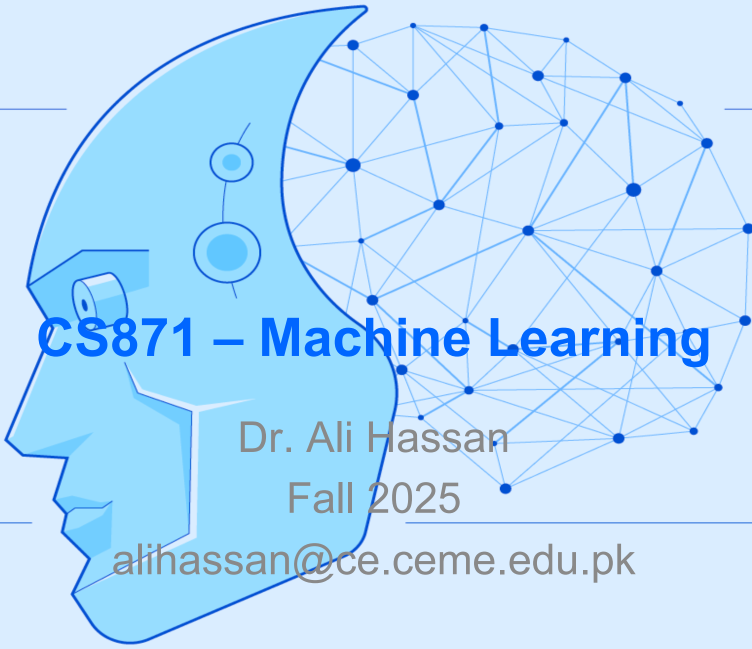


CS871 – Machine Learning

Dr. Ali Hassan
Fall 2025
alihassan@ce.ceme.edu.pk

1

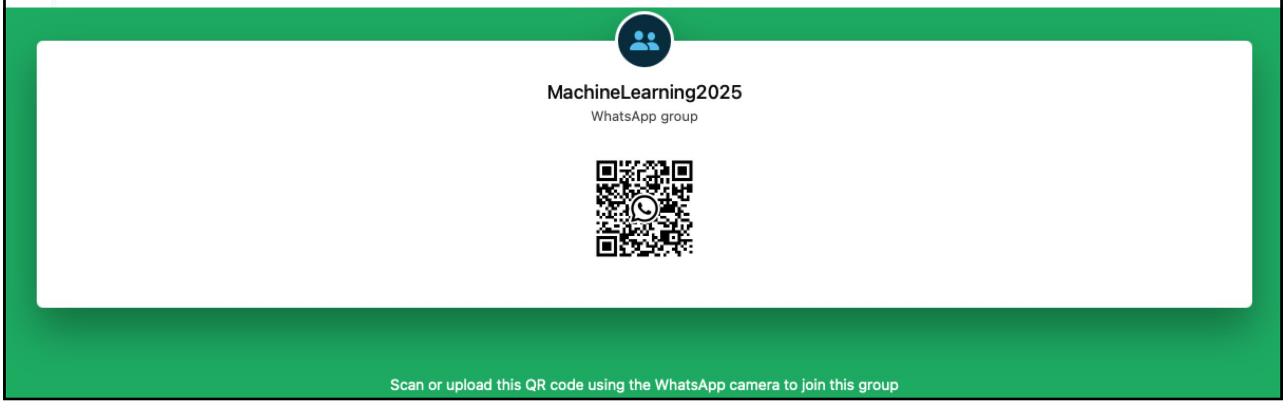
LMS Enrollment Code

- Enrollment Code :
  - 426708913

2

WhatsApp Group

- Follow this link to join WhatsApp group MachineLearning2025:
- https://chat.whatsapp.com/GfU7KrUxXDI332yysFEp9Q



3

Announcements

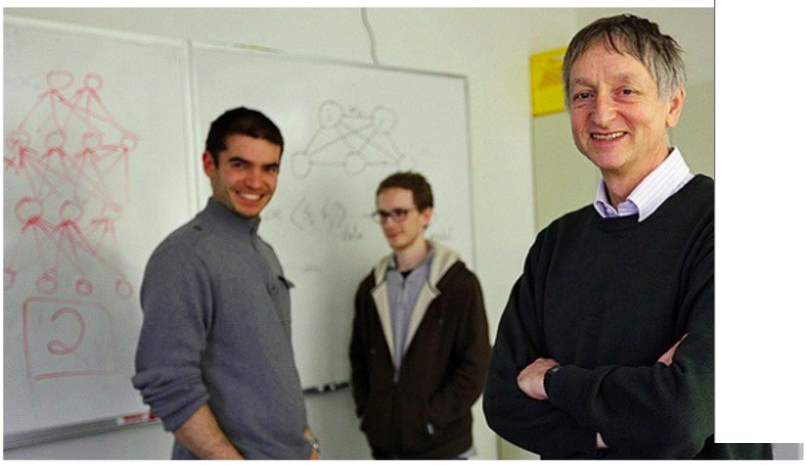
- Assignment 1 is online
  - Deadline: **24 Sep 2025 at 2359**

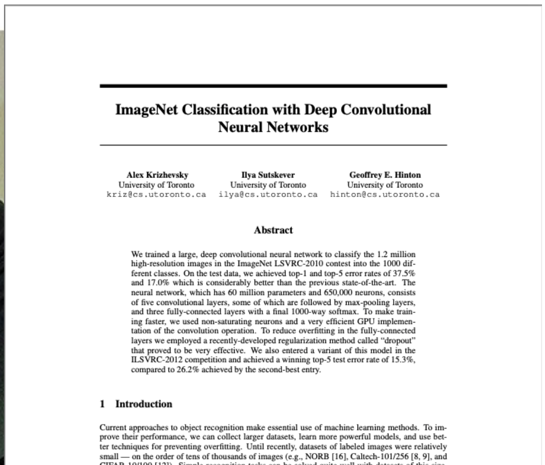
4

Outline

- Types of Machine Learning (ML)
- Example:
  - A ML Problem
- A Typical ML system design cycle/ Pipeline
- Different Types of Learning
- Reading Work !!!
- ChatGPT/Python Demo

Dedication





ImageNet Classification with Deep Convolutional Neural Networks

Alex Krizhevsky
University of Toronto
kriz@cs.utoronto.ca

Ilya Sutskever
University of Toronto
ilya@cs.utoronto.ca

Geoffrey E. Hinton
University of Toronto
hinton@cs.utoronto.ca

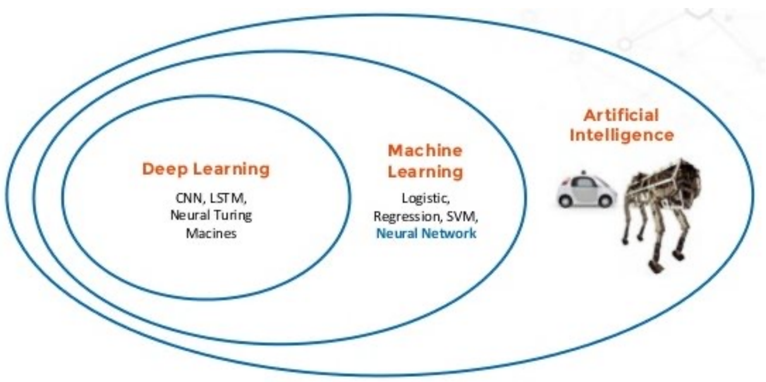
Abstract

We trained a large, deep convolutional neural network to classify the 1.2 million high-resolution images in the ImageNet LSVRC-2010 contest into the 1000 different classes. On the test data, we achieved top-1 and top-5 error rates of 37.5% and 17.0% which is considerably better than the previous state-of-the-art. The neural network, which has 60 million parameters and 650,000 neurons, consists of five convolutional layers, some of which are followed by max-pooling layers, and three fully-connected layers with a final 1000-way softmax. To make training faster, we used non-saturating neurons and a very efficient GPU implementation of the convolution operation. To reduce overfitting in the fully-connected layers we employed a recently-developed regularization method called "dropout" that proved to be very effective. We also entered a variant of this model in the ILSVRC-2012 competition and achieved a winning top-5 test error rate of 15.3%, compared to 26.2% achieved by the second-best entry.

1 Introduction

Current approaches to object recognition make essential use of machine learning methods. To improve their performance, we can collect larger datasets, learn more powerful models, and use better techniques for preventing overfitting. Until recently, datasets of labeled images were relatively small — on the order of tens of thousands of images (e.g., NORB [16], Caltech-101/256 [8, 9], and

AI vs ML vs DL



7

Definition of Machine Learning

- Machine Learning is the science (and art) of programming computers so they can *learn from data*.
- *[Machine Learning is the] field of study that gives computers the ability to learn without being explicitly programmed.*

  —Arthur Samuel, 1959

- *[Machine Learning is the] field of study that gives computers the ability to learn without being explicitly programmed.*

  —Arthur Samuel, 1959

8

Definition of Machine Learning

- *A computer program is said to learn from experience E with respect to some task T and some performance measure P, if its performance on T, as measured by P, improves with experience E.*

  -- Tom Mitchell, 1997

What is Artificial Intelligence

- Field of study which studies how to create computing systems that are capable of intelligent behavior
- FUN Fact:
  - Fun Fact about AI is that after a problem is fully solved it's not called intelligent anymore... (ie: Make a computer play chess was the highest display of intelligence, now people don't consider that anymore)

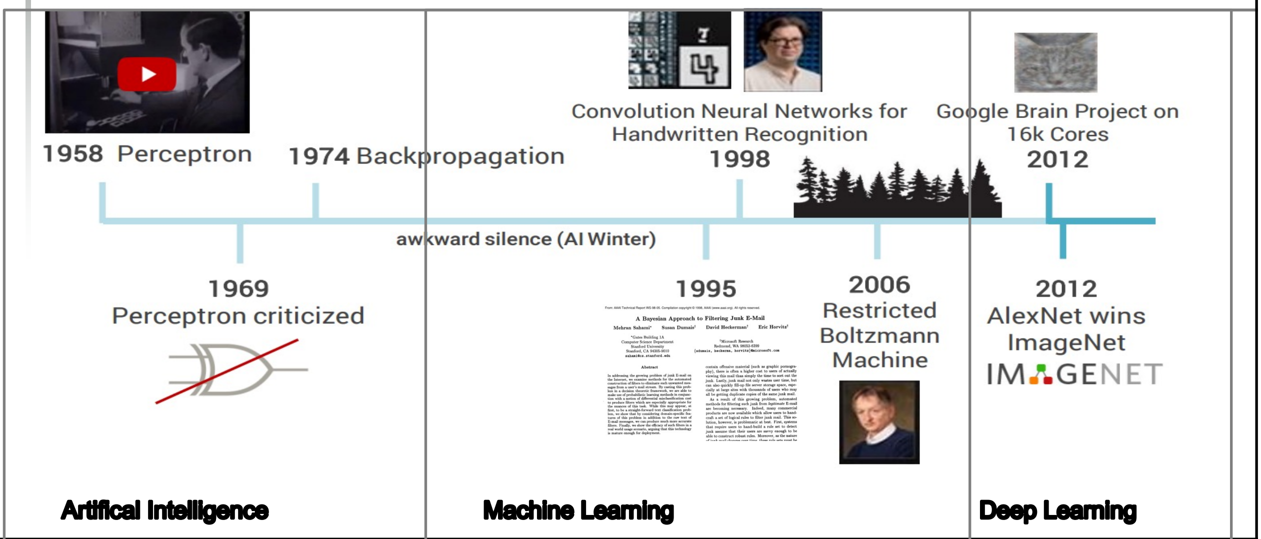
Machine Learning

- Machine learning (ML) is a field of study in artificial intelligence concerned with the development and study of statistical algorithms that can learn from data and generalize to unseen data and thus perform tasks without explicit instructions

