DATA 1202
Spring 2024

Lecture 1

Introduction

Welcome to DATA 1202

Koloud N. Al Khamaiseh, PhD

kalkhama@mtu.edu

Rekhi Hall 310

Office Hours: TR 9:30 am-11:00 am, and by appointment, or by email

Data Science

Learning about the world from data using computation

Why Data Science?

1 How much data is generated every minute?

41,666,667
messages shared
by WhatsApp users

1,388,889
video / voice calls made
by people worldwide

404,444
hours of video streamed
by Netflix users

347,222
stories posted by Instagram users

150,000
messages shared by Facebook users

147,000
photos shared by Facebook users

2 TRILLION
searches on Google by the end of 2021

1.134 TRILLION MB
volume of data created every day

3,026,626
emails sent every second, 67% of which are spam

278,108 PETABYTES
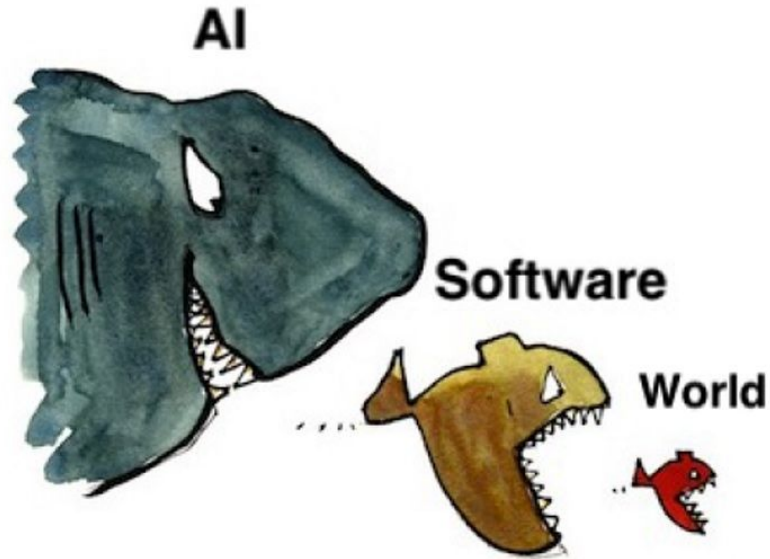global IP data per month by the end of 2021

230,000
new malware versions created every day

82%
share of video in total global internet
traffic at the end of 2021

Why Data Science?

What is Data Science?

Learning about the world from data using computation

- **Exploration**
  - Identifying patterns in data
  - Uses visualizations
- **Inference**
  - Using data to draw reliable conclusions about the world
  - Uses statistics
- **Prediction**
  - Making informed guesses about unobserved data
  - Uses machine learning

Course Content: Questions?

Course Website & Textbook

Canvas: https://mtu.instructure.com/courses/1478675
Book: https://inferentialthinking.com/chapters/intro.html
Ed (discussion): https://edstem.org/us/courses/51530/

Parts of DATA 1202

- Attendance is important, Introduce new topics, go through examples.
- [Textbook](#) is free, online, and written by faculty @ UCBerkeley
- Weekly lab section **(lab starts this week)**
- Weekly homework assignments & two projects
- Office hours: TR 9:30 am-11:00 am
- Midterm ~Week7
- Final exam in finals week

Details are in the Syllabus and on Canvas

Collaboration

Asking questions is greatly encouraged

- Discuss questions with each other (except exams)
- Submit lab assignments with partner
- Submit homework individually, but feel free to discuss
- Submit projects individually or with a partner

The Limits of collaboration

- Don't share solutions with each other (except lab/project partners)
- Do not search for solutions online or use AI tools to create your work
- Copying or other dishonesty will result in severe penalties

Getting Help

- Ask a friend, lab partner, lab colleague
- Ask on Ed
- Come to office hours
- Send an email

Example

(Demo)

Next Steps

- We'll see you in class on Wednesday!

  - And Lab on Thursday!