COGS138: Neural Data Science

Lecture 1

C. Alex Simpkins, PhD
UCSD Dept. of Cognitive Science, Spring 2023
http://casimpkinsjr.radiantdolphinpress.com/pages/cogs138_sp23

rdprobotics@gmail.com | csimpkinsjr@ucsd.edu

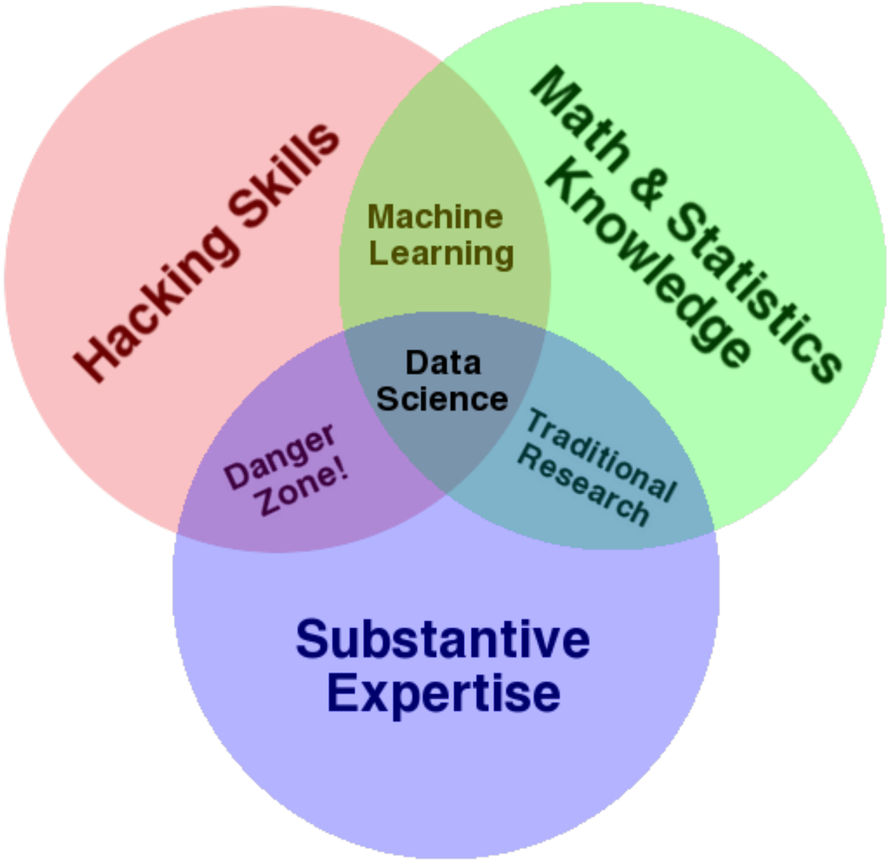
Plan for today

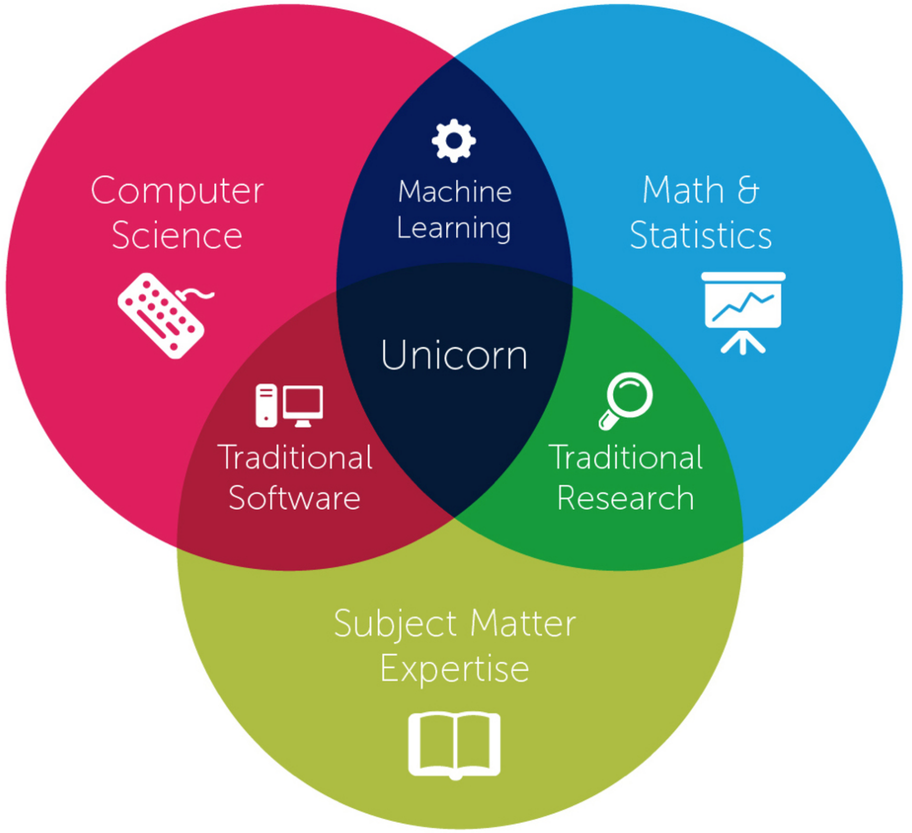
- Introduction
- Course outline and structure
- What is data science?
- What is neural data science?
- What will we be doing?
- Part I review of datahub, jupyter

Course structure

- Course structure
- Lectures (weekly) - at times one lecture one workshop day,
  - Location: SOLIS 109
  - Time: TuTh 12:30-1:50
- Readings + reading quizzes (5-10)
- Assignments (~5, datahub)
- Project (1, github)
- Lecture quizzes (~5, canvas)
- No final
- Piazza

What is data science?





http://drewconway.com/zia/2013/3/26/the-data-science-venn-diagram

Defining Data Science

a "concept to unify statistics, data analysis, machine learning and their related methods" in order to "understand and analyze actual phenomena" with data.[3] *It employs techniques and theories drawn from many fields within the context of mathematics, statistics, information science, and computer science.* -Wikipedia

"This coupling of scientific discovery and practice involves the collection, management, processing, analysis, visualization, and interpretation of vast amounts of heterogeneous data associated with a diverse array of scientific, translational, and interdisciplinary actions." -David Donoho ("*50 years of Data Science)*

"an emerging discipline that draws upon knowledge in statistical methodology and computer science to create impactful predictions and insights for a wide range of traditional scholarly fields" - from a panel Rafael Irizarry moderated, shared on SimplyStatistics ("*The role of academia in data science education"*)

"an umbrella term used by organizations to describe the processes used to extract value from data" -Rafael Irizarry's personal definition in "*The role of academia in data science education"*

"The study of how the quantification of observable phenomena can lead to human understanding of the processes giving rise to those phenomena—or even the ability to predict future outcomes absent human understanding—and why certain phenomena require more or less data to lead to human understanding and/or prediction accuracy". -Brad Voytek's definition

"The scientific process of extracting value from data"

Data scientists ask interesting questions & answer them with data

What is NEURAL Data Science?

What will we be doing?

Learn how to:

- Think from a “data first” perspective: what data *would* you need to answer your scientific questions of interest?
- Develop hypotheses specific to big data environments in neuroscience.
- Work with many different neuroscience data types that might include data on behavior, brain structure and connectivity, single-unit spiking, field potential, gene expression, and even text-mining of the peer-reviewed neuroscientific literature.
- Read and analyze data stored in standard formats (*e.g.*, Neurodata Without Borders and Brain Imaging Data Structure).
- Integrate multiple heterogeneous datasets in scientifically meaningful ways.
- Choose statistical model(s) informed by the underlying data.
- Design a big data experiment and integrate data from multiple open data sources.
- Consider alternative hypotheses and assess for spurious correlations and results.

