Intro to the course

Marco Morales
marco.morales@columbia.edu

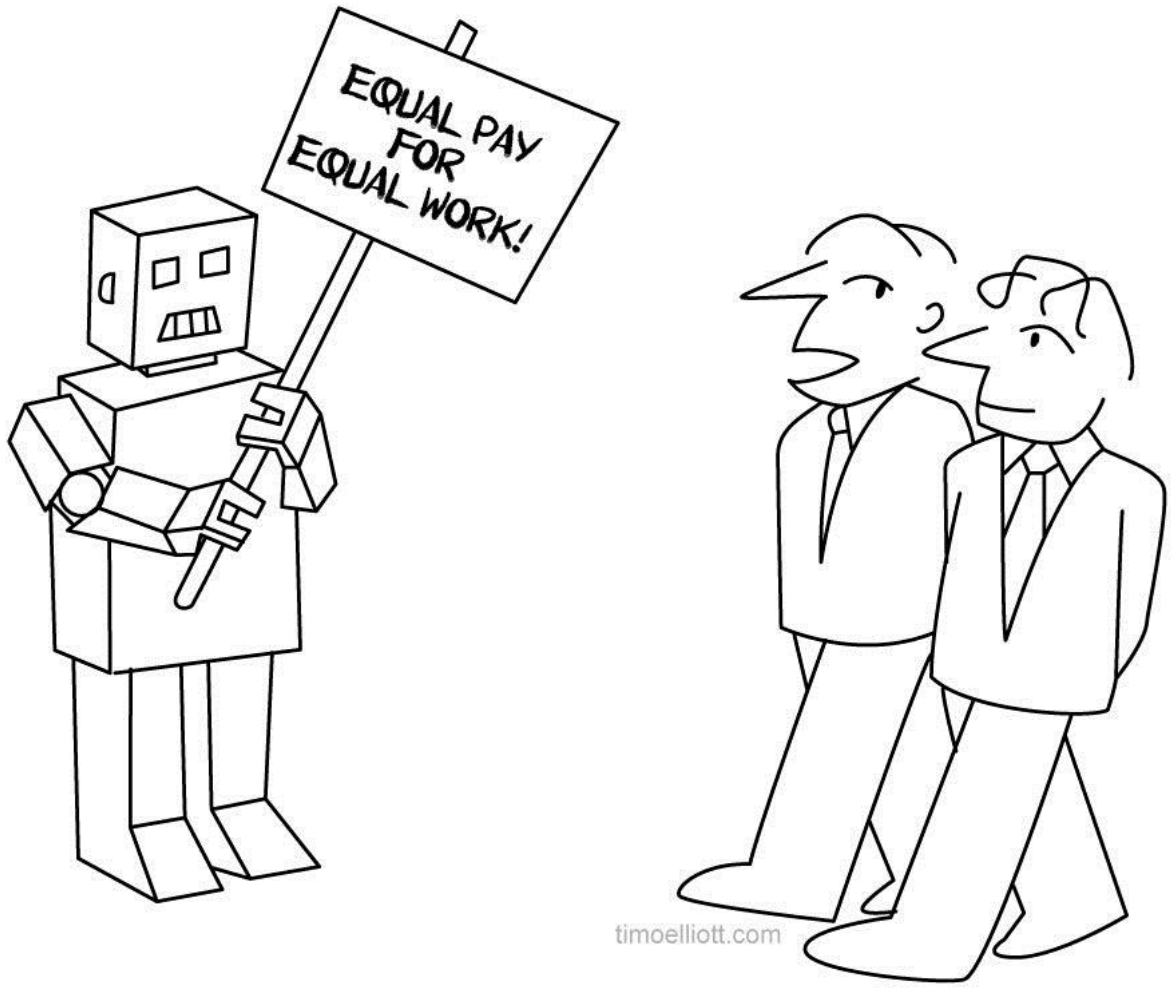
GR5074: Projects in
Advanced Machine Learning

Spring 2022
Columbia University

About us

Instructor:	**Marco Morales**
email:	marco.morales@columbia.edu
Office:	509E International Affairs Building
Office Hours:	appointments (via Calendly)
TA:	**Rachel Lee**
email:	yl3751@tc.columbia.edu
TA:	**Heinrich Peters**
email:	heinrichpeterz@gmail.com

About this course...



