CS 181AI
Lecture 1

Overview

Arthi Padmanabhan

Jan 18, 2023

Hi, I’m Prof. Arthi

- I go by Prof. Arthi
- She/her/hers
- Pomona -> Microsoft -> UCLA -> Harvey Mudd
- I enjoy solving puzzles and playing tennis
- I would like to get to know you!

What this course covers

- A bit of background on ML + lots of hands-on experience building and running ML models

- Considerations and improvements when running ML models in real-world scenarios

- Current challenges and research directions related to systems for machine learning

- We can be flexible!

Goals

- Confidently approach building and deploying machine learning models
- Gain skills that prepare you for research/grad school
  - Paper reading, working on open-ended problems, writing technical reports
- Good mental health

Course website

- https://www.cs.hmc.edu/~apadmanabhan/CS181AI-S23/

Schedule

- Section 1: MW 9:35 – 10:50
- Section 2: MW 11 – 12:15
- Office hours: McGregor 321; F 2:00 – 4:00, M 4:00-5:00
  - At least for now, office hours will be in person
  - Or by appointment

Lectures

- In person; recordings TBD
- Please raise your hand and ask questions if anything is unclear
- There is reading component to this course – we will spend some time at the beginning of each class on paper discussions

Paper Reading

- There will be one paper for each lecture this semester (starting 1/30)
- Each section will be divided into 4 groups: A, B, C, and D – each group takes turns reading the paper of the day
- We'll spend the the beginning of class in groups, where one member from the group that read the paper explains it to the others and answers questions. We'll then convene as a group for final thoughts
- The group members should meet 1-2 days before to discuss the main points that they plan to convey when teaching and send me the main points

Slack

- We’ll be using Slack for discussions
- Feel free to post questions and other students and I can try to help you (but do not give away assignment answers)
- In person as well as on Slack, please be respectful and inclusive of all classmates

Assignments

- Roughly due every two weeks, released Monday and due Friday 5pm of the following week
- Will involve writing code in Google colab and writing a short report in LaTeX about your work
- Assignments will be released and submitted on Gradescope
- Only one assignment after spring break – then you can focus on final projects
- 5 “life happens” days – see me for extenuating circumstances

Final Project

- We’ll talk about ideas as semester progresses
- I’ll give some rough ideas and let you fill out a survey that will help me match you to a project + group or you can suggest your own
- Consists of a proposal, final presentation (Apr 24, 26) and final report due May 5th

Grading

- Assignments (5 x 12%) = 60%
- Paper Reading Discussions (10% for leading + 3% for participating in others): 13%
- Final Project (Proposal (5%) + Presentation (10%) + Report (15%)): 25%
- Drop by office hours some time before spring break: 2%

Presence in Class

- If you know you will be away on a day when your reading group is leading, please let me know ASAP and I'll switch you ahead of time
- I generally expect you to be in class. Life happens, so just give me a heads up when possible if you can't make it

Academic Honesty

- Please do work and study together. Discuss the readings. Post questions on slack. Given how new these topics are, the best information is available online through tutorials and forums. Please use these and feel free to explore.

- When writing up your final solutions, do not simply copy-paste code from a tutorial or peer.

- Please cite any sources you used at the bottom of your assignments

Acknowledgments

- No required textbook for this class
- I have used this as reference:
  - Designing Machine Learning Systems: An Iterative Process for Production-Ready Applications, by Chip Huyen

