ASSIGNMENT 01

Team No. – 14

Team Member- Debendra Kumar Pal (CS24M016) & Manglesh Patidar (CS24M025)

Subject- CS5691: Pattern recognition and Machine Learning

Programming Assignment 1

DATASET – 01:

1-D input data, we have two training datasets,

– Training Dataset 1(a): 10 examples, Training Dataset 1(b): 50 examples

Aim: We wanted to fit a best fit polynomial curve such that the error ERMS is the least.

Process:

1. From input data set X which contain all the data points x_i, where i={1,2,3,...,N} create Design Matrix Φ of dimensions (N,M) where M is the polynomial degree.
2. We want an W* of dimensions (M,1) which is optimal weight vector for which the polynomial curve has best fit.
   We Know
   $$W^* = (\Phi^T\Phi)^{-1}\Phi^T t \qquad \text{(eq. 1)}$$
   And in case of regularization it will be,
   $$W^* = (\lambda I + \Phi^T\Phi)^{-1}\Phi^T t) \qquad \text{(eq. 2)}$$
   Here t is the Y_actual of training dataset.
3. From the design matrix of Training data set we estimate W* using the above formula
4. Using the W* obtained from training dataset, we estimate Y_pred of test dataset.
   $$\text{Y_pred} = ((W^*) * \text{X_test}) \qquad \text{(eq. 3)}$$
5. Calculate
   $$\text{ERMS} = \sqrt{(1/N) * (Y_Pred - Y_actual)^2} \qquad \text{(eq. 4)}$$
6. We will use validation data set of get the best hyperparameter like Regression Constant, Model Complexity M (the degree of polynomial).

Required Common Function for Regression:

We have created few common functions which is common for any type of regression.

1. W_min(dMatrix,y): Takes the input as Design Matrix of dimension (N,M) and y is the output Y of the training dataset, and it returns the optimal W* satisfying the above equation eq. 1.
2. Wmin_reg(dMatrix, y, lam): Computes the minimum-norm least squares solution with regularization. Takes an extra input to Wmin() i.e. lam which is the regularization constant. It uses equation eq. 2.
3. predict(dMatrix,w): Takes the Design Matrix as dMatrix and optimal Weight vector W* as w and return the expected Y i.e. Y_pred using equation eq. 3
4. ERMS(y,y_pred): Calculate the ERMS error using eq 4.

Note: The plot for below are as follows: a) predicted vs true values of Y

b) scatter plot for curve fitting of training data

c) scatter plot for curve fitting of testing data

The only change in each case of regression is The calculation of Design Matrix and the way to load data in X and Y matrices other functions remain the same.

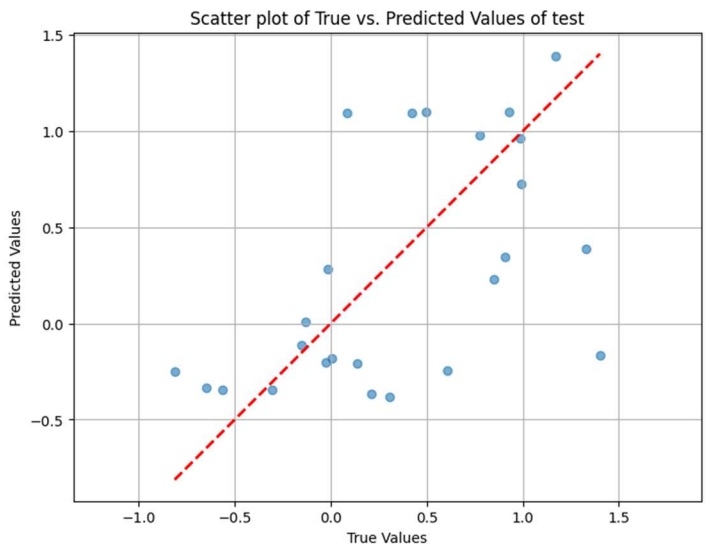
For the dataset 1a , we have considered the cases of M= 3,6,9 without regression and M=9 with regression as lambda = 0.001, 0.1, 1

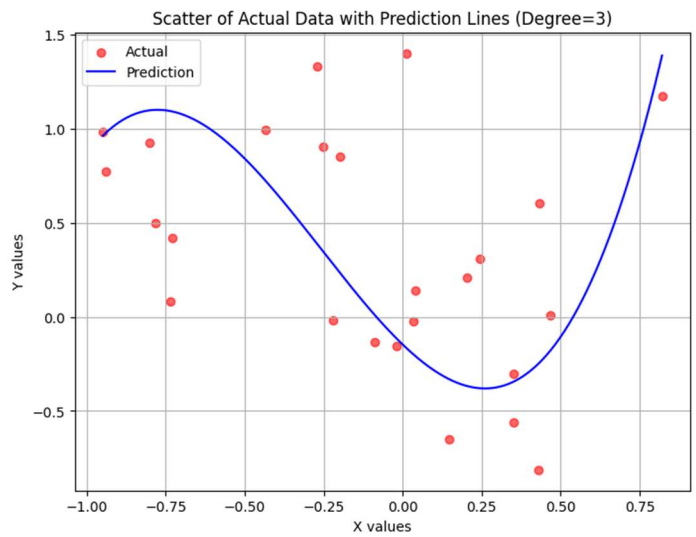
	Degree	λ	ERMS (Train)	ERMS (Val)	ERMS (Test)
0	3	None	0.283378	0.669069	0.578373
1	6	None	0.253710	27.785588	42.201101
2	9	None	0.074845	65448.685458	41671.182016
3	9	0.001	0.280719	0.539594	0.906058
4	9	0.1	0.306639	0.560843	0.662793
5	9	1	0.350065	0.635140	0.567244

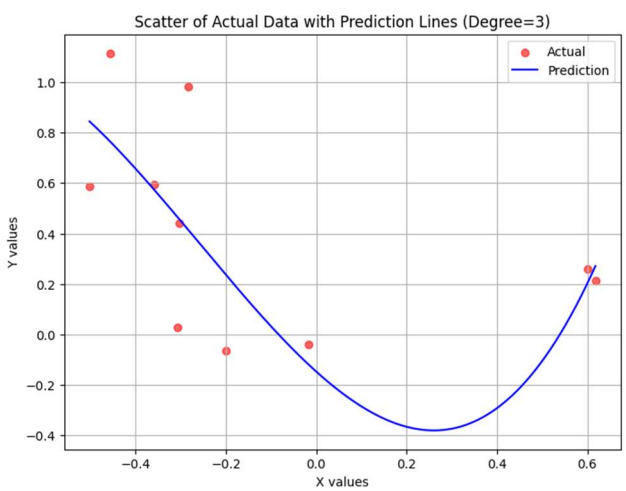
Best Model is with M=9 and lambda=1 with test error of 0.5672438301246967

