00102892: Statistical Learning
Fall 2021

Lectures: Wednesdays 3:10–5:00 pm and even Fridays 8:00–9:50 am, 108 Geology Building

Instructor: Wei Lin (weilin@math.pku.edu.cn)
Office hours: Fridays 10:00–11:30 am, 1547 Science Building 1

Teaching Assistants: Yalong Lyu (lyuyl@stu.pku.edu.cn), Xinpeng Ni (nxp@stu.pku.edu.cn)
Office hours: Wednesdays 5:00–7:00 pm, 118 Building 21

Course Description:
This introductory statistical machine learning course is designed for graduate and advanced undergraduate students in statistics and probability, applied mathematics, and other fields involving learning from data. The course covers fundamental principles and methodology of machine learning, including model selection and regularization, classification and regression, support vector machines, kernel methods, boosting, clustering, and dimension reduction. If time permits, selected topics from the following will also be discussed: random forests, graphical models, ranking, online learning, and reinforcement learning. The PAC/nonasymptotic framework for machine learning theory will be introduced and developed throughout the semester.

Primary Texts:
1. Hastie, T., Tibshirani, R. and Friedman, J. (2009). *The Elements of Statistical Learning: Data Mining, Inference, and Prediction* (2nd ed.). Springer.
2. Mohri, M., Rostamizadeh, A. and Talwalkar, A. (2018). *Foundations of Machine Learning* (2nd ed.). The MIT Press.

Supplementary Texts:
1. Shalev-Shwartz, S. and Ben-David, S. (2014). *Understanding Machine Learning: From Theory to Algorithms*. Cambridge Univ. Press.
2. Vershynin, R. (2018). *High-Dimensional Probability: An Introduction with Applications in Data Science*. Cambridge Univ. Press.
3. Wainwright, M. J. (2019). *High-Dimensional Statistics: A Non-Asymptotic Viewpoint*. Cambridge Univ. Press.

Homework:
There will be about five homework assignments due in class. If you missed the class, contact the TAs to turn in your homework by the end of the day. No late homework will be accepted.

Exams:
There will be one two-hour, closed-book midterm exam on Wednesday, December 1. There will be no final exam, but students need to form groups of two to four and complete a final project on a topic of their own choice. Proposals of project topics are due by Wednesday, November 24. Oral presentations of projects are scheduled on Wednesday, December 29 and Friday, December 31.

Grading:
The course grade breaks down as follows: homework 30%, midterm exam 30%, and project 40%.

Website:
Lecture topics and homework assignments will be posted on the course website at http://www.math.pku.edu.cn/teachers/linw/2892f21.html.