CSCI 566 – Deep Learning and Its Applications

Yan Liu
Thomas Lord Department of Computer Science
University of Southern California

Credits to previous versions of USC CSCI566, CMU 10601/701, Stanford CS 229, 231n

Welcome to CSCI 566!

- This class will teach you some exciting developments in Machine Learning, Computer Vision, NLP, Robotics, and other AI-related fields in the last decade!

- **New cool stuff:** guest lectures from industry, academia, and more to discuss their life in ML and Data Science – this session will be interactive, and I hope it will be helpful for your career planning.

Self-Introduction

- Tenure at USC: Aug 2020 - Present
- Teaching
  - CSCI 567 – Machine Learning
  - CSCI 688 – Advanced Data Analytics
- First time teaching CSCI 566
  - Prepare the transition to a sequel of two classes on machine learning

~~Prerequisite~~ (CSCI 567 – Machine Learning)

- **Do you know the following..?**
  - Probability and Statistical Learning
    - Density function, loss function, cross-validation
  - Supervised Learning
    - Nearest Neighbor, Kernels, Random Forest
  - Unsupervised Learning
    - Clustering, PCA, SVD

- Ideally yes, but this semester we are transitioning

New Offering of Machine Learning Classes – Foundation ML Classes

- **CSCI 566 - Machine Learning 1: Mathematics Foundations and Deep Learning (**official name change: Fall 2025**)**
- **CSCI 567 - Machine Learning 2: Algorithms and Models (**official name change: Fall 2025**)**
  - During the discussion of the AI curriculum committee, it was recommended that we offer a sequel of 2 classes for the introductory core ML class for MS and PhD students.
  - ML 1 introduces the basic maths concepts/skills before we discuss the algorithms and models in detail in ML 2. Students who can pass a screening exam do not need to take ML1 and can proceed with ML 2 directly.
  - Outcome: After the class, the students will understand the fundamentals of ML, and will be able to apply and modify existing ML packages to solve a real world problem.

New Offering of Machine Learning Classes – Advanced Topics

This series of advanced classes will provide MS and PhD students the necessary skills to conduct research in ML. Depending on their interest, the students can take 1-4 of these classes.

- 678 - **Machine learning theory** (everything about the theory of ML)
- 699 - **Advanced topics in deep learning** (advanced topics in deep learning, such as deep generative models, generative adversarial networks, graph neural networks, variational auto-encoders, meta-learning etc; Notes: basic models in deep learning will be introduced in 566 and 567
- 699 - **Deep reinforcement learning** (traditional reinforcement learning models, deep reinforcement learning, and their applications)
- 699 - **Fairness, robustness and interpretability** (foundations and recent development in AI fairness, robustness and interpretability)
- After taking the advanced class, the students will be able to develop novel ML models and theories for research and publish papers.

Teaching Team

- Teaching assistants:
  - James Enouen
  - Samuel Griesemer
  - James Yipeng Huang
  - Ayush Jain
  - Leticia Pinto Alva
  - Bingjie Tang
  - Duygu Nur Yaldiz
  - Jesse Zhang
  - Wen Ye
- Graders (possibly)

Logistics

- Time and Location
  - 6-9:20pm Tue; THH 101
- All materials will be distributed at Piazza (passcode: csci566)
  - https://piazza.com/usc/fall2024/csci566
- Instructor Office Hour:
  - Time: Wed 10am
  - Office Hours Sign-Up Link: https://docs.google.com/spreadsheets/d/1GwNLTGGbPZAByjI8_b6-lHQXt48PhHLsfc1QBMMBko4/edit?usp=sharing
- TA Office Hour: TBD based on your project assignment
- Review Syllabus

Grading Scale from Previous Years

Assignment Submission Policy

All assignments and project reports need to be submitted in an electronic form by **11:59 pm PST** of the due date. There are NO late days for these.