https://en.wikipedia.org/wiki/Machine_learning

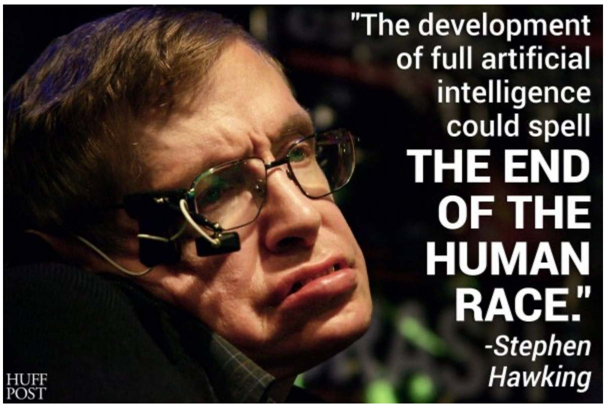
11

History

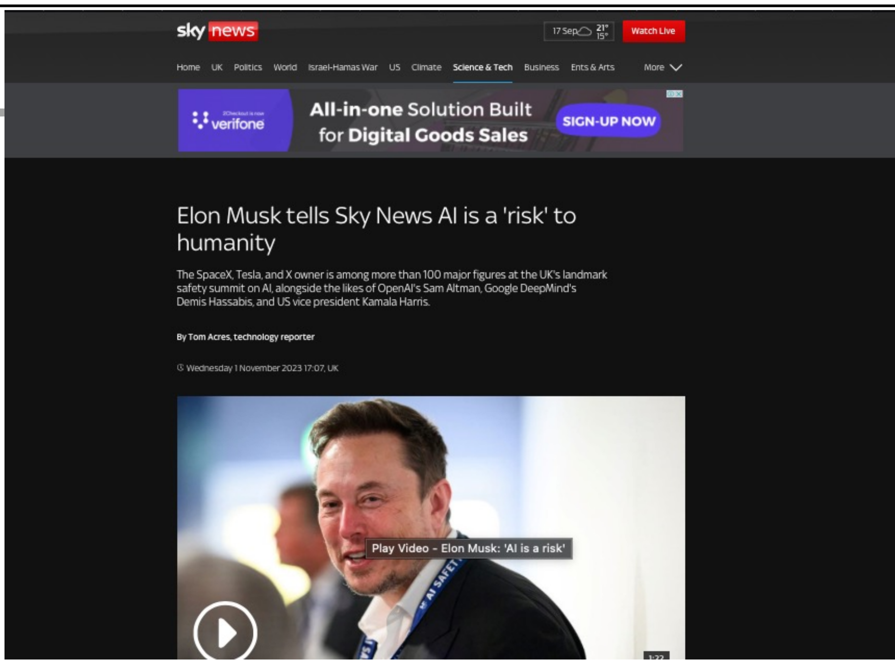


12

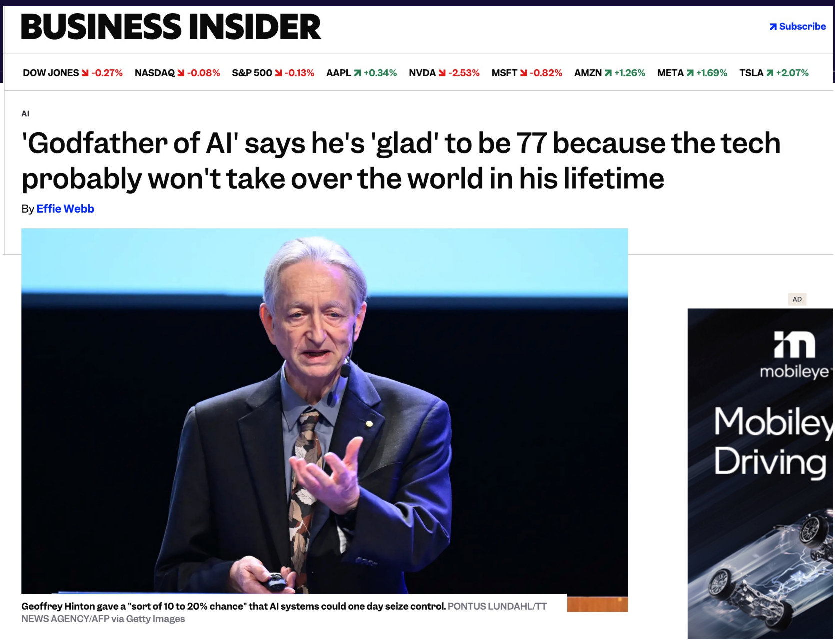
Mix of hype/fear



13



14



BUSINESS INSIDER

Subscribe

DOW JONES -0.27% NASDAQ -0.08% S&P 500 -0.13% AAPL +0.34% NVDA -2.53% MSFT -0.82% AMZN +1.26% META +1.69% TSLA +2.07%

AI

'Godfather of AI' says he's 'glad' to be 77 because the tech probably won't take over the world in his lifetime

By Effie Webb

Geoffrey Hinton gave a "sort of 10 to 20% chance" that AI systems could one day seize control. PONTUS LUNDAHL/TT NEWS AGENCY/AFP via Getty Images

Dr. Ali Hassan, DCSE, College of E&ME, NUST

15

Why Now

- The AI/ ML algorithms (Even the deep ones) are there for decades so why we have now this buzzword?
- Due to the advance of **computing power through** (GPUs, multi-core CPU systems, and FPGAs) and the availability of data (**Big data**) through internet.

Dr. Ali Hassan, DCSE, College of E&ME, NUST 16

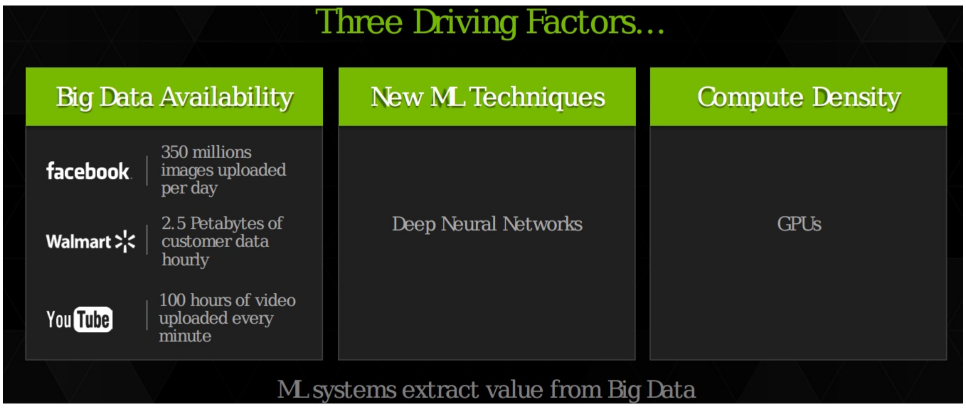
16

Why Now

- Also the amount of data that need to be classified nowadays become too big to be handled manually,
- So big companies Google, Microsoft, Facebook, start to invest heavily on the data.
- Data is the new gold rush

Dr. Ali Hassan, DCSE, College of E&ME, NUST 17

17



Dr. Ali Hassan, DCSE, College of E&ME, NUST 18

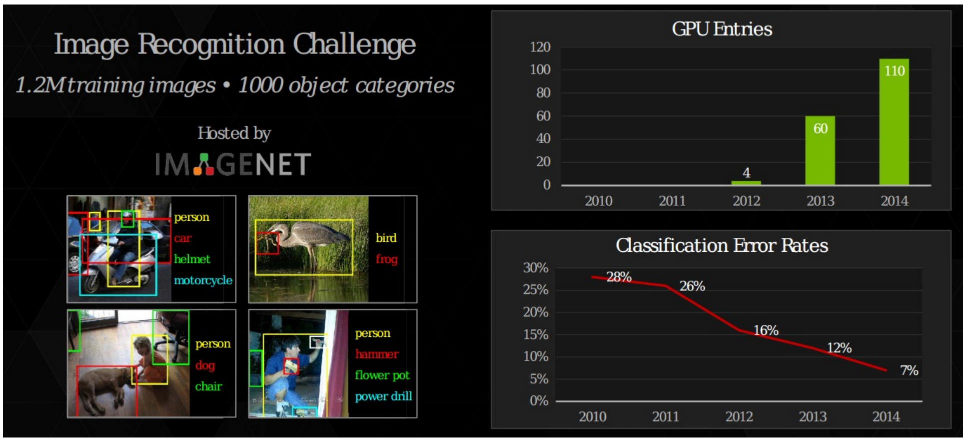
18

Batch Size	Training Time CPU	Training Time GPU	GPU Speed Up
64 images	64 s	7.5 s	8.5X
128 images	124 s	14.5 s	8.5X
256 images	257 s	28.5 s	9.0X

The new Hype

- The last years AI/ML models are surpassing human intelligence
  - Speech and natural language processing
  - Face Recognition
  - Image Classification, object detection
  - Self Driving Car
  - Playing complex games (Alpha Go)
  - Control strategies (Control engineering)
  - Live translation

Examples …



21

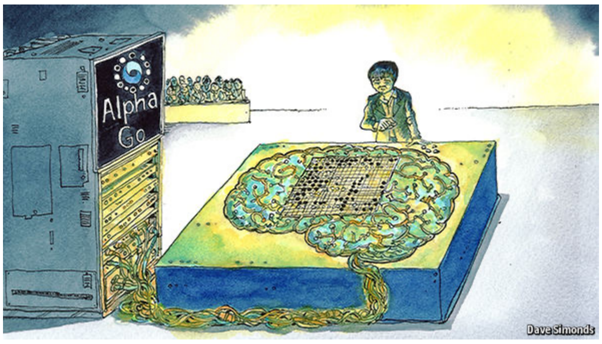
Computing power Comparison

Type	Name	Flops	Cost
Mobile	Raspberry Pi 1st Gen, 700 Mhz	0,04 Gflops	$35
Mobile	Apple A8	1,4 Gflops	$700 (in iPhone 6)
CPU	Intel Core i7-4930K (Ivy Bridge), 3.7 GHz	140 Gflops	$700
CPU	Intel Core i7-5960X (Haswell), 3.0 GHz	350 Gflops	$1300
GPU	NVidia GTX 980	4612 Gflops (single precision), 144 Gflops (double precision)	$600 + cost of PC (~$1000)
GPU	NVidia Tesla K80	8740 Gflops (single precision), 2910 Gflops (double precision)	$4500 + cost of PC (~1500)

24

Deepmind hardware

- This picture is the hardware used to play against one of the best Go players in the world.



What is Machine Learning

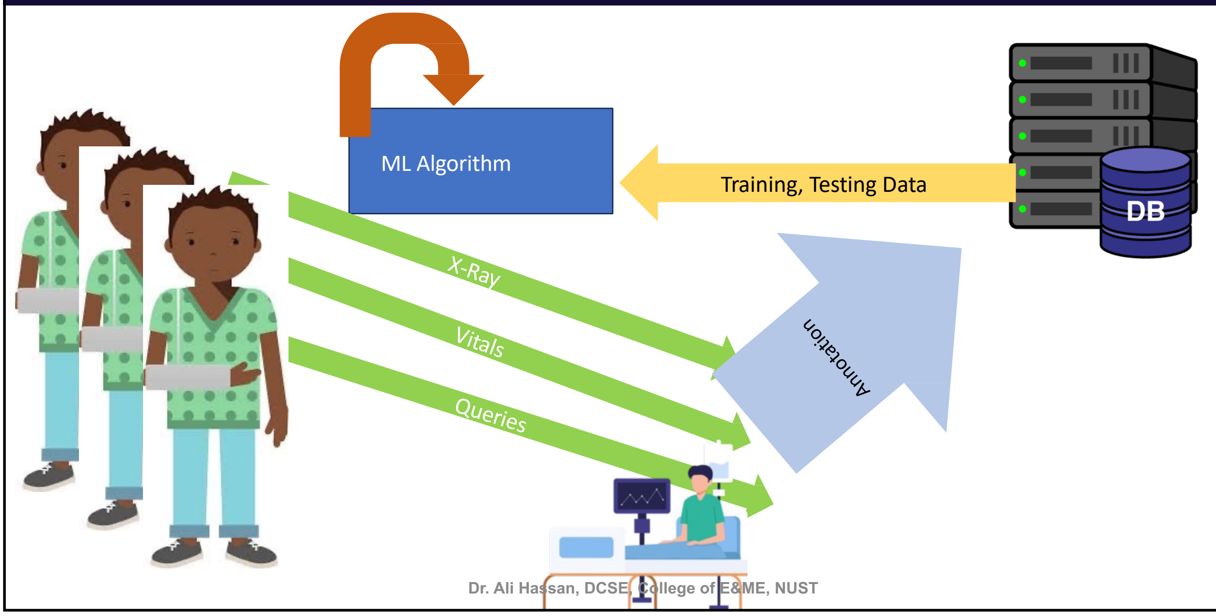
- Machine learning is all about using your computers to "**learn**" how to deal with problems without "**explicit programming**"
- We need ML on cases that would be difficult to program by hand all possible variants of a classification/prediction problem
- The basic Idea of ML is to make the computer **learn** something from the **data**

What is Machine Learning …

- ML is a way to make the computers create a program that gives some output with a known input and that latter give an intelligent output to a different but similar input.
  - Take some data,
  - Train a model on that data, and
  - Use the trained model to make predictions on new (UNSEEN) data.

