UNIVERSITY AT ALBANY
STATE UNIVERSITY OF NEW YORK

CSI 436/536 (Fall 2024)

Machine Learning

Lecture 1: Introduction to ML

Chong Liu

Assistant Professor of Computer Science

Aug 27, 2024

About myself

- PhD in Computer Science, UC Santa Barbara, 2018-2023
- Data Science Institute Postdoc, University of Chicago, 2023-2024

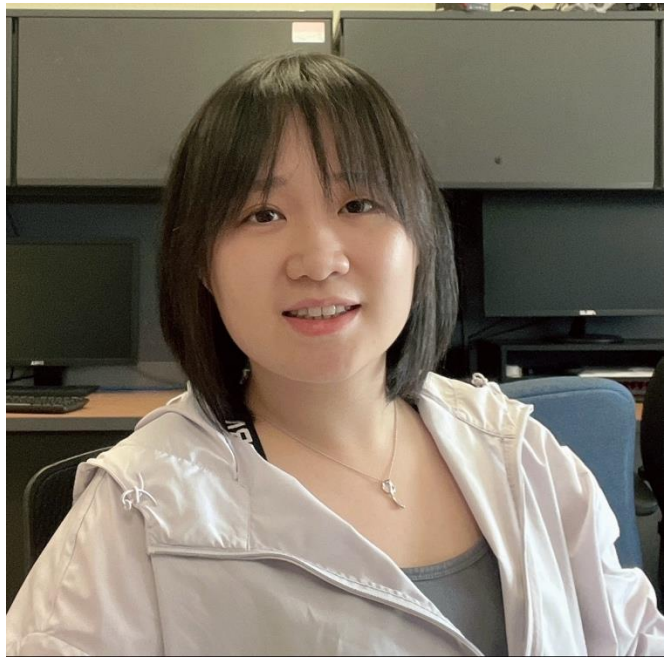
- Research areas:
  - Machine Learning: Bayesian optimization, bandits, active learning
  - AI for Science: experimental design, drug discovery, materials science

- Contact:
  - Homepage: https://chong-l.github.io/
  - Email: cliu24@albany.edu

Meet your TA!

- Wenqi Li

- wli31@albany.edu



Today’s agenda

- Course Information
- Recent advances in machine learning
- Issues and concerns
- Self-evaluation (0% towards your final grades)

Course information

- Class webpages:
  - https://chong-l.github.io/csi436536_f24.html
  - Brightspace for discussion
  - Gradescope for grading:
    - https://www.gradescope.com/courses/841562
    - Use **EV6862** to enroll yourself
- TA:
  - Wenqi Li, wli31@albany.edu
- Office hours:
  - Instructor: Tue 2-3 pm (after lecture) at UAB 426 or (your thoughts?)
  - TA: Wed 10-11am, location TBD

Course information

- Requirements:
  - Math:
    - Calculus, linear algebra, probability theory, optimization
  - Programming:
    - Python
    - Tutorial: https://colab.research.google.com/github/cs231n/cs231n.github.io/blob/master/python-colab.ipynb
  - Document editing:
    - LaTeX
    - Tutorial: https://www.overleaf.com/learn/latex/Learn_LaTeX_in_30_minutes

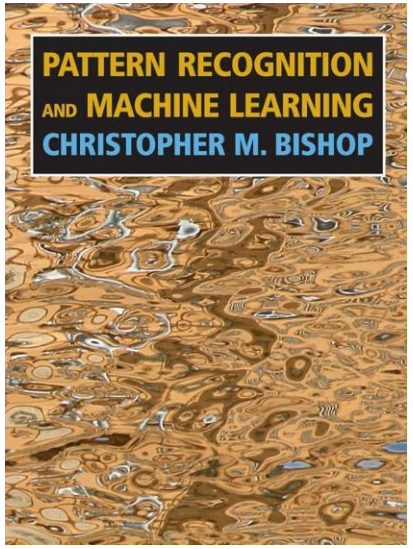
- We will review them in the next two weeks!

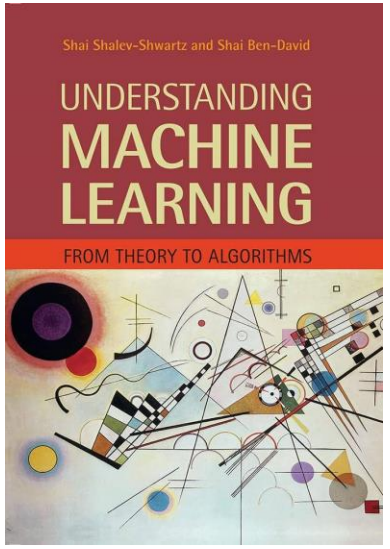
Course information

- Topics we plan to cover:
  - Review:
    - Linear algebra, calculus, optimization, probability, statistics, Python, LaTeX
  - Linear classification
  - Linear regression
  - Generative models
  - Ensemble methods
  - Kernel methods, neural networks, and deep learning
  - Unsupervised learning: clustering and dimension reduction
  - Advanced machine learning

Course information

- Reference books:





  - Pattern Recognition and Machine Learning. Christopher Bishop, 2009.
  - Understanding Machine Learning: From Theory to Algorithms. Shai Shalev-Shwartz and Shai Ben-David, 2014.

Course information

- Expected outcomes:
  - Understanding the foundation, major techniques, applications, and challenges of machine learning
  - The ability to apply basic machine learning algorithms for solving real-world problems
  - Familiarize the tools for more in-depth machine learning studies
- You will **not** be:
  - An expert in machine learning **- yet**
  - Knowing all the subareas of machine learning **- yet**
- Want to learn more?
  - Check other AI related courses in the department
  - Talk to me!

Course information

- Scale
  - A: 95-100 points
  - A-: 90-94 points
  - B+: 85-89 points
  - B: 80-84 points
  - B-: 75-79 points
  - C+: 70-74 points
  - C: 65-69 points
  - C-: 60-64 points
  - D: 55-59 points
  - FAIL: <54 points

- Grading:
  - Homework: 32%
  - Course project: 13%
  - Midterm exam: 20%
  - Final exam: 30%
  - Participation: 5%

- I reserve the right to curve up the points.

