

# Machine Learning

## Lecture # 1 Introduction & Fundamentals

# My Introduction

- Name: Muhammad Zubair Khan
- Ph.D. Computer Science (USA)
  - Area of research: Medical imaging, Machine Learning, Deep Learning
  - Email: [mzkhan@ucmo.edu](mailto:mzkhan@ucmo.edu)
  - Office #: MIC D255

# Technology Requirements

- The latest version of GitHub Desktop at  
<https://desktop.github.com/>
- The latest version of Python IDE (Desktop or Web)
- A current word processing software
- A headset with microphone
- A webcam
- Firefox, Chrome, or Safari browser.
- A Broadband Internet connection is preferred. Examples of broadband Internet connection are high-speed DSL or a Cable modem

- Course Material
  - Lectures slides, assignments (computer/written), projects, and announcements will be uploaded on blackboard.

# Before we start

How many of you  
are good in  
probability and  
linear algebra?

How many of you  
are familiar with  
Python and  
Machine  
Learning?

# What is Machine Learning?

- Machine Learning
  - Study of algorithms that
  - improve their performance
  - at some task
  - with experience
- Optimize a performance criterion using example data or past experience.
- Role of Statistics: Inference from a sample
- Role of Computer science: Efficient algorithms

# What is Machine Learning?

- Adapt to / learn from data
  - To optimize a performance function

Can be used to:

- Extract knowledge from data
- Learn tasks that are difficult to formalise
- Create software that improves over time

# An Introduction to Machine Learning

- According to Herbert Simon, learning is, “Any **change in a System** that allows it to **perform better the second time** on repetition of the same task or on another task drawn from the same population.” [G. F. Luger and W.A. Stubblefield, *Artificial Intelligence: Structures and Strategies for Complex Problem Solving*, The Benjamin/Cummings Publishing Company, Inc. 1989.]

# Why “Learn”?

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

# Machine Learning

## Applications:

- Retail: Market basket analysis, Customer relationship management (CRM)
- Finance: Credit scoring, fraud detection
- Manufacturing: Optimization, troubleshooting
- Medicine: Medical diagnosis
- Telecommunications: Quality of service optimization
- Web mining: Search engines

# Growth of Machine Learning

- Machine learning is preferred approach to
  - Speech recognition, Natural language processing
  - Computer vision
  - Medical outcomes analysis
  - Robot control
  - Computational biology
- This trend is accelerating
  - Improved machine learning algorithms
  - Improved data capture, networking, faster computers
  - New sensors / IO devices

# Categories

- Association Analysis
- Supervised Learning
  - Classification
  - Regression/Prediction
- Unsupervised Learning
- Reinforcement Learning

# Learning Associations

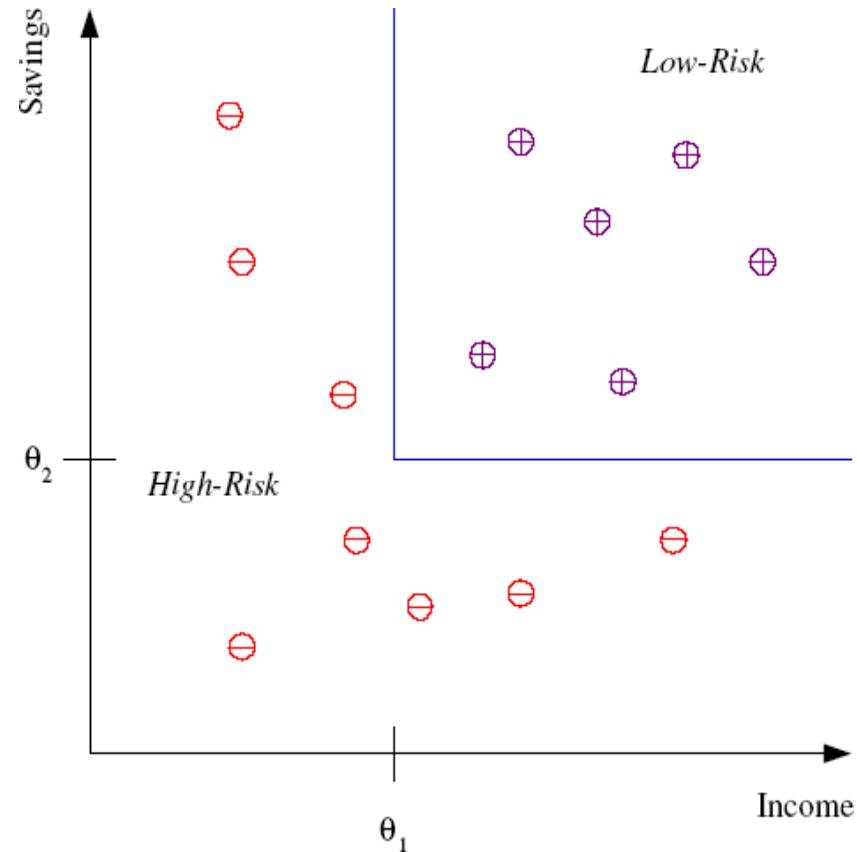
- Basket analysis:

$P(Y|X)$  probability that somebody who buys  $X$  also buys  $Y$  where  $X$  and  $Y$  are products/services.

Example:  $P(\text{bread} | \text{cold drink}) = 0.7$

# Classification

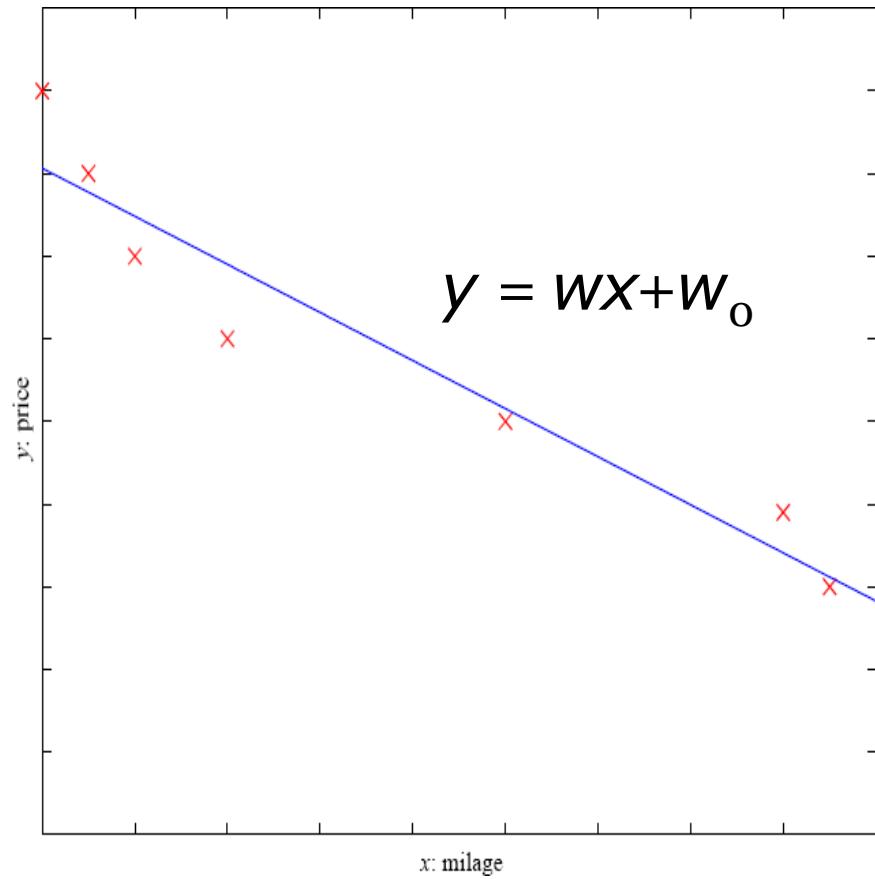
- Example: Credit scoring
- Differentiating between **low-risk** and **high-risk** customers from their *income* and *savings*



**Discriminant:** IF  $income > \theta_1$  AND  $savings > \theta_2$   
THEN **low-risk** ELSE **high-risk**

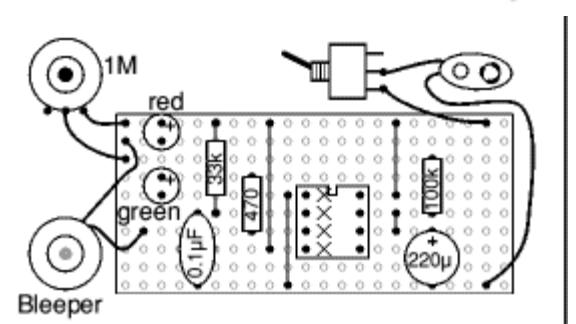
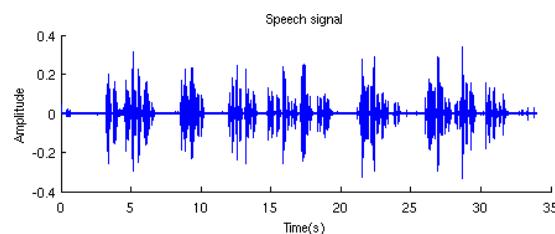
# Prediction: Regression

- Example: Price of a used car
  - $x$ : car attributes  
 $y$ : price
- $$y = g(x | \theta)$$
- $g()$  model,  
 $\theta$  parameters



# Pattern

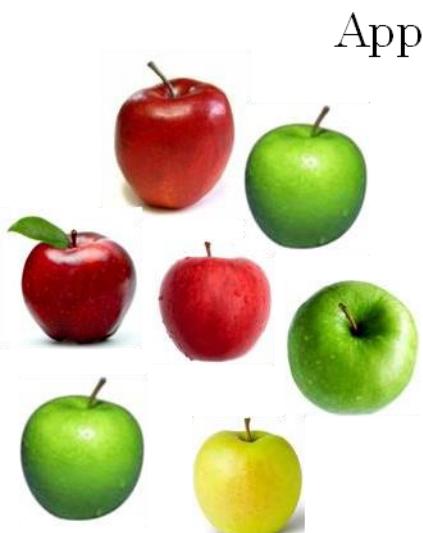
A pattern is the **opposite of a chaos**, it is an entity that can be given a name



# Recognition

- Identification of a pattern as a member of a category

# Classification



Apples



Oranges



# Classification

- You had some training example or '*training data*'
- The examples were '*labeled*'
- You used those examples to make the kid '*learn*' the difference between an apple and an orange



What is this???



Its an  
apple!!!



# Classification



Apple

Pear

Tomato

Cow

Dog

Horse



Given: training images and their categories

What are the categories  
of these test images?

# Pattern Recognition

Given an input pattern, **make a decision** about  
the “category” or “class” of the pattern

# Unsupervised Learning

- Learning “what normally happens”
- No output
- Clustering: Grouping similar instances
- Other applications: Summarization, Association Analysis
- Example applications
  - Image compression: Color quantization
  - Bioinformatics: Learning motifs

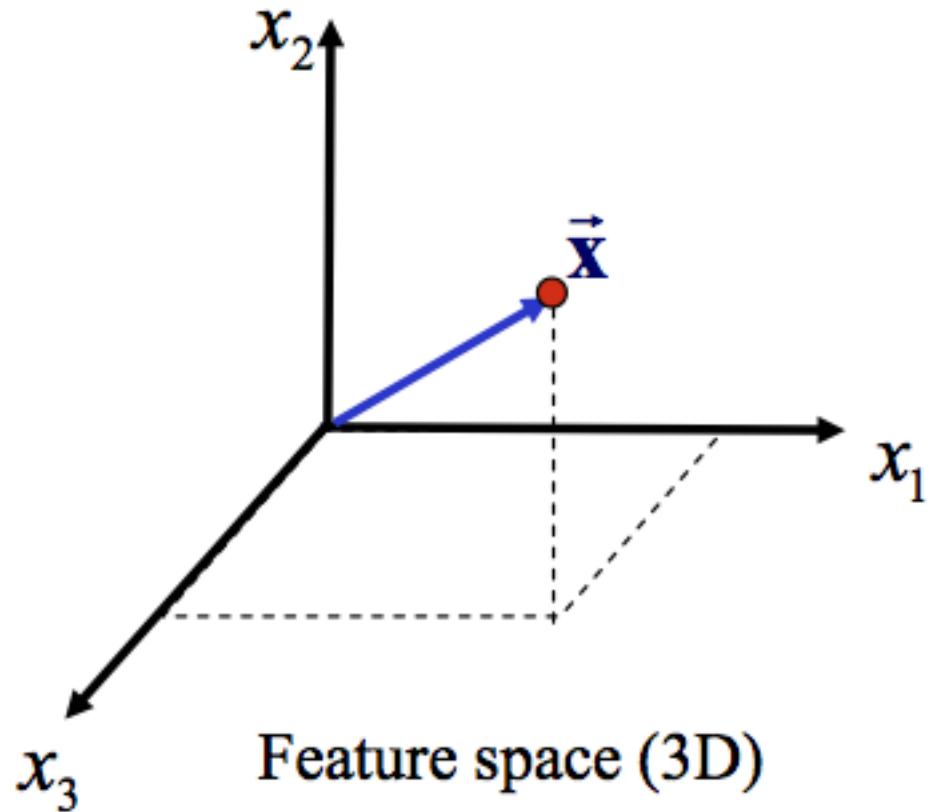
# Model Choice

- 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?

