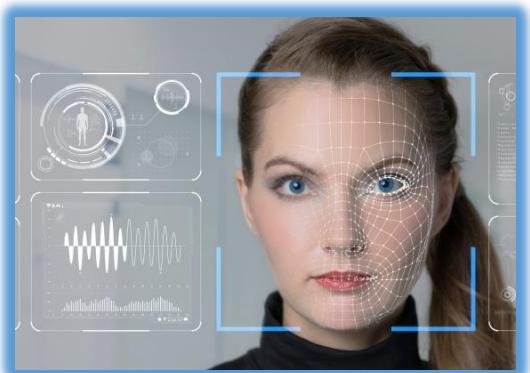
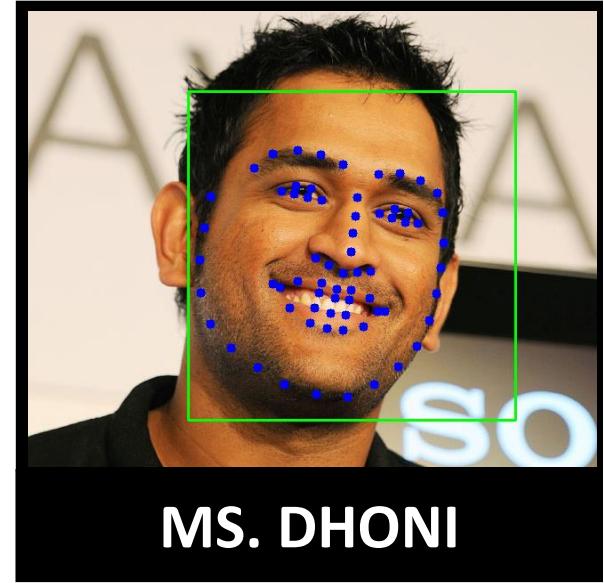


Face Recognition

Power of Data Science



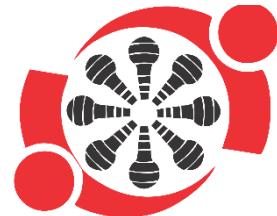
Who are we?



INNOMATICS
TECHNOLOGY HUB

Innovation.Automation.Analytics

- We are Data Analytics-driven Digital Transformation firm
- We provide:
 - ✓ Corporate Trainings
 - ✓ Meet ups
 - ✓ Webinars
 - ✓ College Workshops



Innomatics Technology Hub

Innovation. Automation. Analytics

Our Major Focus on:

✓ **DATA SCIENCE**

- ✓ MACHINE LEARNING
- ✓ DEEP LEARNING
- ✓ NATURAL LANGUAGE PROCESSING, IMAGE PROCESSING



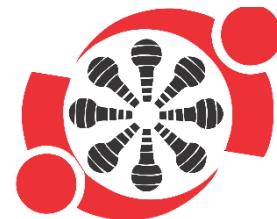
✓ **BIG DATA**

- ✓ HADOOP, SPARK, SCALE

✓ **CLOUD COMPUTING**

- ✓ AWS
- ✓ DEVOPS

✓ **PYTHON**



Innomatics Technology Hub

Innovation. Automation. Analytics

We are Education Partners of



KenMinds
...committed to solution excellence

KenMinds IT Solutions Pvt Ltd.

“Kenminds is a technology Consulting, Staffing and IT services company that helps Business to transform their Business applications from a legacy platform to the today highly effective, flexible and highly capable digital platforms.”



KenMinds IT Solution Pvt Ltd

- Our Products

- Docnizer
- Retail Pace
- Cerebral Palsy



Agenda

01

What is Data Science ?

02

Image Processing Basics

03

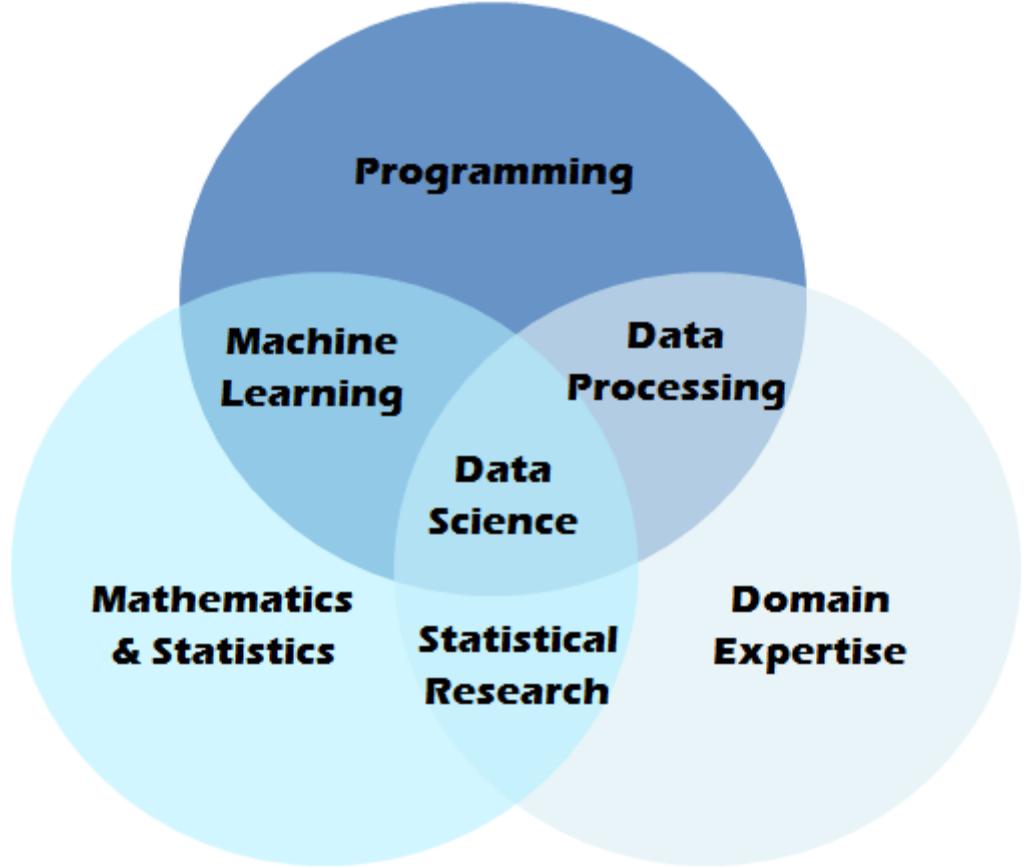
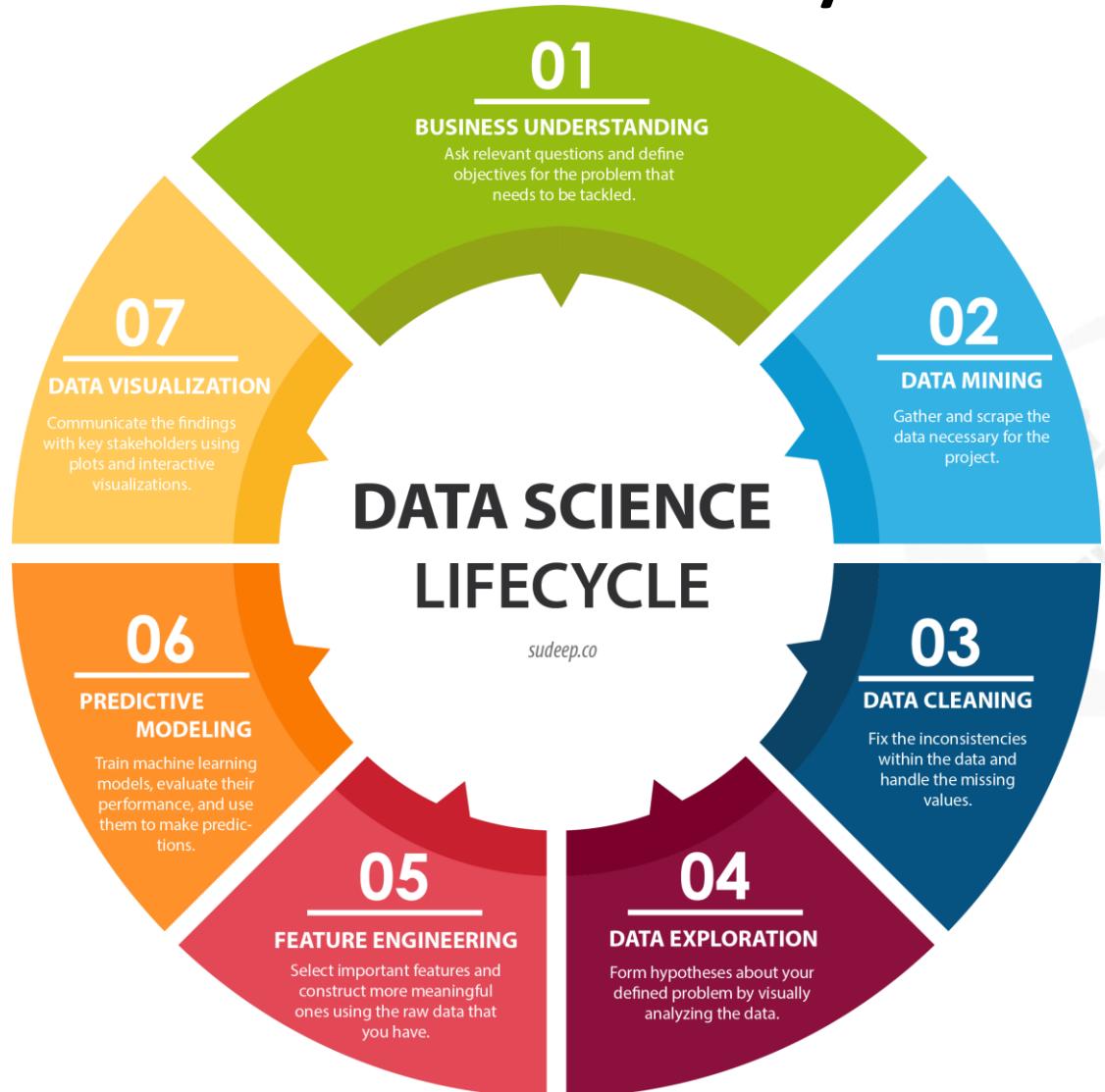
Build Machine Learning model for face recognition



Who is Data Scientist?

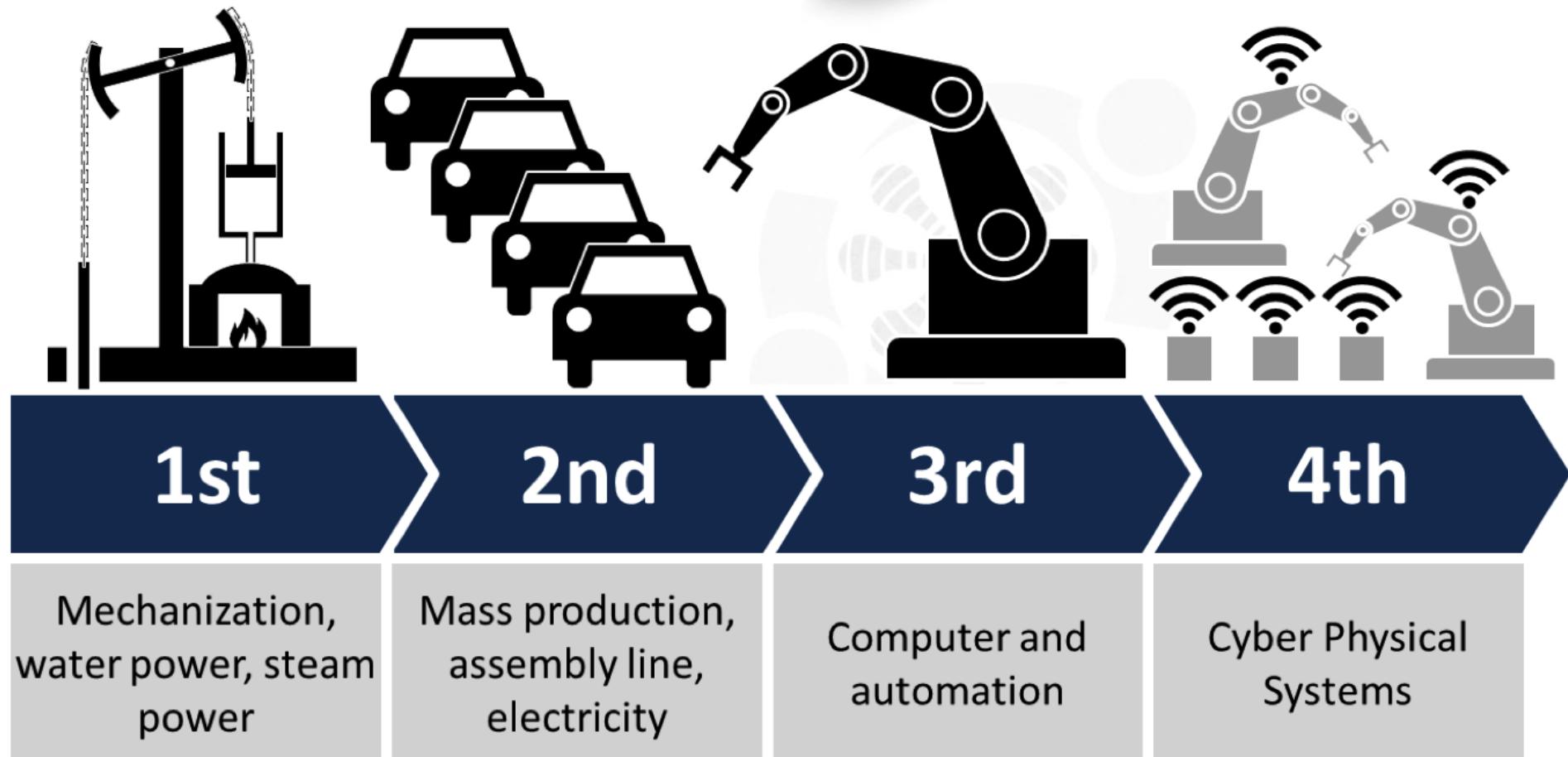
- *“Data scientists use their data and analytical ability to find and interpret rich data sources; manage large amounts of data despite hardware, software, and bandwidth constraints; merge data sources; ensure consistency of datasets; create visualizations to aid in understanding data; build mathematical models using the data; and present and communicate the data insights/findings.”*