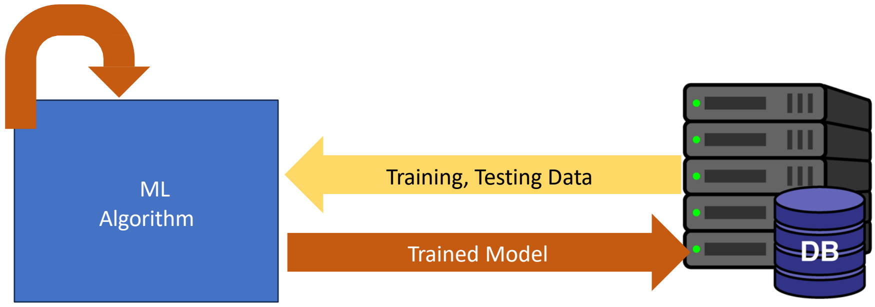
27

Machine Learning



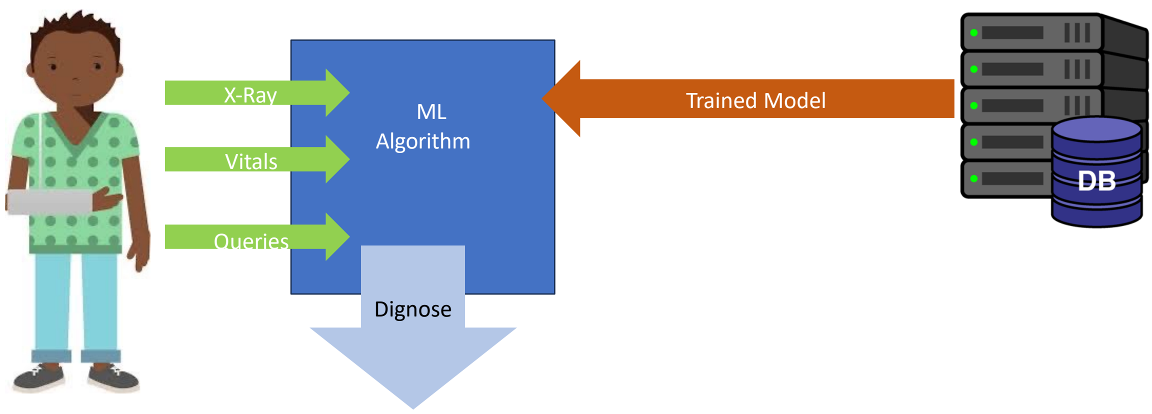
28

Machine Learning Training



29

Machine Learning -- Testing



30

Tr Data labels

- Function F

$f : X \rightarrow Y$

Training Data

$\{(x_i, y_i) \in X, Y\}$

Learning

find $f' \in F$

s.t $y_i \approx f'(x_i)$

Predict Actual

New Data

Prediction $y = f'(x)$ x

Machine Learning Methods

TEMPLATE MATCHING

STATISTICAL APPROACH

SYNTACTIC APPROACH

NEURAL NETWORKS

TEMPLATE MATCHING

Template Matching

- A template (typically a 2D shape) or a prototype of the pattern to be recognized is available.
- Compute the similarity between the template and the pattern to be matched.
- Take into account pose (rotation, translation) and scale changes.

A typical template matching problem

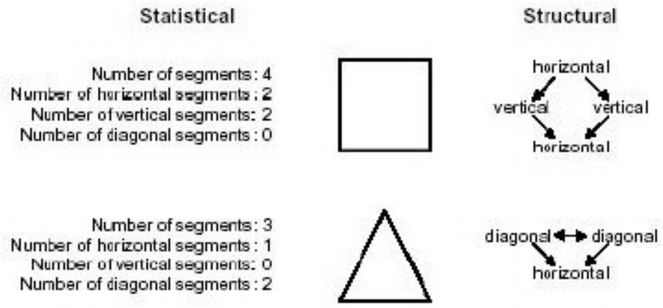
Issues of concern

- Take an example of an image of size 28x28 pixels.
- You have to use all 28x28=784 pixels
- Computational complexity
- Rigidity assumption (use deformable template models)
- Choice of template

STATISTICAL APPROACH

Statistical Approach

- Each pattern is represented in terms of **d- features**, and is viewed as a point in a d-dimensional space
- The goal is to choose those features that allow pattern vectors belonging to different categories to occupy compact and disjoint regions in a d-dimensional feature space.



39

Issues of concern

- Usually d << D (dimensions of image/signal)
- What should be the value of d ??
- We might encounter curse of dimensionality

40

SYNTACTIC APPROACH

Syntactic Approach

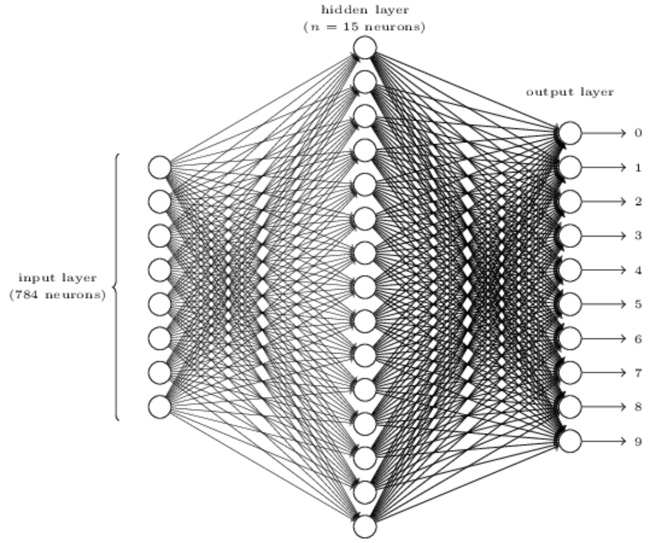
- Use hierarchical structures to represent complex patterns.
- The simplest unit is called: *primitives*
- Complex pattern is represented in terms of the interrelationships (*grammars*) between the primitives.
- Learn Grammatical rules from Data
- Trend analysis using tweets
  - NLP, NLG

Issues of concern

- However, it's usually difficult to segment noisy patterns and infer grammar from the training set.
- May yield a combinatorial explosions of possibilities to be investigated.

NN BASED APPROACHES

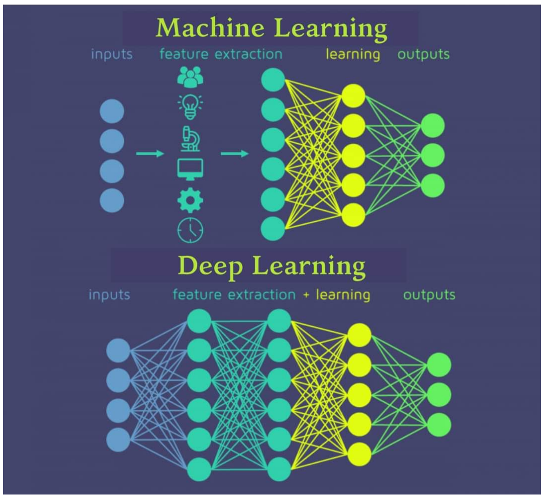
Neural Networks



Neural Networks

- Massively parallel computing unit consisting of an extremely large number of simple processors with many interconnections.
- Can learn complex *non-linear* input-output relationships.
- Feed-forward networks such as multilayer perceptron and Radial Basis Function network are useful for pattern classification.

Machine vs Deep Learning



47

Spam Filter

MACHINE LEARNING EXAMPLE # 1

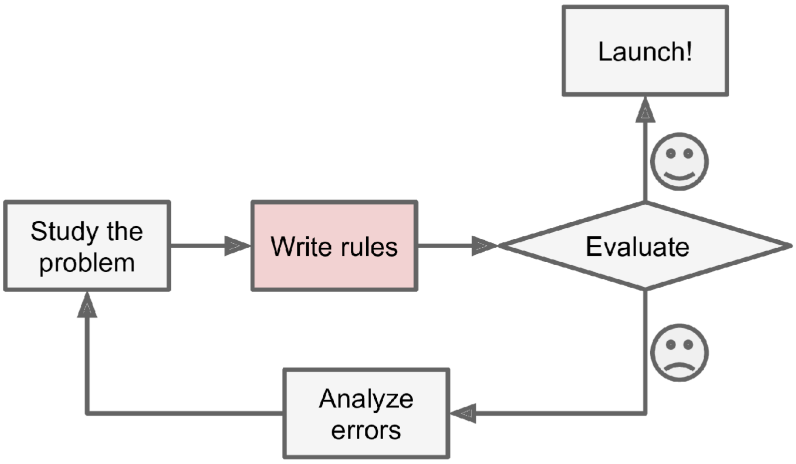
48

What is spam filter

- Your spam filter is a Machine Learning program that, given examples of spam emails (e.g., flagged by users) and examples of regular (nonspam, also called “ham”) emails, can learn to flag spam
- How will you go about it?

49

Traditional Approach



50

Rules

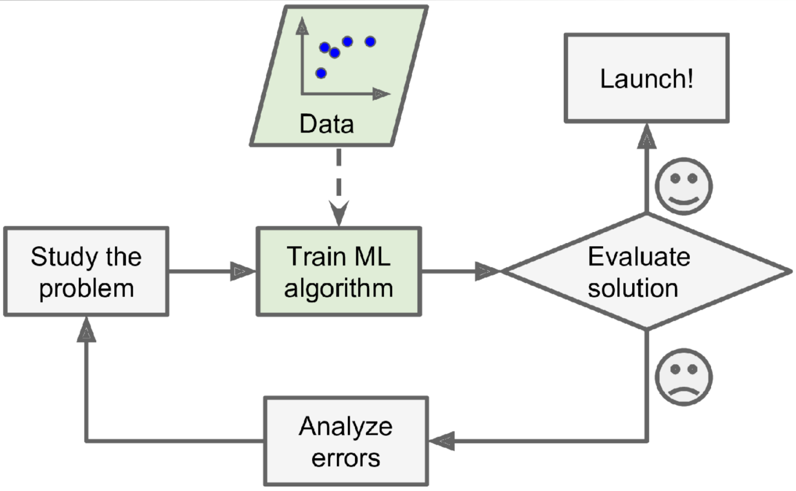
- Consider what spam typically looks like. (such as “4U,” “credit card,” “free,” and “amazing”)
- You would write a detection algorithm for each of the patterns that you noticed, and your program would flag emails as spam
- You would test your program and repeat steps 1 and 2 until it was good enough to launch.
- Any Issues ??

