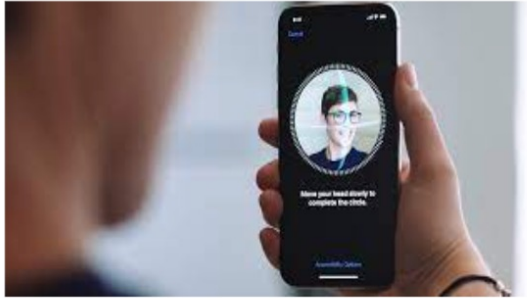


# Machine Learning: *Past, Present and Future*

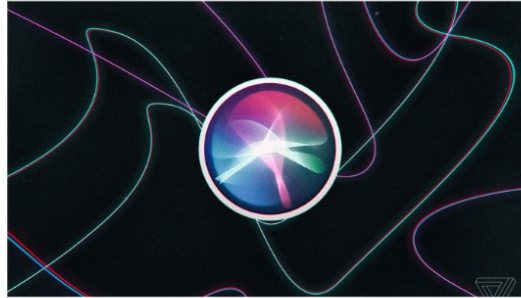
Michael Soltys

March 14, 2022 @ CSUCI

*Sabbatical Presentation*



**Image Recognition**



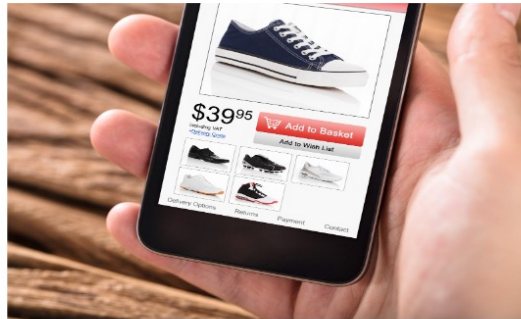
**Speech Recognition**



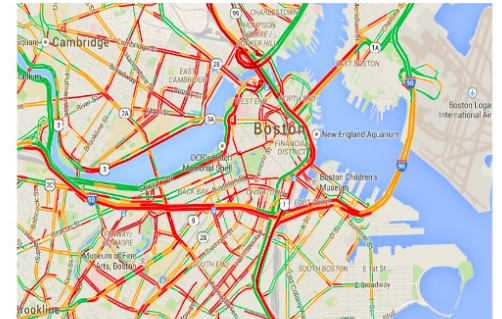
**Fraud Detection**



**Self-Driving Cars**



**Product Recommendation**



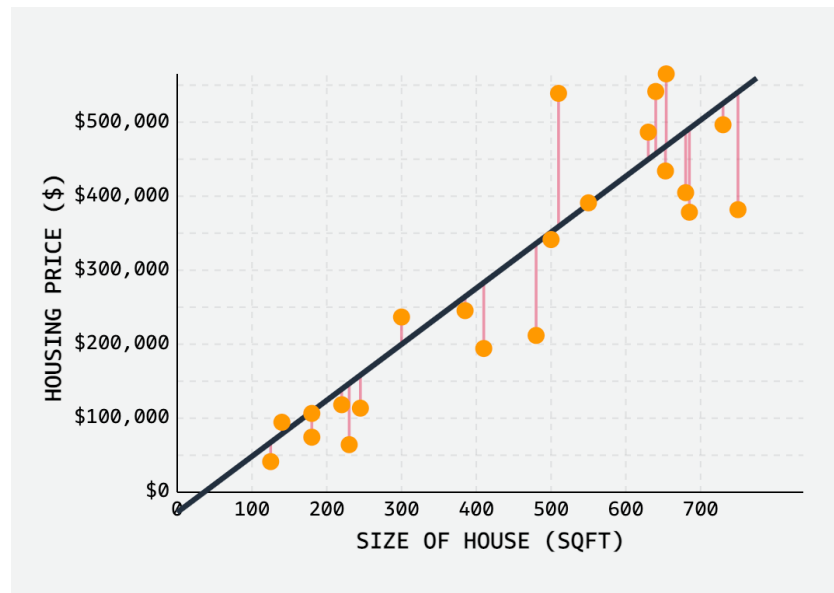
**Traffic Prediction**

# What is ML?

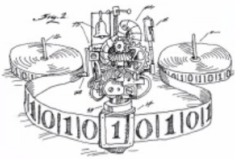
- ML is data-driven (as opposed to rule-driven) computation
- It is a subfield of AI (Artificial Intelligence)

# Example: Linear Regression

- <https://mlu-explain.github.io/linear-regression/>



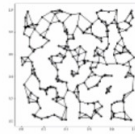
*Alan Turing proposed  
The Turing Test*



*The Dartmouth Summer  
Research Project on AI.*



*The “nearest neighbor”  
algorithm is created, allowing  
computers to use basic  
pattern recognition.*



*IBM's Deep Blue, a chess-  
playing computer program,  
defeated the reigning chess  
world champion*



*Google Brain is  
developed, a neural  
network able to discover  
and categorize objects.*



1950

1952

1956

1957

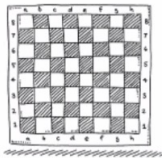
1967

1979

1996

2011

2012



*Arthur Samuel wrote the first  
computer learning program: the  
game of checkers.*



