ASSIGNMENT 1

CS5691 Pattern Recognition and Machine Learning

CS5691 Assignment 1

Team Members:

CS24M039	Sagar Biswas
CS24M043	Saurabh Kumar Sahu

Indian Institute of Technology, Madras

1 Dataset 1

1.1 Polynomial Regression

The data for uni-variate polynomial regression is prepared by raising it to the required degree. In case of uni-variate polynomial regression of degree *d*, the dependent variable, of size (*d,* 1) is assumed to have the form

$$\vec{y}_{n\times 1} = \varphi_{n\times d} W_{d\times 1} \tag{1}$$

The weights corresponding to a given degree is then calculated by using the closed form solution for uni-variate polynomial regression:

$$W = (\varphi^T \varphi + \lambda I)^{-1} \varphi^T \vec{y} \tag{2}$$

Where, λI is the regularization term.

- **Training Dataset 1a (Size-10)**

1.2 Erms Without Regularization on Train, Test and validation data

Degree	Train Erms	Validation Erms	Test Erms
3	0.24364	0.50227	0.50216
6	0.19741	0.49929	0.42895
9	0.00000	187.28234	141.01847

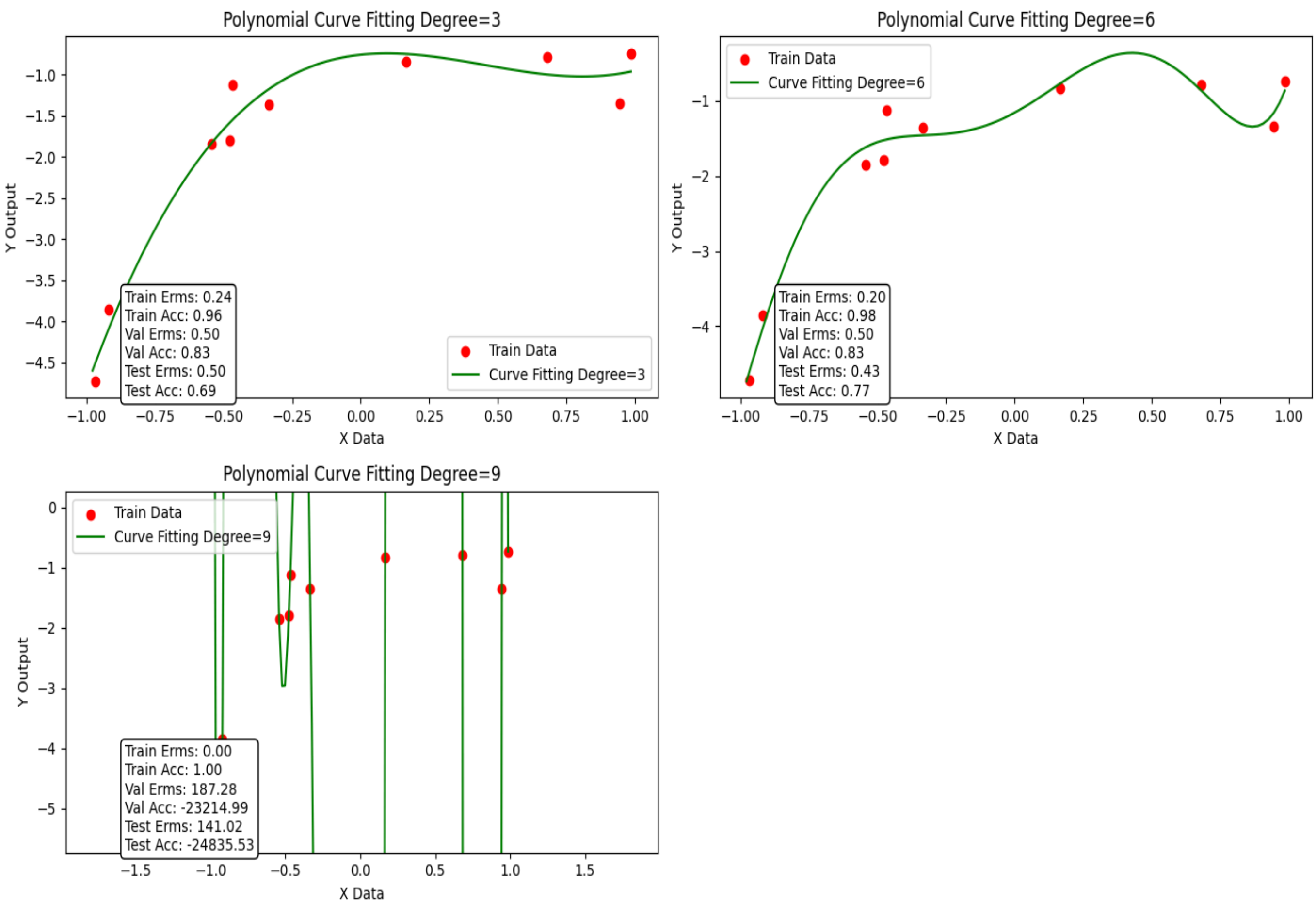


Figure 1: Curve fitting for different degree without regularization, Sample size: 10

1.3 Erms With Regularization on Train, Test and validation data and Degree=9

λ	Train Erms	Validation Erms	Test Erms
0.001	0.17	0.49	0.43
0.1	0.21	0.46	0.43
1	0.43	0.49	0.74

