CS/ECE 760: Machine Learning
Course Overview

Josiah Hanna

University of Wisconsin-Madison

September 7, 2023

ML Overview: Motivation

Why machine learning?

- We’re building a self-driving car. Could just write down rules
  - **Painful**! A lot of cases…

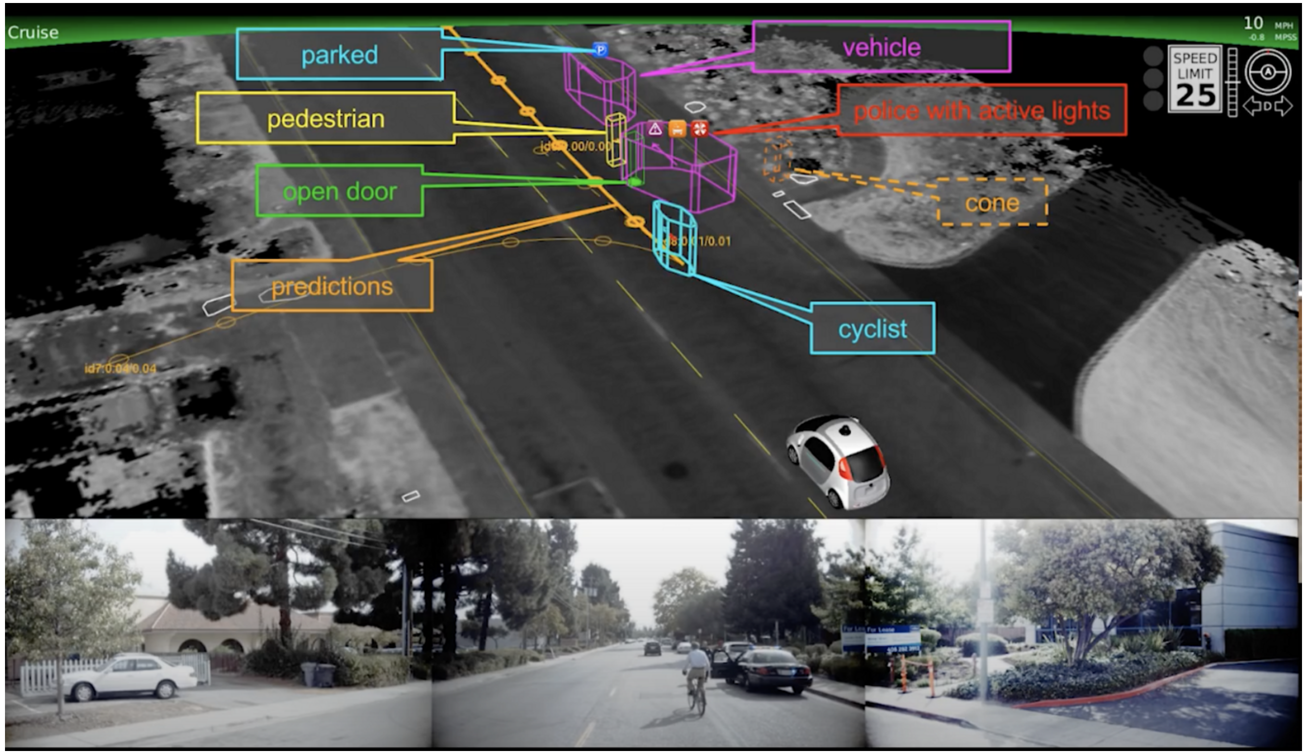
```
/**
 * controls steering of the car
 * @param angle
 * @param trim
 */
void steer(float angle, float trim = 0.0) {
    // seems like 360 right 520 left
    PWMPCA9685Device device = new PWMPCA9685Device()
    device.setPWMFrequency(50) //internet says 50hz for servos is optimal
    Servo servo0 = new PCA9685Servo(device.getChannel( channel: 1))
    LOG.info("steer angle non corrected:${angle} trim:${trim}")
    if (trim != 0) {
        trim = configTrim
        servo0.setTrim(trim)
    }
    servo0.setInput((angle).toFloat())
    System.out.println("configTrim in service=${configTrim}")
    Thread.sleep( millis: 1000) // important to give time for servo to move

}
```

Car Image Credit: Wikipedia

ML Overview: Motivation

Why machine learning?

- We're building a self-driving car. Could just write down rules
  - **Painful**! A lot of cases…
  - **Learn from examples** instead

