Math 156: Machine Learning

Winter 2024

Instructor:	Jona Lelmi	**Time:**	MWF 3:00 pm – 3:50 pm
Email:	jona.lelmi [AT] math.ucla.edu	**Place:**	MS 5127.
TeachAsst:	Colin Ni	**Time:**	R 3:00 am – 3:50 am
Email:	colinni [AT] ucla.edu	**Place:**	MS 5118.

Course Pages:

1. BruinLearn and Gradescope
2. GitHub repository for notebooks and other things `https://github.com/jonalelmi/MATH156_winter24`

Prerequisites: 115A, 164, 170E (or 170A or Statistics 100A) and Programming in Computing 10A of Computer Science 31. Strongly recommended requisite: Program in Computing 16A or Statistics 21. If you did not take 16A, you will need to familiarize yourself with Python during the quarter.

Objectives: This course serves as an introduction to machine learning and pattern recognition. We will touch upon several topics in unsupervised and supervised learning. In particular, we will discuss algorithms for data clustering, dimensionality reduction, regression and classification. You will learn the core ideas behind widely used pattern recognition methods: in particular, we will focus on a solid understanding of the basics.

Office Hours:

- **Instructor**. W 4:00 pm - 5:00 pm @MS7360, or by appointment.
- **TeachAsst**. TBD.

Main References: The content of the course will be based on several sources specified every time. General references on the topics covered during the lectures are:

- Christopher M. Bishop, *Pattern Recognition and Machine Learning*, Springer, 2006.
- Gareth James, Daniela Witten, Trevor Hastie, Robert Tibshirani, Jonathan Taylor, *An Introduction to Statistical Learning: with Applications in Python*, Springer, 2023.
- Richard S. Sutton, Andrew G. Barto, *Reinforcement Learning: an introduction*, A Bradford Book, 2018.

Grading Policy: The course will consist of four homework sheets, one midterm and a final project. If $\mathbf{H}$ is the homework grade, $\mathbf{M}$ is the midterm grade, and $\mathbf{P}$ is the final project grade, the final score will be calculated using the best of the following two schemes:

- **Scheme I**: $\mathbf{H}$ (25%) + $\max(\mathbf{M}, \mathbf{P})$ (45%) + $\min(\mathbf{M}, \mathbf{P})$ (30%)
- **Scheme II**: $\mathbf{H}$ (35%) + $\mathbf{P}$ (65%)

A final average of 90% or higher will guarantee you at least an A-, a final average of 80% or higher will guarantee you at least a B-, and a final average of 70% or higher will guarantee you at least a C-. These letter grade cut-offs may be adjusted downward at the end of the quarter, so that more students receive higher grades. But under no circumstances will the grade cut-offs be higher than stated above: this policy can only help you, not hurt you.

Homework:

- Homework sheets will be distributed on Mondays every two weeks, the hand in deadline is at 3:00 pm of the Monday two weeks after distribution. For instance, if the first homework sheet is distributed Monday 8 January 2024, the deadline for the submission of the solutions is Monday 22 January 2024 at 3:00 pm.
- **Under no circumstances late submissions will be accepted**.
- Students are allowed to submit their homework in groups of at most three people.
- Homework solutions must be submitted on Gradescope.

Midterm:

- The tentative date for the midterm is **Friday 16 February 2024, 3:00 pm**.
- Grading **Scheme II** is meant to help in case you miss the midterm, so **no** make up midterm will be offered.

Final Project:

- By the end of the second week of the class, students will form groups with no more than 4 students (these groups need not be the same as homework groups). Email the instructor to communicate the names of the members of your group. At the end of the second week of class, students without a group will be randomly assigned to groups of three or four members.
- Each group will choose a project to work on: the goal of the project is to apply some of the techniques discussed in class to a simple problem. The problem to work on can be chosen independently by the groups, some suggestions are:
  - Image segmentation using spectral clustering.
  - Customer segmentation using k-means.
  - A character level language model with RNN.
  - A music improvisation tool using RNN.
  - A simple game playing agent with RL.
  - Image classification problem with CNN.

  Some of these suggestions require more work than others. In particular, some of these project will require you to learn new concepts on your own: the material covered in the course will give you the necessary foundation to navigate through additional needed tools.
- The chosen problem should be challenging enough so that you do not complete it in just one day, in particular, it should not be a straightforward modifications of examples being discussed in the lectures.
- By the end of week three, groups will submit a project proposal describing objectives and methodologies. The proposal should not exceed one page. I will read your proposal and suggest modifications/improvements before you start working on the project.

- Groups are required to submit a final report and a Python notebook reproducing the results of the report by **Friday 15 March, 4:00 pm**.
- The final submission should contain:
  - Final Report. A report describing:
    * *Objectives.* Carefully describe the objective of the project. Why is this interesting? What are other approaches used?
    * *Description of the data (if used).*
    * *Methodologies.* A description of the tools used, and the motivations for choosing them.
    * *Theoretical justifications.* Why the chosen approach should work? In which circumstances could it fail?
    * *Results.* A presentation of the obtained results.
    * *Discussion on possible future extension.*
    * *References.* Include a list of references you use.
  - Python Notebook. A Jupyter notebook and all the Python code needed to reproduce the results of the report.
- In the final week of the course, each group is encouraged to give an oral presentation of the results of the project. The oral presentation is meant to expose the rest of the class to the problems and methodologies used in the project. The oral presentation is *not* mandatory and is *not* graded.

Tentative Course Outline:

- **Week 1** Overview, review of probability, the Gaussian distribution and the curve fitting problem.
- **Week 2** Bias Variance Tradeoff, k-means clustering, Gaussian mixture model.
- **Week 3** Spectral clustering, PCA, linear regression.
- **Week 4** Logistic regression, optimization with gradient descent, Naive Bayes classifier, Artificial Neural Network, backpropagation.
- **Week 5** Regularization and initialization, other optimization algorithms, recurrent neural networks.
- **Week 6** Recurrent neural networks, GRU, LSTM.
- **Week 7** Convolutional neural networks.
- **Week 8** Introduction to RL: MDP, Bellmann expectation equation, Bellman optimality equation, policy evaluation, policy iteration, value iteration.
- **Week 9** Introduction to RL: Monte Carlo methods, temporal learning, Q-learning.
- **Week 10** Project presentations.

Important Dates:

Midterm .... **Friday 16 February 2024, 3:00 pm**
Final Project Due .. **Friday 15 March 2024, 4:00 pm**