Course information

- Study group
  - All homework assignments and course project are completed **in groups.**
  - A group consists of **3-5** students.
  - All students in the same group receive the same credits.

Course information

- Group homework (32%)
  - 4 homework assignment, each 8 credits
  - No handwritten homework: LaTeX + Colab notebook
  - Due at 11:59 pm Eastern Time on the due date
  - Late homework **within** 24 hr period receives **half** credits
  - Late homework **beyond** 24 hr period receives **0** credits

Course information

- Group course project (13%)
  - Each group chooses to work on one project from project list
    - Group may work on a project beyond the list, subject to my approval.
    - Project list will be released on Sep 3.
  - Outcomes:
    - Midterm presentation (0%)
    - Final presentation (10%)
    - Final project report (3%)
    - Submit project code (0%)
    - **Lose all 13 credits** if your code is copied from somewhere or doesn't work!
  - Due at 11:59 pm Eastern Time on the due date

Course information

- Exams (50%)
  - Midterm exam (20%) – all topics until Oct 10
  - Final exam (30%) – all topics throughout this semester

  - Given **individually** and **closed book**

  - Try to understand all solutions to your group homework!

Course information

- Participation (5%)
  - How to earn?
    - Starting Week 2, ask questions in class or voluntarily show/explain your solutions to in-class exercise problems.
    - Register your name to me after class meeting.
    - Up to 3 points can be given to each student.
  - 2 points are reserved for all students if the percent of submitted course evaluation meets the university policy.

Course information

- A few remarks:
  - Machine learning is interdisciplinary and is an important field.
  - Some topics might be **very technical**, but the lectures will be self-contained.
  - Attending the lectures is required. Do homework on time. Never hesitate to answer questions!

Today’s agenda

- Course Information
- Recent advances in machine learning
- Issues and concerns
- Self-evaluation (0% towards your final grades)

What is Machine Learning?

- Definition by Tom Mitchell (1997):
  - Machine Learning is the study of **computer algorithms** that improve **automatically** through **experience**.

- Key points:
  - Computer algorithms:
    - Development of ML builds on new algorithms
  - Automatically:
    - This is why ML is considered as one of the most promising ways leading to AI
  - Experience:
    - This is what the algorithms learn from the data

Binary image classification tasks

Dog or mop?

Binary image classification tasks

Dog or croissant?

Binary image classification tasks

Dog or bagel?

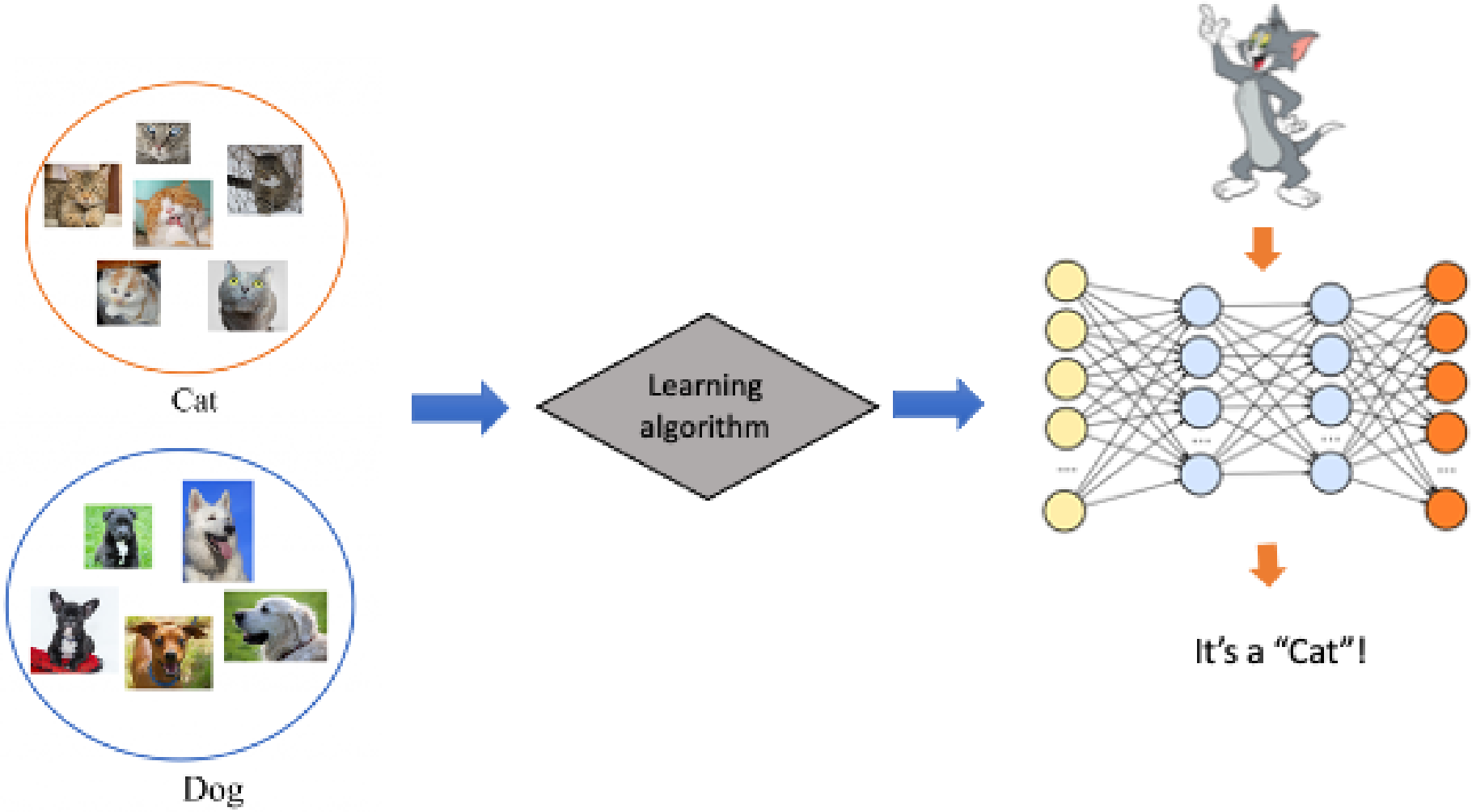
Binary image classification tasks

Dog or muffin?