Grading Scale

A	93-100
A-	90-92
B+	87-89
B	83-86
B-	80-82
C+	77-79
C	73-76
C-	70-72
D+	67-69
D	63-66
D-	60-62
F	59 and below

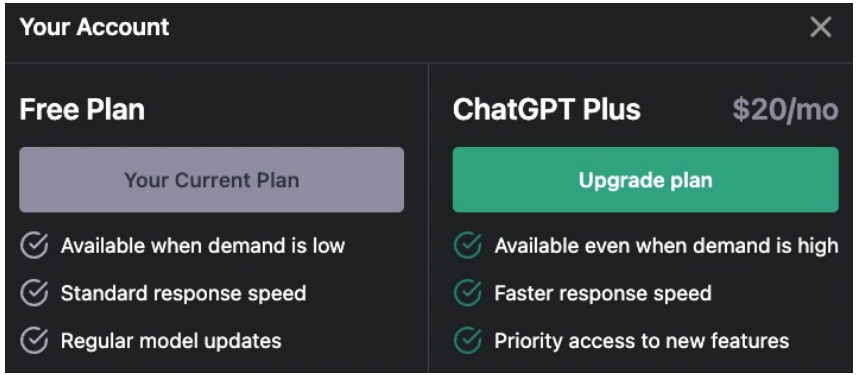
Letter grades are decided by rounding floating point grades up to the nearest whole number (e.g., 92.2 -> A; 59.8 -> D-).

Required Readings and Supplementary Materials

- Deep Learning (MIT Press) by Ian Goodfellow, Yoshua Bengio, and Aaron Courville.
  - A free online version is available at http://www.deeplearningbook.org/
- Mathematics for Machine Learning (Cambridge University Press) by Marc Peter Deisenroth, A. Aldo Faisal, and Cheng Soon Ong.
  - A free online version is available at https://mml-book.github.io/

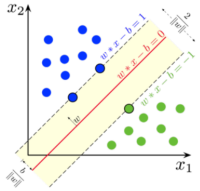
Machine Learning and Deep Learning

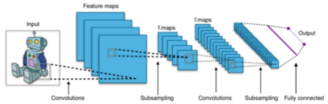
ML and DL Applications

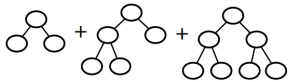


Why Did not See These Happen Earlier?

1958 – 2000: Research







Perceptron Algorithm	Backprop	Support Vector Machine (SVM)	ConvNet	Gradient Boosting Machine (GBM)
1958	1986	1992	1998	1999

Many algorithms we use today are **created before 2000**

Credit to CMU 15884 by Tianqi Chen

Why Did not See These Happen Earlier?

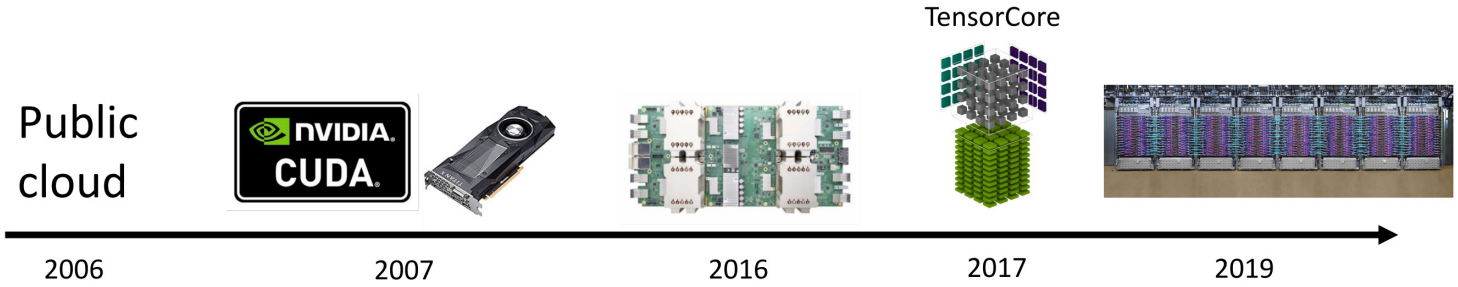
2000 – 2010: Arrival of Big Data

Data serves as fuel for machine learning models

Credit to CMU 15884 by Tianqi Chen

Why Did not See These Happen Earlier?

