Machine Learning
1UCEC701

Kavita Bathe

Assistant Professor

K J Somaiya Institute of Engineering and Information Technology

Outline

- Introduction to Machine learning
- Types of Machine Learning
- Issues in Machine Learning
- Application of Machine Learning
- Steps in developing a Machine Learning Application

Maths

- **Matrix**

$$A = \begin{bmatrix} a_{1,1} & a_{1,2} & a_{1,3} \\ a_{2,1} & a_{2,2} & a_{2,3} \end{bmatrix}$$

- **Types**
- **Row Matrix**
- **Column Matrix**
- **Square**
- **Identity Matrix**
- Its symbol is the capital letter **I**

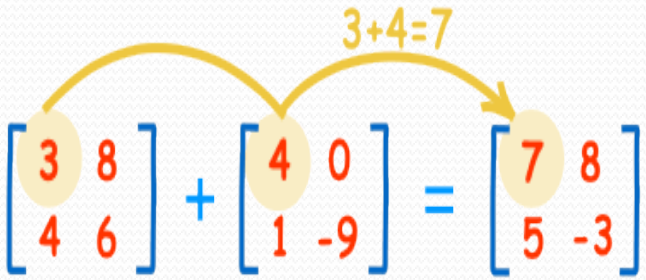
$A \times I = A$

$I \times A = A$

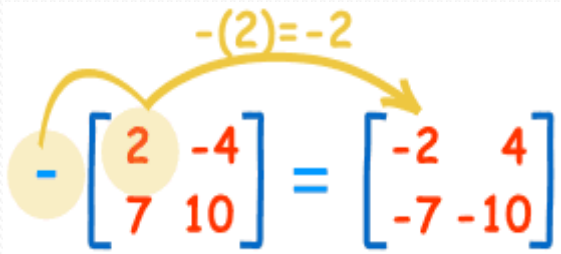
- **Diagonal Matrix**
- **Scalar Matrix**
- **Triangular Matrix**
- **Lower triangular , Upper triangular**
- **Zero Matrix (Null Matrix)**
- **Transpose**

Matrix Operations

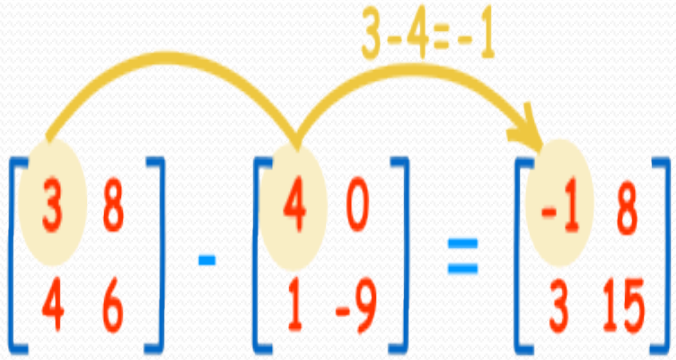
- **Addition**



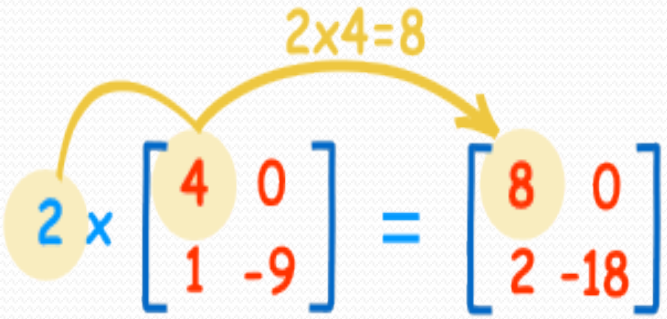
- **Negative**



- **Subtracting**

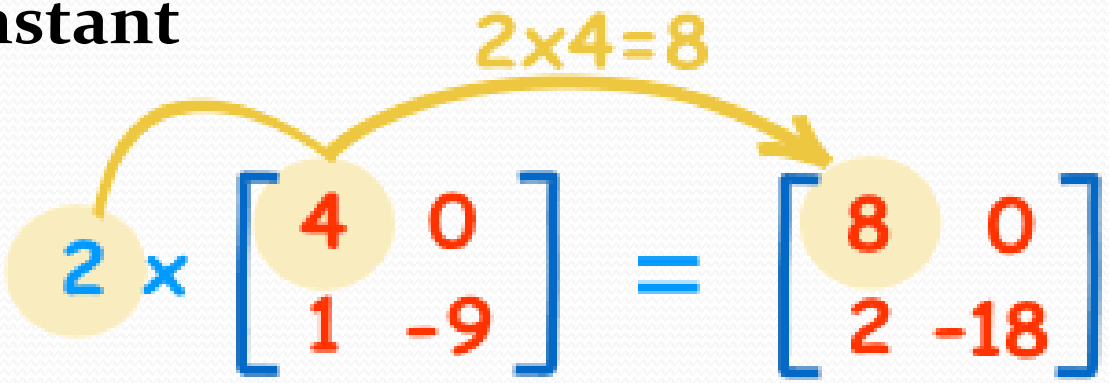


- Multiplication



$$\begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix} \times \begin{bmatrix} 7 & 8 \\ 9 & 10 \\ 11 & 12 \end{bmatrix} = \begin{bmatrix} 58 & 64 \\ 139 & 154 \end{bmatrix} \checkmark$$

- **Multiply by a Constant**



- **Transposing**

$$\begin{bmatrix} 6 & 4 & 24 \\ 1 & -9 & 8 \end{bmatrix}^T = \begin{bmatrix} 6 & 1 \\ 4 & -9 \\ 24 & 8 \end{bmatrix}$$

- **Determinant of a Matrix**
- First of all the matrix must be **square**
- **For a 2×2 Matrix**
- $|A| = ad - bc$

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$$

Example:

$|B| = 4 \times 8 - 6 \times 3$

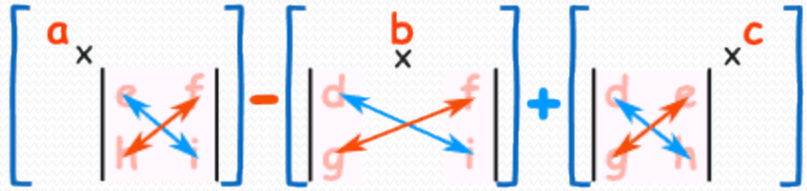
$= 32 - 18$

$= 14$

$$B = \begin{bmatrix} 4 & 6 \\ 3 & 8 \end{bmatrix}$$

- **For a 3×3 Matrix**

$$A = \begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix}$$



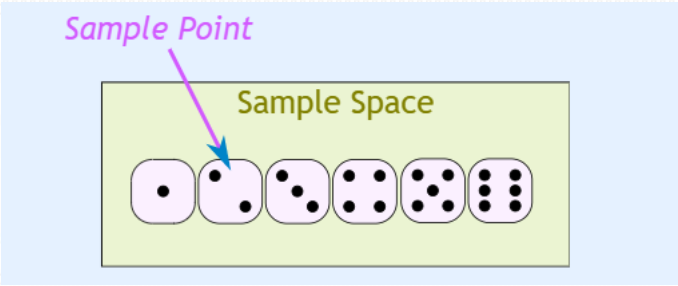
$$|A| = a(ei - fh) - b(di - fg) + c(dh - eg)$$

$$C = \begin{bmatrix} 6 & 1 & 1 \\ 4 & -2 & 5 \\ 2 & 8 & 7 \end{bmatrix}$$

$$|C| = 6\times(-2\times7 - 5\times8) - 1\times(4\times7 - 5\times2) + 1\times(4\times8 - (-2\times2))$$
$$= 6\times(-54) - 1\times(18) + 1\times(36)$$
$$= \mathbf{-306}$$

Probability

- What is probability?
- Tossing a Coin
- Throwing Dice

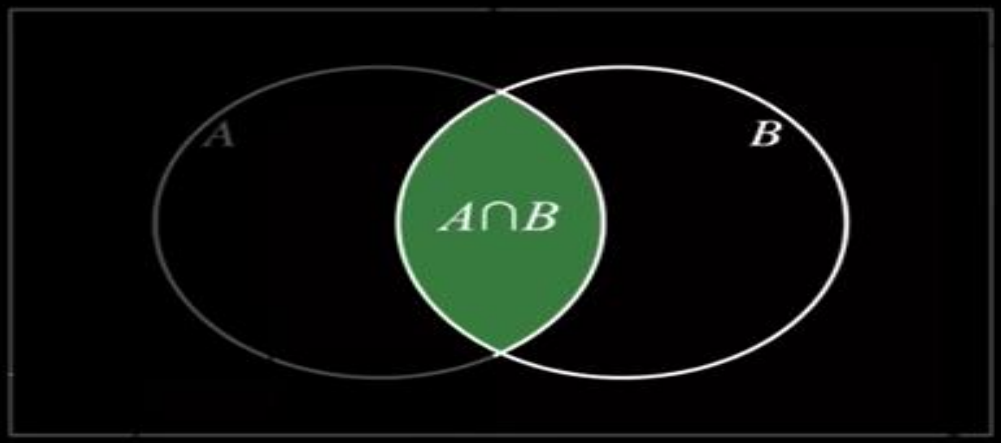


$$\text{Probability of an event happening} = \frac{\text{Number of ways it can happen}}{\text{Total number of outcomes}}$$

Example: the chances of rolling a "4" with a die

Conditional Probability Definition

$$P(A|B) = \frac{P(A \cap B)}{P(B)} \quad \text{for } P(B) > 0$$



$$S = \{1, 2, 3, 4, 5, 6\}$$

Back to rolling a die. Consider the two events:

$A = \{1, 2, \underline{3, 4, 5}\}$

$B = \{\underline{3, 4, 5}, 6\}$

$$P(A|B) = \frac{3}{4}$$

What is the conditional probability of A, given B?

