Lecture 01. Course Overview

Nakul Gopalan

Post-doctoral Fellow,

College of Computing, Georgia Tech

This course is made with slides developed from multiple sources including Mahdi Roozbahani, Rodrigo Borela Valente, Eric Eaton, Michael Littman, Byron Boots. I will try to attribute the sources as often as I can.

Meet the Instructor

Nakul Gopalan – Nakul would do
Rant: Not a professor. I am Dr. Gopalan, but Nakul is preferred by a large margin.

He / him / his

Email: ngopalan3@gatech.edu

Research areas: Language grounding, planning, reinforcement learning, language understanding

Meet the TAs

Head TA:
Vidisha Goyal

TAs:
Zheng Zhang
Christopher J Banks
Aryan J Pariani
Kevin Y Li
Yaru Niu

How to succeed in this course?

Ask questions
It is your fundamental right as a student to ask questions.
Be inquisitive. I am not known for delivering monologues.

Staying motivated in the lectures over time. You have the opportunity to ask questions directly to me, and you should use it.

I need a response **full of energy** from **all** students.

It gives me a positive interactive attitude and shows me you are here to learn new exciting things. That way I stay motivated.

Refer to:

Class Website

For anything (updates, lectures, logistics, and so on) related to this class

I verbally ask questions in the class

Sometimes I ask questions about previous lectures or something that you have already learnt.

Answer the question even if you think you are wrong (nobody loses point answering the question wrong). It will help you and other students to understand concepts much better.

piazza (the best source for Q/A)

Ask your questions in piazza (make it public to other students), and also please see other questions in piazza, it might answer your question. (Please do not send me or TAs Emails regarding hw questions, exams, and other logistics – you can also ask “private question” on piazza) =>Class participation

Bonus points: Undergrad and grad

Blue Jeans interface

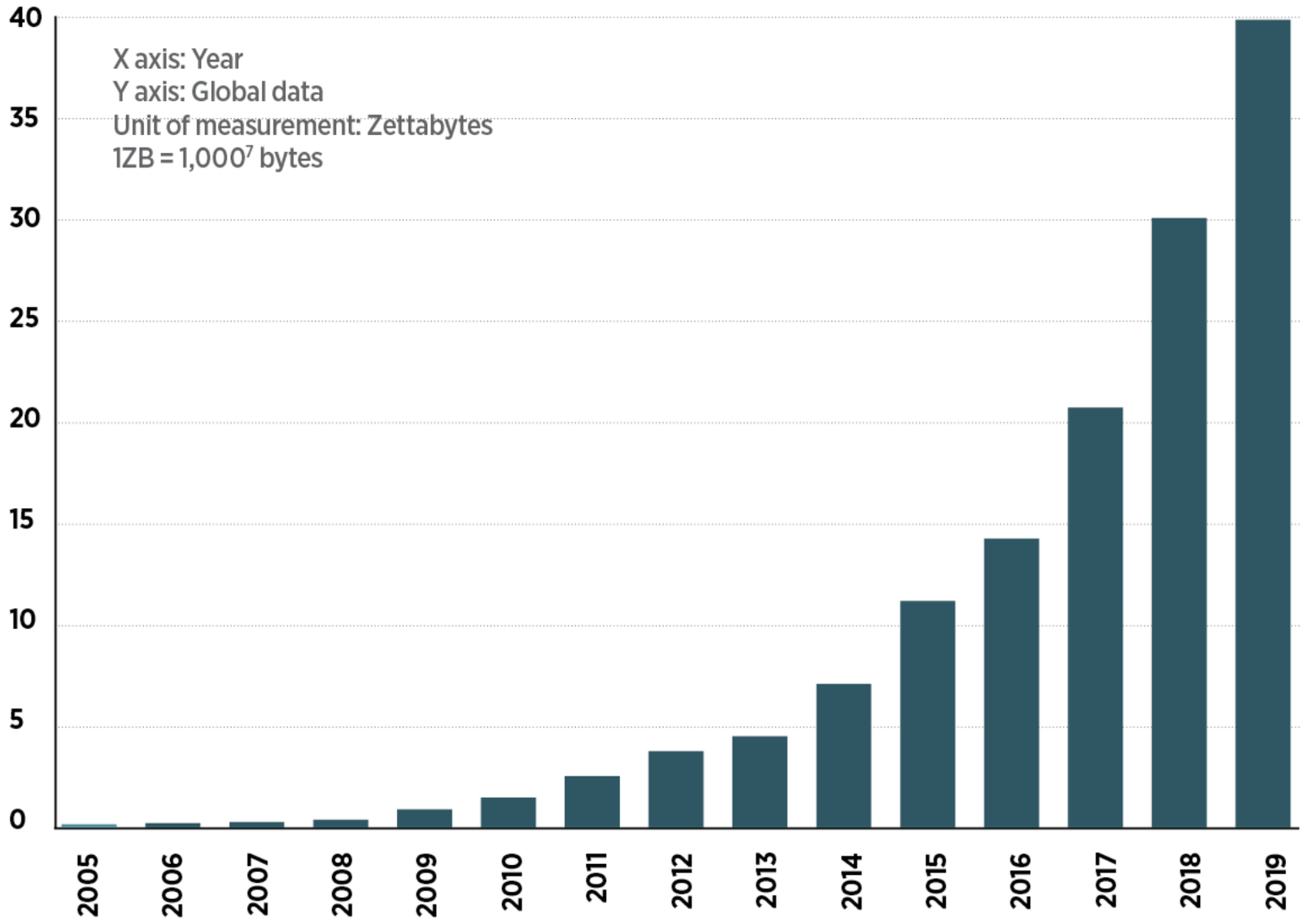
- You can ask questions in the Q&A pane
- Answer polls
- Talk to me on chat
- Raise your hands
- Expect delays from my end as I **do not get a notification** if you ask a question, so I will repeatedly have to check them on my own ☹
- I will try to add volunteers as class questioner / critic to clear doubts and ask questions verbally.

Please answer the poll out right now!!!!

Machine Learning / Why are we here?

