# CMSC 422 Introduction to Machine Learning

Furong Huang / furongh@umd.edu / Department of Computer Science

UNIVERSITY OF MARYLAND

# What is Machine Learning?

Machine Learning studies representations and algorithms that allow machines to improve their performance on a task from experience.

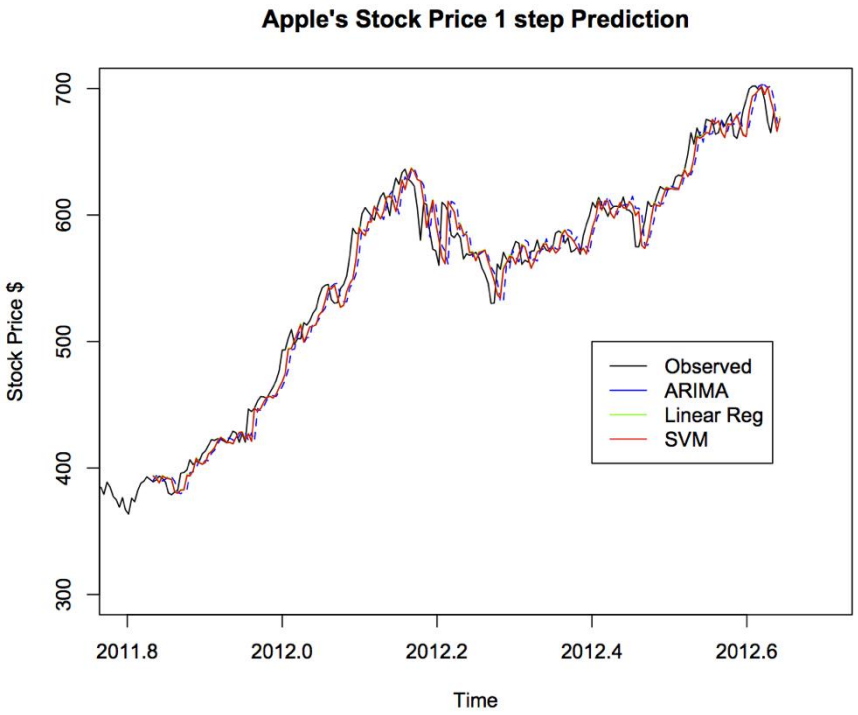
# What is Machine Learning?

- Data driven science
  - Finding and exploiting patterns in data
- Algorithms
  - Data independent algorithms
- Prediction
  - Ultimate goal

# What can ML do?


Analyze text & speech

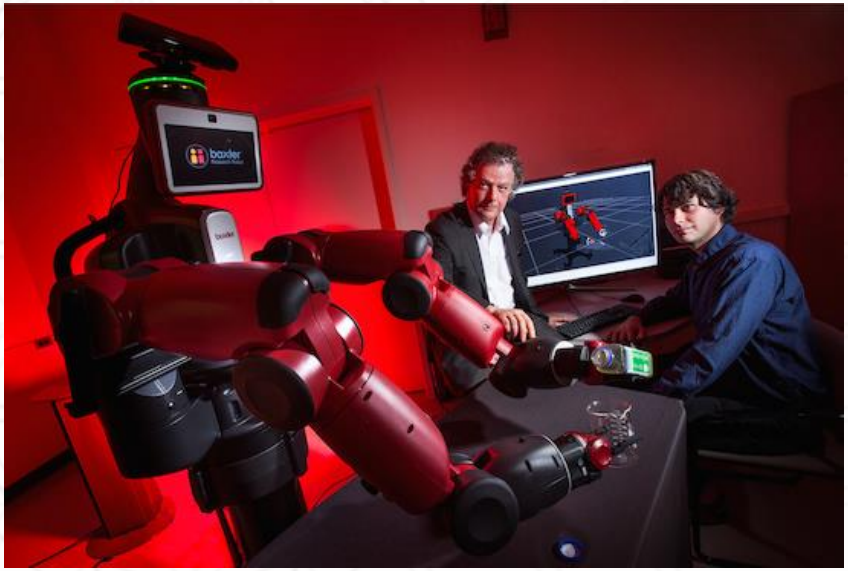
# What can ML do?




