

# Machine Learning I

## DATS 6202 - 10, Spring 2024

### 1 Meeting Time and Location

- Meeting time: Tuesday, 12:45 PM - 3:15 PM
- Location: Corcoran Hall 104

### 2 Instructor

- Yuxiao (James) Huang
  - Email: [yuxiaohuang@gwu.edu](mailto:yuxiaohuang@gwu.edu)
  - Github: <https://github.com/yuxiaohuang>
  - Website: <https://sites.google.com/view/yuxiaohuang>
  - Office address: <https://gwu.webex.com/join/yuxiaohuang>
  - Office hours:
    - \* Friday, 12:00 PM - 2:00 PM
    - \* Note: If you would like to meet during my office hours, please send email (including your Webex meeting room address) to set up an appointment, so that we can have a scheduled time slot for one-on-one meeting.

### 3 Teaching Assistant

- Ei Tanaka
  - Email: [ei.tanaka@email.gwu.edu](mailto:ei.tanaka@email.gwu.edu)
  - Office address: <http://gwstudent.webex.com/meet/ei.tanaka>
  - Office hours:
    - \* Monday, 4:00 PM to 6:00 PM
    - \* Tuesday, 3:30 PM to 5:30 PM
    - \* Wednesday, 4:00 PM to 6:00 PM
    - \* Thursday, 4:00 PM to 6:00 PM

### 4 Course Prerequisites

- DATS 6101
- DATS 6103

## 5 Course Description

- In this course we will discuss the idea, practice and math of popular Machine Learning methods.
- While we will dive deep into the math behind some shallow and deep models, the real focus of this course is to teach students how to use popular Machine Learning tools to solve real-world problems.
- In this course we will use Jupyter Notebook for coding and Google Colaboratory for running the code.

## 6 Learning Outcomes

As a result of completing this course, students will be able to:

- understand the idea, practice and math of popular machine learning methods
- use popular Machine Learning tools to solve real-world problems

## 7 Course Github Repository

The syllabus, slides, code examples and case studies can be found in the course github repository.

## 8 Textbook

The following publicly available books are highly recommended but not required:

HML Hands-On Machine Learning with Scikit-Learn, Keras, and TensorFlow (2nd Edition)

ESL The Elements of Statistical Learning (2nd Edition)

NND Neural Network Design (2nd Edition)

## 9 Average Minimum Amount of Out-Of-Class or Independent Learning Expected Per Week

- Going over the math and coding covered in class is integral for success in this course.
- You should spend at least 5 hours of out-of-class or independent learning per week.

## 10 Homework

- There will be 6 homework, which will only include coding questions.
- Each homework has 8 points (hence 48 points in total).
- Homework **must** be completed individually.

## 11 Exam

- There will be 2 exams (a Take-Home Midterm and Take-Home Final), which will only include coding questions.
- Each exam has 15 points (hence 30 points in total).

## 12 Final Project

- The final project is a great opportunity for you to apply what you have learned in class to solve real-world problems.
- While each team can choose a problem of their interest, you are strongly encouraged to work on Kaggle Competitions. The bottom line is, you **must** use real-world data. Please talk to the instructor if you are not sure about the nature of the data.
- The final project should be completed by teams with no more than three students.

### 12.1 Deliverables

- Report (a ipynb file should be submitted to blackboard), with 12 points

### 12.2 Report

The report should be in ipynb form. It should include:

- Title
- Introduction (including the problem and motivation)
- Experiment (including the code and the discussion of empirical results)
- Conclusions

### 12.3 Presentation

- Each team will present their final project.
- A presentation should be no longer than 10 minutes (and no shorter than 8 minutes).
- All team members should present.

## 13 Submission

- Homework, exams, final project report and video will be due for submission through blackboard by 11:59 PM (Eastern time). See more details in sec 21.
- **Submission will no longer be accepted after the deadline, and will receive a grade of 0.**

## 14 Grading Scheme

- 48% Homework (6)
- 22% Final project (1)
  - 12% Report
  - 10% Presentation
- 30% Exams
  - 15% Midterm
  - 15% Final

## 15 Grading Rubrics

We will apply the following grading rubrics for each Homework and Exam.

- If only part of the pipeline is implemented, the student shall receive 60% of the full mark (assigned to the homework or exam).
- If the whole pipeline is implemented:

- if the implementation is wrong (so that either no results can be generated or the generated results are wrong), the student shall receive 70% of the full mark
- if the implementation is correct (so that results can be generated and the generated results are correct):
  - \* if the validation / test score is lower than or the same as the baseline (i.e., a validation / test score suggested in the write-up), the student shall receive 80% of the full mark
  - \* if the validation / test score is higher than the baseline, the student shall receive 100% of the full mark

## 16 Grade Appeals

- A grade becomes permanent one week after you receive the grade.
- Grade appeals and questions must be raised in writing (email) within one week after the day on which the grade was received.

## 17 Letter Grade Distribution

[93, 100]	A
[90, 93)	A-
(87, 90)	B+
[83, 87]	B
[80, 83)	B-
(77, 80)	C+
[73, 77]	C
[70, 73)	C-
<70	F

## 18 University Policies

### 18.1 University Policy on Observance of Religious Holidays

In accordance with University policy, students should notify faculty during the first week of the semester of their intention to be absent from class on their day(s) of religious observance. For details and policy, see: <https://provost.gwu.edu/policies-procedures-and-guidelines>.