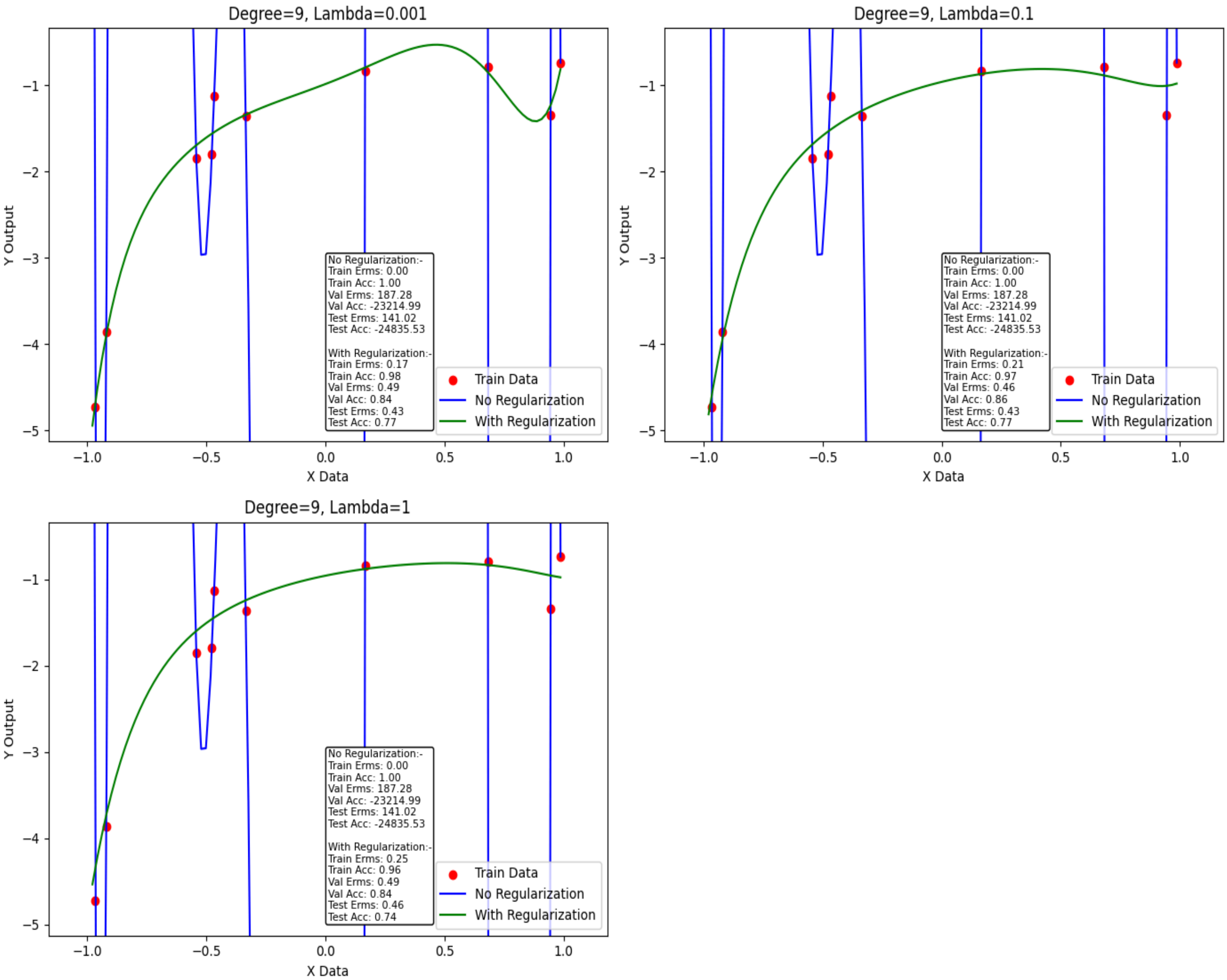


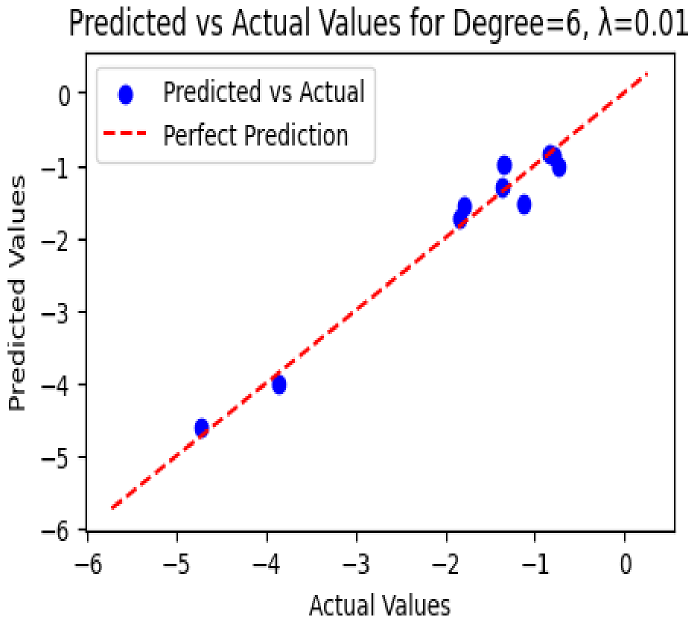
Figure 2: Curve fitting with regularization for degree=9, Sample size: 10

We see that the best fit for the data is obtained for degree: 6 and when we analysis (using validation ERMS) the model for given *λ values and different Degree(given M values) we see that* for degree= 6 and *λ* : 0.01 model gives the best performance (least ERMS in test data)

Conclusion: ERMS for best performing model: (M=6 and *λ* : 0.01)

Train Data Erms: 0.22005
Test Data Erms: 0.43787

The best fit for training data and test data, M : 6 and λ : 0.01 is visualized as follows:



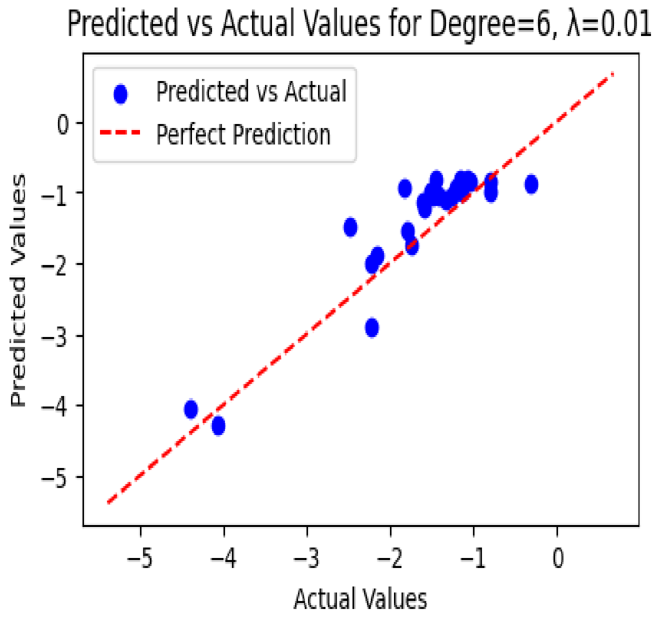


Fig: Predicted value Vs Actual Value curve for Train Data

Fig:Predicted value Vs Actual Value curve for Test data

- **Training Dataset 1b (Size-50)**

1.4 Erms Without Regularization(λ=0) on Train, Test and validation data

Degree	Train Erms	Validation Erms	Test Erms
3	0.39609	0.45764	0.41277
6	0.33425	0.57906	0.43421
9	0.33389	0.60082	0.43620

