

Learning From Data

Lecture 1

The Learning Problem

Introduction

Motivation

Credit Default - A Running Example

Summary of the Learning Problem

M. Magdon-Ismail

CSCI 4100/6100

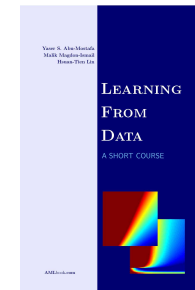
Resources

1. Web Page: www.cs.rpi.edu/~magdon/courses/learn.php
 - course info: www.cs.rpi.edu/~magdon/courses/learn/info.pdf
 - slides: www.cs.rpi.edu/~magdon/courses/learn/slides.html
 - assignments: www.cs.rpi.edu/~magdon/courses/learn/assign.html

2. Text Book:

Learning From Data

Abu-Mostafa, Magdon-Ismail, Lin



3. Piazza

4. TAs.

5. Professor.

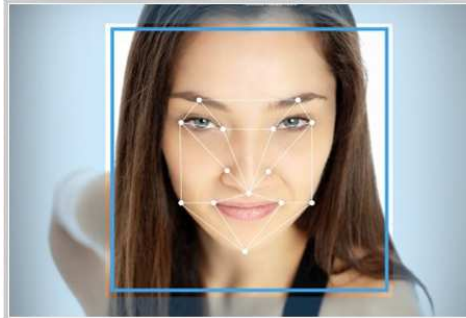
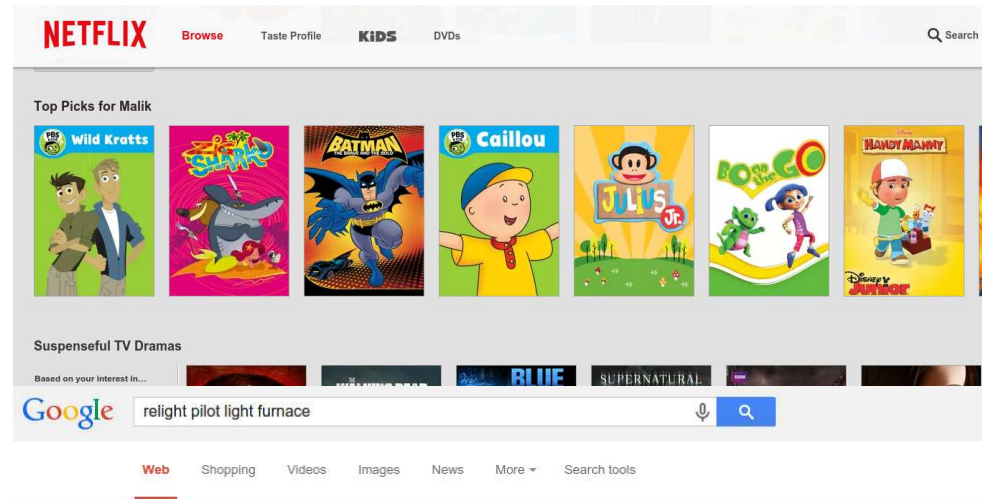
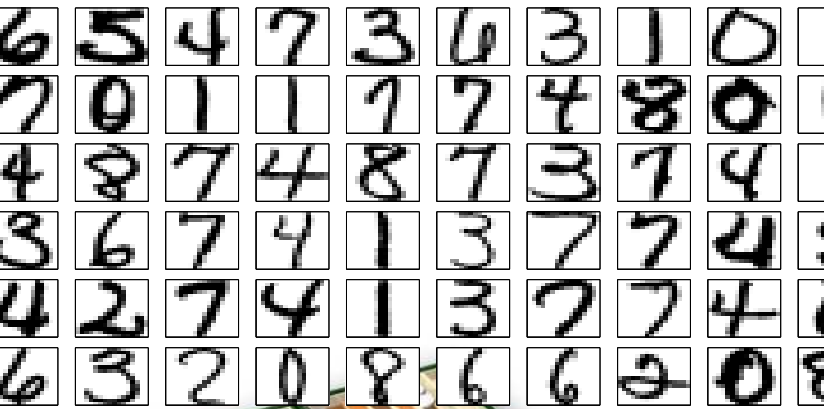
6. Prerequisites? assignment #0

The Storyline

1. What is Learning?
2. Can We do it?
3. How to do it?
4. How to do it well?
5. General principles?
6. Advanced techniques.
7. Other Learning Paradigms.