Let us check the result of each case briefly,

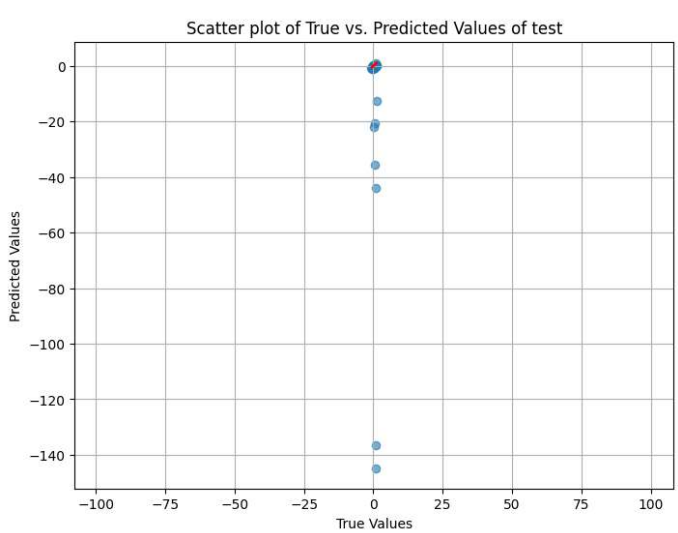
Case: <u>M=3 (No regularization)</u>

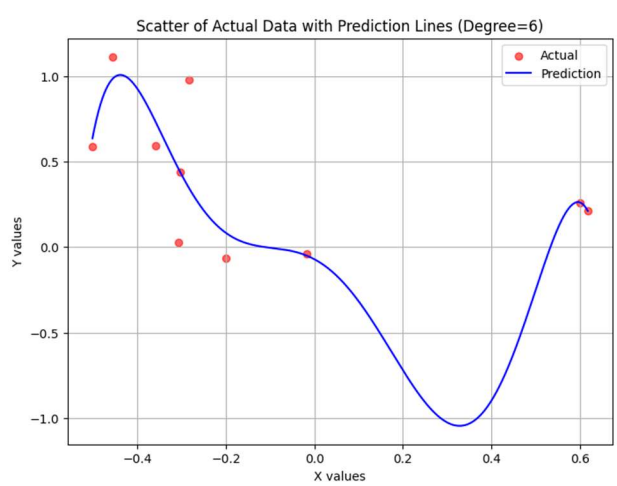


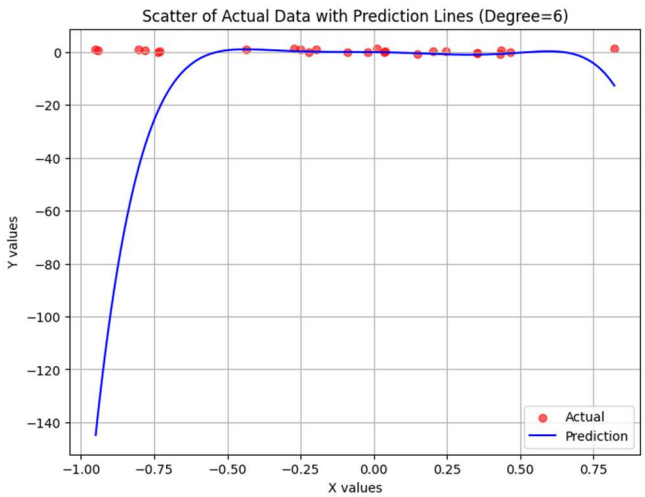




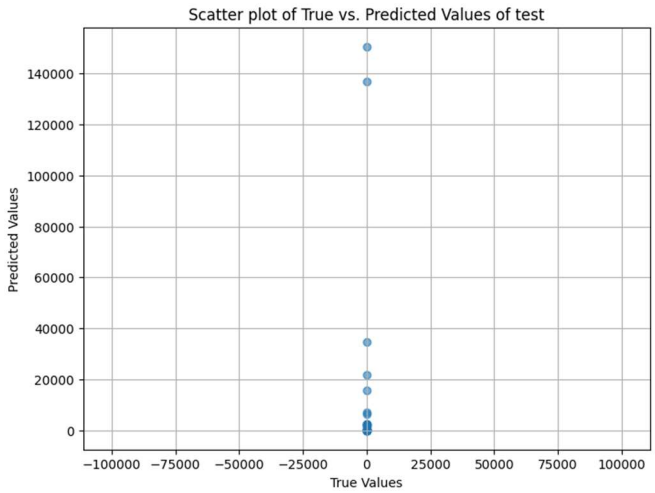
Case: M=6, lambda =0

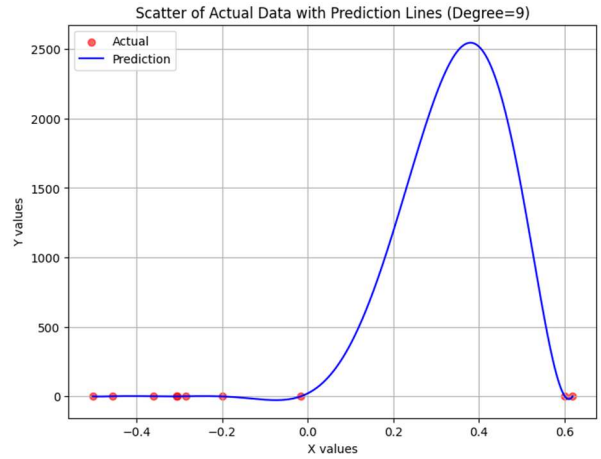


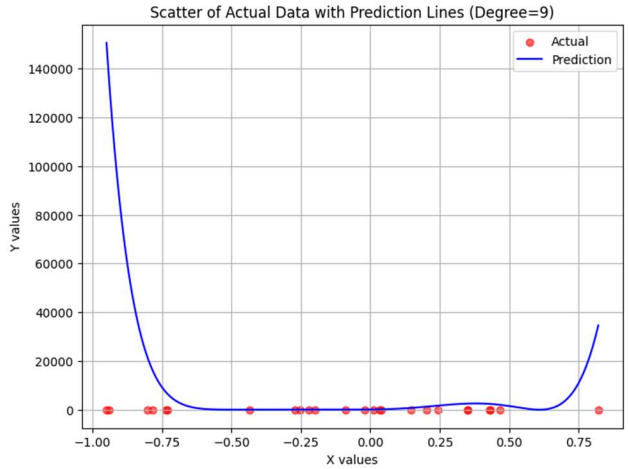




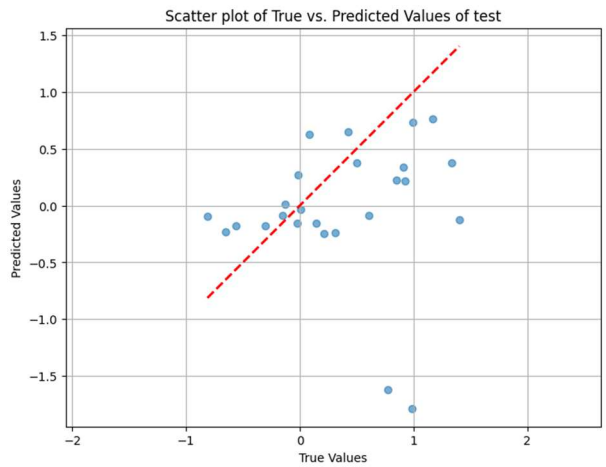
Case: M=9, lambda=0

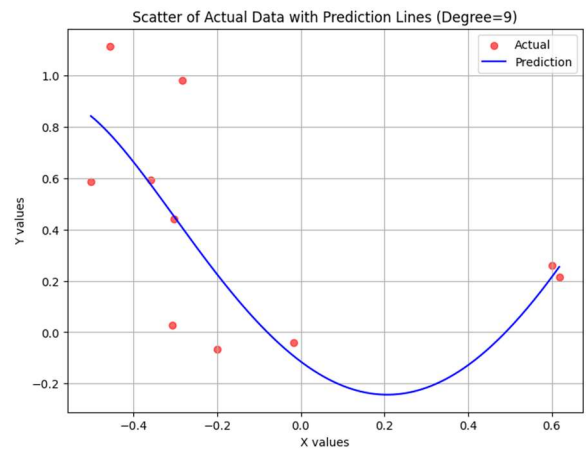


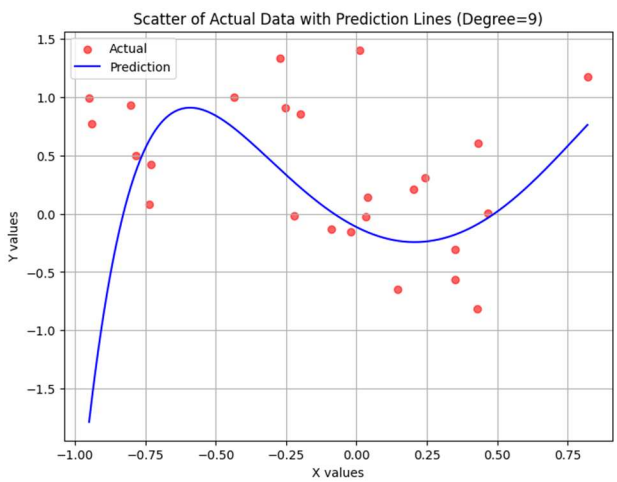




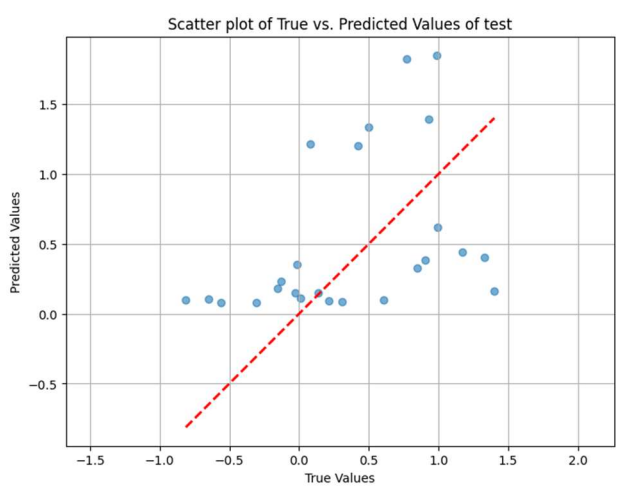
Case: M= 9, lambda= 0.001

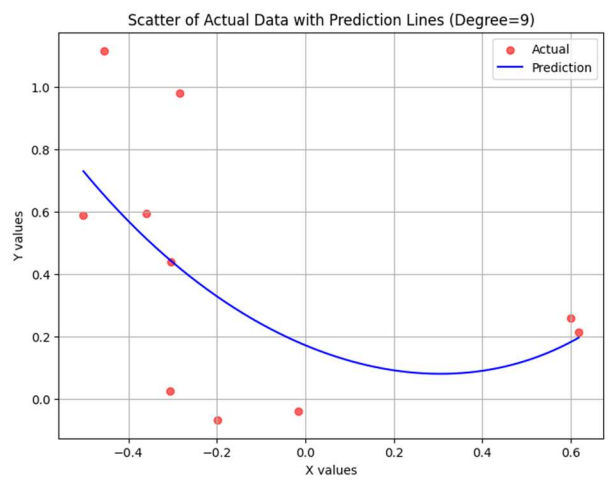


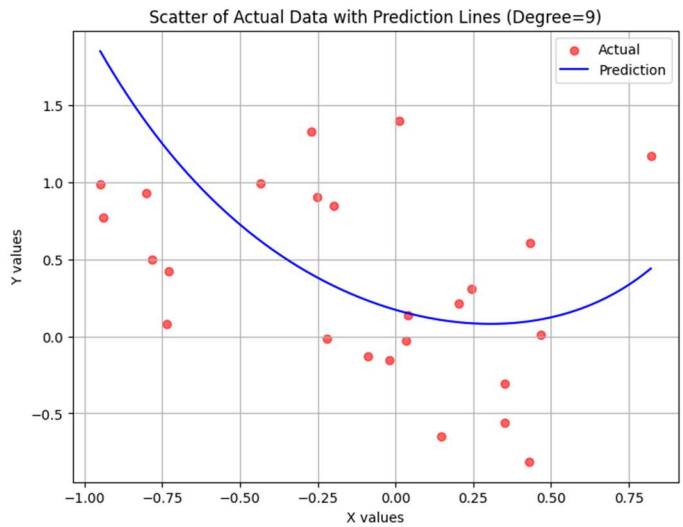




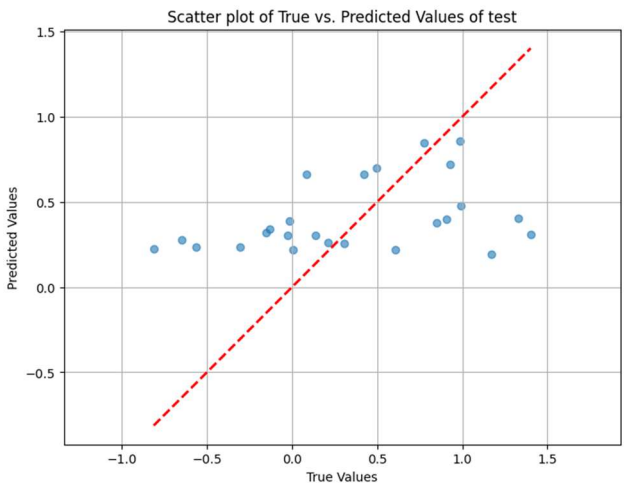
Case: M= 9, lambda=0.1

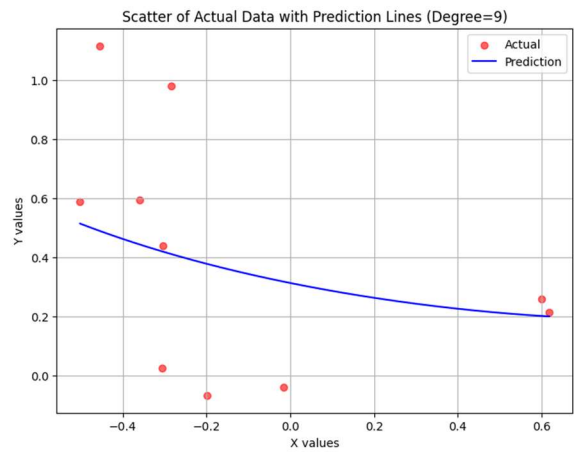


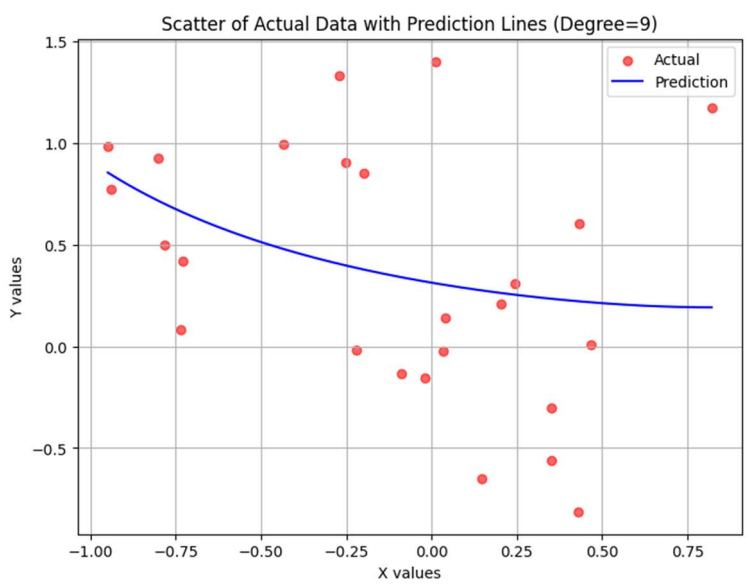




Case: M= 9, lambda= 1







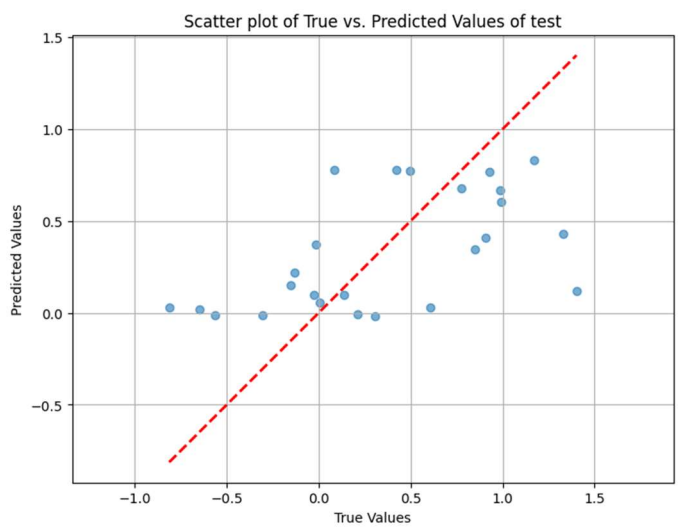
Result: Hence we can see from the above analysis that the best result we get for with M=9 and lambda=1 with test error of 0.567243830124697 for dataset train_1a as a training dataset since the data is too low we are facing underfitting and overfitting problem too.

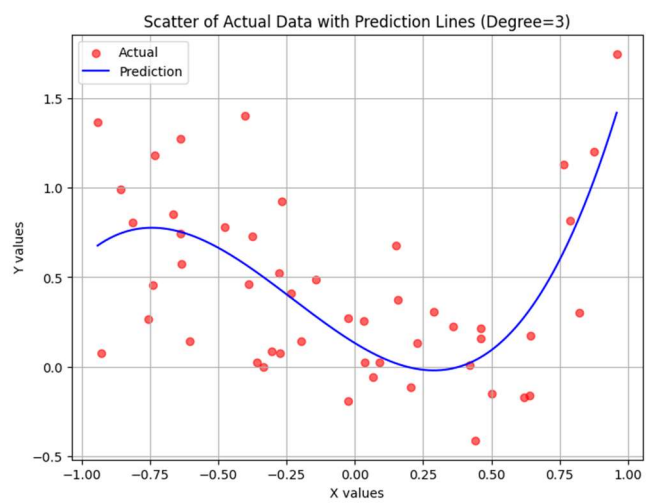
For **train_1b** let's analyze,

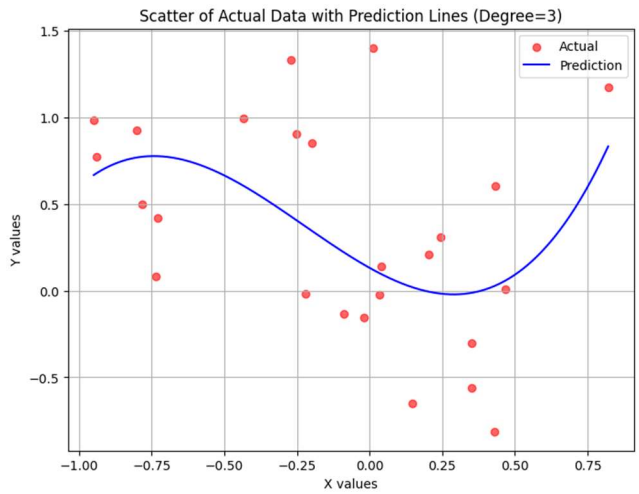
	Degree	λ	ERMS (Train)	ERMS (Val)	ERMS (Test)
0	3	None	0.355957	0.400610	0.508245
1	6	None	0.343583	0.414590	0.512770
2	9	None	0.335026	0.411640	0.516775
3	9	0.001	0.338586	0.411519	0.511183
4	9	0.1	0.344825	0.403183	0.513352
5	9	1	0.352835	0.369344	0.527465

Best Model is with M=3 and lambda=None with test error of 0.5082445393396905

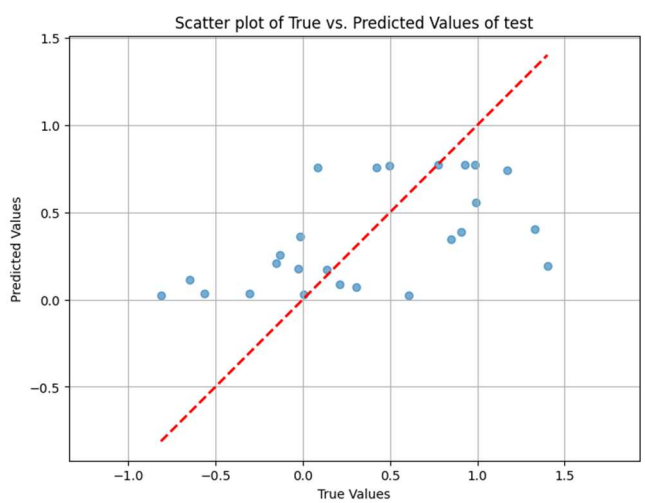
Case: M =3, lambda=0

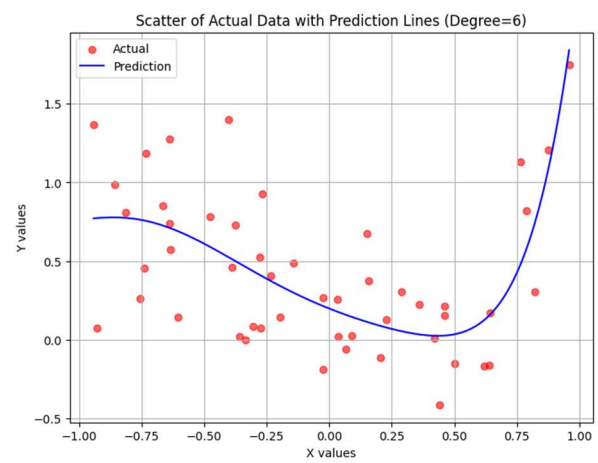


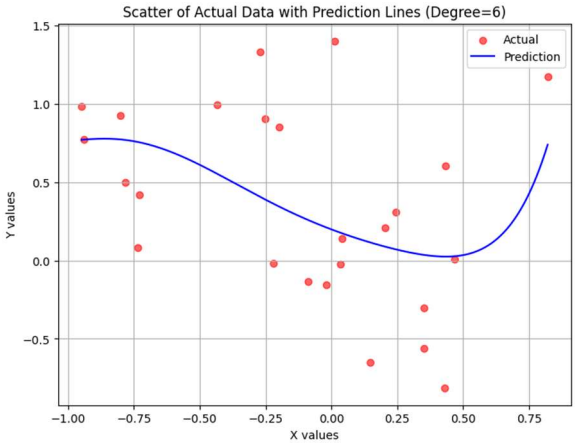




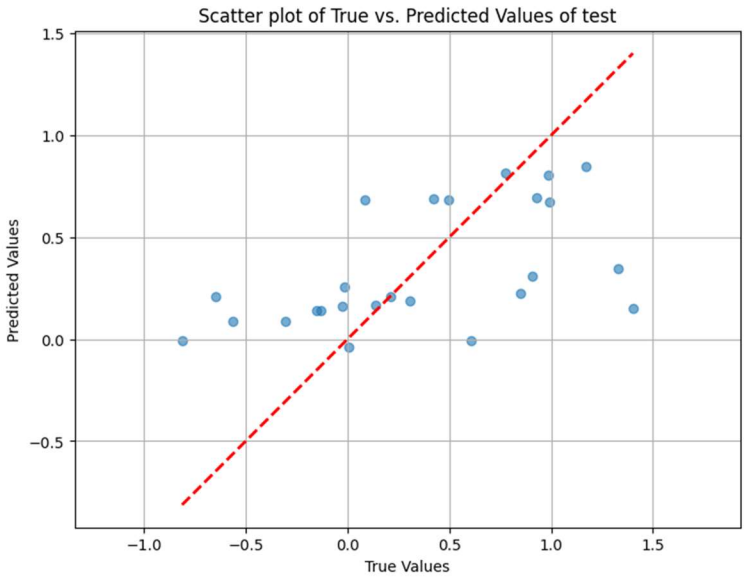
Case: M=6, Lambda=0:

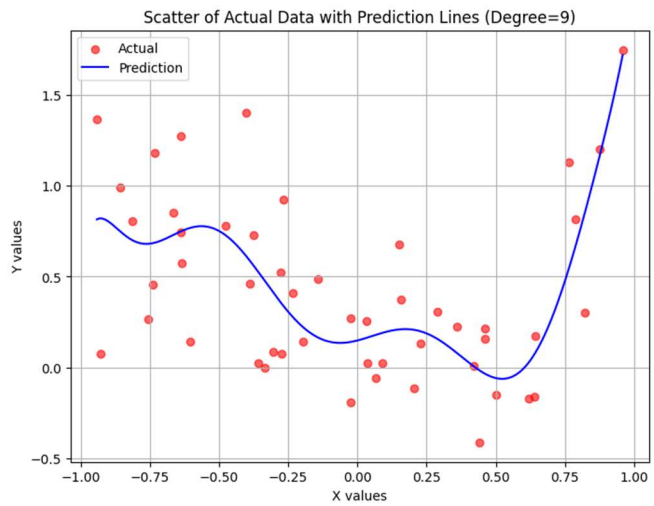


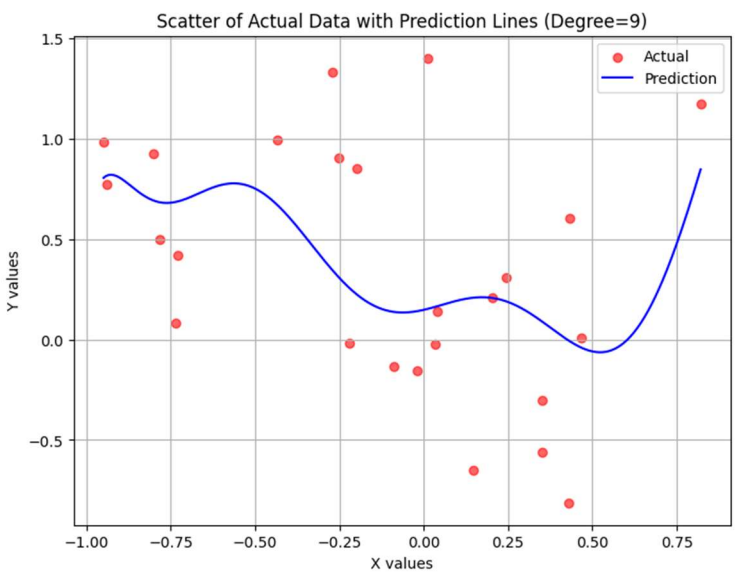




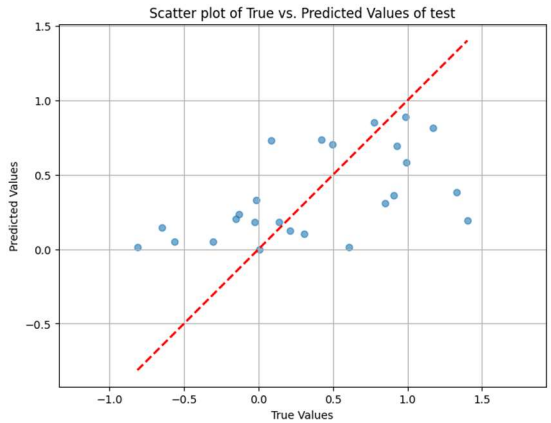
Case: M=9, Lambda = 0

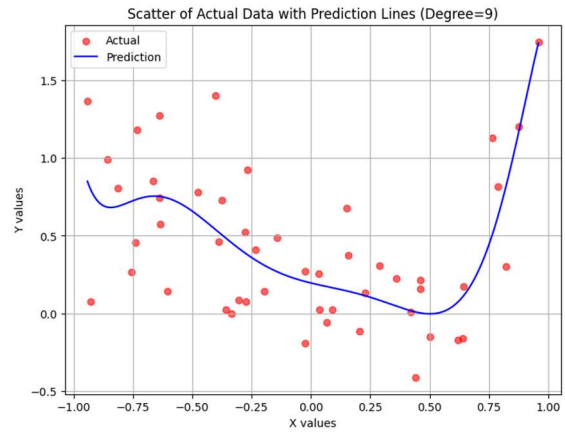


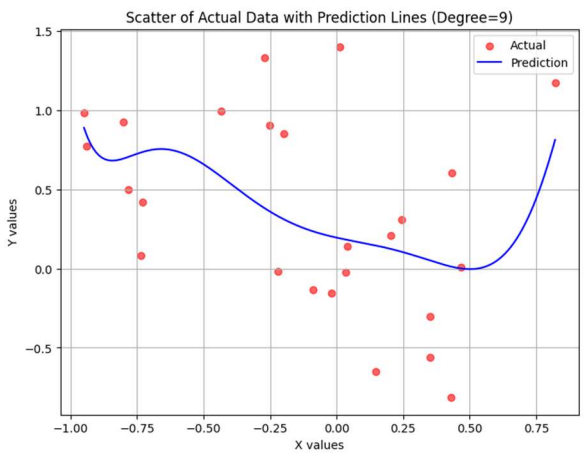




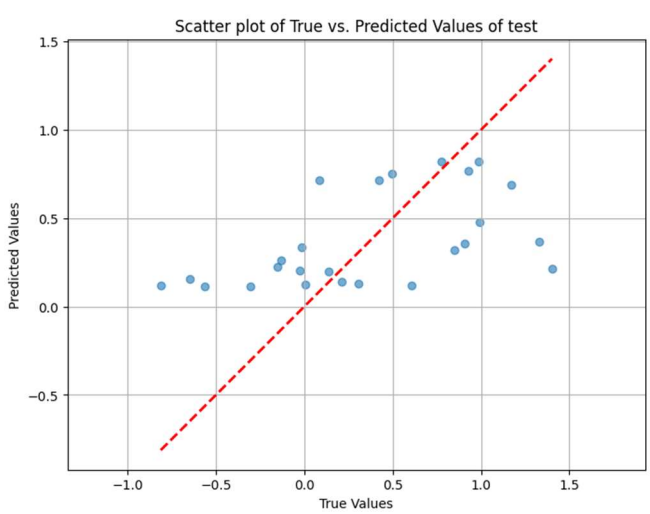
Case Plot For M=9 and Lambda=0.001

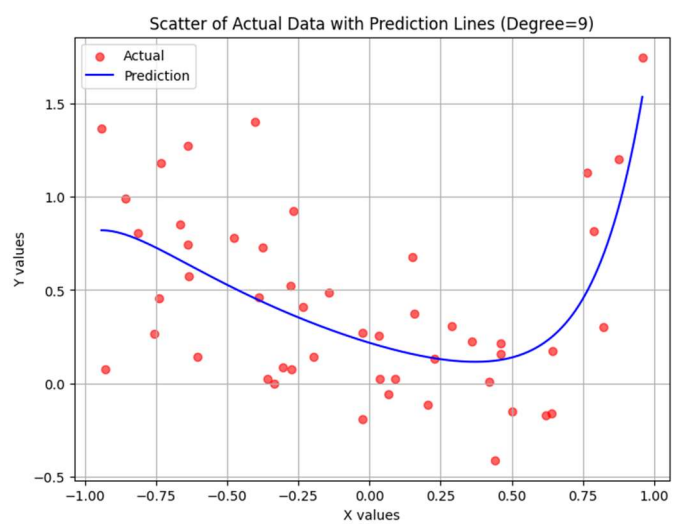


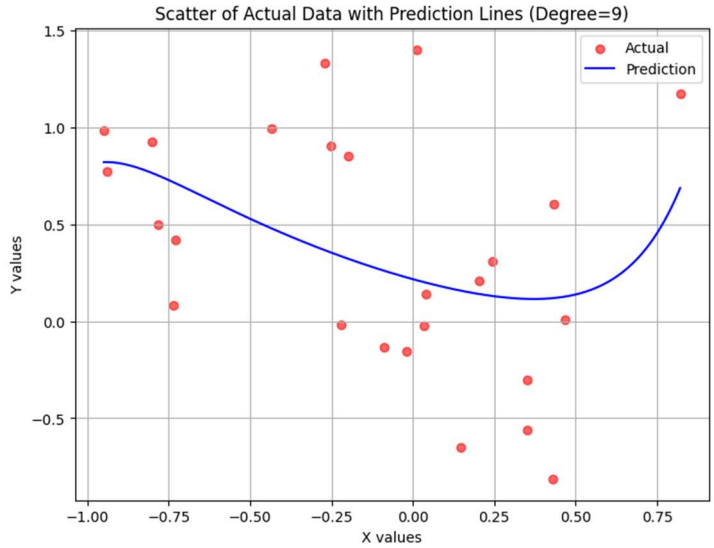




CASE: Plot For M=9 and Lambda=1







<u>Result for Dataset 1b as training :</u>

Best Model is with M=3 and lambda=None with test error of 0.5082445393396905

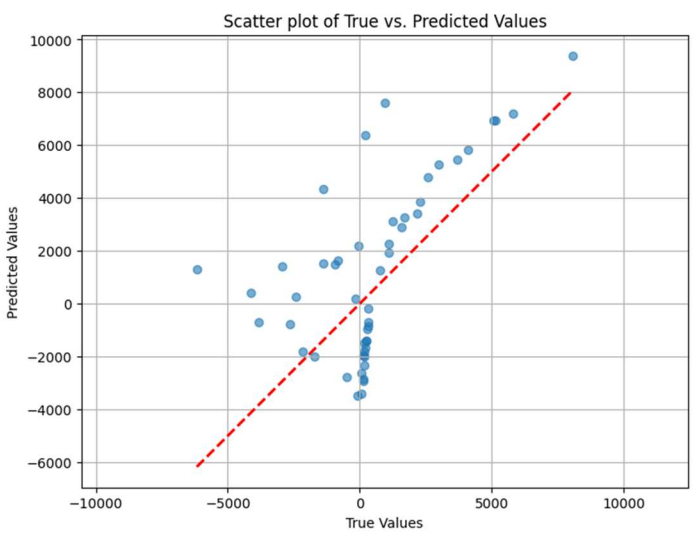
DATASET 2:

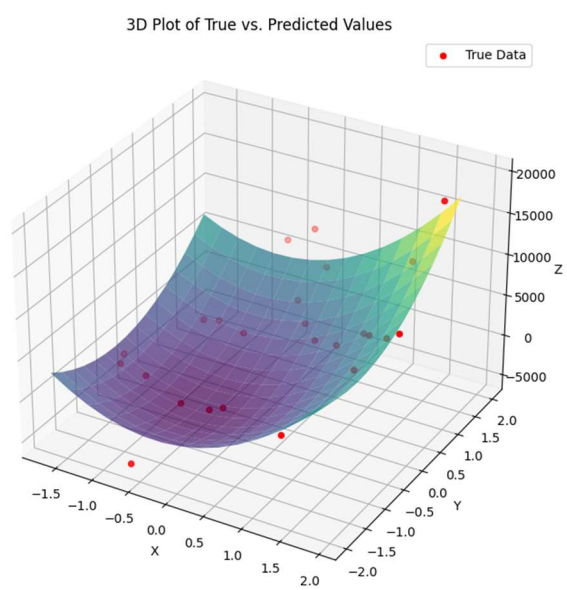
2-dimensional (Bivariate) input data – Training Dataset 2(a): 25 examples, Training Dataset 2(b): 100 examples

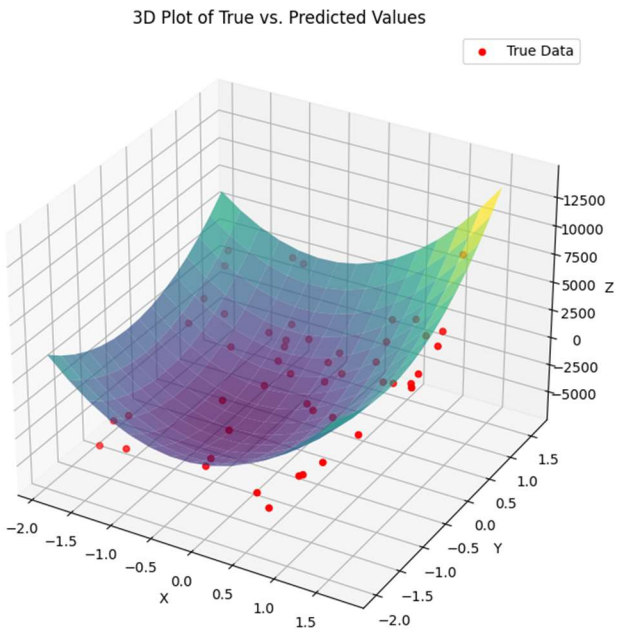
For Training Dataset as 2a.

```
     Dataset  Degree of Polynomial Regularization Coefficient λ  \
0  Dataset 1                    2                        None
1  Dataset 1                    4                        None
2  Dataset 1                    6                        None
3  Dataset 1                    6                       0.001
4  Dataset 1                    6                         0.1
5  Dataset 1                    6                           1

   ERMS (Training)  ERMS (Validation)  ERMS (Test)
0     2.035613e+03        2835.767905  2783.305077
1     1.658333e+02         863.185446   657.437961
2     5.057736e-08        2011.972413  1752.238769
3     2.737724e+00         939.041906   804.232975
4     7.837625e+00         297.405016   259.160574
5     2.705650e+01         395.986030   399.751655


Best Model is with M=6 and lambda=0.1 with test error of 259.1605739725069
```