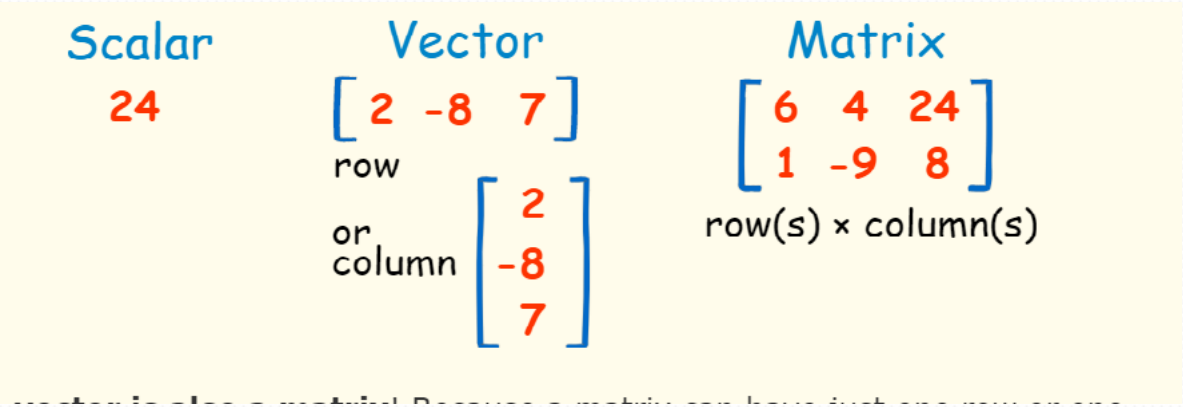
$$A \cap B = \{3, 4, 5\}$$

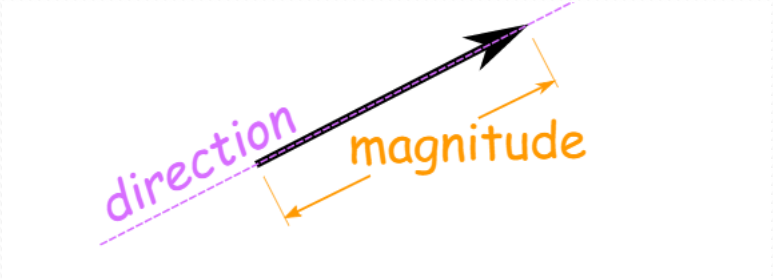
$$P(A|B) = \frac{P(A \cap B)}{P(B)} = \frac{3/6}{4/6} = \frac{3}{4}$$

$$P(A \cap B) = \frac{3}{6}$$

Scalars, Vectors and Matrices

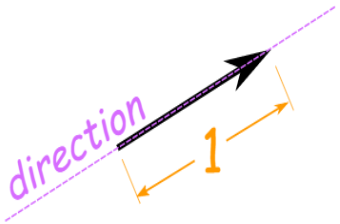
- A **scalar** is a number, like **3, -5, 0.368, etc**,
- A **vector** is a **list** of numbers (can be in a row or column),
- A **matrix** is an **array** of numbers (one or more rows, one or more columns).

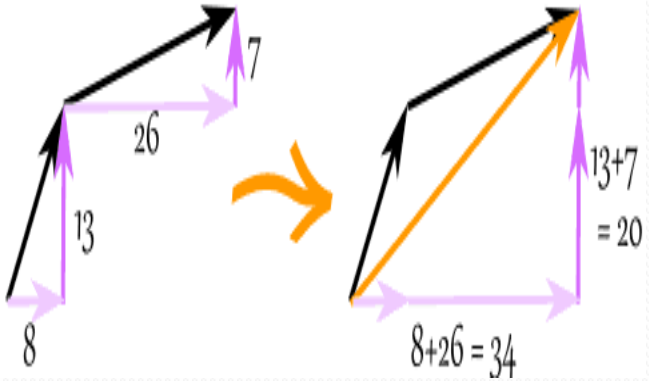




Unit Vector

A **Unit Vector** has a magnitude of **1**:





Example: add the vectors **a** = (8, 13) and **b** = (26, 7)

c = **a** + **b**

c = (8, 13) + (26, 7) = (8+26, 13+7) = (34, 20)

Magnitude of a Vector

- $|\mathbf{a}|$
- $|\mathbf{a}| = \sqrt{(x^2 + y^2)}$
- Example: what is the magnitude of the vector **b** = (6, 8) ?
- $|\mathbf{b}| = \sqrt{(6^2 + 8^2)} = \sqrt{(36+64)} = \sqrt{100} = 10$

opulation Covariance Formula

$$Cov(x,y) = \frac{\Sigma(x_i-\bar{x})(y_i-\bar{y})}{N}$$

ample Covariance

$$Cov(x,y) = \frac{\Sigma(x_i-\bar{x})(y_i-y)}{N-1}$$

Notations in Covariance Formulas

- x_i = data value of x
- y_i = data value of y
- $\bar{x}$ = mean of x
- $\bar{y}$ = mean of y
- N = number of data values.

Relation Between Correlation Coefficient and Covariance Formulas

Correlation=$Cov(x,y)\sigma x*\sigma y$

Here, Cov (x,y) is the covariance between x and y while σ_x and σ_y are the standard deviations of x and y.

- **Question:** The table below describes the rate of economic growth (xi) and the rate of return on the S&P 500 (y_i). Using the covariance formula, determine whether economic growth and S&P 500 returns have a positive or inverse relationship. Before you compute the covariance, calculate the mean of x and y.

Economic Growth % (x_i)	S&P 500 Returns % (y_i)
2.1	8
2.5	12
4.0	14
3.6	10

- **Solution:**
- $\bar{x} = \sum xin$
- $\bar{x} = 2.1+2.5+4+3.64$
- $\bar{x} = 12.24$
- $\bar{x} = 3.1$
- $\bar{y} = \sum xin$
- $\bar{y} = 8+12+14+104$
- $\bar{y} = 444$
- $\bar{y} = 11$

growth and S&P 500 returns.

xi	yi	xi − x̄	yi − ȳ
2.1	8	-1	-3
2.5	12	-0.6	1
4.0	14	0.9	3
3.6	10	0.5	-1

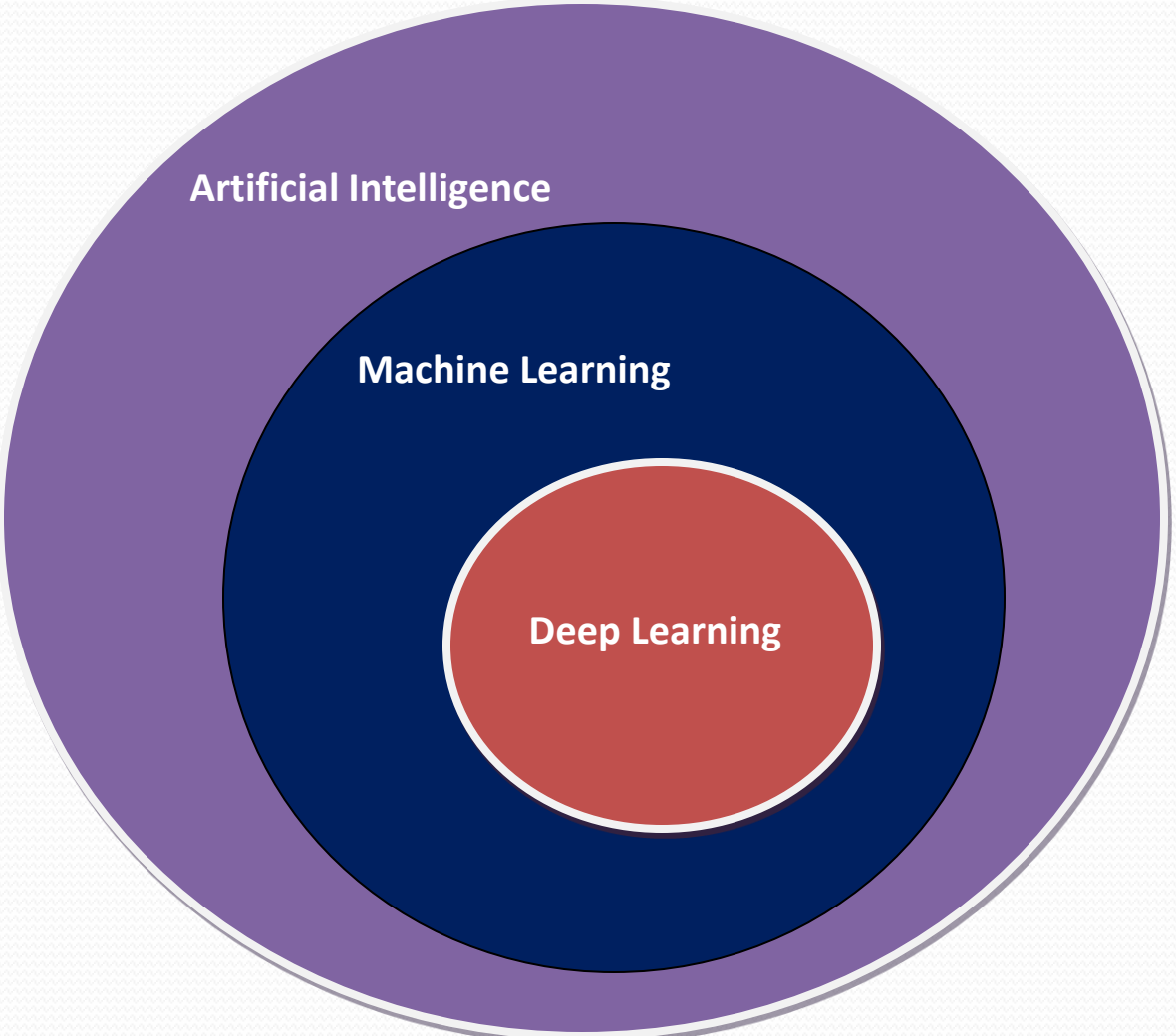
$$Cov(x,y) = \frac{(-1)(-3)+(-0.6)1+(0.9)3+(0.5)(-1)}{4-1} = \frac{3-0.6+2.7-0.5}{3} = \frac{4.6}{3} = 1.533$$

