1

Table 1: Classification accuracies on training and validation data

(a) DATA SET 1

Model	**training**	**validation**
1	48.15	48.5
2	96.8	96.3
3	96.8	96.3
4	96.8	96.3
5	96.8	96.3

(b) DATA SET 2

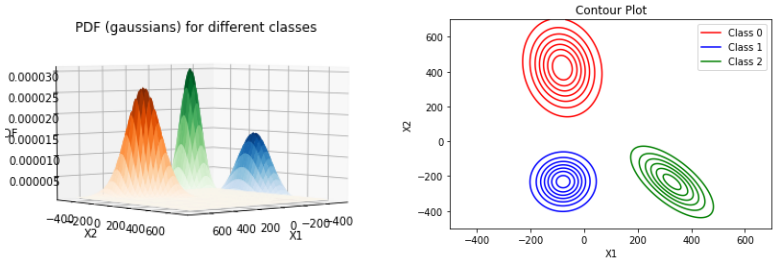
Model	**training**	**validation**
1	92.88	91.69
2	93.49	92.88
3	94.03	93.17
4	93.59	92.14
5	94.79	93.92

Table 2: Confusion matrix on test set -Model -5,DATA SET 1.

Class0	207	6	8	93.66
	30.71	0.89	1.19	6.34
Class1	4	226	4	96.58
	0.59	33.53	0.59	3.42
Class2	4	2	213	97.26
	0.59	0.30	31.60	2.74
	96.28	96.58	94.67	95.84
	3 .72	3.42	5.33	4.16
	Class0	Class1	Class2	

Table 3: Confusion matrix on test set -Model -5,DATA SET 2.

Class0	197	4	7	94.71
	29.22	0.59	0.10	5.29
Class1	6	237	1	97.13
	0.89	35.16	0.59	2.87
Class2	12	8	202	91
	1.78	1.18	29.97	9
	91.63	95.18	96.19	94.36
	8.37	4.72	3.81	5.64
	Class0	Class1	Class2	



(a) pdf,DATA SET 1 (b) contourplots,DATA SET 1.

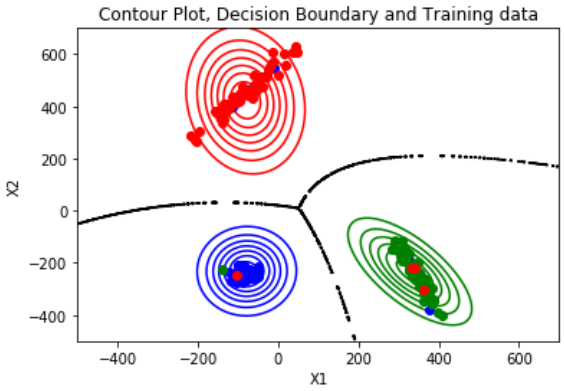


Figure 2: Decision boundary and Decision surface,DATA SET 1

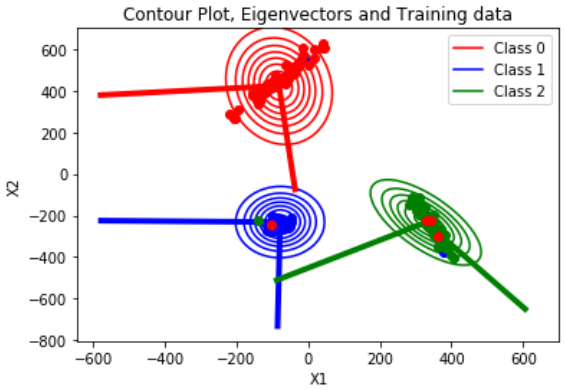
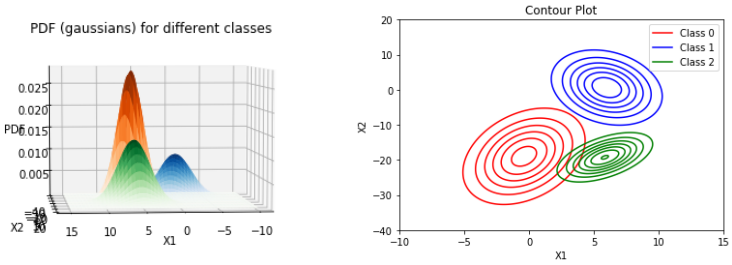


Figure 3: eigenvectors,DATA SET 1



(a) pdf,DATA SET 2 (b) contourplots,DATA SET 2.

