DISCOVERING PATTERNS

Module 6.1 – Introduction

Dr Rita Chakravarti
Institute of Systems Science
National University of Singapore
Email: rita@nus.edu.sg

Pattern Recognition

What is Pattern Recognition?

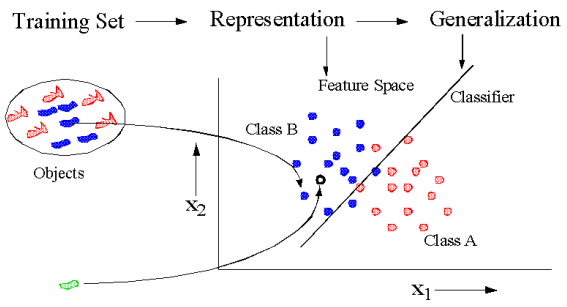
- **Statistical pattern recognition** relates to the use of statistical techniques for analysing data measurements in order to extract information and make justified decisions. It is a very active area of study and research, which has seen many advances in recent years. Applications such as data mining, web searching, multimedia data retrieval, face recognition, and cursive handwriting recognition, all require robust and efficient pattern recognition techniques.
- Wikipedia definition is as follows**: Pattern recognition** is a branch of machine learning that focuses on the recognition of patterns and regularities in data, although it is in some cases considered to be nearly synonymous with machine learning. Pattern recognition systems are in many cases trained from labeled "training" data (supervised learning), but when no labeled data are available other algorithms can be used to discover previously unknown patterns (unsupervised learning).

To summarize

It is a technology useful for automatic detection of shapes, forms and classification of patterns in data

A scientific discipline whose goal is the classification of objects into a number of categories or classes

Examples of Patterns



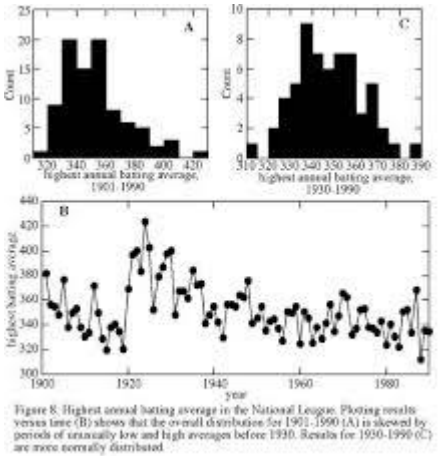
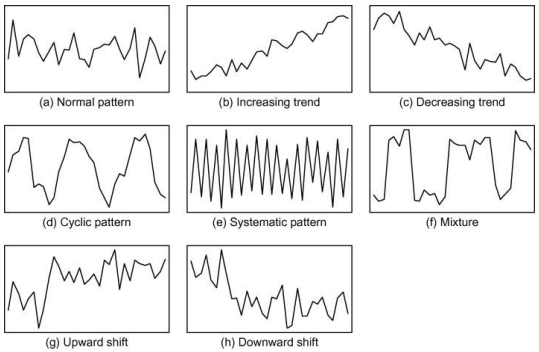
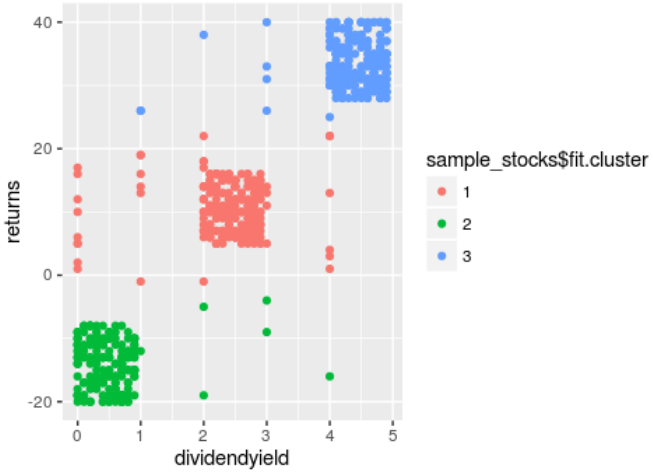
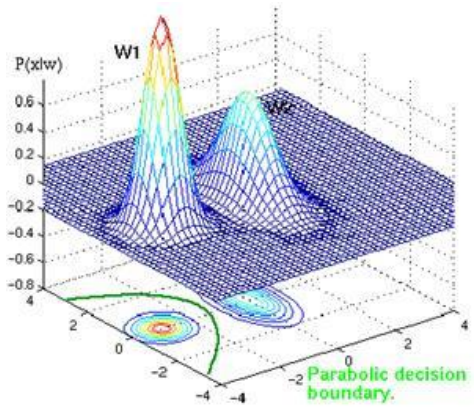


Figure 8. Highest annual batting average in the National League. Plotting results versus time (B) shows that the overall distribution for 1901-1990 (A) is skewed by periods of unusually low and high averages before 1930. Results for 1930-1990 (C) are more normally distributed.







Topics Covered in this Genre

- Five topics covered are:

  - **Topic 1:**
    - ✓ Introduction to Pattern Recognition or Discovering Patterns
    - ✓ Factor Analysis (Principle Component Analysis (PCA)) – Dimension Reduction
  - **Topic 2:**
    - ✓ Cluster Analysis
  - **Topic 3:**
    - ✓ Bayesian Nets
  - **Topic 4:**
    - ✓ Association Mining
  - **Topic 5:**
    - ✓ Recommender Systems

Assessment of Unit 5

Part 1 (Data Mining Methods)	
Exams	: 30%
Assignments	: 20%
Part 2 (Discovering Patterns)	
Exams	: 30%
Assignment 1 (PCA/Factor/Cluster Analysis)	: 10 %
Assignment 2 (Bayesian/Recommender Systems/Association Mining)	: 10%
Total	100%

Question & Answer Session