

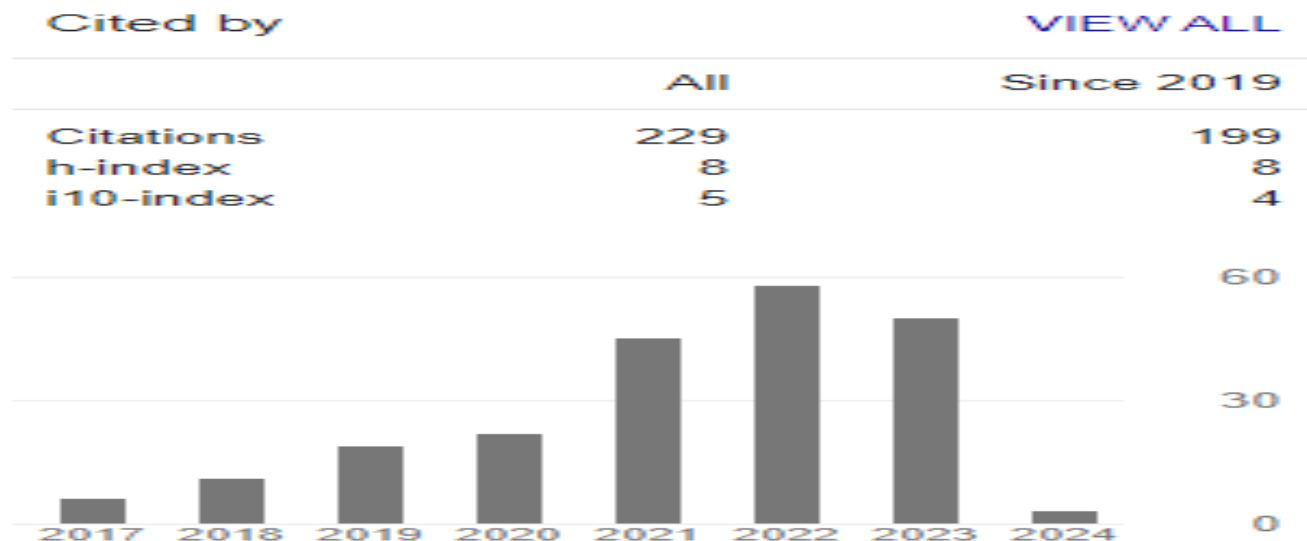
# Machine Learning



# About the Instructor

- Dr. Muhammad Safyan
  - PhD(2018) – AI & Knowledge Graph (Activity Recognition)
  - Research Output
    - Number of publications -28

## Google Scholar:



# About the Course

- About the Course
  - Machine Learning
  - Pre-requisite and background
  - Applications
  - Applied Side
- Course Outline
- Teaching Method
  - Assignment
- Research Focus
  - Term Paper discussion

# Potential Applications- OpenAI API

## Prompt Engineering

<https://platform.openai.com/examples>



### Grammar correction

Convert ungrammatical statements into standard English.



### Parse unstructured data

Create tables from unstructured text.



### Calculate time complexity

Find the time complexity of a function.



### Keywords

Extract keywords from a block of text.



### Python bug fixer

Find and fix bugs in source code.



### Tweet classifier

Detect sentiment in a tweet.



### Mood to color

Turn a text description into a color.



### Summarize for a 2nd grader

Simplify text to a level appropriate for a second-grade student.



### Emoji Translation

Translate regular text into emoji text.



### Explain code

Explain a complicated piece of code.



### Product name generator

Generate product names from a description and seed words.



### Spreadsheet creator

Create spreadsheets of various kinds of data.



### Airport code extractor

Extract airport codes from text.



### VR fitness idea generator

Generate ideas for fitness promoting virtual reality games.

# Image Tagging

## Results

### Tags

- sledding
- mushing
- snow
- snows
- treed

### Nearest Caption in the Training Dataset

a white dog is running fast on a trail covered by snow .