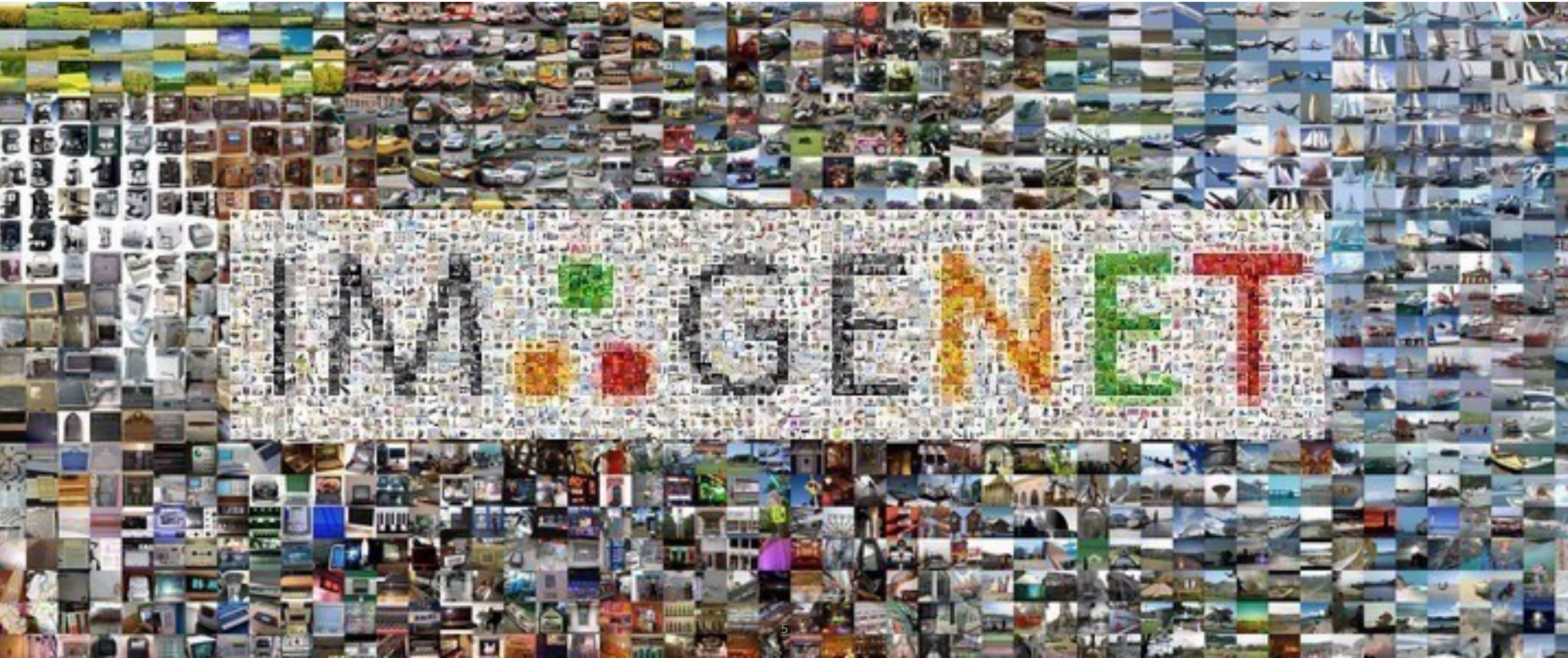


Waymo

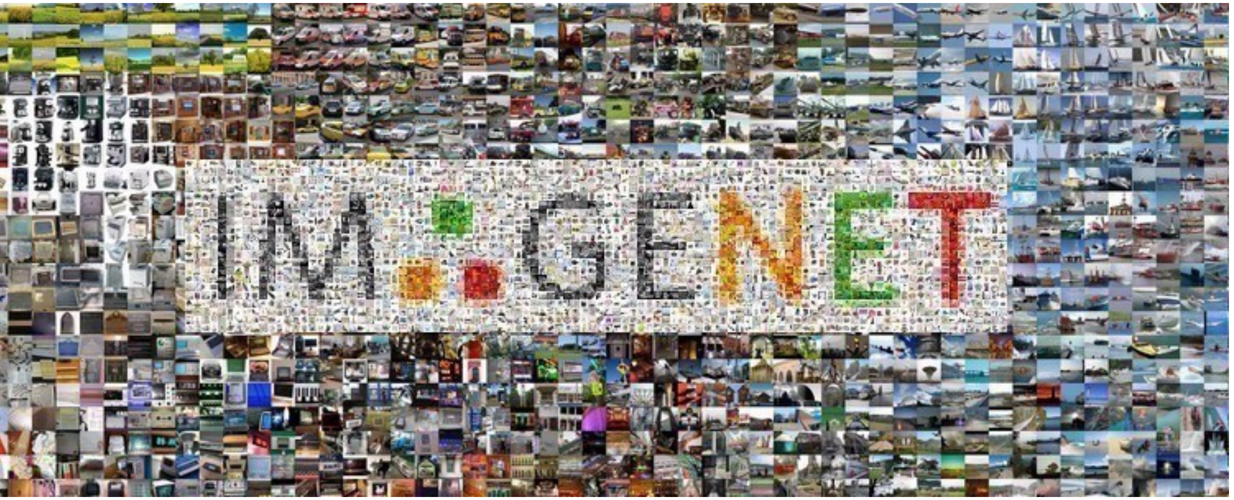
Lion or Tiger?