■ concepts
■ theory
■ practice

our language will be mathematics ...
...our sword will be computer algorithms



- How to relight the pilot on the gas furnace - YouTube**
www.youtube.com/watch?v=4x4v1WAYrFY
 Dec 4, 2013 - Uploaded by grayfurnaceman
 This one covers how to **light the pilot** on the **gas furnace**. This video is part of the heating and cooling series ...
- How to turn your furnace pilot light on - YouTube**
www.youtube.com/watch?v=b4BiZhrqHYQ
 Oct 30, 2010 - Uploaded by MrOzcar82
 How to **relight the pilot** on the **gas furnace** by grayfurnaceman
 45,466 views ... How To Light The Pilot ...
- How to Light a Furnace Pilot Light - HowStuffWorks**
home.howstuffworks.com/how-to-repair-gas-furnaces2.htm
 The **pilot light** on a **gas furnace** can go out because of drafts. To **relight the pilot**, follow the manufacturer's instructions exactly; they are usually fastened to the ...
- How to Light a Gas Furnace Standing Pilot - Relighting**
homerepair.about.com/od/heatingcoolingrepair/ss/pilot_light_3.htm
 This easy to understand tutorial describes how to light a **gas furnace** standing pilot and tips for troubleshooting the **pilot light** and thermocouple.
- 3 Ways to Light a Pilot Light - wikiHow**
www.wikihow.com/Light-a-Pilot-Light - wikiHow
 Although newer **gas furnaces**, boilers, and appliances have electronic starters, there are many older models in service ... This article will tell you how to light a **pilot light** on a natural gas appliance or **furnace**. ... 4. **Relight** your pilot with a match.

Shop for relight pilot light furnace on Google

Robertshaw 785-001 24/120 V. Relight Kit
 \$48.95 - SupplyHouse.com
 Free Shipping On Orders Over \$49

Crisafulli Bros. Heating
www.crisbro.com/
 "Best Of" 8 Years In A Row. BBB A+ Over 75 years, Coupons 24/7 Service
 520 Livingston Avenue, Albany, NY (518) 449-1782

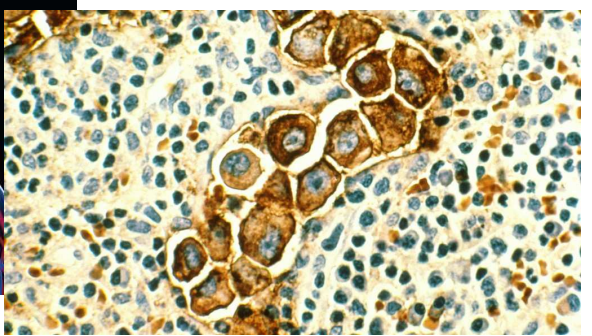
Relight Pilot Light Furnace
www.com/Relight+Pilot+Light+Furnace
 Search Relight Pilot Light Furnace
 Look Up Quick Results Now!
 See your ad here »

Machine learning is like walking forward while looking backward. Explain

The analogy "Machine learning is like walking forward while looking backward" highlights a key aspect of how machine learning algorithms work:

- Walking Forward:** In machine learning, the "forward" direction represents the model's primary task, which is to make predictions or decisions based on input data. This is akin to moving forward toward a goal.
- Looking Backward:** The "looking backward" part refers to the learning process. During training, machine learning models analyze historical or past data (looking backward) to identify patterns, trends, and relationships. It learns from this historical data to improve its ability to make accurate predictions in the future (walking forward).

In essence, machine learning algorithms use past data to inform their future actions. They draw insights from what has happened before (looking backward) to make informed decisions or predictions about new, unseen data (walking forward). This iterative learning process is fundamental to the concept of machine learning.