### Generated Captions

- a dog running down the road next to a hill .
- a small dog is running away from a hill covered in snow .
- a small dog is running down the hill .
- a dog standing on the hill runs down the road



# Image Tagging

## Results

### Tags

- feminist
- sleepwear
- pregnancy
- mommies
- lesbianism

### Nearest Caption in the Training Dataset

a little girl carries a younger girl on the sidewalk .

### Generated Captions

- a girl is standing outside with her hand on the ground .
- a little boy and young girl sit on the street .
- a girl in a blue shirt holds a small child on the sidewalk .
- a young girl is holding a cup on the



# News Clustering

Google News

news.google.com

Web Images Videos Maps News Shopping Gmail more

andrewyantakng@gmail.com | Web History | Settings | Sign out

Google news

Search News Search the Web

Advanced news search


U.S. edition Add a section

## Top Stories

Deepwater Horizon  
Fed meeting  
Foreign exchange market  
Lindsay Lohan  
IBM  
Tom Brady  
Toronto International Film Festival  
Paris Hilton  
Iran  
Hurricane Igor

### White House official denies Tea Party-focused ad campaign


Christine O'Donnell »  
CNN International - Ed Henry - 1 hour ago  
Democratic sources say the White House is not considering an ad campaign tying Republicans to the Tea Party. Washington (CNN) -- A top White House official sharply denied a report that claims President Obama's political advisers are weighing a national ...  
Tea Party is misplacing the blame, former President Bill Clinton claims  
New York Daily News  
GOP tea party backer defends Christine O'Donnell The Associated Press  
Atlanta Journal Constitution - Politics Daily - MyFox Washington DC - Salon all 726 news articles »



CNN Interna...

### US Stocks Climb After Recession Called Over, Homebuilders Gain

MarketWatch - Kristina Peterson - 16 minutes ago  
NEW YORK (MarketWatch) -- US stocks climbed Monday, gaining speed after a key nonprofit organization officially called the recession over, giving investors a boost of confidence in the gradual economic recovery.  
Longest recession since 1930s ended in June 2009, group says  
Los Angeles Times  
Downturn Was Longest in Decades, Panel Confirms New York Times  
Wall Street Journal - AFP - CNN - USA Today all 276 news articles »



MyFox Phila...

## Recent

Recession officially ended in June 2009  
CNMoney - Chris Isidore - 39 minutes ago

Hurricane Igor lashes Bermuda  
USA Today - Gerry Broome - 5 minutes ago

'Explain what you want from us.' reads front-page editorial  
msnbc.com - Olivia Torres - 10 minutes ago

Crisis response: Pakistan floods

San Francisco Bay Area - Edit

Clorox »  
Bay Biz Buzz: Clorox close to selling STP, Armor All  
San Jose Mercury News - 48 minutes ago - all 24 articles »

Google's official beekeeper keeps the company buzzing with excitement  
San Jose Mercury News - Bruce Newman - 1 hour ago