"***We are drowning in information but starved for knowledge.***"

— John Naisbitt



The Booming Age of Data

30 trillion Web pages

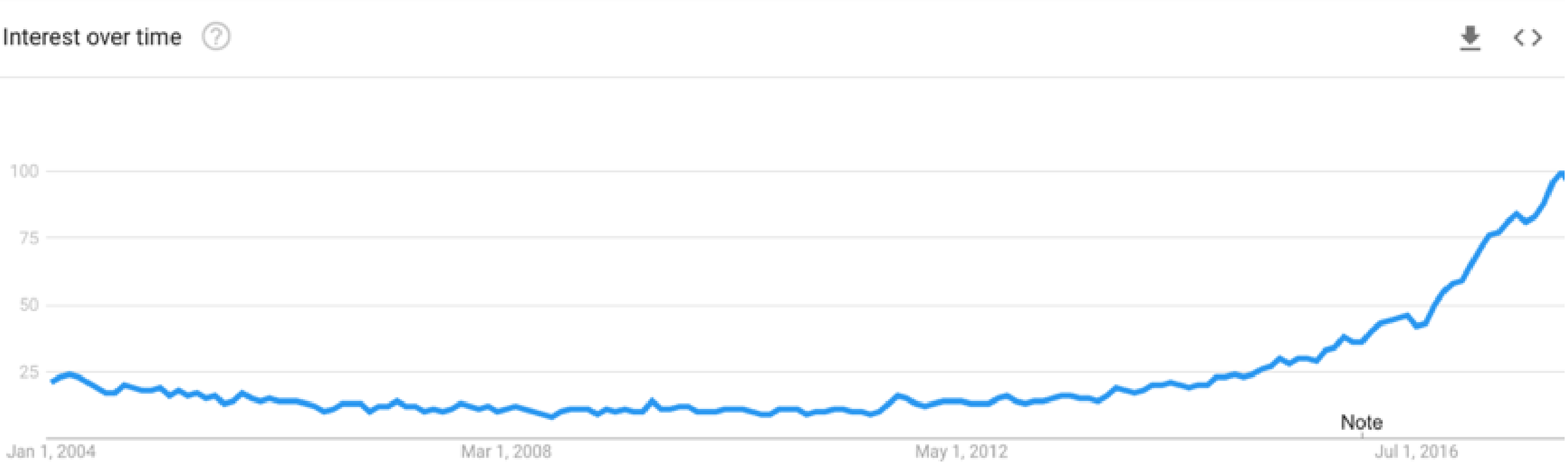
500 million tweets per day

2.27 billion monthly active users

1.8 billion images uploaded to Internet per day

2.9 billion base pairs in human genome

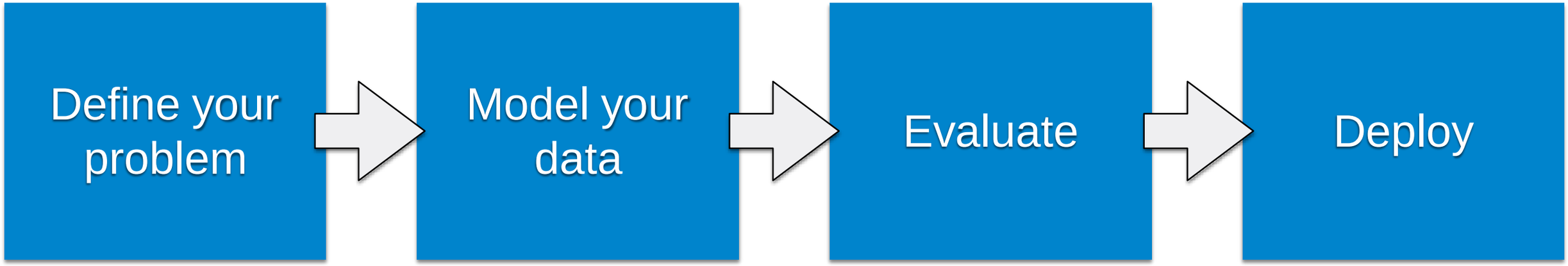
Interest in machine learning



Google trends, “Machine Learning (field of study)”.

Machine Learning

Machine Learning is the process of **turning data into actionable knowledge** for **task support** and **decision making**.