Mean, Median, Mode, and Range

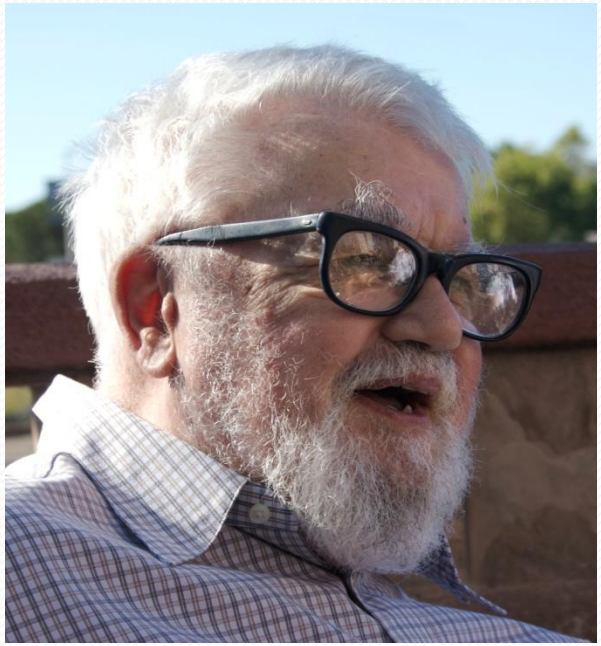
- ***Find the mean, median, mode, and range for the following list of values:***
- **13, 18, 13, 14, 13, 16, 14, 21, 13**
- Mean? 15
- Median:13, 13, 13, 13, 14, 14, 16, 18, 21 :14
- Mode:13
- range is 21 – 13 = 8.
- ***Find the mean, median, mode, and range for the following list of values:***
- ***1, 2, 4, 7*** **mean: 3.5**
  median: 3
  mode: none
  range: 6

- ***Find the mean, median, mode, and range for the following list of values:***
- ***8, 9, 10, 10, 10, 11, 11, 11, 12, 13***
- **mean: 10.5**
  median: 10.5
  modes: 10 and 11
  range: 5
- **Outliers**
- Outliers are values that "**lie out**side" the other values.
- They can change the mean a lot, so we can either not use them (and say so) or use the median or mode instead.
- Example: 3, 4, 4, 5 and 104
- (3+4+4+5+104) / 5 = **24**
- Without the 104 the mean is: (3+4+4+5) / 4 = **4**
- But please tell people you are not including the outlier.

Introduction to Machine Learning



Artificial Intelligence

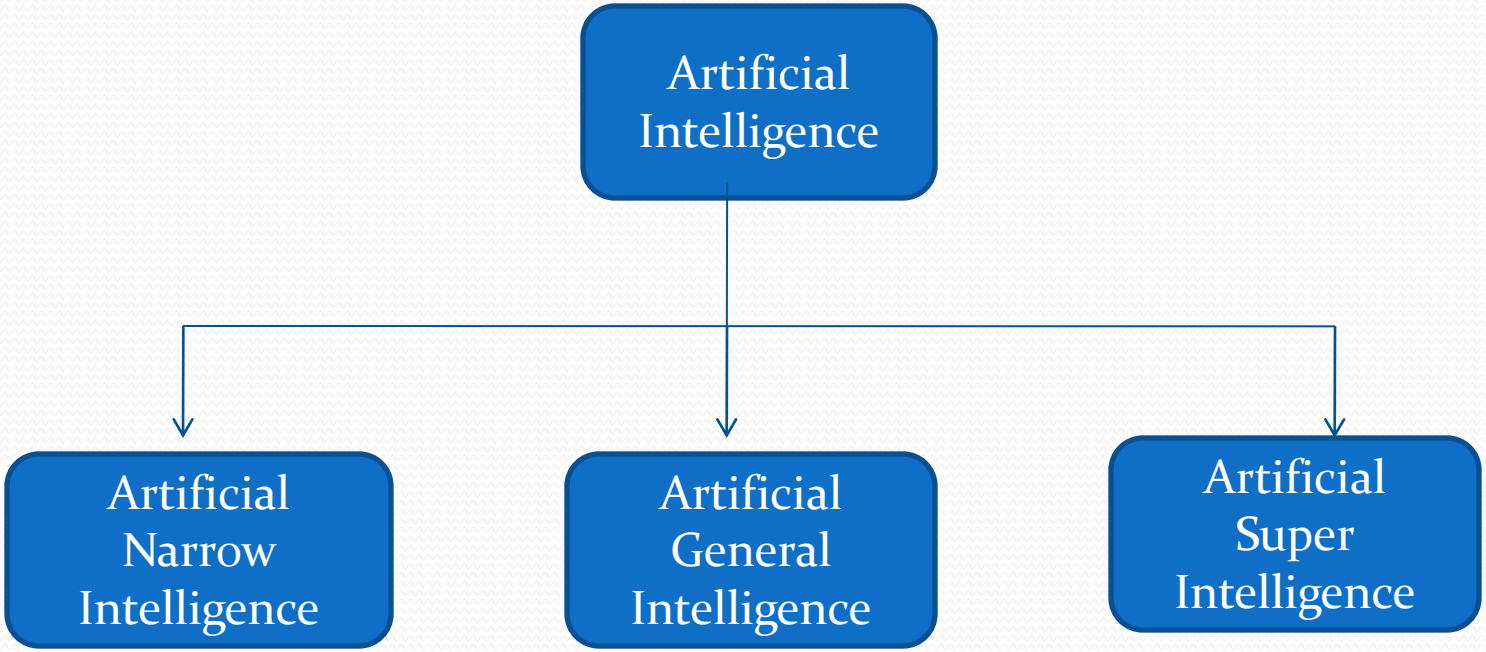


John McCarthy

Image Source:Wikipedia

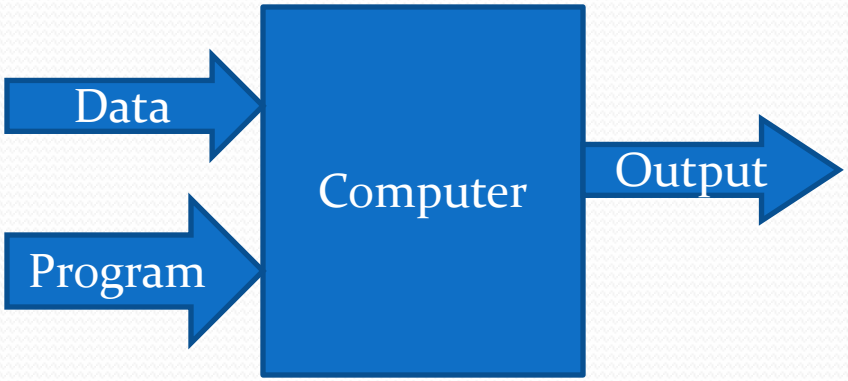
Artificial Intelligence is the science and engineering of making intelligent machines, especially intelligent computer programs

Artificial Intelligence

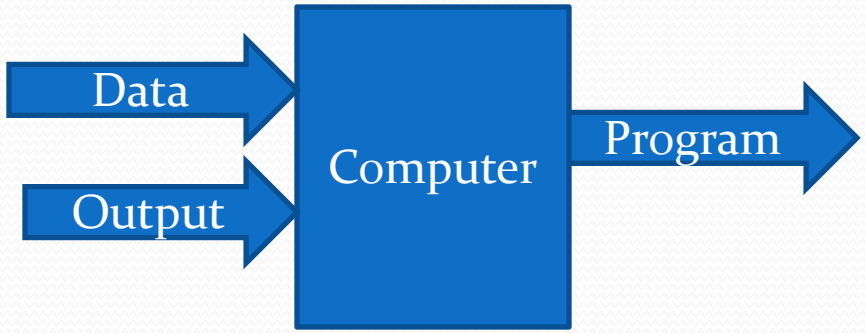


Machine Learning

Traditional Programming



Machine Learning



Machine Learning

- A computer program is said to learn from experience E with respect to some class of tasks T and performance measure P, if its performance at tasks in T, as measured by P, improves with experience E.

