IE 4910/6910: Engineering Data Analytics **Fall 2025**

Tue/Thur, 11:00 AM–12:15 PM
078 Freeman Hall

Instructor

Name: Cheoljoon Jeong, Ph.D.
Office: 134 Freeman Hall
Email: cjeong@clemson.edu
OH: Thur, 1:00 PM–3:00 PM (In-person)

Teaching Assistant

Name: Yan Wu
Office: Remote (Mon) & 278 Freeman Hall (Tue)
Email: yw8@clemson.edu
OH: Mon, 1:00 PM–2:00 PM (Zoom) & Tue, 9:45 AM–10:45 AM (In-person)

Course Description

This course focuses on fundamental concepts and computational methods for modern data analytics and statistical/machine learning, covering both theoretical foundations and practical applications, with a particular emphasis on engineering. Topics include linear regression; classification methods such as logistic regression, discriminant analysis, nearest neighbor classifiers, support vector machines, decision trees, and ensemble methods; principal component analysis and other dimensionality reduction techniques; clustering techniques including k-means, hierarchical clustering, model-based methods, and spectral clustering.

Course Objectives

By the end of this course, students will be able to:

- Understand and explain the fundamental concepts of modern data analytics and statistical/machine learning, including their theoretical underpinnings and practical relevance to engineering;
- Apply appropriate supervised and unsupervised learning methods to analyze and interpret engineering data;
- Evaluate and compare models using proper statistical and computational criteria, and make informed decisions about model selection for specific engineering problems.

Required Background

Students must have a good working knowledge of linear algebra, probability, and statistics, including the ability to read, understand, and carry out mathematically rigorous arguments. Proficiency in Python is also required. Prior exposure to numerical computing, optimization, and related applications is helpful but not required.
Prerequisites: IE 3600, IE 3610.

Textbook

The primary materials for this course will be the lecture notes and slides. Recommended books include:

- *An Introduction to Statistical Learning with Python*, by Gareth James, Daniela Witten, Trevor Hastie, and Robert Tibshirani (available online)
- *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*, by Trevor Hastie, Robert Tibshirani, Jerome Friedman (available online)
- *Pattern Recognition and Machine Learning*, by Christopher Bishop (available online).
- *Linear Algebra Done Right*, by Sheldon Axler (available online)

Lectures

Attendance and engagement in lectures are expected. Students are also expected to demonstrate courtesy toward the instructor and their peers by arriving on time and silencing all electronic devices prior to the start of each lecture. When using laptops or tablets for note-taking, students must ensure that their devices do not cause distractions. Students are responsible for independently reviewing and mastering any material covered in lectures they do not attend.

Lectures will incorporate a combination of projected slides and iPad-based writing. Slides will be made available for download in advance of each lecture. These materials are intended to support effective note-taking and include figures

that are difficult to reproduce by hand. Students are expected to come prepared to take notes while following along with the slides. Annotated slides will also be posted after each class session.

Grading

Course assessments will be based on 4–5 major homework assignments, one group-based final project (3–4 students for one team), weekly or semi-weekly assignments, and one individual report. Each homework assignment will include conceptual questions, analytical problems, and coding exercises. The weekly or semi-weekly assignments will be graded based on completion. Additional details regarding the project and individual report will be provided during the semester. The distribution of the final grade is as follows:

Component	IE 4910	IE 6910
Homework	40%	35%
Project	55%	55%
Weekly Assignments	5%	5%
Individual Report		5%

Grading scale will be:

- IE 4910: 90–100 (A), 80–89 (B), 70–79 (C), 60–69 (D), <60 (F).
- IE 6910: 93–100 (A), 90–92 (A-), 87–89 (B+), 83–86 (B), 80–82 (B-), 77–79 (C+), 73–76 (C), 70–72 (C-), <70 (F).

Homework

Homework assignments will be posted approximately two weeks prior to their due dates and will collectively account for 35% or 40% of the course grade, depending on your registration status. Each assignment is due by 11:55 PM on the specified date, and all submissions must be uploaded through the course website (email submissions will not be accepted).

A 12-hour grace period will follow each deadline. During this time, students may confirm that the correct version of their work has been uploaded and, if necessary, replace it with an updated file containing minor corrections. Submissions made during the grace period will not incur a penalty. After the grace period has concluded, solutions will be posted and the assignment will be considered closed. No further submissions will be accepted under any circumstances, and no instructor questions related to that assignment will be addressed. Students are advised to plan accordingly.