"I guess they really are getting smarter!..."

About this course...

- Continuation of **Machine Learning for the Social Sciences (GR5073)**
  - Linear Regression (and regularizations)
  - Logistic Regression (and regularizations)
  - KNN
  - SVM
  - Decision Trees (and ensembles)

About this course...

Objectives:

1. extend your Machine Learning toolkit through
   - exposure to **Deep Learning** models
     - Recurrent Neural Networks (RNNs)
     - Object detection models
     - Convolutional Neural Networks (CNNs)
   - experimentation with architectures for these models
2. create a portfolio of projects (**AI Model Share Initiative**)

Focus: gaining **intuition** + practical **experience**

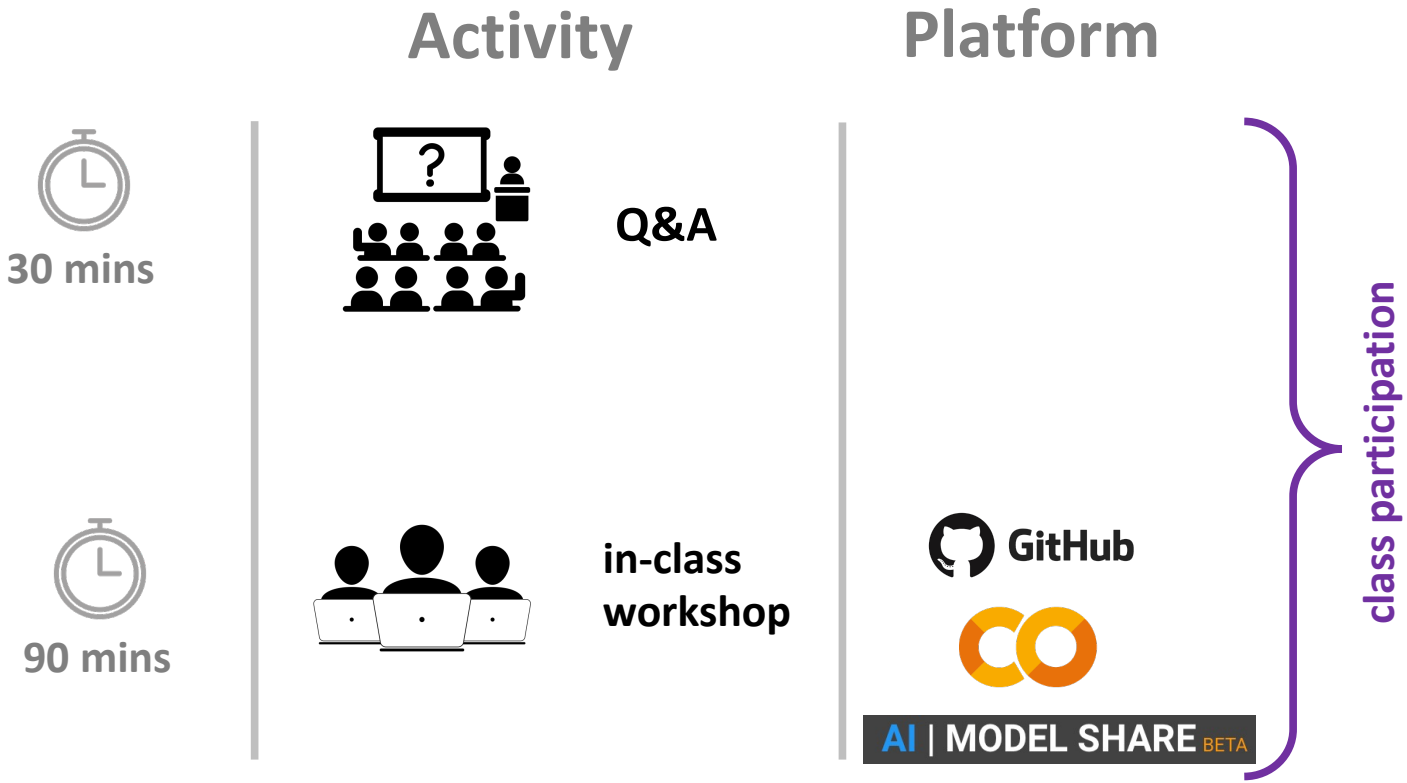
Course Dynamics: the **asynchronous** component

in your own time

Materials	Activity	Submission
assigned readings	read	questions/ comments — slack
Classroom — team projects	team collaboration	submit — Classroom, AI | MODEL SHARE BETA
Classroom — individual project	code	submit — Classroom, AI | MODEL SHARE BETA

