Machine Learning

Shahid Mahmood Awan, PhD

AI.Meetup

31st March 2018

University of Management and Technology, Lahore

Can you read it?

IN73LL1G3NC3
15 7H3 4B1L17Y
7O 4D4P7 7O
CH4NG3.

- 573PH3N H4WK1NG

Can you recognize this?

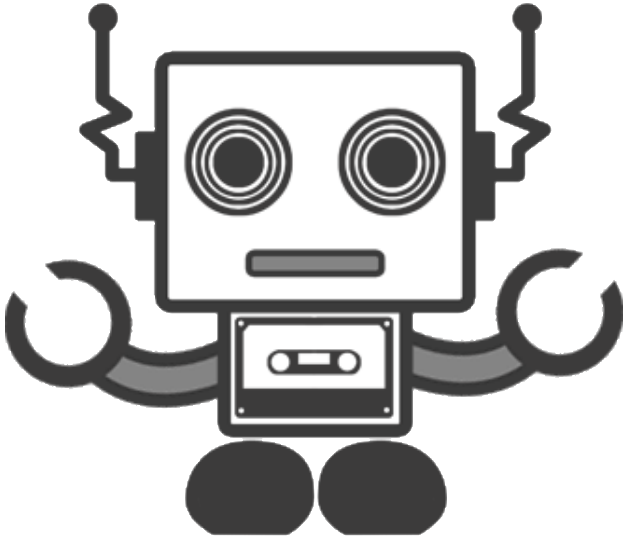
9

Recognize the Fruit…

What is Learning?

What is Machine Learning?

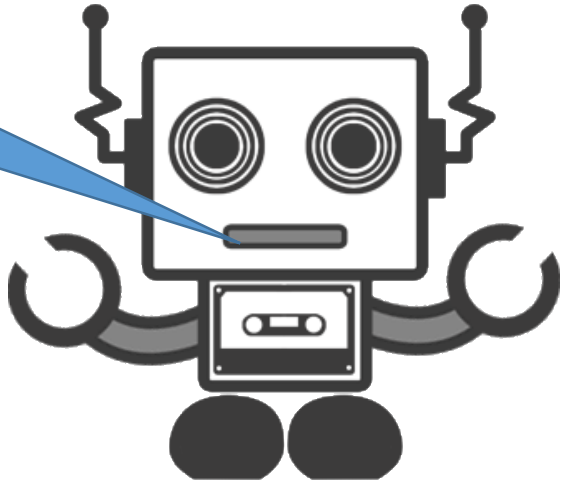
How do we learn?



What can machines do?

What is Machine Learning?

Can we make them learn from Experience?



Experience = Data

Model that fits the data

Follow Instructions

Learn from Experience

Make a mental model out of observed data and the re-tune it after seeing more and more data?

What is Machine Learning?

Traditional Programming

Machine Learning

Why do we need to learn?

- Example Problem: **Face Recognition**

- **Can you hand write all the rules to recognize a particular face?**

Why do we need to learn?

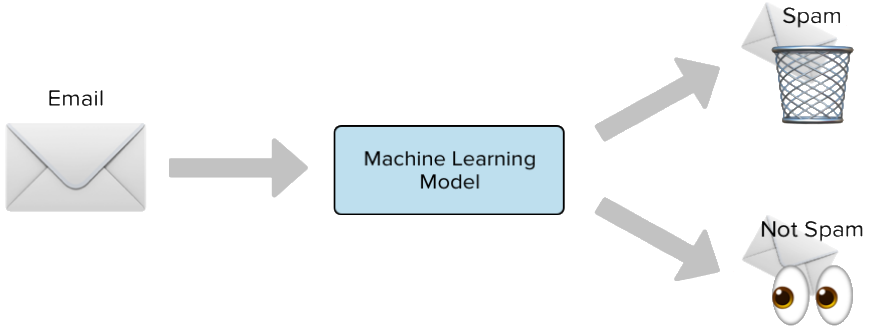
- Example Problem: **Face Recognition**

- **Training data:** a collection of images and labels (names)
- **Evaluation criterion:** correct labeling of new images

Why do we need to learn?

- Example problem: spam detection.
- Data consists of information from 4601 email messages, in a study to try to predict whether the email was “spam”. The data were collected in Hewlett-Packard labs and donated by George Forman.

- **Objective**: design an automatic spam detector.
- **Supervised learning problem**: class variables email/spam classification problem.



Why do we need to learn?

- Machine learning is programming computers to optimize a performance criterion using example data or past experience.

- **There is no need to "learn" to calculate payroll**

- Learning is used when:
  - Human expertise does not exist (navigating on Mars),
  - Humans are unable to explain their expertise (speech recognition)
  - Solution changes in time (routing on a computer network)
  - Solution needs to be adapted to particular cases (user biometrics)

What is Machine Learning?

- Learning is any process by which a system improves performance from experience.” ~ Herbert Simon

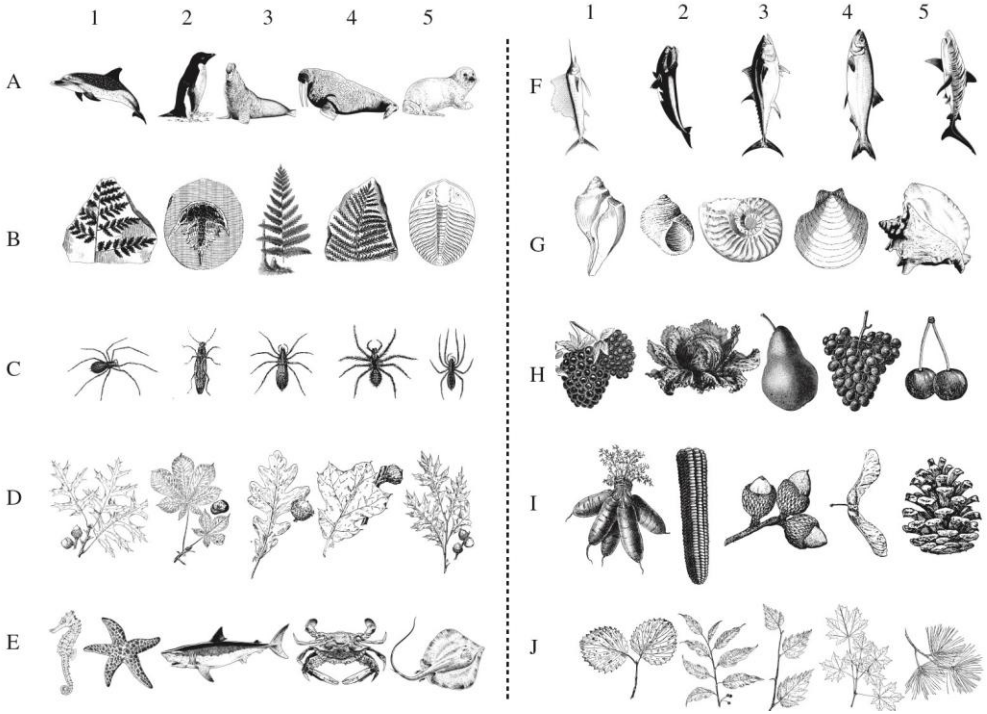
- Definition by Tom Mitchell (1998):

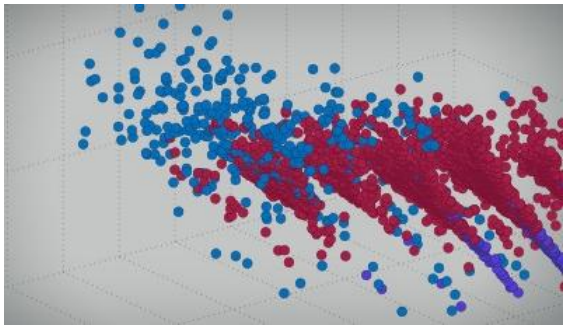
  Machine Learning is the study of algorithms that
  - improve their performance P
  - at some task T
  - with experience E.

  A well-defined learning task is given by <P, T, E>.

What is Machine Learning?