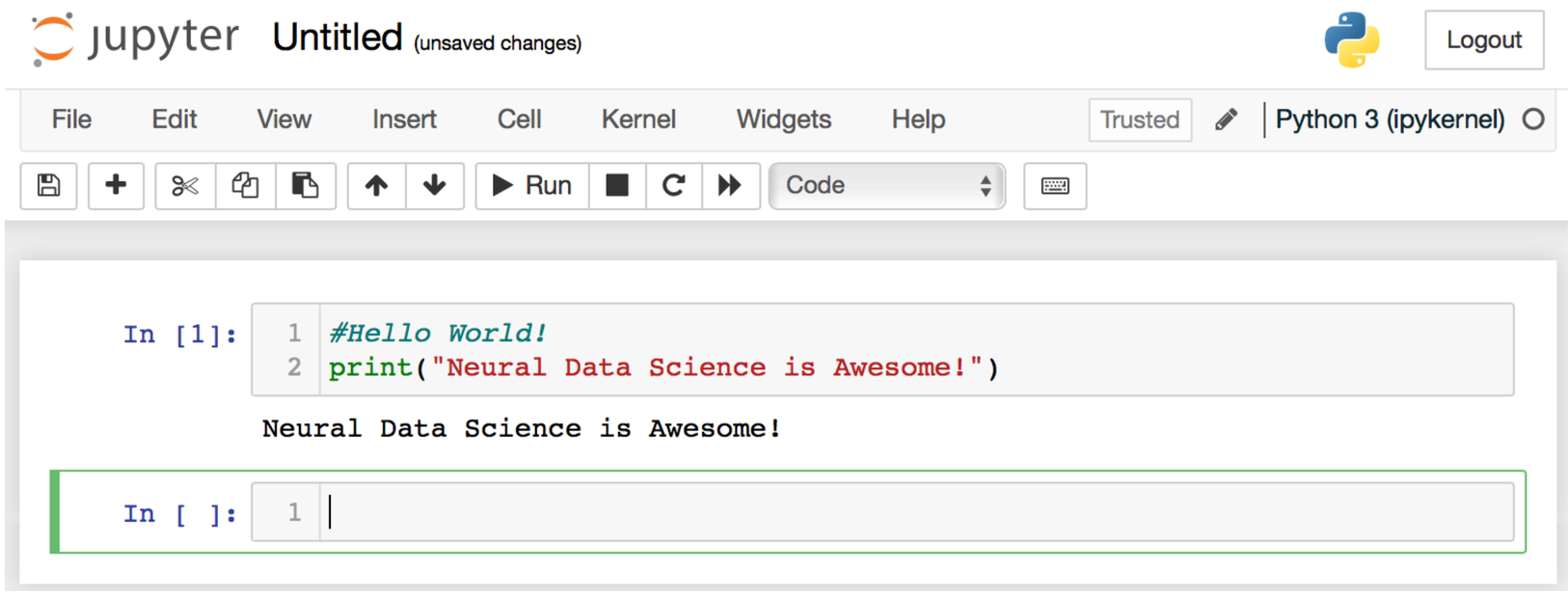
Data Science as the Extended Mind

Introduction to DataHub

- datahub.ucsd.edu
- Logging in
- Navigating
  - Intro to file structures and how they relate to your computer
- Upload
- Download
- Rename files
- Make folders
- Delete
- Submitting assignments, fetching assignments
- Validating!

