

# Machine Learning

## Lect 1 : Introduction

Dr. PRERANA MUKHERJEE

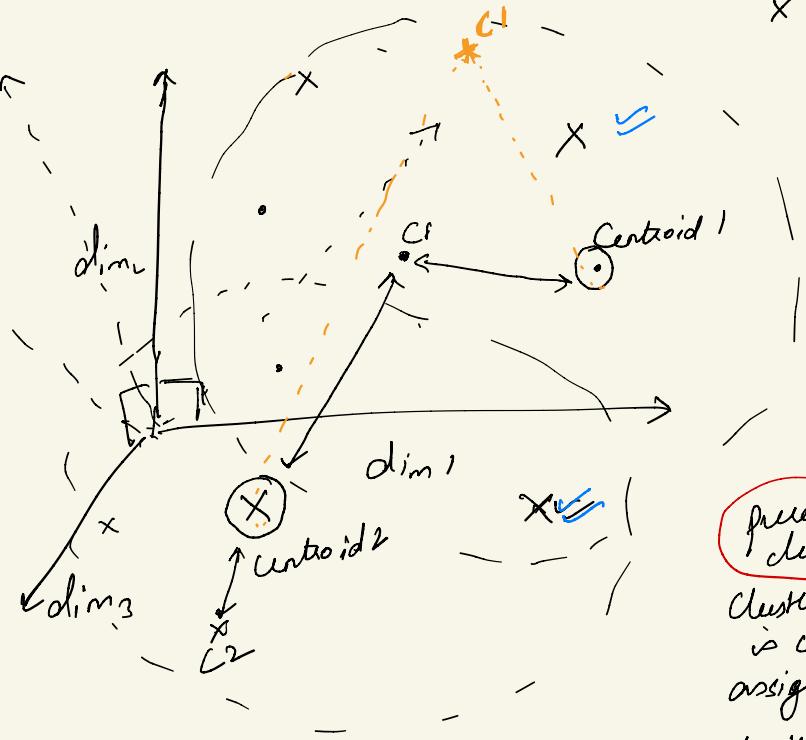


Training samples  $\downarrow$   $\downarrow$   $\downarrow$  features  
 $\rightarrow \mathcal{X}_1 = \{d_1, d_2, \dots, d_n\} \in \mathbb{R}^n$   
 $\rightarrow \mathcal{X}_2 = \{d_2, d_4, \dots, d_n\}$   
 $\rightarrow \mathcal{X}_3 = \{d_3, d_4, \dots, d_n\}$   
 $\vdots$   
 $\mathcal{X}_n = \{d_5, d_6, \dots, d_n\}$

wiki  
 {Distance metrics  
Unsupervised algorithms

k-means Input "k" = # clusters

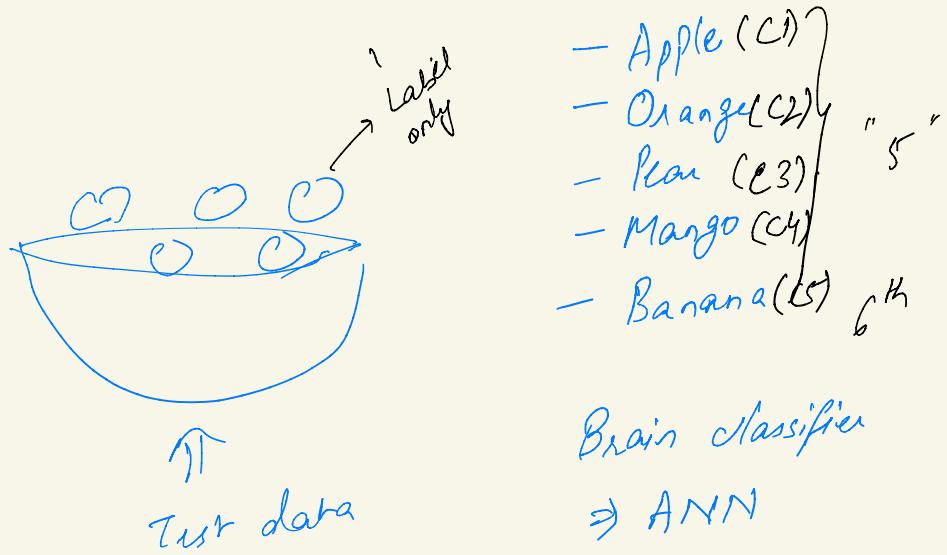
- Cluster 1
- Cluster 2



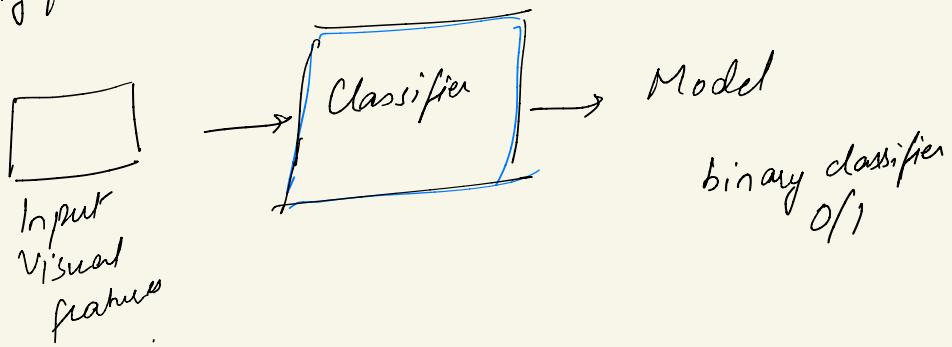
stopping criteria  $\Rightarrow \checkmark$