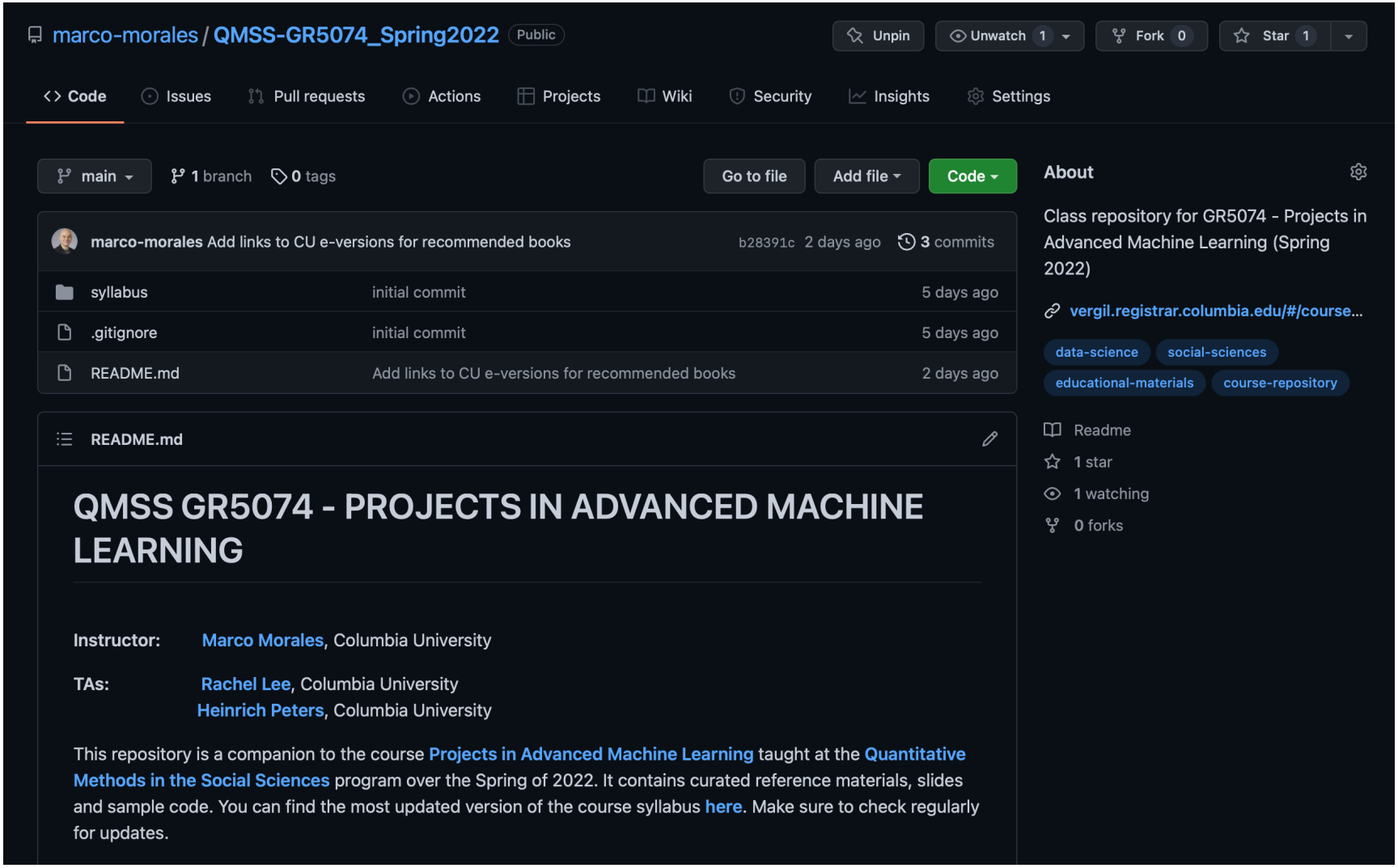
Course Dynamics: the **synchronous** component