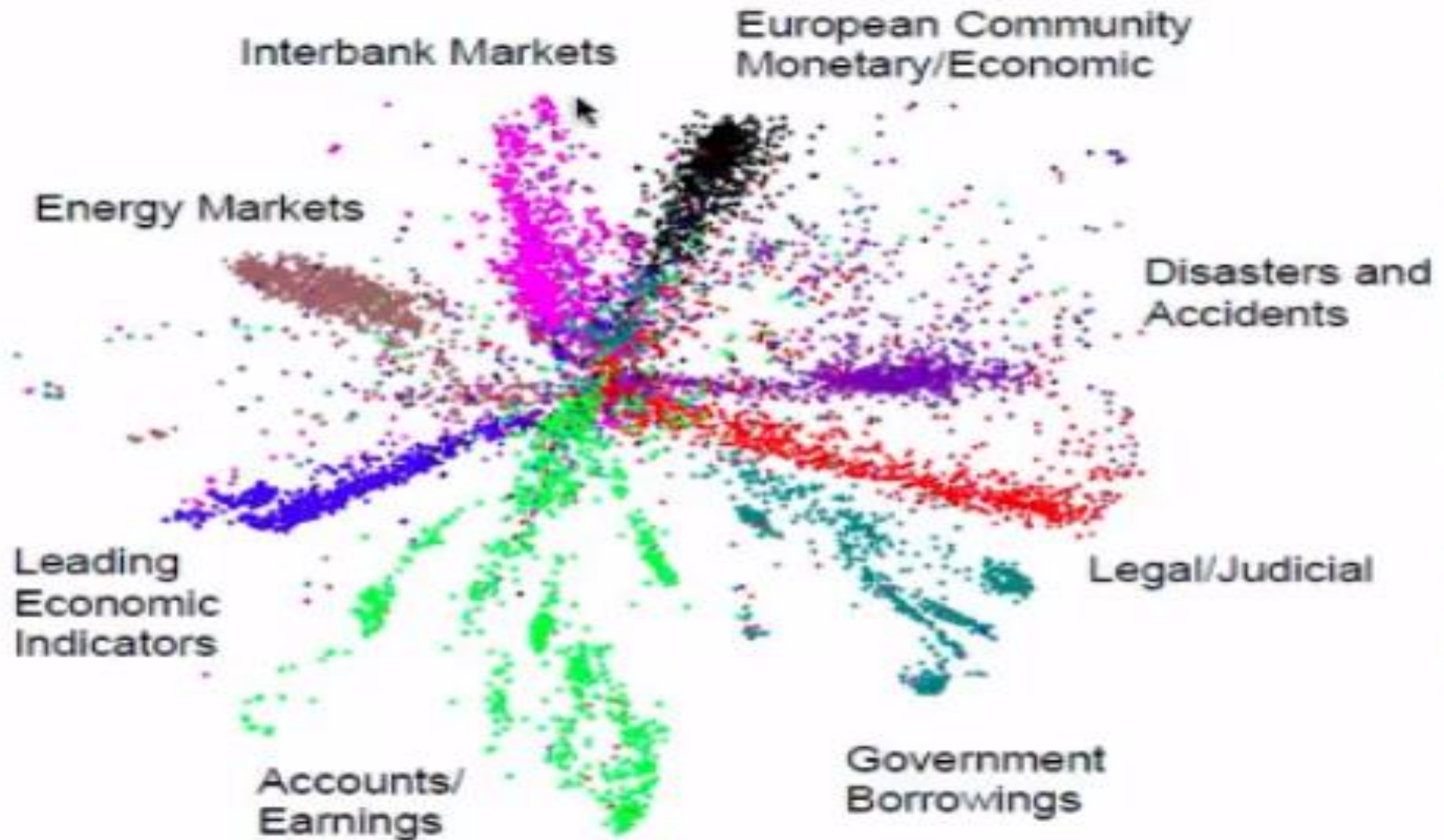
Jon Sylvia »

