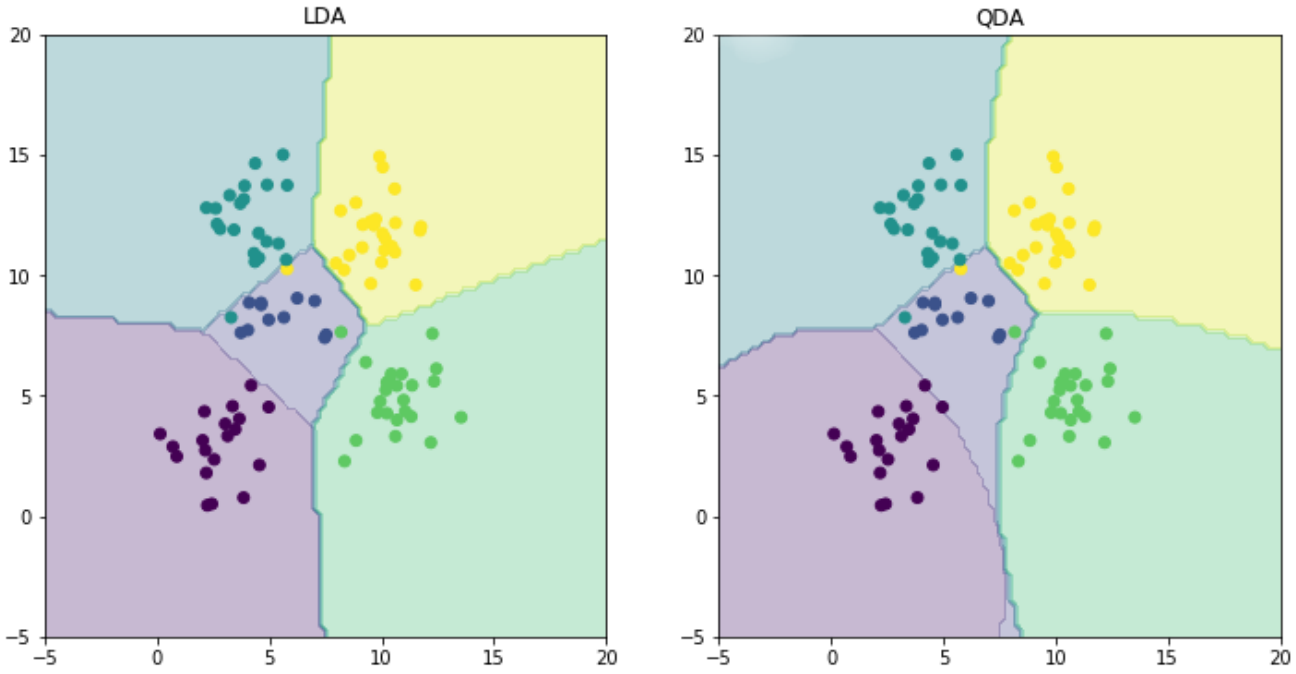
CSE 574 PROJECT 1 REPORT

Project completed by SHUBHAM SHAILESH PANDEY, UBID – spandey4, UB # - 50245725

REPORT 1. Accuracy obtained from Linear Discriminant Analysis was 97%

Accuracy obtained from Quadratic Discriminant Analysis was 96%

```
LDA Accuracy = 0.97
QDA Accuracy = 0.96
```



In the ldalearn method, we searched for unique values of y and found that the data can be classified into 5 classes. The boundaries for LDA and QDA are different because, for LDA we assume the same covariance matrix for all the classes. This leads to linear decision surfaces for the classes. But for QDA, we compute a different covariance matrix for each class and that leads to quadratic decision surfaces.

REPORT 2.

After performing linear regression on the training and test data, these were the observations for mean squared errors:

- MSE for training data without intercept was 19099.44684457055
- MSE for test data without intercept was 106775.36155354623
- MSE for training data with intercept was 2187.1602949303892
- MSE for test data with intercept was 3707.840181277104

Our aim is to reduce the mean squared error as much as possible so that the line fits the data as closely as possible. We see that for both training and test data, the MSE is significantly reduced if make use of the intercept. Hence, using the intercept is the better choice.

REPORT 3.

