**ANN Regressor Assignment**

**Goal:**

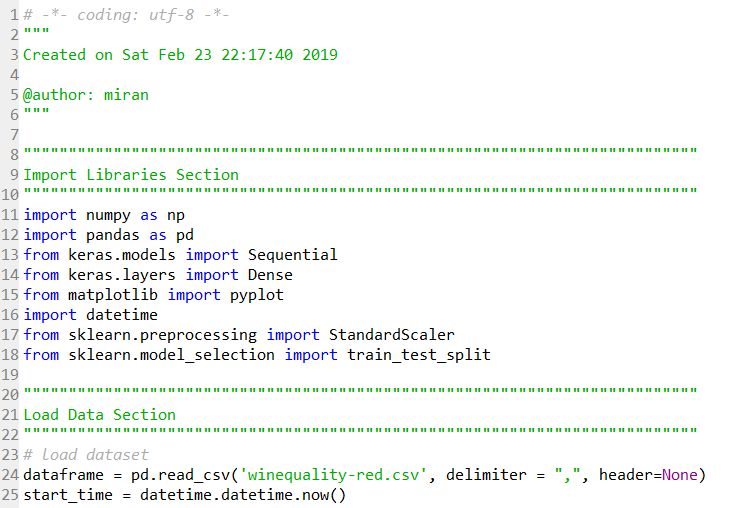
Produce, following the code we have developed in class and using Keras for the neural network functions, a regressor neural network that successfully predicts wine quality.

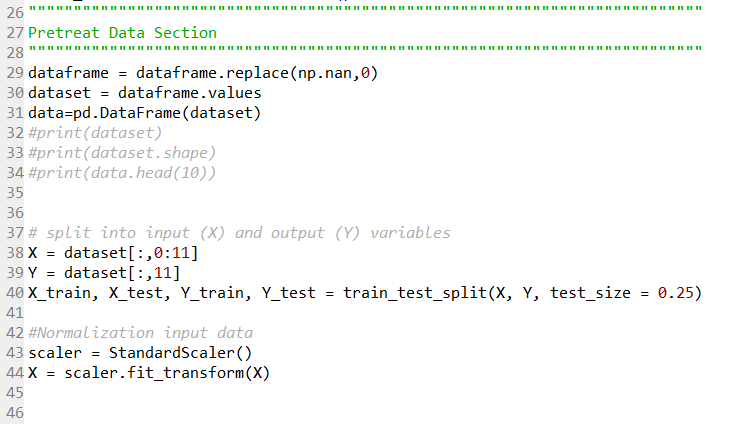
Dataset website: https://archive.ics.uci.edu/ml/datasets/Wine+Quality.

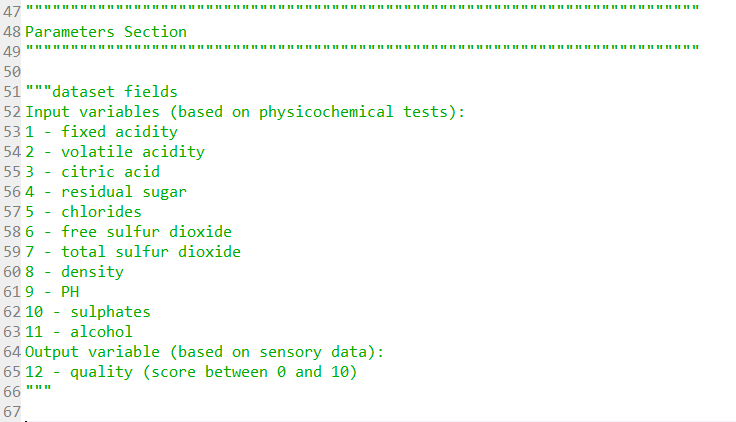
**1. Your objective should be stated in written form. What are you trying to accomplish? Predicting a number? Classifying? Your objective must reference the context of the problem, specifically.**

