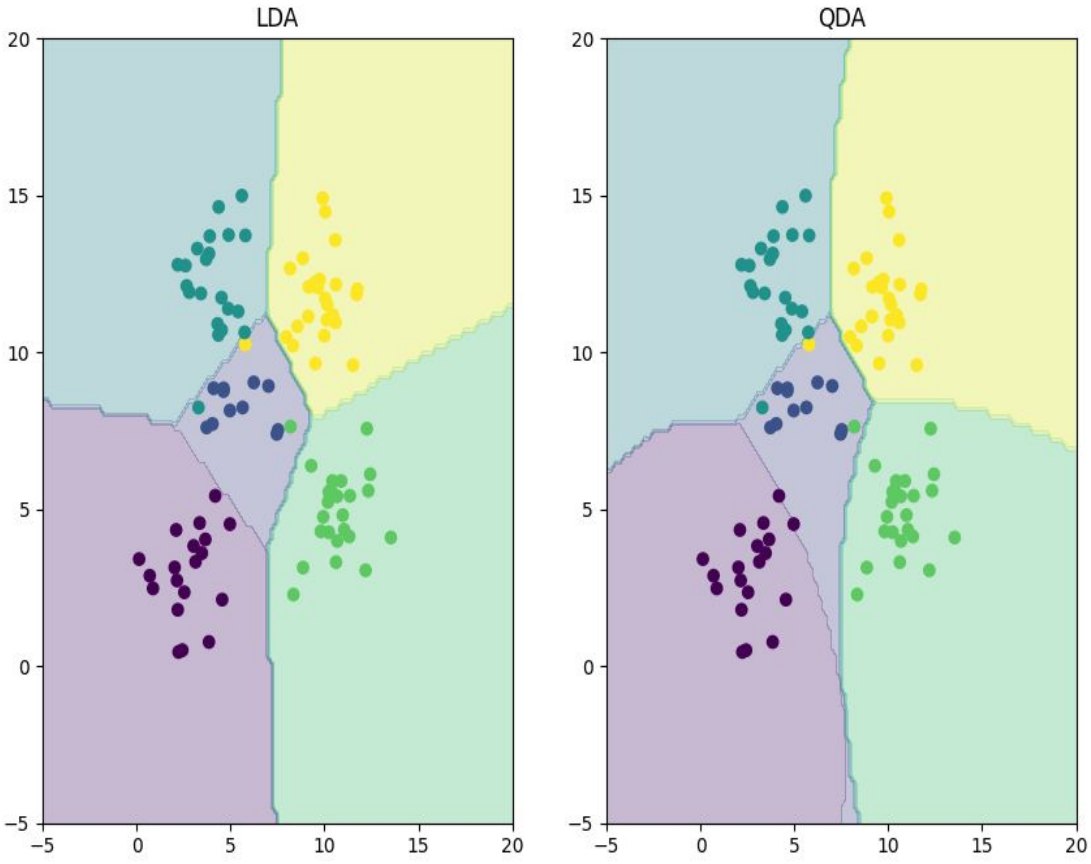
Assignment 1

Group Members:
Shreyas Pimpalkar - spimpalk@buffalo.edu

Objective:
Performing Classification and Regression analysis on given sample.pickle and diabetes.pickle datasets.

Problem 1:
Performing classification using Linear Discriminant Analysis(LDA) and Quadratic Discriminant Analysis(QDA) on sample.pickle dataset.
After creating a classification LDA and QDA model using training data, the accuracy of classification prediction of test data is:

- 97% for Linear Discriminant Analysis(LDA).
- 96% for Quadratic Discriminant Analysis(QDA).



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

Observation:
It is observed that Linear Discriminant Analysis(LDA) have linear decision boundaries, whereas Quadratic Discriminant Analysis (QDA) has parabolic decision boundaries and LDA has better accuracy than QDA model.

Problem 2:
Implementing ordinary least squares method to estimate regression parameters by minimizing least squares loss of diabetes.pickle data.
The regression model is built using a training dataset of diabetes training data with and without intercept.
Test and train MSE with and without intercept.

```
MSE without intercept 106775.36155855583
MSE with intercept 3707.840181528809
MSE without intercept for train data 19099.446844570586
MSE with intercept train data 2187.1602949303906
```

Test MSE without intercept is 106775.36155
Test MSE with intercept is 3707.84018
Train MSE without intercept is 19099.44684
Train MSE with intercept is 2187.16029

Observation:
We can see that linear regression with intercept gives much less MSE(Mean squared error) than linear regression without intercept. Therefore linear regression with intercept performs much better.
The performance of training and test data improves after adding intercept.

