TRINITY COLLEGE DUBLIN

School of Computer Science and Statistics

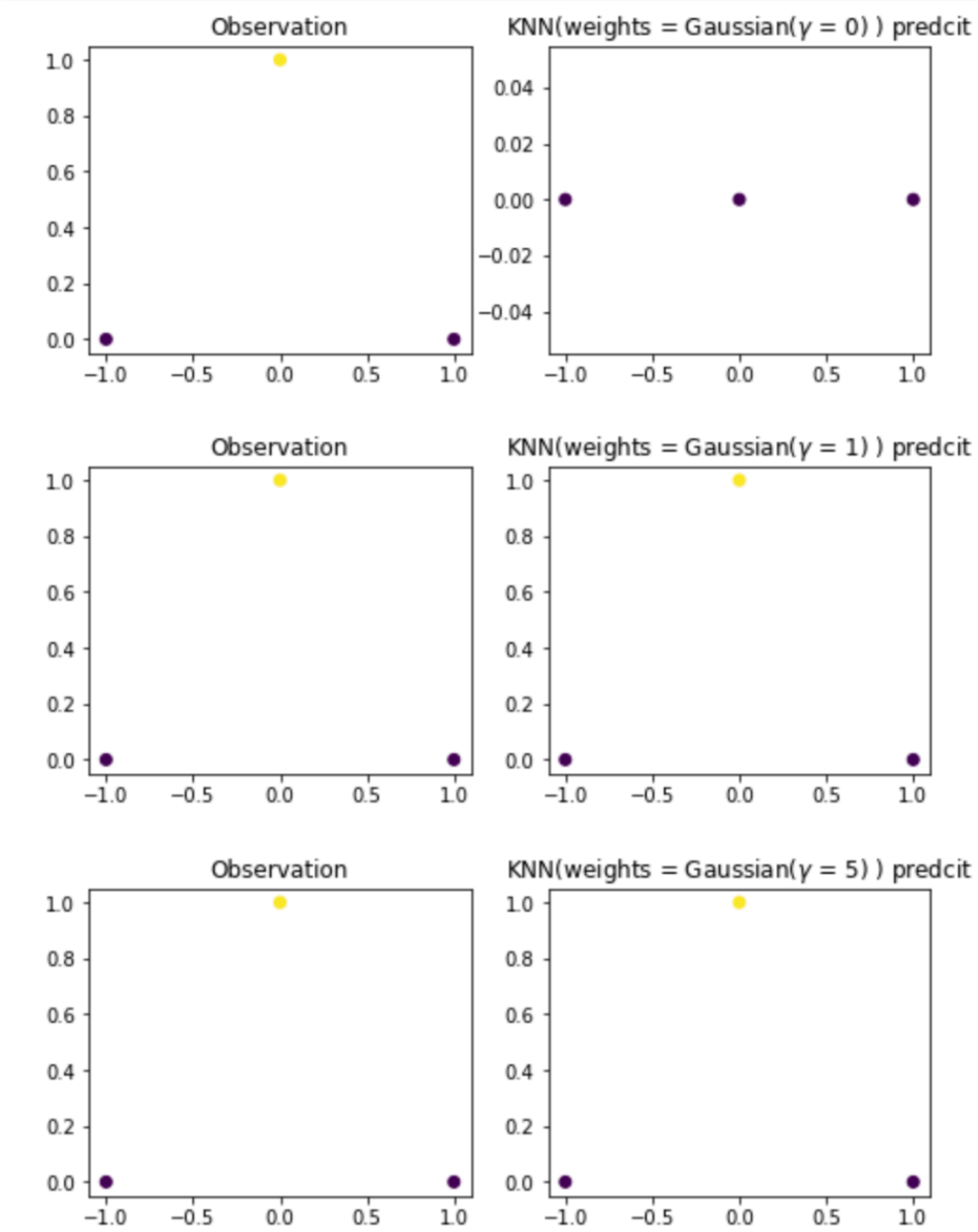
Week 6 Assignment CS7CS4/CSU44061 Machine Learning

