

CS 584 - Machine Learning

Instructor: Steve Avsec

Email: savsec@iit.edu

Due: 3 October at 11:59PM CDT (Thursday night)

Project 1

Your task in this project is to implement *one of the* following algorithms *from first principles*:

1. Linear regression with ElasticNet regularization (combination of L^2 and L^1 regularization)
2. Regularized Linear Discriminant Analysis (as described in section 4.3.1 of Elements of Statistical Learning, 2nd Edition)

Questions

You will need to attach a README to your implementation. It *needs* to answer the following questions. Otherwise, your README should include things any other README would include: basic usage examples and other information about how you have structured your implementation and why. If you would like examples of READMEs in the wild, please ask your instructor.

1. What does the model you have implemented do and when should it be used?
2. How did you test your model to determine if it is working reasonably correctly?
3. What parameters have you exposed to users of your implementation in order to tune performance?
4. Are there specific inputs that your implementation has trouble with? Given more time, could you work around these or is it fundamental to the model?

Rules

1. You must turn in code that executes against any number of data sets with a specified format and is transparently easy to use. (For example, you could implement "fit" and "predict" methods like a SciKit Learn model.)
2. You may work in groups of *at most* 4. You do not need to tell us which group you're working with; when you turn in the assignment, at the top of your write-up, provide *all group members names and CWIDs (A#s)*.

3. You must turn in a write-up answering the questions above.
4. If you are caught using ChatGPT or any other LLM-based "copilot" to write code or answer the questions above, you will receive a 0 on this project.
5. *From first principles* means do not use SKLearn, Statsmodels, or any other "prebuilt" implementation of these models. You may use NumPy or SciPy to handle numerical calculations and Pandas if you feel you need it (shouldn't be necessary). Any other 3rd party libraries must be approved by your instructor.
6. If you are not planning to use Python to complete this project, please let your instructor know so you can be sure you are following the "from first principles" guideline.
7. Submission will be via GitHub with specific instructions to follow.
8. Please do not feel constrained by the basic objective; you should implement at least what is being asked, but if you would like to extend your implementation beyond the exact models you are being asked for, you should feel free. Document everything you do in your README and when grading, we will consider bonus points for groups who have gone above and beyond.