Issue

- What if spammers notice that all their emails containing “4U” are blocked? They might start writing “For U” instead

Solution

- Learn from Data
- A spam filter based on ML techniques automatically notices that “For U” has become unusually frequent in spam flagged by users, and it starts flagging them



53

Separating Salmon from Sea Bass

MACHINE LEARNING EXAMPLE # 2

54

An Example

- "Sorting incoming Fish on a conveyor according to species using optical sensing"

Species → Sea bass

Species → Salmon

55

Machine Learning: Overview

- Function F

$f: X \rightarrow Y$

Training Data

$\{(x_i, y_i) \in X, Y\}$

Learning

find $f' \in F$

s.t $y_i \approx f'(x_i)$

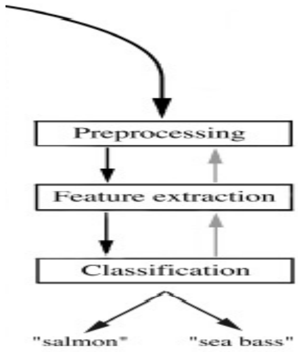
New Data

x

Prediction

$y = f'(x)$

56



57

Pre-Processing

- Use a segmentation operation to isolate fishes from one another and from the background
- Information from a single fish is sent to a feature extractor whose purpose is to reduce the data by measuring certain features
- The features are passed to a classifier or decision maker

58

(a) (b)

The objects to be classified; (a). Sea bass,
and (b). salmon.

59

Problem Analysis

- Set up a camera and take some sample images to extract features
  - Length
  - Lightness
  - Width
  - Number and shape of fins
  - Position of the mouth, etc…
  - This is the set of all suggested features to explore for use in our classifier!

60

System Design Process

- Domain knowledge:
  - a sea bass is generally longer than a salmon
- Selected Feature
  - Select the length of the fish as a possible feature for discrimination

$$x = [x_1] \quad \text{, where } x_1 \text{ is length}$$

- Model:
  - Sea bass have some typical length, and this is greater than the length of a salmon

Dr. Ali Hassan, DCSE, College of E&ME, NUST 61

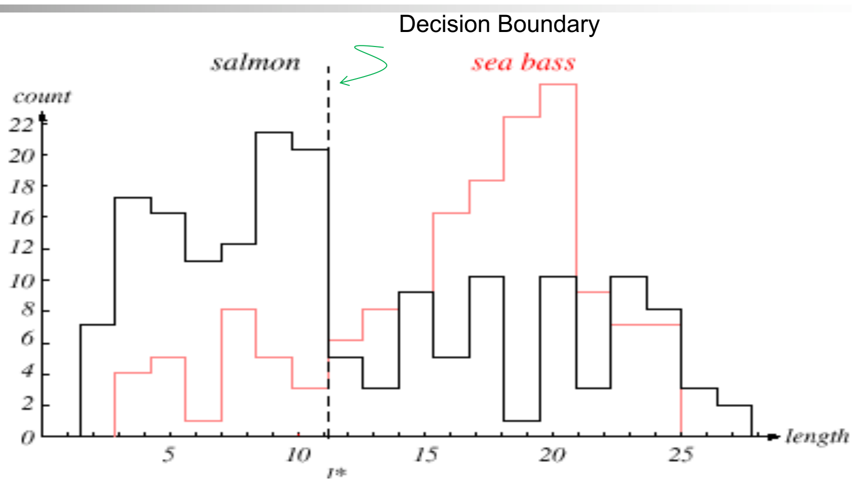
61

System Design Process

- Classification Rule
  - $f' = \begin{cases} \text{If } \textit{Length} >= l^* & \text{then sea bass} \\ & \text{otherwise salmon} \end{cases}$
- How to choose l* ???
- Use *Training Data*

Dr. Ali Hassan, DCSE, College of E&ME, NUST 62

62

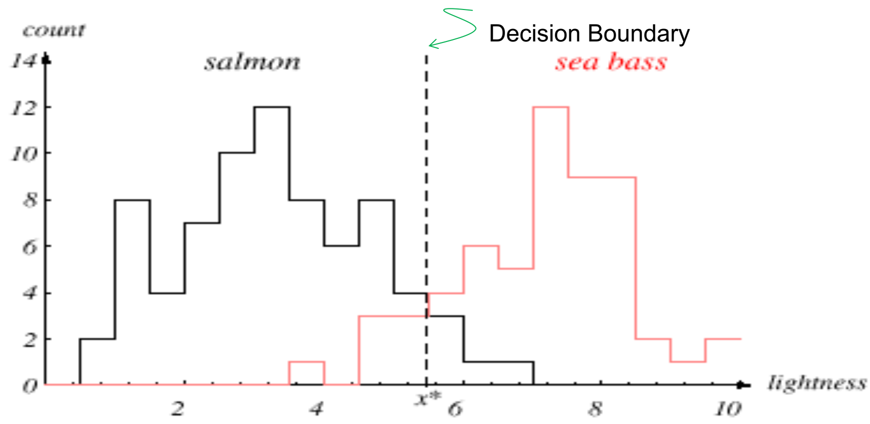


63

Analysis

- The length is a poor feature alone!
- New Feature:
  - Select the lightness of fish as a possible feature

64

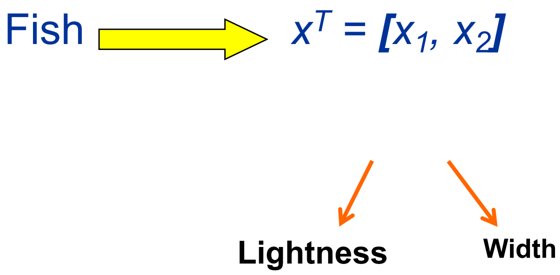


Dr. Ali Hassan, DCSE, College of E&ME, NUST 65

65

Adding More Features

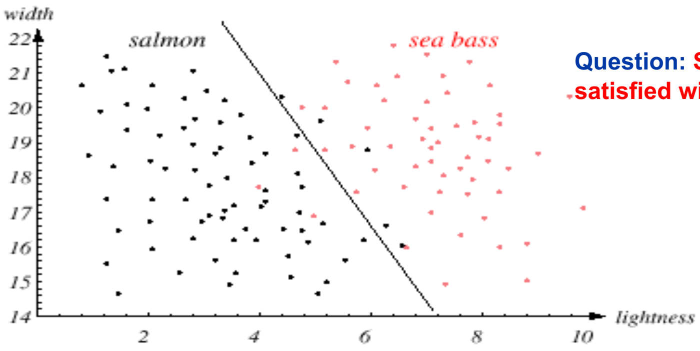
- Adopt the lightness and add the width of the fish



Dr. Ali Hassan, DCSE, College of E&ME, NUST 66

66

Decision rule: Use the linear boundary

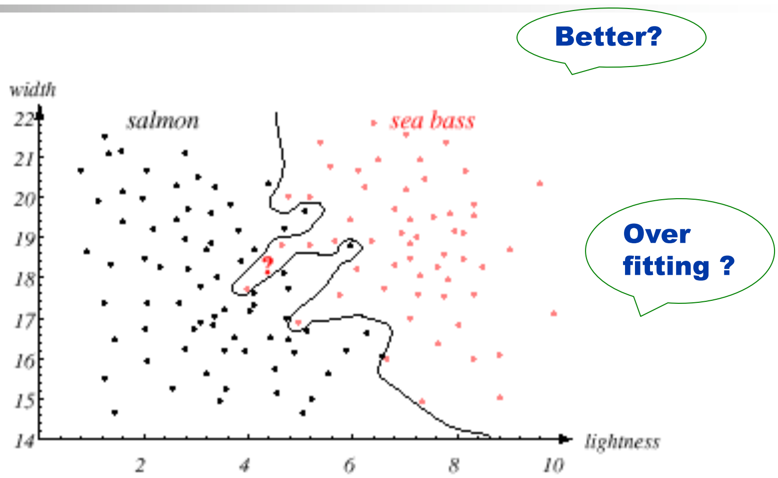


Question: Should we be satisfied with this result?

Simple Linear Boundary – training error > 0

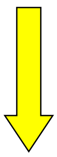
- We might add other features that are not correlated with the ones we already have. A precaution should be taken not to reduce the performance by adding such “noisy features”
- Ideally, the best decision boundary should be the one which provides an optimal performance
- Use a more complex model

Optimal Separation Boundary



Non-Linear Complex Boundary – training error = 0

- However, our satisfaction is premature because the central aim of designing a classifier is to correctly classify novel input

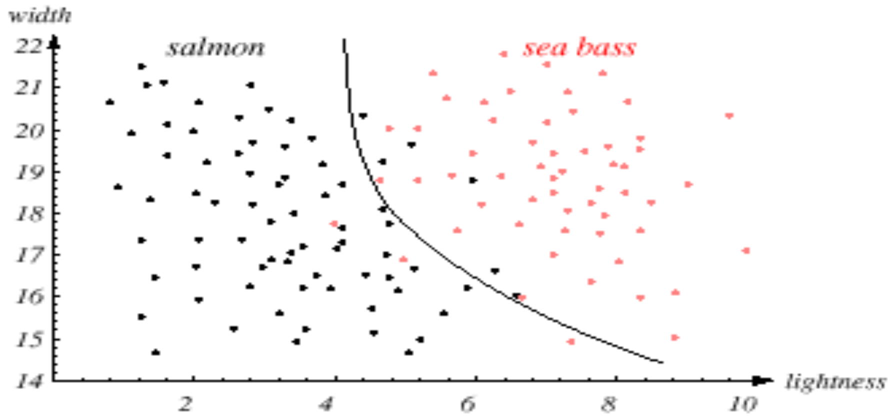


Issue of generalization!

Generalization

- The classifier should capture the underlying characteristics of the categories
- The classifier should NOT be tuned to the specific (accidental) characteristics of the training data
- A good classifier should be able to generalize, i.e. perform well on unseen data
- Training data in practice contain some noise

71



Simpler non-Linear Boundary – training error > 0

72

Decision Costs

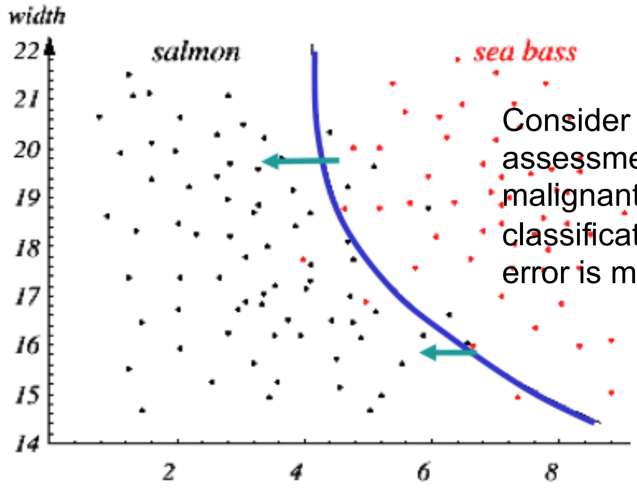
- A classifier, intuitively, is designed to minimize classification error, the total number of instances (fish) classified incorrectly.
  - Is this the best objective (cost) function to minimize?
  - What kinds of error can be made? Are they all equally bad? What is the real cost of making an error?
    - Sea bass misclassified as salmon: Pleasant surprise for the consumer, tastier fish/ manufacturer lose money for selling expensive fish for the cost of inexpensive fish
    - Salmon misclassified as sea bass: Customer upset, paid too much for inferior fish

Decision Costs …

- Threshold decision boundary and cost relationship
  - Move our decision boundary toward smaller values of lightness in order to minimize the cost (reduce the number of sea bass that are classified salmon!)

↓

Task of decision theory



Consider a similar risk assessment for malignant/benign tumor classification? Which error is more costly...?