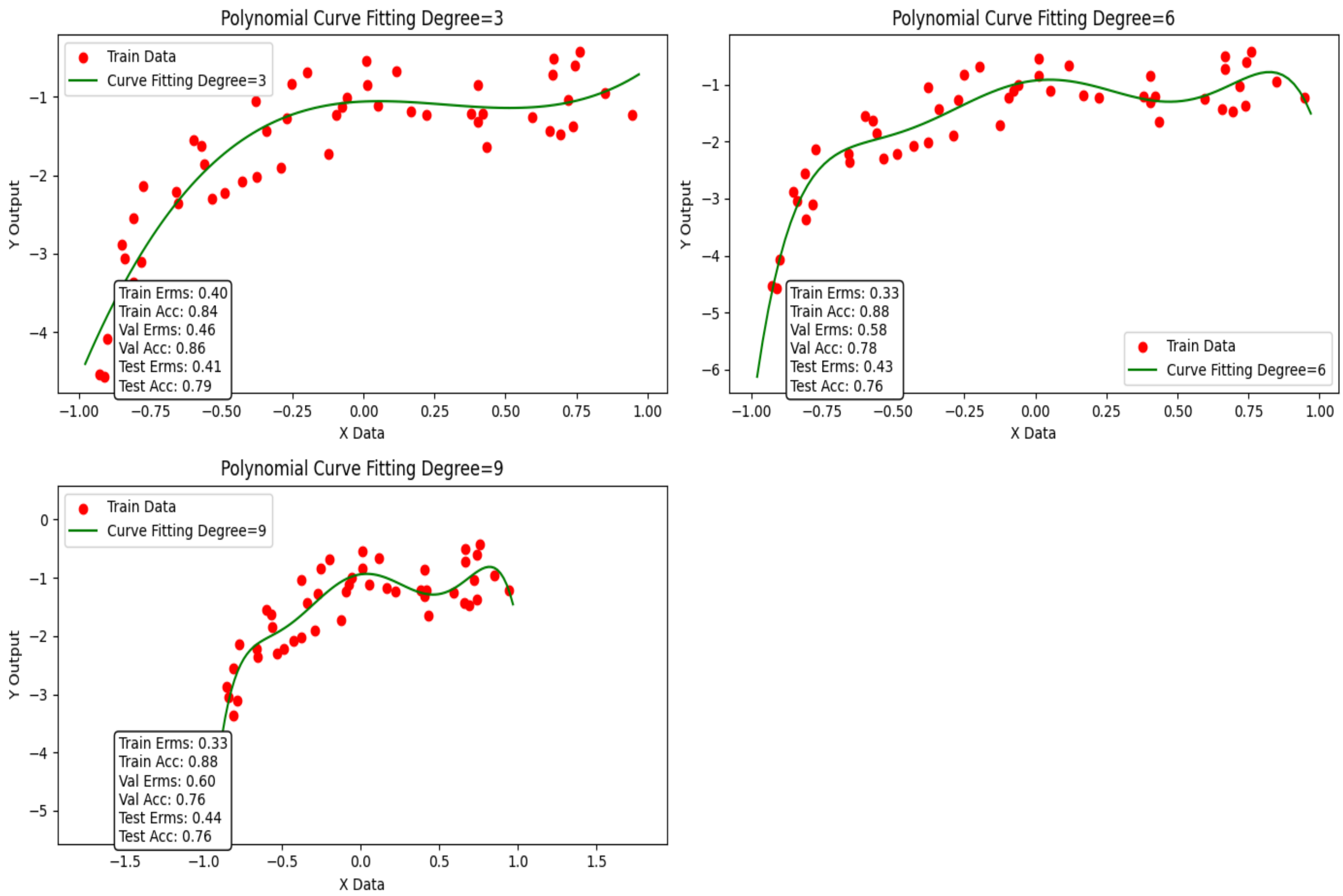


Figure 4: Curve fitting for different degree without regularization, Sample size: 50

1.5 Erms With Regularization on Train, Test and validation data and Degree=9 and Sample Size=50

λ	Train Erms	Validation Erms	Test Erms
0.001	0.33	0.61	0.44
0.1	0.37	0.44	0.38
1	0.38	0.42	0.37

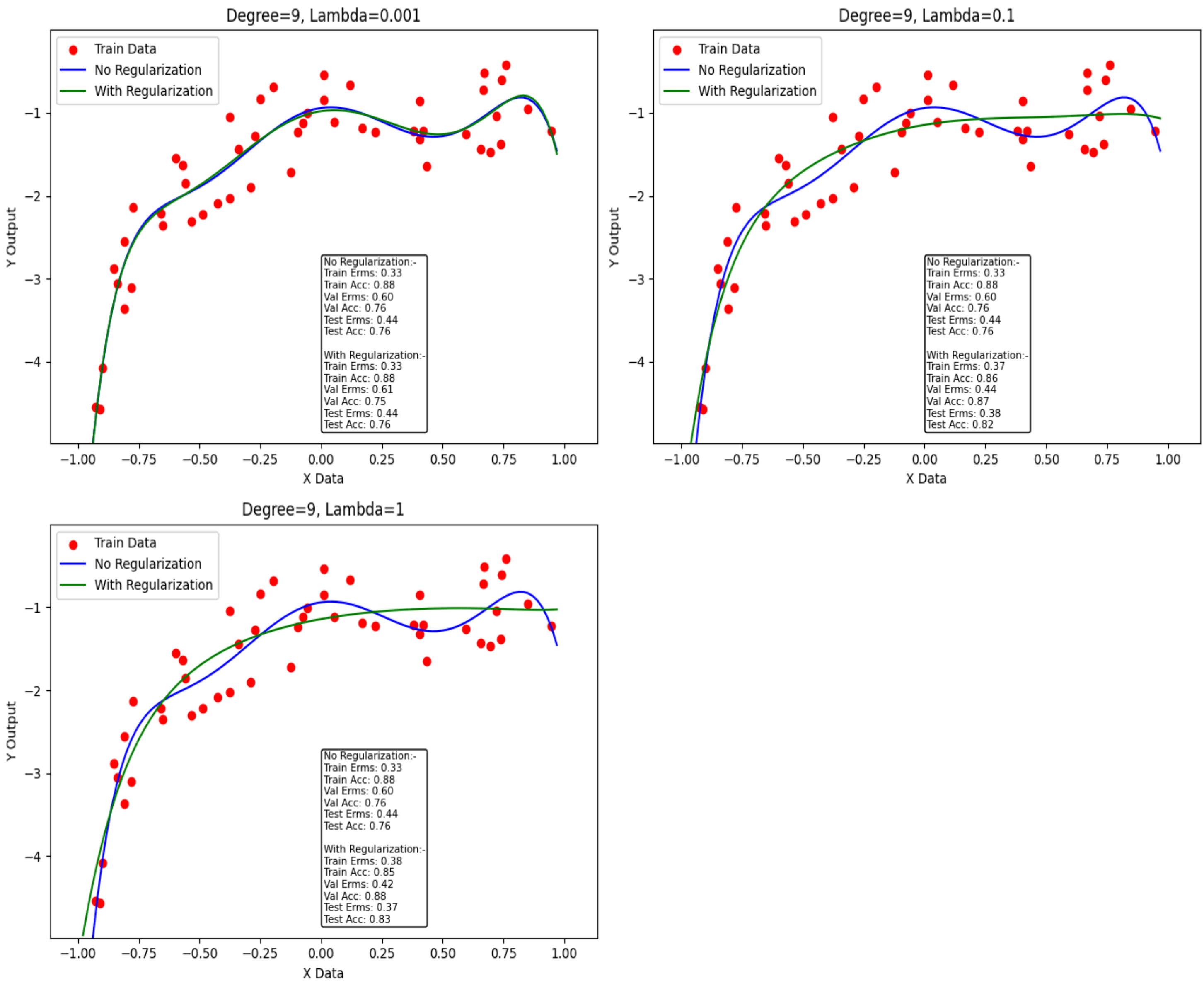


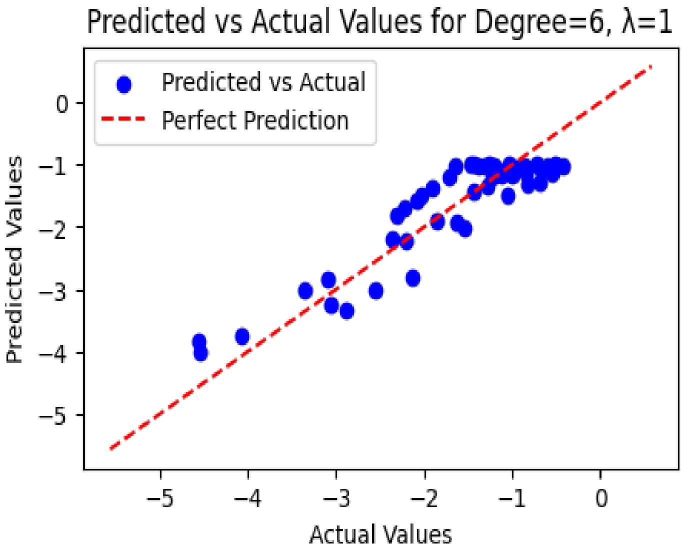
Figure 5: Curve fitting with regularization(different λ values) for degree=9, Sample size: 50