*Frank Rosenblatt  
designed the first neural  
network for computers:  
the perceptron.*



*Stanford University students  
invent the “Stanford Cart,”  
which can navigate obstacles  
in a room on its own.*

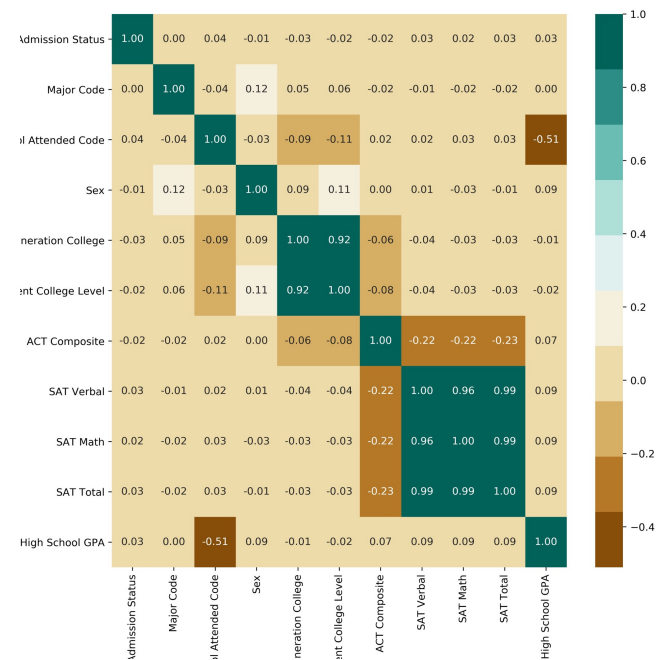
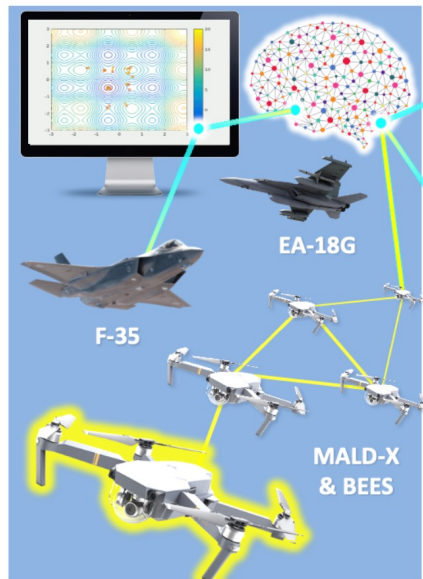


