Mid-semester feedback

CSC 461: Machine Learning

Fall 2022

Prof. Marco Alvarez
University of Rhode Island

Prerequisites

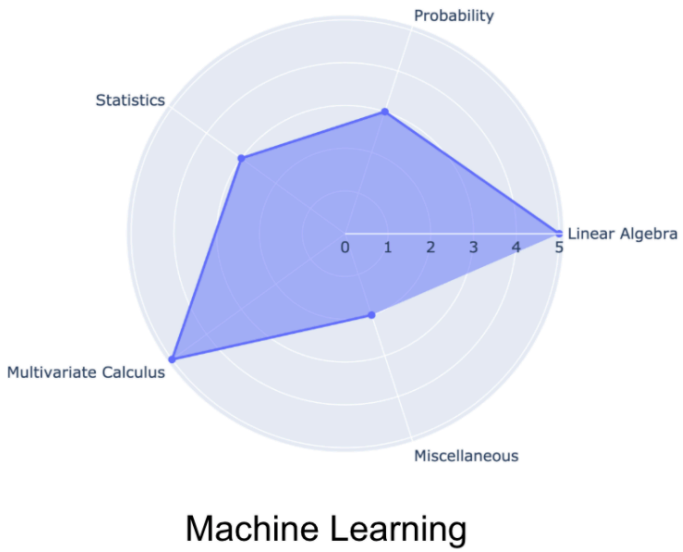
Should I take this class?

- Requires more **math** than traditional CS courses
- Programming experience is **required**

  consider taking this course at a later time if necessary
- Less emphasis on "how to use this library"
  - ✓ focus on understanding/implementing major algorithms
- High grades require high effort
  - ✓ long and challenging assignments/exams

Math resources

- Linear Algebra Review and Reference
  - ✓ http://cs229.stanford.edu/summer2019/cs229-linalg.pdf
- Review of Probability Theory
  - ✓ http://cs229.stanford.edu/summer2019/cs229-prob.pdf
- Computational Linear Algebra for Coders
  - ✓ https://github.com/fastai/numerical-linear-algebra
- Mathematics for Machine Learning
  - ✓ https://gwthomas.github.io/docs/math4ml.pdf



Machine Learning

https://www.analyticsvidhya.com/blog/2019/10/mathematics-behind-machine-learning/

Python/Numpy resources

- Google's Python Class
  - ✓ https://developers.google.com/edu/python/
- From Python to Numpy
  - ✓ https://www.labri.fr/perso/nrougier/from-python-to-numpy/
- Python Numpy Tutorial
  - ✓ http://cs231n.github.io/python-numpy-tutorial/

Grading

Programming assignments

- Still working on it …
  - ✓ should finish soon
  - ✓ grading takes significantly more work
- Kaggle and assignment 3
  - ✓ very low number of submissions
  - ✓ not really exciting results
  - ✓

(old) Grading

- Assignments (25%)
  - ✓ ~6 Homework Assignments
- Midterm Exam (25%)
- ~~Technical Presentation (20%)~~
  - ✓ ~~groups of 2~~
- Final Project
  - ✓ progress report (5%)
  - ✓ final report (15%)
  - ✓ presentation (10%)

(new) Grading

- Assignments (30%)
  - ✓ ~6 Homework Assignments
- Midterm Exam (30%)
  - ✓ second Midterm Exam (keep highest score)
  - ✓ covers everything until the week before the exam
  - ✓ applied on Nov 21st
- Final Project (40%)
  - ✓ progress report (5%)
  - ✓ final report (15%)
  - ✓ poster presentation (10%)
  - ✓ GitHub repo with the source code (10%)

Final Projects

Final project

- Group composition
  - ✓ 2-3 students per group
- Deliverables
  - ✓ progress report (Nov 16th)
  - ✓ final report (Dec 20th)
  - ✓ poster presentation (Dec 21st)
  - ✓ GitHub repository with all the code (Dec 20th)

Topics

- Implement some interesting machine learning application
- Reproduce an academic paper
- Tweak an existing ML algorithm to solve a particular type of problem
- Your own graduate/undergraduate research

Final Project

- Considerations
  - ✓ pick a publicly available dataset (can also collect your own)
  - ✓ define your ML goals and methods
  - ✓ pick a good framework and learn it
- Inspiration
  - ✓ visit cs229 project list (http://cs229.stanford.edu/projects.html)
  - ✓ visit SOTA (https://paperswithcode.com/sota)

Outstanding projects

- Demonstrated knowledge of Machine Learning (theory/practice)
- Significant implementation effort (awesome final product, produces at least one WOW)
- Novelty (can't find this online)
- Use of interactive tools such as:
  - ✓ huggingface, gradio, etc.

straight A in this course

Progress report

- Title
- Team members
- Introduction
  - ✓ provide context and existing work for the problem
- Problem Definition
  - ✓ precisely define what is the ML problem
- Data
  - ✓ provide detailed description of data
- Methods
  - ✓ provide a clear pipeline of the methods used for solving the problem
- Preliminary Results (optional)

Final report

- Title
- Team members
- Introduction
- Problem definition
- Data
- Methods
- Experiments
- Conclusion