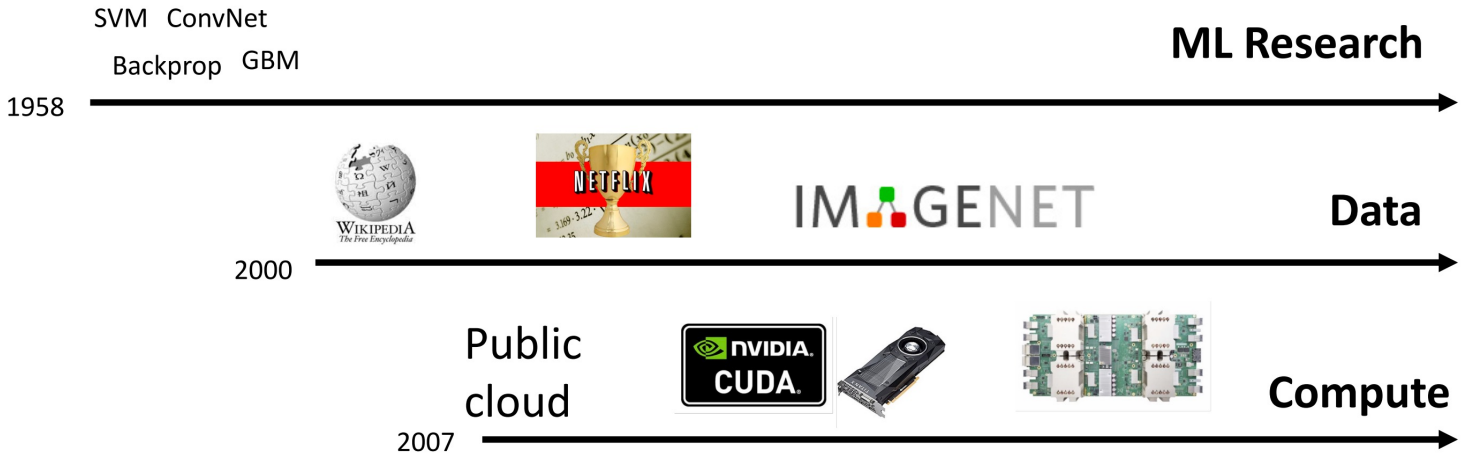
2006 – Now: Compute and Scaling



Compute scaling

- Credit to CMU 15884 by Tianqi Chen

Three Pillars of ML Applications



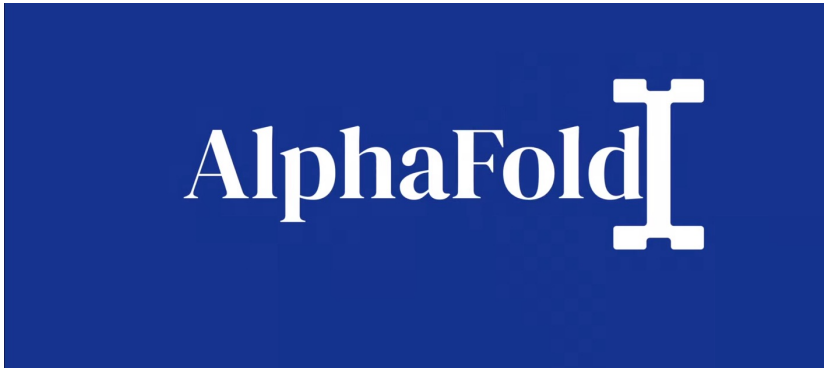
- Credit to CMU 15884 by Tianqi Chen

It Depends on Both ”Soft”- and Hard-ware Advance

- **Opportunity with GPUs**:

1. **Success in deep neural networks** with advancement in hardware

Visual object recognition (2012)

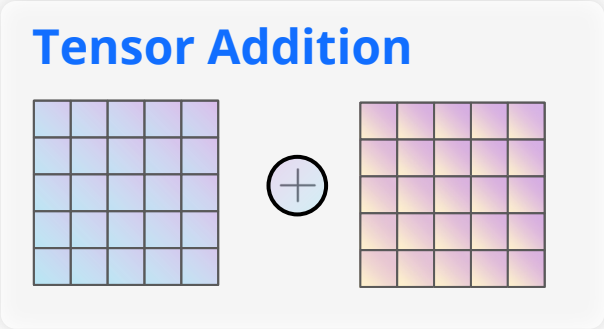


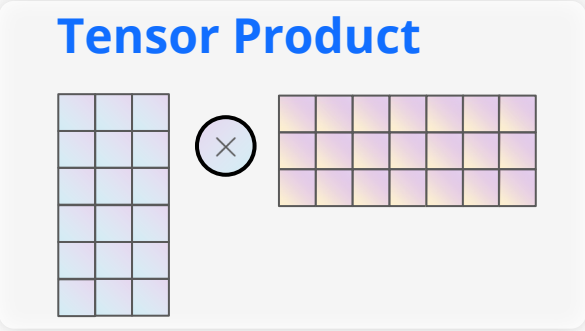
Protein structure prediction (2021)

It Depends on Both ”Soft”- and Hard-ware Advance

- **Opportunity with GPUs**:

1. Success in deep neural networks with advancement in hardware
    - **Fast computation** with **tensor operations/algebra**





It Depends on Both ”Soft”- and Hard-ware Advance

- **Opportunity with GPUs**:

1. **Success in deep neural networks** with advancement in hardware
2. **Maturity in learning systems**

