Introduction to
MACHINE LEARNING

N. Rich Nguyen, PhD

CS 4774

Fall 2019

Machine Learning Meme

what society thinks I do

what my friends think I do

what my parents think I do

$$L_p = \frac{1}{2}\|\mathbf{w}\|^2 - \sum_{i=1}^{l} \alpha_i y_i (\mathbf{x}_i \cdot \mathbf{w} + b) + \sum_{i=1}^{l} \alpha_i$$

$$\alpha_i \geq 0, \forall i$$

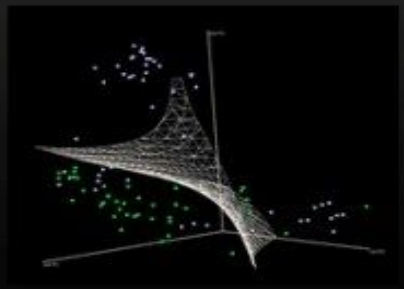
$$\mathbf{w} = \sum_{i=1}^{l} \alpha_i y_i \mathbf{x}_i, \sum_{i=1}^{l} \alpha_i y_i = 0$$

$$\nabla \hat{g}(\theta_t) = \frac{1}{m} \sum_{i=1}^{m} \nabla \ell(x_i, y_i; \theta_t) + \nabla r(\theta_t).$$

$$\theta_{t+1} = \theta_t - \eta_t \nabla \ell(x_{i(t)}, y_{i(t)}; \theta_t) - \eta_t \cdot \nabla r(\theta_t)$$

$$\mathbb{E}_{i(t)}[\ell(x_{i(t)}, y_{i(t)}; \theta_t)] = \frac{1}{n} \sum_i \ell(x_i, y_i; \theta_t).$$

what other programmers think I do



what I think I do

```
>>> from sklearn import svm
```

what I really do

Machine Learning Definition

"Machine Learning is the field of study that gives computers the ability to learn without being explicitly programmed."

-- Arthur Samuel, Gaming and AI Pioneer, 1959

"A computer program is said to learn from experience E with respect to some task T and some performance measure P, if its performance on T, as measured by P, improves with experience E."

-- Tom Mitchell, Computer Scientist, 1997

ML Difference

Traditional Programming

Machine Learning (ML)

ML Timeline

ARTIFICIAL INTELLIGENCE

Early artificial intelligence stirs excitement.