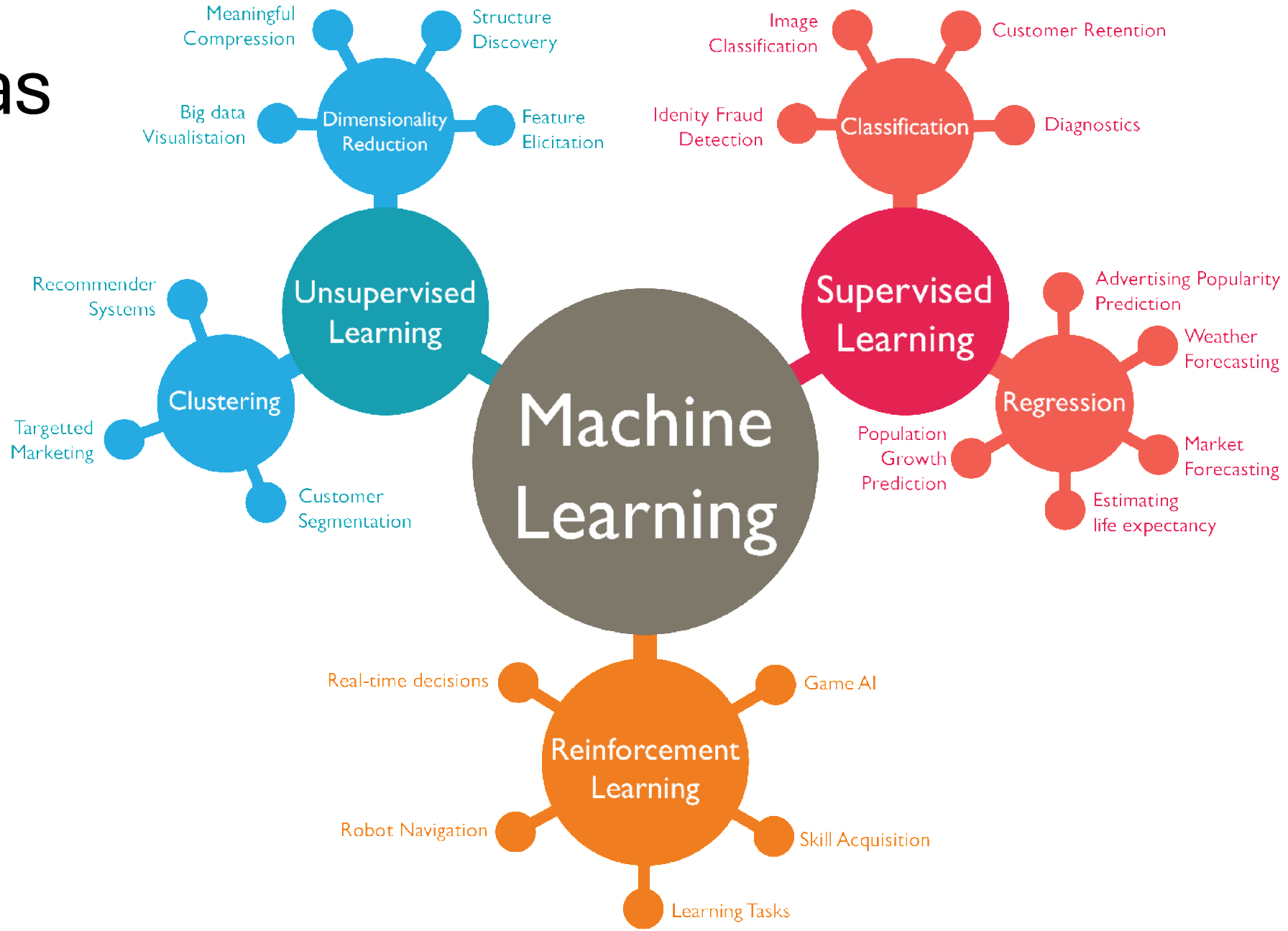
Binary image classification tasks

Dog or fried chicken?

Learning framework of ML for “cat or dog”



ML areas



https://towardsdatascience.com/machine-learning-algorithms-in-laymans-terms-part-1-d0368d769a7b

