Artificial Neural Networks

AI/ML Overview

Samuel Cheng

With many slides borrowed from

Stanford CS 231n

Hinton's coursera course

General Information

- Instructor: Samuel Cheng
- Discussion forum (Discord) is available. Please check the signup link through Canvas
- Office hours:
  - None. Please just contact me through Discord
- Quick logistics:
  - Canvas (private) + Discord (semi-private) + website (public) + Zoom (video lecture)

Your Research Directions?

- Image Processing/Computer Vision/Imaging?
- Medical stuff?
- Natural language processing (NLP)?
- Others?

- HW 0: signup Discord and say a few words about projects you are working on

Deep learning courses online

- Hinton’s coursera course
- Stanford CS 231n
- Oxford’s course
- U Toronto’s course
- Larochelle’s course
- Lecun’s NYU course
- Fastai course
- Others?

What is this course about?

- Look into some history, techniques, applications of neural networks
- Will explore quite a bit on so-called ``deep learning". A hyped-word of artificial neural networks
  - Things advance extremely fast. We will explore the topics together
- Many students applied what was learned to their research directly last several years
- Hope this will help your research!

Goal of this course: expected outcomes

- Able to use open source packages (in particular PyTorch) to construct neural network models to solve your research problems
- Build a sufficiently strong theoretical foundation to understand current ANN research

Quick Logistics

- “Homework” ~ 30%
- Presentation ~ 20%
- Final Project ~ 30%
- Literature review essay ~ 20%
- Extra credit (in-class/Discord participation) ~ 10%
- Textbook: Ian Goodfellow, Yoshua Bengio and Aaron Courville, Deep Learning, MIT Press
  - I won’t follow it closely but it is a good reference in general
- Course website: http://www.samuelcheng.info/deeplearning_2021/index.html

Exposure to deep learning packages?

- PyTorch/Torch:
- Tensorflow:
- Caffe/Caffe2:
- Mxnet:
- Theano (obsolete):
- Keras:
- Lasagne (obsolete):
- Matconvnet:
- Others:

Programming Languages?

- Python/Numpy (required but quite easy)
- C/C++
- Lua
- Matlab
- Java
- Others
- None at all

Platforms?

- Linux
- Windows
- Mac
- GPUs?

Coverage (tentative)

- Overview of machine learning
- Overview of supervised learning
- Neural network basics
- Backpropagation algorithm
- Regularization
- Optimization methods
- **CNN**
- Weight visualization
- Recurrent neural networks
- *Deep learning/neural networks packages? (presented by you all)*
- Applications: computer vision, natural language processing
- ~~Restricted Boltzmann machine~~, autoencoder, deep belief networks
- Introduction to deep reinforcement learning
- Capsule networks
- Attention networks and transformers (GPT!)
- Self-supervised learning
- Graph networks

