

Outline

- Course Logistics
- Course Overview

Outline

- Course Logistics
- Course Overview

Course Info

- Instructor: Gautam Kamath (gckamath@uwaterloo.ca)
 - Office hours: Tuesday and Friday at 2 PM Waterloo time (or by appointment)
- TAs: Theo Hu, Stan Ivashkevich, Udhav Sethi, Augustus Turnbull, Ruixue Zhang
- **Website:** <http://www.gautamkamath.com/courses/CS480-wi2021.html>
 - slides, notes, assignments, policy, etc.
- **Piazza:** piazza.com/uwaterloo.ca/winter2021/cs480680/
 - Announcements, questions, discussions, etc.
- **Learn:** <https://learn.uwaterloo.ca>
 - Assignments, solutions, grades, etc.
- **Teams**
 - Office Hours

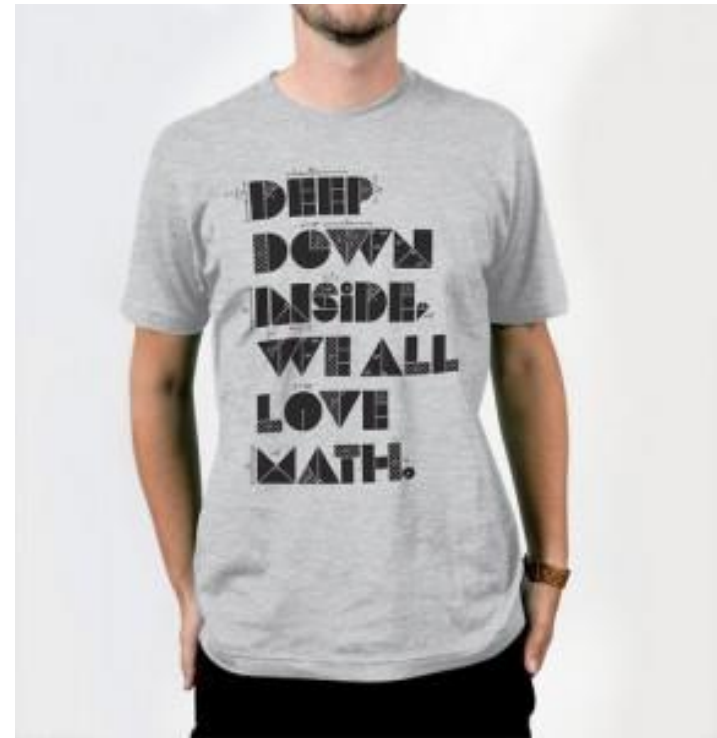
Prerequisites

- CM 339/CS 341 or SE 240; STAT 206 or 231 or 241
 - Basic probability, statistics, linear algebra, calculus
- Programming