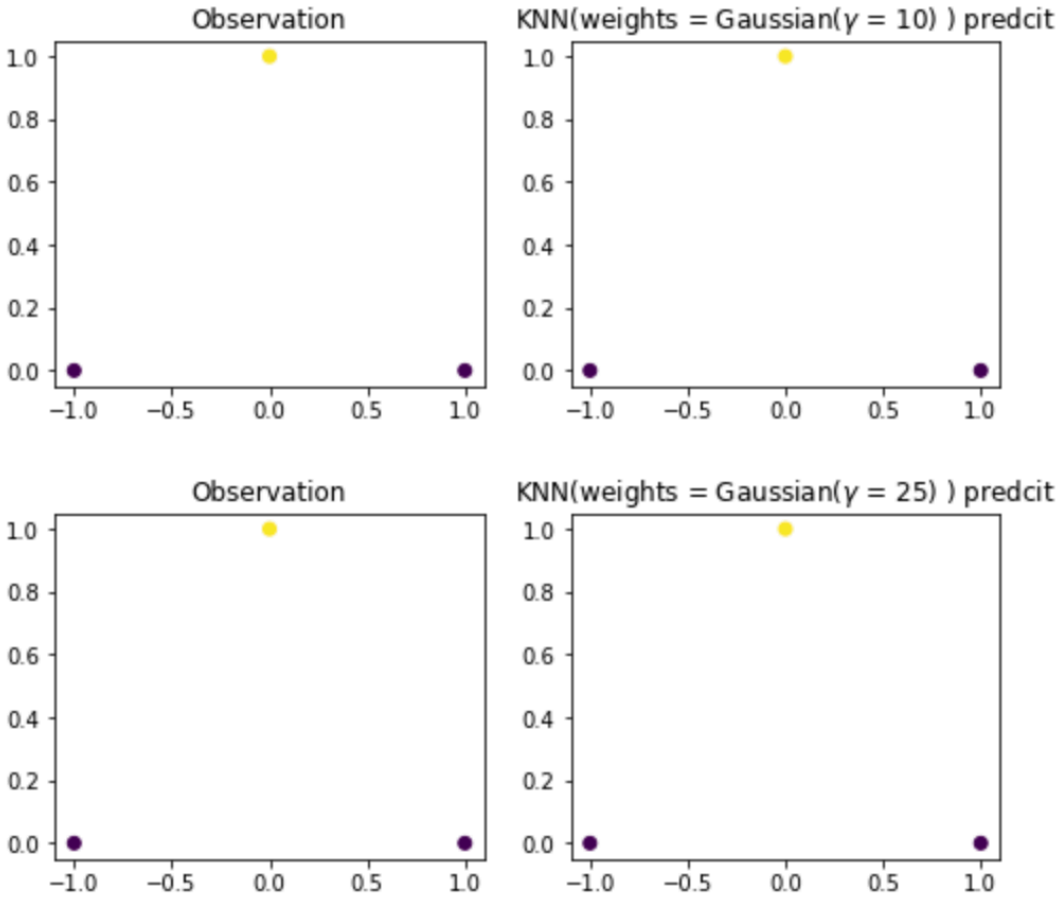
Min Zhang

(i)

(a)

When gamma=0, it can be seen that KNN(3) predicted 0 labels in our three customized samples. With the increase of gamma, when gamma= 1,5,10,25, KNN(3) predicted 100% correctly in our three customized samples.

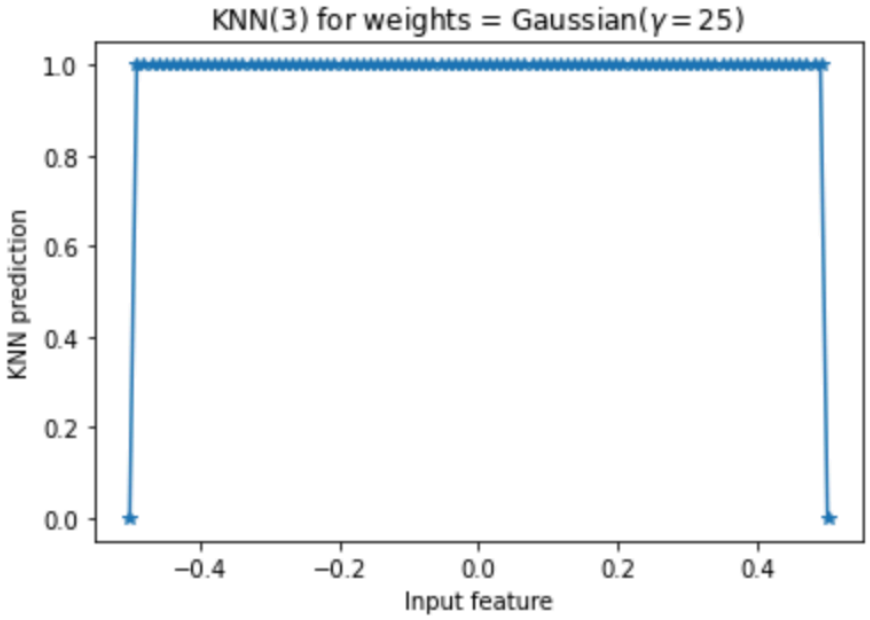



Observation
KNN(weights = Gaussian(γ = 10)) predcit
Observation
KNN(weights = Gaussian(γ = 25)) predcit
1.0
0.8
0.6
0.4
0.2
0.0
−1.0
−0.5
0.0
0.5
1.0