Machine Learning is only possible because there is a structure/pattern in this world





Types of Learning

- **Supervised** (inductive) learning
  – Given: training data + desired outputs (labels)

- **Unsupervised** learning
  – Given: training data (without desired outputs)

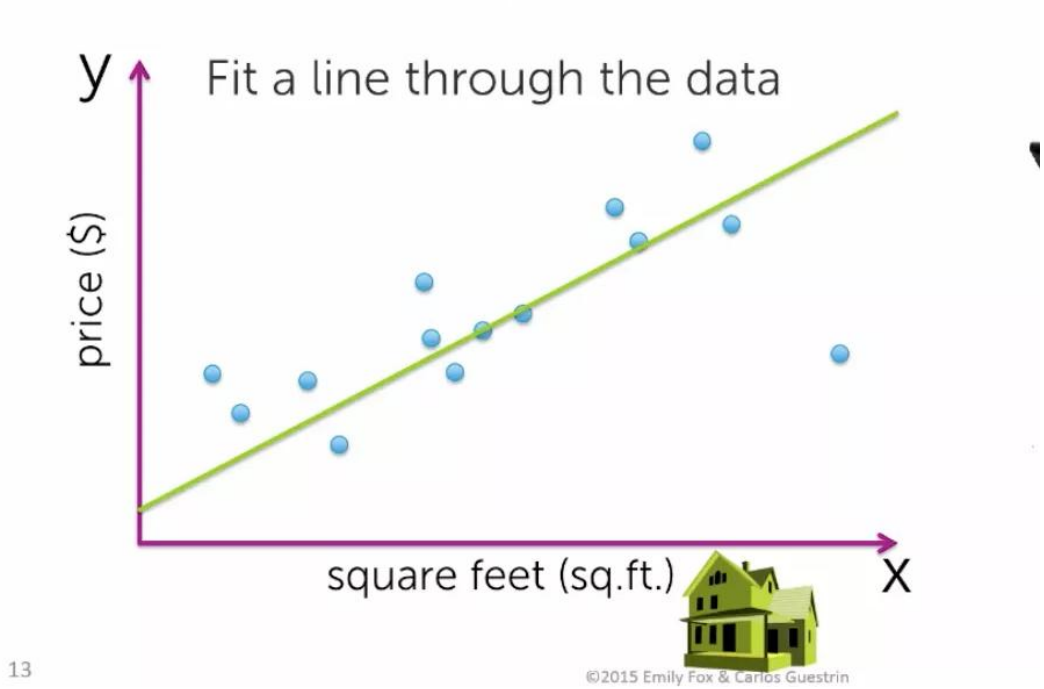
- **Semi-supervised** learning
  – Given: training data + a few desired outputs

- **Reinforcement** learning
  – Rewards from sequence of actions

Training data = Feature Vectors extracted from the raw data

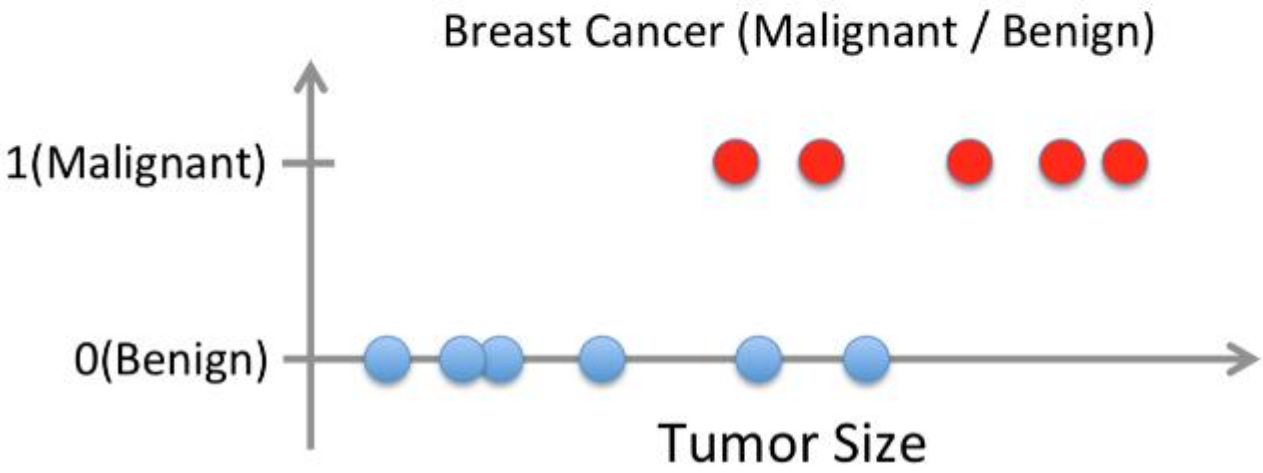
Supervised Learning: Regression

- Given $(x_1, y_1), (x_2, y_2), \ldots, (x_n, y_n)$
- Learn a function $f(x)$ to predict y given x
  - y is real-valued == regression

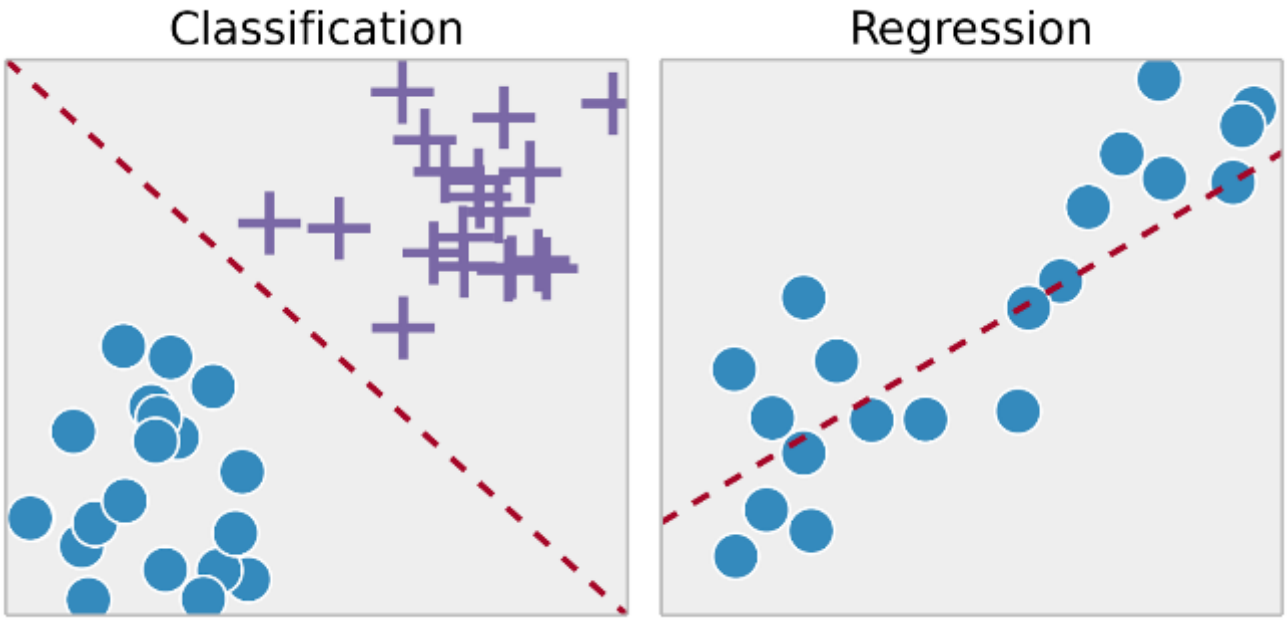


Supervised Learning: Classification

- Given $(x_1, y_1), (x_2, y_2), \ldots, (x_n, y_n)$
- Learn a function $f(x)$ to predict y given x
  - y is categorical == classification