## Starred

San Francisco Bay Area  
World  
U.S.  
Business  
Sci/Tech  
More Top Stories

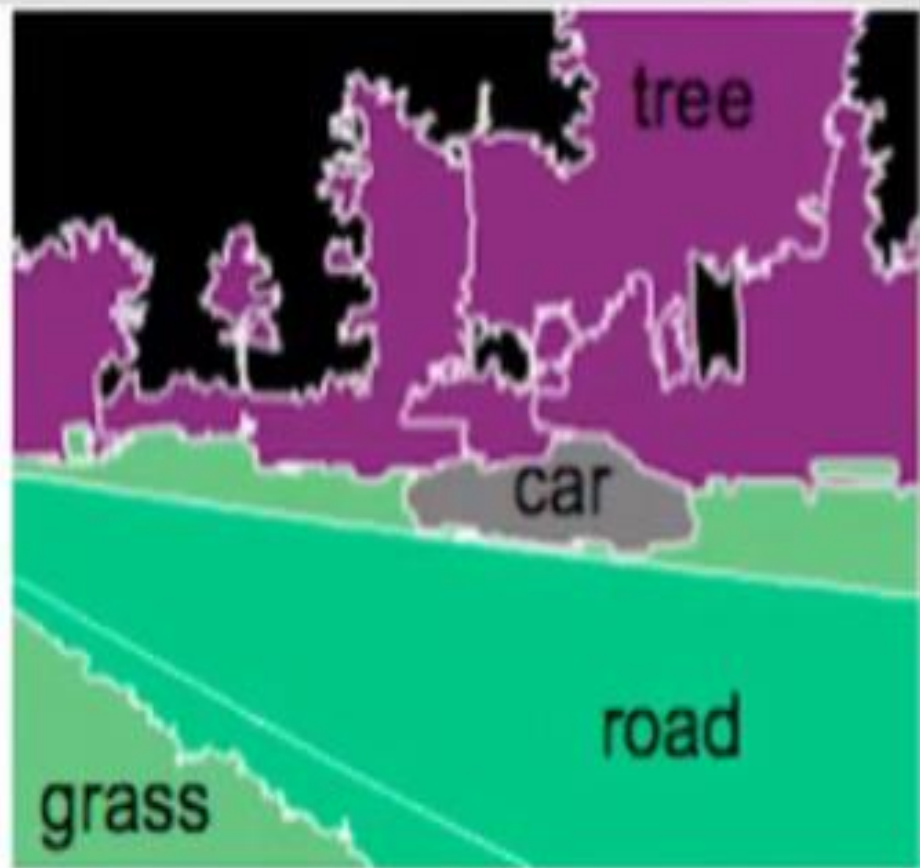


# Market Segmentation

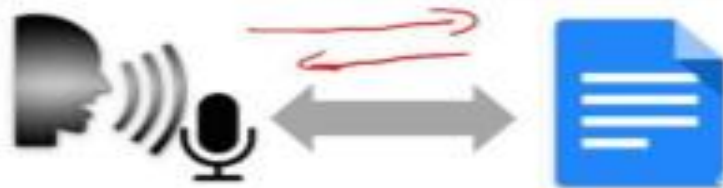




# Image Detection and Localization



# Daily used Applications



# Daily used Applications





# Daily used Applications



# Applications



# Historical Background

## **Early Foundations (1940s - 1950s):**

Warren McCulloch and Walter Pitts: first neural networks, artificial neurons inspired by biological brains

## **AI Winter (1970s - 1980s):**

faced significant challenges and a period known as the "AI Winter"  
lack of funding and limited practical success

## **Reemergence (1990s):**

resurgence in the 1990s, driven by advances in computing power and algorithmic improvements.

Decision trees, support vector machines

## **Rise of Deep Learning (2000s - Present):**

deep neural networks, particularly deep convolutional neural networks (CNNs) for image recognition.

better optimization algorithms

availability of large datasets, like ImageNet,

increased computing power (GPUs)



# Historical Background

## **Deep Learning's Dominance (2010s - Present):**

dominant approach for image and speech recognition, natural language processing, and reinforcement learning  
recurrent neural networks (RNNs), long short-term memory networks (LSTMs), and Transformers  
Frameworks like TensorFlow and PyTorch

## **Applications and Real-World Impact:**

healthcare (diagnosis and drug discovery), autonomous vehicles, finance (algorithmic trading), and more.  
Natural language processing models like GPT-3 and BERT

## **Future Directions (Beyond 2021):**

evolve, with ongoing research into improving the efficiency, robustness, and generalization capabilities of deep learning models.

Exploring quantum computing's potential for machine learning and addressing ethical AI issues will be critical in the coming years.

# Machine Learning Guru's

<b>Geoffrey Hinton:</b>	<p>affiliated with Google Brain as a Chief Scientific Advisor and a Professor Emeritus at the University of Toronto.</p> <p>Godfather of Deep Learning," Geoffrey Hinton made significant contributions to neural networks, backpropagation, and deep learning architectures, including Boltzmann Machines and the development of convolutional neural networks (CNNs).</p>
<b>Yann LeCun:</b>	<p>renowned for his work in convolutional neural networks (CNNs) and is a pioneer in the field of computer vision. He's also known for his contributions to the development of the backpropagation algorithm.</p> <p>Yann LeCun was a Chief AI Scientist at Facebook AI and a Professor at New York University (NYU).</p>
<b>Andrew Ng:</b>	<p>Professor at Stanford University and the co-founder of deeplearning.ai, an online learning platform.</p> <p>Andrew Ng is a co-founder of Google Brain and is known for his work on online education, including the creation of the popular machine learning course on Coursera. He has made significant contributions to the field of deep learning.</p>

