EE 559 **Mathematical Pattern Recognition** **Spring 2020**
B. Keith Jenkins **Course Syllabus**
v1.0, 1/12/2020

Corequisites

- EE 503, Probability for Electrical and Computer Engineers
- EE 510, Linear Algebra for Engineering

Required technical proficiency

- Familiarity with python (preferred) or MATLAB is required. While you may use either language for homework problems (and your project), everyone is encouraged to learn python if they don't already know it, and use it for class work.

Date, time, location

- Lecture: MW 3:30-4:50 PM, OHE 122
  and over DEN@Viterbi
- Discussion: Th 5:30–6:20 PM, OHE 122
  and over DEN@Viterbi

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: Foundations and Methods), offered each fall semester.
- EE 559 is a prerequisite for EE 660.

Related courses

- To browse other courses related to pattern recognition, machine learning, and data science:

  http://ee.usc.edu/info/studentresources/Data%20Science.pdf

Course Content

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

Course Description (extended): This course covers concepts and algorithms for pattern recognition using machine learning, with an emphasis on pattern classification and decision theory. The course will stress an understanding of different supervised-learning algorithms at both theoretical and practical levels, as well as their advantages and disadvantages. Topics include: distribution free classification and learning (*e.g.*, perceptron, pseudoinverse/least squares, and support vector machines); artificial neural networks (single and multiple layer feedforward networks, including supervised learning and capabilities); feature reduction and cross-validation; statistical classification and learning (Bayesian, parametric, and nonparametric). Treatment will include relevant classical techniques, underlying fundamentals, and current techniques. The course will include a moderately sized course project in the second half of the semester, to give the student an opportunity to design and implement a pattern recognition system on real-world data using topics from class.

Contact Information

Instructor: Prof. B. Keith Jenkins
jenkins@sipi.usc.edu [2]
EEB 404A, phone 213-740-4149, fax 213-740-6618
Office hours: Tue 2:00–4:00 PM.

T.A.'s: Zihang Cheng
zihangch@usc.edu [2]
Location to be determined
Office hours: Mon 10:00-11:30 AM

Pratyusha Das
daspraty@usc.edu [2]
PHE 320 (to be confirmed)
Office hours: Thu 10:00-11:30 AM

Fernando Valladares Monteiro
fvallada@usc.edu [2]
EEB 322
Office hours: Wed 1:30-3:00 PM

Graders: TBA

Notes:

[1] For technical questions on the course material over the internet, please post to piazza.

[2] For all email correspondence relating to this class, please include "EE 559" in the subject line.

[3] About help with coding: generally you are expected to use your own resources for help with coding (this can include help files, internet searches, and discussions with other students - on piazza or in person). If this isn't sufficient and you want personalized help, please contact Pratyusha or Fernando for help with python, or any TA for help with MATLAB.

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.

Course Texts and Other Books of Interest

- Required text: R. O. Duda, P. E. Hart, and D. G. Stork, *Pattern Classification*, Second Edition (Wiley-Interscience, John Wiley and Sons, Inc., New York, 2001)
- Optional books (if you'd like to explore course topics further on your own; get a different perspective on course topics; or explore related topics not covered in this class):

  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).

  I. Goodfellow, Y. Bengio, A. Courville, *Deep Learning* (MIT Press, 2016)

  Kevin Murphy, *Machine Learning: A Probabilistic Perspective*, (MIT Press, Cambridge, 2012).

Homework Assignments

- There will be approximately one homework assignment per week. Homework assignments will generally include theory and analytical problems, as well as computer problems. For the computer problems you may use MATLAB or python.

  Access to MATLAB is provided on campus and can be downloaded by all registered students at: itservices.usc.edu/software . Python is open source and available for free download (*e.g.*, as a distribution like www.anaconda.com).

