

Fall 2019, Statistics 873 Statistical Learning and Data Mining

Time: M W F, 9:10–10:00 am	Place: C506 Wells Hall
Instructor: Luda Sakhanenko	Office: C441 Wells Hall
Telephone: (517)432-9795	E-Mail: sakhanen@msu.edu
Office Hours: Tu 9:10-10:10 am	or by appointment

Website: <http://stt.msu.edu/Academics/ClassPages/>

Resources:

Bousquet, O., S. Boucheron and G. Lugosi. Introduction to Statistical Learning Theory. Advanced Lectures on Machine Learning, Lecture Notes in Artificial Intelligence 3176, 169-207. (Eds.) Bousquet, O., U. von Luxburg and G. Ratsch, Springer, Heidelberg, Germany, 2004.

http://www.kyb.mpg.de/fileadmin/user_upload/files/publications/pdfs/pdf2819.pdf

Hastie, T., Tibshirani, R., and Friedman, J. The Elements of Statistical Learning: Data Mining, Inference, and Prediction. 2nd edition. Springer, 2013.

<http://statweb.stanford.edu/~tibs/ElemStatLearn/>

Devroye, L., Györfi, L., and Lugosi, G. A Probabilistic Theory of Pattern Recognition. Springer, 1997.

Prerequisites: STT 868 and STT 872 or equivalent courses.

The specific topics of the course include statistical methods focusing on machine learning and data mining, modern regression and classification techniques, support vector machines, boosting, kernel methods and ensemble methods, clustering dimension reduction, manifold learning, and selected topics.

The course will cover selected chapters in the resources and lecture notes.

Examination:

There will be **one** Midterm Exam on October 14, worth 100 points, given during the lecture time. If you are unable to take an exam, you must contact the instructor on or before the day of the exam. All excuses must be verifiable. The make-up exams will be given only under exceptional circumstances.

Presentation: The last 1-2 weeks of the course will involve students' presentations. Students are required to read journal papers related to the course topics. The list of papers will be provided by the instructor. All students will be asked to present their assigned articles in the class.

Homework: There will be **five** homework assignments worth 20 points each that will be given during the semester as we cover the corresponding material. In addition, during each class some reading and suggested homework problems from the resources will be assigned. These problems will not be collected.

Course Grade: The final course grade will be based on the total number of points earned by a student during the term in

Homework	$5 \times 20 = 100$ points
Midterm Exam	$1 \times 100 = 100$ points
Presentation	200 points
Total	400 points

The grading scale will be as follows:

90–100%	4.0	60–64%	2.0
80–89%	3.5	55–59%	1.5
70–79%	3.0	50–54%	1.0
65–69%	2.5	0–49%	0.0

Academic Honesty: The Department of Statistics and Probability adheres to the policies of academic honesty as specifies in the General Student Regulations 1.0, Protection of Scholarships and Grades, and in the All-University of Integrity of Scholarship and Grades which are included in *Spartan Life: Student Handbook and Resource Guide*. Student who plagiarize will receive a grade 0.0 on the assignment.

Note: The instructor reserves the right to make any changes she deems academically advisable.

Important Dates for Fall Semester 2019

- Aug. 28 First day of Classes
- Sept. 2 Labor Day, No Classes
- Sept. 4 Close of Adds
- Sept. 23 End of 100% Refund
- Oct. 16 Middle of the semester
- Nov. 28-29 Thanksgiving holiday, No Classes
- Dec. 6 Last day of Classes