Deep optimization & **Neural nets on GPUs**

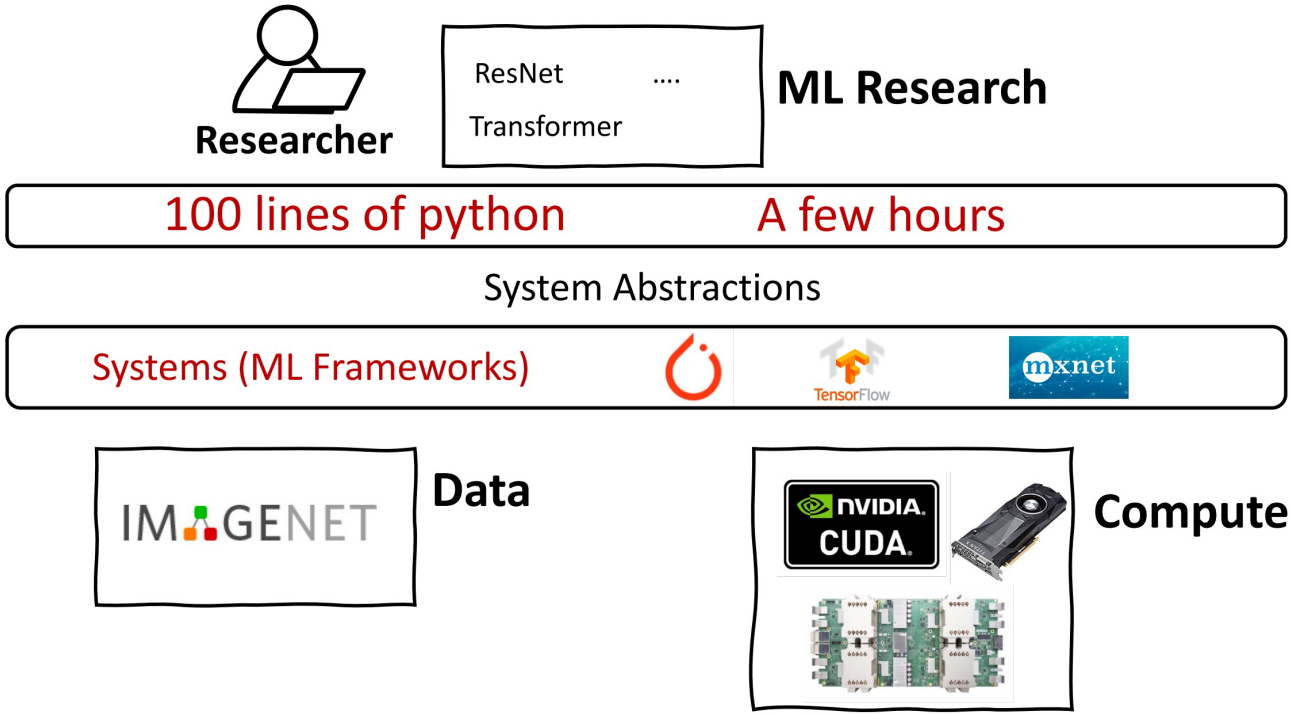
Case Study: Ingredient of AlexNet

Year 2012

Methods	Data	Compute
SGD Dropout ConvNet Initialization	IMAGENET 1M labeled images	Two GTX 580 Six days