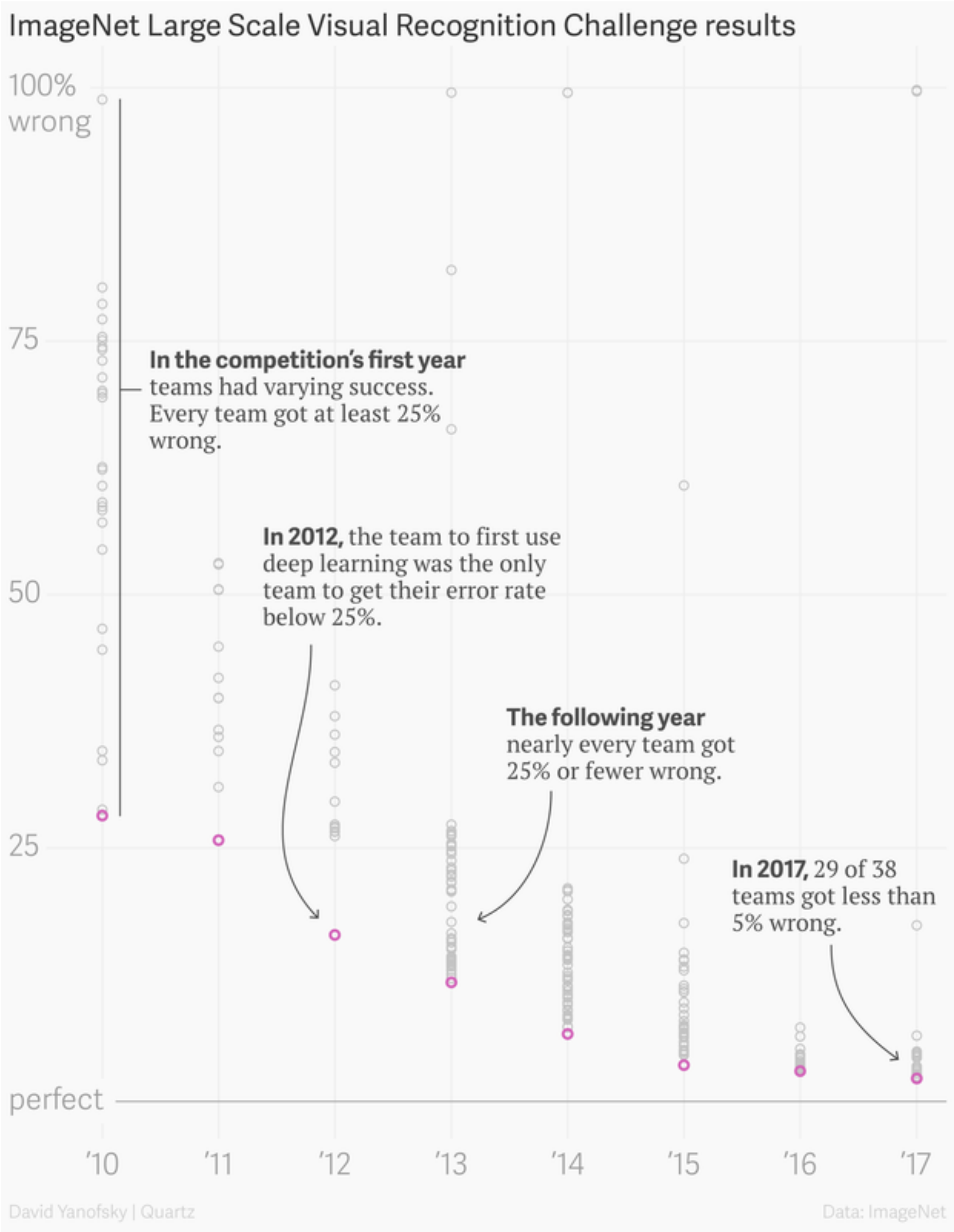
Classify Images

http://www.image-net.org/



Classify Images





Detect and Segment Objects

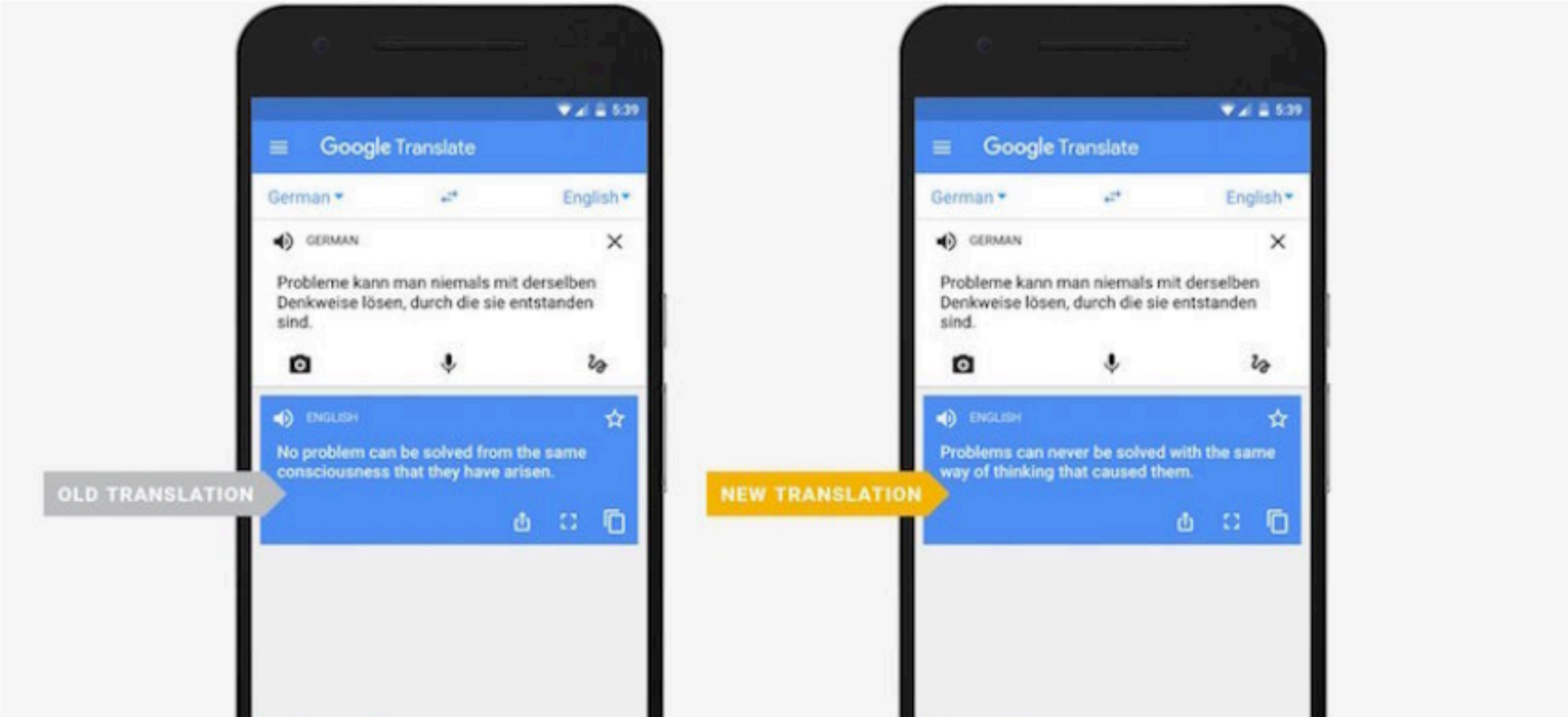
https://github.com/matterport/Mask_RCNN

Machine Translation

Google Expands Neural Networks for Language Translation

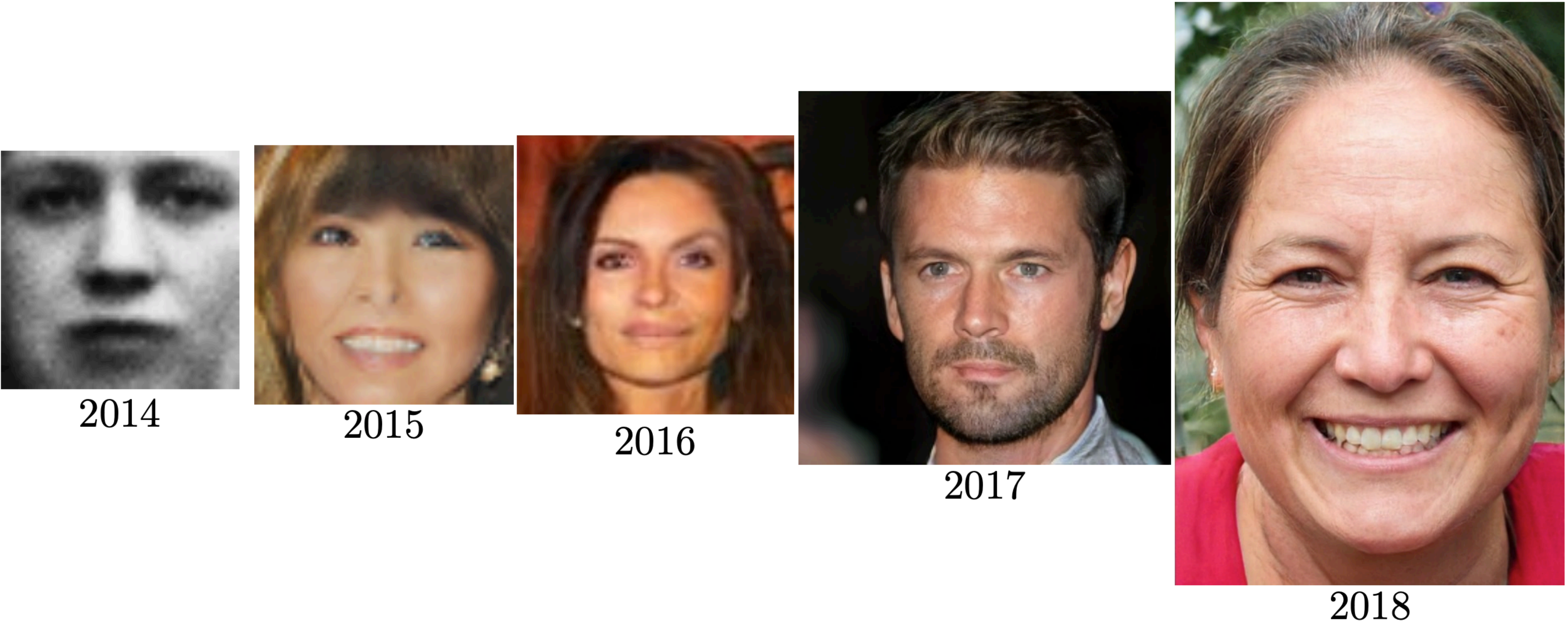
The new system can translate whole sentences at a time, rather than just phrases.

By Tom Brant November 15, 2016



Generative Modeling

4.5 years of face generation



http://www.whichfaceisreal.com/methods.html

Open AI's ChatGPT

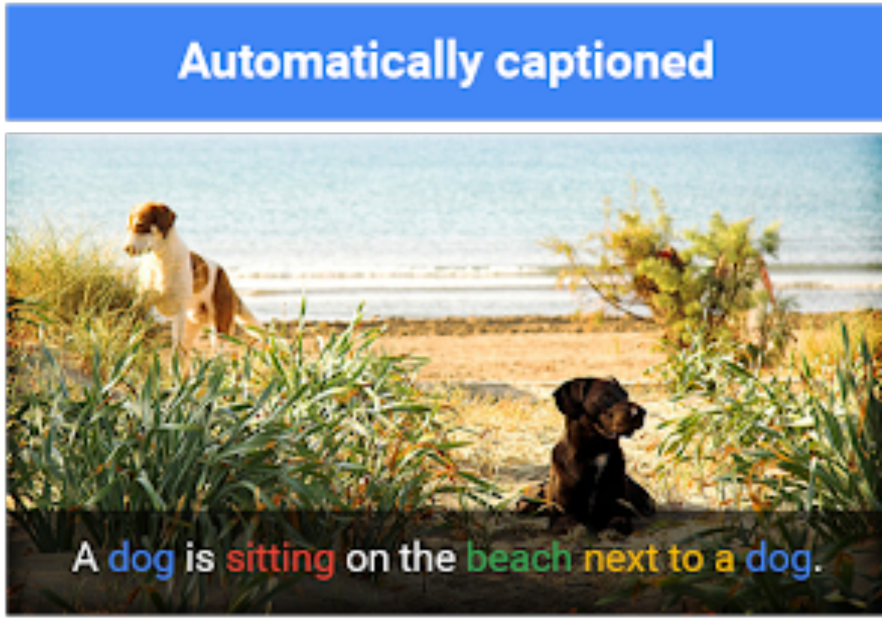
HA

Why should graduate students take a machine learning course?

Graduate students in various fields can benefit significantly from taking a machine learning course for several reasons:

1. **Relevance to Modern Research and Industry**: Machine learning techniques are becoming increasingly essential in various research domains and industries. Whether you're pursuing a career in academia or industry, having a strong foundation in machine learning can make you more competitive and relevant.
2. **Interdisciplinary Applications**: Machine learning is not limited to computer science. It finds applications in fields like healthcare, finance, natural language processing, computer vision, and more. Graduate students from diverse backgrounds can leverage machine learning to enhance their research or work.
3. **Problem-Solving Skills**: Machine learning courses teach valuable problem-solving skills that can be applied to a wide range of complex problems. These skills include data analysis, modeling, and algorithm design, which are beneficial in various research and professional contexts.

Image Captioning



Defeat Human Players at Board Games

https://qz.com/639952/googles-ai-won-the-game-go-by-defying-millennia-of-basic-human-instinct/

Machine learning used in many other fields

- Other areas of computer science
  - Distributed systems, computer architecture etc.
- Electrical engineering
- Industrial engineering
- Physics
- Materials science
- Drug discovery
- Finance & economics
- E-commerce
- ….

Motivation from Artificial Intelligence

The ability to learn from experience is one of the first capabilities we consider when contemplating the nature of intelligence.

Outline for today

- Overview of machine learning
- Who should take this class? Prerequisites and expectations
- Syllabus and class logistics

Overview of machine learning

ML Overview: Definition

What is machine learning?