Data Science Life Cycle

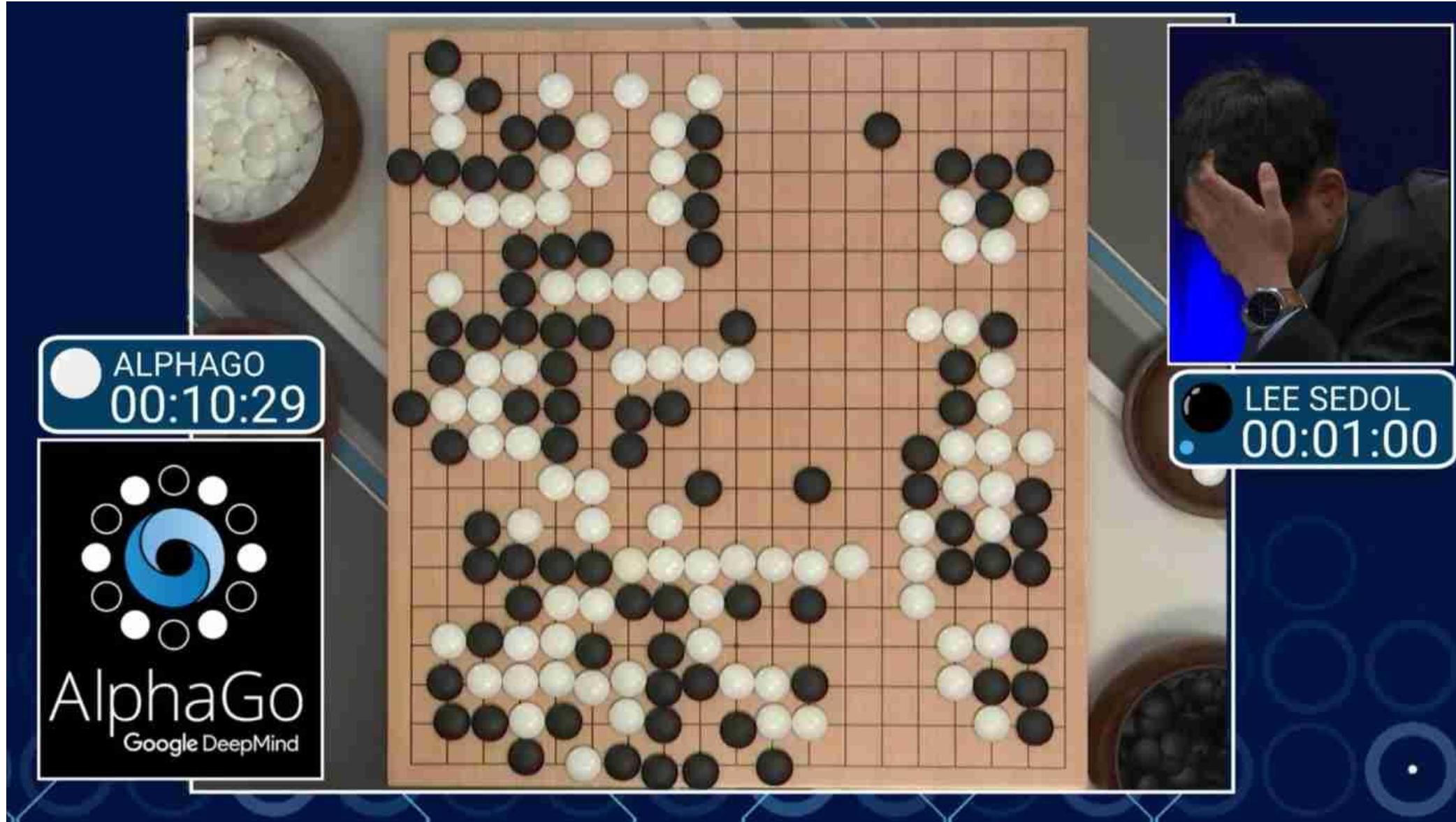




Data Science ?



"Machine Learning algorithm beats Go world Champion"

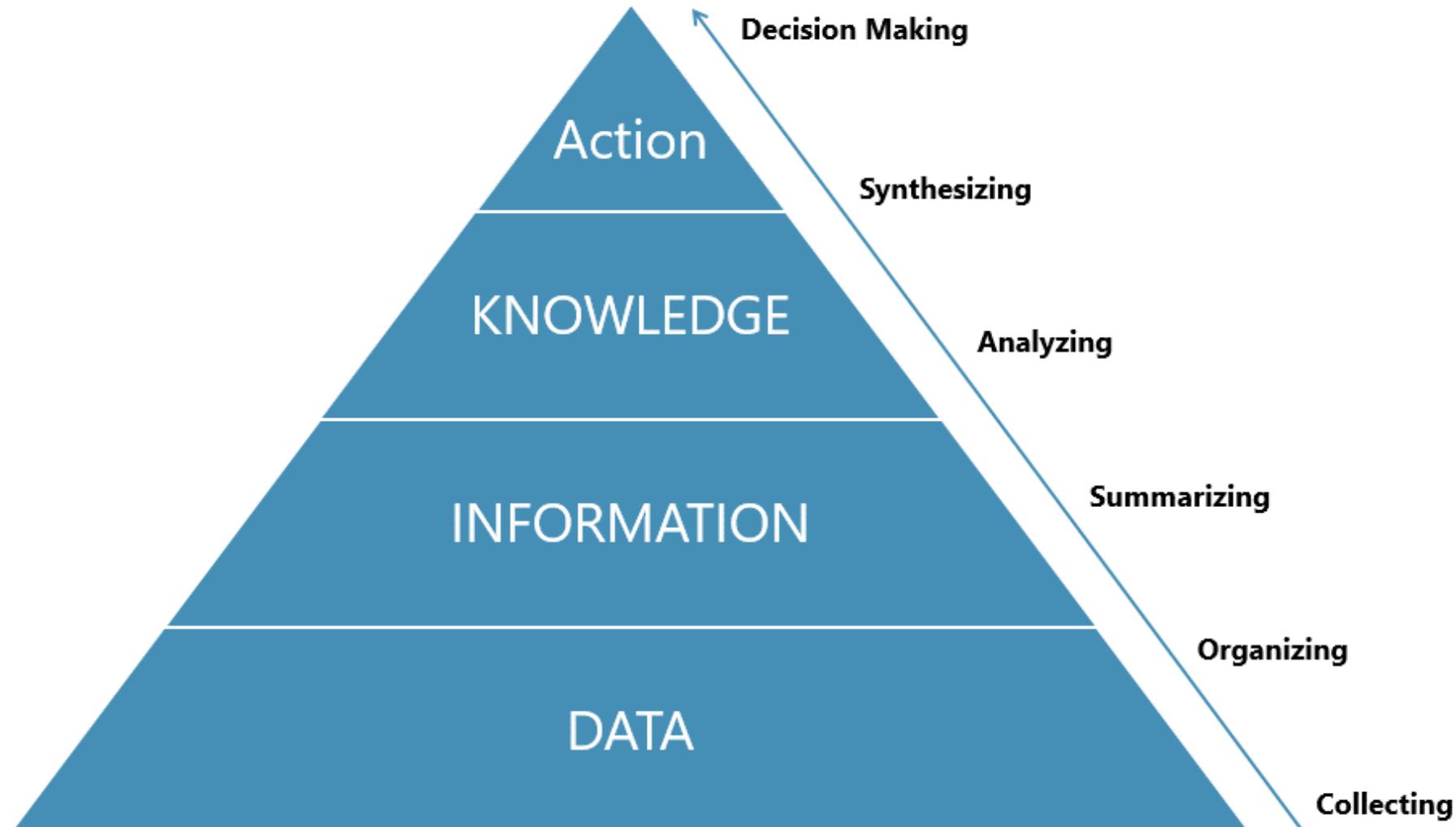




Data

In computing, data is **information** that has been translated into a form that is efficient for movement or processing. Relative to today's computers and transmission media, data is information converted into binary digital form. It is acceptable for data to be used as a singular subject or a plural subject. Raw data is a term used to describe data in its most basic digital format.