Stock price prediction

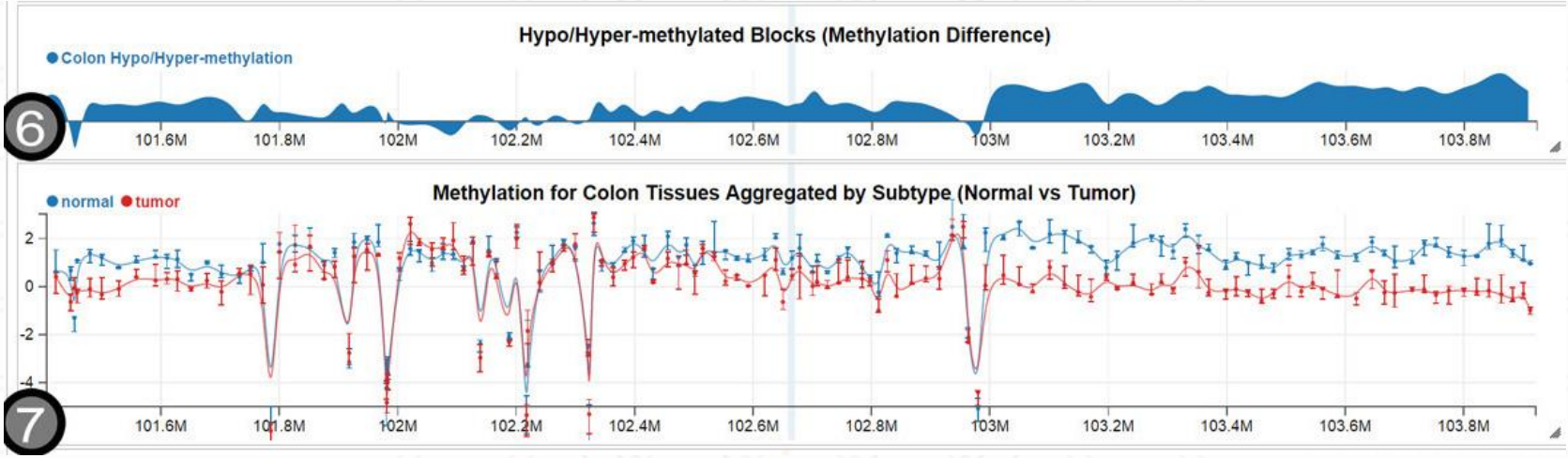
# What can ML do?


Recognize objects in images



Teach robots how to cook from youtube videos

# What can ML do?




Analyze genomics data

exploration and understanding of correlations between various genome features

# What can ML do?

- Intelligent Machine
  - As smart as humans
    - Mechanical turk evaluation, cloud sourcing
  - Smarter than human
    - Computation Power
    - Memory

# What can ML do?

nature International weekly journal of science

Home | News & Comment | Research | Careers & Jobs | Current Issue | Archive | Audio & Video | For