# 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.

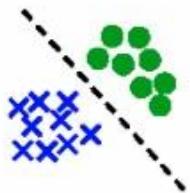
# Features



# 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?

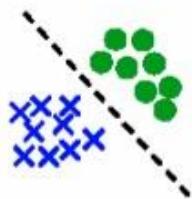
# Features



*"Good" features*



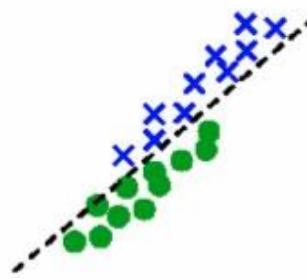
*"Bad" features*



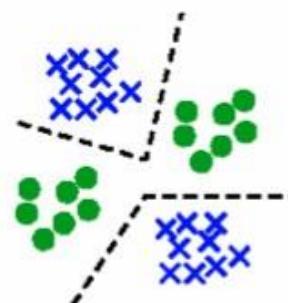
*Linear separability*



*Non-linear separability*



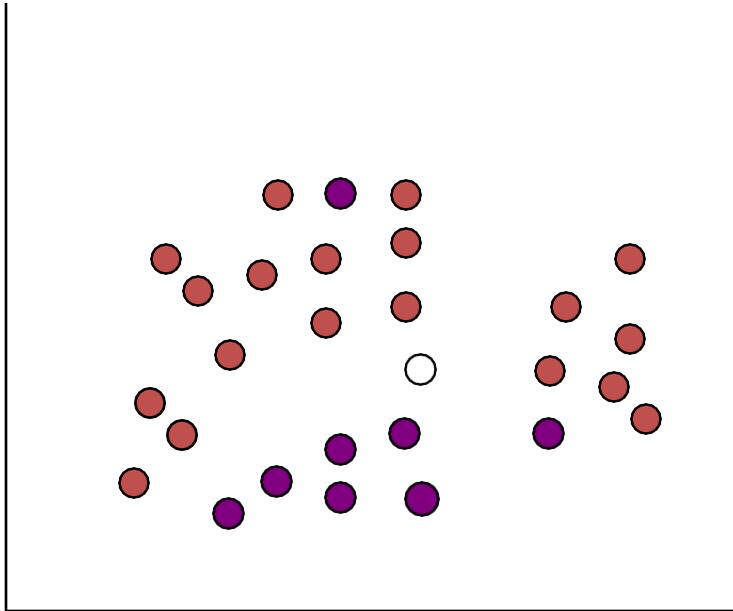
*Highly correlated features*



*Multi-modal*

# Classification

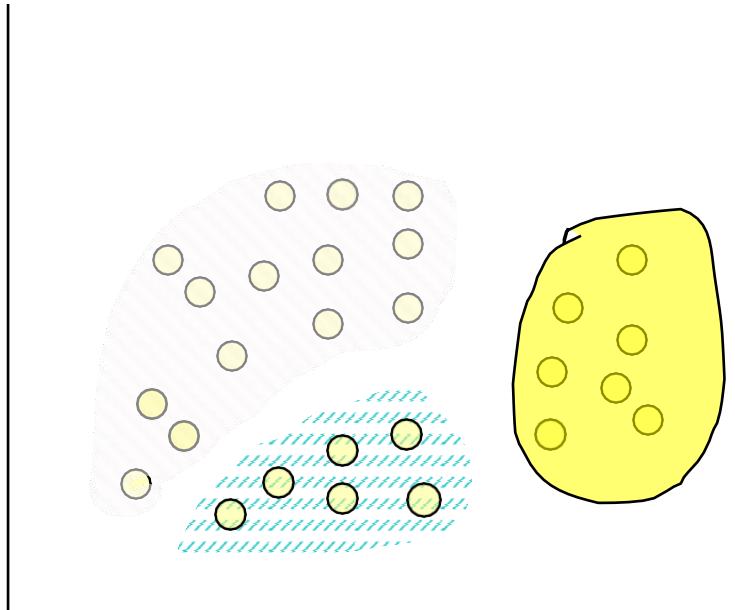
**Learn a method for predicting the instance class from pre-labeled (classified) instances**



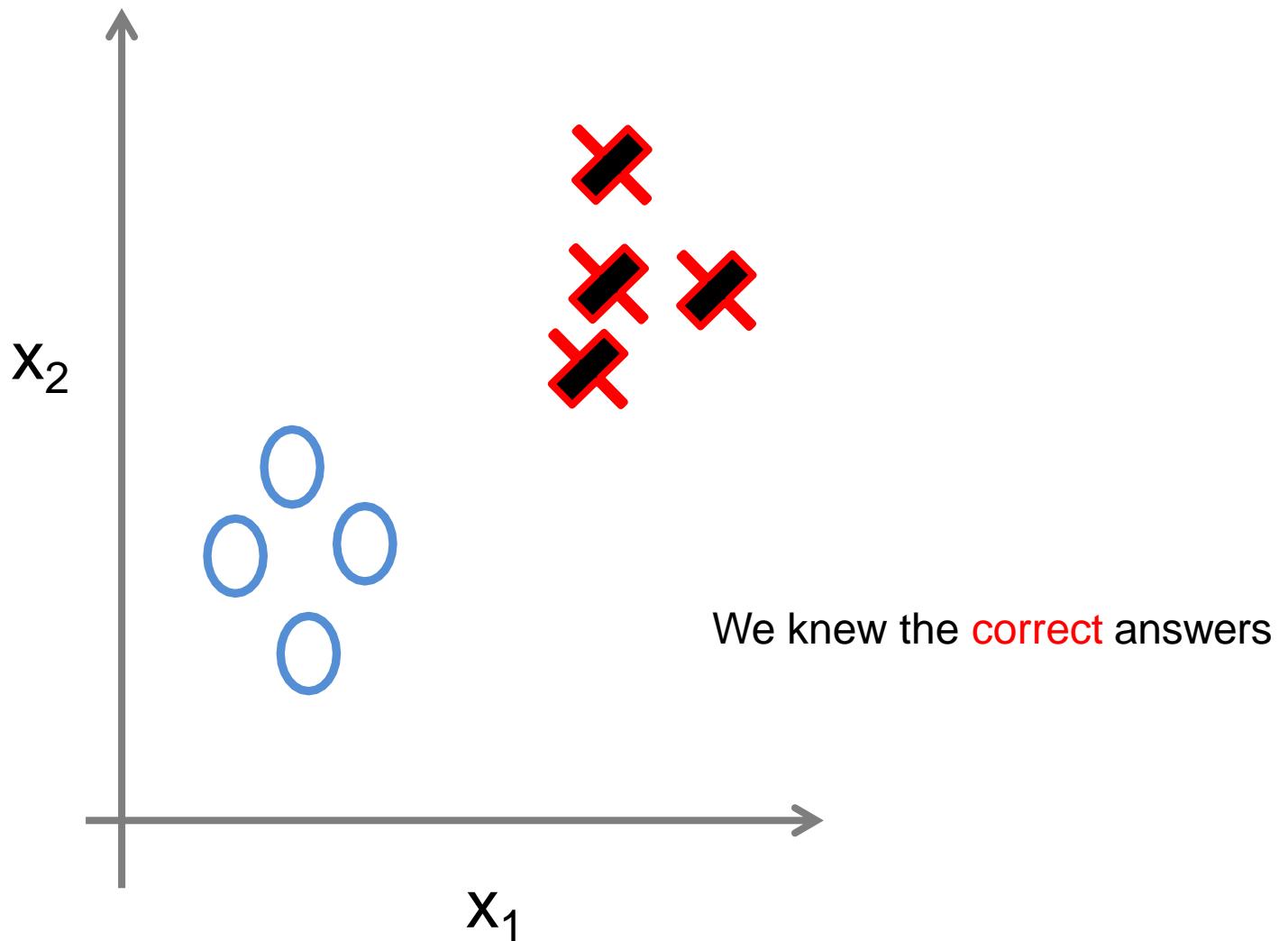
Many approaches:  
Statistics,  
Decision Trees, Neural  
Networks,  
...

# Clustering

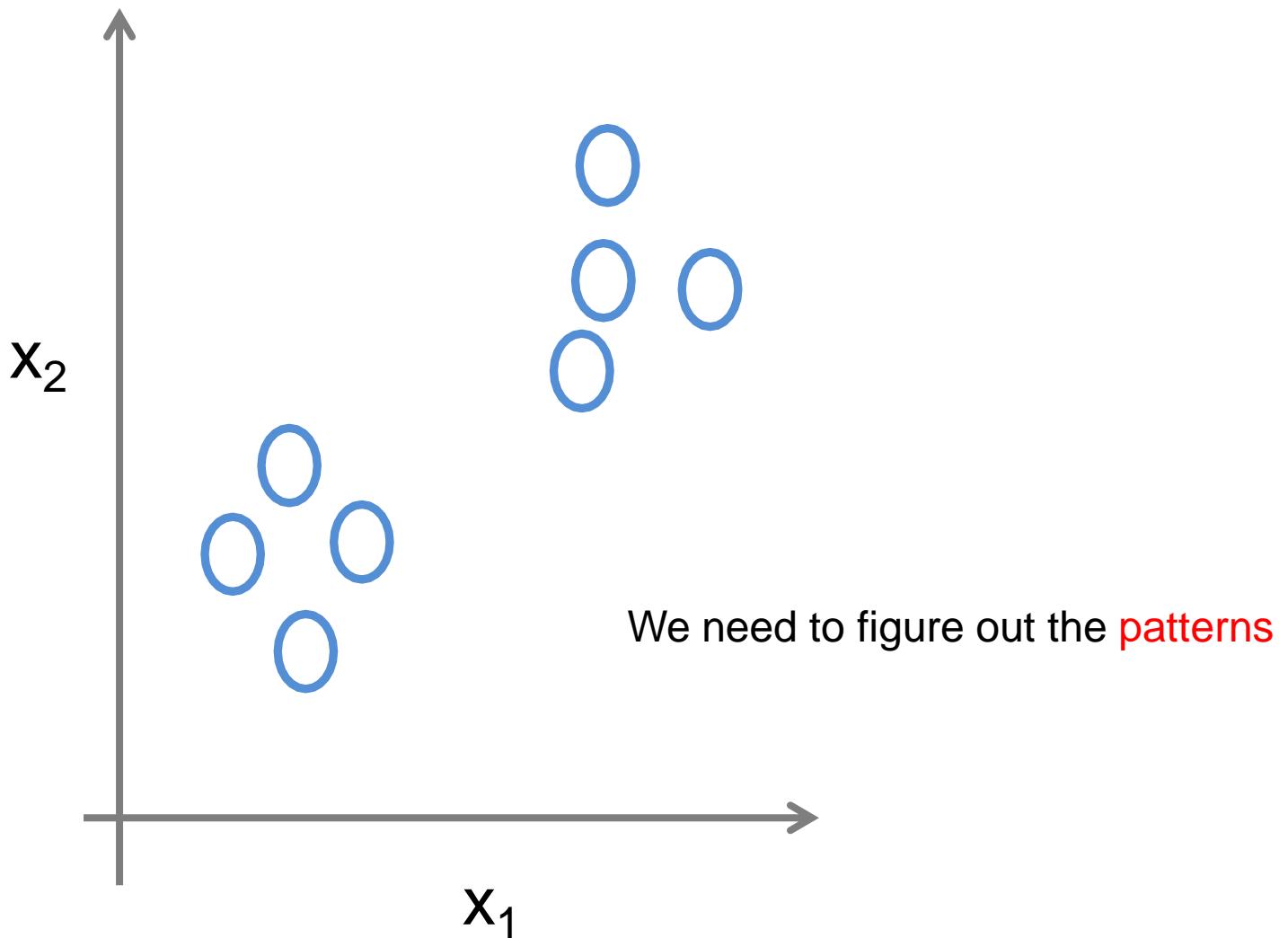
Find “natural” grouping of instances  
given un-labeled data



