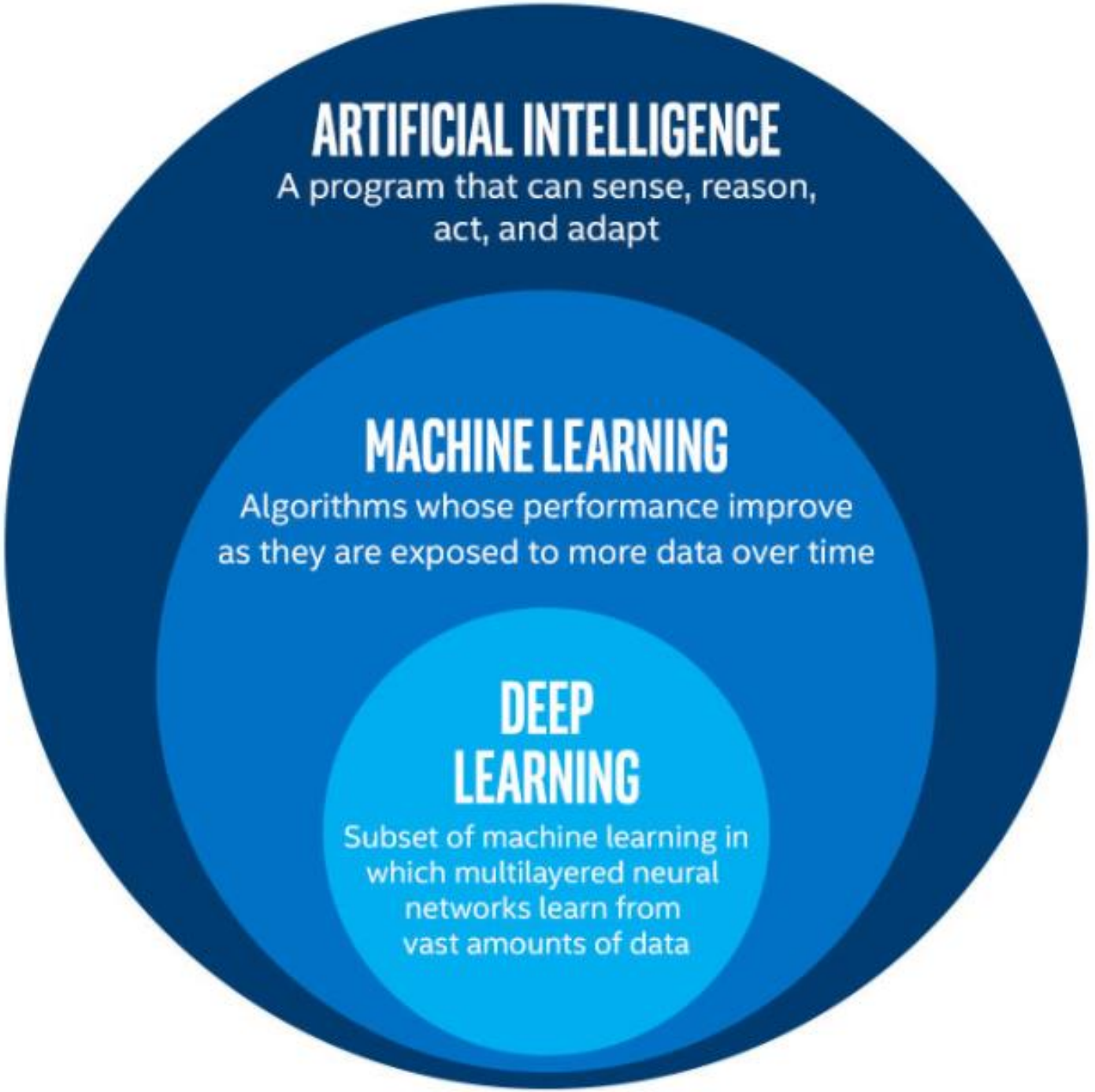
FREE
WORKSHOP

MACHINE LEARNING

& IMAGE PROCESSING

Thakshila Dasun
BSc. Eng in Mechatronics Eng
CIMA (UK)

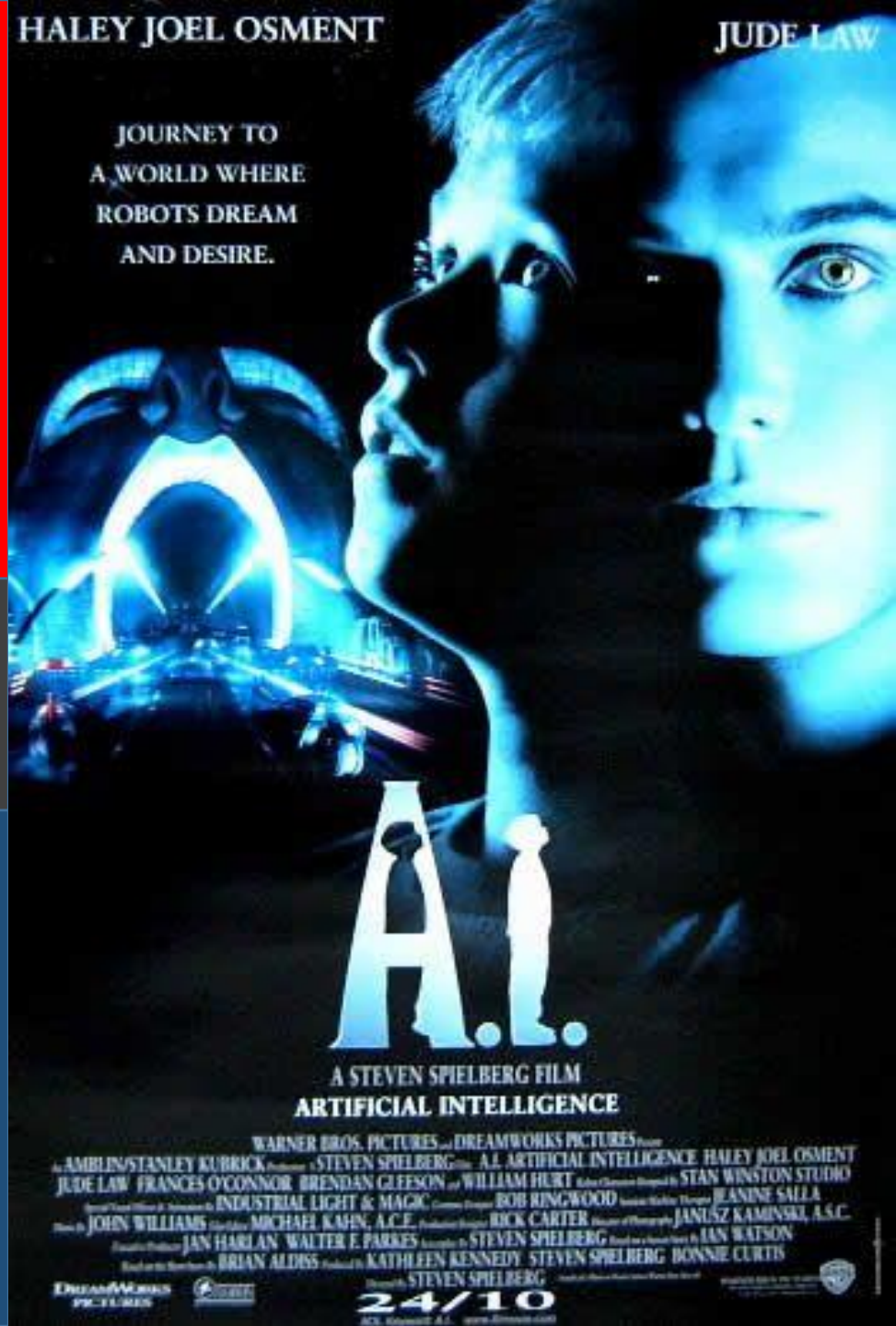
AI vs ML vs DL


ARTIFICIAL INTELLIGENCE
A program that can sense, reason, act, and adapt
MACHINE LEARNING
Algorithms whose performance improve as they are exposed to more data over time
DEEP LEARNING
Subset of machine learning in which multilayered neural networks learn from vast amounts of data

AI

IS NOT ONLY

ROBOTS



AI for Business

Finance, Marketing, Automobile, HRM, Healthcare, Stock Market, Investments, Scheduling

