For Summer 2021: to get the full 4 units of credit, please also register for 1 unit of directed research (EE 590):

1. Through myviterbi.usc.edu, Directed Research, Ming Hsieh Dept. of ECE, choose EE 590 (1 unit);
2. For supervisor/instructor: Mohammad Reza Rajati; for topic/project description: “EE 559 supplement”;
3. You will then be given D clearance for EE 590 (1 unit)
4. At the end of the semester, you will receive a letter grade for EE 559 and CR/NC for EE 590.

=> If you don’t register for EE 590, you will only get 3 units of credit.

EE 559: Machine Learning I: Supervised Methods (including Mathematical Pattern Recognition)(Summer 2021)

Units: 4
Instructor: Mohammad Reza Rajati, PhD
Office Location: Online
rajati@usc.edu – Include EE 559 in subject
Office Hours: Thursday 1:00 –2:00 PM (by appointment)

TA(s): Amirhesam Abedsoltan
abedsolt@usc.edu – Include EE 559 in subject
Office Hours: TBA
Office Location: Online, by appointment

Grader(s): Zhenyang Li
lizhenya@usc.edu – Include EE 559 in subject

Lecture(s): Tuesday, Wednesday, Thursday 11:00 am - 1:00 pm online
Discussion(s): Friday, 2:00-2:50 pm online
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.).
–**Note:** Students need to be familiar with Python programming
or be willing to learn Python.
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). Often times, if one student has a question/comment, other also have a similar question/comment. Use private Piazza posts with the professor, TA, graders only for issues that are specific to your individually (e.g., a scheduling issue or grade issue). Try minimizing the use of email to the course staff.

Catalogue Description: Distribution-free and probabilistic methods for supervised classification and regression; learning algorithms; optimization techniques; feature-space transformations; parametric and nonparametric methods; Bayes decision theory; artificial neural networks.

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

- Broadly understand major algorithms used in supervised machine learning.
- Understand the difference between supervised and unsupervised learning techniques.
- Understand regression techniques.
- Understand resampling methods, including cross-validation and bootstrap.
- Understand methods of evaluation of classifiers and regression models.
- Understand statistical and distribution-free pattern recognition techniques.
- Understand density estimation techniques
- Understand kernel methods for regression and classification.
- Understand dimensionality reduction, feature creation, and regularization.
- Understand unsupervised learning methods that serve as pre-processing for supervised methods.
- Understand feedforward neural networks and deep learning.

Exam Dates:

- **Midterm Exam:** Friday, June 25, 11:00 AM- 12:50 PM (may be changed to afternoon of the same day)
- **Final Exam:** Tuesday, July 27, 11:00 AM- 12:50 PM (may be changed to afternoon of the same day)

Textbooks:

- **Required Textbooks:**
  1. *The Elements of Statistical Learning*, 2nd Edition
     Authors: Trevor Hastie, Robert Tibshirani, and Jerome Friedman; Springer, 2009. **ISBN-13:** 978-0-387-84857-0
  2. *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
- **Recommended 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
  3. *Machine Learning: An Algorithmic Perspective*, 2nd Edition
     Author: Stephen Marsland; CRC Press; 2014. **ISBN-13:** 978-1-4614-7137-0
  4. *Pattern Recognition and Machine Learning*, 1st Edition
     Author: Christopher Bishop; Springer; 2006. **ISBN-13:** 978-0-387-31073-2
  5. *Pattern Recognition*, 1st Edition
     Author: Sergio Theodoridis; Academic Press; 2009. **ISBN-13:** 978-1-597492720
  6. *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
  7. *Deep Learning*, 1st Edition
     Authors: Ian Goodfellow and Yoshua Bengio; Springer, 2009. **ISBN-13:** 978-0-262-03561-3
  8. *Neural Networks and Learning Machines*, 3rd Edition
     Author: Simon Haykin; Pearson; 2008. **ISBN-13:** 978-0131471399

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.
  - Your lowest homework grade and half of your second lowest homework grade will be dropped from the final grade. For example, if you received 90, 85, 10, 95, 65, 80, 100 your homework score will be $\frac{0.5\times65+80+85+90+95+100}{5.5} = 87.72$ instead of $\frac{10+65+80+85+90+95+100}{7} = 75$.