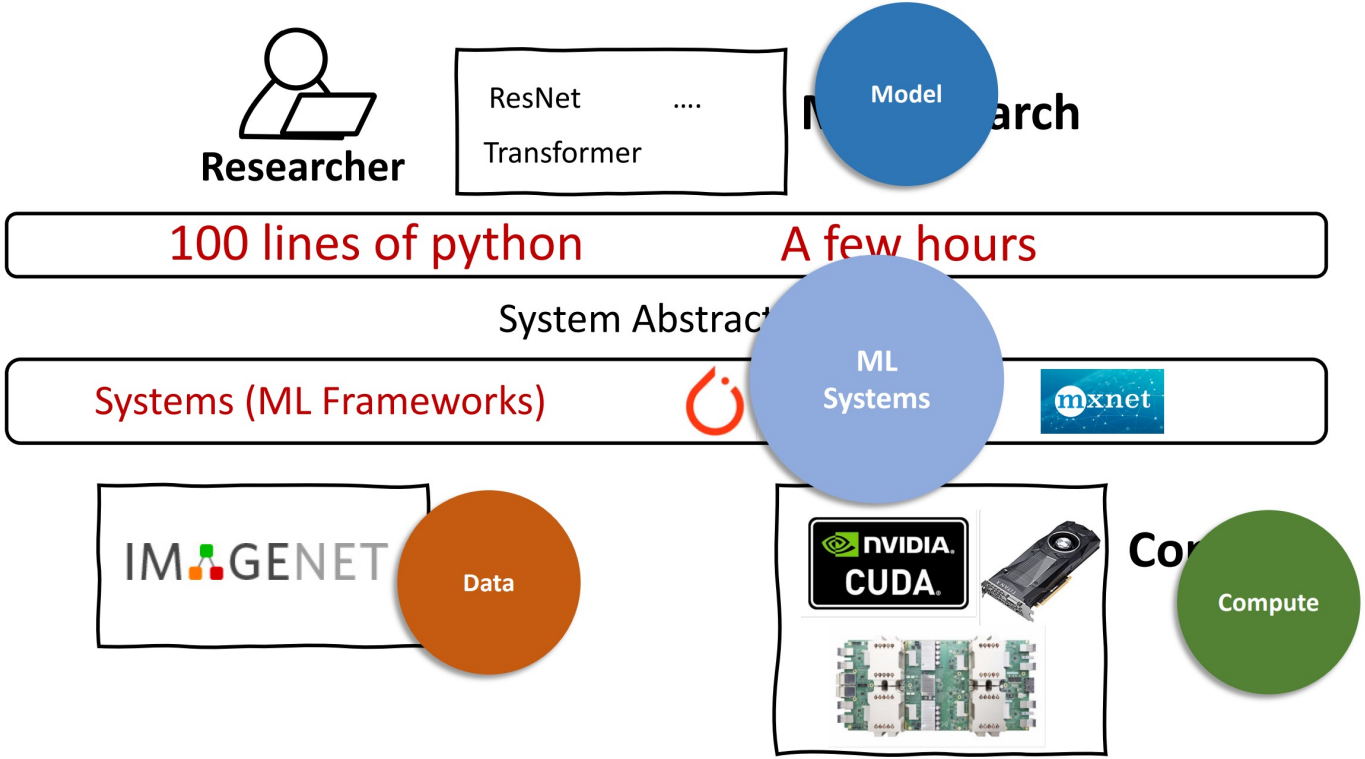
- Credit to CMU 15884 by Tianqi Chen

Case Study: Ingredient of AlexNet



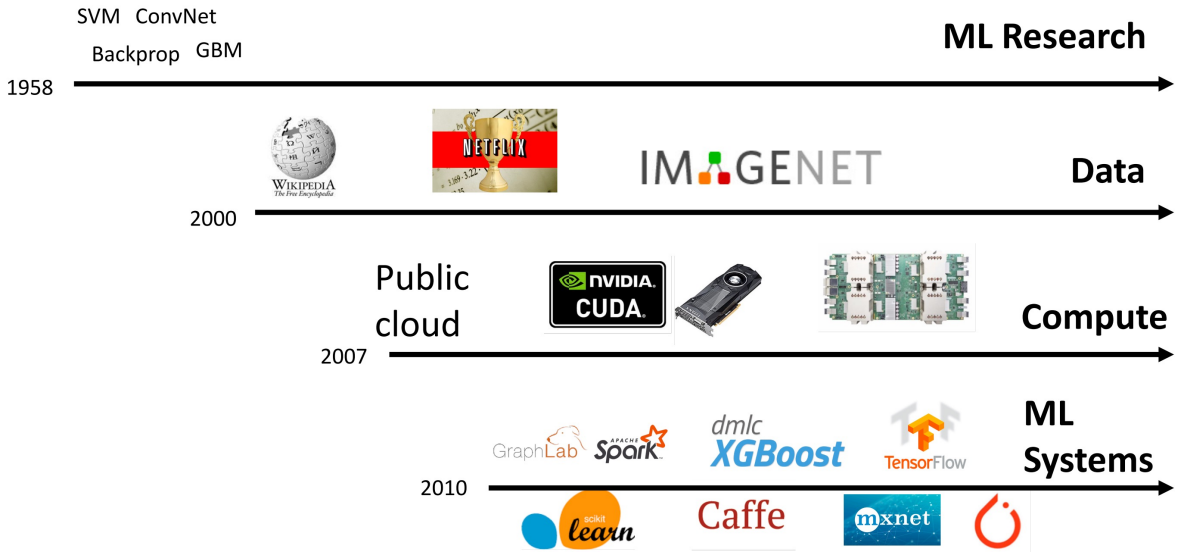
- Credit to CMU 15884 by Tianqi Chen

Case Study: Ingredient of AlexNet



- Credit to CMU 15884 by Tianqi Chen

Evolution of ML



- Credit to CMU 15884 by Tianqi Chen

Machine Learning: A Key Driver of the Modern Era

- **Ubiquitous Applications**

- From speech recognition to predictive analytics

- **Expansive Data Sets**

- Datasets have grown exponentially: ImageNet, for instance, now has over 14 million images across thousands of categories

Machine Learning: A Key Driver of the Modern Era

- **Enhanced Algorithms and Computational Power**

- Deep Learning breakthroughs: Transformers, GANs, and reinforcement learning.
- High-performance computing: GPUs and TPUs facilitating faster model training.

- **Advanced Development Paradigms**

- The rise of AutoML and neural architecture search simplifies the creation of complex models.

Thank you !