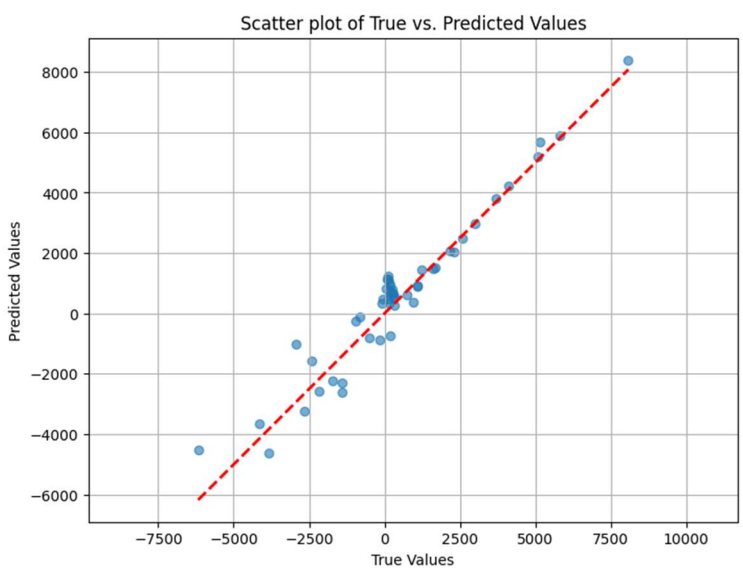
Case : M=2 and No regularization

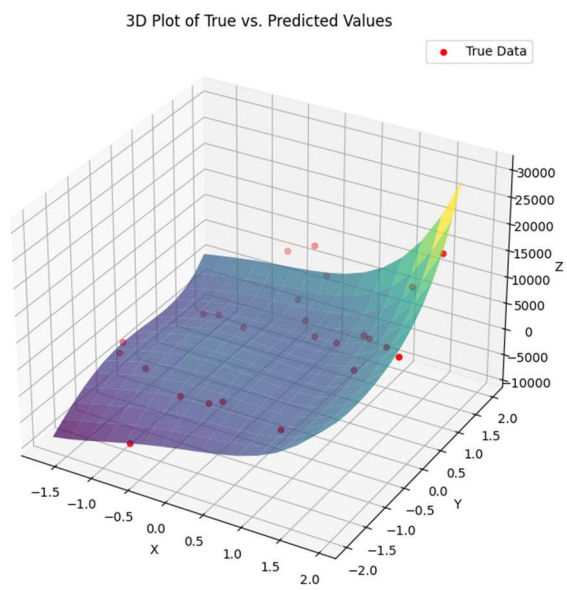


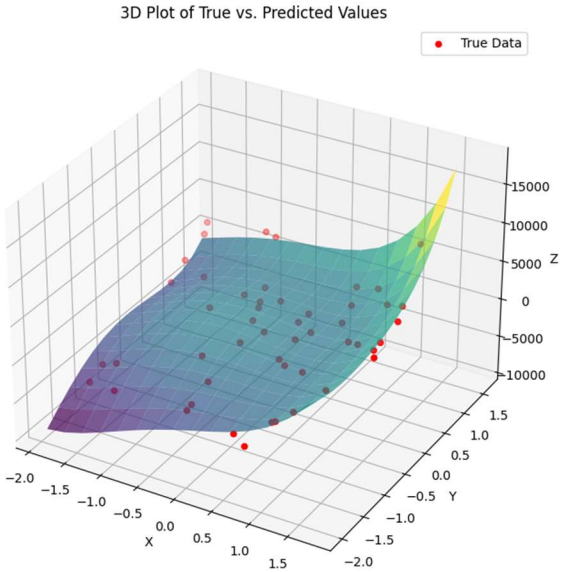




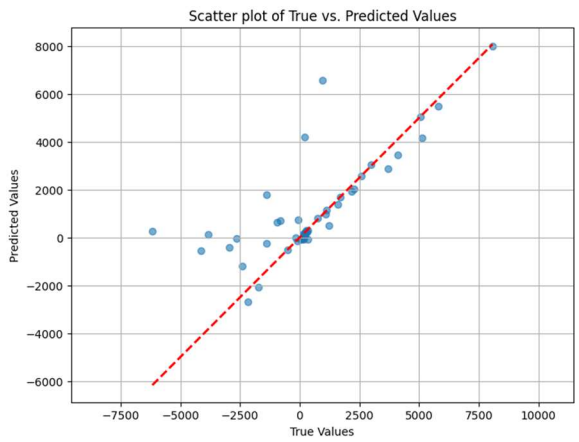
Case: M=4 and Lambda =0

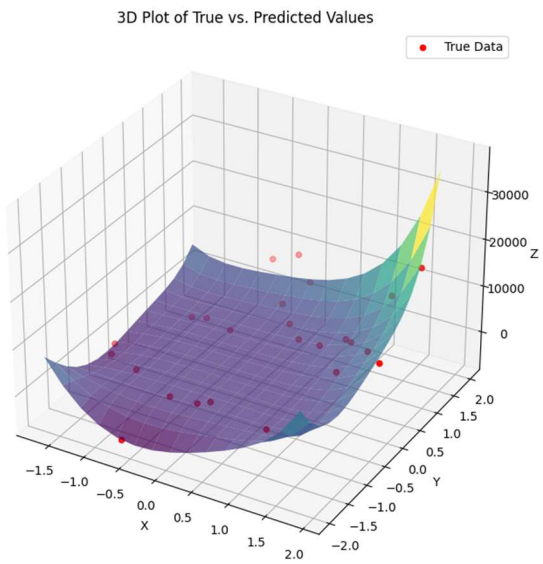


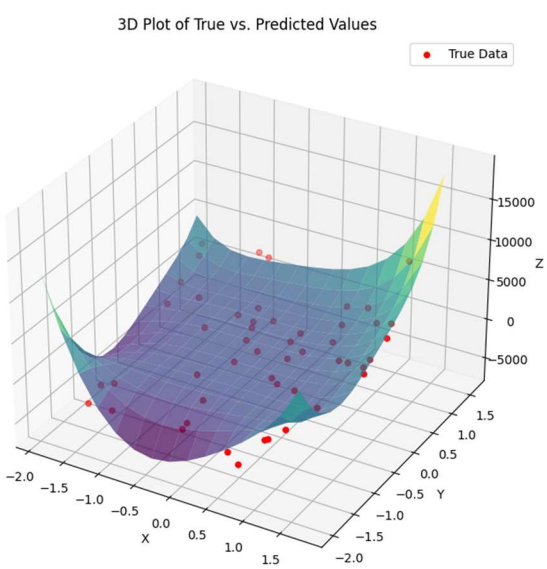




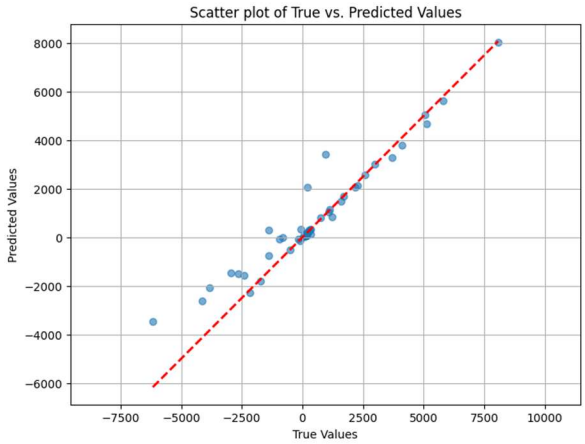
Case M=6 and lambda =0

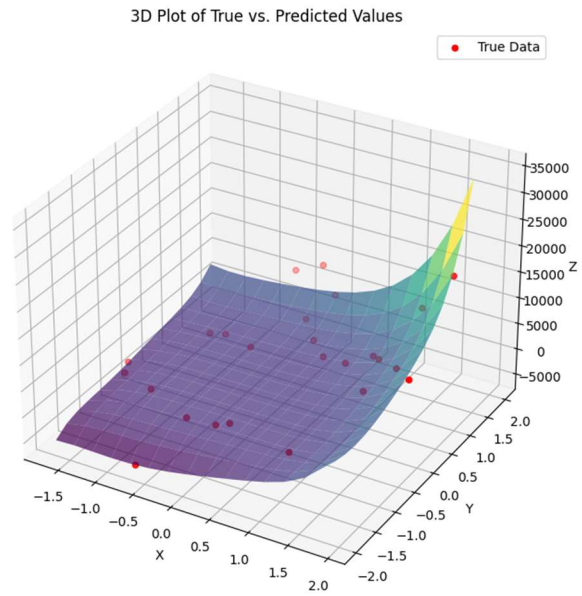


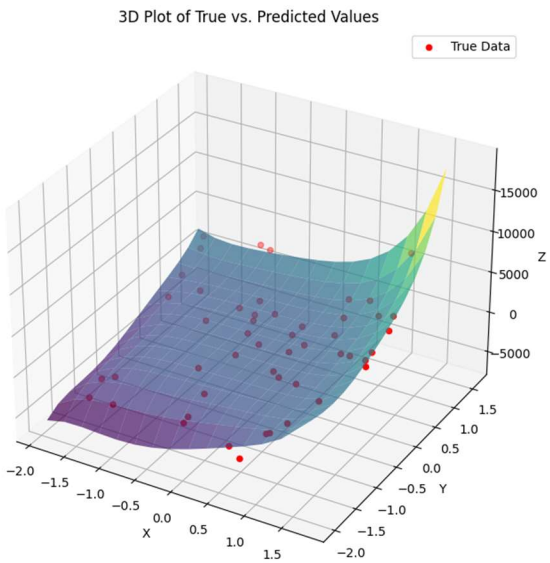




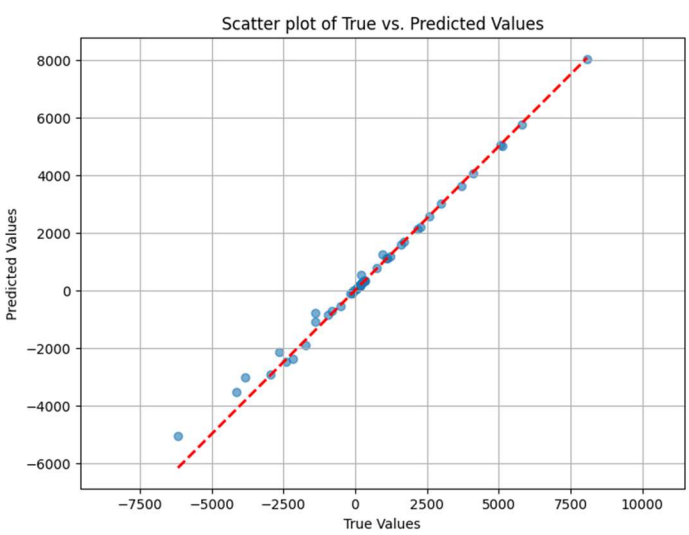
Case M=6 and lambda = 0.001

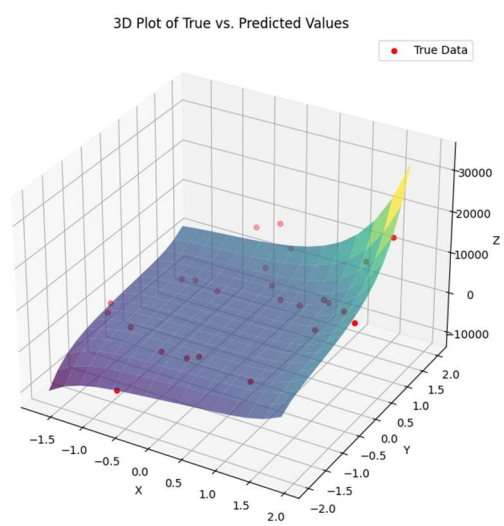


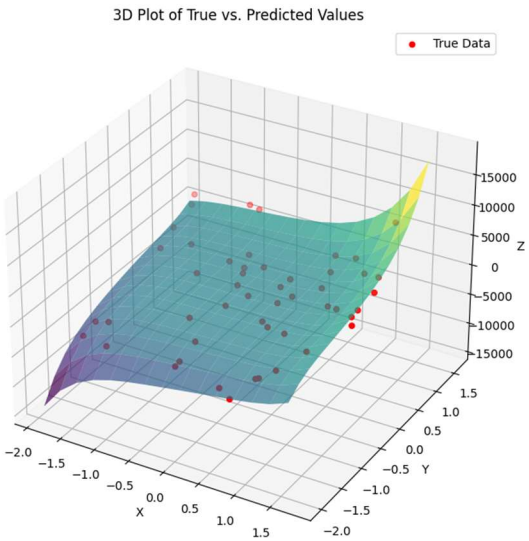




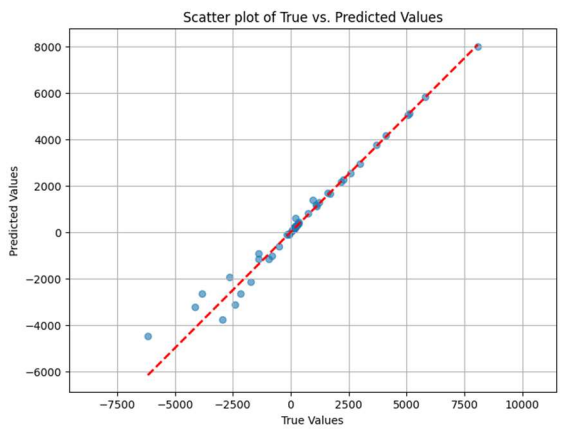
Case M=6 and lambda = 0.1

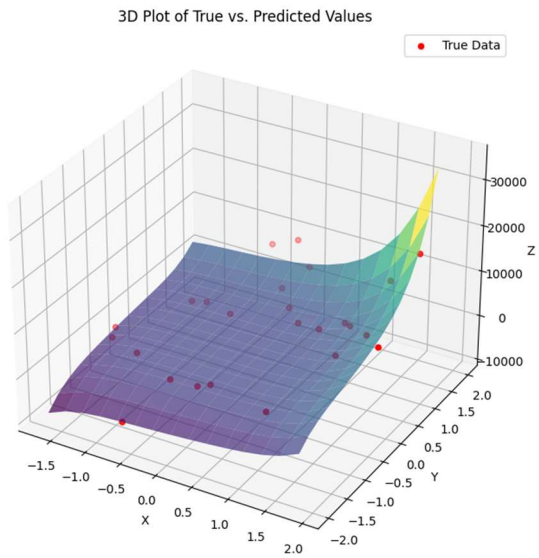


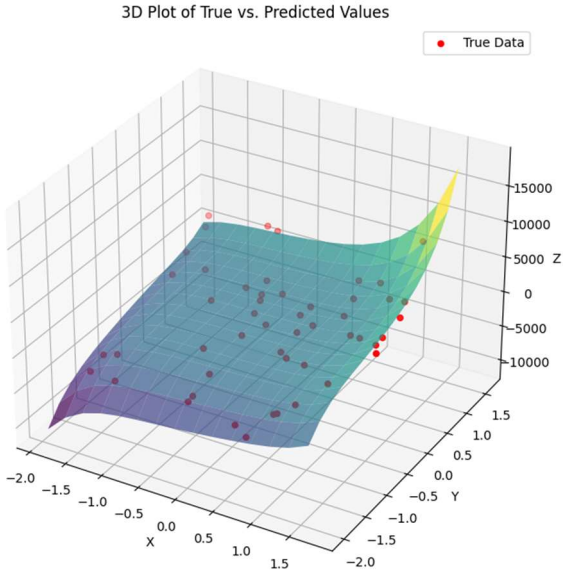




Case M=6 and lambda = 1







Result for dataset 2a:

Best Model is with M=6 and lambda=0.1 with test error of 259.1605739725069

Similarly for dataset 2b:

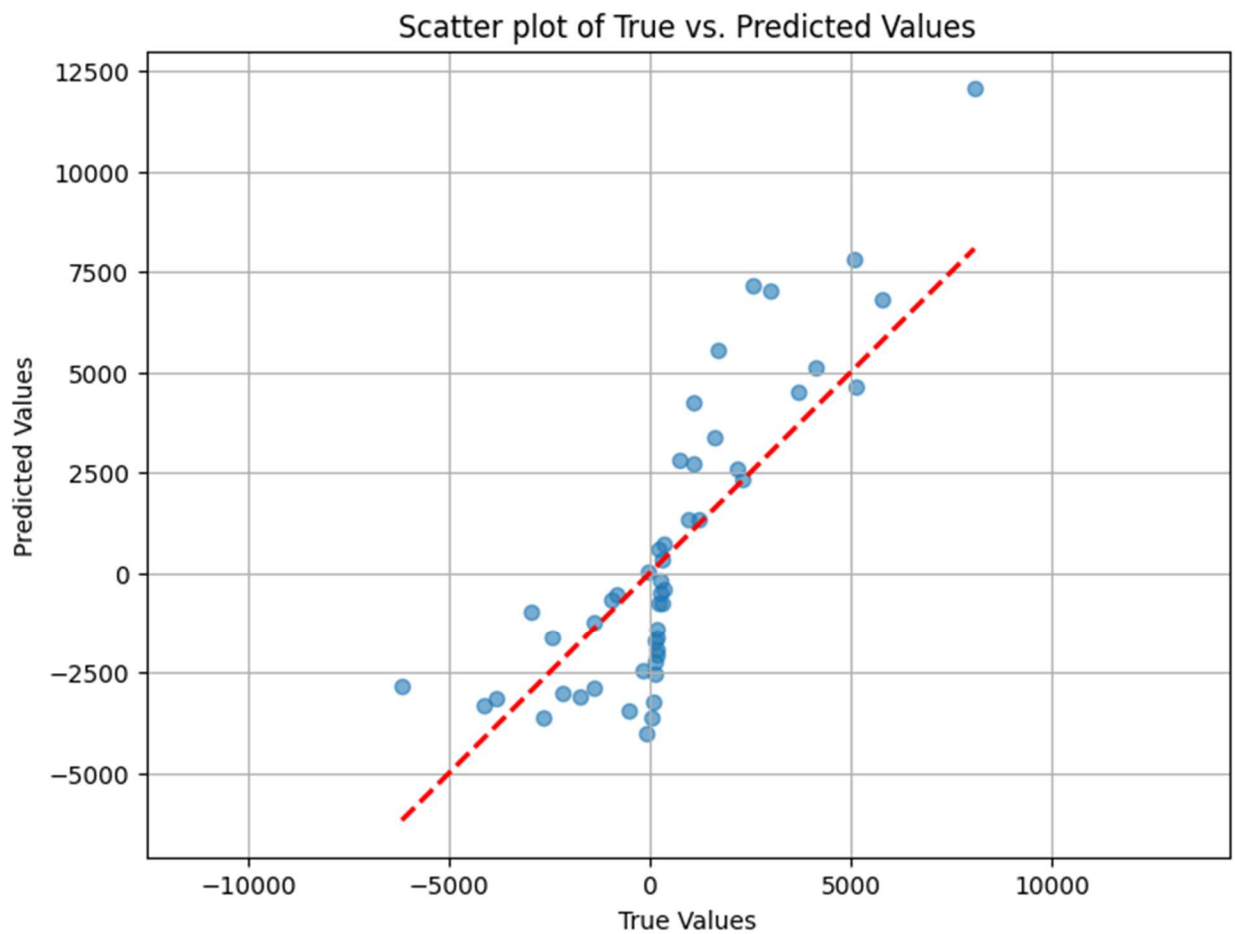
```
     Dataset  Degree of Polynomial Regularization Coefficient λ  \
0  Dataset 1                     2                        None
1  Dataset 1                     4                        None
2  Dataset 1                     6                        None
3  Dataset 1                     6                       0.001
4  Dataset 1                     6                         0.1
5  Dataset 1                     6                           1

   ERMS (Training)  ERMS (Validation)  ERMS (Test)
0      2766.975033        2270.600245  2049.713579
1       357.631952         427.806601   410.411206
2        13.493367           7.481293     8.052611
3        13.493370           7.486627     8.052756
4        13.528313           8.105409     8.162154
5        16.080582          16.442681    13.731523


Best Model is with M=6 and lambda=None with test error of 8.052610760286651
```