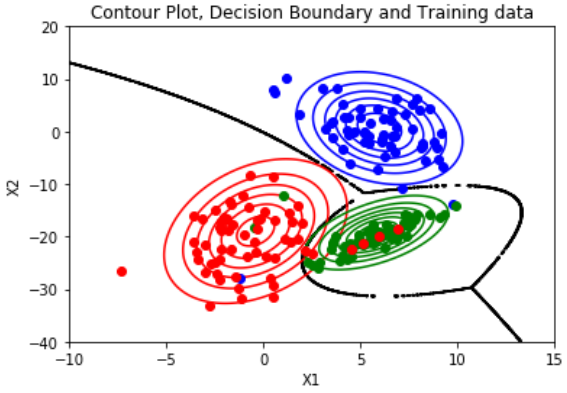


Figure 5: Decision boundary and Decision surface,DATA SET 2

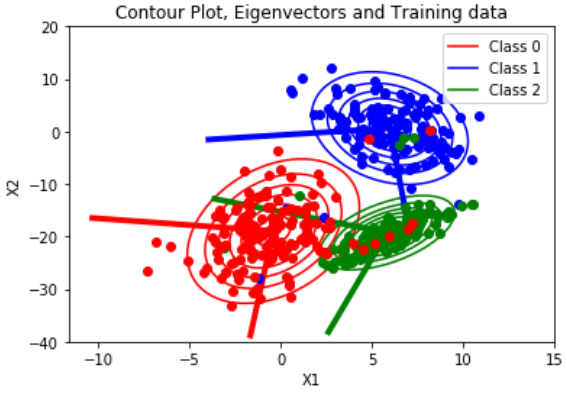


Figure 6: eigenvectors,DATA SET 2

2

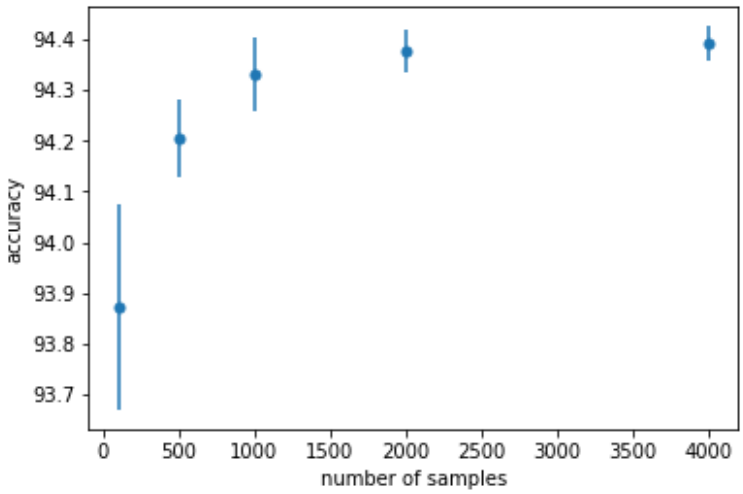


Figure 7: Averaged over 20 replicas varying data size vs accuracy

As training sample size increases the standard error in accuracy decreases
Number of training samples needed to get 85 percent is 20