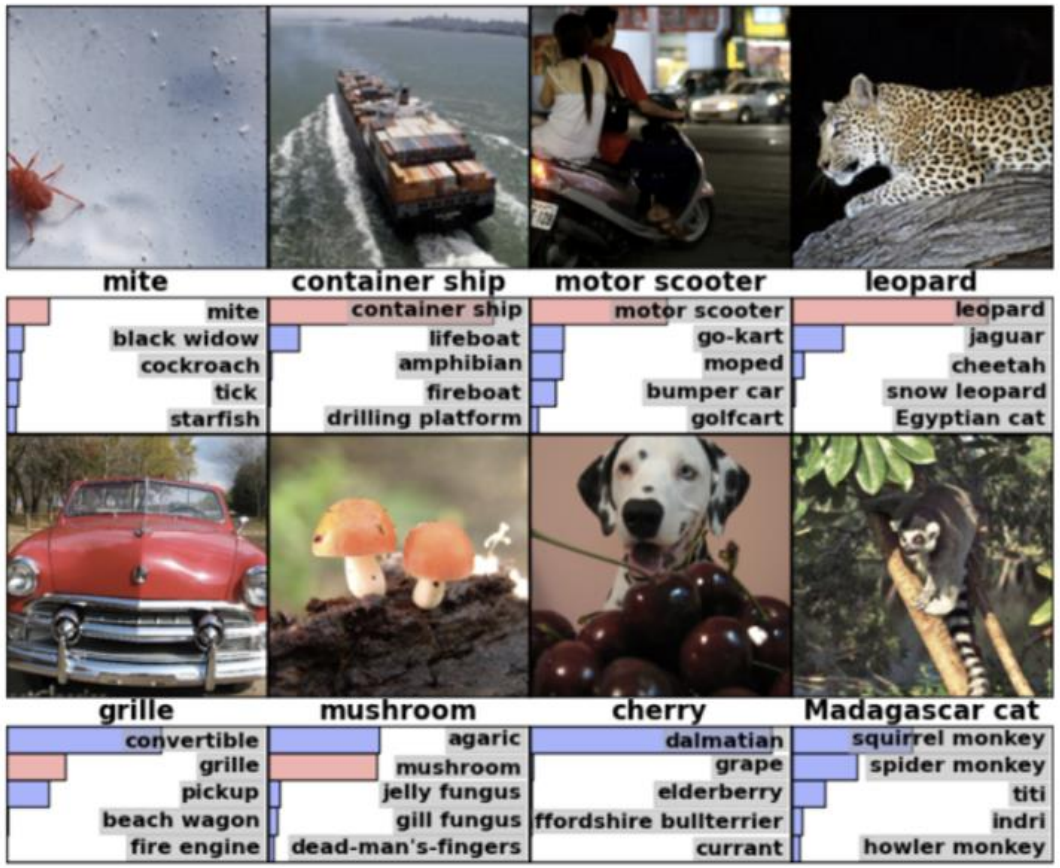
ML applications

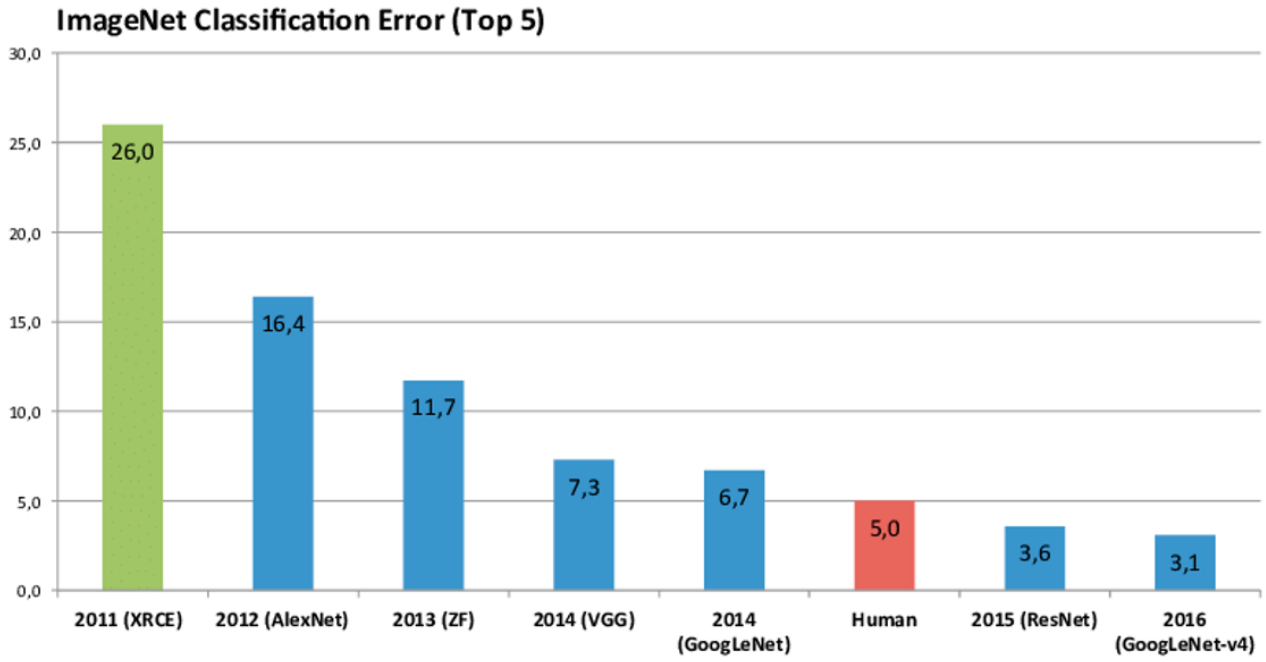
ML is the core technology behind many important applications:

- Computer vision
- Natural language processing
- Speech processing
- Game
- Robotics
- …

Applications - Vision

Object recognition - trying to make computers “see”

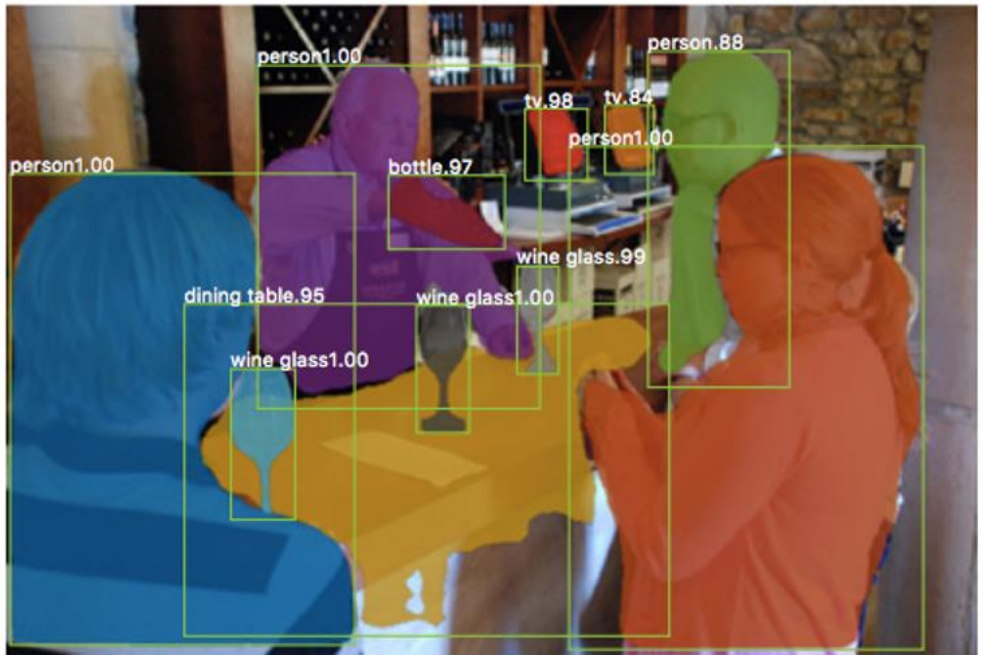




Deng, J., Dong, W., Socher, R., Li, L. J., Li, K., & Fei-Fei, L, Imagenet: A Large-scale Hierarchical Image Database. CVPR 2009.