Case M=2 and Lambda =1

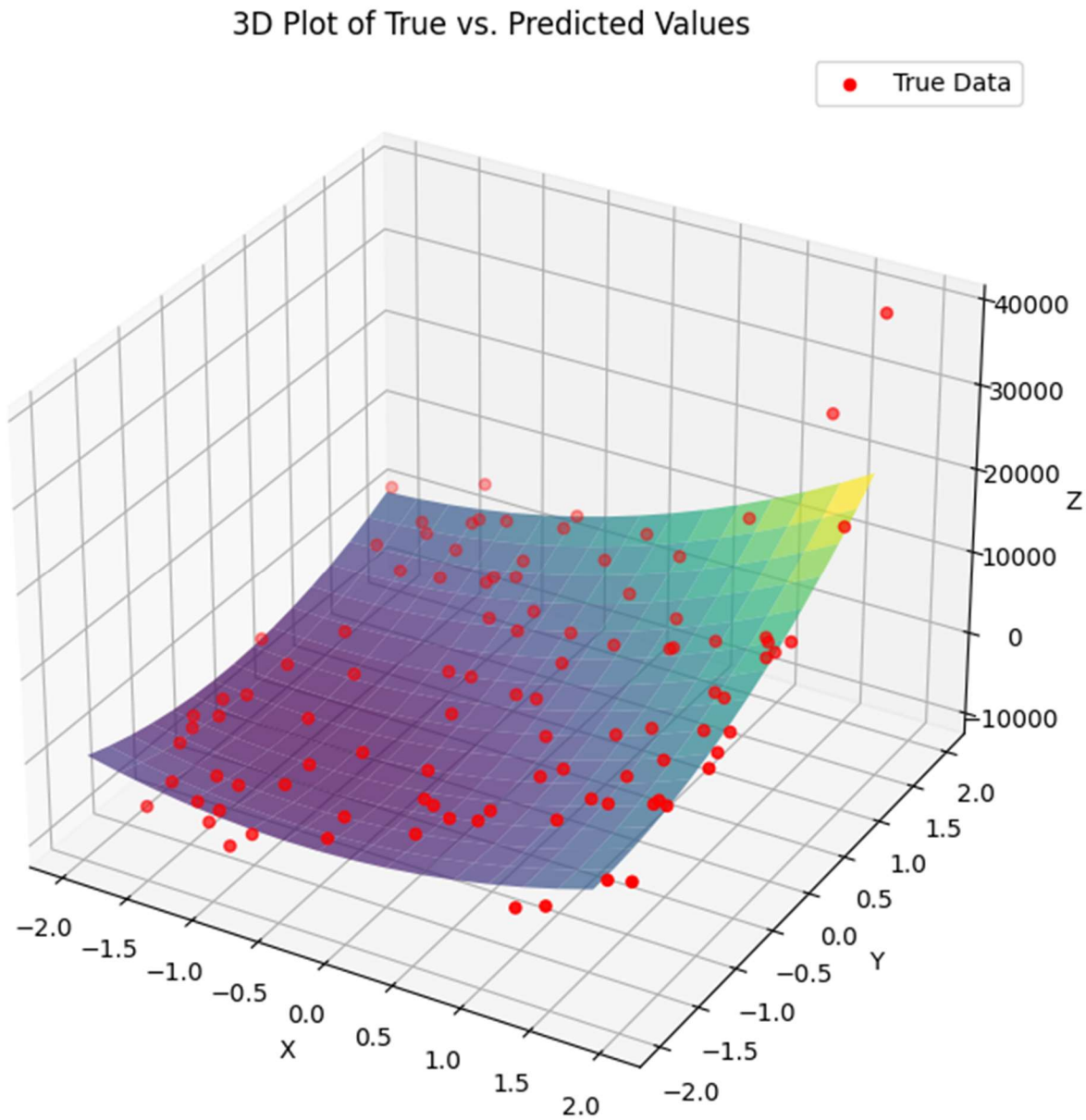
Train_err: 2766.9750334349096

val_err 2270.6002453024857

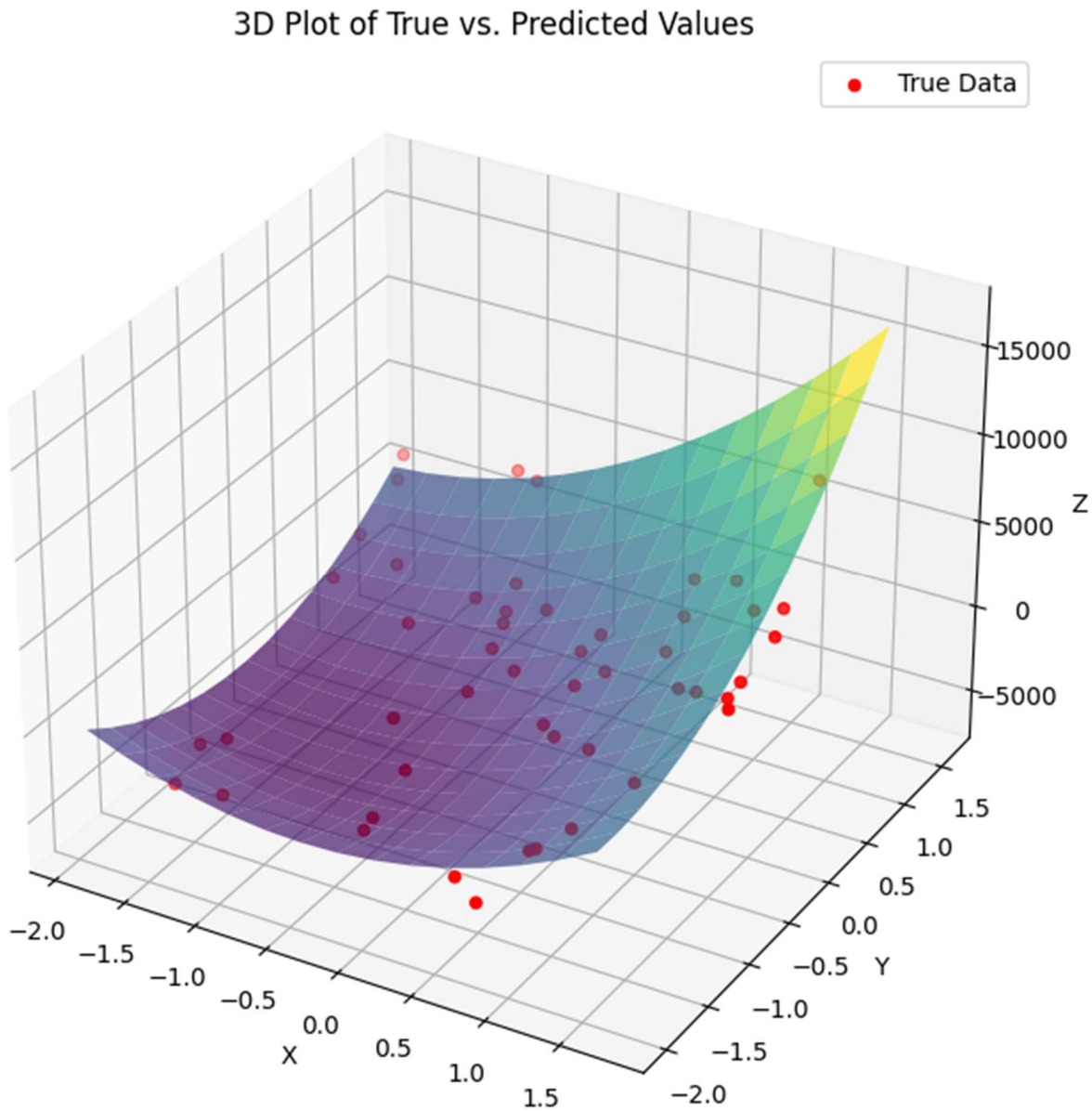
test_err 2049.7135793430493



3D plot for Train data



3D plot for Test data

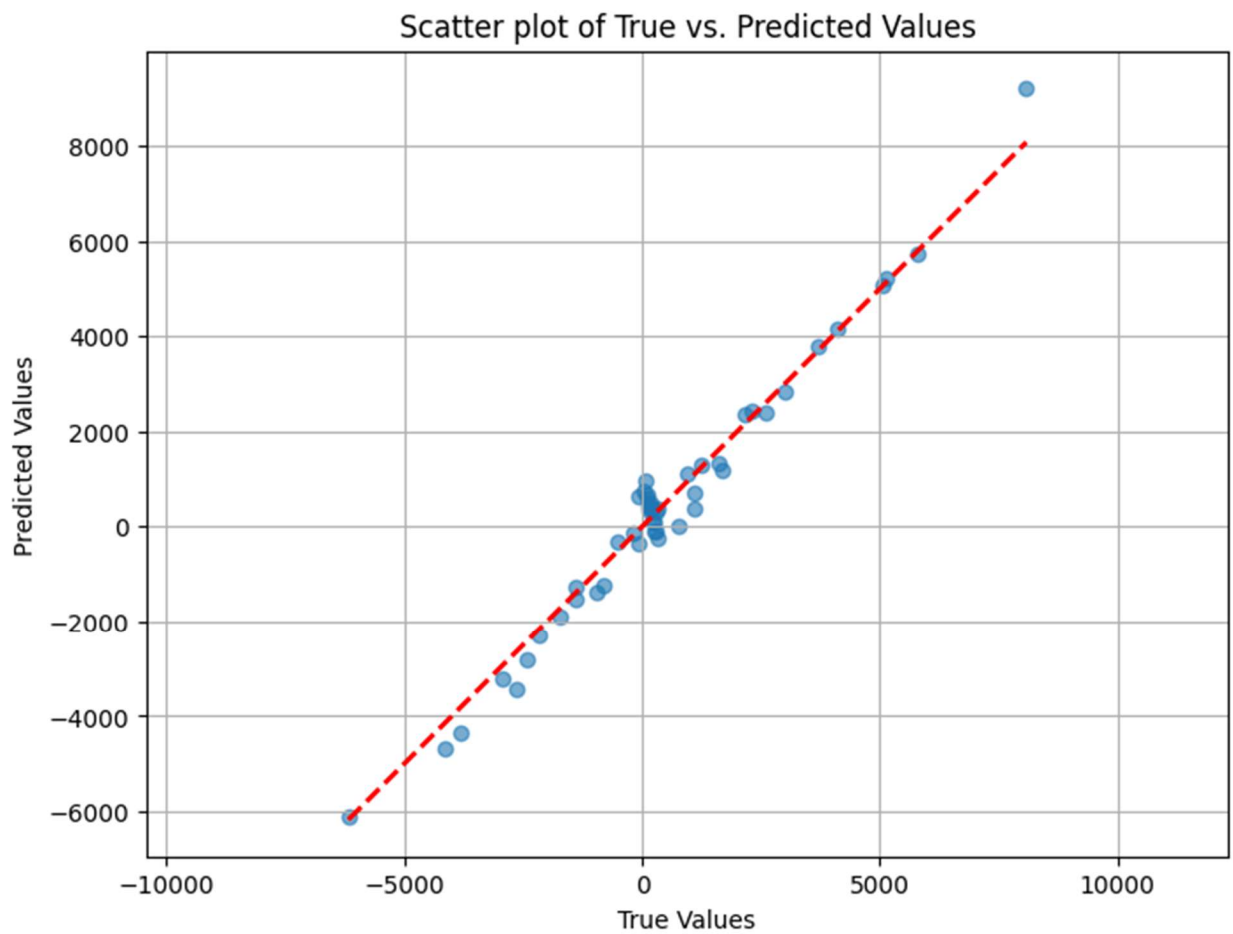


Case M=4 and Lambda =1

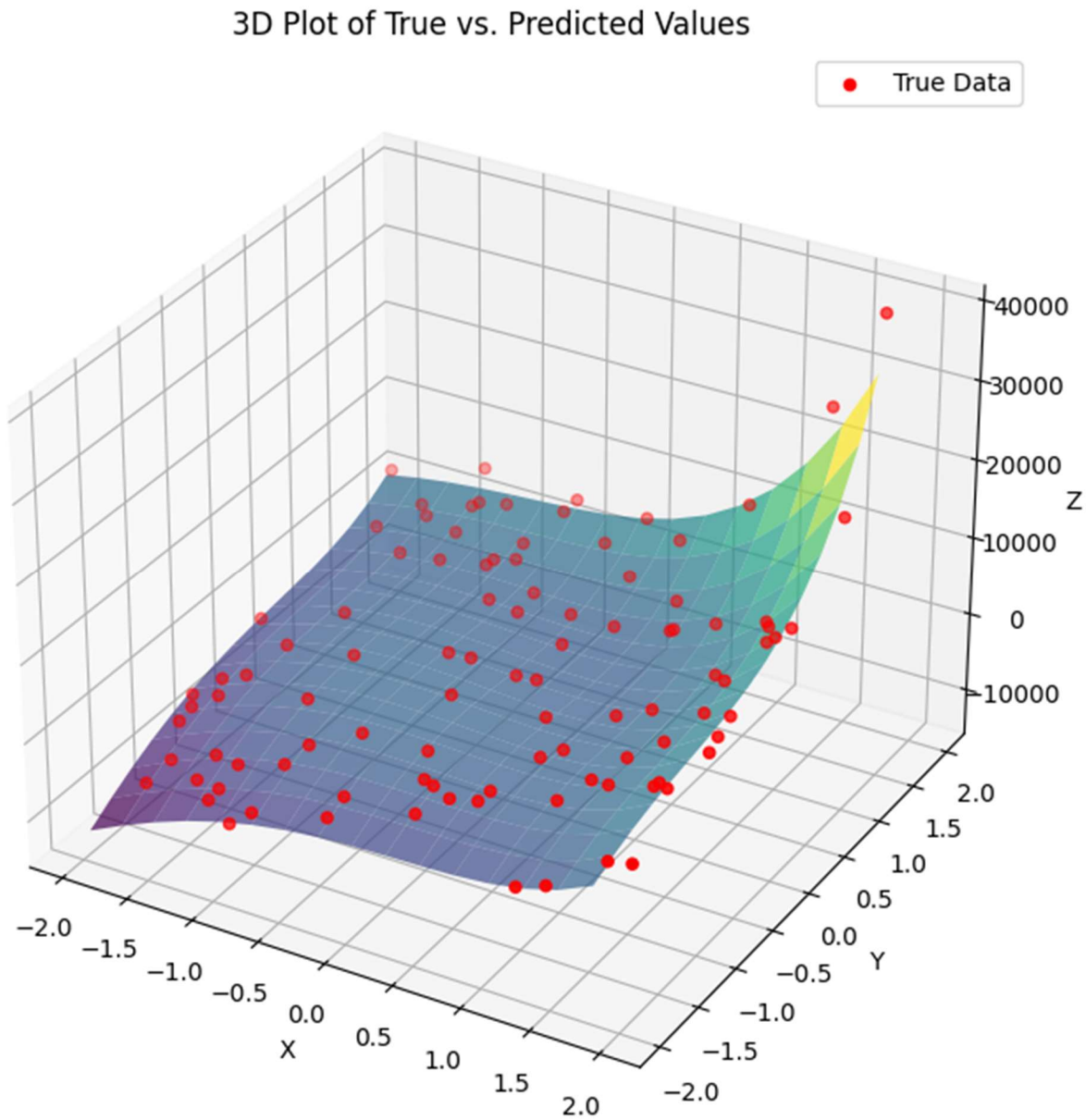
Train_err: 357.63195231413323

val_err 427.806601445848

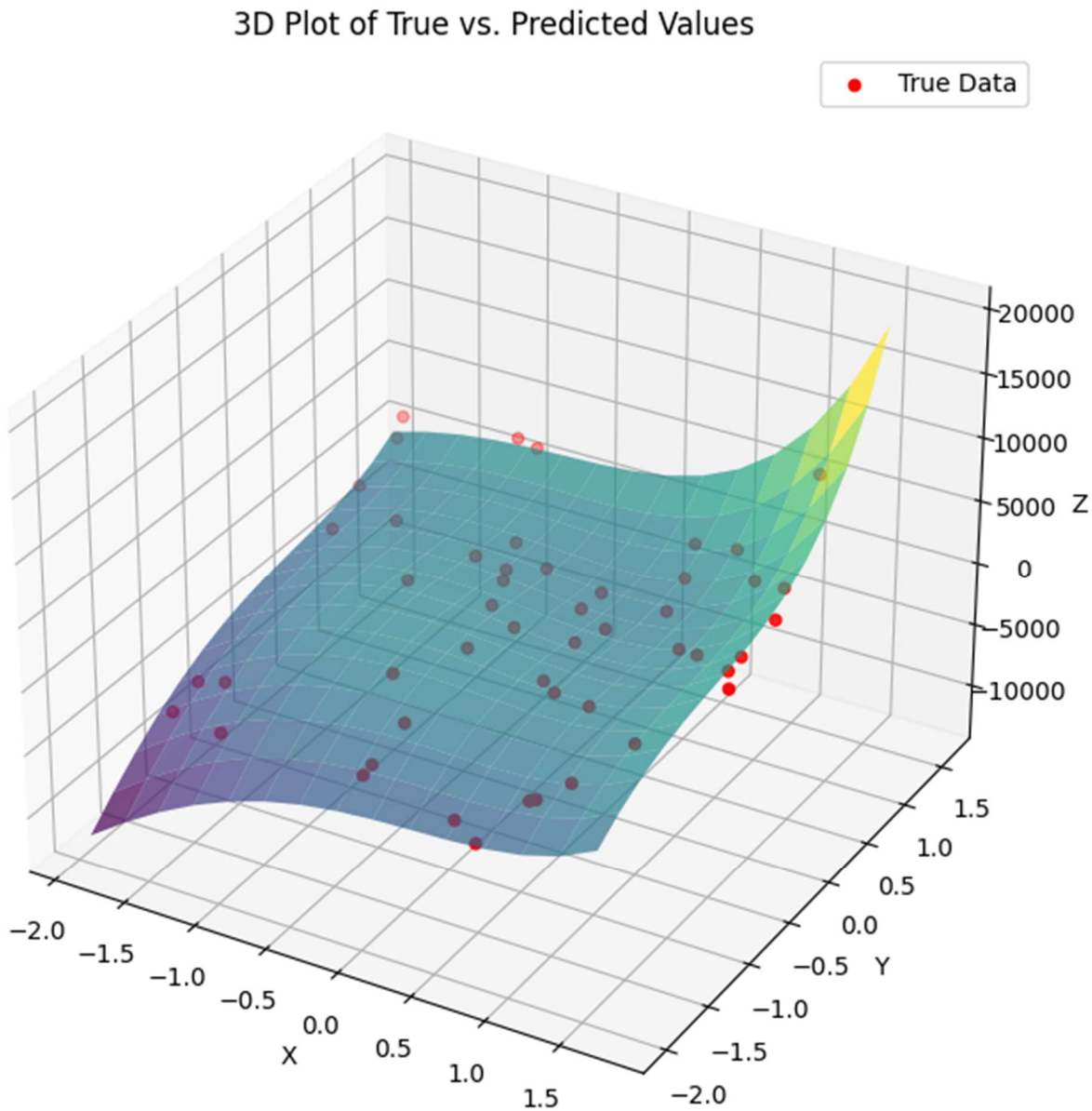
test_err 410.41120586838



3D plot for Train data



3D plot for Test data

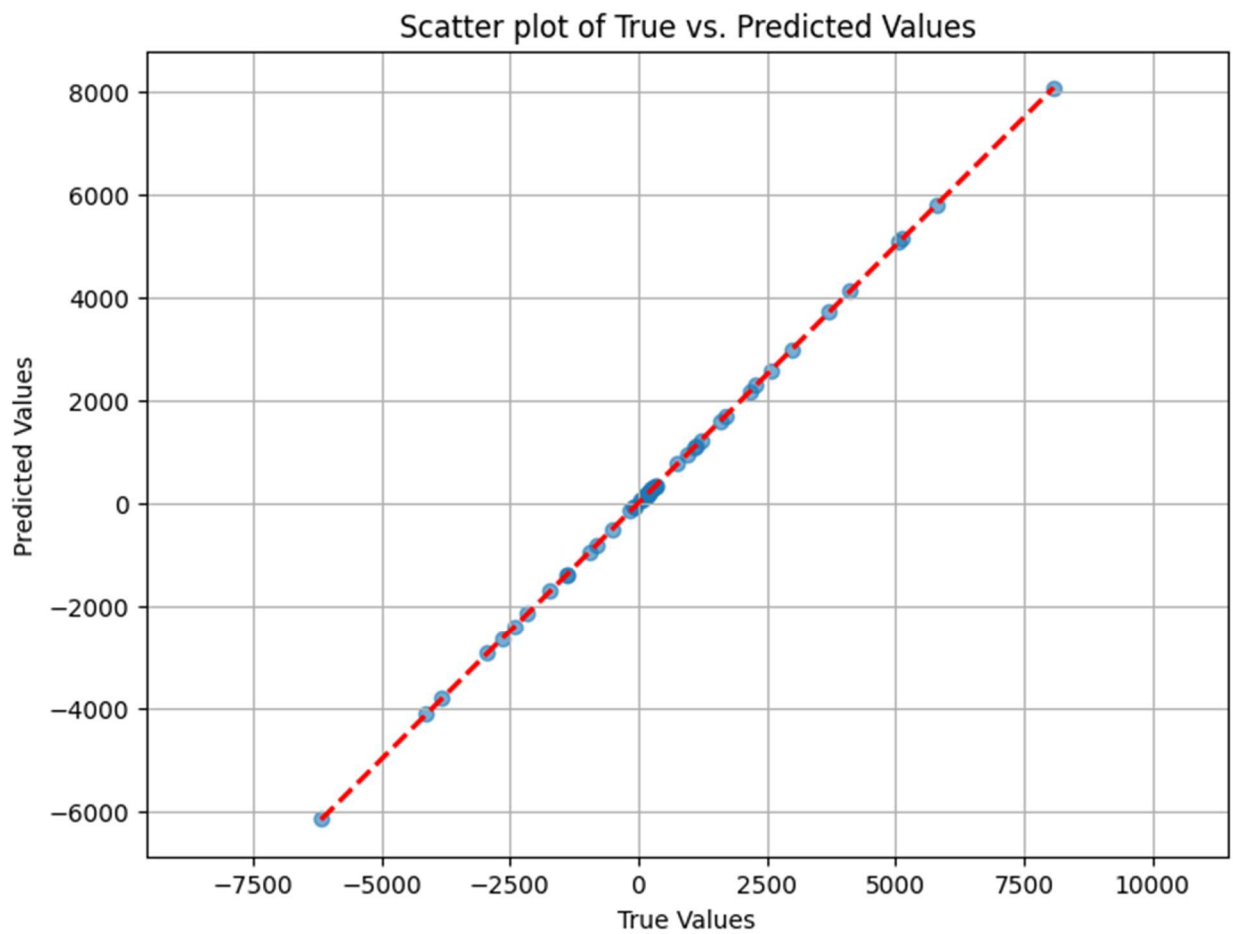


Case M=6 and Lambda =1

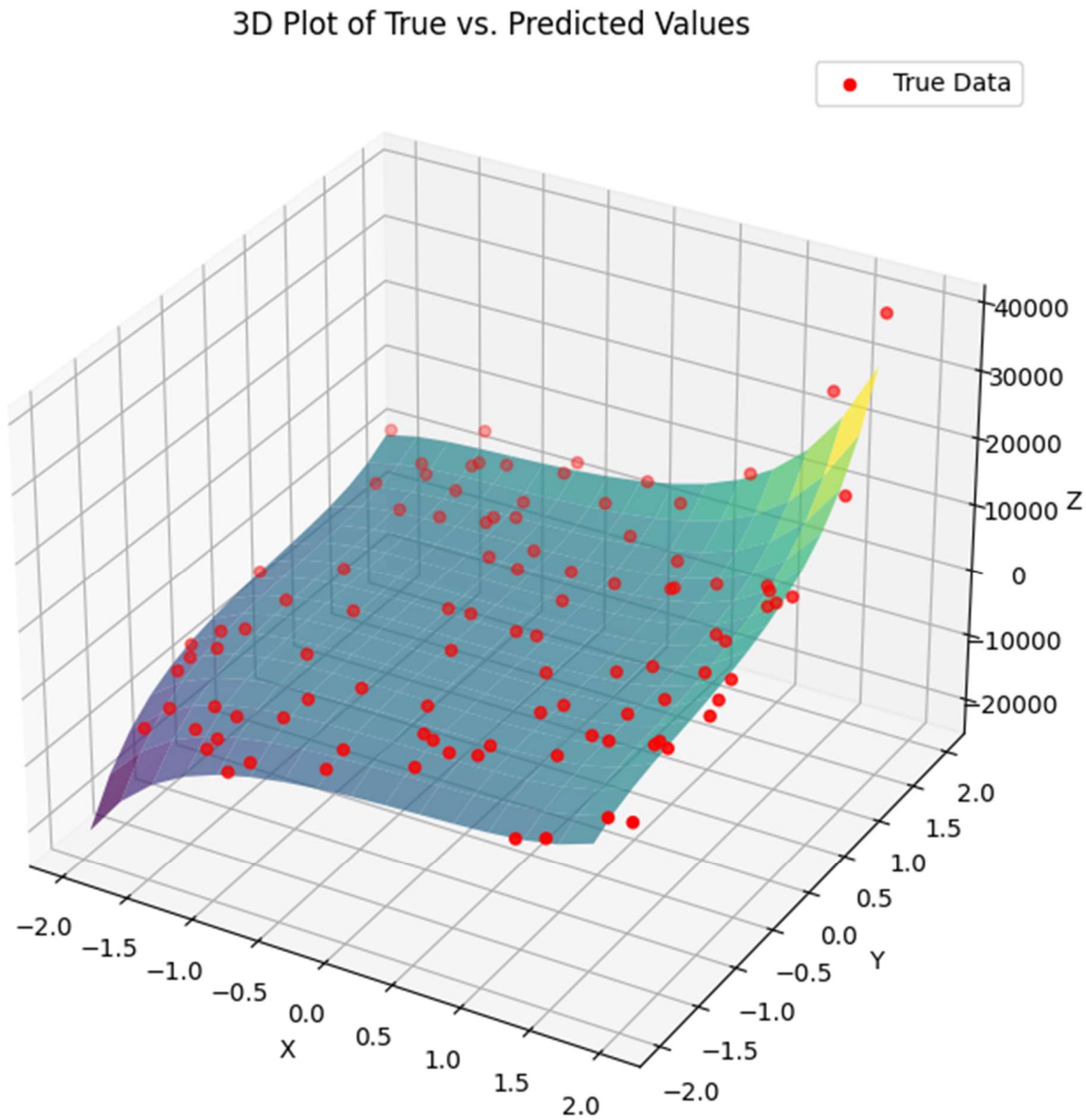
Train_err: 13.49336681997728

val_err 7.481292869227013