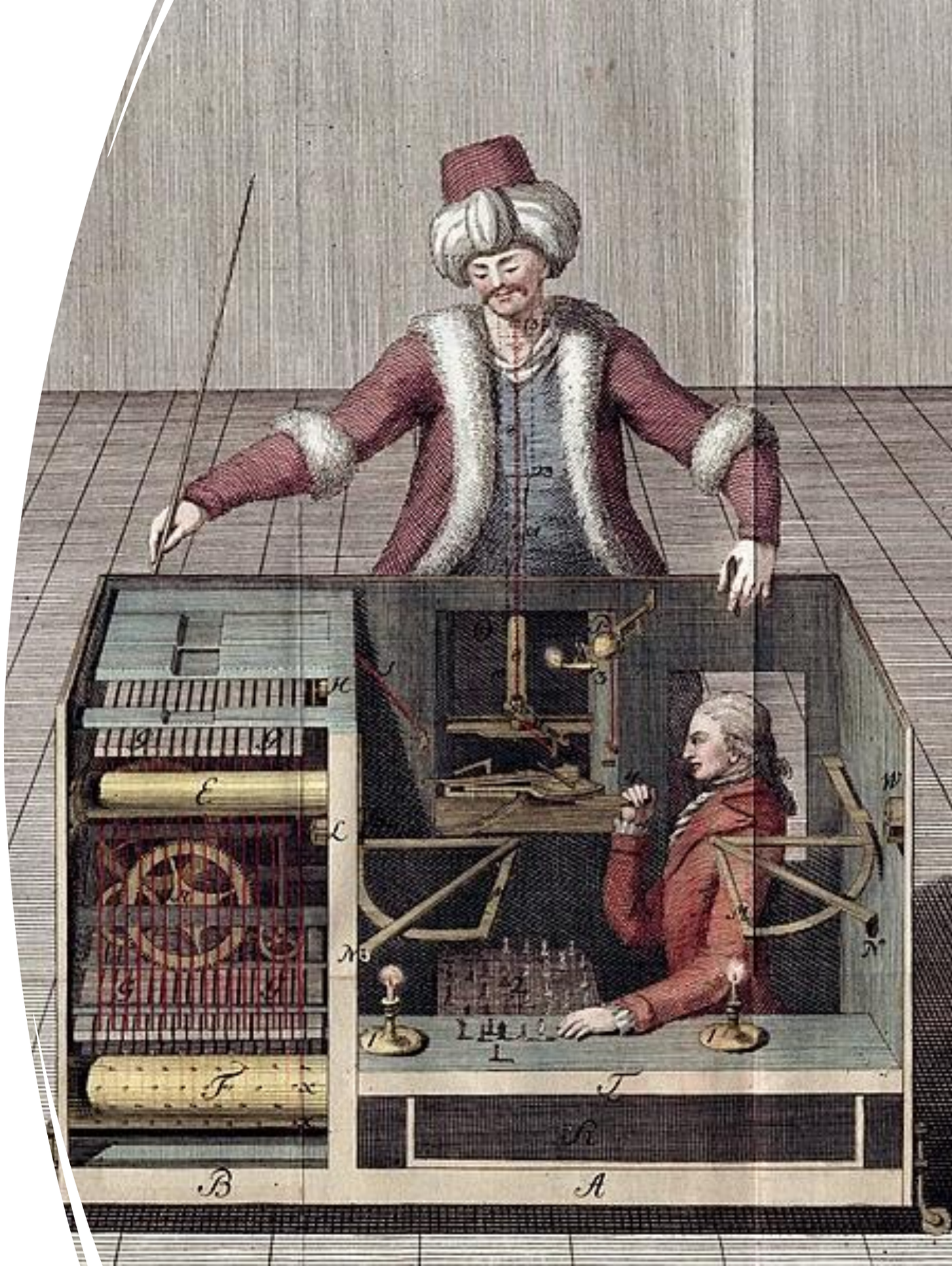


Course Objectives

- Introduce to you the **discipline of Machine Learning: Theory and Practice!**
- Help you understand **major machine learning algorithms**
- Help you learn to **apply tools** for **real data analysis problems**
- Encourage you to **do research** in data science and machine learning

Brief History of Machine Learning

- Mechanical Turk or Automaton Chess Player made by Wolfgang von Kempelen. image from Wikipedia



Brief History of Machine Learning

1950s

- **Samuel's checker player**
- Selfridge's Pandemonium
- **Dartmouth conference**
- **Turing test**
- **Neural networks: Perceptron**

1960s:

- **Connectionism**
- Pattern recognition Learning in the limit theory
- **Minsky and Papert prove limitations of Perceptron**

1970s:

- Symbolic concept induction Winston's arch learner
- Expert systems and the knowledge acquisition bottleneck Quinlan's ID3
- Michalski's AQ and soybean diagnosis Scientific discovery with BACON
- **Mathematical discovery with AM (Automated Mathematician)**

Slide Credit: Ray Mooney

Brief History of Machine Learning

1980s:

- **Advanced decision tree and rule learning**
- Explanation-based Learning (EBL) Learning and planning and problem solving Utility problem
- Analogy
- **Cognitive architectures**
- **Resurgence of neural networks (Connectionism 2 with backpropagation)**
- **Valiant's PAC Learning Theory**
- Focus on experimental methodology

1990s

- **Data mining**
- Adaptive software agents and web applications Text learning
- Reinforcement learning (RL) Inductive Logic Programming (ILP)
- **Ensembles: Bagging, Boosting, and Stacking**
- **Bayes Net learning**

Slide Credit: Ray Mooney

Brief History of Machine Learning

2000s:

Support vector machines
Kernel methods
Graphical models
Statistical relational learning
Transfer learning
Sequence labeling
Collective classification and structured outputs
Computer Systems Applications
Learning in robotics and vision

2010s:

Deep learning / Connectionism 3
Reinforcement learning
Generative models
Adversarial learning
Muti-task learning
Learning in NLP, CV, Robotics, …

Based on Ray Mooney's slide

Syllabus

Part I: Basic math for computational data analysis

- Probability, statistics, linear algebra

Part II: Supervised learning for predictive analysis

- Tree-based models, linear classification/regression, neural networks

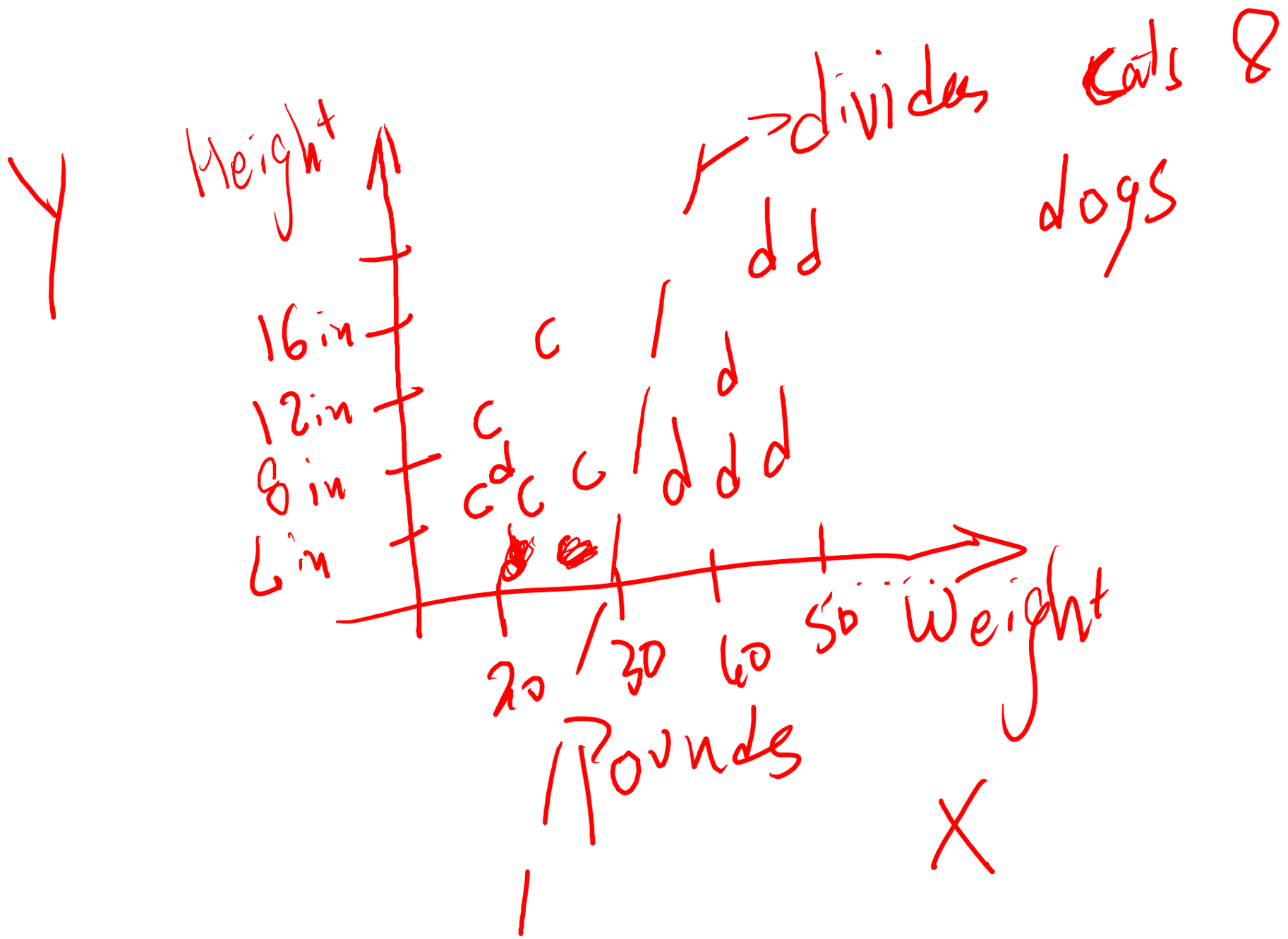
Part III: Unsupervised learning for data exploration