<https://www.python.org/>

- Mathematical maturity

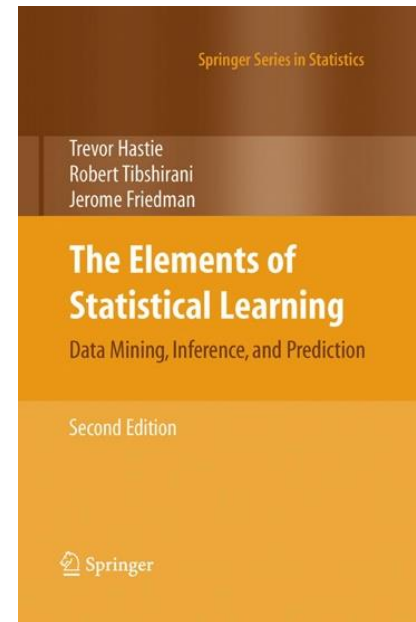
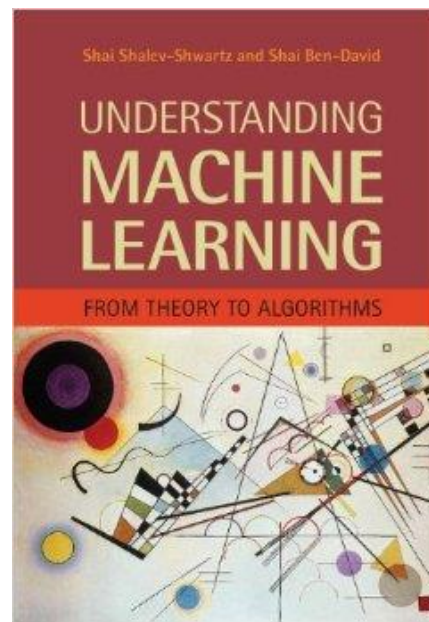
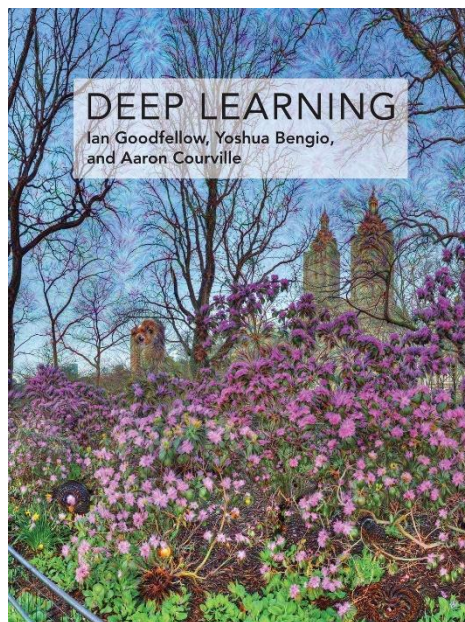
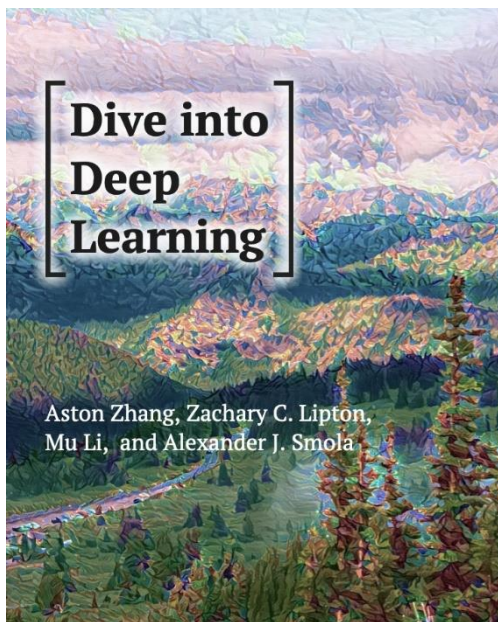


What to expect



Textbook

- **No** required textbook
- Notes, slides & videos will be posted on course web
- Some fine textbooks:

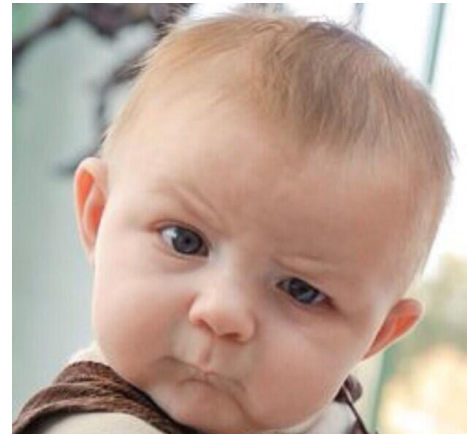


Workload

- Roughly 25 lectures, each lasting about an hour
- Expect roughly 5 assignments, approx. 1 every other week
 - 20 points each
 - Only top 4 will be calculated into your grade
- Submit on **LEARN**. **Submit early and often.**
 - Typeset using LaTeX is recommended