Course Project (in second half of semester)

- For the course project, you may use MATLAB, python, and/or C/C++. If you want to use any other language, please check with the instructor or TA first.
- You will be given a real-world dataset and nontrivial problem to solve (or possibly a choice of 2 or 3 datasets with associated problems to solve). Your goal is to design and demonstrate a good-performing pattern recognition system, to compare some different approaches, and to understand and explain the results you obtain. To do so, you will use your choice of preprocessing, feature extraction, dimensionality adjustment, and classifiers, etc., to solve the problem; as well as cross validation, training and test datasets as appropriate. You will be given some guidelines and suggestions, as well as some requirements of items your project must include. Near the end of the semester, you will turn in a written report and a file with your code; detailed instructions will be posted.
- Your course project will be graded by the following criteria: inclusion of required elements; understanding and interpretation (of approach, algorithms used, and results); technical soundness and final performance; quantity and quality of effort; and report write-up (clarity, conciseness, and completeness).

Submission of Homework Assignments and Late Policy

Homework assignments will be submitted by uploading one pdf file of your solution, and one pdf file of your code, to the Assignment Dropbox on the D2L website. Graded assignments and comments will be available on the same website, approximately one week later.

Due date and time will be given with each homework assignment. The late submission policy will be posted. Late penalties will be uniformly applied to everyone, and there will be no exceptions. Your two lowest homework scores will be dropped to accommodate special circumstances, without the need for requesting an extension.

Grading and exam schedule

- Homework (throughout semester) 20%
- Course project 25%
- Midterm exam (Monday, March 9, 2020, 3:30 – 4:50 PM PDT) 25%
- Final exam (Friday, May 8, 2020, 2:00 - 4:00 PM PDT) 30%
- Extra credit problems (throughout the semester) will be tallied separately from the above scores, and will determine students' grades for borderline cases (*e.g.*, students near border between A and A-, etc.).

Policy on Collaboration and Individual Work 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 homework assignments, computer problems, or course projects, from any source including other students, before the assignment is turned in, is not permitted.

For class projects, general collaboration to resolve issues, or to clarify technical material, is allowed. Use of internet as well as journal and conference literature is encouraged. However, each student (or team) does their own work and writes up their own report. The author(s) of the report are presenting themselves as having done the work described in the report. Any reported work, explanations, information, or code that is obtained from others must be cited as such; instructions for doing this will be given with the project assignment. Including such work in the report without citing it amounts to plagiarism.

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
            - Optimization by gradient descent and stochastic descent
        - Pseudoinverse/ minimum mean-squared error
            - Optimization by algebraic formula
        - Support vector machine
            - Lagrange optimization with constraints
        - Classifier complexity, degrees of freedom, and introduction to VC dimension
3. Validation and data reduction
    - Validation and cross-validation; test
    - Feature selection; dimensionality reduction and expansion
        - Principal Components Analysis
        - Other more optimal techniques (Fisher's linear discriminant, MDA)
        - Nonlinear mappings
4. Statistical Classification
    - Statistics are known: Bayes decision theory
        - Optimal solutions for minimum-error and minimum-risk criteria
    - Statistics are unknown: Nonparametric methods
        - Histogram techniques
        - Parzen Windows
        - *k*-Nearest Neighbor classification
    - Statistics are partially known: Parameter estimation
        - Maximum Likelihood, Maximum A Posteriori, Bayesian Estimation

5. Artificial Neural Networks
    - Single layer (feedforward) networks
        - Supervised learning algorithms
    - Multiple layer (feedforward) networks
        - Gradient descent and back-propagation learning
    - Capabilities and limitations