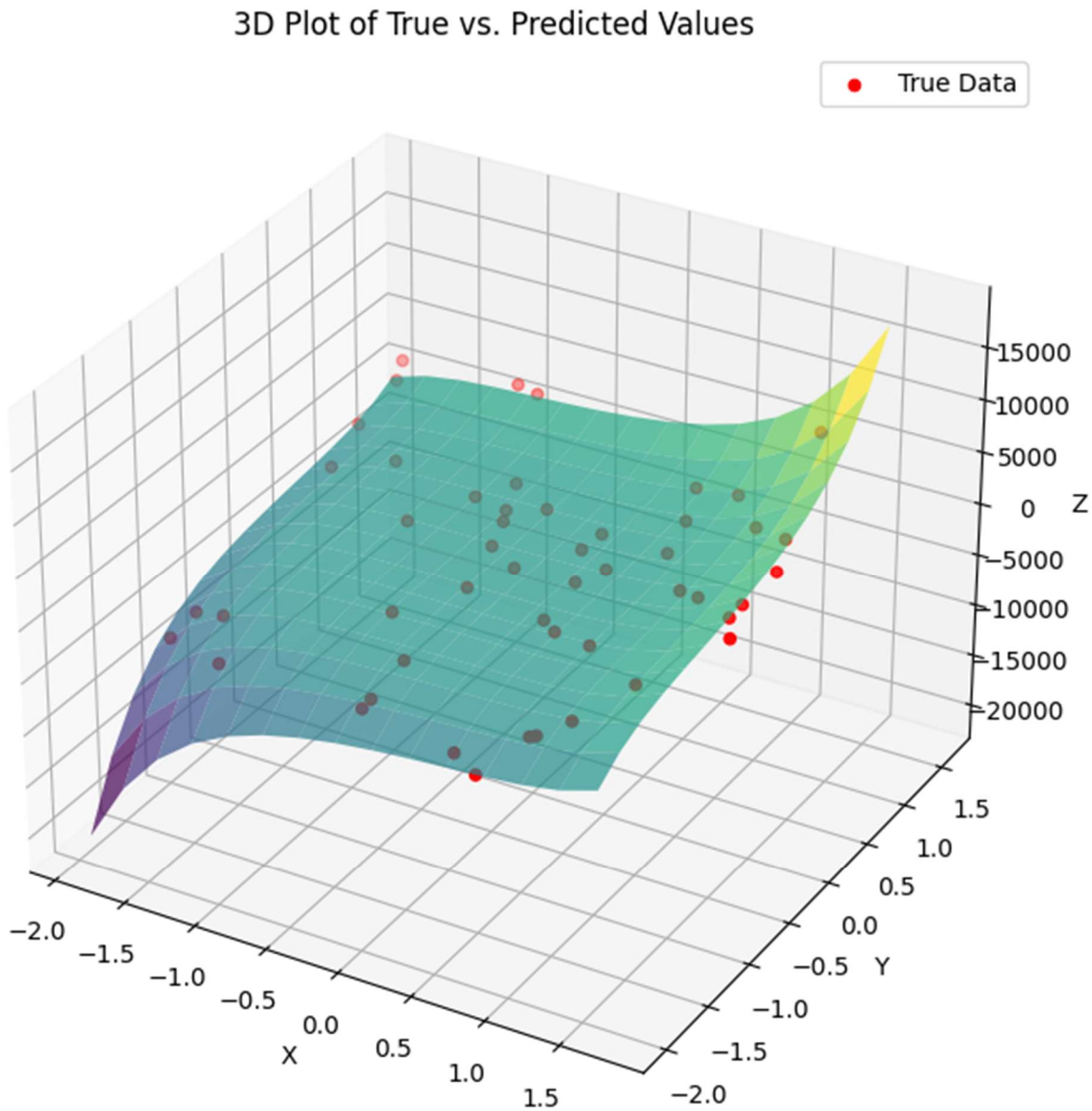
test_err 8.052610760286651



3D plot for Train data



3D plot for Test data

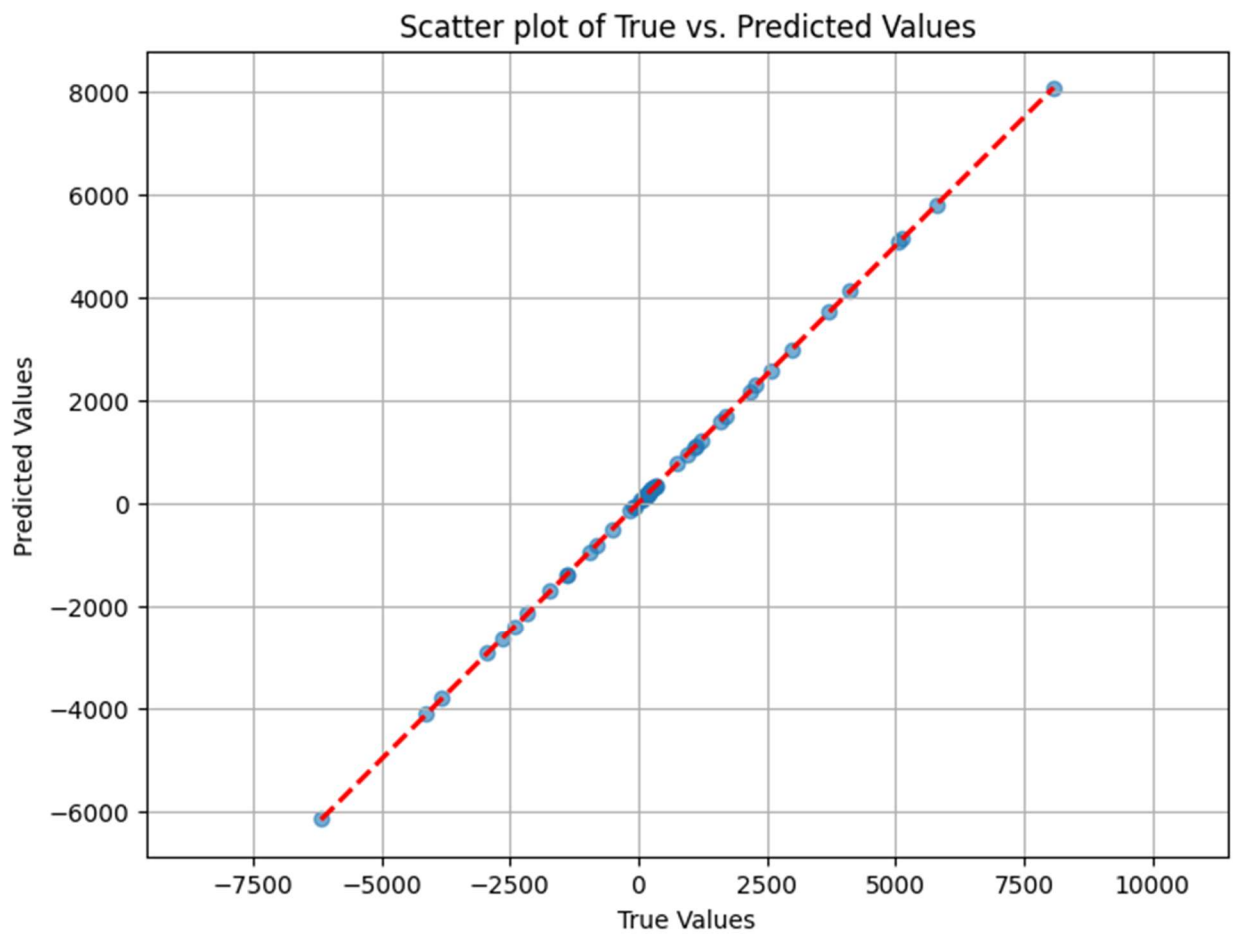


Case M=6 and Lambda =0.001

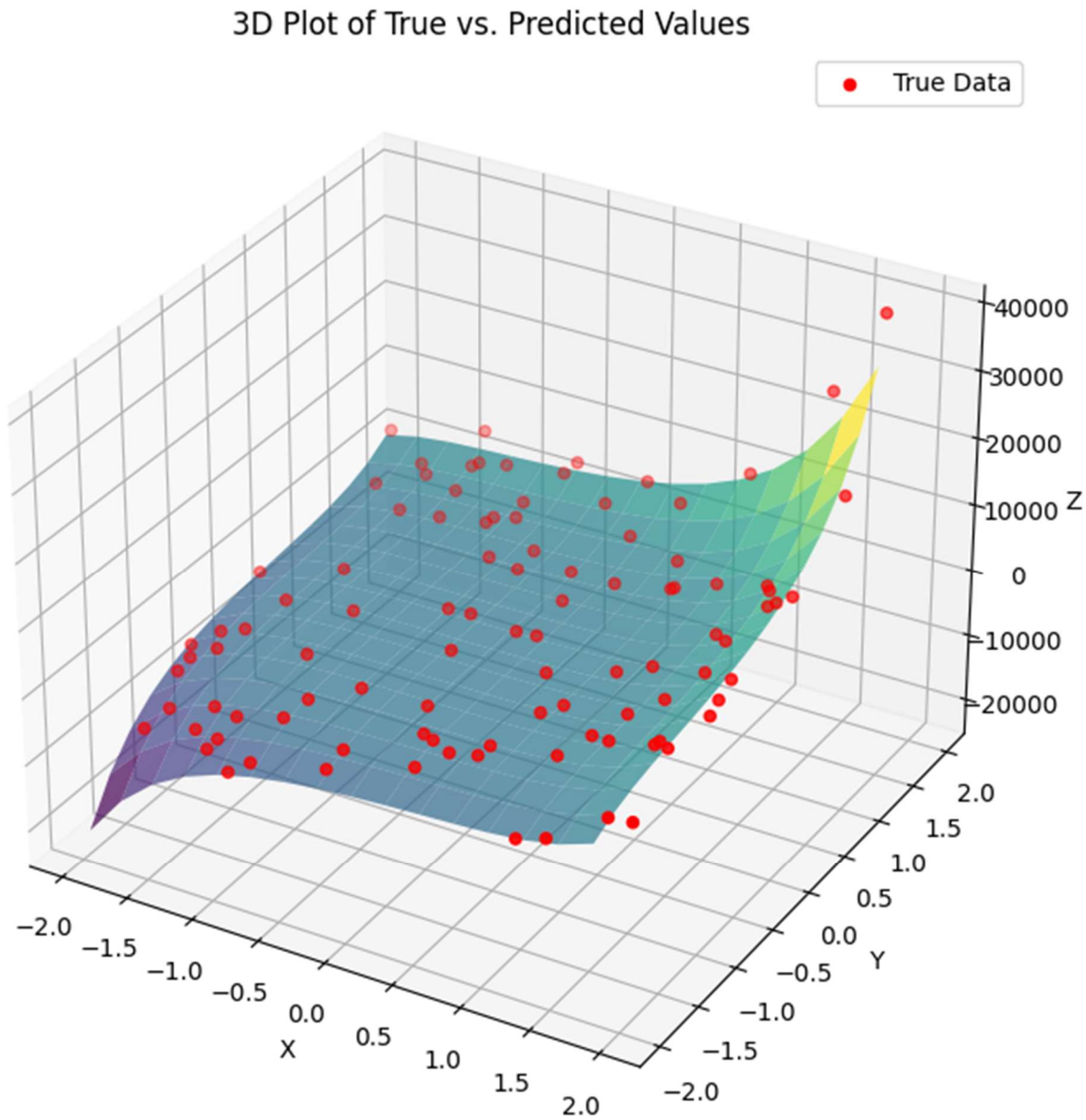
Train_err: 13.493370463787427

val_err 7.486626803407363

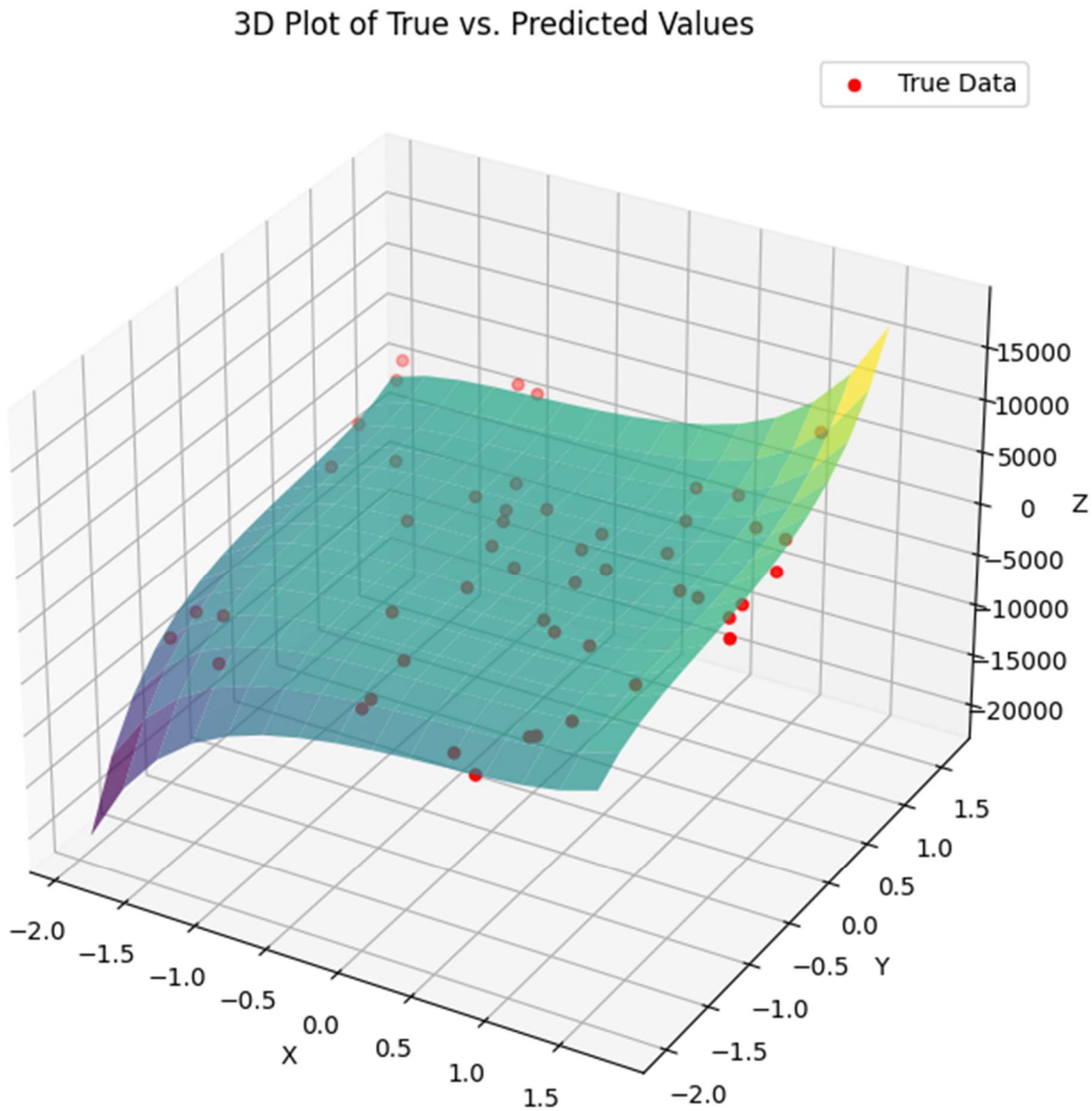
test_err 8.052756250754877



3D plot for Train data



3D plot for Test data

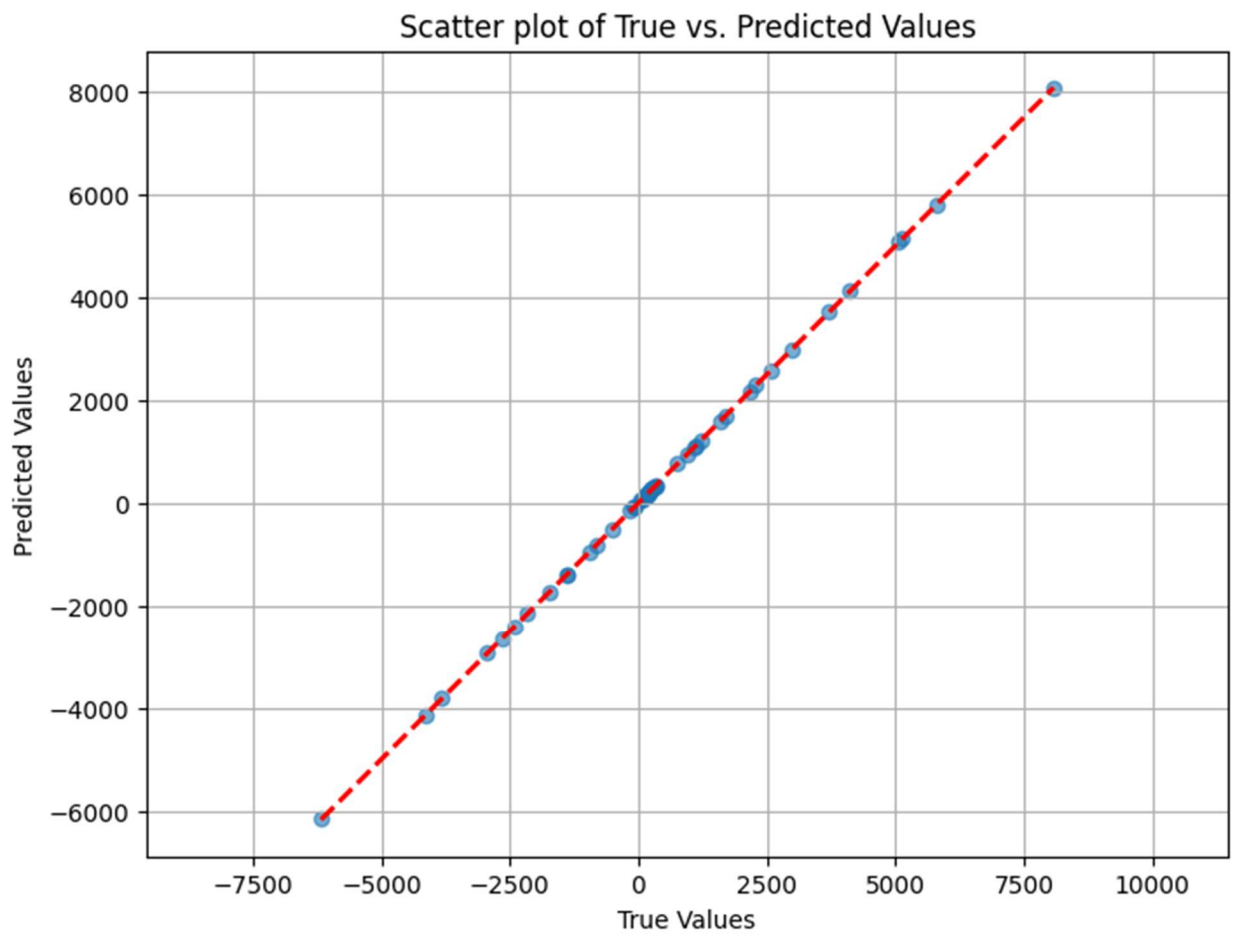


Case M=6 and Lambda =0.1

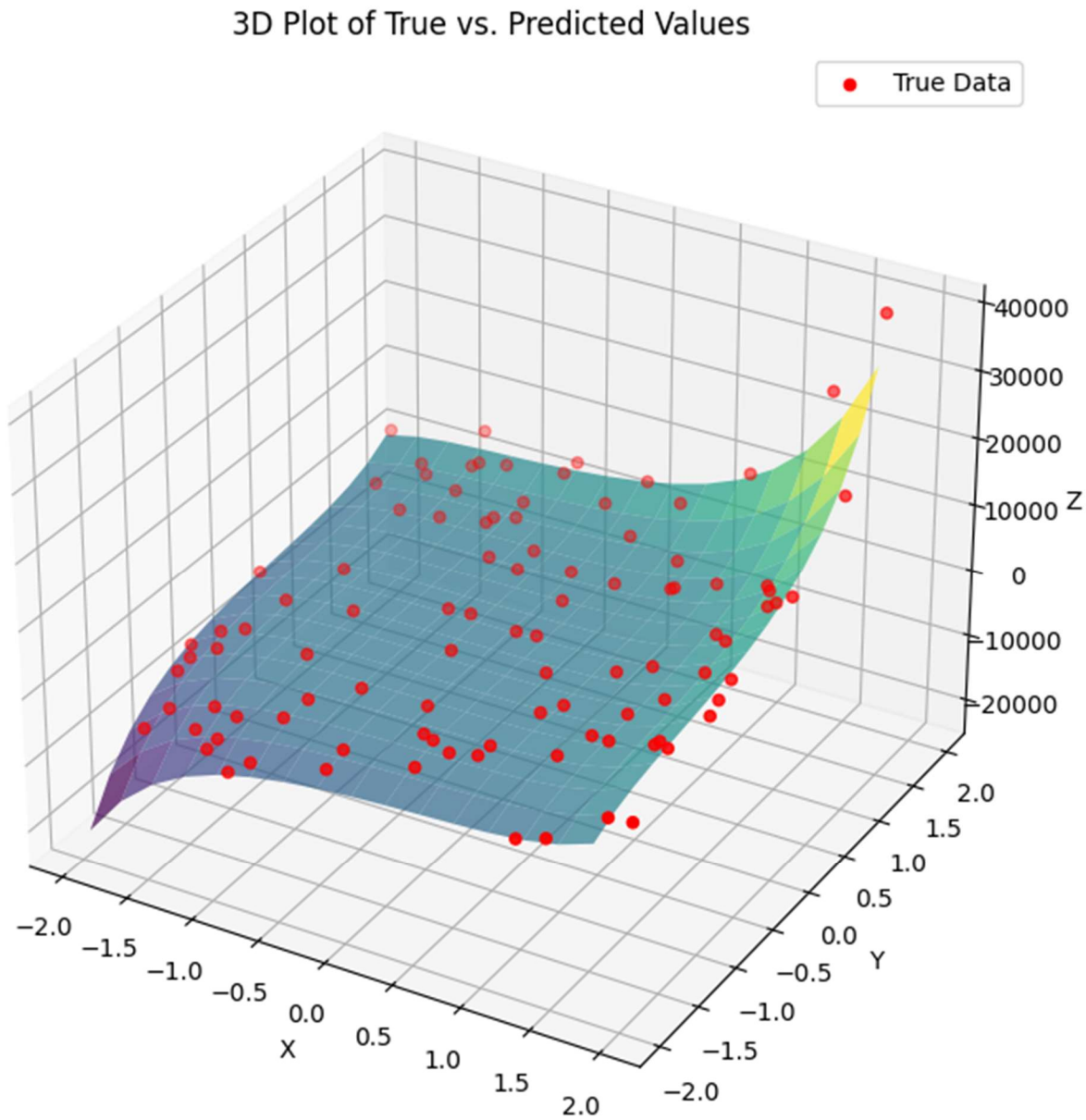
Train_err: 13.528313476359504

val_err 8.10540878096103

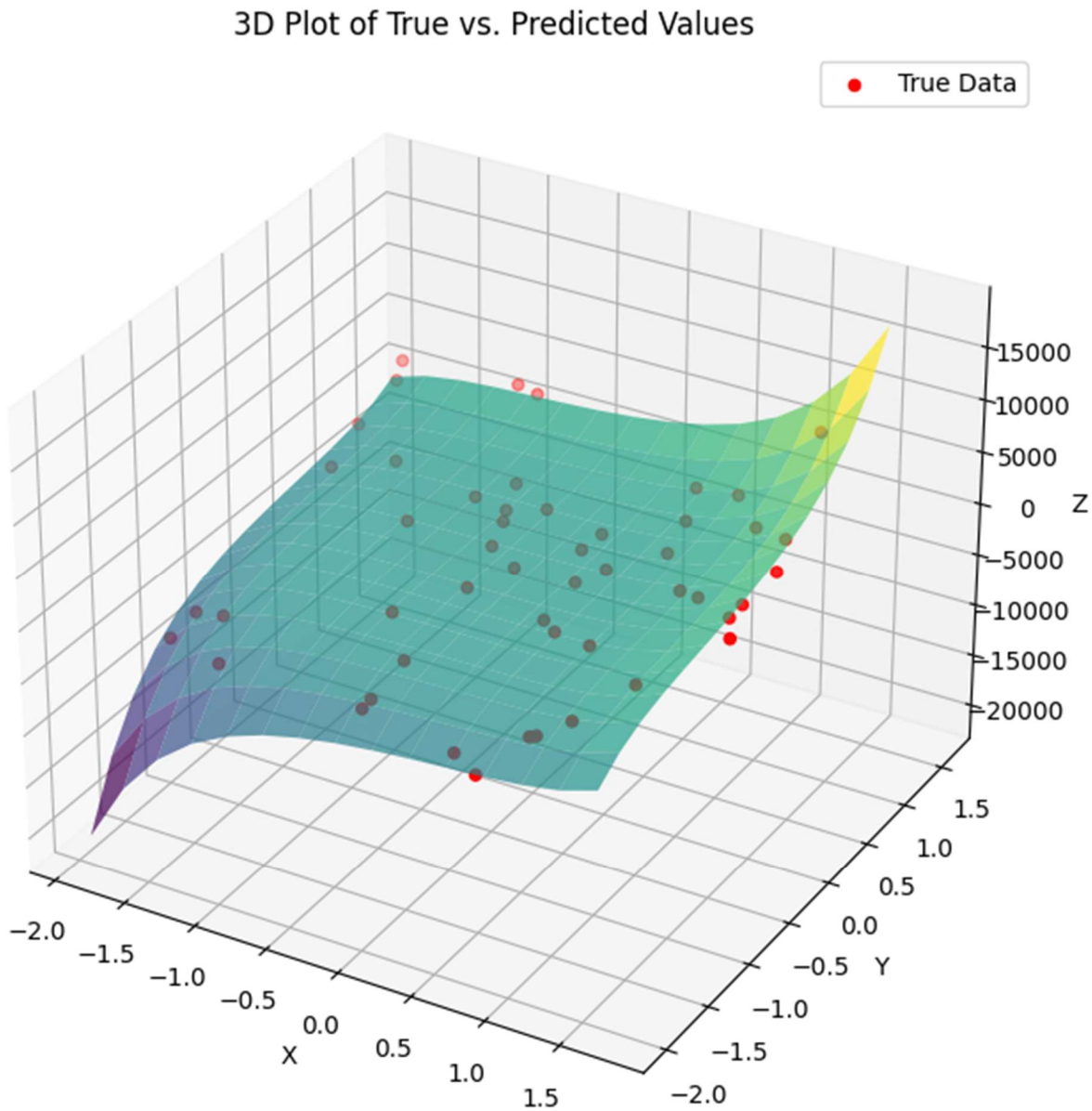
test_err 8.162154090199921



3D plot for Train data



3D plot for Test data