Let's *Define* a Tree?

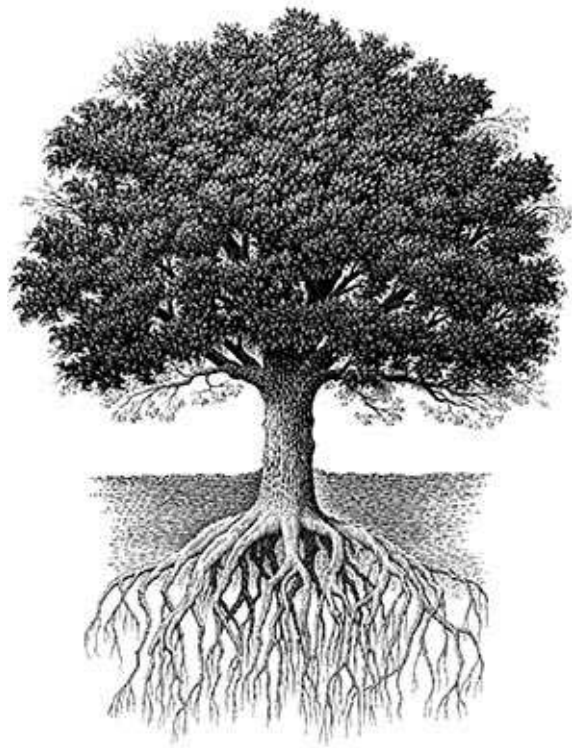


Let's *Define* a Tree?



A brown *trunk* moving upwards and *branching* with *leaves* ...

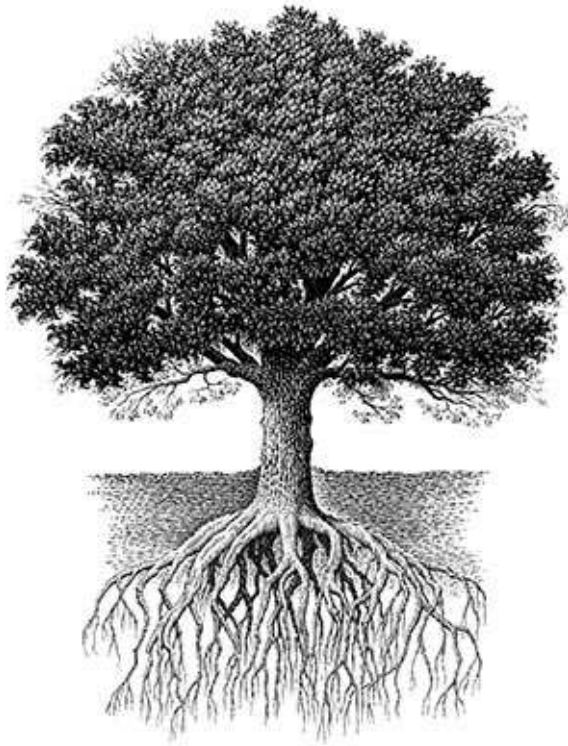
Are These Trees?



Learning “What are Trees” is ‘Easy’



Defining is Hard; Recognizing is Easy



Hard to give a complete mathematical definition of a tree.

Even a 3 year old can tell a tree from a non-tree.