Jupyter notebooks review

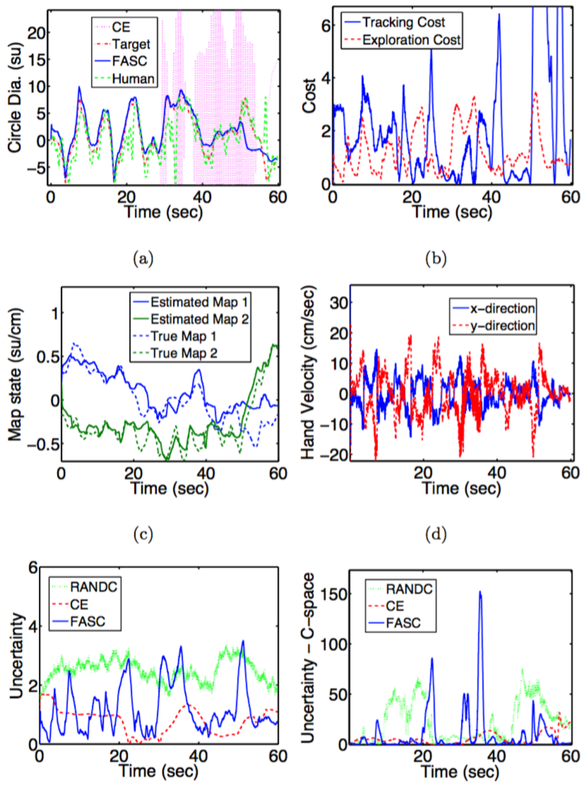
- Installing anaconda
- https://github.com/COGS108/Tutorials
- https://github.com/NeuralDataScience/Tutorials
- Correcting common issues
- Up to students to correct and resubmit so grading can be timely

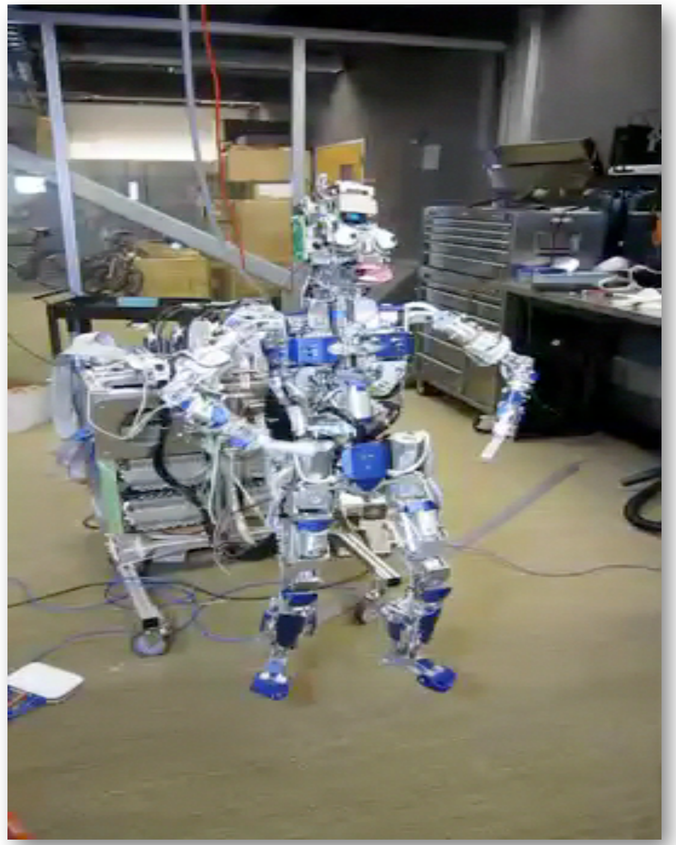


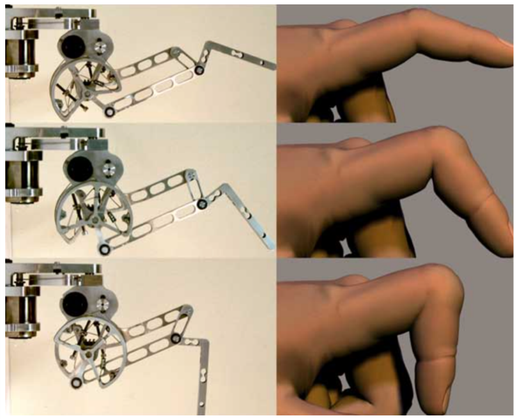
Who am I?