-Tom Mitchell

1.Predicting price of house

2. Weather prediction

3. CAT Recognition

Machine Learning

Experience

CAT Model

CAT Recognition (Task)

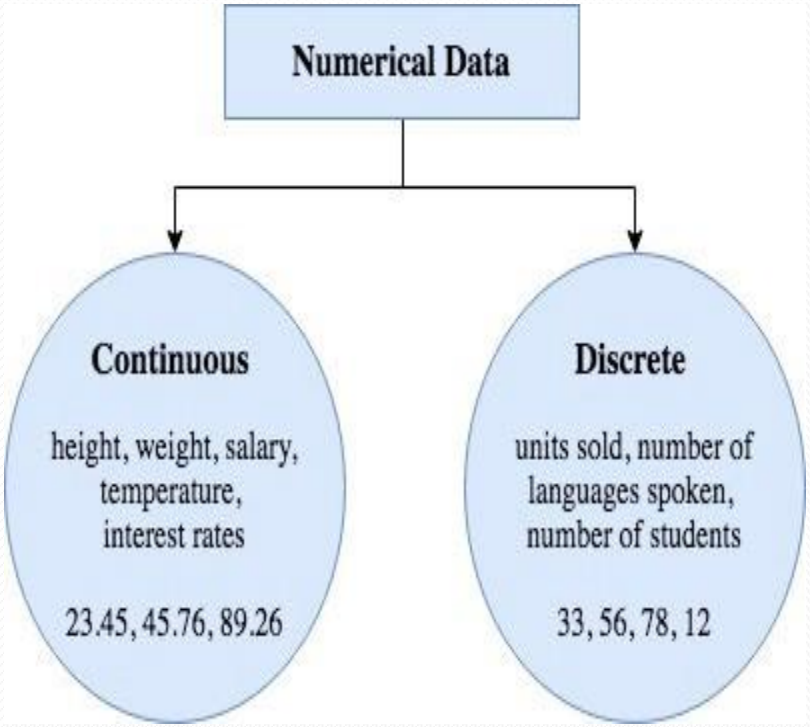
CAT :90% (Performance)

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. In this setting, what is E?
- What is T,P,E?

Learning Perspective With Examples

Numerical Data	**Categorical Data**
Time Series Data	**Text**

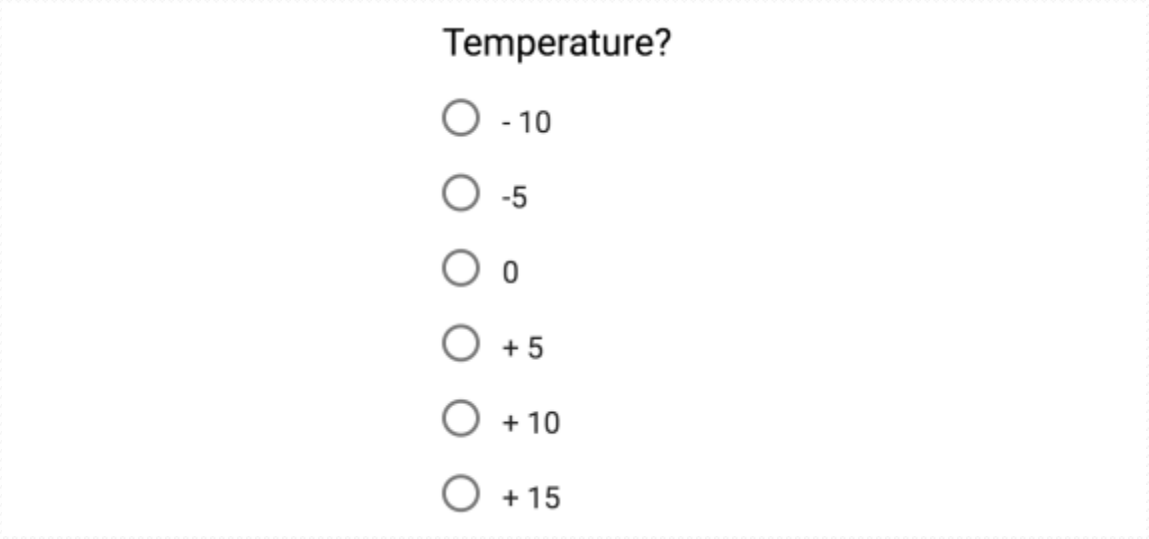


Numerical Data

- Discrete Data
- You can check by asking the following two questions whether you are dealing with discrete data or not:
- Can you count it?
- Can it be divided up into smaller and smaller parts?
- Continuous Data
- Continuous Data represents measurements and therefore their values can't be counted but they can be measured. Ex.height of a person

Continuous Data

- **Interval Data**
- Interval values represent ordered units that have the same difference

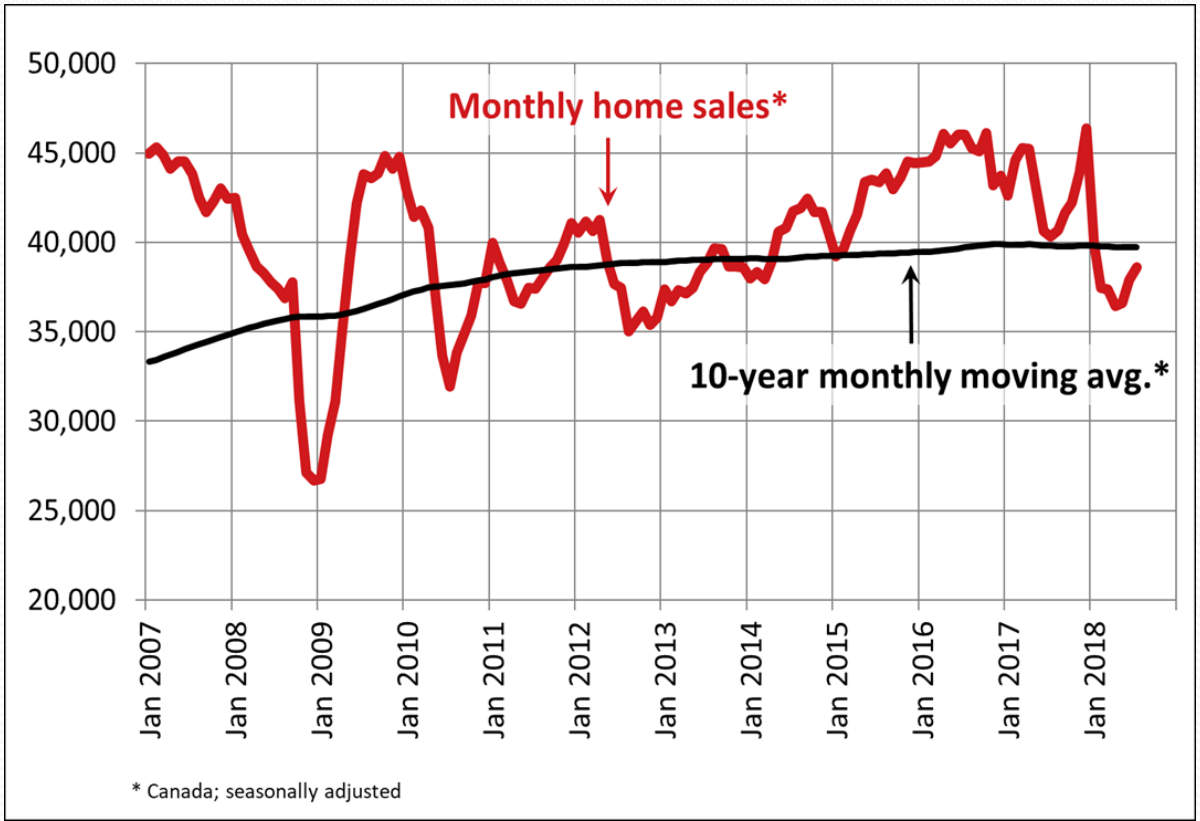


Categorical Data

- **Nominal Data**
- Nominal values represent discrete units and are used to label variables, that have no quantitative value.
- Ex.What is your Gender?
- What languages do you speak?
- **Ordinal Data**
- Ordinal values represent discrete and ordered units. It is therefore nearly the same as nominal data, except that it's ordering matters.
- Ex.What is your educational background?