# Machine Learning Guru's

<b>Fei-Fei Li:</b>	is known for her work in computer vision and co-authored the ImageNet dataset, which played a pivotal role in the development of deep learning for image
<b>Richard Sutton:</b>	Richard Sutton is a prominent figure in reinforcement learning. His work includes the development of the temporal difference learning algorithm and his
<b>Ian Goodfellow:</b>	Ian Goodfellow is the author of the widely used textbook "Deep Learning." He's known for his work on generative adversarial networks (GANs) and has made significant contributions to deep learning research.

# Programming Languages

## **Python:**

most popular programming language for machine learning and deep learning due to its extensive libraries (NumPy, SciPy, pandas) and the availability of frameworks like TensorFlow and PyTorch.

## **R:**

R often used in statistical modeling and data analysis. It has a rich ecosystem of packages for machine learning

## **Julia:**

an emerging programming language known for its speed and is increasingly being used in scientific computing and machine learning.

# Programming Languages

## **JavaScript:**

JavaScript is used for implementing machine learning models in web applications and browser-based AI.

## **MATLAB:**

MATLAB is used in academic and research settings for machine learning and signal processing applications.

# Frameworks/API's for Machine Learning

## **TensorFlow:**

Developed by Google, TensorFlow is one of the most widely used deep learning frameworks. It offers a high level of flexibility and supports both CPU and GPU computations.

## **PyTorch:**

PyTorch, developed by Facebook's AI Research lab (FAIR), is known for its dynamic computational graph and is favored by researchers for its ease of use and flexibility.

## **Keras:**

Keras is a high-level neural networks API that can run on top of TensorFlow, Theano, or Microsoft Cognitive Toolkit (CNTK). It's known for its user-friendly interface.

## **Scikit-learn:**

Scikit-learn is a popular machine learning library for Python, providing tools for classical machine learning algorithms and data preprocessing.



# Frameworks/API's for Machine Learning

## **Caffe:**

Caffe is a deep learning framework developed by the Berkeley Vision and Learning Center. It's known for its speed and efficiency in image classification tasks.

## **MXNet:**

MXNet is an open-source deep learning framework known for its scalability and support for multiple programming languages, including Python and Julia.