*IBM's Watson  
computer beat  
two champions  
on Jeopardy.*

# Example

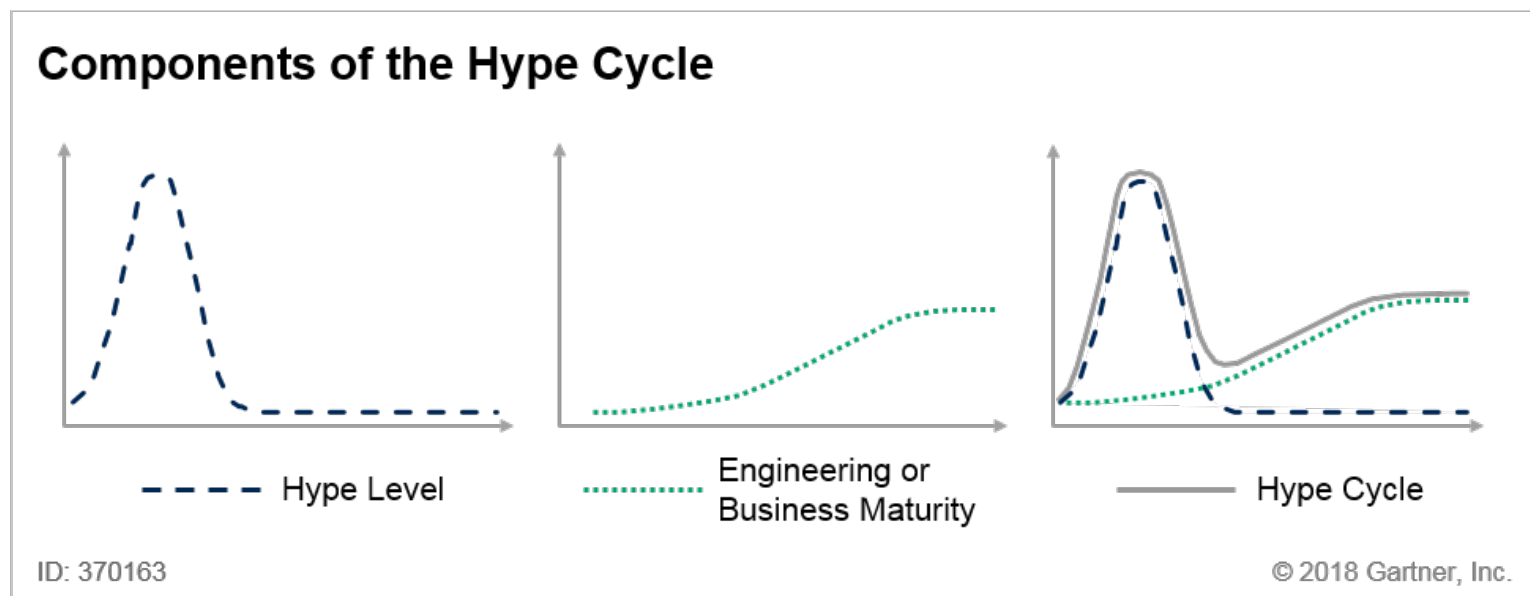


AI/Machine Learning Technologies |

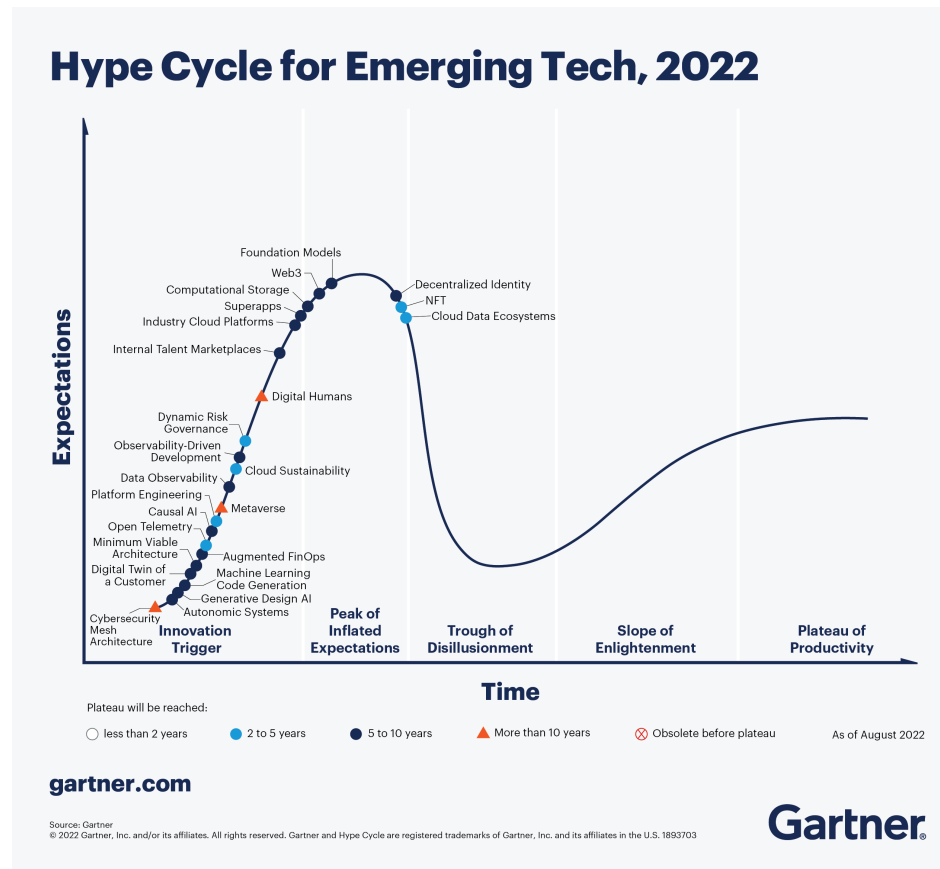


<https://github.com/michaelsoltys/sagemaker-enrollment>

# Gartner Hype Cycle



# Where is ML in the hype cycle?





Past

*Theory and bespoke code*

# University of Toronto



Geoffrey Hinton was pioneering deep learning (1990-2015)

The New York Times.

In Greater New York, 2 Elsewhere  
Jersey City and Newark, (Two Call

NEW YORK, FRIDAY, OCTOBER 9, 1903.—FOURTEEN PAGES.

ONE CENT

**MAY SUSPEND BALKAN REVOLT. TO EXTRADITE GRAFTERS**

Yesterday Part of Manchuria Should  
Have Been Evacuated, but No  
Move Has Been Made by  
the Russians There.

—TIEN-TSUN, Oct. 8 — A telegram from Yixiow, under date of Oct. 8, states that the Russians are taking no steps to evacuate NIAO-CHUNG.

The members of parts of Manchuria by the Russian troops passed without a blow and were toward the straits.

In the meantime the negotiations regarding Korea and Manchuria are proceeding at Tokyo between the Japanese Cabinet and the Russian Minister, Gromov de Russie. The latter is writing under these instructions from Admiral Aleksei Viceroy of the Far East. The prospect of an agreement is considered doubtful, while Russia makes substantial concessions.

The rumor of the Japanese intention of the Russian demands is officially denied.

**MAY SUSPEND BALKAN REVOLT. TO EXTRADITE GRAFTERS**

**London Times—New York Times**  
 Same Column.

**LONDON, Oct. 8-11** is assigned from Constantinople through official channels, says the Times correspondent of The Times, that the Mohammedan Insurrection is reported to be growing into a successful rebellion. That, in cooperation of the Porte against the efforts to suppress the rising, they have decided to suspend the intervention until next year.

There are something like 100,000 cars on roads and parked in lots at the proposed De Witt Clinton Expressway interchange. The methods to be used are:

  - 1. To make the interchange, the expressway, that Elmore would need some light of the current situation of advancing the expressway.
  - 2. To make the interchange, the expressway, that Elmore would need some light of the current situation of advancing the expressway.

