Machine learning algorithms

Introduction

2020–09-23

CSCI 471 / 571, Fall 2020

Kameron Decker Harris

Welcome!

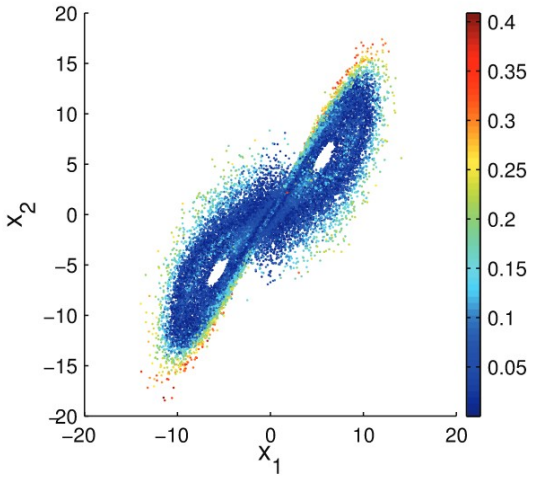
- Plan for today:
  - Introductions
  - Read & discuss syllabus
  - What is machine learning (ML)?

A little about me

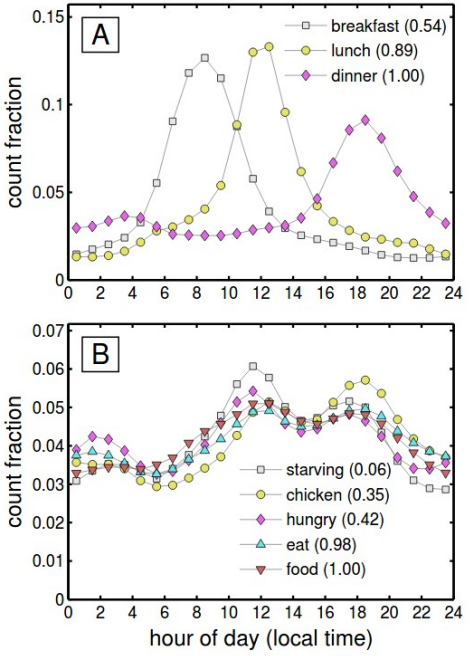
- Grew up in small-town Vermont
- He/him
- “Professor Harris”

University of Vermont

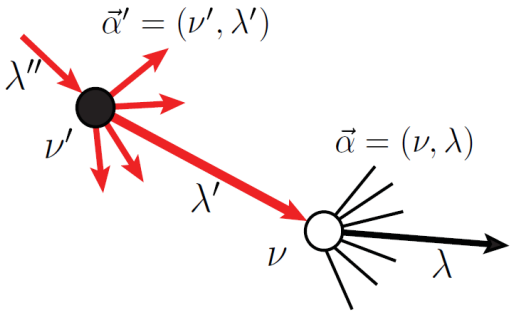
- Undergrad in physics & math, MSc in math