Cluster 1 is correctly assigned  
 Cluster 2 is correctly assigned

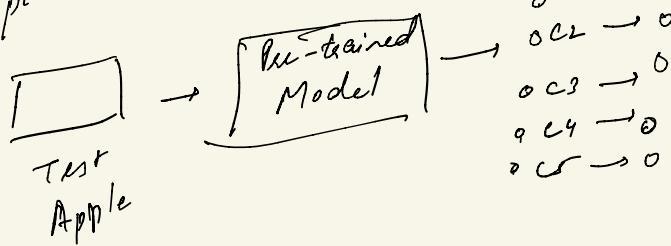
Supervised Setting  
Train a Classifier  $\Rightarrow$  SVM, kNN, Bayes, DNN, ...  
CNN, LSTM

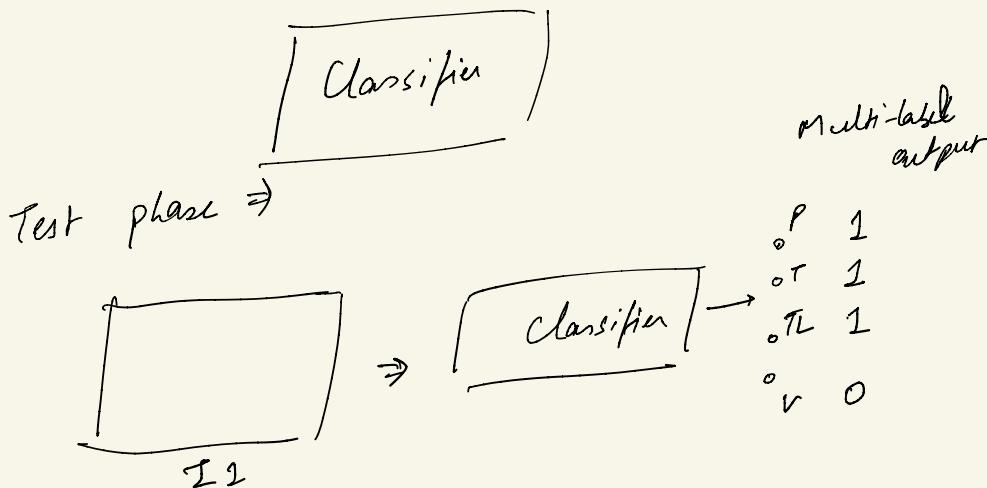
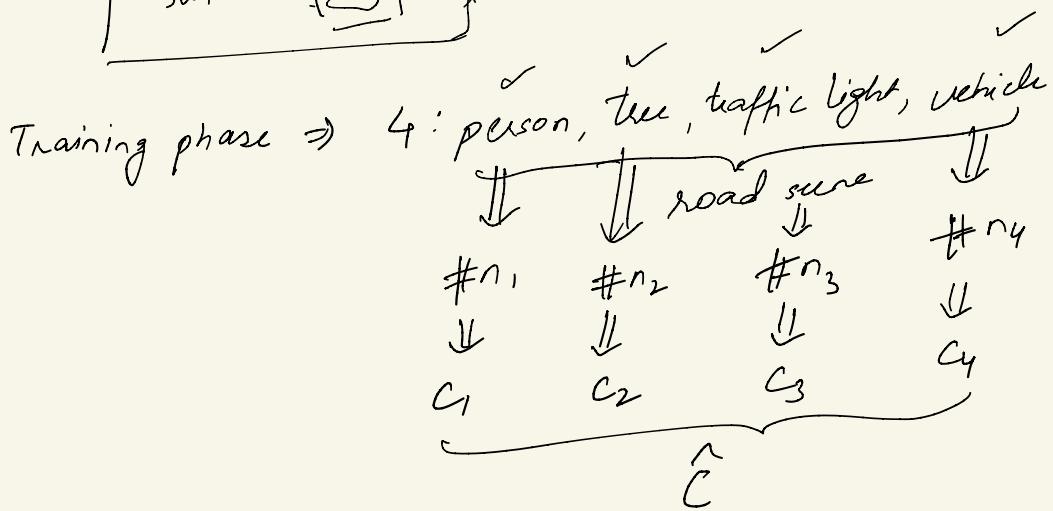
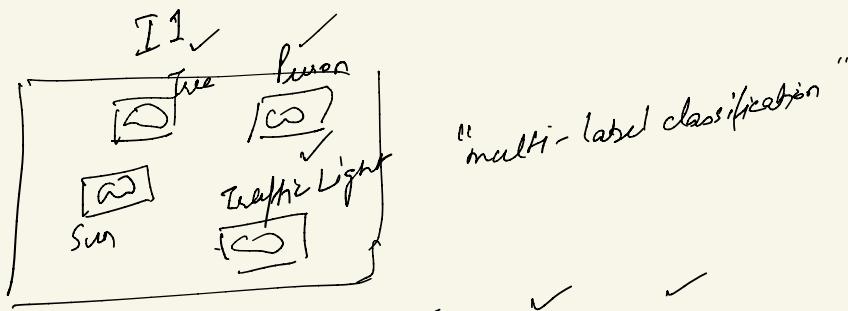


