**EE 559** **Mathematical Pattern Recognition** **Spring 2017**
**B. Keith Jenkins** **Course Syllabus**
v1.1, 1/7/2017

## Corequisites

- EE 441, Applied Linear Algebra for Engineering
- EE 503, Probability for Electrical and Computer Engineers
- Corequisites may be waived by passing the corresponding placement exam.

  Both placement exams are offered by the EE Department, on or near the beginning of Spring semester. You must be on campus to take it on that day if you want to place out of either (or both) corequisite(s). (Distant students intending to take EE 559 over the DEN@Viterbi system: if you want to take a placement exam at your remote location, please contact denexam@usc.edu to make arrangements.) More information can be found at:

  http://ee.usc.edu/students/placement-exams.htm

  Note that "EE 441" and "EE 559" are the exams to sign up for. ("EE 559" refers to EE 503 material for entrance to EE 559.)

  For additional information, contact the EE Student Services Office.

## Prerequisite

- Basic familiarity with Matlab.

## Follow-on course

- EE 559 is the first course of a 2-course EE sequence in Machine Learning. The second course is EE 660 (Machine Learning from Signals), offered each fall semester. (While the courses can be taken in reverse order, it is recommended to take EE 559 first.)

## Date, time, location

Lecture: MW 3:30-4:50 PM, OHE 132 (and over DEN@Viterbi)

Discussion (NEW TIME): Th 5:00–5:50 PM, OHE 132 (and over DEN@Viterbi)

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

## Course Content Descriptions

**Course Description (catalog):** Distribution free classification, discriminant functions, training algorithms; statistical classification, parametric and nonparametric techniques; artificial neural networks.

**Course Description (extended):** 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, incorporating machine learning techniques. The course will stress an understanding of different algorithms at both theoretical and practical levels, as well as their advantages and disadvantages. Topics include: statistical classification and learning (Bayesian, parametric, and nonparametric); distribution free classification and learning (*e.g.*, perceptron, pseudoinverse/least squares, and support vector machines); 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.

## Administrative Information

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

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

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

  http://courses.uscden.net/d2l/home
- Course materials (daily lecture 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 a few 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 (if you'd like to explore topics further on your own; or get a different perspective on some topics):

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

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

  Simon Haykin, *Neural Networks and Learning Machines*, 3rd Edition, (Pearson, 2009).

## Computer Packages and Languages

- You will use Matlab and a Pattern Classification toolbox for some of the homework problems and for some of the course project. Access to Matlab is provided on campus and can be downloaded by all registered students at: software.usc.edu. The Matlab "Pattern Classification" toolbox can be downloaded from the Wiley website, using a URL and password given in the computer manual listed above. Other toolboxes (freely downloaded) will also be available for the course project.
- For portions of the computer project that involve coding on your own, you can use any language you are comfortable with, including Matlab, Python, C++, etc.

**Homework, Exams, and Grading** (subject to minor changes before semester begins)

- Homeworks (throughout semester) 22%
- Class participation (online discussion forum, lectures) 3%
- Course project 25%
- Midterm exam (TBA; during one regular lecture day and time, sometime during 2/27 - 3/10) 25%
- Final exam (Friday, May 5, 2017, 2:00 - 4:00 PM PDT) 25%

## Collaboration on assignments 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 course projects, from any source including other students, before the assignment is turned in, is not permitted. Of course, collaboration on exams is not permitted. Please also see the last page of this syllabus for additional policies that apply to all USC classes.

# 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 and optimization for supervised learning
        - Perceptron (and optimization by gradient descent)
        - Pseudoinverse/ minimum mean-squared error
        - Support vector machine (and Lagrange optimization with constraints)
        - 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. Validation and data reduction
    - Validation and cross-validation
    - Feature selection and reduction
5. Artificial Neural Networks
    - Single layer networks
    - Multiple layer networks
    - Supervised learning
    - Capabilities and limitations

## Sample Applications of Pattern Recognition

- Remote Sensing
  - Environment monitoring
  - Extraterrestrial exploration
  - 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 or videos)
  - Flexible and adaptive industrial automation
  - Robotics
  - Autonomous vehicle control
- Signal Analysis
  - Infrared imaging, 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)

## Contact Information

**Instructor:** Prof. B. Keith Jenkins, EEB 404A
Email: jenkins@sipi.usc.edu (please include "EE 559" in the subject line)
Phone 213-740-4149; fax: 213-740-6618
Office hours: Mon., Fri. 11:00 AM - 12:00 PM.

**T.A.:** Yinwen Cao, EEB B20
Email: yinwenca@usc.edu
Phone (during office hours): 213-740-1488
Office hours: Tues. 10:00 AM - 12:00 PM

***Grader:*** TBA

**Distance Education Network (DEN@Viterbi) students:**

For help with DEN@Viterbi 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 site:

http://courses.uscden.net/d2l/home

click on "Support" in the navigation bar, then click on "DEN@Viterbi Online Students Only" in the left column.

Some of the contact information is listed below for your convenience:

| | | |
|---|---|---|
| General technical support (including homework submission, posted notes and videos, other course materials) | dentsc@usc.edu | 213-740-9356 |
| General administrative support | masters@gapp.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 |

## Statement on Academic Conduct and Support Systems

### Academic Conduct

Plagiarism – presenting someone else's ideas as your own, either verbatim or recast in your own words – is a serious academic offense with serious consequences. Please familiarize yourself with the discussion of plagiarism in *SCampus* in Section 11, *Behavior Violating University Standards* https://scampus.usc.edu/1100-behavior-violating-university-standards-and-appropriate-sanctions. Other forms of academic dishonesty are equally unacceptable. See additional information in *SCampus* and university policies on scientific misconduct, http://policy.usc.edu/scientific-misconduct.

Discrimination, sexual assault, and harassment are not tolerated by the university. You are encouraged to report any incidents to the *Office of Equity and Diversity* http://equity.usc.edu or to the *Department of Public Safety* http://capsnet.usc.edu/department/department-public-safety/online-forms/contact-us. This is important for the safety of the whole USC community. Another member of the university community – such as a friend, classmate, advisor, or faculty member – can help initiate the report, or can initiate the report on behalf of another person. *The Center for Women and Men* http://www.usc.edu/student-affairs/cwm/ provides 24/7 confidential support, and the sexual assault resource center webpage http://sarc.usc.edu describes reporting options and other resources.

### Support Systems

A number of USC's schools provide support for students who need help with scholarly writing. Check with your advisor or program staff to find out more. Students whose primary language is not English should check with the *American Language Institute* http://dornsife.usc.edu/ali, which sponsors courses and workshops specifically for international graduate students. *The Office of Disability Services and Programs* at http://sait.usc.edu/academicsupport/centerprograms/dsp/home_index.html provides certification for students with disabilities and helps arrange the relevant accommodations. If an officially declared emergency makes travel to campus infeasible, *USC Emergency Information* *http://emergency.usc.edu* will provide safety and other updates, including ways in which instruction will be continued by means of blackboard, teleconferencing, and other technology.