STA552-Applied Statistical Machine Learning

Instructor: Cheng Peng

Spring 2025

Contact

- Email: cpeng@wcupa.edu
- Office: UNA 107
- Phone: 610-436-2369
- Office Hours:

Day	Time	Location
Teusday	9:30 AM - 11:30 AM	UNA107
Wednesday	4:00 PM - 5:00 PM	UNA107
Thursday	9:30 AM - 11:30 AM	UNA107
By appointment		ZOOM or UNA107

- Class Meeting:

Day	Time	Location
Wednesday	5:45 PM - 8:30 PM	UNA127

- ZOOM Link: Available on the course web page.
- Course Web Page: https://pengdsci.github.io/STA552/

Course Description

This introductory course on machine learning methodologies focuses on both algorithm design and real-world applications. While covering traditional machine learning algorithms, the course places special emphasis on those rooted in traditional statistics. Additionally, emerging algorithms addressing challenging problems—such as imbalanced learning and learning from data with partially accurate labels—will be explored. A scripting programming language will be used for case studies, assignments, and term projects.

Learning Objectives and Outcomes

- **Course Objectives**

  The objectives of this course are to equip students with basic skills and strategies to

  - Understand foundational concepts and technical underpinnings of machine learning algorithms and key concepts such as overfitting, bias-variance tradeoff, and regularization.
  - Develop practical skills to implement common machine learning algorithms, including regression, classification, clustering, and dimensionality reduction.

  - Apply machine learning techniques by applying supervised and unsupervised learning methods to solve real-world problems and evaluate model performance using appropriate metrics and validation techniques.
  - Handle complex datasets through preprocessing, cleaning, transforming raw data for effective analysis, and manage high-dimensional data and apply feature engineering techniques.
  - Understand theoretical and practical tradeoffs between computational efficiency, interpretability, and predictive power of different models.
  - Communicate results effectively by visualizing data and model outcomes to support decision-making and presenting findings in a clear and accessible manner to diverse audiences.
  - Engage with advanced topics such as ensemble learning, deep learning, or probabilistic models and gain insight into recent developments and emerging trends in statistical machine learning.

- **Learning Outcomes**

  After finishing this course, students will be able to

  - Identify real-world problems that align with canonical machine learning tasks.
  - Develop features using heuristic methods and model-based approaches.
  - Design and implement effective solutions for prediction or classification tasks.
  - Create and apply basic clustering and dimensionality reduction algorithms.
  - Evaluate predictive models using appropriate KPI metrics and compare various evaluation methods.
  - Develop and implement strategies for data preprocessing, partitioning datasets into training and validation sets, and tuning hyperparameters.

Logistics and Required Materials

- **No textbook** is required for this class. Some of the optional texts are listed below:
  - *An Introduction to Statistical Learning*, by Daniela Witten, Gareth James, Robert Tibshirani, and Trevor Hastie, Springer, 2013. Free Ebook available at https://www.statlearning.com/
  - *Elements of Learning*, by Jerome H. Friedman, Robert Tibshirani, and Trevor Hastie, Springer, 2001. Free Ebook at https://hastie.su.domains/ElemStatLearn/
  - *Understanding Machine Learning: From Theory to Algorithms*, Shai Shalev-Shwartz, Shai Ben-David, Cambridge University Press, 2014.
  - *An Introduction to Machine Learning*, by Ethem Alpaydın, MIT Press, 2004.
  - *Pattern Recognition and Machine Learning*, by Chris Bishop, Springer, 2006.
  - *Introduction to Machine Learning*, by Alex Smola and S.V.N. Vishwanathan, Cambridge University Press, 2018.
  - *Machine Learning: A Probabilistic Perspective*, by Kevin P. Murphy, MIT Press, 2012.
  - *Machine Learning*, by Tom Mitchell, McGraw-Hill, 1997.
- **Class Notes:** Class notes will be provided.
- **Programming and Technical Tools:**
  - Programming languages and Software: R (primary) or Python (acceptable)
  - Platforms: RStudio, Anaconda, Github
  - Computation Notebooks: RMarkdown (R Notebook), Jupyter Notebook, Quarto (for R and Python)
  - Relevant Markup Languages: Markdown, LaTex, HTML, CSS.

- **Coverage:** See the list of tentative topics

Assessments

- **Assessment Components**: The course grade consists of the following components:
  - Assignments (50%):
  - Class attendance and participation (10%):
  - Term project report (40%):
- **Grade Scales**: The final course grade will be calculated based on the above components. A letter grade will be assigned according to the following scale.

Grade	Quality Points	Percentage Equivalents	Interpretation
A	4.00	[93%, 100%]	Superior graduate attainment
A-	3.67	[90%, 93%)	
B+	3.33	[86%, 90%)	Satisfactory graduate attainment
B	3.00	[83%, 86%)	
B-	2.67	[80%, 83%)	
C+	2.33	[76%, 80%)	Attainment below graduate expectations
C	2.00	[73%, 76%)	
C-	1.67	[70%, 73%)	
F	0	$< 70\%$	Failure

D grades are not used. Refer to the Graduate Catalog for the description of NG (No Grade), W, & other grades.

