A3 assigned on Feb ~~25~~27	
Mar 1 - Mar 5	Explaining CNNs: Visualization Methods -- V38 -- Slides Explaining CNNs: Early Methods -- V39 -- Slides Explaining CNNs: Class Attribution Map Methods -- V40 -- Slides Explaining CNNs: Recent Methods -- V41 -- V42 -- Slides Going Beyond Explaining CNNs [OPTIONAL] -- V43 -- Slides		Q3 on Mar 6
Mar 8 - Mar 12	CNNs for Object Detection-I -- V44 -- V45 -- Slides CNNs for Object Detection-II -- V46 -- Slides CNNs for Segmentation -- V47 -- Slides CNNs for Human Understanding: Faces -- V48 -- V49 -- Slides CNNs for Human Understanding: Human Pose and Crowd -- V50 -- Slides CNNs for Other Image Tasks -- V51 -- Slides	A3 due on Mar ~~9~~11; A4 assigned on Mar ~~10~~11	
Mar 15 - Mar 19	Recurrent Neural Networks: Introduction -- V52 -- Slides Backpropagation in RNNs -- V53 -- Slides LSTMs and GRUs -- V54 -- Slides Video Understanding using CNNs and RNNs -- V55 -- Slides		Q4 on Mar 20
Mar 22 - Mar	Buffer Week		

26			
Mar 28 - Apr 1	Attention in Vision Models: An Introduction -- V56 -- Slides Vision and Language: Image Captioning -- V57 -- Slides **[OPTIONAL]** Beyond Captioning: Visual QA, Visual Dialog -- V58 -- Slides **[OPTIONAL]** Other Attention Models -- V59 -- Slides Self-Attention and Transformers -- V60 -- Slides	A4 due on Mar 30; A5 released on Mar 31	
Apr 4 - Apr 8	Deep Generative Models: An Introduction -- V61 -- Slides Generative Adversarial Networks -- V62 -- V63 -- Slides Variational Autoencoders -- V64 -- Slides **[OPTIONAL]** Combining VAEs and GANs -- V65 -- Slides **[OPTIONAL]** Beyond VAEs and GANs: Other Deep Generative Models -- V66 -- V67 -- Slides		Q5 on Apr 10
Apr 11 - Apr 15	GAN Improvements -- V68 -- Slides **[OPTIONAL]** Deep Generative Models across Multiple Domains -- V69 -- Slides **[OPTIONAL]** VAEs and Disentanglement -- V70 -- Slides Deep Generative Models: Image Applications -- V71 -- Slides **[OPTIONAL]** Deep Generative Models: Video Applications -- V72 -- Slides	A5 due on Apr 12; A6 released on Apr 13	
Apr 18 - Apr 22	Few-shot and Zero-shot Learning -- V73 -- V74 -- Slides Self-Supervised Learning -- V75 -- Slides Adversarial Robustness -- V76 -- Slides **[OPTIONAL]** Pruning and Model Compression -- V77 -- Slides **[OPTIONAL]** Neural Architecture Search -- V78 -- Slides Course Conclusion -- V79 -- Slides		Q6 on Apr 24
Apr 25 - Apr 29	Buffer Week	A6 due on Apr 26	

GRADING CRITERIA:

(If you choose to not write the final exam and want to consider best 3 of 6 assignments -- DEFAULT CHOICE)

- 20%: Quizzes (Best 4 of 6)
- 80%: Homework/Assignments

(If you choose to not write the final exam and want to consider best 4 of 6 assignments)

- 15%: Quizzes (Best 4 of 6)
- 85%: Homework/Assignments

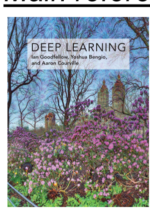
(If you choose to write the final exam)

- 15%: Quizzes (Best 4 of 6)
- 70%: Homework/Assignments (Best 3 of 6)
- 15%: Exam

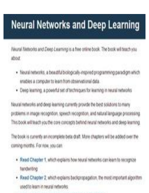
REFERENCES:

Deep learning is a rapidly evolving field, and we will hence use multiple sources of references, including books, blogs and articles, each of which will be pointed out at the end of each topic. However, there are three recent books that ground a lot of the fundamentals. In particular, the book by Goodfellow, Bengio and Courville is highly recommended, not only for the quality of its discussions, but also given that it has widest coverage of topics. It is also the most up-to-date and will be followed in most of the lectures.

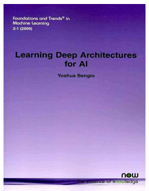
Main references:



Deep Learning By Ian Goodfellow and Yoshua Bengio and Aaron Courville, *MIT Press, 2016*



Neural Networks and Deep Learning, By Michael Nielsen, *Online book, 2016*



Learning Deep Architectures for AI (slightly dated) By Yoshua Bengio, *NOW Publishers, 2009*

...and lots of research papers, blogs, web posts

Tools

We will use PyTorch for our assignments just for ease and uniformity of evaluation.

Other useful references:

- Bishop, Christopher. *Neural Networks for Pattern Recognition*. New York, NY: Oxford University Press, 1995. ISBN: 9780198538646.
- Bishop, Christopher M. *Pattern Recognition and Machine Learning*. Springer, 2006. ISBN 978-0-387-31073-2
- Duda, Richard, Peter Hart, and David Stork. *Pattern Classification*. 2nd ed. New York, NY: Wiley-Interscience, 2000. ISBN: 9780471056690.
- Mitchell, Tom. *Machine Learning*. New York, NY: McGraw-Hill, 1997. ISBN: 9780070428072.