Dynamical systems & chaos



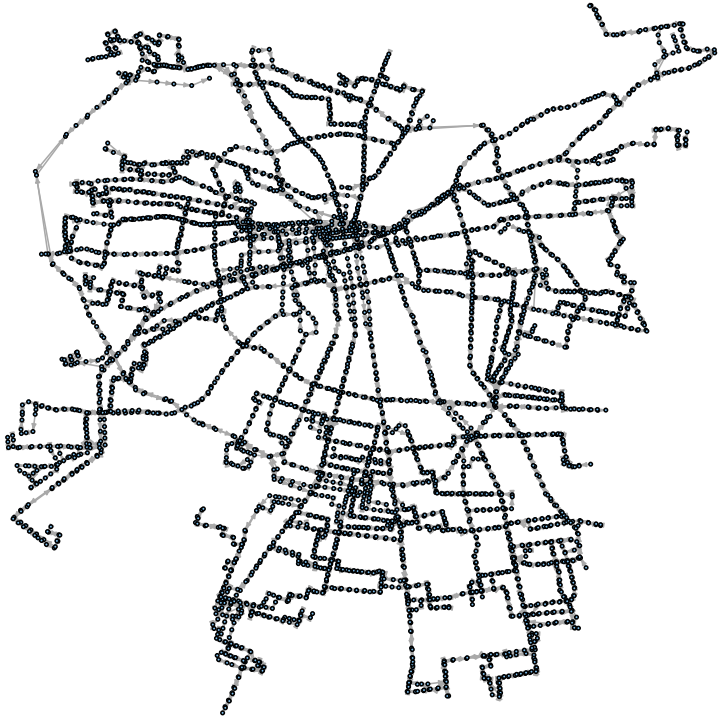
Twitter data

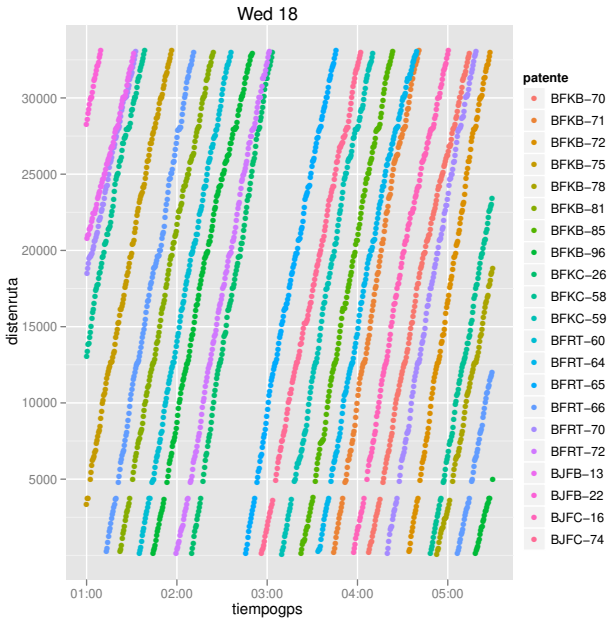


Networks & spreading

Year in Chile

- Optimization to improve Santiago bus scheduling





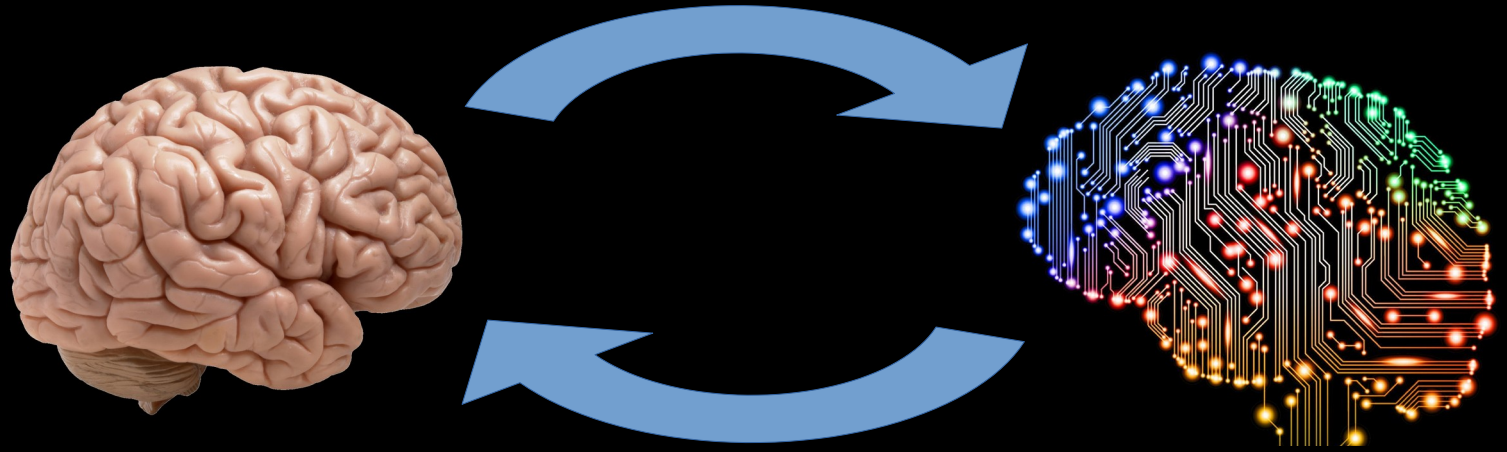
Moved to Washington in 2012

Now-wife Meira

Computational neuroscience

- PhD in applied math, postdoc CSE & biology, U. Washington

Brain-inspired algorithms
Emphasis on dynamics
Alternate models of computation



Neuro

Big data management
Analysis of experiments
Artificial neural networks

CS

Round of introductions

- Tell us your name, preferred pronouns
- **Briefly** share something about yourself, e.g.
  - Something you learned recently (non-school)
  - Something you did over summer