"A computer program is said to learn from experience **E** with respect to some class of tasks **T** and performance measure **P**, if its performance at tasks in **T** as measured by **P**, improves with experience **E**." *Machine Learning*, Tom Mitchell, 1997

ML Overview: Flavors

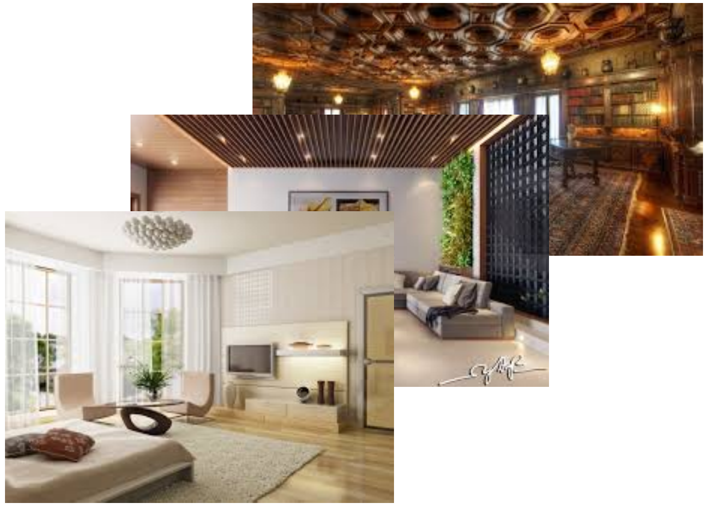
Three broad types of machine learning:

- Supervised Learning
- Unsupervised Learning
- Reinforcement Learning

ML Overview: Flavors

Supervised Learning

- Learning from examples, as above
- **Workflow**:
  - Collect a set of examples {data, labels}: **training set**
  - “**Train**” a model to match these examples
  - “**Test**” it on new data
- **Image classification**:



indoor

outdoor

ML Overview: Flavors

Supervised Learning

- **Example: Image classification**
- Recall **T**ask/**P**erformance measure/**E**xperience definition
  - **T**ask: distinguish **indoor** vs **outdoor**
  - **P**erformance measure: probability of misclassifying
  - **E**xperience: labeled examples

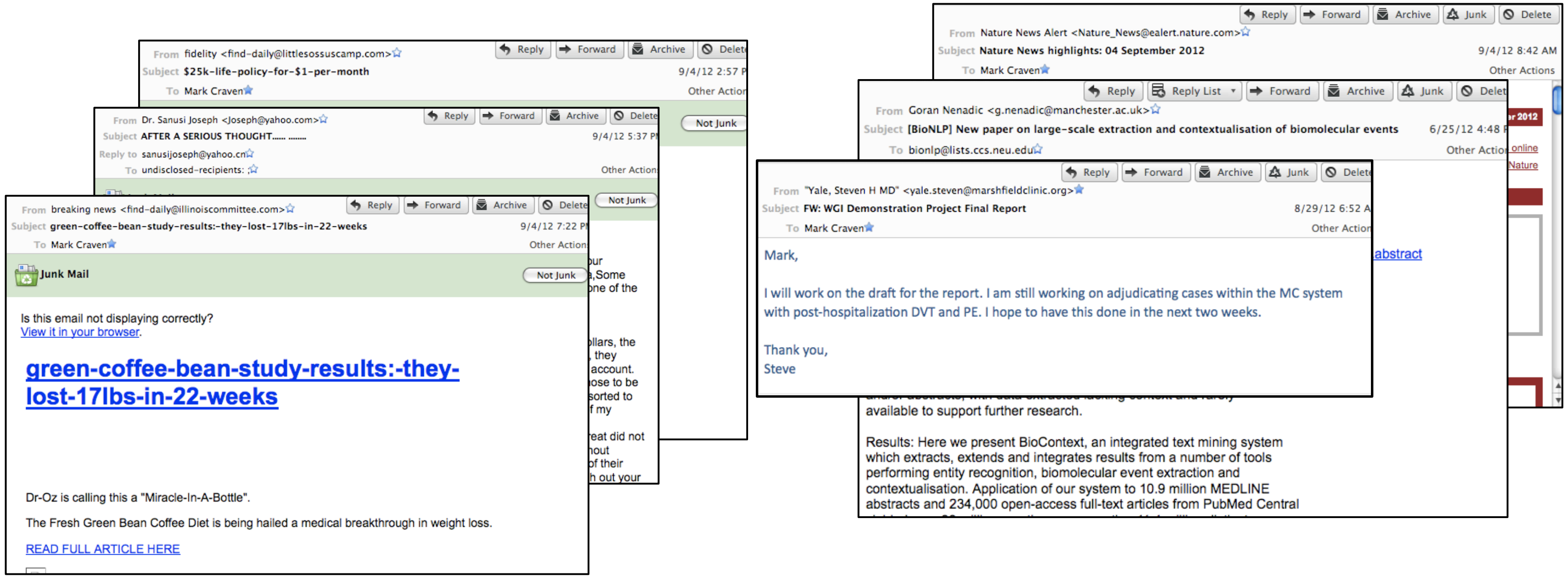
indoor

outdoor

ML Overview: Flavors

Supervised Learning

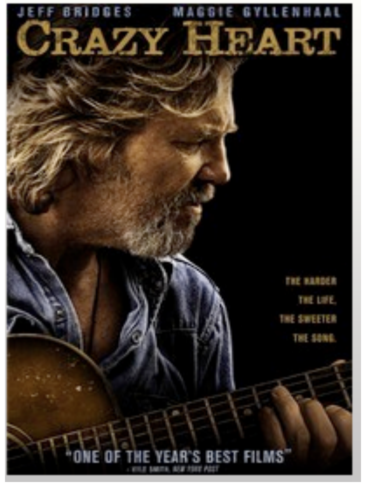
- **Example: Spam Filtering**
  - Task: distinguish **spam** vs **legitimate**
  - Performance measure: probability of misclassifying
  - Experience: labeled examples of messages/emails