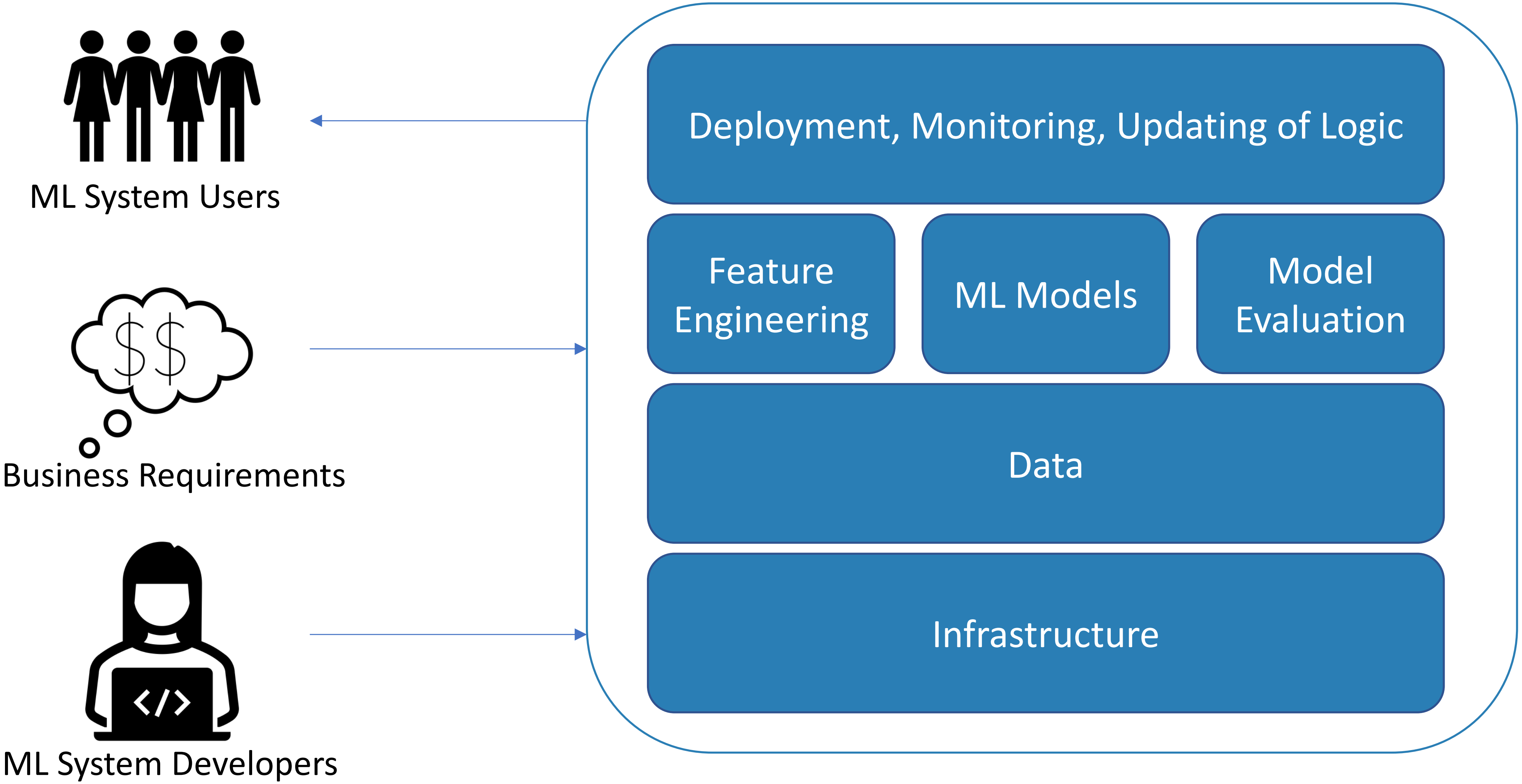
Accessibility

- *HMC is committed to providing an inclusive learning environment and support for all students. Students with a disability (including mental health, chronic or temporary medical conditions) who may need accommodations in order to fully participate in this class are encouraged to contact the Office of Accessible Education at access@g.hmc.edu to request accommodations. Students from the other Claremont Colleges should contact their home college's Accessible Education officer.*

Please Provide Feedback

- I want you to get the most out of this course
- Open Feedback form

ML Systems


ML System Users
Business Requirements
ML System Developers
Deployment, Monitoring, Updating of Logic
Feature Engineering
ML Models
Model Evaluation
Data
Infrastructure

To use ML or not to use ML

- ML is not a magical tool that can solve all your problems!
- Machine learning is an approach to learn complex patterns from existing data and use these patterns to make predictions on unseen data

To use ML or not to use ML

- ML is not a magical tool that can solve all your problems!
- Machine learning is an approach to **learn** complex patterns from existing data and use these patterns to make predictions on unseen data

To use ML or not to use ML

- ML is not a magical tool that can solve all your problems!
- Machine learning is an approach to learn **complex patterns** from existing data and use these patterns to make predictions on unseen data

To use ML or not to use ML

- ML is not a magical tool that can solve all your problems!
- Machine learning is an approach to learn complex patterns from **existing data** and use these patterns to make predictions on unseen data

To use ML or not to use ML

- ML is not a magical tool that can solve all your problems!
- Machine learning is an approach to learn complex patterns from existing data and use these patterns to **make predictions** on unseen data

To use ML or not to use ML

- ML is not a magical tool that can solve all your problems!
- Machine learning is an approach to learn complex patterns from existing data and use these patterns to make predictions on **unseen data**

What Makes ML a Good Idea?

- Repetitive
- At-scale
- Cost of wrong predictions is cheap*

What Makes ML a Bad Idea?

