

Spring 2019

ST 578: Topic course - Statistical Learning in Data Science

Instructors Annie Qu, Professor of Statistics, University of Illinois at Urbana-Champaign

Office: Illini Hall 116B; Phone: 217-244-8334; E-mail: anniequ@illinois.edu

Course information: meet twice a week, Tue and Thur. 11-12:20. **Location:** 243 Mechanical Engineering

TA: Yujia Deng, Email: yujiad2@illinois.edu

Office Hours for Instructor or TA: by appointment.

Course outlines: This course will train students to analyze large complex data using advanced statistical learning methods and algorithms. The main topics in the course include: machine learning tools in data science; large data processing; regularization methods; recommender system; network and graphics, deep learning, text mining; and imaging analysis. Students will gain practical skills of data mining and knowledge discovery in various applications such as business, political Science, biology and medicine.

Data: The following data have been discussed and analyzed in the class or students' projects: Presidential election data, IRi marketing data, MovieLens data, Millions Song Dataset (MSD), Last.fm data, ADHD data, Trip Advisor data, Yelp data, lung cancer imaging data, ADNI data (Alzheimer Disease), ABIDE data (Autism Disorder).

Credit hours: 4 credits.

Software used in class: R, Python

Prerequisites: ST510/410 and ST425.

Workload: Homework (30% of your grade), a final written project (50%) and a final presentation (20%) by each student.

Objective: After taking this course, the student should gain working knowledge of using machine learning and data mining techniques in the field of data science.