**LIFE INSURANCE AGENTS SOLD.**

**R. H. Smith, of Guyan, says he is**  
**ready to sell his business to the**  
**State of New York.**

Smith, of Guyan, says he is ready to sell his business to the State of New York. Smith, of Guyan, says he is ready to sell his business to the State of New York.

[illegible]

**PRESIDENT GIVEN REVISIONS.**  
Discrepancies Found in Accounts of Lancaster County (Penn.) Railway Light and Power Plant.  
LANCASTER, Pa., Oct. 8.—A 30-framer which has been carried for some time attempting the relations of WILLIAM B. GIBBS, President of the Lancaster County Railway and Light and Power Co., and the late President of the Pennsylvania Electric Co., is submitted today by the association that

**\$150,000,000 MORTGAGE FILED.**  
Given by Levisa Valley Railroad to the

This document is in the name of general consolidated bonds, which were being issued by the Directors. The bonds are to be issued from time to time, and will eventually amount to \$100,000, and will eventually amount to \$100,000, and will eventually amount to \$100,000.

Went to Milwaukee, and Wedded  
a Kid de Law Girl.

MILWAUKEE, Wis., 6.—ARTHUR C. NADOL, who absconded from Indiana with \$2500, on June 28 and who was removed on extradition to Milwaukee, continued to prosper in Chicago today.

The young man worked for a Campbell bank, and it was his duty to carry large sums of money from one bank to another. He will be paid.

Heard arrested in Milwaukee shortly after leaving Dallas, and made the acquaintance of a girl named "Lulu" de Law, Wis. Two weeks later they were married, the wedding being performed in Chicago.

The young couple, who are now in the city, then took a wedding trip to the Pacific Coast and returned to Milwaukee almost three months ago.

THE FATHERS OF THE THREE BROTHERS, who were arrested in Dallas shortly after the escape of their son, are now in Chicago, and are giving aid, as far as they are able, to the fatherless child.

It is reported that arrested men on extradition, after the 30-day term has expired, may be held in jail for 60 days.

The money was not deposited and an hour later, when the Third National Bank notified the police once again to find Nard, a heading company settled with the bank.

**Minister of Defense Says She Will Have Force of 150,000 Men.**

OTTAWA, Oct. 8.—Sir Frederick Borden, Minister of Militia and Defence, in the House today, on a vote of \$1,000,000 for arms and equipment, said that "it was his intent to have a force of 150,000 men to be called up in the case of an emergency. Of these, 75 per cent. will be called upon, and nearly 50,000 will be called upon for general drill. The other 40 per cent. will be equipped with arms and ammunition and will be ready in case of trouble."

**DENIAL BY MORGAN & CO.**

Statement of Connection With  
Shipbuilding Trust.

---

HOW SHARES WERE SOLD

---

Mr. Undermyer Says There is no  
Channel with Building Firm—

**Servicing Plan.**

"During the first 2 of 20 August 1942, my partner and I had my meeting with the Inspection, explanation, or financing of the United States Stabilizing Company, and days that they received or would pay of the securities and cash of the United States Steel Corporation. We took the Corporation Steel Company stock and a purchase could be found. It was received from the United States Steel Corporation on exact cost, and the preferred and common stock which were received in full of interest and earnings of the corporation. The proceeds were paid for the stock, and the proceeds were distributed as paid of the profits of the United States Steel Corporation."

"BY EXHIBITION that we were parties to 237 agreement for selling the stock in 1942, within."

"The witness and two bonded promissory documents by the banking house of 2. F. 364-

Mr. Dreesen had stated that the stolen TV question had been given to Mr. Kozlowski with bonds worth \$200,000.00 for the Bank.

After the formal statement had been issued by the bankers earlier in the afternoon, a meeting of the firm's directors and officers was held at the New York City Club. Supplementing it by making clear the \$5,000,000 mentioned in the statement represented what had been received for the \$100,000 par value preferred and common stock of the misnamed company. When this stock was sold, the net proceeds, less the low offer was paid to individuals in return.

The same member of the firm advised

Mr. Deane as President of the Trust Company at the Republic for the benefit of Mr. Chalmers, who was at that time the company's cashier. This report was (1894-1895).

[illegible]

And the other one will be considered an irregularity. If the bank is to make an error, it is to be made in the direction of the customer's benefit. It is not to be made in the direction of the bank's benefit. It is not to be made in the direction of the bank's benefit. It is not to be made in the direction of the bank's benefit.

One day later Mr. Alexander ("Buffy") moved to Col. McCook that he had further business of the system of the past. "This is a matter of some importance," he said, "and I shall not leave it behind the following postal time if called."

Alford said there came a radio dispatch from John C. Young to Louis Brand, the president of the American Friends of the Bazaar, an Alexandria organization, on behalf of Mammatah Young as it called. Paris correspondents there it said, Mammatah being a "slight duck," called in July of this he had seemed it to be known that a leading English baron had been in the city, and he had been in the house. He added, "What is true, of course." Two days later, it was a visit to Alexander, he said. He said, "I am not sure whether they were the same, according to my appointment, but that he may had discussed hospital that was taken under the

## M'LAUGHLIN DECLARES WAR ON M'GARDEN

**James Shevin Calls the Senator a Political Acrobat.**  
**EFFECT OF LITTLETON SPEECH**  
Right and Perhaps Ten of Brooklyn's  
Twenty-two Districts Counted as  
Against Wiloughby Street.