Dr. Ali Hassan, DCSE, College of E&ME, NUST 75

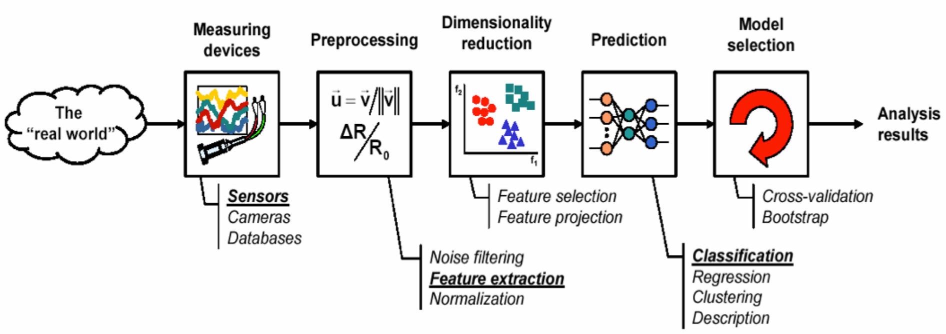
75

MACHINE LEARNING DESIGN CYCLE

76

76

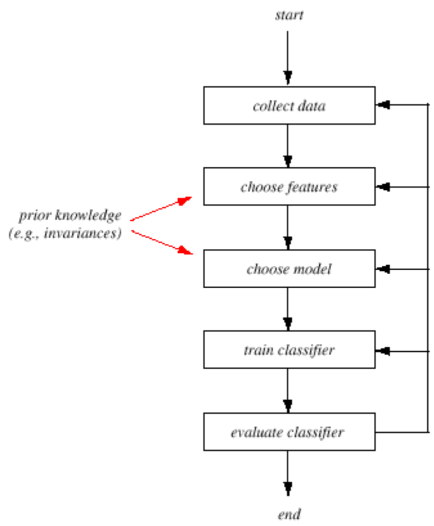
The Design Cycle



The Design Cycle

- Data collection
- Feature Choice
- Model Choice
- Training
- Evaluation
- Computational Complexity

*figure taken from Pattern Classification, Duda Hart



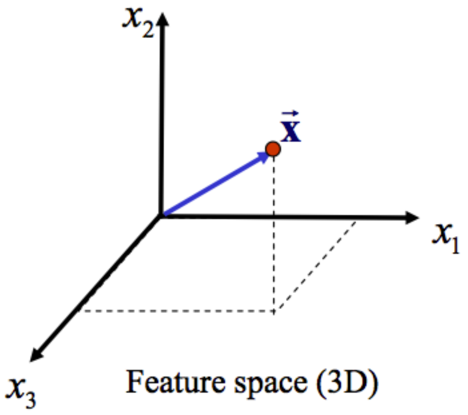
- Data Collection
  - How do we know when we have collected an adequately large and representative set of examples for training and testing the system?
- Pre-Processing
  - Filtering
  - Normalisation
  - How many examples of each class should be obtained

Features

- **Features**: a set of variables believed to carry discriminating and characterizing information about the objects under consideration
- **Feature vector**: A collection of d features, ordered in some meaningful way into a d- dimensional column vector, that represents the signature of the object to be identified.
- **Feature space**: The d-dimensional space in which the feature vectors lie. A d-dimensional vector in a d-dimensional space constitutes a point in that space.

Features

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_d \end{bmatrix} \begin{array}{l} \textit{feature 1} \\ \textit{feature 2} \\ \\ \textit{feature d} \end{array}$$



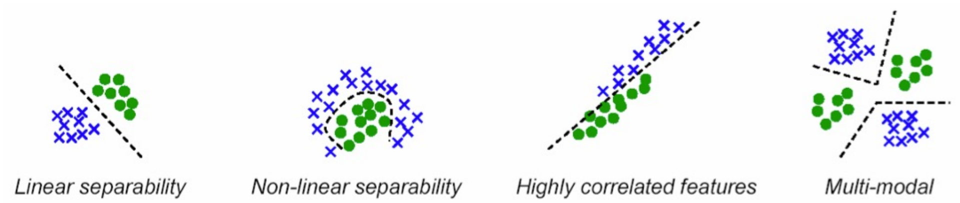
Feature space (3D)

Features

- Feature Choice
  - Good Features
    - Ideally, for a given group of patterns coming from the same class, feature values should all be similar
    - For patterns coming from different classes, the feature values should be different.
  - Bad Features
    - irrelevant, noisy, outlier?
- Bottom Line
  - In any PR problem, you get what you put in:
  - garbage in – garbage out (GIGO)!

Dr. Ali Hassan, DCSE, College of E&ME, NUST 83

83

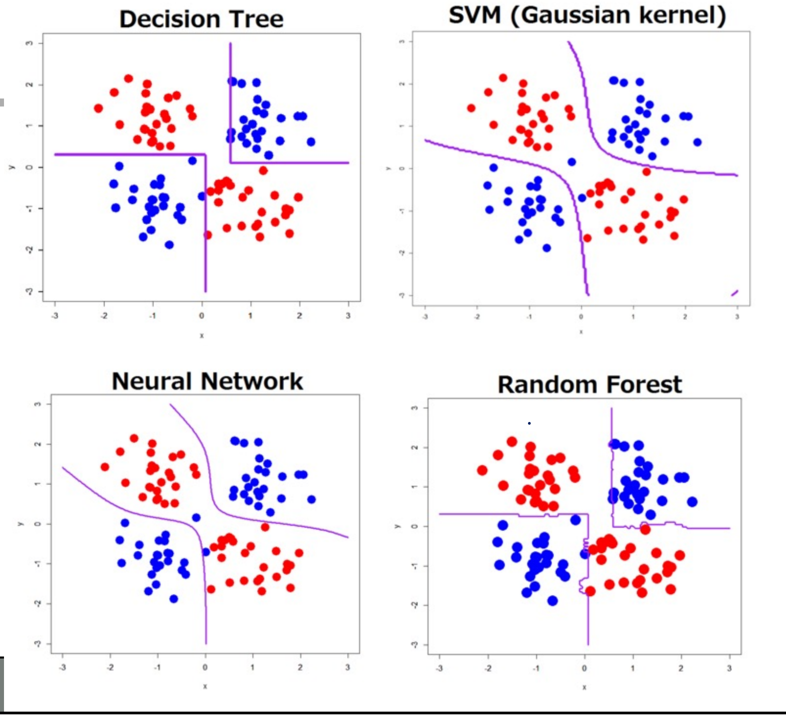


Dr. Ali Hassan, DCSE, College of E&ME, NUST 84

84

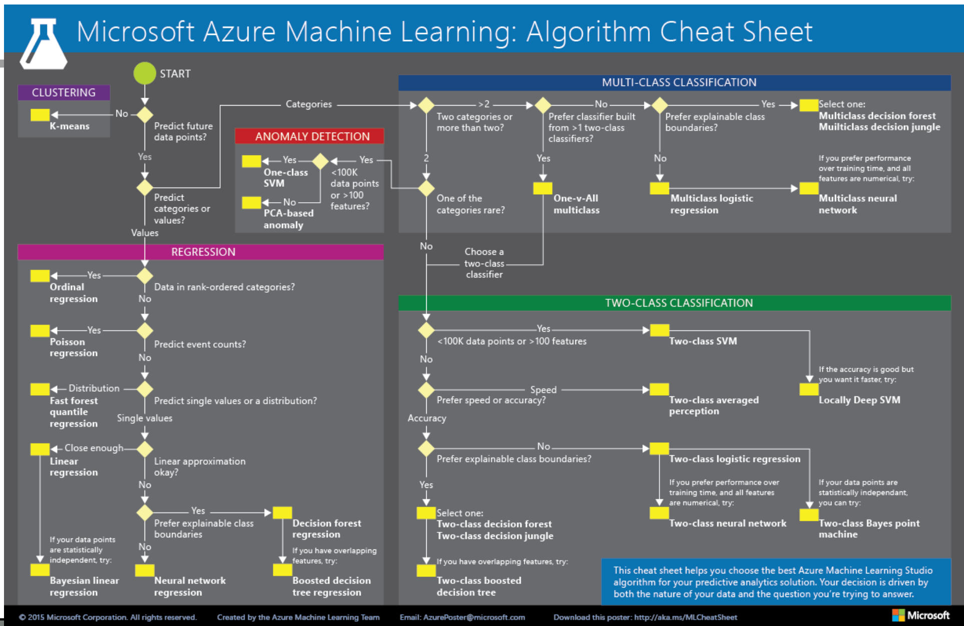
- Model Choice
  - What type of *classifier* shall we use? How shall we select its parameters? Is there best classifier...?
  - How do we train...? How do we adjust the parameters of the model (*classifier*) we picked so that the model fits the data?

- Training
  - Use data to determine the *classifier*. Many different procedures for training classifiers and choosing models



87

87



Dr. Ali Hassan, DCSE, College of E&ME, NUST

88

88

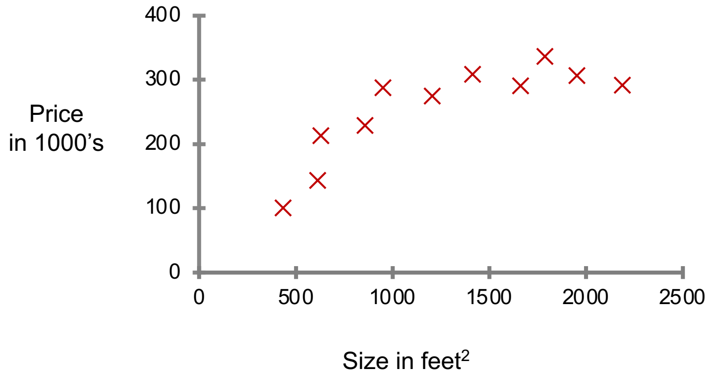
- Evaluation
  - Measure the error rate (or performance) and switch from one set of features to another one

- Computational Complexity
  - What is the trade-off between computational ease and performance?
  - (How an algorithm scales as a function of the number of features, patterns or categories?)

SUPERVISED LEARNING

91

Housing price prediction.

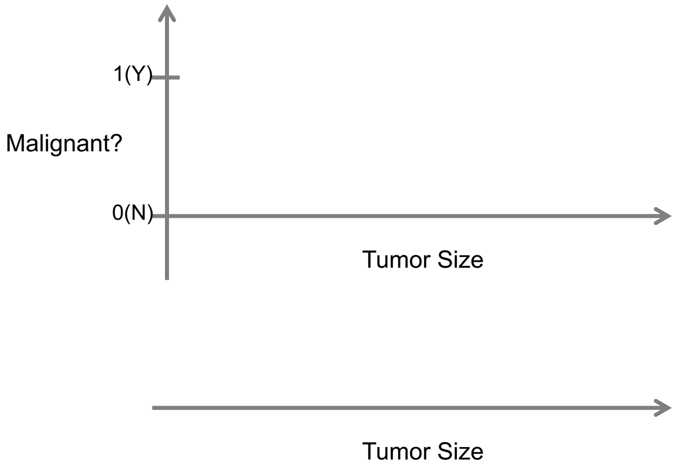


Supervised Learning
"right answers" given

Regression: Predict
continuous valued output
(price)

92

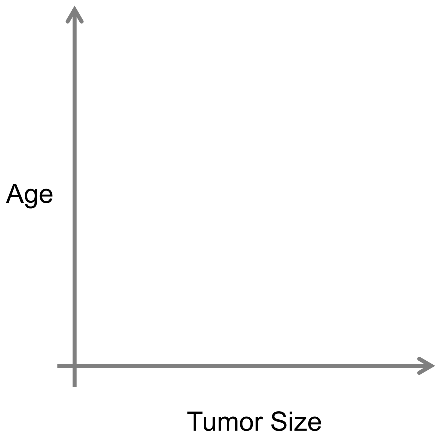
Cancer Problem (malignant, benign)



<u>Classification</u>

Discrete valued output (0 or 1)

93



- Clump Thickness
- Uniformity of Cell Size
- Uniformity of Cell Shape

…

94

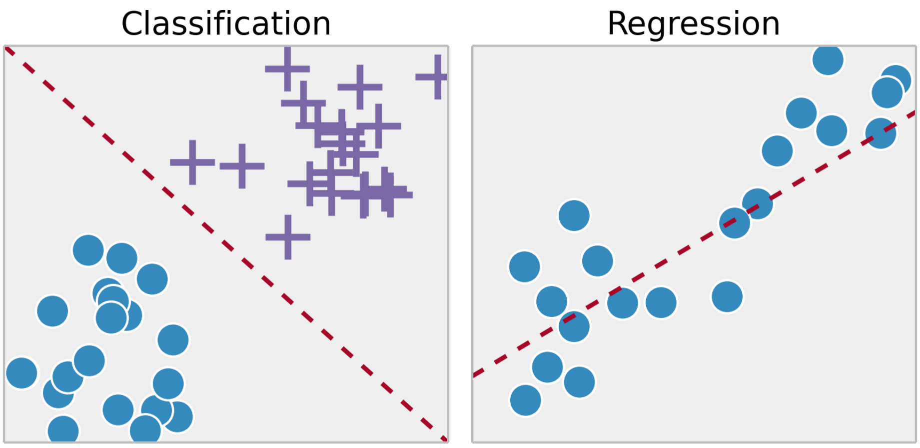
Examples of Supervised Learn.

