EE 559 **Mathematical Pattern Recognition** **Spring 2013**
B. Keith Jenkins **Course Syllabus**
Draft v0.9, 10/30/2012

Prerequisites

EE 441 Linear Algebra (co-req.) [May be waved by passing the EE placement exam.]

EE 464 Probability (pre-req.) [May be waved by passing the EE placement exam.]

Basic familiarity with Matlab

Date, time, location

Lecture: MW 3:30-4:50 PM, OHE 122 (and over DEN)

Discussion: F 12:00-12:50 PM, OHE 100B (and over DEN)

Course Description

Mathematical pattern recognition can be defined as: the categorization of input data into identifiable classes, via the extraction of significant features or attributes of the data, using mathematical techniques[1].

This course covers concepts and algorithms for pattern recognition, with an emphasis on pattern classification and decision theory. The course will stress an understanding of different algorithms at both theoretical and practical levels, as well as their advantages and disadvantages. Topics will include: statistical classification (Bayesian, parametric, and nonparametric); distribution free classification (*e.g.*, perceptron, pseudoinverse/least squares, and support vector machines); supervised and unsupervised learning; and an introduction to artificial neural networks for pattern recognition. Treatment will include a sampling of relevant classical techniques, underlying fundamentals, and current techniques. The course will include a moderately sized course project in the second half of the semester using Matlab toolboxes, to give the student an opportunity to apply concepts from class to real-world data.

Reference: 1. Definition is adapted from Tou and Gonzalez, *Pattern Recognition Principles*.

Course Outline

1. Introduction
    - Basic concepts in pattern recognition
    - A paradigm for pattern recognition
2. Distribution-Free Classification
    - Classifier design - different techniques
        - Discriminant functions
        - Linear, nonlinear
        - 2-class, multiclass
    - Training algorithms for supervised learning
        - Perceptron
        - Pseudoinverse/ minimum mean-squared error
        - Support vector machine
        - Others*
3. Statistical Classification
    - Statistics are known: Bayes decision theory
        - Optimal solutions for minimum-error and minimum-risk criteria
    - Statistics are partially known: Parameter estimation
        - Maximum Likelihood, Maximum A Posteriori, Bayesian Estimation
    - Statistics are unknown: Nonparametric techniques
        - Histogram, Parzen Windows, k-Nearest Neighbor classification
        - Techniques for reducing data and computational complexity
    - Supervised learning
4. *Feature selection and reduction
5. Unsupervised Learning
    - Statistical techniques
    - Nonstatistical clustering techniques
    - Criterion functions and optimization

Outline continues on next page…

6. Artificial Neural Networks
    - Single layer networks
    - Multiple layer networks*
    - Supervised learning
    - Capabilities and limitations
7. *Other topics of interest
    - Other optimization techniques (Conjugate gradient, Expectation Maximization, etc.)
    - Regularization
    - Algorithm-independent machine learning (e.g., boosting and AdaBoost)
    - Semi-supervised learning
    - Examples in object recognition, character recognition, text classification, and others

* Coverage depends on time available and student interest

Administrative Information

- General information about USC's Distance Education Network program for graduate courses and degrees: http://gapp.usc.edu/graduate-programs/den

EE 559 Course Materials (lecture notes, handouts, and homeworks)

- The main web site for all course materials can be accessed from:

  http://www.usc.den.net

- Course materials (lecture notes, course notes, handouts, homework assignments, etc.) will be available to all registered students at this site. Live lecture broadcasts and video archives of lectures can also be accessed from this site.
- Daily lecture notes (written out in real time during lecture) will be available after class, at the same web site (please allow 24 hours after each class for posting). Class notes for some lectures will be prepared in advance and will be available on the web site.