# Supervised Learning

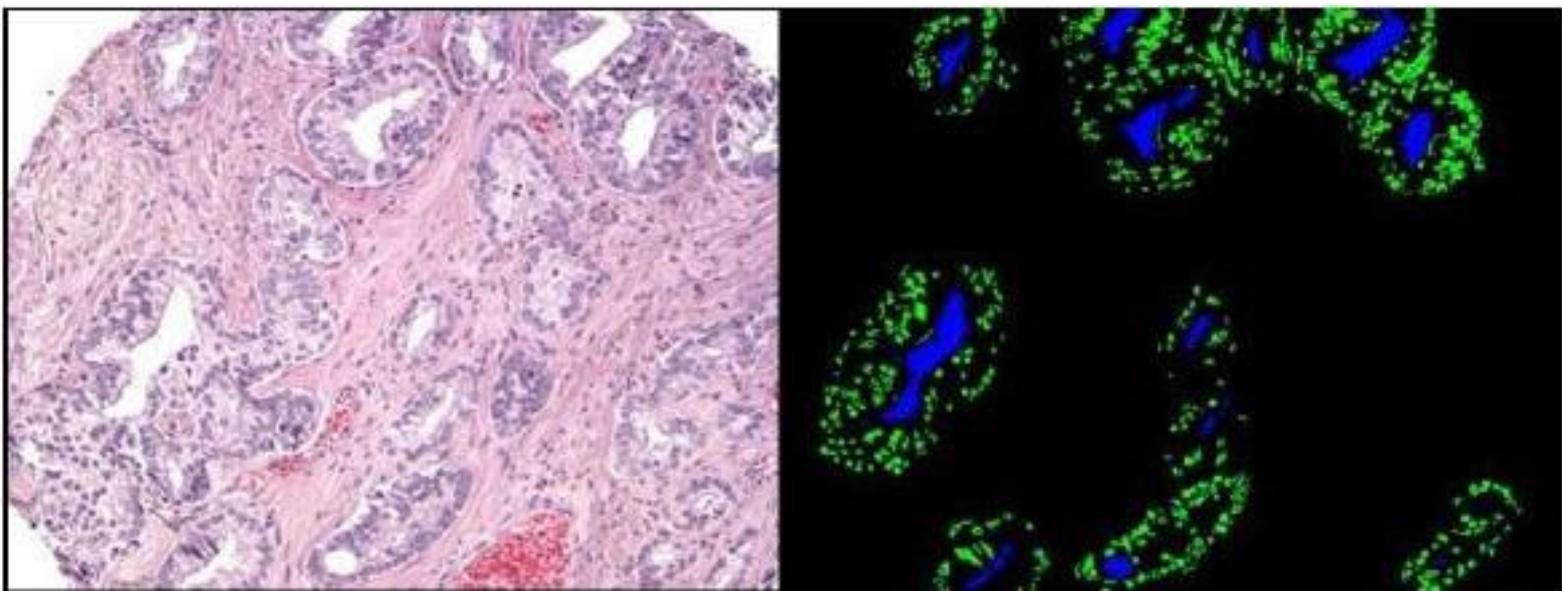


# Unsupervised Learning



# Example Applications

# Examples: Medicine

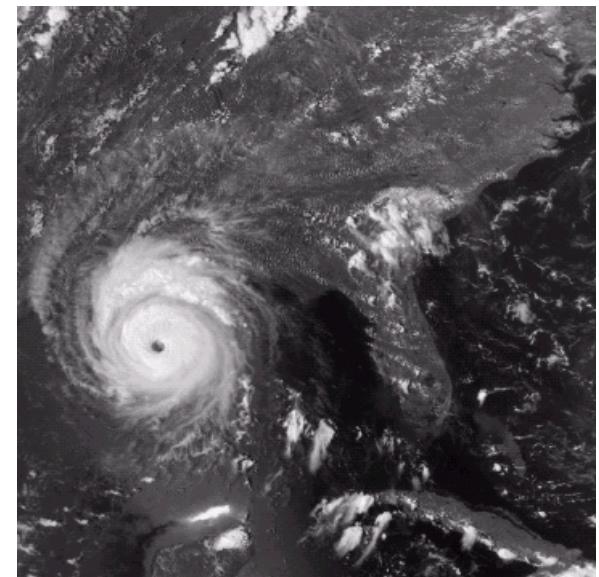
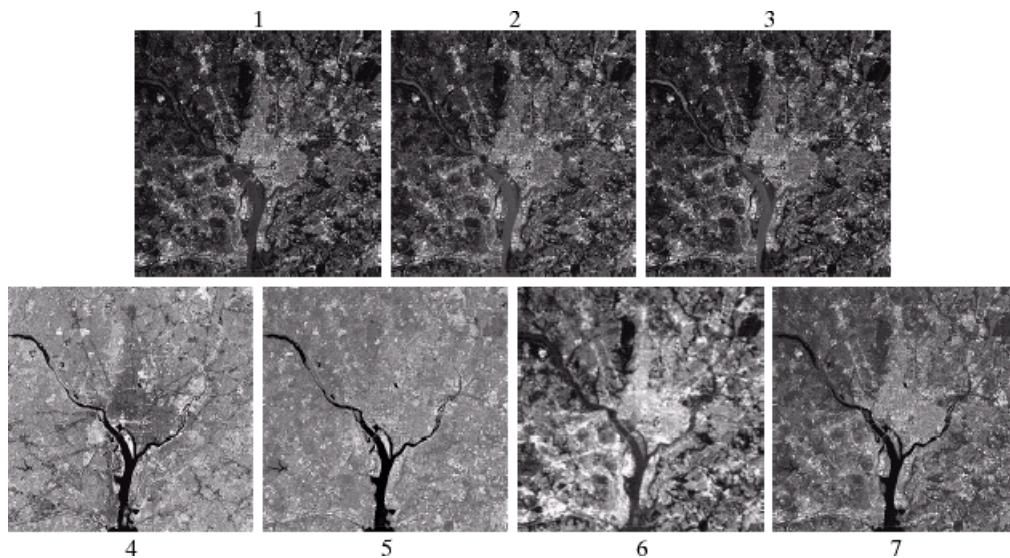


Microscopic tissue data - Cancer Detection

# Examples: GIS

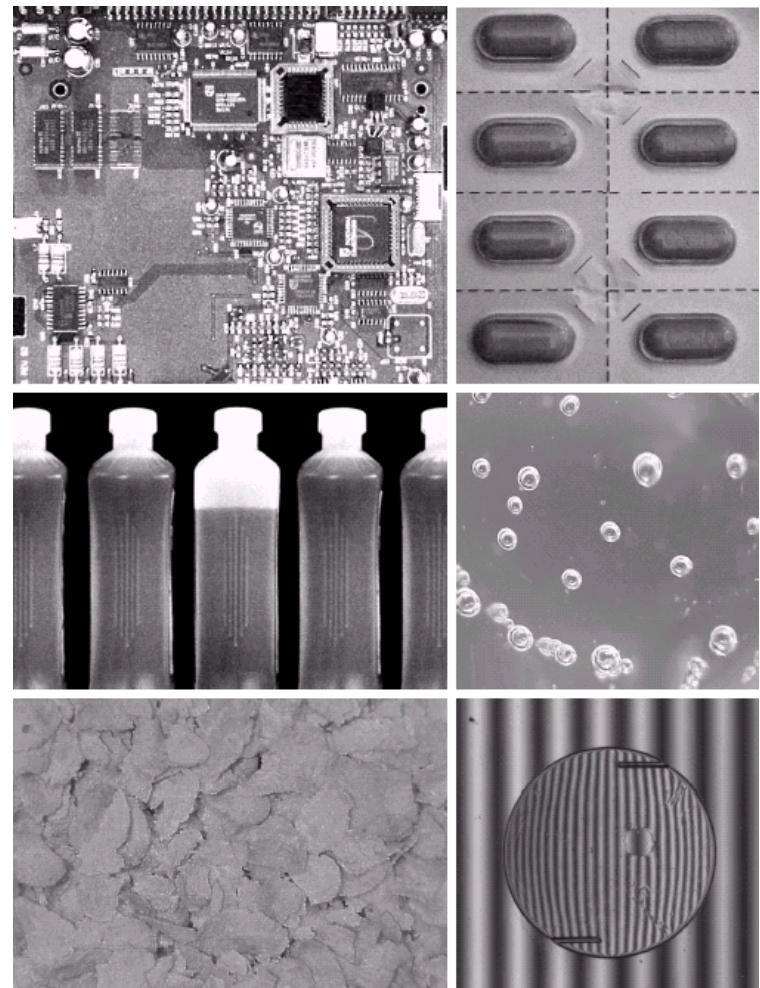
## ◆ Geographic Information Systems

- Manipulation of Satellite Imagery
- Terrain Classification, Meteorology



# Examples: Industrial Inspection

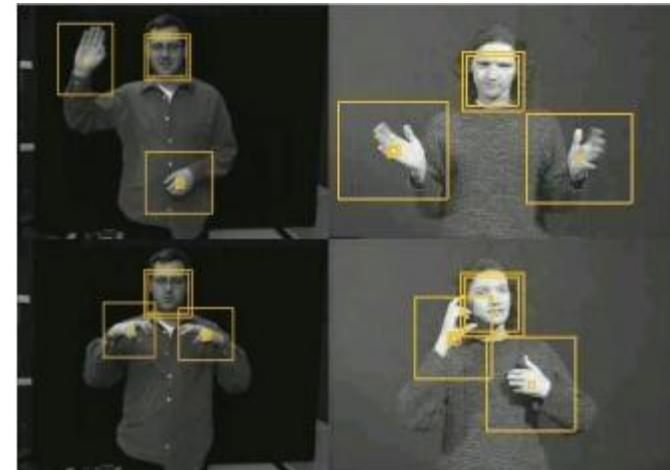
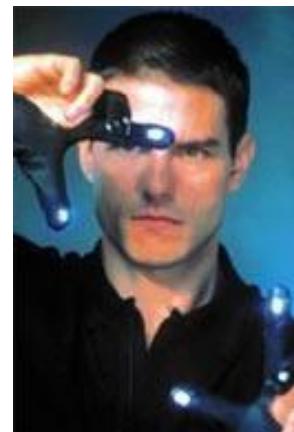
- ◆ Human operators are expensive, slow and unreliable
- ◆ Make machines do the job instead



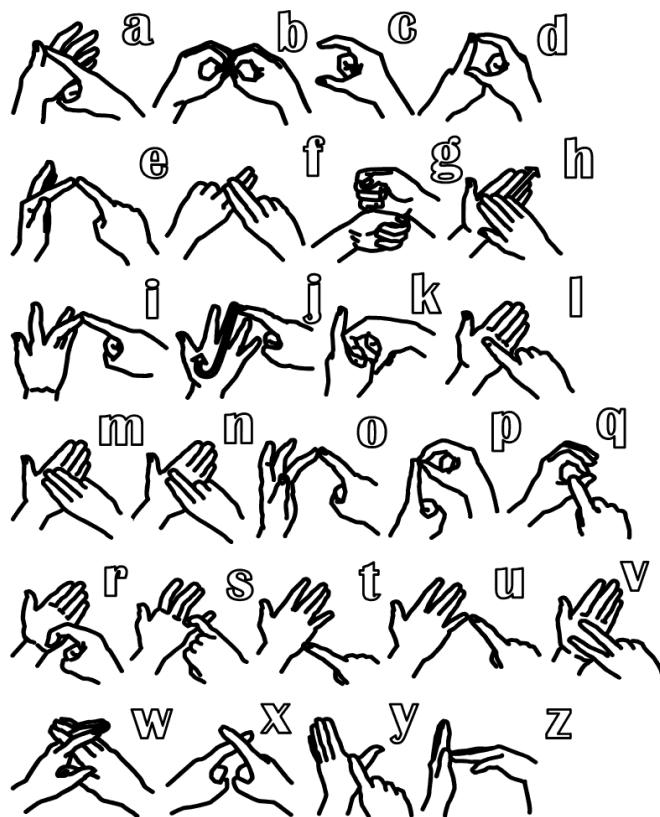
# Examples: HCI

◆ Try to make human computer interfaces more natural

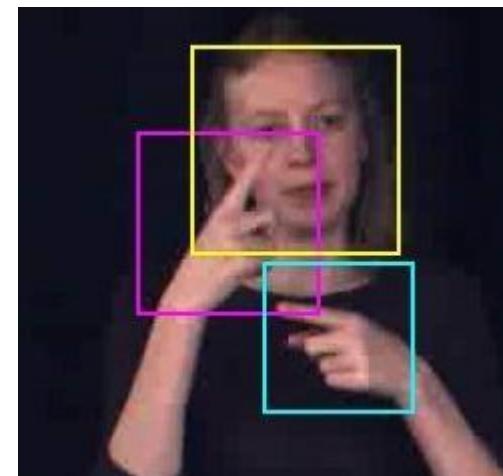
- Gesture recognition
- Facial Expression Recognition
- Lip reading



# Examples: Sign Language/Gesture Recognition



British Sign Language Alphabet



# Examples: Facial Expression Recognition

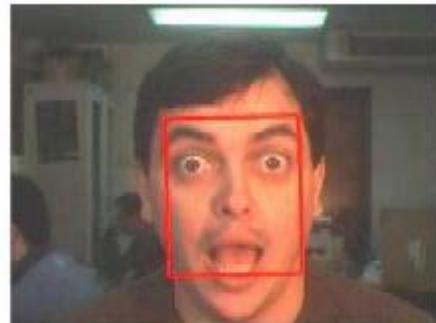
## ■ Implicit customer feedback

Upper Face Action Units		
AU4	AU1+4	AU1+2
		
Brows lowered and drawn together	Medial portion of the brows is raised and pulled together	Inner and outer portions of the brows are raised
AU5	AU6	AU7
		
Upper eyelids are raised	Cheeks are raised and eye opening is narrowed	Lower eyelids are raised

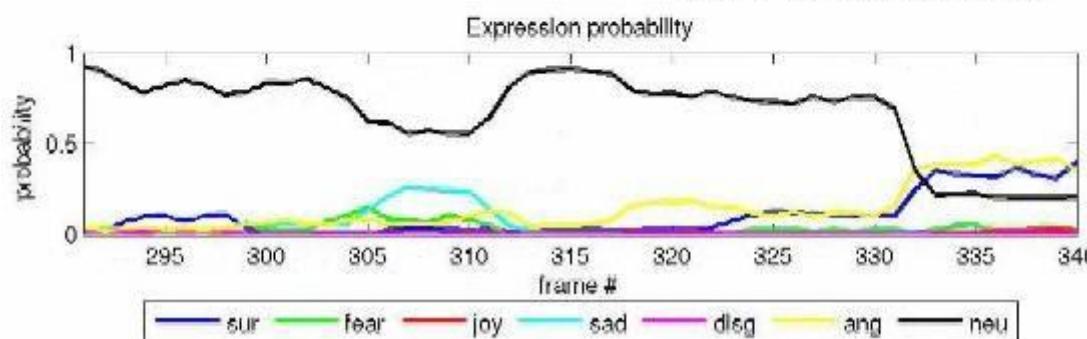
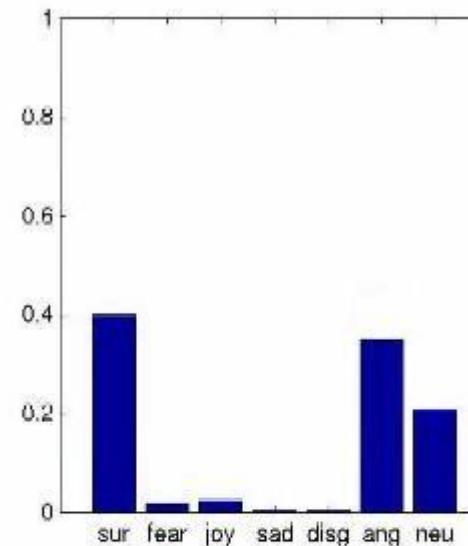
Lower Face Action Units		
AU25	AU26	AU27
		
Lips are relaxed and parted	Lips are relaxed and parted; mandible is lowered	Mouth is stretched open and the mandible pulled down
AU12	AU12+25	AU20+25
		
Lip corners are pulled obliquely	AU12 with mouth opening	Lips are parted and pulled back laterally

# Examples: Facial Expression Recognition

- Implicit customer feedback

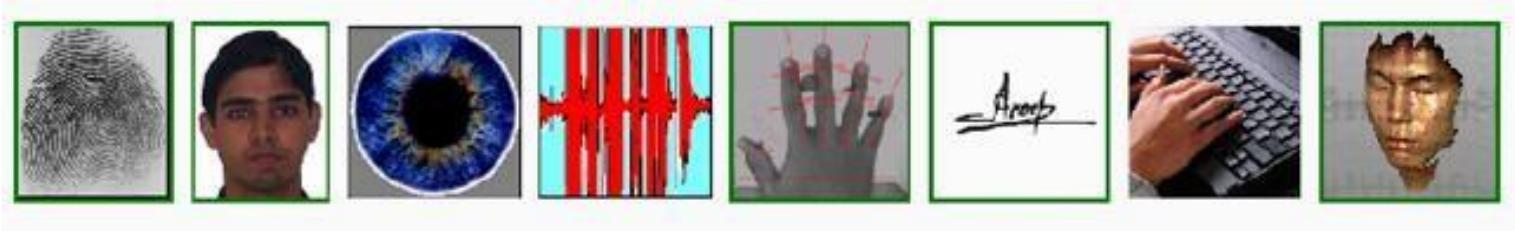


#340



# Examples: Biometrics

- ◆ Biometrics - Authentication techniques
- ◆ Physiological Biometrics
  - Face, IRIS, DNA, Finger Prints
- ◆ Behavioral Biometrics
  - Typing Rhythm, Handwriting, Gait



