Machine Learning

6th Semester B.Tech. (CSE)
Course Code: CS31002

Dr. Jasaswi Prasad Mohanty
School of Computer Engineering
Bhubaneswar, India

MODULE I

A Few Quotes

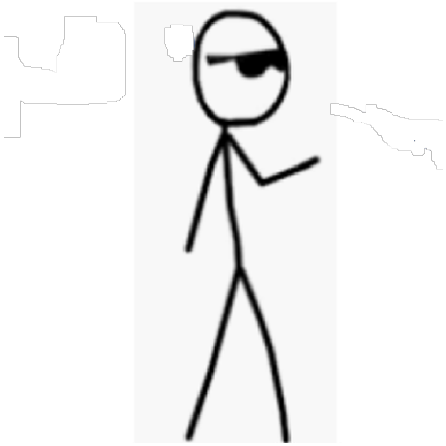
- A breakthrough in machine learning would be worth 10 Microsofts
  ----- Bill Gates, Chairman, Microsoft
- Machine learning is the next Internet
  ----- Tony Tether, Director, DARPA
- Machine learning is the hot new thing
  ----- John Hennessy, President, Stanford
- Web rankings today are mostly a matter of machine learning
  ----- Prabhakar Raghavan, Dir. Research, Yahoo
- Machine learning is going to result in a real revolution
  ----- Greg Papadopoulos, CTO, Sun
- Machine learning is today's discontinuity
  ----- Jerry Yang, CEO, Yahoo

What is Machine Learning?

Human can learn from past experience
and make decision of its own

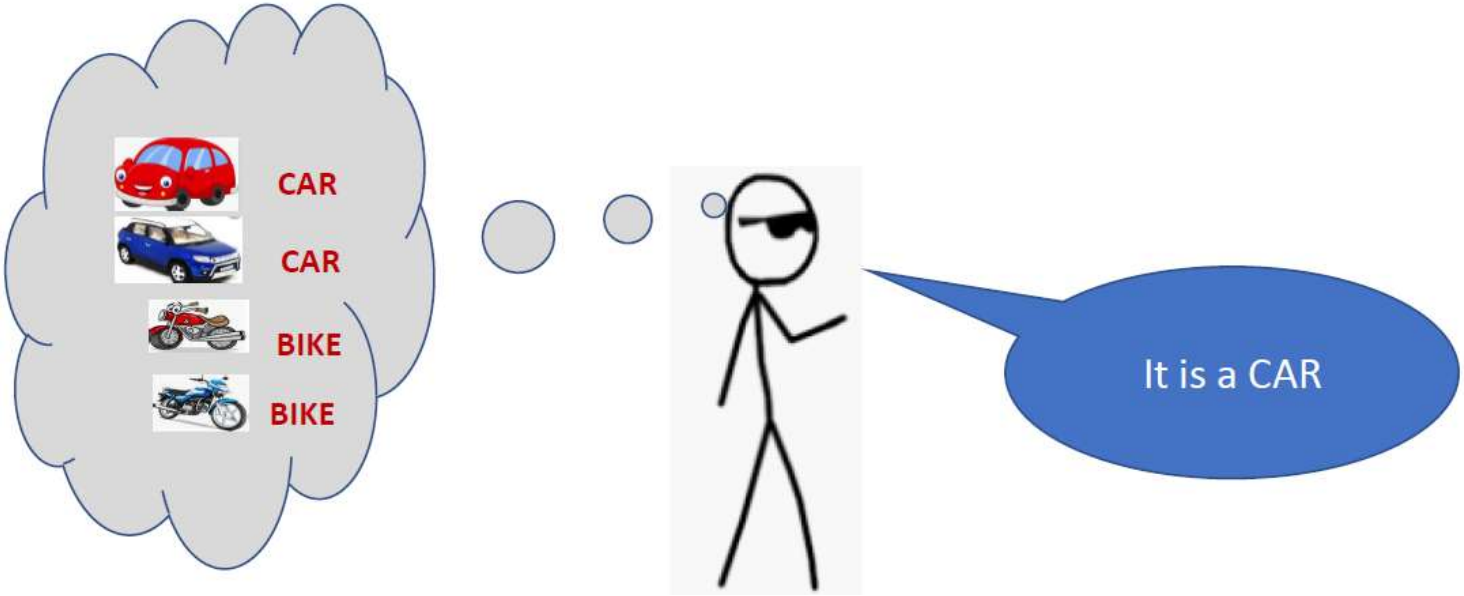
What is Machine Learning?

What is this object?



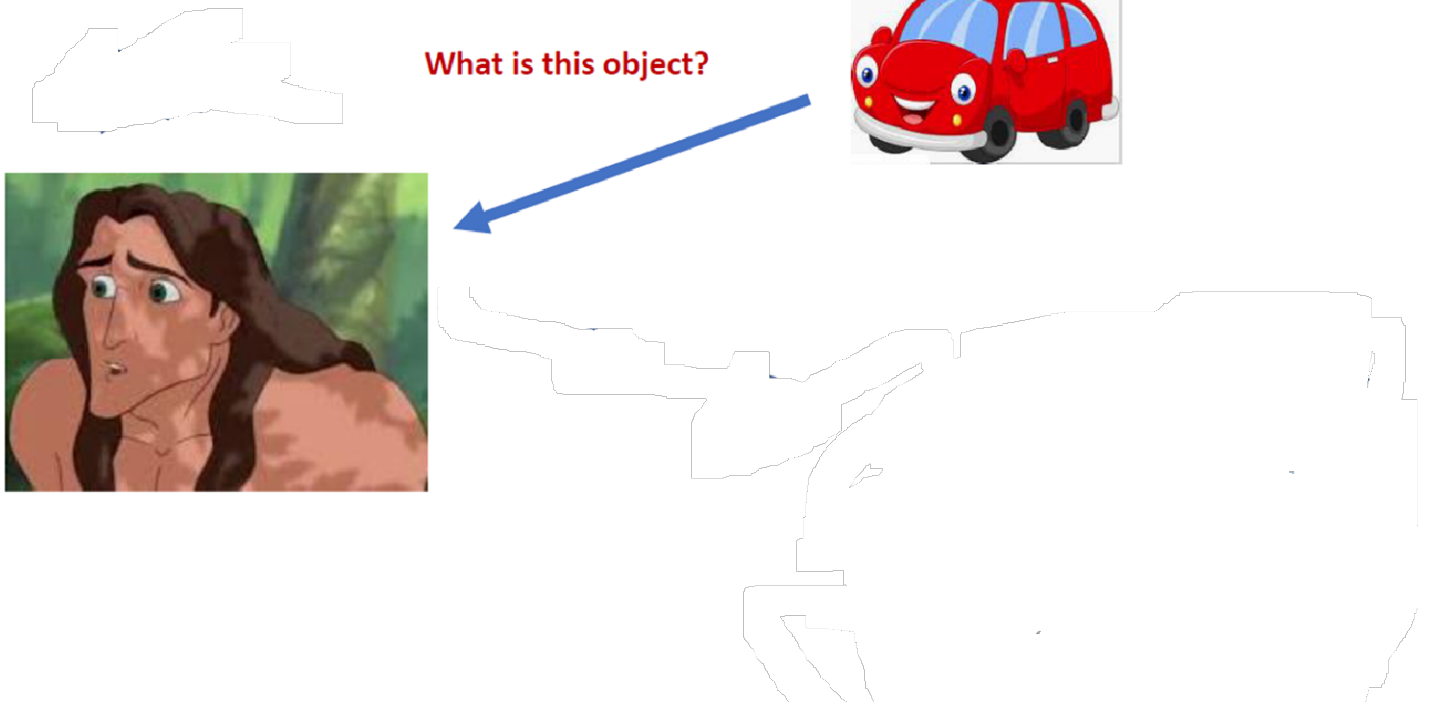
What is Machine Learning?

What is this object?



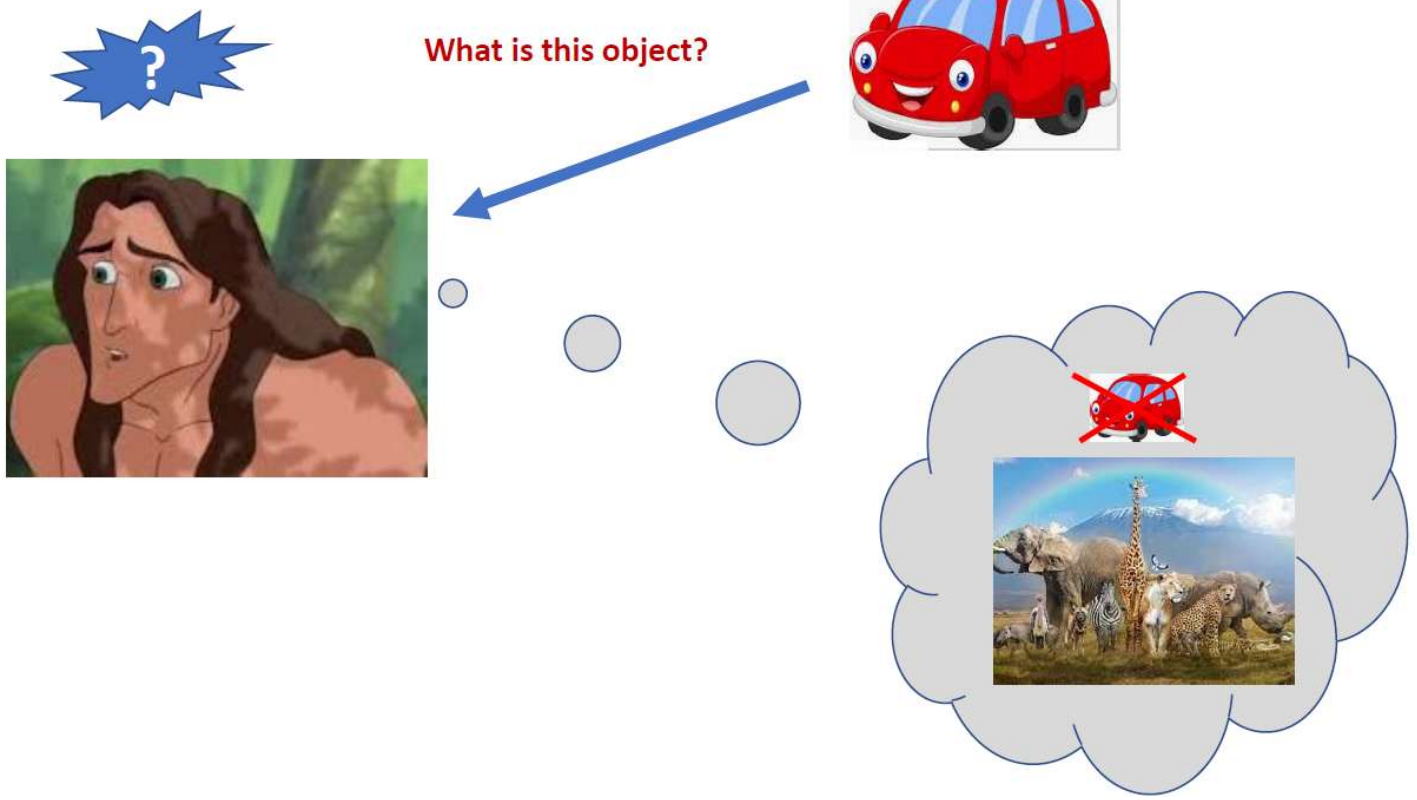
What is Machine Learning?

Let us ask the same question to him



What is Machine Learning?

Let us ask the same question to him



What is Machine Learning?