Learning from Data



Dealing with unstructured and structured data

Data Science

- Process of extracting knowledge from data

- Preparation
- Analyze
- Models





Raw Data



Preparation

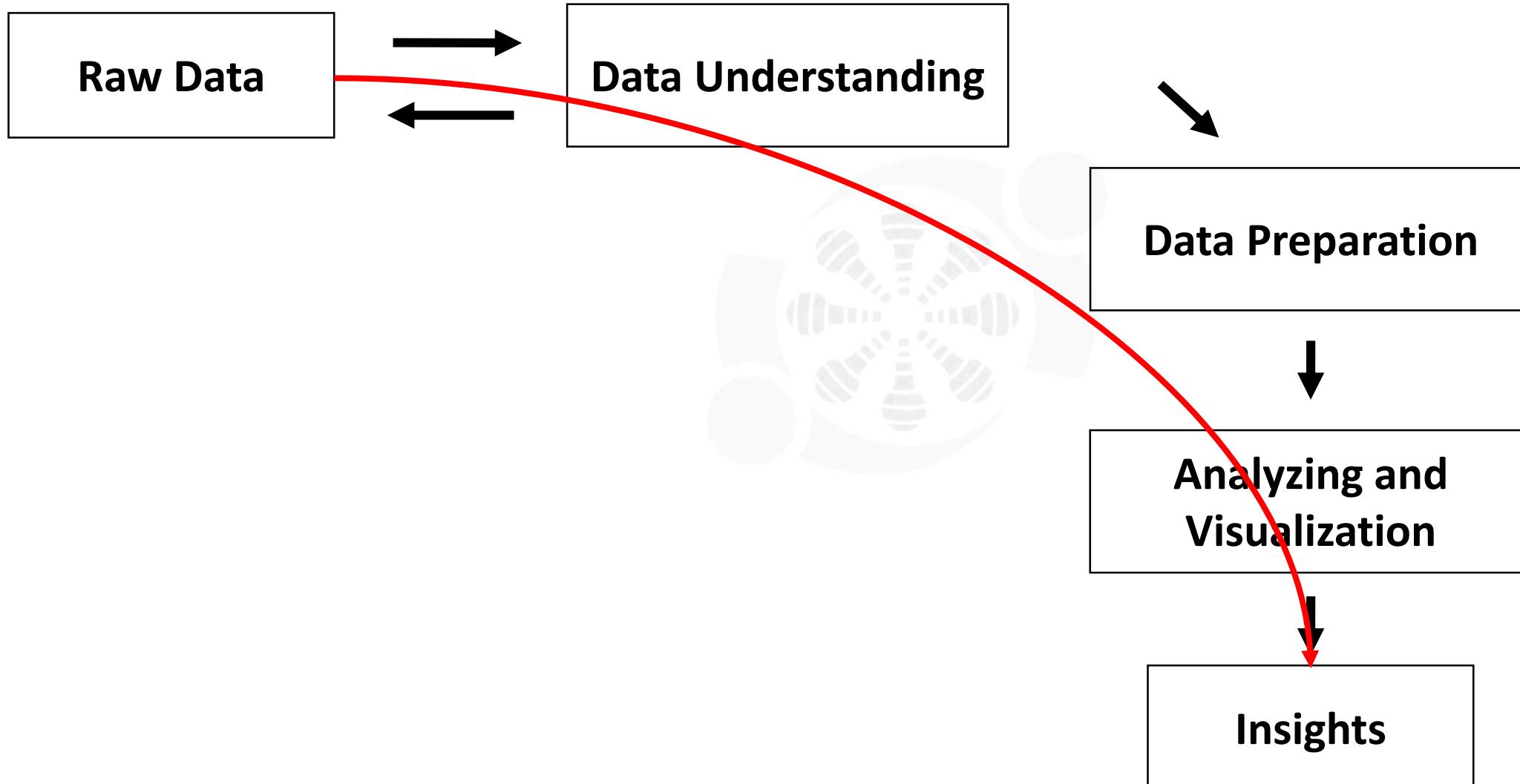


Analysis & Visualization

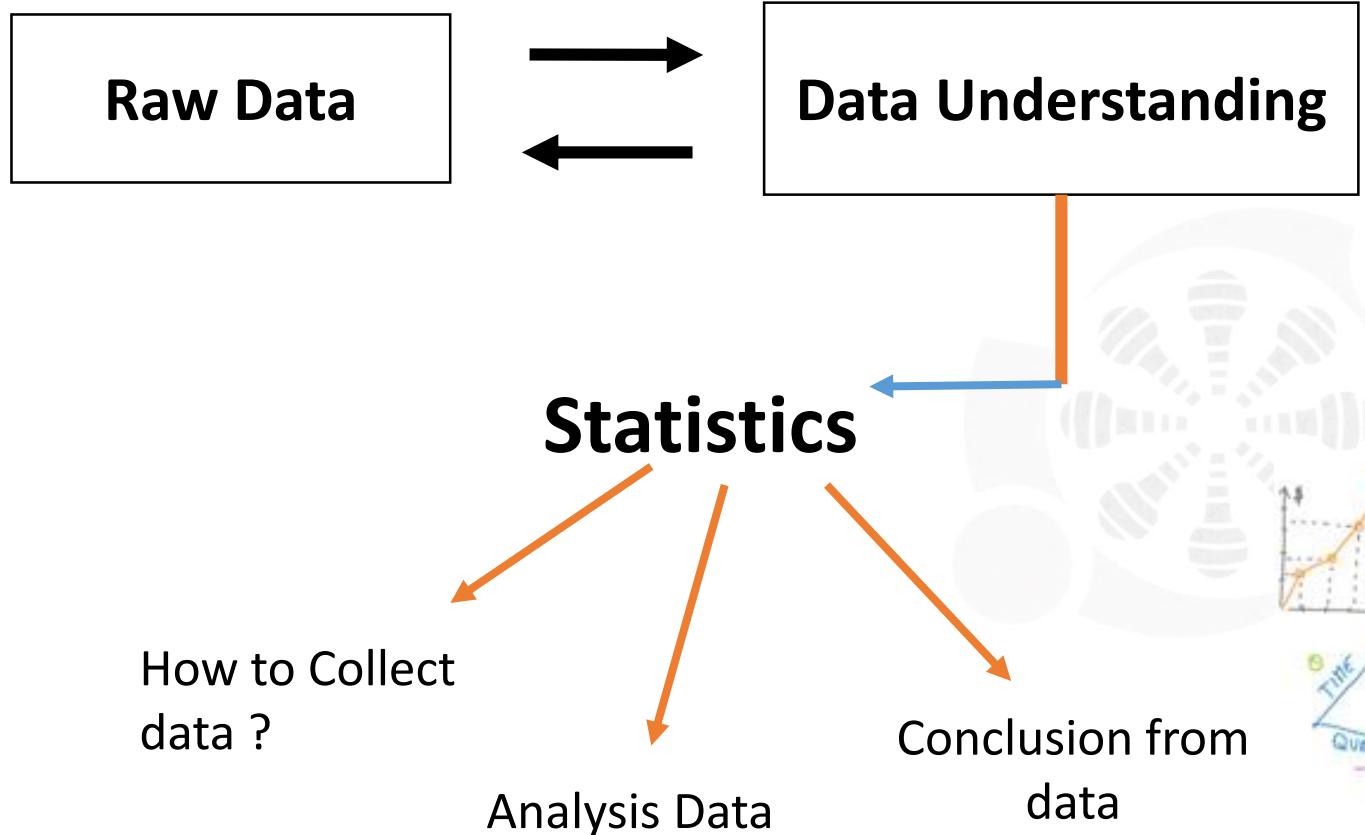


Insights

Data Analysis Life cycle



Statistics for Data understanding



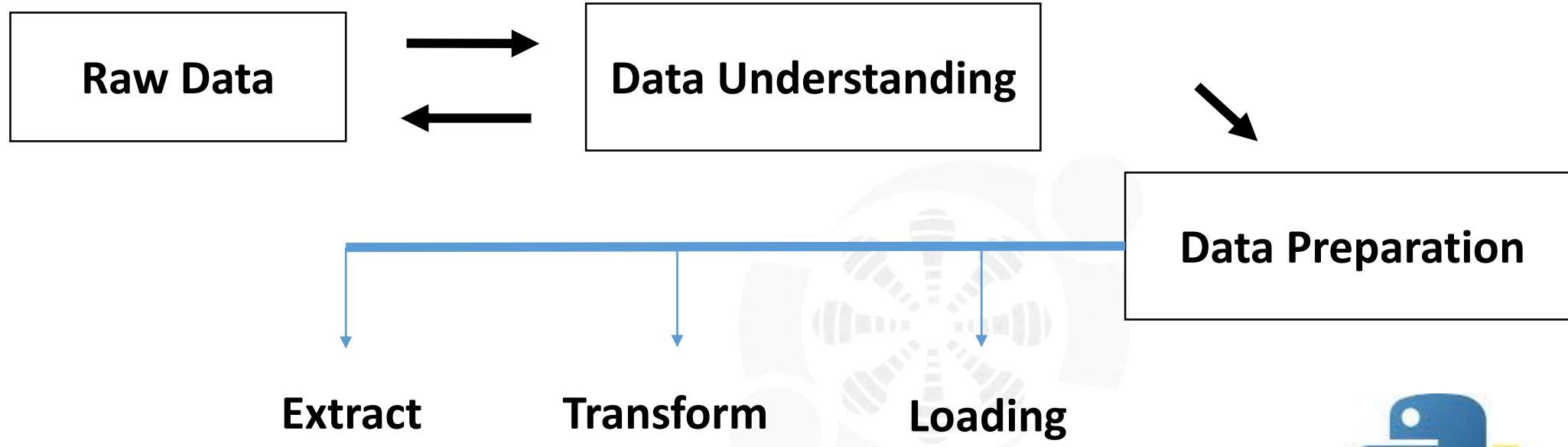
Data Science

Dealing with unstructured and structured data

- Process of extracting knowledge from data
- Using knowledge predict the unknown



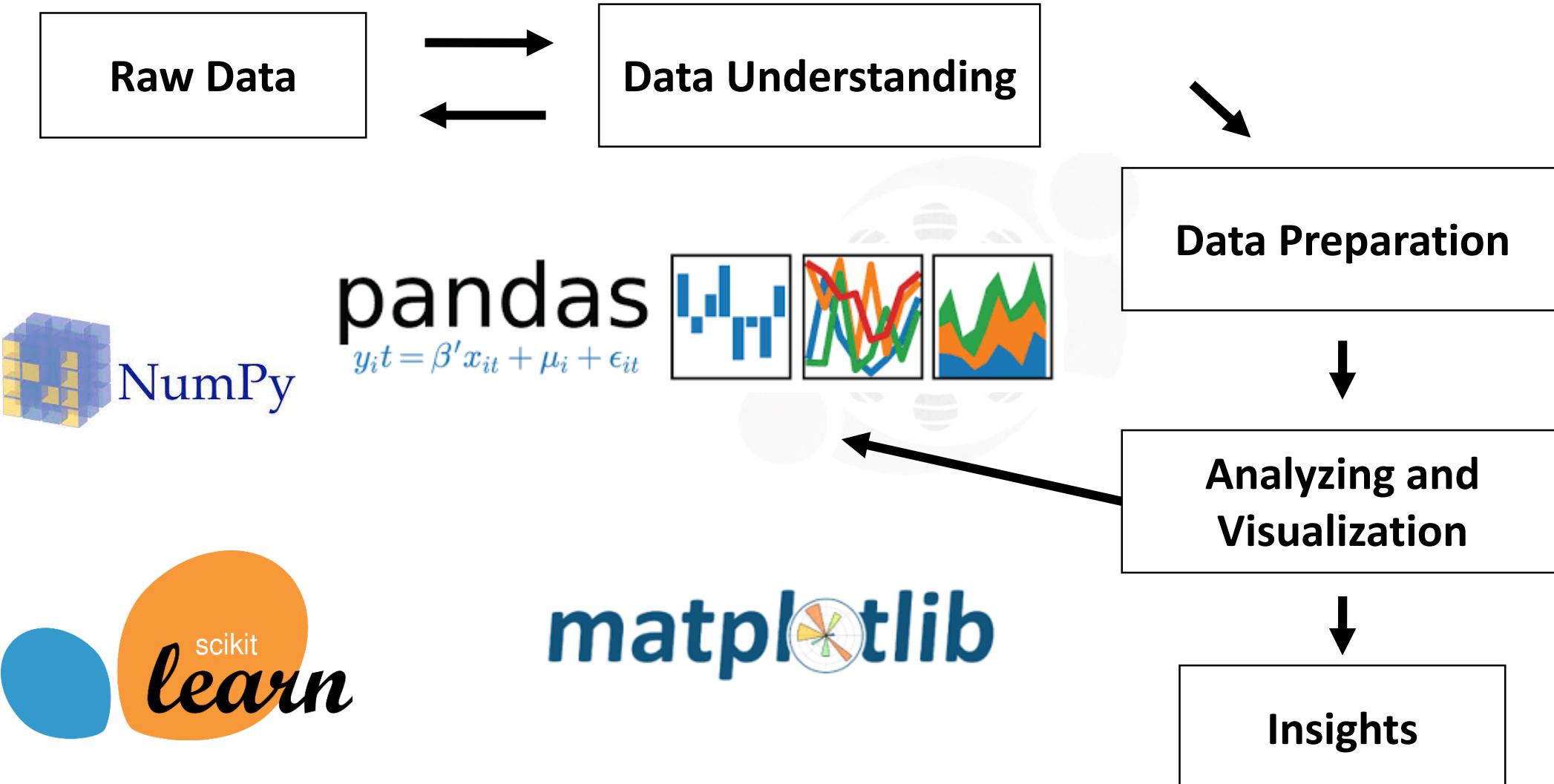
Statistics for Data understanding



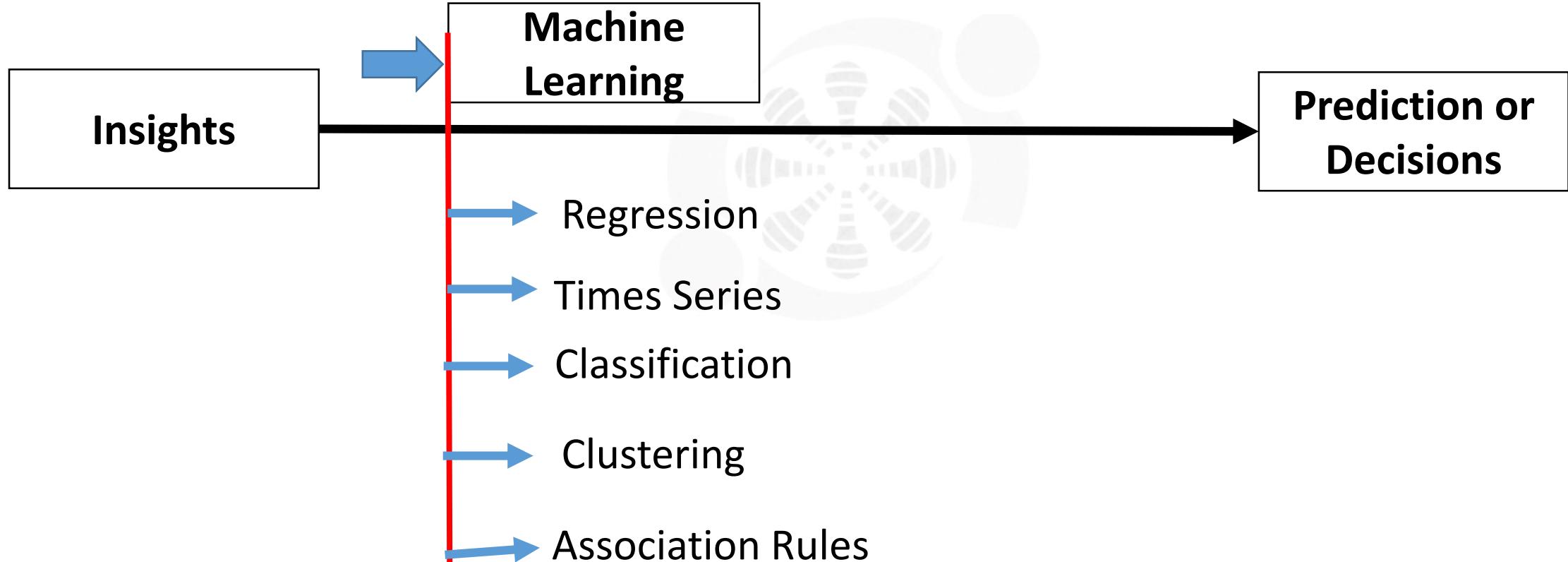
python™

Description	King of Data Science Programming Language	Golden Child of Data Science
Purpose	General-purpose language which is known for : 1.Simple syntax 2.Compatibility	Open-Source programming language and very beneficial for statistical computing.
Features	1.. Broadness 2. Efficient 3. Extensible 4. Can be mastered Easily	1. Open source 2. All in one analysis toolkit 3. Robust 4. Powerful package ecosystem
Libraries	1. NumPy/ SciPy 2. Pandas 3. Matplotlib / Seaborn	1. CARET 2. STRINGR 3. GG PLOT2
Data Handling Capability	Python uses data analysis packages for data handling. So Python good handles capability	R computes everything in memory and hence capabilities are limited by RAM Size.
Speed	Python outshines R in terms of Speed in majority cases	Performance of R is slow.