Course Texts

- Required texts: R. O. Duda, P. E. Hart, and D. G. Stork, "Pattern Classification", Second Edition (Wiley-Interscience, John Wiley and Sons, Inc., New York, 2001)

  David G. Stork and Elad Yom-Tov, "Computer Manual in MATLAB to accompany Pattern Classification" (Wiley-Interscience, 2004)

- Optional texts: C. M. Bishop, "Pattern Recognition and Machine Learning" (Springer, 2006)

  S. Marsland, "Machine Learning: An Algorithmic Perspective" (CRC Press, 2009)

  T. Hastie, R. Tibshirani, and J. Friedman, "The Elements of Statistical Learning: Data Mining, Inference, and Prediction", Second Edition (Springer, 2009)

Computer Packages and Languages

- You will use a Matlab toolbox for some of the homework problems and for the course project(s). You can download the toolbox from the Wiley website, using a URL and password given in the computer manual listed above. Other toolboxes will also be available for the course project.
- For some computer homework problems and portions of the computer project, you will need to generate your own code rather than using the Matlab toolbox. For this, you can use any language you are comfortable with, including Matlab.

Homework, Exams, and Grading

- Homeworks (throughout semester) 20%
- Course project 25%
- Midterm exam (during regular class time; date TBA; probably Wed. 3/6/2013 or Wed. 3/13/2013) 25%
- Final exam (Friday, May 10, 2013, 2:00 - 4:00 PM PDT) 30%

Policy on Academic Integrity

All students are expected to abide by the USC student conduct code, as well as apply common sense as to what behavior is reasonable and fair to other students. Violations will be dealt with in accordance with university guidelines.

In this class, collaboration on techniques for solving homework assignments and computer problems is allowed, and can be helpful; however, each student is expected to work out, code, and write up his or her own solution. Use of other solutions to homeworks, computer problems, or computer projects, from any source including other students, before the assignment is turned in, is not permitted. Of course, collaboration on exams is not permitted.

For more information, please see the documents at:

http://www.usc.edu/student-affairs/SJACS/acadresources.html

especially the "Academic Integrity Overview" and the three documents listed under "Publications".

Contact Information

Instructor: Prof. B. Keith Jenkins, EEB 404A
Phone 213-740-4149; fax: 213-740-6618
Email: jenkins@sipi.usc.edu (please include "EE 559" in the subject line)
Office hours: TBA

T.A.: TBA

Grader: TBA

Distance Education Network (DEN):

For help with DEN web site access, transferring of course materials (e.g., turning in and receiving homeworks from remote sites), and viewing downloaded files and viewing video lectures, consult the help function and service/contact info on the DEN web sites:
http://www.usc.den.net and http://gapp.usc.edu/graduate-programs/den .
Some of the contact information is listed below for your convenience:

General technical problems (online services, webcasts, software)	webclass@usc.edu	213-821-1321
General administrative problems	denadmin@usc.edu	213-740-4488
Master Control (Class broadcasting, classroom telephones)	networkcontrol@den.usc.edu	213-740-0130
Exams and proctoring:	denexam@usc.edu	213-821-3136
Homework submissions, records, and delivery (remote students):		213-740-9356
denhw@usc.edu	Fax submission:	213-740-9121

Sample Applications of Pattern Recognition

- Remote Sensing
  - Environment monitoring
  - Exploration of other planets
  - Water, crop, and forest resource management
- Fingerprint Identification
- Text
  - Optical character recognition
  - Categorization of topics from text
- Speech Recognition
- Image Analysis
  - Object recognition (from pictures)
  - Flexible and adaptive industrial automation
  - Robotics
  - Autonomous vehicle guidance
- Signal Analysis
  - Radar and sonar
  - Seismic
  - Communications
- Multimedia
  - Recognition of objects, actors, words, or voices in video clips or movies
- Human-Computer Interface
  - Face, expression, and gesture recognition
  - Recognition of objects in a scene (e.g., hand against background)
  - Recognition of brain signals acquired for brain-computer interface
- Biomedical and bioinformatics
  - Gene analysis
  - DNA sequencing
  - Analysis of large amounts of data
  - EKG, EEG, CT, MRI, fMRI, PET, NIRS data
- Finance
  - Investments, including stock market analysis and prediction
  - Economic analysis (economic indicators)
  - Banking (loan risk, signature verification)