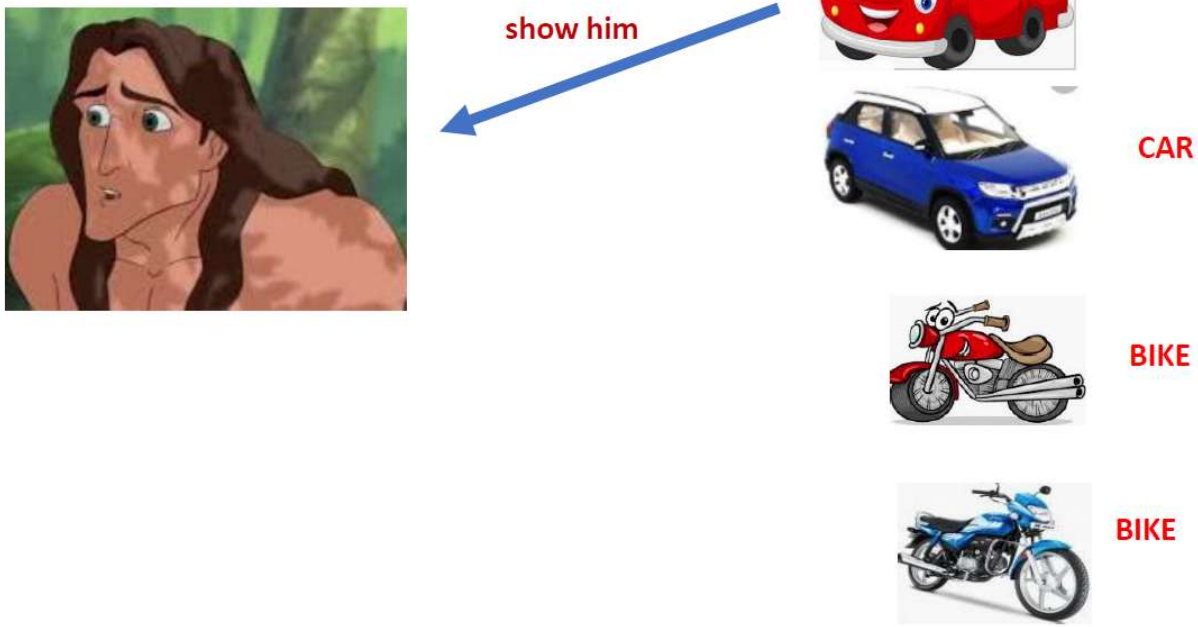
But, he is a human being. He can observe and learn.

Let us make him learn.

What is Machine Learning?

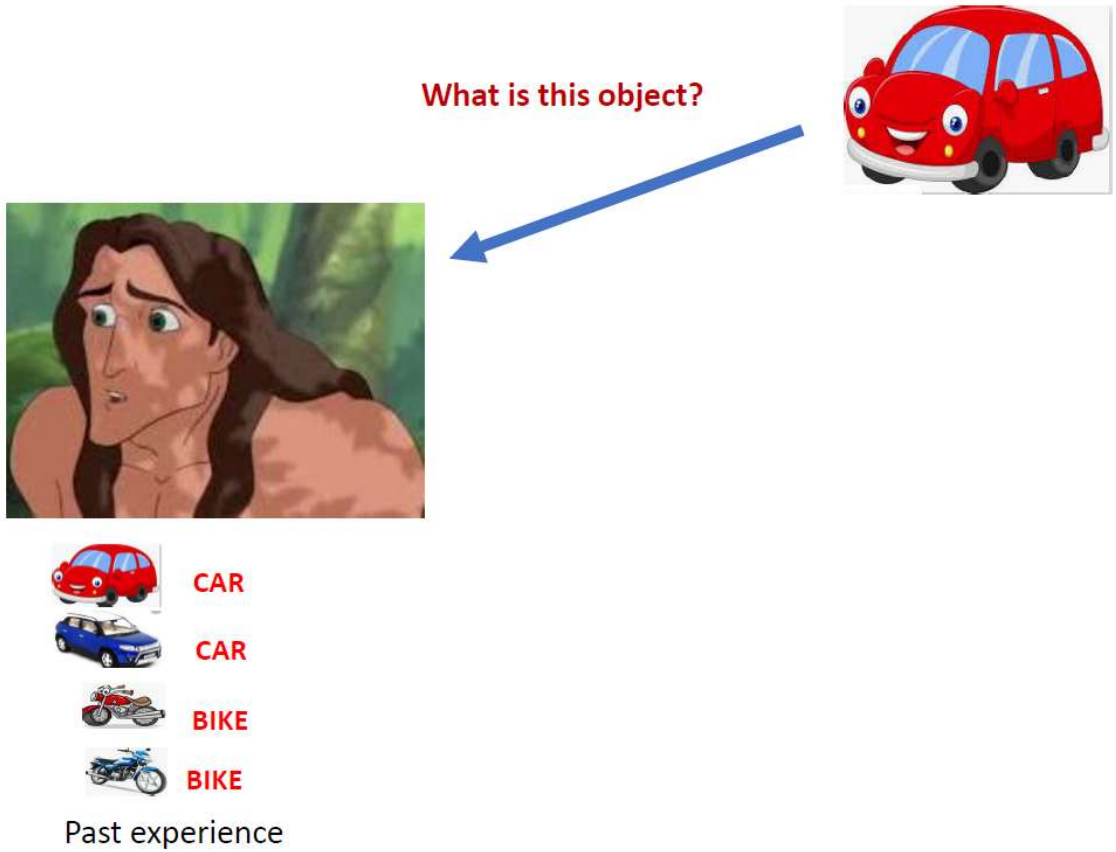
But, he is a human being. He can observe and learn.

Let us make him learn.



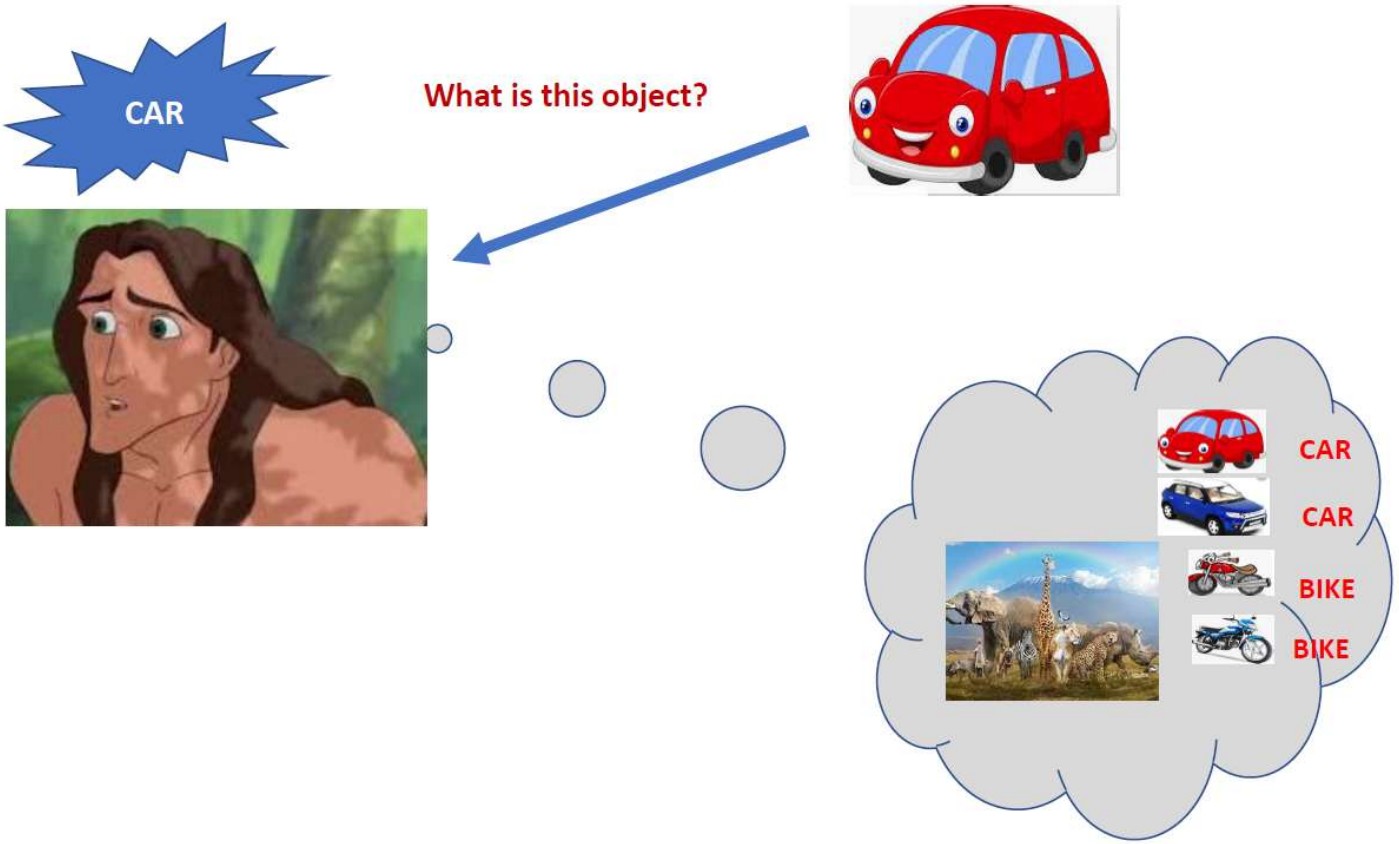
What is Machine Learning?

Let us ask the same question now.



What is Machine Learning?

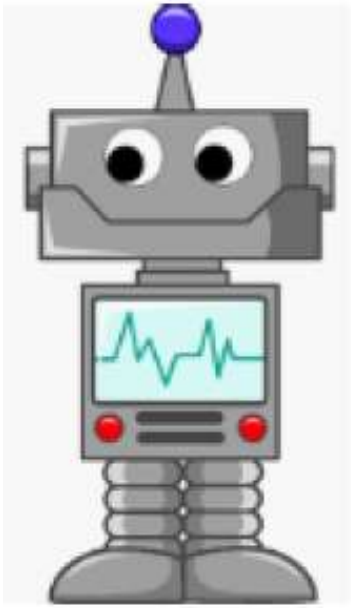
Let us ask the same question now.



What is Machine Learning?

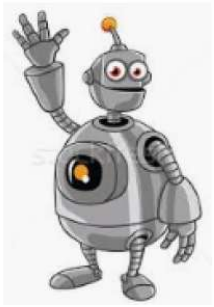
What about a machine?

- Machines follow instructions.
- It can not take decision of its own.
- We can ask a machine to perform an arithmetic operations such as:
  - Addition
  - Multiplication
  - Division
  - Comparison
  - Print
  - Plotting a chart

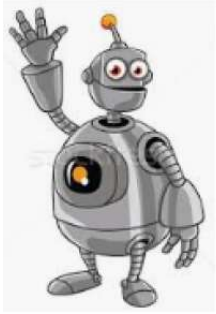


What is Machine Learning?

We want a machine to act like a human.

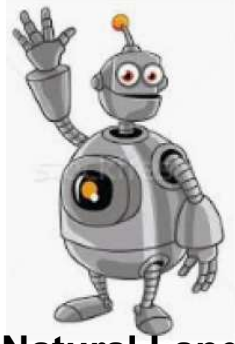


To identify this object.



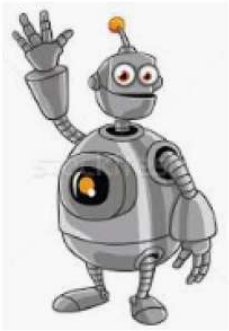
Price in 2025?

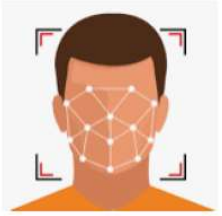
To predict the price in future.



I ~~made~~ met him yesterday

Natural Language understand, and correct grammar.