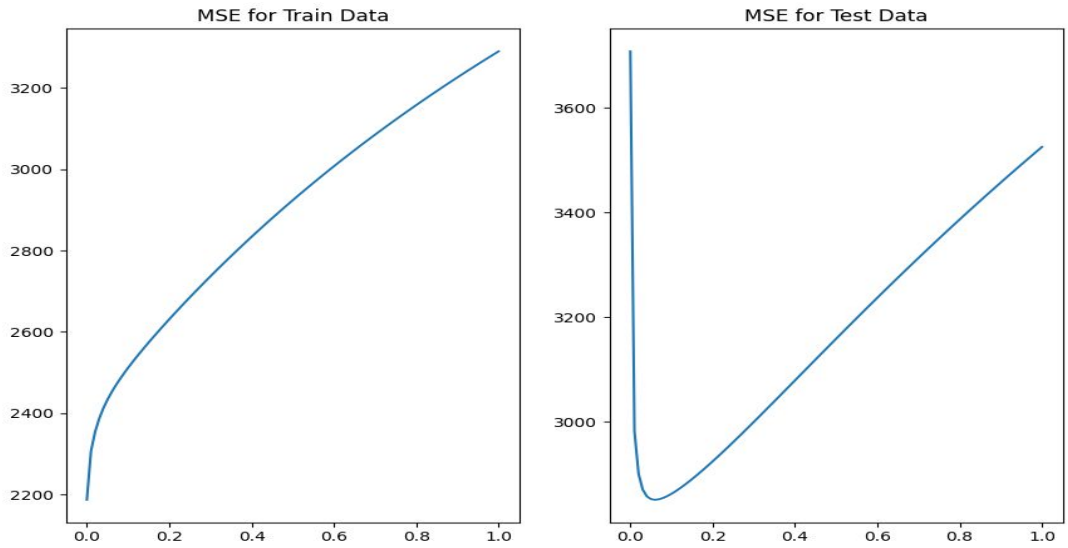
Problem 3:
Calculating MSE for training and test data using ridge regression. Value of lambda is considered between 0 and 1, training and test error is plotted. Calculating optimal value of lambda.

i) The MSE for different lambdas between 0 and 1 for training and test data using ridge regression is shown in figure below.

lambdas	mse train	mse test
0.5	2923.630092425277	3159.0140358219687
0.51	2932.260443922271	3166.9213242097485
0.52	2940.8271930941482	3174.8132914496114
0.53	2949.331064726021	3182.688908375535
0.54	2957.7727769869493	3190.5472153271176
0.55	2966.153041367787	3198.3873177745268
0.56	2974.472562592325	3206.2083822477957
0.5700000000000001	2982.7320385114676	3214.0096325461227
0.58	2990.932159988126	3221.790346205168
0.59	2999.073610778748	3229.549851202394
0.6	3007.157067416048	3237.287522882277
0.61	3015.1831990963647	3245.0027810848796
0.62	3023.15266757422	3252.69508746267
0.63	3031.0661270658675	3260.363942971775
0.64	3038.9242241631177	3268.008885525034
0.65	3046.727597758221	3275.6294877952187
0.66	3054.4768789802392	3283.2253551577696
0.67	3062.17269114309	3290.7961237632408
0.68	3069.8156497051673	3298.341458730404
0.6900000000000001	3077.406362240333	3305.8610524516866
0.7000000000000001	3084.945428419916	3313.3546230032393
0.71	3092.4334400052776	3320.82191265256
0.72	3099.87098085042	3328.262686457067
0.73	3107.25862691409	3335.676730947573
0.74	3114.5969462808066	3343.063852890999
0.75	3121.88649919019	3350.423878127147
0.76	3129.127838074031	3357.756650474659
0.77	3136.321507600479	3365.062030701684
0.78	3143.468044724789	3372.339895557086
0.79	3150.5679787460667	3379.5901368583145
0.8	3157.621831369462	3386.8126606323167
0.81	3164.6301167733045	3394.0073863061725
0.8200000000000001	3171.5933416806815	3401.1742459442917
0.8300000000000001	3178.5120054350027	3408.3131835293007
0.84	3185.3866000790986	3415.42415428389
0.85	3192.217610437456	3422.507124031102
0.86	3199.0055142011856	3429.5620685907106
0.87	3205.75078201537	3436.5889732094806
0.88	3212.4538775684564	3443.587832023271
0.89	3219.1152576833747	3450.5586475490545
0.9	3225.7353724100844	3457.501430205073
0.91	3232.3146651193033	3464.4161978574357
0.92	3238.8535725971396	3471.302975391614
0.93	3245.352525140424	3478.161794307369
0.9400000000000001	3251.811946652516	3484.9926923357166
0.9500000000000001	3258.232254739393	3491.795713076666
0.96	3264.6138608058623	3498.570905656542
0.97	3270.9571701517034	3505.318324403715
0.98	3277.2625820676244	3512.038028541751
0.99	3283.530489930882	3518.7300818989156
1.0	3289.7612813004457	3525.3945526331786