Sample Applications of Pattern Recognition

- Remote Sensing
  - Environment monitoring
  - Extraterrestrial exploration
  - Water, crop, and forest resource management
- Fingerprint and Face Identification
- Text
  - Optical character recognition
  - Natural language processing
  - Categorization of topics or emotions from text
- Audio
  - Speech recognition
  - Emotion recognition
  - Music recognition
- 2D and 3D Image and Video Analysis
  - Recognition of objects, people, actions, obstacles, diseases
  - Flexible and adaptive industrial automation
  - Robotics
- Autonomous vehicle operation
- Signal Analysis
  - Infrared imaging, lidar, radar, sonar
  - Health, medical, and fitness related signals, including from wearable devices
  - Seismic, including for earthquake analysis and prediction
  - Communications
- Multimedia
  - Recognition of objects, actors, words, or voices in video clips or movies
- Human-Computer Interface
  - Face, expression, and gesture recognition
  - Recognition of brain signals acquired for brain-computer interface
- Biomedical and bioinformatics
  - Gene analysis
  - DNA sequencing
  - Analysis of large or huge amounts of data
  - EKG, EEG, CT, MRI, fMRI, PET, NIRS data
- Finance
  - Investments, including stock market analysis and prediction
  - Economic analysis (economic indicators)
  - Banking (credit risk, ID authentication)

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 Part B, Section 11, "Behavior Violating University Standards" policy.usc.edu/scampus-part-b. Other forms of academic dishonesty are equally unacceptable. See additional information in SCampus and university policies on scientific misconduct, policy.usc.edu/scientific-misconduct.

Support Systems:

Student Health Counseling Services - (213) 740-7711 – 24/7 on call
engemannshc.usc.edu/counseling
Free and confidential mental health treatment for students, including short-term psychotherapy, group counseling, stress fitness workshops, and crisis intervention.

National Suicide Prevention Lifeline - 1 (800) 273-8255 – 24/7 on call
suicidepreventionlifeline.org
Free and confidential emotional support to people in suicidal crisis or emotional distress 24 hours a day, 7 days a week.

Relationship and Sexual Violence Prevention Services (RSVP) - (213) 740-4900 – 24/7 on call
engemannshc.usc.edu/rsvp
Free and confidential therapy services, workshops, and training for situations related to gender-based harm.

Office of Equity and Diversity (OED) | Title IX - (213) 740-5086
equity.usc.edu, titleix.usc.edu
Information about how to get help or help a survivor of harassment or discrimination, rights of protected classes, reporting options, and additional resources for students, faculty, staff, visitors, and applicants. The university prohibits discrimination or harassment based on the following protected characteristics: race, color, national origin, ancestry, religion, sex, gender, gender identity, gender expression, sexual orientation, age, physical disability, medical condition, mental disability, marital status, pregnancy, veteran status, genetic information, and any other characteristic which may be specified in applicable laws and governmental regulations.

Bias Assessment Response and Support - (213) 740-2421
studentaffairs.usc.edu/bias-assessment-response-support
Avenue to report incidents of bias, hate crimes, and microaggressions for appropriate investigation and response.
The Office of Disability Services and Programs - (213) 740-0776
dsp.usc.edu

Support and accommodations for students with disabilities. Services include assistance in providing readers/notetakers/interpreters, special accommodations for test taking needs, assistance with architectural barriers, assistive technology, and support for individual needs.

USC Support and Advocacy - (213) 821-4710
studentaffairs.usc.edu/ssa
Assists students and families in resolving complex personal, financial, and academic issues adversely affecting their success as a student.

Diversity at USC - (213) 740-2101
diversity.usc.edu
Information on events, programs and training, the Provost's Diversity and Inclusion Council, Diversity Liaisons for each academic school, chronology, participation, and various resources for students.

USC Emergency - UPC: (213) 740-4321, HSC: (323) 442-1000 – 24/7 on call
dps.usc.edu, emergency.usc.edu
Emergency assistance and avenue to report a crime. Latest updates regarding safety, including ways in which instruction will be continued if an officially declared emergency makes travel to campus infeasible.

USC Department of Public Safety - UPC: (213) 740-6000, HSC: (323) 442-120 – 24/7 on call
dps.usc.edu
Non-emergency assistance or information.