

## COMS W4721: Machine Learning for Data Science HW3

Abhay Pawar (UNI: asp2197)

### Problem 1

b)

sigma squared b	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1
5	1.9663	1.9331	1.9234	1.9222	1.9248	1.9292	1.9346	1.9406	1.9468	1.9532
7	1.9202	1.9049	1.9081	1.9159	1.9248	1.9337	1.9423	1.9504	1.9581	1.9654
9	1.8976	1.9025	1.9176	1.9325	1.9457	1.9572	1.9674	1.9765	1.9847	1.9923
11	1.8905	1.9150	1.9388	1.9579	1.9732	1.9858	1.9964	2.0056	2.0138	2.0213
13	1.8958	1.9356	1.9646	1.9855	2.0013	2.0139	2.0243	2.0333	2.0413	2.0486
15	1.9096	1.9595	1.9908	2.0119	2.0274	2.0395	2.0495	2.0581	2.0658	2.0730

c)

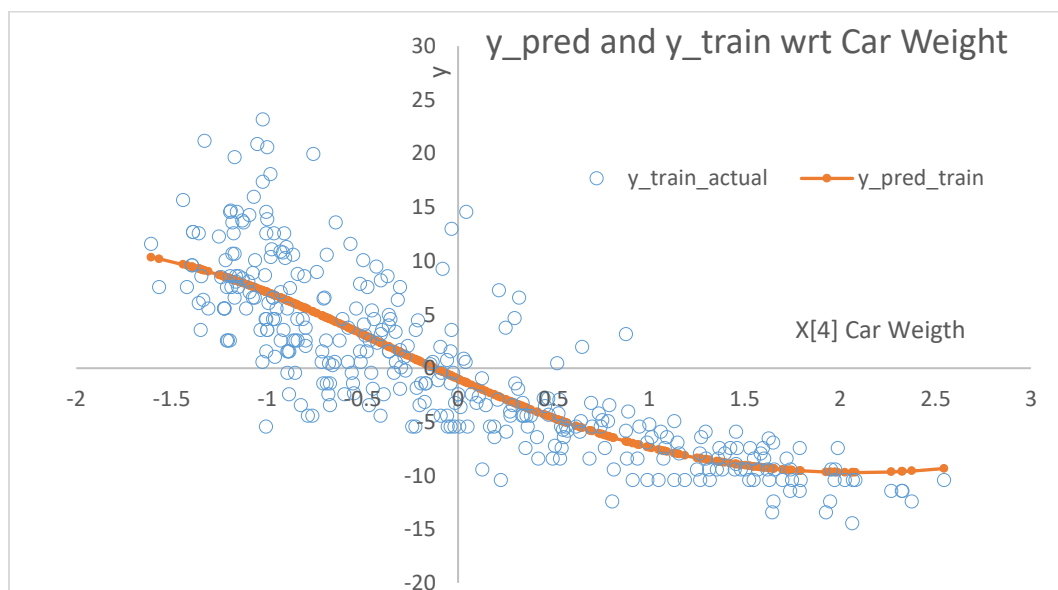
**Sigma squared = 0.1 and b = 11** works the best. **RMSE is 1.8905** for these settings.

For HW1, lambda=21 and p=2 gave the best RMSE of 2.19. Thus, Gaussian process works well.

#### Drawbacks:

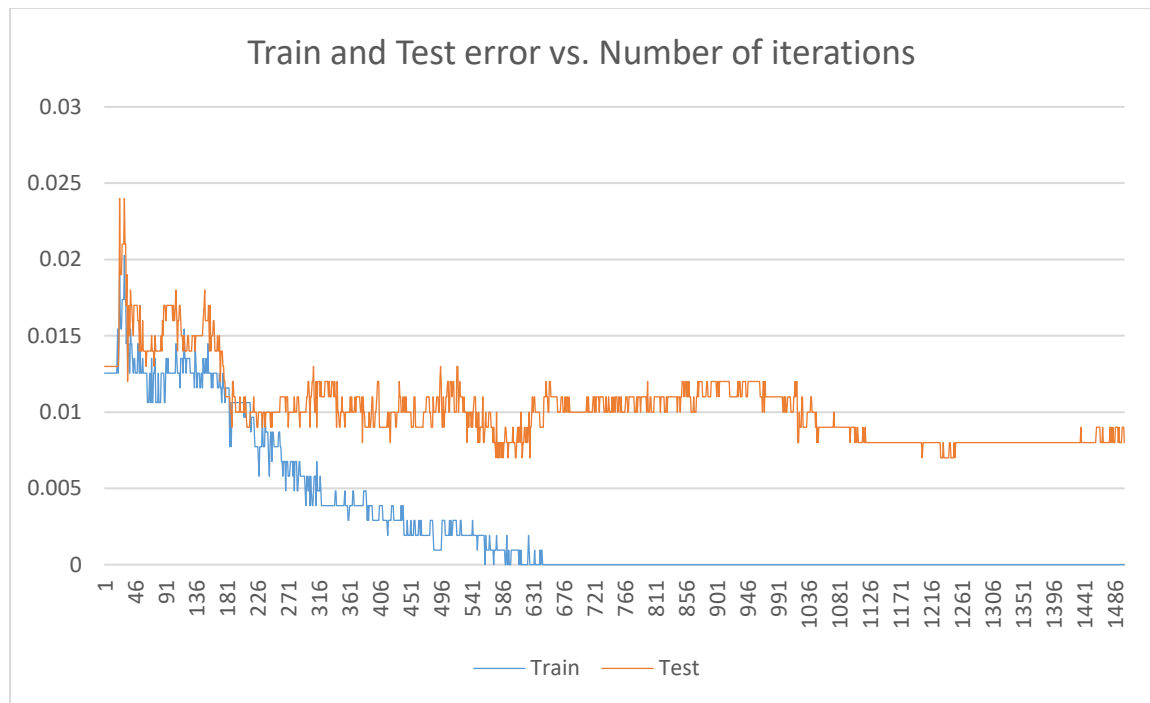
1. Unlike linear regression, Gaussian process doesn't give any coefficients. Hence, feature importance can't be inferred.
2. Gaussian process won't scale well for large datasets as kernel computations will take large time.

