PURDUE UNIVERSITY | College of Science

STAT598 – Statistical Machine Learning

Instructor

Name: Xiao Wang
Office: Math514
Email: wangxiao@purdue.edu
Virtual office hour: Wednesday 9 am --10 am (EST) at https://purdue.webex.com/meet/wangxiao or by appointment

Teaching Assistant

Name: Haoyun Yin
Office: Math517
Email: yin164@purdue.edu
Virtual office hour: Friday 11:30 am – 12:30 pm (EST) at https://purdue.webex.com/meet/yin164 or by appointment

Course Description

This course provides a comprehensive treatment of modern statistical machine learning topics. These include linear regression, kernel method, trees, boosting, and deep neural networks. We understand these subjects from a statistical perspective with some mathematical rigorousness.

This course is also a 14-week fully online course. Each week includes 3 lectures.

Prerequisite

- Knowledge of master level of probability and mathematical statistics is required.
- Basic programming skills using Python is required.

Course Outcomes

- CO1: Understand classical supervised learning methods such as linear regression, nonlinear regression, random forest, and boosting
- CO2: Understand classical unsupervised learning methods such as K-means clustering and PCA
- CO3: Understand modern AI techniques using deep neural networks, such as manifold learning and deep generative models
- CO4: Practice complex and high dimensional data analysis using the methodology and software packages

Technical Requirements

The following information has been provided to assist you in preparing to use technology successfully.

- A reliable internet connection - capable of consistently streaming video and stable enough to finish homework and project without dropping connection.
- Access to Purdue's Brightspace Learning Management System - All course content, course readings, and exams will be accessed online through Brightspace.

Learning Resources & Texts

Required Textbook:

- *The Elements of Statistical Learning* by Hastie, Tibshirani, Friedman (Its pdf is available online)

Suggested Textbook:

- *Pattern Recognition and Machine Learning* by Christopher M. Bishop
- *Probabilistic Machine Learning: Advanced Topics*, by Kevin P. Murphy (pdf available online)
- *Deep Learning* by Goodfellow, Bengio, Courville
- *Deep Learning with PyTorch* by Stevens, Antiga, Viehmann

Instructor's Online Hours

I will be available and respond to student questions as soon as I am available generally 48 hours during the M-F work week. Student inquiries made during the weekend may experience a delayed response time. Questions about the course content, assignments, or lectures should be asked in the **Course Q & A** forum provided on the discussion boards. Students are encouraged to answer the questions their peers ask on the **Course Q & A** forum. Email should only be used for personal questions. When emailing me, please place the course number in the subject line of the email. This will help me tremendously in locating your emails quicker.

Virtual Office Hours are a synchronous session through Webex to discuss questions related to the course content. Please check the course site on Brightspace for detailed information about the virtual office hours.

Assignments

Homework: There will be four written homework assignments to be collected and graded. Late homework will NOT be accepted without prior permission. Details on these assignments, and guidelines on discussion participation and evaluation will be posted on Brightspace. The due dates for the assignments posted on Brightspace are in Eastern Standard Time (the local time zone of West Lafayette, Indiana).

Final Project: The final project must be typed, double spaced and should be 8-10 pages in length (including tables and graphics). Late projects will be graded down one letter grade.

Assignments	Points
Assignment 1	15 %
Assignment 2	15 %
Assignment 3	15 %
Assignment 4	15 %
Final Project	40 %
Total	100%

Participation and Assignment policies

Some collaboration on homework is acceptable, but each student must do his/her own write-up of the solution to show their full understanding. Direct copying of another student's solution will result in grade of zero for both students.

Please remember that an illustration of effort is in itself a creditable achievement. We will be awarding partial credits for all problems depending on approach and degree of completeness.

Residential students are expected to attend the lectures in-person, complete homework and final project on time. Online students are expected to watch the lecture videos, complete homework and final project on time. Late submission will cause late penalty on grades. Please check the requirements for each of the assignment on the course site in Brightspace.

