

# PA 470 - ML/AI for the Public Sector - Syllabus

**Instructor:** Divij Sinha

**Email:** [divij@uic.edu](mailto:divij@uic.edu)

**Venue:** Room 2236

**Office Hours:** 5:30pm to 6:30pm, Thursdays (Before the lecture)

**Course Website/GitHub:** <https://github.com/uic-cuppa-pa470-spring2025/course-website>

## Course Description:

This course will explore the applications of machine learning (ML) and artificial intelligence (AI), particularly focusing on their use within the public sector. We will cover fundamental ML concepts, the current state of AI (including Large Language Models or LLMs), and the policy implications of these technologies. A significant portion of the course will involve practical application using the R programming language, culminating in a final project.

## Course Objectives:

This course will take a bi-directional approach to ML/AI, focusing on both theory and practice. You will be able to produce ML models, as well as consume and critique them.

Upon completion of this course, students will be able to:

1. Understand the fundamental concepts of machine learning.
2. Grasp the current landscape of AI, including LLMs and their capabilities.
3. Analyze the policy and ethical implications of AI in public administration.
4. Utilize the R programming language for data analysis and ML model building.
5. Apply R programming skills relevant to public sector data challenges.
6. Develop simple machine learning models for real-world problems.
7. Effectively use LLMs for relevant tasks.

## Required Materials:

- Access to a computer with internet connectivity.
- RStudio installation (R <https://cran.rstudio.com> ; RStudio <https://posit.co/download/rstudio-desktop/>).
- GitHub account.
- (Optional but recommended) Textbooks or online resources on introductory statistics and machine learning.
  - R for Data Science (2e), Hadley Wickham, Mine Çetinkaya-Rundel & Garrett Grolmund <https://r4ds.hadley.nz>
  - Data Science for Public Policy, Jeffrey C. Chen , Edward A. Rubin , Gary J. Cornwall <https://link.springer.com/book/10.1007/978-3-030-71352-2>

## Course Structure:

Each week's lecture will be split into consist of two ~75 minute sessions:

- **Part 1:** “Lab” sessions involving code-along exercises, statistical concepts, ML fundamentals, and AI theory.
- **Break:** A 10-minute break between sessions.
- **Part 2:** Discussions on real-world applications of AI/ML in the public sector, emphasizing practical implications and ethical considerations. There will be a 10-minute break between sessions.

## Course Topics:

- Week 1 - January 16:
  - Introduction to course + R programming + github + modelling + AI/ML in the public sector
- Week 2 - January 23:
  - Part 1: Tidyverse + Data Wrangling (`tidyverse`)
  - Part 2: Can AI solve a problem?
  - *Assignment 1 Due*
- Week 3 - January 30:
  - Part 1: Data Visualization (`ggplot2`)
  - Part 2: Should AI solve a problem?
  - *Assignment 2 Due*
- Week 4 - February 6:
  - Part 1: Census data + Public Data (`tidycensus`)
  - Part 2: How does involving AI change the problem being asked?
- Week 5 - February 13:
  - Part 1: Linear Regression (`lm`, `glm`, `tidymodels`)
  - Part 2: How to convince others AI should (or shouldn't) solve a problem?
  - **Proposal for Final Project Due**
- Week 6 - February 20:
  - Part 1: Classification ()
  - Part 2: Is my AI solution solving the right problem?
  - *Assignment 3 Due*
- Week 7 - February 27:
  - Part 1: Clustering ()

- Part 2: How is AI currently being used in and around the public sector?
- Week 8 - March 6:
  - Part 1: Deeper Dive into ML in R ()
  - Part 2: How is the government thinking about AI?
  - *Assignment 4 Due*
- Week 9 - March 13:
  - Part 1: Text Analysis ()
  - Part 2: Should AI exist?
  - **Update 1 on Final Project Due**
- Week 10 - March 20:
  - Part 1: Image Analysis ()
  - Part 2: Who owns AI?
- Week 11 - March 27: **SPRING BREAK**
- Week 12 - April 3:
  - Part 1: Model Evaluation ()
  - Part 2: Should I believe AI?
  - *Assignment 5 Due*
- Week 13 - April 10:
  - Part 1: Model Deployment ()
  - Part 2: AI is out and about
  - *Assignment 6 Due*
- Week 14 - April 17:
  - Part 1: Fundamentals of Deep Learning ()
  - Part 2: What are the tangential impacts of AI?
  - **Update 2 on Final Project Due**
- Week 15 - April 24:
  - Part 1: LLMs (+ others) in practice ()
  - Part 2: What is the future of AI?
  - *Assignment 7 Due*
- Week 16 - May 1:
  - Final Project Presentations

### **Assessment:**

Your final grade will be calculated as follows:

- **Assignments (30%):** These will consist of coding and writing assignments. Code will be graded on effort, thoughtfulness, and approach rather than perfect correctness. Writing will be evaluated based on argument strength, clarity, and understanding of the subject matter, not solely on language mechanics. Allowed use of tools like ChatGPT will be on a per question/assignment basis. Lowest assignment grade will be dropped. Assignments will always be released right after class, you will have one or two weeks to complete them, depending on the complexity. The final project will be due at the end of the semester, with a presentation in the last week.
- **Final Project (45%):** This project will involve applying ML/AI techniques in R to a real-world problem relevant to the public sector. The project can be a proof of concept or a fully functional application. Grading will consider the ambition and effort invested, and we will discuss project scoping throughout the semester. A proposal is due in Week 5, with updates in Weeks 9 and 14. The final submission must also include a critical analysis of your project and its limitations.
- **Engagement (25%):** Active participation in class discussions is essential. The classroom is structured as a dialogue; those comfortable should speak up! If you feel like you are falling behind in grades elsewhere, this is a good place to make up for it. Participation will be graded on the quality of your contributions. This includes asking questions, answering questions, and engaging with your peers. I will be happy to provide extra credit for those who go above and beyond in their participation. If possible, please let me know if you will be absent.

For the first late submission, no penalty will be applied. For subsequent late submissions, please try to ask me for an extension at least 24 hours before the deadline. Extensions will be granted on a case-by-case basis. I will be quite lenient, but please do not abuse this policy.

### **Code Management:**

All code will be managed through GitHub. I will share my code on GitHub, and you will submit your assignments through the same platform. You will need RStudio and GitHub; no other software is required. For Windows users, using the Windows Subsystem for Linux (WSL) or Git SCM for Windows is recommended.

**Schedule:** (A detailed weekly schedule with topics, readings, and assignments will be posted on GitHub.)

**Academic Integrity:** This is a tricky statement to write for a class around AI/ML. I will trust you to use the resources you need to complete your assignments. However, you must cite any code or resources you use. If you are unsure about what constitutes plagiarism, please ask me. I will be happy to help you understand what is and is not acceptable. If you are caught cheating, you will receive a 0 on the assignment and be reported to the university. If you are unsure about what constitutes cheating, please ask me. I will be happy to help you understand what is and is not acceptable. Working together is encouraged, but you must write your code

and assignments independently. University policies on academic integrity will supersede any course policies.

**Disability Accommodations:** Please reach out to me/your advisor/UIC services in case you need any accommodations. I will do my best to accommodate your needs.