Class Policies

- **Attendance and Participation**

  Attendance in the class is mandatory. Actively participating in class discussion is required in this class and is one of the components of the final course grade.

- **Late Homework and Assignments**

  Late assignments will be accepted. However, all late assignments will be subject to a small penalty deduction.

Tentative Topics

The following topics are subject to adjustment. I am also open to suggestions for additional topics relevant to machine learning that are of practical importance for the class.

1. Platforms and tools for machine learning
2. Visual analytics in machine learning
3. Data preparation for ML - multiple imputation
4. Feature engineering for machine learning
5. Methods of cross-validation for regression and classification
6. Overview of machine learning algorithms
7. Linear and nonlinear support vector machines (SVM)
8. Ensemble Learning - Random Forest

9. Neural networks algorithms and applications
10. Linear and kernel principal component analysis
11. Algorithmic-based clustering
12. Gaussian-mixture model-based clustering method
13. Methods for imbalance learning
14. Smoothing methods in machine learning

University Policies and Resources

Policy Statements

- **ACADEMIC & PERSONAL INTEGRITY**: It is the responsibility of each student to adhere to the university's standards for academic integrity. Violations of academic integrity include any act that violates the rights of another student in academic work, that involves misrepresentation of your own work, or that disrupts the instruction of the course. Other violations include (but are not limited to): cheating on assignments or examinations; and plagiarizing, which means copying any part of another's work and/or using ideas of another and presenting them as one's own without giving proper credit to the source; selling, purchasing, or exchanging of term papers; falsifying of information; and using your own work from one class to fulfill the assignment for another class without significant modification. Proof of academic misconduct can result in automatic failure and removal from this course. For questions regarding Academic Integrity, the No-Grade Policy, Sexual Harassment, or the Student Code of Conduct, students are encouraged to refer to the Department Graduate Handbook, the Graduate Catalog, the Ram's Eye View, and the University website at www.wcupa.edu.
- **STUDENTS WITH DISABILITIES**: If you have a disability that requires accommodations under the Americans with Disabilities Act (ADA), please present your letter of accommodations and meet with me as soon as possible so that I can support your success in an informed manner. Accommodations cannot be granted retroactively. If you would like to know more about West Chester University's Services for Students with Disabilities (OSSD), please visit them at 223 Lawrence Center. The OSSD hours of Operation are Monday – Friday, 8:30 a.m. – 4:30 p.m. Their phone number is 610-436-2564, their fax number is 610-436-2600, their email address is ossd@wcupa.edu, and their website is at www.wcupa.edu/ussss/ossd.
- **EXCUSED ABSENCES POLICY**: Students are advised to carefully read and comply with the excused absences policy, including absences for university-sanctioned events, contained in the WCU Graduate Catalog. In particular, please note that the "responsibility for meeting academic requirements rests with the student," that this policy does not excuse students from completing required academic work, and that professors can require a "fair alternative" to attendance on those days that students must be absent from class in order to participate in a University-Sanctioned Event.
- **REPORTING INCIDENTS OF SEXUAL VIOLENCE**: West Chester University and its faculty are committed to assuring a safe and productive educational environment for all students. In order to comply with the requirements of Title IX of the Education Amendments of 1972 and the University's commitment to offering supportive measures in accordance with the new regulations issued under Title IX, the University requires faculty members to report incidents of sexual violence shared by students to the University's Title IX Coordinator. The only exceptions to the faculty member's reporting obligation are when incidents of sexual violence are communicated by a student during a classroom discussion, in a writing assignment for a class, or as part of a University-approved research project. Faculty members are obligated to report sexual violence or any other abuse of a student who was, or is, a child (a person under 18 years of age) when the abuse allegedly occurred to the person designated in the University Protection of Minors Policy. Information regarding the reporting of sexual violence and the resources that are available to victims of sexual violence is set forth at: https://www.wcupa.edu/_admin/diversityEquityInclusion/sexualMisconduct/default.aspx

- **INCLUSIVE LEARNING ENVIRONMENT AND ANTI-RACE STATEMENT**: Diversity, equity, and inclusion are central to West Chester University's mission as reflected in our Mission Statement, Values Statement, Vision Statement and Strategic Plan: Pathways to Student Success. We disavow racism and all actions that silence, threaten, or degrade historically marginalized groups in the U.S. We acknowledge that all members of this learning community may experience harm stemming from forms of oppression including but not limited to classism, ableism, heterosexism, sexism, Islamophobia, anti-Semitism, and xenophobia, and recognize that these forms of oppression are compounded by racism.

