**World Quant University**

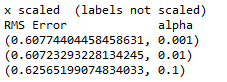
**Professor: Tiberiu Stoica**

**Machine Learning**

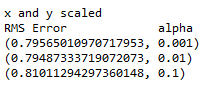
Nikolas Lippmann Pareschi - [nikolaslippmann@gmail.com](mailto:nikolaslippmann@gmail.com)

**Unit 2 Project**

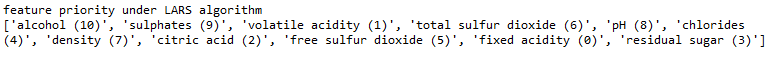
The project implements the Lasso Regression considering different values for Alphas. The smallest values make the Lasso Regression similar to the Linear Regression. The coefficients tend to zero as we use larges alphas, but we have less variance. For the values of alpha in the code, i.e. 0.001, 0.01 and 0.1 we have these RMS errors:



We also scaled y, which improved the predictions:

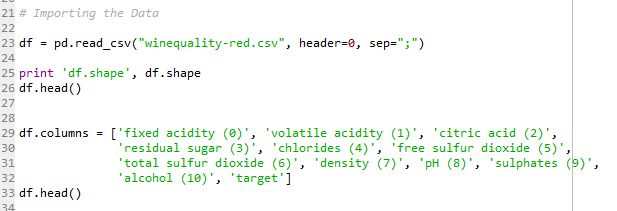


This is the order of importance of our predictors:

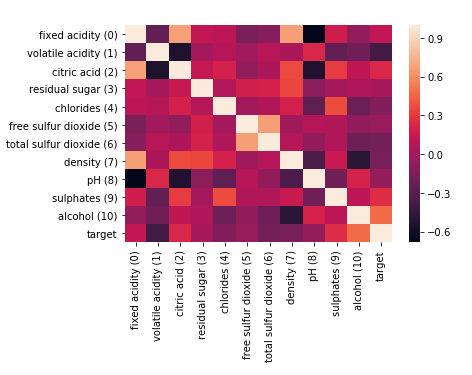


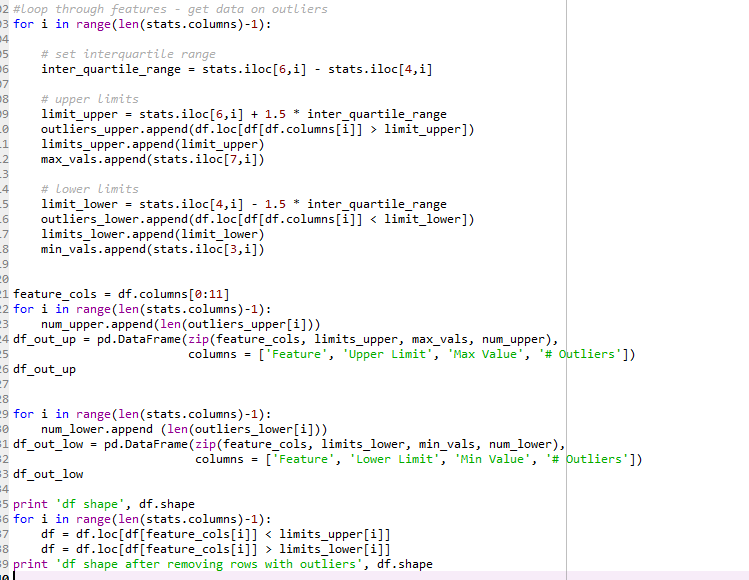
1. Conduct data manipulations

Codes:



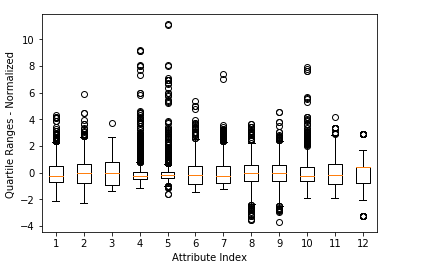
A seaborn map of correlation to help in the interpretation of the data:



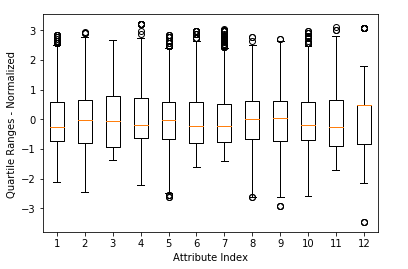


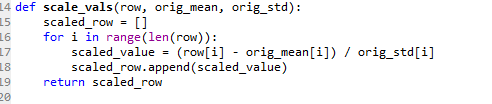
1. Conduct normalization on labels

Data before normalization:

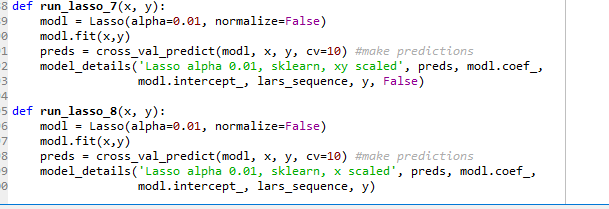


Data after normalization:

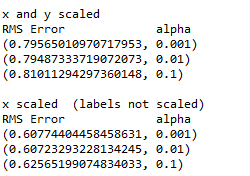




3.  Use LassoCV. Try out a 10 fold cross validation.



Results:



For this project I used these references:

<http://scikit-learn.org/stable/auto_examples/linear_model/plot_lasso_coordinate_descent_path.html#sphx-glr-auto-examples-linear-model-plot-lasso-coordinate-descent-path-py>

<http://methodmatters.blogspot.com/2017/04/analyzing-wine-data-in-python-part-1.html>

<https://www.r-craft.org/r-news/analyzing-wine-data-in-python-part-1-lasso-regression/>

<https://github.com/peterco877/datascience/blob/master/earlier%20code/Wine_Taste_Quality.ipynb>