3

Table 4: Classification accuracies on training data

(a) DATA SET 3

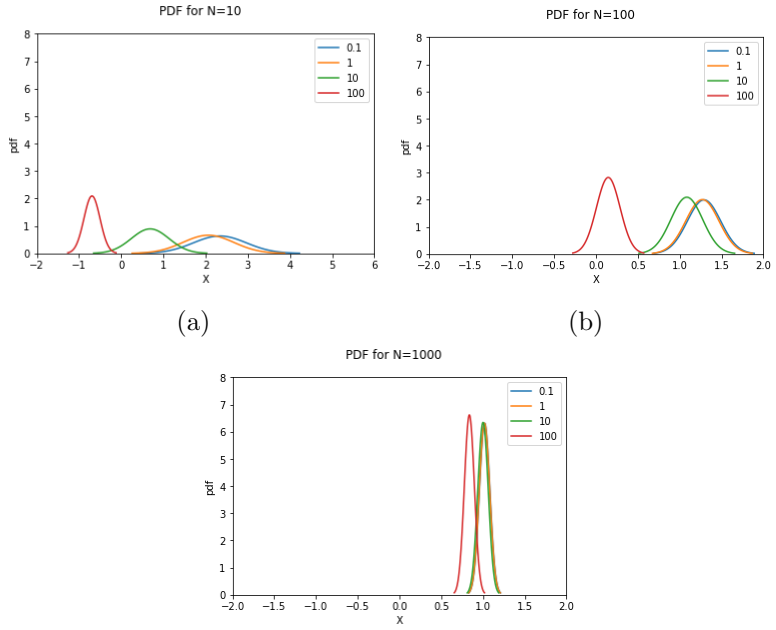
features	**accuracy**	**error**
1	83.50	16.50
2	87.03	12.97
3	31.03	68.97

(b) DATA SET 4

features	**accuracy**	**error**
1	76.67	23.33
2	80.50	19.50
3	83.77	16.23

We can see a drop in accuracy when we considered all the three features of data set 3 As the number of features increases to train the model,we need more training data but here we are using a fixed training data set,so it is possible that error increases with more number of features However in case of data set 4 accuracy increases as we consider all the three features.

4

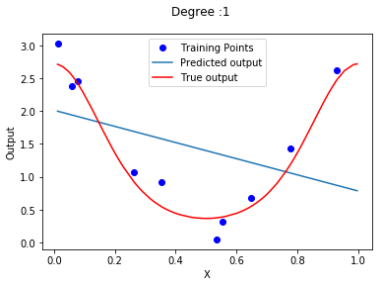


With each addditional observation the uncertaininty about the true value of mean decreases,In our plot we see the shift of mean and decrease in varaince

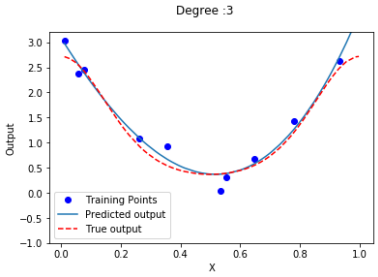
As the ratio of variances increases the mean shifts toward prior mean and updated variance decreses

5

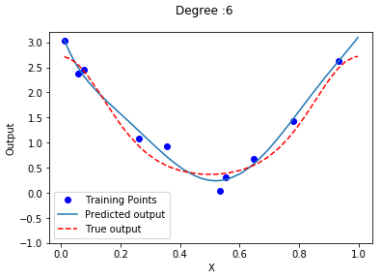
Normal regression



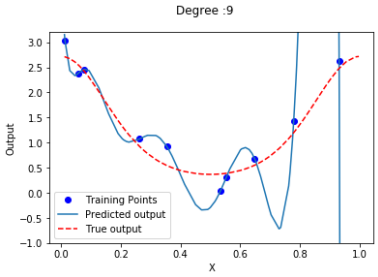
(a) degree 1.



(b) degree 3.



(c) degree 6.

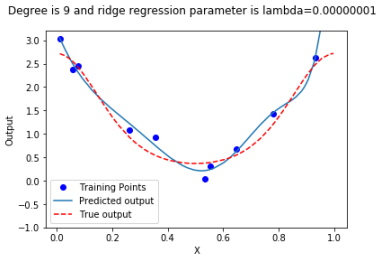


(d) degree 9.