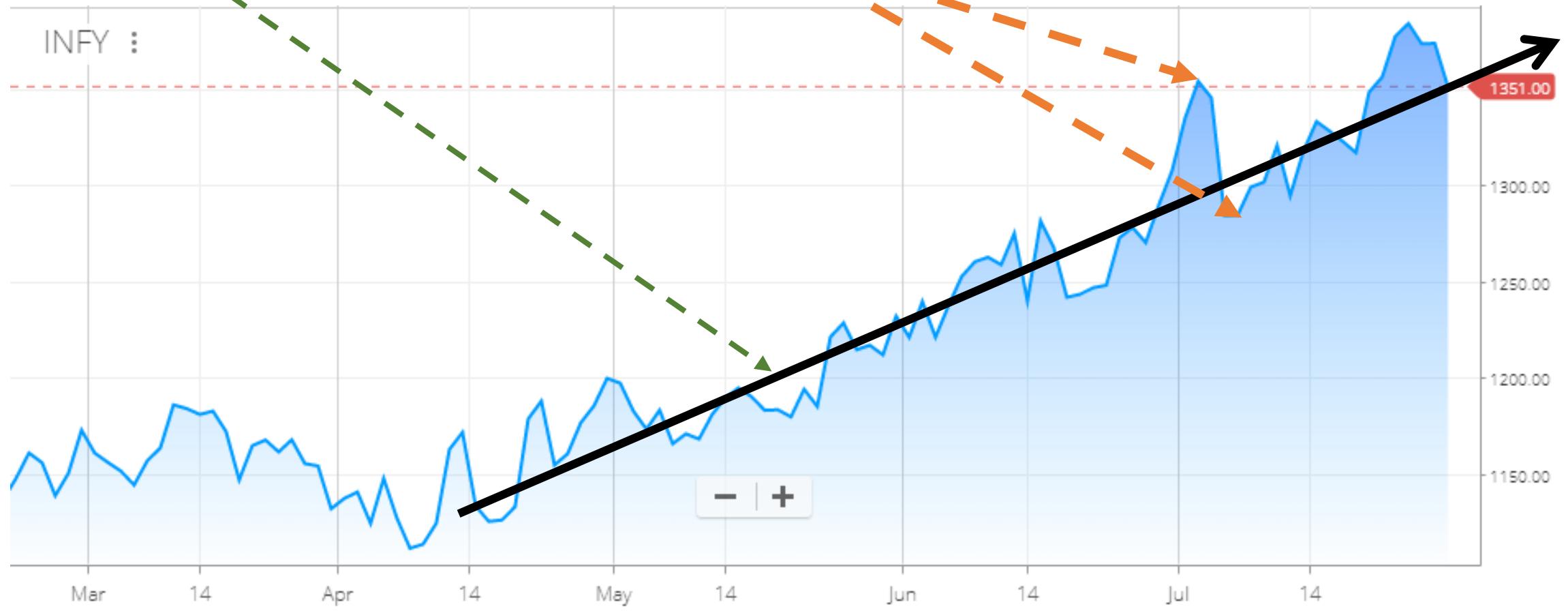
Data Analysis Life cycle



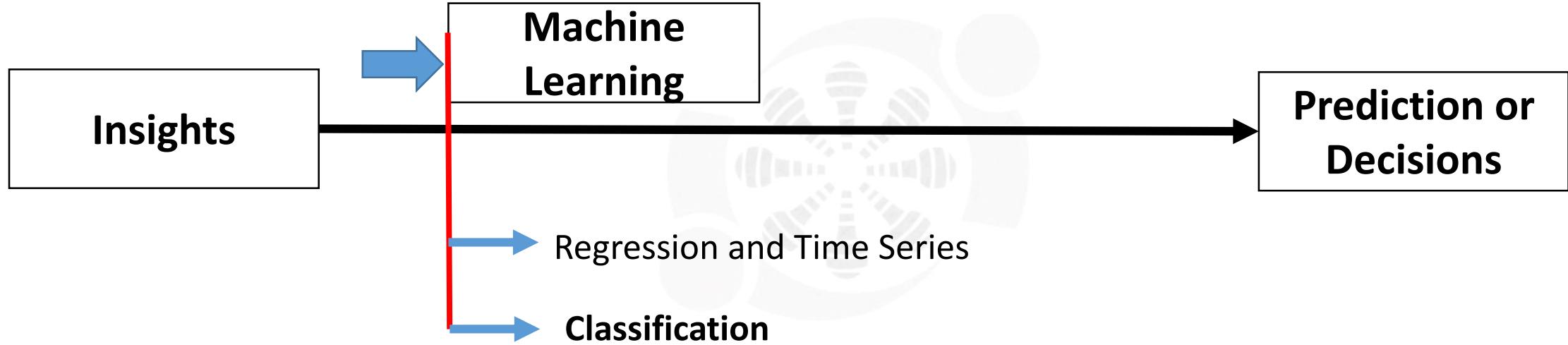
Where it is used



Regression and Time series for prediction



Where it is used

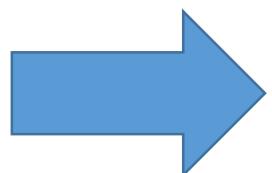
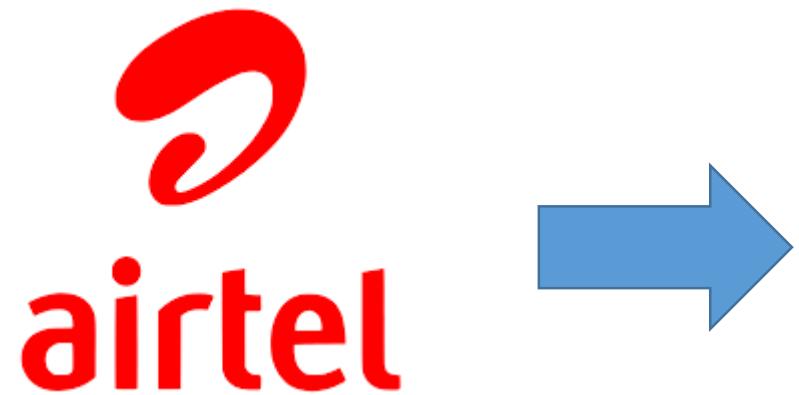


Domain	Question
Telecom	Is a customer likely to leave the network? (churn prediction)
Retail	Is he a prospective customer? that is likelihood of purchase vs. non-purchase?
Insurance	To issue insurance should a customer be sent for a medical checkup?
Insurance	Will the customer renew the insurance?
Banking	Will a customer default on the loan amount?
Banking	Should a customer be given a loan?
Manufacturing	Will the equipment fail?
Health Care	Is the patient infected with a disease?
Health Care	What type of disease does a patient have?
Entertainment	What is the genre of music?

Classification

Retail Market

Eg: Churn Problem

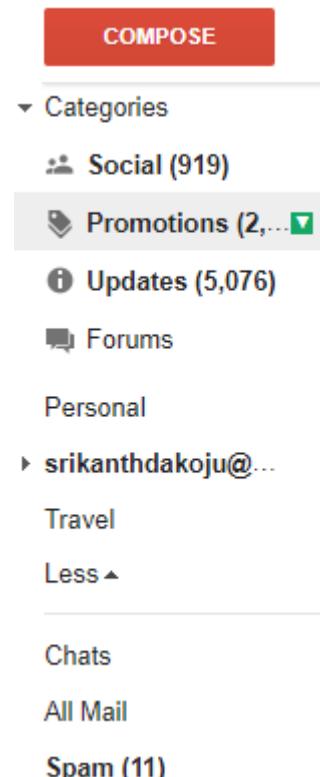


1. Which Kind of people are moving from airtel to jio
2. Why ?
3. What are the preventive measure to be taken to retain their customers

Classification

Medical Filed
Cancer Detection

Text Classification
Eg. Gmail Spam detection

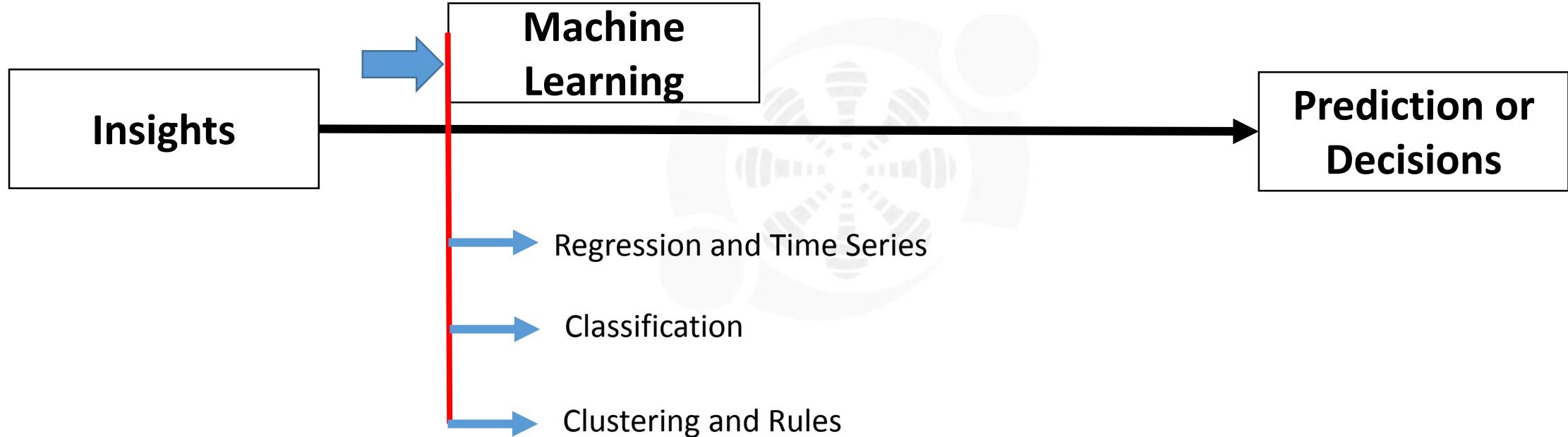


Face Recognition



Many more ...

Where it is used



Clustering and Rules – Recommendation system

If you purchase one product you will purchase one more product



If you watch one movie you will watch one more movie



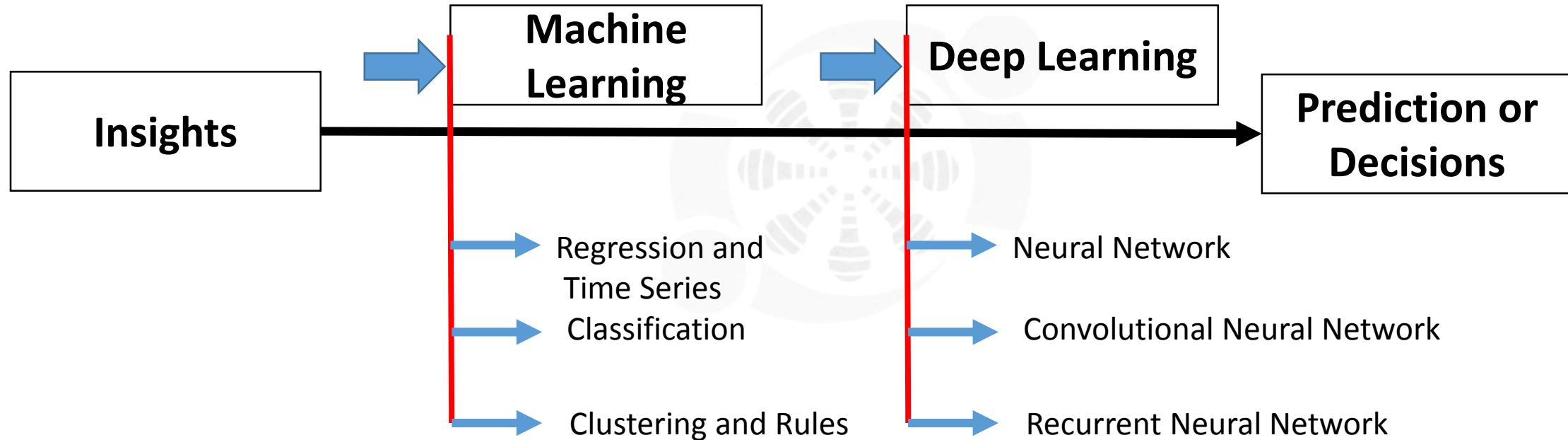
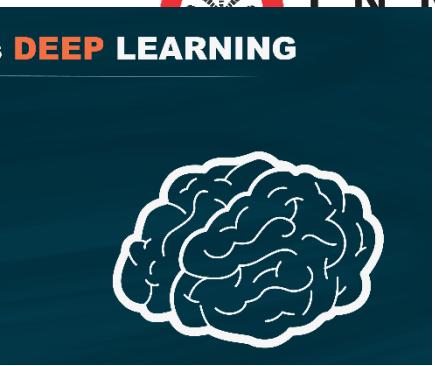
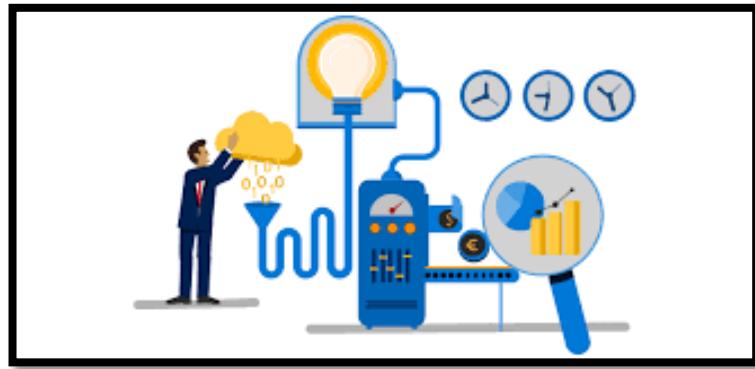
YouTube

Recommended phones

amazon

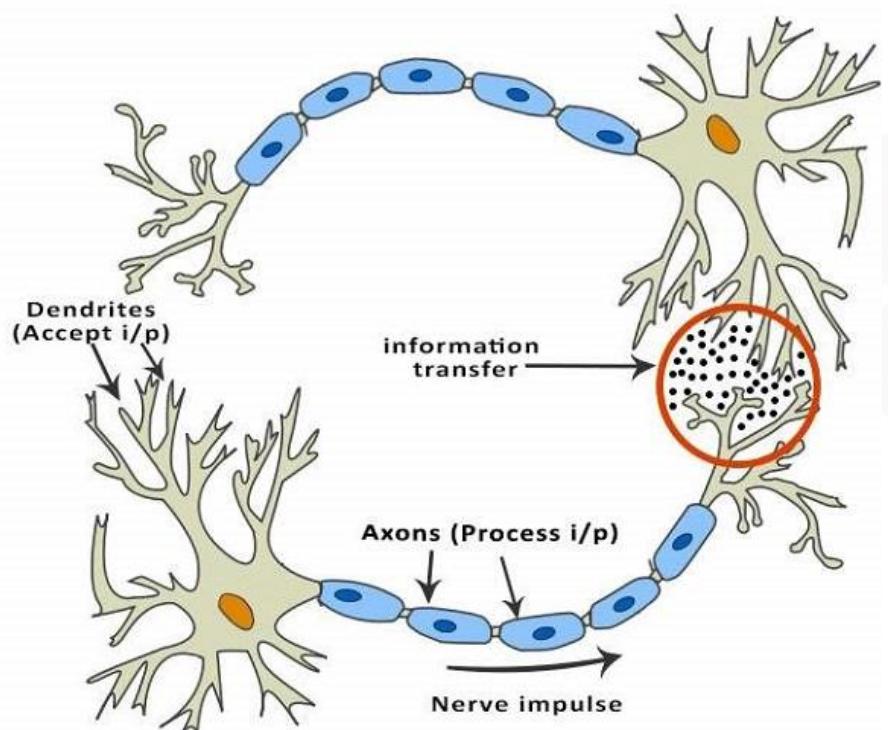


OS	Android 8.0 Oreo
RAM	3 GB
Weight	150 g
Dimensions	14.6 x 0.8 x 7.2 cm
Batteries	1 Lithium ion batteries required. (included)
Processor	Xperia R1 Plus Dual
Communication	Bluetooth, WiFi Hotspot
Technologies	4G LTE, GPRS, WiFi
Sensors	Dual SIM, GPS, FM Radio, Proximity sensor, eCompass, Accelerometer, Light sensor, Hall sensor, Gyro sensor, E-mail
Cameras	8MP
Type	Touchscreen Phone
Weight	150 Grams
Color	Black
Screen Size	2620
Accessories	Handset, Quick Charger, Type-C Data Cable, Startup Guide, Screen Guard and Stereo Headphones

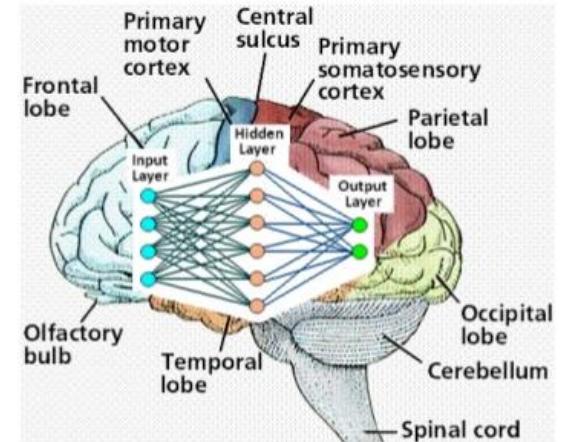


Neural Network

- Neuron In Human Brain



Neural Network



The human brain is essentially a large and unimaginably complex **Neural Network**. We can also think of the brain as an organized series of interconnected subsections of Neural Networks. We will look at how nature has implemented the Neural Network, and then look at the workings of the most common artificial Neural Network, the **Perceptron**.