MACHINE LEARNING

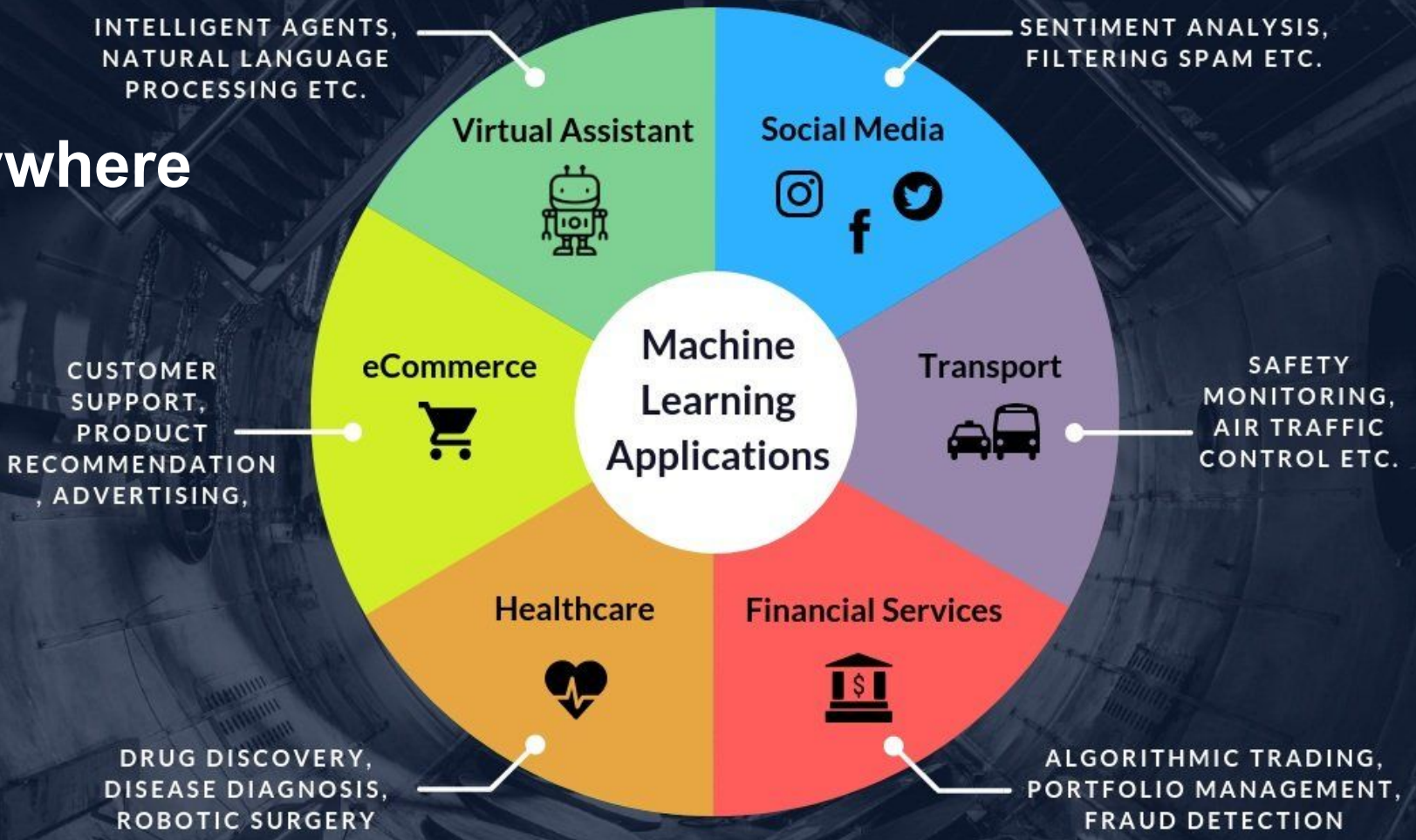
Machine learning begins to flourish.