Lambda = 0.0	Training Data MSE = 2187.16029493	Test Data MSE = 3707.84018128
Lambda = 0.01	Training Data MSE = 2306.83221793	Test Data MSE = 2982.44611971
Lambda = 0.02	Training Data MSE = 2354.07134393	Test Data MSE = 2900.97358708
Lambda = 0.03	Training Data MSE = 2386.7801631	Test Data MSE = 2870.94158888
Lambda = 0.04	Training Data MSE = 2412.119043	Test Data MSE = 2858.00040957
Lambda = 0.05	Training Data MSE = 2433.1744367	Test Data MSE = 2852.66573517
Lambda = 0.06	**Training Data MSE = 2451.52849064**	**Test Data MSE = 2851.33021344**
Lambda = 0.07	Training Data MSE = 2468.07755253	Test Data MSE = 2852.34999406
Lambda = 0.08	Training Data MSE = 2483.36564653	Test Data MSE = 2854.87973918
Lambda = 0.09	Training Data MSE = 2497.74025857	Test Data MSE = 2858.44442115
Lambda = 0.1	Training Data MSE = 2511.43228199	Test Data MSE = 2862.75794143
Lambda = 0.11	Training Data MSE = 2524.60003852	Test Data MSE = 2867.63790917
Lambda = 0.12	Training Data MSE = 2537.35489985	Test Data MSE = 2872.96228271
Lambda = 0.13	Training Data MSE = 2549.77688678	Test Data MSE = 2878.64586939
Lambda = 0.14	Training Data MSE = 2561.92452773	Test Data MSE = 2884.62691417
Lambda = 0.15	Training Data MSE = 2573.84128774	Test Data MSE = 2890.85910969
Lambda = 0.16	Training Data MSE = 2585.55987497	Test Data MSE = 2897.30665895
Lambda = 0.17	Training Data MSE = 2597.10519217	Test Data MSE = 2903.94112629
Lambda = 0.18	Training Data MSE = 2608.49640025	Test Data MSE = 2910.73937213
Lambda = 0.19	Training Data MSE = 2619.74838623	Test Data MSE = 2917.68216413
Lambda = 0.2	Training Data MSE = 2630.8728232	Test Data MSE = 2924.75322165
Lambda = 0.21	Training Data MSE = 2641.87894616	Test Data MSE = 2931.93854417
Lambda = 0.22	Training Data MSE = 2652.77412633	Test Data MSE = 2939.22592987
Lambda = 0.23	Training Data MSE = 2663.56430077	Test Data MSE = 2946.60462378
Lambda = 0.24	Training Data MSE = 2674.25429667	Test Data MSE = 2954.06505602
Lambda = 0.25	Training Data MSE = 2684.84807809	Test Data MSE = 2961.59864341
Lambda = 0.26	Training Data MSE = 2695.34893502	Test Data MSE = 2969.19763677
Lambda = 0.27	Training Data MSE = 2705.75962912	Test Data MSE = 2976.85500119
Lambda = 0.28	Training Data MSE = 2716.0825067	Test Data MSE = 2984.56432079
Lambda = 0.29	Training Data MSE = 2726.31958674	Test Data MSE = 2992.31972181
Lambda = 0.3	Training Data MSE = 2736.4726296	Test Data MSE = 3000.11580946
Lambda = 0.31	Training Data MSE = 2746.54319109	Test Data MSE = 3007.94761559
Lambda = 0.32	Training Data MSE = 2756.53266482	Test Data MSE = 3015.81055453
Lambda = 0.33	Training Data MSE = 2766.44231574	Test Data MSE = 3023.70038563
Lambda = 0.34	Training Data MSE = 2776.27330654	Test Data MSE = 3031.61318093
Lambda = 0.35	Training Data MSE = 2786.02671854	Test Data MSE = 3039.54529713
Lambda = 0.36	Training Data MSE = 2795.70356824	Test Data MSE = 3047.49335111
Lambda = 0.37	Training Data MSE = 2805.30482034	Test Data MSE = 3055.45419817
Lambda = 0.38	Training Data MSE = 2814.83139806	Test Data MSE = 3063.42491285
Lambda = 0.39	Training Data MSE = 2824.28419133	Test Data MSE = 3071.40277169
Lambda = 0.4	Training Data MSE = 2833.66406312	Test Data MSE = 3079.38523776
Lambda = 0.41	Training Data MSE = 2842.97185452	Test Data MSE = 3087.36994673
Lambda = 0.42	Training Data MSE = 2852.2083886	Test Data MSE = 3095.35469418
Lambda = 0.43	Training Data MSE = 2861.3744735	Test Data MSE = 3103.33742413
Lambda = 0.44	Training Data MSE = 2870.47090474	Test Data MSE = 3111.31621849
Lambda = 0.45	Training Data MSE = 2879.49846701	Test Data MSE = 3119.28928746

