Intro to AI-ML

Yogesh Kulkarni

ISACA
Pune Chapter 25+ YEARS

ITS
ISACA Pune Chapter Conference
2025

Gold Sponsor
ShieldByte
Infosec

Introduction to Artificial Intelligence

We have a Problem!!

- Along with some softer words like "disruption", "passionate", "excited"...
- If you don't have word "innovation" in your talk/speech/conversation it's a BIG problem
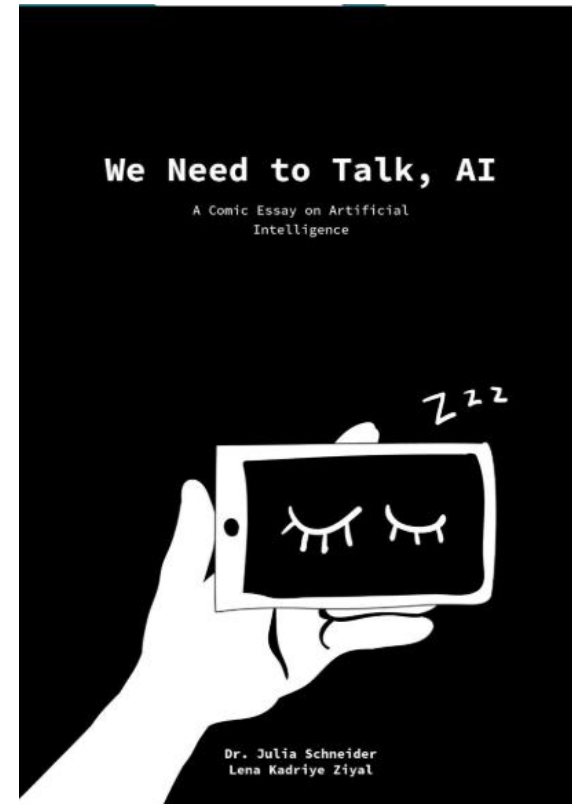- Irrespective of fields. You can be Corporate, Political, Social, etc.

And there is an addition of one more word, which is a must in every talk...and that is?

The Problem is

Every company is claiming to be working in AI-ML

- Is it really so?
- What exactly is AI (ML)?
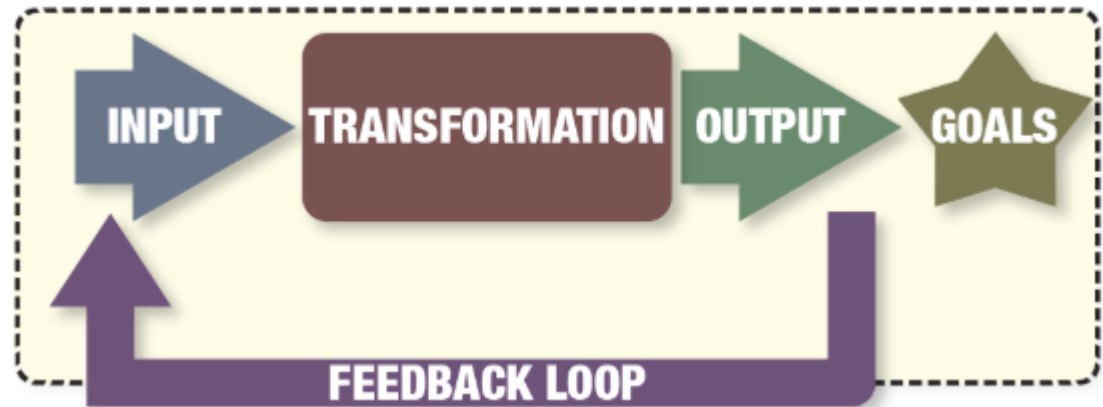- What is not AI?

Or is it just a plain BIG hype?

What is the Core Idea?

- Problem solving?
- Writing software algorithms?
- Solving research problems?

Desire

- To find a "function"
- To find a relation
- To find a transformation
- To build a model
- From given inputs to desired outputs.

That's it.

ISACA
Pune Chapter 25+ YEARS

ITS
ISACA Pune Chapter Conference
2025

Gold Sponsor
ShieldByte
Infosec

Simple Functions

- Some functions are straight forward
- "In summer, ice-cream sale goes up"
- Cause and effect
- Relation (function, Mathematical model) is found out
- Here, simple rule based programming suffices