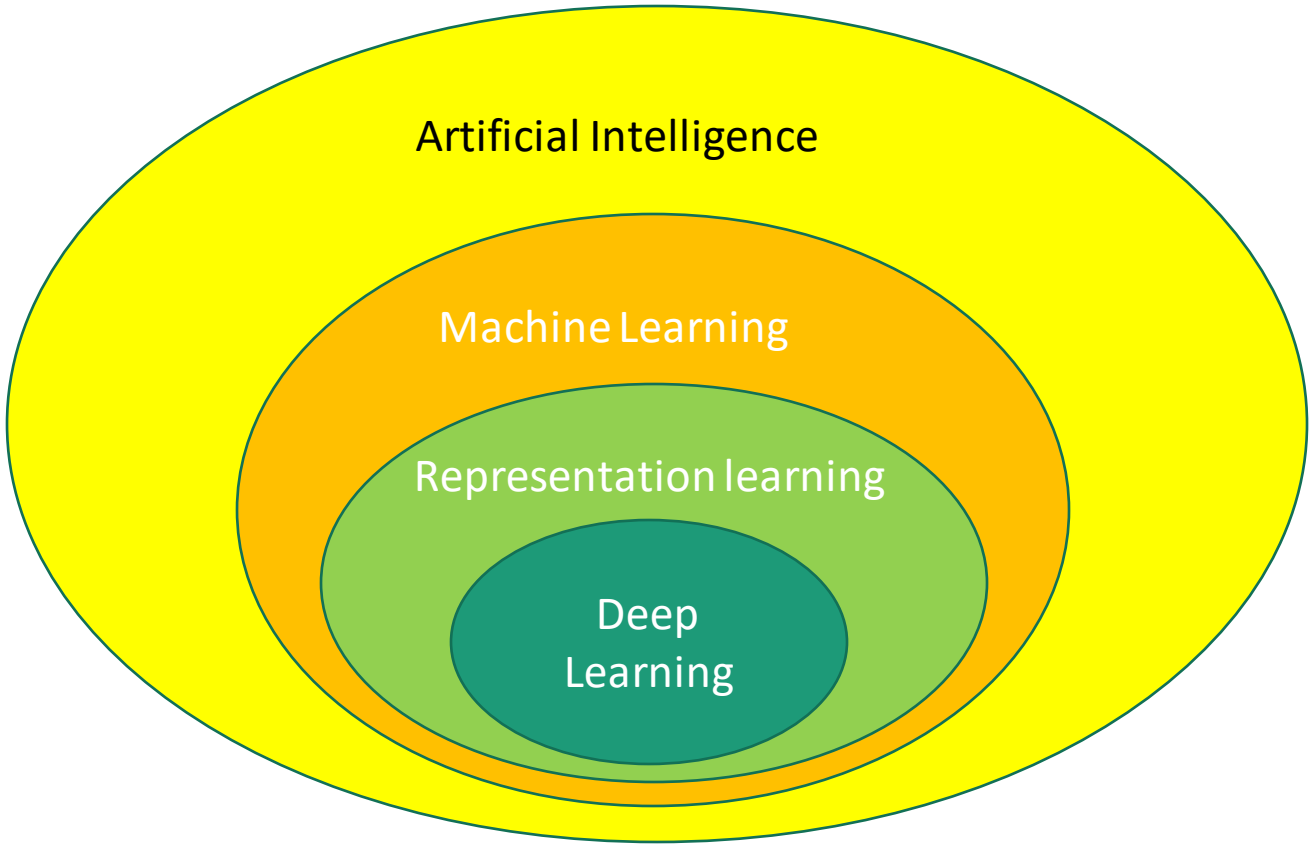
Prerequisite

- Python proficiency
  - Homework are all based on Python. Actually all common packages (Tensorflow, PyTorch, etc) have Python interfaces
  - If you have programmed before and are familiar with one of those high-level languages, you should be fine
  - Check out this for a quick tutorial: http://cs231n.github.io/python-numpy-tutorial/
- College calculus and linear algebra
  - Won't explore proof much but better math foundation definitely eases understanding of many topics

Deep learning in 1 slide

- Wide sense
  - (machine) **learning** goes **deep** (with layers of representation)
- Narrow sense
  - (Artificial) neural networks with more than one hidden layers
- Why so popular?
  - Probably the most powerful machine learning technique at this moment
  - Won machine learning competition across wide categories with large margins

AI Taxonomy


Artificial Intelligence
Machine Learning
Representation learning
Deep
Learning

AI

- Strong-AI (Artificial general intelligence): fully human-like
  - Consciousness?
  - Common sense?
  - Reasoning?
  - Check out this video podcast by Lex Fridman
  - *… not what we try to study here*
- Weak-AI
  - Trying to tackle some (very) specific tasks

AI Tasks

- Playing chess :
- Playing strategy games (e.g., Starcraft) :
- Playing GO :
- Solve Sudoku :
- Find a cat in a picture :
- Describe a picture :
- Solve a calculus problem :
- *Understand* a story:

AI Tasks

- Playing chess : medium->easy
- Playing strategy games (e.g., Starcraft) : hard->medium
- Playing GO : hard->easy
- Solve Sudoku : easy
- Find a cat in a picture : hard to achieve human performance
- Describe a picture : ditto
- Solve a calculus problem : easy
- *Understand* a story: very hard