- Unethical
- Simpler solutions do the trick*
- Not cost-effective*

ML Common Use Cases (Enterprise)

- Recommendation Systems
- Fraud Detection
- Price optimization
- Churn prediction
- Sentiment Analysis
- Healthcare

ML Stakeholders: Restaurant Recommendations

- **ML engineers**
- **Sales team**
- **Product team**
- **ML infra team**
- **Org director**

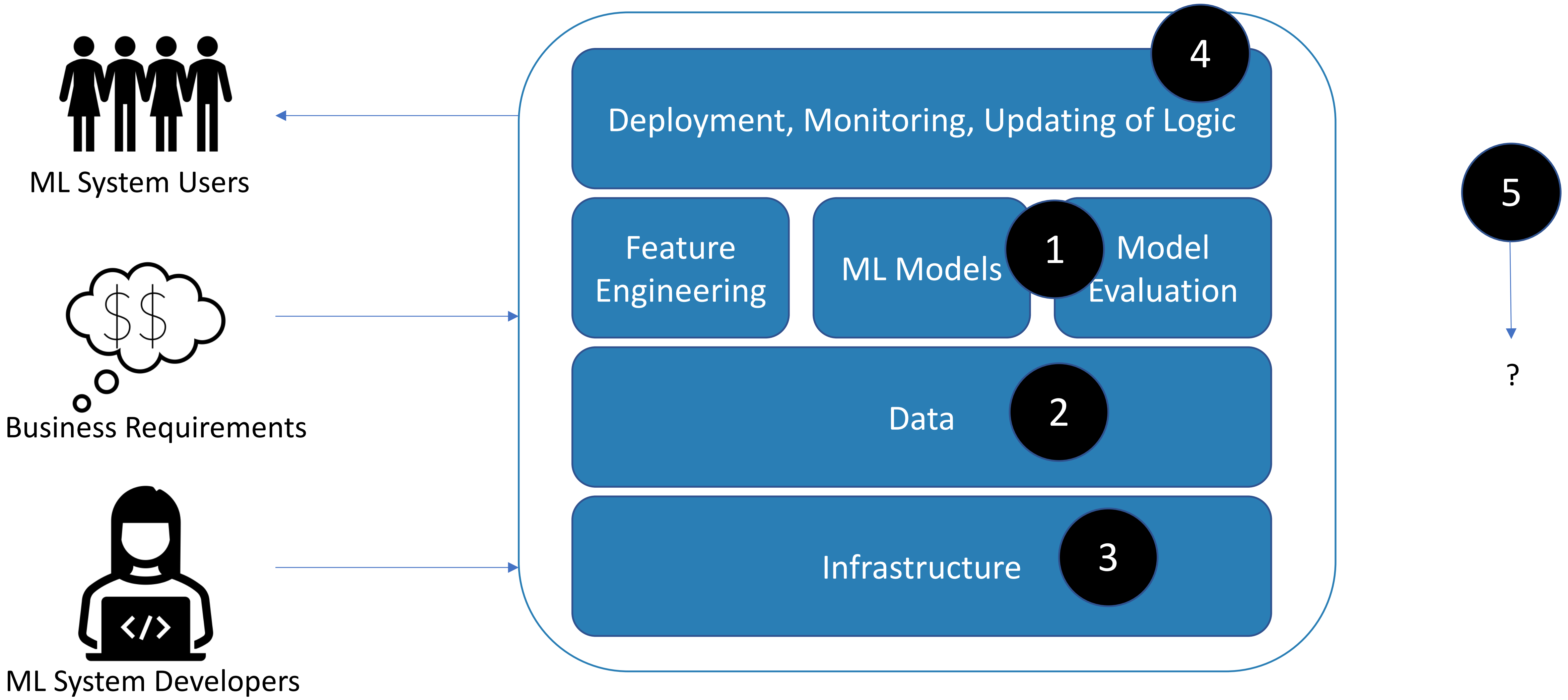
ML Stakeholders: Restaurant Recommendations

- **ML engineers**: want a model that recommends restaurants that users will most likely order from
- **Sales team**: want to recommend more expensive restaurants because they bring in higher service fees
- **Product team**: notices that every increase in latency leads to a drop in orders through the service, so they want a model that can return the recommended restaurants in less than 100 ms
- **ML infra team**: as traffic grows, they keep getting woken up in the middle of the night because of problems with scaling and the service needing to be constantly available, so they want to hold off on model updates to prioritize improving the ML platform
- **Org director**: wants to maximize margins, which might mean downsizing the ML team

ML in Research vs. Production

- Different objectives -> often the “best” research models aren’t used in production

Rough Plan


ML System Users
Business Requirements
ML System Developers
Deployment, Monitoring, Updating of Logic
4
Feature Engineering
ML Models
1
Model Evaluation
Data
2
Infrastructure
3
5
?

Next Week

- Bring laptops if you have them (for others, we’ll team up)
- Hands-on practice with using Google colab, PyTorch, and setting up and running an existing model