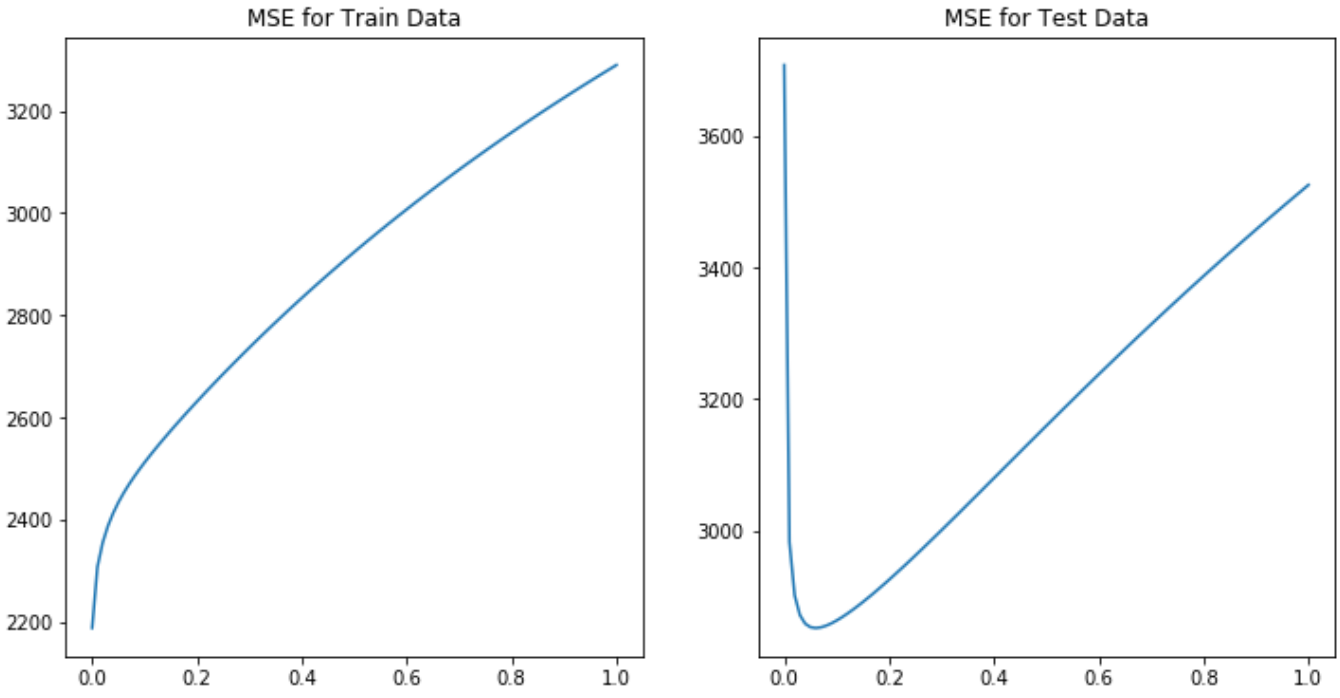
Lambda = 0.46	Training Data MSE = 2888.45793552	Test Data MSE = 3127.25496075
Lambda = 0.47	Training Data MSE = 2897.35007697	Test Data MSE = 3135.21167941
Lambda = 0.48	Training Data MSE = 2906.17565032	Test Data MSE = 3143.15798839
Lambda = 0.49	Training Data MSE = 2914.93540723	Test Data MSE = 3151.09252966
Lambda = 0.5	Training Data MSE = 2923.63009243	Test Data MSE = 3159.01403582
Lambda = 0.51	Training Data MSE = 2932.26044392	Test Data MSE = 3166.92132421
Lambda = 0.52	Training Data MSE = 2940.82719309	Test Data MSE = 3174.81329145
Lambda = 0.53	Training Data MSE = 2949.33106473	Test Data MSE = 3182.68890838
Lambda = 0.54	Training Data MSE = 2957.77277699	Test Data MSE = 3190.54721533
Lambda = 0.55	Training Data MSE = 2966.15304137	Test Data MSE = 3198.38731777
Lambda = 0.56	Training Data MSE = 2974.47256259	Test Data MSE = 3206.20838225
Lambda = 0.57	Training Data MSE = 2982.73203851	Test Data MSE = 3214.00963255
Lambda = 0.58	Training Data MSE = 2990.93215999	Test Data MSE = 3221.79034621
Lambda = 0.59	Training Data MSE = 2999.07361078	Test Data MSE = 3229.5498512
Lambda = 0.6	Training Data MSE = 3007.15706742	Test Data MSE = 3237.28752288
Lambda = 0.61	Training Data MSE = 3015.1831991	Test Data MSE = 3245.00278108
Lambda = 0.62	Training Data MSE = 3023.15266757	Test Data MSE = 3252.69508746
Lambda = 0.63	Training Data MSE = 3031.06612707	Test Data MSE = 3260.36394297
Lambda = 0.64	Training Data MSE = 3038.92422416	Test Data MSE = 3268.00888553
Lambda = 0.65	Training Data MSE = 3046.72759776	Test Data MSE = 3275.6294878
Lambda = 0.66	Training Data MSE = 3054.47687898	Test Data MSE = 3283.22535516
Lambda = 0.67	Training Data MSE = 3062.17269114	Test Data MSE = 3290.79612376
Lambda = 0.68	Training Data MSE = 3069.81564971	Test Data MSE = 3298.34145873
Lambda = 0.69	Training Data MSE = 3077.40636224	Test Data MSE = 3305.86105245
