MODULE 1
INTRODUCTION

LESSON 0

Background

Keywords: History of Pattern Recognition, Prerequisites, Notation and Convention, Table of Contents

Preface

Pattern Recognition is an important topic and finds applications in diverse areas such as healthcare, education, agriculture, environment, and transportation. It is closely associated with topics like *machine learning* and *data mining*.

The primary aim of this course is to attract the reader towards pattern recognition and provide a platform for understanding the basics and gaining an insight into several important topics associated with machine recognition. Keeping this in mind, we have concentrated more on working out numerical examples and providing a collection of exercises at the end of each module. In order to appreciate the intricacies further the reader may have to go through the references provided at the end of each module. This material is useful to undergraduate students and teachers interested in pattern recognition, machine learning, and data mining.

We are extremely thankful to *Professor G. Nagaraja*, Dept. of CSE, IIT Bombay, Mumbai and *Professor V. S. Ananthanarayana*, Dept. of IT, NITK Suratkal for providing critical inputs on an earlier version which helped in enhancing the quality of the course material.

M. Narasimha Murty
V. Susheela Devi

Prerequsites

1. **Calculus:** Limits, differentiation, integration, maxima and minima.
2. **Probability:** Independent events, conditional probability, Bayes theorem, and random variables.
3. **Matrices and Linear Algebra:** Determinant, rank of a matrix, inverese, eigenvalues and eigenvectors.
4. **Geometry:** Dot products, distance between a point and a plane.

Most of the material is easily accessible to the readers with a knowledge of basic mathematics. Readers interested in getting exposed to the prerequisites further may refer to books dealing with engineering mathematics. A good reference is the book *Advanced Engineering Mathematics* by Erwin Kreyszig, Wiley Eastern, 2006.

Notation and Conventions

$\mathcal{X}$ *or* $\mathcal{D}$: Set of patterns or documents
X: Pattern Vector
x: Scalar pattern
n: Number of patterns
d: Number of features
f_i: Feature i
C_i: Class i or Cluster i
cl_i: Classifier i
m_i: Mean of class i
$d_i(X, Y)$: Distance between vectors X and Y
w^t: Transpose of the weight vector w
'X', 'O', '+': Class labels
Other symbols are explained in the places where they are used.

Historical Perspective

- One of the early algorithms for pattern recognition is attributed to Fisher [1] in the form a *linear discriminant* in 1936.
- Another early classifier in the form of a simple neural network is the *Perceptron* proposed by Rosenblatt in 1957[2].
- The book by Minsky and Papert [3] gave an excellent formal characterization of parallel distributed computing based on group theory.
- Cover and Hart[4] provided bounds on the performance rates of NNC and KNNC and contributed the Condensed Nearest Neighbour Classification Algorithm.
- Early work on decision trees appeared in 1966; Quinlan[5] is responsible for a variety of classifiers based on decision trees.
- Artficial neural networks based on backpropagation[6] were responsible for increase in the interest in neural networks; the contribution was made in 1986.
- Vapnik[7] is associated with the early work on statistical learning theory, Vapnik-Chervonenkis (VC) dimension, and support vector machines (SVMs).
- Freund[8] is an early researcher on boosting and a proponent of AdaBoost.
- The book by Xu and Wunsch II[9] is a good reference on clustering.
- Han and Kamber's book on data mining[10], Bishop's book[11] on machine learning, and Duda, Hart and Stork's book[12] on pattern recognition are of contemporary nature and are useful references.

References

1. R. A. Fisher, The use of Multiple Measurements in Taxonomic Problems. *Annals of Eugenics*, Vol. 7, pp. 111-132, 1936.

2. F. Rosenblatt, *Principles of Neurodynamics: Perceptrons and the theory of Brain Mechanisms*, Spartan Books, Washington DC, 1962.

3. M. Minsky and S. Papert, *Perceptrons: An Introduction to Computational Geometry*, MIT Press, 1969.

4. T. M. Cover and P. E. Hart, Nearest Neighbor Pattern Classification, *IEEE Transactions on Informnation Theory*, Vol. 13, pp. 21-27, 1967.

5. J. Ross Quinlan, *C4.5: Programs for Machine Learning*, Morgan Kauffmann, San Francisco, 1993.

6. D. E. Rumelhart and J. L. McCleland, Editors, *Parallel Distributed Processing*, MIT Press, 1986.

7. V.N. Vapnik, *The Nature of Statistical Learning Theory*, Springer-Verlag, New York, 1995.

8. Y. Freund, Boosting a Weak Learning Algorithm by Majority, *Information and Computation*, vol. 121, pp.256-285, 1995.

9. R. Xu and D. C. Wunsch, II, *Clustering*, IEEE Press, 2009.

10. J. Han and M. Kamber, *Data Mining: Concepts and Techniques*, Morgan Kauffmann, 2008.

11. C. M. Bishop, *Pattern Recognition and Machine Learning*, Springer, 2006.

12. R. O. Duda, P. E. Hart, and D. G. Stork, *Pattern Classification*, John Wiley & Sons, 2004.

Table of Contents

- Module09: Different Approaches to Prototype Selection
  - Lesson15: Minimal Distance Classifier
  - Lesson16: Condensed Nearest Neighbour Classifier
  - Lesson17: Modified Condensed Nearest Neighbour Classifier
- Module10: Bayes Classifier
  - Lesson18: Bayes Classifier
  - Lesson19: Naive Bayes Classifier
  - Lesson20: Bayesian Belief Networks
- Module11: Decision Trees
  - Lesson21: Introduction to Decision Trees
  - Lesson22: Construction of Decision Trees
  - Lesson23: Axis-Parallel and Oblique Decision Trees
  - Lesson24: Learning Decision Trees
- Module12: Linear Discriminnat Functions
  - Lesson25: Introduction to Discriminant Functions
  - Lesson26: Characterization of the Decision Boundary
  - Lesson27: Learning the Discriminant Function
- Module13: Support Vector Machines
  - Lesson28: Introduction to Support Vector Machines
  - Lesson29: Training Support Vector Machines
- Module14: Clustering
  - Lesson30: What is Clustering?
  - Lesson31: Representation of Patterns and Clusters
  - Lesson32: Clustering Process
  - Lesson33: Clustering Algorithms

- Module15: Clustering Large Datasets
  - Lesson34: Incremental Clustering
  - Lesson35: Divide-and-Conquer Clustering
- Module16: Combination of Classifiers
  - Lesson36: Introduction to Combining Classifiers
  - Lesson37: AdaBoost for Classification
  - Lesson38: Combination of Homogeneous Classifiers
  - Lesson39: Schemes for Combining Classifiers
- Module17: Application - Document Recognition
  - Lesson40: Document Processing
  - Lesson41: Document Classification and Retrieval