# Face Recognition

Training examples of a person

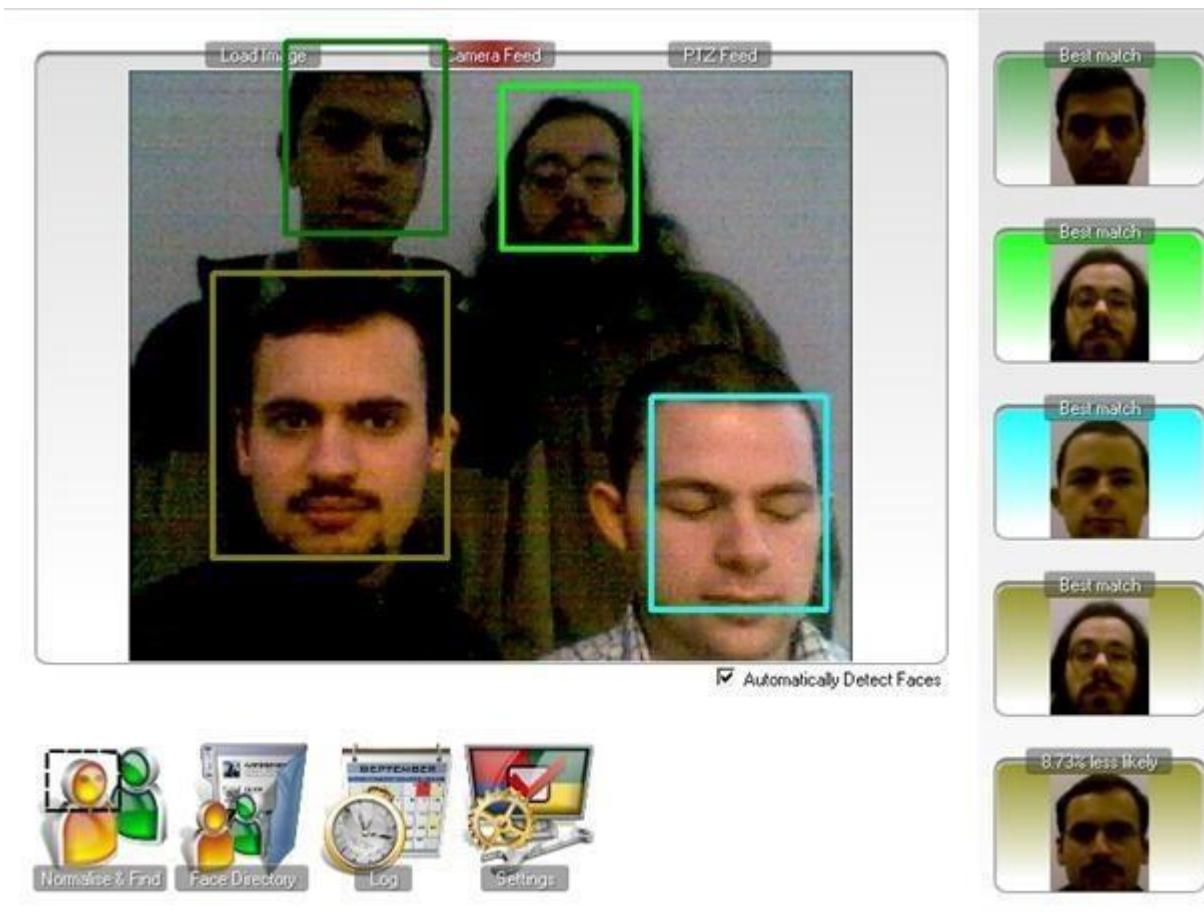


Test images



AT&T Laboratories, Cambridge UK  
<http://www.uk.research.att.com/facedatabase.html>

# Examples: Biometrics – Face Recognition



# Faces and Digital Cameras



Setting camera focus via face detection

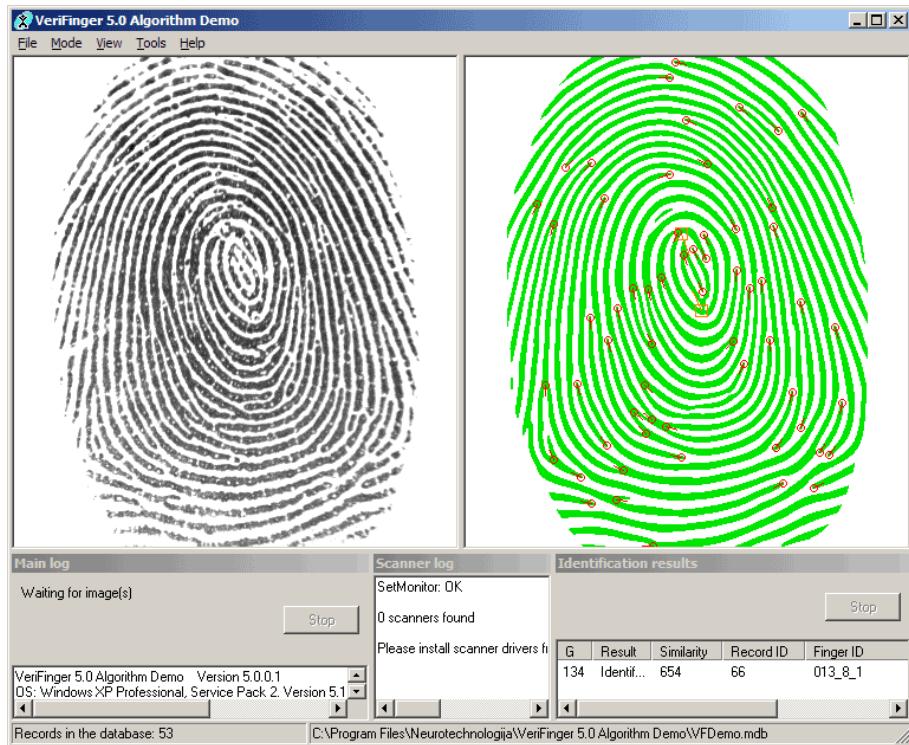


Camera waits for everyone to smile to take a photo [Canon]

Automatic lighting correction based on face detection



# Examples: Biometrics – Finger Print Recognition



# Examples: Biometrics – Signature Verification



# Examples: Robotics



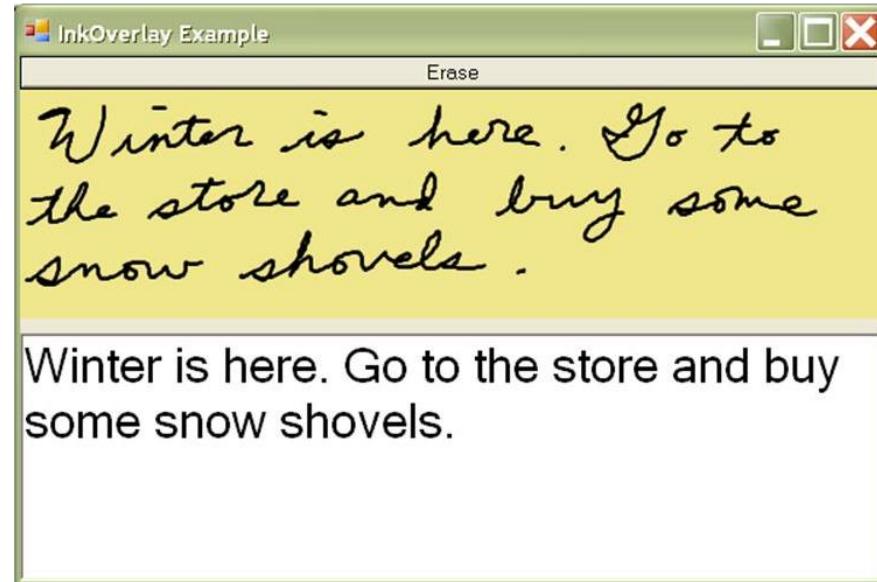
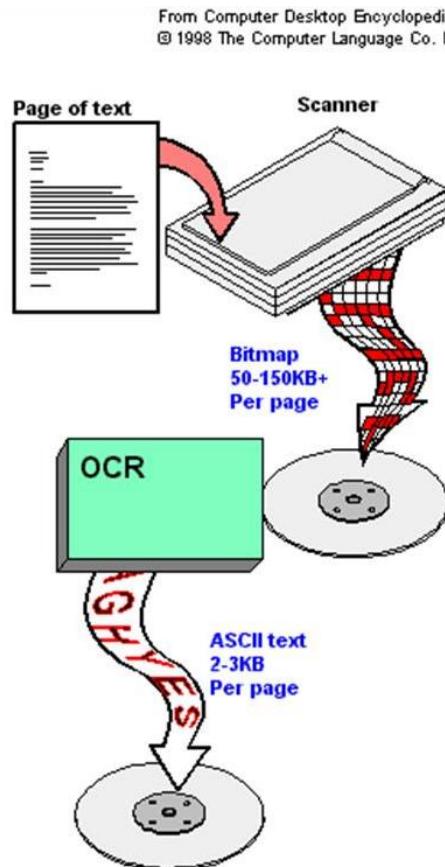
# Examples: Robotics

- ◆ AIBO



# Examples: Optical Character Recognition

- ◆ Convert document image into text



# Examples: Optical Character Recognition

## ◆ Indexing and Retrieval

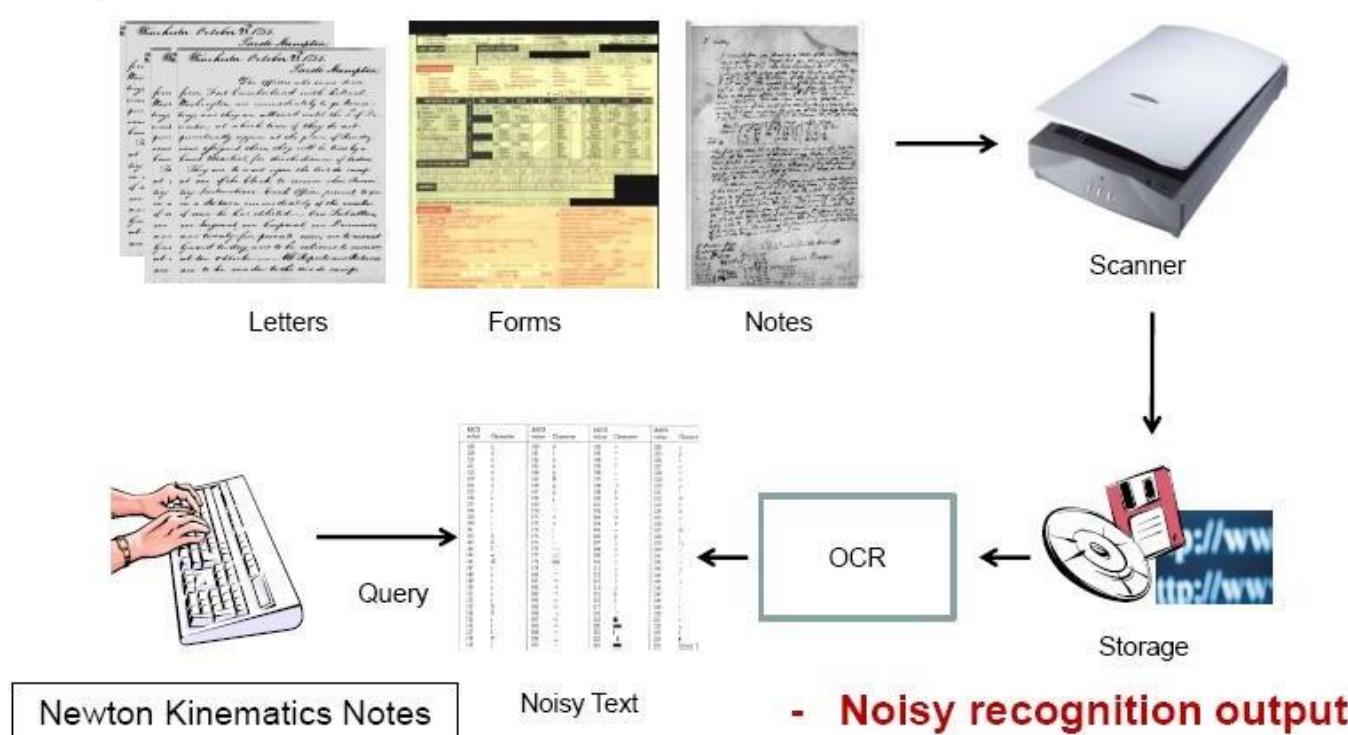


Image Source: CEDAR

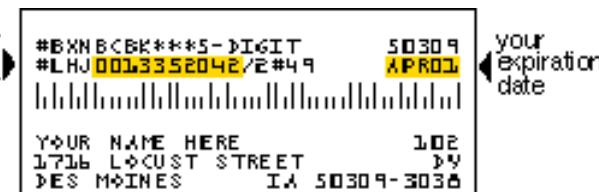
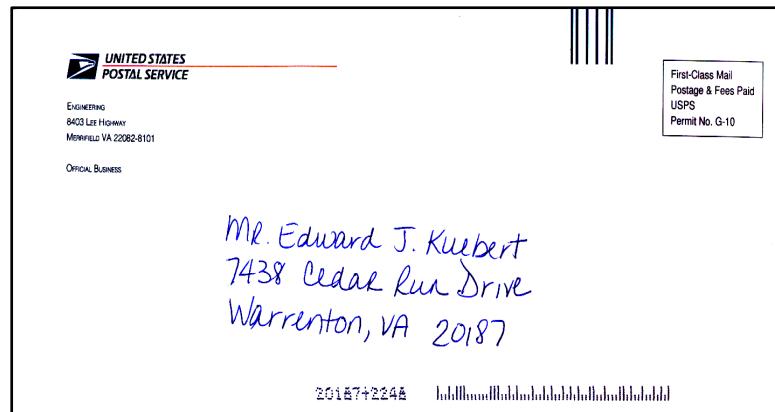
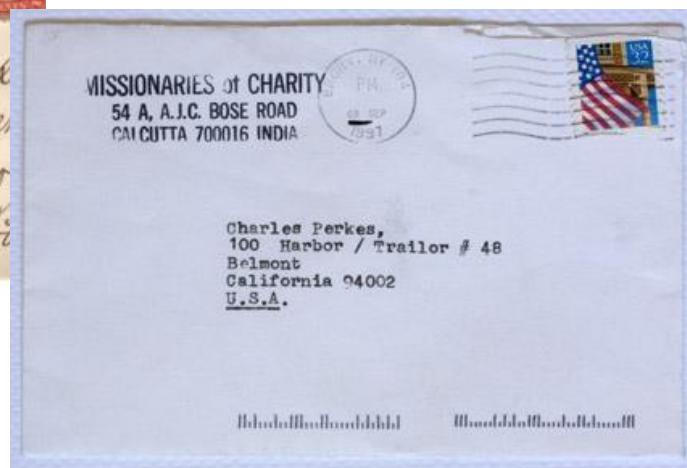
# Examples: Optical Character Recognition

