USC Viterbi School of Engineering

EE559: Mathematical Pattern Recognition (Summer 2018)

Units: 3

Instructor: Mohammad Reza Rajati, PhD
PHE 414
rajati@usc.edu – Include 559 in subject

Office Hours: Monday 1:00 –3:00

TA(s): Mehrdad Kiamari
kiamari@usc.edu – Include EE 559 in subject

Office Hours: Friday 3:00-5:00 PM

Office Location: PHE 320

Grader(s): Nivedita Suresh
nsuresh@usc.edu – Include EE 559 in subject
Balasubramanian Thiagarajan
bthiagar@usc.edu@usc.edu – Include EE 559 in subject

Lecture(s): Monday, Wednesday, 9:00 - 10:55 am in OHE 132 (Section 30585R)
Monday, Wednesday, 9:00 - 11:50 DEN@Viterbi (Section 30589D)

Discussion(s): Friday, 9:00-9:50 am in OHE 132

Webpages: Piazza Class Page for everything except grades
and USC DEN Class Page for grades
and GitHub for code submission
– All HWs, handouts, solutions will be posted in PDF format
– *Student has the responsibility to stay current with webpage material*

Prerequisites: No formal pre-requisites.
Prior courses in multivariable calculus, linear algebra, and probability.
– This course is a prerequisite to EE 660.

Corequisites: EE 503, EE 510

Other Requirements: Basic computer skills (e.g., plotting, Python, Matlab, R, etc.).

Tentative Grading: Assignments 45%
Midterm Exam 25%
Final Exam 30%
Participation on Piazza* 5%

Letter Grade Distribution:

$\geq$ 93.00	A	73.00 - 76.99	C
90.00 - 92.99	A-	70.00 - 72.99	C-
87.00 - 89.99	B+	67.00 - 69.99	D+
83.00 - 86.99	B	63.00 - 66.99	D
80.00 - 82.99	B-	60.00 - 62.99	D-
77.00 - 79.99	C+	$\leq$ 59.99	F

Disclaimer: Although the instructor does not expect this syllabus to drastically change, he reserves every right to change this syllabus any time in the semester.

Note on e-mail vs. Piazza: If you have a question about the material or logistics of the class and wish to ask it electronically, please post it on the piazza page (not e-mail). You may post it anonymously if you wish. Often times, if one student has a question/comment, other also have a similar question/comment. Use e-mail with the professor, TA, graders only for issues that are specific to your individually (e.g., a scheduling issue or grade issue).

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

Course Objectives: Upon successful completion of this course a student will

- Broadly understand major algorithms used in pattern recognition.
- Understand the difference between supervised and unsupervised learning techniques.
- Understand resampling methods, including cross-validation and bootstrap.
- Understand methods of evaluation of classifiers.
- Understand statistical and distribution-free pattern recognition techniques.
- Understand density estimation techniques
- Understand dimensionality reduction, regularization, and kernel methods.
- Understand feedforward neural networks and deep learning.

Exam Dates:

- **Midterm Exam:** Friday, 29 June, 9:00 - 10:50 AM
- **Final Exam:** Monday, July 23, 9:00 - 11:50 AM

Textbooks:

- **Required Textbooks:**

  1. *Pattern Classification*, 2nd Edition
     Authors: Richard O. Duda, Peter E. Hart, and David G. Stork; Wiley, 2001. **ISBN-13:** 978-81-265-1116-7

2. *Applied Predictive Modeling*, 1st Edition
   Authors: Max Kuhn and Kjell Johnson; Springer; 2016. **ISBN-13:** 978-1-4614-6848-6