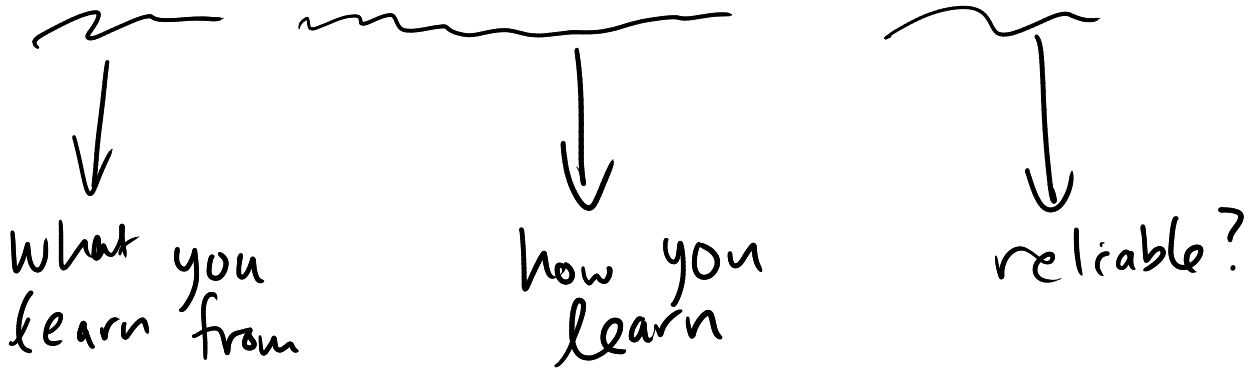
Syllabus

1st assignment

- Read the course webpage:
  - https://glomerul.us/teaching/CSCI-471/2020-Fall/
- Can access ^ from Canvas syllabus page
- Breakout into small groups for 15 min
  - Questions?
  - Suggested additions, norms?

What is ML?

- Data + Optimization + Statistics → Predictions



Examples of ML applications

- Let’s list some examples together

Famous recent ML successes

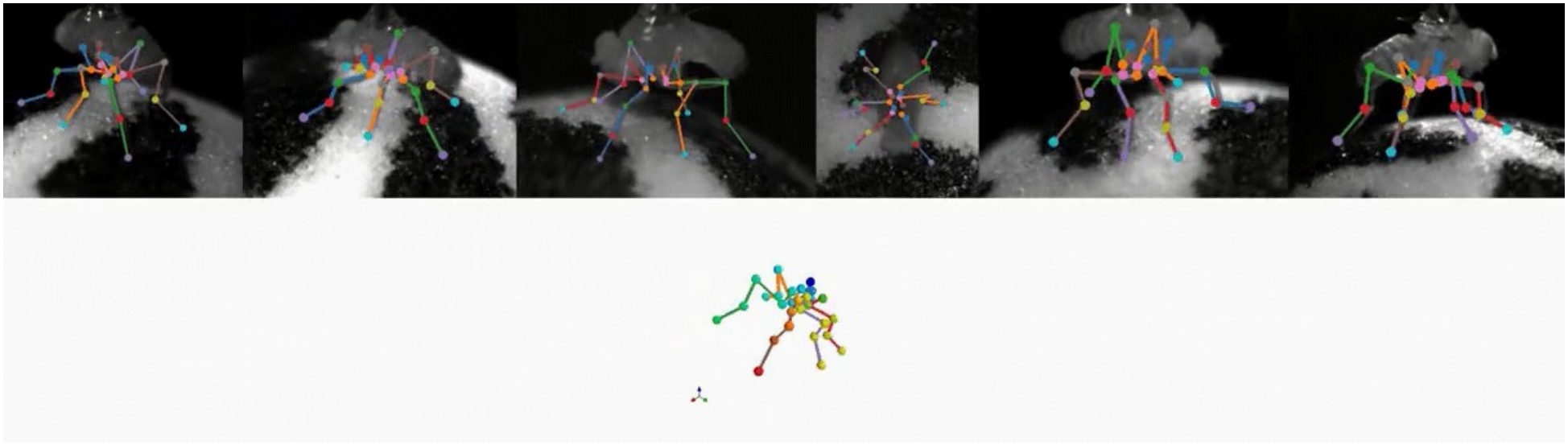
Image classification

(CIFAR 100 data)