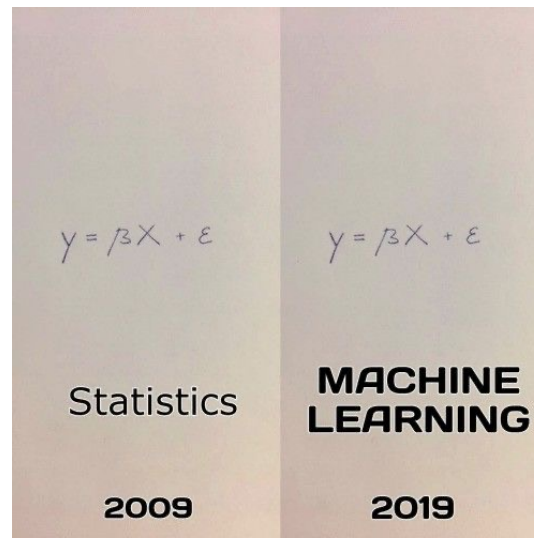
High Country, and the organization is known County, and with McLaughlin, aided by McLaughlin, John L. Egan, and Herbert W. Little, who were the main organizers of the movement. McLaughlin was the main organizer, because of his own influence in the County. McLaughlin was the main organizer of the movement, because of his own influence in the County. McLaughlin was the main organizer of the movement, because of his own influence in the County.

crucially on the advice of Sherida, who is related to the victim by marriage, and of whom McLaughlin is especially fond. Sherida generally is credited with having advised the huge demands on the Treasury Dept. management, together with the statement as "graft" and "Two-faced" which McLaughlin issued against Murphy, which together brought about the open break between Taft-Hartley and Wm. Hughes.

It is demanding that the entire Democratic Party ticket be suspended, Senator McCarran said the backing of several of the members of the district board in the Kings County organization. Congressman George H. Brown's suggestion in the Pittsburgh District, of which he is the leader, already has inspired the entire matter, so it would be the organization in Senator McCarran's Northwest District, Cuyahoga County, Cleveland, which has taken action. The South District, of which McCarran is a member, is also active, he said. The whole thing is being

1903 NYT  
weeks before  
Wright  
brothers flew  
for first time:  
*man will not  
fly for 10  
million years*

# Mathematics



#10yearchallenge

# Stephen Cook





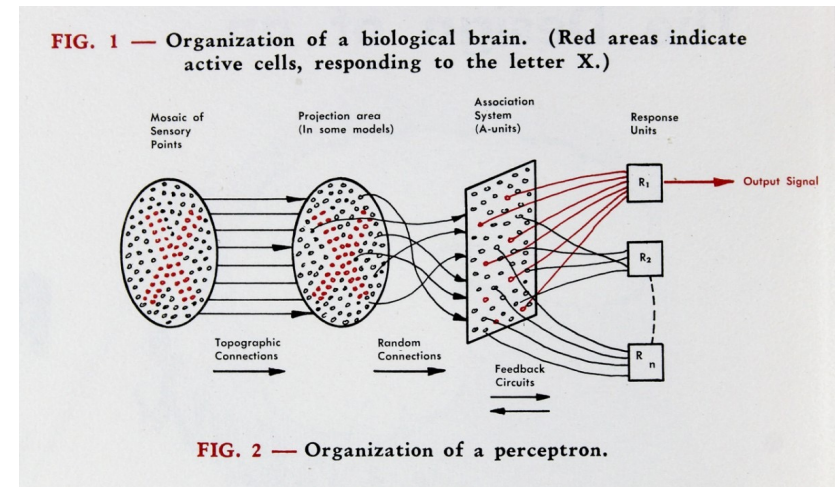
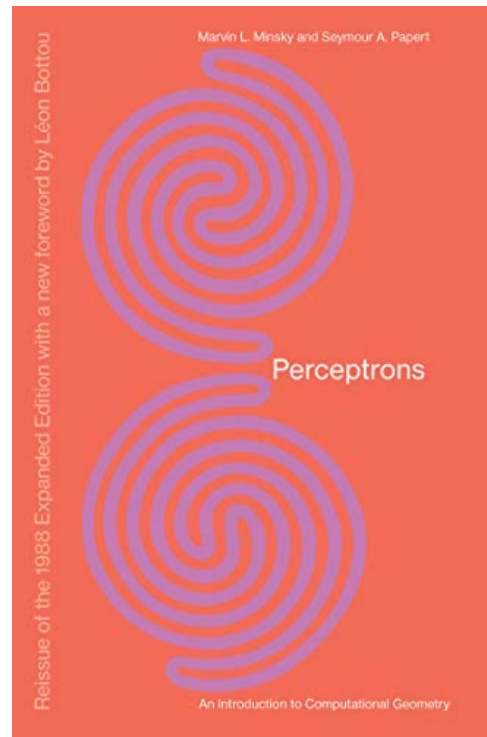
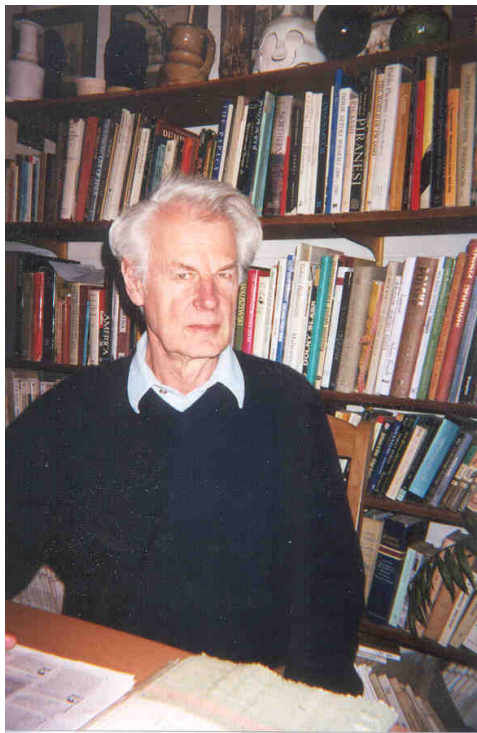
# McMaster University