- ◆ License Plate Recognition



# Examples: Optical Character Recognition

- Automatic Mail Sorting



# Example: Hand-written digits

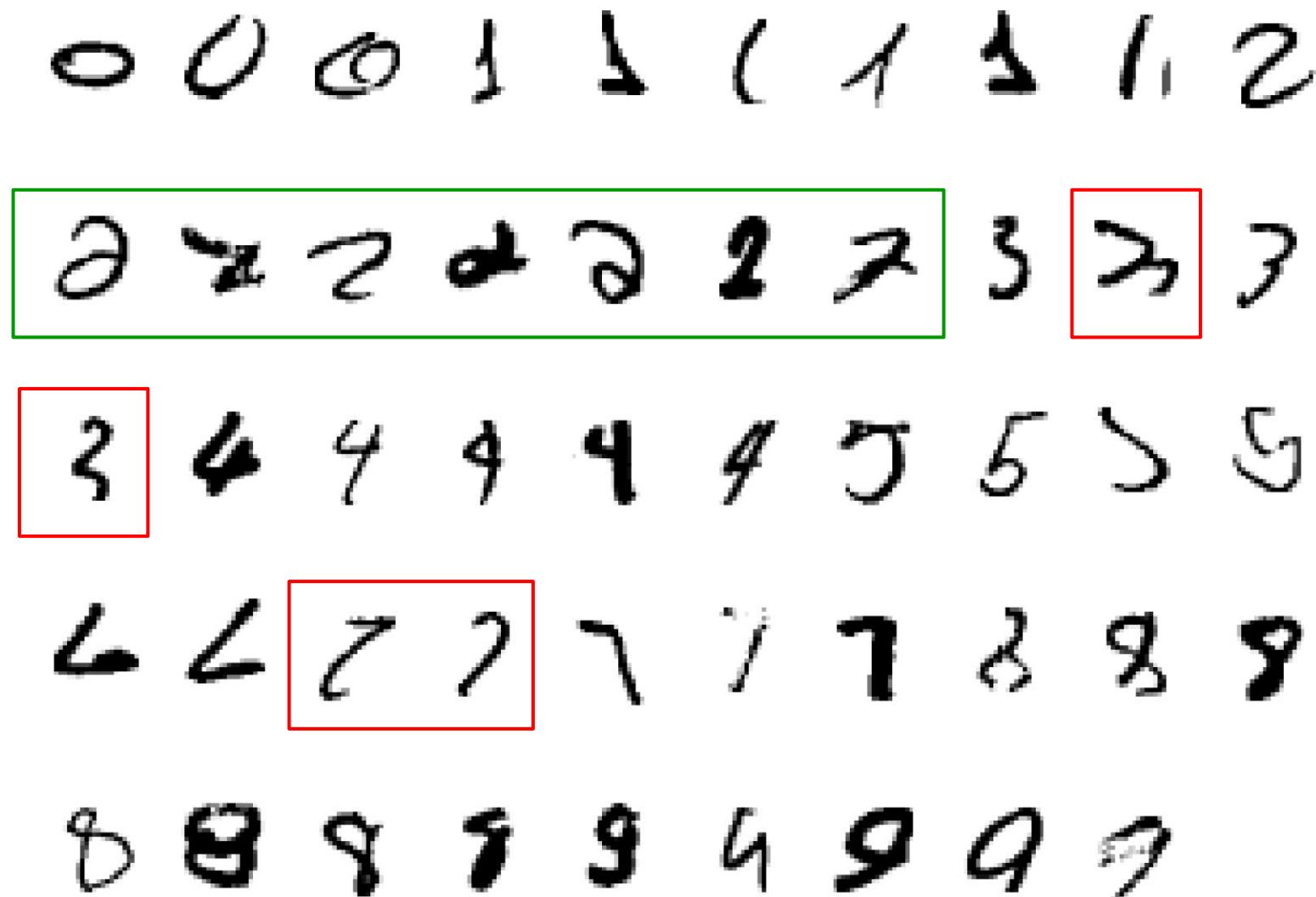
Data representation: Greyscale images

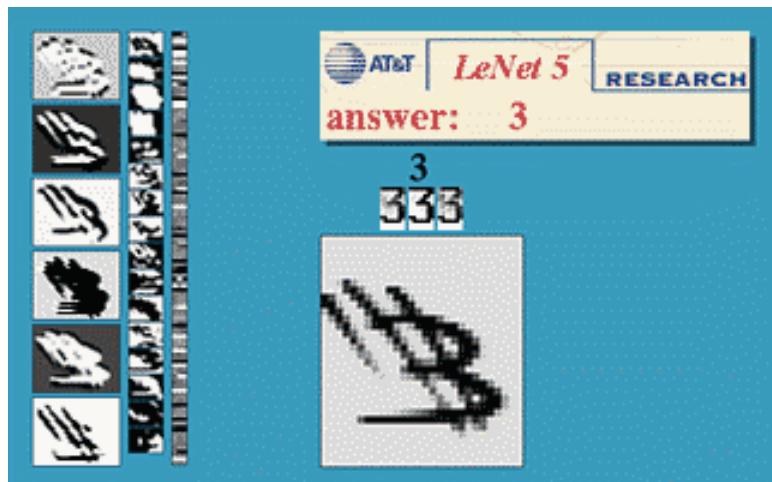
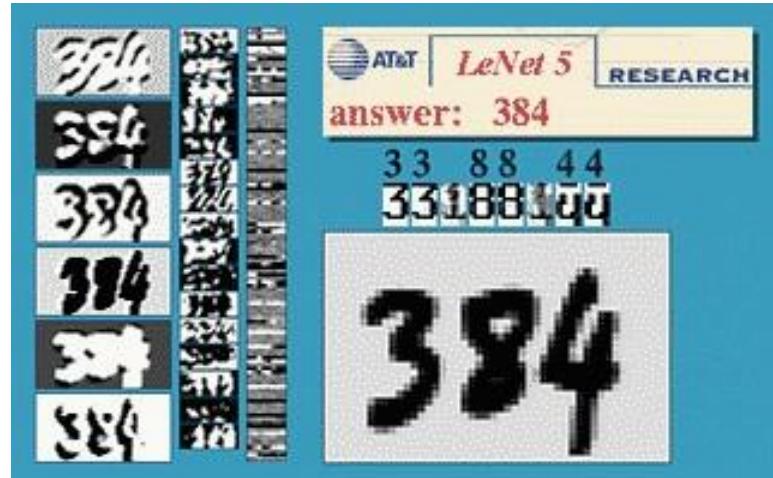
Task: Classification (0,1,2,3.....9)

Problem features:

- Highly variable inputs from same class including some “weird” inputs,
- imperfect human classification,
- high cost associated with errors so “don’t know” may be useful.

A classic example of a task that requires machine learning:  
It is very hard to say what makes a 2





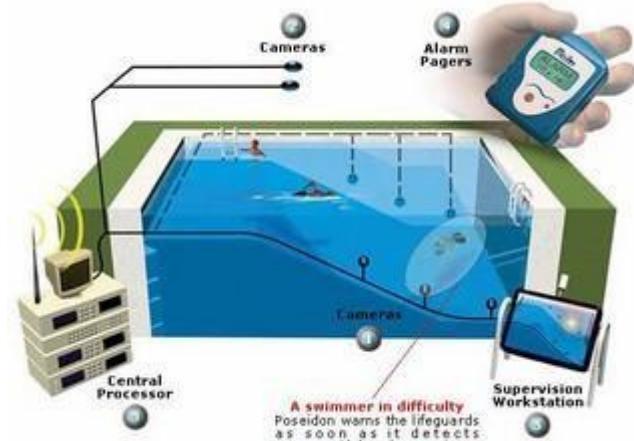
# Safety and Security



Autonomous robots



Driver assistance



Monitoring pools  
(Poseidon)



Pedestrian detection  
[MERL, Viola et al.]



Surveillance

# Summary of Applications

<b>Problem Domain</b>	<b>Application</b>	<b>Input Pattern</b>	<b>Output Class</b>
Document Image Analysis	Optical Character Recognition	Document Image	Characters/words
Document Classification	Internet search	Text Document	Semantic categories
Document Classification	Junk mail filtering	Email	Junk/Non-Junk
Multimedia retrieval	Internet search	Video clip	Video genres
Speech Recognition	Telephone directory assistance	Speech waveform	Spoken words
Natural Language Processing	Information extraction	Sentence	Parts of Speech
Biometric Recognition	Personal identification	Face, finger print, Iris	Authorized users for access control
Medical	Computer aided diagnosis	Microscopic Image	Healthy/cancerous cell
Military	Automatic target recognition	Infrared image	Target type
Industrial automation	Fruit sorting	Images taken on conveyor belt	Grade of quality
Bioinformatics	Sequence analysis	DNA sequence	Known types of genes

# Resources: Journals

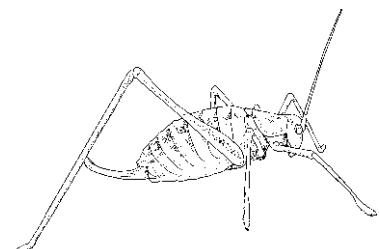
- Journal of Machine Learning Research  
[www.jmlr.org](http://www.jmlr.org)
- Machine Learning
- IEEE Transactions on Neural Networks
- IEEE Transactions on Pattern Analysis and Machine Intelligence
- Annals of Statistics
- Journal of the American Statistical Association
- ...

# Resources: Conferences

- International Conference on Machine Learning (ICML)
- European Conference on Machine Learning (ECML)
- Neural Information Processing Systems (NIPS)
- Computational Learning
- International Joint Conference on Artificial Intelligence (IJCAI)
- ACM SIGKDD Conference on Knowledge Discovery and Data Mining (KDD)
- IEEE Int. Conf. on Data Mining (ICDM)

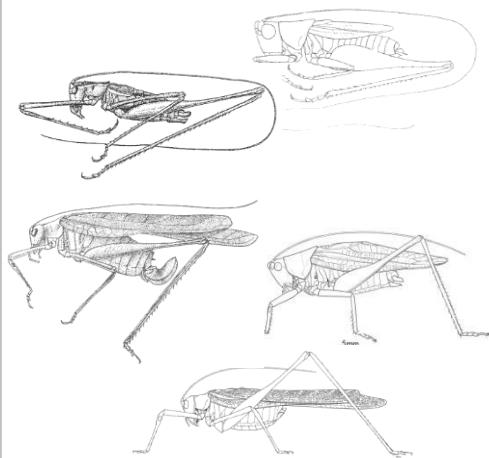
# A Classification Problem Example

Given a collection of annotated data. In this case 5 instances of **Katydid** and five of **Grasshoppers**, decide what type of insect the unlabeled example is.

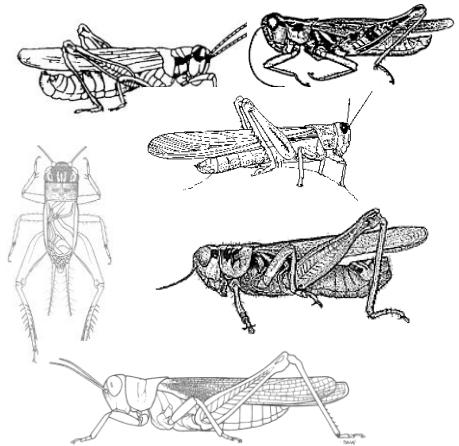


**Katydid or Grasshopper?**

## Katydid



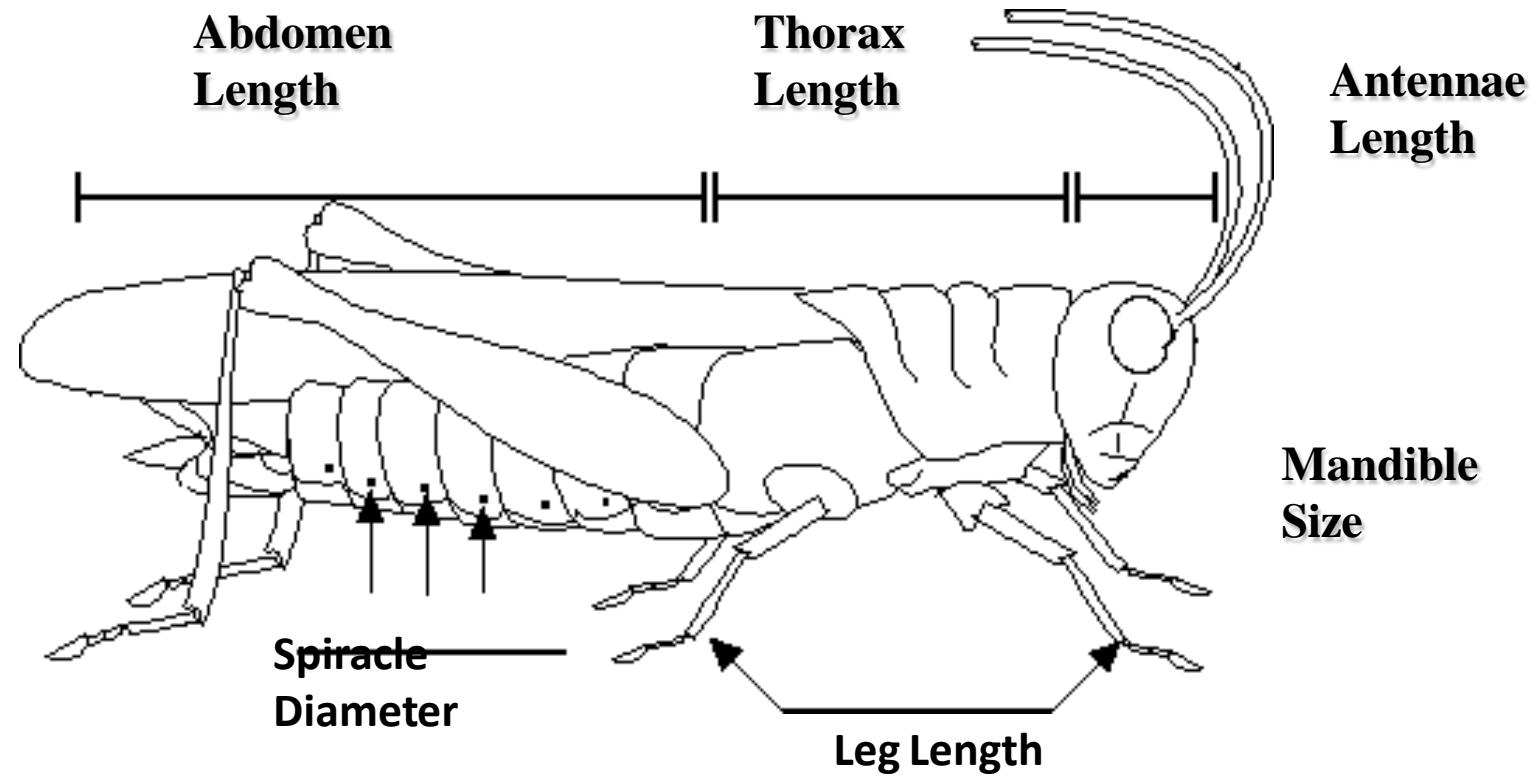
## Grasshoppers



For any domain of interest, we can measure *features*

Color {Green, Brown, Gray, Other}

Has Wings?