DEEP LEARNING

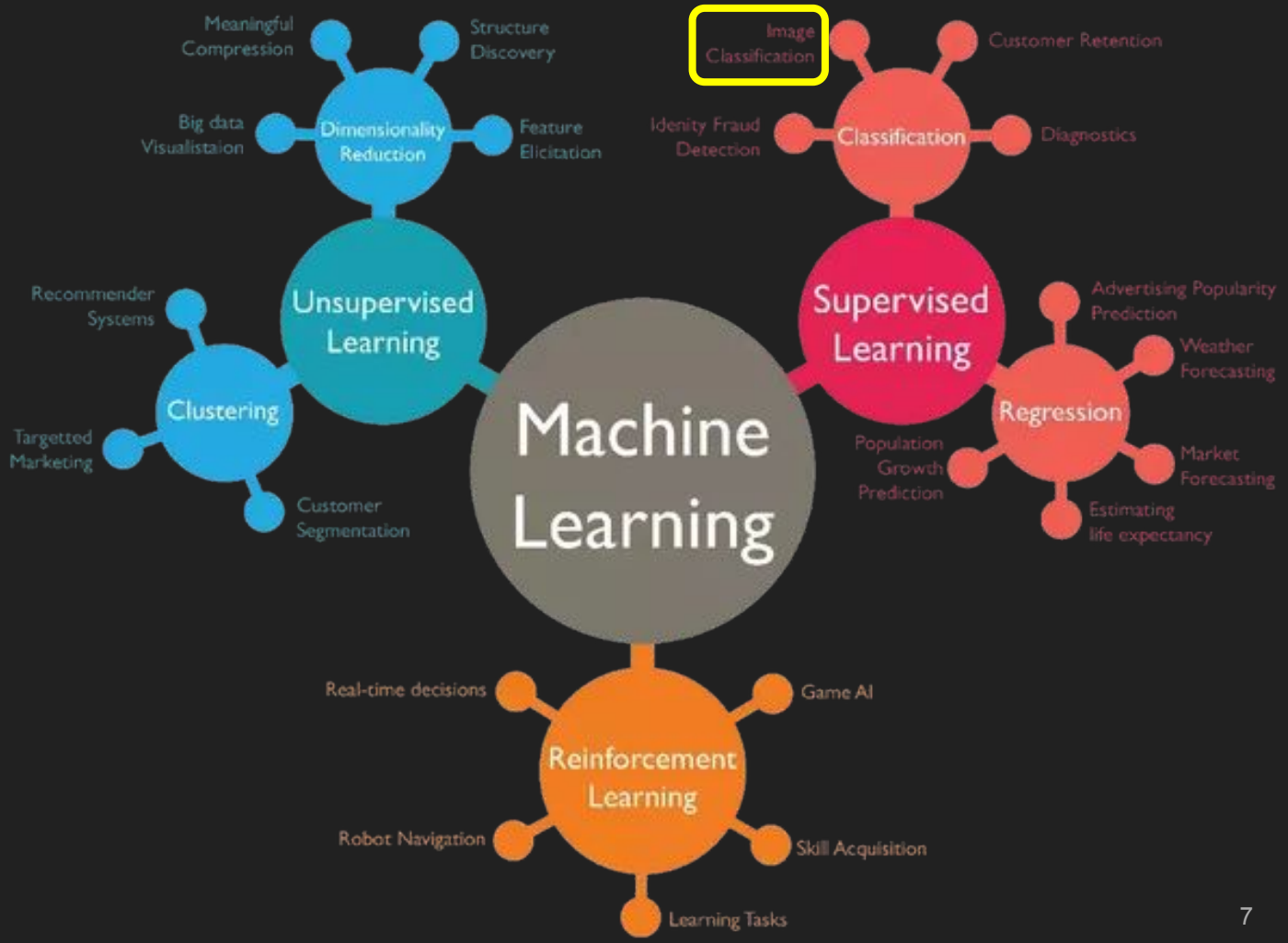
Deep learning breakthroughs drive AI boom.