Project

The tentative project due dates are as follows (last updated: August 20, 2025):

- Project Proposal: Thursday, October 9, 2025
- Milestone Report: Thursday, November 6, 2025
- Presentation: Tuesday, November 25, 2025; Tuesday, December 2, 2025; Thursday, December 4, 2025
- Final Report: Monday, December 8, 2025

Note: The schedule is subject to change, so students should check Canvas announcements regularly.

Individual Report

Students (IE 6910) will be expected to complete an individual, in-depth written assignment (1–2 pages, single spacing, 1 inch margin, times new roman, and fontsize=11–12). The assignment may focus on discussing any statistics, data science, or data analytics methodology related to the student's area of interest and should include an introduction, a review of recent literature, key challenges, and a discussion of how researchers or practitioners might address these challenges. The specific topic is at the student's discretion but must be approved by the instructor.

Submissions

All homework, project, and individual reports submissions must be typed rather than handwritten. Typed solutions are

easier to read and grade, and they allow students to review and edit their work to improve clarity, notation, explanations, and derivations. Students are strongly encouraged to learn the basics of LaTeX (via Overleaf) and to use it in preparing their submissions. A sample file and a set of predefined commands for typesetting mathematical notation will be made available on Canvas.

Regrading Policy

If you believe an error has been made in the grading of your assignments and project, you must submit a written re-grade request to the instructor **within one week of the posting of grades**. The request must include a clear explanation for the additional points being requested. Please note that the entire assignment will be regraded, which may result in **either an increase or a decrease in the overall score**.

Class Attendance and Participation

Attendance at lectures is not mandatory but is expected. However, each student is responsible for all information distributed and/or presented during lecture periods. Also, the students are authorized to leave if the instructor or his assistant is more than 15 minutes late for a lecture. However, every possible effort will be made to inform you for any classes that are canceled due to an unscheduled conflict or emergency.

Independent Work Policy

Students may use the course textbook, class notes, and solutions to prior assignments distributed during this semester as references for homework. General knowledge not explicitly discussed in the textbook (e.g., concepts from calculus, analysis, probability theory, or linear algebra) may also be cited. However, students must not consult other textbooks, journals, websites, solution manuals, or individuals outside of the course for solutions or hints to assigned problems.

Collaboration is permitted only at the level of conceptual discussion with other students currently enrolled in the course. Direct assistance may be obtained only from the instructor or teaching assistant. All submitted work—including writing, calculations, computer code, and outputs—must represent the student's own independent effort and may not be copied from or based on material produced by others. Each submission must include the names of all classmates with whom the student discussed any aspect of the assignment.

Redistribution of Course Materials

You are **NOT** permitted to share with any individuals, websites, or repositories any documents obtained from Canvas or Gradescope for this class. This includes, but is not limited to, homework assignments and their solutions, as well as any publications or manuscripts shared on Canvas (Note that the textbooks are already available on the public web). Posting such documents or files constitutes a violation of copyright law. Furthermore, although sharing your own answers to homework or any others may not constitute a copyright violation, it does represent a violation of academic integrity, as it facilitates academic dishonesty by others.

Email

When emailing the instructor, please include the following text in the subject line: "[IE4910]" or "[IE6910]" and allow for two business days to receive a response. Note that the instructor checks Canvas messages far less often than email. If you would like reach out to him, please use email rather than Canvas.

Canvas

To access the Canvas content management system for this course:

1. Open your web browser and navigate to: http://clemson.instructure.com.
2. Use your Clemson University username and password to log in.
3. The Welcome page will list your current courses.

4. Select this course and enter the canvas classroom.

If you were registered with a non-Clemson email, go to the direct URL: https://clemson.instructure.com/login/canvas to log into Canvas and access your course.

AI Policy

The use of artificial intelligence (AI) tools, such as ChatGPT, may be permitted for assignments and projects only under limited circumstances in this course. **You are strongly encouraged to write code from scratch and complete homework independently.** However, if you find it necessary to use AI tools, their use must be ethical, properly documented, and appropriately cited, and you must be able to demonstrate understanding of the content of any work submitted. Please be aware that you are accountable for the responsible, ethical use of these tools. Concerns regarding potential violations of this policy will be reviewed fairly through the appropriate university academic integrity process.

Academic Integrity

As members of the Clemson University community, we have inherited Thomas Green Clemson's vision of this institution as a "high seminary of learning." Fundamental to this vision is a mutual commitment to truthfulness, honor, and responsibility, without which we cannot earn the trust and respect of others. Furthermore, we recognize that academic dishonesty detracts from the value of a Clemson degree. Therefore, we shall not tolerate lying, cheating, or stealing in any form.

