# Pattern Classification and Clustering

**Spring 2006**

**Time:** **Room:**


**Instructor:** Yingen Xiong
**Office:** 621 McBryde
**Office Hours:**
**Phone:** 231-4212
**Email:** yxiong@cs.vt.edu
**URL:** http://www.cs.vt.edu/~yxiong/PCC/


**Detailed Description:** The course introduces to classical and modern computational approaches to pattern classification and clustering. Topics covered include some or all of the following: the probability and statistical basis for pattern classification and clustering, Bayesian classification decision theory, density and parameter estimation, dimensionality reduction, nonparametric estimation and classification, linear discriminant functions, feature extraction, parametric and nonparametric clustering algorithms, principal component analysis, and classification using artificial neural networks. Emphasis will be on the applications to digital video and speech analysis and classification, target tracking

**Course Objectives:**

- Introduce the basic mathematical and statistical techniques commonly used in pattern classification and clustering
- Provide the students with a variety of pattern classification and clustering algorithms and methods which they can apply to real-world problems.

**Prerequisites:**

- Basic knowledge of Linear Algebra, Probability and Statistics
- Some knowledge of signal/image/video/speech processing.
- Experience with MATLAB and C++ Programming is desirable.


**Textbook**:

- R.O. Duda, P.E. Hart, and D.G. Stork, Pattern Classification, 2nd Edition, John Wiley and Sons, New York, 2001 (ISBN 0-471-05669-3).

**References**:

- C. M. Bishop, Neural Networks for Pattern Recognition, Oxford University Press, 1995.
- K. Fukunaga, Introduction to Statistical Pattern Recognition, 2nd ed., Academic Pr, 1990.
- A.R. Webb, Statistical Pattern Recognition, 2nd Edition, John Wiley and Sons, New York, 2002 .
- R. J. Shalkoff, Pattern Recognition: Statistical, Structural, and Neural Approaches, John Wiley and Sons, 1992

- S.M. Kay, Fundamentals of Statistical Signal Processing Estimation Theory, Prentice-Hall, Inc. Englewood Cliffs, NJ, 1993.
- B. Widrow, S.D. Stearns, Adaptive Signal Processing, Englewood Cliffs, N.J. Prentice-Hall, 1985.

**Course Outline:**

- **Introduction to Pattern Classification and Clustering**
  Objective of Pattern Classification, Model of the pattern classification process, linear decision function, minimum-distance classification, approaches to pattern classification and clustering: statistical, neural and structural.

- **Review of Some Basic Knowledge**
  Probability and statistics: probability theory, conditional probability and Bayes rule, Random vectors, expectation, correlation, covariance. Linear algebra, linear transformations

- **MATLAB Tutorial**
  Review of some tools which need to be used to complete programming assignments. Students are highly encouraged to use MATLAB to implement their assignments and projects.

- **Bayesian Classification Decision Theory**
  Bayesian decision rules, Minimum error-rate classification, discriminant functions and decision boundaries, Bayes classifier for Gaussian patterns, linear and quadratic classifiers.

- **Density and Parameter Estimation**
  Maximum-likelihood estimation, Bayesian estimation

- **Dimensionality Reduction**
  The curse of dimensionality, principal component analysis, linear discriminants analysis.

- **Nonparametric Estimation and Classification**
  Parzen windows, K-nearest-neighbor classification, Non-parametric classification, density estimation, Parzen estimation.

- **Linear Discriminant functions**
  Linear discriminant, Perceptron learning, optimization by gradient descent, Support Vector Machine

- **Clustering Algorithms**
  Maximum-likelihood estimation and unsupervised learning, Mixture of Gaussian, K-means algorithm, hierarchical clustering, component analysis.

- **Introduction to Classification Using Artificial Neural Networks**
  Single-layer networks, multilayer neural networks, feedforward operation, backpropagation algorithm, learning curves, neural networks classifiers.

**Grading**: The course grade will be the weighted sum of four grades. Grading will be straight scale (90-100 A, 80-89 B, 70-79 C, 60-69 D, below 60 F).