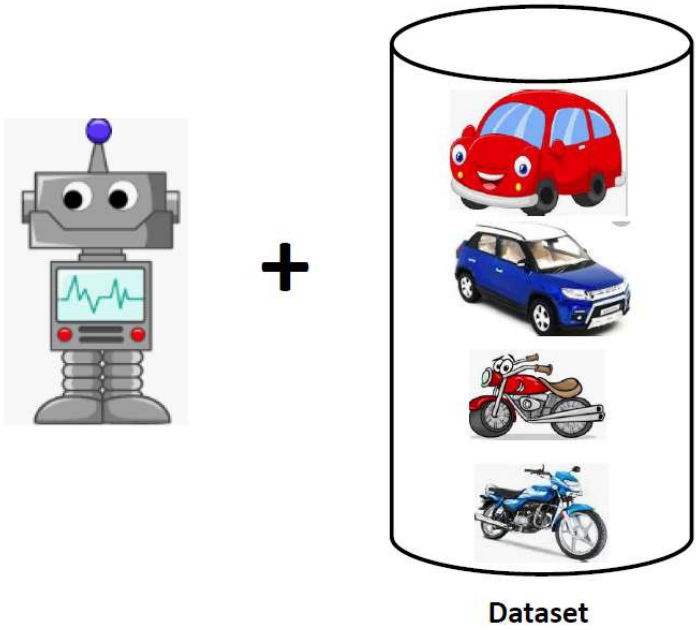
recognize face

To recognize faces.

What is Machine Learning?

What should we do to achieve all these?

- Just like, what we did to human, we need to provide experience to the machine.



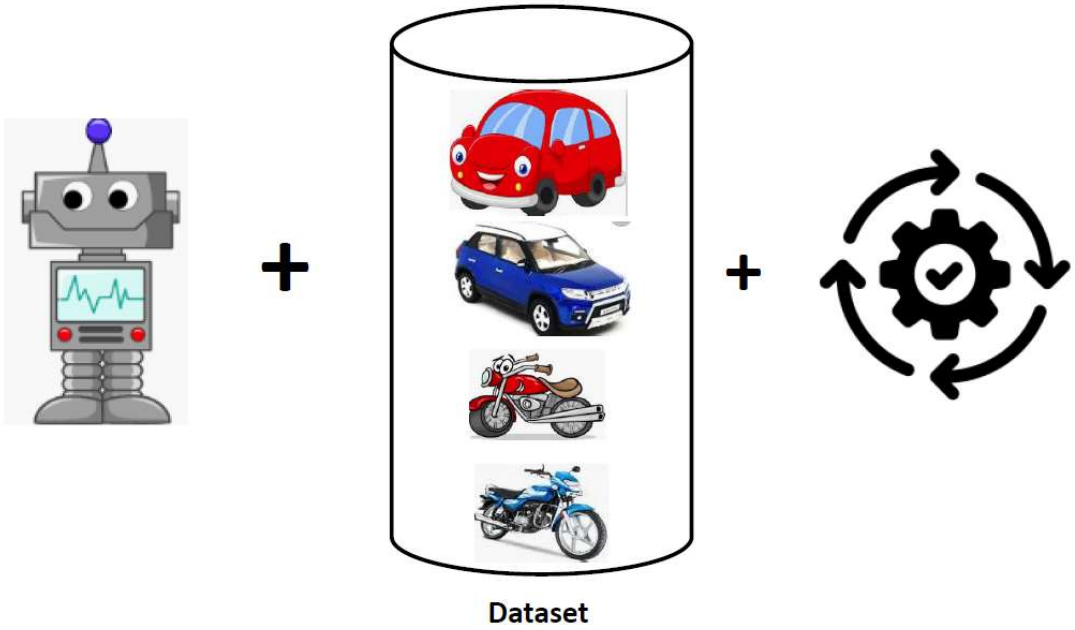
This what we called as Data or Training dataset.

So, we first need to provide training dataset to the machine.

What is Machine Learning?

What should we do to achieve all these?

- Just like, what we did to human, we need to provide experience to the machine.

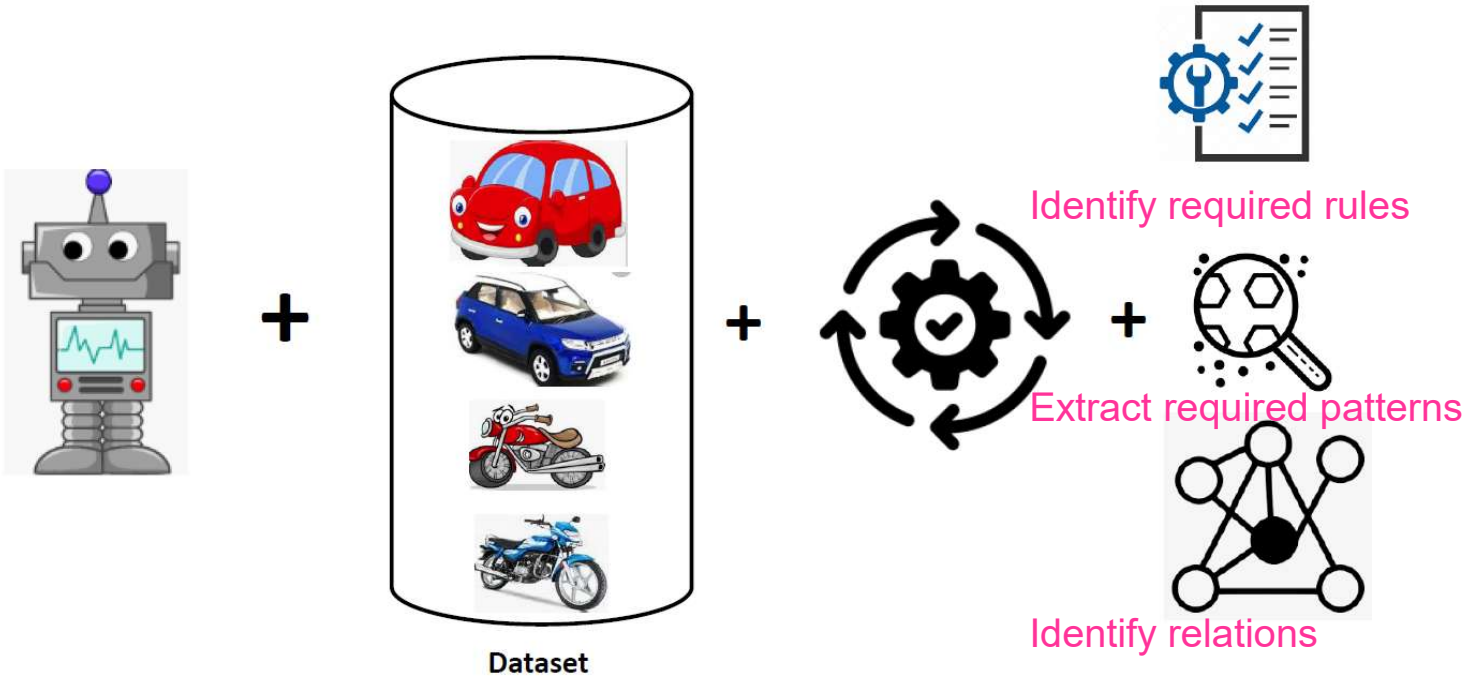


Then, devise algorithms and execute programs on the data with respect to the underlying target tasks.

What is Machine Learning?

What should we do to achieve all these?

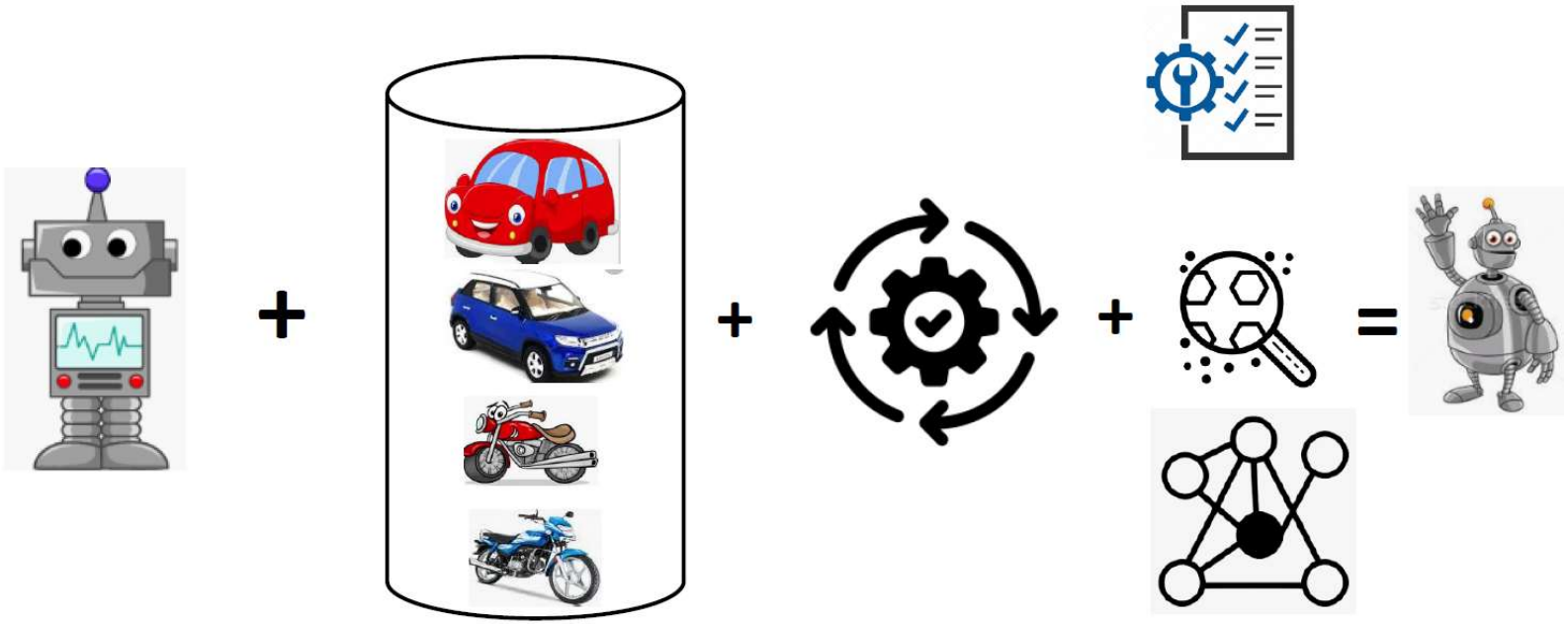
- Just like, what we did to human, we need to provide experience to the machine.



What is Machine Learning?

What should we do to achieve all these?

- Just like, what we did to human, we need to provide experience to the machine.

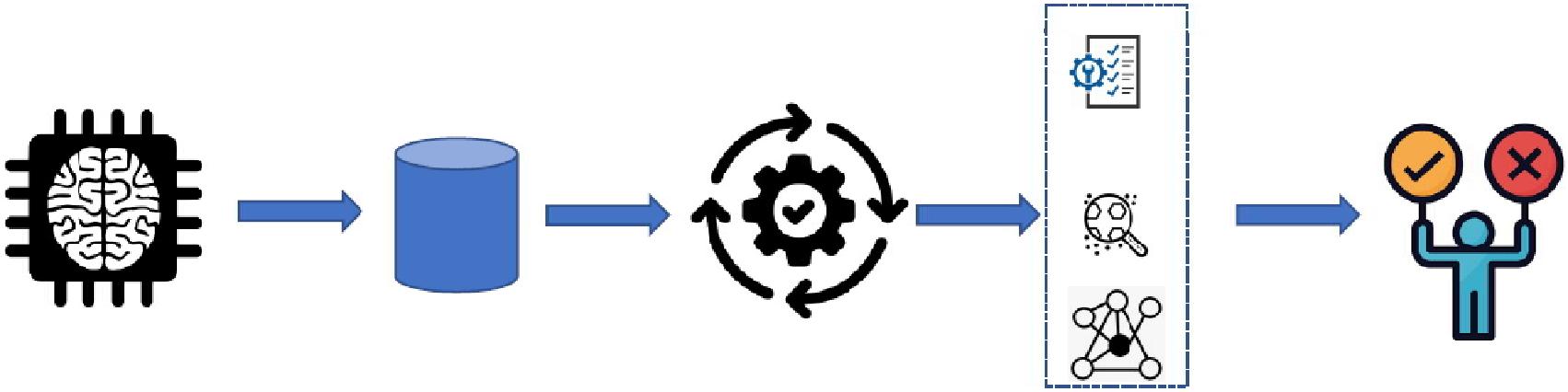


[So that machine can derive inferences from the data]

What is Machine Learning?