d)

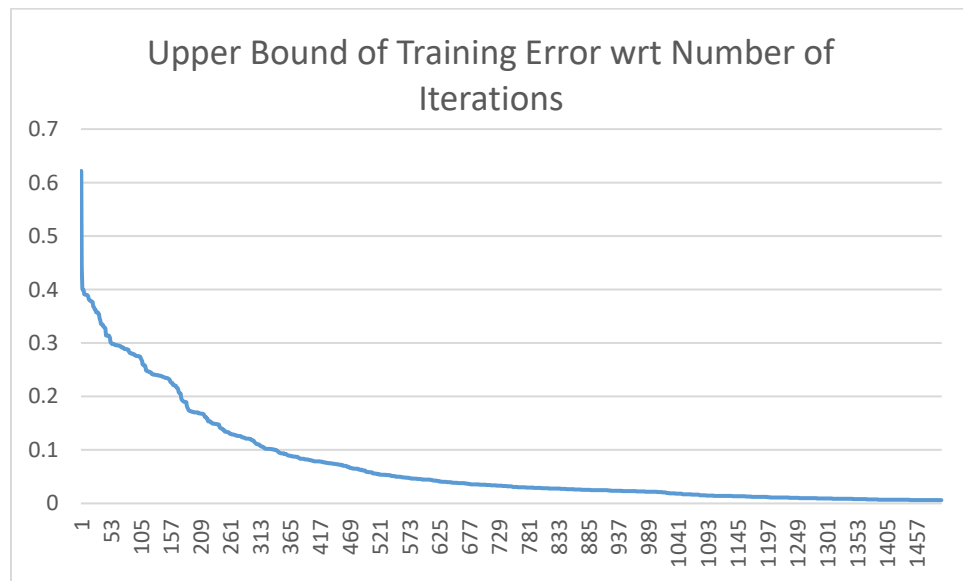


## Problem 2

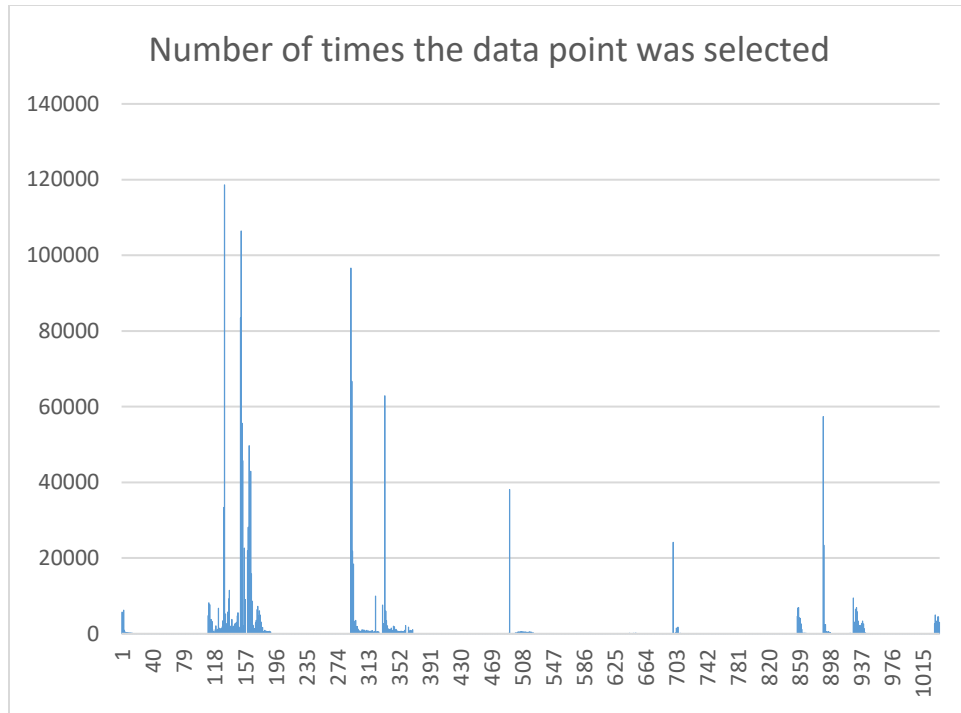
a)



b)



c)



d)

