Homework 2: Due Fri 09-07-2018
Total Points (34 pts)

1. (Flatten) (2 pts) Consider the matrix M given below. Predict the result of `M.flatten()` and check your answer using numpy.

   $M = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix}$

   ```
   import numpy as np

   M = np.array([[1,2,3],[4,5,6]])
   ```

2. (Curve Fitting) (12 pts) Use the normal equations to fit a polynomial with specified degree (1, 2, or 3) to the three points given below:

   $$(-1, 2), (0, 1), (2, 5).$$

   For each polynomial do the following:
   I. Write out the three linear equations that need to be solved and specify the data matrix $\mathbf{X}$ and right hand side $\mathbf{y}$.
   II. Specify if there are i) more equations than unknowns ii) the same number of equations as unknowns or iii) fewer equations than unknowns.
   III. Specify the polynomial.
   IV. Compute RMSE.

   (a) $y = co + c_1 x$.
   (b) $y = co + c_1 x + c_2 x^2$.
   (c) $y = co + c_1 x + c_2 x^2 + c_3 x^3$.

3. (Linear Regression for White Wine) (10 pts)

   (a) Load the data set `wine_quality_white.csv` into a dataframe. Assume wine quality is the feature to be predicted. How many input features does the data set contain? How many data points?

   (b) Determine the feature weights and bias that minimize MSE. Hint: Add a column of 1's to the data matrix $\mathbf{X}$ to represent the bias. Then solve the normal equations to obtain the optimal weights and bias.

   (c) Compare RMSE of the linear regression network with the RMSE of a simple bias network.

4. (L^1 Error and the Median) (10 pts) Let $y_1, y_2, \ldots, y_n$ be n target values. Show that the L^1 error $\frac{1}{n}\sum_{i=1}^{n} |b - y_i|$ of a bias network is minimized if b equals the median of the target values. To simplify the problem, assume $b \neq y_i$, $i = 1, 2, \ldots, n$. Recall that the median is any number that separates the lower and upper halves of the data.

   Hint: When $b > y_i$, $|b - y_i| = b - y_i$ and when $b < y_i$, $|b - y_i| = y_i - b$. Separating the data this way allows the absolute value operations to be removed making it possible to take derivatives.