- Image Classification:
  - Your train with images/labels. Then in the future you give a new image expecting that the computer will recognise the new object (Classification)
- Market Prediction:
  - You train the computer with historical market data and ask the computer to predict the new price in the future (Regression)

Dr. Ali Hassan, DCSE, College of E&ME, NUST 95

95

Examples of Supervised Learn.



Dr. Ali Hassan, DCSE, College of E&ME, NUST 96

96

Learning from Data

- Convert real world data into features
- We are learning from data
- Different from Traditional AI—rule based
- Data is typically high dimensional (i.e. many features/attributes)

IRIS Dataset

- Historic Data collected by R. A. Fisher
- Determining iris species
- A multi class classification Problem


Setosa


Versicolor


Virginica

IRIS Attributes

- Attributes
  - sepal length in cm
  - sepal width in cm
  - petal length in cm
  - petal width in cm

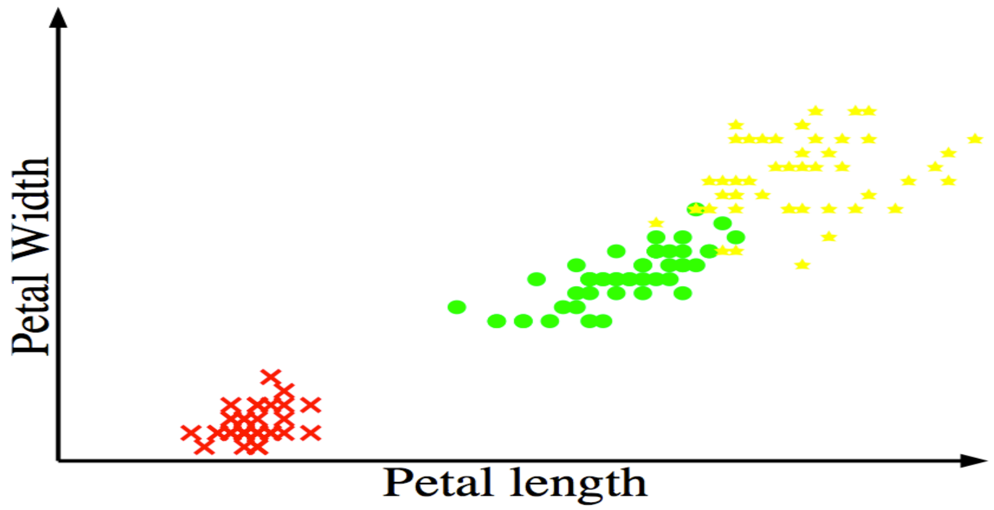
	Min	Max	Mean	SD	Class Correlation	
sepal length:	4.3	7.9	5.84	0.83	0.7826	
sepal width:	2.0	4.4	3.05	0.43	-0.4194	
petal length:	1.0	6.9	3.76	1.76	0.9490	(high!)
petal width:	0.1	2.5	1.20	0.76	0.9565	(high!)

Data

- 50 measurements from each species

  5.1,3.5,1.4,0.2,Iris-setosa
  4.9,3.0,1.4,0.2,Iris-setosa
  4.7,3.2,1.3,0.2,Iris-setosa
  4.4,2.9,1.4,0.2,Iris-setosa
  7.0,3.2,4.7,1.4,Iris-versicolor
  6.4,3.2,4.5,1.5,Iris-versicolor
  6.9,3.1,4.9,1.5,Iris-versicolor
  5.5,2.3,4.0,1.3,Iris-versicolor
  6.3,3.3,6.0,2.5,Iris-virginica
  5.8,2.7,5.1,1.9,Iris-virginica
  …

What Data Looks Like



HANDLING NON-NUMERIC DATA

Non-Numeric Data

- Inferring Salary from attributes
  - Age
  - Workclass: Private, Self-emp-not-inc, Self-emp-inc, Federal-gov, Local-gov, State-gov, Without-pay, Never-worked.
  - Education: Bachelors, Some-college, HS-grad, Prof-school, Masters, Doctorate
  - marital-status: Married-civ-spouse, Divorced, Never-married, Separated, Widowed, Married-spouse-absent, Married-AF-spouse

Data

39, State-gov, 77516, Bachelors, 13, Never-married, Adm-clerical, Not-in-family, White, Male, 2174, 0, 40, US, <=50K

50, Self-emp-not-inc, 83311, Bachelors, 13, Married-civ-spouse, Exec-managerial, Husband, White, Male, 0, 0, 13, US, <=50K

38, Private, 215646, HS-grad, 9, Divorced, Handlers-cleaners, Not-in-family, White, Male, 0, 0, 40, US, <=50K

53, Private, 234721, 11th, 7, Married-civ-spouse, Handlers-cleaners, Husband, Black, Male, 0, 0, 40, US, <=50K

28, Private, 338409, Bachelors, 13, Married-civ-spouse, Prof-specialty, Wife, Black, Female, 0, 0, 40, Cuba, <=50K

52, Self-emp-not-inc, 209642, HS-grad, 9, Married-civ-spouse, Exec-managerial, Husband, White, Male, 0, 0, 45, US, >50K

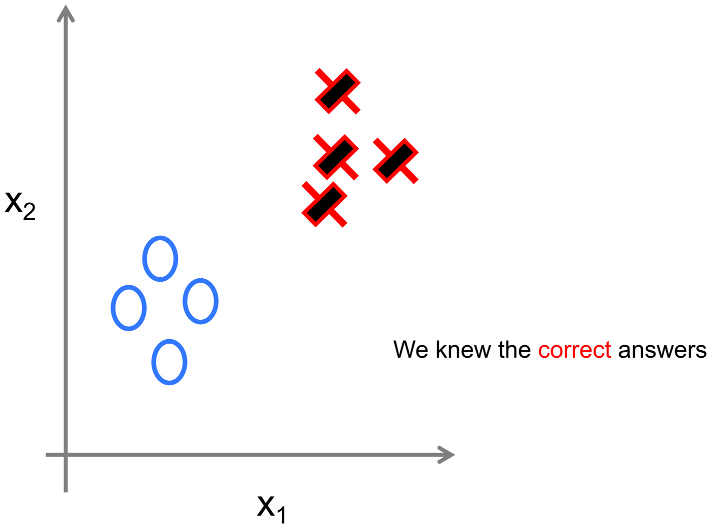
31, Private, 45781, Masters, 14, Never-married, Prof-specialty, Not-in-family, White, Female, 14084, 0, 50, US, >50K

19, ?, 37332, HS-grad, 9, Never-married, ?, Own-child, White, Female, 1055, 0, 12, US, <=50K

UNSUPERVISED LEARNING

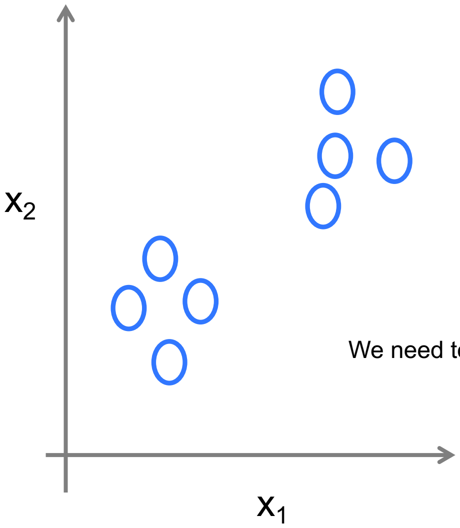
105

Supervised Learning



106

Unsupervised Learning



We need to figure out the patterns

107

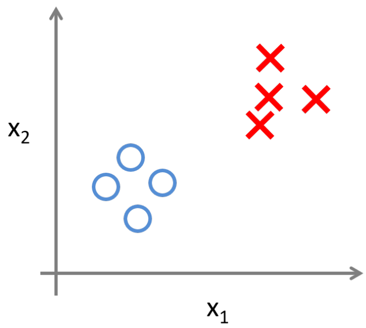
Types of ML

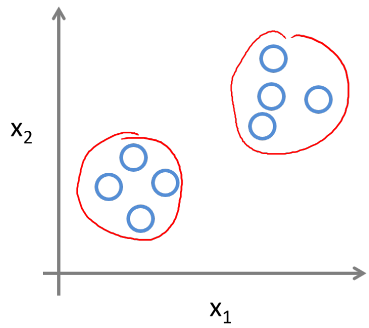
- Supervised Learning:
  - You give to the computer some pairs of inputs/outputs, so in the future new when new inputs are presented you have an intelligent output.
- Unsupervised Learning:
  - You let the computer learn from the data itself without showing what is the expected output.

