Assignment: Perceptron and Shallow Neural Network

Problem 1 (30 points)

For this problem, you use the data file wine.csv. You will create a model that will use an optimized set of features to predict the wine quality.

1. Use lasso regression to effectively choose the best set of input features. Report these set of features, the value of the coefficients and Eout.
2. Debias your model. Report the value of the coefficients and Eout. Which one the two models performed better on out of sample points?
3. Report the final model coefficients. Plot the predicted values vs. actual values.

   In solving the above questions:

   - Set seed to 42.
   - Use the first 30 rows for training and the rest for test.
   - Set up a pipeline that uses linear normalization.
   - Use a grid search with 5 fold cross validation for this regularization parameter range: $\alpha = np.logspace(-1, 0, 3)$
   - Use neg-median-absolute-error for scoring.

Problem 2 (40 points)

In this problem you will use the data in *siCoData.csv* file to train a neural network. Use the backpropagation algorithm to train a 3-layer (input, hidden, output) neural network. Use stochastic gradient decent (SGD) technique and assume that the activation function for the hidden layer and output layer are tanh and linear, respectively. (You must write *your own code* for BP and SGD).

1. The stopping criteria for training in this problem should be a combination of achieving a minimum in-sample error

$$E_{in} = \frac{1}{N} \sum_{n=1}^{N} e_n$$

   and reaching a maximum number of epochs (In this expression N is the number of observations in the data set and e_n is the error corresponding to each individual training point). Report the minimum E_{in} that you could achieve along with the related weights and number of iterations.

2. Graph the original data (y vs. x) and the predicted values ($\hat{y}$ vs. x) on two separate scatter plots.