Current Issue > Articles > Article

*NATURE* | ARTICLE

日本語要約

## Mastering the game of Go with deep neural networks and tree search

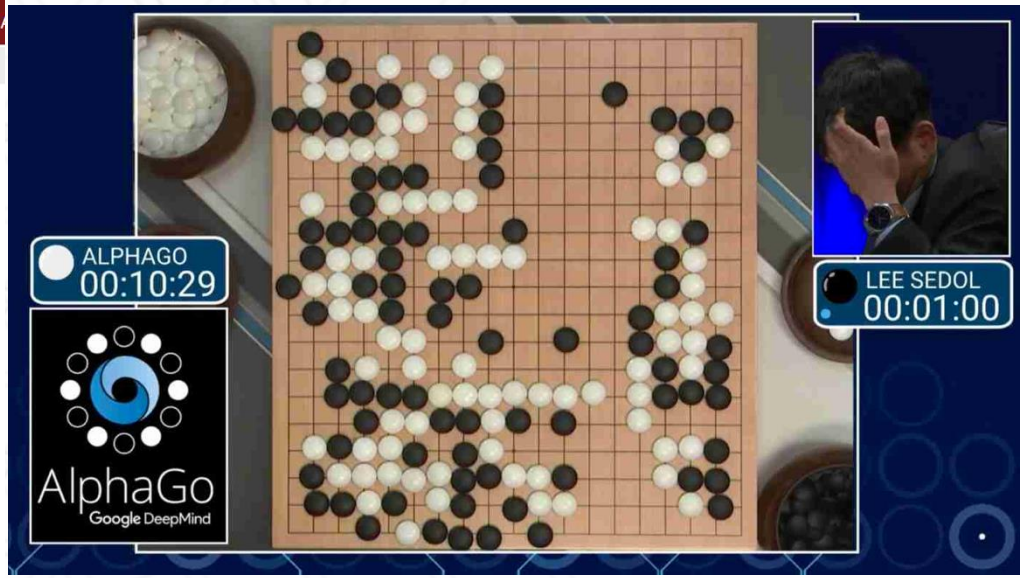
David Silver, Aja Huang, Chris J. Maddison, Arthur Guez, Laurent Sifre, George van den Driessche, Julian Schrittwieser, Ioannis Antonoglou, Veda Panneershelvam, Marc Lanctot, Sander Dieleman, Dominik Grewe, John Nham, Nal Kalchbrenner, Ilya Sutskever, Timothy Lillicrap, Madeleine Leach, Koray Kavukcuoglu, Thore Graepel & Demis Hassabis

Affiliations | Contributions | Corresponding authors

*Nature* **529**, 484–489 (28 January 2016) | doi:10.1038/nature16961

Received 11 November 2015 | Accepted 05 January 2016 | Published online 27 January 2016

PDF | Citation | Reprints | Rights & permissions | Article metrics



# What can ML do?

Question Answering system beats Jeopardy champion Ken Jennings at Quiz bowl!

# Other ML applications?

# Machine Learning

- Paradigm: "Programming by example"
  - Replace "human writing code" with "human supplying data'"
- Most central issue: generalization
  - How to abstract from "training'" examples to "test" examples? Memorization vs learning?
  - Analogy with human learning?

# Machine Learning vs Programming

- Algorithm
  - ML: usually with uncertainty
  - Programming: usually deterministic
- Goal
  - ML: prediction, do better in the future using data in the past
  - Programming: computing
- Data dependency
  - ML: universal for all or a wide class of data
  - Programming: data specific

# A fast-moving field

- Broad applicability
  - Daily life: finance, natural language processing, computer vision, robotics, healthcare, medicine, biology
- Open/young field
  - Deep learning, gap between theory and practice
- Fear of AI/ML
  - Understand what's in the black box

# CMSC 422 Goal

- This is a broad overview of existing methods for machine learning and an introduction to adaptive systems in general.

# CMSC 422 Goal

- By the end of the semester, you should be able to
  - Look at a problem
  - Identify if ML is an appropriate solution
  - If so, identify what types of algorithms might be applicable
  - Apply those algorithms

- This course is not
  - A survey of ML algorithms
  - A tutorial on ML toolkits such as Pytorch, TensorFlow, …

# CMSC 422 Topics

**Foundations of Supervised Learning**

- Decision trees and inductive bias
- Geometry and nearest neighbors
- Perceptron
- Practical concerns: feature design, evaluation, debugging
- Beyond binary classification

**Advanced Supervised Learning**

- Linear models and gradient descent
- Support Vector Machines
- Naive Bayes models and probabilistic modeling
- Neural networks
- Kernels
- Ensemble learning

# CMSC 422 Topics

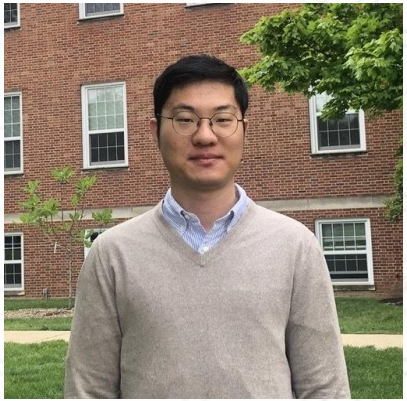
**Unsupervised learning**

- K-means
- PCA
- Expectation maximization

**Selected advanced topics (as time permits)**

- Expectation maximization
  - –Online learning
  - –Markov decision processes
- Imitation learning