Lambda = 0.7	Training Data MSE = 3084.94542842	Test Data MSE = 3313.354623
Lambda = 0.71	Training Data MSE = 3092.43344001	Test Data MSE = 3320.82191265
Lambda = 0.72	Training Data MSE = 3099.87098085	Test Data MSE = 3328.26268646
Lambda = 0.73	Training Data MSE = 3107.25862691	Test Data MSE = 3335.67673095
Lambda = 0.74	Training Data MSE = 3114.59694628	Test Data MSE = 3343.06385289
Lambda = 0.75	Training Data MSE = 3121.88649919	Test Data MSE = 3350.42387813
Lambda = 0.76	Training Data MSE = 3129.12783807	Test Data MSE = 3357.75665047
Lambda = 0.77	Training Data MSE = 3136.3215076	Test Data MSE = 3365.0620307
Lambda = 0.78	Training Data MSE = 3143.46804472	Test Data MSE = 3372.33989556
Lambda = 0.79	Training Data MSE = 3150.56797875	Test Data MSE = 3379.59013686
Lambda = 0.8	Training Data MSE = 3157.62183137	Test Data MSE = 3386.81266063
Lambda = 0.81	Training Data MSE = 3164.63011677	Test Data MSE = 3394.00738631
Lambda = 0.82	Training Data MSE = 3171.59334168	Test Data MSE = 3401.17424594
Lambda = 0.83	Training Data MSE = 3178.51200544	Test Data MSE = 3408.31318353
Lambda = 0.84	Training Data MSE = 3185.38660008	Test Data MSE = 3415.42415428
Lambda = 0.85	Training Data MSE = 3192.21761044	Test Data MSE = 3422.50712403
Lambda = 0.86	Training Data MSE = 3199.0055142	Test Data MSE = 3429.56206859
Lambda = 0.87	Training Data MSE = 3205.75078202	Test Data MSE = 3436.58897321
Lambda = 0.88	Training Data MSE = 3212.45387757	Test Data MSE = 3443.58783202
Lambda = 0.89	Training Data MSE = 3219.11525768	Test Data MSE = 3450.55864755
Lambda = 0.9	Training Data MSE = 3225.73537241	Test Data MSE = 3457.50143021
Lambda = 0.91	Training Data MSE = 3232.31466512	Test Data MSE = 3464.41619786
Lambda = 0.92	Training Data MSE = 3238.8535726	Test Data MSE = 3471.30297539
Lambda = 0.93	Training Data MSE = 3245.35252514	Test Data MSE = 3478.16179431