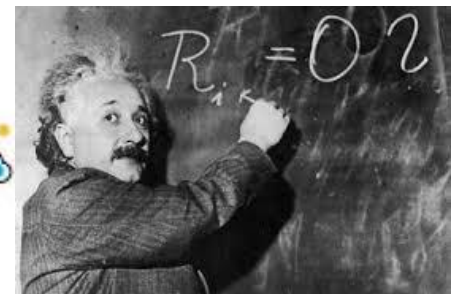
Policy

- **Do your own work** *independently and individually*
 - Discussion is fine, but no sharing of text or code
 - **Explicitly acknowledge** any source that helped you
- Ignorance is no excuse!
 - Good online discussion, more on course website
- Serious offence will result in expulsion...
- **NO late submissions!**
 - Except hospitalization, family emergency, ...
- Appeal within two weeks



Remaining 20% grade

- Kaggle competition for CS480
- Individual or pair project for CS680
 - Relate to machine learning (obviously)
 - Allow you to learn something new (and hopefully significant)
 - Be **interesting** and **nontrivial**
- 2-4 pages proposal and ≤ 8 pages report



Enrollment

- If you already enrolled
 - Good for you!
- If you are not enrolled yet
 - Permission numbers will be sent later this week

Questions?



Outline

- Course Logistics
- Course Overview

What is Machine Learning (ML)?

- “Machine learning is the field of study that gives computers the ability to learn without being explicitly programmed.” --- Arthur Samuel (1959).



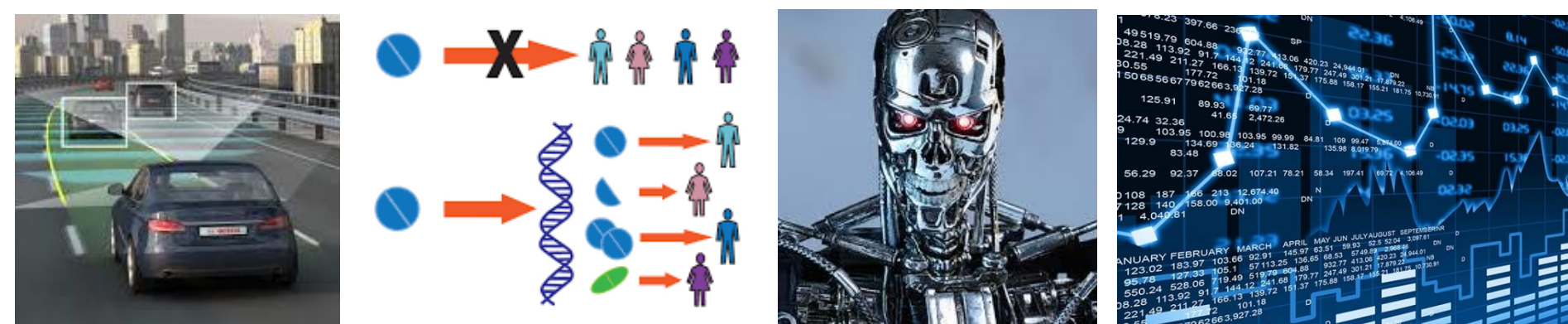
- “A computer program is said to learn from **experience E** with respect to **some class of tasks T** and **performance measure P**, if its performance at tasks in T, as measured by P, **improves** with experience E.” --- Tom Mitchell (1998)

Why is ML important for YOU?