# Teaching Assistants



Minghui Liu

Pankayaraj

**Answer questions on Piazza and during their office hours**

# What you can expect from the instructors

- Introducing concepts from multiple perspectives
  - Lecture, reading material, office hours, online discussion (Piazza)
- Providing opportunities to practice
  - Homework, programming projects

# What I expect from you

- Math background
  - Calculus, linear algebra, probability
- Come to class prepared
  - **Pre-class Readings required**
- Attend the lectures
  - Ask questions, take notes if necessary
- Complete HW independently
- Complete Projects collaboratively
- Be active on Piazza

# Course Logistics - Grading

- Homework (30%)
  - Released on canvas weekly, usually due **Thursday 10:30am (subject to change)**
  - **No late submissions**
- Three projects (30%)
  - Form groups of two or three
- Midterm exam (20%)
  - **March 12 12:30 pm - 1:45 pm**, open book, no internet access
- Final exam (20%)
  - **May 12 12:30 pm - 1:45 pm**, open book, no internet access

# Course Logistics

Quick Links (also in the syllabus on Canvas/Elms)

- Pre-class Reading and Slides: Website
- Textbook: A Course in Machine Learning
- Assignment release and submission: Canvas
- Grades: Canvas
- After class discussions/surveys, ask questions and everything else: Piazza
  - **Please use Piazza instead of email**

# University Rules

# Class Rules

- Nothing may be handed in late without:
  - prior arrangements (at least one week in advance), or
  - a doctor's note (or equivalent), modulo university regulations
- You are encouraged to ask questions in class or on Piazza
- Add/drop deadlines
  - University policy: http://registrar.umd.edu/current/registration/Schedule%20Adjustment.html
  - If you're going to drop, please please please do it now!
- Cheating: see academic integrity policy
- ADA/DSS:
  - Letter of accommodation to me within the first two week of class
  - Please also see ADA/DSS policy

# Policy on class behavior

The open exchange of ideas, the freedom of thought and expression, and respectful scientific debate are central to the aims and goals of a this course. These require a community and an environment that recognizes the inherent worth of every person and group, that fosters dignity, understanding, and mutual respect, and that embraces diversity. **Harassment and hostile behavior are unwelcome in any part of this course.** This includes: speech or behavior that intimidates, creates discomfort, or interferes with a person's participation or opportunity for participation in the course.

We aim for this course to be an environment where harassment in any form does not happen, including but not limited to: harassment based on race, gender, religion, age, color, national origin, ancestry, disability, sexual orientation, or gender identity. Harassment includes degrading verbal comments, deliberate intimidation, stalking, harassing photography or recording, inappropriate physical contact, and unwelcome sexual attention. Please contact an instructor or CS staff member if you have questions or if you feel you are the victim of harassment (or otherwise witness harassment of others).

*Edited from the NAACL Policy: http://naacl.org/policies/anti-harassment.html*

# Academic Integrity Policy

Any assignment or exam that is handed in must be your own work (unless otherwise stated). However, talking with one another to understand the material better is strongly encouraged. Recognizing the distinction between cheating and cooperation is very important. If you copy someone else's solution, you are cheating. If you let someone else copy your solution, you are cheating (this includes *posting solutions online in a public place)*. If someone dictates a solution to you, you are cheating.

Everything you hand in must be in your own words, and based on your own understanding of the solution. If someone helps you understand the problem during a high-level discussion, you are not cheating. We strongly encourage students to help one another understand the material presented in class, in the book, and general issues relevant to the assignments. When taking an exam, you must work independently. Any collaboration during an exam will be considered cheating. Any student who is caught cheating will be given an F in the course and referred to the University Office of Student Conduct. Please don't take that chance - if you're having trouble understanding the material, please let us know and we will be more than happy to help.