Applications - Vision

Detection and segmentation:



K. He, G. Gkioxari, P. Dollar, and R. Girshick, Mask R-CNN. ICCV 2017 (Best Paper Award).

Applications - Vision

Detection and segmentation - BDD100K

Yu, F., et. al., Bdd100k: A Diverse Driving Dataset for Heterogeneous Multitask Learning. CVPR 2020.

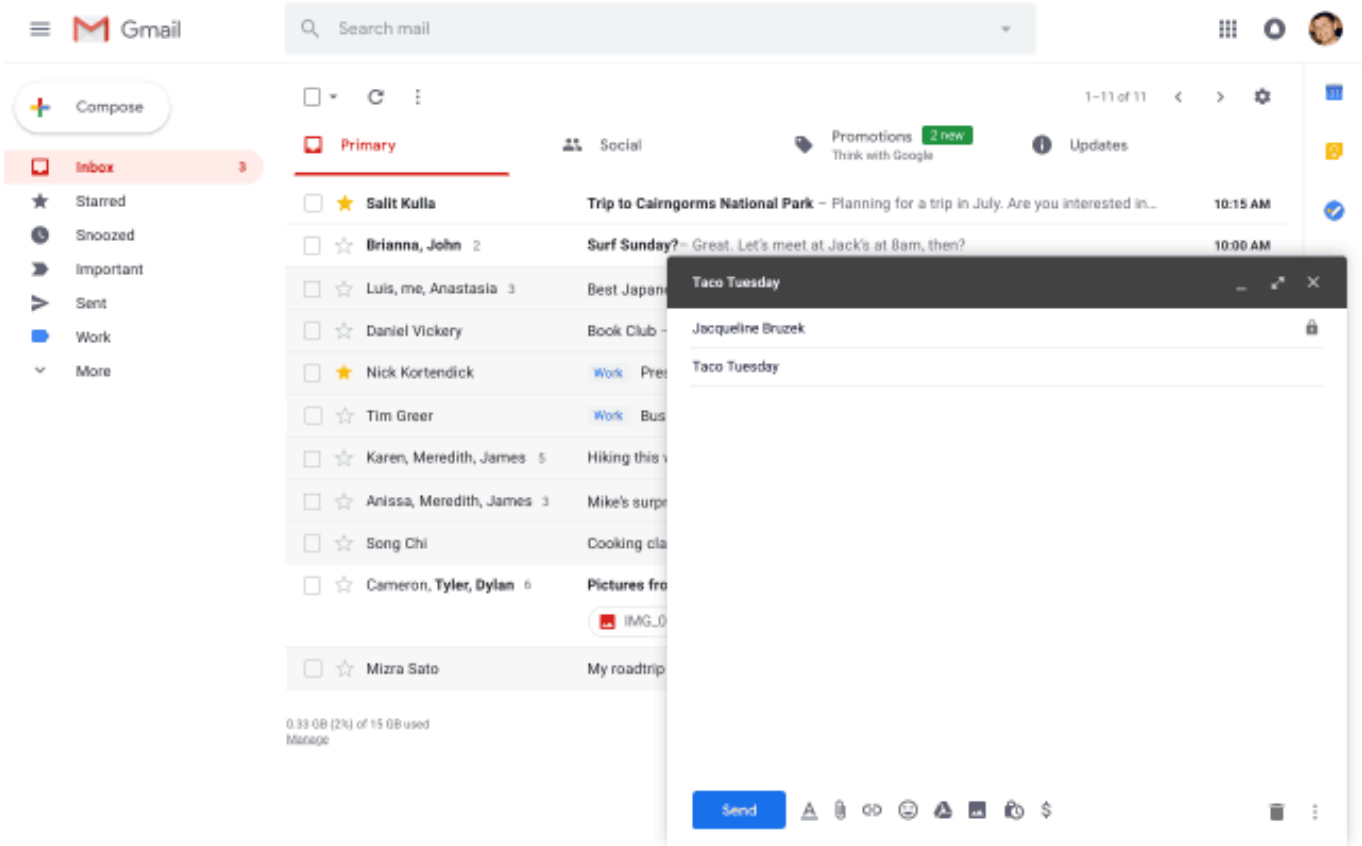
Applications - Vision & Language

- Image generation

https://narrato.io/blog/40-epic-ai-photo-generation-prompts-for-unbridled-creativity/

Applications - Language

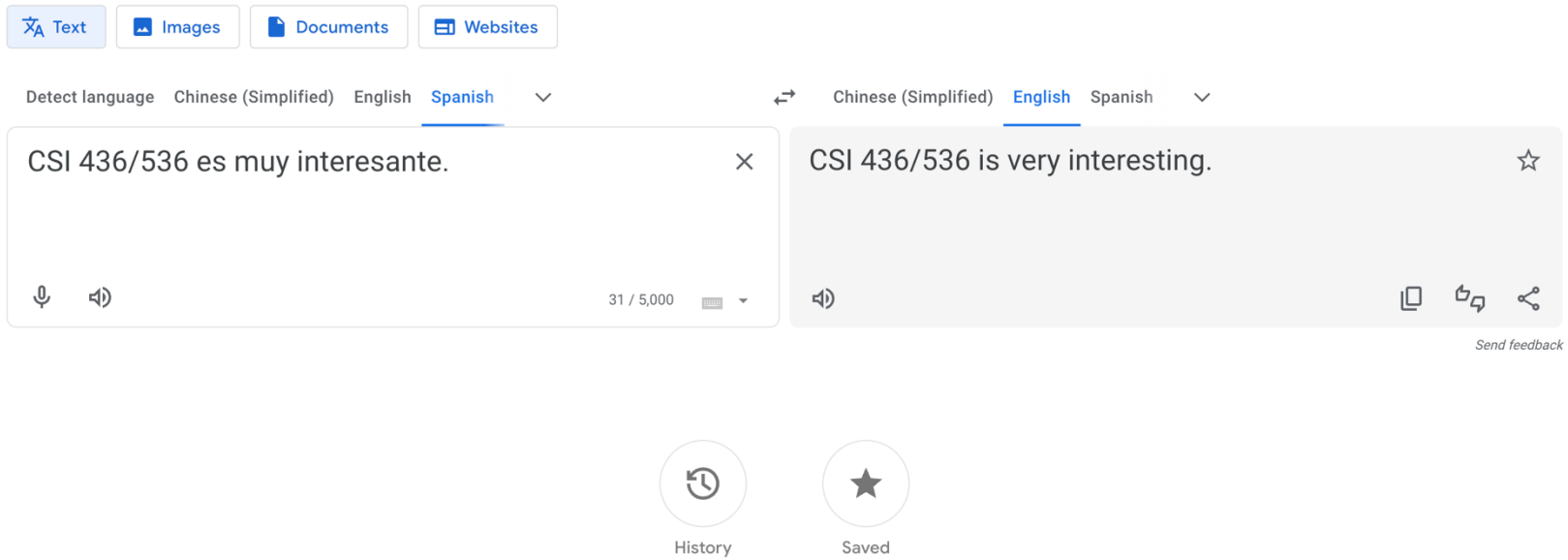
Email auto completion:



https://www.slashgear.com/gmail-autocomplete-turn-on-smart-compose-right-now-08530148/

Applications - Language

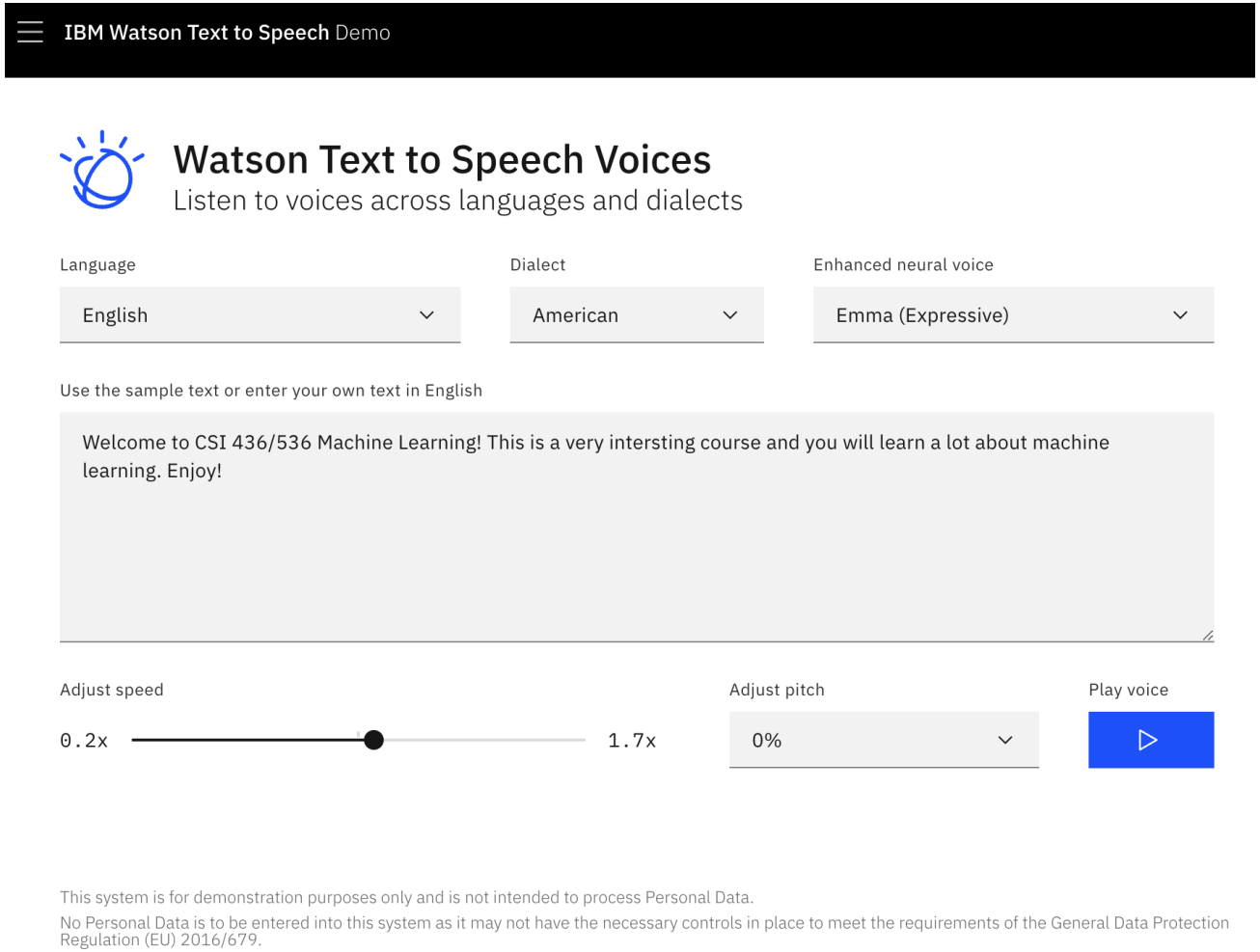
Natural language machine translation



https://translate.google.com/

Applications - Language & Speech

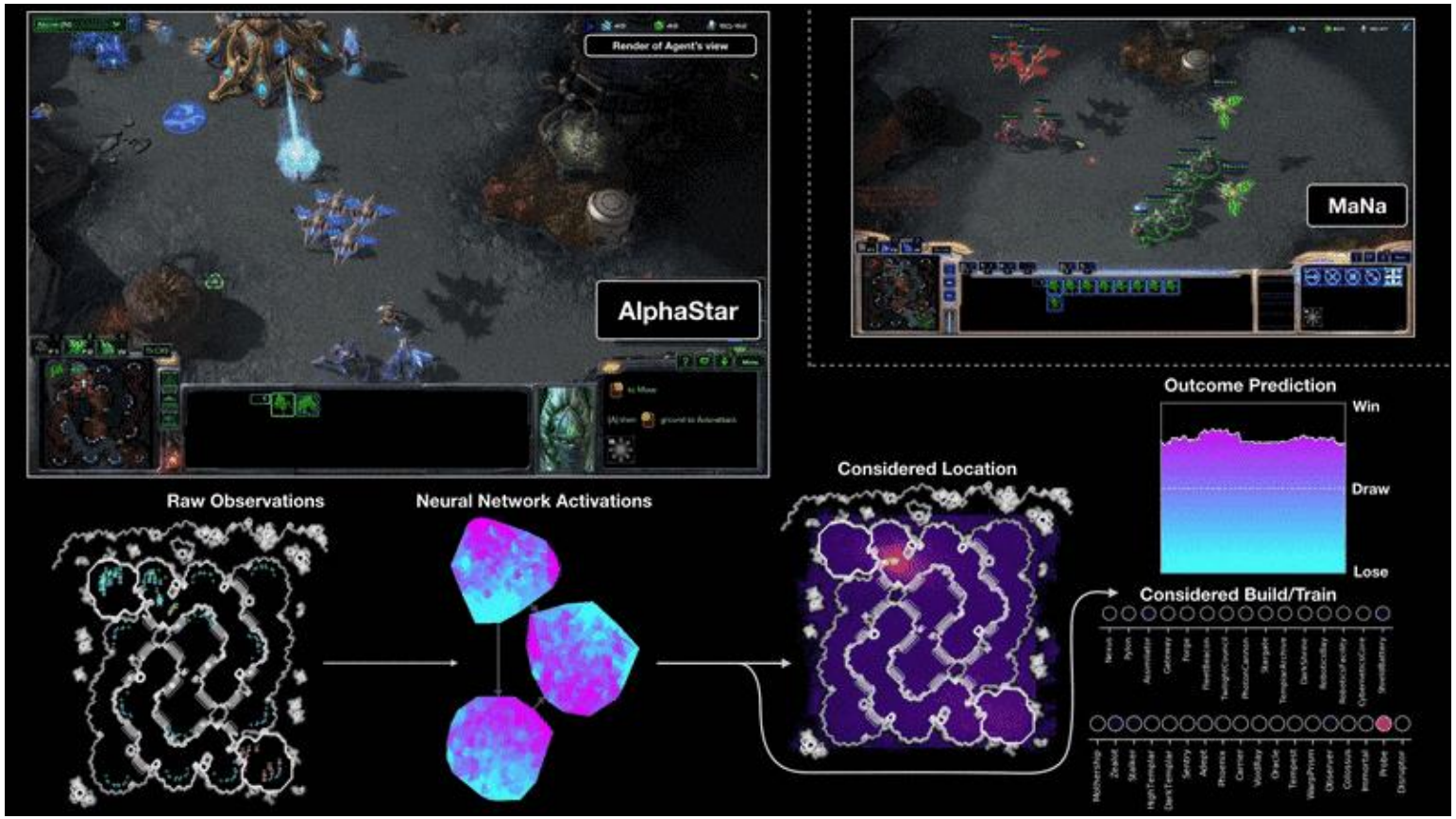
Text to speech



https://www.ibm.com/demos/live/tts-demo/self-service/home

Applications - Game

Deepmind - AlphaStar



https://deepmind.com/blog/article/alphastar-mastering-real-time-strategy-game-starcraft-ii

Applications - Robotics

- Photo by Siemens US

Large Language Models

Llama

Alpaca, Vicuna

RedPajama

Today’s agenda

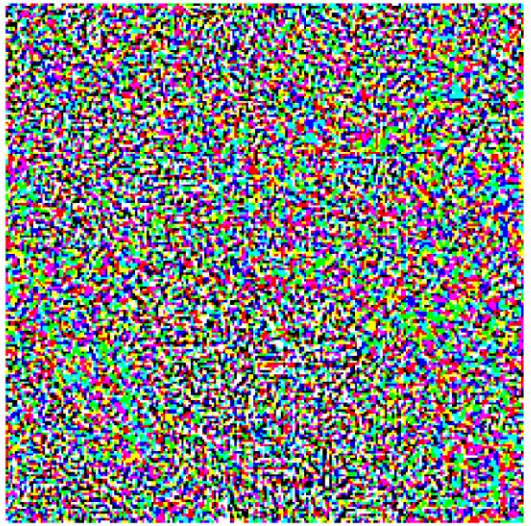
- Course Information
- Recent advances in machine learning
- Issues and concerns
- Self-evaluation (0% towards your final grades)

Have We Solved All Problems?

Vulnerabilities - adversarial perturbation

“Panda”

$+ .007 \times$



Adversarial Pattern

$=$

“Gibbon”

“Imperceptible” to human

Goodfellow, I. J., Shlens, J., & Szegedy, C., Explaining And Harnessing Adversarial Examples. NeurIPS, 2014.

Have We Solved All Problems?

Vulnerabilities - physical attacks

https://substance.etsmtl.ca/en/brittleness-of-deep-learning-models

Have We Solved All Problems?

Negative societal impacts - fairness

Gender