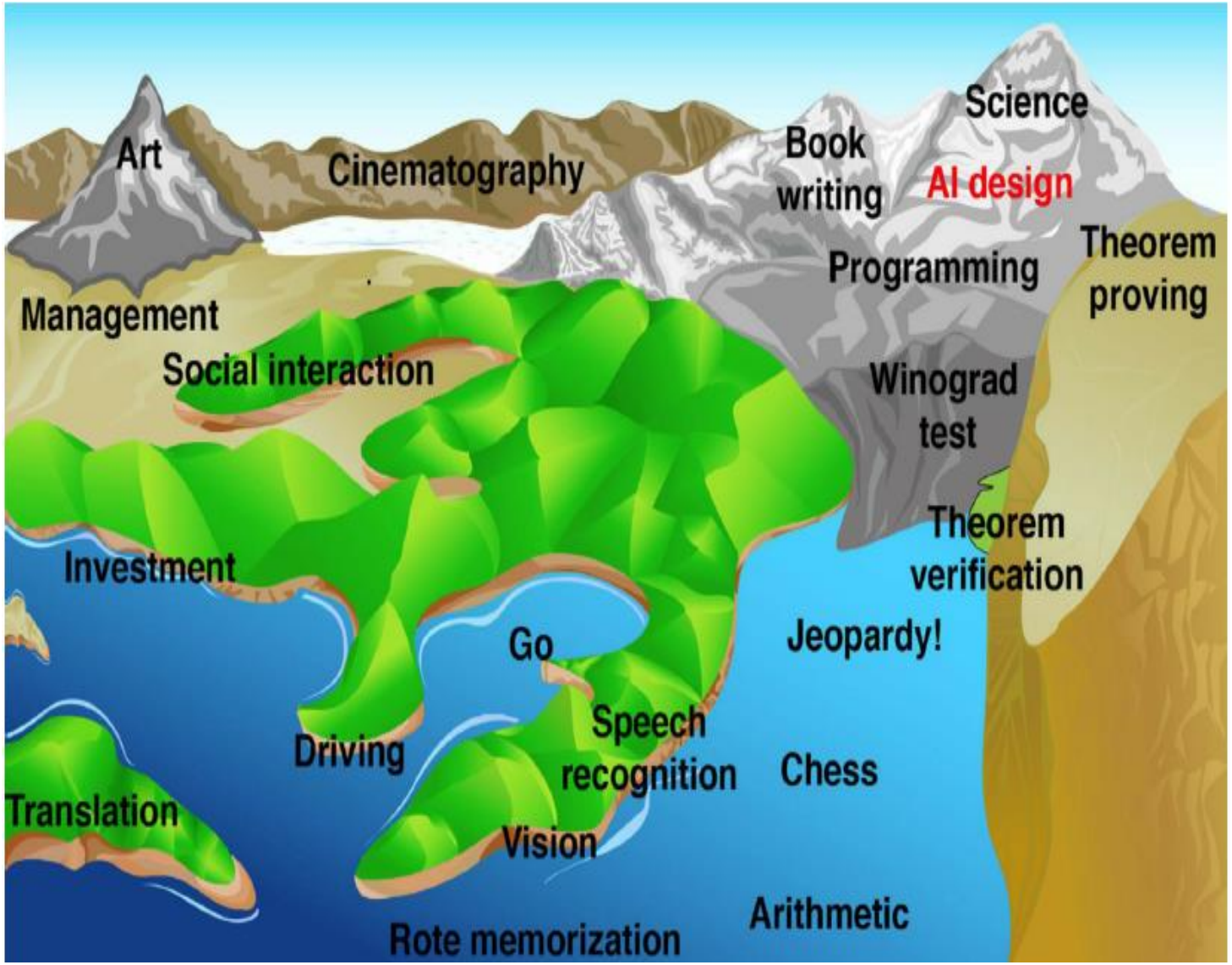
AI Tasks

Machine better

- Solve Sudoku : easy
- Solve a calculus problem : easy
- Playing chess : medium->easy
- Playing GO : hard->easy
- Find a cat in a picture (?) : hard to achieve human performance
- Describe a picture : ditto
- Playing strategy games (e.g., Starcraft) : hard
- *Understand* a story: very hard

Human better

Hans Moravec's landscape of Human competence



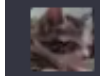
Slide credit to Max Tegmark

Winograd Schema Test

- Winograd Schema questions simply require the resolution of anaphora: the machine must identify the antecedent of an ambiguous pronoun in a statement
- Example
  - Given
    - The city councilmen refused the demonstrators a permit because they feared violence
    - The city councilmen refused the demonstrators a permit because they advocated violence
  - Question
    - Does the pronoun “they” refer to councilmen or demonstrators?

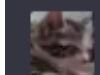
A bad Winograd question

- "The women stopped taking pills because they were [pregnant/carcinogenic]. Which individuals were [pregnant/carcinogenic]?"
- The question does not need reasoning
  - Only women but not pills can get pregnant
  - Only pills but not women can be carcinogenic
- Winograd questions need to be carefully constructed to ensure that common sense reasoning is required



The women stopped taking pills because they were pregnant. What are "they" refering to in the statement?