ISACA
Pune Chapter 25+ YEARS

ITS
ISACA Pune Chapter Conference
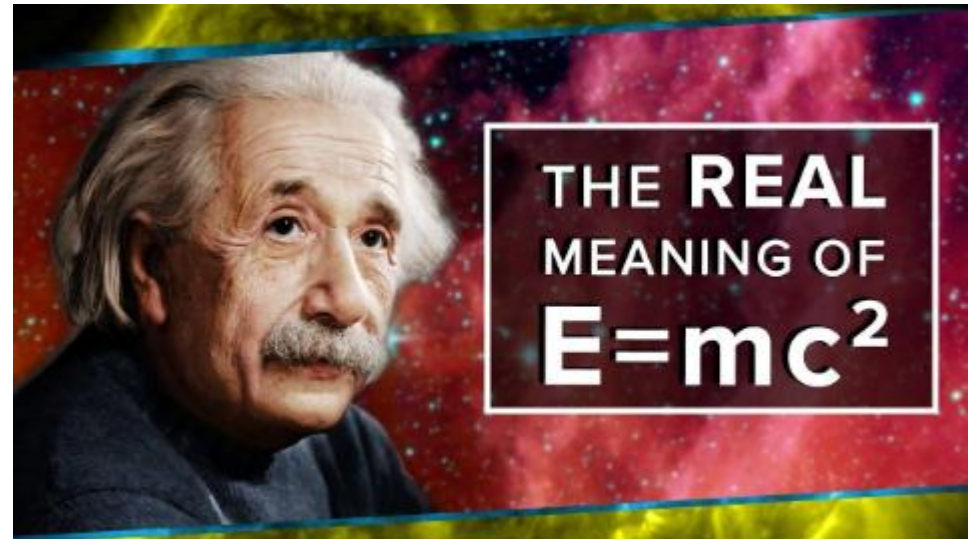2025

Gold Sponsor
ShieldByte
Infosec

Complex Functions

- But some functions are complex
- "More you put efforts, your business flourishes."
- Cause and effect again, but the relation is far too complex
- Too many variables
- Here, simple rule based programming not humanly possible
- Lots of research needed to come up with equations
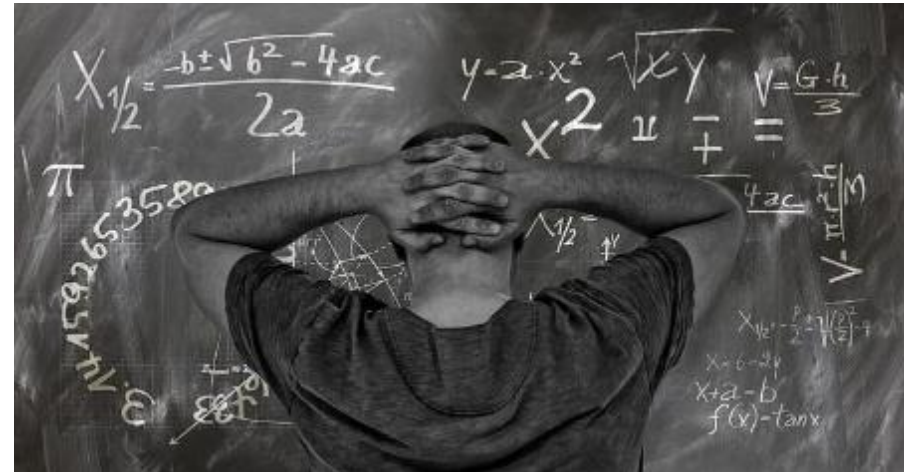
Example: $E = mc^2$

- What's this? a function?
- Input variable(s)?
- Output variable(s)?
- Parameters?
- How's the relation? linear?

THE REAL MEANING OF $E=mc^2$

ISACA. Pune Chapter 25+ YEARS

ITS
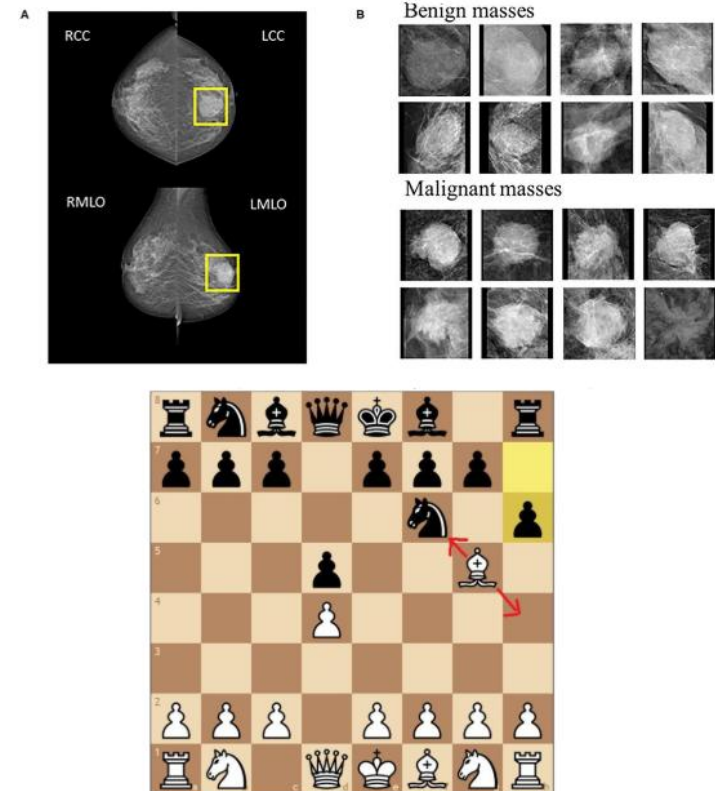ISACA Pune Chapter Conference 2025

Gold Sponsor
ShieldByte
Infosec

Controversial Example

- Even astrology is a model, based on the past cases
- Could claim empirical evidence
- Given this planetary position, it predicts
- Represented by "Horoscope"
- Got weights for each planets (real or fictitious)
- Reliable??
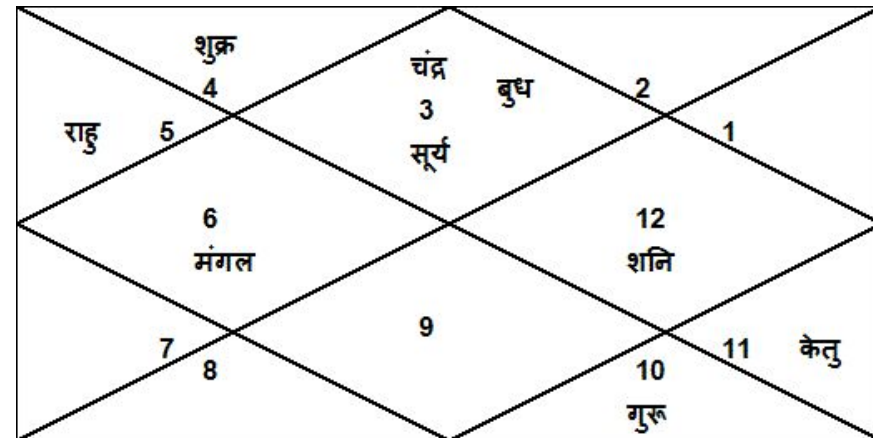
Real-life Functions

- But most real-life functions are not deterministic
- Some are probabilistic, some non-linear
- "Detecting if the tumor is benign or malignant"
- "At any state in the game of chess, what's the next move?"

![ISACA Pune Chapter 25+ YEARS]

ITS
ISACA Pune Chapter Conference
2025

Gold Sponsor
ShieldByte
Infosec

Chess: Next Move?

- Needs extreme expertise
- Needs "intelligence"
- How do you get that?
  a. Built by lots of training
  b. By studying lots of past games
- This is how Humans build intelligence

ISACA.
Pune Chapter 25+ YEARS

ITS
ISACA Pune Chapter Conference
2025

Gold Sponsor
ShieldByte
Infosec

Intelligence

- Can machine (software/program) also do the same?
- Can it play chess?
- Can it build intelligence?
- By looking at past experiences (data)
- Training Data: games played, moves used, etc.

Yes, it can!! That's Artificial Intelligence.

What is Artificial Intelligence?