Time Series Data

- Time series data is a sequence of numbers collected at regular intervals over some period of time

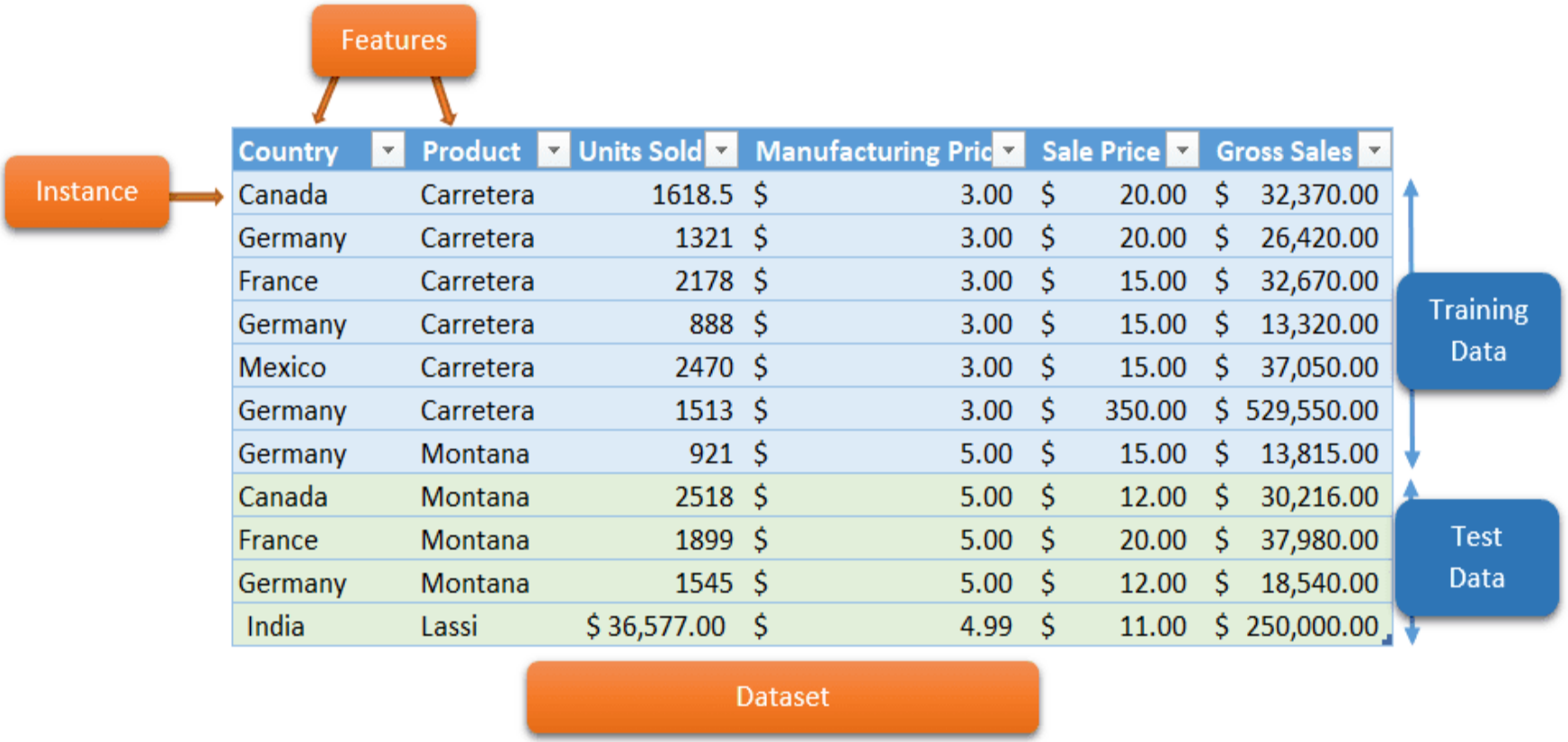


Text

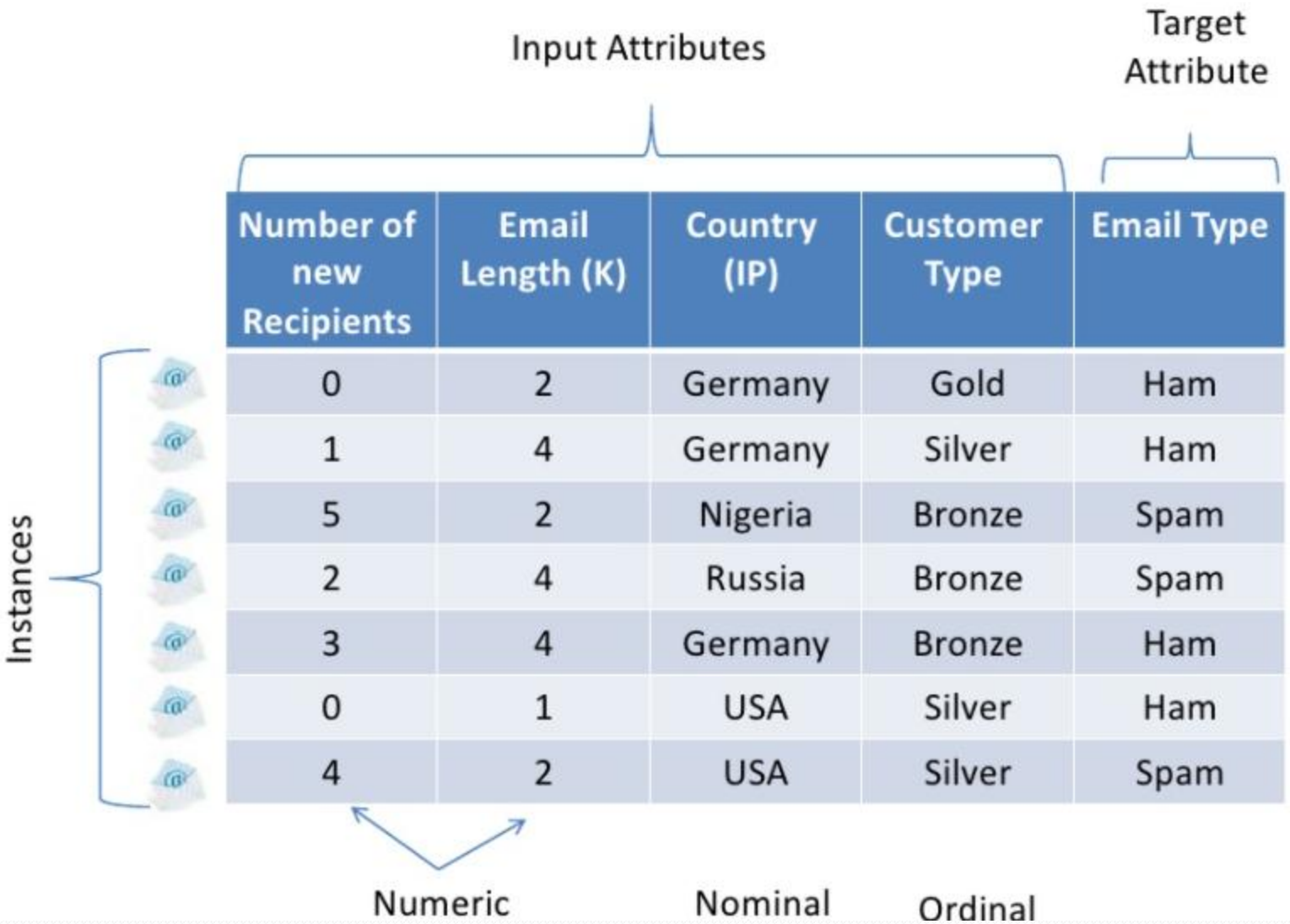
- Text data is basically just words. A lot of the time the first thing that you do with text is you turn it into numbers using some interesting functions like the bag of words formulation.
-

Terminologies in Machine Learning

- **Dataset**
- **Instance/Samples/Observations/Indivduals**
- **Features/Variables/Attributes/Field**
- **Target**
- **Dimension**
- **Size**
- **prediction**
- **model**
- **Training Dataset**
- **Testing Dataset**


Features
Instance
Country
Product
Units Sold
Manufacturing Pric
Sale Price
Gross Sales
Canada Carretera 1618.5 $ 3.00 $ 20.00 $ 32,370.00
Germany Carretera 1321 $ 3.00 $ 20.00 $ 26,420.00
France Carretera 2178 $ 3.00 $ 15.00 $ 32,670.00
Germany Carretera 888 $ 3.00 $ 15.00 $ 13,320.00
Mexico Carretera 2470 $ 3.00 $ 15.00 $ 37,050.00
Germany Carretera 1513 $ 3.00 $ 350.00 $ 529,550.00
Germany Montana 921 $ 5.00 $ 15.00 $ 13,815.00
Canada Montana 2518 $ 5.00 $ 12.00 $ 30,216.00
France Montana 1899 $ 5.00 $ 20.00 $ 37,980.00
Germany Montana 1545 $ 5.00 $ 12.00 $ 18,540.00
India Lassi $ 36,577.00 $ 4.99 $ 11.00 $ 250,000.00
Training Data
Test Data
Dataset