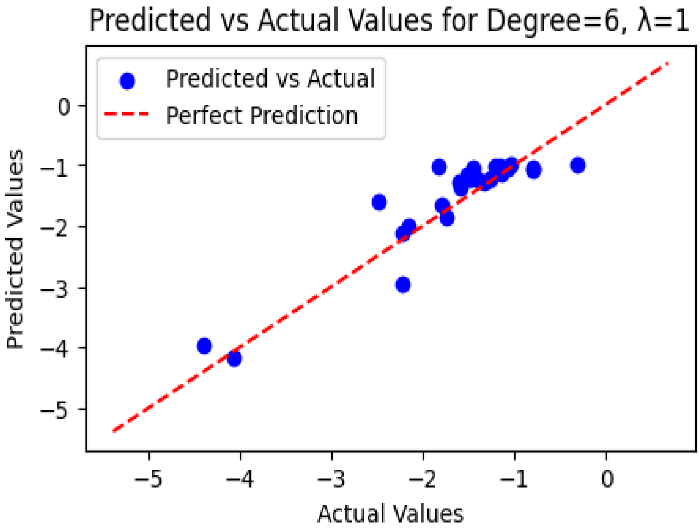
We see that the best fit for the data is obtained for degree: 6 and when we analysis (using validation ERMS) the model for given *λ values and different Degree(given M values) we see that* for degree= 6 and λ : 1 model gives the best performance (least ERMS in test data)

Conclusion: ERMS for best performing model: (M=6 and λ : 1)

Train Data Erms: 0.38641
Test Data Erms: 0.37123

The best fit for training data and test data, M : 6 and λ : 1 is visualized as follows





2 Dataset-2

- **Training Dataset 2a (Size-25)**

2.1 Erms Without Regularization(λ=0) on Train, Test and validation data and Sample Size=25

Degree	Train Erms	Validation Erms	Test Erms
2	1575.06	3680.08	3094.67
4	72.22	2540.76	3383.84
6	2937.45	7944.78	10059.44

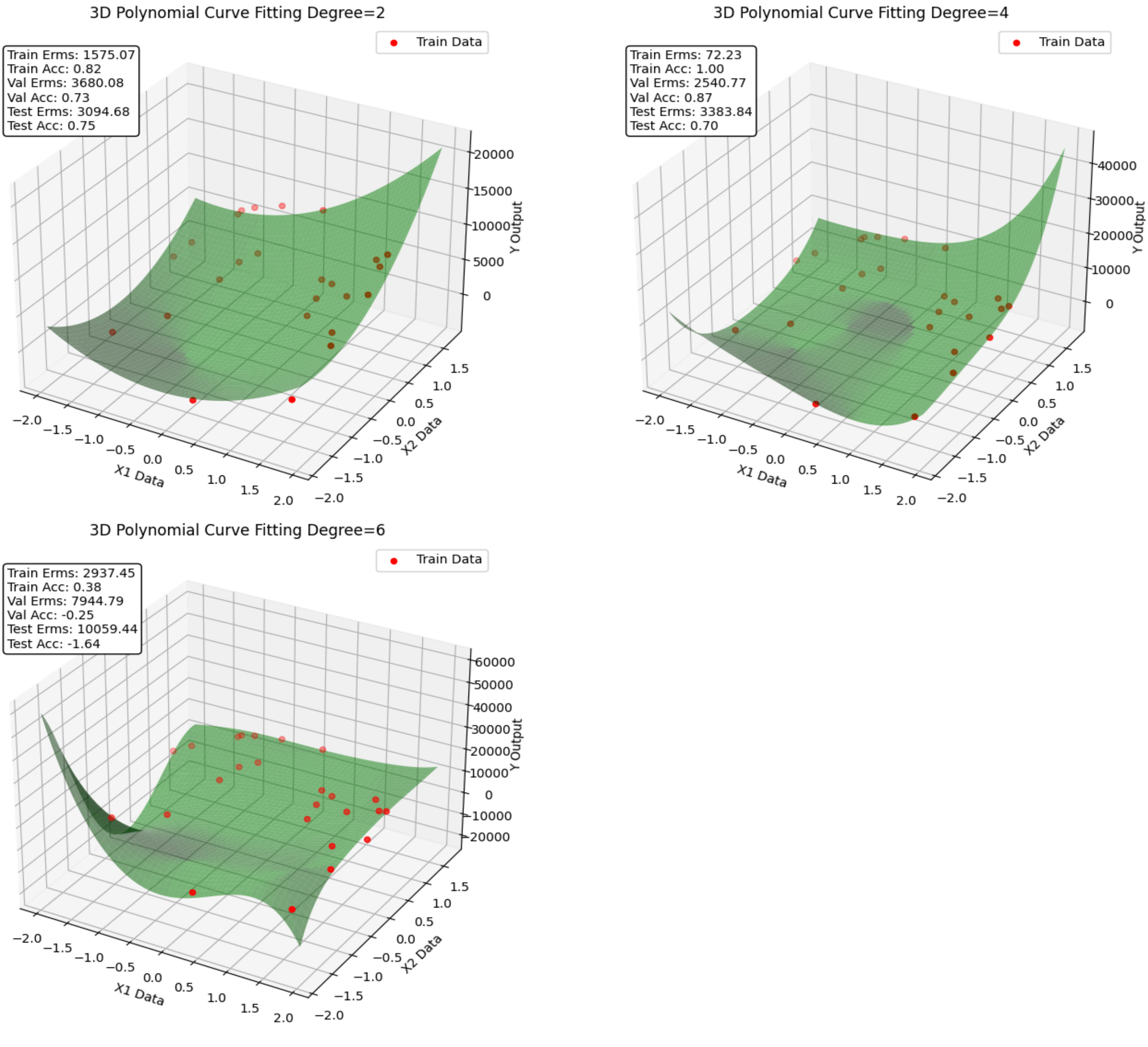
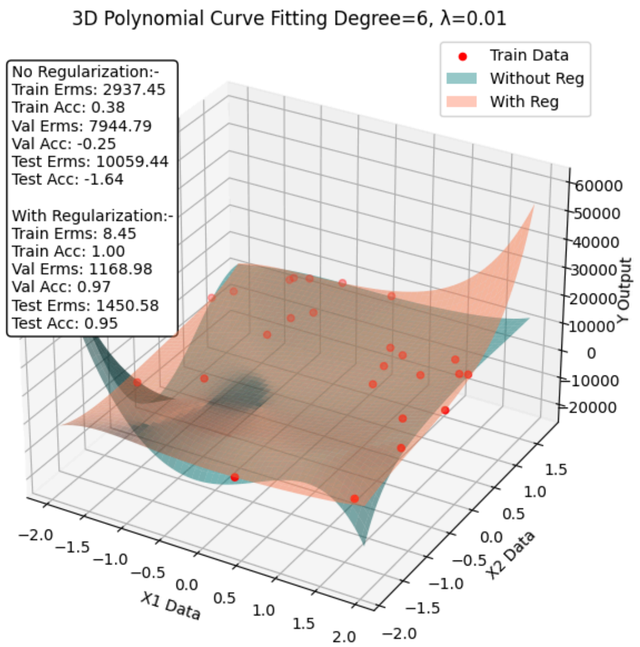
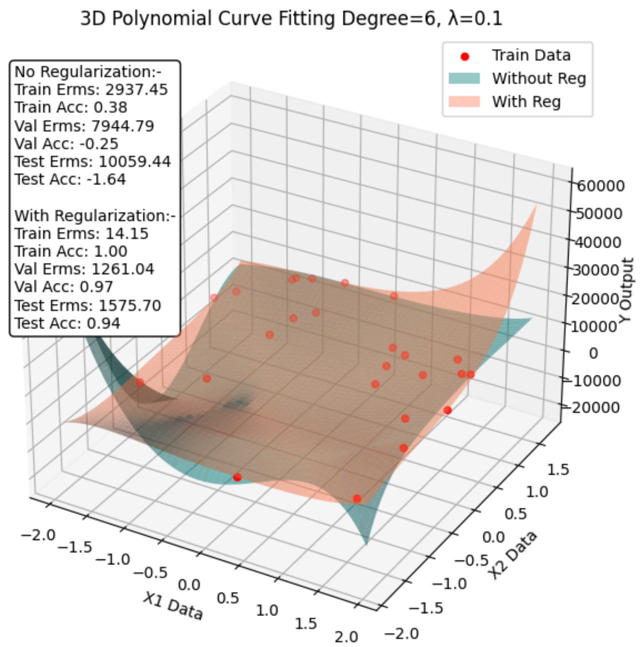