- First off, you use ML **everyday**



- Lots of cool applications



- Excellent for job-hunting

Learning Categories

Supervised



- Classification
- Regression
- Ranking

Teacher provides answer

Reinforcement



- Control
- Pricing
- Gaming

Teacher provides motivation

Unsupervised



- Clustering
- Visualization
- Representation
- Density Estimation

Surprise, surprise

Supervised learning

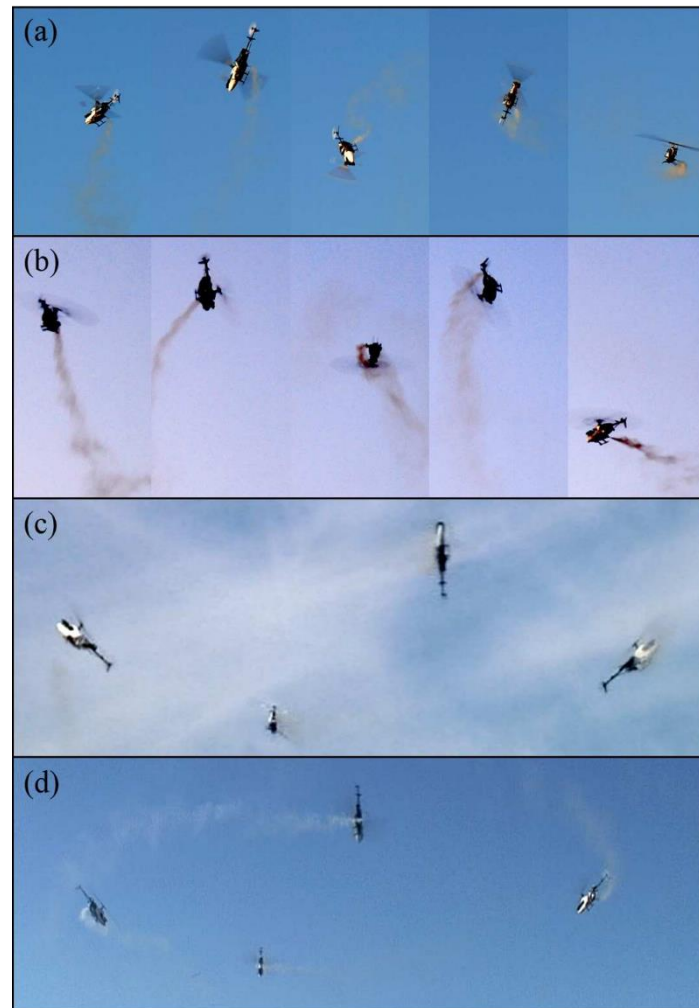


Reinforcement learning

- Not in this course...
- CS486/686; CS885



Silver et al., Nature'16



Abbeel et al., NIPS'06



Unsupervised learning



Context (human-written): In a shocking finding, scientist discovered a herd of unicorns living in a remote, previously unexplored valley, in the Andes Mountains. Even more surprising to the researchers was the fact that the unicorns spoke perfect English.

GPT-2: The scientist named the population, after their distinctive horn, Ovid's Unicorn. These four-horned, silver-white unicorns were previously unknown to science.

Now, after almost two centuries, the mystery of what sparked this odd phenomenon is finally solved.

Dr. Jorge Pérez, an evolutionary biologist from the University of La Paz, and several companions, were exploring the Andes Mountains when they found a small valley, with no other animals or humans. Pérez noticed that the valley had what appeared to be a natural fountain, surrounded by two peaks of rock and silver snow.

Pérez and the others then ventured further into the valley. "By the time we reached the top of one peak, the water looked blue, with some crystals on top," said Pérez.

