CS 790: Selected Topics in Computer Science
Introduction to Pattern Recognition
Winter 2002
Time: Mon-Wed 5:35-6:50
Room: 302 Russ Engineering Center

Instructor:	Ricardo Gutierrez-Osuna
Office:	401 Russ Engineering Center
Office Hours:	MW 4:00-5:30PM
Phone:	775-5120
Email:	rgutier@cs.wright.edu
URL:	http://www.cs.wright.edu/~rgutier/

Catalog Description: Lectures on and study of selected topics in current research and recent developments in computer science.

Prerequisites: MTH 255, MTH 355, STT 363 or equivalent. Basic knowledge of Linear Algebra, Probability and Statistics: algebra of matrices, geometry of Euclidean space, vector spaces and subspaces, basis, linear independence, linear transformations, eigenvalues and eigenvectors, mean, variance, probability and distributions. Programming experience in some high-level language is required.

Textbook:
- R. O. Duda, P. E. Hart and D. G. Stork, 2001, Pattern Classification, 2nd ed., Wiley.

Recommended:
- D. Hanselman and B. Littlefield, 2001, Mastering MATLAB 6, Prentice Hall.

References:
- C. M. Bishop, 1995, Neural Networks for Pattern Recognition, Oxford University Press.
- K. Fukunaga, 1990, Introduction to statistical pattern recognition, 2nd ed., Academic Pr.
- B. D. Ripley, 1996, Pattern recognition and neural networks, Cambridge University Pr.

Course Objectives: The objectives of this course are:
- Introduce the fundamental concepts in pattern recognition
- Provide the students with a toolbox of methods and algorithms they can use for practical pattern recognition problems

Course Outcomes: Upon satisfactory completion of the course, the student will be able to:
- Analyze a pattern recognition problem and present a valid formulation
- Propose and evaluate possible methods to solve the problem
- Implement a number of algorithms in a high-level language
- Design an experiment to validate formulation and implementation

Course Outline

- Introduction to pattern recognition (1 lecture)
  - What is pattern recognition?
  - Approaches to pattern recognition: statistical, neural and structural
- Overview of background material (2)
  - MATLAB®
  - Random variables and Probability
  - Linear Algebra
- Decision Theory (1)
  - Likelihood Ratio Test
  - Probability of error, Bayes Risk
- Dimensionality Reduction (2)
  - The curse of dimensionality
  - Principal Components Analysis
  - Linear Discriminants Analysis
- Statistical Classifiers (2)
  - Linear and quadratic classifiers
  - The K Nearest Neighbor (KNN) classifier
- Density Estimation (2)
  - Parameter estimation: Maximum Likelihood
  - Non-parametric density estimation: histograms, Parzen windows, KNN
  - Optimal and Naïve Bayes classifiers
- Unsupervised Learning (2)
  - Hierarchical clustering: divisive, agglomerative
  - K-means and ISODATA
  - Kohonen Self-Organizing Maps
- Feature Selection (2)
  - Search strategies: exhaustive, sequential, randomized
  - Evaluation strategies: filters, wrappers
- Validation (1)
  - Holdout, cross-validation, bootstrap
  - Data splits
- Classification using Multilayer Perceptrons (2)
  - Historical overview
  - Learning: back-prop and enhancements
  - The role of hidden and output units

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). These numeric thresholds may be lowered due to clustering, but will not be raised.

- **Homework**: There will be three homework assignments, distributed approximately every two weeks during the first half of the quarter. Homework assignments will emphasize the implementation (programming) of material presented in class. Homework assignments must be done individually.
- **Tests**: There will be a midterm exam and a final exam. All tests will be closed-books, closed-notes. One double-sided, hand-written sheet (8.5 x 11") will be allowed. Tests will have an emphasis on new material from the class notes or the reading assignments.
- **Project**: Students will work on a term project. The instructor will provide several sample projects to choose from, but the students will be allowed to propose projects related to their own research. Students can work individually or form groups of up to three people. Projects will be graded by their content (70%) and the quality of a classroom presentation (30%) at the end of the quarter.

	Weight (%)
Homework	30
Project	30
Midterm	20
Final Exam	20

Missed Tests: Missed tests can only be made up in case of emergency or work conflicts, and will require supporting documentation. Whenever possible, these issues should be discussed with the instructor prior to the conflicting date.

Collaboration vs. Academic Dishonesty: Students are encouraged to exchange ideas and form study groups to discuss the course material, prepare for homework assignments and tests. However, discussions regarding homework assignments should be kept at the conceptual level. Academic dishonesty will not be tolerated in homework assignments, projects or tests. For a list of examples of cheating see Section X in the Code of Student Conduct in the online Wright State University Student Handbook. (http://www.wright.edu/studsvcs/handbook/03_02.html.)

Course Schedule

	Date	Topic	Reading (chapters)	Assignments
Week 1	12/31	(No class)		
	1/2	Course introduction	1	
Week 2	1/7	Random variables, Probability	A.4, A.5	
	1/9	Linear Algebra, MATLAB®	A.2	
Week 3	1/14	Bayesian Decision Theory	2.1-3	HW1 assigned
	1/16	Dimensionality reduction: Principal Components Analysis	3.7, 3.8.1	
Week 4	1/21	Martin Luther King, Jr. Day (No class)		
	1/23	Dimensionality reduction: Linear Discriminants Analysis	3.8.2	
Week 5	1/28	Linear and quadratic classifiers	2.4-7	HW1 due HW2 assigned
	1/30	The K Nearest Neighbors classifier	4.5-6	
Week 6	2/4	Midterm		
	2/6	Parameter estimation, histograms, KNN	4.1-2, 4.4	
Week 7	2/11	Kernel Density Estimation	4.3	HW2 due HW3 assigned
	2/13	Unsupervised learning: statistical clustering	10.6-9	
Week 8	2/18	Unsupervised learning: Competitive Learning, Kohonen SOM	10.11, 10.14	
	2/20	Feature selection I: objective functions, sequential FS		Project proposal due
Week 9	2/25	Feature selection II: exponential and randomized FS	7.2.1-2 7.5-6	HW3 due
	2/27	Validation	9.1-2, 9.4 9.6.1-3	
Week 10	3/4	Multi-layer perceptrons: history, back-prop, enhancements	6.1-4	
	3/6	Multi-layer perceptrons: the role of hidden and output units	6.5-8	
Week 11	3/11	Final Exam		
	3/13	Project Presentations 5:30-7:30 PM, RC 302		Project report due