HW3

Name: Wenbo Song
UNI:ws2505

Problem 1 (Gaussian process coding)

1. Write code to implement the Gaussian process and to make predictions on test data.

Answer: For this part, I choose b = 5 and variance = 0.1 to make the prediction, after implementing the Gaussian process, the prediction result (part) is shown as below

```
[[-5.10175072]]
[[-8.2766852]]
[[ 7.00110395]]
[[-6.07445388]]
[[-3.76972584]]
[[ 1.38008861]]
[[-7.12689087]]
[[-1.8819534]]
[[-11.06468146]]
[[ 14.28407279]]
[[-7.50951133]]
[[ 14.44245973]]
[[-9.89165025]]
[[-3.97763394]]
[[-0.87859151]]
[[-9.60471176]]
[[-2.30669723]]
[[-0.61945446]]
[[ 12.56691028]]
[[-5.38229495]]
[[-10.28929725]]
[[ 13.67613411]]
[[-3.86354796]]
[[-6.58671316]]
[[-9.83128865]]
[[-8.33437583]]
[[ 3.73529152]]
[[ 0.86455461]]
[[-2.83591531]]
[[-4.74292164]]
```

2. For the given b and variance, use the mean of the Gaussian process at the test point as the prediction, the result table is shown as below:

result2.txt

```
[ 1.96627564  1.93313496  1.92341988  1.92219731  1.92476887  1.92921232  1.93463387  1.94058292  1.94681977  1.95321235]
[ 1.92016252  1.90487602  1.90807995  1.91590138  1.92480383  1.93370119  1.94225358  1.95037987  1.95809292  1.9654379 ]
[ 1.89764785  1.90251864  1.91764736  1.93251408  1.94569928  1.95723463  1.96740312  1.97649163  1.98474067  1.99234112]
[ 1.89050623  1.91498061  1.93884854  1.95793608  1.9732157   1.98576411  1.99637506  2.00560315  2.01383539  2.02134475]
[ 1.89584777  1.93558563  1.96459718  1.9855019   2.00131421  2.01387841  2.02431035  2.03330676  2.04131748  2.04864154]
[ 1.9096027   1.95954869  1.99080353  2.01191544  2.02737028  2.03946517  2.04946339  2.05810491  2.06584529  2.07297608]
```

The y axis is b from 5 to 15 and the x axis is variance from 0.1 to 1

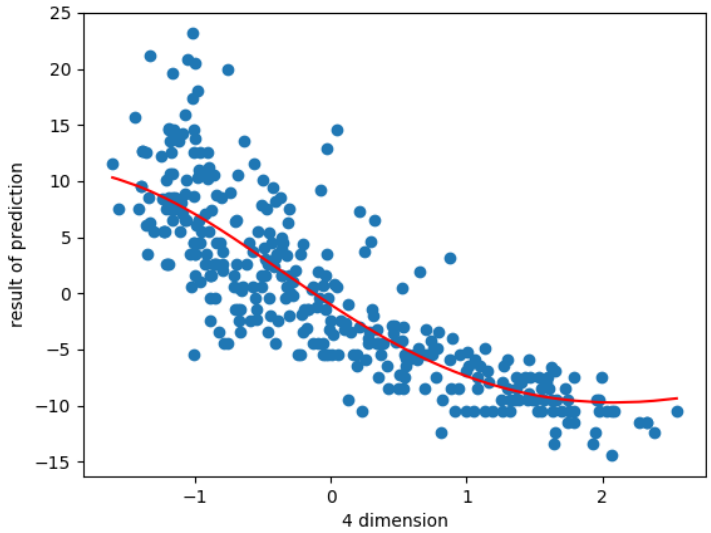
3. Which value was the best and how does this compare with the first homework? What might be a drawback of the approach in this homework

We can see from the table that when b = 11 and variance = 0.1 rems can reach its minimum 1.8905. In the first homework, the minimum rems is about 2.2, so using the Gaussian

process for prediction has a better performance than using the ridge regression and polynomial regression.

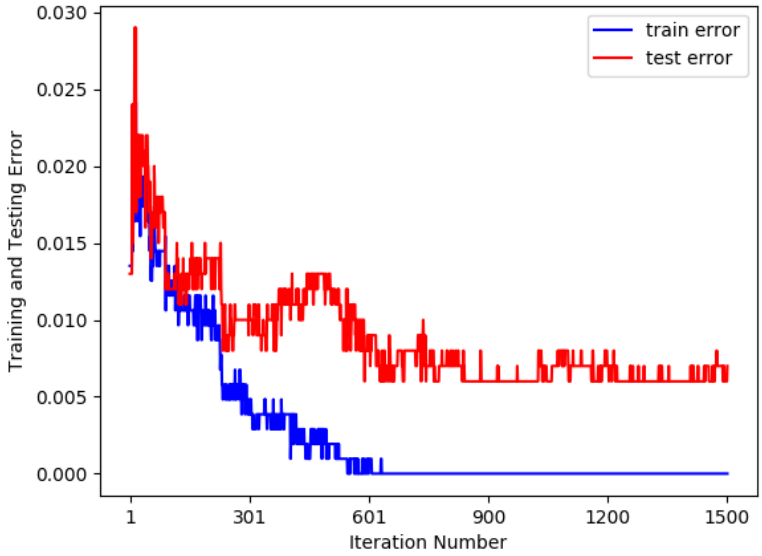
The drawback of Gaussian process compared to ridge regression is the computation time, the computation time can scale cubically, so it is hard to work with more than a few thousand points in real practice.