All infractions of academic dishonesty by undergraduates must be reported to Undergraduate Studies for resolution through that office. In cases of plagiarism instructors may use the Plagiarism Resolution Form. See the Undergraduate Academic Integrity Policy website for additional information and the current catalogue for the policy. For graduate students, see the current Graduate School Handbook for all policies and procedures.

Accessibility

Clemson University values the diversity of our student body as a strength and a critical component of our dynamic community. Students with disabilities or temporary injuries/conditions may require accommodations due to barriers in the structure of facilities, course design, technology used for curricular purposes, or other campus resources. Students who experience a barrier to full access to this class should let the instructor know and are encouraged to request accommodations through SAS (Student Accessibility Services) as soon as possible. To request accommodations through SAS, please see this link: https://www.clemson.edu/academics/studentaccess/register.html. You can also reach out to SAS with questions by calling 864-656-6848, visiting SAS at the ASC Suite 239, or stopping by the office as a drop-in appointment.

The Clemson University Title IX Statement Regarding Non-Discrimination

Clemson University is committed to a policy of equal opportunity for all persons and does not discriminate on the basis of race, color, religion, sex, sexual orientation, gender, pregnancy, national origin, age, disability, veteran's status, genetic information or protected activity in employment, educational programs and activities, admissions and financial aid. This includes a prohibition against sexual harassment and sexual violence as mandated by Title IX of the Education Amendments of 1972. This Title IX policy is located on the Access and Equity website. Ms. Alesia Smith is the Clemson University Title IX Coordinator, and the Assistant Vice President of Equity Compliance. Her office is located at 223 Brackett Hall, 864-656-3181 and her email address is alesias@clemson.edu. Remember, email is not a fully secured method of communication and should not be used to discuss Title IX issues.

Clemson University aspires to create a diverse community that welcomes people of different races, cultures, ages, genders, sexual orientation, religions, socioeconomic levels, political perspectives, abilities, opinions, values and experiences.

Emergency Preparation

Emergency procedures have been posted in all buildings and on all elevators. Students should be reminded to review these procedures for their own safety. All students and employees should be familiar with guidelines from Clemson University

Public Safety. Clemson University is committed to providing a safe campus environment for students, faculty, staff, and visitors. As members of the community, we encourage you to take the following actions to be better prepared in case of an emergency:

1. Familiarize yourself with all possible exits, safer locations, and other key information on the emergency evacuation maps in this building, and those that you visit regularly.
2. Make a plan for how you would Run, Hide, and Fight in case of an active threat in this building, and those that you visit regularly. For example:
    a. Run: What are all the possible exits in this building, and the routes to them?
    b. Hide: What are the potential hiding locations in this room and building that are out of sight of doors and windows, how do you lock the door(s), how would you barricade the door(s) and windows, where do you turn off the lights?
    c. Fight: What tools are available in this room and building, should you have to fight?
3. Ensure you are signed up for emergency alerts. Alerts are only sent when there is a potential threat to safety, a major disruption to campus services, and once-monthly tests.
4. Download the Rave Guardian app to your phone.
5. Learn what you can do to prepare yourself for the hazards that affect our locations.

List of Topics

1. Review of Linear Algebra.
    - Basic matrix/vector operations.
    - Euclidean vector spaces (notions of linear independence, basis, span, rank, etc).
    - SVD, Eigendecomposition, and PSD matrices; Induced and element-wise norms.
2. Linear Regression.
    - Least-squares estimation.
    - Regularized linear regression (Lasso and Ridge).
    - Cross-validation.
3. Classification.
    - Logistic regression.
    - Discriminant analysis.
    - Nearest neighbor classifiers.
    - Support vector machines (primal and dual viewpoint).
4. Nonlinear Models.
    - Splines.
    - Generalized additive model.
5. Tree-Based Methods.
    - Decision trees.
    - Ensemble methods: bagging, random forests, boosting.
6. Principal Component Analysis.
    - Dimensionality reduction.
    - Principal component analysis (PCA).
    - Different variants of PCA (robust PCA, sparse PCA, etc.).
7. Clustering.
    - K-means, hierarchical clustering, model-based methods, spectral clustering.
8. Graphical Models.
    - Bayesian networks, Markov random fields.
    - Inference in graphical models.

Note: The topics covered are subject to revision at any time. It is your responsibility stay informed of any changes.