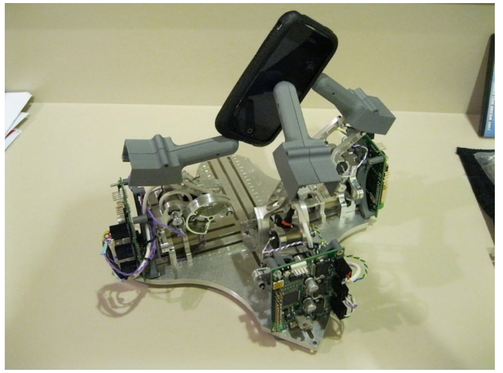
- C. Alex Simpkins Jr. Ph.D.
- BS/BS/MS/PhD UCSD Psyc, AMES, MAE, MAE, 2 postdocs UCSD Cogs and UW CSE
- Taught as TA ~20 times as a student, taught COGS109 as a grad student, taught at SDSU for a year in ME in Design, came back to UCSD last quarter - COGS100 and 108
- Been involved in teaching for over 30 years teaching martial arts

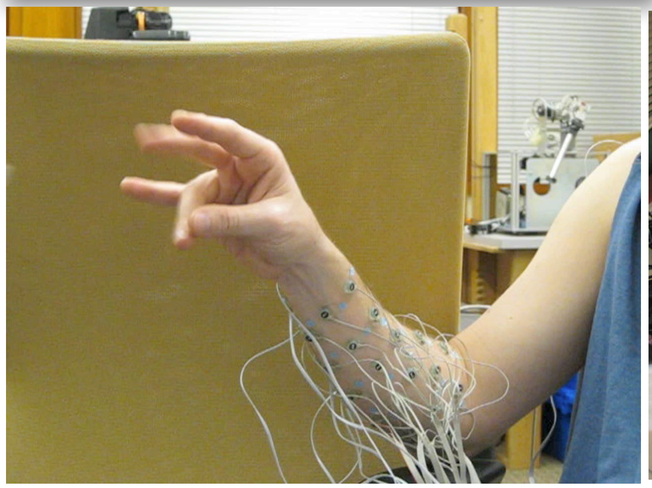
- PhD in Control Theory (and Design) is based on cybernetics
- Robotics