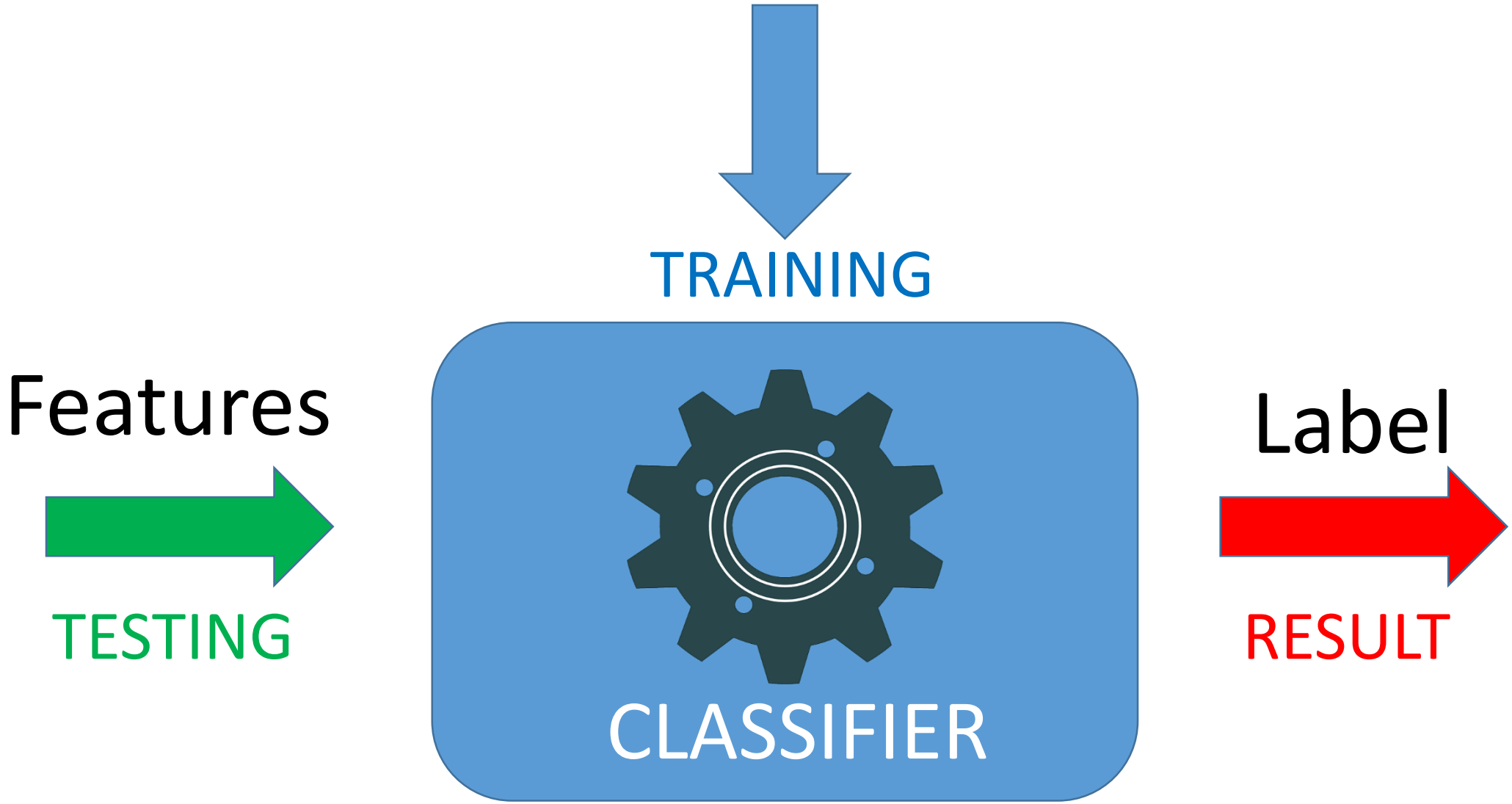
AI for Business

- Big Data analysis
- Data Science
- Statistics
- AI based IOT and Mobile App Development
- Finance and HRM
- Stock Market Data Analysis, share value Prediction
- Automobile Industry
- Medical Industry
- Music Composers
- Marketing & Advertising
- Investments & Cost Reduction

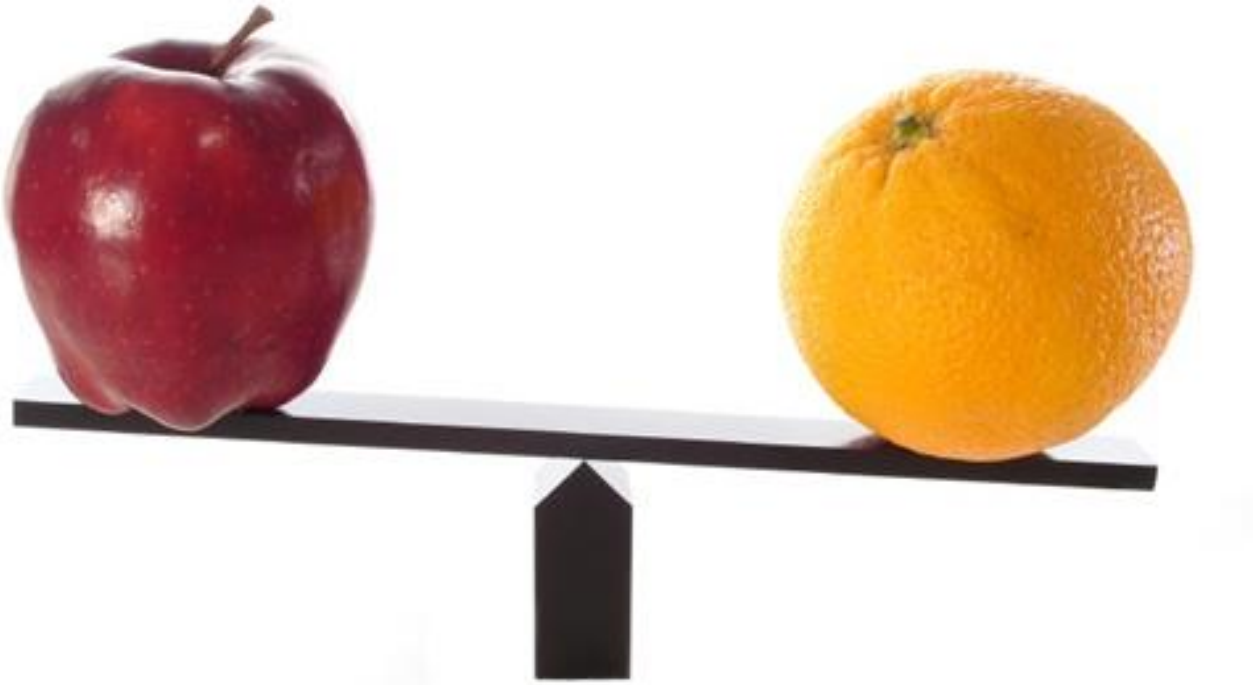
MACHINE LEARNING (ML)

- Machine learning is an application of artificial intelligence (AI) that provides systems the ability to automatically learn and improve from experience without being explicitly programmed.
- **Machine learning focuses on the development of computer programs** that can access data and use it learn for themselves.
- The process of learning begins with observations or data, such as examples, direct experience, or instruction, in order to look for patterns in data and make better decisions in the future based on the examples that we provide.
- **The primary aim is to allow the computers learn automatically** without human intervention or assistance and adjust actions accordingly.

Features & Labels



A Simple Example



How you are going to differentiate an Apple and an Orange?

What are the Critical Features ?

1. Color
2. Weight
3. Texture
4. Radius
5. Volume

Prepare a dataset from good features

FEATURES		**LABEL**
Weight (g)	Texture	
120	Smooth	Apple
140	Rough	Orange
135	Smooth	Orange
115	Smooth	Apple
170	Rough	Orange
165	Rough	Apple

(120,smooth), Apple
(140,Rough),Orange
(135,Smooth),Orange
(115,Smooth),Apple
(170,Rough),Orange
(165,Rough),Apple

TRAINING

(150,rough)

TESTING

CLASSIFIER

Apple/Orange

RESULT

MACHINE LEARNING

Allow the computers learn automatically without human intervention or assistance and adjust actions accordingly

Supervised ML

learned in the past to new data using labeled examples to predict future events

Unsupervised ML

Used when the information used to train is neither classified nor labeled. Unsupervised learning studies how systems can infer a function to describe a hidden structure from unlabeled data

Semi-Supervised ML

Use both labeled and unlabeled data for training. useful when only incomplete labels are available

MAIN TYPES

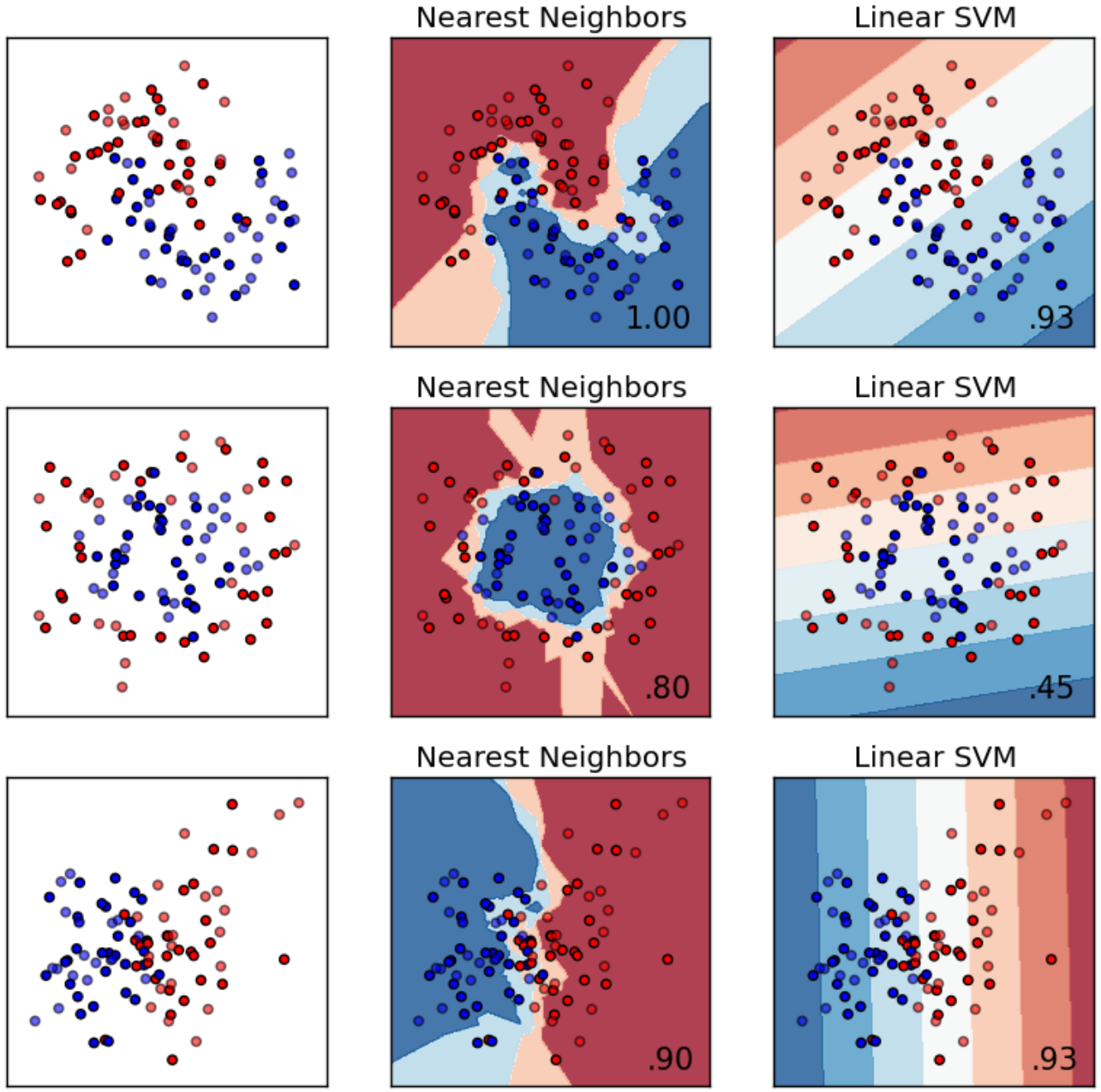
- **Supervised machine learning algorithms:**
  - An apply what has been learned in the past to new data using labeled examples to predict future events. Starting from the analysis of a known training dataset, the learning algorithm produces an inferred function to make predictions about the output values.
- **Unsupervised machine learning algorithms**
  - Used when the information used to train is neither classified nor labeled. Unsupervised learning studies how systems can infer a function to describe a hidden structure from unlabeled data. The system doesn't figure out the right output, but it explores the data and can draw inferences from datasets to describe hidden structures from unlabeled data.