AlphaGo

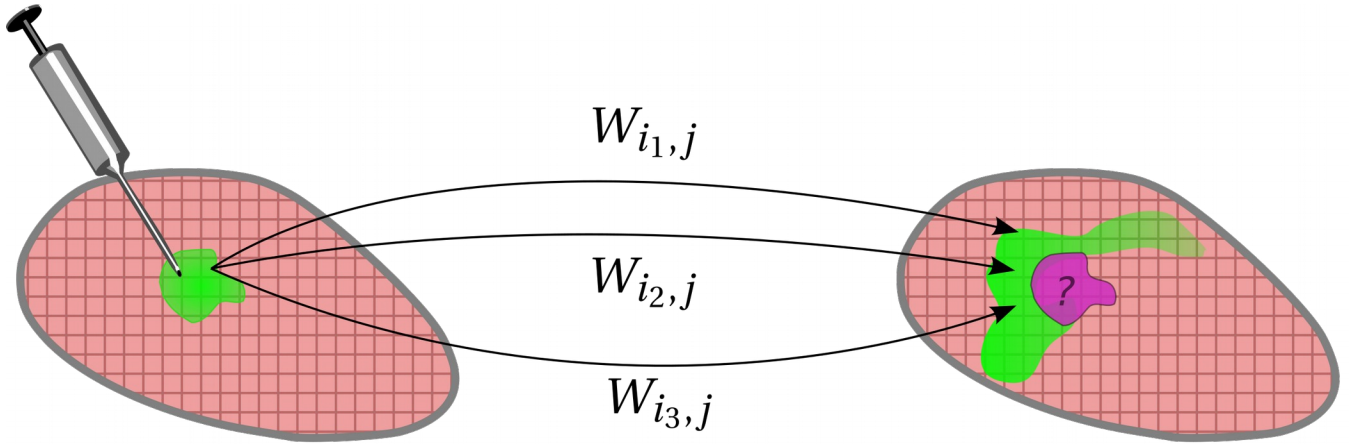
Wikimedia commons Dilaudid

ML in data analysis



Karaschuk et al., 2020

Ex: network reconstruction



x: source expression

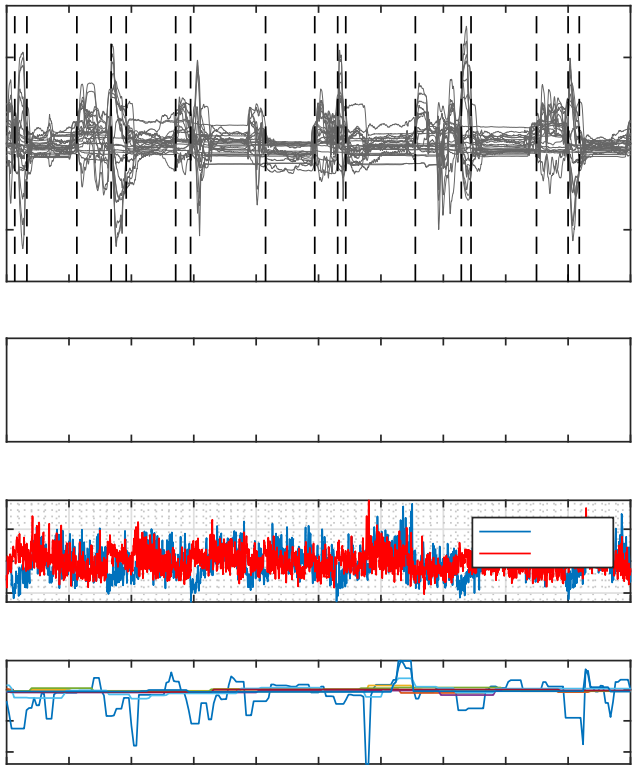
y: target expression

Goal

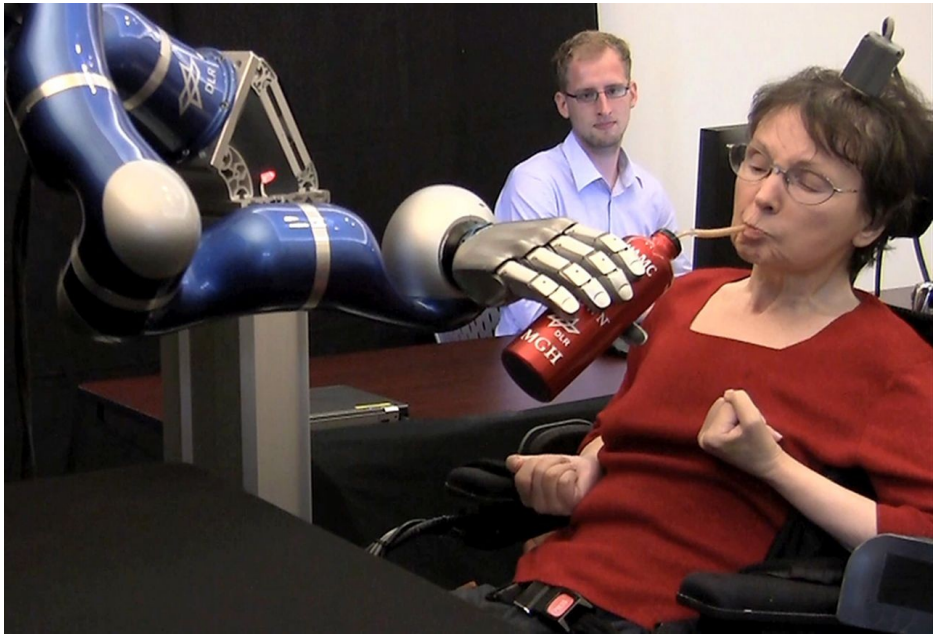
Find unknown weight matrix *W* so

$$\mathbf{y} \approx W\mathbf{x}$$

ML for neuroscience



Harris et al., 2020



Hochberg et al., (2012)

Goals for the quarter

Goals for the quarter

- Understand important, existing algorithms
  - Theoretical grounding
  - Implementation in code

Goals for the quarter

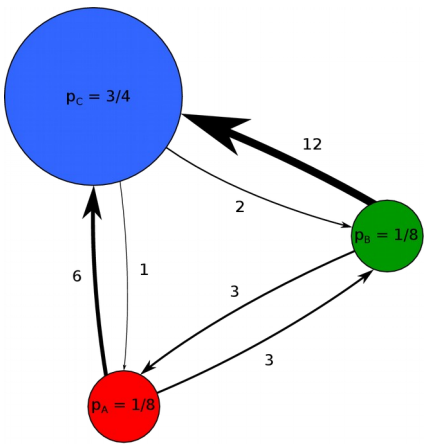
- Understand important, existing algorithms
  - Theoretical grounding
  - Implementation in code
- General principles of ML
  - Tradeoffs, scalability, uncertainty
  - Building blocks of cutting-edge algorithms

ML Taxonomy

Data

F	G	H	I	J
STNAME	CTYNAME	CENSUS2000POP	ESTIMATESBASE2000	POPESTIMATE2000
Alabama	Alabama	4447100	4447382	4451849
Alabama	Autauga County	43671	43671	43872
Alabama	Baldwin County	140415	140415	141358