Figure 6: Surface plot for different degree without regularization, Sample size: 25

2.2 Erms With Regularization on Train, Test and validation data and Degree=6 and Sample Size=25

λ	Train Erms	Validation Erms	Test Erms
0.001	5.37106	1591.6955	1855.9731
0.1	14.1513710	1261.042214	1575.70352
1	40.47980	1573.01998	1929.4827





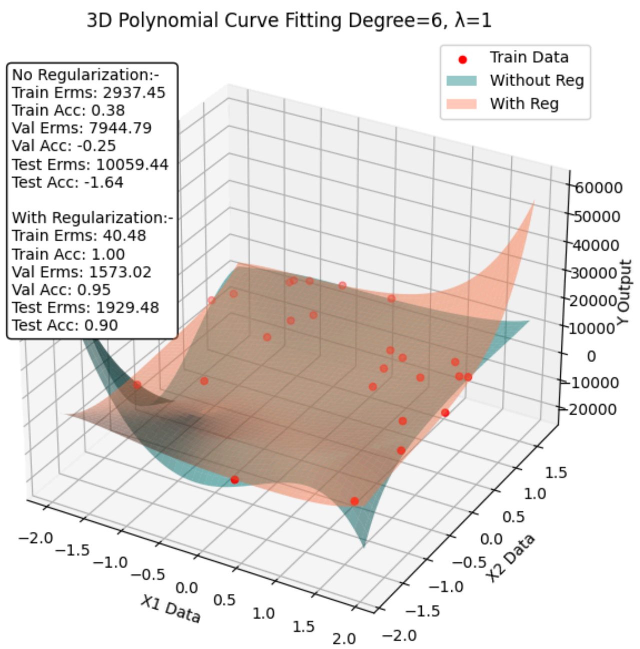


Figure 7: Surface plot for different **λ values** (two plot shown on training data set with and without regularization) , Sample size: 25

We see that the best fit for the data is obtained for degree: 4 and when we analysis (using validation data ERMS) the model for given *λ values and different Degree(given M values). we see that* for degree= 4 and λ : 1 model gives the best performance (least ERMS in test data)

Conclusion: ERMS for best performing model: (M=4 and λ : 1)

Train Data Erms: 191.8518
Test Data Erms: 1851.88464

The best fit for training data and test data, M : 4 and λ : 1 is visualized as follows

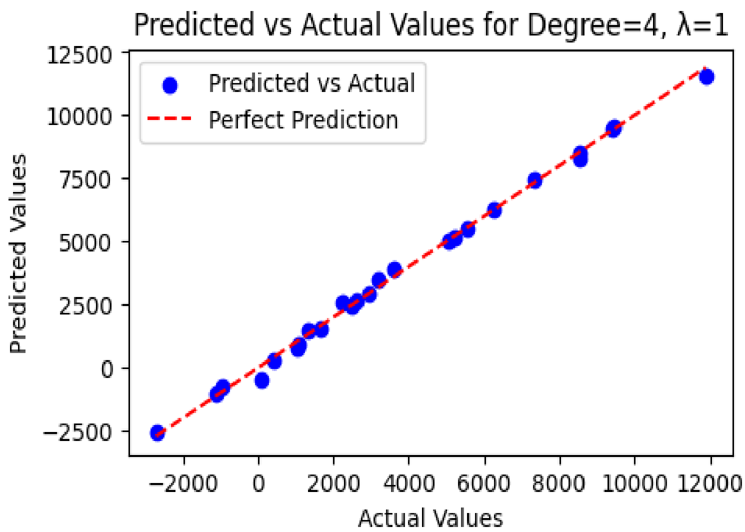


Fig: Predicted value Vs Actual Value curve for Train Data

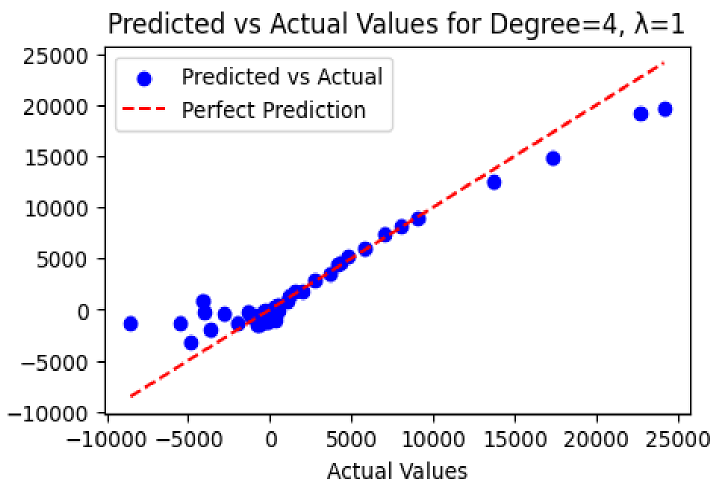


Fig: Predicted value Vs Actual Value curve for Test Data

- **Training Dataset 2b (Size-100)**

2.3 Erms Without Regularization(λ=0) on Train, Test and validation data and Sample Size=25

Degree	Train Erms	Validation Erms	Test Erms
2	3318.4021	3200.0273	2603.2541
4	372.45563	630.83930	668.49619
6	13.627326	13.877274	15.35854

