Introduction to Machine Learning

How to Humans Learn ?

Ever happen to do this in your childhood?

This is what happens

- We are repeatedly shown samples of objects
- Correct answers also shown
- We make a model of this in our minds
- Models are made based on features of the object
- Later on we use this model to identify

Machine Learning

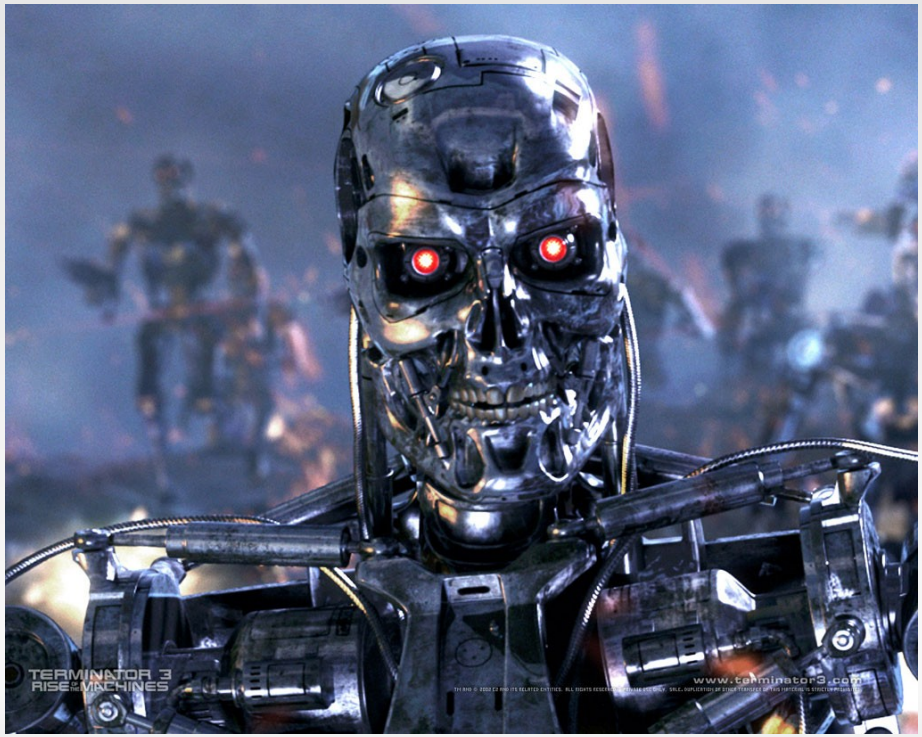
- Give machine training samples
  - Images with correct answers
- Extract features of objects
- Make a model
- Use this model to predict, when shown new sample

*Machine Intelligence is not the same as artificial intelligence

**”Pattern Recognition”, ”Machine Learning”, ”Machine Intelligence” all refer to the above process

Machines taking over Humans?

- Just a myth
- I hope at end of this intro you'll know why



Distinquishing Mango and Banyan Tree Leaves

Mango Leaf

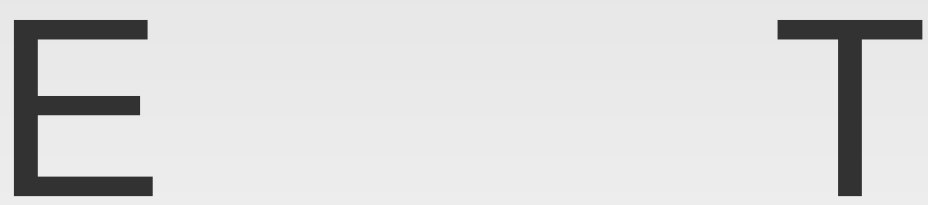
Banyan Tree Leaf

What are Features ?

- Quantities which characterize an object
- To be extracted from segmentation results
- Shape features, texture features, color features

Machine Learning Process

Application: Character Recognition



What features distinquish E & T ??

Application: Character Recognition

- Get multiple sample and segment characters
- Features :
  - Area
  - Perimeter
  - Number of right angles
- Create model

How to Create Models ?