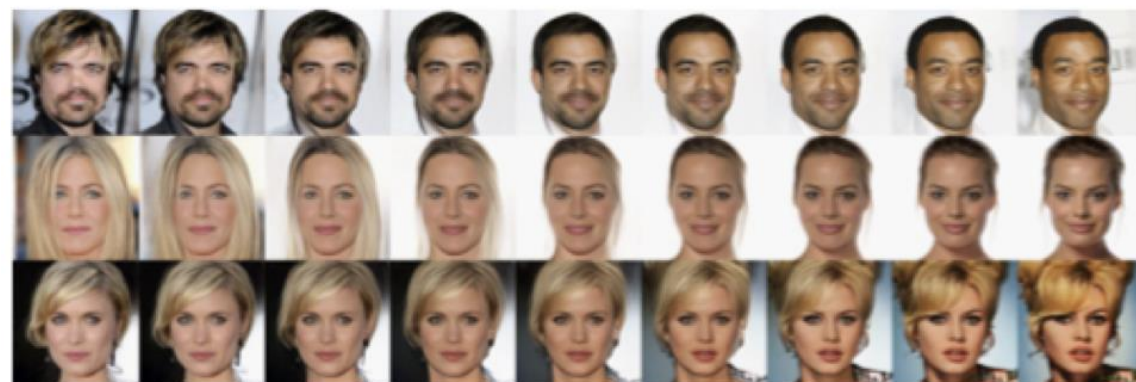
Pérez and his friends were astonished to see the unicorn herd. These creatures could be seen from the air without having to move too much to see them – they were so close they could touch their horns.

While examining these bizarre creatures the scientists discovered that the creatures also spoke some fairly regular English. Pérez stated, "We can see, for example, that they have a common 'language,' something like a dialect or dialectic."

Dr. Pérez believes that the unicorns may have originated in Argentina, where the animals were believed to be descendants of a lost race of people who lived there before the arrival of humans in those parts of South America.

While their origins are still unclear, some believe that perhaps the creatures were created when a human and a unicorn met each other in a time before human civilization. According to Pérez, "In South America, such incidents seem to be quite common."

However, Pérez also pointed out that it is likely that the only way of knowing for sure if unicorns are indeed the descendants of a lost alien race is through DNA. "But they seem to be able to communicate in English quite well, which I believe is a sign of evolution, or at least a change in social organization," said the scientist.



Kingma et al., NeurIPS'18

Radford et al., 2019



Focus of ML research

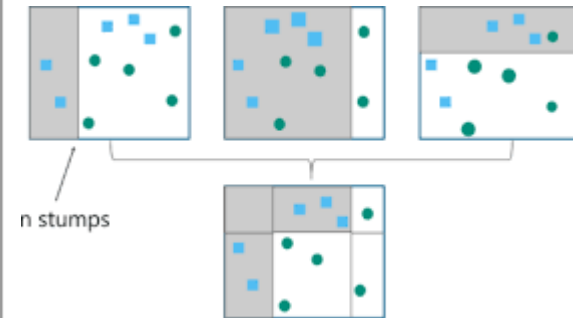
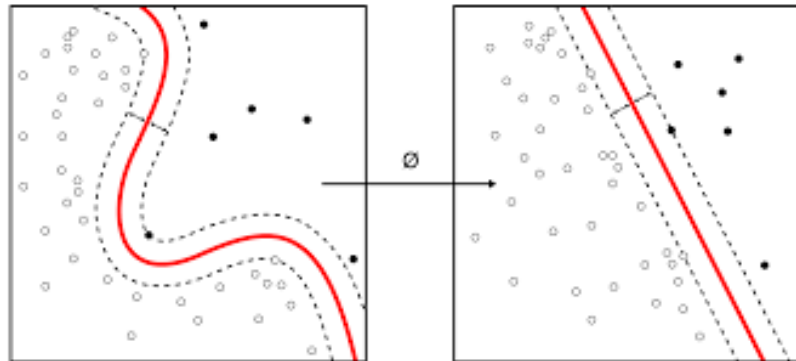
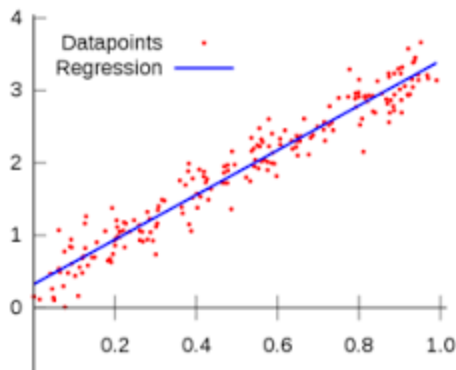
- Representation and Interpretation
 - How to represent the data? How to interpret result?
- Generalization
 - How well can we do on test data? On a different domain?
- Complexity
 - How much time and space?
- Efficiency
 - How many samples?
- Applications