(b)

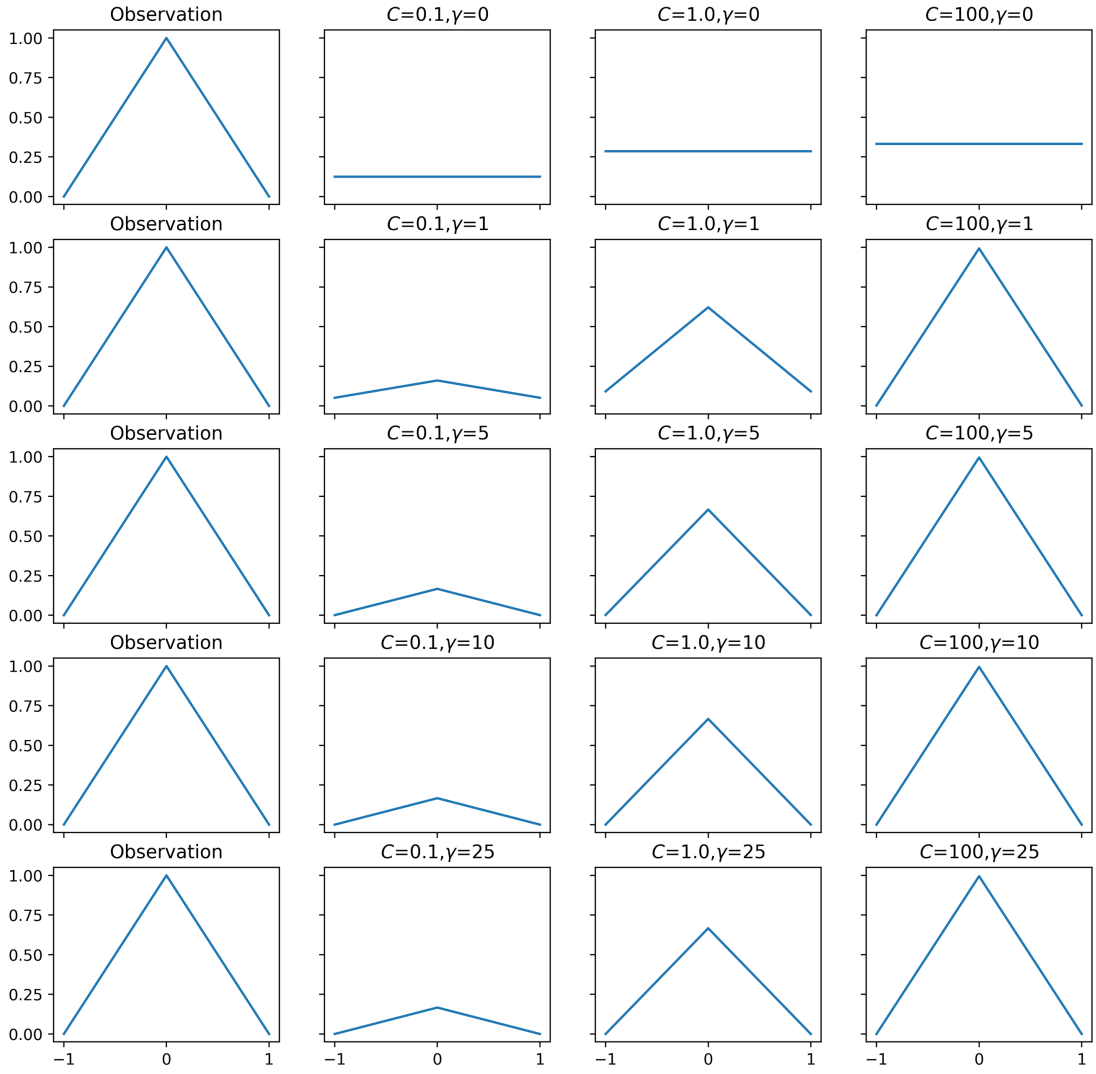
When gamma is larger and larger, the farther away the training sample is from the test sample, the smaller the weight of the training sample is. Even if the training sample is far enough away from the test sample, their label has little influence on the label of the test sample. Taking KNN (3) in gamma=25 as an example, in our three customized training samples, when the input features of the test samples are less than -0.5 or greater than 0.5, these samples are far from the training samples (0,1).