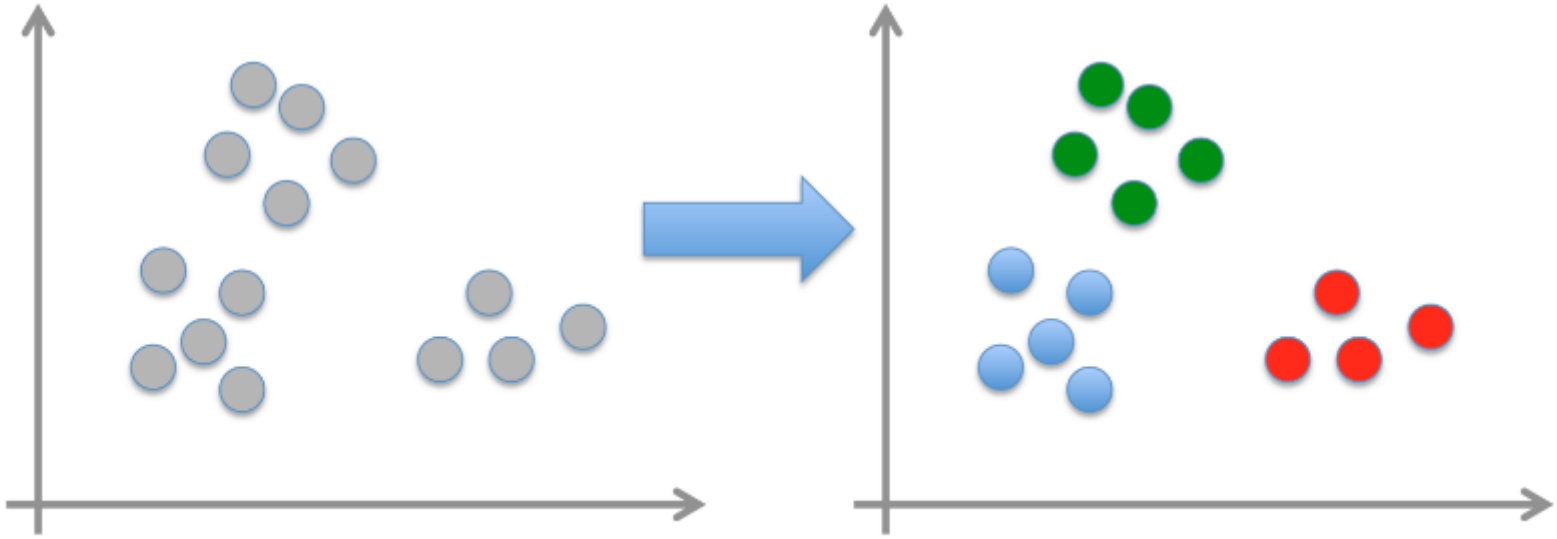


Supervised Learning



Unsupervised Learning

- Given $x_1, x_2, \ldots, x_n$ (without labels)
- Output hidden structure behind the x's
  - E.g., clustering

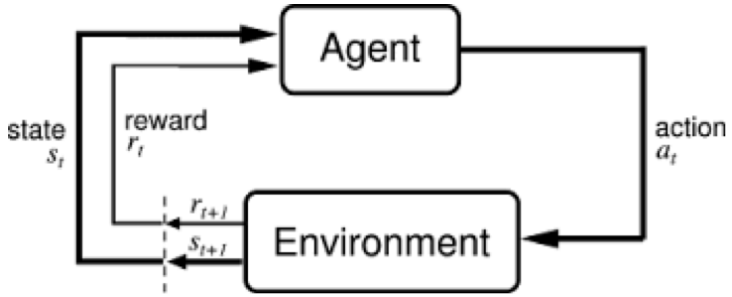


Reinforcement Learning

- Given a sequence of states and actions with (delayed) rewards, output a policy
  - Policy is a mapping from states -> actions that tells you what to do in a given state

- Examples:
  - Credit assignment problem
  - Game playing
  - Robot in a maze
  - Balance a pole on your hand

Reinforcement Learning



ML in a Nutshell

- Every ML algorithm has three components:

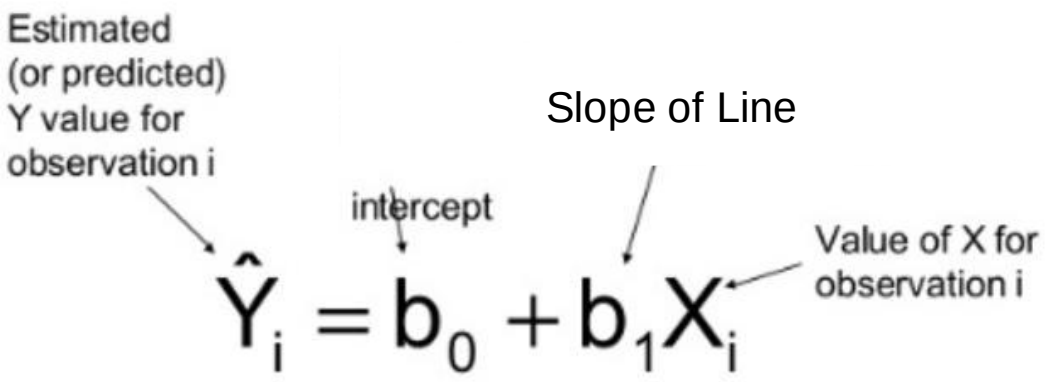
- **Representation**
  - (Linear Regression, Neural Networks, SVM, Decision Trees, Naïve Bayes, etc.)

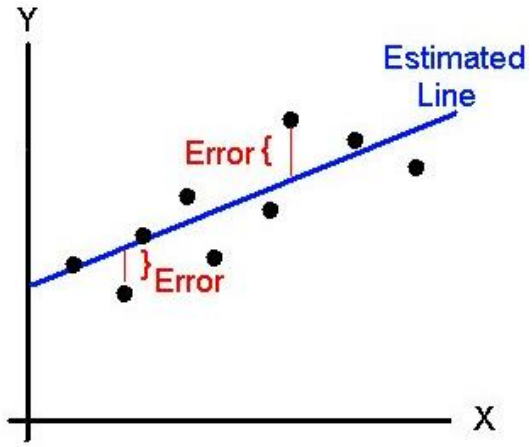
- **Optimization**
  - (Gradient Descent, Dynamic Programming, Divide and Conquer, Evolutionary Computation, etc.)

- **Evaluation**
  - (Accuracy, Precision, Recall, Cost/Utility, etc.)

Machine Learning Task

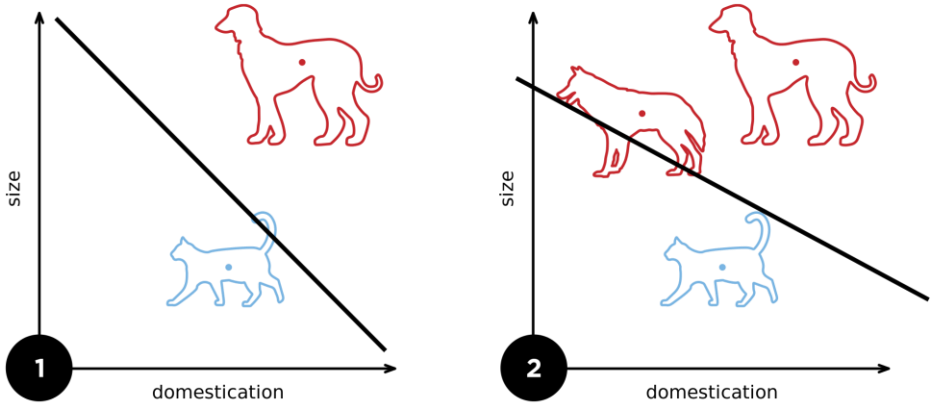
- At the very core, ML tries to fit a model on the data or finds out a decision boundary to separate the data belonging to various classes