We can store features in a database.

The classification problem can now be expressed as:

- Given a training database (**My\_Collection**), predict the **class** label of a previously unseen instance

**My\_Collection**

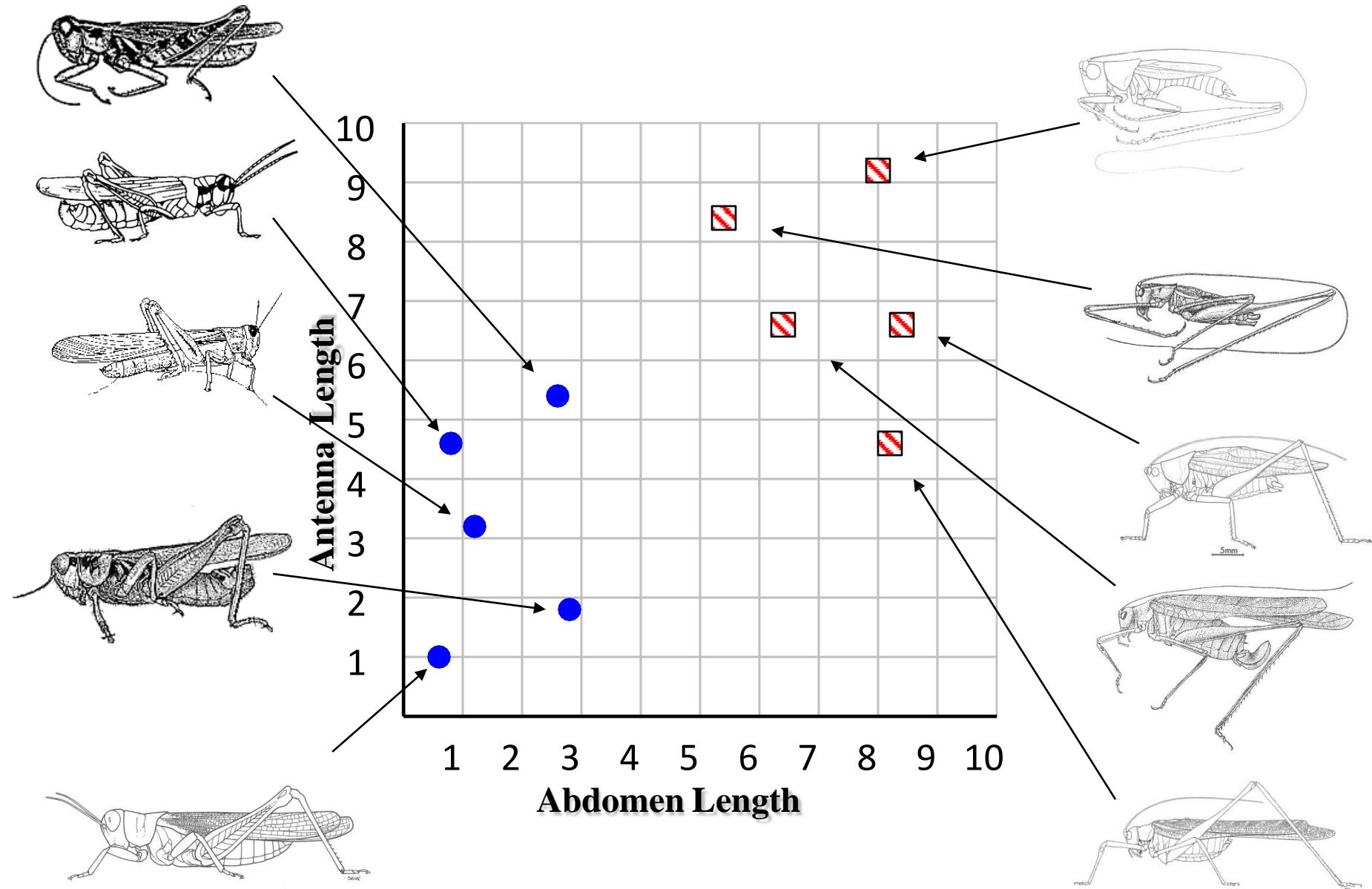
Insect ID	Abdomen Length	Antennae Length	Insect Class
1	2.7	5.5	Grasshopper
2	8.0	9.1	Katydid
3	0.9	4.7	Grasshopper
4	1.1	3.1	Grasshopper
5	5.4	8.5	Katydid
6	2.9	1.9	Grasshopper
7	6.1	6.6	Katydid
8	0.5	1.0	Grasshopper
9	8.3	6.6	Katydid
10	8.1	4.7	Katydids

previously unseen instance =

11	5.1	7.0	???????
----	-----	-----	---------

# Grasshoppers

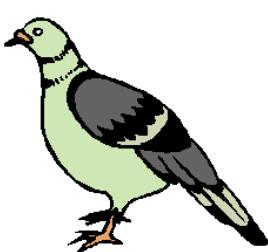
# Katydid





We will return to the previous slide in two minutes. In the meantime, we are going to play a quick game.

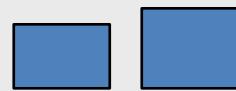
I am going to show you some classification problems which were shown to pigeons!



Let us see if you are as smart as a pigeon!

# Pigeon Problem 1

Examples of class A



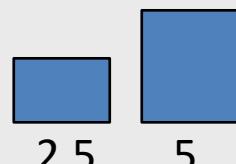
3      4



1.5      5

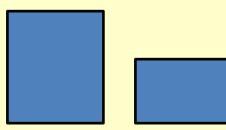


6      8

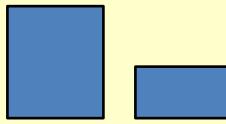


2.5      5

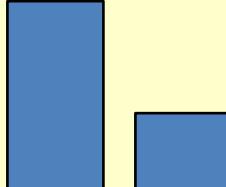
Examples of class B



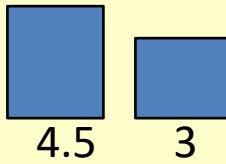
5      2.5



5      2



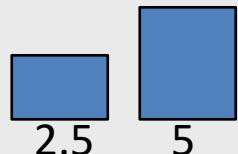
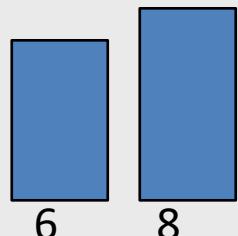
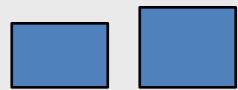
8      3



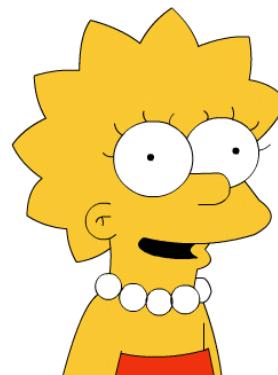
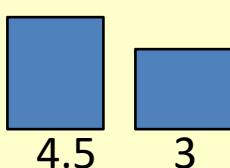
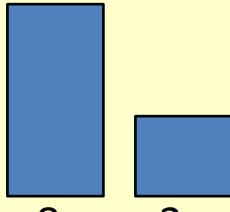
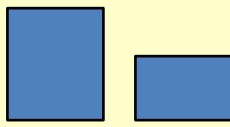
4.5      3

# Pigeon Problem 1

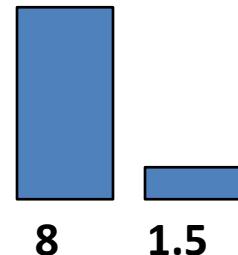
Examples of class A



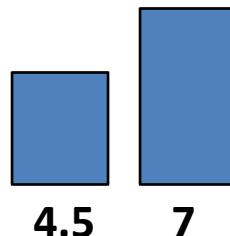
Examples of class  
B



What class is this object?

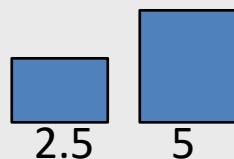
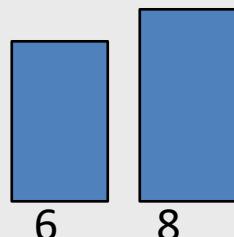
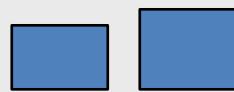


What about this one,  
A or B?

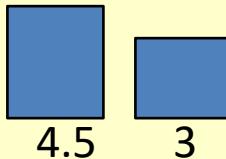
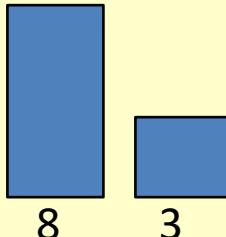
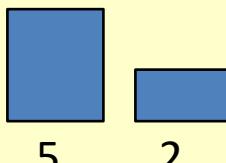
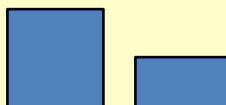


# Pigeon Problem 1

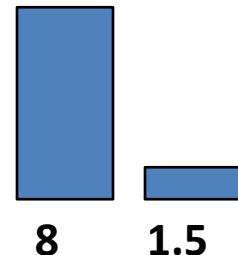
Examples of class A



Examples of class B



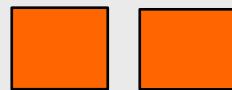
This is a B!



Here is the rule.  
If the left bar is  
smaller than the  
right bar, it is an A,  
otherwise it is a B.

# Pigeon Problem 2

Examples of class A



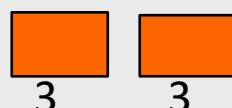
4      4



5      5

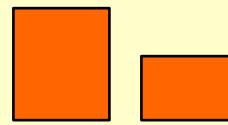


6      6

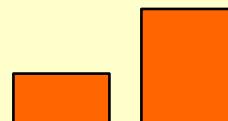


3      3

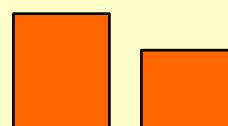
Examples of class B



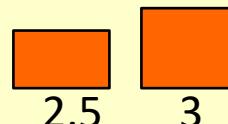
5      2.5



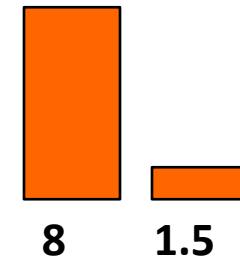
2      5



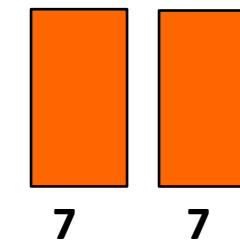
5      3



2.5      3



8      1.5



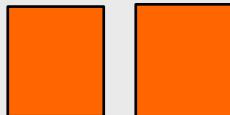
7      7

# Pigeon Problem 2

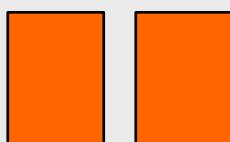
Examples of class A



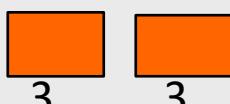
4      4



5      5

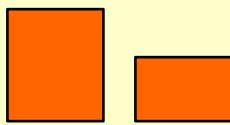


6      6

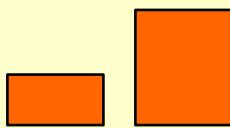


3      3

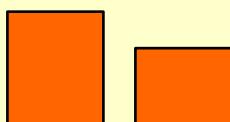
Examples of class  
B



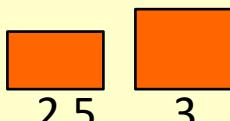
5      2.5



2      5



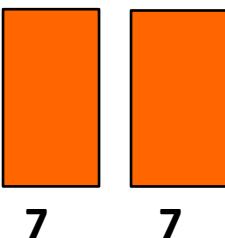
5      3



2.5      3

The rule is as follows, if the two bars are equal sizes, it is an **A**. Otherwise it is a **B**.

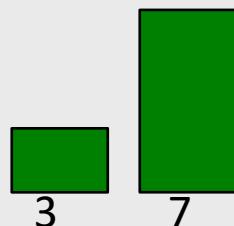
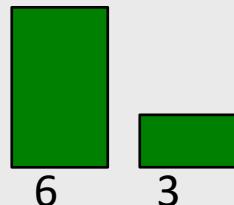
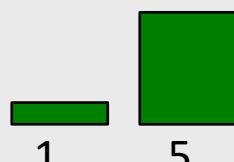
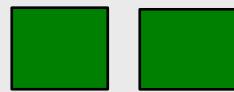
So this one is an **A**.