ML Overview: Flavors

Supervised Learning

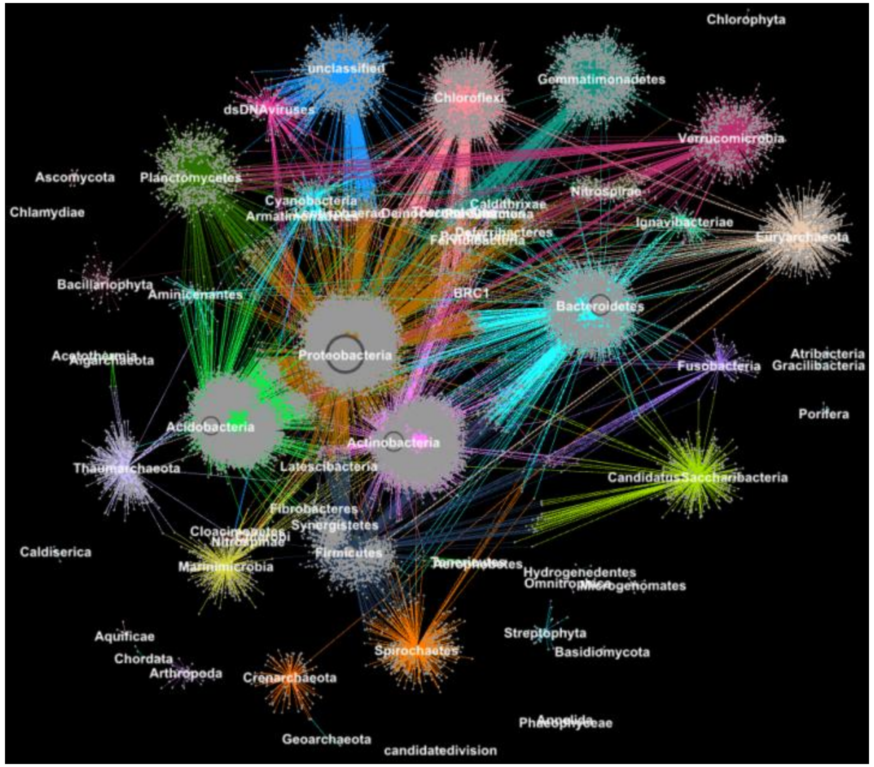
- **Example: Ratings/Recommendations**
  - **T**ask: predict how much a user will like a film
  - **P**erformance measure: difference between prediction and user's true rating
  - **E**xperience: previous ratings



ML Overview: Flavors

Unsupervised Learning

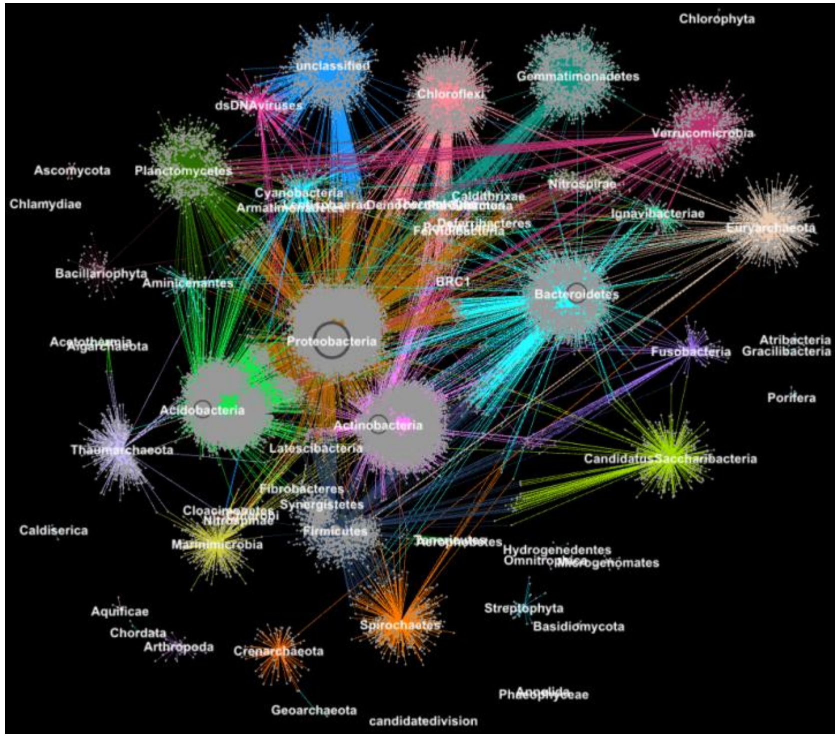
- Data, but no labels. No input/output.
- Goal: "find something": structure, hidden information, etc
- **Workflow**:
  - Collect a set {data}
  - Perform some algorithm on it and draw insights about data
  - Sometimes: test on new data



ML Overview: Flavors

Unsupervised Learning

- **Example: Clustering**
  - **T**ask: divide datapoints into distinct groups, aka clusters.
  - **P**erformance measure: closeness to underlying structure
  - **E**xperience: available datapoints



ML Overview: Flavors

Unsupervised Learning

- **Example: Generative Models**
  - **T**ask: produce artificial images of faces
  - **P**erformance measure: photorealism
  - **E**xperience: available images

StyleGAN2 (Kerras et al ’20)

ML Overview: Flavors

Reinforcement Learning

- Agent interacting with the world; gets rewards for actions
- Goal: learn to perform some activity
- **Workflow**:
  - Create a training environment, define reward
  - **Train**: train a policy to maximize rewards
  - **Deploy** in new environment

ML Overview: Flavors

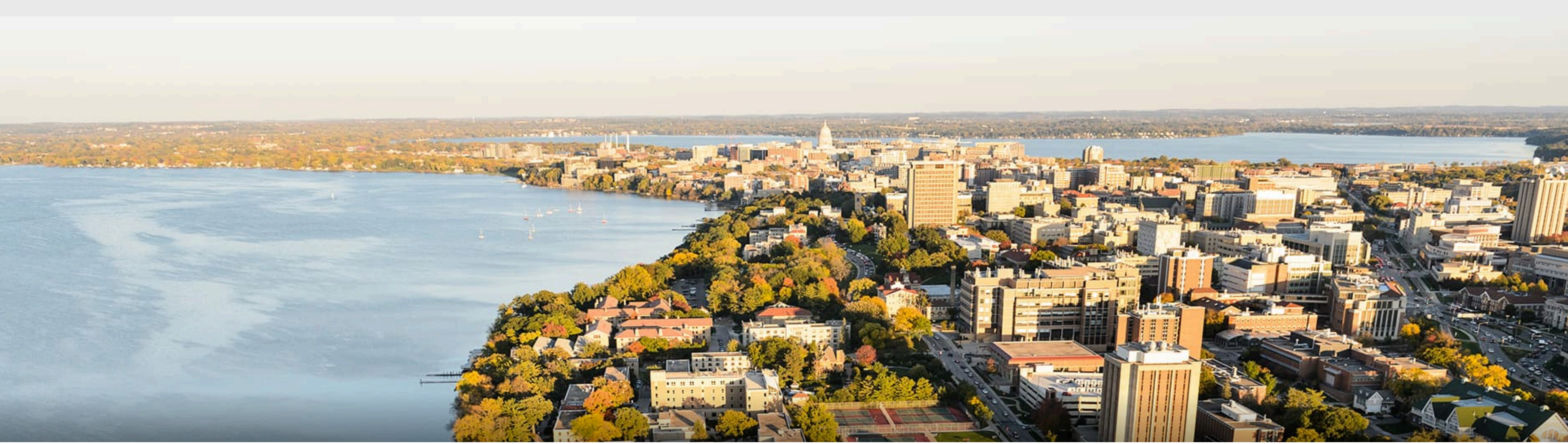
Reinforcement Learning

- **Example: Controlling aircraft**
  - **T**ask: keep the aircraft in the air, steer towards a desired goal
  - **P**erformance measure: reward for reaching goal quickly
  - **E**xperience: data (state/action/reward) from previous flights

ML Overview: Flavors

Reinforcement Learning

- **Example: Playing video games**
  - **T**ask: play Atari arcade games
  - **P**erformance measure: winning/advancing
  - **E**xperience: state/action/reward from previous gameplay episodes



Who should take this class?