In the statement "The women stopped taking pills because they were pregnant. What are they here?" "they" is referring to the women who stopped taking pills because they were pregnant.



The women stopped taking pills because they were carcinogenic. What are "they" refering to in the statement?

In the statement "The women stopped taking pills because they were carcinogenic. What are they here?" "they" is referring to the pills that the women stopped taking because they were carcinogenic, meaning that they had a potential to cause cancer.

The city councilmen refused the demonstrators a permit because they feared violence. Who are "they" referring to here?

In the statement "The city councilmen refused the demonstrators a permit because they feared violence. Who are "they" referring to here? "They" is referring to the city councilmen who refused the demonstrators a permit because they feared that violence may occur.

The city councilmen refused the demonstrators a permit because they avocated violence. Who are "they" referring to here?

In the statement "The city councilmen refused the demonstrators a permit because they advocated violence. Who are "they" referring to here? "They" is referring to the demonstrators who the city councilmen believed advocated for violence, that's why the city councilmen refused them a permit.

Machine learning overview

Machine learning overview

- As the name suggests, have **machine learned** (from data)
- Machine learning is only a part of AI
  - Other such as planning and search
  - But machine learning is one key component (tool)

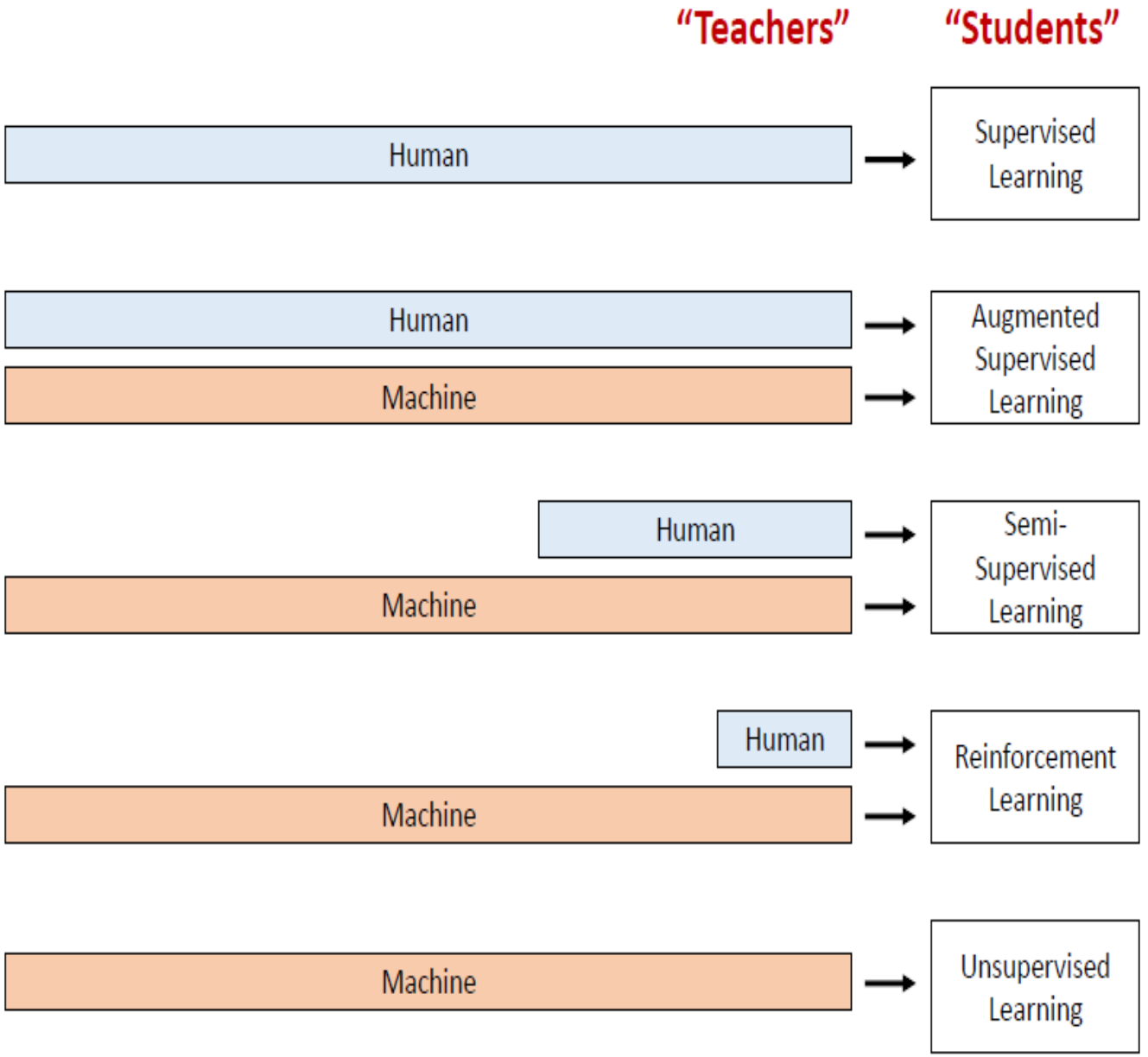
Why and when do we need machine learning?