My definition:

"If machines (or computer programs) start doing some/all of these 'intelligent' tasks, then that's Artificial Intelligence"

ISACA.
Pune Chapter 25+ YEARS

ITS
ISACA Pune Chapter Conference
2025

Gold Sponsor
ShieldByte
Infosec

Intelligence: the differentiation

- Ability to think various domains
- Ability produce something new
- Ability to detect the unseen
- Ability to enhance knowledge (rules, patterns)

All these, AI has started doing. The AI era has arrived!!

ISACA.
Pune Chapter 25+ YEARS

ITS
ISACA Pune Chapter Conference
2025

Gold Sponsor
ShieldByte
Infosec

Everyday Usage

Artificial intelligence seems to have become ubiquitous.

- Replying to our emails on Gmail
- Learning how to drive our cars
- Sorting our holiday photos
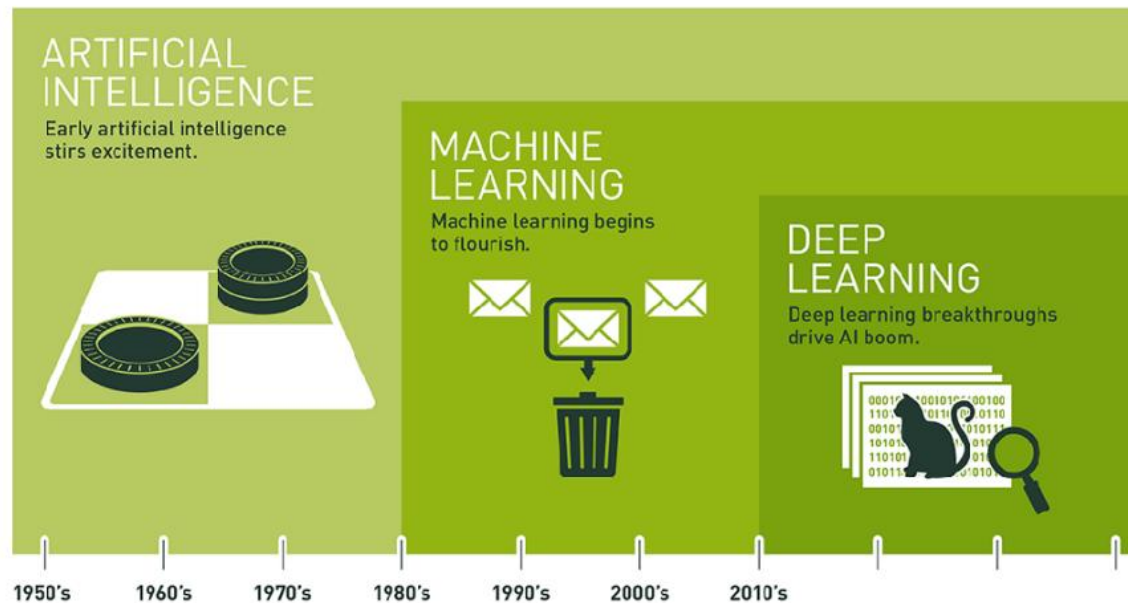- etc.

Too good to be true, isn't it, sort of Magical!!

ISACA Pune Chapter 25+ YEARS

ITS
ISACA Pune Chapter Conference
2025

Gold Sponsor
ShieldByte
Infosec

But then...

- When it's too good, you start suspecting
- Is it for real!!
- How can such thing happen?
- How far will it go?

The next thing you know, people are worrying about exactly how and when AI is going to doom humanity.
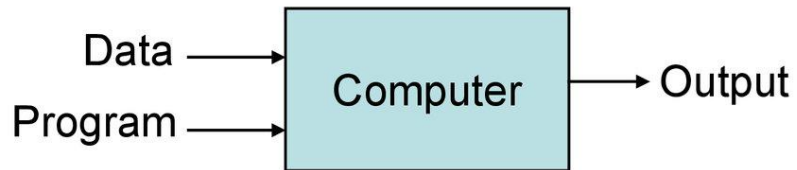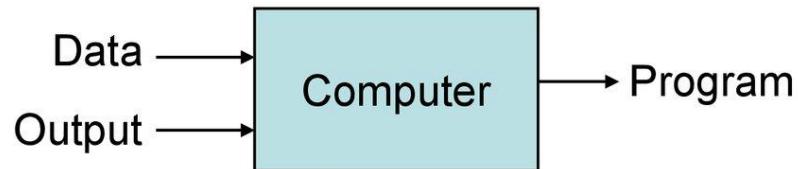
AI, ML, DL... Same?

A Paradigm Shift

ISACA
Pune Chapter 25+ YEARS

ITS
ISACA Pune Chapter Conference
2025

Gold Sponsor
ShieldByte
Infosec

Why Machine/Deep Learning?

- Problems with High Dimensionality
- Hard/Expensive to program manually
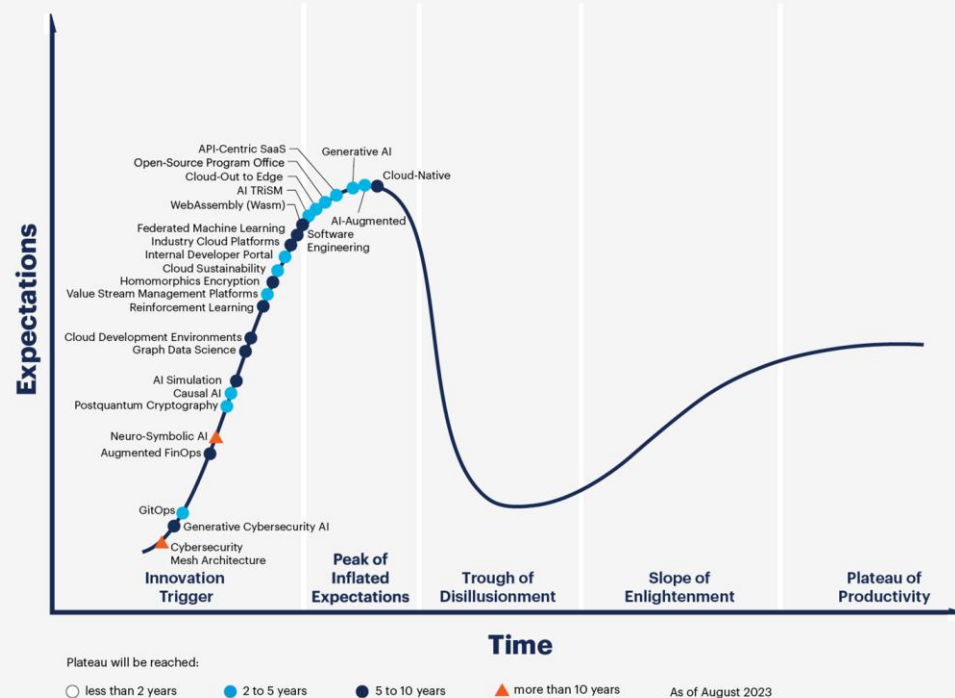- Techniques to model 'ANY' function given 'ENOUGH' data
- Job $$$