Data Set

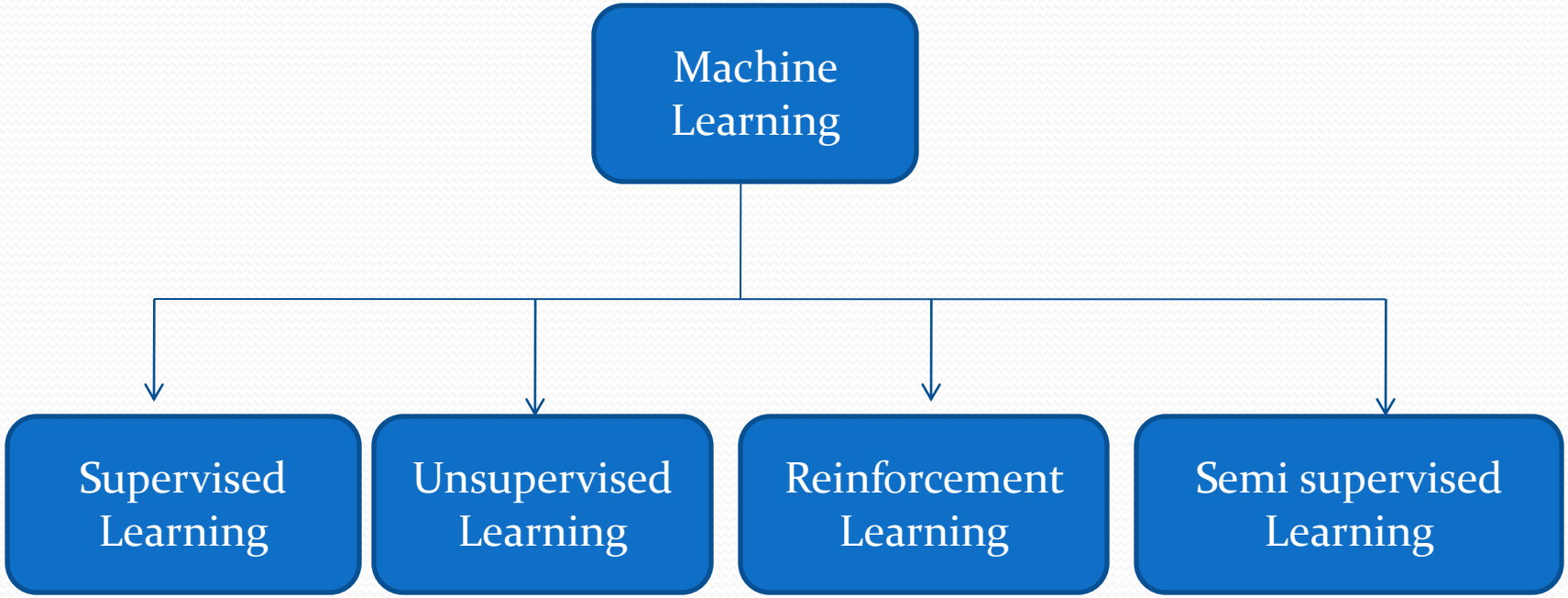


Number of new Recipients	Email Length (K)	Country (IP)	Customer Type	Email Type
0	2	Germany	Gold	Ham
1	4	Germany	Silver	Ham
5	2	Nigeria	Bronze	Spam
2	4	Russia	Bronze	Spam
3	4	Germany	Bronze	Ham
0	1	USA	Silver	Ham
4	2	USA	Silver	Spam

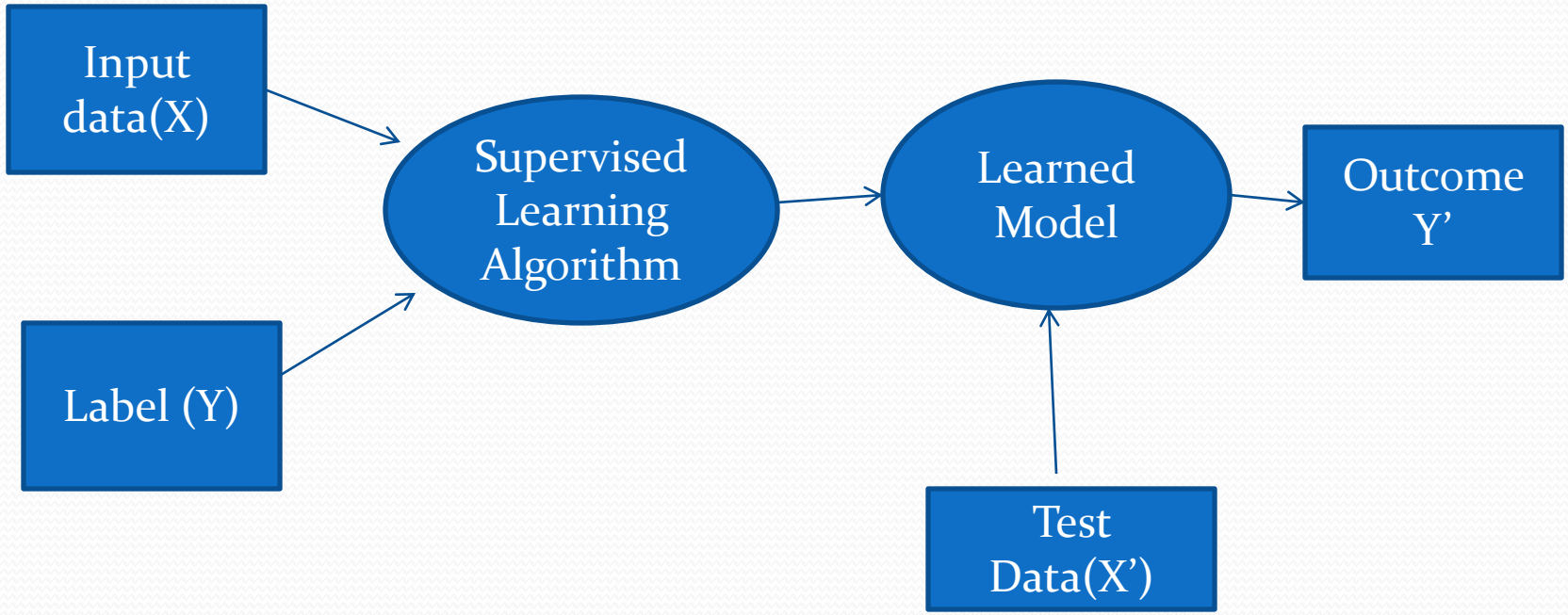
- **Representation: Feature Engineering**
- The process of creating features from raw data is called feature engineering.

- **Classification**
- **Regression**
- **Clustering**
- **Classification Accuracy**
- **Confusion Matrix**