4. The visualization result of the 4^{th} dimension of x(car weight) can be seen as follow

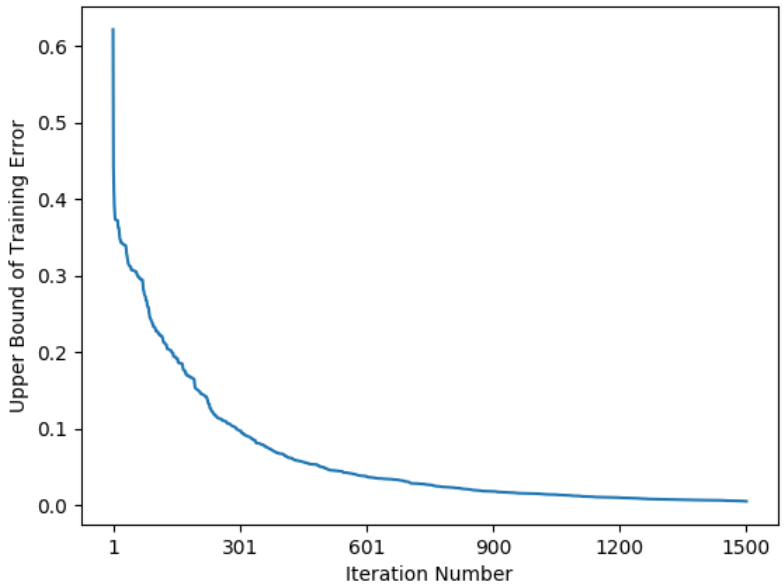


Problem 2

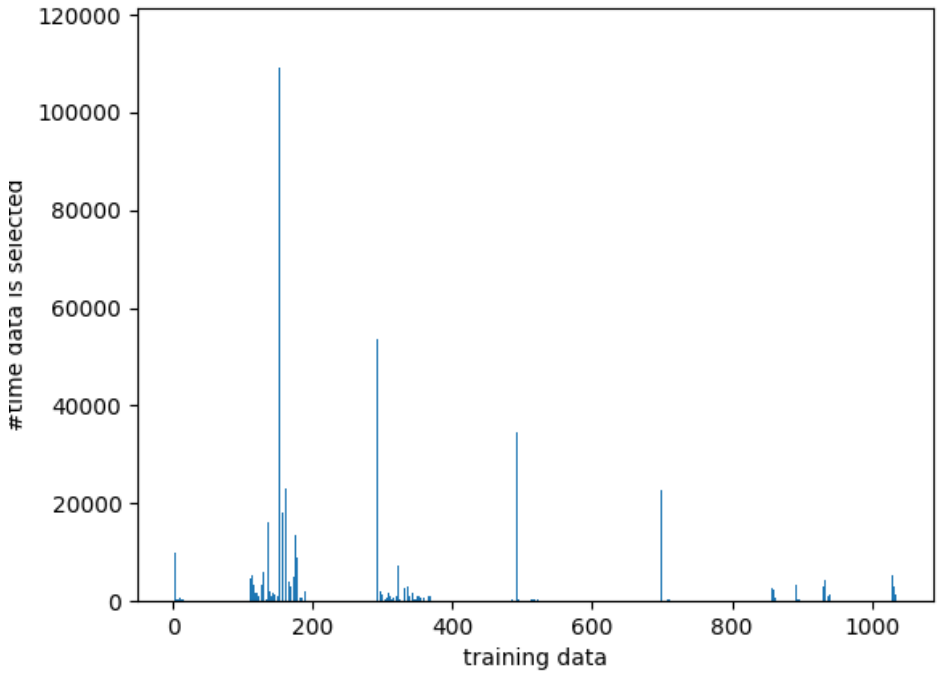
1 Run boosted LS classifier for T = 1500 rounds. The training and testing error of f_{boost} is shown as follow



2. The upper bound on the training error as function of t is shown as follow



3. The histogram of total number of times each training data point was selected is shown as below



4. The eps and alpha as a function of t can be seen as follow separately