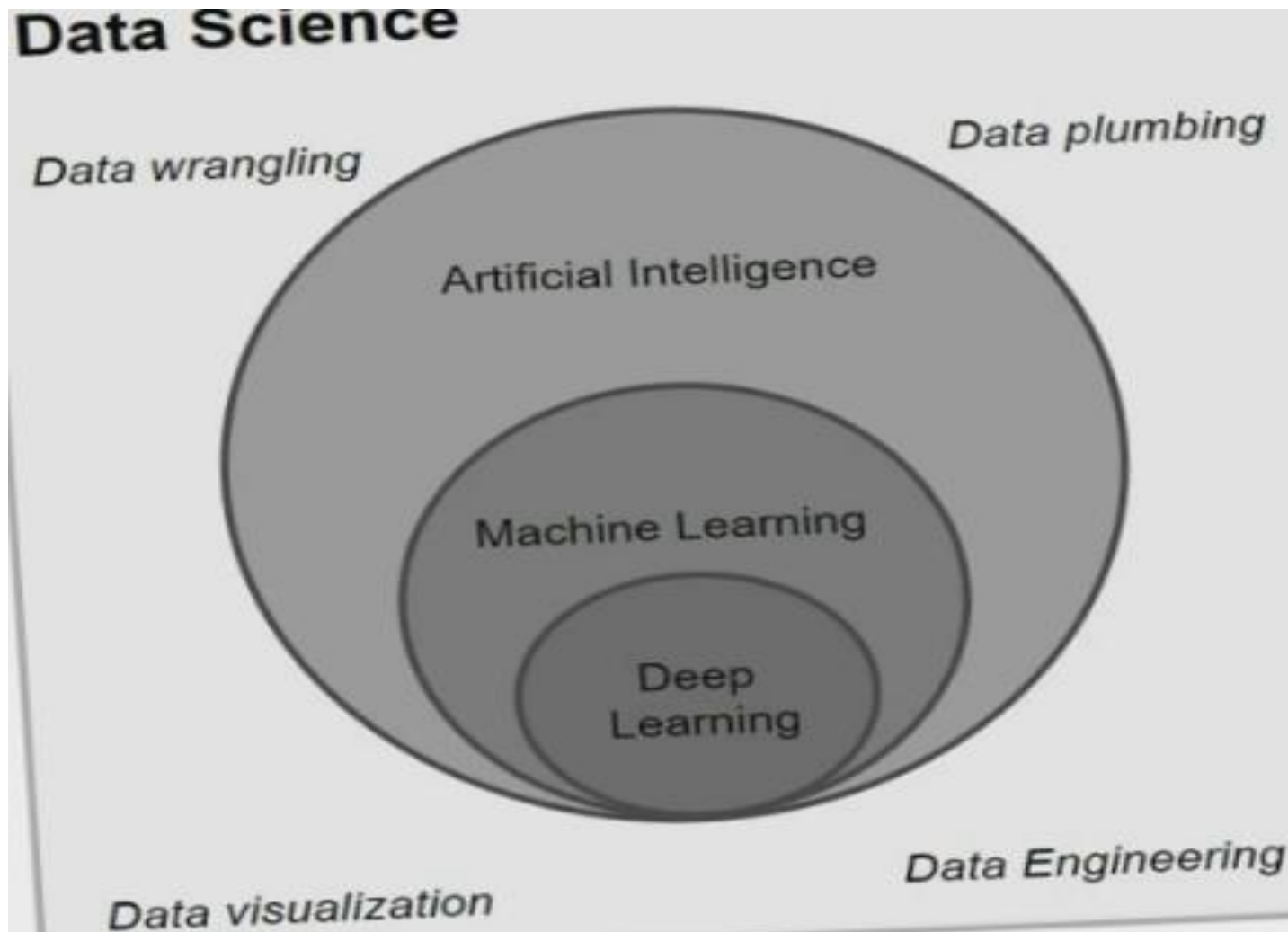
## **Theano:**

Although its development has slowed down, Theano was influential in the early days of deep learning and served as a precursor to other frameworks.

## **Caffe2:**

Developed by Facebook, Caffe2 was designed for mobile and large-scale distributed training and has been integrated into PyTorch as part of the PyTorch 1.0 release.

# Demystifying DS, AI, ML and DL



## **Artificial Intelligence**

A branch of computer science that deals with providing intelligence to machines artificially; and hence the name. They could be simple rule based systems, knowledge driven databases etc.

## **Machine Learning**

A subfield within AI that deals with making machines intelligent without explicit programming. It is in fact the most successful branch of AI

## **Deep Learning**

A subfield of machine learning that deals with making machine intelligent and are concerned with algorithms inspired by the structure and function of the brain called artificial neural networks.

## **Data Science**

It deals with the entire journey of extracting, cleaning, transforming, visualizing, mining and developing AI/ML products from data.

# Machine Learning Definition

- Philosophical learning of meaning
- Tom Michel Definition ML as
  - A computer is said to be Learn from **experience E** with respect to some **task T** and some performance **measure P**, if its performance on T, as measured by P, improve with **Experience E**.