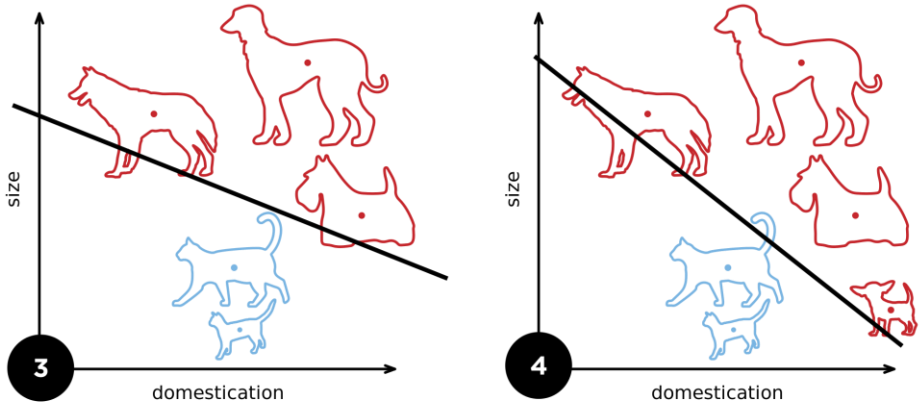


Linear Classifier/Expert

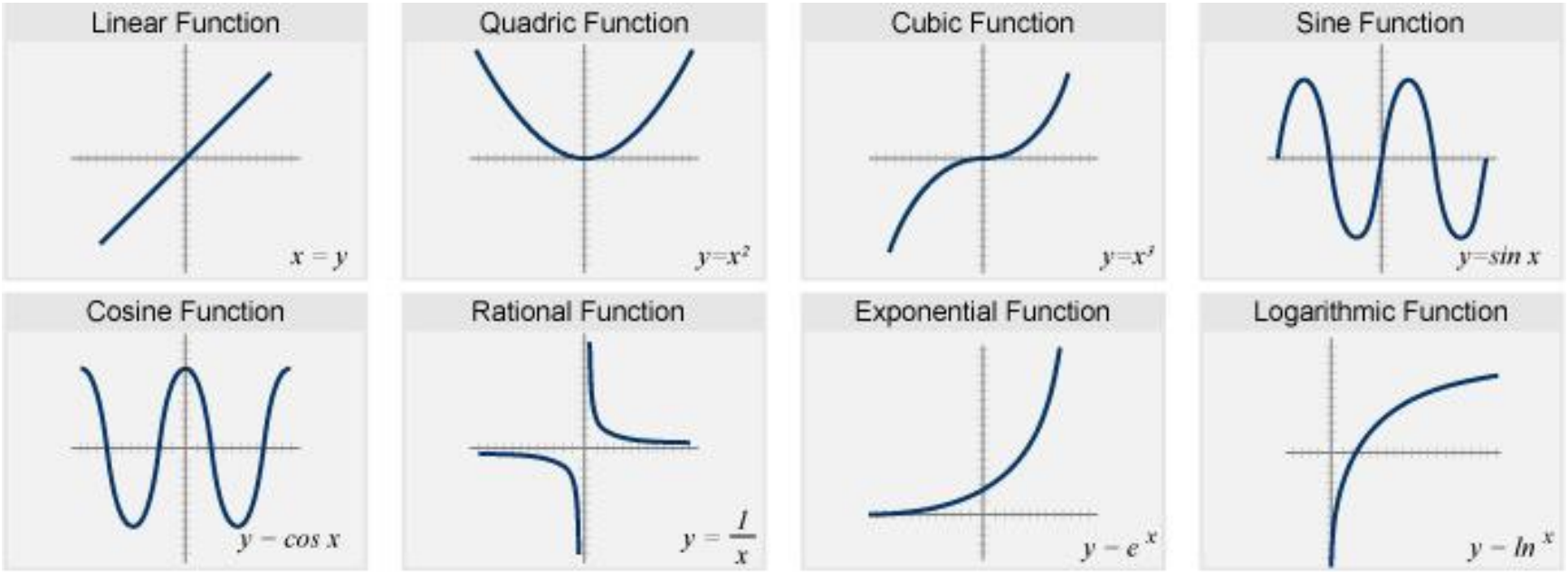
- It is all about finding the “right Hyper-plane” to separate vectors that belong to different classes