- Program a machine to recognize a cat
  - We don't know how we manage to do it
  - So there is no clue how to program such machine
  - Even if we know how to do it, the program is probably extremely complicated

```
def predict(image):
    # ????
    return class_label
```

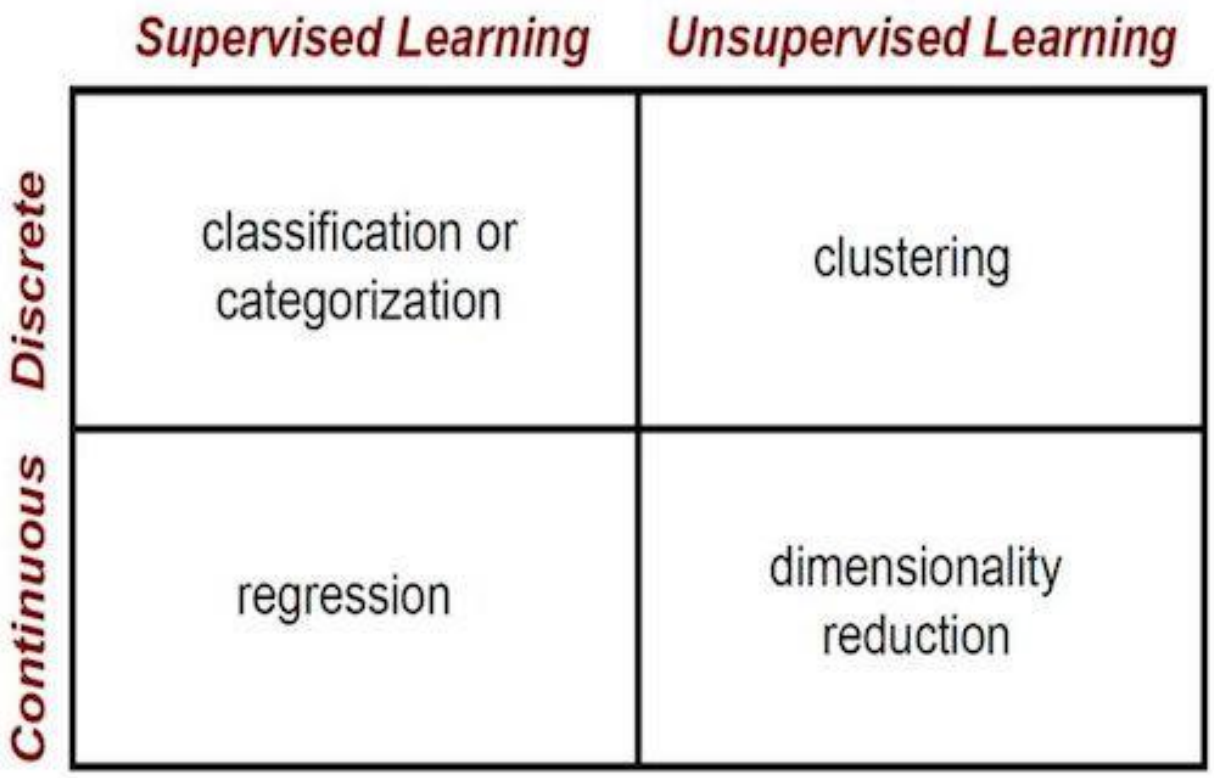
No obvious way to implement this function!

- Compute the probability of a credit card fraud
  - No one simple rule. Need to combine many weak rules
  - It is a moving target



deeplearning.mit.edu

Supervised vs unsupervised learning



https://towardsdatascience.com/supervised-vs-unsupervised-learning-14f68e32ea8d