ISACA. Pune Chapter 25+ YEARS

ITS
ISACA Pune Chapter Conference
2025

Gold Sponsor
ShieldByte
Infosec

Why now?

- Flood of data (Internet, IoT)
- Increasing computational power
- Easy/free availability of algorithms
- Increasing support from industries

Hype Cycle for Emerging Technologies, 2023

ISACA. Pune Chapter 25+ YEARS

ITS
ISACA Pune Chapter Conference
2025

Gold Sponsor
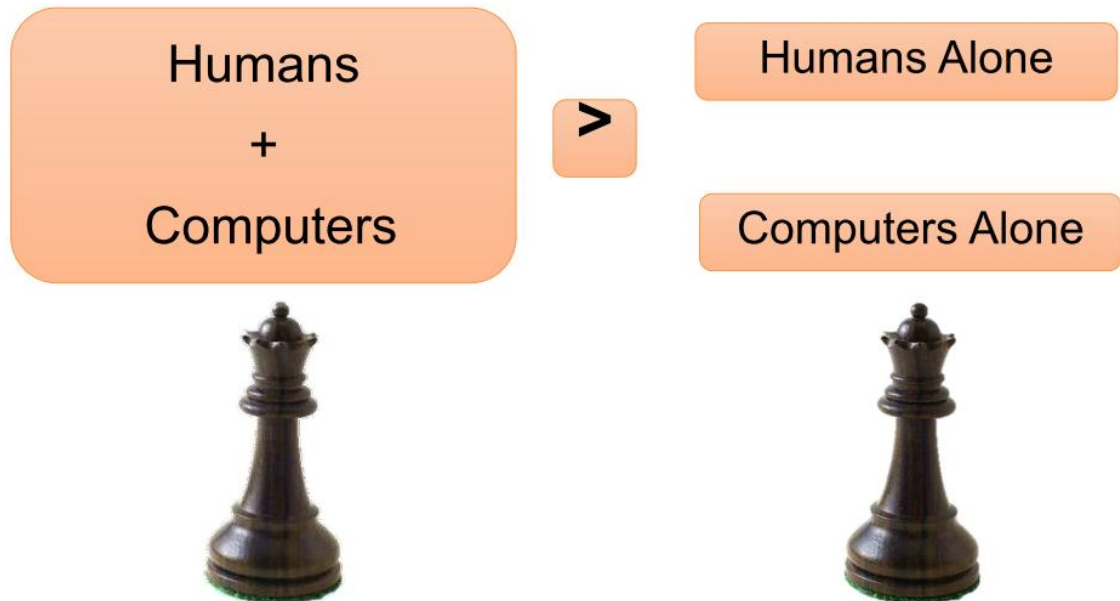ShieldByte
Infosec

Is AI a threat?

- If you believe in what Elon Musk says, then YES.
- If you believe in these movies, then YES.
- Well, AI based War robots are not impossible anymore.

Elon Musk recently commented on Twitter that artificial intelligence (AI) is more dangerous than North Korea

Fear: Are we being replaced?

- Yes, in tasks that are repetitive, But not which require complex thinking and creativity
- Mostly, Technology Enhancing (Not Replacing) Humans

Humans
+
Computers

>

Humans Alone

Computers Alone

Limits on Artificial Intelligence

- Many things still beyond the realm of AI
- No thinking computers
- No Abstract Reasoning
- Often AI systems Have Accuracy Limits
- Many things difficult to capture in data
- Sometimes Hard to interpret Systems

Introduction to Machine Learning

How do we learn?

- What do we do when we have to prepare for an examination?
- Study. Learn. Imbibe. Take notes. Practice mock papers.
- Thus, prepare for the unseen test.

What is Learning?

"Learning is any process by which a system improves performance from experience." - Herbert Simon, Turing Award 1975, Nobel in Economics 1978.

ISACA
Pune Chapter 25+ YEARS

ITS
ISACA Pune Chapter Conference
2025

Gold Sponsor
ShieldByte
Infosec

What is Machine Learning?

Machine learning is a type of artificial intelligence (AI) which:

- Learns function without being explicitly programmed
- Can grow and change when exposed to new data

So, What is Machine Learning?

- Ability of computers to "learn" from "data"
- Learn: Discover patterns, underlying structure
- Data: Comes from sensors, transactions, etc.

Mathematical Definition of Machine Learning

Machine Learning comes up with a Model given inputs and targets.

- Input data is available
- Input data is transformed to get output
- Output: something that needs to be predicted or estimated
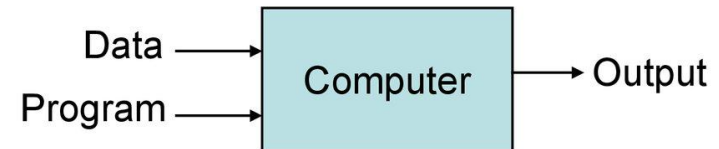- Transformation engine is called Model or function

Model entities

For income = $c + \beta_0 \times$ education $+ \beta_1 \times$ experience

- Inputs: Education and experience, also called as features or attributes or dimensions or variables
- Mathematical entities added to input data, are Parameters. β_0 and β_1 are parameters
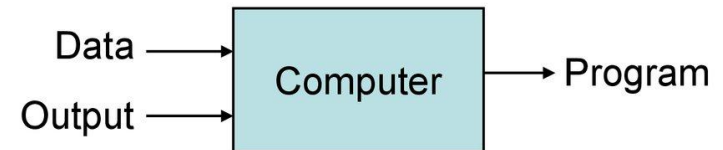- Income is target, also called as outcome or class

Traditional vs. Machine Learning?