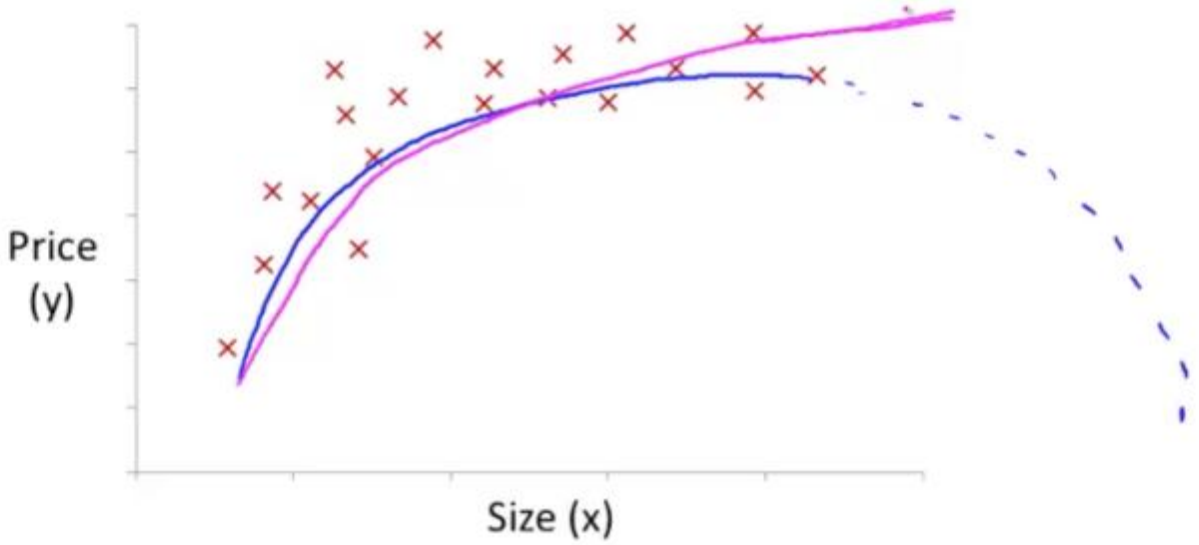
Playing with intercept and slope



Math Functions

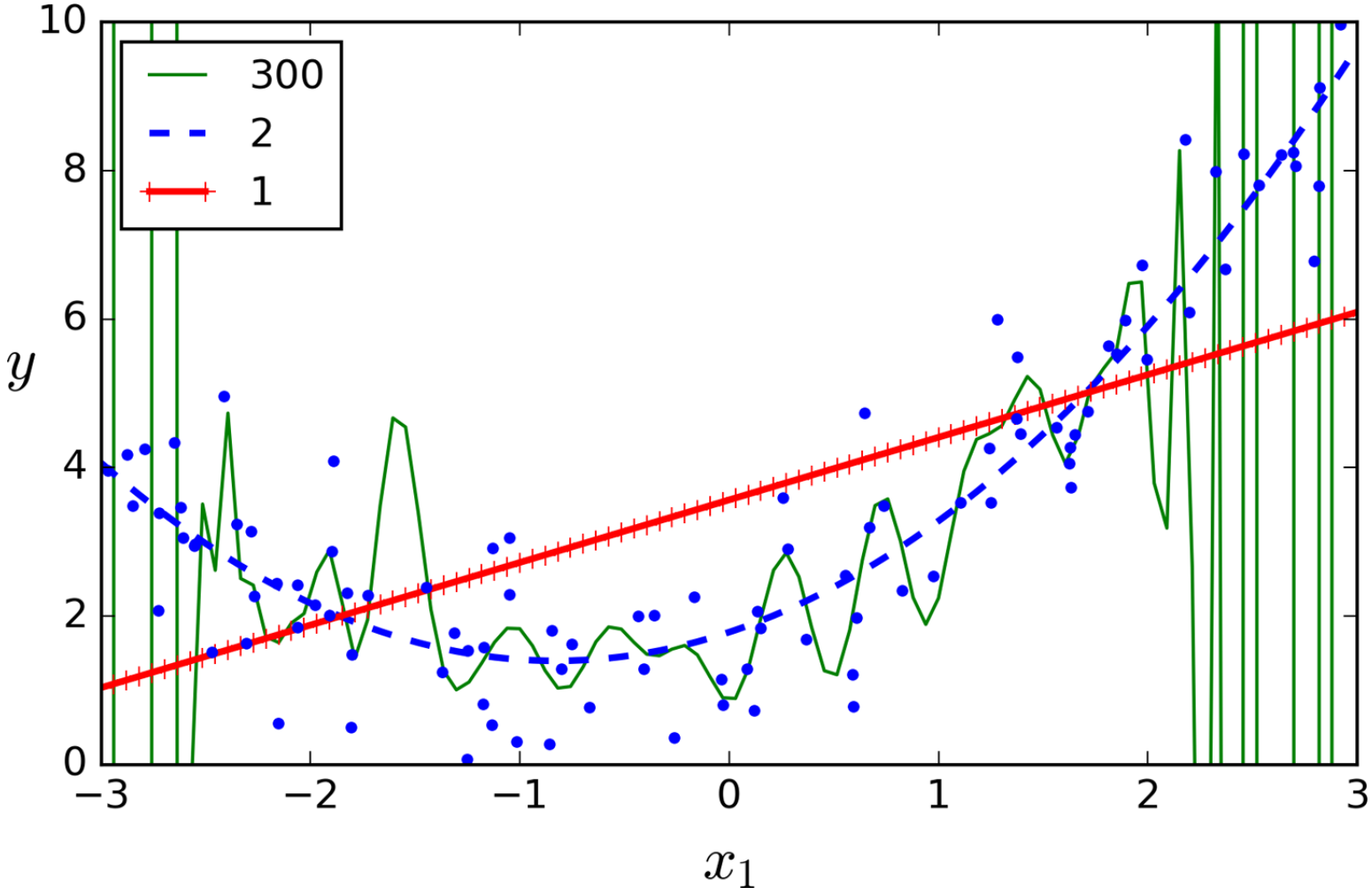


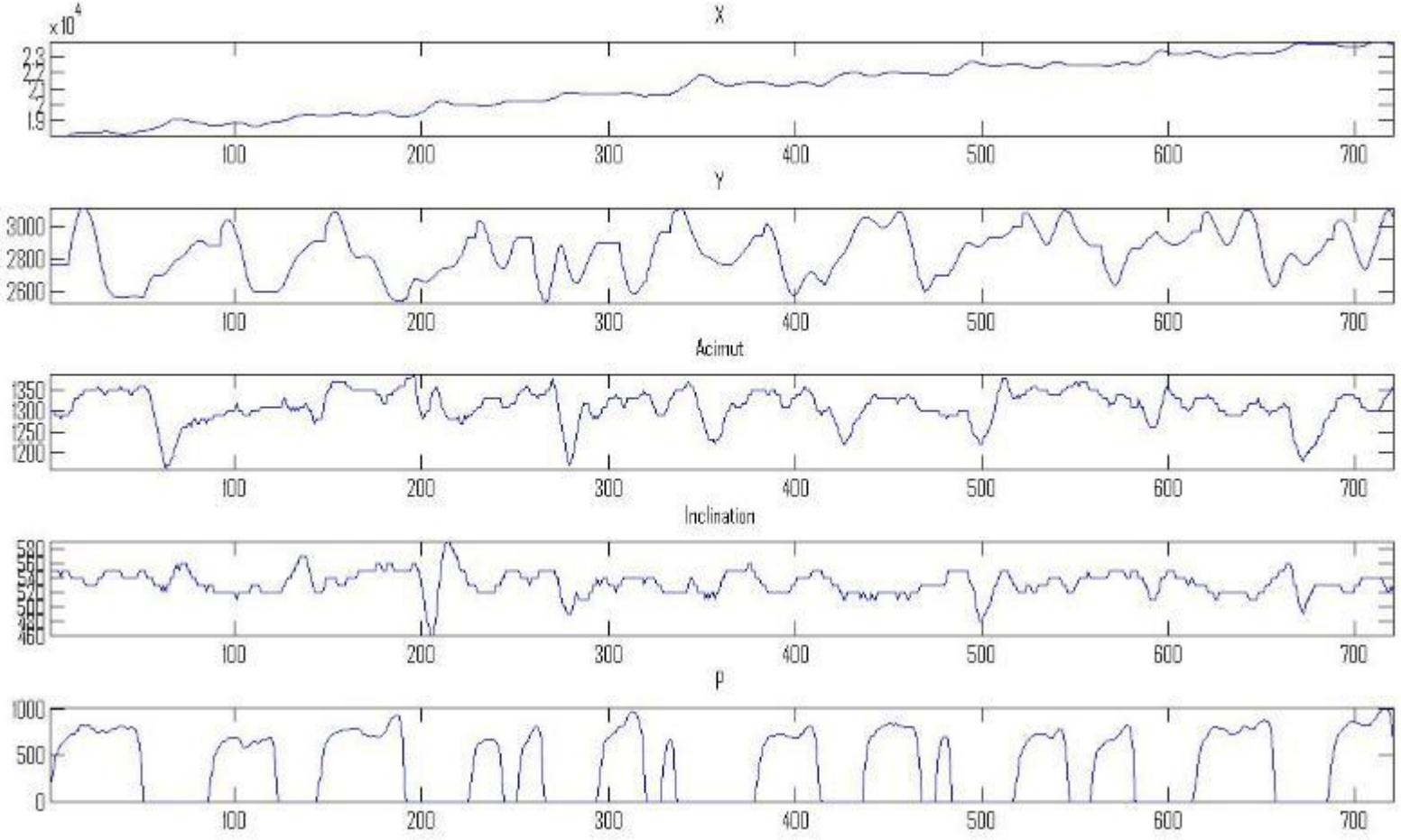
Choice of features

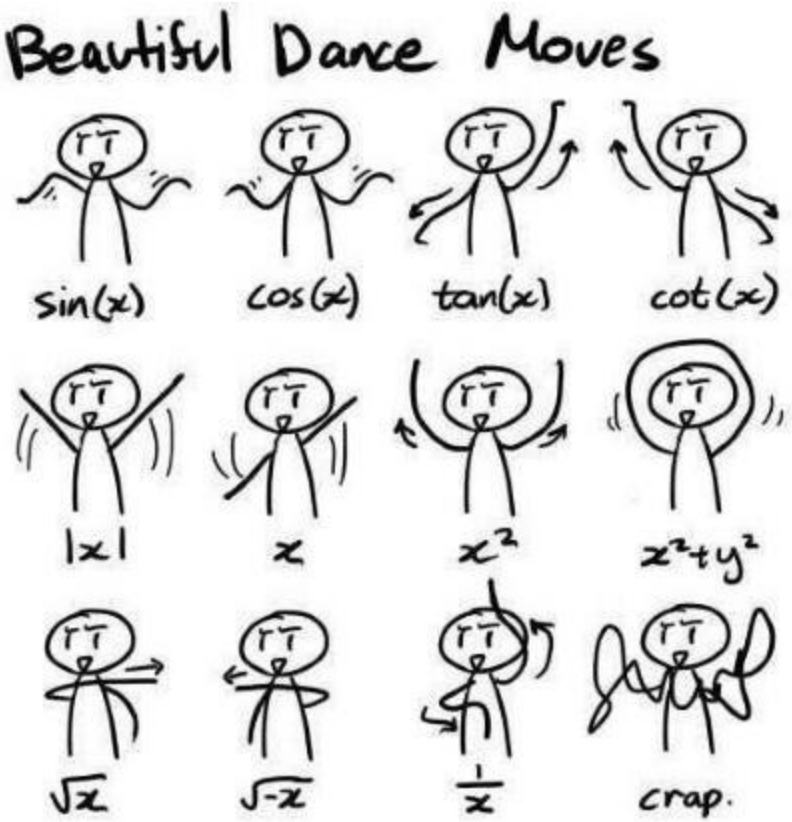


$$h_\theta(x) = \theta_0 + \theta_1(size) + \theta_2(size)^2$$

$$h_\theta(x) = \theta_0 + \theta_1(size) + \theta_2\sqrt{(size)}$$

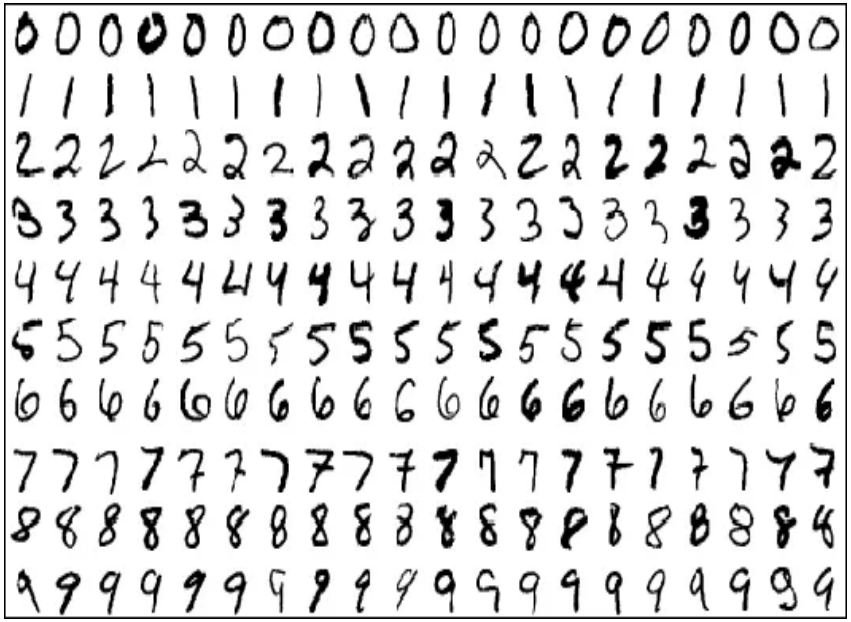

300
2
1
10
8
6
4
2
0
y
−3
−2
−1
0
1
2
3
x_1

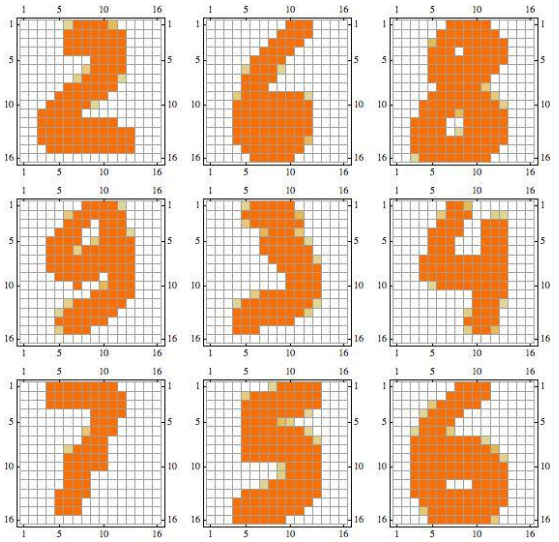

X
x 10^4
2.3
2.2
2.1
2
1.9
100
200
300
400
500
600