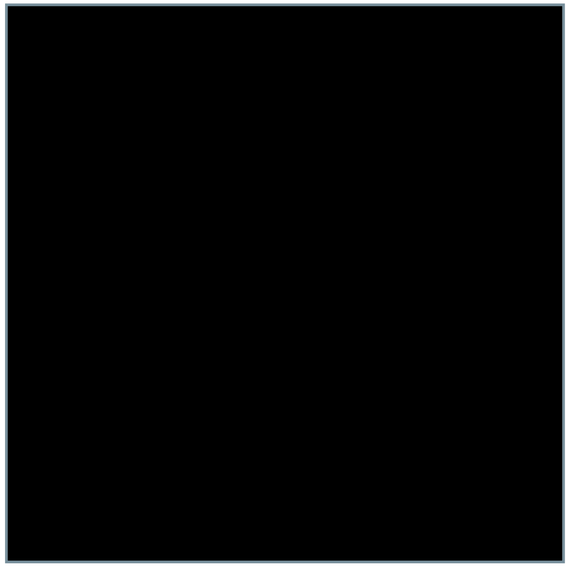






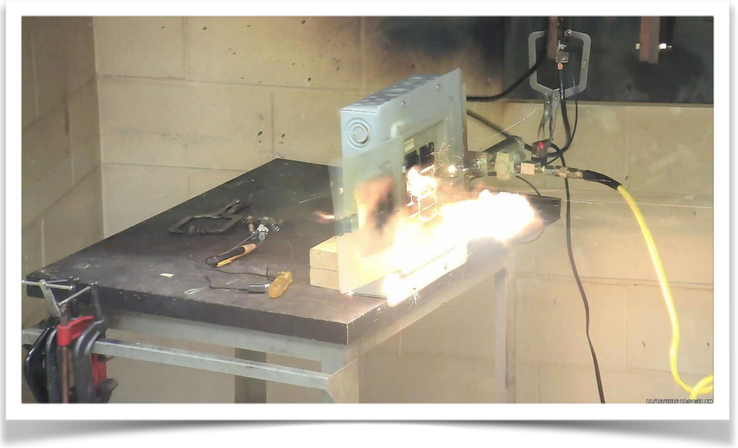


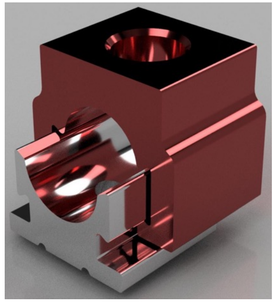


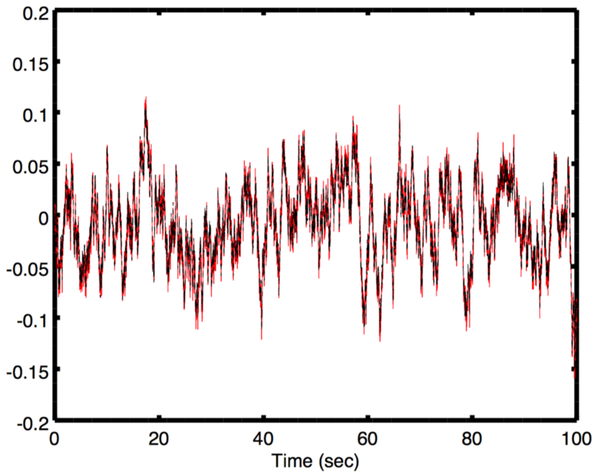


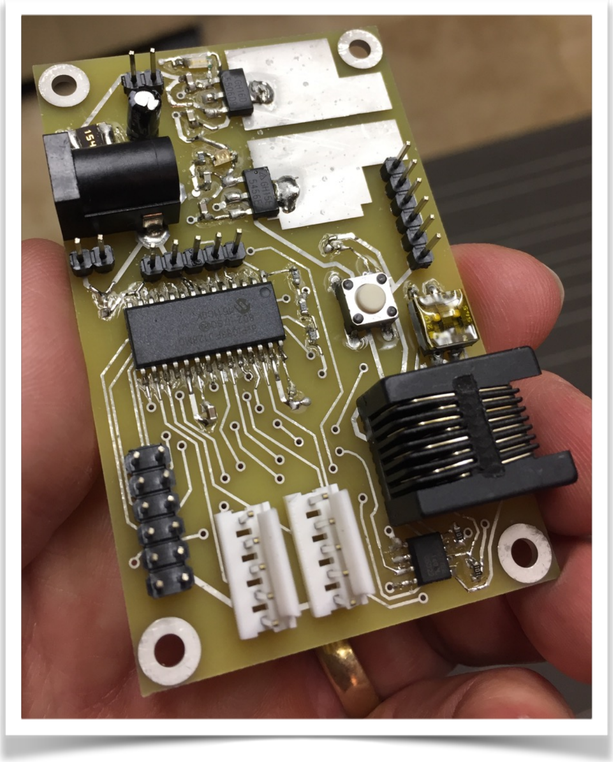
Who am I?

- Many industrial projects - robotics, AI, drones
- Consulting
- Entrepreneurial startup work
- Work with larger companies
- Research - COVID, robotics, AI, control









What to expect in this class

- What is my (our) role?
  - I am (we are) here to help you to learn and succeed, to open the door
  - NOT here to weed anybody out
  - NOT here to compete with you
  - Mutual respect

What to expect in this class II

- What is your role?
  - Learn! Open your mind
  - Put in the effort - you must walk through the door
    - Watch/attend lectures, do the readings, complete assignments and tests, and think about it all
  - Treat each other well, help each other to succeed (but do your own work of course)