- Traditional Programming:
  a. Explicit rules and logic coded by programmers
  b. Fixed decision-making process
- Machine Learning:
  a. System learns patterns from data
  b. Adaptive decision-making based on experience

Traditional Programming

Data → Computer → Output
Program → Computer

Machine Learning

Data → Computer → Program
Output → Computer

Why Machine Learning?

- Problems with High Dimensionality
- Hard/Expensive to program manually
- Techniques to model 'ANY' function given 'ENOUGH' data
- Job $$$

Why now?

- Flood of data (Internet, IoT)
- Increasing computational power
- Easy/free availability of algorithms
- Increasing support from industries

The storm: The Big Data is coming

- In 2012, HBR put Data Scientists on the radar
- "The Sexiest Job of the 21st Century"
- Industry, trying to be data-driven, than manual

(Big) Data Characteristics

Volume	Velocity	Variety	Veracity*
Data at Rest	**Data in Motion**	**Data in Many Forms**	**Data in Doubt**
Terabytes to exabytes of existing data to process	Streaming data, milliseconds to seconds to respond	Structured, unstructured, text, multimedia	Uncertainty due to data inconsistency & incompleteness, ambiguities, latency, deception, model approximations

(Image Credit: http://www.rosebt.com/blog/data-veracity)

What's the answer?

AI-ML-DL

- Machines showing intelligence of Humans
- Machine Learning: part of AI
- Logic is not programmed by hand
- Gets emerged in training with data

ISACA.
Pune Chapter 25+ YEARS

ITS
ISACA Pune Chapter Conference
2025

Gold Sponsor
ShieldByte
Infosec

A Puzzle

Math Quiz #1 - Teacher's Answer Key

1) 2 4 5 = 3   5) 6 2 2 = 10

2) 5 2 8 = 2   6) 3 1 1 = 2

3) 2 2 1 = 3   7) 5 3 4 = 11

4) 4 2 2 = 6   8) 1 8 1 = 7

How different is Machine Learning?

Maths Puzzle: Given a pattern in numbers, can you find the rule?

Maths Puzzle

- Letting the computer work out that relationship for you
- 'Learn' to solve such problems
- 'Test' with any other problem of the same type!

Types of Machine Learning
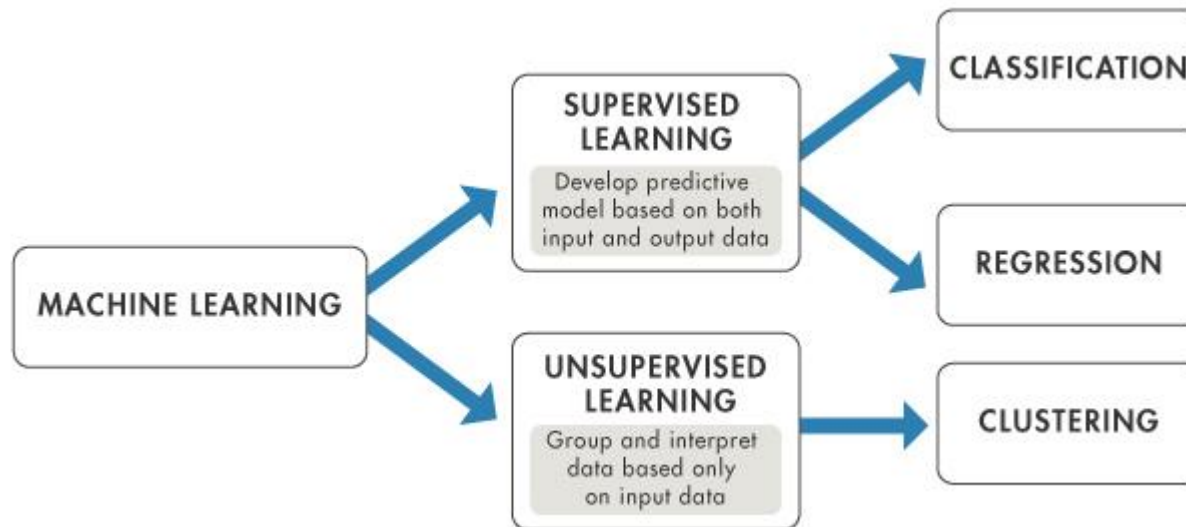
Some types of algorithms

- Prediction: predicting a continuous variable from data
- Classification: assigning records to predefined groups
- Clustering: splitting records into groups based on similarity
- Association learning: seeing what often appears together

Machine Learning Algorithms Categories

- Supervised Learning:
  a. Classification
  b. Regression
- Unsupervised Learning:
  a. Clustering
  b. Dimensionality Reduction

Questions Data Science Answers

- Is this A or B? → Classification algorithms
- Is this weird? → Anomaly detection algorithms
- How much—or—How many? → Regression algorithms
- How is this organized? → Clustering algorithms, Dimensionality reduction
- What should I do next? → Reinforcement learning algorithms

Classification

- Description: Identifying the category an object belongs to
- Applications: Spam detection, Image recognition
- Algorithms: SVM, nearest neighbors, random forest, Logistic Regression

Regression

- Description: Predicting a continuous-valued attribute associated with an object
- Applications: Drug response, Stock prices
- Algorithms: Linear Regression

Clustering

- Description: Automatic grouping of similar objects into sets
- Applications: Customer segmentation, Grouping experiment outcomes
- Algorithms: k-Means

Dimensionality Reduction

- Description: Reducing the number of random variables to consider
- Applications: Visualization, Increased efficiency
- Algorithms: PCA, Singular Value Decomposition