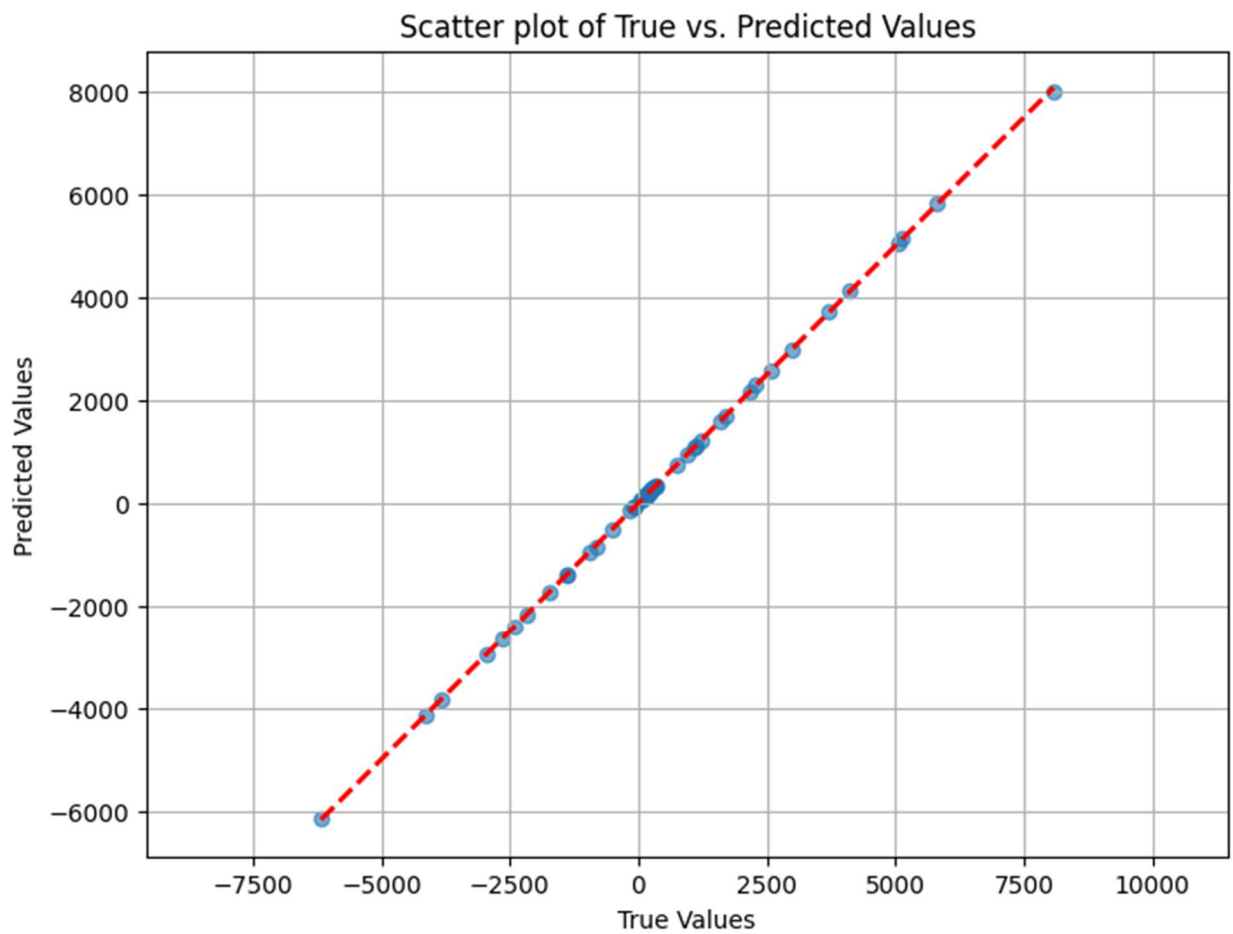


Case M=6 and Lambda =1

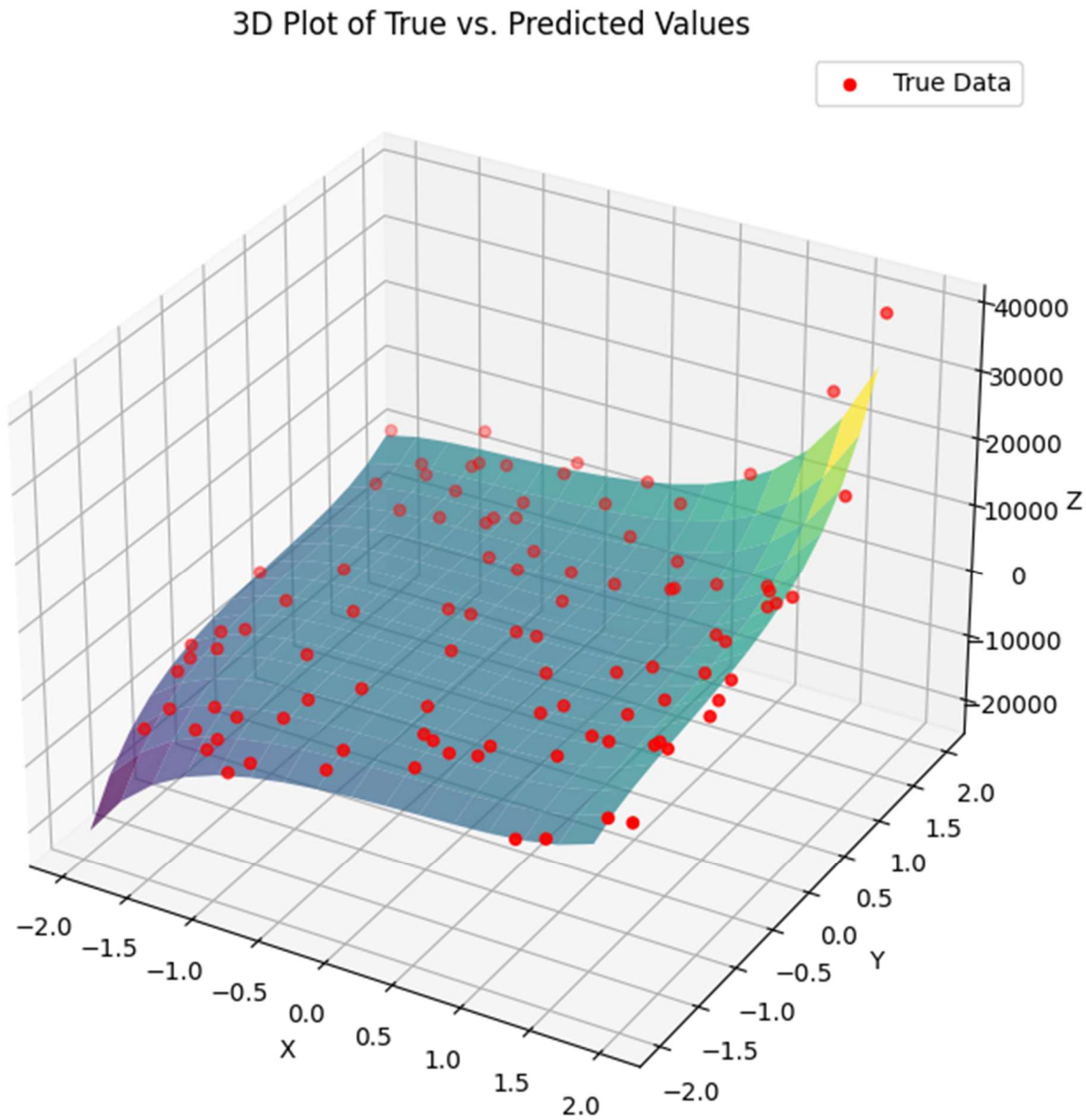
Train_err: 16.08058210402508

val_err 16.442681033435044

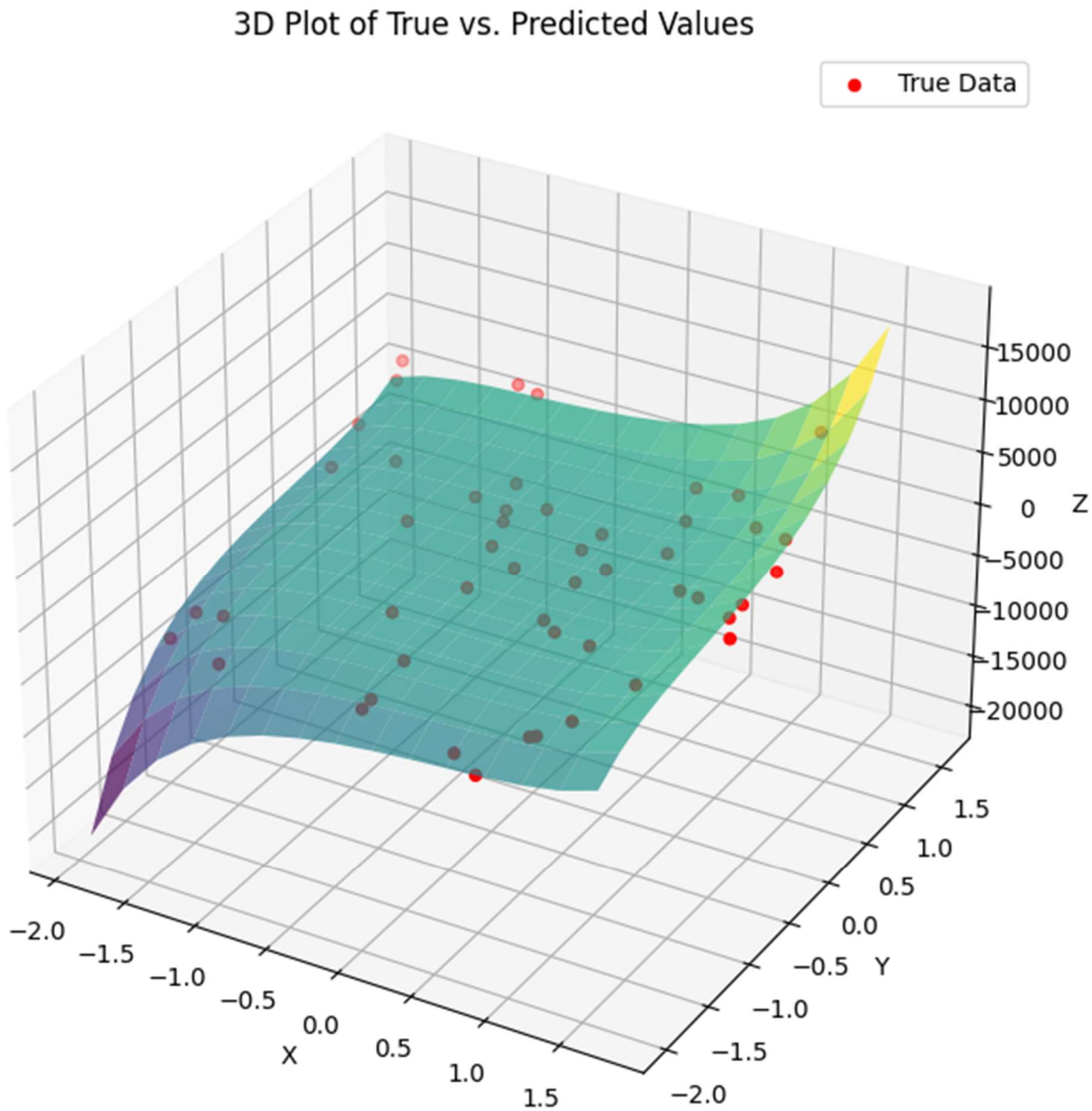
test_err 13.731522936123316



3D plot for Train data



3D plot for Test data


3D Plot of True vs. Predicted Values
True Data
15000
10000
5000
0
−5000
−10000
−15000
−20000
Z
−2.0
−1.5
−1.0
−0.5
0.0
0.5
1.0
1.5
X
1.5
1.0
0.5
0.0
−0.5
−1.0
−1.5
−2.0
Y

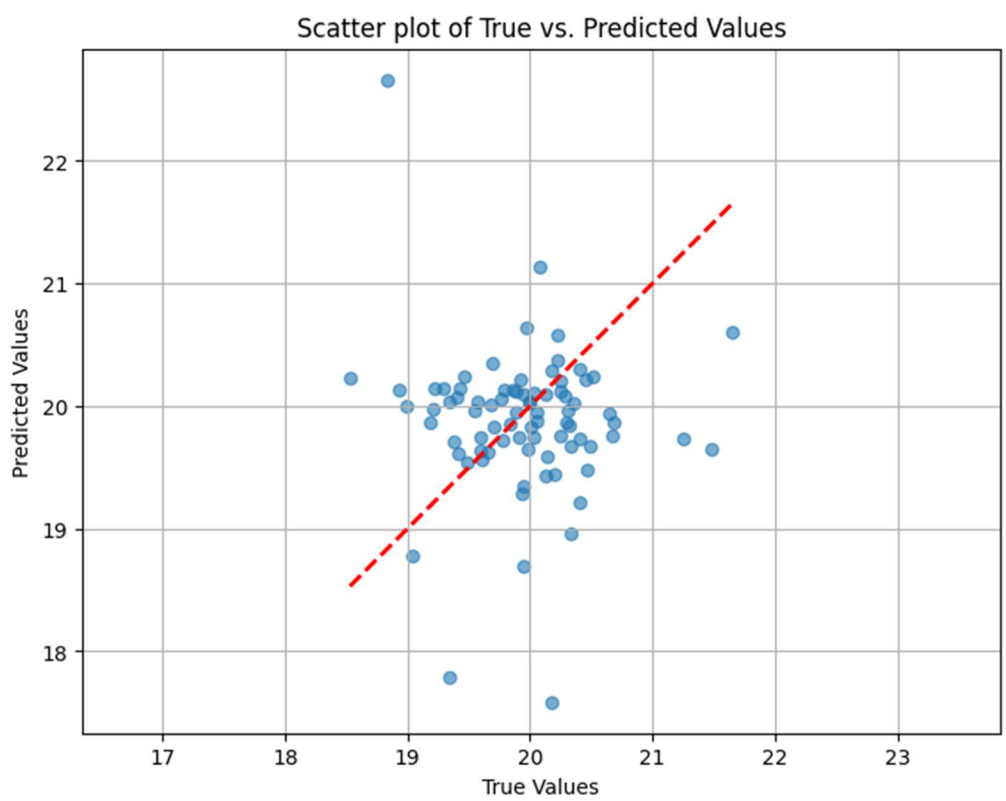
Dataset 3:

Multivariate data

Dataset	Degree of Polynomial	Regularization Coefficient λ	ERMS (Training)	ERMS (Validation)	ERMS (Test)
Dataset 1	2	None	0.462419	0.493614	0.583119
Dataset 1	2	0.001	0.462632	0.492805	0.589583
Dataset 1	2	None	0.462419	0.493614	0.583119
Dataset 1	2	0.1	0.463901	0.493517	0.598979
Dataset 1	2	None	0.462419	0.493614	0.583119
Dataset 1	2	1	0.464921	0.498695	0.588663
Dataset 1	3	None	0.37013	0.832919	0.830145
Dataset 1	3	0.001	0.370136	0.81074	0.833053
Dataset 1	3	None	0.37013	0.832919	0.830145
Dataset 1	3	0.1	0.371821	0.923476	0.847176
Dataset 1	3	None	0.37013	0.832919	0.830145
Dataset 1	3	1	0.37481	0.910188	0.842792

Best Model is with M=2 and lambda=0 with test error of 0.5831193638743538

The Plot for the above data is given below:



Result for Dataset 3:

Best Model is with M=2 and lambda=0 with test error of 0.5831193638743538