

# Machine Learning I

## DATS 6202 - 10, Fall 2018

### 1 Meeting Time and Location

- Meeting time: Tuesday, 6:10 PM - 8:40 PM
- Location: 1957 E Street 212

### 2 Instructor

- Name: Yuxiao Huang
- Email: yuxiaohuang@gwu.edu
- Office address: Samson Hall, Room 314
- Office hours: Wednesday, 1:00 PM - 3:00 PM
- Note: If you cannot make my scheduled office hours and need to talk outside of class, please send email to set up an appointment. I will try to respond within 24 hours. Please be aware that I may be unable to answer emails about Homework and Final project before the deadline, if they are received less than 24 hours before they are due.

### 3 Teaching Assistant

- Name: Purvi Jayant Thakor
- Email: purvithakor@gwmail.gwu.edu
- Office address: Samson Hall, Room 304 (the conference room)
- Office hours: Monday / Tuesday 4:30 PM - 5:30 PM

### 4 Course Description

- This course is an introduction to machine learning and its core models and algorithms for students in the Data Science Program
- The objective of the course is to provide students an overview of machine learning techniques to explore, analyze, and leverage data
- This course cover: supervised, unsupervised, and reinforcement learning
- This course will use `Python` to demonstrate machine learning technics
- Students will also use `Python` to complete the Homework and Final project through the course

## 5 Learning Outcomes

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

- understand machine learning algorithms and concepts
- demonstrate machine learning technics using **Python**
- train, test and validate machine learning models
- understand the core concept of supervise, unsupervised, and reinforcement learning
- apply the machine learning concepts to real world problems

## 6 Textbook

The following book is recommended but not required:

- Raschka S. and Mirjalili V. (2017). *Python Machine Learning. 2nd Edition.*

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

- Going over key concepts and doing lots of problems, beyond what is assigned 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

## 8 Homework

- There will be 6 Homework assignments, which will only include programming questions
- Homework **must** be completed individually

## 9 Quiz

- There will be 6 Quizzes, which will only include written questions

## 10 Exam

- There will be 2 exams (midterm and final), which will only include programming questions

## 11 Final Project

The Final project is a good opportunity for you to apply machine learning methods to complex, real-world problems. It will be completed by teams of 1, 2, or 3 students. Each team can choose a problem in the domain of their interest.

### 11.1 Deliverables

- Project proposal
- Data and Code
- Final report
- Presentation

### 11.2 Proposal

The project proposal is 1-page maximum. It should include:

- The title of the project
- The problem definition and motivation
- The proposed method, language, and package you will need for the implementation
- The link to the data
- The responsibility of each team member

### 11.3 Data and Code

- Each team **must** use real-world data. Simulated data are not allowed. Please talk to the instructor if you are not sure about the nature of the data. There are many publicly available datasets. For example, UCI and Kaggle provide repositories that could be useful for you project:
  - UCI: <http://www.ics.uci.edu/~mllearn/MLRepository.html>
  - Kaggle: <https://www.kaggle.com/datasets>
- Each team must submit the code with a readme file describing how to run the code
- For full consideration, experiments must be reproducible given the (link to the) data, code, and the readme file

### 11.4 Final Report

The Final report is 3-4 pages. It must include:

- Title
- Introduction (including problem definition and motivation)
- Proposed method and the idea behind it (e.g. why should the method work)
- Experimental results and analysis (e.g. why the results look like this)
- Conclusions

### 11.5 Presentation

- Each team will present their Final project in the last week of class
- A presentation should be no longer than 10 minutes, and will be followed by a Q & A session (no longer than 2 minutes)
- All team members should present

## 12 Submission

- Homework **must** be completed by individual students. Final project will be completed by groups of 1, 2, or 3 students. Both the Homework and Final project should be submitted to Blackboard.
- Homework and Final project will be due for submission through Blackboard by Tuesday at 11:59 PM (Eastern time). **Submission will no longer be accepted after the deadline, and will receive a grade of 0.**

## 13 Grading Scheme

- 30% Homework (6)
- 18% Quiz (6)
- 22% Final project (1)
  - 5% Proposal
  - 6% Data and Code
  - 5% Final report (3-4 pages)
  - 6% Presentation (10 minutes)
- 30% Exams
  - 10% Midterm Examination
  - 20% Final Examination

## 14 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

## 15 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

## 16 University Policies

### 16.1 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://students.gwu.edu/accommodations-religious-holidays>

## **16.2 Academic Integrity Code**

Academic dishonesty is defined as cheating of any kind, including misrepresenting one's own work, taking credit for the work of others without crediting them and without appropriate authorization, and the fabrication of information. For details and complete code, see: <https://studentconduct.gwu.edu/code-academic-integrity>

## **16.3 Safety and Security**

In the case of an emergency, if at all possible, the class should shelter in place. If the building that the class is in is affected, follow the evacuation procedures for the building. After evacuation, seek shelter at a predetermined rendezvous location.

# **17 Support for Students Outside the Classroom**

## **17.1 Disability Support Services (DSS)**

Any student who may need an accommodation based on the potential impact of a disability should contact the Disability Support Services office at 202-994-8250 in the Rome Hall, Suite 102, to establish eligibility and to coordinate reasonable accommodations. For additional information see: <https://disabilitysupport.gwu.edu/>

## **17.2 Mental Health Services 202-994-5300**

The University's Mental Health Services offers 24/7 assistance and referral to address students' personal, social, career, and study skills problems. Services for students include: crisis and emergency mental health consultations confidential assessment, counseling services (individual and small group), and referrals. For additional information see: <https://counselingcenter.gwu.edu/>

## 18 Tentative Schedule

Class Date	Topic	Assignment Given	Assignment Due
08/28	Review of Machine Learning Data Preprocessing		
09/04	Optimization		
09/11	Linear Regression	Homework 1 Given	
09/18	Quiz 1 Logistic Regression	Homework 2 Given	Homework 1 Due
09/25	Quiz 2 Neural Network	Homework 3 Given	Homework 2 Due
10/02	Quiz 3 Neural Network (continued)		
10/09	Fall Break (no classes)		Homework 3 Due
10/16	Midterm		Final Project Proposal Due
10/23	Decision Tree Random Forest	Homework 4 Given	
10/30	Quiz 4 Naive Bayes K-Nearest Neighbor	Homework 5 Given	Homework 4 Due
11/06	Quiz 5 Support Vector Machine		
11/13	Hyperparameter Tuning Model Selection	Homework 6 Given	Homework 5 Due
11/20	Quiz 6 Unsupervised Learning		Homework 6 Due
11/27	Review		Final project Data, Code, and Report Due
12/04	Presentation		
12/11	Final exam		