Popular Algorithms in Machine Learning

- Linear, Logistic Regression
- Decision Trees
- SVM - Support Vector Machines, Naive Bayes
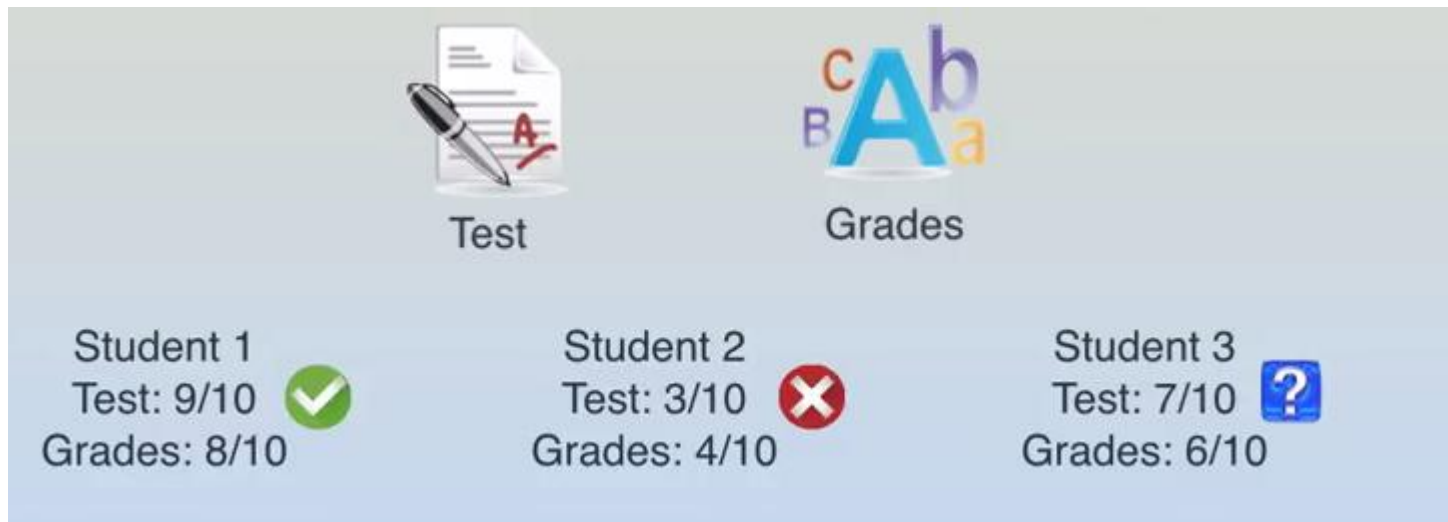- K-Means

Recommendation

Is Gender the decider or Age?

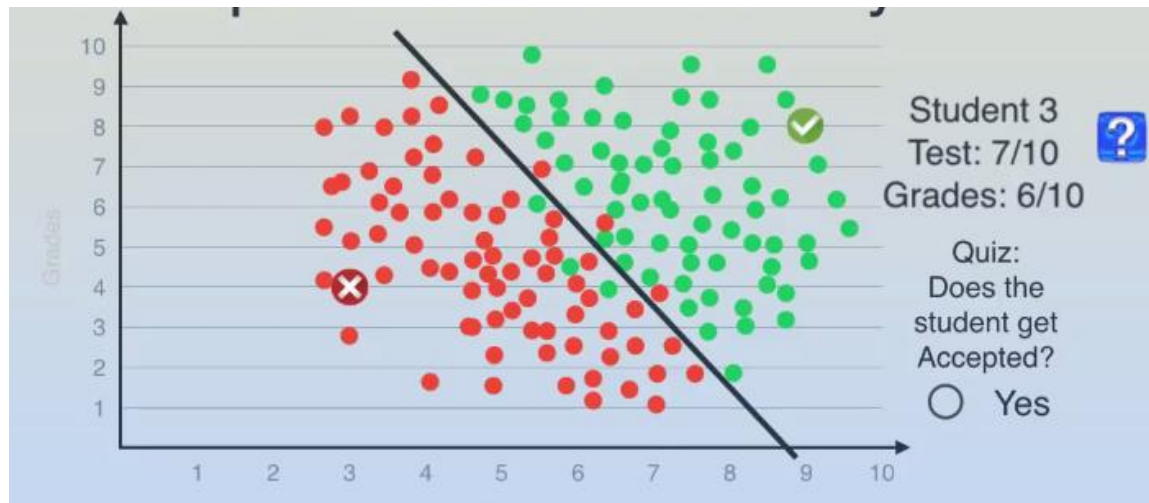(Image Credit: A Gentle Introduction To Machine Learning; SciPy 2013 Presentation - Kastner, Kyle)

Classification

Will Student 3 get accepted?

(Image Credit: A Gentle Introduction To Machine Learning; SciPy 2013 Presentation - Kastner, Kyle)

ISACA Pune Chapter 25+ YEARS

ITS
ISACA Pune Chapter Conference
2025

Gold Sponsor
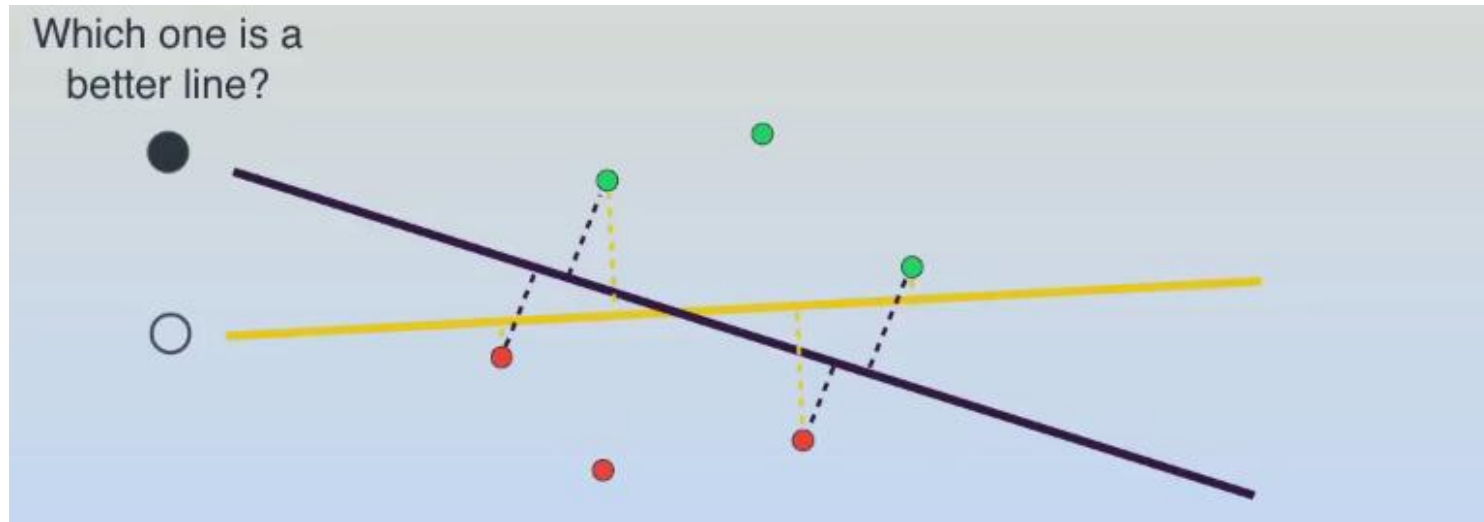ShieldByte
Infosec

Classification

After separating line, Student 3's fate can be predicted. That's Logistic
Regression.