- *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 an approximately weekly basis. A one-day grace period can be used for each homework with 10% penalty. *Absolutely no late homework will be accepted after the grace period. A late assignment results in a zero grade.*
  - In case of *documented illness* or *grave family* situations, exceptions can be made to the late submission policy.
  - Poor internet connection, failing to upload properly, or similar issues are **NOT** acceptable reasons for late submissions. If you want to make sure that you do not have such problems, submit homework *eight* hours earlier than the deadline. Please do not ask the instructor to make individual exceptions.
  - 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, preferably in Jupyter Notebooks.
  - 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 are allowed depending in the 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	Thursday	Tuesday
May 19th **1** **Introduction to Statistical Learning** Motivation: Big Data Supervised vs. Unsupervised Learning	20th **2** **Introduction to Statistical Learning** Regression, Classification The Regression Function Nearest Neighbors	25th **3** **Introduction to Statistical Learning** Model Assessment The Bias-Variance Trade-off No Free Lunch Theorem
26th **4** **Linear Regression** Estimating Coefficients Estimating the Accuracy of Coefficients Variable Selection and Hypothesis Testing	27th **5** **Linear Regression** The Gauss-Markov Theorem Multiple Regression Analysis of Variance and The F-test Qualitative Variables	June 1st **6** **Linear Regression** Pseudo-Inverse Learning and Linear Regression
2nd **7** **Classification** Some Simple Classification Problems The Bias-Variance Trade-off Overfitting Multi-Class and Multi-Label Classification Class Imbalance SMOTE Confusion Matrices and Hypothesis Testing Accuracy Other Metrics Receiver Operational Curve (ROC) Remedies for Class Imbalance	3rd **8** **Resampling Methods** Model Assessment Validation Set Approach Cross-Validation The Bias-Variance Trade-off for Cross Validation The Bootstrap Bagging Classifiers Bootstrap Confidence Intervals*	8th **9** **Bayesian Decision Theory** Maximum Likelihood and Maximum A Posteriori Decisions Minimum Risk Decision Rule Minimum Error Rate Decision*
9th **10** **Bayesian Decision Theory** Discriminant Functions and Decision Surfaces Linear Discriminant Analysis Quadratic Discriminant Analysis* Conditional Independence Assumption	10th **11** **Parameter Estimation for Classification** Maximum Likelihood Estimation Maximum A Posteriori Estimation Naïve Bayes' Classifier Feature Creation for Text Data TF-IDF features	15th **12** **Parameter Estimation for Classification** Data Imputation Logistic Regression Multinomial Regression Generative and Discriminative Models Comparison with K-Nearest Neighbors Bayesian Estimation* Expectation Maximization*

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

Wednesday	Thursday	Tuesday
16th **13** **(Linear) Model Selection, Regularization, and Feature Creation** Subset Selection Shrinkage Methods	17th **14** **(Linear) Model Selection, Regularization, and Feature Creation** Dimension Reduction Methods and Principal Component Analysis (PCA) Fisher's Linear Discriminant Analysis	22nd **15** **Non-parametric Methods and Density Estimation** Histograms Parzen Windows K-Nearest Neighborhood Method for Density Estimation Kernel Density Estimation
23rd **16** **Discriminant Functions** Linear Discriminants and Decision Surfaces Multi-Class and Multi-Label Problems One vs. One and One vs. All Classification Perceptrons	24th **17** **Discriminant Functions** Minimizing Perceptron Criterion Minimum Squared Error Learning Pseudo-Inverse Learning Gradient Descent Formulation Widrow-Hoff Algorithm Generalized Linear Discriminants	29th **18** **Optimization for Discriminative Models** Lagrange Constrained Optimization **Support Vector Machines** Maximal Margin Classifier Support Vector Classifiers
30th **19** **Support Vector Machines** Support Vector Machines The Kernel Trick Mercer's Kernels and Mercer's Theorem	July 1st **20** **Support Vector Machines** L1 Regularized SVMs Multi-class and Multilabel Classification Using SVMs The Vapnik-Chervonenkis Dimension Multi-Label Classification Metrics for Assessing Multi-Label Problems **Support Vector Regression**	6th **21** **Radial Basis Function Neural Networks*** Training RBF Networks and Function Approximation Relationship with Least Squares, Pseudo-Inverse Learning, and Linear Regression

Wednesday	Thursday	Tuesday
7th **22** **Radial Basis Function Neural Networks*** Relationship with Kernel Methods, and Kernel Linear Regression	8th **23** **Neural Networks and Deep Learning** Feedforward Neural Networks and Perceptrons Representation of Feedforward Neural networks The Universal Approximation Theorem*	13th **24** **Optimization for Discriminative Models** Gradient Descent (Batch, Sequential, Stochastic, Mini-batch) Newton and Quasi-Newton Methods* Conjugate Gradient* The Levenberg-Marquardt Algorithm* Nelder-Mead Algorithm* Broyden-Fletcher-Goldfarb-Shanno (BFGS) Algorithm*
14th **25** **Neural Networks and Deep Learning** Backpropagation and Gradient Descent Overfitting Regularization Early Stopping and Dropout	15th **26** **Neural Networks and Deep Learning** Autoencoders and Deep Feedforward Neural Networks* Convolutional Neural Networks* Adversarial Training*	20th **27** **Unsupervised Learning** K-Means Clustering Hierarchical Clustering Unsupervised Learning as Pre-Processing for Supervised Learning Training RBF Networks via Clustering
21st **28** **Active and Semi-Supervised Learning** Semi-Supervised Learning Self-Training Co-Training Yarowsky Algorithm Refinements Active vs. Passive Learning Stream-Based vs. Pool-Based Active Learning Query Selection Strategies	22nd **29** **Guest Lecture or Fuzzy Systems***	27th **30** **Final Exam**

Notes:

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

Homework and Exam Due Dates

Monday	
May 24th -	**1**
31st **Homework 0 Due (not graded)**	**2**
June 7th **Homework 1 Due**	**3**
14th **Homework 2 Due**	**4**
21st **Homework 3 Due**	**5**
28th -	**6**
July 5th **Homework 4 Due (moved to Tuesday July 6th)**	**7**
12th **Homework 5 Due**	**8**
19th **Homework 6 Due**	**9**
26th **Homework 7 Due**	**10**

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.