# Notation

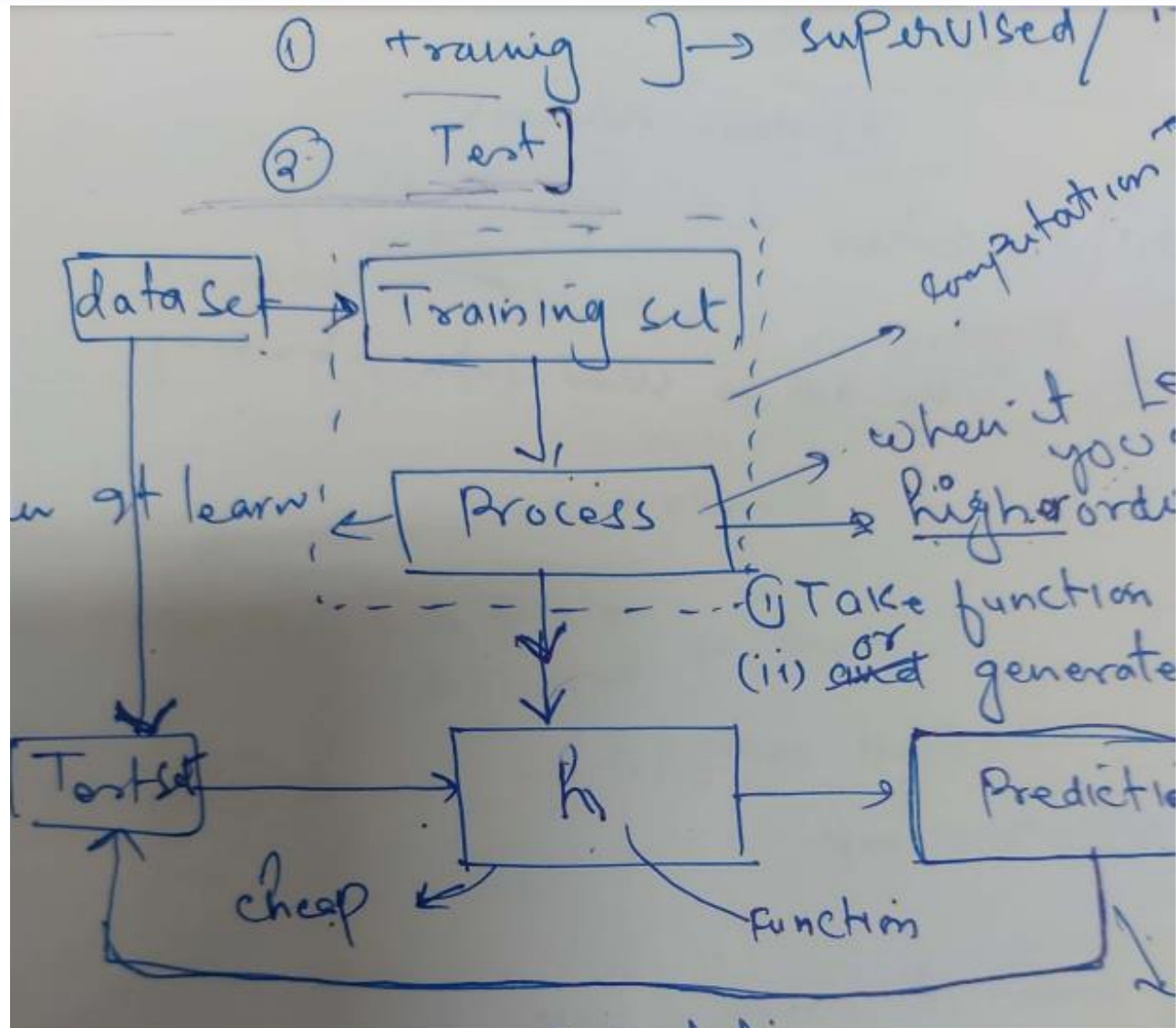
- Variable
- Random Variable
- Vector

# Machine Learning Models

- Supervised Learning
  - Like Teach , Supervisor
  - Human knows the ground truth
    - Regression
    - Classification
- Unsupervised Learning
  - Clustering
  - Auto-encoder
  - Dimensionality Reduction
- Reinforcement Learning