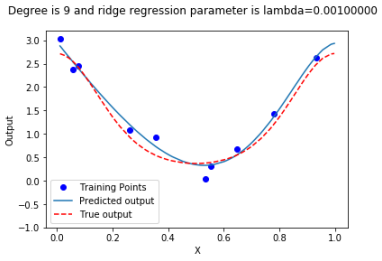
Degree	Train error	Valid error	Test error
1	0.9283008115778021	0.971499902634403	0.5649857203143979
3	0.16130934063028565	1.1021676649503447	1.0271045078692982
6	0.12138157352252413	1.1617315830580974	0.9754174639925292
9	0.0014145153022794915	7.619038944554978	59.045916792231274

(e)

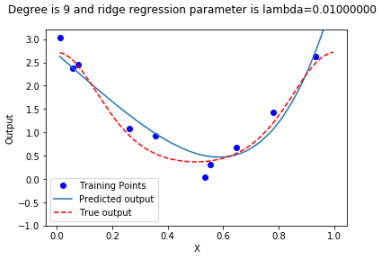
Regularisation degree 9



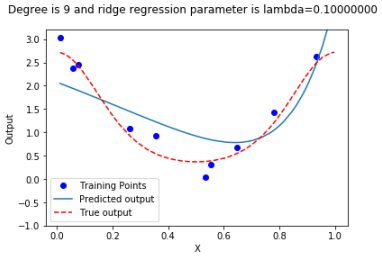
(a)



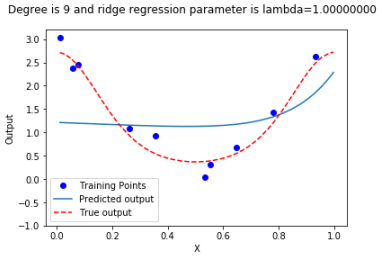
(b)



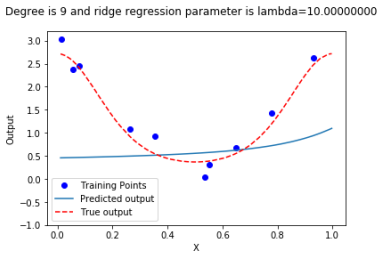
(c)



(d)



(e)



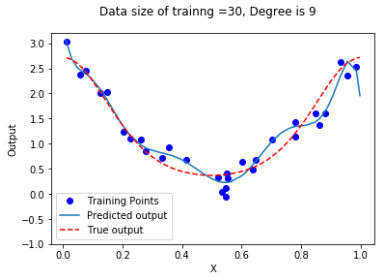
(f)

Ridge regression

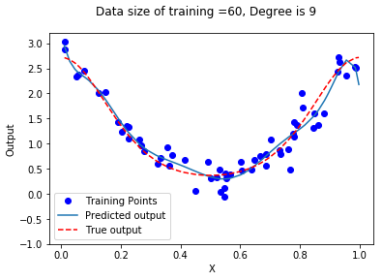
Lambda	Train error	Valid error
10^{-8}	0.11058	1.07982
10^{-3}	0.143255	1.15079
0.01	0.263423	1.06158
0.1	0.53419	0.957797
1	0.950917	0.820512
10	1.35736	0.201919

(g)

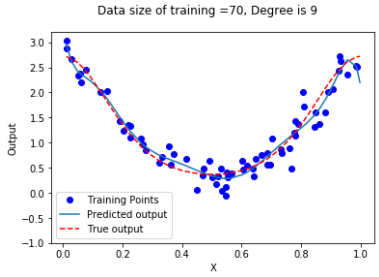
Varying data size



(a)



(b)

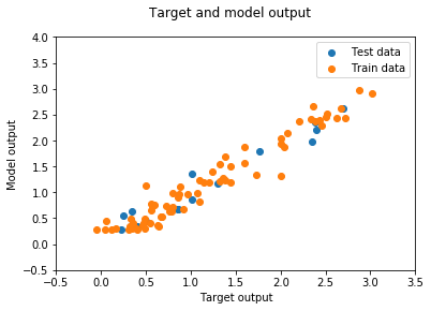


(c)