Copy and paste is a design error.  
-David Parnas



# Jan Mycielski



***Perceptrons: an intro to computational geometry*** by Marvin Minsky and Seymour Papert, 1969. An edition with handwritten corrections released in the early 1970s.

# How was ML done

- Code was *bespoke*
  - Written *de novo* each time
- But by the early 2000s:
  - Shared theoretical core of knowledge:
    - Backpropagation
    - Statistical Learning Theory
- What was taught:
  - Theory of neural networks and limitation of learning algorithms
  - How to code them by hand

```
# This code uses for loops to implement backpropagation for a two-layer fully connected sigmoid network
# The network has 2 inputs, 2 hidden units, and 1 output unit

def sigmoid(x):
    return 1.0 / (1.0 + math.exp(-x))

def derivative_sigmoid(x):
    return x * (1.0 - x)

def learn():
    # Inputs
    inputs = [[1, 2], [2, 3], [3, 1], [4, 3], [5, 3], [6, 2]]
    targets = [[0], [0], [0], [1], [1], [1]]

    # Define network
    n_inputs = 2
    n_hidden = 2
    n_outputs = 1

    # Initialize weights
    weights_input_to_hidden = [[0.15, 0.2, 0.25], [0.4, 0.45, 0.5]]
    weights_hidden_to_output = [[0.6, 0.7], [0.65, 0.8], [0.8, 0.9]]

    # Train network
    for i in range(500):
        # Forward pass
        hidden_layer_in = [0, 0]
        for j in range(n_inputs):
            for k in range(n_hidden):
                hidden_layer_in[k] += inputs[j][k] * weights_input_to_hidden[j][k]
        hidden_layer_out = [sigmoid(x) for x in hidden_layer_in]
        output_layer_in = [0, 0]
        for j in range(n_hidden):
            for k in range(n_outputs):
                output_layer_in[k] += hidden_layer_out[j] * weights_hidden_to_output[j][k]
        output_layer_out = [sigmoid(x) for x in output_layer_in]
        # Backward pass
        output_errors = [0, 0]
```

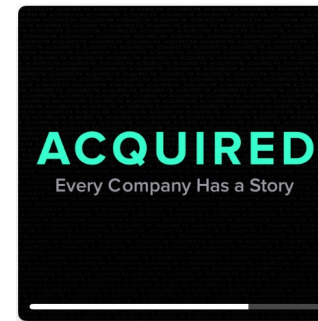


Present  
*Powerful tools*

# The Cloud as enabler

- Specs of an AWS SageMaker instance:

- `ml.g5.48xlarge`:
  - 8 NVIDIA A10G Tensor Core GPUs
  - 192 vCPUs!
  - 768GiB storage



MAR 27, 2022 · S10 E5 · 29 MIN LEFT

**Nvidia: The GPU Company**

Acquired

▶ Resume

<https://podcasts.apple.com/us/podcast/acquired/id1050462261?i=1000558142063>

- But Cloud is *not* the solution for everything:  
read [this post](#) on the Stack Overflow architecture

# Proliferation of Packages

- 2007 – Theano
- 2010 – Scikit Learn
- 2014 – Jupyter Notebooks
- 2014 – XGBoost
- 2015 – Tensorflow, Keras
- 2016 – PyTorch, MXNet