(Image Credit: A Gentle Introduction To Machine Learning; SciPy 2013 Presentation - Kastner, Kyle)

Classification

The one with max separation in the middle. That's Support Vector Machine.

(Image Credit: A Gentle Introduction To Machine Learning; SciPy 2013 Presentation - Kastner, Kyle)

ISACA Pune Chapter 25+ YEARS

ITS
ISACA Pune Chapter Conference
2025

Gold Sponsor
ShieldByte Infosec

Decision Tree

Any new person can walk through the Tree and predict.

ISACA. Pune Chapter 25+ YEARS

ITS
ISACA Pune Chapter Conference
2025

Gold Sponsor
ShieldByte
Infosec

Clustering

Wish to put 3 pizza shops in a city. Customers are plotted. What are the best locations?
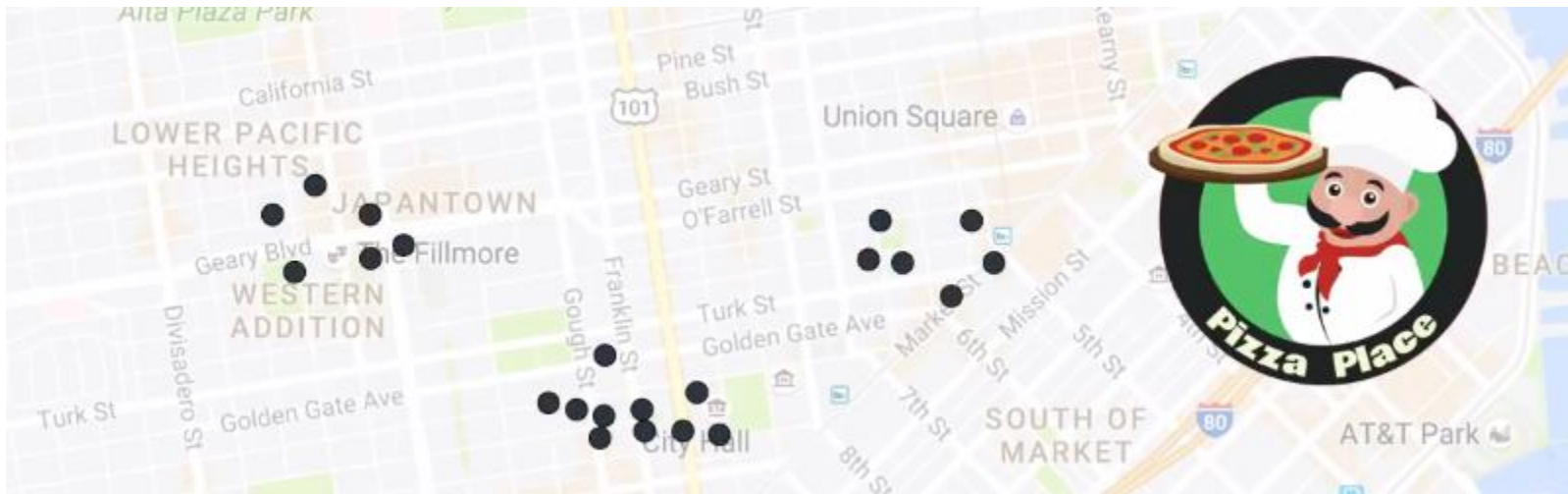
(Image Credit: A Gentle Introduction To Machine Learning; SciPy 2013 Presentation - Kastner, Kyle)

Clustering

Start with random locations and set ownership. Update locations. Repeat.

(Image Credit: A Gentle Introduction To Machine Learning; SciPy 2013 Presentation - Kastner, Kyle)

Applications of Machine Learning

- Face Recognition
- Spam recognition in Emails
- Recommender Systems
- Feelings Analysis, Sentiments
- Natural language: Translate a sentence from Hindi to English, question answering, etc.
- Speech: Recognize spoken words, speaking sentences naturally
- Game playing: Play games like chess
- Robotics: Walking, jumping, displaying emotions, etc.
- Driving a car, flying a plane, navigating a maze, etc.

ISACA Pune Chapter 25+ YEARS

ITS
ISACA Pune Chapter Conference
2025

Gold Sponsor
ShieldByte
Infosec

Example

ISACA. Pune Chapter 25+ YEARS

ITS
ISACA Pune Chapter Conference
2025

Gold Sponsor
ShieldByte
Infosec

Predicting House Price

Bedrooms	Sq. feet	Neighborhood	Sale price
3	2000	Normaltown	$250,000
2	800	Hipsterton	$300,000
2	850	Normaltown	$150,000
1	550	Normaltown	$78,000
4	2000	Skid Row	$150,000

Bedrooms	Sq. feet	Neighborhood	Sale price
3	2000	Hipsterton	???

Traditional way

Find formula, as simple as linear equation, weighed sum!!

Price = W_0 + W_1 x Bedrooms + W_2 x SqFeet (ignoring location for now)

BUT, how to decide which weights value to put? PRAY!!!

