ISE 529 Predictive Analytics
Summer 2023 TF 3 – 5:40 p.m.
Location: ZHS 252

Instructor: Cesar Acosta-Mejia
Office: GER 216
Office Hours: TBD
Contact Info: acostame@usc.edu

Teaching Assistant: Adhithya Bhaskar
Office: online
Office Hours: M 11 a.m.
Contact Info: adhithya@usc.edu

IT Help:
Hours of Service:
Contact Info:

Course Description

This course focus is on building models for prediction and classification. The standard multiple linear regression model is the basic prediction model. This model is extended to shrinkage models (ridge and lasso regression) for improved accuracy and dimension reduction. Overfitting, bias, cross validation, and AIC are used to evaluate the performance of these models.

The course also focuses on models for classification. Logistic regression, KNN, Trees and ensembles of trees (random forests, bagging, and gradient boosting), and artificial neural networks are reviewed.

Learning Objectives and Outcomes

- To understand the Data Analytics levels: Descriptive, Predictive, and Prescriptive Analytics.
- To understand the difference between supervised and unsupervised learning methods.
- To learn the most common data aggregation operations (cross tabulation and pivot tables).
- To build models for prediction and classification.
- To understand key concepts for predictive analytics (overfitting, shrinkage, regularization, R^2, adjusted R^2, VIF, mean square prediction error, Cross-validation).
- To estimate the performance of Analytics models.
- To compare the performance of different prediction and classification models.
- To build models to classify observations into two or more classes (categories).

Prerequisite(s): An undergraduate course on Statistics and knowledge of a programming language (any).

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

Course Notes
The course material is available on Blackboard.

Technological Proficiency and Hardware/Software Required
Python programming language is used throughout the course. Prior knowledge of Python is not required. Students will learn Python and work with it throughout the term. Jupyter Notebook is used as the main interface for documenting the scripts and get the results. It is also used to create reports for assignments and exams.

Supplementary Materials (for reference)

- Navlani M., *Python Data Analysis, 3ed, Packt,* 2021, ISBN 978-1-78995-524-8
- VanderPlas, *Python Data Science Handbook*, O'Reilly, 2017
- Muller, *Introduction to Machine Learning with Python,* O'Reilly, 2017

PDF files
- Jupyter-shortcuts
- Python basics
- Pandas basics
- VanderPlas – tour-of-python

Description and Assessment of Assignments

All assignments and examinations are on-line. Unless otherwise noted the assignments are individual. All homework assignments are released and distributed one-week in advanced. Dates are shown in the Course schedule on page 4. Submit on to Blackboard by the due date. No late homework will be accepted.

Grading Policy

Assignment	Points	% of Grade
Homework	100 each (6⁺ homework assignments)	30
Midterm	100	30
Final	100	40
TOTAL		100

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

A	94-100	B-	80-82.9	D+	67-69
A-	90-93.9	C+	77-79	D	63-66.9
B+	87-89	C	73-76.9	D-	60-62.9
B	83-86.9	C-	70-72.9	F	59.9 and below

Assignment Submission Policy
Assignments should be typewritten and clean. Email submissions and late submissions are not allowed. No make-up exams are considered.

Timeline and Rules for submission
Assignments are to be returned the week after submission. Solutions will be released soon after the homework submission date.

Course Schedule: A Weekly Breakdown

	Date	Topics/Daily Activities	Deliverables	slides	Files
1	**Jun 30**	**Introduction** to Analytics Descriptive, Predictive and Prescriptive Analytics. Python and Jupyter Notebook (JN) setup. Python review, Numpy library.		overview.ppt analytics.ppt python.ppt numpy.ppt Python basics.pdf	Intro.ipynb numpy.ipynb numpyreg.ipynb Odometer.csv
2	**July 4** (recorded)	Pandas library, data structures. Most Common Data Operations. Pivot tables and cross tabulation.	HW1 Pandas	Pandas .ppt Pandas basics.pdf	Example4.ipynb Cars93.csv
3	**July 7**	Financial Data Visualization with the pandas-datareader library.		returns3.ppt	finance13.ipynb
4	**July 11**	**Linear Regression.** OLS vs. linear regression. Libraries sklearn and statsmodels	HW1 due HW2 OLS	slr1.ppt mlr4.ppt	slr3.ipynb finished3.ipynb Cars93.csv
5	**July 14**	**Linear Regression.** Categorical variables. Interaction terms. Label encoding and one-hot encoding		categoricals.ppt small.csv part2.csv homes.csv	plots3.ipynb part2c.ipynb homes.ipynb example1d,2.ipynb
6	**July 18**	Overfitting and Cross validation, Training/test sets, mean square prediction error (MSPE).	HW2 due **Midterm** (regression, cv, cross validation and categ. Vars.	cv3.ppt part1.ppt part2.ppt	Polynomial4.ipynb feature-cv7.ipynb itertools.ipynb
7	**July 21**	**Classification Problems.** Logistic Regression vs KNN.		classification2.ppt logistic3.ppt KNN2.ppt	cancerlogistic.ipynb cancer_knn.ipynb sklogis3.ipynb
9	**July 25**	**Shrinkage Methods and Regularization.** Ridge regression and the LASSO.	Midterm due HW3 RR	rr5.ppt hyperparam3.ppt	ridge8.ipynb Hitters.csv
10	**July 28**	**Classification and Regression Trees (CART) -** Examples		Trees3.ppt categ2.ppt	regression4.ipynb cart4.ipynb
11	**Aug 1**	**Ensemble Methods.** Random Forest, Bagging, and Boosting.		ensembles2.ppt	ensemblereg4.ipynb ensembcancer3.ipynb polyboosting5.ipynb
	Aug 2		HW3 due		
12	**Aug 4**	**Introduction to Neural Networks (NN).** Data representations for NN, tensors. Layers, Loss functions, optimizers.	HW4 due Final Exam released	nn4.ppt means.csv	perceptron5.ipynb multilayerp4.ipynb gradient4.ipynb mnist.ipynb
13	**Aug 8**	**Final Exam** (Schedule TBD)			

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.