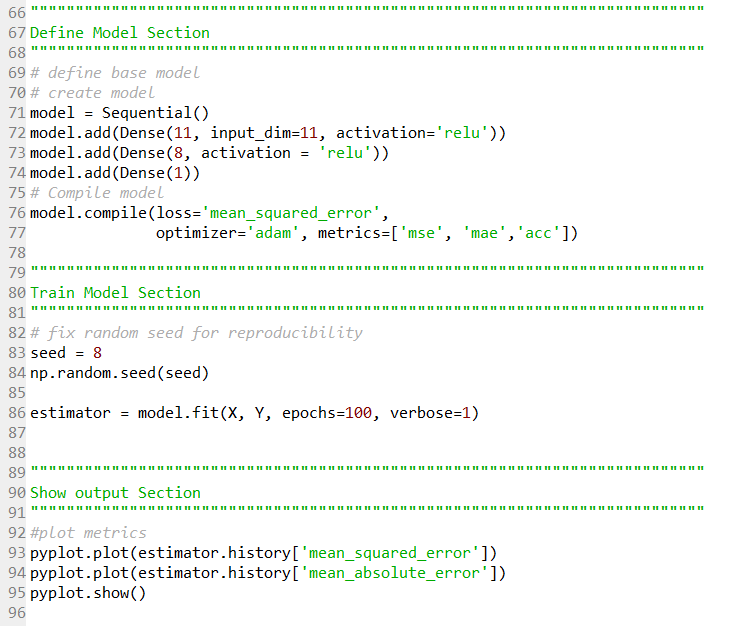
My objective is to build a regressor model to predict the wine quality based on the variables given including fixed acidity, volatile acidity, citric acid, residual sugar, chlorides, free sulfur dioxide, total sulfur dioxide, density, pH, and sulphates.

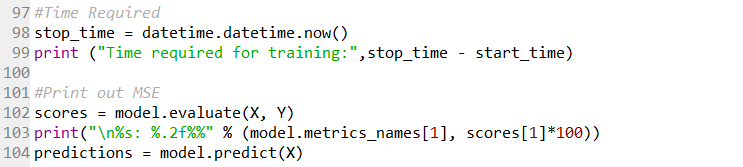
**2. Your final ANN model, in code.**











**3. Your final model and training algorithm, in words.**

My final model is regression model and my regression algorithm is regression algorithm.

The regression model has 11 units for the first layer, 8 hidden units for the intermediate layer, and 1 unit in the final layer. The final output model is 1-dimensional. The last layer is purely linear, the network is free to learn to predict values in any range. Relu(rectified linear unit) activation function is applied. The loss functions are MSE, MAE, ACC. **textbook states that the ACC is not a good indicator in regression model. I just put it there. )**

**4. Your experimental plan for arriving at the final model**

Answer: The experimental plan is to change epochs, change batch size, add intermediate layer, change unit in layers. After rerunning the code, the MSE, MAE, ACC will be changed.

I change the epochs to 50, 100, 200, 300, 500, 1000. The final epoch is 1000.

I change the batch size to 20, 32, 64, 128, 256, 512. The final batch size is 32(default).

I add a intermediate layer and set the unit to 8.

**5. How long it took to run all the models in your experimental plan**

Time required for training: 0:02:26.928059

**6. An explanation of the input variables and any preprocessing steps you took**

"""dataset fields

Input variables (based on physicochemical tests):

1 - fixed acidity

2 - volatile acidity

3 - citric acid

4 - residual sugar

5 - chlorides

6 - free sulfur dioxide

7 - total sulfur dioxide

8 - density

9 - PH

10 - sulphates

11 - alcohol

Output variable (based on sensory data):

12 - quality (score between 0 and 10)

"""

**7. An explanation of your metrics and justification for your choice.**

The loss functions are MSE, MAE, ACC **(Even though textbook states that the ACC is not a good indicator in regression model. I just put it there. )**

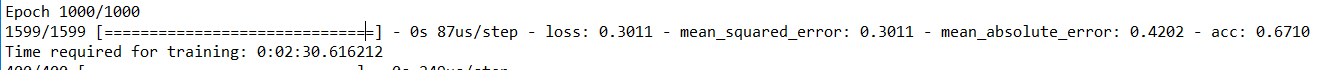
MSE loss function—Mean Squared Error, the square of the different between the predictions and the targets, a widely used loss function for regression problems.

MAE loss function—Mean Absolute Error. It is simply the absolute value of the difference between the predictions and the targets.

**8. An explanation of your method to validate the model**

To use test dataset to validate the train dataset. Use MSE metrics to evaluate the result. MAE and ACC can also to evaluate the result.

**9. Your results in terms of appropriate metrics for the objective and problem**



**The MSE is 0.3011**, MAE is 0.42, and the ACC is 0.67.

The most appropriate metrics for this progression problem is MSE.

========================================================================

Metrics Preference from textbook.

Reference from textbook (P110 & P88 ):