lambdas	mse train	mse test
0.0	2187.1602949303906	3707.840181528809
0.01	2306.8322179337324	2982.4461197118903
0.02	2354.0713439338283	2900.973587082239
0.03	2386.780163097977	2870.9415888843955
0.04	2412.1190430007478	2858.0004095733952
0.05	2433.1744367023966	2852.665735165676
0.06	2451.528490643497	2851.3302134438477
0.07	2468.0775525260105	2852.349994057722
0.08	2483.3656465308623	2854.8797391758376
0.09	2497.7402585657906	2858.4444211485747
0.1	2511.432281988939	2862.757941425696
0.11	2524.6000385245284	2867.6379091670997
0.12	2537.3548998459173	2872.9622827114276
0.13	2549.776886783926	2878.6458693869204
0.14	2561.9245277254972	2884.6269141677913
0.15	2573.841287742291	2890.8591096903638
0.16	2585.559874972394	2897.306658951088
0.17	2597.1051921675826	2903.9411262909825
0.18	2608.496400254902	2910.739372130538
0.19	2619.7483862258196	2917.6821641327806
0.2	2630.872823196496	2924.7532216474046
0.21	2641.8789461590886	2931.9385441674413
0.22	2652.774126329711	2939.2259298658405
0.23	2663.5643007697795	2946.604623783517
0.24	2674.2542966714577	2954.065056016316
0.25	2684.848078094598	2961.59864340977
0.26	2695.348935022924	2969.1976367703487
0.27	2705.7596291193145	2976.8550011879192
0.28	2716.0825067040846	2984.5643207941216
0.29	2726.319586736426	2992.3197218087876
0.3	2736.47262960395	3000.1158094622183
0.31	2746.543191088131	3007.947615588462
0.32	2756.5326648173937	3015.8105545342087
0.33	2766.4423157366027	3023.7003856324823
0.34	2776.2733065362486	3031.613180925095
0.35000000000000003	2786.026718543498	3039.545297133645
0.36	2795.7035682425085	3047.4933511105864
0.37	2805.3048203356966	3055.4541981735047
0.38	2814.831398061087	3063.4249128540246
0.39	2824.2841913289553	3071.402771689596
0.4	2833.6640631229566	3079.385237760068
0.4100000000000003	2842.9718545187498	3087.3699467275546
0.42	2852.2083886008577	3095.354694182061
0.43	2861.374473501618	3103.3374241295423
0.44	2870.4709047410274	3111.316218486249
0.45	2879.498467010675	3119.2892874647277
0.46	2888.457935516476	3127.2549607541823
0.4700000000000003	2897.350076972327	3135.211679411953
0.48	2906.1756503186175	3143.157988394309
0.49	2914.935407225021	3151.0925296642677

ii) Plot of errors for training and test data is shown below.



From the values of mse train and mse test we can see that:

- Optimal value of λ for train data is at 0.0 where mse is 2187.1602 which is the same as linear regression with intercept.
- Optimal value of λ for test data is at 0.06 where mse is 2851.3302.
- Therefore, we can conclude that the optimal value of λ is 0.06, Since test data is considered for evaluation.

