

Spring 2026 ML4P, Problem Set #1

(Non-)Linear Regression

Submission Policy

- Submit your code via a link to a git repository (e.g. `GitHub`) on Brightspace.
- If you collaborate with anyone and/or use online/textbook resources, you must cite them. You are allowed to use AI chatbots, but you must cite them.
- You are graded on effort, not correctness. To receive full credit, you must make a genuine attempt at each problem and clearly explain your reasoning, even if your final answer is wrong. This means just submitting raw chatbot dialogue will not count as full credit.

Dataset All questions on this problem set use data from <https://www.sdss4.org/dr17/irspec/>.

1 Regression

1. Use the notebook [\[here\]](#) to download the data. Note we've had issues downloading the data, so use the data in the Google Drive.
2. Split the features & labels into train, validation, test.
3. Predict the labels (say LOGG) using the spectra features using (a) linear regression, (b) K-Nearest-Neighbors, (c) Multi-Layer Perceptron.

Every method has many choices; choose something reasonable, and convince us your implementation is working. Hint: plot something.

2 Hyperparameter Tuning

1. For each method (linear, KNN, MLP), play around with your hyperparameters to get better performance on the validation. No need to be exhaustive, just try a couple.
For the KNN & MLP, rerun the algorithm but with various RNG seeds. Do you converge to a similar answer?
2. Once you're satisfied, justify your choices of hyperparameters.
3. Post your performance on the test set with a blurb about your methodology.

Hint: Read <https://karpathy.github.io/2019/04/25/recipe/>. And then after, you can google this non-exhaustive list of buzzwords to improve model performance. LeakyRELU, layer norm, weight initialization, learning rate scheduling, gradient clipping, feature engineering (normalization), batch norm.

3 Final Project: First Steps

The final project will be in the form of a slide deck or short paper (we haven't decided). But to get started, give us a few sentences about what you wanna work on!