all together



All class materials in the course's GitHub repo

clone and `pull` before each class



Course communications

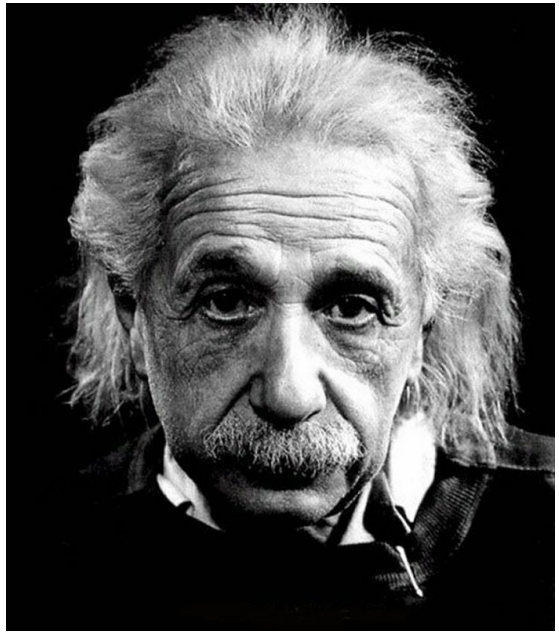
- **Slack** is the **preferred method of communication** for this course
- post questions in appropriate **channels**
- DM classmates & instructors / TAs
- **aim: collaboration –** if you have a question, likely others have it also (and perhaps an answer to share)
- we will **reserve email for official communications**

Tech stack for this class

Course Requirements

- **Team Projects (60%)**
- **Individual Project (25%)**
- **Attendance & Class Participation (15%)**

Why Deep Learning?



"The most incomprehensible thing about the world is that it is at all comprehensible."

- Albert Einstein

Why Deep Learning?

- the world is **compositional**: can be represented by stacked hierarchies increasing in abstraction
- **Deep Learning** allows learning at each level of the hierarchy (modules)
- **Deep Learning** replaces hand-engineered feature-extractors with (stacked) trainable modules
  - traditional ML relies on hand-engineered feature-extractors to expand representations of inputs
- but ... need care to match architecture with the nature of the data (for generalization)

