Summer 2021 MATH 495: Prediction and Statistical Learning: Reading Class

Instructor: David Olive

Text: James, G., Witten, D., Hastie, T., and Tibshirani, R. (2013), *An Introduction to Statistical Learning With Applications in R*, Springer, New York, NY. ISBN: 978-1-4614-7137-0 $\approx$ \$60 at www.addall.com and www.amazon.com. The library has the text online and the last 2 authors may have the pdf file available online. The online notes Olive, D.J. (2021), *Prediction and Statistical Learning*, (http://parker.ad.siu.edu/Olive/slearnbk.htm), will be used.

Office: (J.W.) Neckers 261 (wing A) *Phone:* (618)-453-6566

email: dolive@siu.edu

Office hours:

I am also available by appointment and on a walkin basis. I will often be in my office around 10:00 and 12:00.

The *prerequisite* for this class is Math 483 or equivalent (calculus based course in probability and statistics, CS 480 may be adequate) and Math 221 or equivalent (Linear Algebra: matrix multiplication, inverse matrix, eigenvalues). A course in regression such as Math 484, 473 or 485 is recommended (CS 473 or 438 may be adequate).

Statistical Learning could be defined as the statistical analysis of multivariate data. Machine learning, data mining, big data, analytics, business analytics, data analytics, and predictive analytics are synonymous terms. The techniques are useful for Data Science and Statistics, the science of extracting information from data. The R software will be used. Often there is a response variable Y of interest, and a $p \times 1$ vector $\boldsymbol{x} = (x_1, ..., x_p)^T$ used to predict Y. The data are $(Y_1, \boldsymbol{x}_1), ..., (Y_n, \boldsymbol{x}_n)$. Multiple linear regression (MLR) and classification will be covered. Standard regression methods such as least squares (OLS) and linear discriminant analysis need n/p large. Statistical learning methods are often useful when n/p is not large. MLR methods will include OLS forward selection, lasso, relaxed lasso, elastic net, ridge regression, partial least squares, and principal component regression. Prediction intervals will be used to compare these methods, and prediction regions will be used for bootstrap hypothesis tests. Classification will consider linear discriminant analysis (LDA), lasso for LDA, K-nearest neighbors, classification trees, and support vector machines. These are supervised learning methods. Unsupervised learning methods include clustering methods and principal component analysis.

There may be overlap with Machine Learning courses such as CS437, 533, and 586. The text is also used for the SOA and CASACT actuarial statistics exams. Google "siu big data", "siu analytics", etc.

Final sometime in the week of August 1.

The grading and schedule below are tentative. (Drop days are ? Friday, , Sunday ?.)

3 homeworks may be turned in one class period late (ie on Friday) with no penalty. A fourth late will be accepted with 25% penalty. 2 quizzes may be taken late before the next class period (ie on Monday). At least two sheets of notes are allowed on quizzes more for exams. A calculator is permitted.

Grading:

HW	300		Quizzes	100	
exam1	100	exam 2	100	exam 3	100
final	300	or project		total	1000
min. grade	points	min. grade	points	min. grade	points
A	900-1000	B	800-899	C	700-799
D	550-699				

O1.1 refers to Olive (2021, ch. 1, section 1), while J2.3 and J3 refer to for James et al. (2013, section 2.3 and chapter 3).

Week of	MON	WED	Th	FRI
Jun 7	J1,J2.1,J2.2,O1.1,O1.2,O1.3	O1.2,O1.3	O1.3	O1.3, HW1,
Jun 14	O1.3, O1.4,Q1	01.4,O2.1,HW2	O2.1,O2.2, Q2	O2.2,O2.3, HW3
Jun 21	O2.3,O3.1,J5.2,J3,Q3	Exam 1	O3.2,O3.3,J6.1	J6.2,J6.3,HW4
Jun 28	O3.4,O3.5,J6.2,J6.3,Q4	O3.7,O3.8,HW5	O3.8,O3.9,Q5	O3.10-3.12,J5.1,HW6,Q6
Jul 5	no class	Exam2	Och4,J7.1-7.7,	O5.1-5.7,J4.1-4.5,HW7
Jul 12	J2.2,4.6,5.1 Q7	O5.6-5.8,J5.1,HW8	O6.1-6.2,O7.1,J10.3,Q8	O7.1,J8.1,HW9
Jul 19	J8.2,Q9	J8.2,HW10	J9.1-9.4,Q10	Exam3
Jul 26	J9.5 HW11	J10.2?,Q11	final?	final?

References: Lorayne and Lucas (2000), *The Memory Book* is useful for **memorization**.

Hastie, T., Tibshirani, R., and Friedman, J. (2009), *The Elements of Statistical Learning: Data Mining, Inference and Prediction*, 2nd ed., Springer, New York, NY.

Hastie, T., Tibshirani, R., and Wainwright, M. (2015), *Statistical Learning with Sparsity: the Lasso and Generalizations*, CRC Press Taylor & Francis, Boca Raton, FL.

Kuhn, M., and Johnson, K. (2013), *Applied Predictive Modeling*, Springer, New York, NY.

Abu-Mostafa, Y.S., Magdon-Ismail, M., and Lin, H.T. (2012), *Learning From Data*, AMLBook, Pasadena, CA.

Berk, R.A. (2008), *Statistical Learning From a Regression Perspective*, Springer, New York, NY.

Witten, I.A., and Frank, E. (2005), *Data Mining: Practical Machine Learning Tools and Techniques*, 2nd ed., Elsevier, San Francisco, CA.