MAIN TYPES cntd

- **Semi-supervised machine learning algorithms**
    - fall somewhere in between supervised and unsupervised learning, since they use both labeled and unlabeled data for training – typically a small amount of labeled data and a large amount of unlabeled data.

Classification & Regression

2 Types of Predictions in Machine Learning, Qualitative and Quantitative

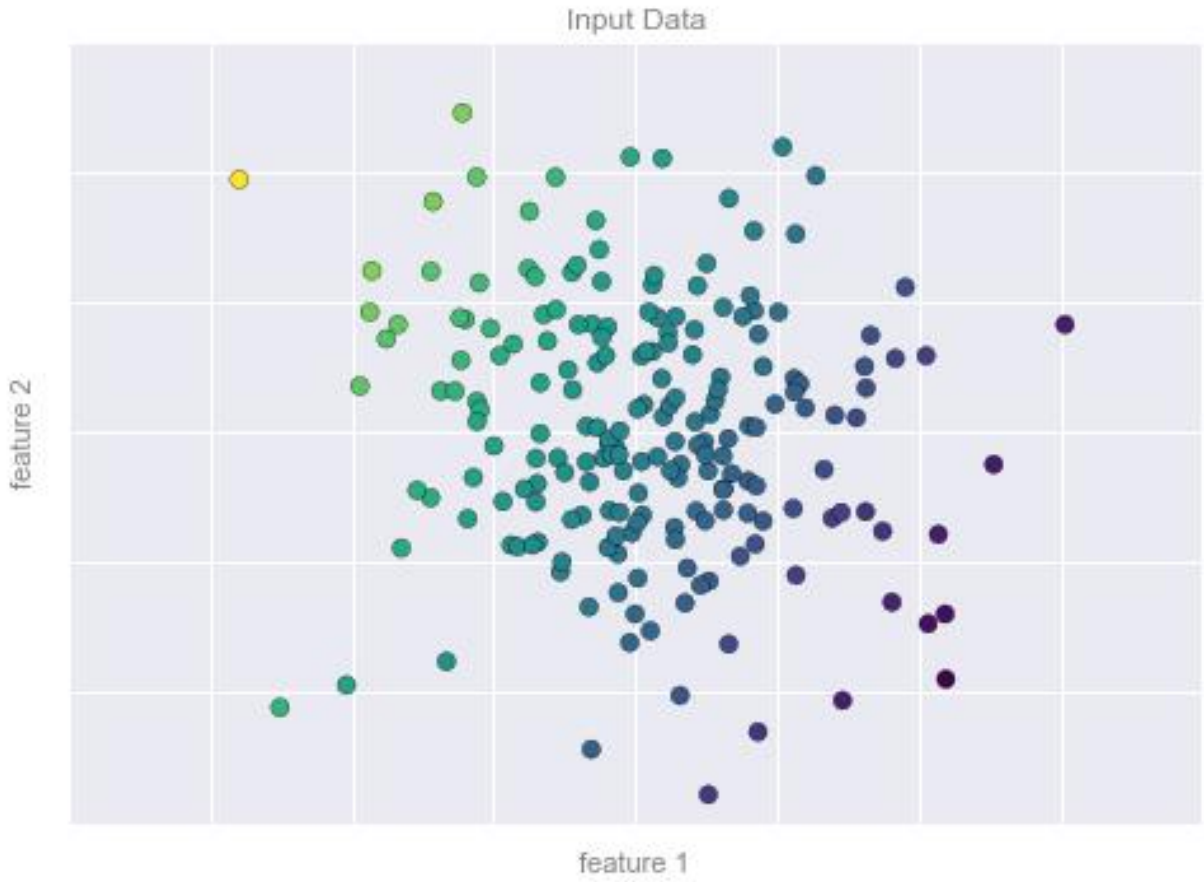


1. Classification: Predicting discrete labels

- “Classification" indicates that the data has discrete class label.
- Classification predictive modeling is the task of approximating a mapping function (f) from input variables (X) to discrete output variables (y) or classes.
- The output variables are often called labels or categories. The mapping function predicts the class or category for a given observation

2. Regression: Predicting continuous labels

- Regression predictive modeling is the task of approximating a mapping function (f) from input variables (X) to a continuous output variable (y).
- A continuous output variable is a real-value, such as an integer or floating point value. These are often quantities, such as amounts and sizes.
- For example, a house may be predicted to sell for a specific dollar value, perhaps in the range of $100,000 to $200,000.



Algorithms (1)

- Classification Algorithms (Supervised):
    1. Support vector machines
    2. Nearest Neighbors
    3. Decision Trees and Random Forests
    4. Gaussian naive Bayes

- Classification Algorithms (unsupervised):
    1. K-Means Clustering.
    2. Means Shift Clustering
    3. Gaussian Mixture Models
    4. Spectral Clustering

Algorithms (2)

- Dimensionality Reduction Algorithms
    1. Manifold Learning.
    2. Principal Component Analysis

- Regression Algorithms
    1. Linear regression
    2. Support Vector Machines
    3. Random forest regression

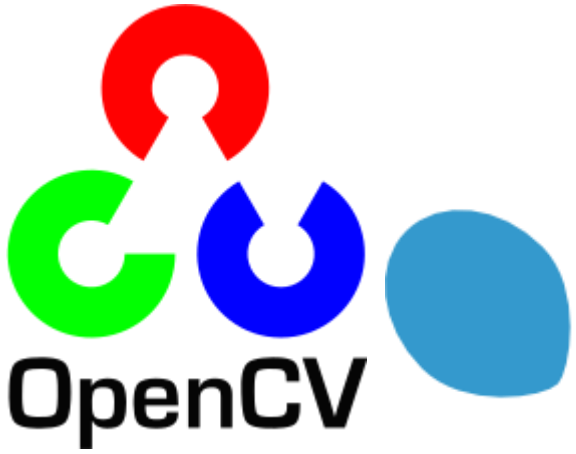
- High Level Classifiers (Deep Learning)
    1. Neural Networks

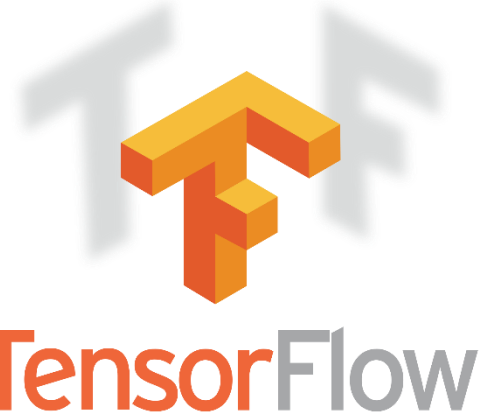
Why You Should Learn ML

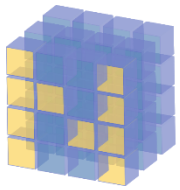
- It is the Future Industry
- Research Opportunities for Undergraduates / Post Graduates
- Data Scientist
- Big Data analysis
- AI Engineers
- AI based IOT and Mobile App Development
- Finance and Statistics Sector
- Stock Market Data Analysis, share value Prediction
- Automobile Industry
- Medical Industry

Why Python

- Interactive
- Interpreted
- Modular
- Dynamic
- Object-oriented
- Portable
- High level