In summary

- Given a machine learning problem
  - Identify and create the appropriate dataset
  - Perform computation to learn required rules, pattern and relations
  - Output the decision



Motivation for ML

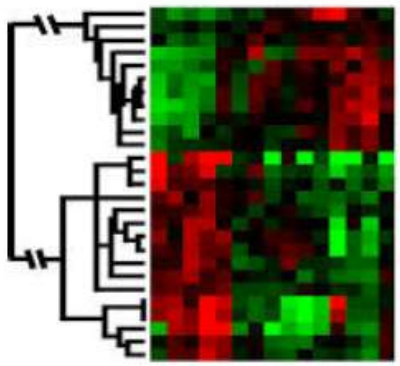
- It is hard to write programs for certain tasks
  - Human face or handwriting recognition
  - Playing complex games like chess
  - Recommending movies that a person will like
- Why ML?
  - We do not ourselves know how to solve
  - Algorithm will be too complicated
  - Too many instances of the program needed (e.g., one for every user)
  - Growing number of online data and their availability
  - Computational power is available
  - Progress in theory and algorithms

When Do We Use Machine Learning?

ML is used when:

- Human expertise does not exist (navigating on Mars)
- Humans can't explain their expertise (speech recognition)
- Models must be customized (personalized medicine)
- Models are based on huge amounts of data (genomics)



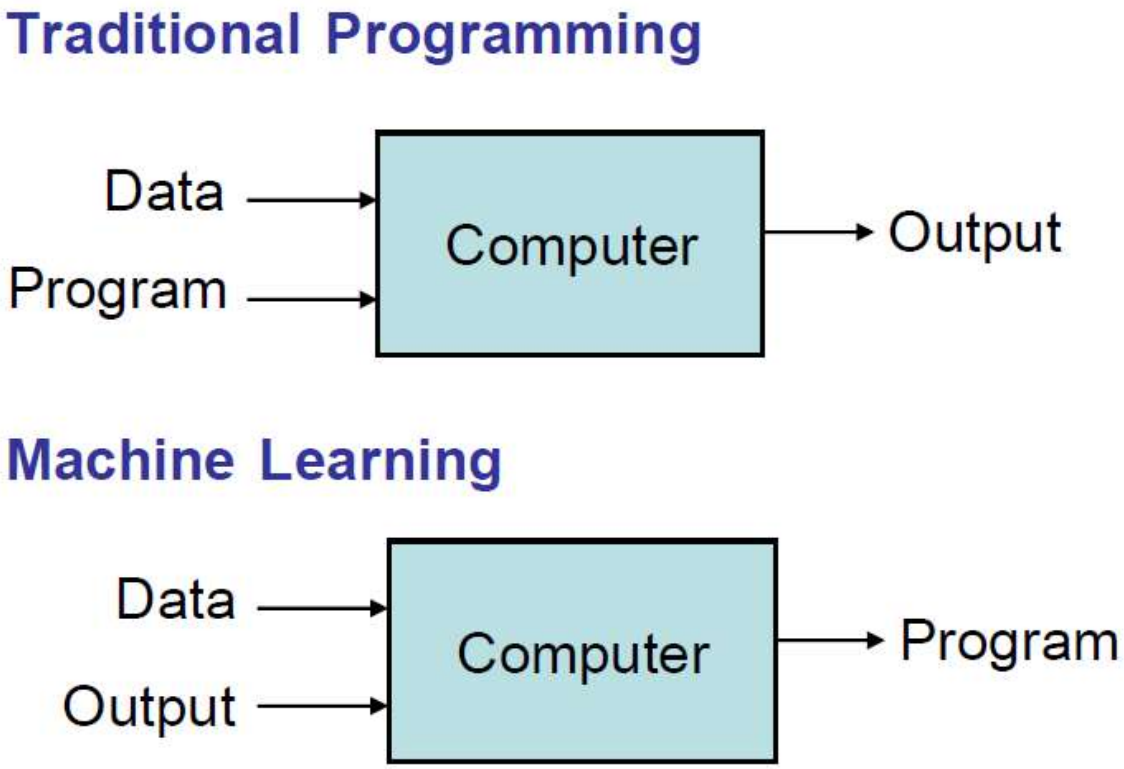
Example

A classic example of a task that requires machine learning: It is very hard to say what makes a 2

Motivation for ML

- Instead of writing a program by hand, collect lots of examples that specify the correct output for a given input.
- A machine learning algorithm takes these examples and produces a program that does the job.
- If done right, the program works for new cases as well as the ones we trained it on.

Traditional Programming vs. Machine Learning



Machine Learning Definition

- **Arthur Samuel (1959):**

  Machine Learning is a field of study that gives computers the ability to learn without being explicitly programmed.

- **Herbert Simon:**

  Learning is any process by which a system improves performance from experience.

- **Tom Mitchell (1998):**

  Machine Learning is a well-posed learning problem defined as: A computer program is said to learn from experience E with respect to some task T and some performance measure P, if its performance on T, as measured by P, improves with experience E.

  Example: Suppose your email program watches which emails you do or do not mark as spam, and based on that learns how to better filter spam.

  Here T: Classifying emails as spam or not spam

  E: Watching you label emails as spam or not spam.

  P: The number (or fraction) of emails correctly classified as spam/not spam.

Defining the Learning Task: More Examples

Improve on task T, with respect to performance metric P, based on experience E.

1. T: Playing checkers
   P: Percentage of games won against an arbitrary opponent
   E: Playing practice games against itself
2. T: Recognizing hand-written words
   P: Percentage of words correctly classified
   E: Database of human-labeled images of handwritten words
3. T: Driving on four-lane highways using vision sensors
   P: Average distance traveled before a human-judged error
   E: A sequence of images and steering commands recorded while observing a human driver.

Machine Learning Definition – contd…

- **Ethem Alpaydin:**

  Machine learning is programming computers to optimize a performance criterion using example data or past experience.

- **Kevin P. Murphy:**

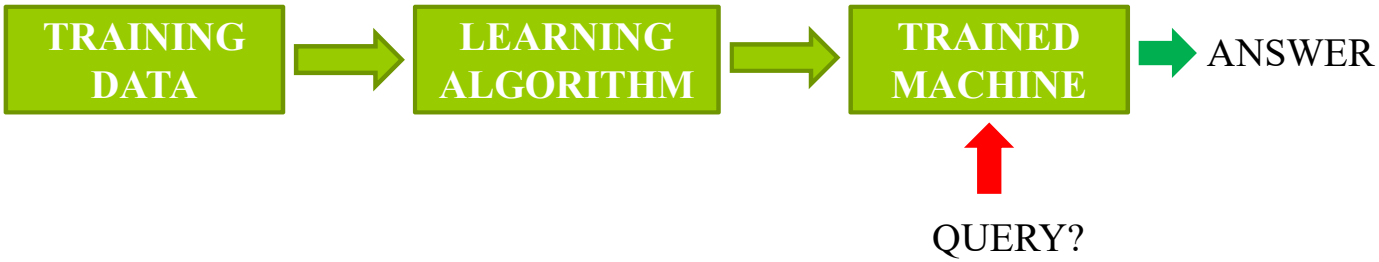
  The goal of machine learning is to develop methods that can automatically detect patterns in data, and then to use the uncovered patterns to predict future data or other outcomes of interest.

- **Christopher M. Bishop:**

  The field of pattern recognition is concerned with the automatic discovery of regularities in data through the use of computer algorithms and with the use of these regularities to take actions.

What is ML?

- Endow computers with the ability to ‘learn’ from data.



Quiz

A computer program is said to learn from experience E with respect to some task T and some performance measure P if its performance on T, as measured by P improves with experience E. Suppose we feed a learning algorithm a lot of historical weather data, and have it learn to predict weather. What would be a reasonable choice for P?

a) The weather prediction task

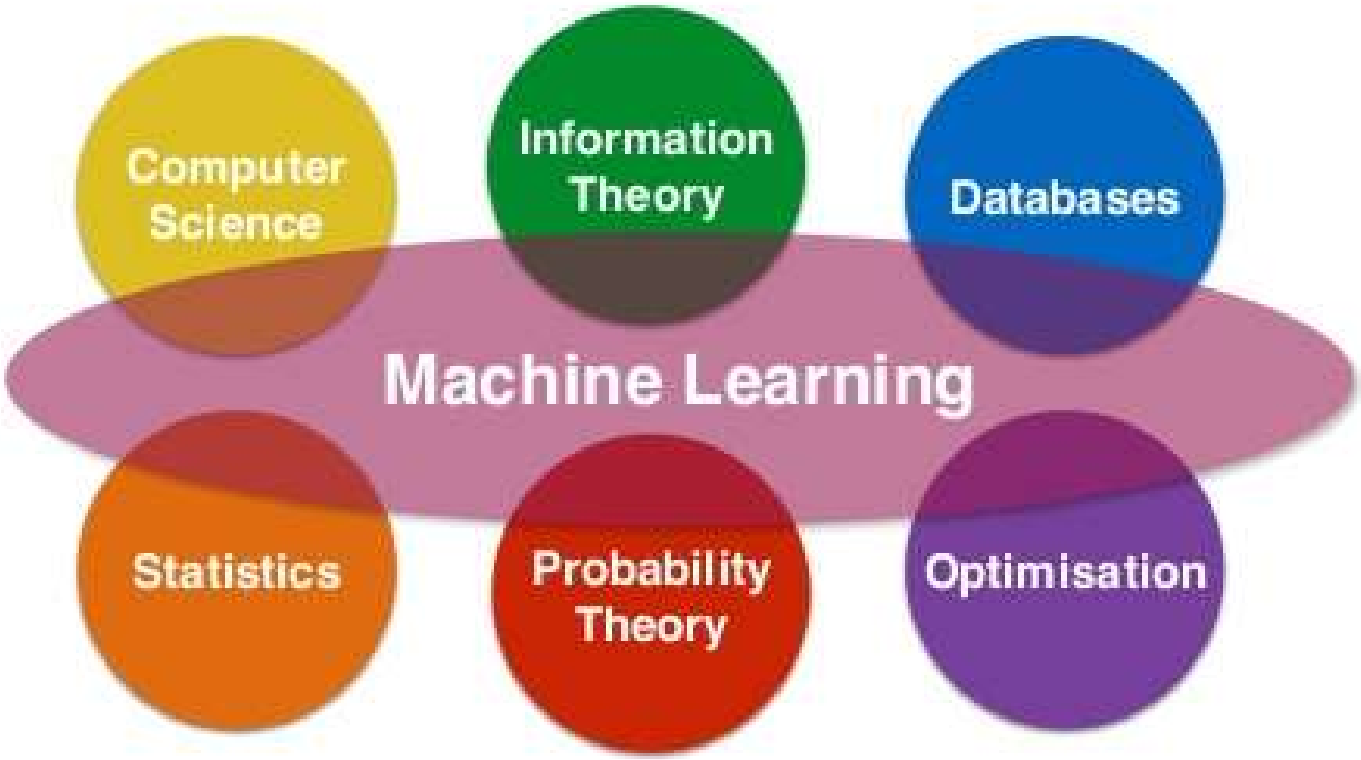
b) The probability of correctly predicting future date's weather

c) The process of the algorithm examining a large amount of historical weather data.

d) None of these

Answer: b)