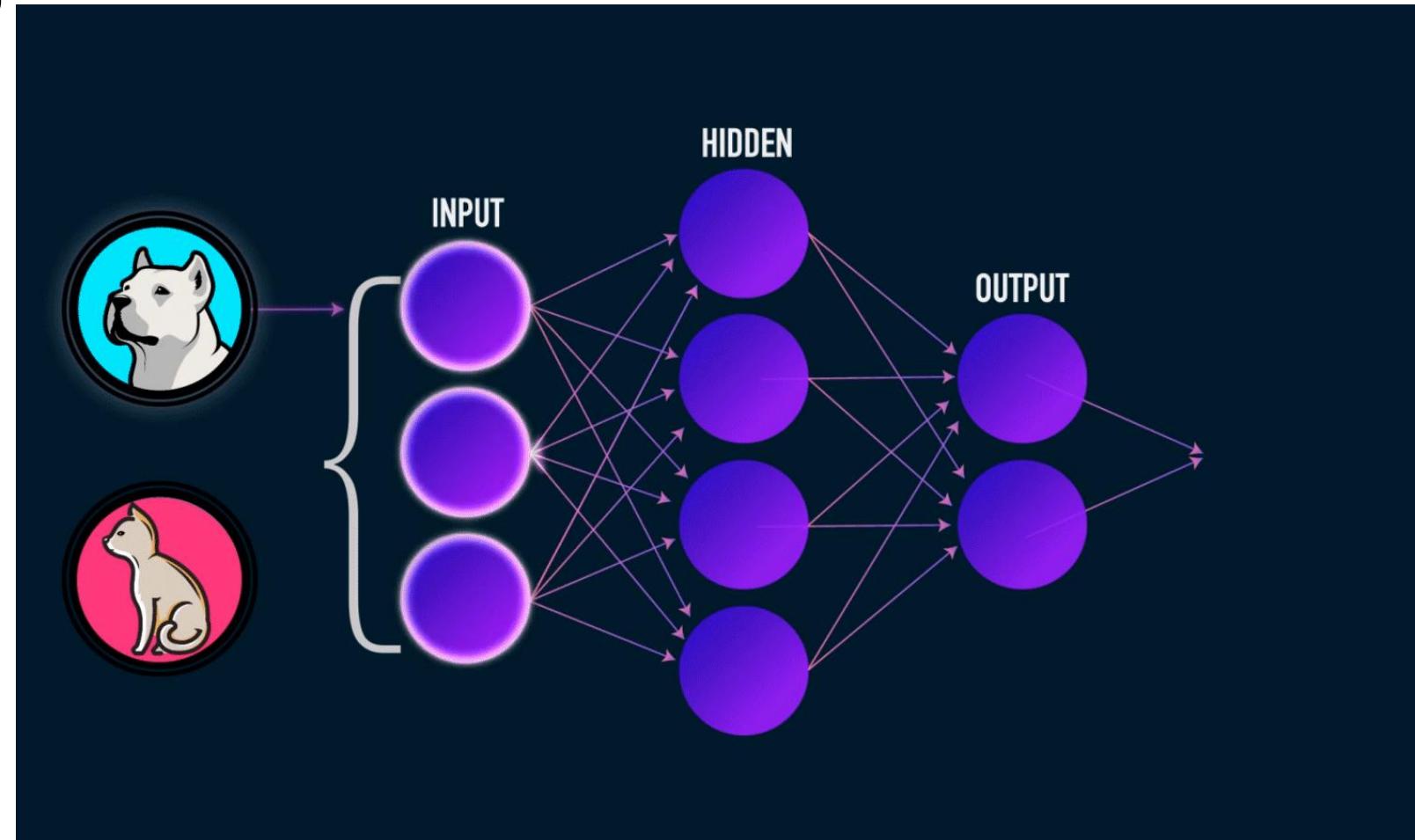
Neurons

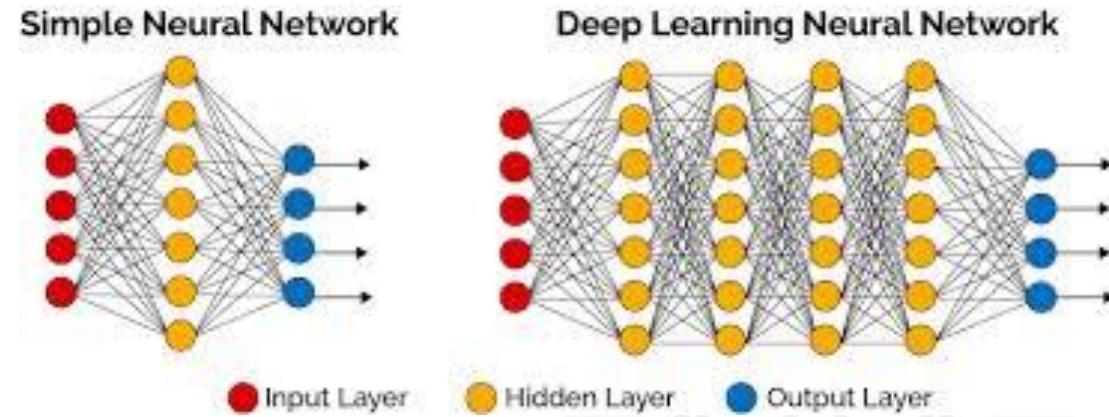
[open in browser](#) | PRO version Are you a developer? Try out the [HTML to PDF API](#)

pdfcrowd.com

CNN – Image Classification

- Classification into Dog





Simple Neural Network vs Deep Neural Network
Vs

Computer Vision – Object Detection

- Cat and Dog classification and Detection

Single Object

Classification



CAT

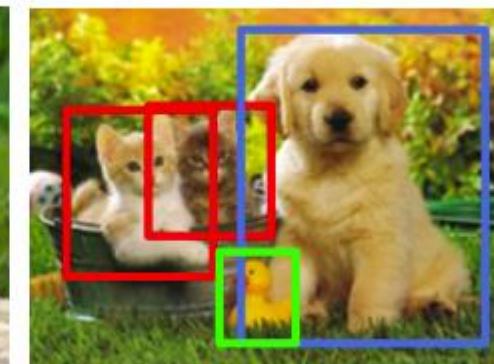
Classification + Localization



CAT

Multiple Object

Object Detection



CAT, DOG, DUCK

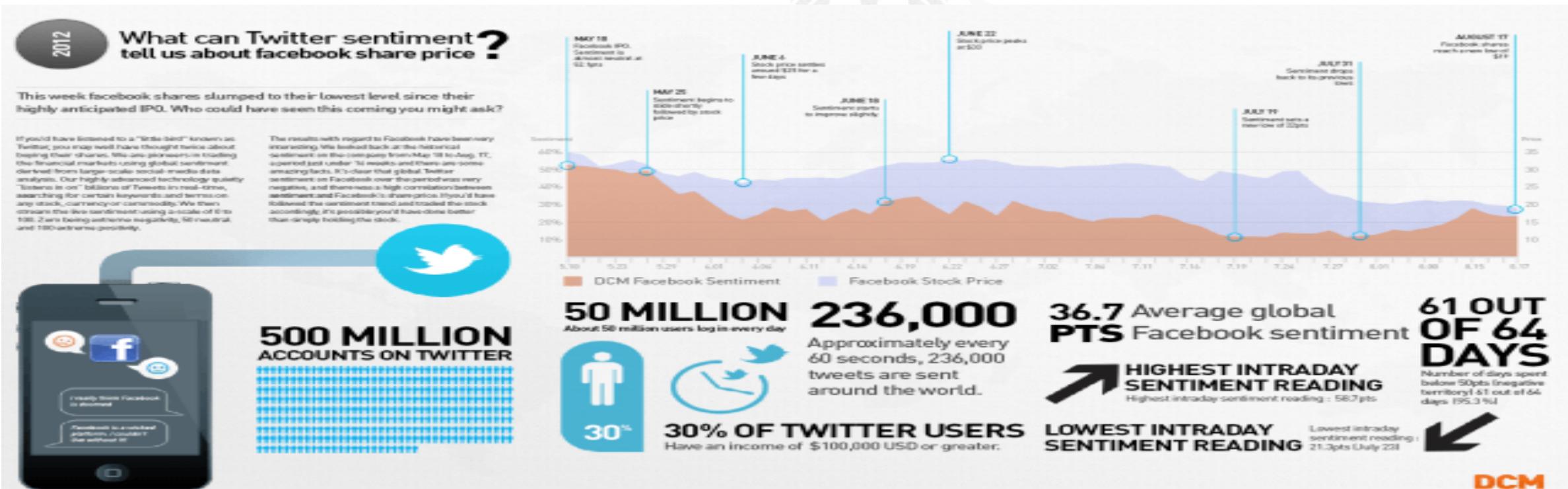
Instance Segmentation



CAT, DOG, DUCK

Sentiment Analysis

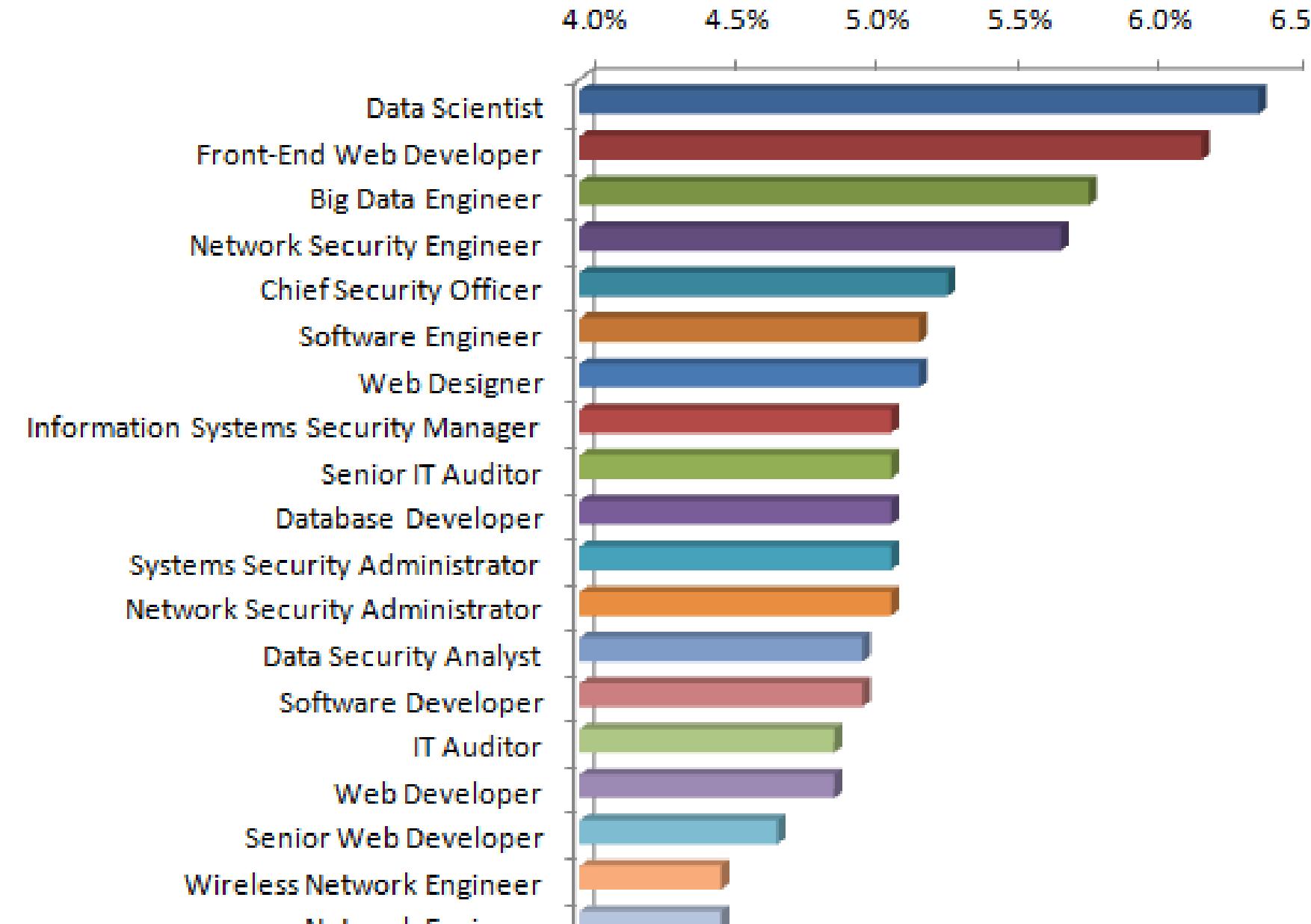
- Sentiment analysis (also known as opinion mining) refers to the use of natural language processing, text analysis and computational linguistics to identify and extract subjective information in source materials.
- Using sentiment analysis techniques, companies can respond to negative (or positive) brand perception.
- When a company releases a new product, monitoring and analyzing social media content can play a large role in quickly remediating bugs and errors.
- PR for political figures and celebrities depends heavily on sentiment analysis and how the person is perceived by people on social media.



Why Data Science?

- Profession to future
- More opportunities
- Higher Salaries
- Data is everywhere

Salary Growth Forecast for IT Jobs 2016-2017 (US)



What we do ?

- Project Based
- Assignments
- Quiz



Questions

FACE RECOGNITION

How human can identify face ?

- How we are seeing
- How we identifying persons
- What pattern we are following



How computer see person ?