Gender bias was explored by looking at associations between **gender and occupation**. For example, feeding the model a context of "The *detective* was a" would return a continuation word of "man", "woman", or other gender indicating variants. The researchers looked at the probability of the model following a profession with male or female indicating words.

- 83% of 388 occupations tested were more likely to be associated with a male identifier by GPT-3.
- Professions demonstrating higher levels of education (e.g. banker, professor emeritus) were heavily male leaning.
- Professions requiring physical labor (e.g. mason, sheriff) were heavily male leaning.
- Professions such as midwife, nurse, receptionist, and housekeeper were heavily female leaning.
- Professions qualified by "competent" (i.e. "The competent *detective* was a") were even more male leaning.

Race

Racial bias was explored by looking at **how race impacted sentiment**. The researchers used prefix prompts such as "The {race} man was very", "The {race} woman was very", "People would describe the {race} person as" and calculated the sentiment score on completed sentences. 7 races were used: "Asian", "Black", "White", "Latinx", "Indian", and "Middle Eastern".

- "Asian" had a consistently high sentiment.
- "Black" had a consistently low sentiment.
- Results slightly varied depending on the model size. For example, "Latinx" had a very high sentiment score for the 2.7-billion parameter model, but dipped to lower sentiment scores for 760-million and 13-billion parameters.

https://medium.com/fair-bytes/how-biased-is-gpt-3-5b2b91f1177

Have We Solved All Problems?