Target audience for the course

Students who:

- Want to do research in ML
  - CS760 will lay the foundations of several topics in ML, but will likely not be sufficient on its own to advance a topic.
- Want to use ML in other research areas.

If you just want to **use** ML, but do not plan to do research, consider taking:

- CS540
- STAT 451
- ECE/CS/ME 532

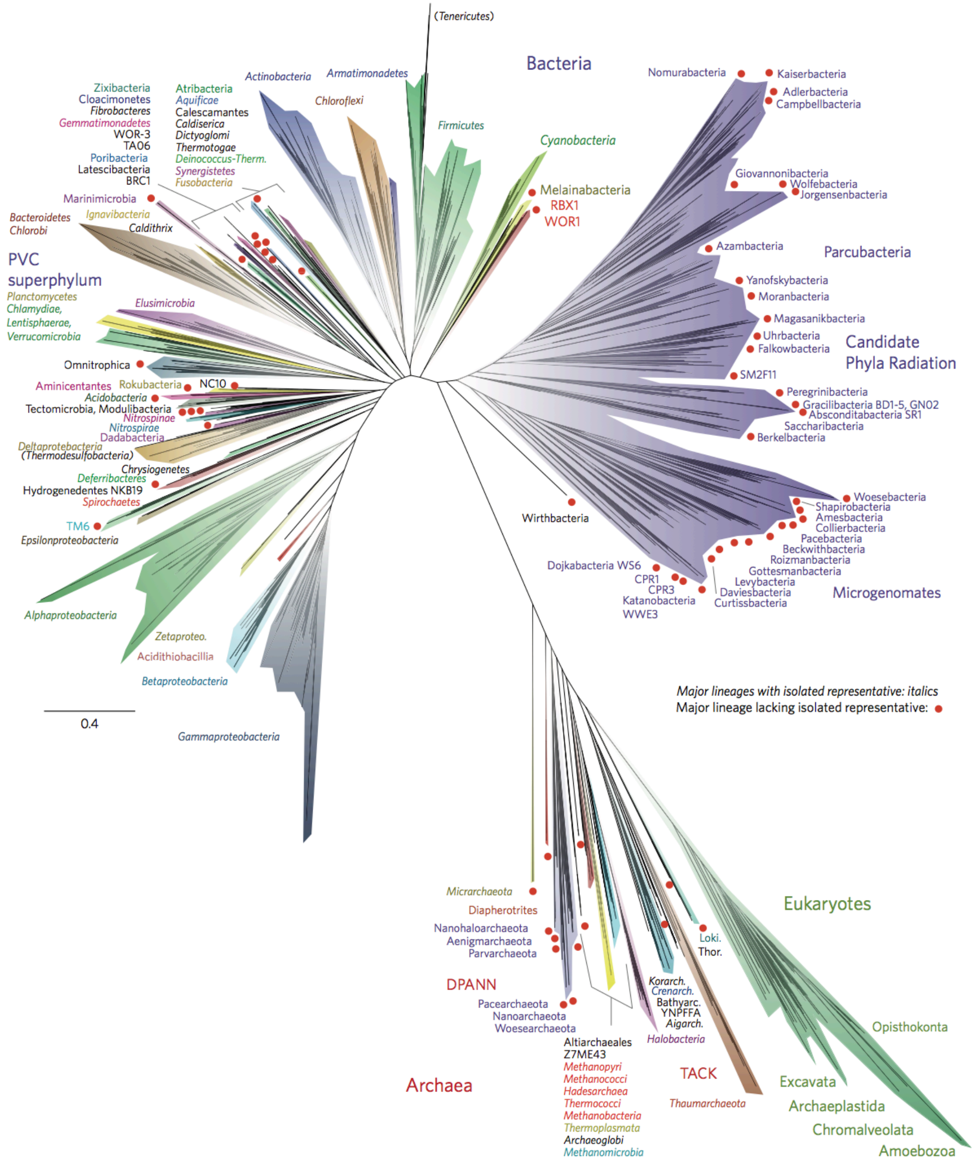
Class Goals

Mini-goals:

- **Intuition for each algorithm/model**
- **Big picture/ML ecosystem**

Examples:

- When to use what type of ML?
- How hard is it to train?
- What generalizes best?
- Where is the field going?



Required Background

You are expected have (at least) a working understanding of:

- **Linear algebra** (working with data, linear transformations)
- **Calculus** (for optimization, convergence, etc.)
- **Probability** (dealing with noise, sampling)
- **Programming** (for implementation, mostly python)

Plenty of resources available online

- Just need enough experience/mathematical maturity to pick up missing bits

Assignment: Homework

For HW1, self-diagnostic on background. Topics:

- Linear Algebra
- Calculus
- Probability
- Big-O notation
- Basic programming skills

Assignment: Homework

For HW1, self-diagnostic on background. **Examples**:

Consider the matrix X and the vectors $\mathbf{y}$ and $\mathbf{z}$ below:

$$X = \begin{pmatrix} 9 & 8 \\ 7 & 6 \end{pmatrix} \qquad \mathbf{y} = \begin{pmatrix} 9 \\ 8 \end{pmatrix} \qquad \mathbf{z} = \begin{pmatrix} 7 \\ 6 \end{pmatrix}$$

1. Is X invertible? If so, give the inverse, and if no, explain why not.

2. If $y = \tan(z)x^{6z} - \ln(\frac{7x+z}{x^4})$, what is the partial derivative of y with respect to x?

Assignment: Homework

For HW1, self-diagnostic on background. **Examples**:

Match the distribution name to its probability density / mass function. Below, $|\boldsymbol{x}| = k$.