NumPy

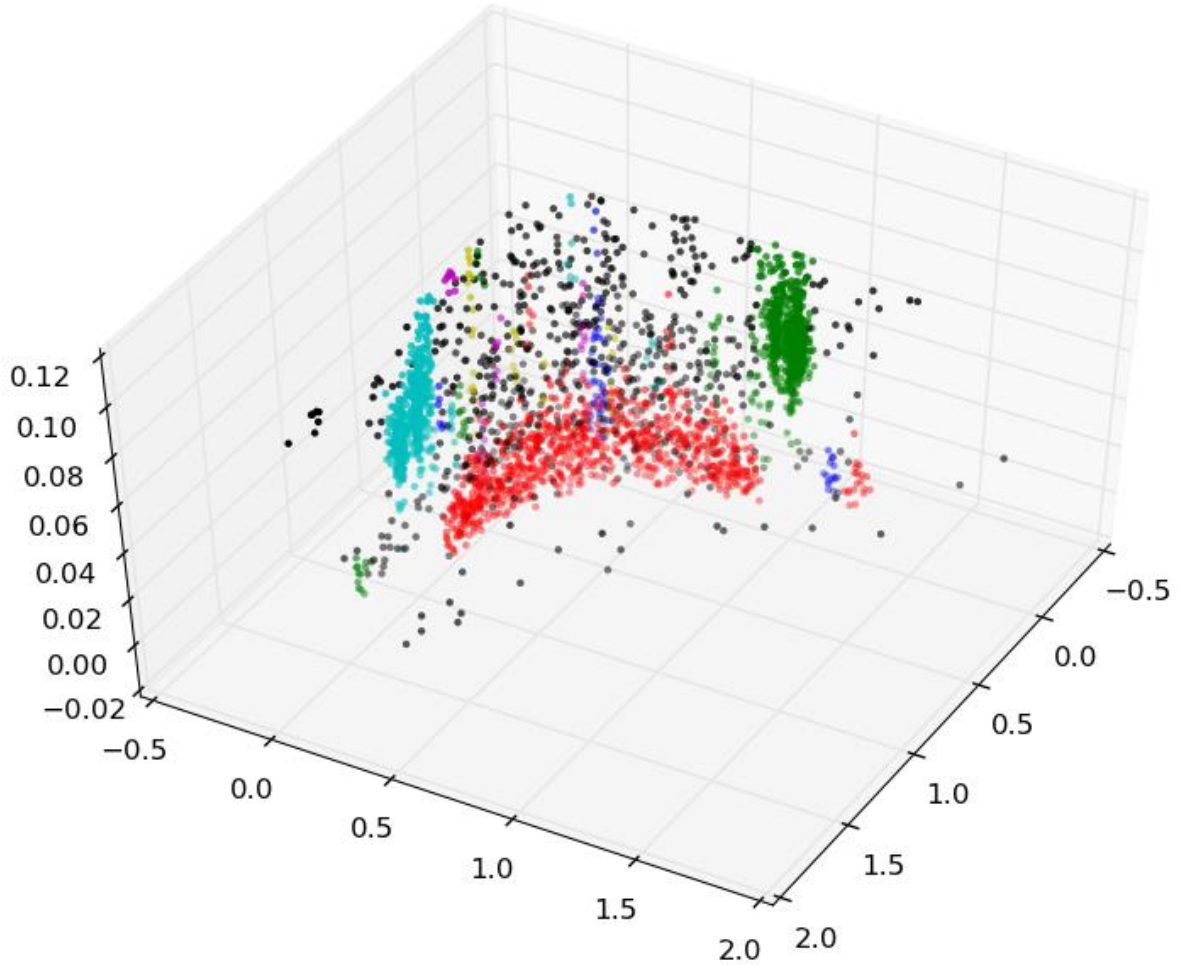
SciPy

Quandl

- Open Source advanced contributed Modules

Let’s get Started

1st Machine Learning classifier



Iris flower data set (1)

- The Iris flower data set or Fisher's Iris data set is a multivariate data set introduced by the British statistician and biologist Ronald Fisher in his 1936 paper The use of multiple measurements in taxonomic problems as an
- The data set consists of 50 samples from each of three species of Iris (Iris setosa, Iris virginica and Iris versicolor).
- Features from each sample: the length and the width of the sepals and petals, in centimeters.

Iris flower data set (2)

- The dataset contains a set of 150 records under five attributes - petal length, petal width, sepal length, sepal width and species.

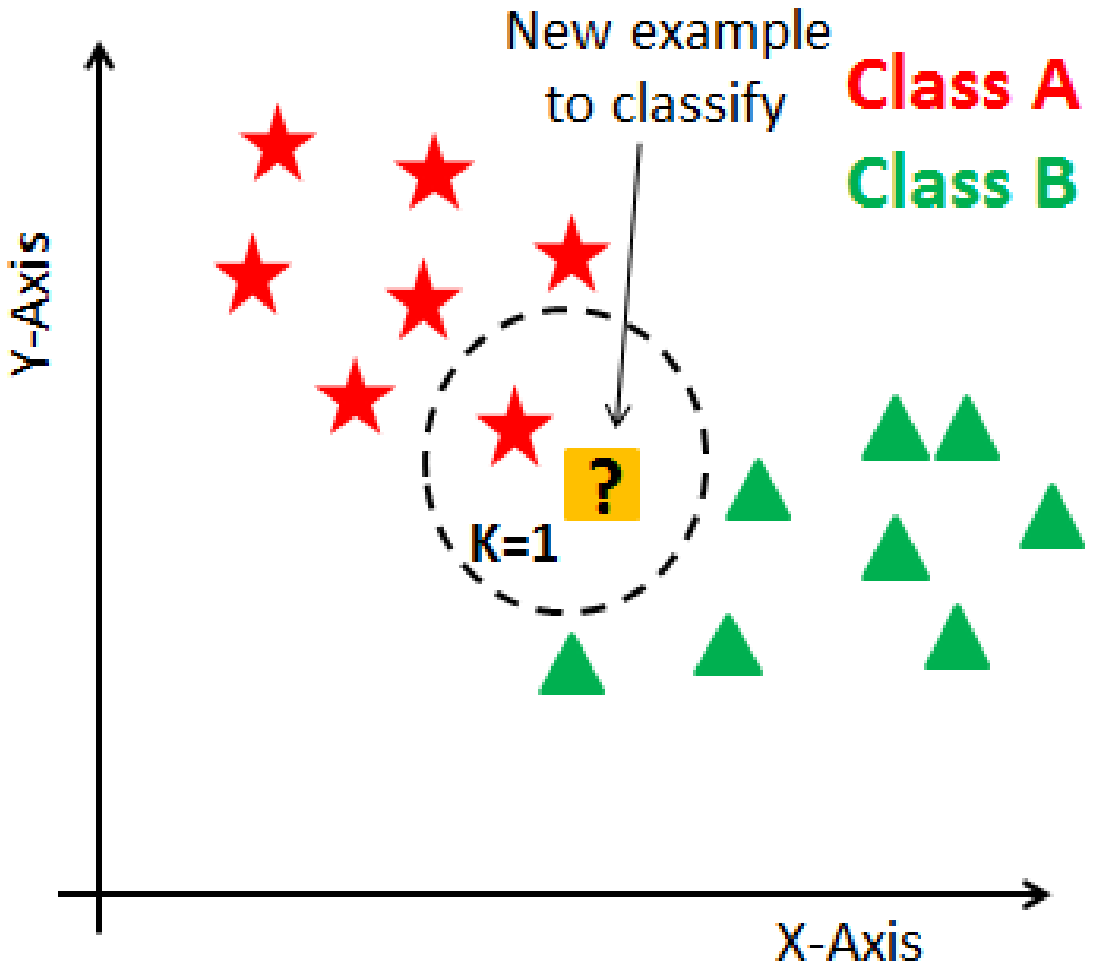
Dataset Order	Sepal length	Sepal width	Petal length	Petal width	Species
1	5.1	3.5	1.4	0.2	*I. setosa*
2	4.9	3.0	1.4	0.2	*I. setosa*
3	4.7	3.2	1.3	0.2	*I. setosa*
4	4.6	3.1	1.5	0.2	*I. setosa*
5	5.0	3.6	1.4	0.3	*I. setosa*
6	5.4	3.9	1.7	0.4	*I. setosa*
7	4.6	3.4	1.4	0.3	*I. setosa*
8	5.0	3.4	1.5	0.2	*I. setosa*
9	4.4	2.9	1.4	0.2	*I. setosa*

KNN Classifier(1)

- K Nearest Neighbor(KNN) is a very simple, easy to understand, versatile and one of the topmost machine learning algorithms.
- KNN used in the variety of applications such as finance, healthcare, political science, handwriting detection, image recognition and video recognition.
- In Credit ratings, financial institutes will predict the credit rating of customers
- In loan disbursement, banking institutes will predict whether the loan is safe or risky. In political science, classifying potential voters in two classes will vote or won’t vote.
- KNN algorithm used for both classification and regression problems. KNN algorithm based on feature similarity approach.