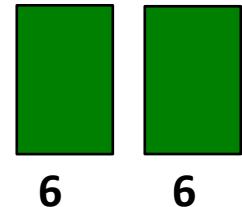
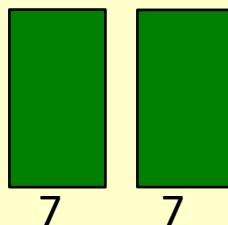
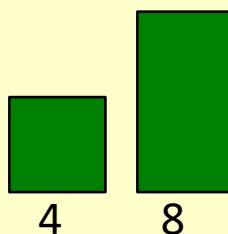
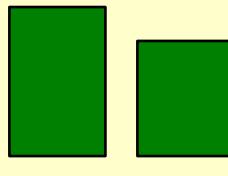
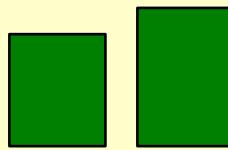
7      7

# Pigeon Problem 3

Examples of class A



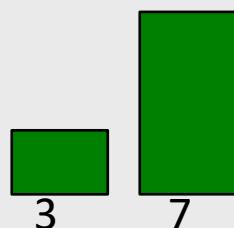
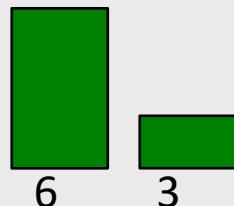
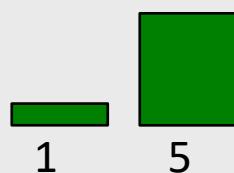
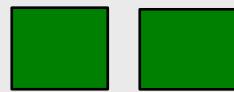
Examples of class B



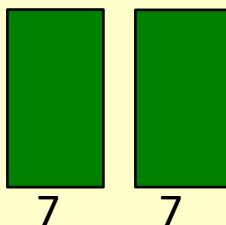
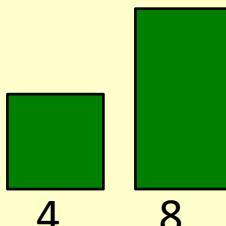
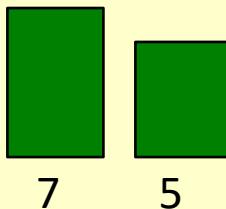
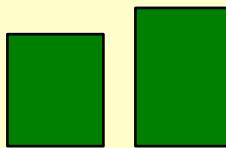
This one is really hard!  
What is this, A or B?

# Pigeon Problem 3

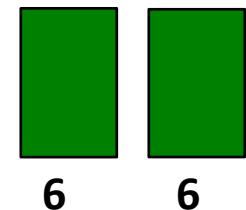
Examples of class A



Examples of class B



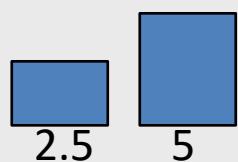
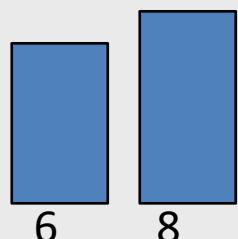
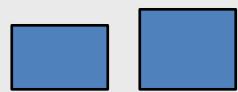
It is a B!



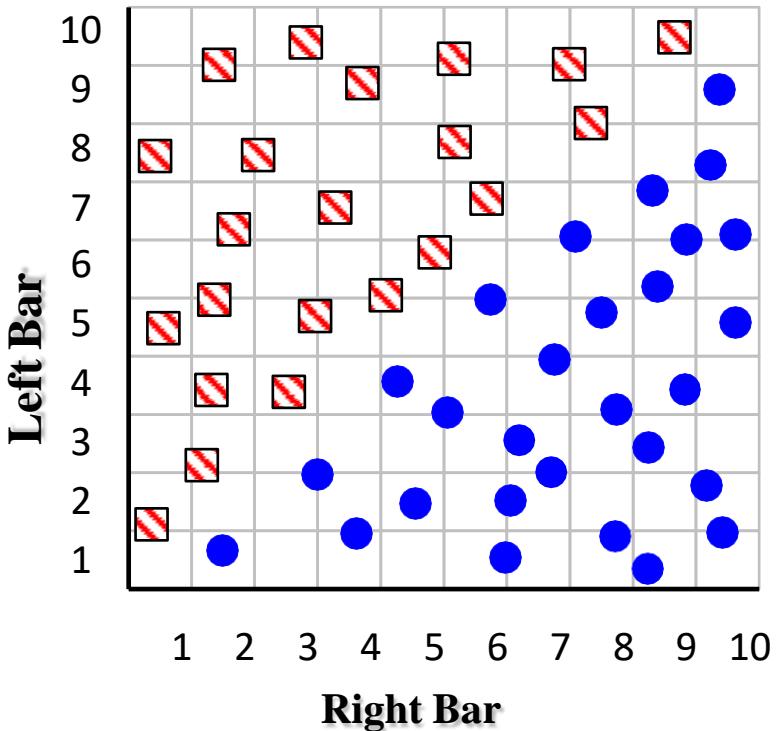
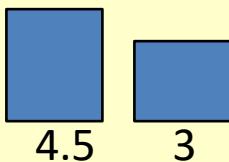
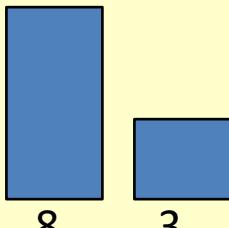
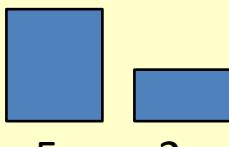
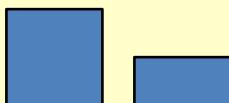
The rule is as follows, if the sum of the two bars is less than or equal to 10, it is an A. Otherwise it is a B.

# Pigeon Problem 1

Examples of class A



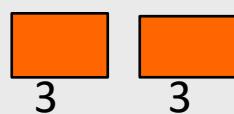
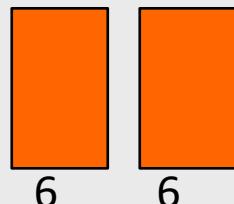
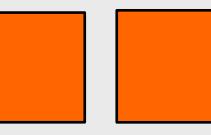
Examples of class B



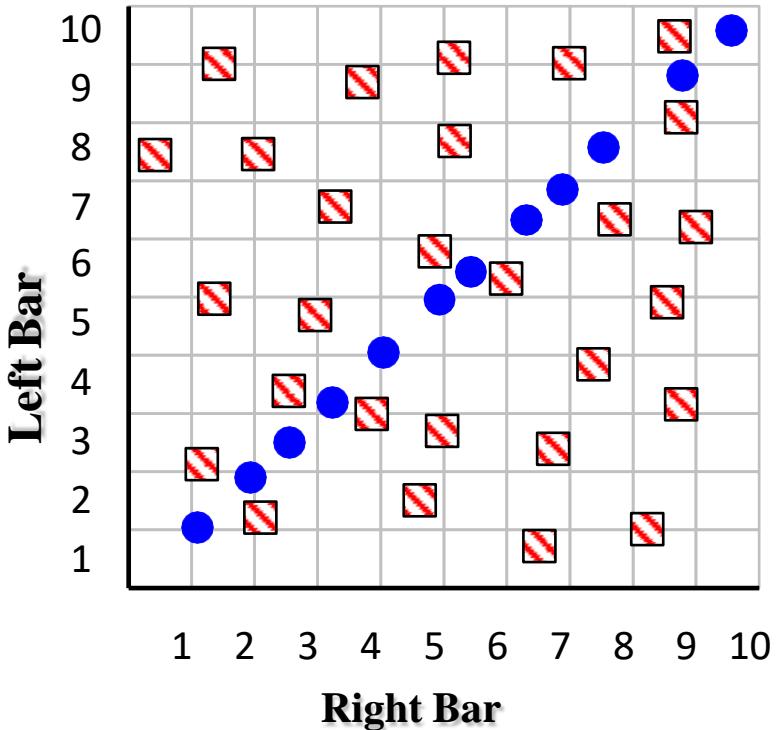
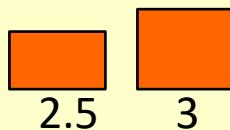
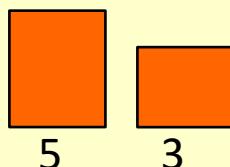
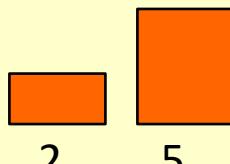
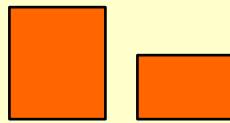
Here is the rule again.  
If the left bar is smaller than the right bar, it is  
an **A**, otherwise it is a **B**.

# Pigeon Problem 2

Examples of class A



Examples of class B

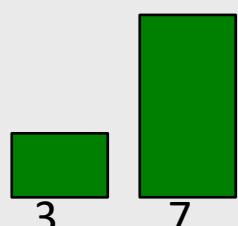
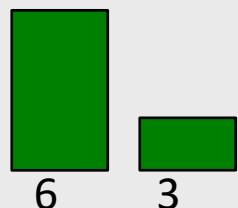
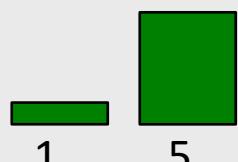


Let me look it up... here it is..  
the rule is, if the two bars  
are equal sizes, it is an A.  
Otherwise it is a B.

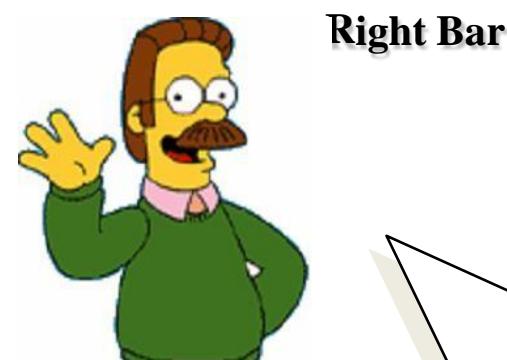
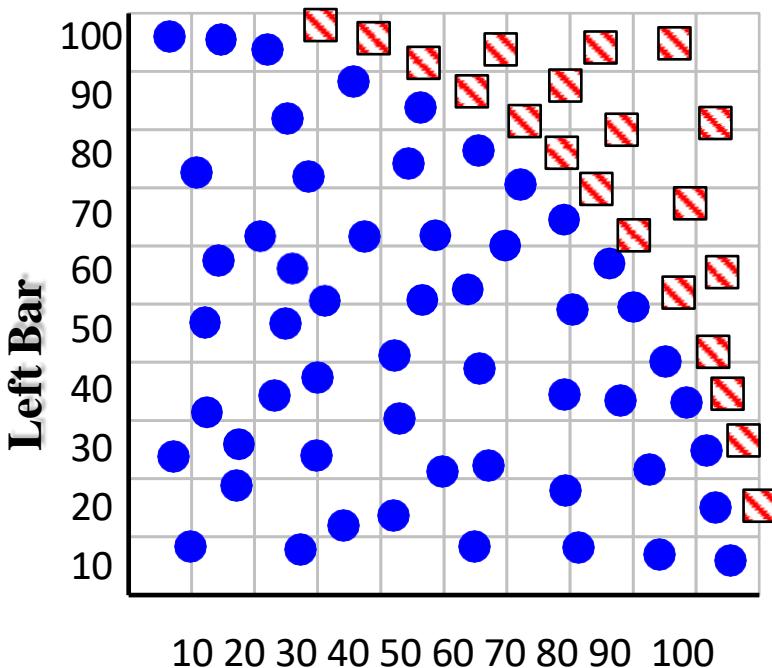
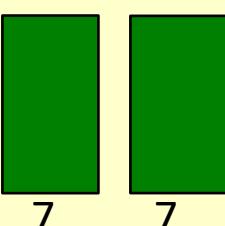
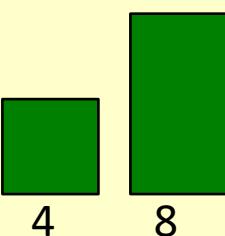
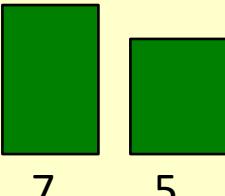
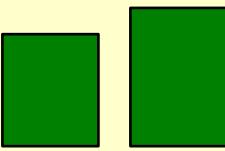