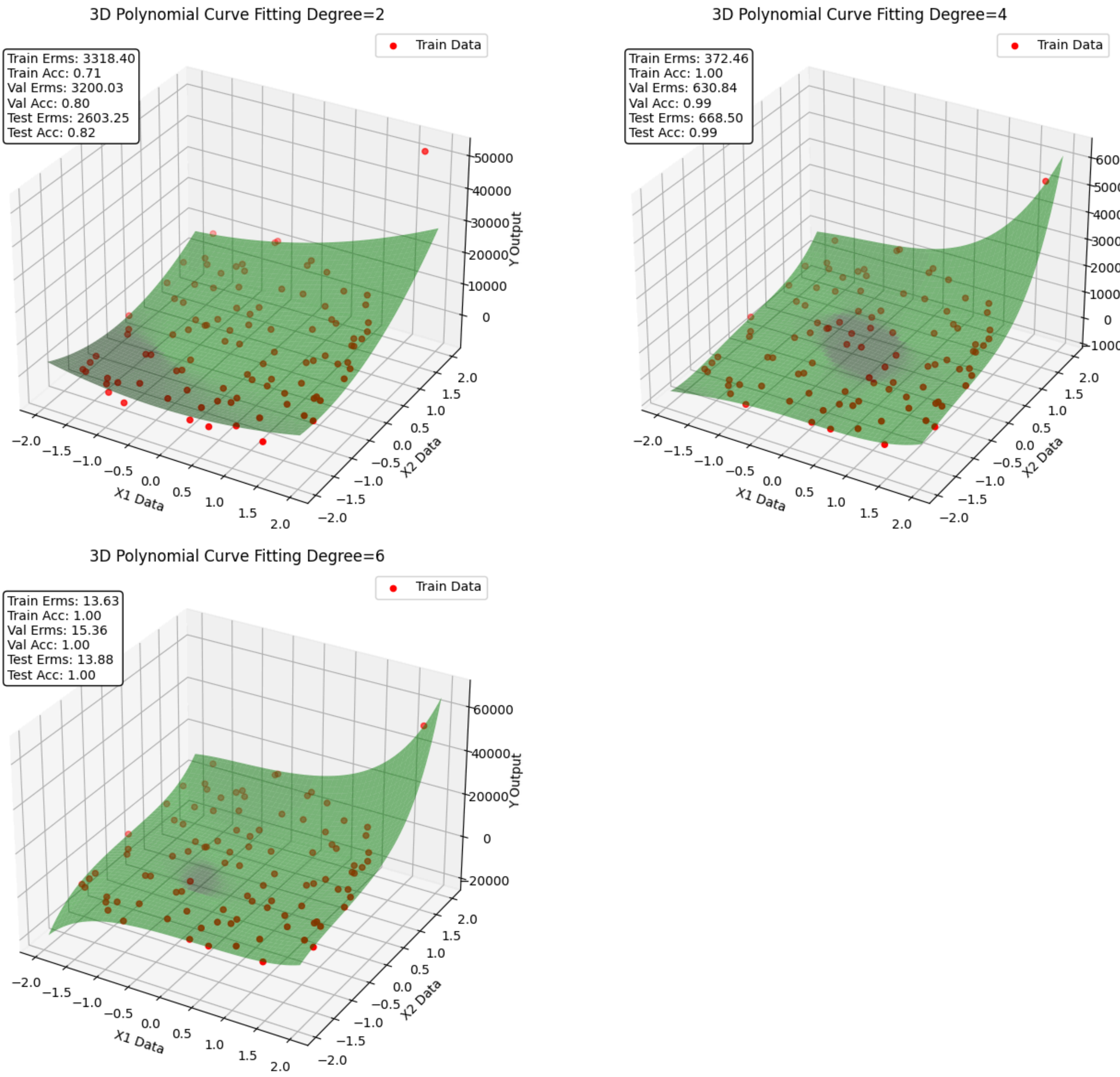


Figure 8: Surface plot for different degree without regularization, Sample size: 100

2.4 Erms With Regularization on Train, Test and validation data and Degree=6 and Sample Size=25

λ	**Train Erms**	**Validation Erms**	**Test Erms**
0.001	13.627	15.36130	13.89426
0.1	13.74305	16.01942	15.81274
1	18.73140	22.031694	29.15260

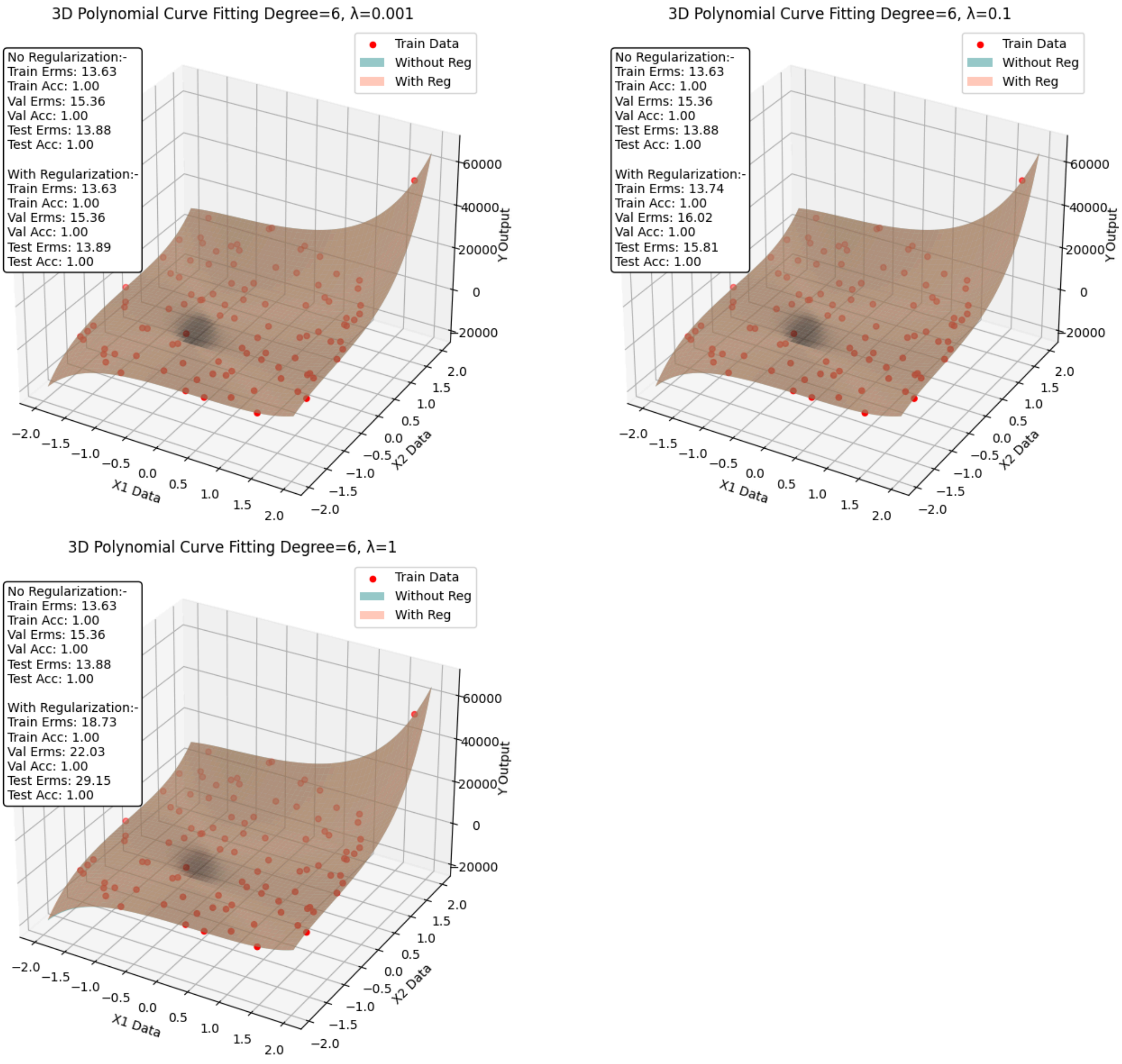


Figure 9: Surface plot for different **λ values** (two plot overlapped on training data set with and without regularization) , Sample size: 100

We see that the best fit for the data is obtained for degree: 6 and when we analyze (using validation data ERMS) the model for given *λ values and different degrees(given M values). we see that* for degree= 6 and *for λ=0.1* values , model gives the best performance (least ERMS for test data)

Conclusion: ERMS for best performing model: (M=6 and λ : 0.1)

Train Data Erms: 13.7430
Test Data Erms: 15.812

The best fit for training data and test data, M : 6 and λ : 0.1 is visualized as follows

