CS 559 Machine Learning: Fundamentals and Applications

Fall 2024
Week 0: Orientation

Orientation Outline

- Instructor Information
- Course Information
- Reading information
- Course Prerequisites and Goals
- Lecture Structure Information
- Course Structure Information
- Grading Procedure Information
- Late Submission Information
- The Honor Code Information
- Ways to Success the Course

Instructor Information

- Instructor: In Suk Jang
- Contact Info: ijang@stevens.edu
- Course Schedule: Online
- Virtual Office Hours: Friday 11 AM-12 PM

Course Information

- Prerequisite Course: MA 222 Probability Theory
- Computational Language: Python (higher than 2.7)
- Working Environment: Jupyter Notebook via
  - Anaconda: https://www.anaconda.com/products/distribution
  - Visual Studio Code: Install Jupyter Notebook directly

Reading Information

- Bishop, Christopher M., 2006
  Pattern Recognition and Machine Learning
- Ian Goodfellow, Yoshua Bengio, and Aaron Courville, 2016
  Deep Learning
- T. Hastie, R. Tibshirani, and J. Friedman, 2016
  The Elements of Statistical Learning

Course Prerequisites and Goals

Before this course, you should know…
- Programming - Python
- Linear Algebra – Vector, Matrix, Projection, Eigenvalues, ….
- Probability and Optimization – distributions, expectation/variance, etc…

At the end of this course, you should......
- Be more proficient in math and programming.
- Be able to implement general ML techniques for a variety of problem types.
- Be able to recognize when and how a new problem can be solved with an existing technique.

Lecture Structure Information

- **Course Mode:** Synchronous
- **Materials:**
  - Videos and lecture notes (pdf and notebook) will be posted on Monday before noon.
- **Timelines (~150 minutes):**
  - Announcements and Review (10 - 20 minutes) if needed
  - Current week's topic (100 – 120 minutes)
    - Includes coding practices and numerical examples
  - Overview and summary (10 – 20 minutes)
- Videos:
  - Separated by subtopics, each may be about 30 min. to 1 hour

Course Structure Information

1. Quizzes (10 %):

- There will be nine weekly quizzes via LockDown Browser, and each quiz may take about 10 to 15 minutes.

- Every quiz will be published on Tuesday and is due on Saturday in the same week by 11:59 PM.

- Question formats: Multiple Choice(s), True/False, Calculations

- Questions are theoretical and application questions, and no coding questions will be asked.

Course Structure Information

2. Assignments (30%):

- There will be 5 bi-weekly assignments.

- Each assignment will be published on Tuesday.

- Each assignment may have up to 4 or 5 questions.

- Typical questions will be

  - Model implementation using the Python Scikit-Learn

  - ML Algorithm implementation using NumPy

- Late submission policy: 10% deduction for every 24 hours.

Course Structure Information

3. Project (25%):

- A group project: Groups will be assigned during the 9th week (10/29-11/4), and the topic will be announced during the 10th week (11/5-11/11).

- The project will be a class competition.

- The project will require a video presentation and a written report.

Course Structure Information

4. Midterm Exam (15%):

- 6th Week (10/15 – 10/21)

- Online closed exam via LockDown Browser

- Topics from weeks 1 to 5 (Linear Regression to Neural Networks)

5. Final Exam (20%):

- Same format as the midterm exam.

- 15th Week (12/10– 12/17)

- Cumulative closed exam via LockDown Browser.

Grading Procedure Information

- **Exceed expectations**: The writer demonstrates the quality of work a reader expect from an ML professional within an organization.
  - Shows the ability to assess an organization and communicate the writer's ideas for the application of different methods of analytics to solve business problems.
  - The writer's analysis is thorough and well-written.
  - Assignments in this category usually are A-range.

- **Meets expectations**: Readers can clearly understand the writer's analysis.
  - Can communicate ideas with gaps minor enough that readers can figure out what the writer wanted to say.
  - Arguments and calculations are consistent.
  - Assignments in this category usually are B-range.

Grading Procedure Information

- **Nearly meets expectations**: There are some significant gaps in responding to the assessment criteria.
  - Some of the writer's explanations may not be articulated clearly.
  - Some confusion in the writer's assessment of the application of different methods of analytics to solve business problems.
  - Assignments in this category usually are C-range.
- **Does not meet expectations**: It is clear the writer does not understand the concepts being described in this course.
  - Multiple frames have not been addressed or are addressed insufficiently.
  - The writer's descriptions and analysis are not written clearly.
  - Assignments in this category usually are D-range.
- **No credit**: No credit was earned for this assignment, or no work was submitted.
  - Assignments in this category usually are F-range.

Late Submission Information

- A **10 % late submission penalty** will be applied every 24 hours from the due date.

- **All assignments must be submitted for passing consideration**.

- Late assignment submissions must have my permission; otherwise, the penalty will be applied.

- All written assignments must be submitted to Canvas; email submissions or re-submissions are not accepted.

- **No late submission requests on projects and exams.**

The Honor Code Information

- Do not share any solutions or codes.

- Quote any use of references if necessary.

- Any tendency of violations:

  - Students (both providers and receivers) will get zero.

  - Students will be reported to the University on the 2nd attempt and receive an "F" for the semester.

Ways to success the course

1. Use office hours.
   - Make appointments if there are any schedule conflicts.
2. Use discussion forums.
   - Students are free to post any course-related questions, and other students and I can answer and/or guide them through to solve the problems.
   - I will use topics discussed in the forum for the review.
3. Form a study group
4. Self-Study for more information
   - This course will help you to start applying machine learning algorithms.
   - More hands-on machine learning is extensive and open-ended.
   - There are lots of resources outside, e.g., Stack Overflow, toward data science, etc.