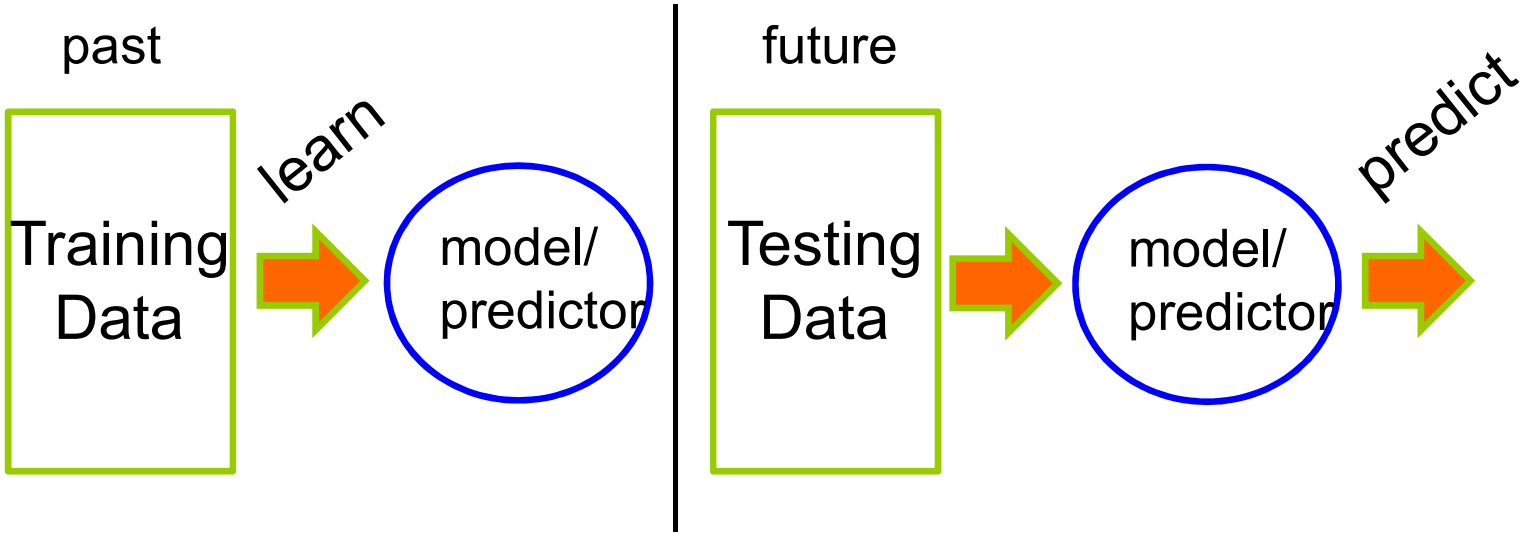
ML founding disciplines



Machine Learning is…

- **Hal Daume:**

  Machine learning is about predicting the future based on the past.



Is ML a Magic?

No, more like gardening

- **Seeds** = Data
- **Nutrients** = Outputs
- **Gardener** = You
- **Plants** = Programs

Resources: Journals & Conferences

- Journal of Machine Learning Research www.jmlr.org
- Machine Learning
- Neural Computation
- Neural Networks
- IEEE Trans on Neural Networks and Learning Systems
- IEEE Trans on Pattern Analysis and Machine Intelligence
- Journals on Statistics/Data Mining/Signal Processing/Natural Language Processing/Bioinformatics/...

- International Conference on Machine Learning (ICML)
- European Conference on Machine Learning (ECML)
- Neural Information Processing Systems (NIPS)
- Uncertainty in Artificial Intelligence (UAI)
- Computational Learning Theory (COLT)
- International Conference on Artificial Neural Networks (ICANN)
- International Conference on AI & Statistics (AISTATS)
- International Conference on Pattern Recognition (ICPR)

Tasks solved by using a learning algorithm
(Real World Applications)

- Recognizing patterns
  - Facial identities or facial expressions
  - Handwritten or spoken words
  - Medical images
- Generating patterns
  - Generating images or motion sequences
- Recognizing anomalies
  - Unusual sequences of credit card transactions
  - Unusual patterns of sensor readings in a nuclear power plant or unusual sound in your car engine.
- Prediction
  - Future stock prices or currency exchange rates

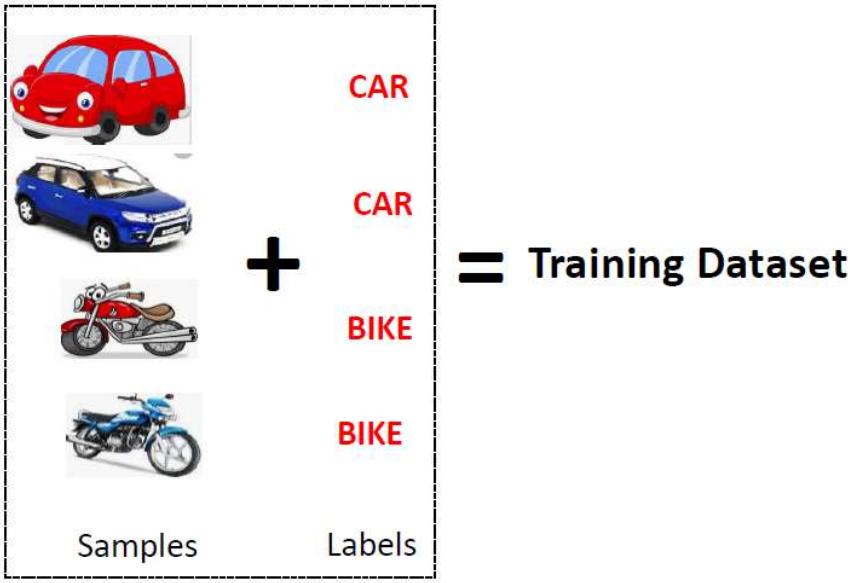
- Spam filtering, fraud detection
  - The enemy adapts so we must adapt too.
- Recommendation systems
  - Need to adapt to millions of individuals. E.g., Youtube video, Netflix movies recommendation
- Information retrieval
  - Find documents or images with similar content.
- Data Visualization
  - Display a huge database in a revealing way

Tasks solved by using a learning algorithm
(Real World Applications)

- Speech recognition
- Computer Vision
  - face recognition, classification of Microscope image of cells, handwritten digit/text recognition, Image denoising, etc.
- Robot control
- Medical Diagnosis / Neuro science
- Biosurvelliance
- Genomic analysis
- Stock Market Analysis
- Detecting Credit Card Fraud
- Detecting Network Intrusion
- Detecting latent factors (hidden relationship and correlations)
- Bioinformatics

Types of Learning

1. **Supervised (inductive) learning**
   - In supervised learning, we need some thing called a Labelled Training Dataset
   - Training data includes desired / correct outputs or labels
   - Learn by examples
   - Trying to predict a specific quantity or category
   - Can measure accuracy directly.
   - Given a labelled dataset, the task is to devise a function which takes the dataset, and a new sample, and produces an output value.



$f($  $)=$ CAR

Types of Learning

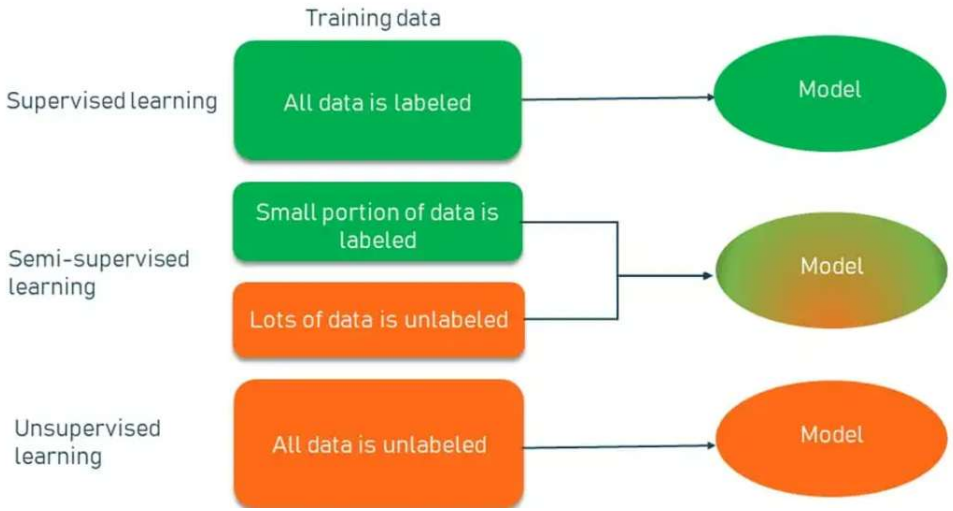
2. Unsupervised Learning

- In the unsupervised learning, we do not need to know the labels or ground truth values.
- Training data does not include desired outputs.
- Tries to identify similarities between input.
- Trying to “understand” the data
- Looking for structure or unusual patterns.
- Not looking for something specific (supervised).
- Evaluation usually indirect or qualitative.

Types of Learning – contd…

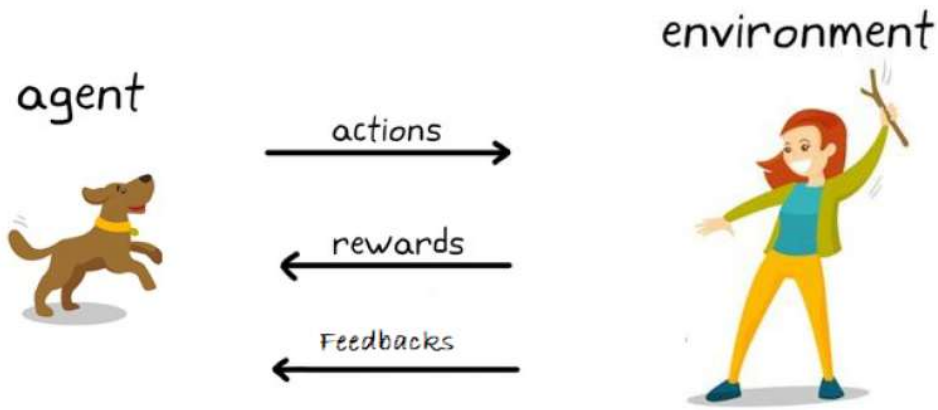
3. Semi-supervised learning

- Usually training data includes a few labeled examples plus a lot of unlabeled examples
- Training data has desired outputs, but for a different (related) task.
- Using unsupervised methods to improve supervised algorithms.