Lambda = 0.94	Training Data MSE = 3251.81194665	Test Data MSE = 3484.99269234
Lambda = 0.95	Training Data MSE = 3258.23225474	Test Data MSE = 3491.79571308
Lambda = 0.96	Training Data MSE = 3264.61386081	Test Data MSE = 3498.57090566
Lambda = 0.97	Training Data MSE = 3270.95717015	Test Data MSE = 3505.3183244
Lambda = 0.98	Training Data MSE = 3277.26258207	Test Data MSE = 3512.03802854
Lambda = 0.99	Training Data MSE = 3283.53048993	Test Data MSE = 3518.7300819
Lambda = 1.0	Training Data MSE = 3289.7612813	Test Data MSE = 3525.39455263

From the table above, we see that the smallest MSE for test data is observed at **lambda = 0.06**. Hence, 0.06 is the optimal value for lambda and the same is evident from the plots below**.**

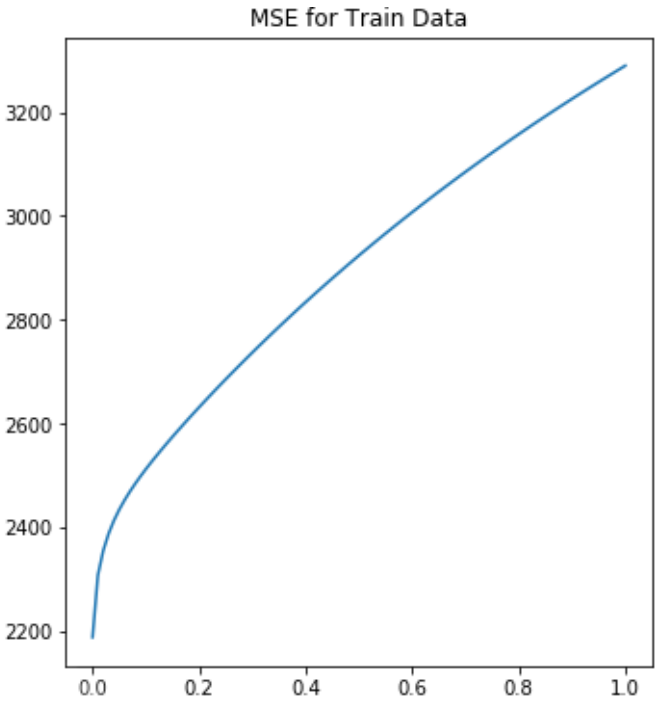
- MSE for training data from OLE is **2187.1602949303892**.
- MSE for Ridge regression training data at lambda = 0.06 is **2451.52849064**
- MSE for test data from OLE is **3707.840181277104**.
- MSE for Ridge regression test data at lambda = 0.06 is **2851.33021344**.

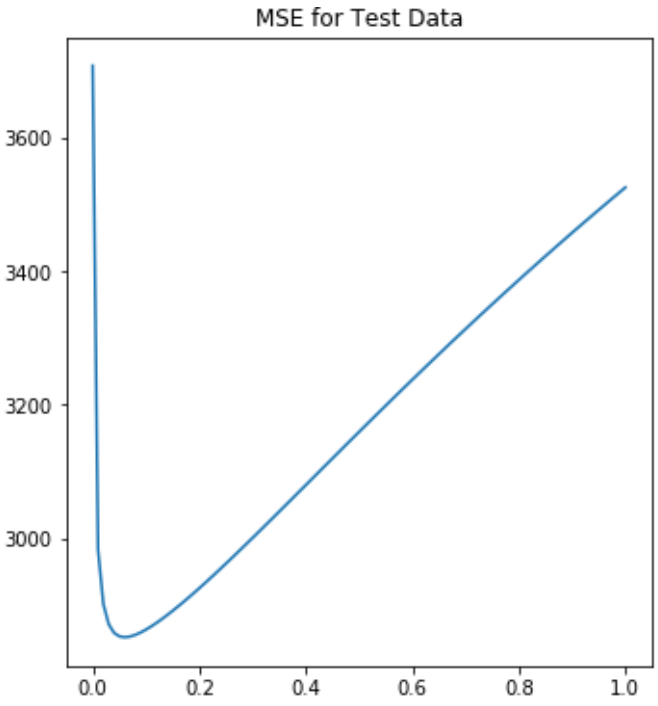
MSE value is smaller for Ridge regression than OLE at lambda = 0.06 and hence Ridge regression is the better approach for this problem.