# PyTorch implementation

```
class Net(nn.Module):
    def __init__(self):
        super(Net, self).__init__()
        self.fc1 = nn.Linear(2, 2)
        self.fc2 = nn.Linear(2, 1)

    def forward(self, x):
        x = F.sigmoid(self.fc1(x))
        x = self.fc2(x)
        return x
```

```
# This code uses for loops to implement backpropagation for a two-layer fully connected sigmoid neural network
# The network has 2 inputs, 2 hidden units, and 1 output unit

def sigmoid(x):
    return 1.0 / (1.0 + math.exp(-x))

def derivative_sigmoid(x):
    return x * (1.0 - x)

def learn():
    # Inputs
    inputs = [[1, 2], [2, 3], [3, 1], [4, 3], [5, 3], [6, 2]]
    targets = [[0], [0], [0], [1], [1], [1]]

    # Define network
    n_inputs = 2
    n_hidden = 2
    n_outputs = 1

    # Initialize weights
    weights_input_to_hidden = [[0.15, 0.2, 0.25], [0.4, 0.45, 0.5]]
    weights_hidden_to_output = [[0.6, 0.7], [0.65, 0.8], [0.8, 0.9]]

    # Train network
    for i in range(500):
        # Forward pass
        hidden_layer_in = [0, 0]
        for j in range(n_inputs):
            for k in range(n_hidden):
                hidden_layer_in[k] += inputs[j][k] * weights_input_to_hidden[j][k]
        hidden_layer_out = [sigmoid(x) for x in hidden_layer_in]
        output_layer_in = [0, 0]
        for j in range(n_hidden):
            for k in range(n_outputs):
                output_layer_in[k] += hidden_layer_out[j] * weights_hidden_to_output[j][k]
        output_layer_out = [sigmoid(x) for x in output_layer_in]

        # Backward pass
        output_errors = [0, 0]
        for j in range(n_outputs):
            error = targets[j][0] - output_layer_out[j]
            output_errors[j] = error * derivative_sigmoid(output_layer_out[j])
        for k in range(n_hidden):
            error = weights_hidden_to_output[j][k] * error
            weights_hidden_to_output[j][k] += hidden_layer_out[k] * error * derivative_sigmoid(hidden_layer_out[k])
        hidden_errors = [0, 0]
        for j in range(n_hidden):
            error = 0
```



# Impact

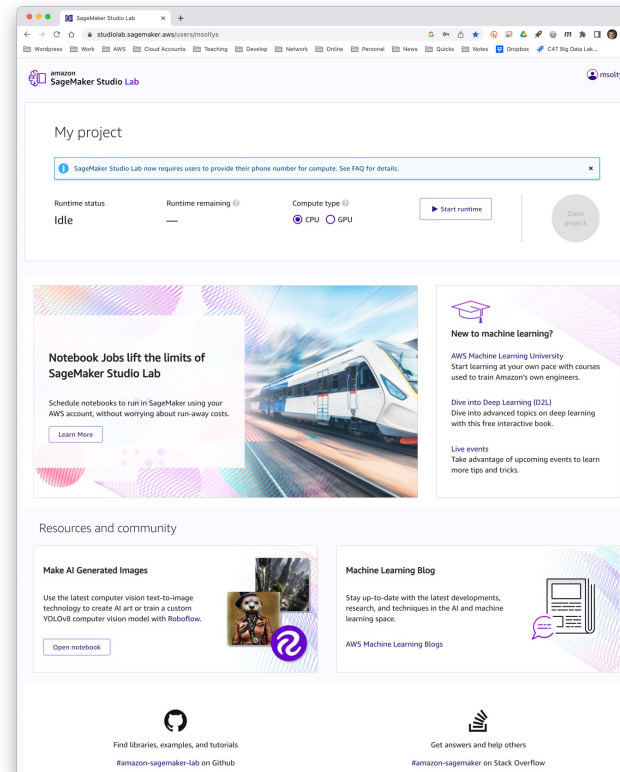
- New tools allowed practitioner to go up one level of abstraction:
  - Before: “How do I take all this math and write it in code?”
  - Now: “How can I structure this network to solve my problem?”
  - Or Even: “How do I organize my data/problem so a model can train on it?”
- Entry bar was high (PhD!), but now:
  - Moving ML from research to production with emphasis on tooling
  - Open Source tools like AutoGluon: <https://auto.gluon.ai>

```
from autogluon.tabular import TabularPredictor

predictor = TabularPredictor(label="label").fit(train_data="train.csv")
predictions = predictor.predict("test.csv")
```

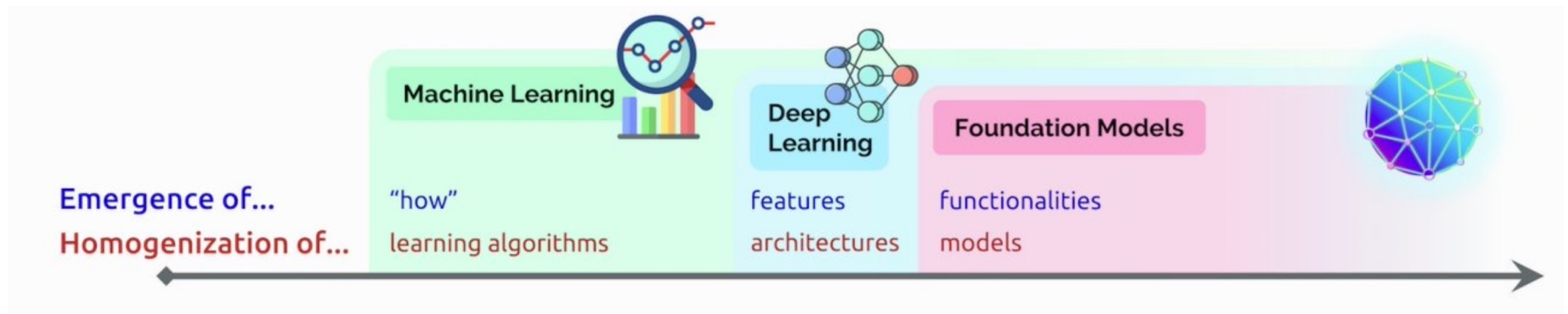
# SageMaker Studio Lab

- <https://studiolab.sagemaker.aws>
- **Free**
- Takes about a week to be approved for account
- Linked to GitHub  with lots of examples
- Community on Stack Overflow 