# Pigeon Problem 3

Examples of class A



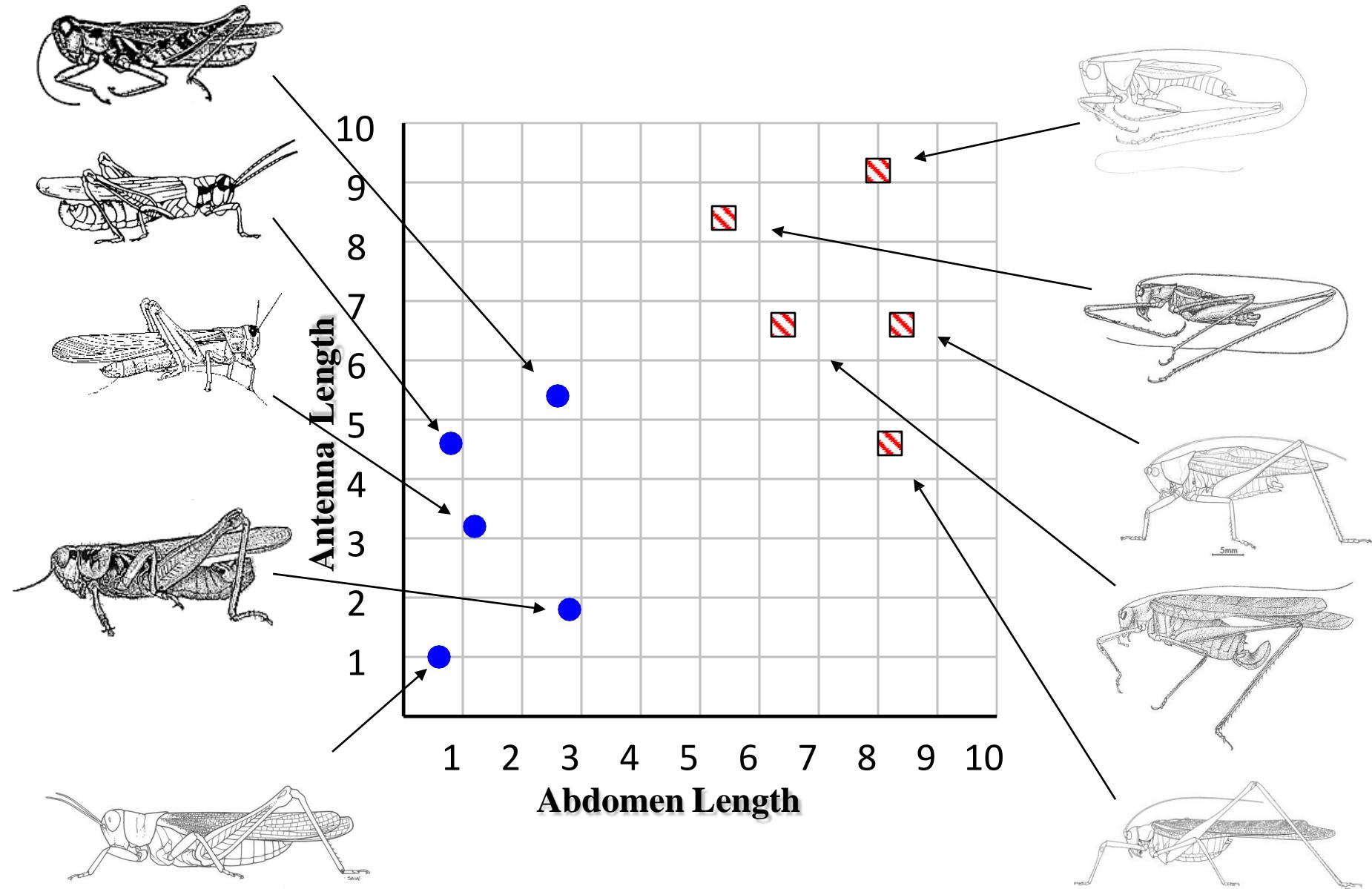
Examples of class B



The rule again:  
if the square of the sum of the  
two bars is less than or equal  
to 100, it is an A. Otherwise it  
is a B.

# Grasshoppers

# Katydid



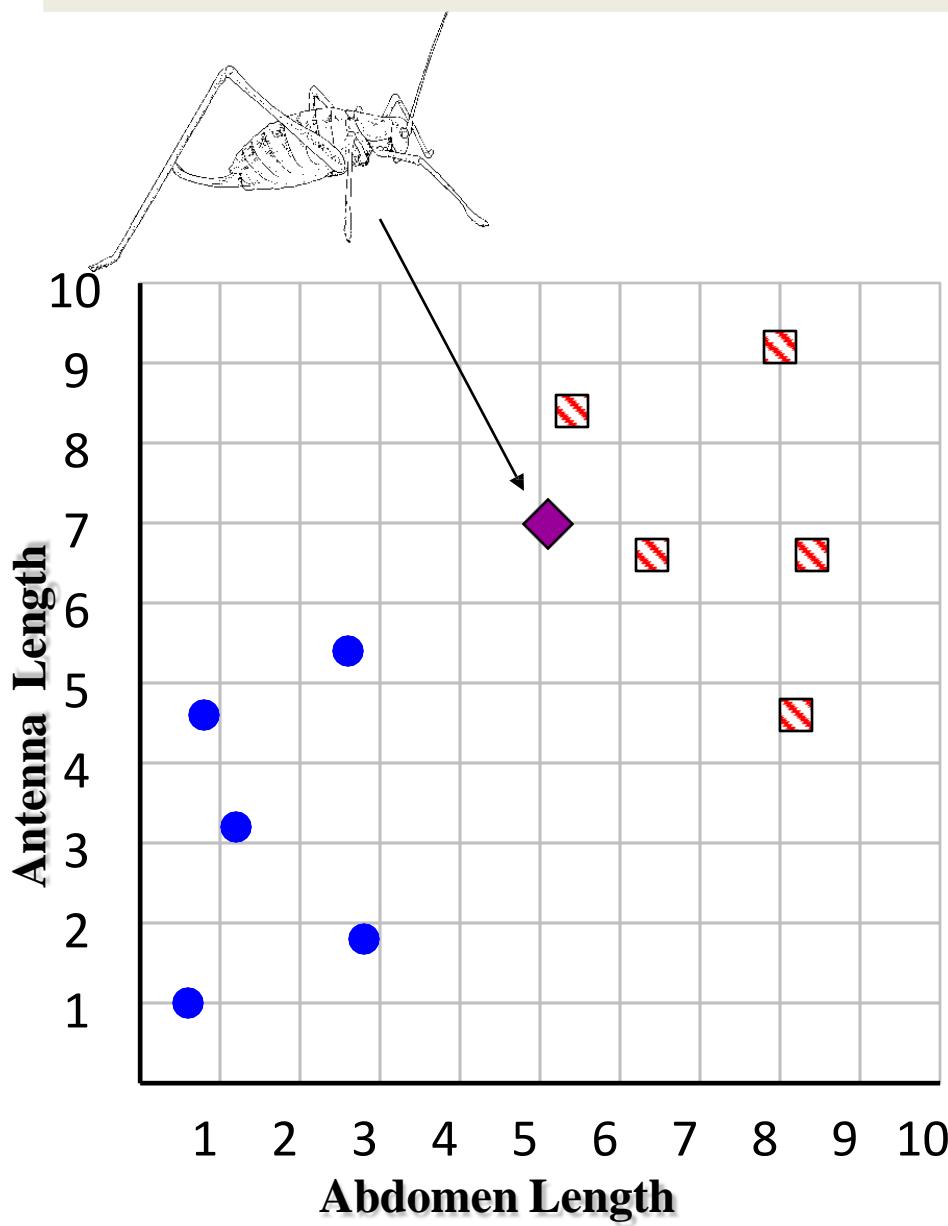
previously unseen instance =

11

5.1

7.0

???????



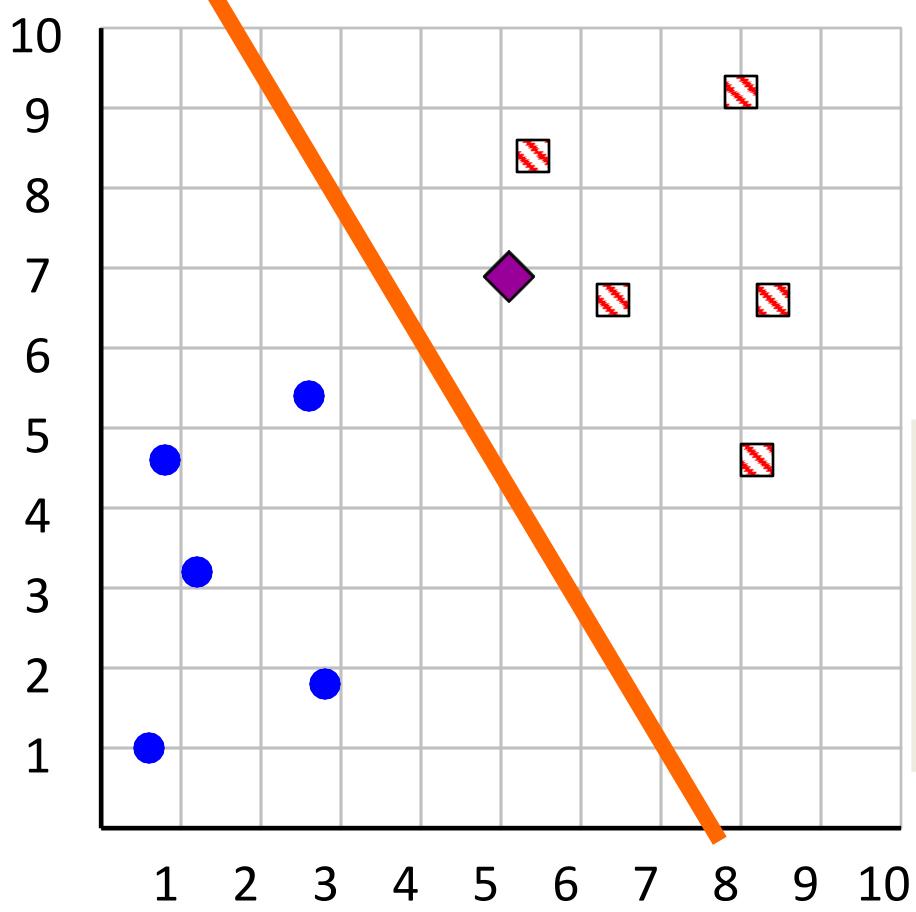
- We can “project” the previously unseen instance into the same space as the database.
- We have now abstracted away the details of our particular problem. It will be much easier to talk about points in space.

▣ **Katydid**  
● **Grasshoppers**

# Simple Linear Classifier



R.A. Fisher  
1890-1962



If previously unseen instance above the line  
then

class is **Katydid**

else

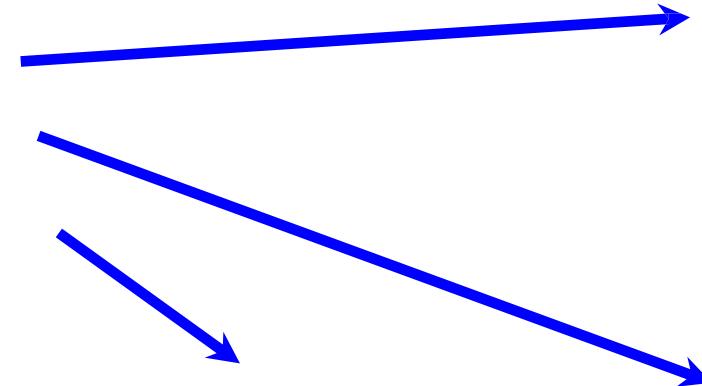
class is **Grasshopper**

■ **Katydid**

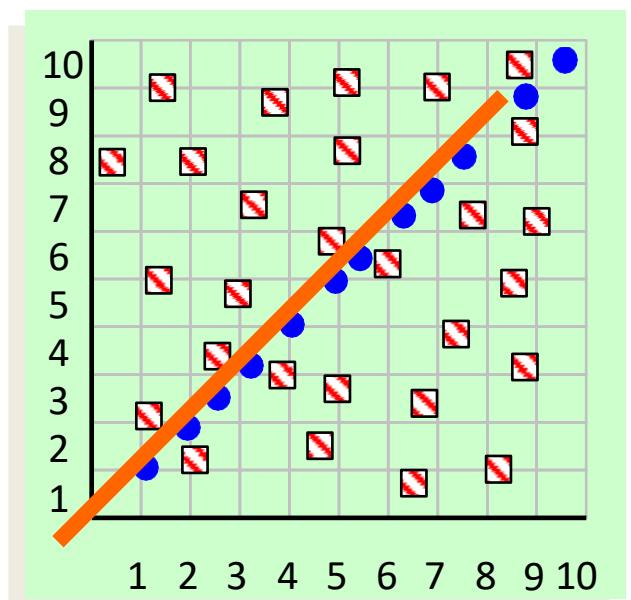
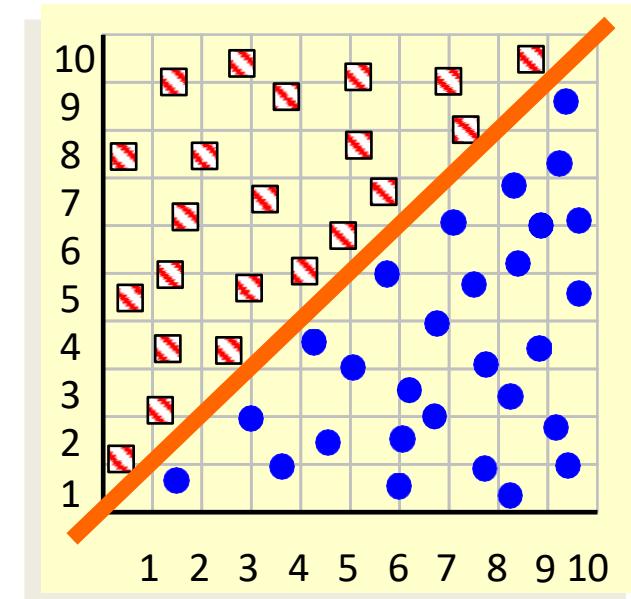
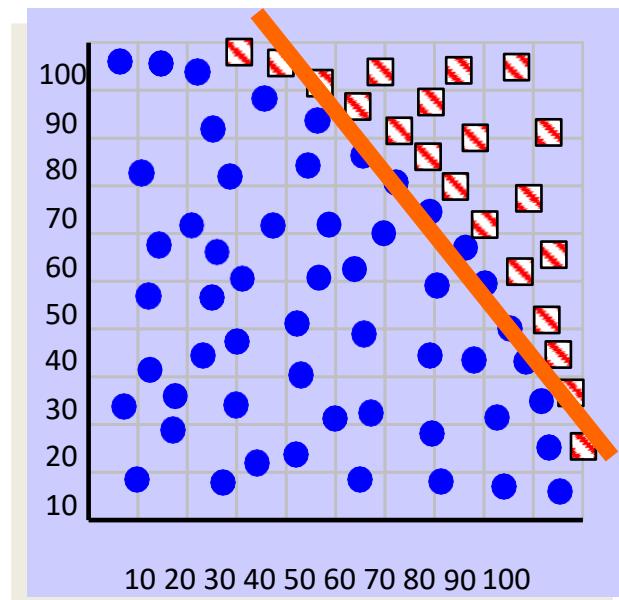
● **Grasshopper**

Which of the “Pigeon Problems” can be solved by the Simple Linear Classifier?

- 1) Perfect
- 2) Useless
- 3) Pretty Good



Problems that can be solved by a linear classifier are called **linearly separable**.



# Acknowledgements

- ◆ Introduction to Machine Learning, Alpaydin
- ◆ Statistical Pattern Recognition: A Review – A.K Jain et al., PAMI (22) 2000
- ◆ Pattern Recognition and Analysis Course – A.K. Jain, MSU
- ◆ *Pattern Classification*" by Duda et al., John Wiley & Sons.