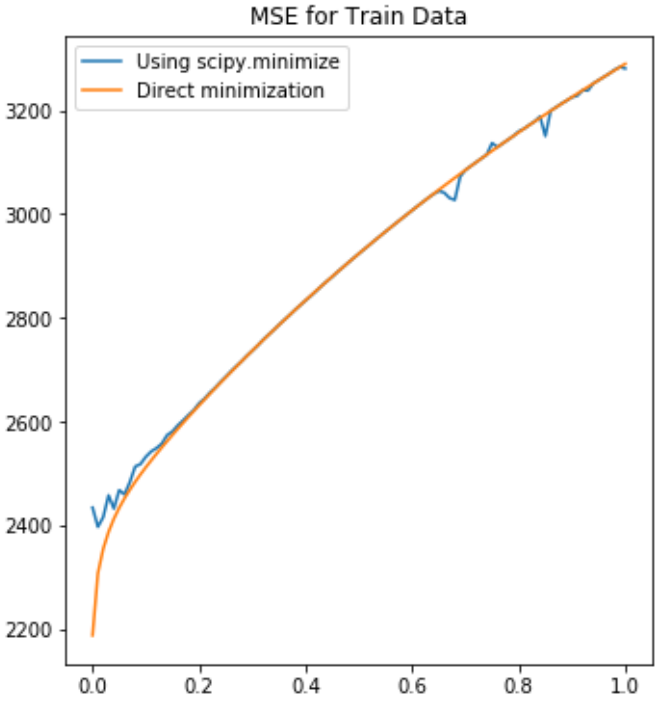


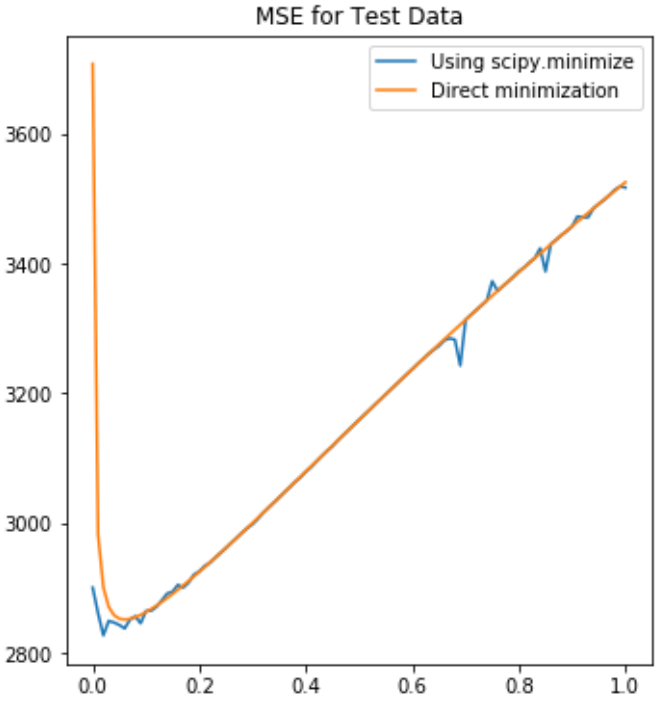
REPORT 4.

As we can see from the graphs below, the MSE functions are quite similar for Ridge regression and Ridge regression with gradient descent based learning. However, the minimize function in the gradient descent approach takes some time to converge to a point. Also, it has some visible peaks and troughs. Hence, for our problem Ridge regression using intercepts is the better approach.


MSE for Train Data
3200
3000
2800
2600
2400
2200
0.0
0.2
0.4
0.6
0.8
1.0


MSE for Test Data
3600
3400
3200
3000
0.0
0.2
0.4
0.6
0.8
1.0


MSE for Train Data
Using scipy.minimize
Direct minimization
3200
3000
2800
2600
2400
2200
0.0
0.2
0.4
0.6
0.8
1.0


MSE for Test Data
Using scipy.minimize
Direct minimization
3600
3400
3200
3000
2800
0.0
0.2
0.4
0.6
0.8
1.0

REPORT 5.