700
Y
3000
2800
2600
Acimut
1350
1300
1250
1200
Inclination
P
1000
500
0


Beautiful Dance Moves
sin(x)
cos(x)
tan(x)
cot(x)
|x|
x
x²
x²+y²
√x
√-x
1/x
crap.

Formulation as a ML Task

- **Handwritten Digits Recognition**



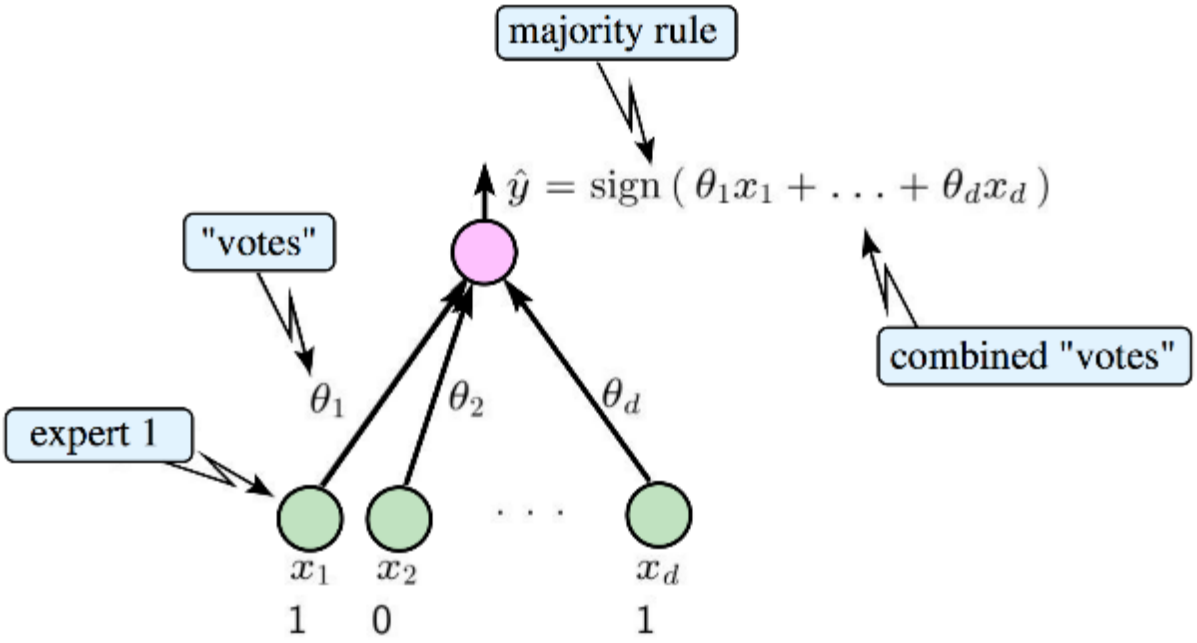


Linear Classifier/Expert

- We can understand the simple linear classifier

$$\hat{y} = f(x; \theta) = \text{sign}(\theta \cdot x) = \text{sign}(\theta_1 x_1 + \cdots + \theta_d x_d),$$

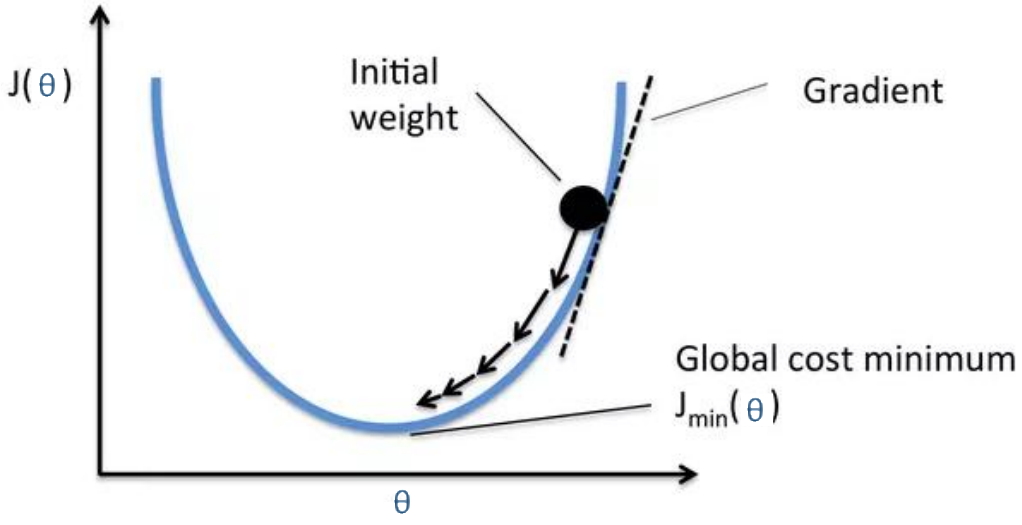
as a way of combining expert opinion (binary features)