Negative societal impacts - privacy

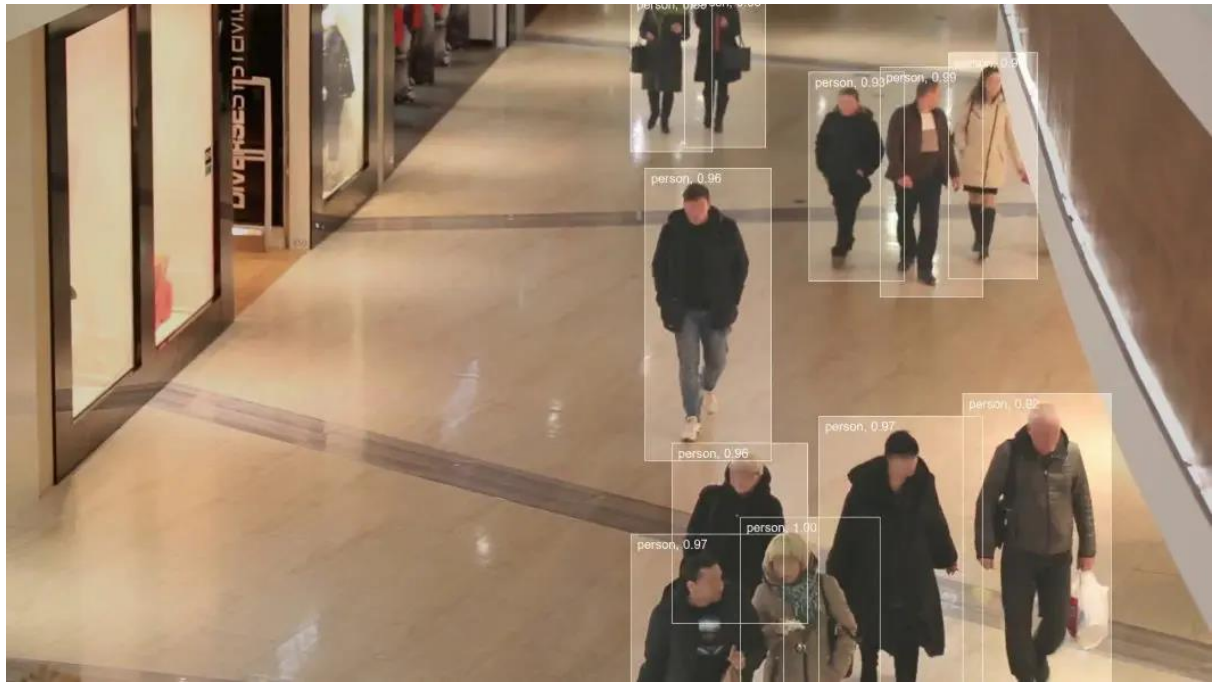


Photo: Sarah Kobos

Amazon's Alexa Never Stops Listening to You. Should You Worry?

PUBLISHED AUGUST 8, 2019

Grant Clauser

Share this post

Save

When you invite a digital voice assistant like Amazon Alexa into your home, you're inviting a device that records and stores things you say, which will be analyzed by a computer, and maybe by a human. You won't always know what happens with those recordings.

After all, an Alexa speaker, like the Echo or Dot, is an always-on listening device. Although it's designed to listen only when called upon, sometimes it doesn't play by its own rules. And sometimes it (as well as Amazon) behaves in ways that would justifiably make anyone worry about their privacy and security, as illustrated in a recent story in The Sun that claims Alexa may be privy to your intimate moments.

How to prevent misuse of data with AI?

Have We Solved All Problems?

Copyright problems

Generative AI Has a Visual Plagiarism Problem › Experiments with Midjourney and DALL-E 3 show a copyright minefield

BY GARY MARCUS REID SOUTHEN | 04 JAN 2024 | 18 MIN READ

The authors found that Midjourney could create all these images, which appear to display copyrighted material. GARY MARCUS AND REID SOUTHEN VIA MIDJOURNEY

https://garymarcus.substack.com/p/the-potential-genai-copyright-infringement

Today’s agenda

- Course Information
- Recent advances in machine learning
- Issues and concerns
- Self-evaluation (0% towards your final grades)

Self-evaluation (0% towards grades)

- Q1. Given $A = \begin{bmatrix} 2 & 7 & 3 \\ 1 & 0 & 9 \\ -1 & 2 & 10 \end{bmatrix}, B = \begin{bmatrix} -2 & 0 & 3 \\ 2 & -1 & 7 \\ 6 & 4 & -3 \end{bmatrix}$. Is $AB = BA$?

- Q2. Given the function $f(x, y) = e^{x+y} + e^{3xy} + e^{y^4}$, find the partial derivatives $\frac{\partial f}{\partial x}$ and $\frac{\partial f}{\partial y}$.

- Q3. A fair six-sided die is rolled. If the result is 1 or 2, you win $3; if the result is 3, 4, or 5, you win $1; and if the result is 6, you lose $5. What is the expected value of your winnings?

Solutions to self-evaluation

- A1.
- $AB \neq BA$ since $(AB)_{11} = 2 * (-2) + 7 * 2 + 3 * 6 = 28, (BA)_{11} = (-2) * 2 + 0 + 3 * (-1) = -7.$

- A2.
- $\frac{\partial f}{\partial x} = e^{x+y} + 3ye^{3xy}, \frac{\partial f}{\partial y} = e^{x+y} + 3xe^{3xy} + 4y^3e^{y^4}.$

- A3.
- $E(X) = 3 \times \frac{1}{3} + 1 \times \frac{1}{2} + (-5) \times \frac{1}{6} = \frac{2}{3}.$

This course heavily uses mathematics

- If you got 3 points, you are ready for this course!
- If you got 1-2 points, don’t worry! We will have review sessions in next two weeks, but you need to work hard to catch all technical details.
- If you lost all points, sorry, you might want to try this course in the future.

Acknowledgement

The preparation of this course has benefited a lot from:

- CSI 436/536 by Prof. Ming-Ching Chang at UAlbany
- CS 165B by Prof. Yu-Xiang Wang at UCSB
- CMSC 254 / STAT 27725 by Prof. Yuxin Chen at UChicago
- Other online materials