- **Homework**: There will be 3-5 homework assignments and will require students to implement some of the algorithms covered during the semester and apply them. Homework assignments must be done individually. No collaboration on homework is allowed. Homework assignments will be done in MATLAB

- **Exam**: There will be a midterm exam and a final exam. All tests will be closed-books, closed-notes. The final exam may cover material from the entire course, but will emphasize material not covered on the mid-term.

- **Project**: The term project is due at the end of the semester and accounts for 40% of the course grade. Students will choose their own problem topic. Students will write a short proposal for the purpose of approval and feedback. It can be a comprehensive literature review or the implementation of the algorithms covered during the semester. Students are encouraged to propose projects related to their own research. To facilitate the completion of the project in a semester, it is advised that teams of 2-3 students work together. Students are highly encouraged to use MATLAB to implement their projects. Projects will be graded by their content (75%) and the quality of a classroom presentation (25%) at the end of the semester.

| | |
|---|---|
| Homework | 30% |
| Project | 40% |
| Midterm | 10% |
| Final Exam | 20% |

# Course Schedule

| Week | Date | Topics | Readings | Assignments/activities |
|---|---|---|---|---|
| 1 | | Introduction to Pattern Classification and Clustering: Problem, Model, Decision Function, and Approaches | DHS Ch.1 | |
| | | Review of Statistics and Probability | DHS A.4 | Homework#1 assigned |
| 2 | | Review of Random Vectors, Expectation, Correlation, Covariance | DHS A4, notes | |
| | | Review of Linear Algebra, Linear Transformations | DHS A2 | |
| 3 | | MATLAB Tutorial: Tool Box and Programming | Notes | |
| | | Bayesian Decision Rules, Minimun Error-rate Classification, Discriminant Functions and Decision Boundary | DHS Ch. 2 | |
| 4 | | Bayes Classifiers for Gaussian Pattern, Linear and Quadratic Classifiers | DHS Ch. 2 | Homework#1 due |
| | | Density and Parameter Estimation: Maximum-Likelihood Estimation | DHS Ch.3 | Homework#2 assigned |
| 5 | | Density and Parameter Estimation: Bayesian Estimation | DHS Ch. 3 | |
| | | The Curse of Dimensionality, Fisher Linear Discriminant Analysis | DHS Ch. 3 | |
| 6 | | Principal Component Analysis | DHS Ch. 3 | |
| | | Nonparametric Density Estimation | DHS Ch.4 | |
| 7 | | Parzen Window, K-nearest Neighbor Estimation | DHS Ch. 4 | Homework#2 due |
| | | Nonparametric Classification, Parzen Estimation | DHS Ch. 4 | Homework#3 assigned |
| 8 | | Midterm | | Midterm |
| | | Linear Discriminant, Percepton Learning | DHS Ch. 5 | |
| 9 | | Optimization by Gradient Descent, Support Vector Machine | DHS Ch. 5 | |
| | | Mixture of Gaussian, Maximum-likelihood Estimation and Unsupervised Learning | DHS Ch. 10 | |
| 10 | | K-means Algorithm | DHS Ch. 10 | Homework#3 due |
| | | Hierarchical Clustering | DHS Ch. 10 | Term project proposal due |
| 11 | | Componen Analysis | DHS Ch. 10 | |
| | | Single Layer Networks | DHS Ch. 6 | |
| 12 | | Multilayer Neural Networks | DHS Ch. 6 | |
| | | Neural Networks Classifiers | DHS Ch. 6 | |
| 13 | | Parameter Optimization Algorithm II | CMB Ch. 7 | |
| | | Parameter Optimization Algorithm I | CMB Ch. 7 | |
| 14 | | Project Presentation I | | Project presentation I |
| | | Project Presentation II | | Project presentation II |
| 15 | | Course Review | | |
| | | Final Exam, 2 hours | | Final Exam |

**Note:**

1. DHS--- R.O. Duda, P.E. Hart, and D.G. Stork, Pattern Classification, 2nd Edition, John Wiley and Sons, New York, 2001

2. CMB---C. M. Bishop, Neural Networks for Pattern Recognition, Oxford University Press, 1995