Alabama	Barbour County	29038	29038	29035
Alabama	Bibb County	20826	19889	19936
Alabama	Blount County	51024	51022	51181
Alabama	Bullock County	11714	11626	11604
Alabama	Butler County	21399	21399	21313
Alabama	Calhoun County	112249	112243	111342
Alabama	Chambers County	36583	36614	36593
Alabama	Cherokee County	23988	23986	24053

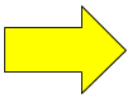
Color
Red
Red
Yellow
Green
Yellow



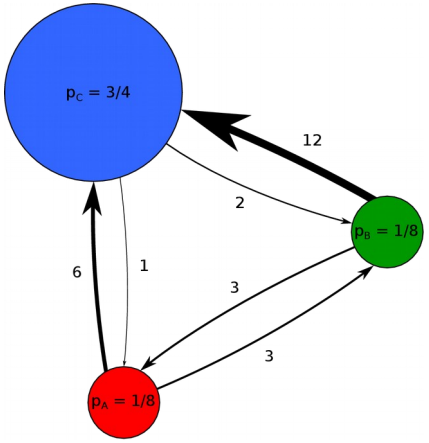
Data

F	G	H	I	J
STNAME	CTYNAME	CENSUS2000POP	ESTIMATESBASE2000	POPESTIMATE2000
Alabama	Alabama	4447100	4447382	4451849
Alabama	Autauga County	43671	43671	43872
Alabama	Baldwin County	140415	140415	141358
Alabama	Barbour County	29038	29038	29035
Alabama	Bibb County	20826	19889	19936
Alabama	Blount County	51024	51022	51181
Alabama	Bullock County	11714	11626	11604
Alabama	Butler County	21399	21399	21313
Alabama	Calhoun County	112249	112243	111342
Alabama	Chambers County	36583	36614	36593
Alabama	Cherokee County	23988	23986	24053

Color
Red
Red
Yellow
Green
Yellow



Red	Yellow	Green
1	0	0
1	0	0
0	1	0
0	0	1



Data as vectors