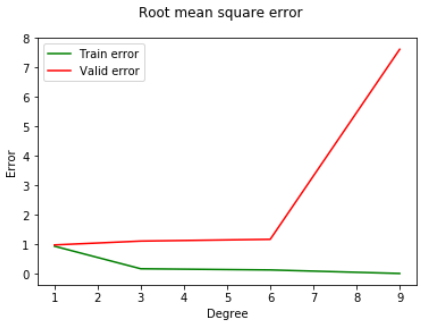
Data Size	Train error	Valid error
10	0.0014145153022794915	7.619038944554978
30	0.14359859648995235	0.9906333781862198
60	0.20996586692904454	0.9952349440997669
70	0.20056975079060063	0.9866803983214962

(d)

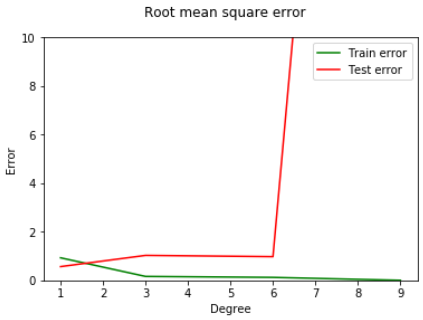
Training Data size 70,degree 9 is chosen as best model on basis of tables



(a) target-model



(b) RMS error-train vs valid



(c) RMS error-train vs test

Coefficient of	1 degree	3 degree	6 degree	9 degree
w^9				-297983
w^8				1.13025e+06
w^7				-1.77808e+06
w^6			143.813	1.50374e+06
w^5			-477.462	-740726
w^4			584.435	215263
w^3		3.68456	-320.335	-35669.9
w^2		6.32403	86.0133	3091.45
w^1	-1.22698	-9.50108	-16.5657	-124.955
w^0	2.01277	3.08138	3.21685	4.19054

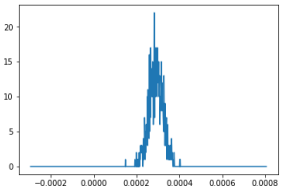
(a)

Coefficient of	ln(λ)= −18.42	ln(λ)= −6.9077	ln(λ)= −4.605	ln(λ)= −2.3025	ln(λ)= 0	ln(λ)= 2.3025
w^9	276.76	-2.35335	-0.462206	0.621865	0.200221	0.0479843
w^8	-290.3	-1.52781	-0.100236	0.685354	0.210766	0.0522241
w^7	-232.859	-0.440062	0.365758	0.745059	0.218586	0.0569209
w^6	147.792	0.881629	0.927632	0.786794	0.220697	0.0621078
w^5	258.223	2.30068	1.53208	0.780796	0.211638	0.0677987
w^4	-138.462	3.47236	2.01729	0.665298	0.18138	0.0739988
w^3	-46.8632	3.67343	1.95876	0.3107	0.11149	0.0808511
w^2	43.226	1.40446	0.284229	-0.553412	-0.0290924	0.0897927
w^1	-14.4715	-7.43556	-5.28043	-2.28264	-0.233213	0.116394
w^0	3.19854	2.96045	2.68583	2.07551	1.21495	0.456602

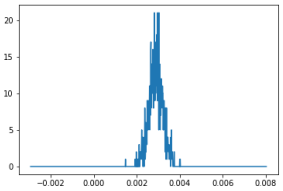
(b)

