Statistical Machine Learning for Data Science **Spring 2019**

Instructor: Zaid Harchaoui, zaid@uw.edu **TA**: Corinne Jones, cjones6@uw.edu

Time and Place: W 5:00-9:20pm in DEM 024

Office Hours:

- Zaid Harchaoui: Thursday 8:00-9:00pm Online
- Corinne Jones: Monday 8:00-9:00pm Online

Course Web Page: https://canvas.uw.edu/courses/1325582

Textbooks:

- *The Elements of Statistical Learning* by Hastie, Tibshirani, and Friedman https://web.stanford.edu/~hastie/ElemStatLearn/
- *An Introduction to Statistical Learning* by James, Witten, Hastie, and Tibshirani http://www-bcf.usc.edu/~gareth/ISL/index.html

Additional Reading

- *Machine Learning: A Probabilistic Perspective* by Kevin Murphy https://www.amazon.com/Machine-Learning-Probabilistic-Perspective-Computation/dp/0262018020/
- *All of Statistics* by Larry Wasserman https://www.amazon.com/All-Statistics-Statistical-Inference-Springer/dp/0387402721/

Prerequisites
Previous courses in the statistics sequence of the Master of Science in Data Science curriculum. Students should be able to complete programming assignments using programming languages, such as Python, R, or Matlab.

Grading
Grades will be assigned based on homework assignments (40%), a midterm (30%), a data competition (20%), and class participation (10%).

Homework
Homework will usually be posted before the week-end and will be due Friday by 11:59pm the following week. Homework due date will anyway be explicitly stated on each homework assignment.

Each student is allowed 3 deadline extensions over the quarter. The extension is 48hrs, hence from Friday 11:59pm to Sunday 11:59pm. You do not need to explicitly request an extension; Canvas will keep track of them. **Homework turned in late will receive 0 points**.

Tests and Projects
There will be a midterm exam around week 6 or week 7. The final exam will consist of the participation in a data competition, with the final submission to the competition leaderboard plus a written report plus the computer programs, all due during the finals week.

Honesty and integrity
Students are expected to practice high standards of academic and professional honesty and integrity, as defined in University of Washington Student Governance Policy, Chapter 209 Section 7.C. See the website: `https://www.washington.edu/cssc/facultystaff/academic-misconduct/`

Homework must be done individually. Each student must hand in their own answers. In addition, each student must submit their own code in the programming part of the assignment. It is acceptable, however, for students to collaborate in figuring out answers and helping each other solve the problems. In such cases, you must indicate on each homework with whom you collaborated.

No exceptions will be given to the grading policies. If you are not able to comply with the late homework policy, due to travel, conferences, other deadlines, or any other reason, do not enroll in the course.

As we sometimes reuse problem set questions from previous years, covered by papers and webpages, we expect the students not to copy, refer to, or look at the solutions in preparing their answers. Referring to unauthorized material is considered a violation of the honor code. Similarly, we expect students not to google directly for answers. The homework is to help you think about the material, and we expect you to make an honest effort to solve the problems.