Machine Learning I
DATS 6202 - O10, Spring 2019

1 Meeting Time and Location

- Meeting time: Monday, 12:45 PM - 3:15 PM
- Location: 1776 G St, NW C-106

2 Instructor

- Name: Yuxiao Huang
- Email: yuxiaohuang@gwu.edu
- Office address: Samson Hall, Room 314
- Office hours:
  Monday, 3:30 PM - 4:30 PM;
  Tuesday, 2:20 PM - 3:20 PM;
  Wednesday, 5:00 PM - 6:00 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: Pedro Uria Rodriguez
- Email: pedrouriar@gwmail.gwu.edu
- Office address: Samson Hall, Room 304 (the conference room)
- Office hours: Monday, 3:30 PM - 5:30 PM

4 Course Description

- This course offers an introduction to machine learning for students in Data Science program
- While we will dive deep into the math underpinning some of the simple (i.e., linear) models, the real focus of this course is to teach students hands-on coding skills to apply machine learning to solve real word problems
- This course will use `Jupyter notebook` for coding

5 Learning Outcomes

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

- understand the math underlying the linear models
- implement the linear models (using `numpy`)
- conduct data preprocessing, hyperparameter tuning, and model selection (using `scikit-learn`)

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 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

8 Homework

- There will be 6 homework, which will only include coding questions
- Homework must be completed individually

9 Quiz

- There will be 10 Quizzes, which will only include written questions

10 Exam

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

11 Final Project

- The final project is a great opportunity for you to apply what you have learned in class to solve real-world machine learning problems.
- While each team can choose a problem in the domain 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 of 1, 2, or 3 students.

11.1 Deliverables

- Proposal (a `pdf` file should be submitted to `blackboard`)
- Report (a `ipynb` file should be submitted to `blackboard`)
- Presentation (the slides are not required to be submitted)

11.2 Proposal

The project proposal is 1-page maximum, in `pdf` form. It should include:

- The title of the project
- The problem and motivation
- The link to the data
- The responsibility of each team member

11.3 Report

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

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

11.4 Presentation

- Each team will present their final project
- A presentation should be no longer than 10 minutes (and no shorter than 8 minutes), and will be followed by a Q & A session (no longer than 2 minutes)
- All team members should present

12 Submission

- Homework and final project (proposal and report) will be due for submission through `blackboard` by Monday 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)
- 10% Quiz (10)
- 30% Final project (1)
  - 5% Proposal
  - 15% Report
  - 10% Presentation
- 30% Exams
  - 10% Midterm
  - 20% Final

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	Release Date	Due Date
01/14	Inclement Weather ~~(no classes)~~		
01/21	Martin Luther King Day ~~(no classes)~~		
01/28	Introduction to Machine Learning Data Preprocessing	Homework 1	
02/04	Linear Regression	Quiz 1 Homework 2	Homework 1
02/11	Linear Regression (continued)	Quiz 2 Homework 3	Homework 2
02/18	President's Day ~~(no classes)~~		
02/25	Logistic Regression	Quiz 3 Homework 4	Homework 3
03/04	Logistic Regression (continued)	Quiz 4	Project proposal
03/11	Spring Break ~~(no classes)~~		Homework 4
03/18	Midterm		
03/25	Neural network	Quiz 5 Homework 5	
04/01	Neural network (continued)	Quiz 6 Homework 6	Homework 5
04/08	Neural network (continued)	Quiz 7	
04/15	Decision Tree & Random Forest	Quiz 8	Homework 6
04/22	Clustering Analysis	Quiz 9	
04/29	Review	Quiz 10	Project report
05/01	Final project presentation		
05/06	Final exam		