  Our core commitment as an institution of higher education shapes our expectation for behavior within this learning community, which represents diverse individual beliefs, backgrounds, and experiences. Courteous and respectful behavior, interactions, and responses are expected from all members of the University. We must work together to make this a safe and productive learning environment for everyone. Part of this work is recognizing how race and other aspects of who we are shape our beliefs and our experiences as individuals. It is not enough to condemn acts of racism. For real, sustainable change, we must stand together as a diverse coalition against racism and oppression of any form, anywhere, at any time.

  Resources for education and action are available through WCU's Office for Diversity, Equity, and Inclusion (ODEI), DEI committees within departments or colleges, the student ombudsperson, and centers on campus committed to doing this work (e.g., Dowdy Multicultural Center, Center for Women and Gender Equity, and the Center for Trans and Queer Advocacy).

  Guidance on how to report incidents of discrimination and harassment is available at the University's Office of Diversity, Equity and Inclusion.

- **EMERGENCY PREPAREDNESS**: All students are encouraged to sign up for the University's free WCU ALERT service, which delivers official WCU emergency text messages directly to your cell phone. For more information, visit www.wcupa.edu/wcualert. To report an emergency, call the Department of Public Safety at 610-436-3311.

- **ELECTRONIC MAIL POLICY**: It is expected that faculty, staff, and students activate and maintain regular access to University-provided e-mail accounts. Official university communications, including those from your instructor, will be sent through your university e-mail account. You are responsible for accessing that mail to be sure to obtain official University communications. Failure to access will not exempt individuals from the responsibilities associated with this course.

Resources

- **COMMUNICATION VIA NAVIGATE**
  - West Chester University uses Navigate as a communication tool between faculty, students, and student support offices. Throughout the term, you may receive text messages and/or WCU emails from Navigate regarding your academic progress. These messages will have information to promote your success and direct you to support services that will best fit your needs—please open and read them. Additionally, Navigate has useful features such as your course schedule, campus resources with contact and location information, and a feature called Study Buddies that lets you connect with other students in your class sections who are interested in studying together outside of class.
  - To access Navigate (http://www.wcupa.edu/Navigate), you will need to download the Student App to your phone by visiting the Navigate website (use the QR code).
  - For technical assistance visit the Help Desk (https://www.wcupa.edu/infoServices/serviceNow/default.aspx). For additional Navigate documentation, visit the Navigate Training Site (https://www.wcupa.edu/academicEnterpriseSystems/training/navigateTraining/student.aspx).
- **TECHNICAL SUPPORT**
  - The IS&T Help Desk (https://www.wcupa.edu/infoServices/serviceNow/) is the central point of contact for all technology-related hardware, software, and system support such as Office 365 and D2L at West Chester University. Located in Anderson Hall 119, the Help Desk provides first-level customer support for students and employees in locations such as administrative offices, classrooms, residence halls, remote sites & satellite campuses using ServiceNow.

  - Submit a support ticket: Login to ServiceNow (https://wcupaprod.service-now.com/sp) Phone: 610-436-3350
  - Hours for support may vary. Check the website for times of operation.
- **ACADEMIC SUPPORT**
  - The Academic Support Resources website (https://www.wcupa.edu/universityCollege/academicSupportServices.aspx) is a good place to start looking for academic support services.
  - The Learning Assistance and Resource Center (https://www.wcupa.edu/universityCollege/larc/default.aspx) provides tutoring services for free, but availability is limited.
  - WCU Library (https://library.wcupa.edu/home) resources are available remotely.
  - You are strongly encouraged to review the Navigating Digital Learning (https://d2l.wcupa.edu/d2l/home/2513024) Orientation D2L module (in the Orientation tab of your My Courses page).
- **STUDENT SUPPORT**
  - The WCU Student Support Center (https://www.wcupa.edu/Support/) is a good place to start looking for support services.
  - WCU Counseling & Psychological Services (https://www.wcupa.edu/_services/counselingCenter/) offers counseling, referrals, and emergency resources.
  - In the event of a health-related absence, the Health and Wellness Student Assistance webpage (https://www.wcupa.edu/_services/STU/health-and-wellness/student-assistance.aspx) can help you notify instructors.
  - Need a Laptop? Loaner equipment may be available; see the Laptop and Equipment Loaner Program (https://library.wcupa.edu/techhelp/computers).
- **APSCUF**
  - APSCUF (The Association of Pennsylvania State College and University Faculties) is the union that represents the faculty and coaches who have devoted themselves to providing quality higher education for Pennsylvania students at the ten universities in the Pennsylvania State System of Higher Education, including West Chester University. To learn more about APSCUF, see www.apscuf.org.
- **WRITING CENTER**
  - The West Chester University Writing Center provides tutoring **free of charge** for any student on any writing project. To learn more about the WCU Writing Center and to make an appointment, go to www.wcupa.edu/writingcenter.