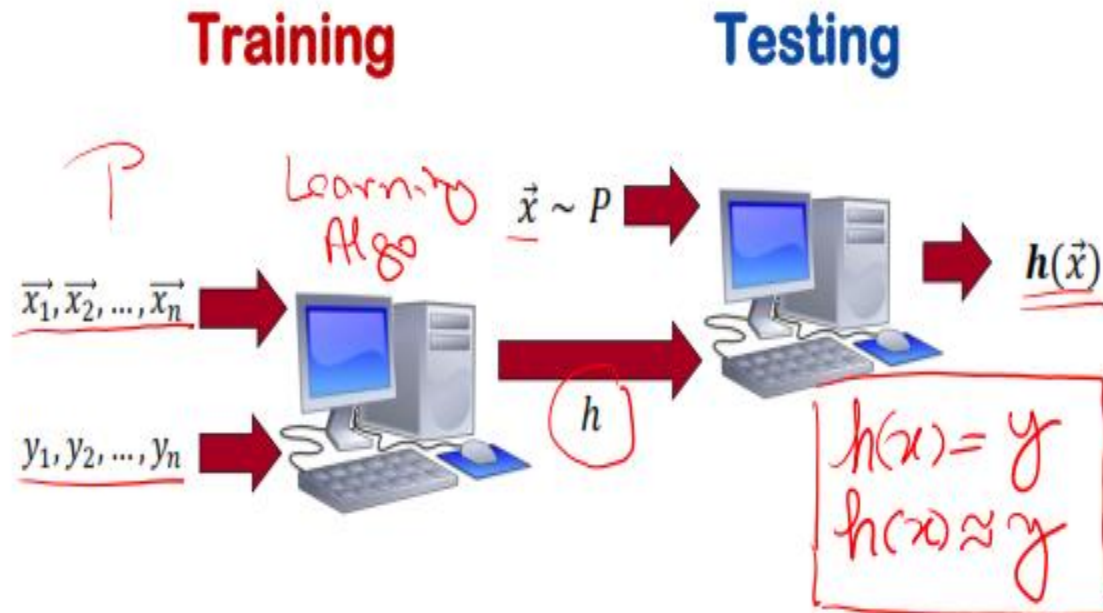


# ML Model



# Machine Learning Models

## Training and testing



# ML Model

.....

Features					Label
#	Height (inches)	Weight (kgs)	B.P. Sys	B.P. Dia	Heart disease
1	62	70	120	80	No
2	72	90	110	70	No
3	74	80	130	70	No
4	65	120	150	90	Yes
5	67	100	140	85	Yes
6	64	110	130	90	No
7	69	150	170	100	Yes
8	66	125	145	90	?
9	74	67	110	60	?

Feature vector (4-dimensional)

Label vector

Training Data

Test Data

# Machine Learning Models

- Rules Vs. Learning
- Supervised Learning
  - DATA(features, Labels, Train, Test, Dimension)
  - Feature Space
  - Label Space
  - Hypothesis space

# Model Evaluation

## – Loss functions

- Calculate the average error of  $h$  in predicting  $y$ .
- Smaller is better
- 0 loss: No error
- 100% loss: Could not even get one instance right
- 50% loss: Your  $h$  is as informative as a coin toss

# Model Evaluation

## 0/1 Loss

$$\underline{L_{0/1}(h)} = \frac{1}{n} \sum_{i=1}^n \underline{\delta_{h(x_i) \neq y_i}}, \text{ where } \underline{\delta_{h(x_i) \neq y_i}} = \begin{cases} \underline{1}, & \text{if } h(x_i) \neq y_i \\ \underline{0}, & \text{otherwise} \end{cases}$$

## Squared loss

$$L_{sq}(h) = \frac{1}{n} \sum_{i=1}^n (\underline{h(x_i)} - \underline{y_i})^2$$


## Absolute loss

$$L_{abs}(h) = \frac{1}{n} \sum_{i=1}^n \underline{|h(x_i) - y_i|}$$



# Linear Regression

## Example:



Area in Square Feet	Price
100	100,000
150	140,000
200	250,000
...	...

## Notation:

$x$ 's = "input" variable / features

$m$  = Number of training examples

$y$ 's = "output" variable / "target" variable

# Plot/Shape-up Data

**Example:**