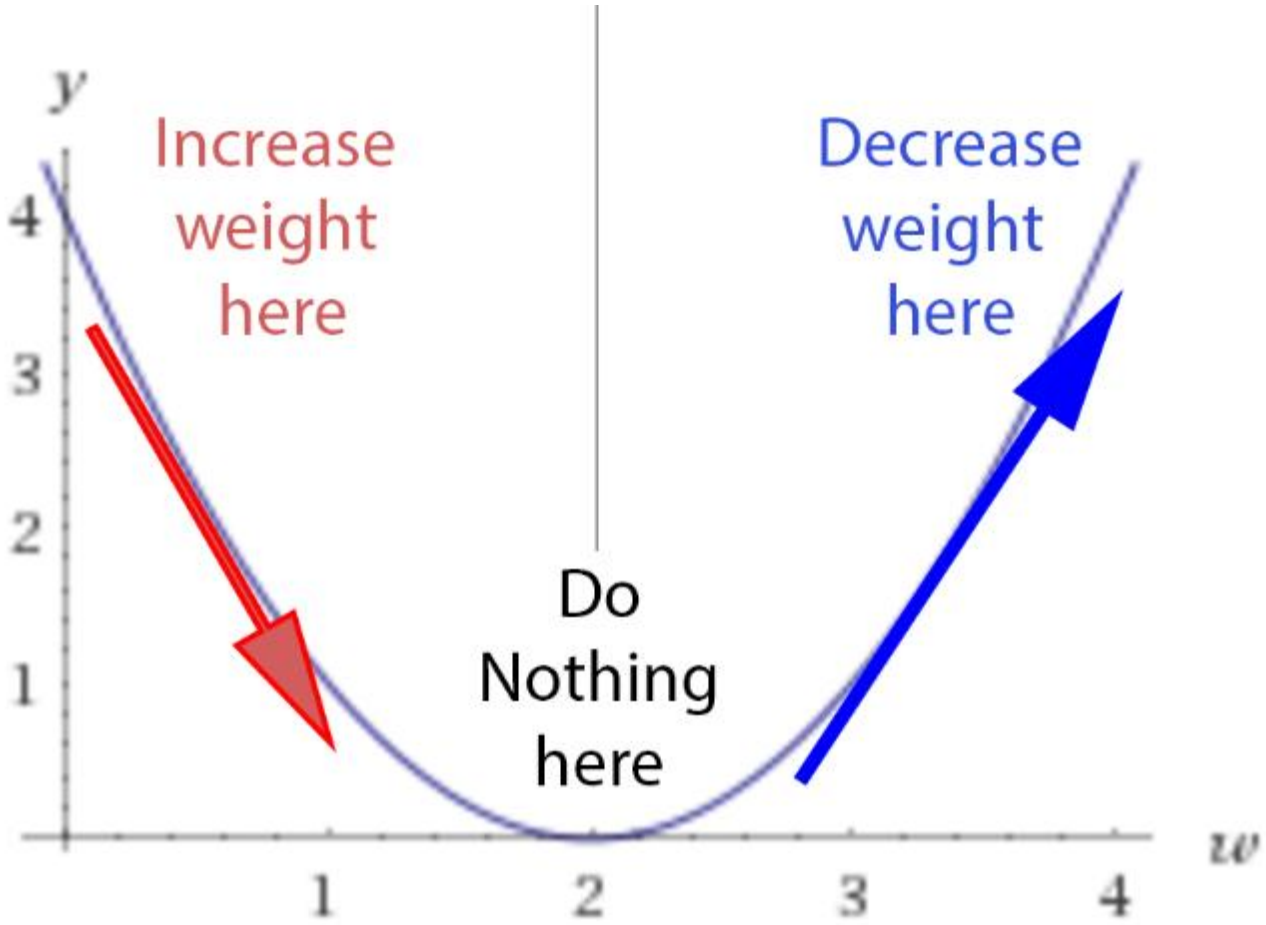


Linear Classifier/Expert

- How do we **adjust the parameters θ** based on the labeled examples?

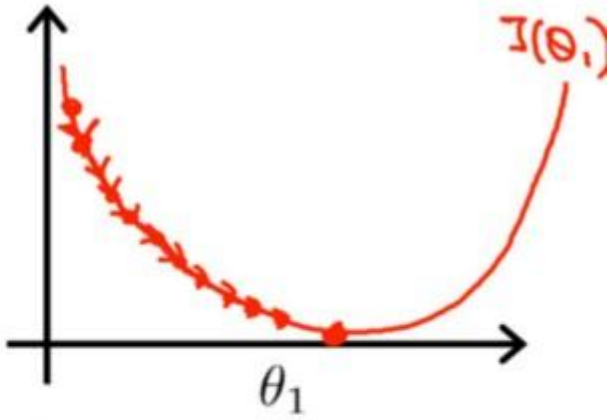
- **Gradient Descent**
- Standard loss/cost/objective function measures the squared error between y and the true value t
  - $J(\theta) = \mathbf{\Sigma} (y - y\hat{})^2$
- we can update the parameters:
- $\theta_{new} = \theta + \lambda \, . \, \partial J(\theta)/\partial\theta$, where λ = Learning Rate, J(θ) = disparity b/w target and actual value



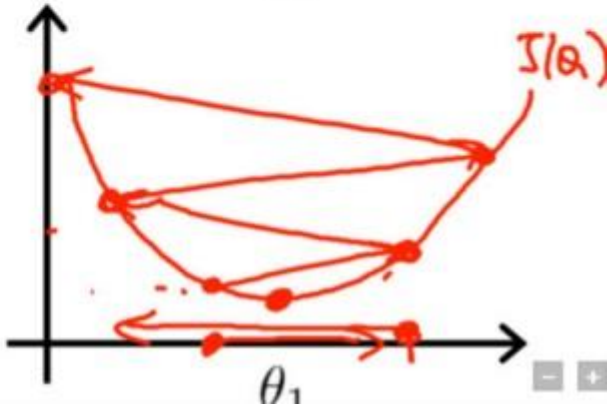

y
Increase weight here
Decrease weight here
4
3
2
1
Do Nothing here
1
2
3
4
w

$$\theta_1 := \theta_1 - \alpha \frac{\partial}{\partial \theta_1} J(\theta_1)$$

If α is too small, gradient descent can be slow.



If α is too large, gradient descent can overshoot the minimum. It may fail to converge, or even diverge.

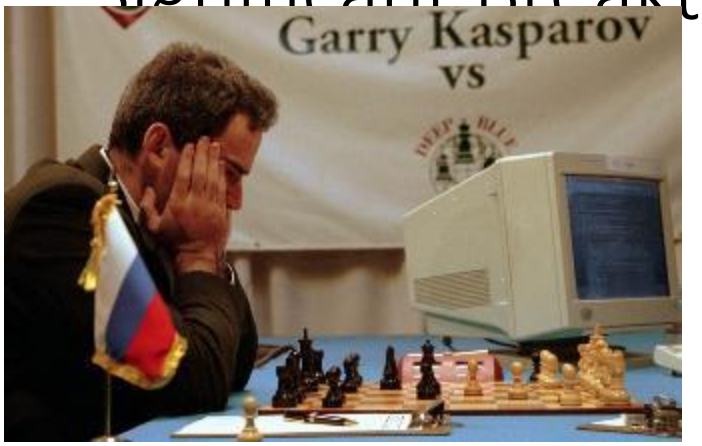


What can ML do today?

- Play a decent game of table tennis?
- Play the game of jeopardy?
- Drive safely along a curving mountain?
- Drive safely at Chowburjy Chowk?
- Buy weekly grocery on the web?
- Buy weekly grocery in Hyper star?
- Converse successfully with a person for an hour?
- Perform surgical operation?
- Put away the dishes and fold the laundry?
- Translate spoken Chinese into English at real time?
- Write an intentional funny story?

[Pieter Abbeel, Dan Klein, et. al]

Significant Breakthroughs



Deep Blue (1997)

IBM WATSON (2011)