KNN Classifier(2)

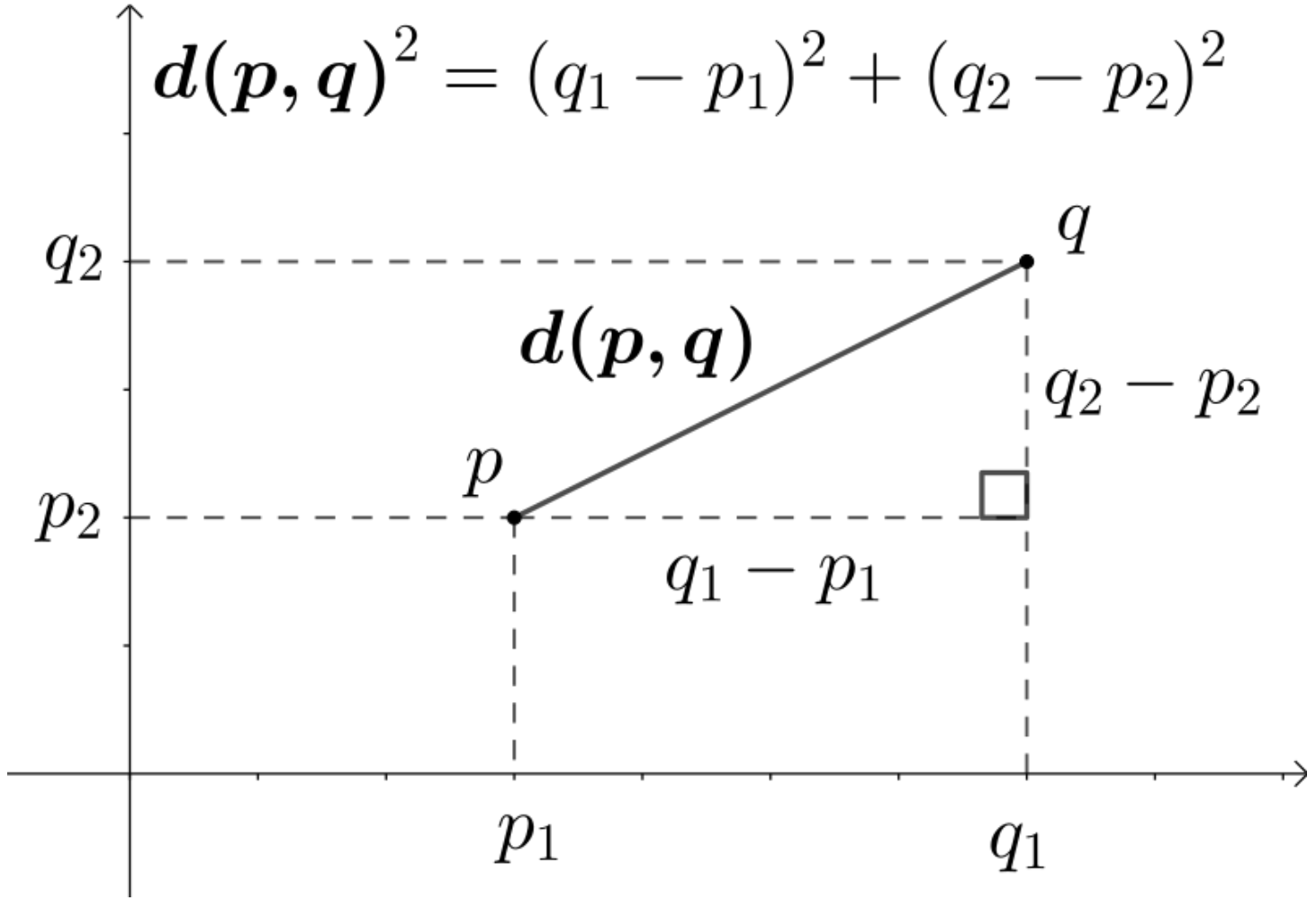
- In KNN, K is the number of nearest neighbors. The number of neighbors is the core deciding factor.
- K is generally an odd number if the number of classes is 2. When K=1, then the algorithm is known as the nearest neighbor algorithm. This is the simplest case.
- Suppose P1 is the point, for which label needs to predict. First, you find the one closest point to P1 and then the label of the nearest point assigned to P1.



KNN Classifier(3)

- Suppose P1 is the point, for which label needs to predict.
- First, you find the k closest point to P1 and then classify points by majority vote of its k neighbors.
- Each object votes for their class and the class with the most votes is taken as the prediction.
- For finding closest similar points, you find the distance between points using distance measures such as Euclidean distance, Hamming distance, Manhattan distance and Minkowski distance.

Euclidean Distance (1)



Euclidean Distance (2)

Three dimensions [edit]

In three-dimensional Euclidean space, the distance is

$$d(\mathbf{p}, \mathbf{q}) = \sqrt{(p_1 - q_1)^2 + (p_2 - q_2)^2 + (p_3 - q_3)^2}.$$

n dimensions [edit]

In general, for an *n*-dimensional space, the distance is

$$d(\mathbf{p}, \mathbf{q}) = \sqrt{(p_1 - q_1)^2 + (p_2 - q_2)^2 + \cdots + (p_i - q_i)^2 + \cdots + (p_n - q_n)^2}.$$

KNN Classifier(4)