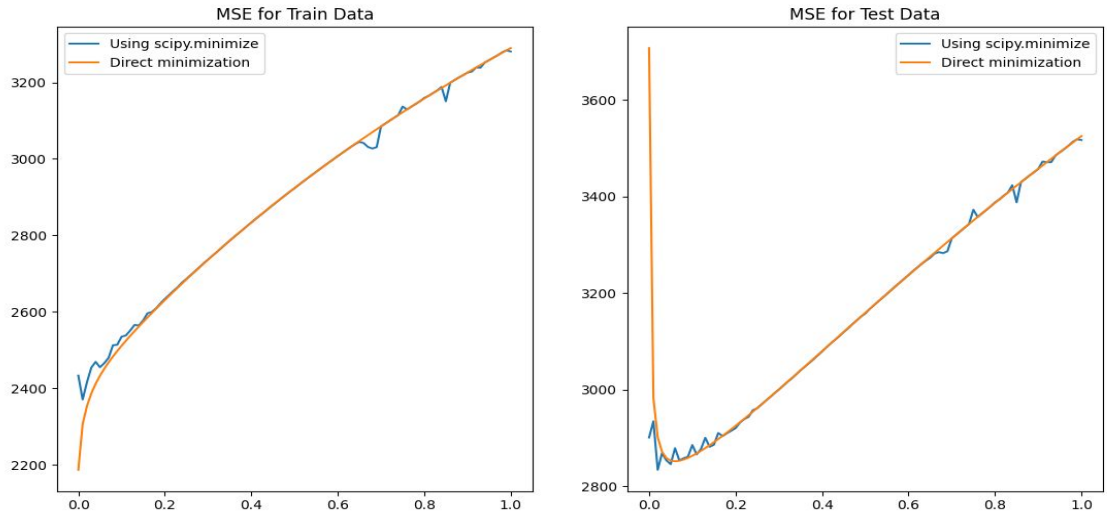
Comparing Results:

Method	Train MSE	Test MSE
Linear Regression Without Intercept	19099.44684	106775.36155
Linear Regression With Intercept	2187.1602	3707.84018
Ridge Regression For λ = 0.06	2451.528490	2851.330213
Lowest	2187.1602 (Linear Regression With Intercept)	2851.330213 (Ridge Regression for λ=0.06)

We can see from the above values that:
i) training MSE for linear regression is the lowest.
ii) Test MSE for Ridge regression is the lowest.

Therefore, Ridge regression performs better for test data than linear regression.

Problem 4:
Using Gradient descent for ridge regression.
Plot of MSE for train data and test data by using gradient descent is by varying lambda is shown below.



From the figure above we can see that the value of MSE for test data at some places is further minimised due to use of gradient descent.

For gradient descent values of Train and test MSE is shown below.

lambda	MSE train	MSE test
0.0	2433.6514329471315	2900.550626346172
0.01	2370.9776353058046	2934.2881689963024
0.02	2416.286868690745	2833.5376087813147
0.03	2454.3385068937587	2868.2677617915488
0.04	2469.5531595583925	2853.2653183274674
0.05	2455.3303182301784	2845.2680390147875
0.06	2466.0573686872353	2878.506900068766
0.07	2480.0497746240276	2853.1861401726114
0.08	2512.857851956945	2857.5946174339333
0.09	2514.109263499708	2861.0284378166
0.1	2535.2062738980485	2885.037249978651
0.11	2538.5573427332456	2865.5227759973823
0.12	2551.3896400096596	2876.6016401046963
0.13	2565.893913765323	2900.0310072210814
0.14	2564.6898007668706	2881.3043884598465
0.15	2577.814286886473	2885.5670488354917
0.16	2596.4278509742912	2909.9367981470864
0.17	2599.714901975988	2903.422500846647
0.18	2609.6369457458422	2909.2976254167547
0.19	2622.3086621030907	2914.500250068399
0.2	2633.9573920962143	2919.993558946964

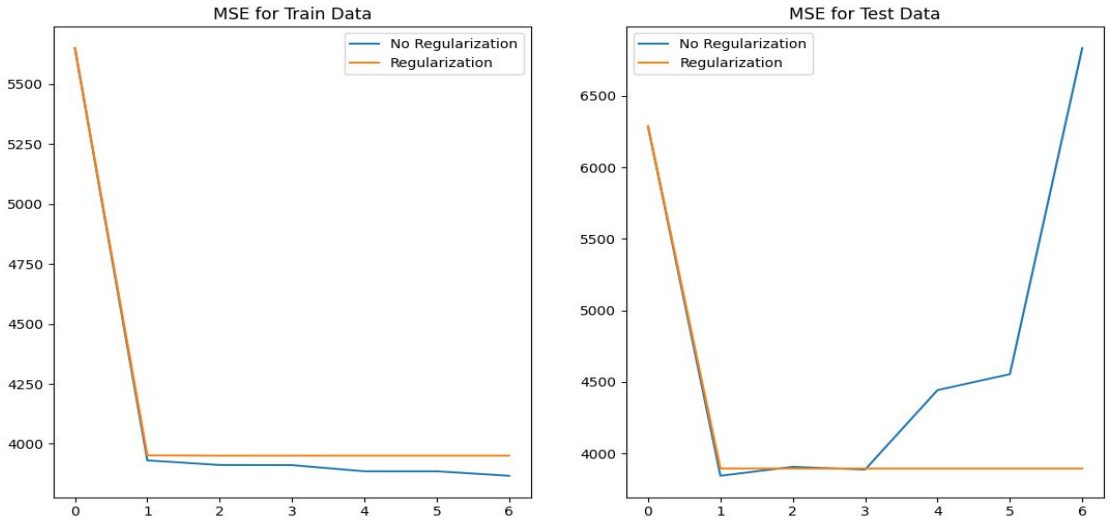
We can see that for λ = 0.05 the value of MSE is minimum for gradient descent, and is less than MSE from ridge regression MSE(λ = 0.06).