	Date	Topic	Slides	Notes	Videos	Assignments
00	May 12, 2020	Introduction	pdf , ppt		mp4	
01	May 12, 2020	Perceptron	pdf , ppt	pdf	mp4	pdf , tex
02	May 14, 2020	Linear Regression	pdf , ppt	pdf	mp4	
03	May 19, 2020	Optimization Basics				
04	May 21, 2020	Statistics Basics				
05	May 26, 2020	K Nearest Neighbors				
06	May 28, 2020	Logistic Regression				
07	Jun 02, 2020	Hard-margin SVM				
08	Jun 04, 2020	Soft-margin SVM				
09	Jun 09, 2020	Reproducing Kernels				
10	Jun 11, 2020	Decision Trees				
11	Jun 16, 2020	Boosting				
12	Jun 18, 2020	Multilayer Perceptron				
13	Jun 23, 2020	Deep Networks				
14	Jun 25, 2020	Convolutional NNs				
15	Jun 30, 2020	Recurrent NNs				
16	Jul 02, 2020	Graph NNs				
17	Jul 07, 2020	Mixture Models				
18	Jul 09, 2020	Graphical Models				
19	Jul 14, 2020	GANs				
20	Jul 16, 2020	Flows				
21	Jul 21, 2020	Robustness				
22	Jul 23, 2020	Attention				
23	Jul 28, 2020	Learning to learn				
24	Jul 30, 2020	Interpretability				
25	Aug 04, 2020	Causality				

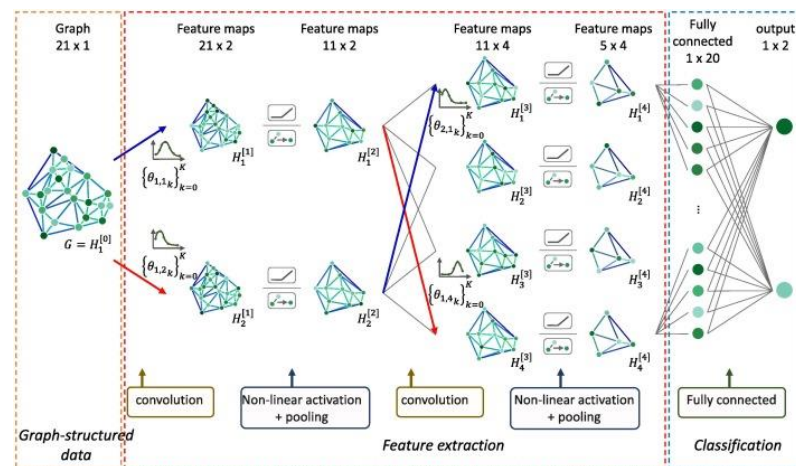
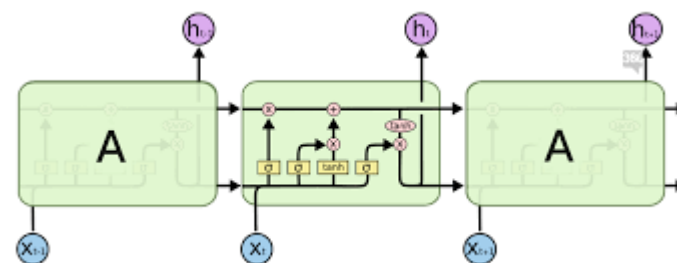
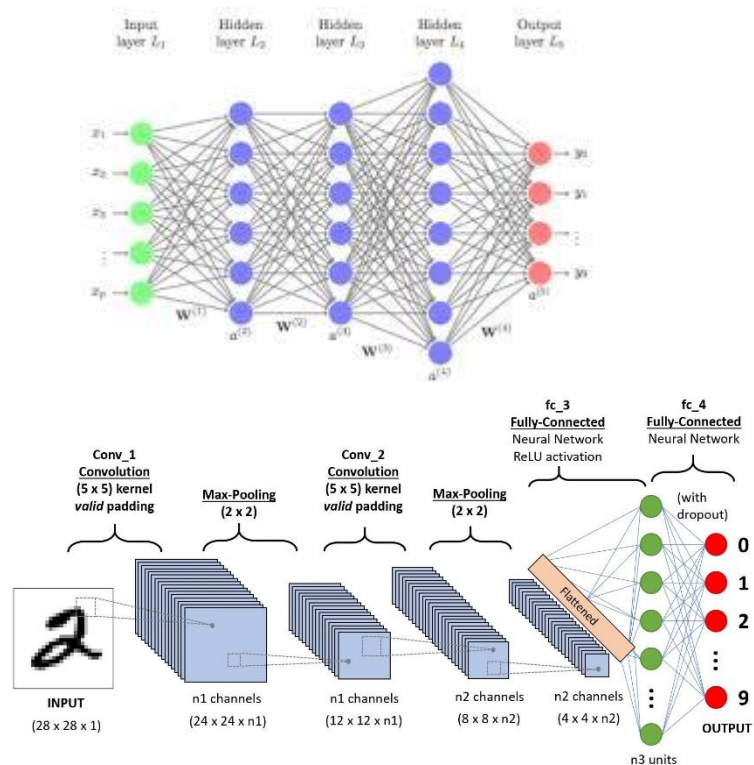
Classic