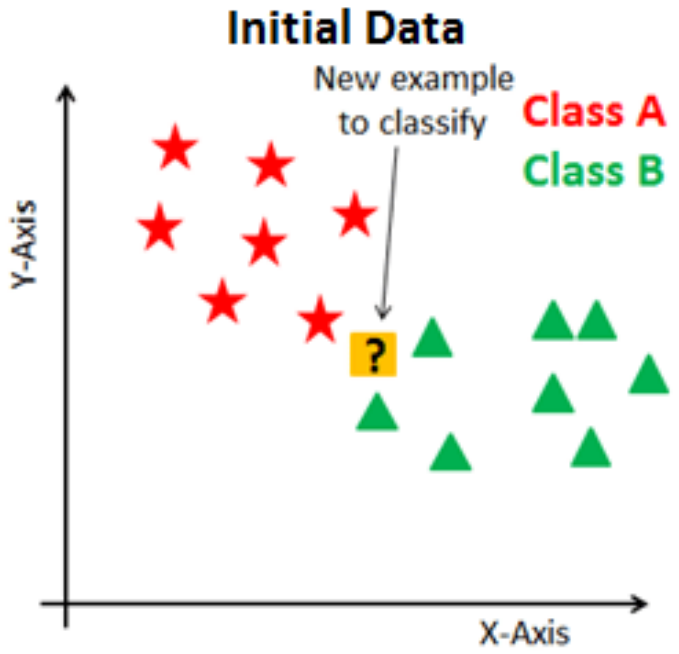
Calculate Distance

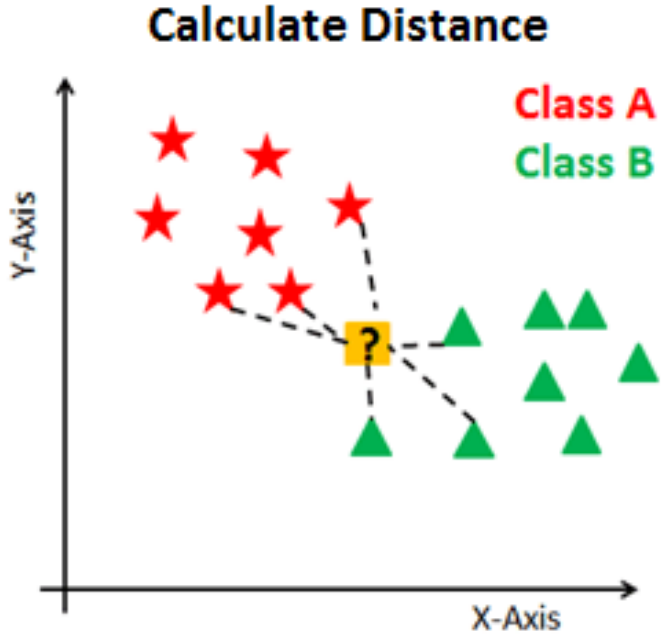
↓

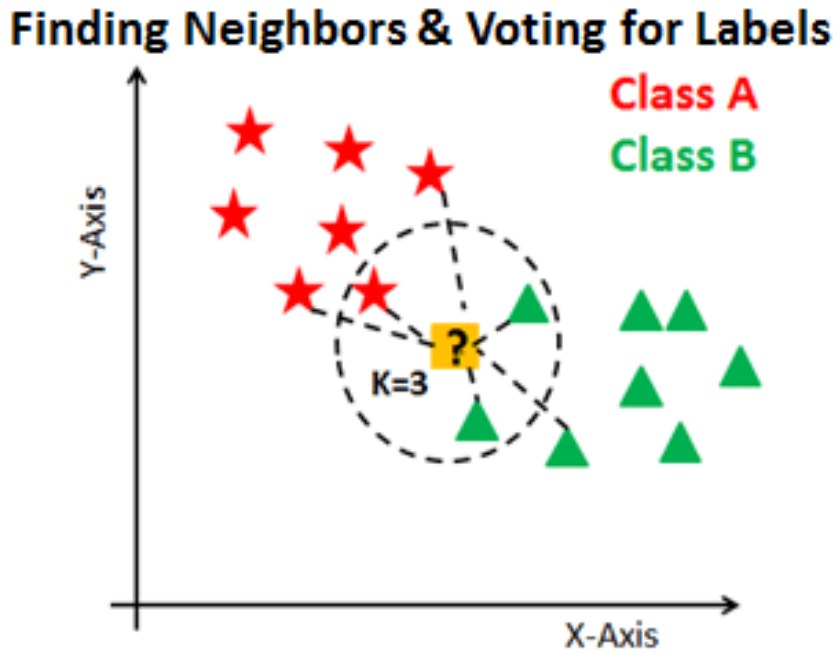
Find Closest Neighbors

↓

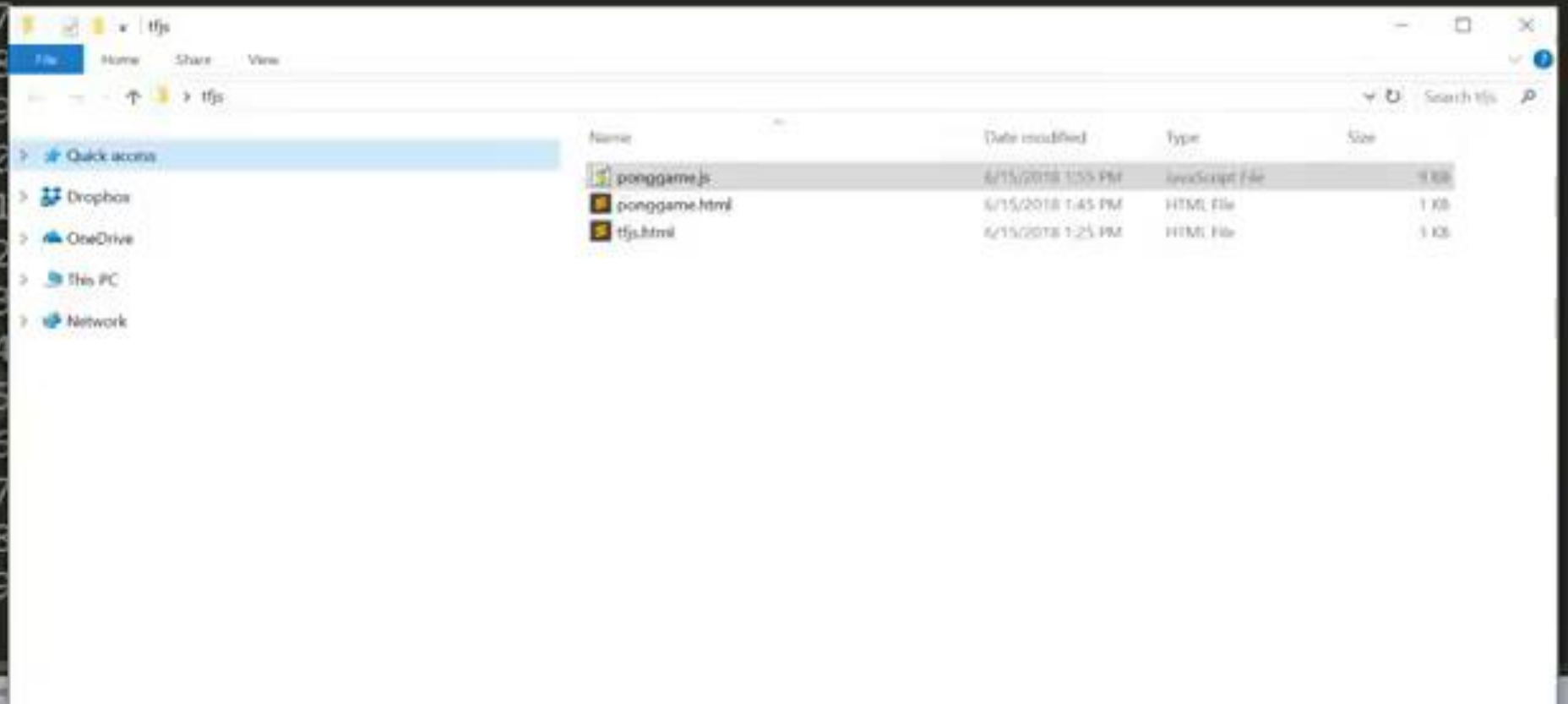
Vote for Labels







```
309         //do prediction here
310         //return -1/0/1
311         prediction = model.predict(tf.tensor([this.last_data_object]));
312         return tf.argMax(prediction, 1).dataSync()-1;
313     }
314
315 }
316
317
318
319
320
321
322
323
324
325
326
327
328
329
```




```
ep 13 2015, 02:

nse()" for more

-gamging\video-

s
ds
ds
s
ds
ds
Frame took 0.21426701545715332 seconds
```

```
        PressKey(W)
        ReleaseKey(A)
        ReleaseKey(D)

def left():
        PressKey(A)
        ReleaseKey(W)
        ReleaseKey(D)
```

Image Processing

& Computer Vision

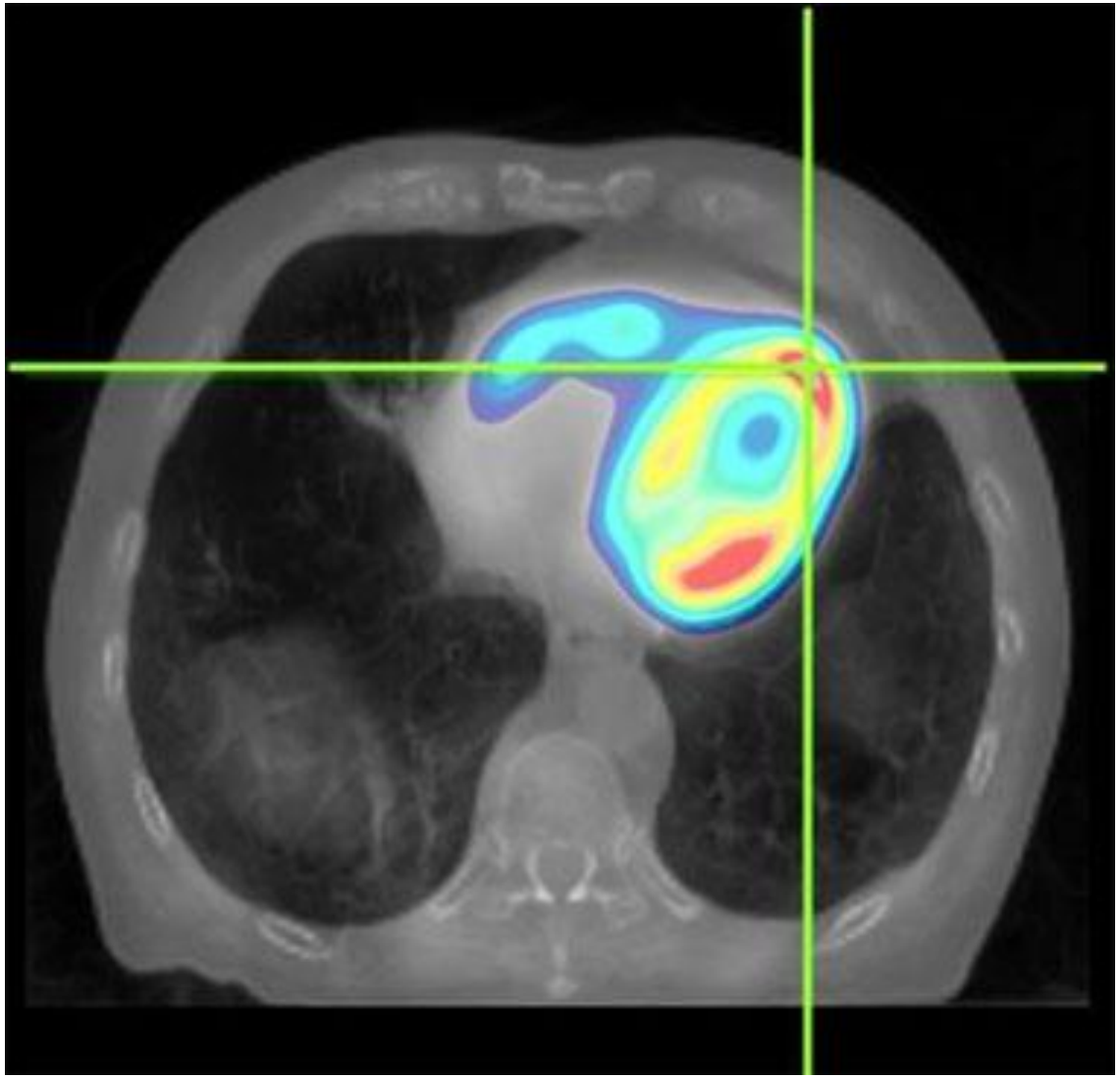


Image Representation

- A digital image is composed of M rows and N columns of pixels each storing a value
- Pixel values are most often grey levels in the range 0-255(black-white) or R,G,B values
- images can easily be represented as matrices