- Clustering analysis, dimensionality reduction, kernel density estimation

Part IV: Advanced topics for learning behaviors

- Reinforcement learning, Hidden Markov Models

The classic question Cat or Dog



Supervised learning

	Weight(lb)	Height(cm)	Fur color	Eye color		Label
Point 1	10	20	w	g		cat
Point 2	50	100	br	bl		dog
Point 3	8	15	bl	bl	=	dog
Point 4	12	25	w	bl		cat
Point 5	14	10	bl	g		dog
				$X_{n\times d}$		$Y_{n\times 1}$

$n \times 1$

Supervised just focuses on $X_{n\times d}$ and $Y_{n\times 1}$

Supervised learning

Domain set

Label set

	Weight(lb)	Height(cm)	Fur color	Eye color		Label
Point 1	10	20	w	g		cat
Point 2	50	100	br	bl		dog
Point 3	8	15	bl	bl	=	dog
Point 4	12	25	w	bl		cat
Point 5	14	10	bl	g		dog
				$X_{n\times d}$		$Y_{n\times 1}$

Supervised just focuses on $X_{n\times d}$ and $Y_{n\times 1}$

Supervised learning

Missing pieces:

- A loss function that measures risk.
- An optimizer that gets us to the solution to the problem
- Data

Will discuss these over the semester!!

Syllabus: Supervised Learning

Tree-based models

- Decision tree
- Ensemble learning/Random forest

Linear classification/regression models

- Linear regression
- Naive Bayes
- Logistic regression
- Support vector machine

Neural networks

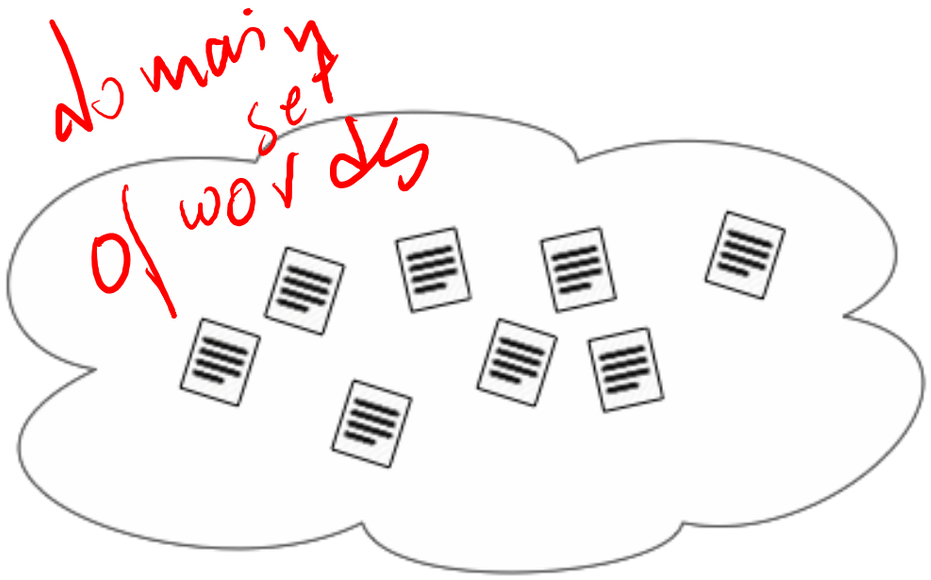
- Feedforward neural networks and backpropagation analysis

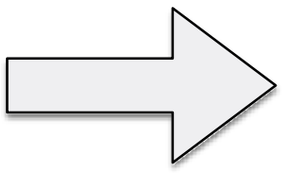
News Classification

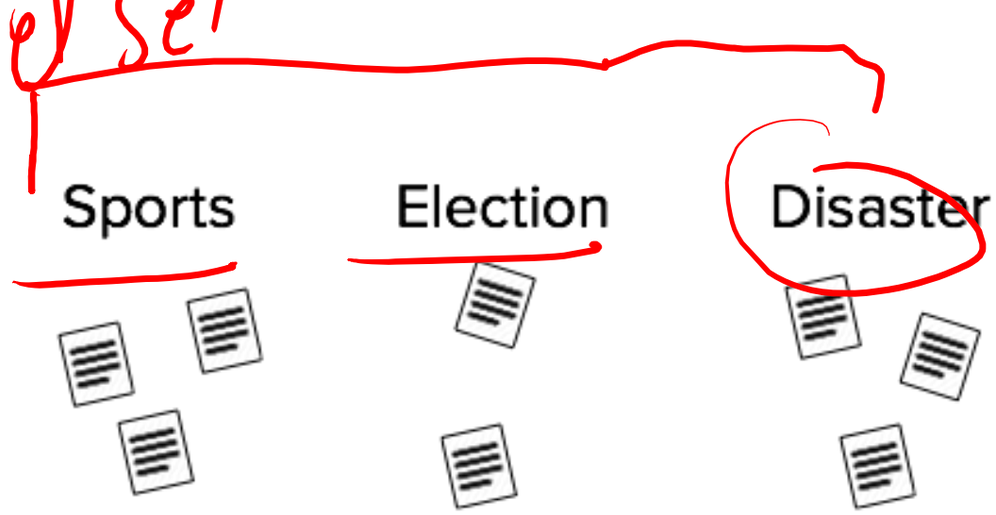
What are the inputs and how to represent them?

What are the desired outputs?