# Future *Foundation Models*

# Foundation models



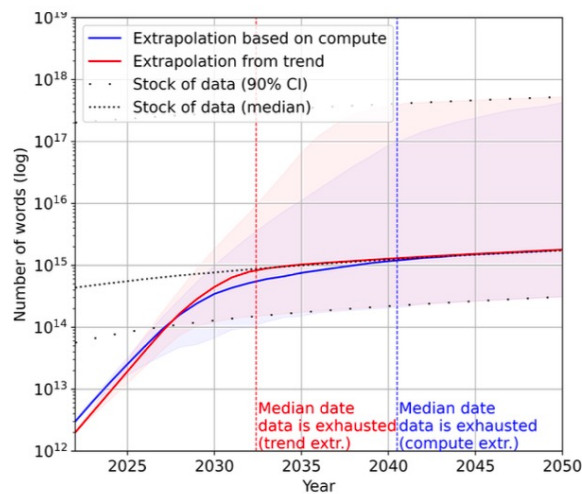
- Increased standardization of models:
  - Code Whisperer
  - GPT-3
  - Stable Diffusion
  - Chat GPT



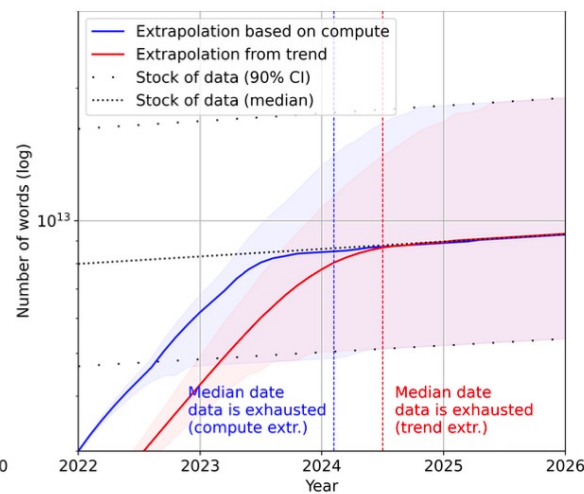
# Characteristics of Foundation Models

- Often trained “self-supervised”
- Predict portions of data from other portions with no explicit labels
  - Eg., fill in blanked out word in text, or fill in missing portion of image
  - Use rich data source (say most text written in history of humanity)
- Expensive, requiring millions \$ to train
- Made once, then reused by many without modification of any kind
- Interact by making a sentence where the only way to fill the blank is with answer you want:
  - Eg., “George Washington was born in the year \_\_\_\_”

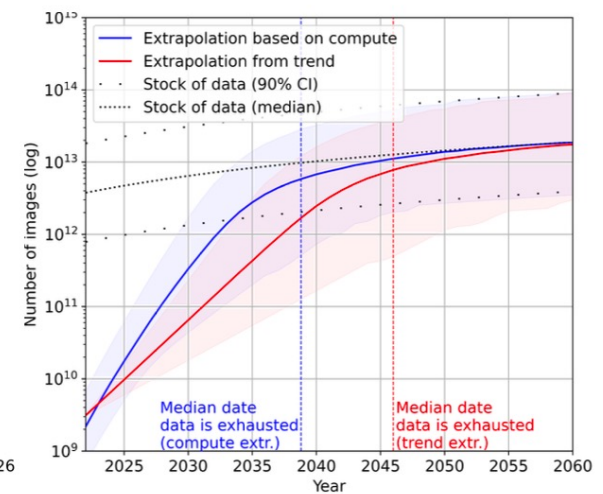
# Running out of data



(a) Projections for low-quality language data



(b) Projections for high-quality language data



(c) Projections for vision data

## Will we run out of data? An analysis of the limits of scaling datasets in Machine Learning

Pablo Villalobos\*, Jaime Sevilla<sup>†</sup>, Lennart Heim<sup>§</sup>, Tamay Besiroglu<sup>\*‡</sup>, Marius Hobbhahn <sup>\*¶</sup>, Anson Ho\*

# Explain-ability and Ethics

- How to demonstrate (prove) that a model is correct?
  - Why is model training so successful?
- How to demonstrate that a model is not biased?
- How to protect human beings?

# Important but not intellectually “elegant”

- CI/CD aspect of ML
  - In industry Git is one of the most important tools
  - Understanding the mathematical foundations is probably the least important
- Documentation has to be superb, and it seldom is
- It doesn't work for a long time ... , until it finally works a little bit
- Interpretation of data – what does 0.3 likelihood of coming to CI mean?
- Communications of methodology and findings – super important! Listen to customer, do not push your fav technology; what is business need?
- Politics of data:
  - No one wants to share their data, even within the same organization; negotiating for data and terms of usage (e.g., access) takes 50% of time of entire effort
  - Hard to reach agreement on “goodness” of data
  - Even harder to reach agreement on “conclusion” and how to craft policy based on the data

**AI**  
**SUPER-**  
**POWERS**  
**CHINA,**  
**SILICON VALLEY,**  
AND THE  
**NEW WORLD ORDER**  
**KAI-FU LEE**



Man has made his match ... now it's his problem



Skynet is a fictional neural network-based AI system that animates the Terminator