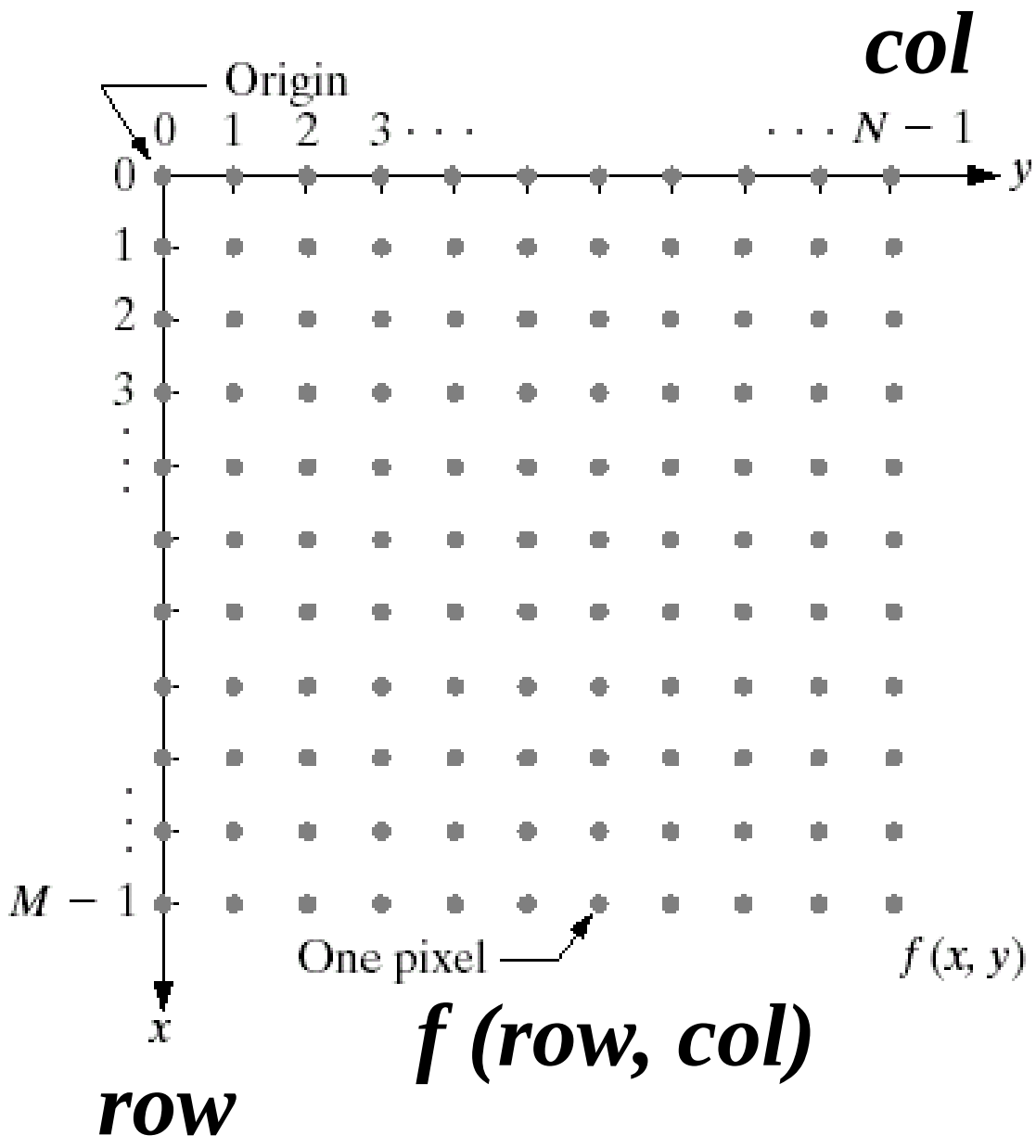
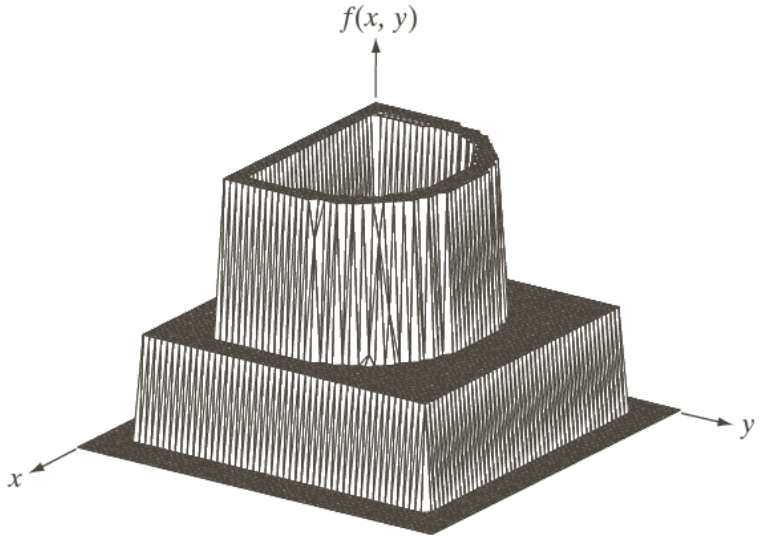
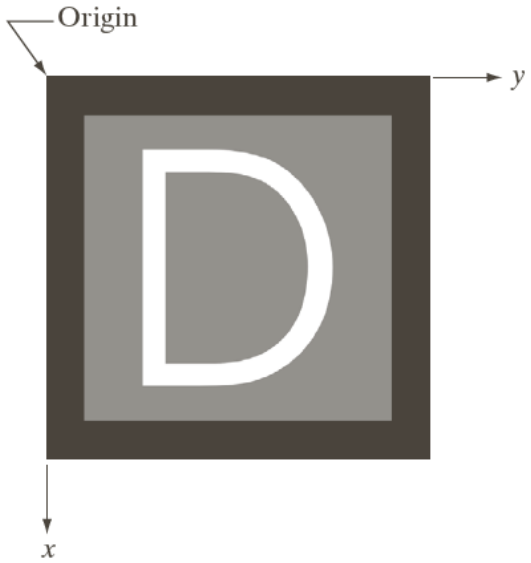


Image Representation





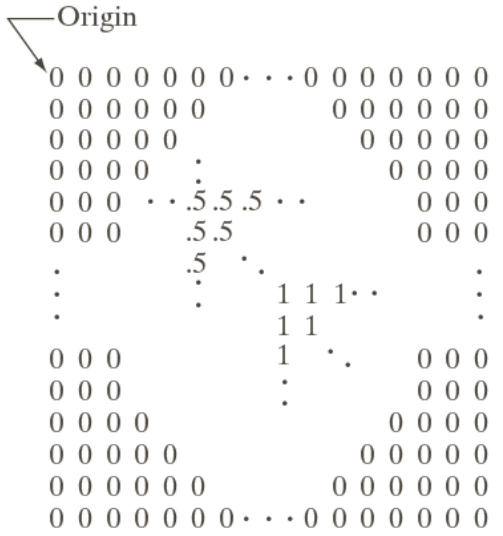


Image Representation

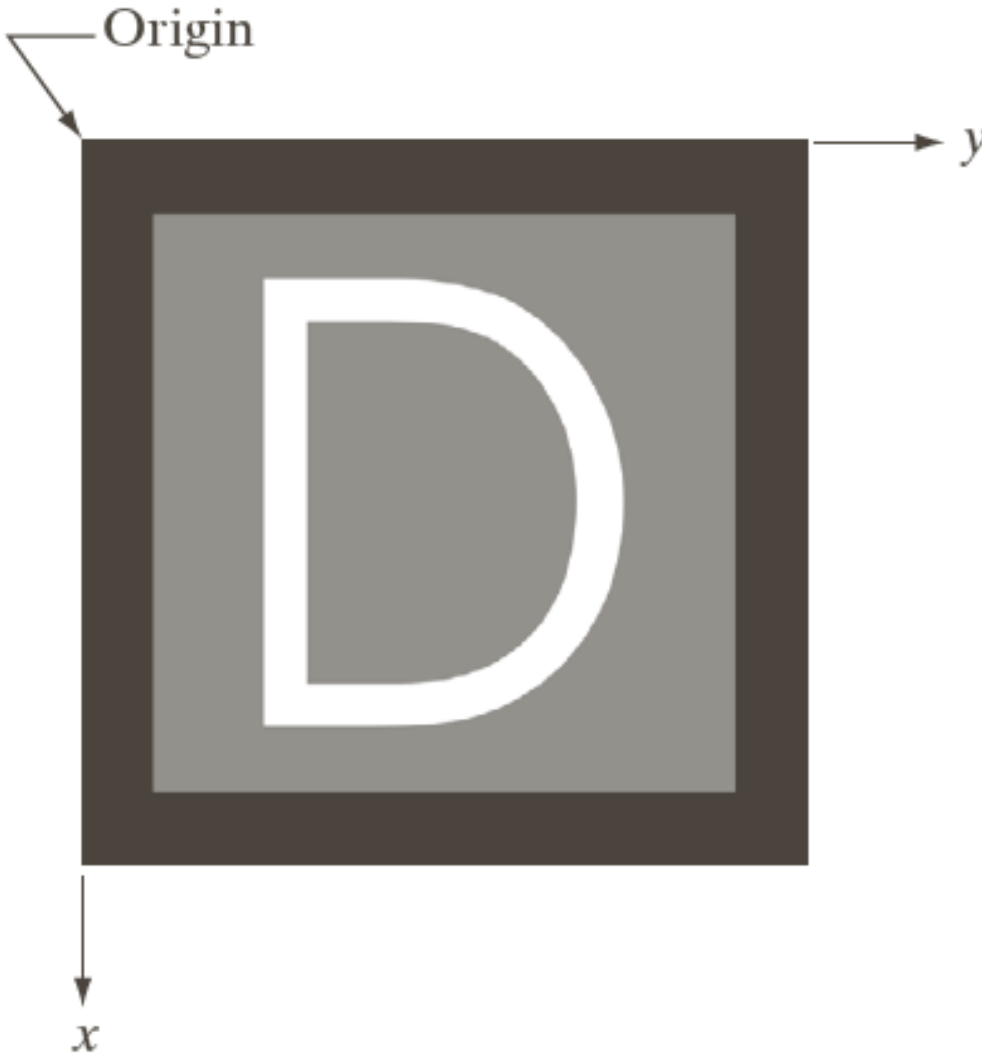
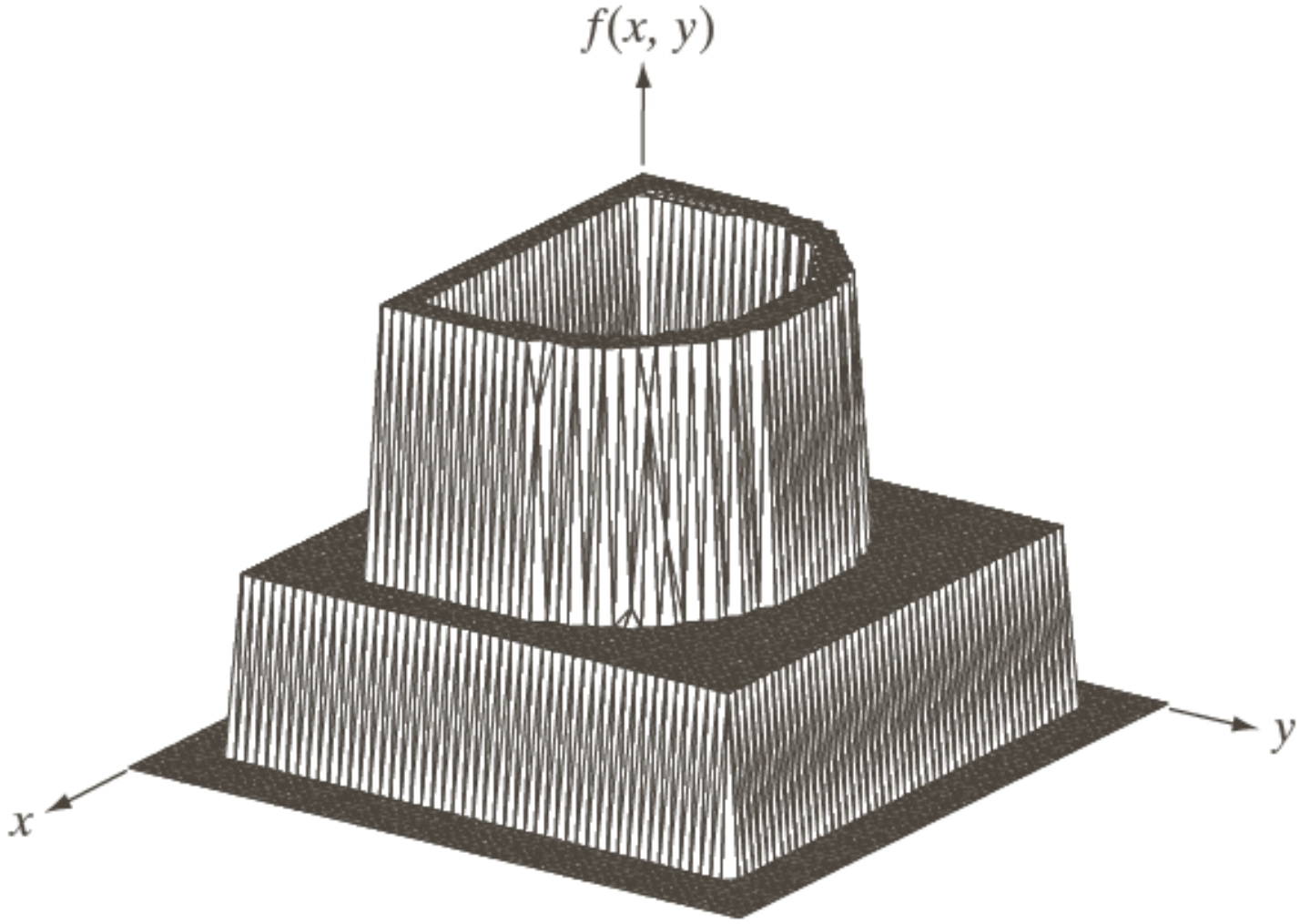