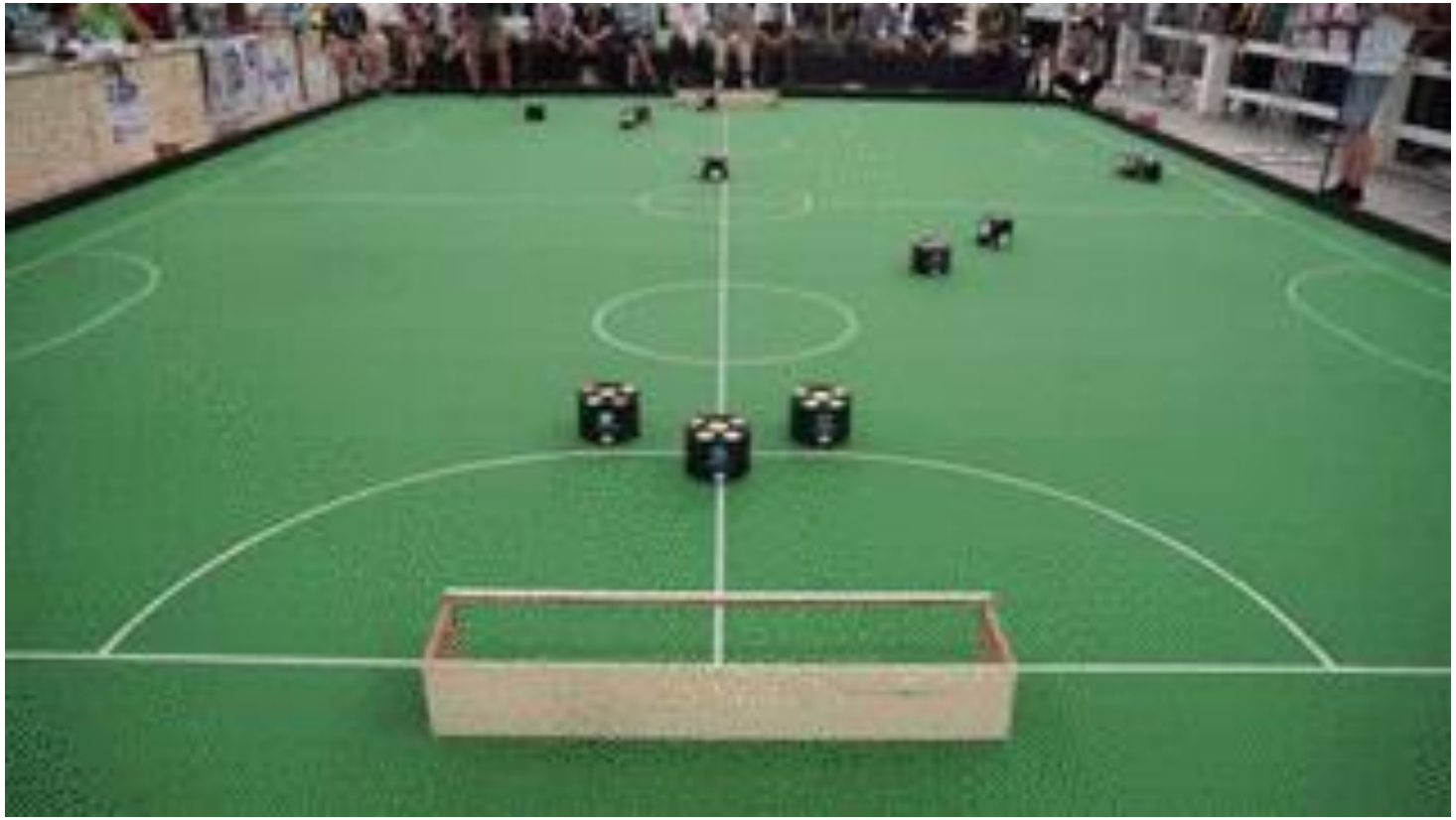
Human versus machine

Playing table tennis

Autonomous driving

Autonomous Robots

robocup



Machine Translation

Recommender systems

Customers who viewed this item also viewed these products

Dualit Food XL1500 Processor

$560

Add to cart

Kenwood kMix Manual Espresso Machine

★★★★☆

$250

Select options

Weber One Touch Gold Premium Charcoal Grill-57cm

$225

Add to cart

NoMU Salt Pepper and Spice Grinders

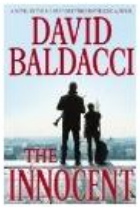
$3

View options

Amazon !!

Today's Recommendations For You

Here's a daily sample of items recommended for you. Click here to **see all recommendations**.



The Innocent (Kindle Edition) by David Baldacci

$14.99

Fix this recommendation

The Expats: A Novel (Kindle Edition) by Chris Pavone

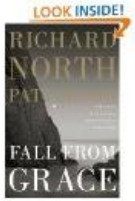
★★★★☆ (65) $12.99

Fix this recommendation

Stolen Prey (Kindle Edition) by John Sandford

$14.99

Fix this recommendation

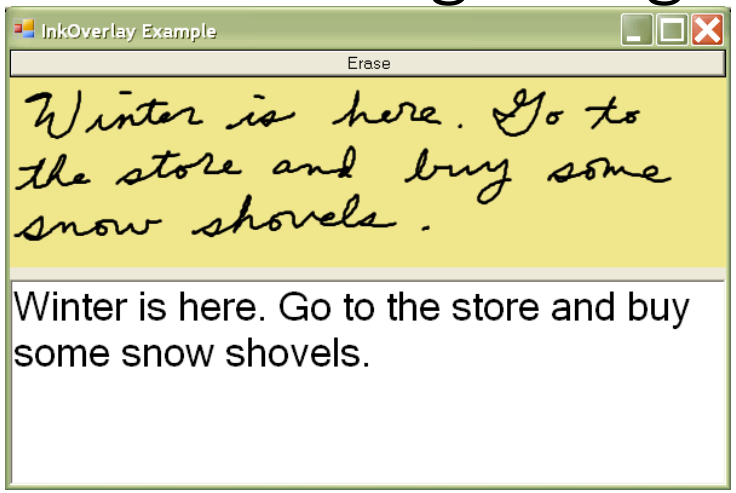


Fall from Grace (Kindle Edition) by Richard North Patterson

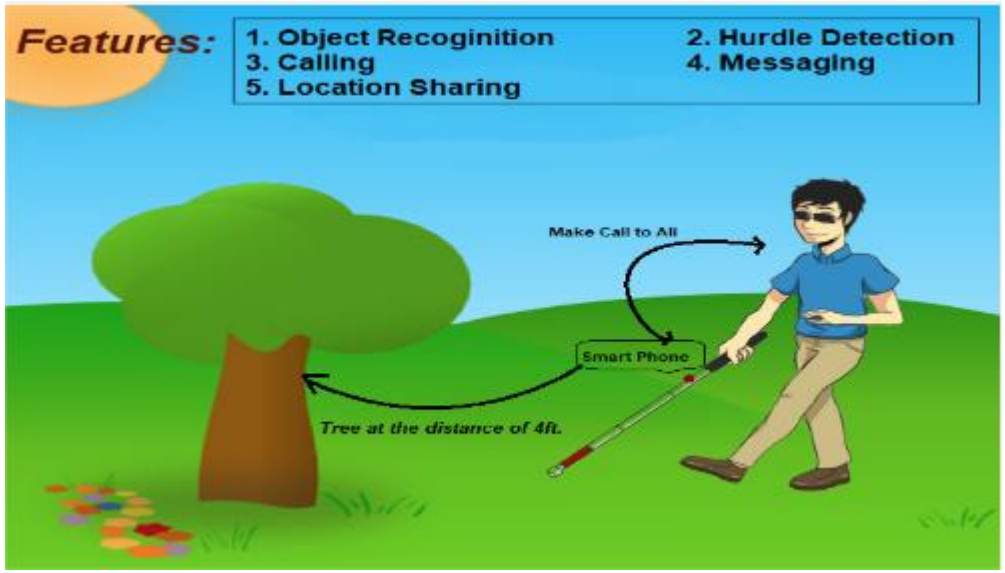
★★★½☆ (32) $12.99

Fix this recommendation

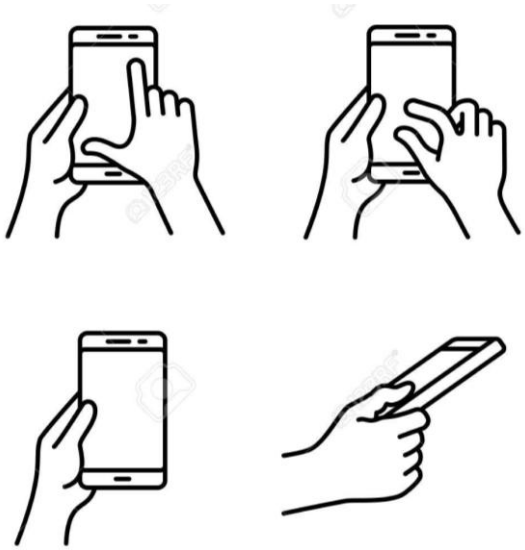
Handwriting recognition



ML at UMT



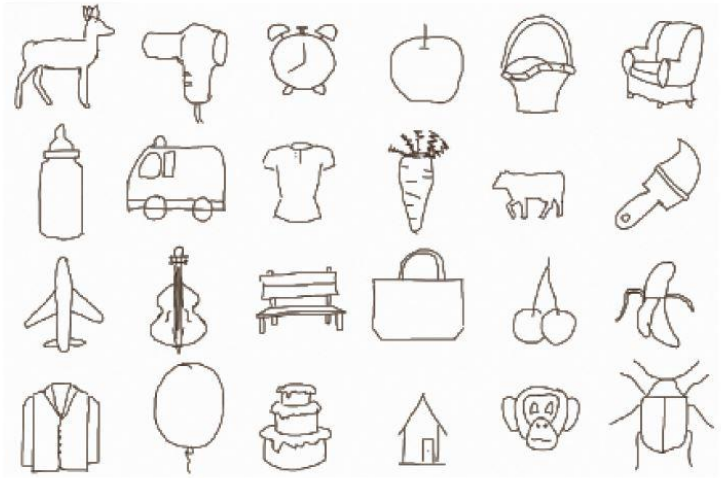
Smart Cane for the Blind



User Identification through
Gesture Recognition on Android

ML at UMT

Activity Recognition in Smart Homes



Object Recognition based on Sketches