108

Types of ML

Supervised Learning | Unsupervised Learning

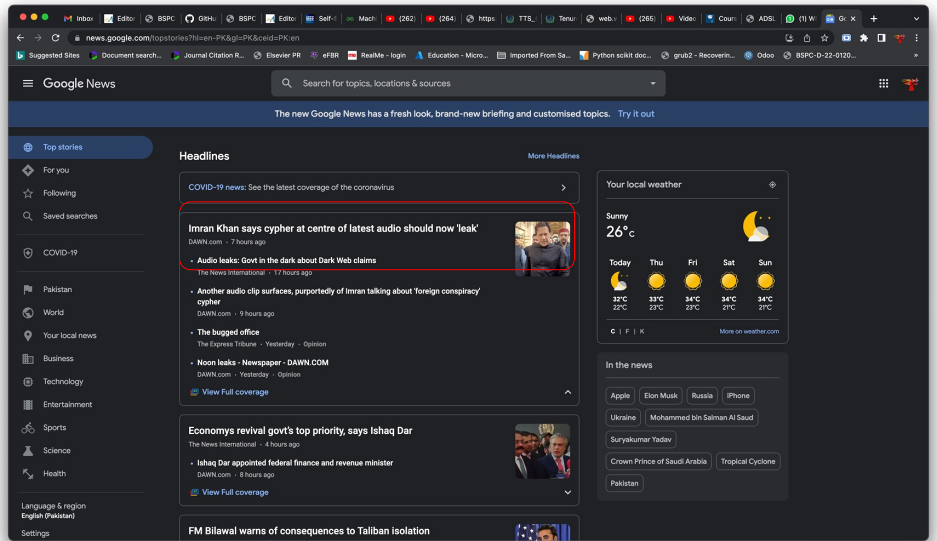




Dr. Ali Hassan, DCSE, College of E&ME, NUST 109

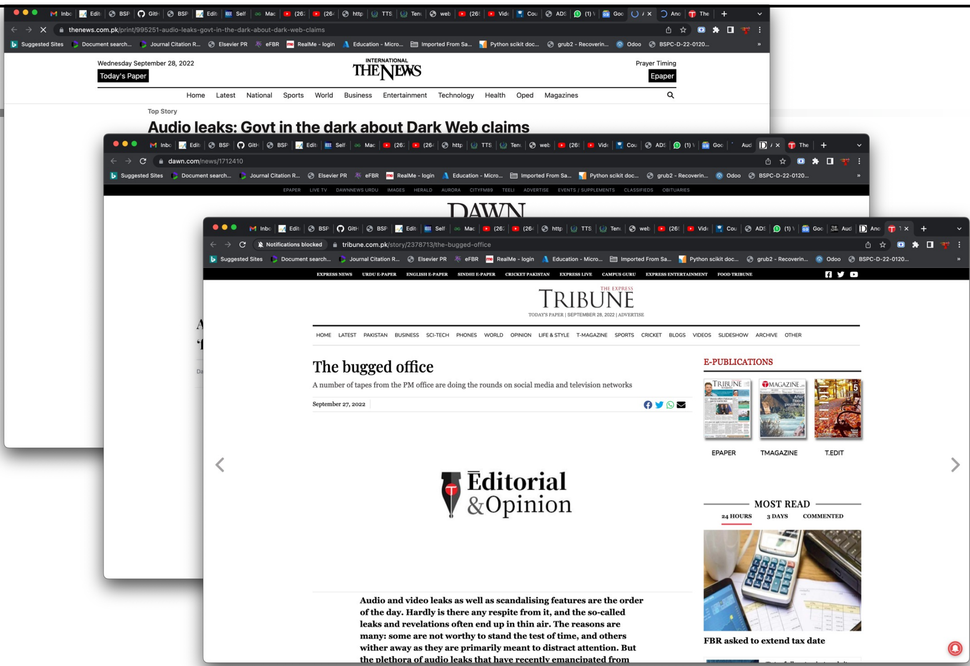
109

An Application of Unsupervised Learning



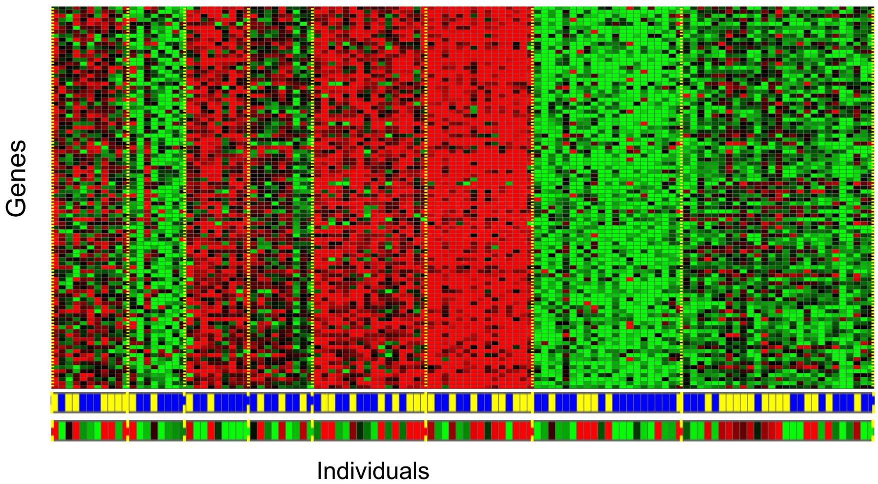
Dr. Ali Hassan, DCSE, College of E&ME, NUST 110

110



Dr. Ali Hassan, DCSE, College of E&ME, NUST 111

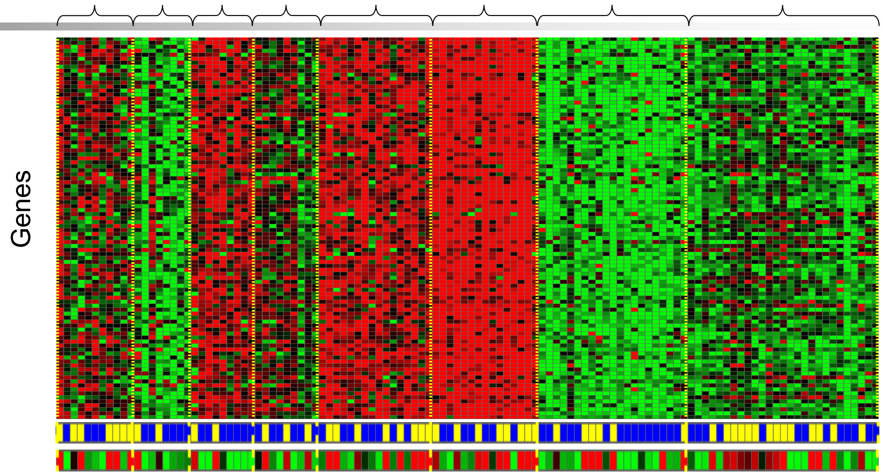
111



[Source: Daphne Koller]

Dr. Ali Hassan, DCSE, College of E&ME, NUST 112

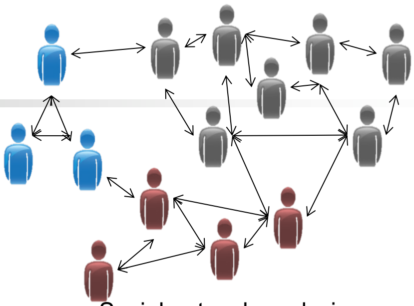
112



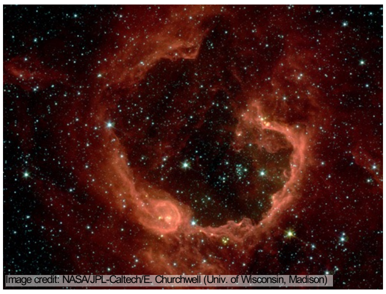
[Source: Daphne Koller]

113

Organize computing clusters



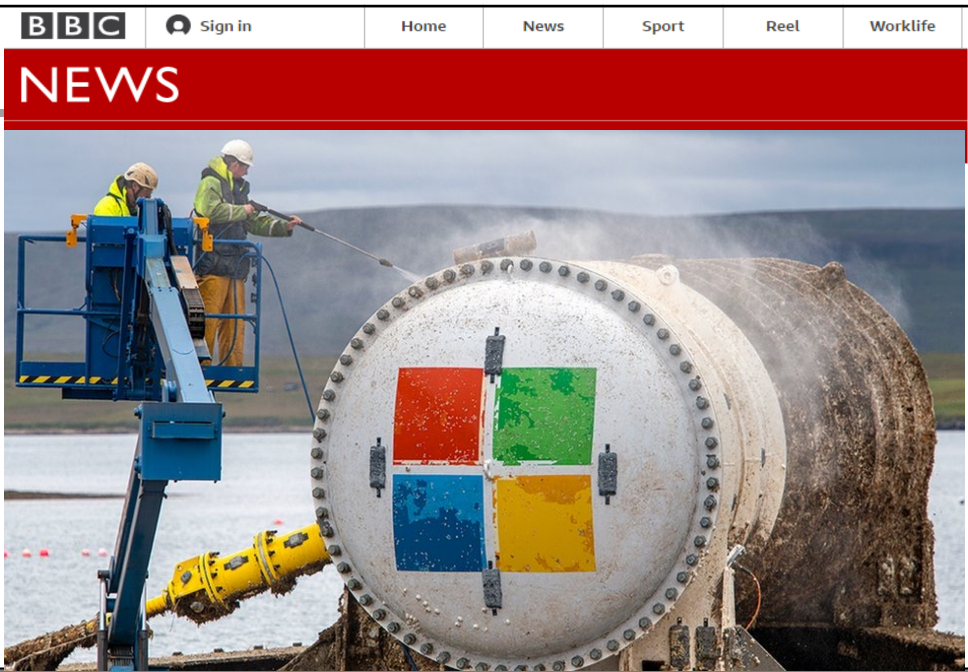
Social network analysis



Market segmentation

Astronomical data analysis

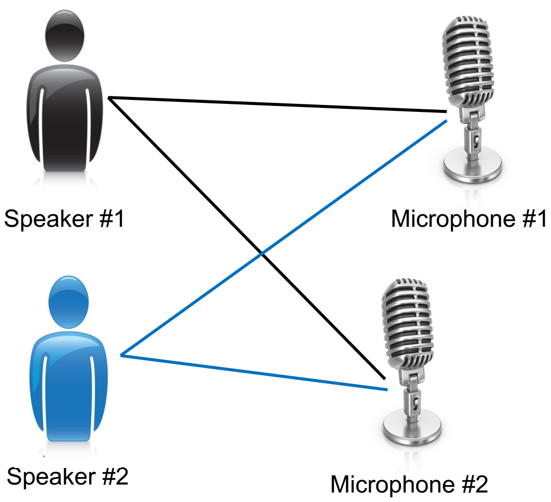
114



Dr. Ali Hassan, DCSE, College of E&ME, NUST 115

115

Cocktail party problem



Dr. Ali Hassan, DCSE, College of E&ME, NUST 116

116

Microphone #1: Output #1:

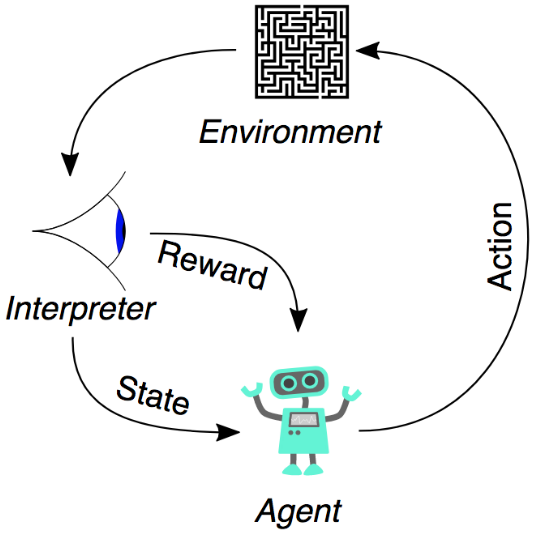
Microphone #2: Output #2:

[Audio clips courtesy of Te-Won Lee.]

Cocktail party problem algorithm

```
[W,s,v] = svd((repmat(sum(x.*x,1),size(x,1),1).*x)*x');
```

Reinforced Learning

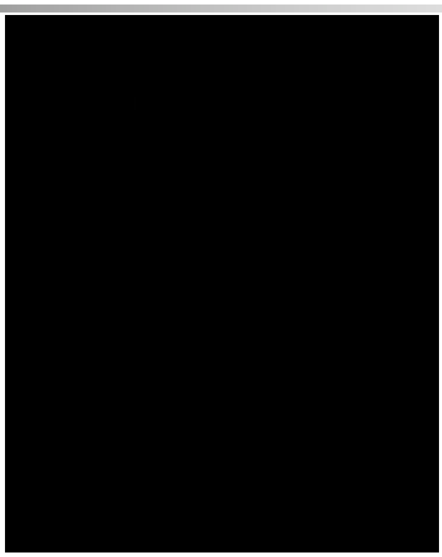


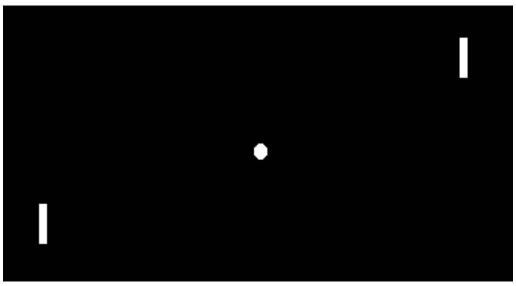
The typical framing of a Reinforcement Learning (RL) scenario: an agent takes actions in an environment, which is interpreted into a reward and a representation of the state, which are fed back into the agent.

119

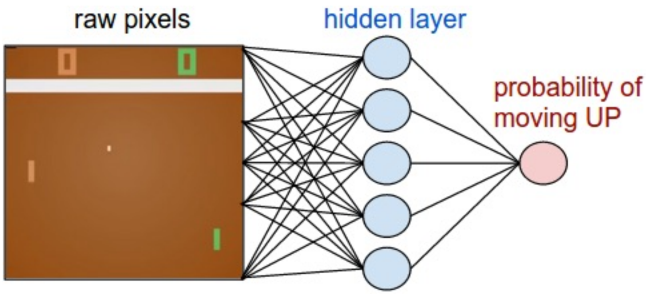
Examples of RL

- How to do this using supervised machine learning?
- How to do this using RL?