ISACA. Pune Chapter 25+ YEARS

ITS
ISACA Pune Chapter Conference
2025

Gold Sponsor
ShieldByte
Infosec

Traditional way

Grab an expert, get weights from him/her!!

```python
def estimate_house_sales_price(num_of_bedrooms, sqft, neighborhood):

        price = 0
        price_per_sqft = 200

        if neighborhood == "hipsterton":
                price_per_sqft = 400
        elif neighborhood == "skid row":
                price_per_sqft = 100

        price = price_per_sqft * sqft
        if num_of_bedrooms == 0:
                price = price - 20000
        else:
                price = price + (num_of_bedrooms * 1000)

        return price
```

ISACA Pune Chapter 25+ YEARS

ITS
ISACA Pune Chapter Conference
2025

Gold Sponsor
ShieldByte
Infosec

Prayer

- Wouldn't it be better if computer figures out?
- Treat it as black box
- Feed Inputs and outputs
- That's it!!

```
def estimate_house_sales_price(num_of_bedrooms, sqft, neighborhood):
    price = <computer, plz do some math for me>
    return price
```

ISACA Pune Chapter 25+ YEARS

ITS
ISACA Pune Chapter Conference
2025

Gold Sponsor
ShieldByte
Infosec

Prayer granted!!

- Notice the magic numbers
- .841, 1231.123, 2.324, and 201.234.
- These are weights.
- Better the weights - better the prediction!
- Done!!

```
def estimate_house_sales_price(num_of_bedrooms, sqft, neighborhood):
        price = 0
        price += num_of_bedrooms * .841
        price += sqft * 1231.123
        price += neighborhood * 2.324
        price += 201.234
        return price
```

How to figure out? A dumb way

- Step 1: Start with each weight set to 1.0:

```
def estimate_house_sales_price(num_of_bedrooms, sqft, neighborhood):
        price = 0
        price += num_of_bedrooms * 1.0
        price += sqft * 1.0
        price += neighborhood * 1.0
        price += 1.0
        return price
```

ISACA
Pune Chapter 25+ YEARS

ITS
ISACA Pune Chapter Conference
2025

Gold Sponsor
ShieldByte
Infosec

How to figure out? A dumb way

- Step 2: Guess/predict for all houses

Bedrooms	Sq. feet	Neighborhood	Sale price	My Guess
3	2000	Normaltown	$250,000	$178,000
2	800	Hipsterton	$300,000	$371,000
2	850	Normaltown	$150,000	$148,000
1	550	Normaltown	$78,000	$101,000
4	2000	Skid Row	$150,000	$121,000

Predictions NOT good, right?

ISACA
Pune Chapter 25+ YEARS

ITS
ISACA Pune Chapter Conference
2025

What to do?

- Actual $250,000, but guessed $178,000
- Off by $72,000 for that single house.
- Diff can be positive or negative, so square it
- Add squared diffs of all houses.
- Total: $86,123,373.
- That's whole error in the system.
- That's how ``wrong'' your function currently is.

What Next?

- Average per house error is ``cost''.
- Get cost to be zero/minimum by playing with the weights.
- That's the Goal!!!

Mathematically

$$J(\theta) = \frac{1}{2m} \sum_{i=1}^{m} \left(h_\theta(x^{(i)}) - y^{(i)} \right)^2$$

Graphically

Plotting cost values, for all possible ranges of weights for number_of_bedrooms and sqft:
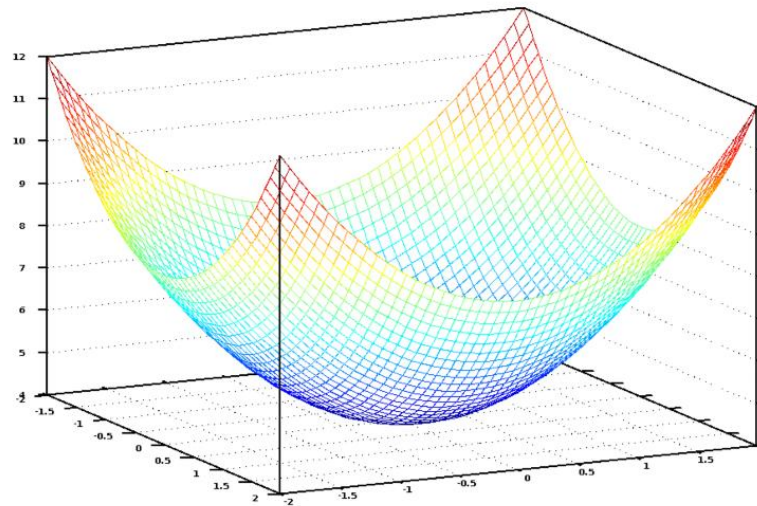
Calculus, anybody?

$$^*W_x = W_x - a\left(\frac{\partial Error}{\partial W_x}\right)$$

Old weight

Derivative of Error with respect to weight

New weight

Learning rate

Graphically

Cost is lowest at lowest point of the surface. Ideal. Weights at that point are the answers!
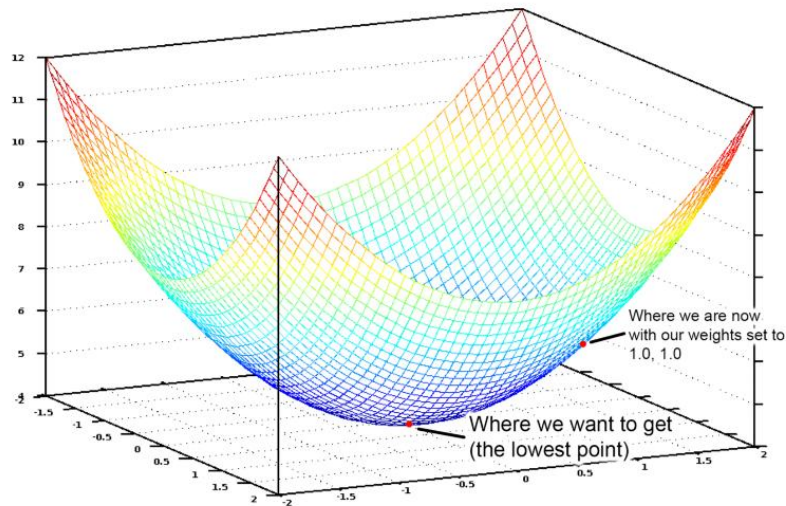
How to find the lowest cost point?

- Start somewhere.
- Find direction (slope? Derivative? Partial?)
- Derivative: tells us which way is downhill for any given point on our graph.
- Move in slope direction.
- Adjust our weights to get to next point
- `walking down hill" towards the lowest point.
- That's gradient descent.
- Scikit Learn does this for you, hushsh!!

Cool-down: Summary

SO...

- What is Machine learning, after-all?
- Its usage in your domain?

Thanks

yogeshkulkarni@yahoo.com