1950's 1960's 1970's 1980's 1990's 2000's 2010's

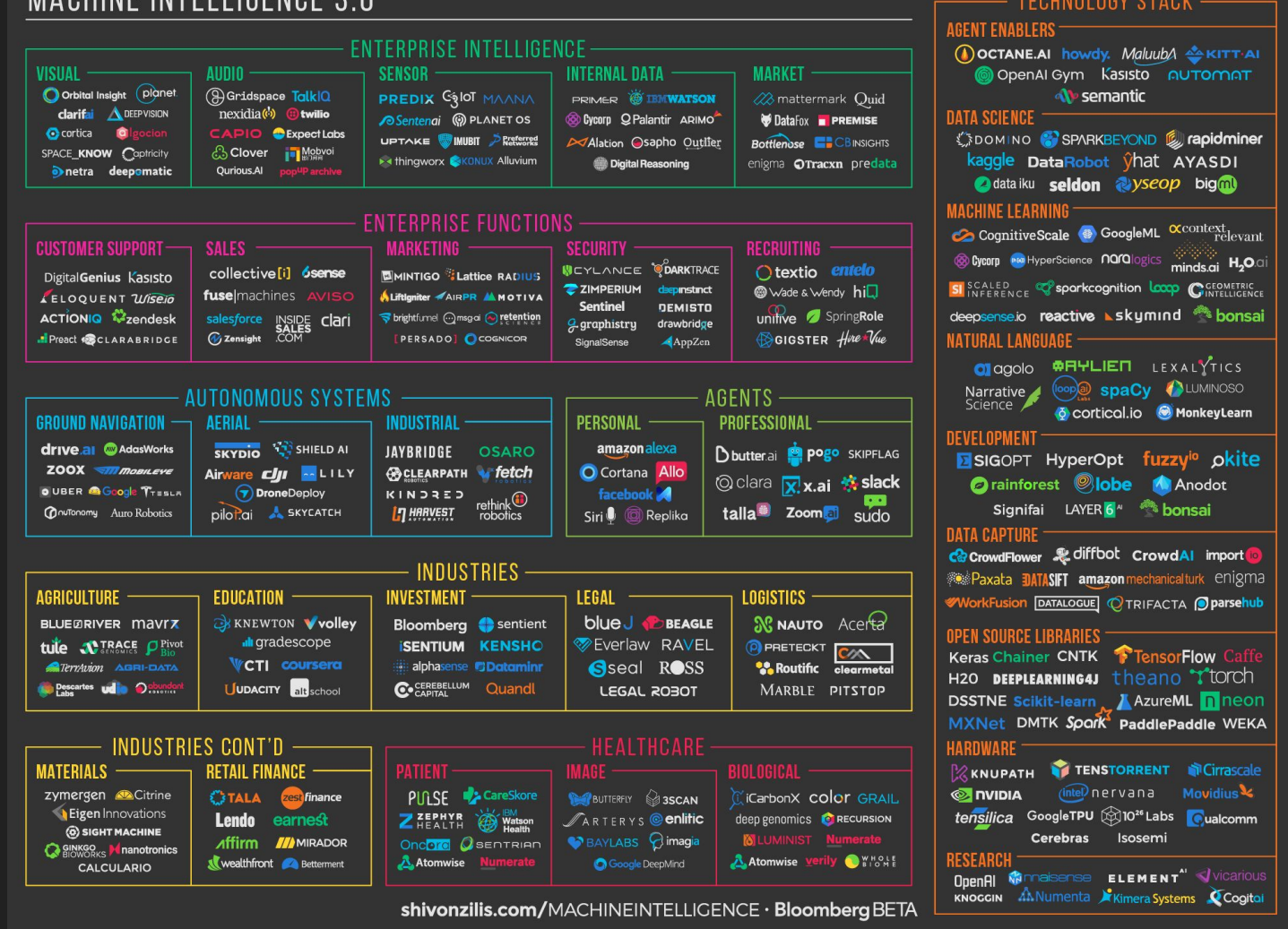
ML is everywhere



ML Taxonomy


Machine Learning
Unsupervised Learning
Dimensionality Reduction
Meaningful Compression
Structure Discovery
Big data Visualistaion
Feature Elicitation
Clustering
Recommender Systems
Targetted Marketing
Customer Segmentation
Supervised Learning
Classification
Image Classification
Customer Retention
Idenity Fraud Detection
Diagnostics
Regression
Advertising Popularity Prediction
Weather Forecasting
Population Growth Prediction
Market Forecasting
Estimating life expectancy
Reinforcement Learning
Real-time decisions
Game AI
Robot Navigation
Skill Acquisition
Learning Tasks

ML Landscape



Why do you study ML?

[your reason here]

It's one of the best careers for the 21st century

Gain abilities to solve large-scale problems

Obtain a data scientist position at a company

Curious to know how it works

CS 4774 ML
by the Numbers

1 instructor

Nhat Rich **Nguyen** **/:win/**, PhD
202 Rice Hall
OH: Tue,Wed,Thu 2-3p
nn4pj@virginia.edu
cs.virginia.edu/~nn4pj

7 TAs

Arjun, Akanksha : Masters' students
Layne, Clara, Johnny, Yuxin, Jeffrey: 4th year students