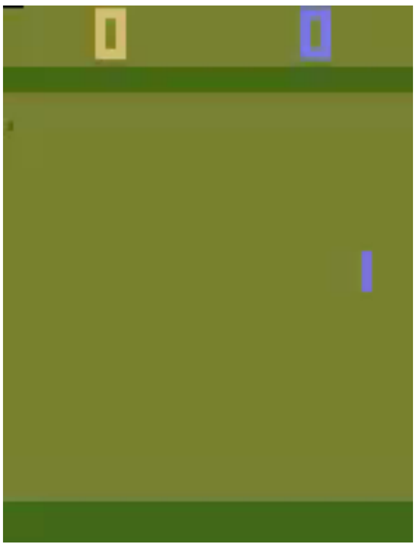


120

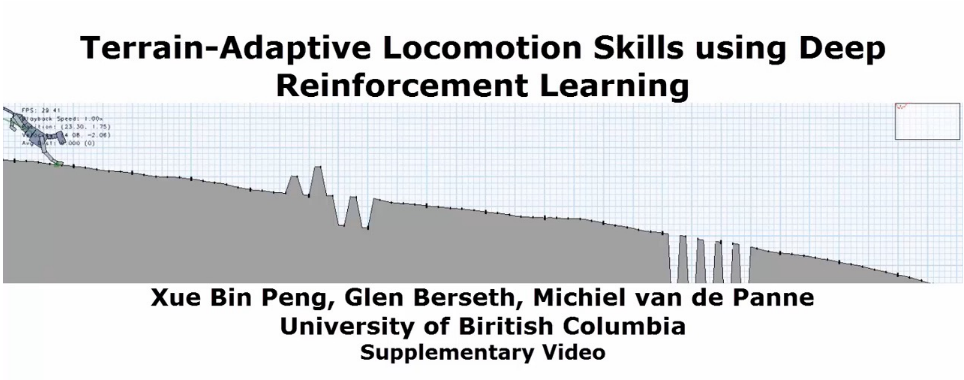


Our policy network is a 2-layer fully-connected net

121



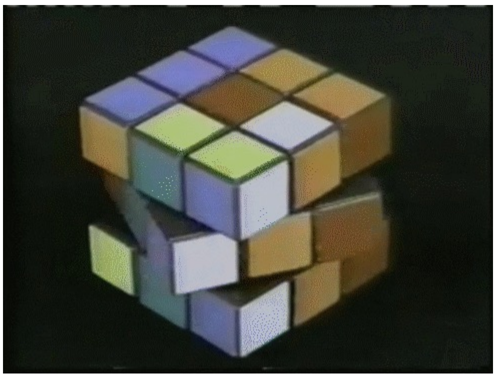
122



Dr. Ali Hassan, DCSE, College of E&ME, NUST 123

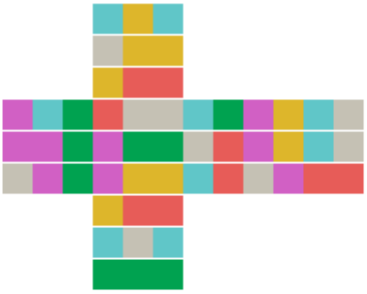
123

Learning To Solve a Rubik's Cube From Scratch using RL

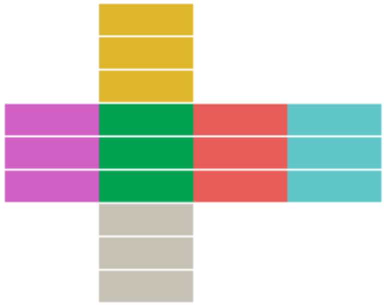


Dr. Ali Hassan, DCSE, College of E&ME, NUST 124

124



Flattened Cube



Flattened solved Cube

Dr. Ali Hassan, DCSE, College of E&ME, NUST 125

125

Solving the Rubik's Cube Without Human Knowledge

Stephen McAleer*
Department of Statistics
University of California, Irvine
smcaleer@uci.edu

Forest Agostinelli*
Department of Computer Science
University of California, Irvine
fagostin@uci.edu

Alexander Shmakov*
Department of Computer Science
University of California, Irvine
ashmakov@uci.edu

Pierre Baldi
Department of Computer Science
University of California, Irvine
pfbaldi@ics.uci.edu

Abstract

A generally intelligent agent must be able to teach itself how to solve problems in complex domains with minimal human supervision. Recently, deep reinforcement learning algorithms combined with self-play have achieved superhuman proficiency in Go, Chess, and Shogi without human data or domain knowledge. In these environments, a reward is always received at the end of the game; however, for many combinatorial optimization environments, rewards are sparse and episodes are not guaranteed to terminate. We introduce Autodidactic Iteration: a novel reinforcement learning algorithm that is able to teach itself how to solve the Rubik's Cube with no human assistance. Our algorithm is able to solve 100% of randomly scrambled cubes while achieving a median solve length of 30 moves — less than or equal to solvers that employ human domain knowledge.

Dr. Ali Hassan, DCSE, College of E&ME, NUST 126

126

LINEAR ALGEBRA

Scalar Vector Matrix

Scalar

24

Vector

$\begin{bmatrix} 2 & -8 & 7 \end{bmatrix}$

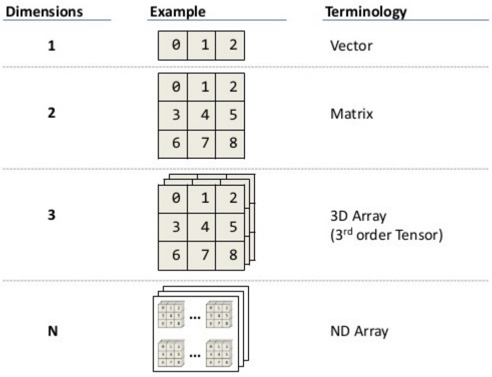
row

or column $\begin{bmatrix} -6 \\ -4 \\ 27 \end{bmatrix}$

Matrix

$\begin{bmatrix} 6 & 4 & 24 \\ 1 & -9 & 8 \end{bmatrix}$

row(s) × column(s)



129

Command Window

```
>> a = 24;
>> b = [2 -8 7]

b =

     2    -8     7

>> c = [-6;-4;27]

c =

    -6
    -4
    27

>> D = [6 4 24; 1 -9 8]

D =

     6     4    24
     1    -9     8

>> D(2,3)

ans =

     8

>> b(3)

ans =

     7

>> size(D)

ans =

     2     3
```

Workspace | MATLAB Coder: C/C++ S

Name	Value	Min	Max
a	24	24	24
ans	[2,3]	2	3
b	[2,-8,7]	-8	7
c	[-6;-4;27]	-6	27
D	[6,4,24;1,-9,8]	-9	24

130

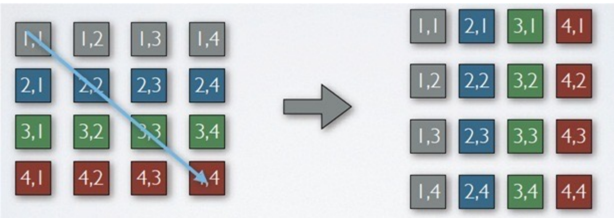
```
leo@monsterpc2:~$ ipython
Python 2.7.6 (default, Jun 22 2015, 17:58:13)
Type "copyright", "credits" or "license" for more information.

IPython 2.3.0 -- An enhanced Interactive Python.
?         -> Introduction and overview of IPython's features.
%quickref -> Quick reference.
help      -> Python's own help system.
object?   -> Details about 'object', use 'object??' for extra details.

In [1]: import numpy as np

In [2]: a = 24

In [3]: b = np.array([[2, -8, 7]])

In [4]: c = np.array([[-6],[-4],[27]])

In [5]: D = np.array([[6,4,24],[1,-9,8]])

In [6]: print(b)
[[ 2 -8  7]]

In [7]: print(c)
[[-6]
 [-4]
 [27]]

In [8]: print(D)
[[ 6  4 24]
 [ 1 -9  8]]

In [9]: D[1,2]
Out[9]: 8

In [10]: D.shape
Out[10]: (2, 3)
```

131

131

Matrix Operations

- Transpose



Dr. Ali Hassan, DCSE, College of E&ME, NUST 132

132

A

$$\begin{bmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{bmatrix}$$

- $\begin{bmatrix} 1 & 2 \end{bmatrix}^{\mathrm{T}} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$
- $\begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}^{\mathrm{T}} = \begin{bmatrix} 1 & 3 \\ 2 & 4 \end{bmatrix}$
- $\begin{bmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{bmatrix}^{\mathrm{T}} = \begin{bmatrix} 1 & 3 & 5 \\ 2 & 4 & 6 \end{bmatrix}$

Dr. Ali Hassan, DCSE, College of E&ME, NUST 133

133

Addition/Subtraction

3+4=7

$$\begin{bmatrix} 3 & 8 \\ 4 & 6 \end{bmatrix} + \begin{bmatrix} 4 & 0 \\ 1 & -9 \end{bmatrix} = \begin{bmatrix} 7 & 8 \\ 5 & -3 \end{bmatrix}$$

3-4=-1

$$\begin{bmatrix} 3 & 8 \\ 4 & 6 \end{bmatrix} - \begin{bmatrix} 4 & 0 \\ 1 & -9 \end{bmatrix} = \begin{bmatrix} -1 & 8 \\ 3 & 15 \end{bmatrix}$$

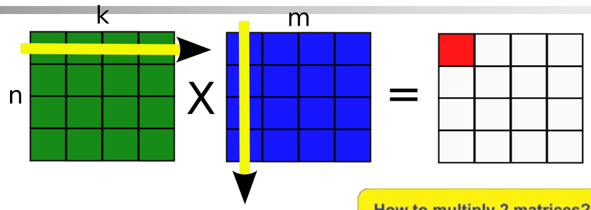
Dr. Ali Hassan, DCSE, College of E&ME, NUST 134

134

Multiply by scalar

$$2 \times \begin{bmatrix} 4 & 0 \\ 1 & -9 \end{bmatrix} \xrightarrow{2x4=8} \begin{bmatrix} 8 & 0 \\ 2 & -18 \end{bmatrix}$$

Matrix Multiplication



How to multiply 2 matrices?

$$\begin{bmatrix} 1 & 2 \\ 2 & 4 \\ 3 & 5 \end{bmatrix} \times \begin{bmatrix} 1 & 2 \\ 2 & 4 \end{bmatrix} \quad 1x1 + 2x2 = 1 + 4 = 5$$

$$\begin{bmatrix} 1 & 2 \\ 2 & 4 \\ 3 & 5 \end{bmatrix} \times \begin{bmatrix} 1 & 2 \\ 2 & 4 \end{bmatrix} \quad 1x2 + 2x4 = 2 + 8 = 10$$

$$\begin{bmatrix} 1 & 2 \\ 2 & 4 \\ 3 & 5 \end{bmatrix} \times \begin{bmatrix} 1 & 2 \\ 2 & 4 \end{bmatrix} \quad 2x1 + 4x2 = 2 + 8 = 10$$

& so on ....

Types of Matrix

$$I = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

137

Google colab

Visual Studio Code

python™

PYTHON DEMO

ANACONDA®

138

Reading …

- Pattern Classification, Duda, Hart, Stork
  - Chapter 1: Introduction
- Deep Reinforcement Learning: Pong from Pixels
  - (http://karpathy.github.io/2016/05/31/rl/)
- Solving the Rubik's Cube Without Human Knowledge
  - https://arxiv.org/pdf/1805.07470.pdf

139

Next Class

- Linear Regression – Uni-variate

140