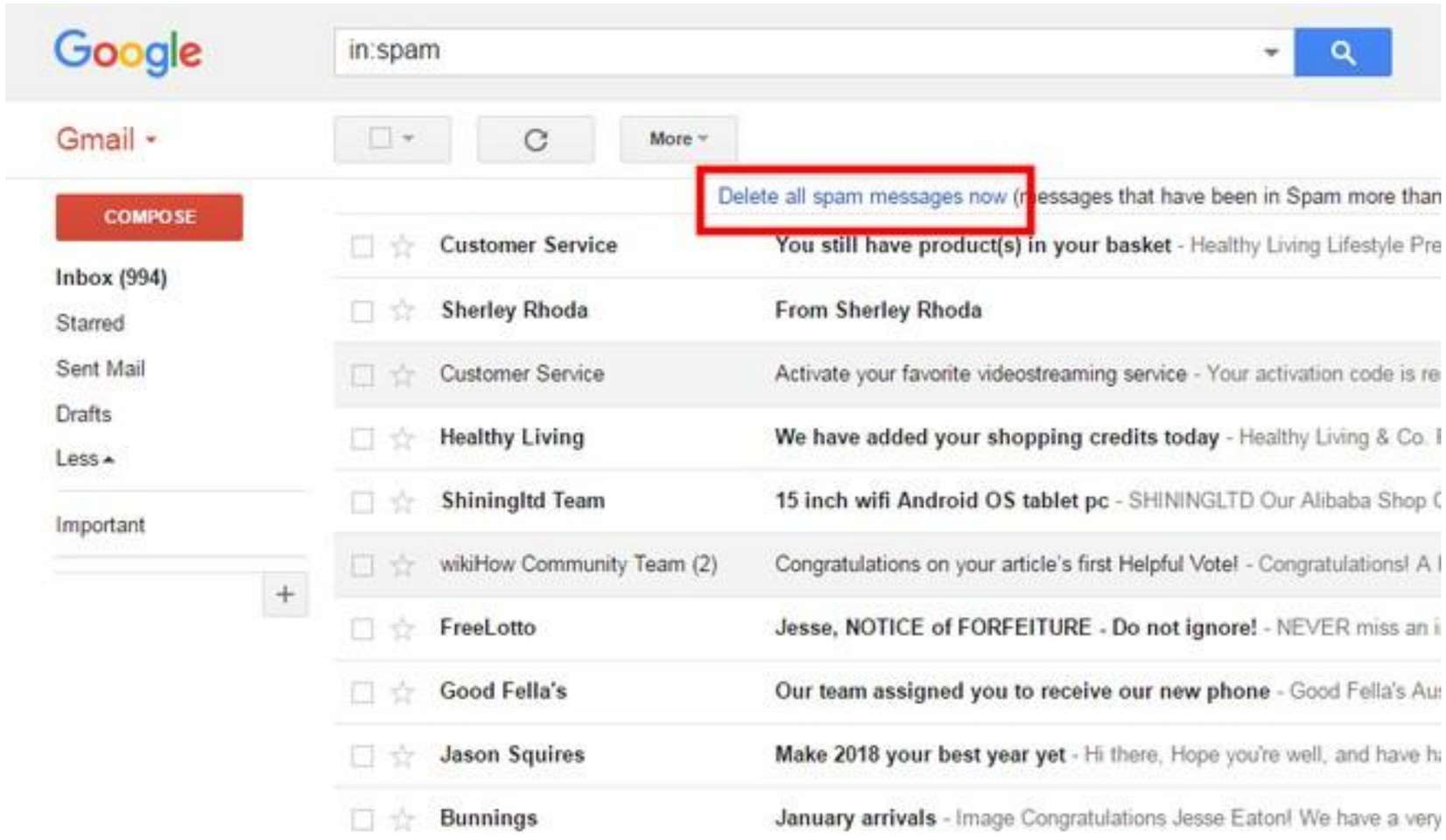
What learning algorithms to choose?

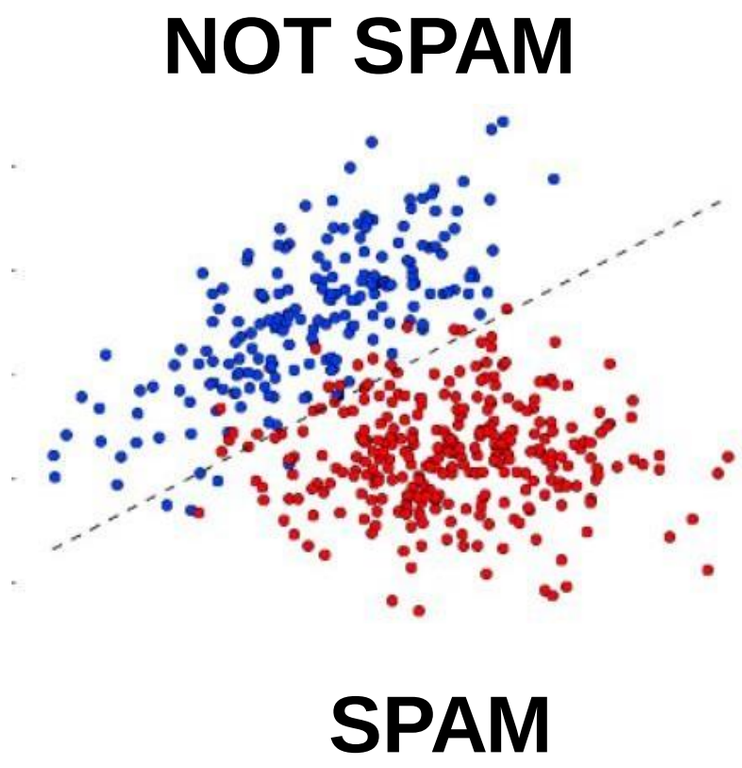






Spam Detection





What are the inputs and how to represent them?

What are the desired outputs?

What learning algorithms to choose?

