DATA SCIENCE: MACHINE LEARNING
CCJ 5934/SYA7933, Spring 2026
Monday 8:30 AM - 11:30 AM
Location: TUR 2350

Instructor: Dr. Nan Li
E-Mail: nan.li@ufl.edu
Office: Turlington Hall Room 3358
Office Hours: Monday & Wednesday 2:00 PM – 4:00 PM, or by appointment

Course Overview

Welcome to Machine Learning for Criminology and Sociology! This graduate-level course introduces you to machine learning as a research method for social science. Rather than approaching machine learning as a purely technical or engineering topic, this course treats it as an analytic framework for understanding social patterns, institutional dynamics, and structural inequality using large and complex datasets.

You will learn how supervised learning, unsupervised learning, semi-supervised learning, and text mining can be applied in criminology and sociology research. Throughout the semester, you will practice connecting theory to modeling decisions, critically evaluating model outputs, and reflecting on ethical and fairness issues surrounding algorithmic systems in social science research. By the end of the course, you will be able to design and implement an independent machine-learning-based research project at the graduate level.

Required Readings

There is no single required textbook. Weekly readings will be posted on Canvas.

Grading

1. **Attendance and Participation – 10%**
   This is a discussion- and lab-based course. Active participation is expected in lectures, discussions, coding exercises, and peer feedback. Regular attendance is required.
2. **Assignments / Labs – 30%**
   You will complete applied assignments throughout the semester. These may include data preprocessing, supervised classification, clustering exercises, HMM modeling, and text analysis.
3. **Final Research Project – 40%**
   You will design and complete an original research project using one or more machine learning techniques covered in this course.
4. **Presentation – 20%**
   You will present your final project during the final weeks of the semester.

Grade scale

A 93.00-100.00	B 83.00-86.99	C 73.00-76.99	D 63.00-66.99

A- 90.00-92.99	B- 80.00-82.99	C- 70.00-72.99	D- 60.00-62.99
B+ 87.00-89.99	C+ 77.00-79.99	D+ 67.00-69.99	F Below 60.00

University Policies

Accessibility

Students with disabilities requesting accommodations should first register with the Disability Resource Center (352-392-8565, https://disability.ufl.edu/) by providing appropriate documentation. Once registered, students will receive an accommodation letter which must be presented to the instructor when requesting accommodations. Students with disabilities need to follow this procedure as early as possible in the semester.

Course Communications

All class activities will take place during our assigned class time. You are expected to be available during those times. Since this class is a face-to-face course, you are expected to stay in Gainesville throughout this semester. You are responsible for all information made available through classes and Canvas. Students are encouraged to communicate with the instructor anytime if any concerns.

Course Evaluation

Students are expected to provide professional and respectful feedback on the quality of instruction in this course by completing course evaluations online via GatorEvals. Guidance on how to give feedback in a professional and respectful manner is available at https://gatorevals.aa.ufl.edu/students/. Students will be notified when the evaluation period opens, and can complete evaluations through the email they receive from GatorEvals, in their Canvas course menu under 2 GatorEvals, or via https://ufl.bluera.com/ufl/. Summaries of course evaluation results are available to students at https://gatorevals.aa.ufl.edu/public-results/.

Recording

Students are allowed to record the class only for 1) personal educational use, 2) complaints to the university, and 3) criminal or civil proceedings. All other purposes are prohibited. This class will not provide a recorded lecture for students who miss a class.

University Honesty policy

University Honesty Policy UF students are bound by The Honor Pledge, which states, "We, the members of the University of Florida community, pledge to hold ourselves and our peers to the highest standards of honor and integrity by abiding by the Honor Code. On all work submitted for credit by students at the University of Florida, the following pledge is either required or implied: On my honor, I have neither given nor received unauthorized aid in doing this assignment." The Honor Code (https://www.dso.ufl.edu/sccr/process/studentconduct-honor-code/) specifies the number of behaviors that violate this code and the possible sanctions. Furthermore, you are obligated to report any condition that facilitates academic misconduct to appropriate personnel. If you have any questions or concerns, please consult with the instructor in this class.

COURSE SCHEDULE

Topic	**Date**	**Assignment**
Week 1: Introduction to Machine Learning	Jan 12	
Holiday, no class	Jan 19	
Week 3: Regression as Supervised Learning	Jan 26	
Week 4: Decision Trees	Feb 2	Assignment 1 due
Week 5: Random Forests	Feb 9	
Week 6: Clustering Methods	Feb 16	Assignment 2 due
Week 7: Principal Component Analysis (PCA)	Feb 23	
Week 8: Introduction to Semi-Supervised Learning	Mar 2	Assignment 3 due
Week 9: Hidden Markov Models	Mar 9	
Spring break, no class	Mar 16	
Week 11: HMM Validation and Interpretation	Mar 23	Assignment 4 due
Week 12: Text as Data	Mar 30	
Week 13: Topic Modeling	Apr 6	Assignment 5 due
Week 14: Text Classification	Apr 13	
Week 15: Student Project Presentations	Apr 20	Assignment 6 due
	Apr 27	Final project due