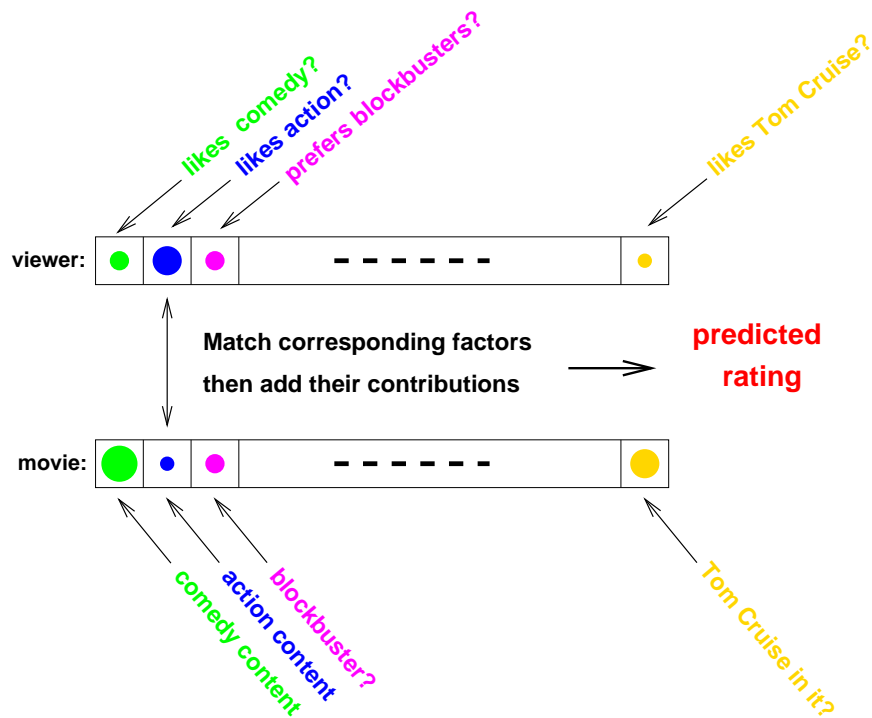
The 3 year old has learned from data.

(Other tasks like graphics or GAN?)

Learning to Rate Movies

- Can we predict how a viewer would rate a movie?
- Why? So that Netflix can make better movie recommendations, and get more rentals.
- **\$1 million** prize for a mere 10% improvement in their *recommendation system*.

Previous Ratings Reflect Future Ratings



- Viewer taste & movie content imply viewer rating.
- No magical formula to predict viewer rating.
- Netflix has data. We can **learn** to identify movie “categories” as well as viewer “preferences”

Class Motto:

A pattern exists. We don't know it. We have data to learn it.

Credit Approval

Let's use a conceptual example to crystallize the issues.

age	32 years
gender	male
salary	40,000
debt	26,000
years in job	1 year
years at home	3 years
...	...

Approve for credit?

Credit Approval

Let's use a conceptual example to crystallize the issues.

- Using salary, debt, years in residence, etc., approve for credit or not.
- No magic credit approval formula.
- Banks have lots of data.
 - customer information: salary, debt, etc.
 - whether or not they defaulted on their credit.

age	32 years
gender	male
salary	40,000
debt	26,000
years in job	1 year
years at home	3 years
...	...

Approve for credit?

A pattern exists. We don't know it. We have data to learn it.

The Key Players

- Salary, debt, years in residence, ...
- Approve credit or not
- True relationship between \mathbf{x} and y
- Data on customers

input $\mathbf{x} \in \mathbb{R}^d = \mathcal{X}$.

output $y \in \{-1, +1\} = \mathcal{Y}$.

target function $f : \mathcal{X} \mapsto \mathcal{Y}$.

(The target f is *unknown*.)

data set $\mathcal{D} = (\mathbf{x}_1, y_1), \dots, (\mathbf{x}_N, y_N)$.

($y_n = f(\mathbf{x}_n)$.)

\mathcal{X} \mathcal{Y} and \mathcal{D} are *given* by the learning problem;
The target f is fixed but unknown.

We learn the function f from the data \mathcal{D} .

Learning

- Start with a set of candidate hypotheses \mathcal{H} which you think are likely to represent f .

$$\mathcal{H} = \{h_1, h_2, \dots, \}$$

is called the hypothesis set or *model*.

- Select a hypothesis g from \mathcal{H} . The way we do this is called a *learning algorithm*.
- Use g for new customers. We hope $g \approx f$.

\mathcal{X} \mathcal{Y} and \mathcal{D} are *given* by the learning problem;

The target f is **fixed but unknown**.

We choose \mathcal{H} and the learning algorithm

This is a very general setup (eg. choose \mathcal{H} to be all possible hypotheses)

Summary of the Learning Setup