Image credit: WikiHow

Unsupervised Learning

$$
\begin{array}{l}
 & \begin{array}{cccc} \text{Weight(lb)} & \text{Height(cm)} & \text{Fur color} & \text{Eye color} \end{array} \\
\begin{array}{l} Point\ 1 \\ Point\ 2 \\ Point\ 3 \\ Point\ 4 \\ Point\ 5 \end{array} & \underbrace{\begin{bmatrix} 10 & 20 & w & g \\ 50 & 100 & br & bl \\ 8 & 15 & bl & bl \\ 12 & 25 & w & bl \\ 14 & 10 & bl & g \end{bmatrix}}_{X_{n\times d}} = \underbrace{\begin{bmatrix} cat \\ dog \\ dog \\ cat \\ dog \end{bmatrix}}_{Y_{n\times 1}}
\end{array}
\quad \text{Label}
$$

Unsupervised Learning

	Weight(lb)	Height(cm)	Fur color	Eye color
Point 1	10	20	w	g
Point 2	50	100	br	bl
Point 3	8	15	bl	bl
Point 4	12	25	w	bl
Point 5	14	10	bl	g

$X_{n \times d}$

Examples in human learning:

- Taxonomy in Biology
- Periodic Table in Chemistry

Unsupervised Learning

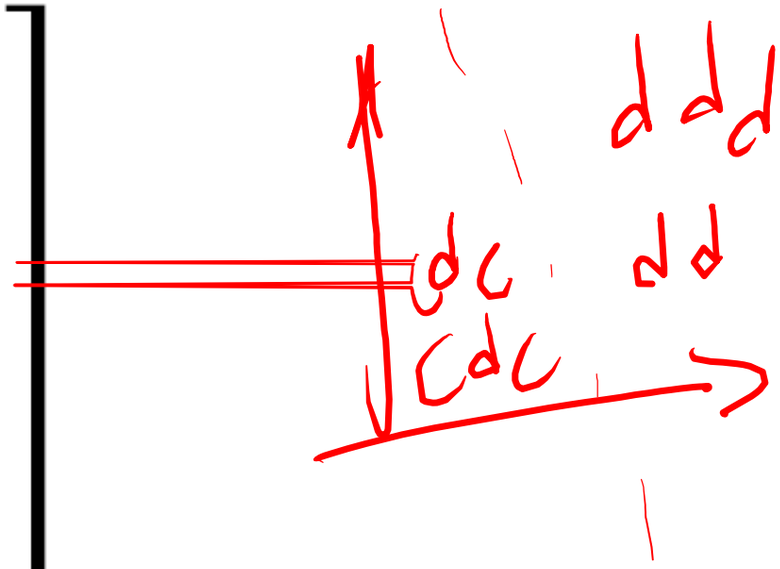
Still Raw from [xkcd]

Planet categorizing run rampant!!

Unsupervised Learning

	Weight(lb)	Height(cm)	Fur color	Eye color
Point 1	10	20	w	g
Point 2	50	100	br	bl
Point 3	8	15	bl	bl
Point 4	12	25	w	bl
Point 5	14	10	bl	g

$X_{n \times d}$



Unsupervised just focuses on $X_{n \times d}$

Unsupervised learning

Missing pieces:

- A loss function that measures risk. What does that mean in this case?
- An optimizer that gets us to the solution to the problem
- Data

Will discuss these over the semester!!

Syllabus: Unsupervised Learning

Clustering Analysis

- K-means
- Gaussian mixture model
- Density-based clustering
- Evaluation of clustering algorithms

Dimension Reduction

- Principal component analysis

Kernel Density Estimation

- Parametric density estimation
- Non-parametric density estimation

Community Detection in Social Networks

What are the inputs and how to represent them?

What are the desired outputs?

What learning algorithms to choose?