Image

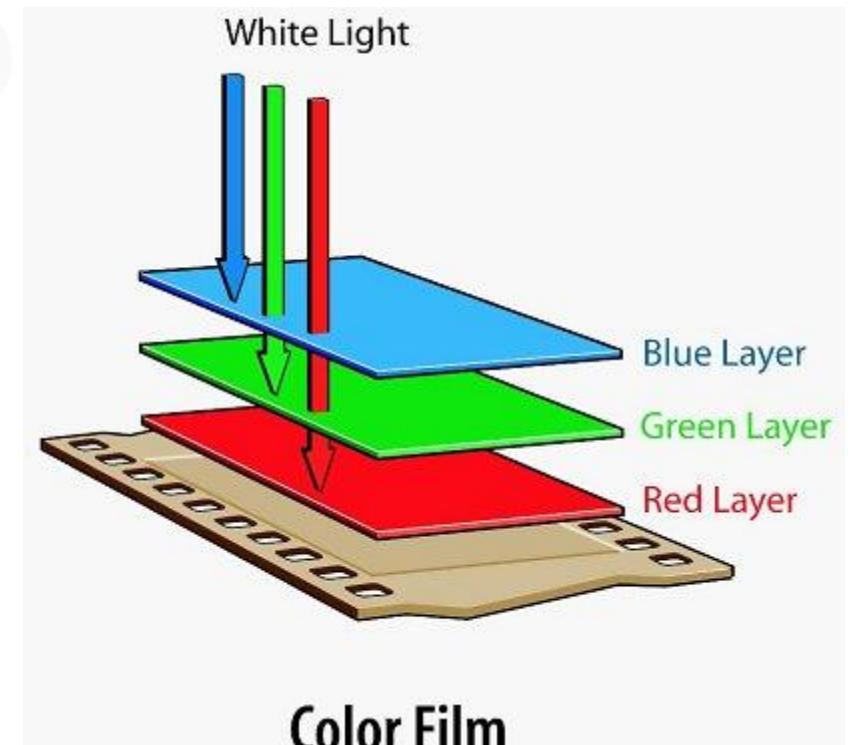
- An image is a picture that has been created by combination of ***vectors*** or ***values***.
- An image can be described in terms of ***vector graphics*** or ***raster graphics***.
- An image stored in raster form is sometimes called a **bitmap**.

Color Space

- Color image is combination of three channel
 - Red Channel
 - Green Channel
 - Blue Channel

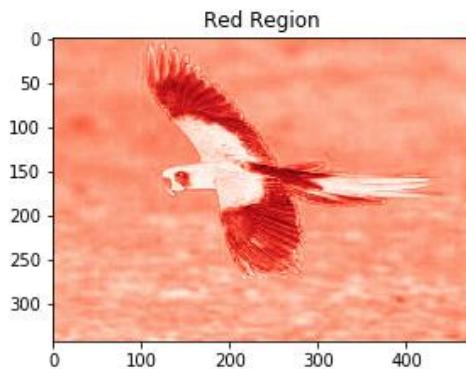


Basics of Image Processing

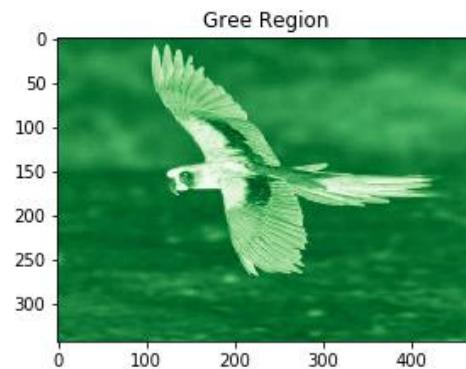




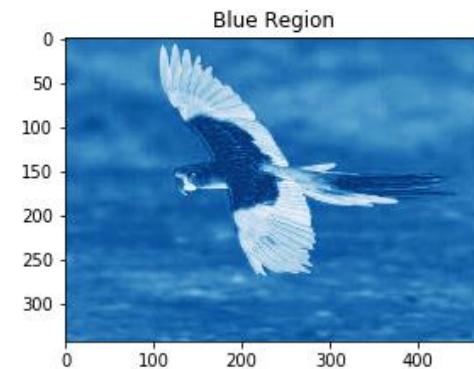
(343, 473 , 3)



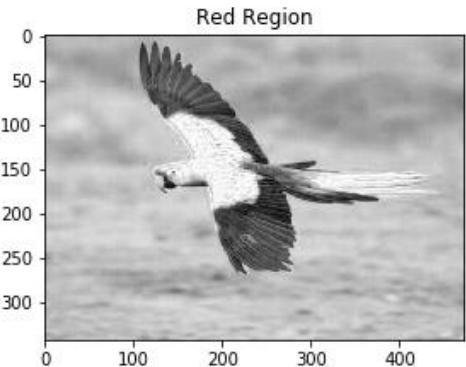
(343, 473)



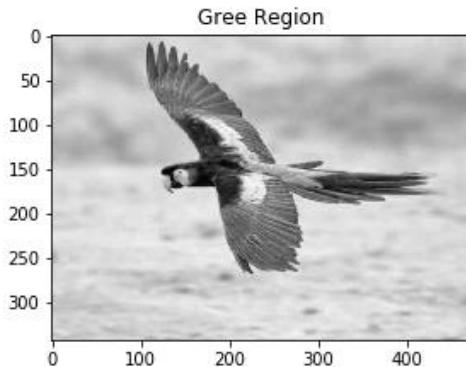
(343, 473)



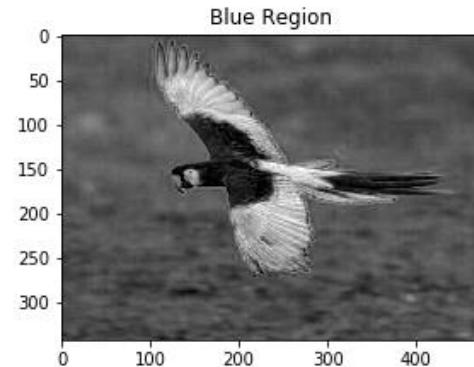
(343, 473)



(343, 473)



(343, 473)



(343, 473)

Use of Grayscale image



(343, 473 , 3)



(343, 473)

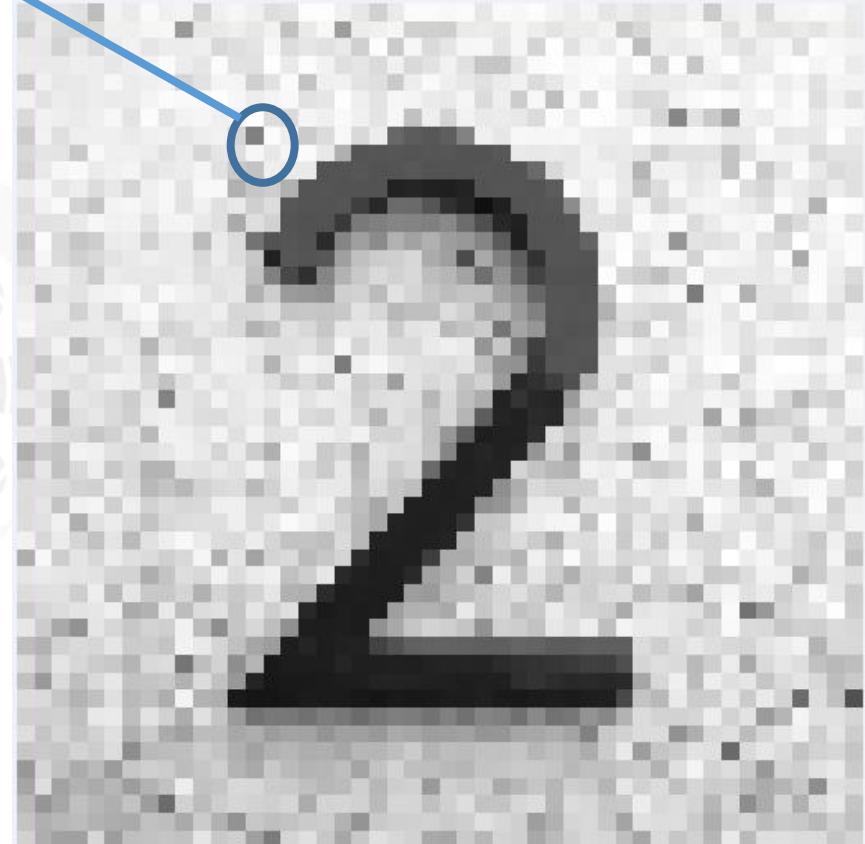
Length of matrix or array reduce by 3 times hence the computation time

Deeper into Image



(1200 , 1200)

PIXELS



PIXELS

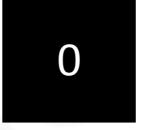
- Pixels are also called values
 - Range from $0 - 2^n$

Eg: for 8 bit image : $n = 8$
range of pixel values is **0 – 255**

PIXELS

- Lower the value ***darker*** the color

Eg: black color for gray scale image



0

- Higher the value ***lighter*** the color

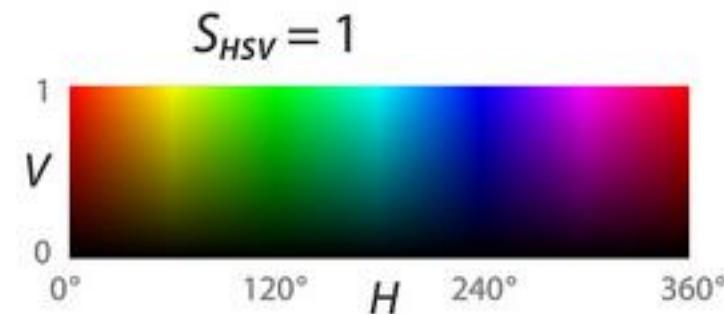
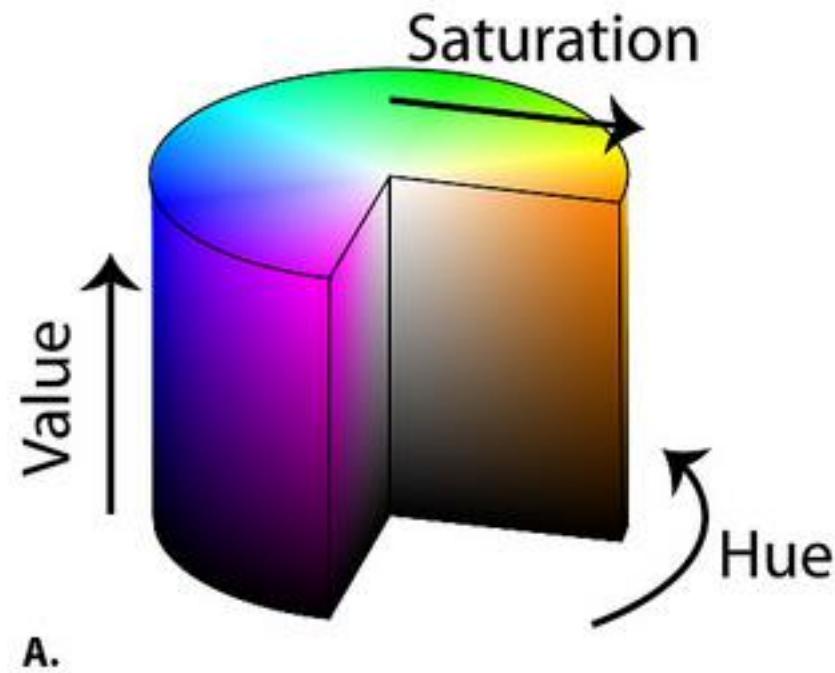
Eg: White color for gray scale image



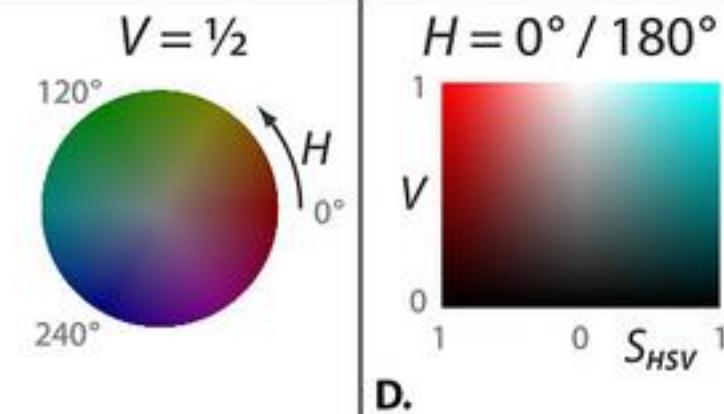
255

Color Space

HSV



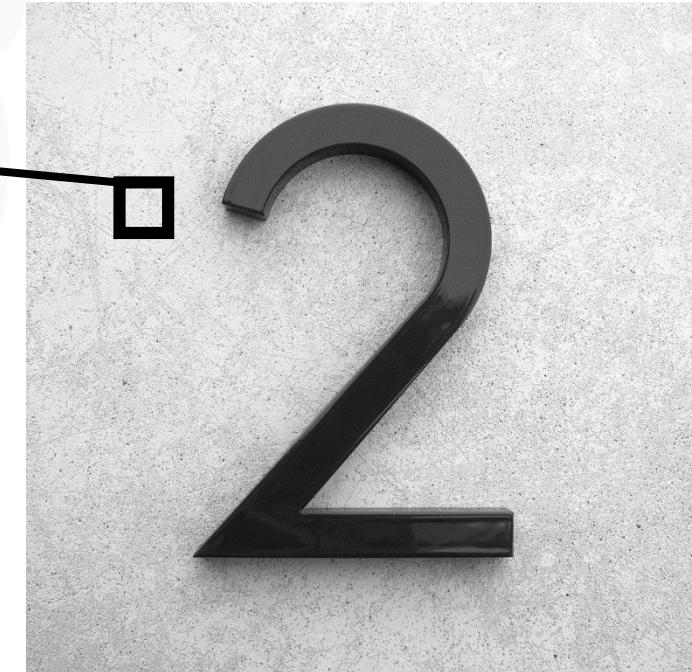
B.



