



ISE 529 Predictive Analytics
Summer 2022 (second summer session)
Location: GFS 101

Instructor: Bruce Wilcox
Office: OHE 310u
Office Hours: Mon/Thurs, 3:00 – 4:00
Contact Info: brucewil@usc.edu

Teaching Assistant: TBA
Office Hours: TBA
Contact Info: TBA

Catalog Description

Analytics for supervised and unsupervised statistical learning. Generalized linear models, discriminant analysis, support vector machines. Nonparametric classification, trees, ensemble methods, k-nearest neighbors. Principal components, clustering.

Learning Objectives and Outcomes

- Develop an advanced level or proficiency with the primary classes of predictive modeling used by data scientists.
- Develop skills in using the Python programming environment and the primary packages and tools currently used by data scientists.
- Understand key concepts for measuring the performance of analytical models and key techniques for enhancing their performance.

Class Delivery Mode: This class will be conducted in a hybrid mode. While the primary mode is in-person lectures, it is possible to take this class remotely via Zoom. However, note that while the mid-term and final can be taken online, they must be taken during the regular class times.

Recommended Preparation: ISE 225 (Engineering Statistics I) or equivalent, working knowledge of a programming language.

Course Notes

Course materials will all be published on Blackboard.

Technological Proficiency and Hardware/Software Required

The course makes extensive use of the Python programming language and several of its key data science packages. These are all open source and can be downloaded by the student for no cost.

Textbooks

This class is based on the following text which is mandatory. It can be downloaded free of charge from the author's website at: <https://www.statlearning.com/>

- James, et. al., *An Introduction to Statistical Learning with Applications in R*, 2nd edition, Springer, 2021 (ISLR)

The ISLR textbook is an excellent treatment of the theory behind the various predictive modeling techniques. We will be augmenting it with a systems view of the methodology for determining the most appropriate model types and configuring and diagnosing the models using materials from the following two textbooks (which are optional for the student to have):

- Harrell, *Regression Modeling Strategies*, 2nd edition, Springer, 2015 (RMS)
- Kuhn, et. L., "Applied Predictive Modeling, Springer, 2016 (APM)

This class will be based on Python and several of major analytics libraries including NumPy, Pandas, Scikit Learn, and Statsmodels. The following references will be used related to this software:

- Heydt M., *Learning Pandas*, Packt, 2017, ISBN 978-1-78712-313-7 (LP)
- VanderPlas, *Python Data Science Handbook*, O'Reilly, 2017 (PDS)
- Muller, *Introduction to Machine Learning with Python*, O'Reilly, 2017 (MLP)

Description and Assessment of Assignments

- **Homework assignments (one per module)** - 40% of final grade
- **Mid-term exam** – 30% of final grade (covering Modules 1 – 4)
- **Final exam** - 30% of final grade

Grading Scale (Course final grades will be determined using the following scale)

A	95-100	B-	80-82	D+	67-69
A-	90-94	C+	77-79	D	63-66
B+	87-89	C	73-76	D-	60-62
B	83-86	C-	70-72	F	59 and below

Assignment Submission Policy

Assignments will all be prepared and submitted using Jupyter Notebook unless otherwise directed. They should be submitted via GradeScope by the due date. Email submissions are not allowed.

Timeline and Rules for Submission

Assignments will be accepted late for two days with a 10% penalty and will not be accepted beyond that time. The lowest homework grade will be dropped.

Course Schedule: A Weekly Breakdown

Class	Date	Topics/Daily Activities	Assignments	References
1	6/30	Module 1: Introduction to Predictive Analytics and Python/Pandas Introduction to Python, Jupyter Notebook <i>Tools: NumPy, Pandas</i>	Module 1 HW Assigned	Course Notes
2	7/7	Module 2: Modeling Introduction. Statistical learning, modeling types, model assessment and selection	Module 1 HW Due Module 2 HW Assigned	ISLR, Chapters 1-2
3	7/11	Module 3: Linear Regression Simple and Multiple Linear Regression models and extensions	Module 2 HW Due	ISLR, Chapter 3
4	7/14	Tools: sklearn, statsmodels	Module 3 HW Assigned	
5	7/18	Module 4: Classification Logistic regression, linear discriminant analysis, and generalized linear models	Module 3 HW Due Module 4 HW Assigned	ISLR, Chapter 4
6	7/21	Linear Models Review Mid-Term (90 Minutes)	Module 4 HW due	
7	7/25	Module 5: Resampling Methods Subset selection, shrinkage methods, dimension reduction methods, high dimensional data		ISLR, Chapter 5
8	7/28	Module 6: Linear Model Selection and Regularization	Module 5-6 HW Assigned	ISLR, Chapter 6
9	8/1	Module 7: Moving Beyond Linearity Basis functions, splines, generalized additive models	Module 5-6 HW Due Module 7 HW Assigned	ISLR, Chapter 7
10	8/4	Module 8: Tree-Based Methods and Ensemble Models Decision trees, forests, gradient boosting	Module 7 HW Due Module 8 HW Assigned	ISLR Chapter 8
11	8/8	Module 9: Introduction to Machine Learning Models Support Vector Machines and Neural Networks	Module 8 HW Due Module 9 HW Assigned	ISLR Chapters 9-10
12	8/11	Course Review Final Exam (120 minutes)	Module 9 HW Due	
Final	TBA	Final Exam		

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.

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:

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.