### 18.2 Academic Integrity Code

Academic Integrity is an integral part of the educational process, and GW takes these matters very seriously. Violations of academic integrity occur when students fail to cite research sources properly, engage in unauthorized collaboration, falsify data, and in other ways outlined in the Code of Academic Integrity. Students accused of academic integrity violations should contact the Office of Academic Integrity to learn more about their rights and options in the process. Outcomes can range from failure of assignment to expulsion from the University, including a transcript notation. The Office of Academic Integrity maintains a permanent record of the violation.

More information is available from the Office of Academic Integrity at <https://studentconduct.gwu.edu/academic-integrity>. The University's "Guide of Academic Integrity in Online Learning Environments" is available at <https://studentconduct.gwu.edu/guide-academic-integrity-online-learning-environments>. Contact information: [rights@gwu.edu](mailto:rights@gwu.edu) or 202-994-6757.

## 19 Support for Students Outside the Classroom

### 19.1 Virtual Academic Support

- A full range of academic support is offered virtually in fall 2020. See <https://coronavirus.gwu.edu/top-faqs> for updates.
- Tutoring and course review sessions are offered through Academic Commons in an online format. See <https://academiccommons.gwu.edu/tutoring>.
- Writing and research consultations are available online. See <https://academiccommons.gwu.edu/writing-research-help>.
- Coaching, offered through the Office of Student Success, is available in a virtual format. See <https://studentsuccess.gwu.edu/academic-program-support>.
- Academic Commons offers several short videos addressing different virtual learning strategies for the unique circumstances of the fall 2020 semester. See <https://academiccommons.gwu.edu/study-skills>. They also offer a variety of live virtual workshops to equip students with the tools they need to succeed in a virtual environment. See [https://library.gwu.edu/events?order=DATE\\_ASC&format=workshop&open\\_to=GWorld&series=&category=&sponsor=&events\\_date\\_start=&events\\_date\\_end=&terms=&page=1](https://library.gwu.edu/events?order=DATE_ASC&format=workshop&open_to=GWorld&series=&category=&sponsor=&events_date_start=&events_date_end=&terms=&page=1).

### 19.2 Writing Center

GW's Writing Center cultivates confident writers in the University community by facilitating collaborative, critical, and inclusive conversations at all stages of the writing process. Working alongside peer mentors, writers develop strategies to write independently in academic and public settings. Appointments can be booked online. See <https://gwu.mywconline.com>.

### 19.3 Academic Commons

Academic Commons provides tutoring and other academic support resources to students in many courses. Students can schedule virtual one-on-one appointments or attend virtual drop-in sessions. Students may schedule an appointment, review the tutoring schedule, or access other academic support resources at <https://academiccommons.gwu.edu>. For assistance contact [academiccommons@gwu.edu](mailto:academiccommons@gwu.edu).

### 19.4 Disability Support Services (DSS) 202-994-8250

Any student who may need an accommodation based on the potential impact of a disability should contact <https://disabilitysupport.gwu.edu> to establish eligibility and to coordinate reasonable accommodations.

### 19.5 Counseling and Psychological Services 202-994-5300

GW's Colonial Health Center offers counseling and psychological services, supporting mental health and personal development by collaborating directly with students to overcome challenges and difficulties that may interfere with academic, emotional, and personal success. See <https://healthcenter.gwu.edu/counseling-and-psychological-services>.

## 20 Safety and Security

- In an emergency: call GWPd 202-994-6111 or 911.
- For situation-specific actions: review the Emergency Response Handbook at <https://safety.gwu.edu/emergency-response-handbook>.

- In an active violence situation: Get Out, Hide Out or Take Out. See <https://www.youtube.com/watch?v=CpBT6tAa0dY&feature=youtu.be>.
- Stay informed: <https://safety.gwu.edu/stay-informed>.

## 21 Tentative Schedule

Date	Topic	Release	Due
01/16	<a href="#">Data Preprocessing</a>		
01/23	<a href="#">Introduction to Machine Learning</a> <a href="#">Data Preprocessing</a>		
01/30	<a href="#">Linear Regression</a>		
02/06	<a href="#">Linear Regression</a> <a href="#">Training Shallow Models</a>		
02/13	<a href="#">Training Shallow Models</a>	Homework 1	
02/20	<a href="#">Logistic Regression</a>		Homework 1
02/27	<a href="#">Shallow Neural Networks</a>	Homework 2	
03/05	<a href="#">Tree Based Models</a>	Homework 3 Midterm Exam	Homework 2
03/12	Spring Break (No Class)		
03/19	<a href="#">Training Deep Neural Networks</a>	Homework 4	
03/26	<a href="#">Deep Neural Networks</a> <a href="#">Training Deep Neural Networks</a>		Homework 3 Midterm Exam
04/02	<a href="#">Convolutional Neural Networks</a> (Computer Vision)		Homework 4
04/09	<a href="#">Convolutional Neural Networks</a> (Computer Vision)	Homework 5	
04/16	<a href="#">Recurrent Neural Networks</a> (Natural Language Processing)		Homework 5
04/23	<a href="#">Recurrent Neural Networks</a> (Natural Language Processing)	Homework 6	
04/30	Final project presentation	Final Exam	Homework 6 Final project report
05/07	No Class		Final exam