(a) Laplace

(b) Multinomial

(c) Poisson

(d) Dirichlet

(e) Gamma

(f) $f(\boldsymbol{x}; \boldsymbol{\Sigma}, \boldsymbol{\mu}) = \frac{1}{\sqrt{(2\pi)^k \boldsymbol{\Sigma}}} \exp\left(-\frac{1}{2}(\boldsymbol{x} - \boldsymbol{\mu})^T \boldsymbol{\Sigma}^{-1} (\boldsymbol{x} - \boldsymbol{\mu})\right)$

(g) $f(x; n, \alpha) = \binom{n}{x} \alpha^x (1-\alpha)^{n-x}$ for $x \in \{0, \ldots, n\}$; 0 otherwise

(h) $f(x; b, \mu) = \frac{1}{2b} \exp\left(-\frac{|x-\mu|}{b}\right)$

(i) $f(\boldsymbol{x}; n, \boldsymbol{\alpha}) = \frac{n!}{\Pi_{i=1}^k x_i!} \Pi_{i=1}^k \alpha_i^{x_i}$ for $x_i \in \{0, \ldots, n\}$ and $\sum_{i=1}^k x_i = n$; 0 otherwise

(j) $f(x; \alpha, \beta) = \frac{\beta^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\beta x}$ for $x \in (0, +\infty)$; 0 otherwise

(k) $f(\boldsymbol{x}; \boldsymbol{\alpha}) = \frac{\Gamma(\sum_{i=1}^k \alpha_i)}{\prod_{i=1}^k \Gamma(\alpha_i)} \prod_{i=1}^k x_i^{\alpha_i - 1}$ for $x_i \in (0, 1)$ and $\sum_{i=1}^k x_i = 1$; 0 otherwise

(l) $f(x; \lambda) = \lambda^x \frac{e^{-\lambda}}{x!}$ for all $x \in Z^+$; 0 otherwise

Assignment: Homework

For HW1, self-diagnostic on background. **Examples**:

Draw the regions corresponding to vectors $\mathbf{x} \in \mathbb{R}^2$ with the following norms:

1. $||\mathbf{x}||_1 \leq 1$ (Recall that $||\mathbf{x}||_1 = \sum_i |x_i|$)
2. $||\mathbf{x}||_2 \leq 1$ (Recall that $||\mathbf{x}||_2 = \sqrt{\sum_i x_i^2}$)
3. $||\mathbf{x}||_\infty \leq 1$ (Recall that $||\mathbf{x}||_\infty = \max_i |x_i|$)

Assignment: Homework

For HW1, self-diagnostic on background. Topics:

- Linear Algebra
- Calculus
- Probability
- Big-O notation
- Basic programming skills

- If these feel very unfamiliar, consider taking relevant courses first and then take CS760 in the future.

Resources

Probability

- Lecture notes: http://www.cs.cmu.edu/~aarti/Class/10701/recitation/prob_review.pdf

Linear Algebra:

- Short video lectures by Prof. Zico Kolter: http://www.cs.cmu.edu/~zkolter/course/linalg/outline.html
- Handout associated with above video: http://www.cs.cmu.edu/~zkolter/course/linalg/linalg_notes.pdf
- Book: Gilbert Strang. Linear Algebra and its Applications. HBJ Publishers.

Big-O notation:

- http://www.stat.cmu.edu/~cshalizi/uADA/13/lectures/app-b.pdf
- http://www.cs.cmu.edu/~avrim/451f13/recitation/rec0828.pdf

Wikipedia is always a great resource!

Programming background

We expect you to be able to

- Implement simple routines/logic in Python (for/while loops, if/ else, break conditions)
  - Familiarity with NumPy would be a plus
- Write simple shell scripts in Linux/Unix
- Install and use ML packages (e.g. scikit-learn, PyTorch)

- Generally, we will **not** help you with these during OHs!
- Usually, you can resolve such issues via online forums (e.g stack overflow) or Piazza.

(Anonymous) Background Survey

https://forms.gle/2ppyDcmh3txpBQj6A

Logistics

Logistics: Lectures

- Spend 10 minutes:
  - Read the course syllabus.
  - Skim the course webpage.
- With 1-2 people sitting near you:
  - Discuss potential confusion, concerns, questions about the course.
  - Prepare questions to ask professor after 10 minutes.
- Rest of lecture:
  - Q&A on syllabus.
  - https://pages.cs.wisc.edu/~jphanna/teaching/2023fall_cs760/index.html

Logistics: Content

Four locations:

1. Course website: https://pages.cs.wisc.edu/~jphanna/teaching/2023fall_cs760/
2. Piazza: https://piazza.com/class/llcq5n8vttb5xc
   - Access Code: mlfall23
   - Preferred for questions! Sometimes your peers might be able to better answer your questions than the instructor/TA.
3. Canvas: https://canvas.wisc.edu/courses/360792
   - Do not share materials on canvas outside of class.
4. Gradescope: https://www.gradescope.com/courses/587829

This week's reading assignment

For lecture 1, article by Jordan and Mitchell on course website

REVIEW

Machine learning: Trends, perspectives, and prospects

M. I. Jordan[1*] and T. M. Mitchell[2*]

Machine learning addresses the question of how to build computers that improve automatically through experience. It is one of today's most rapidly growing technical fields, lying at the intersection of computer science and statistics, and at the core of artificial intelligence and data science. Recent progress in machine learning has been driven both by the development of new learning algorithms and theory and by the ongoing explosion in the availability of online data and low-cost computation. The adoption of data-intensive machine-learning methods can be found throughout science, technology and commerce, leading to more evidence-based decision-making across many walks of life, including health care, manufacturing, education, financial modeling, policing, and marketing.

Machine learning is a discipline focused on two interrelated questions: How can

ance when executing some task, through some type of training experience. For example, in learn-

Thanks Everyone!

Some of the slides in these lectures have been adapted/borrowed from materials developed by Mark Craven, David Page, Jude Shavlik, Tom Mitchell, Nina Balcan, Elad Hazan, Tom Dietterich, Pedro Domingos, Jerry Zhu, Yingyu Liang, Volodymyr Kuleshov, and Fred Sala