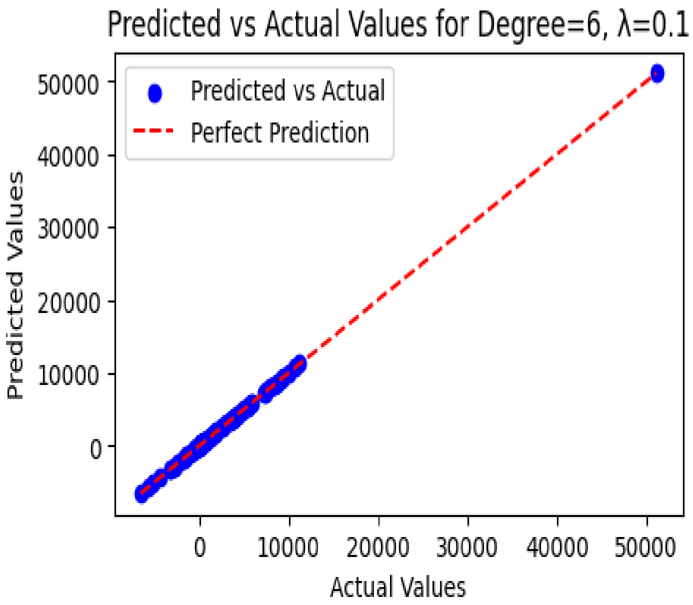


Fig: Predicted value Vs Actual Value curve for Train Data

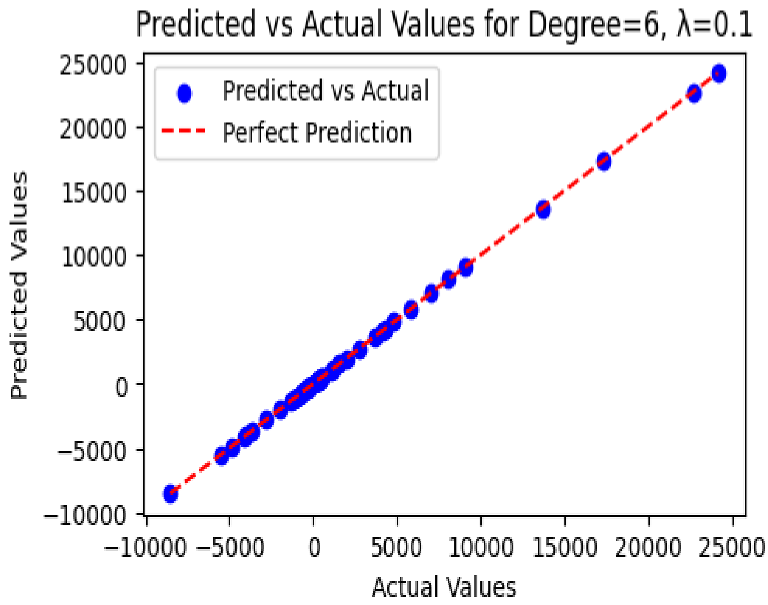


Fig: Predicted value Vs Actual Value curve for Test Data

Dataset-3:

3. Erms Without Regularization on Train, Test and validation data

M	Train Erms	Validation Erms	Test Erms
2	42.98033	65.33244	64.51939
3	265.64875	269.12683	361.71827

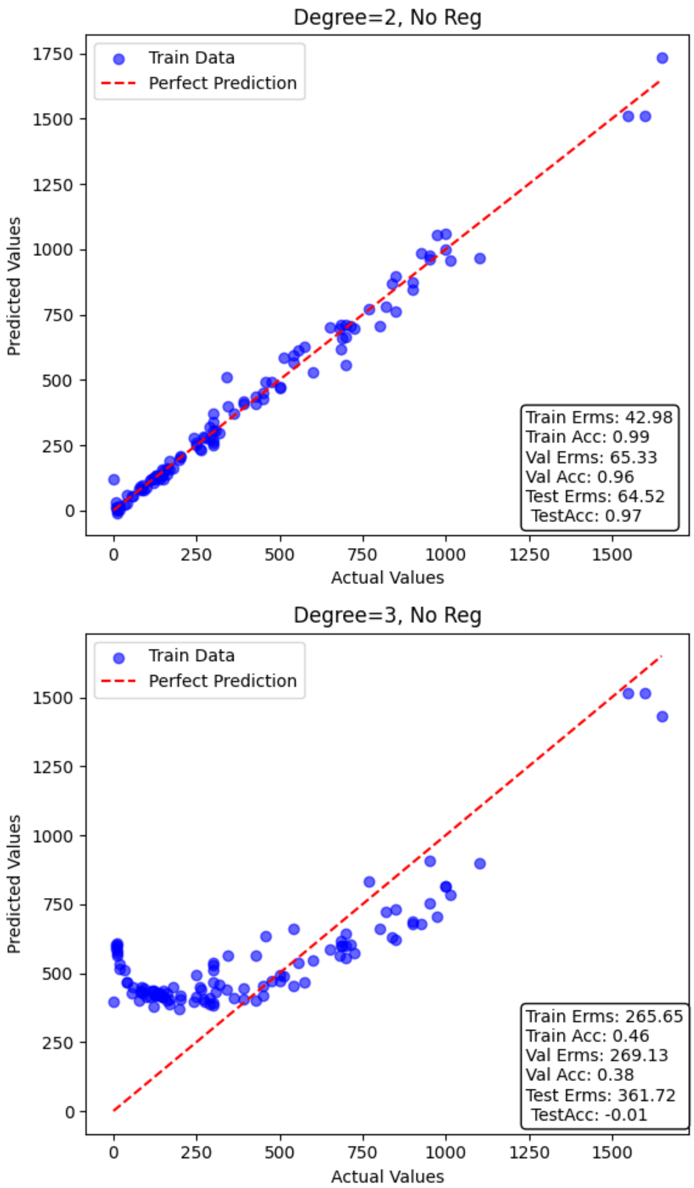


Fig: Predicted value Vs Actual Value curve for Train Data

3. Erms With Regularization on Train, Test and validation data, Degree M = 2

λ	Train Erms	Validation Erms	Test Erms
0.000001	42.98033	65.33240	64.5193
0.0001	42.98034	65.3301	64.5166
0.1	43.07188	64.881743	63.68748