Deep Learning genealogy

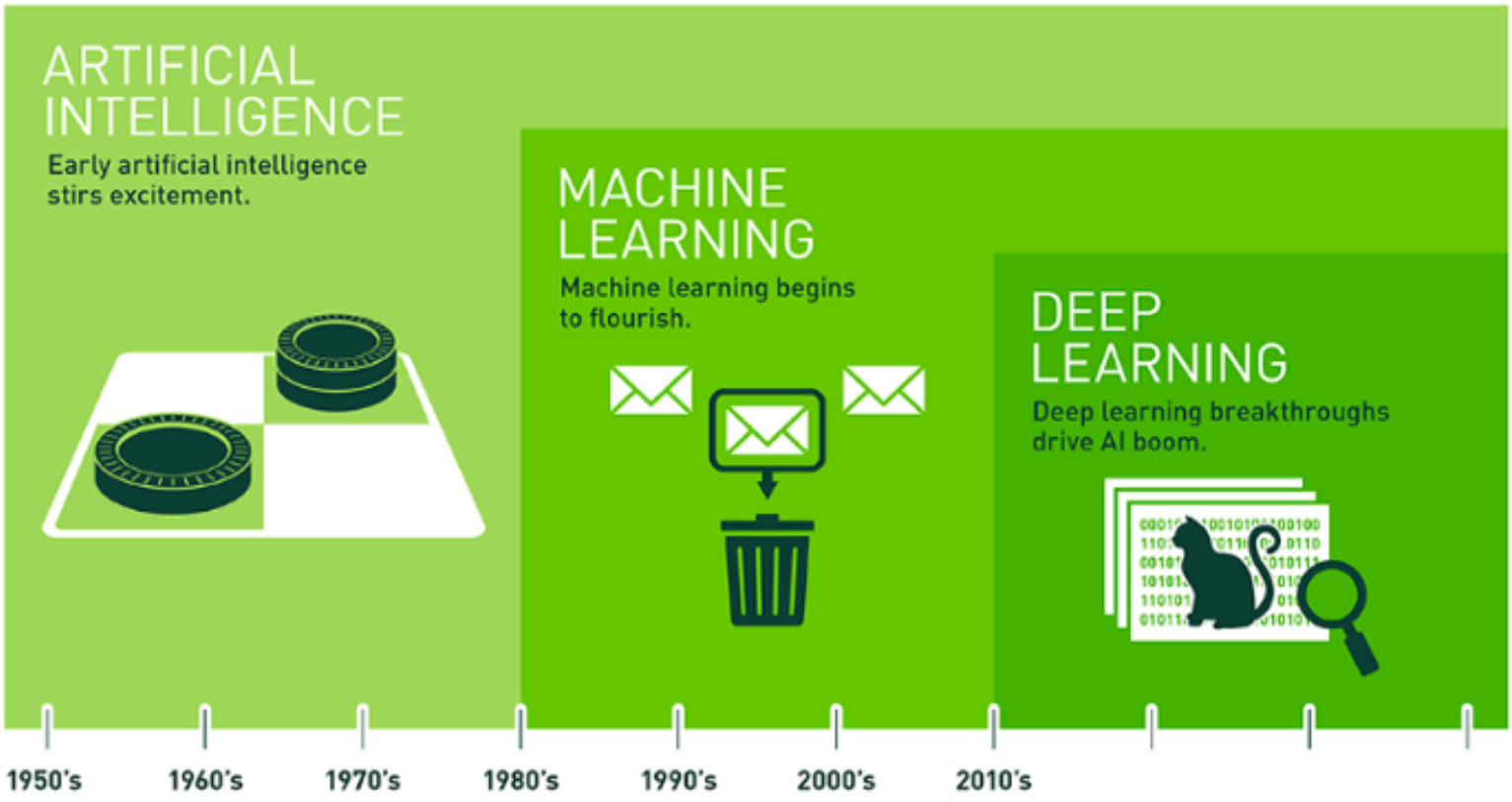


Figure: Cassie Kozyrkov (2018)

Deep Learning genealogy

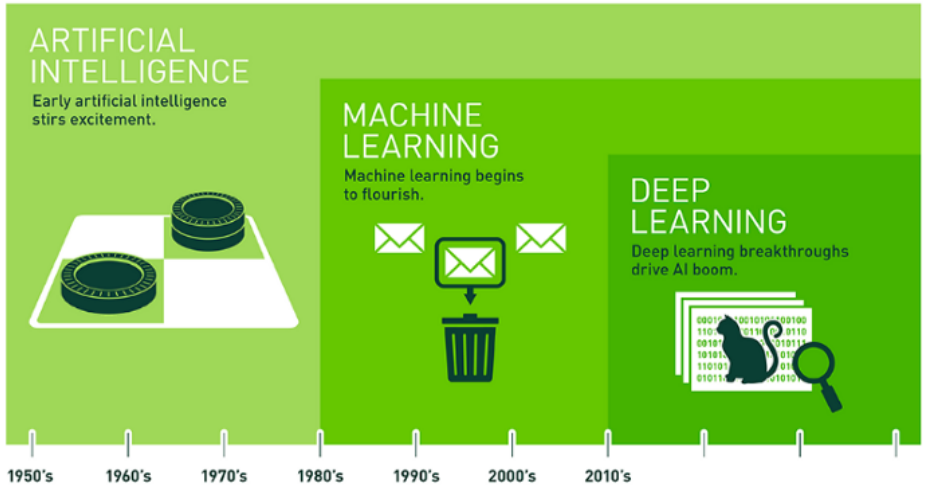


Figure: Cassie Kozyrkov (2018)

- **1940 - 60s**: cybernetics - rules-based
- **1980 - 90s:** back-propagation (“shallow”)
- **2010s - :** back-propagation + (“deep”)

Intro to the course

Marco Morales
marco.morales@columbia.edu

GR5074: Projects in
Advanced Machine Learning

Spring 2022
Columbia University