140 students

a good-size course

introduce yourself to your neighbors

1 pre-assessment

To help me determine the flow and pace of the course

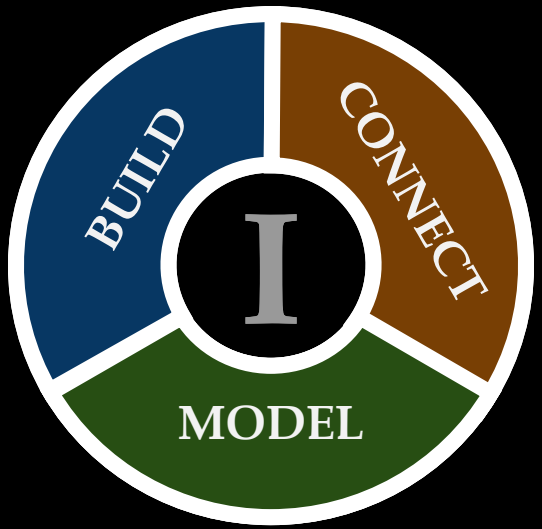
Take it at bit.ly/cs4774pre

29 meetings

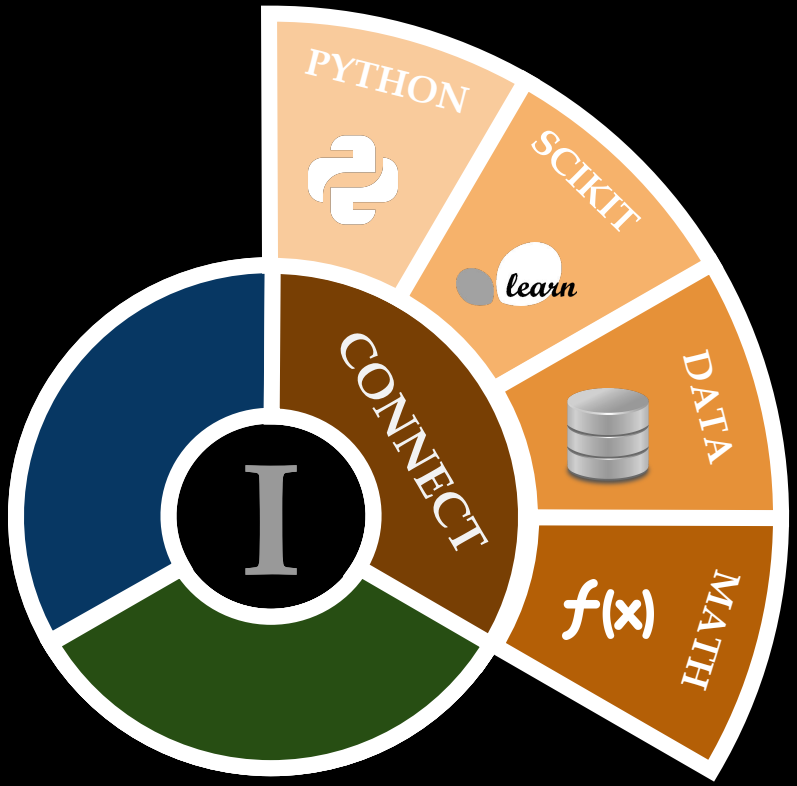
Tue & Thu 12:30p-1:45p

Nau Hall 101

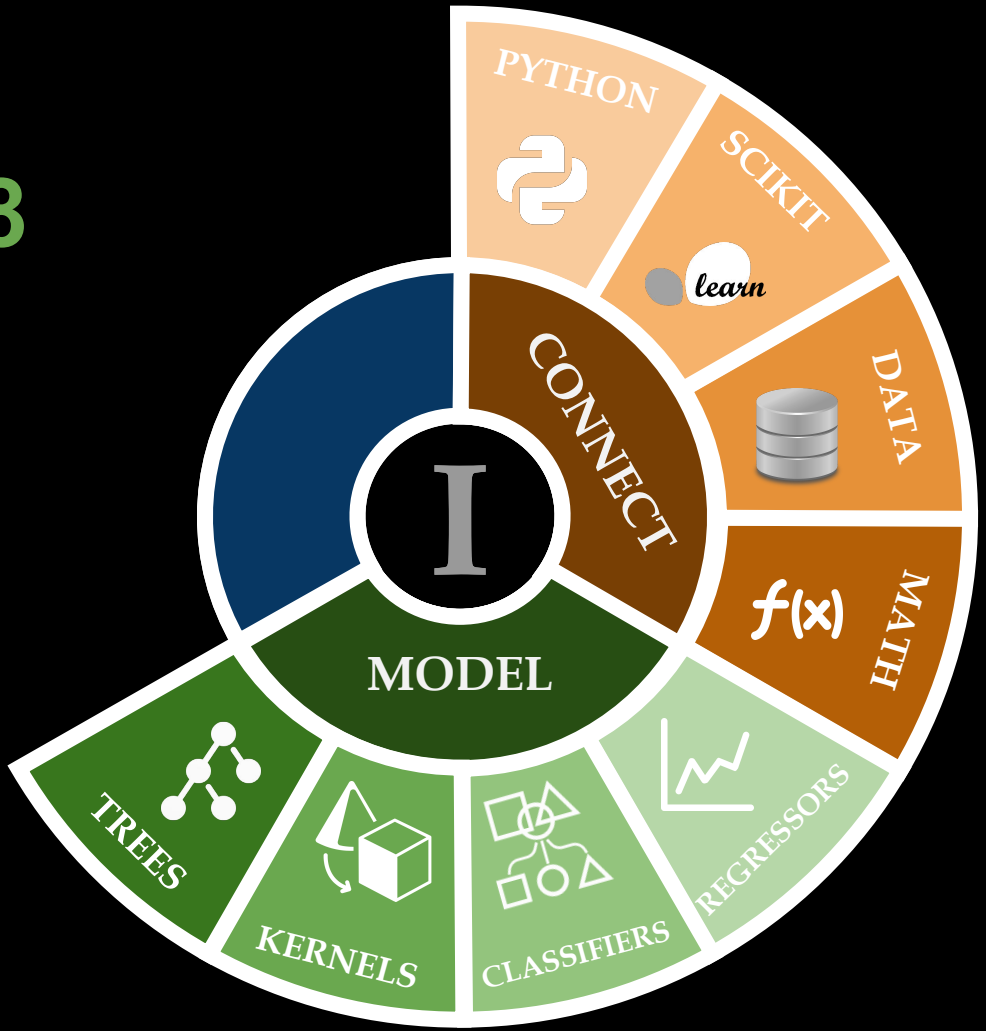
3 Phases



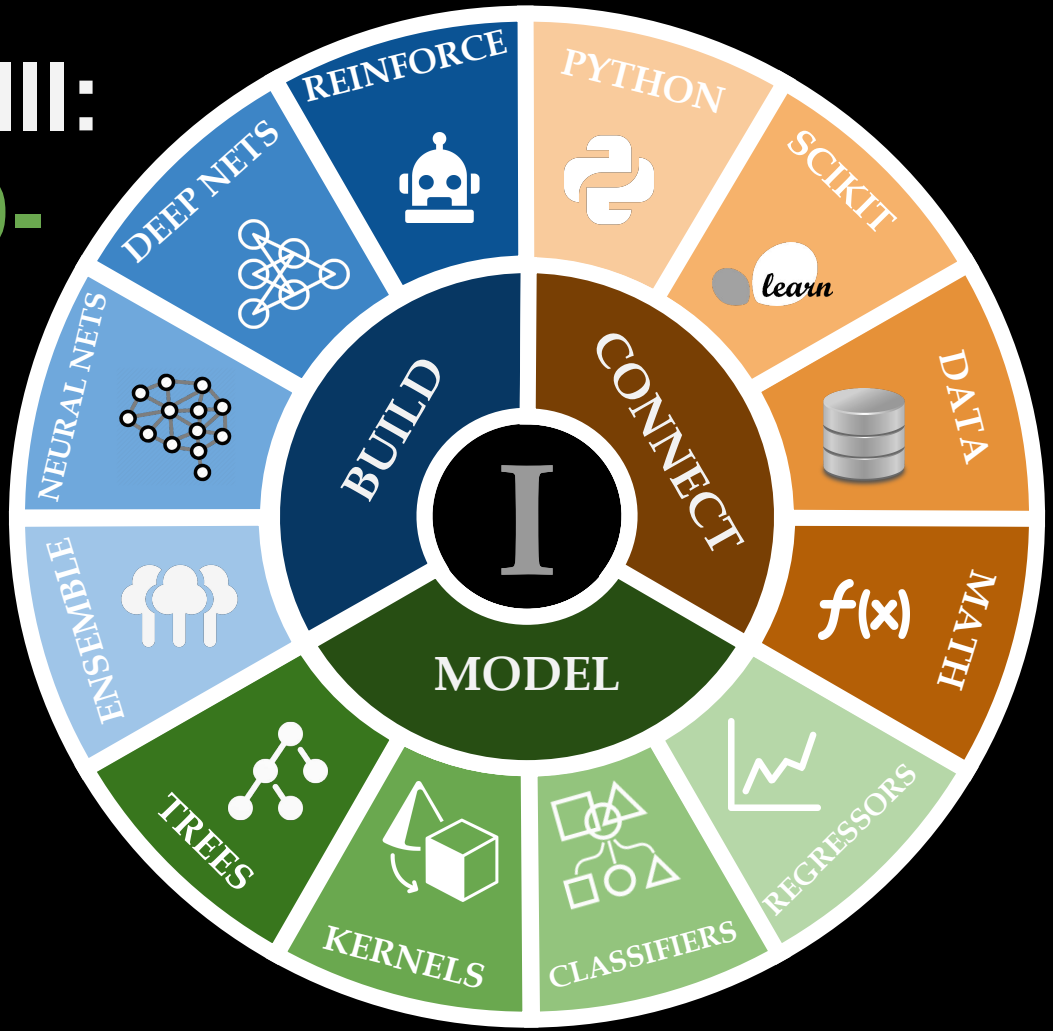
Phase I:
Week 1-3



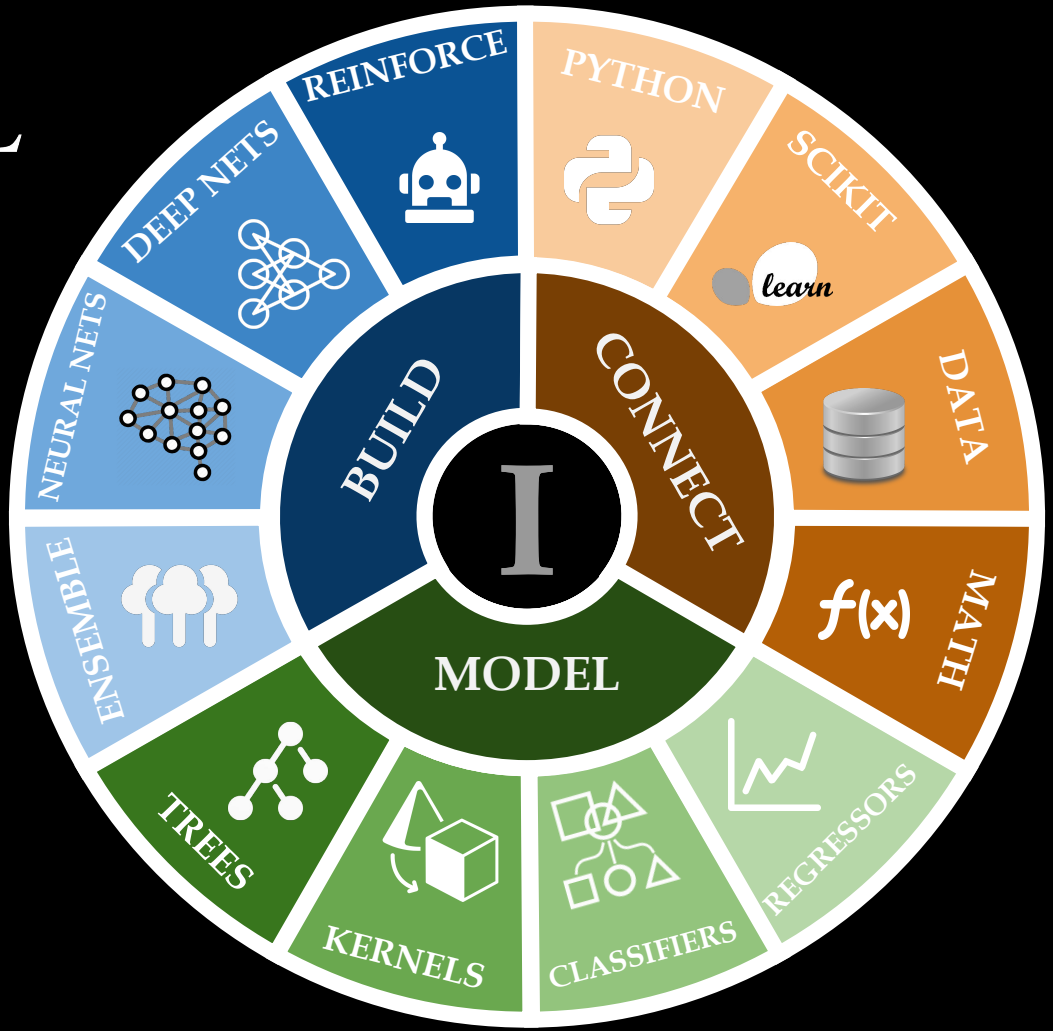
Phase II:
Week 4-8

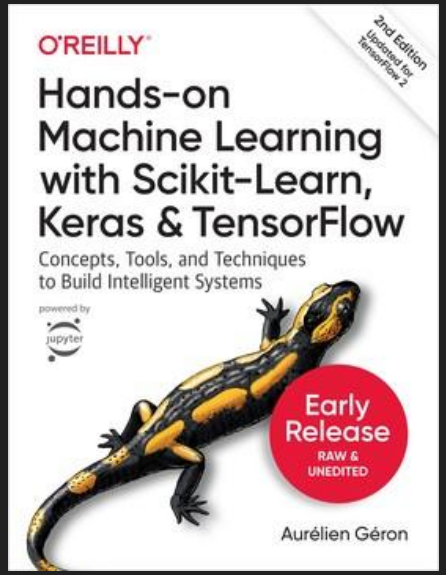


Phase III:
Week 9-



I ♡ ML


I
BUILD
CONNECT
MODEL
REINFORCE
DEEP NETS
NEURAL NETS
ENSEMBLE
TREES
KERNELS
CLASSIFIERS
REGRESSORS
PYTHON
SCIKIT
learn
DATA
MATH
f(x)