Machine Learning Tasks



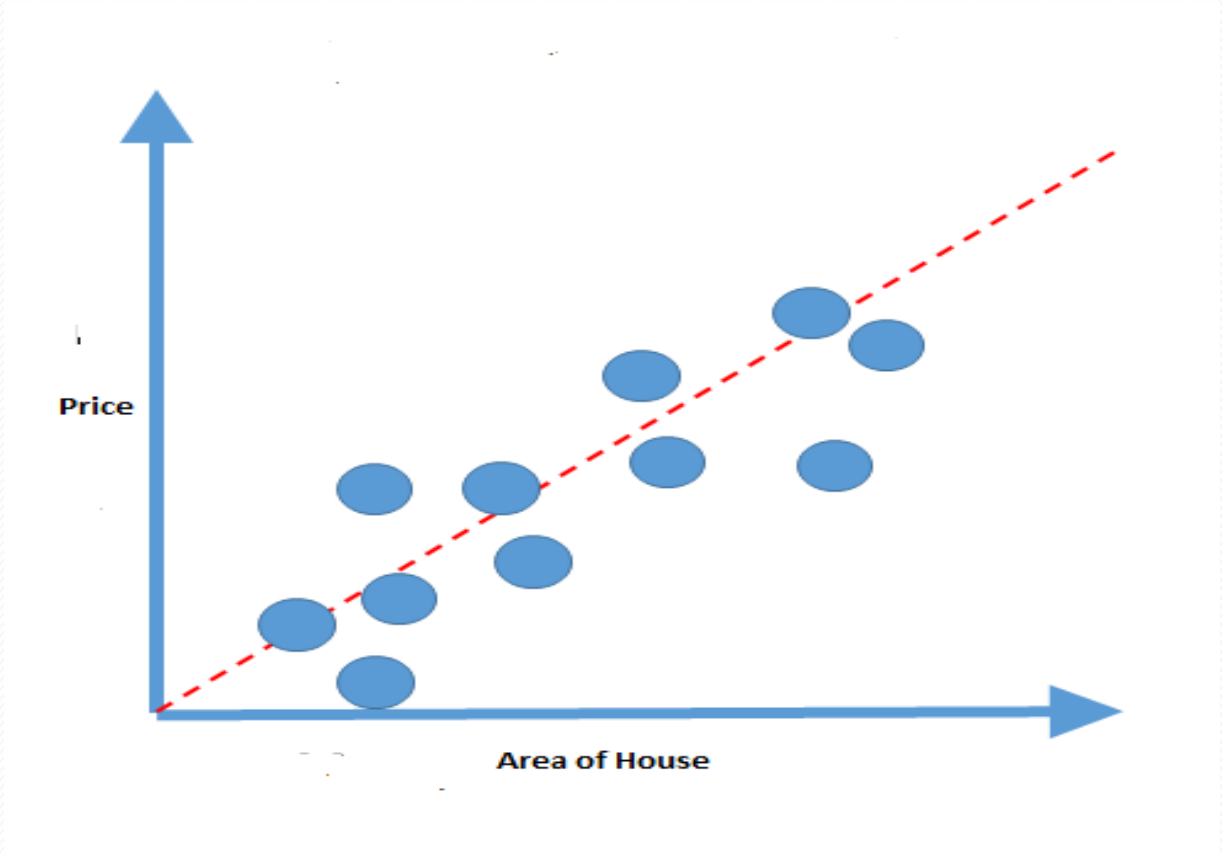
Supervised Learning



Training Data

Regression

- Predicting price of a house



Classification

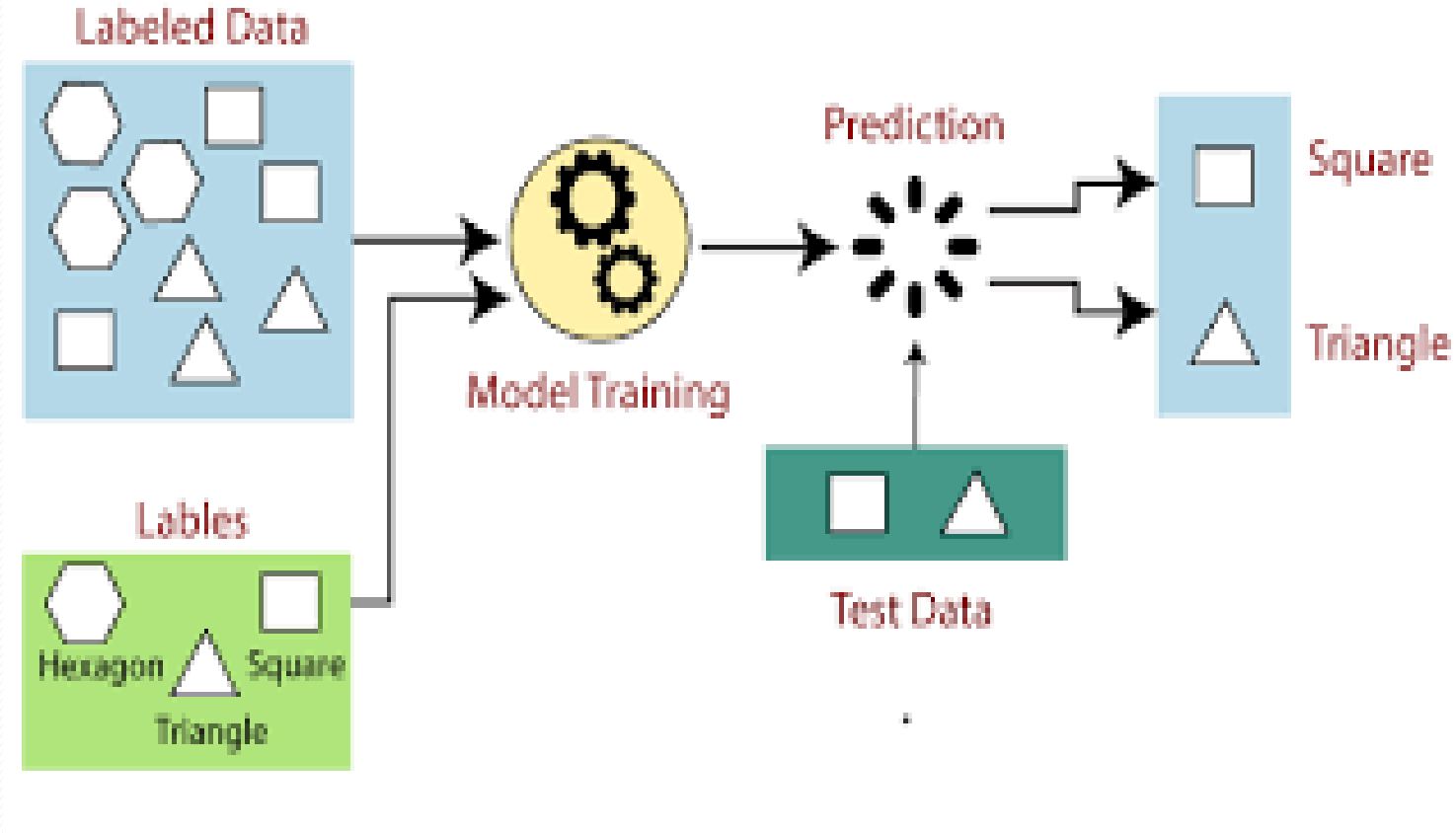
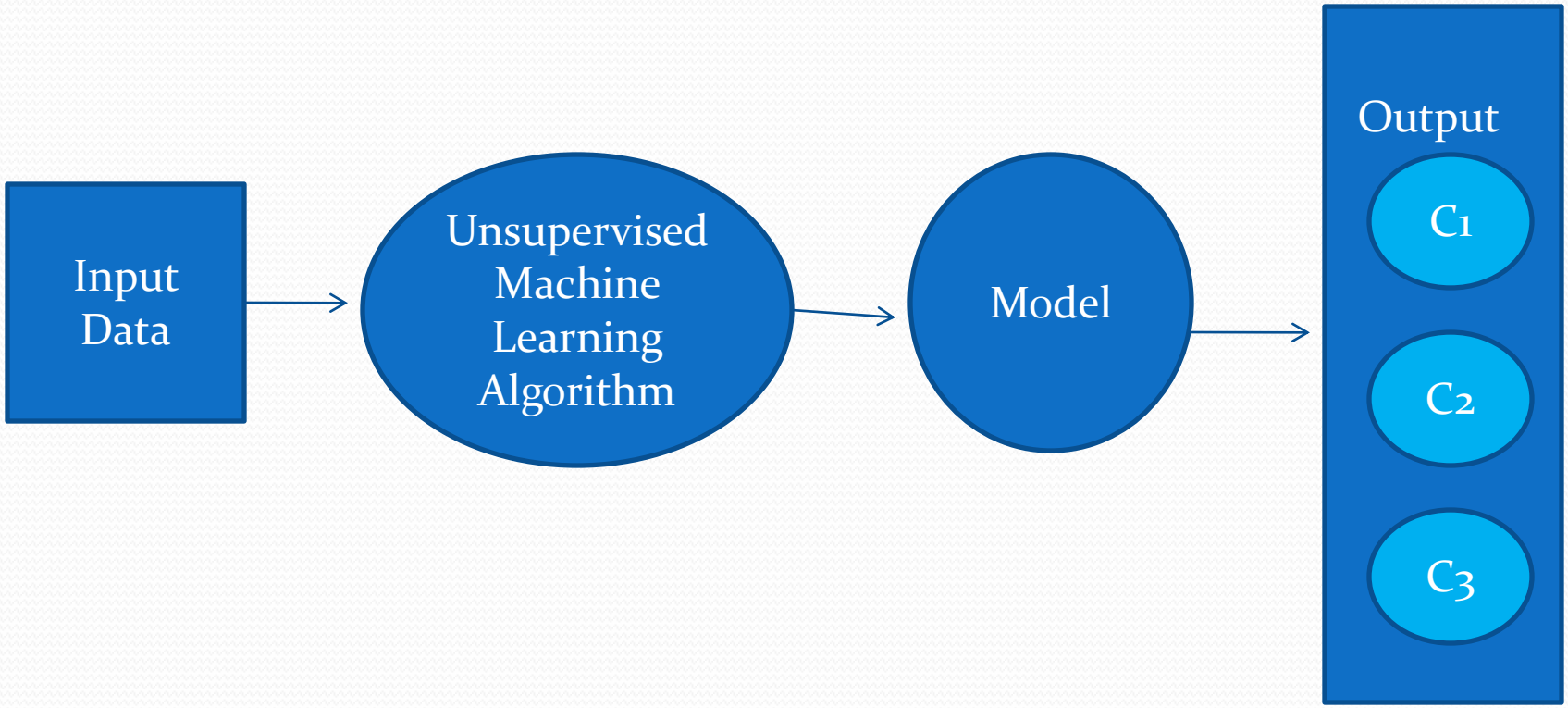