- Bayes Rules
- **K-Nearest Neighboors**
- Neural Networks
- **SVM (Support Vector Machines)**

K-Nearest Neighbours

- Find closeness of the testing feature
- For eg. K=2

Feature Set	Manually Assigned Labels
(100,100,8)	T
(108,90,7)	T
(110,170,12)	E
(120,160,13)	E
(115,160,11)	E

(101,90,8)

*K-Means is very much different from K-NN

Implementing K-NN in Lab

- Mango-banyan tree measures provided
  - Width, height of 40 leaves
- You will be given a new width and height
- Closeness measures : Euclidean distance
- By K-NN you will predict the outcome.
- Take K=5

*K-Means is very much different from K-NN

Support Vector Machines : SVM

- Another method for creating models
- Give very good recall
- In wide use as a black-box
- Code will be provided
- Two modes of operation : Training & Testing

Training Mode : svm_learn()

- Input : Features and labels
- Output : Model

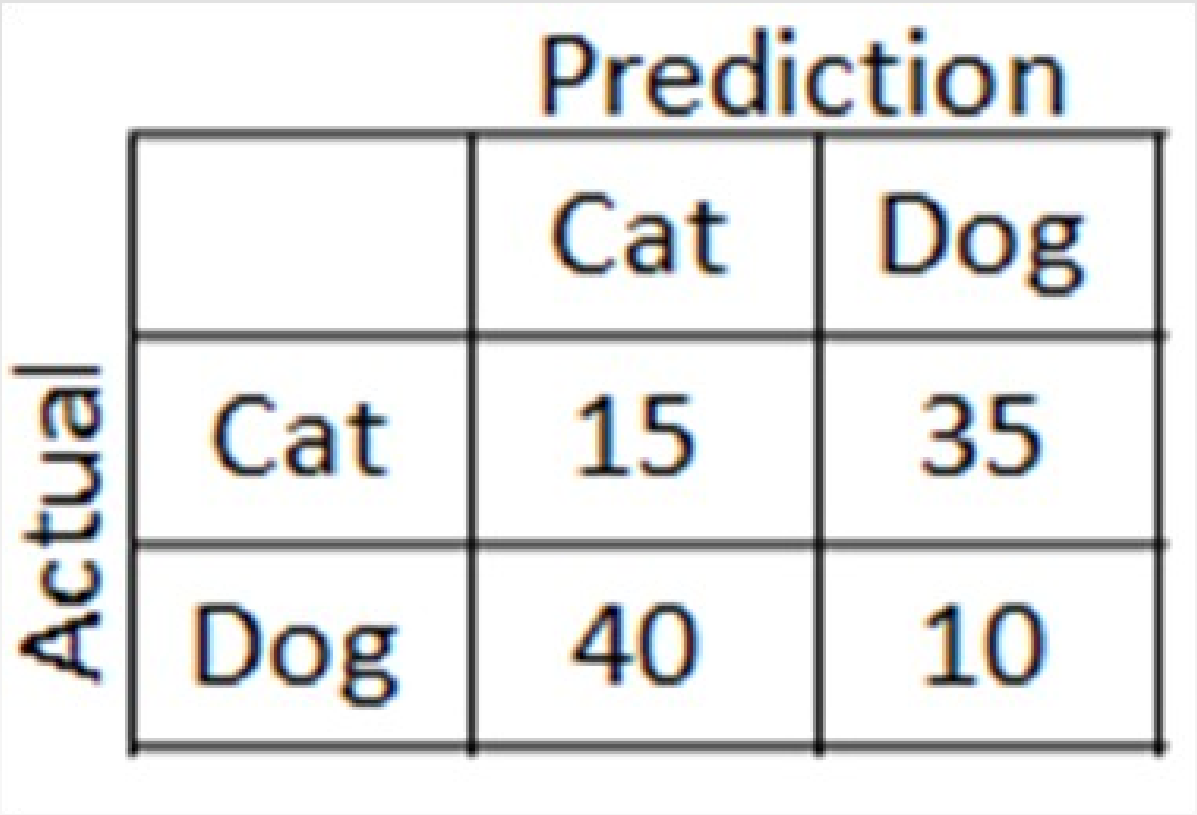
Feature Set	Manually Assigned Labels
(100,100,8)	T
(108,90,7)	T
(110,170,12)	E
(120,160,13)	E
(115,160,11)	E

Testing Mode : svm_classify()

- Input : Model
- Output : predicted label

Evaluating Machine's Intelligence

- Confusion Matrix



Actual \ Prediction	Cat	Dog
Cat	15	35
Dog	40	10

*This example for 2-class classification. All this can be generalize for N-class

Evaluation Metrics

	p' (Predicted)	n' (Predicted)
P (Actual)	True Positive	False Negative
n (Actual)	False Positive	True Negative

Accuracy, Sensitivity etc.

- Sensitivity
  - % of actual positives which are correctly identified
  - TP / (TP + FN)
- Accuracy
  - % of predictions which were correct
  - (TP + TN) / (TP + TN + FP + FN)

http://en.wikipedia.org/wiki/Sensitivity_and_specificity

Evaluating Machine's Intelligence

- Sensitivity : 15 / (15 + 35) = 0.3
- Accuracy : (15 + 10) / (15 + 10 + 40 + 35) = 0.25

		Prediction	
		Cat	Dog
Actual	Cat	15	35
	Dog	40	10

*This example for 2-class classification. All this can be generalize for N-class

Thanks :)