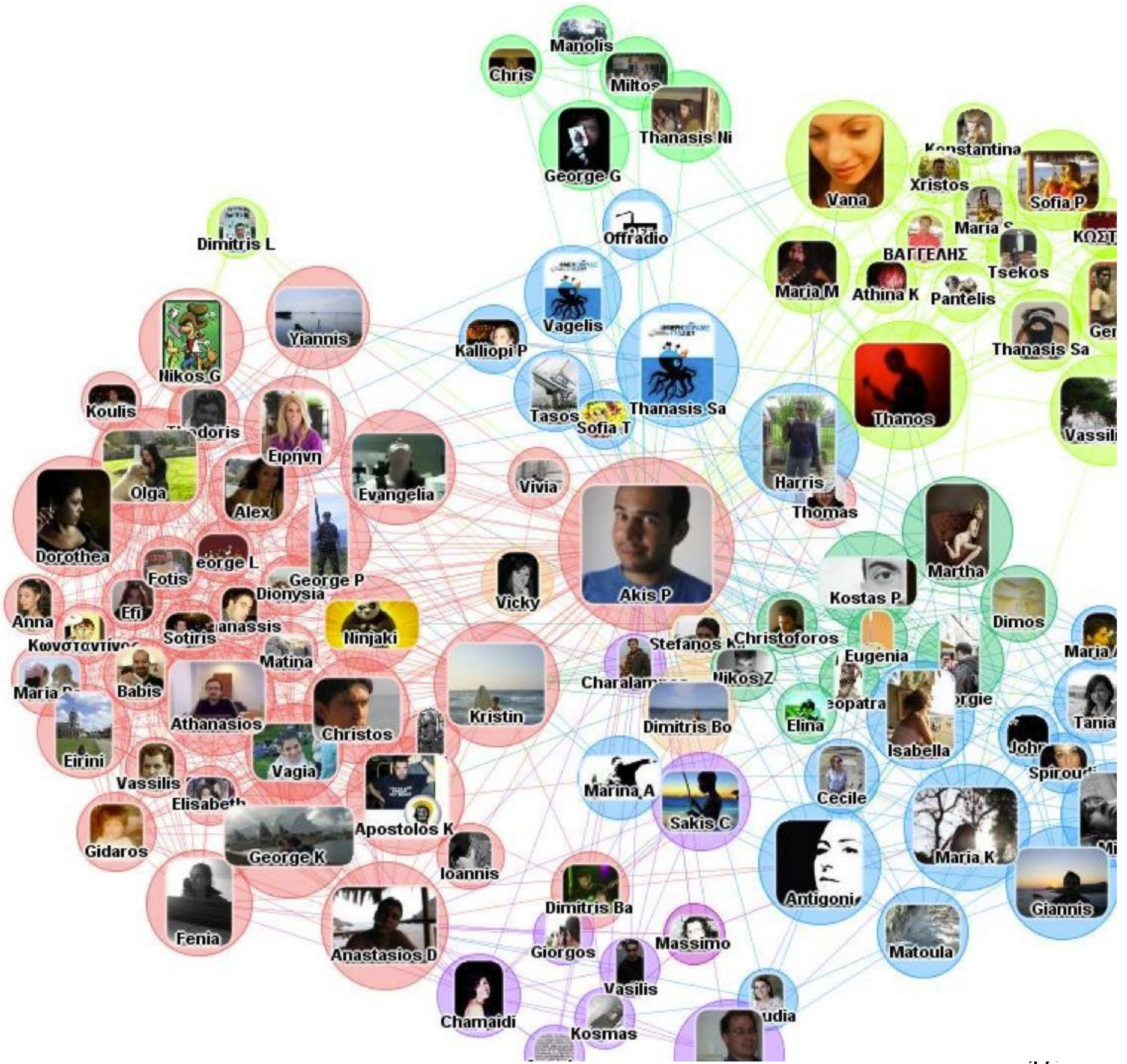
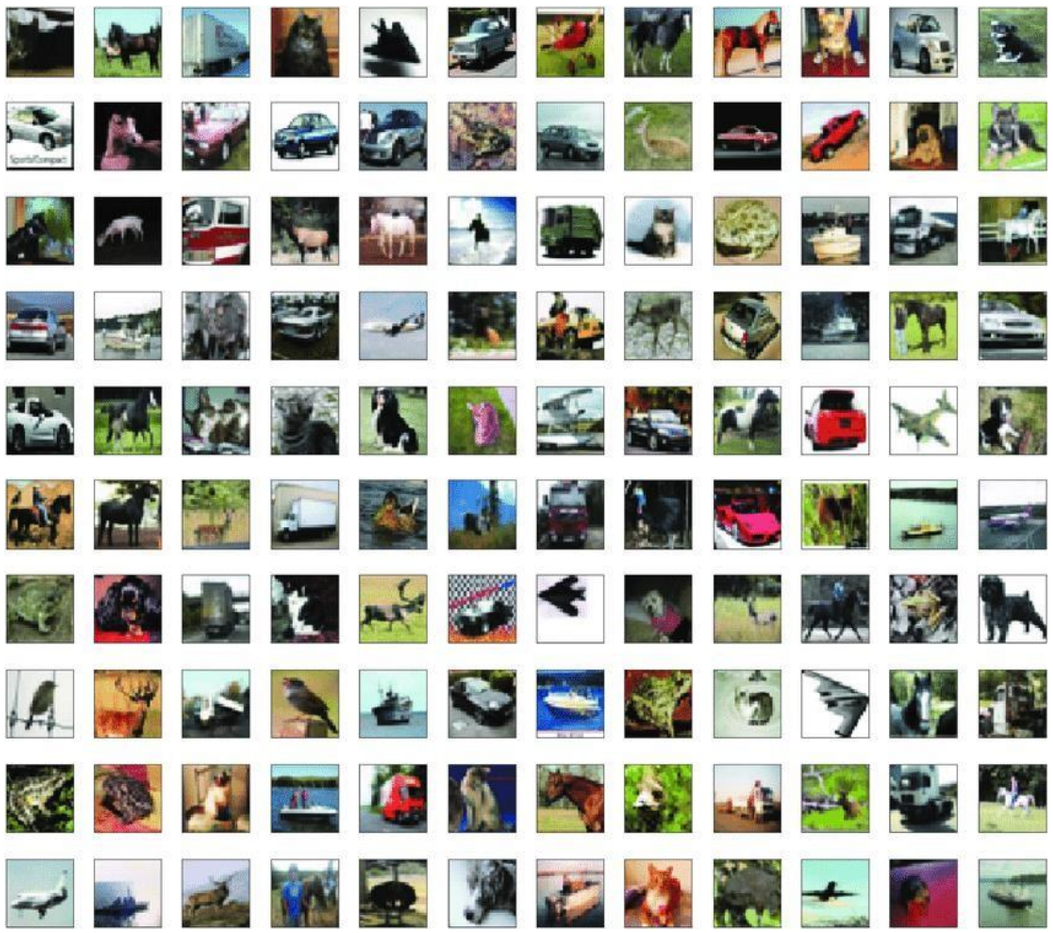
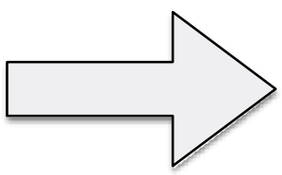
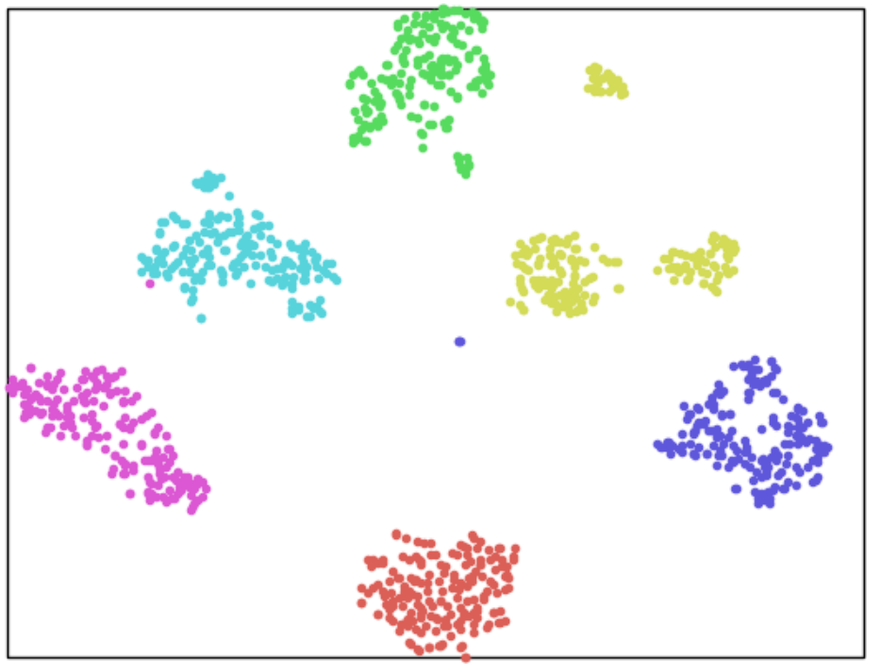