- **Recommended Textbooks:**

1. *An Introduction to Statistical Learning with Applications in R*, 1st Edition
   Authors: Gareth James, Daniela Witten, Trevor Hastie, and Robert Tibshirani; Springer; 2013. **ISBN-13:** 978-1-4614-7137-0
2. *Machine Learning: An Algorithmic Perspective*, 2nd Edition
   Author: Stephen Marsland; CRC Press; 2014. **ISBN-13:** 978-1-4614-7137-0
3. *Pattern Recognition and Machine Learning*, 1st Edition
   Author: Christopher Bishop; Springer; 2006. **ISBN-13:** 978-0-387-31073-2
4. *Pattern Recognition*, 1st Edition
   Author: Sergio Theodoridis; Academic Press; 2009. **ISBN-13:** 978-1-597492720
5. *Computer Age Statistical Inference: Algorithms, Evidence, and Data Science*, 1st Edition
   Authors: Bradley Efron and Trevor Hastie; Cambridge University Press, 2016. **ISBN-13:** 978-1107149892
6. *The Elements of Statistical Learning*, 2nd Edition
   Authors: Trevor Hastie, Robert Tibshirani, and Jerome Friedman; Springer, 2009. **ISBN-13:** 978-0-387-84857-0
7. *Deep Learning*, 1st Edition
   Authors: Ian Goodfellow and Yoshua Bengio; Springer, 2009. **ISBN-13:** 978-0-262-03561-3

Grading Policies:

- The letter grade distribution table guarantees the *minimum* grade each student will receive based on their final score. When appropriate, relative performance measures will be used to assign the final grade, at the discretion of the instructor.
  - Final grades are non-negotiable and are assigned at the discretion of the instructor. If you cannot accept this condition, you should not enroll in this course.
  - Two of your lowest homework grades will be dropped from the final grade.
  - *Participation on Piazza has up to 5% extra credit, which is granted on a competetive basis *at the discretion of the instructor.*

- **Homework Policy**
  - Due to shortness of the summer session, the project of this course is integrated into the assignments. Assignments include theoretical problems as well as application of the algorithms to real-world data.
  - Homework is assigned on a weekly basis. *Absolutely no late homework will be accepted. A late assignment results in a zero grade.*

  - Homework solutions should be typed or *scanned* using scanners or mobile scanner applications like CamScanner and uploaded on the course website (photos taken by cell-phone cameras and in formats other than pdf will NOT be accepted). Programs and simulation results have to be uploaded on the course website as well.
  - Students are encouraged to discuss homework problems with one another, but each student must do their own work and submit individual solutions written/ coded in their own hand. Copying the solutions or submitting identical homework sets is written evidence of cheating. The penalty ranges from F on the homework or exam, to an F in the course, to recommended expulsion.
  - Posting the homework assignments and their solutions to online forums or sharing them with other students is strictly prohibited and infringes the copyright of the instructor. Instances will be reported to USC officials as academic dishonesty for disciplinary action.

- **Exam Policy**

  - **Make-up Exams:** No make-up exams will be given. If you cannot make the above dates due to a class schedule conflict or personal matter, you must drop the class. In the case of a required business trip or a medical emergency, a signed letter from your manager or physician has to be submitted. This letter must include the contact of your physician or manager.
  - Midterms and final exams will be closed book and notes. Calculators may or may not be allowed depending on the exam (the decision will be announced before each exam). No computers and cell-phones or any devices that have internet capability will be allowed. One letter size cheat sheet (back and front) is allowed for the midterms. Two letter size cheat sheets (back and front) are allowed for the final.
  - All exams are cumulative, with an emphasis on material presented since the last exam.

- **Attendance:**

  - Students are required to attend all the lectures and discussion sessions and actively participate in class discussions. Use of cellphones and laptops is prohibited in the classroom. If you need your electronic devices to take notes, you should discuss with the instructor at the beginning of the semester.

Important Notes:

- Textbooks are secondary to the lecture notes and homework assignments.
- Handouts and course material will be distributed.
- Please use your USC email to register on Piazza and to contact the instructor and TAs.