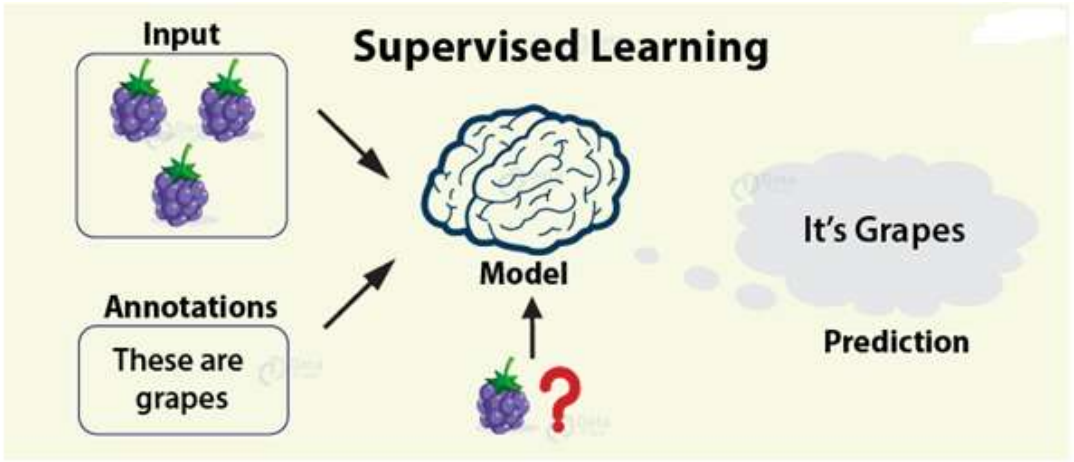


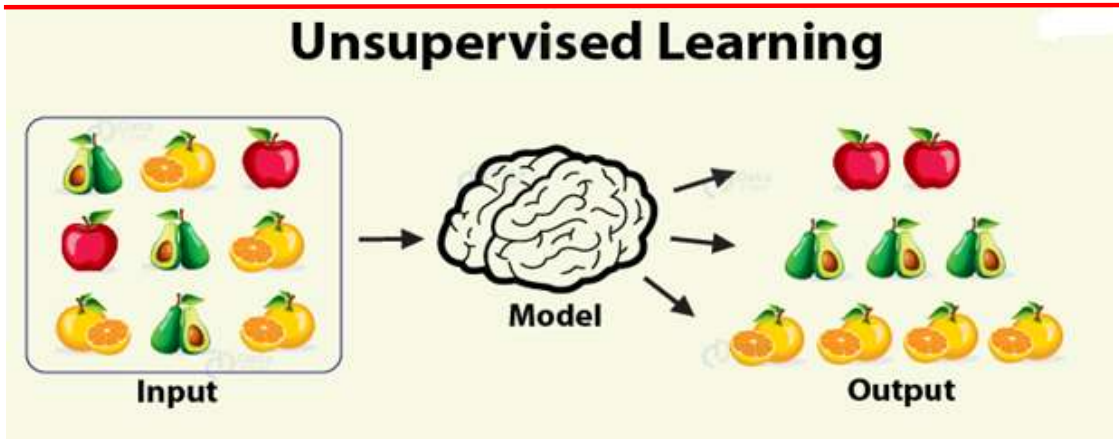
Types of Learning – contd…

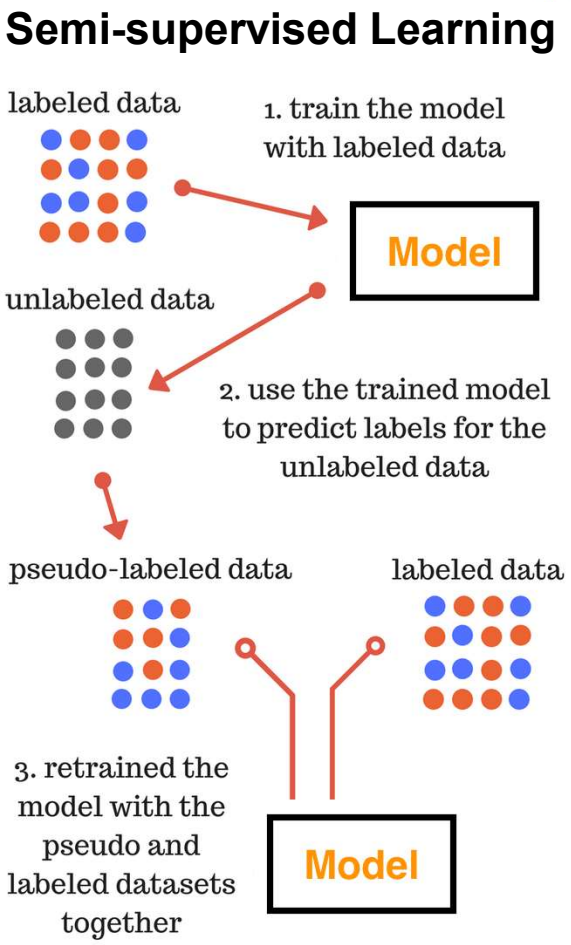
4. **Reinforcement learning**
   - It is also known as learning from trials and errors.
   - Learn by experience
   - Rewards from sequence of actions
   - Algorithm learns to teach to an environment
   - The algorithm is told when the answer is wrong but does not get told how to correct it. It explores possibilities until it gets the answer right.



Types of Learning – contd…



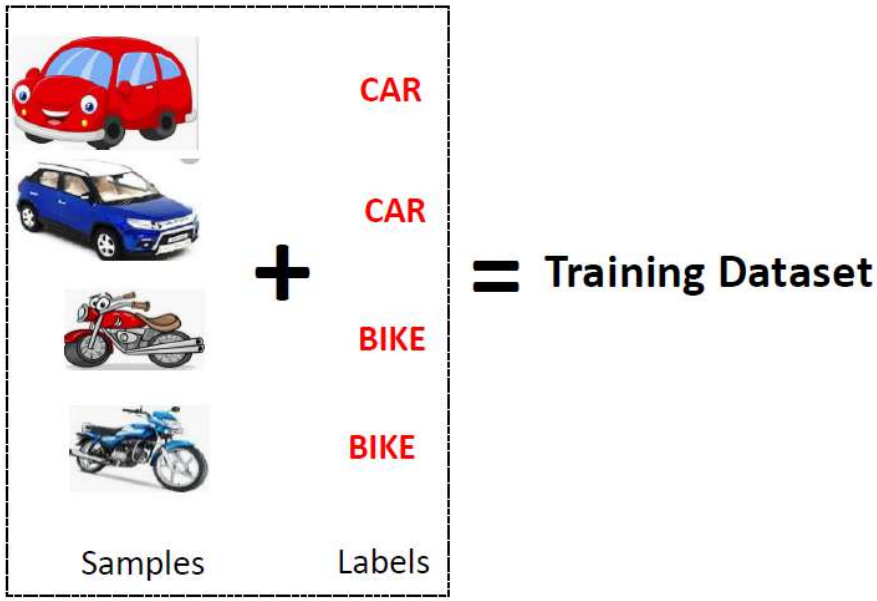




Types of Supervised Learning

a) Classification:

- If the possible output values of the function are predefined and discrete/categorical, it is called Classification.
- Predefined classes means, it will produce output only from the labels defined in the dataset.
  - For example, even if we input a bus, it will produce either CAR or BIKE.



Examples of real-world classification problems

- Predicting whether a patient has a specific disease (e.g., diabetes, cancer) based on symptoms, test results, or imaging data.
- Classifying X-rays or MRIs as normal or showing abnormalities.
- Identifying fraudulent transactions in credit card data.
- Predicting categories of products a customer is likely to buy.
- Categorizing customer reviews or social media posts as positive, negative, or neutral.
- Predicting whether students will pass or fail a course based on attendance and grades.
- Classifying road conditions as clear, congested, or blocked.
- Predicting the winner of a match based on past performance.

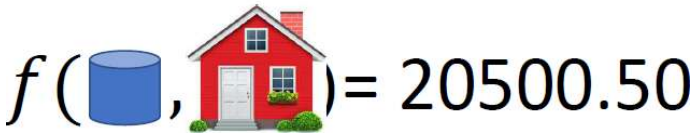
Types of Supervised Learning

b) Regression:

- If the possible output values of the function are continuous real values, then it is called Regression.

NOTE: The classification and Regression problems are supervised, because the decision depends on the characteristics of the ground truth labels or values present in the dataset, which we define as experience.

Dataset



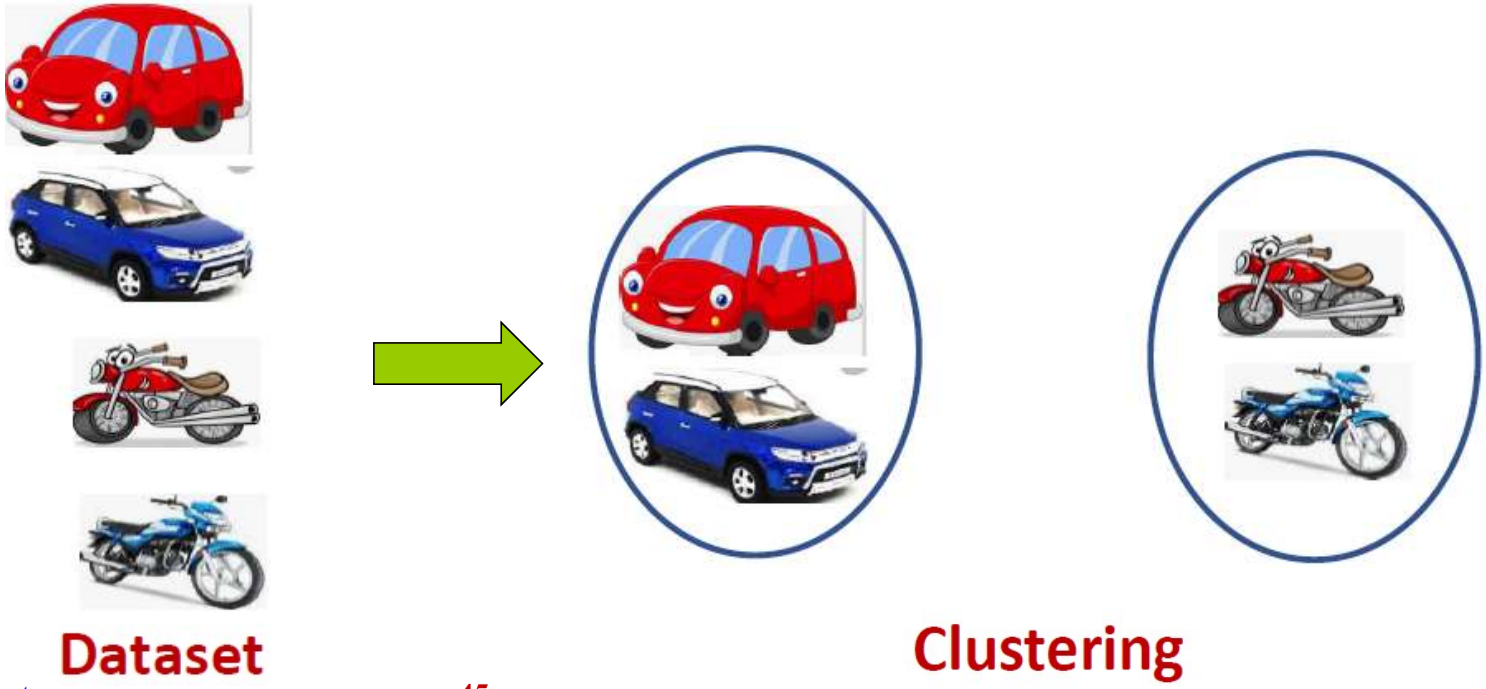
Examples of real-world regression problems

- Predict tomorrow's stock market price given current market conditions and other possible side information.
- Predict the age of a viewer watching a given video on YouTube.
- Predict the location in 3D space of a robot arm end effector, given control signals (torques) sent to its various motors.
- Predict the amount of prostate specific antigen (PSA) in the body as a function of a number of different clinical measurements.
- Predict the temperature at any location inside a building using weather data, time, door sensors etc.
- Predict the mark of a student in a particular subject.