Image Representation



ML & IP 3 Months Course

- **STAGE 1: Python Programming Basics and Essential Modules (W1)**

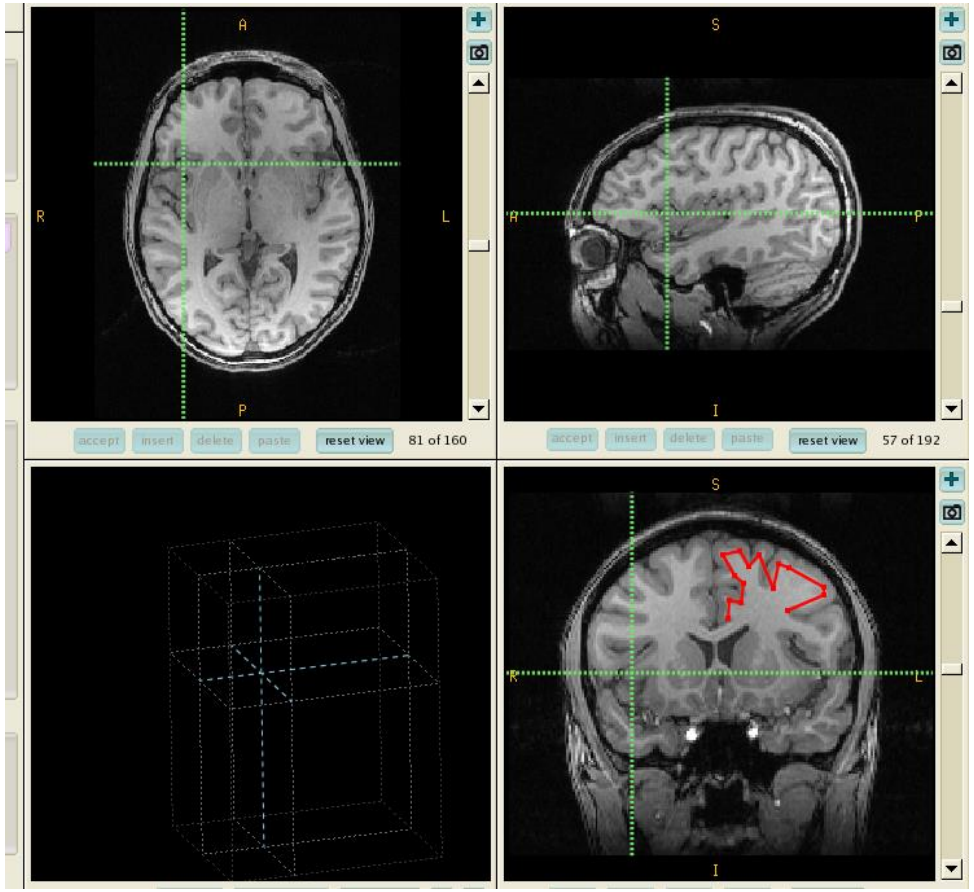
  - (WEEK 1-2)
  - Lists, Tuples, Dictionaries
  - Functions,Modules
  - Files I/O
  - OOP Concepts
  - MatPlotLib
  - Numerical Python (NumPy)
  - Scientific Python (SciPy)

STAGE 2: LOW LEVEL IMAGE PROCESSING & MACHINE LEARNING BASICS (W2-6)

- **Machine Learning Basics**
  - Linear Regression, R Squared Method, Euclidean Method
  - Classification Intro with K Nearest Neighbors
  - Support Vector Machine
  - Kernels, Soft Margin SVM, and Quadratic Programming
  - Decision Tree Classification
  - Creating, Training and Testing Your Own Classifiers
- **Low Level Image Processing**
  - Drawing Functions for Image Processing
  - Basic Operations on Images
  - Arithmetic Operations
  - Color Spaces and Geometric Transformation
  - Image Thresholding and Morphological Operations
  - Feature Matching and Video Analysis

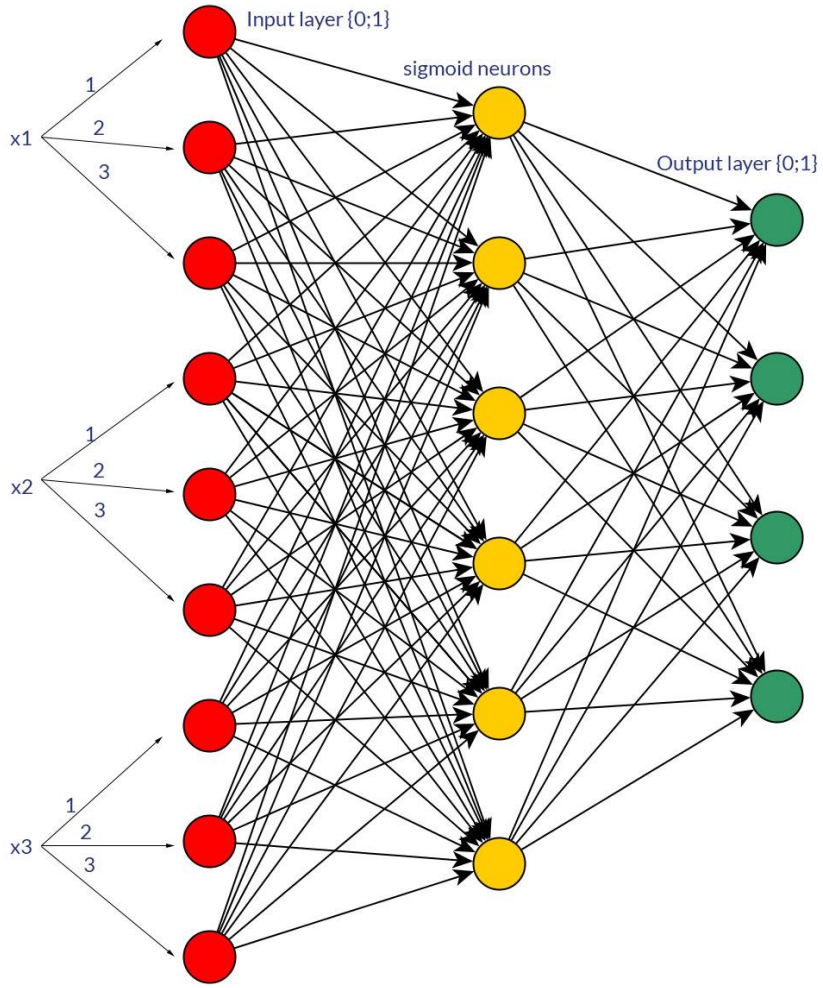
STAGE 3: MACHINE LEARNING ALGORITHMS FOR IMAGE PROCESSING (W6-7)

- OCR of Hand-written Data using kNN
- OCR of Hand-written Data using SVM
- K-Means Clustering
- Cascade Classifiers
- Making your own Cascade Classifier



STAGE 4: DEEP LEARNING (W8-12)

- Introduction to Neural Networks with Tensor-Flow
- Creating the Neural Network Models
- Deep Learning with our own Data
- Image and Basic Pattern Recognition
- Recurrent and Convolutional Neural Networks (CNN) (RNN)
- High Level Abstraction Layerfor TensorFlow
- 3D Convolutional Neural Network
- Unconventional Neural Networks
- Deep Learning Chat-bot
- TensorFlow Object Detection API



COMPULSORY FINAL PROJECT

- Everyone has to involve in a Group Projects
- Topics will be given in the 6th Week
- The Final Grade of the Certificate will be decided by the performances