Course Contents and Schedule

Modules	Lectures	Assignments
Module 1	Overview of Statistical Learning	•
Module 2	Linear and Nonlinear Regression Models	• Assignment 1 • Online Discussions
Module 3	High Dimensional Regression	•
Module 4	Classification	•
Module 5	Regression and Classification Trees	•
Module 6	Boosting	• Assignment 2 • Online Discussions
Module 7	Clustering and K-Means	•
Module 8	Manifold Learning (PCA, MDS, Isomap, t-SNE)	•
Module 9	MCMC	• Assignment 3 • Online Discussions
Module 10	Deep Neural Networks (MLP and CNN)	•
Module 11	Variational Autoencoder and FLOW	•
Module 12	Generative Adversarial Networks and Optimal Transport	• Assignment 4 • Online Discussions
Final Exam	No lecture this week	• Final project

Grading Scales

The Grading Scale in this course follows the Typical A-F structure, with +/- as appropriate. For full detail, this has been provided below:

Percentage	*Letter Grade*
95% or higher	A+
90% - 94%	A
85% - 89%	A-
80% - 84%	B+
75% - 79%	B
70% - 74%	B-

65% - 69%	C
60% - 64%	D
59.9% or lower	F

Netiquette

You are encouraged to comment, question, or critique ideas. However, be mindful that sarcasm and humor can be easily misconstrued in online interactions. Please read the Netiquette rules for this course:

- Give other students the opportunity to join in the discussion.
- Present ideas appropriately.
- Be cautious in using Internet language. For example, do not capitalize all letters since this suggests shouting.
- Avoid using vernacular and/or slang language. This could possibly lead to misinterpretation.
- Keep an "open-mind" and be willing to express even your minority opinion.
- Think and edit before you share (e.g., post or email).
- Ask for feedback.

Course Evaluation

During the last two weeks of the course, you will be provided with an opportunity to evaluate this course and your instructor. Purdue now uses an online course evaluation system. You will receive an official email from evaluation administrators with a link to the online evaluation site. You will have up to two weeks to complete this evaluation. Your participation is an integral part of this course, and your feedback is vital to improving education at Purdue University. I strongly urge you to participate in the evaluation system.

Academic Dishonesty

Purdue prohibits "dishonesty in connection with any University activity. Cheating, plagiarism, or knowingly furnishing false information to the University are examples of dishonesty." [Part 5, Section III-B-2-a, University Regulations] Furthermore, the University Senate has stipulated that "the commitment of acts of cheating, lying, and deceit in any of their diverse forms (such as the use of substitutes for taking examinations, the use of illegal cribs, plagiarism, and copying during examinations) is dishonest and must not be tolerated. Moreover, knowingly to aid and abet, directly or indirectly, other parties in committing dishonest acts is in itself dishonest." [University Senate Document 72-18, December 15, 1972]

Emergency Statement

In the event of a major campus emergency, course requirements, deadlines and grading percentages are subject to changes that may be necessitated by a revised semester calendar or other circumstances beyond the instructor's control. Relevant changes to this course will be posted onto the course website or can be obtained by contacting the instructors or TAs via email or phone. You are expected to read your @purdue.edu email on a frequent basis.

Disability Statement

Students with disabilities must be registered with Disability Resource Center in the Office of the Dean of Students before classroom accommodations can be provided. If you are eligible for academic accommodations because you have a documented disability that will impact your work in this class, please schedule an appointment with me as soon as possible to discuss your needs.

Nondiscrimination

Purdue University is committed to maintaining a community which recognizes and values the inherent worth and dignity of every person; fosters tolerance, sensitivity, understanding, and mutual respect among its members; and encourages each individual to strive to reach his or her own potential. In pursuit of its goal of academic excellence, the University seeks to develop and nurture diversity. Purdue University prohibits discrimination against any member of the University community on the basis of race, religion, color, sex, age, national origin or ancestry, genetic information, marital status, parental status, sexual orientation, gender identity and expression, disability, or status as a veteran. Any student who believes they have been discriminated against may visit www.purdue.edu/report-hate to submit a complaint to the Office of Institutional Equity. Information may be reported anonymously.

Academic Guidance in the Event a Student is Quarantined/Isolated

If you become quarantined or isolated at any point in time during the semester you will have access to an Academic Case Manager who can provide you academic support. Your Academic Case Manager can be reached at acmq@purdue.edu and will provide you with general guidelines/resources around communicating with your instructors, be available for academic support, and offer suggestions for how to be successful when learning remotely. Importantly, if you find yourself too sick to progress in the course, notify your academic case manager and notify me via email or Brightspace. We will make arrangements based on your particular situation. The Office of the Dean of Students (odos@purdue.edu) is also available to support you should this situation occur.