Andrew Tran

CS 1675

Assignment 4 Report

Due: 2/14/2019

1a) There is 1 binary attribute in the set: Charles River dummy variable (= 1 if tract bounds river; 0 otherwise)

1b)

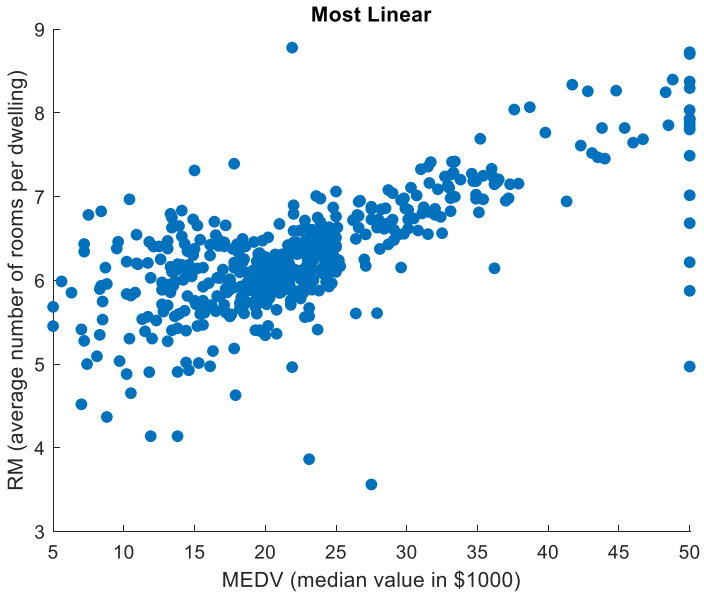
Attribute	Correlation w/ Attribute 14
1. CRIM	-0.3883
2. ZN	0.3604
3. INDUS	-0.4837
4. CHAS	0.1753
5. NOX	-0.4273
6. RM	0.6954
7. AGE	-0.3770
8. DIS	0.2499
9. RAD	-0.3816
10. TAX	-0.4685
11. PTRATIO	-0.5078
12. B	0.3335
13. LSTAT	-0.7377

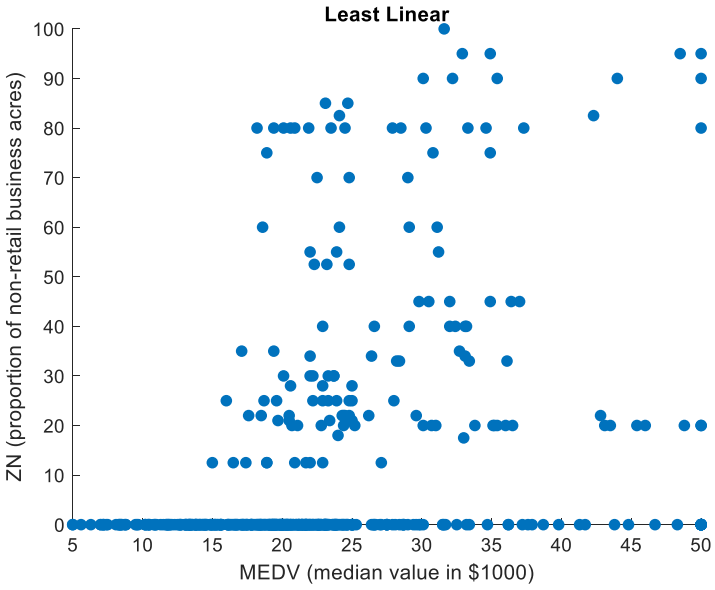
Highest Correlations:

Negative: MEDV vs. LSTAT = -0.7377

Positive: MEDV vs RM = 0.6954

1c)





My decision on which looked most/least linear was not entirely based on the value of their correlation with the target attribute. For example, my choice for most linear (MEDV vs RM) visually shows an obvious linear relationship but has the second strongest correlation. The attribute with the strongest correlation produced a scatter plot that seems more like an exponential distribution.

My choice for least linear (most nonlinear) was MEDV vs ZN because after MEDV = $15000, there seems to be no obvious relationship at all between either attribute.

1d) NOX (nitric oxides concentration in parts per 10 mil) vs DIS (weighted distances to five Boston employment centres)

2d)

Weights: -0.0979, 0.0489, -0.0253, 3.4508, -0.3554, 5.8165, -0.0033, -1.0205, 0.2265, -0.0122, -0.3880, 0.0170, -0.4850

Training MSE: 24.4759

Testing MSE: 24.2922 ← Better

3a) Function: gradient_descent.m

3b)

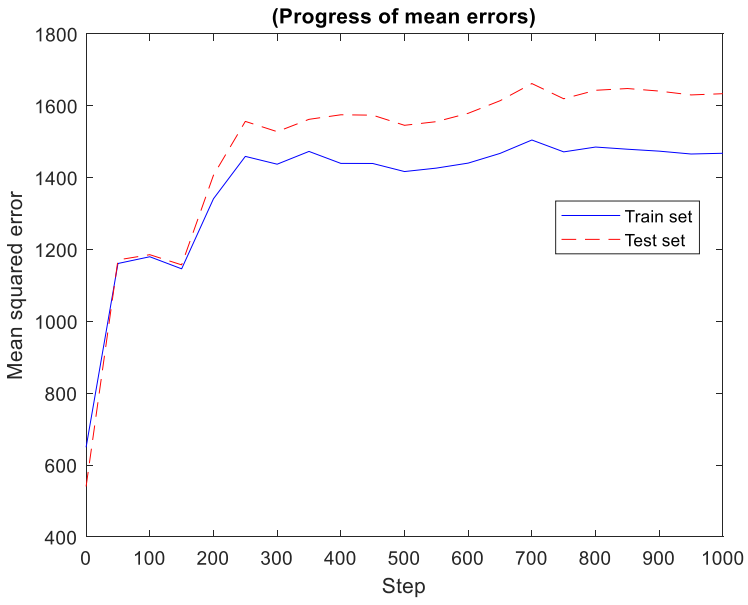
Training MSE: 1467.57

Testing MSE: 1633.40

These mean errors are worse than the results from solving the regression exactly.

3c) The weights and MSEs result in NaN.

3d)



3e) Observations:

- After 300 updates, the mean errors do not seem to change as drastically and stay around 1500 (with 0.05/n learning rate). This happens because the learning rate causes future data to not affect the model as much and after 300 data points, they have little to no effect.
- Making the learning rate =1 causes the MSE to rapidly increase. The MSEs reached inf around 200 updates.
- When setting the learning rate to 0.005, the MSEs increased rapidly at first but decreased after about 350 updates. The final MSEs were lower than the results reported in 3b after 1000 updates.