Week 12.2

Agenda

- Announcements

- Presentation: algorithms in social spaces

- Discussion

Announcements

- Projects and Debates

- Next week

- Discussion about Algorithms in Social Spaces

Definitions

Automated Decision-Making Systems - Systems in which decisions that impact people are made with software or through automated means

Data-Driven Systems - Systems that are informed, determined by, or dependent on the collection or analysis of data

Socio-Technical Systems - Systems that are characterized by interaction between people and technology

The Importance of Computational Algorithms

- Increasingly, computational algorithms decide things like:

  - who should be considered for which jobs (most digital hiring platforms)

  - what shows we watch (Netflix, Youtube, Amazon Prime, etc)

  - who we date (Match, Tinder, OkCupid, Hinge)

  - what content we see (Weibo, Twitter, Instagram, Facebook, Youtube, Amazon, Google/Alphabet)

  - which neighborhoods are more policed (Predpol, police stations)

  - who is more likely to receive harsher criminal justice decisions

Ingredients for Machine Learning

- Data

- Computing Power

- Machine Learning Scientists (who can write)

- Machine Learning Algorithms

Algorithms

A series of steps (or set of rules) for solving or performing a task.

For example, maybe you have a set algorithm for making yourself a sandwich for lunch. Your goal, or output, might be to make a sandwich that will fill you up. The inputs are all of the ingredients that will go into the sandwich, and your algorithm is how you order and arrange these things so that you can get your desired result, or output, of a sandwich for lunch.

Computational Algorithms are the same...

- Bubble Sort: the simplest sorting algorithm that works by repeatedly swapping the adjacent elements if they are in wrong order.

- Haversine: an algorithm implementing a formula for calculating the shortest distance on sphere between two points

- K-nearest Neighbors: simple, supervised machine learning algorithm for solving classification and regression problems

...but different
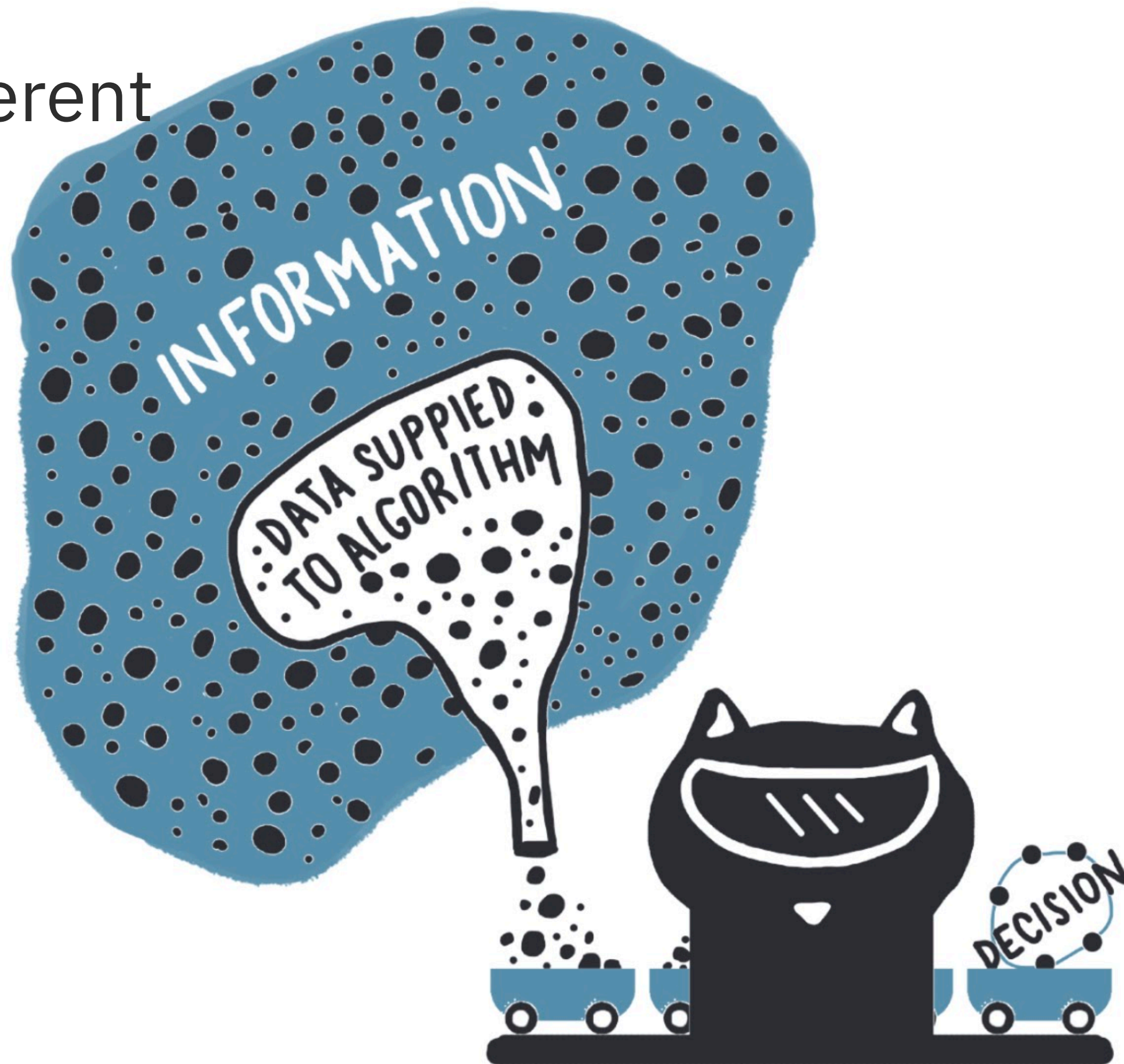
Discussion Prompt

NYU is throwing a high-stakes event taking place that will contain access to important people and opportunities that can impact all of our lives. Only 25 students in the entire university are allowed to attend. The university administration needs to bring the 25 most interesting and fun students to the event. But the university doesn't want to just choose students—to make things fair and replicable, the university is asking our class to create an algorithm that can determine which 25 students in the university are the most fun and interesting.

The catch is that we can only use digital data that the university has about the students. This includes things like where the students are from, what classes they've taken, their grades, any disciplinary marks, etc.

Your job is to figure out what data you will use to determine which people are fun and interesting, and then to use that data to come up with an algorithm for choosing the 25 students