Types of Unsupervised Learning

a) Clustering:

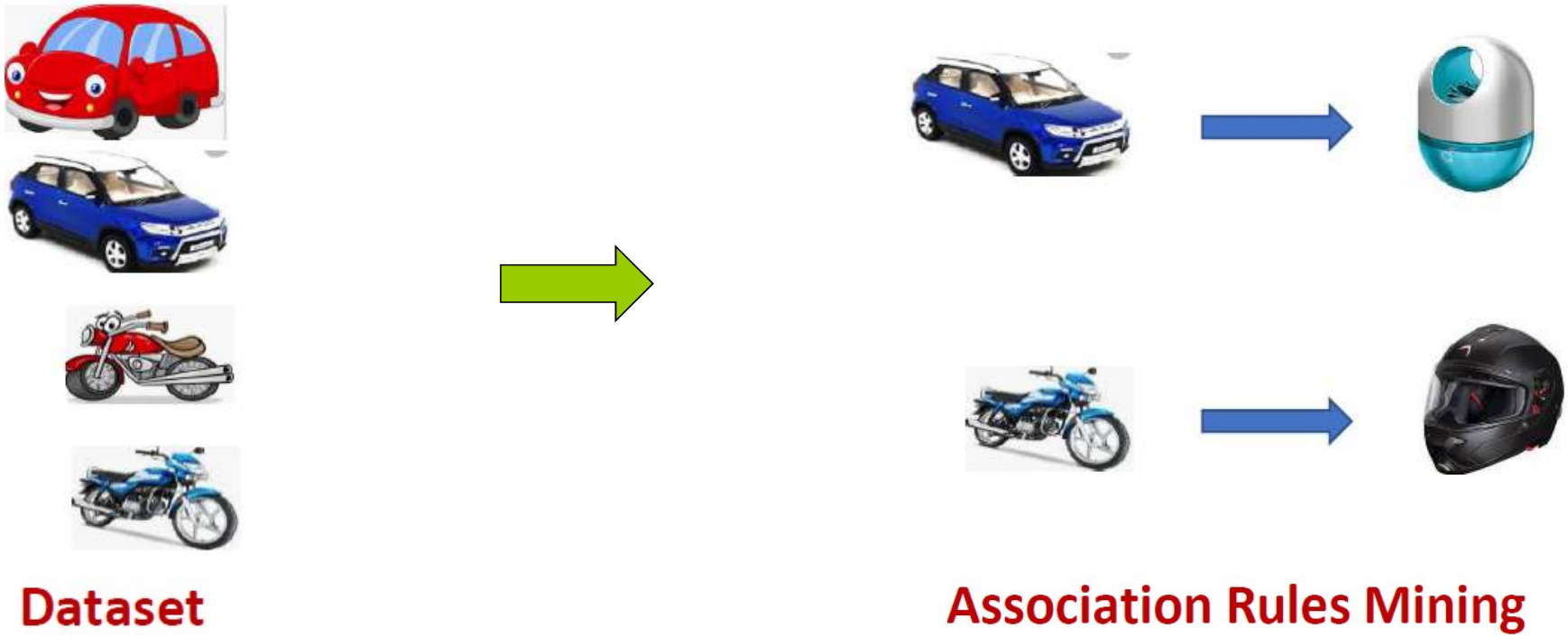
- The task is to identify the patterns like group the similar objects together.



Types of Unsupervised Learning

b) Association Rule Mining:

- It is a technique used in data mining to uncover interesting relationships, patterns, or associations among a set of items in large datasets.



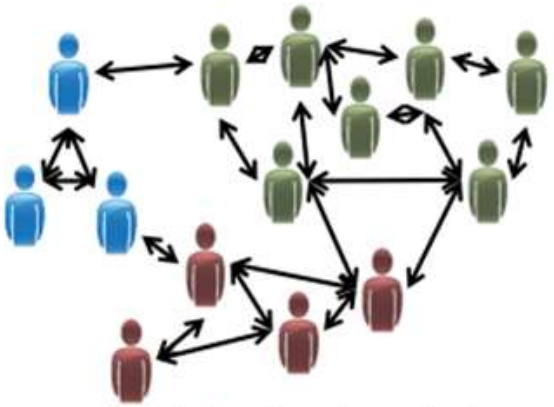
More Examples of Unsupervised Learning

Recommender System

More Examples of Unsupervised Learning

- **Organize Computing Clusters**: Find a group of machines that work together in a data center.
- **Social Network Analysis:** Given the knowledge of which friend you email the most or given your facebook friends. Can we automatically identify which are cohesive group of friends?
- **Market Segmentation:** Looking at the customer dataset can we automatically discover market segments.

Organize computing clusters

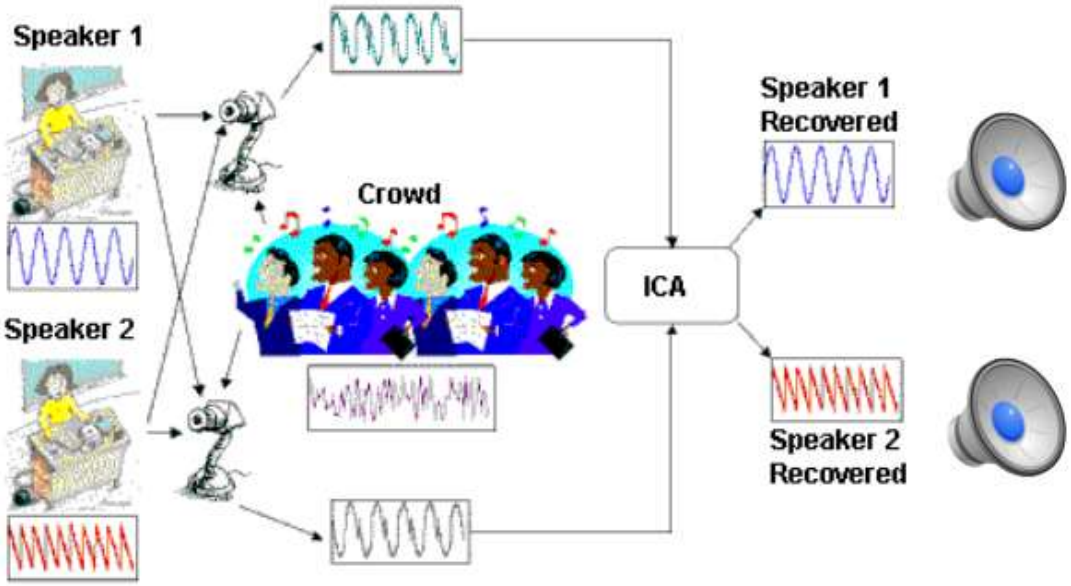


Social network analysis

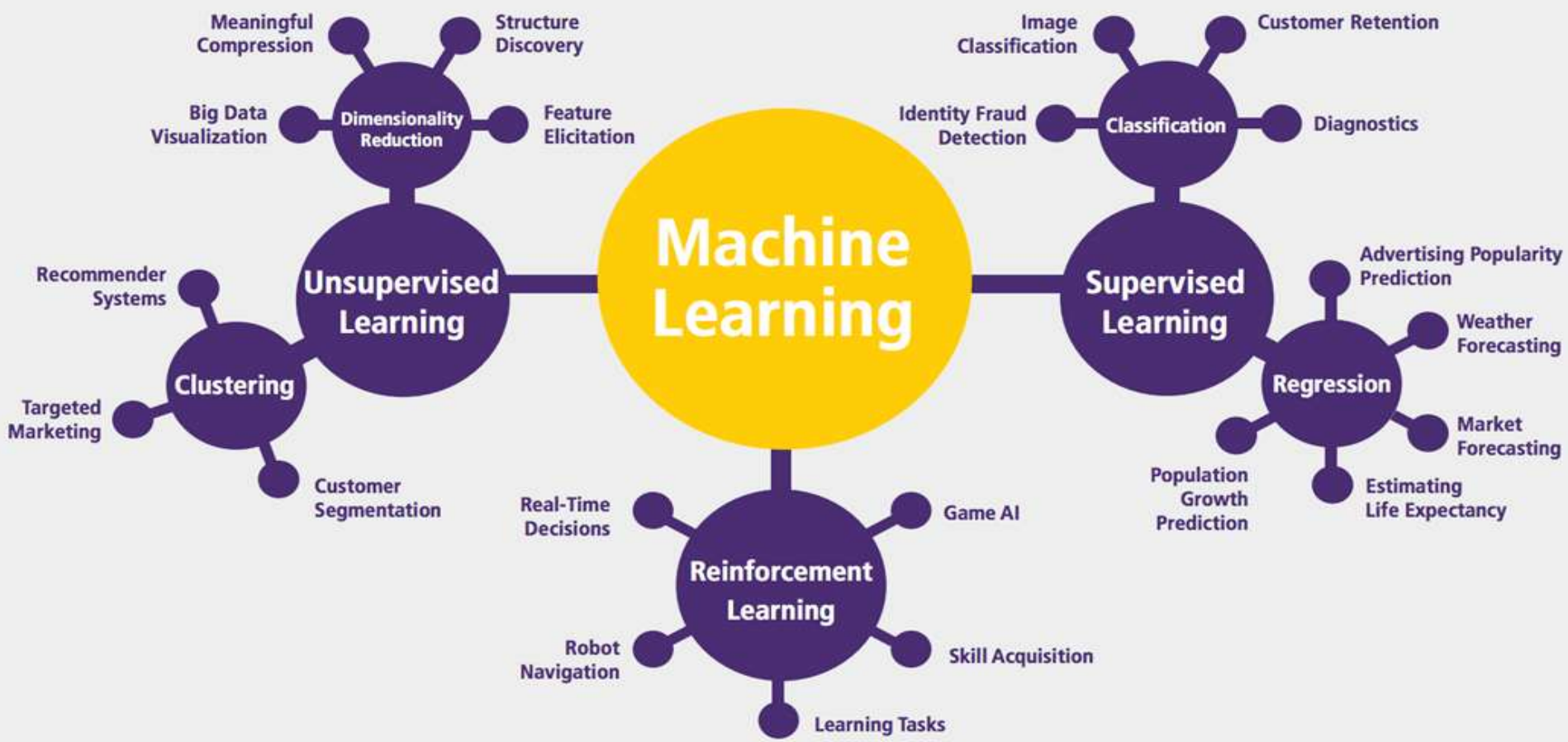
Market segmentation

More Examples of Unsupervised Learning

- **Independent Component Analysis (ICA)**
  - Separate a combined signal into its original sources.
  - There are two persons and two microphones. Each microphone records overlapping combination of music. These two recordings are provided to an unsupervised algorithm which separates out these two audio sources.
  - Also it can separate the music from voices.



Types of Machine Learning algorithms



Problems solved by ML

- **Two Class classification Problem:**
  - Recognising a disease (Diabetes) from past data of patients.
  - Because our answer consists of two options either yes or no.

	Age (X_1)	Weight (X_2)	Blood Pressure (X_3)	Sugar (X_4)	Cholesterol (X_5)	Having Diabetes or not (Y)
1	x_{11}	x_{12}	x_{13}	x_{14}	x_{15}	*yes*
2	x_{21}	x_{22}	x_{23}	x_{24}	x_{25}	*no*
	⋮	⋮	⋮	⋮	⋮	⋮
i	x_{i1}	x_{i2}	x_{i3}	x_{i4}	x_{i5}	*yes*
	⋮	⋮	⋮	⋮	⋮	⋮
1000	$x_{1000,1}$	$x_{1000,2}$	$x_{1000,3}$	$x_{1000,4}$	$x_{1000,5}$	*no*
New Patient	$x_{new,1}$	$x_{new,2}$	$x_{new,3}$	$x_{new,4}$	$x_{new,5}$	?

Problems solved by ML

- **Two Class classification Problem:**
  - Movie Recommendation.
  - Because our answer consists of two options either yes or no.

Movie Name	*Persons*				I want or not (Y)
	$user_1$	$user_2$	…	$user_m$	
m_1	⋮	⋮	…	⋮	*yes*
m_2	⋮	⋮	…	⋮	*yes*
m_3	⋮	⋮	…	⋮	*no*
⋮	⋮	⋮	…	⋮	⋮
m_r	⋮	⋮	…	⋮	*yes*
New Movie	⋮	⋮	…	⋮	?

Problems solved by ML

- **Ten Class classification Problem:**
  - Recognising hand written digits.
  - Because our answer consists of ten options that is digits from 0 to 9.

	Image of Handwritten Digits (Matrices or Column Vectors)	Actual Digit (Y)
1		3
2		7
⋮	⋮	⋮
⋮	⋮	⋮
N		1
New Image		?

Problems solved by ML

- **Regression Problem:**
  - Predicting Sales based on advertising budget on three media: TV, Radio, News Paper.

	Advertising Budget on			Sales (Y)
	TV (X_1)	Radio (X_2)	News Paper (X_3)	
x_1	x_{11}	x_{12}	x_{13}	y_1
x_2	x_{21}	x_{22}	x_{23}	y_2
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
x_i	x_{i1}	x_{i2}	x_{i3}	y_i
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
x_N	x_{N1}	x_{N2}	x_{N3}	y_N
	-	-	-	?