6

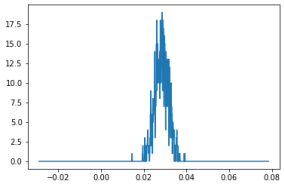
Histograms- zoom to observing variance



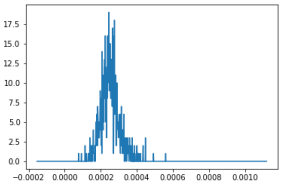
(a) deg-1 lamda-0.001



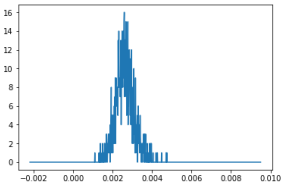
(b) deg-1 lamda-0.01



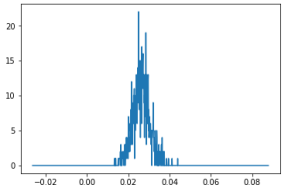
(c) deg-1 lamda-0.1



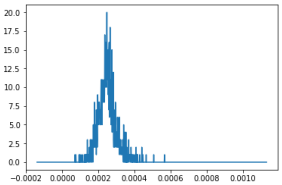
(a) deg-5 lamda-0.001



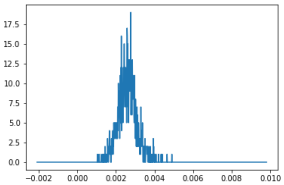
(b) deg-5 lamda-0.01



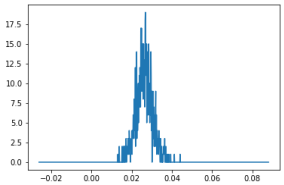
(c) deg-5 lamda-0.1



(a) deg-9 lamda-0.001



(b) deg-9 lamda-0.01

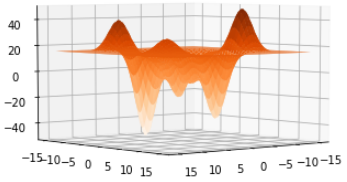


(c) deg-9 lamda-0.1

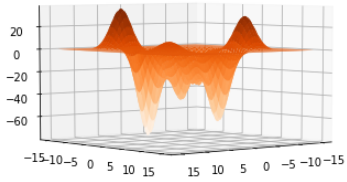
For a given degree with increase in lambda value the values of emperical risk increase from the order 0.0001 to 0.001 to 0.1 and also the variance in the emperical risk increases from 0.0002 to 0.002 to 0.02
For a given degree as lambda increases there is a slight decrease in bias
At a given lambda as the degree increases bias decreases,In linear model as the bias is high variance is very low

7

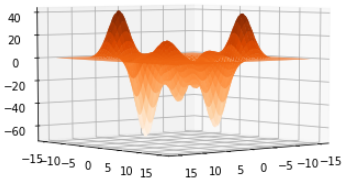
Best model training size 2000,K=10,variance=3



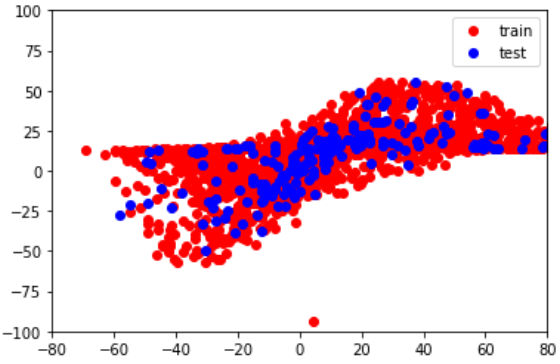
(a) training data size=100 ,K =10 ,variance=3



(b) training data size=1000 ,K =10 ,variance=3



(c) training data size=2000 ,K =10 ,variance=3



(d) training data size=2000 ,K =10 ,variance=3

2.0718083205670914
1.3028668210423784
1.1013914121084512

(a) training data size increases from top to down vs error rms

1.8218922536764157
1.8058381465793523
1.8050728831527947

(b) training data size increases from top to down vs error rms

1.9417863524157424
1.9152217990443392
1.9195257113518152

(c) training data size increases from top to down vs error rms

2.071808320693018
2.0719341732976417
2.0730602417323314
2.0836942962794005
2.151781958186159
2.238533899718736

(a) lamda increases from top to down,train error rms

1.8218922537293674
1.8219451950547019
1.822420636990759
1.8270210376804983
1.8574117182975096
1.8957425828059584

(b) lamda increases from top to down,valid error rms

1.9417863525257484
1.9418962472991146
1.9428754331631657
1.951815859570481
2.0024094284687854
2.0597486884736815

(c) lamda increases from top to down,test error rms