Tentative Course Outline[1]

Wednesday	Monday
May 16th **1** **Introduction** Machine Learning Types of Learning Pattern Classification A Simple Classifier: K Nearest Neighbors	21st **2** **Model Assessment** Accuracy Other Metrics The Bias-Variance Trade-off Confusion Matrices and Hypothesis Testing Class Imbalance Receiver Operational Curve (ROC) Remedies for Class Imbalance **Resampling Methods** Cross-Validation The Bootstrap
23rd **3** **Bayesian Decision Theory** Maximum Likelihood and Maximum A Posteriori Decisions Minimum Risk Decision Rule Minimum Error Rate Decision*	28th Memorial Day
30th **4** **Bayesian Decision Theory** Discriminant Functions and Decision Surfaces Linear Discriminant Analysis Quadratic Discriminant Analysis* Comparison with K-Nearest Neighbors Naïve Bayes' Classifier Text Classification	June 4th **5** **Parameter Estimation for Classification** Maximum Likelihood Estimation Maximum A Posteriori Estimation Expectation Maximization Logistic Regression
6th **6** **(Linear) Model Selection and Regularization** Subset Selection* Shrinkage Methods Dimension Reduction Methods and Principal Component Analysis (PCA)*	11th **7** **(Linear) Model Selection and Regularization** Probabilistic and Information Theoretic Methods* Considerations in High Dimensions*

[1]Special Thanks to Prof. Keith Jenkins for his assistance in designing the syllabus.

Wednesday	Monday
13th **8** **Density Estimation** Histograms Parzen Windows K-Nearest Neighborhood Method for Density Estimation Kernel Density Estimation	18th **9** **Discriminant Functions** Linear Discriminants and Decision Surfaces Multi-Class and Multi-Label Problems One vs. One and One vs. All Classification Metrics for Assessing Multi-Class and Multi-Label Problems* Generalized Linear Discriminants Chain Classifiers
20th **10** **Optimization for Discriminative Models** Lagrange Constrained Optimization Gradient Descent Newton and Quasi-Newton Methods* Nelder-Mead Algorithm* Broyden-Fletcher-Goldfarb-Shanno (BFGS) Algorithm*	25th **11** **Discriminant Functions** Minimizing Perceptron Criterion Minimum Squared Error Learning Pseudo-Inverse Learning Widrow-Hoff Algorithm Generalized Linear Discriminants Radial Basis Function Neural Networks*
27th **12** **Support Vector Machines** Maximal Margin Classifier Support Vector Classifiers Support Vector Machines	July 2nd **13** **Support Vector Machines** The Kernel Trick L1 Regularized SVMs Multi-class and Multilabel Classification Using SVMs The Vapnik-Chervonenkis Dimension
4th Independence Day	9th **14** **Similarity Learning*** Metric Learning
11th **15** **Semi-supervised Learning*** Self-training Co-training Active Learning	16th **16** **Neural Networks and Deep Learning** Feedforward Neural Networks Backpropagation and Gradient Descent Overfitting
18th **17** **Neural Networks and Deep Learning** Autoencoders and Deep Feedforward Neural Networks Regularization Early Stopping and Dropout Adversarial Training*	23rd **18** **Final Exam**

Notes:

- Items marked by * will be covered only if time permits.

Statement on Academic Integrity: USC seeks to maintain an optimal learning environment. General principles of academic honesty include the concept of respect for the intellectual property of others, the expectation that individual work will be submitted unless otherwise allowed by an instructor, and the obligations both to protect one's own academic work from misuse by others as well as to avoid using another's work as one's own. All students are expected to understand and abide by these principles. SCampus, the Student Guidebook, contains the University Student Conduct Code (see University Governance, Section 11.00), while the recommended sanctions are located in Appendix A. See: http://scampus.usc.edu.

Emergency Preparedness/Course Continuity in a Crisis In case of a declared emergency if travel to campus is not feasible, USC executive leadership will announce an electronic way for instructors to teach students in their residence halls or homes using a combination of Blackboard, teleconferencing, and other technologies. See the university's site on Campus Safety and Emergency Preparedness: http://preparedness.usc.edu

Statement for Students with Disabilities: Any student requesting academic accommodations based on a disability is required to register with Disability Services and Programs (DSP) each semester. A letter of verification for approved accommodations can be obtained from DSP. Please be sure the letter is delivered to me (or to TA) as early in the semester as possible. DSP is located in STU 301 and is open 8:30 a.m.5:00 p.m., Monday through Friday. Website: http://sait.usc.edu/academicsupport/centerprograms/dsp/home_index.html

(213) 740-0776 (Phone), (213) 740-6948 (TDD only), (213) 740-8216 (FAX) ability@usc.edu.