Problems solved by ML

- **Regression Problem:**
  - Predicting rainfall in a certain week.
  - Input is an Image and Output is a Real Number.

Image of Cloud		Quality of Rainfall (in cm)
3rd Week of June	4th Week of June	
2020	2020	?
2019	2019	1.5 cm
2018	2018	2.00 cm
2017	2017	2.5 cm
⋮	⋮	⋮
2000	2000	1.003 cm

Problems solved by ML

- **Regression Problem:**
  - Stock Index Prediction.
  - We get the value of Stock Index as a Real Number.

Risk free Interest Rate	Price of Gold	Price of Petrol	USD/ INR Exchange Value	Stock Index (Number)
-	-	-	-	-
-	-	-	-	-
-	-	-	-	-
-	-	-	-	-
-	-	-	-	-
-	-	-	-	?

Machine Learning Techniques / Approaches

- KNN, linear regression, logistic regression
- Neural networks
- Genetic algorithms / Evolutionary algorithms
- Bayesian decision theory
- Decision Trees
- Hidden Markov Models
- Kernel methods – (support vector machines)
- Clustering methods
- Graphical methods
- Ensemble methods (train and combine multiple predictors to obtain more accurate results - begging and boosting) - Random Forest
- Deep Learning (Auto Encoder, Deep Neural Network, CNN, RNN)

Terminologies

- Y: Response(Statistics) / Target(Machine Learning)
- $X_1, X_2, \ldots, X_p$: Predictors(Statistics) / Features(Machine Learning) / Attributes (Database)
- $X_1, X_2, \ldots, X_p$ and Y : Sample / Example / Pattern(Machine Learning)
- x_i is a column vector containing feature values of i^{th} sample.

$$x_i = \begin{pmatrix} x_{i1} \\ x_{i2} \\ \vdots \\ x_{ip} \end{pmatrix}, i = 1, 2, \ldots, N.$$ Here, $x_i \in R^p$

- The value of X_j (j^{th} feature) in the set of samples are denoted as x_{ij}, where x_{ij} is the value of j^{th} feature in the i^{th} sample i.e. in x_{ij}, i represents the sample number and j represent the feature number.

$$i = 1, 2, \ldots, N \text{ and } j = 1, 2, \ldots, p$$

- **Note**: $x_i \in R^p$, $y_i \in R$, $x_{ij} \in R$

M-Class Classification Problem: Definition

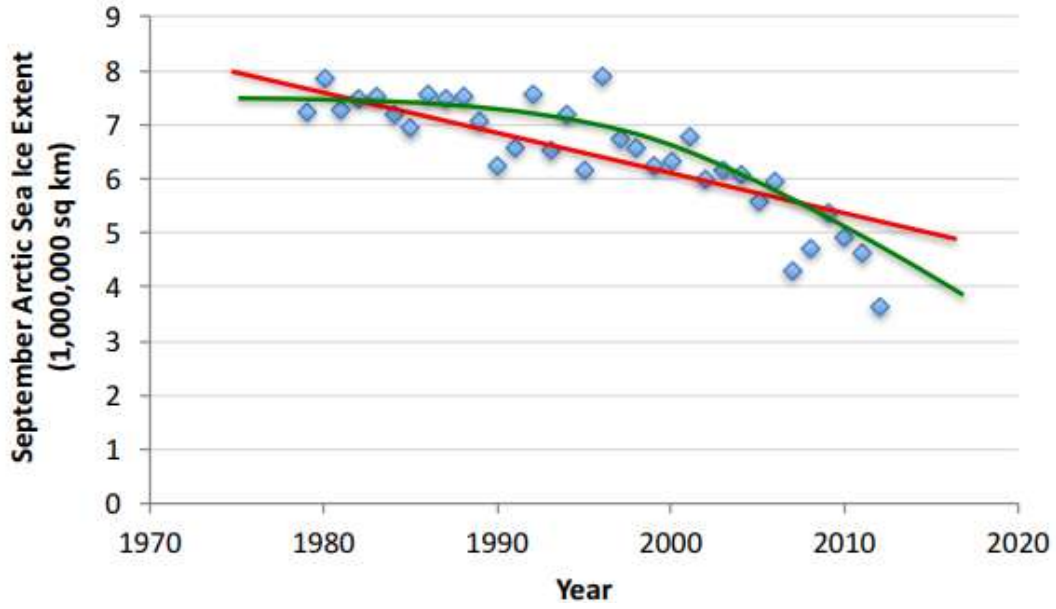
- Given a data set $D = \{(x_i, y_i)\}, i = 1, 2, 3, \ldots, N$ where $x_i \in R^p$, $y_i \in \{1, 2, \ldots, M\}$ and a new sample $x_0 \in R^p$.
- Determine the class to which x_0 is associated with based on D.
- Input: D, x_0
- Output: y_0

Regression Problem: Definition

- Given a data set $D = \{(x_i, y_i)\}, i = 1, 2, 3, \ldots, N$ of N patterns where $x_i \in R^p$, $y_i \in R$ and a new predictor vector $x_0 \in R^p$.
- Determine the response $y_0 \in R$ is associated with the pattern x_0 based on D.
- Input: D, x_0
- Output: y_0

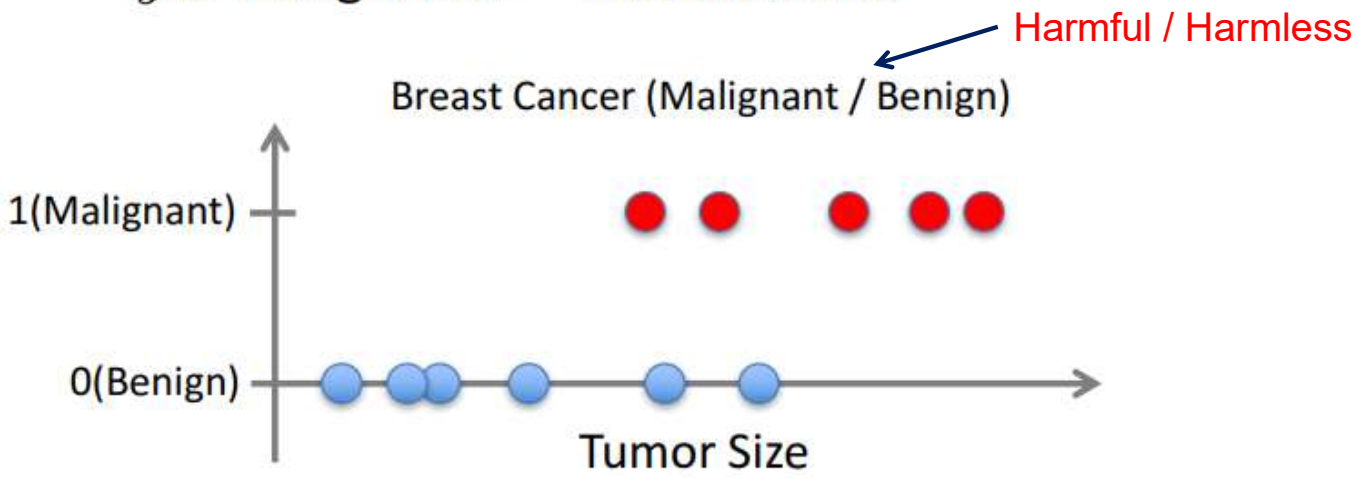
Supervised Learning: Regression

- Given $(x_1, y_1), (x_2, y_2), \ldots, (x_n, y_n)$
- Learn a function $f(x)$ to predict y given x
  - y is real-valued == regression



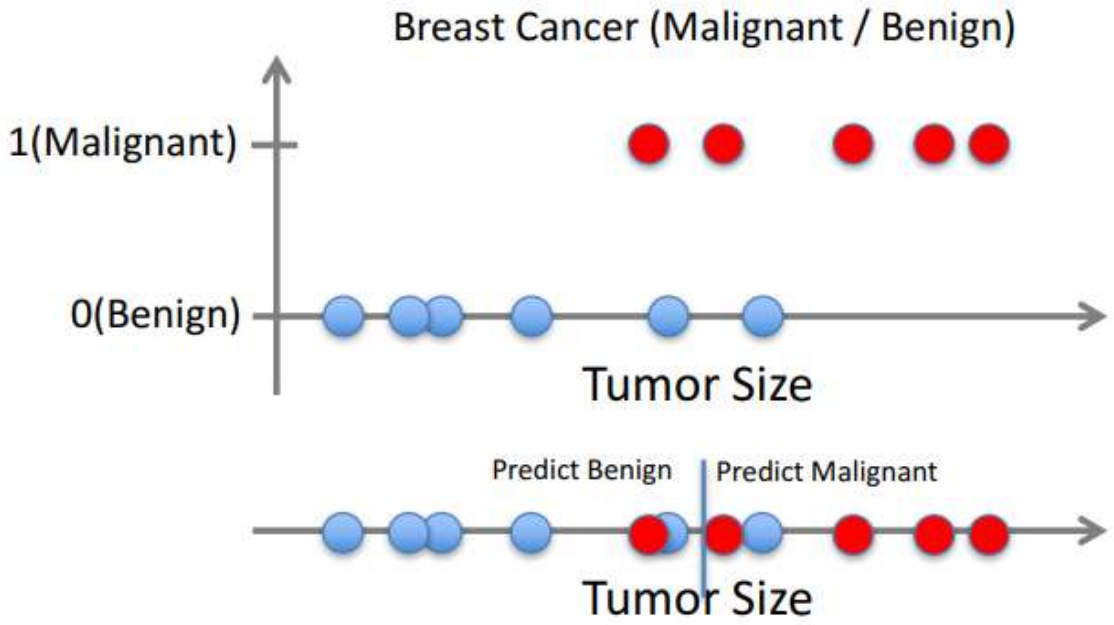
Supervised Learning: Classification

- Given $(x_1, y_1), (x_2, y_2), \ldots, (x_n, y_n)$
- Learn a function $f(x)$ to predict y given x
  - y is categorical == classification



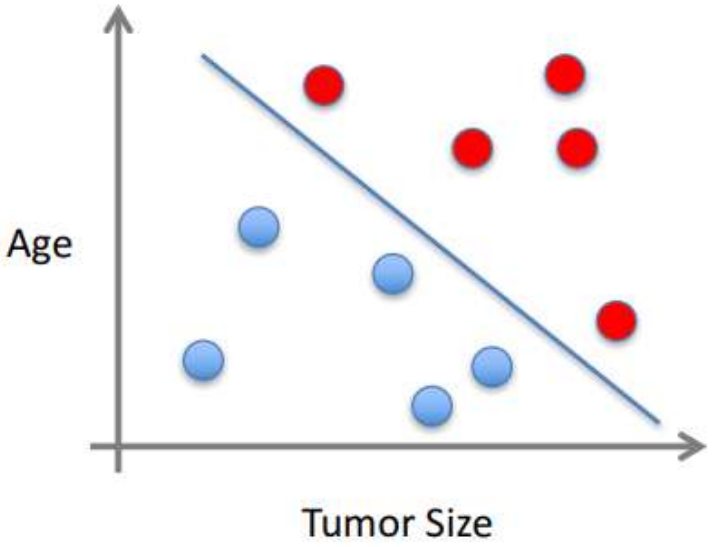
Supervised Learning: Classification

- Given $(x_1, y_1), (x_2, y_2), \ldots, (x_n, y_n)$
- Learn a function $f(x)$ to predict y given x
  - y is categorical == classification



Supervised Learning: Classification

- x can be multi-dimensional
  - Each dimension corresponds to an attribute



- Clump Thickness
- Uniformity of Cell Size
- Uniformity of Cell Shape
- ...

Unsupervised Learning

- Given $x_1, x_2, \ldots, x_n$ (without labels)
- Output hidden structure behind the x's
  - E.g., clustering