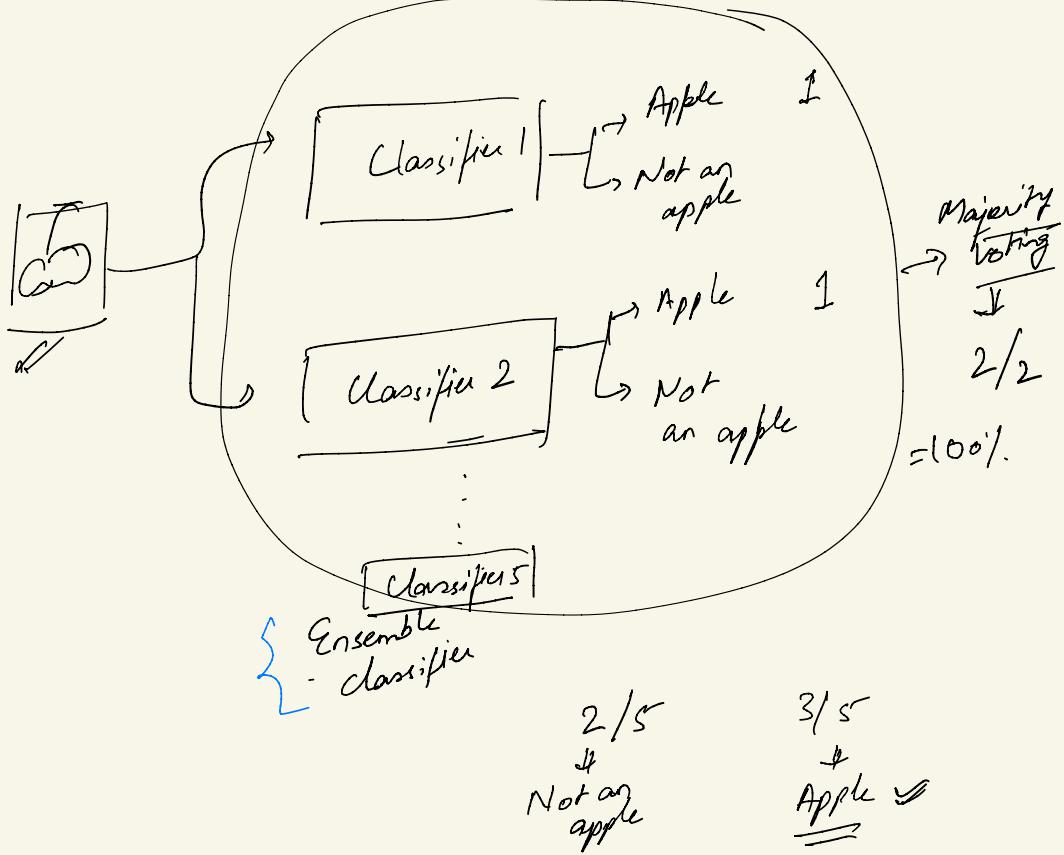
Training phase:



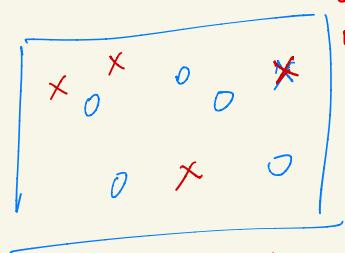
Test phase:







Semi supervised setting  $\Rightarrow$  EM, R.L.



Expectation maximization  $\Rightarrow$  Reinforcement learning

$\Rightarrow$  Regression  $\Leftrightarrow$

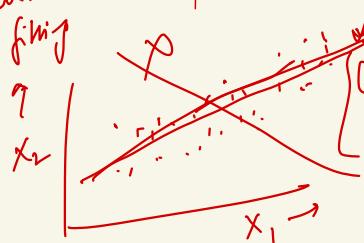
Linear +  
Non-linear

$\Rightarrow$  Classification

Unsupervised  
Supervised  
Semi-supervised



$x_0$   
 $x_1$



Curve fitting  
Line fitting

$x_1$

# # Purchase of Laptop

- Budget ( $30k < \text{INR} < 70k$ )
- cores of processor (CPU) ( $8 < \# \text{cores} < 12$ )
- RAM ( $> 4GB$ )
- Processor generation ( $> i5$ )
- Operating System (Windows 10, Ubuntu)