	Date	Topic	Slides	Notes	Videos	Assignments
00	May 12, 2020	Introduction	pdf , ppt		mp4	
01	May 12, 2020	Perceptron	pdf , ppt	pdf	mp4	pdf , tex
02	May 14, 2020	Linear Regression	pdf , ppt	pdf	mp4	
03	May 19, 2020	Optimization Basics				
04	May 21, 2020	Statistics Basics				
05	May 26, 2020	K Nearest Neighbors				
06	May 28, 2020	Logistic Regression				
07	Jun 02, 2020	Hard-margin SVM				
08	Jun 04, 2020	Soft-margin SVM				
09	Jun 09, 2020	Reproducing Kernels				
10	Jun 11, 2020	Decision Trees				
11	Jun 16, 2020	Boosting				



Neural Nets

12	Jun 18, 2020	Multilayer Perceptron				
13	Jun 23, 2020	Deep Networks				
14	Jun 25, 2020	Convolutional NNs				
15	Jun 30, 2020	Recurrent NNs				
16	Jul 02, 2020	Graph NNs				



Generative Models

17	Jul 07, 2020	Mixture Models				
18	Jul 09, 2020	Graphical Models				
19	Jul 14, 2020	GANs				
20	Jul 16, 2020	Flows				



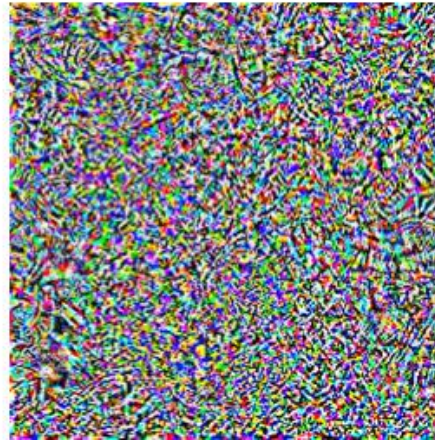
Exotic

21	Jul 21, 2020	Robustness				
22	Jul 23, 2020	Attention				
23	Jul 28, 2020	Learning to learn				
24	Jul 30, 2020	Interpretability				
25	Aug 04, 2020	Causality				

“pig”



+ 0.005 x



=

“airliner”



Questions?