Array in graphical representation

- Image is matrix and representation in graphical manner

0	230	255	53	235
255	20	240	255	0
255	70	255	100	120



Information in Image

- Skill of extraction information or knowledge from data is called Data Science

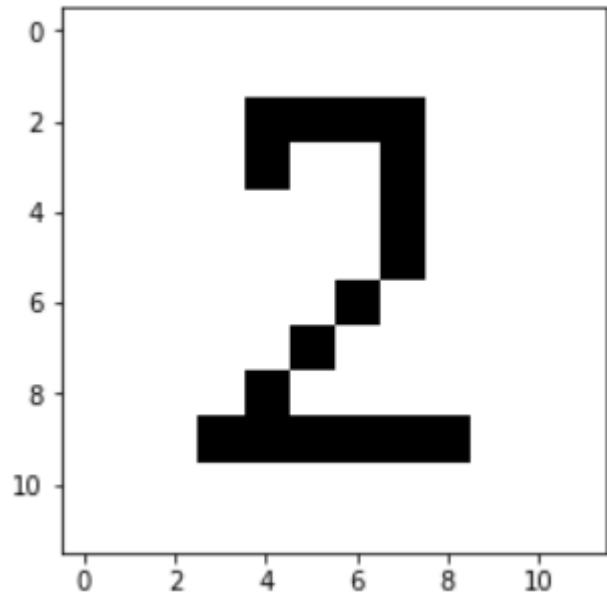
- Transformation
- Analysis
- Modeling



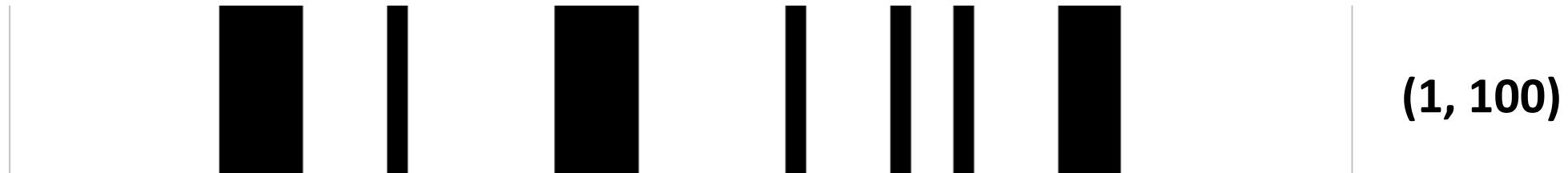
(1200, 1200)

Data Transformation

- Transforming the data into structured
 - Resize {Bring the all image into same shape}



- Reshaping
 - Converting into either row matrix or column matrix



Analysis

- Number of black bars
- Distance between each bars
- Width of each bars
- Position or index of bars



Train images or arrays



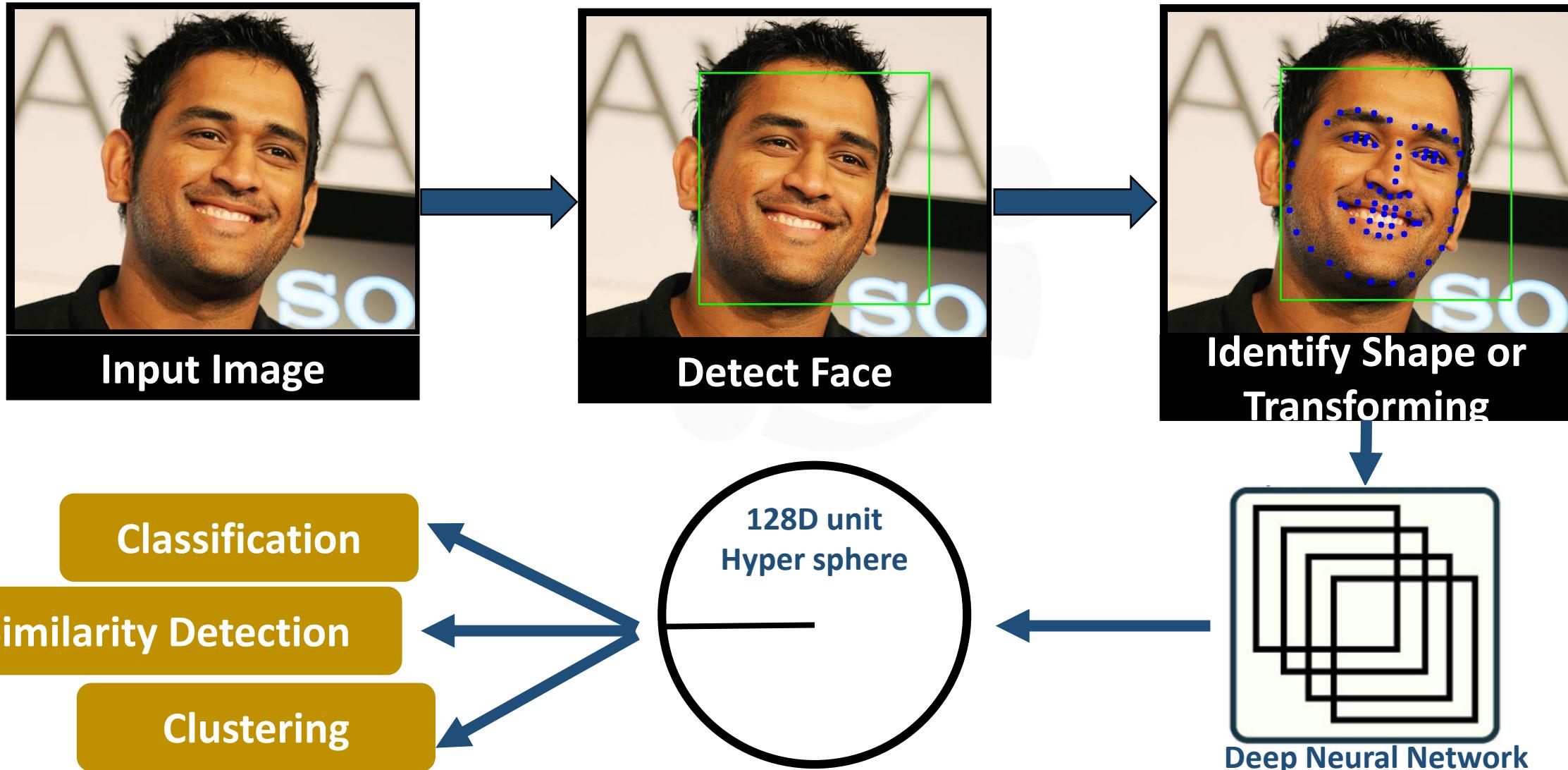
Face ?

- Converting into Grayscale
- Resize (Bring face into same size)
- ~~Reshaping~~
- Transforming
- Building Model





