



ISE 529 Predictive Analytics
4 Units

Day/Time: TBA
Location: TBA

Instructor: Bruce Wilcox
Office: GER 203
Office Hours:
TBA
See Piazza for weekly updates

Contact Info:
Please use Piazza for all course communications
Email: brucewil@usc.edu

Teaching Assistants:
TBA

Course Description

This is a foundational course on predictive analytics and is required for all ISE MS Analytics majors. The course covers the following topics:

- **Module 1: Introduction to Predictive Analytics and Python/Pandas**
- **Module 2: Modeling Introduction.** Statistical learning, modeling types, model assessment and selection
- **Module 3: Linear Regression Introduction.** Model definition and model assessment
- **Module 4: Linear Model Diagnosis.** Resampling methods and model variance
- **Module 5: Linear Model Validation.**
- **Module 6: Linear Model Selection and Regularization.** Subset selection, shrinkage methods, dimension reduction methods, high-dimensional data.
- **Module 7: Classification.** Logistic regression, linear discriminant analysis, and generalized linear models.
- **Module 8: Generalized Linear Models and Poisson Regression**
- **Module 9: Moving Beyond Linearity**
- **Module 10: Tree-Based Methods and Ensemble Models.** Decision trees, forests, gradient boosting.
- **Module 11: Support Vector Machines.**
- **Module 12: Introduction to Neural Networks**

Learning Objectives and Outcomes

- Develop an advanced level of 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 techniques for enhancing their performance.

Class Delivery Mode: This class will be conducted in a fully in-person mode. Lectures will be broadcast and recorded using Zoom to accommodate students who are ill and those who wish to re-view portions of the lecture after the class. The mid-term and final exams must be taken in person.

Prerequisite(s): None

Recommended Preparation: ISE 225 (Engineering Statistics I) or equivalent, working knowledge of Python

Course Notes: All course materials (PowerPoints, assigned readings, etc.) will be distributed via 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)
- We will be augmenting this text 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):
 - 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

- 8-10 homework assignments (approx. one per week) - 50% of final grade
- Mid-term exam (in class) – 20% of final grade
- Final exam - 30% of final grade

Grading Scale

Course final grades will be determined using the following scale

A	95-100
A-	90-94.9
B+	87-89.9
B	83-86.9
B-	80-82.9
C+	75-79.9
C	70-74.9
C-	50-69.0
F	Below 50

Up to two points may be added to the overall grade based on class engagement.

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.

Timeline and Rules for Submission

- Homework assignments will be posted on or shortly after each weekly class and will be due one week after posted
- All assignments must be submitted prior to the due date
- You get one “free” late submission per semester. After that, there is a 10% penalty for late homework submitted within 48 hours of due date
- No submissions will be accepted after 48 hours
- The lowest homework grade will be dropped

Course Communications

- All materials will be uploaded to Blackboard
- Assignments will be submitted through Gradescope
- We will use Piazza as the primary communications mechanism
 - Class announcements will be posted there, and we request that any questions you have be posted there so that other students can benefit from your question and responses from the instructors, TAs, and hopefully other students
 - Students who actively post responses to questions MAY receive extra credit (which could result in an increase by one letter grade in borderline cases)
- I will periodically post “discussion questions” on Piazza. Class engagement credit can be earned by participating in these online discussions

Course Schedule: A Weekly Breakdown

Week	W/E	Topics/Daily Activities	Assignments	References
1	8/25	Module 1: Introduction to Predictive Analytics and Python/Pandas Introduction to Python, Jupyter Notebook <i>Tools: NumPy, Pandas</i>		Course Notes
2	9/1	Module 2: Modeling Introduction. Statistical learning, modeling types, model assessment and selection		ISLR, Chapters 1-2
3	9/8	Module 3: Linear Regression Introduction Model definition and model assessment <i>Tools: scikit-learn, statsmodels</i>		ISLR, Ch. 3.1-3.5
4	9/15	Module 4: Linear Model Diagnosis Resampling methods and model variance		ISLR 3.3.3 Lecture notes
5	9/22	Module 5: Linear Model Validation		ISLR Ch. 5
6	9/29	Module 6: Linear Model Selection and Regularization Subset selection, shrinkage methods, dimension reduction methods, high-dimensional data		ISLR Ch. 6
7	10/6	Mid-Term Module 7A: Classification Logistic regression, linear discriminant analysis, and generalized linear models		ISLR, Chapter 4.1-4.5
8	10/11	Module 7B: Classification <i>No class 10/13</i>		
9	10/20	Module 8: Generalized Linear Models and Poisson Regression Module 9: Moving Beyond Linearity		ISLR, Ch. 4.6 ISLR, Ch. 7
10	10/27	Module 10: Tree-Based Methods and Ensemble Models Decision trees, forests, gradient boosting		ISLR, Ch. 8
11	11/3	Module 11: Support Vector Machines		ISLR Ch. 9
12	11/10	Module 12: Introduction to Neural Networks		ISLR Ch.10
13	11/17			
14	11/22			
15	12/1	Course Review		
		Final Exam – 12/13 8:00AM – 10:00AM		

Notes:

- This schedule is subject to change throughout the semester. This syllabus will not be updated, but the latest schedule will always be available on Piazza.

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/ssu

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.