Image Source: Java Point

Unsupervised Machine Learning



Unsupervised Learning

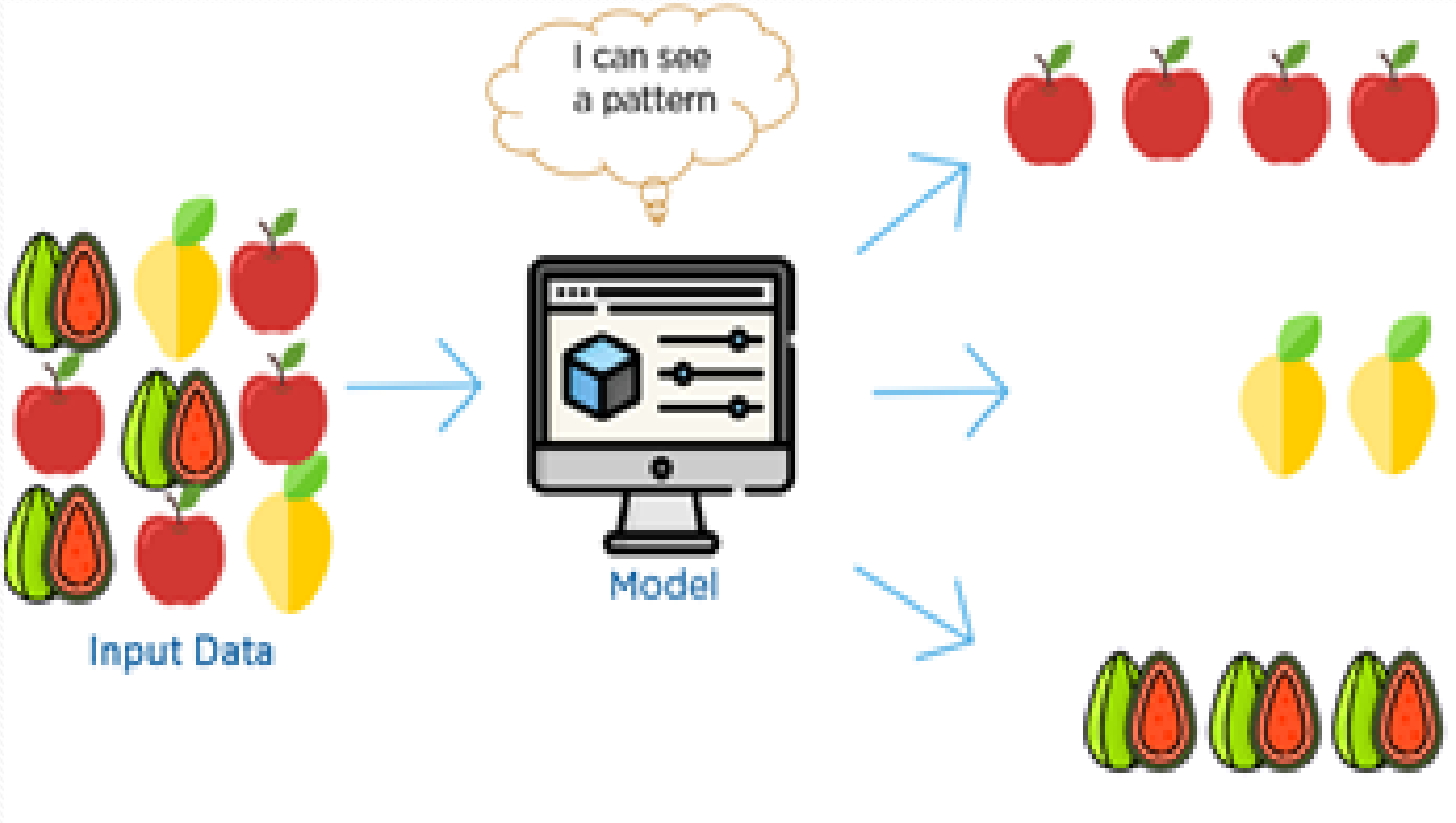
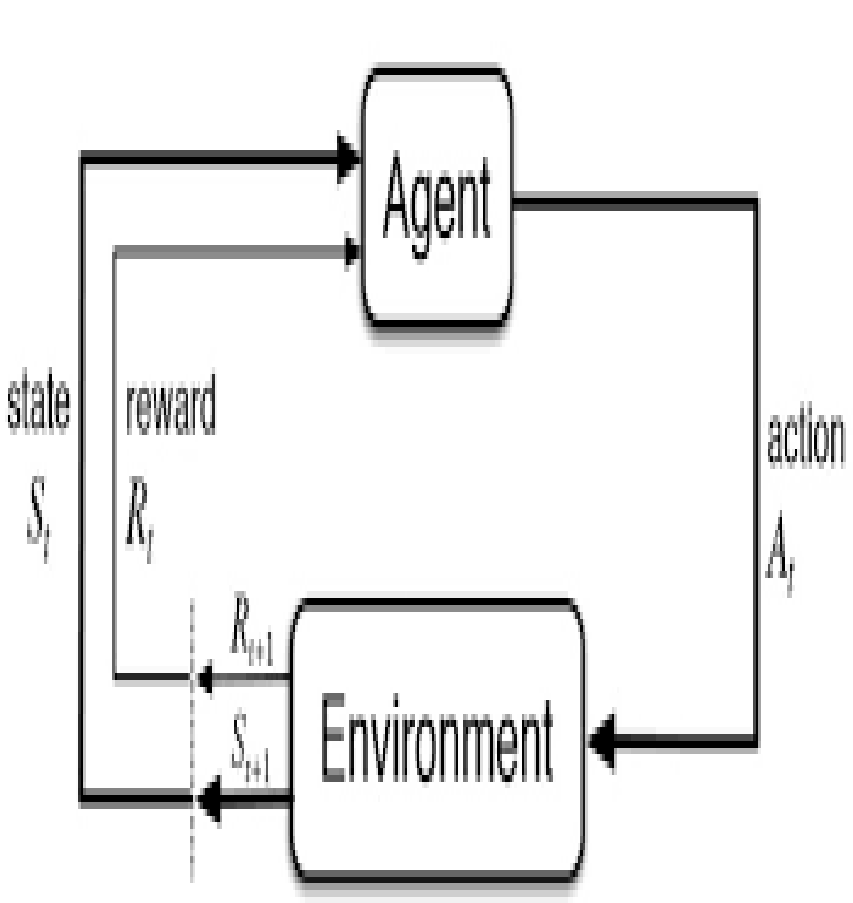


Image Source:Medium.com

Reinforcement Learning

Reinforcement Learning(RL) is a type of machine learning technique that enables an agent to learn in an interactive environment by trial and error using feedback from its own actions and experiences.

Reinforcement Learning



- Baby
- Living Room
- Ice crème
- State

Semi supervised Learning

- Combination of supervised and unsupervised learning
- Pseudo Labeling

	Machine learning	Deep learning
Training dataset	Small	Large
Choose features	Yes	No
Number of algorithms	Many	Few
Training time	Short	Long

Machine Learning Applications

1.Virtual Assistant

Google Assistant, Amazon's Alexa

Microsoft's Cortana,Samsung's Bixby

2. Transportation

Google Maps-Traffic Alerts (Maps)

Dynamic Pricing in Travel

3. Videos Surveillance

4.Social Media

Facebook

Machine Learning Applications

5.Education

-Predict Grade of a student

-MOM generator

-Examination Process

6. Agriculture

-disease detection

-weather forecasting

-Agriculture Bots

7. Web Services

-Spam filtering

-Google Search

-Google Translate

8. Sales and Marketing

-Product Recommendations

-Online Customer Support-Chatbots

9.Healthcare

Anomaly detection using X-ray images

COVID19

10. Automatic email response

11.Tesla

How to approach Machine Learning Problem

- Find the Problem statement

  -Institutions

  -COVID 19 (Ex.Bots,Predict Disease spread)

  -Look at problems of post COVID situation in your organization

  -Client Company

- Understand the problem statement

How to approach Machine Learning Problem-Cont

- Data acquisition
  -public datasets available
  -Client provides the data
  -Figure out the data from websites
  -Sometimes generate the data
- Data preparation
  -Data Preprocessing
  -Feature engineering /Data Transformation
    1.Scaling-height/weight
    2.Decomposing-DateTime
    3.Agrigation

How to approach Machine Learning Problem-Cont

- Choose the algorithm

  -Visualize the data
- Training the model
- Evaluation
- Improving Model -Parameter Tuning
- Deployment of the system

How to choose the right algorithm

- How can you choose which one to use?
- Goal
- Type of data
- Issues in Machine Learning

Technology Stack

- Machine Learning Deep Learning Libraries
    1. Scikitlearn
    2. Tensor flow
    3. Keras
    4. pytorch
    5. Caffe2

Technology Stack

- Computer Vision

  -opencv

  -pillow

- Conversational AI:

  RASA

- NLP

  -NLTK

  -Std nlp libraries

Online Web Service Jupyter Notebook

- Google Colaboratory
- Microsoft azure
- Collaborative Calculation and Data(CoCalc)
- Binder
- Kaggle Karnel
- Datalore

Thank You