# ADA/DSS Policy

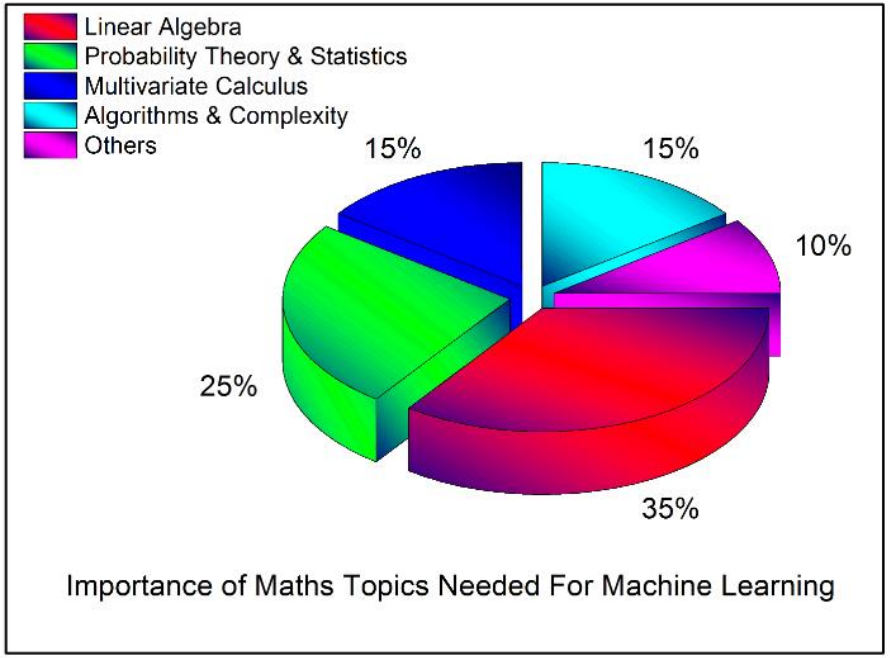
Any student eligible for and requesting reasonable academic accommodations due to a disability is requested to provide, to the instructor in office hours, a letter of accommodation from the Office of Disability Support Services (DSS) within the first two weeks of the semester. You may reach them at 301-314-7682 or by visiting Susquehanna Hall on the 4th Floor.

The CS department does not consider requests for retroactive accommodation to be reasonable. In the same vein, we do not consider it reasonable to ask an instructor to create an alternate assignment of substance. The spirit of our accommodation is to help DSS-advised students find creative ways to meet the high standards we set for all our students.

# Prerequisites

- CMSC351 (Algorithms) and CMSC330 (Programming Languages)
- Recommended: STAT400 (Applied probability and statistics) and Linear Algebra.
- These previous courses require CMSC250 (Discrete Structures), CMSC216 (Computer Systems)
- Which in turn require CMSC131 (Object oriented programming) and MATH141 (Calculus)
- Course is about data, representations, mathematical modeling, and programming

# Math in ML


Linear Algebra
Probability Theory & Statistics
Multivariate Calculus
Algorithms & Complexity
Others
15%
15%
10%
25%
35%
Importance of Maths Topics Needed For Machine Learning

# Reviews
# - Linear Algebra

# Basics

- Sets, vector space
- $\mathbb{R}^N$: N-dimensional **Euclidean space**
- A **vector** $\boldsymbol{a} \in \mathbb{R}^N$ **is a N-tuple** $[a_1, \ldots, a_N]$**, where** $\boldsymbol{a_i} \in \mathbb{R}$
- **Inner/Dot product:**

$$\boldsymbol{a}, \boldsymbol{b} \in \mathbb{R}^{\boldsymbol{N}}, < \boldsymbol{a}, \boldsymbol{b} > = \sum_{\boldsymbol{i=1}}^{\boldsymbol{N}} a_i b_i = \boldsymbol{a}^\top \boldsymbol{b} = \boldsymbol{b}^\top \boldsymbol{a} =< \boldsymbol{b}, \boldsymbol{a} > \in \mathbb{R}$$

- **Euclidean norm:** Induced by the inner product

$$\boldsymbol{a} \in \mathbb{R}^{\boldsymbol{N}}, \|\boldsymbol{a}\|_{\boldsymbol{2}} = \sqrt{< \boldsymbol{a}, \boldsymbol{a} >}$$