Image credit: Symeon Papadopoulos

Dimensionality Reduction



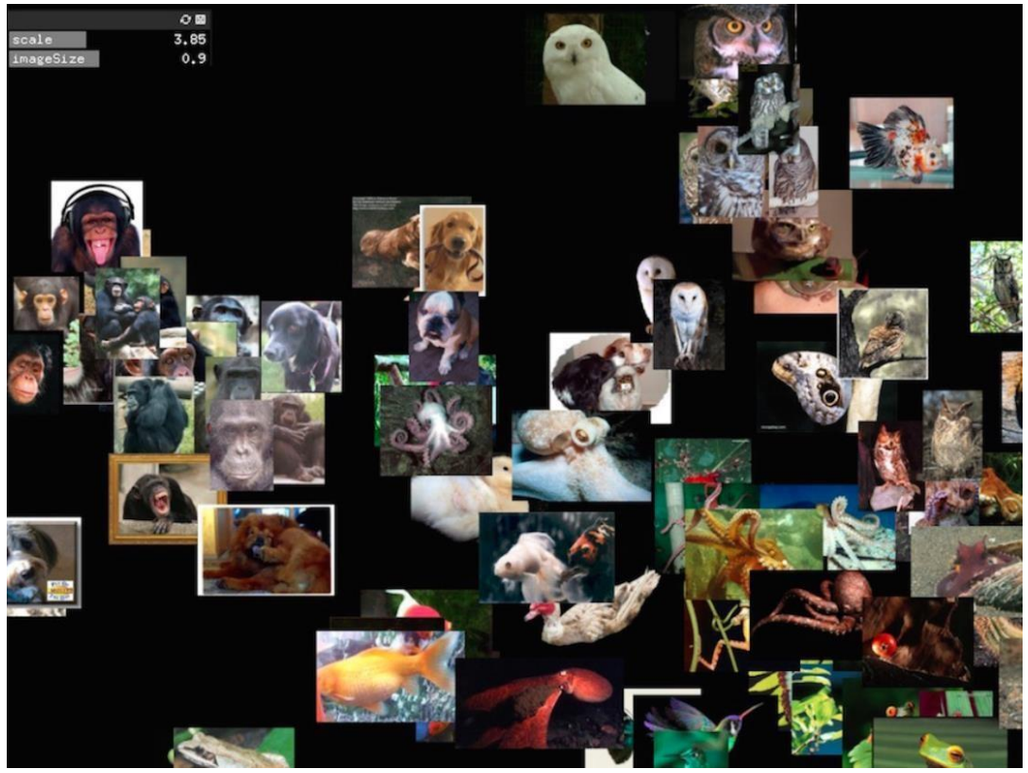




What are the inputs and how to represent them?

What are the desired outputs?

What learning algorithms to choose?



mage credit: Gene Kogan

Images from waymo.com

What is missing??

Images from waymo.com

Images from wikipedia.com

Images from waymo.com

Actions

Image from wikipedia

action

Image from amazon.com

Search

Images from waymo.com

Advanced topics

Reinforcement Learning

How does “an agent” take actions in the world?

Advanced topics

Reinforcement Learning

How does “an agent” take actions in the world?

Video that inspired me to work in robotics and RL: https://youtu.be/5oBAYbOF2Qo?t=50

Incredibly hard problem of learning the model of a robot arm well enough to swing it like that with raw torques.

Kober et al. 2008

Advanced topics

Reinforcement Learning

How does “an agent” take actions in the world?

Hidden Markov Model

How to make predictions about the future?
What is the weather tomorrow like?

Repeated theme over the course

ML algorithm = representation + loss function + optimizer

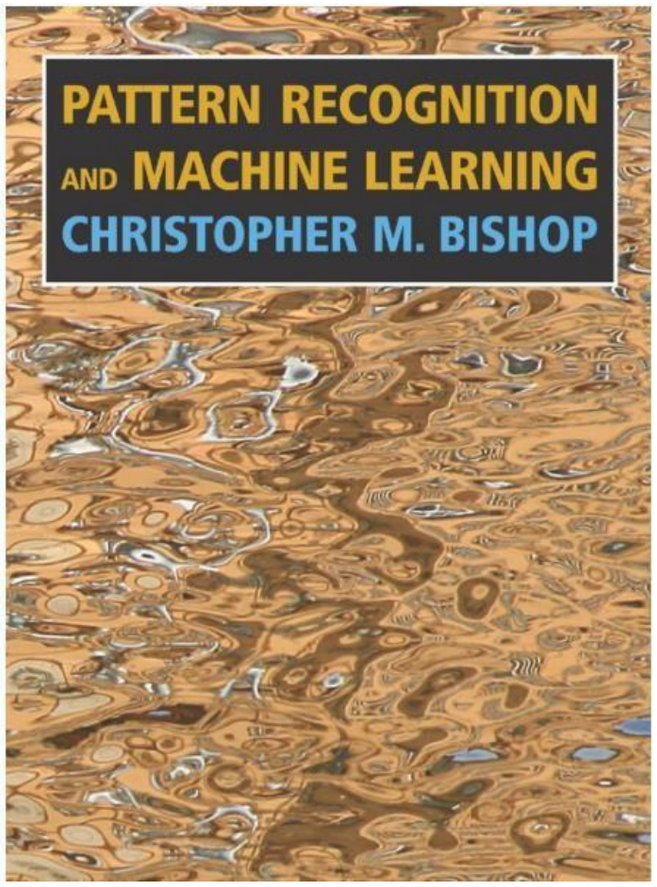
Administrative Details

Prerequisites

Basic knowledge in probability, statistics, and linear algebra

Basic programming skills in Python (Jupyter Notebook)

No background in machine learning is required



Text Books

Pattern Recognition and Machine Learning, by Chris Bishop

↖ Fundamental

→ Preferred book for beginners

Other recommended books:

Probabilistic Machine Learning by Kevin P. Murphy – Free for now

Learning from data, by Yaser S. Abu-Mostafa

Machine learning, by Tom Mitchell

Deep Learning, by Ian Goodfellow, Yoshua Bengio, and Aaron Courville

Assignments

Four assignments (GradeScope)

Each can include written analysis or programming

Late policy

Three (3) late days for the entire semester to be used only on assignments.

Don’t copy

Any academic misconduct is subject to F grade as well as reporting to the Dean of students.

Projects

- Work on a real-life Machine Learning problem
  - What is the problem? What is your method? How do you evaluate it?
- Exactly 5 people in a team (~~Grad and undergrad can~~'t be ~~mixed in a group~~)
- GitHub Pages (index.html)
- Start your projects early
- Ask for comments and feedbacks from the teaching staff

Quizzes

- To test knowledge weekly
- ~7 mins.
- Out every Thursday and needs to be completed on the same day
- About 14 quizzes with a practice quiz
- Practice quiz out next week.
- Top 9 or 10 counted for your final grade.

Grading Policy

Assignments (50%)

Four assignments; programming or written analysis

Project (35%)

5 people in a team (not less and not more); should be done using GitHub Pages)

Quizzes (10%)

About 15 quizzes, we will consider your 10 topmost score

Class participation(5%)

Piazza

- TODO: Sign up for the class on Piazza from Canvas
- Emails can get lost and you can help you peers on Piazza. Do not email us. Try private chats with us on Piazza for serious concerns.