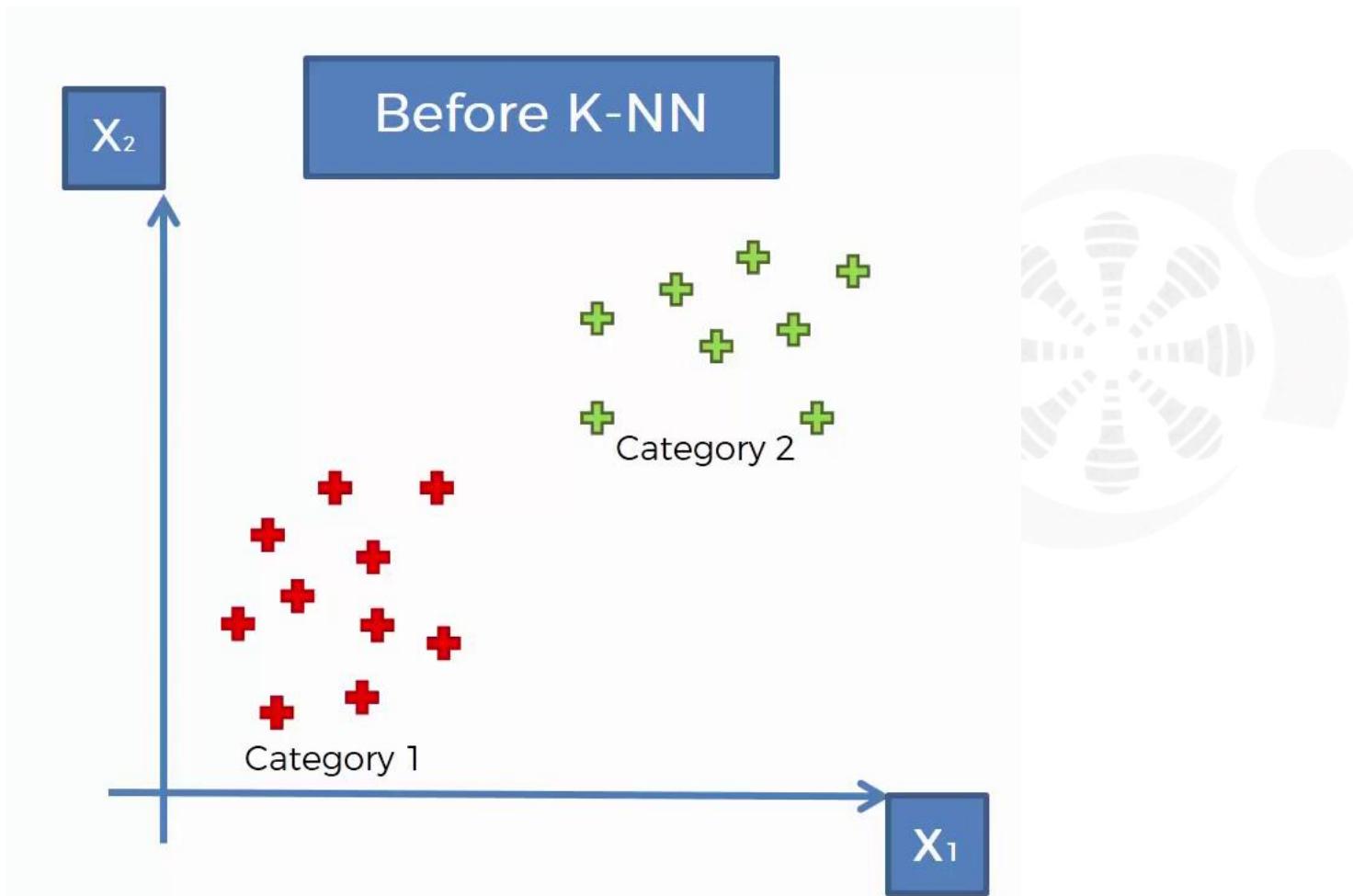
Recognition Algorithm



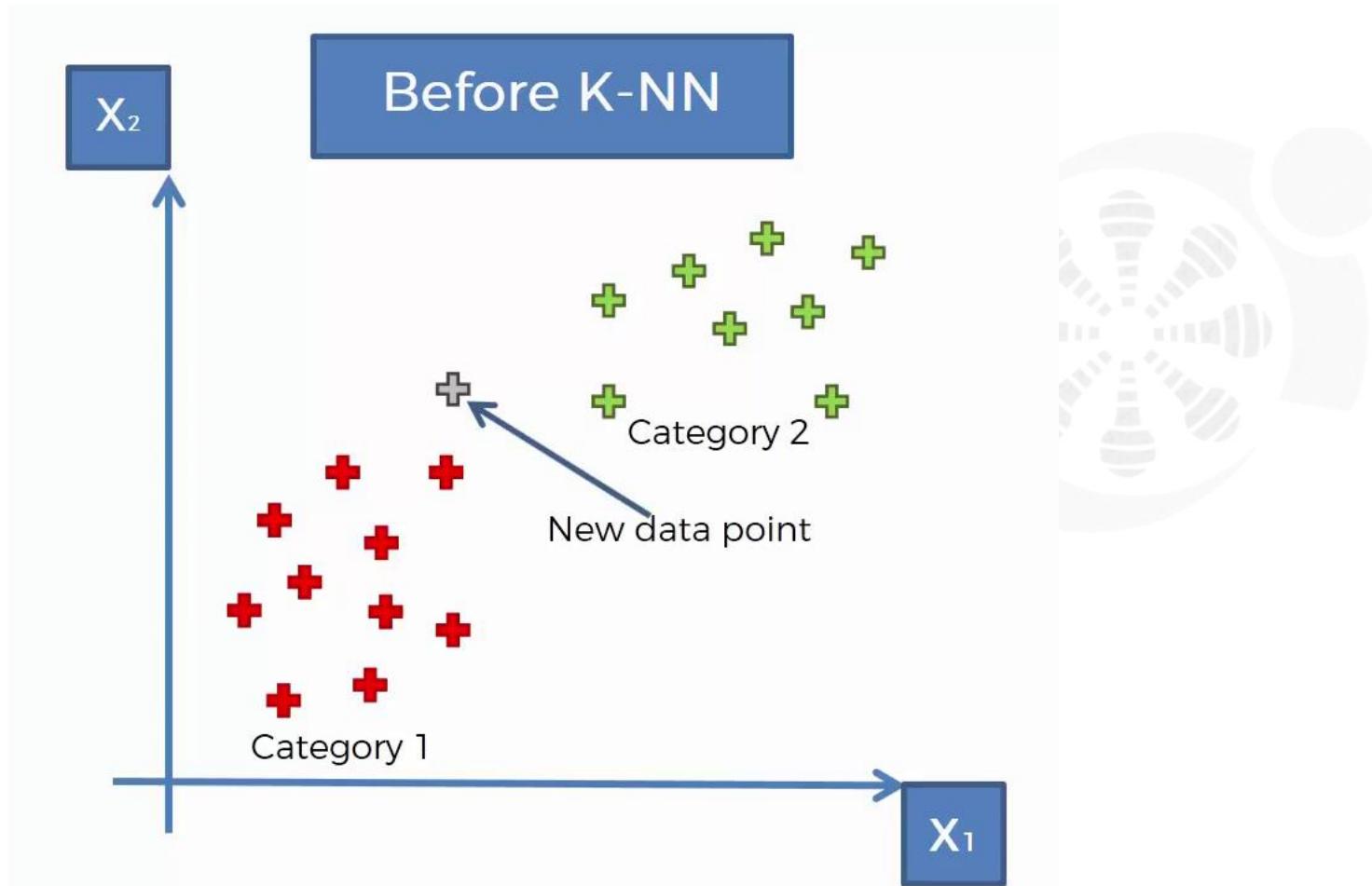
Let's Building Face Recognition Model

K- Nearest Neighbor Algorithm

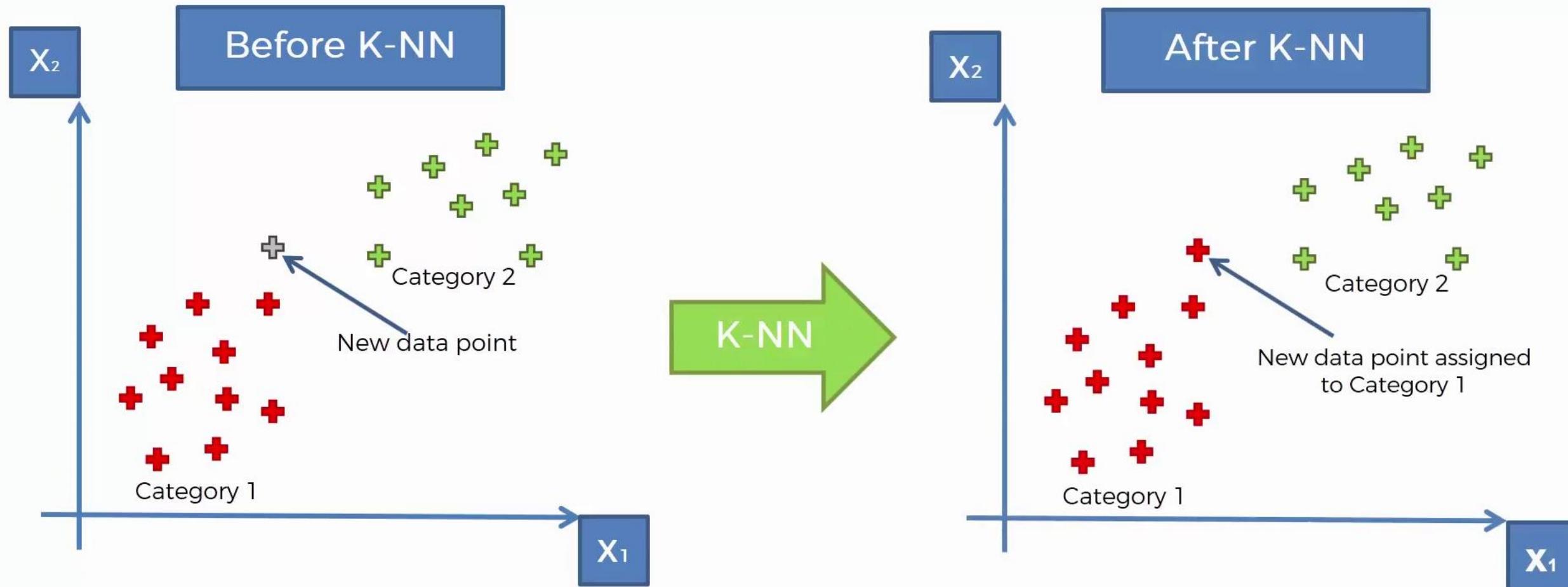
What K-NN does for you



What K-NN does for you



What K-NN does for you



How did it do that

STEP 1: Choose the number K of neighbors



STEP 2: Take the K nearest neighbors of the new data point, according to the Euclidean distance



STEP 3: Among these K neighbors, count the number of data points in each category



STEP 4: Assign the new data point to the category where you counted the most neighbors



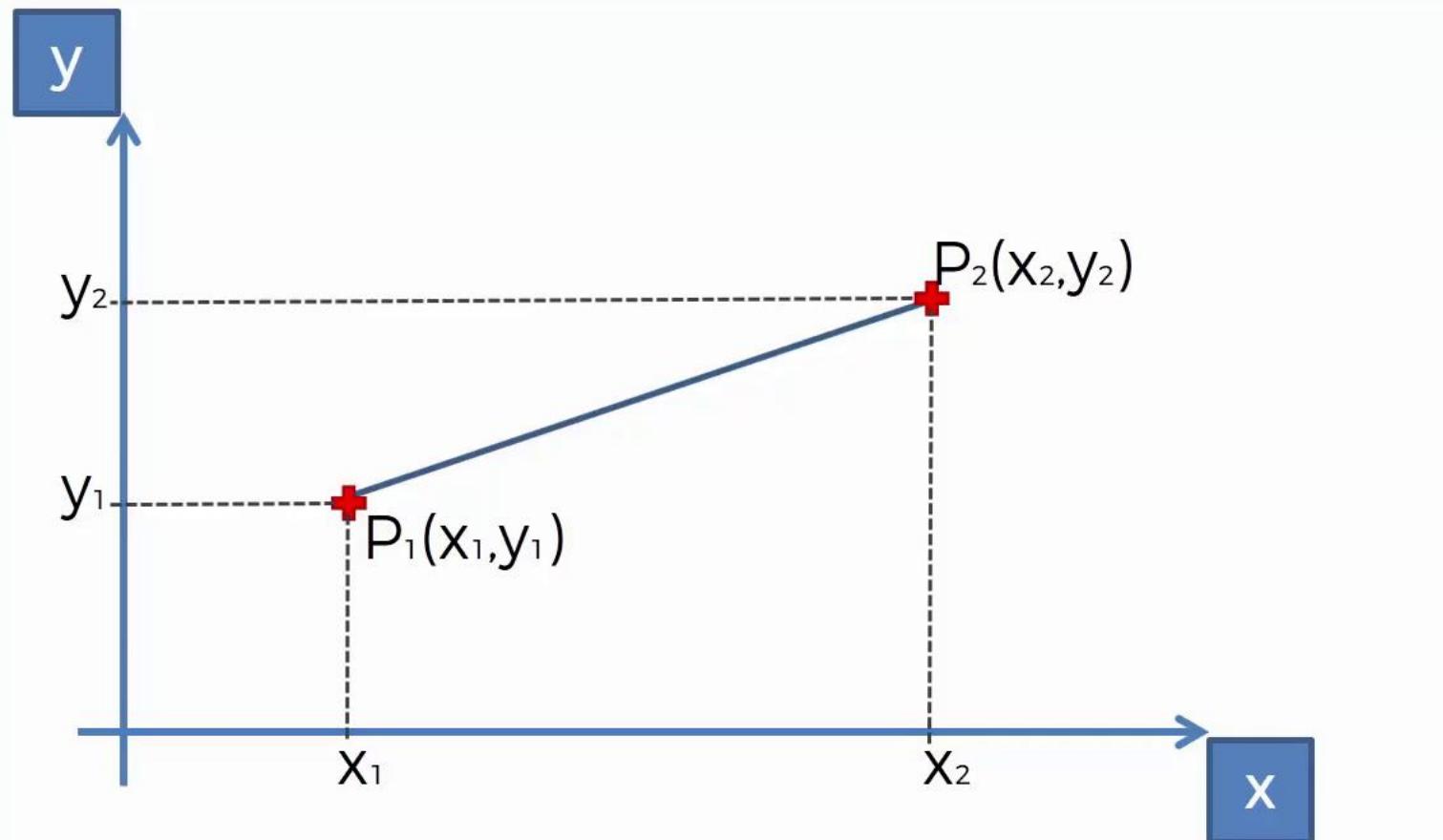
Your Model is Ready

K-NN Algorithm

STEP 1: Choose the number K of neighbors: K = 5



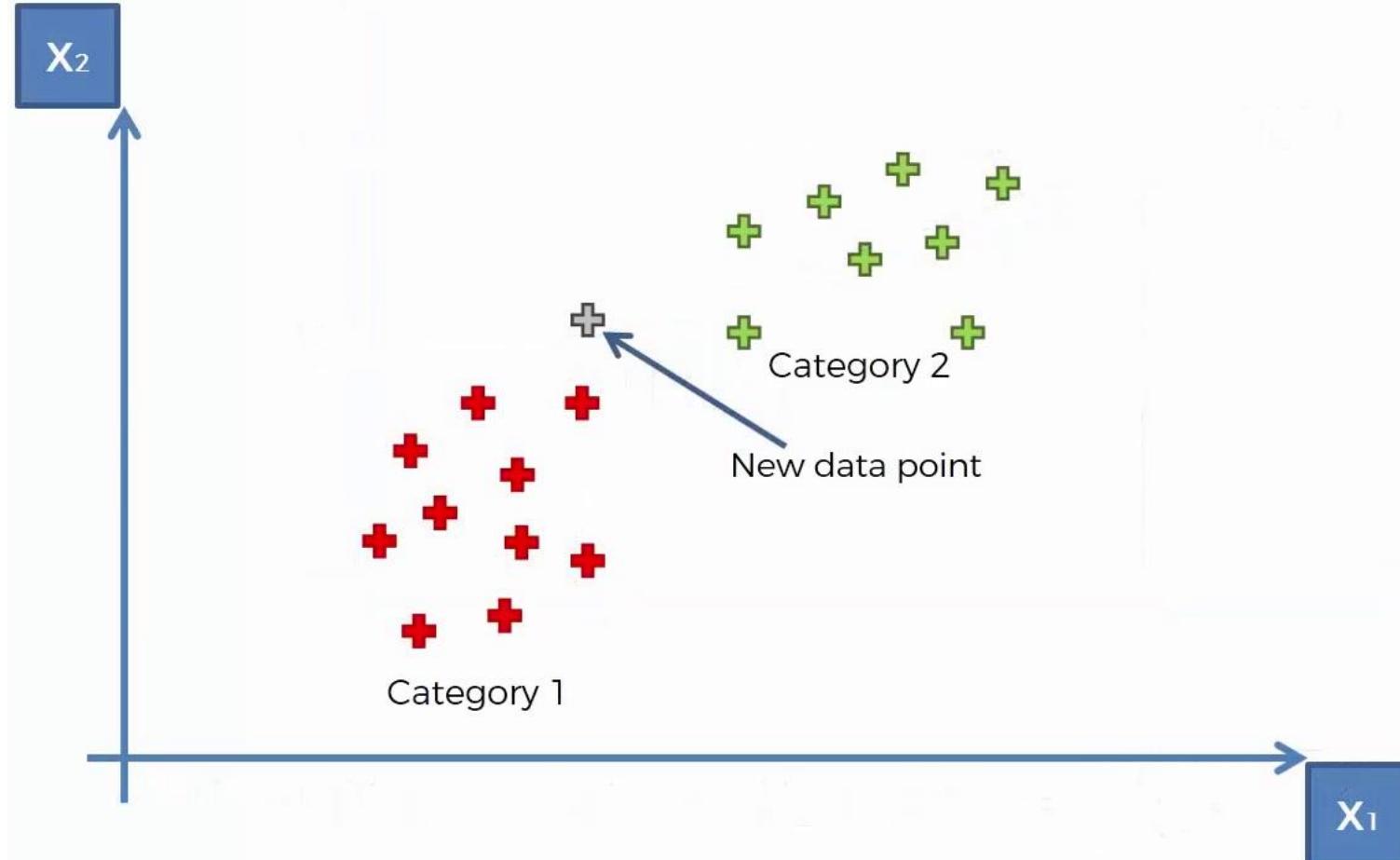
Euclidean Distance



Euclidean Distance between P_1 and P_2 = $\sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2}$

K-NN Algorithm

STEP 2: Take the $K = 5$ nearest neighbors of the new data point, according to the Euclidean distance



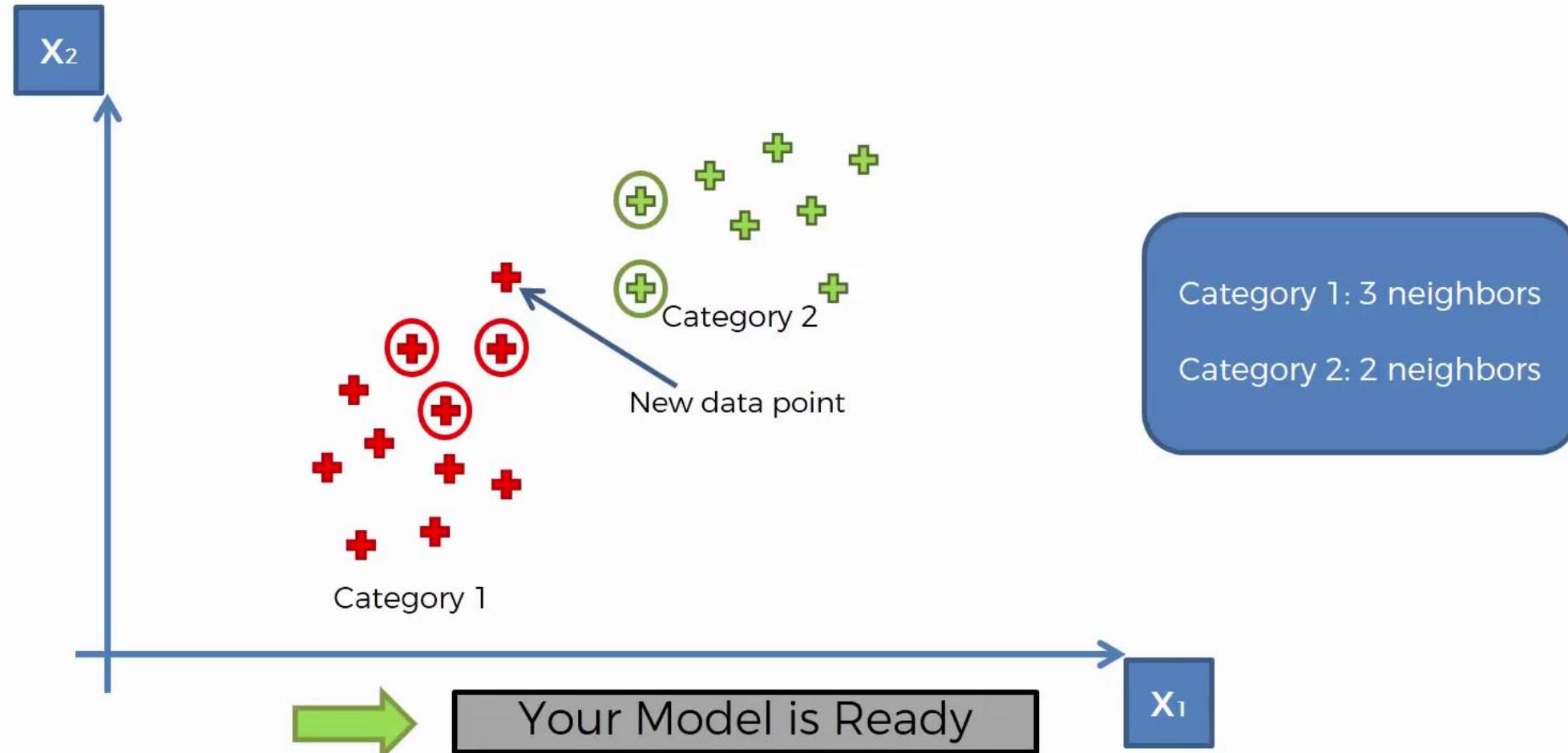
K-NN Algorithm

STEP 3: Among these K neighbors, count the number of data points in each category



K-NN Algorithm

STEP 4: Assign the new data point to the category where you counted the most neighbors



Let's build K-NN for Face Recognition

INNOMATICS TECHNOLOGY HUB

Innovation is our tradition