Comparing Results

METHOD	Train MSE	Test MSE
Ridge Regression For λ = 0.06	2451.528490	2851.330213
Gradient descent on ridge regression (λ = 0.05)	2455.3303	2845.2680390

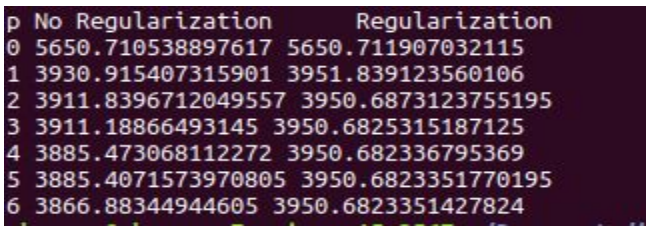
Optimal value of λ = 0.05 for gradient descent on ridge regression.

Problem 5:
Training and Test MSE error for non-linear regression.



P values are from 0 to 6.

- For train MSE, when λ = 0(No regularisation), MSE for train data decreases with increase in degree of polynomial.
- For train data, when, λ = 0.06(regularisation) with increase in degree of polynomial degree MSE value remains stable below 4000.

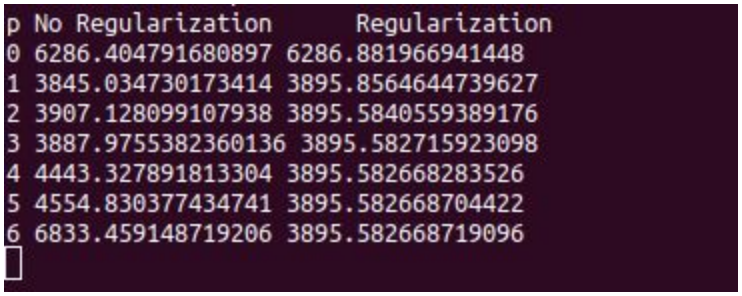


p	No Regularization	Regularization
0	5650.710538897617	5650.711907032115
1	3930.915407315901	3951.839123560106
2	3911.8396712049557	3950.6873123755195
3	3911.18866493145	3950.6825315187125
4	3885.473068112272	3950.682336795369
5	3885.4071573970805	3950.6823351770195
6	3866.88344944605	3950.6823351427824

Optimal Value:

- No regularisation λ = 0 ,MSE = 3866.88344 for p=6.
- Regularisation λ = 0.06, MSE = 3950.68233514 for p = 6.

- For Test MSE, when λ = 0(No regularisation), MSE for test data increases significantly with increase in degree of polynomial above 3 and 5.
- For test MSE, when, λ = 0.06(regularisation) with increase in degree of polynomial degree MSE value remains stable below 4000.



Optimal value of p in terms of test error.

- With No Regularisation λ = 0, p = 1, MSE = 3845.03467
- With Regularisation λ = 0.06, p= 4, MSE = 3895.5826682

Problem 6:

Interpreting Results.

Method	Train MSE	Test MSE
Linear Regression Without Intercept	19099.44684	106775.36155
Linear Regression With Intercept	2187.1602	3707.84018
Ridge Regression For λ = 0.06	2451.528490	2851.330213
Gradient descent on ridge regression (λ = 0.05)	2455.3303	2845.2680390
Non Linear Regression (No Regularisation)	3866.88344	3845.03467
Non Linear Regression Regularisation	3950.68233514	3895.5826682

Since Test MSE shows us how accurate the model is. Comparing test MSE for different methods we see that:

- Test MSE is minimum for ridge regression with Gradient descent for $\lambda = 0.05$. So we get minimum MSE on test data by using ridge regression along with gradient descent.

Therefore, the metric used to choose the best setting must be Test MSE(Mean Squared Error) since it gives the accuracy of the model. We see that Test MSE is minimum for ridge regression with gradient Descent and highest for linear regression without intercept.

Therefore, the best model is ridge regression with gradient descent and the worst performing model is linear regression without intercept.