- Don't confuse the norm $\|\boldsymbol{x}\|_2$ with the absolute value $|x|$

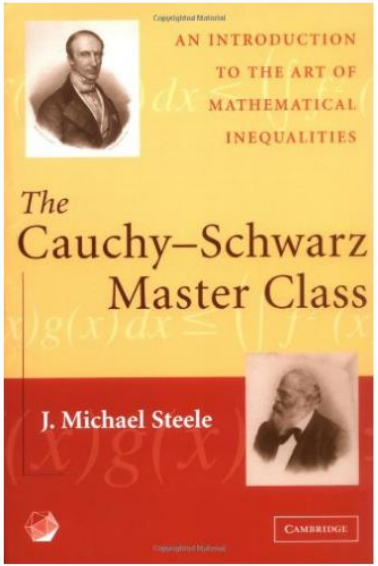
# Cauchy-Schwarz inequality

**Lemma (Cauchy-Schwarz inequality)**

*Given* $\mathbf{a}, \mathbf{b} \in \mathbb{R}^N$,

$$|\langle \mathbf{a}, \mathbf{b} \rangle| \leq \|\mathbf{a}\|_2 \|\mathbf{b}\|_2.$$

- Probably the most important inequality out there!
- There is a book solely devoted to this inequality.
- When does it hold with equality?
- Is used to derive the triangle inequality shown next



# Triangle inequality

**Lemma (Triangle inequality)**

*Given* $\mathbf{a}, \mathbf{b} \in \mathbb{R}^N$,

$$\|\mathbf{a} + \mathbf{b}\|_2 \leq \|\mathbf{a}\|_2 + \|\mathbf{b}\|_2.$$

- When does this inequality hold with equality?
- Reverse (or inverse) triangle inequalities:

$$\|\mathbf{a} + \mathbf{b}\|_2 \geq \|\mathbf{a}\|_2 - \|\mathbf{b}\|_2$$
$$\|\mathbf{a} + \mathbf{b}\|_2 \geq \|\mathbf{b}\|_2 - \|\mathbf{a}\|_2$$

# Collection of vectors, Subspaces

A set of $T$-vectors, $V = \{\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_T\}$

- **Linear combination:** $\sum_{k=1}^{T} \alpha_k \mathbf{a}_k, \alpha_k \in \mathbb{R}$
- **Linearly independent:** No vector in $V$ can be written as linear combination of others
- **Span:** $\text{Span}(V) = \{\mathbf{x} \,|\, \mathbf{x} = \sum_k \alpha_k \mathbf{a}_k, \alpha_k \in \mathbb{R}\}$

# Collection of vectors, Subspaces

## Definition (Subspace)

A collection of vectors $V \subset \mathbb{R}^N$ is a subspace iff it is closed under linear combinations

$$\mathbf{a}, \mathbf{b} \in V \implies \alpha\mathbf{a} + \beta\mathbf{b} \in V, \alpha, \beta \in \mathbb{R}$$

- **Basis of a subspace:** A linearly independent *spanning* set
- **Dimensionality of a subspace:** #elements in a basis

# Matrix

- $A \in \mathbb{R}^{M \times N}$ : A matrix of dimension $M \times N$
- $A = [a_{ij}] = [\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_N], \mathbf{a}_i \in \mathbb{R}^M$
- $\mathrm{rank}(A) =$ largest number of linearly *independent* columns
- $\mathrm{rank}(A) = \mathrm{rank}(A^T) \leq \min(M, N)$
- $A$ is *full-rank* if $\mathrm{rank}(A) = \min(M, N)$.

Matrices are *representations of linear operators.*

$A : \mathbb{R}^N \to \mathbb{R}^M$
$\mathbf{x} \in \mathbb{R}^N \mapsto A\mathbf{x} \in \mathbb{R}^M$

# Your task before next class

- Check out course webpage, Canvas, and Piazza
- **Do the readings**
- Let TAs know dates of religious holiday you observe this semester
- Let me know if you will need DSS arrangements

UNIVERSITY OF
MARYLAND

**Furong Huang**
4204 IRB, College Park, MD 20740
301.405.8010 / furongh@umd.edu