For lambda = 0, we get

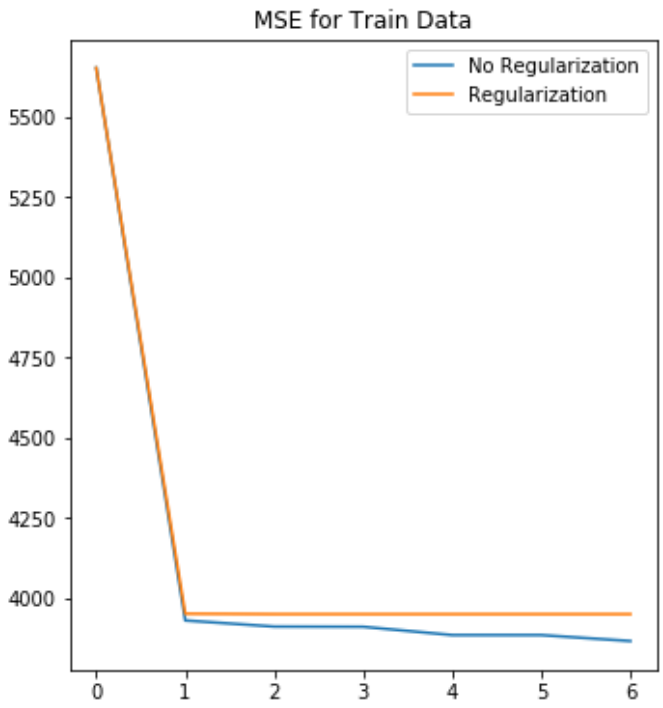
p	MSE for training data	MSE for test data
p = 0	5650.710539	6286.404792
p = 1	3930.915407	3845.03473
p = 2	3911.839671	3907.128099
p = 3	3911.188665	3887.975538
p = 4	3885.473068	4443.327892
p = 5	3885.407157	4554.830377
p = 6	3866.883449	6833.459149

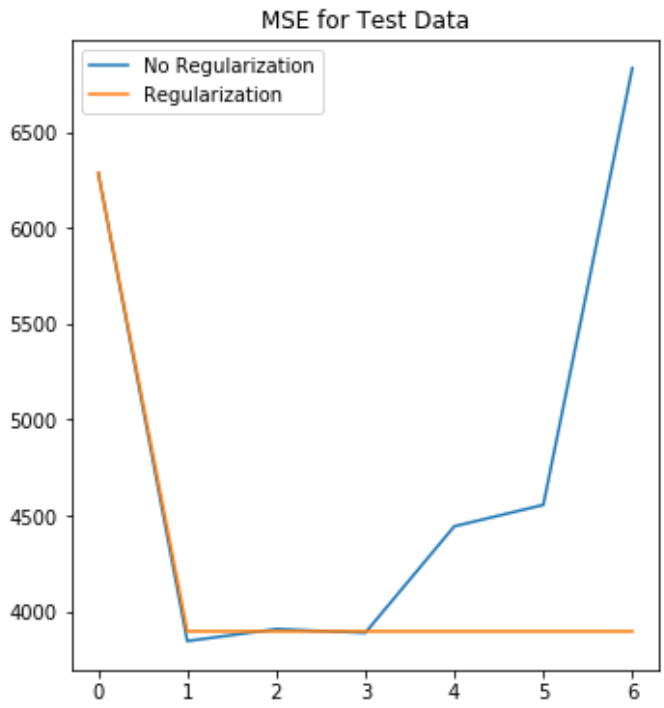
For lambda = 0.06, we get

p	MSE for training data	MSE for test data
p = 0	5650.711907	6286.881967
p = 1	3951.839124	3895.856464
p = 2	3950.687312	3895.584056
p = 3	3950.682532	3895.582716
p = 4	3950.682337	3895.582668
p = 5	3950.682335	3895.582669
p = 6	3950.682335	3895.582669

For lambda = 0, optimal value of p for test data is **1** as that gives the smallest mean squared error. (No regularization)

For lambda = 0.06, optimal value of p for test data is **4** as that gives the smallest mean squared error. (Regularization)





REPORT 6.

Technique Used	**MSE for Training Data**	**MSE for Test Data**
OLE and without intercept	19099.44684	106775.3616
OLE and with intercept	2187.160295	3707.840181
Ridge regression for lambda = 0.06	2451.528491	2851.330213
Ridge regression using Gradient descent	2451.5306897	2851.32857305
Non-linear regression without regularization	383866.883449 for p = 6	3845.03473 for p = 1
Non-linear regression with regularization	3950.682335 for p = 6	3895.582668 for p = 4

For training data, the best technique would be OLE with intercept. For test data it would be Ridge regression using Gradient descent. But, the minimize function in Gradient descent approach takes a lot of time to converge. Hence, the best approach for reducing the error would be Ridge regression for lambda = 0.06. So, for small datasets as provided in the pickle files, Ridge regression using optimal lambda is the approach to use.