

CS 5300, Assignment 2
Back Propagation with Housing Data
Due, Thursday, May 24 in class

You are to build a regression model using Back Propagation, presumably coded in Python. You will use a new housing data set (thanks Ms. Lomeli !), much like the Boston Housing data in that we are predicting house price based on house features. This data set is in Mr. Daudert's repo as `other_housing.csv`.

Your challenges are chiefly twofold:

Code Back Propagation. You have been given several useful links which should provide guidance of various kinds. This is not merely a rite of passage – coding your own version of the algorithm removes any doubt about missing magic, because all the magic has been your own conjuring.

Feature Massaging and Representation – be sure to read the description of the features and build a representation and corresponding architecture

And then of course, there is the experimental work with the model you have built.

In order to standardize the assignment shall we agree to all do the following:

- 1) Your architecture *must* have an input and output layer along with 2 hidden layers. This will force you to write code that captures the essentials of the Back Propagation algorithm. You will almost certainly have to “normalize” and “denormalize” your data.
- 2) Keep some ongoing lab notes of some kind, because it is the progress that is most interesting to include in the submission of your work. You have lots of choices to make and both “Efficient Back Propagation” by LeCuan and the link put on the Slack channel by Mr. Oolomany speak to various choices you might explore. How you make choices and then reject or modify same is always part of the ‘fun’.
- 3) You will ultimately turn in a (yes, printed) report that contains the written word – text that describes the problem, your choices for data manipulation, feature representation, architecture, activation function, learning rates etc etc. Obviously, you will report on the performance of your final work. But if you report (create plots ?) how your network progressed, how you improved performance as you varied particular parameters, etc. the reader can be better informed. But you know that, already. That is, the science of the work.

- 4) And there is a Prize ! We have held out a Super Secret Test Set (SSTS !) against which we will test. Please don't peek.

If you have gotten busy already, then by the time you come in on Monday, we should be able to engage issues large and small to keep you on your way to a Thursday deadline. Getting things delivered in class that day lets us have work in hand, so try to make that deadline, please ? In the meantime, please use the Slack channel, which has some active and helpful participants, and really leverages the various skills in our class.

If there are particular questions please use Slack and email for either Dillon or me.

rgt

5/18/18