1 textbook

Aurelien Geron - **Hands-On Machine Learning with Scikit-Learn, Keras, and TensorFlow**, O'Reilly 2019

Download the code repo
https://github.com/ageron/handson-ml2

1 Google Faculty Award

Textbook Support (you don’t have to spend $$$)

Additional TAs to help with TensorFlow 2.0

1 portal

Collab: https://collab.its.virginia.edu/

1 course page

cs.virginia.edu/~nn4pj/teaching

2 Industry Speakers

To share their experience with students

Sep 12 and **Oct 01**

1000 points

F: 0-600;

D-: 600-629; **D**: 630-669; **D+**: 670-699;

C- : 700-729; **C**: 730-769; **C+**: 770-799;

B-: 800-829; **B**: 830-869; **B+**: 870-899;

A-: 900-929; **A**: 930-969; **A+**: 970-1000

1 Team Project

ML4VA

3 member team

200 pts / member

Project Expo (Dec 3rd)

Award Winning Videos

2 Exams

150 pts each

Midterm: Oct 3rd

Final: Dec 11th

3 Codeathons

20, 30, and 50 pts

To apply ML to real-world problems

48-72 hours to solve

Lightning Talks for bonus pts

4 Programming Assignments

100 pt each

14 days to implement

Code on **Google Colaboratory**

Submit on **UVACollab** as a **Jupyter Notebook**

2+ Extra Credits

by participating in **class discussion** (5-10 pts each)
and/or outside **hackathons** (20 pts)
Up to a maximum of **50 pt**
By the instructor's discretion

1 Piazza forum

Get answers by TAs or other students

Don’t post any code on assignments (that’s considered cheating!)

30% penalty

for **late** assignments
(10% penalty per late day)
if it's overdue more than 3 days, you will receive **zero** point.

10 students

who earn the **highest** points
by the end of the course will be **rewarded**.

and all of your **hard work** will be **recognized**!


Brace Yourselves
ML is coming...

Unused Slides

CS 4774 Checklist

- ❑ Familiar with Python's main scientific library
- ❑ Play with Scikit-learn algorithms efficiently
- ❑ Handle, clean, and prepare large-scale data
- ❑ Review linear algebra and probabilities reasonably well
- ❑ Understand common learning algorithms: regressions, k-nearest neighbors
- ❑ Engineer features, select models, and tune parameters
- ❑ Learn Support Vector Machine and its applications
- ❑ Make predictions with decision trees
- ❑ Learn bagging, pasting, boosting, stacking of voting classifiers
- ❑ Familiar with artificial neurons, perceptrons, and neural networks
- ❑ Explore layers of deep neural nets and tune their parameters

Activity 1:

Assume we are given the task to build a system that can distinguish junk emails.

- What is in a junk email that lets us know that it is junk?
- How can a junk email detected? What if it is not?
- What would you like the computer to do if it detects a junk email delete it automatically, move it t
  just highlight it on the screen?

Activity 2: Self-driving Car

You are given the task of building an automated taxi.

- Define the constraints. What are the inputs and the output?
- How can you communicate with the passenger?
- How to communicate with the other automated cars?

"Speeding, officer? You'll have to ask the self-driving car."

Phase I: CONNECT

PYTHON: Quick Review on Python (Week 1)

SCIKIT-LEARN: ML toolbox step-by-step (Week 2)

DATA: Handle, learn, prep data (Week 2)

MATH: Review some of essential concepts (Week 3)

Phase II: MODEL

REGRESSORS: Train regression models (Week 3,4)

CLASSIFIERS: Tune classification models (Week 5)

KERNEL: Train the kernel-based algorithms (Week 5,6)

TREES: Build decision trees (Week 9)

Phase III: BUILD

ENSEMBLE: Bagging, Boosting, Stacking (Week 9)

NEURAL NETS: Artificial Neural Network (Week 10)

DEEP NETS: Train DNN, CNN, RNN (Week 11,12,13)

REINFORCEMENT: Optimize Rewards (Week 14)