–

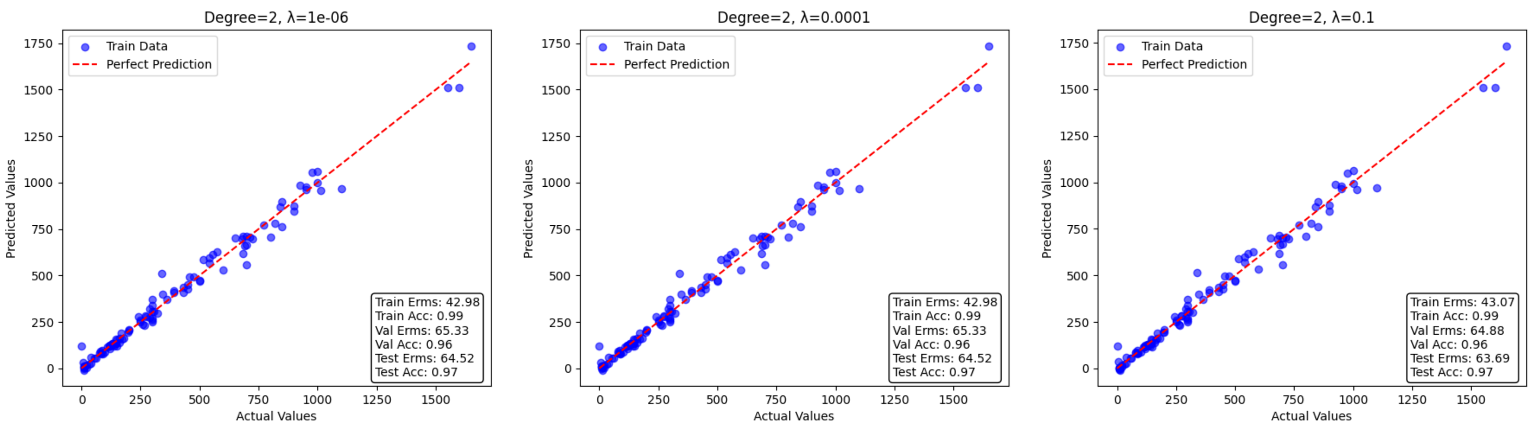


Fig: Predicted value Vs Actual Value curve for Train Data

3. Erms With Regularization on Train, Test and validation data, Degree M = 3

λ	Train Erms	Validation Erms	Test Erms
0.000001	163.28762	170.57828	225.7205
0.0001	96.70559	116.18292	163.10833
0.1	31.22957	59.481627	59.55018

–

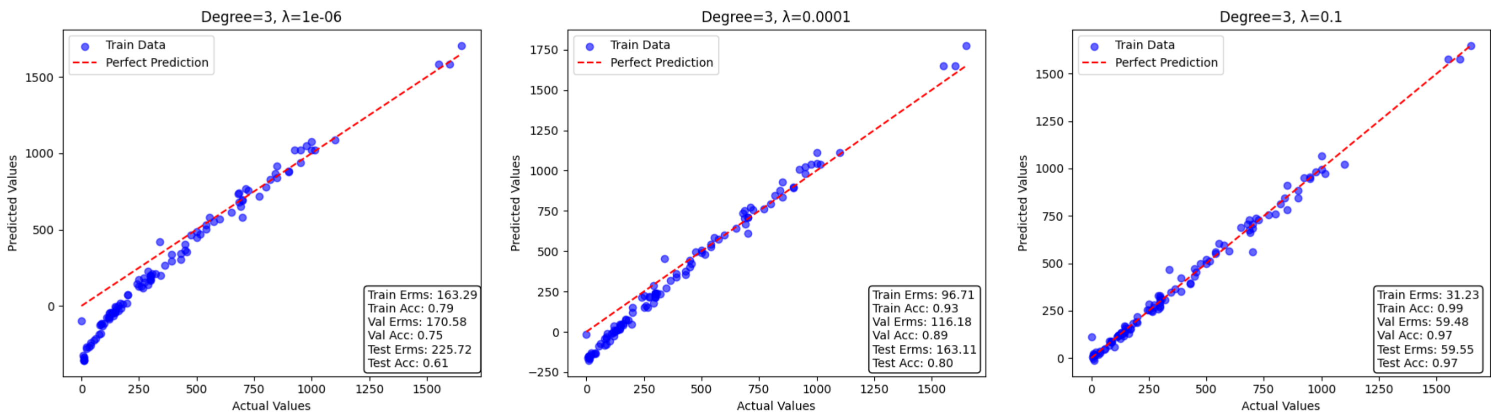


Fig: Predicted value Vs Actual Value curve for Train Data

Figure 7: Surface plot for different **λ values** (two plot shown on training data set with and without regularization) , Sample size: 25

We see that the best fit for the data is obtained for degree: 3 and when we analysis (using validation data ERMS) the model for given *λ values and different Degree(given M values). we see that* for degree= 3 and λ : 0.1 model gives the best performance (least ERMS in test data)

Conclusion: ERMS for best performing model: (M=3 and λ : 0.1)

Train Data Erms: 31.22957
Test Data Erms: 59.55018

The best fit for training data and test data, M : 3 and λ : 0.1 is visualized as follows

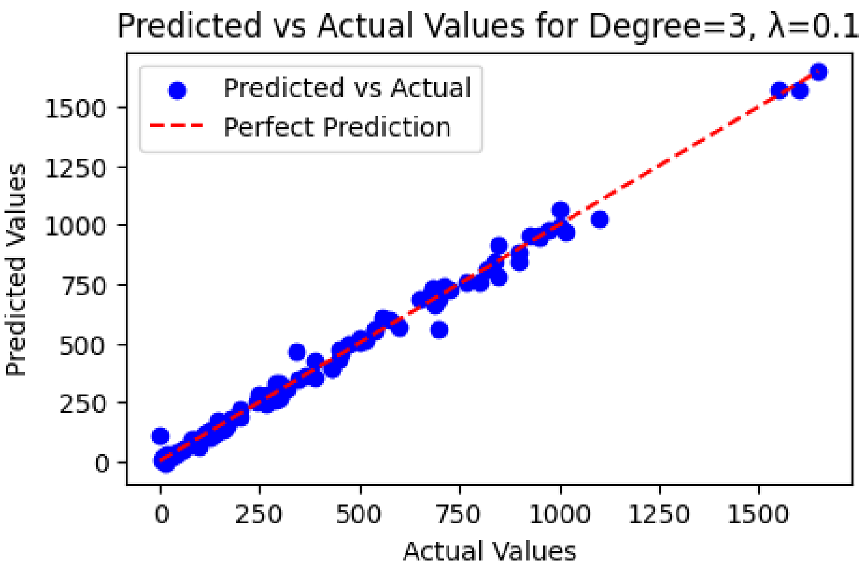


Fig: Predicted value Vs Actual Value curve for Train Data

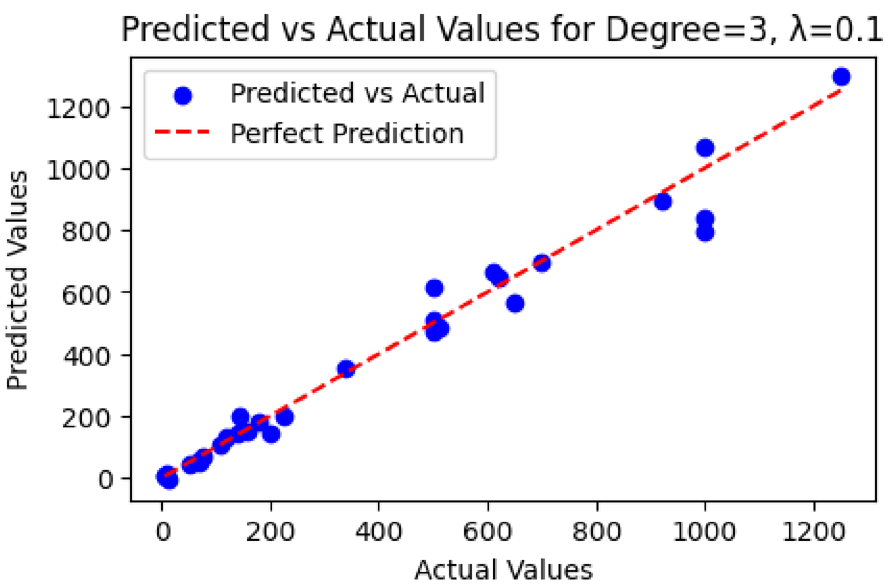


Fig: Predicted value Vs Actual Value curve for Train Data