It predicts that the label is 0. When the test sample is close to the training sample (0,1), the model predicts that the label is 1. Although there are more samples with label 0 in the neighbor recently, but their weight is relatively small.



(c)

For the kernelRidge model with different parameters, the predictions in the three custom training samples are shown in the figure below. It can be observed that the control gamma is constant, and the predicted value is higher and higher as C increases. The value of C is kept constant. As the value of gamma increases, kernelRidge's predicted value for the sample with label 1 becomes larger and larger, while for the sample with label 0, there is no significant change.



The trained parameters of each model are shown as follows:

gamma	C	parameter1	parameter2	parameter3
0	0.1	-2.500000e-02	0.175000	-2.500000e-02
0	1.0	-5.714286e-01	1.428571	-5.714286e-01
0	100.0	-6.655574e+01	133.444260	-6.655574e+01
1	0.1	-1.026472e-02	0.167925	-1.026472e-02
1	1.0	-1.833162e-01	0.756584	-1.833162e-01
1	100.0	-4.854817e-01	1.350445	-4.854817e-01
5	0.1	-1.871657e-04	0.166667	-1.871657e-04
5	1.0	-2.994764e-03	0.666694	-2.994764e-03
5	100.0	-6.671669e-03	0.995114	-6.671669e-03
10	0.1	-1.261109e-06	0.166667	-1.261109e-06
10	1.0	-2.017775e-05	0.666667	-2.017775e-05
10	100.0	-4.494931e-05	0.995025	-4.494931e-05
25	0.1	-3.857762e-13	0.166667	-3.857762e-13
25	1.0	-6.172419e-12	0.666667	-6.172419e-12
25	100.0	-1.375010e-11	0.995025	-1.375010e-11

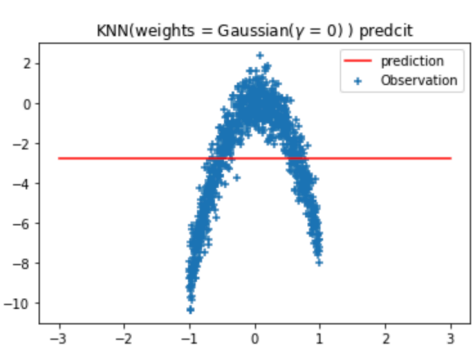
(d)

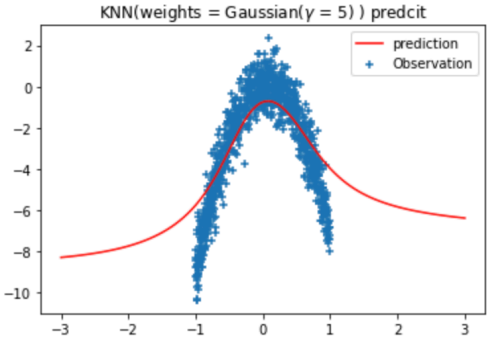
It can be seen from the parameter results in (c) that with the increase of gamma, the penalty of the number of parameters of the loss coefficient will increase, and the first parameter and the third parameter will rapidly decrease to around 0, allowing the second parameter to play a major role in the kernelRidge model. With the increase of parameter C, alpha will decrease, and the second parameter will increase accordingly, thus changing the model prediction. In the customized training sample, when the C and Gamma parameters of kernelRidge model are large, its prediction accuracy can be comparable to KNN (Gaussian kernel weight).

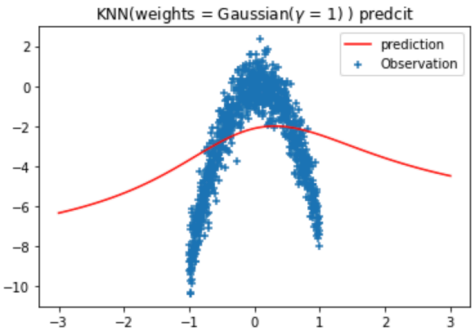
(ii)

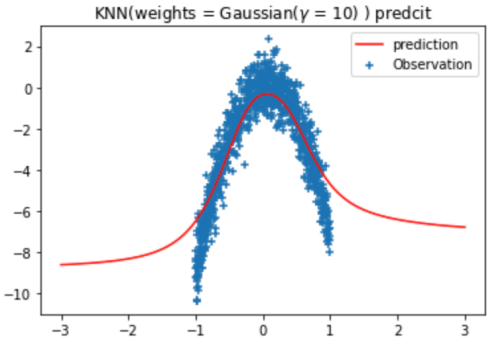
(a)

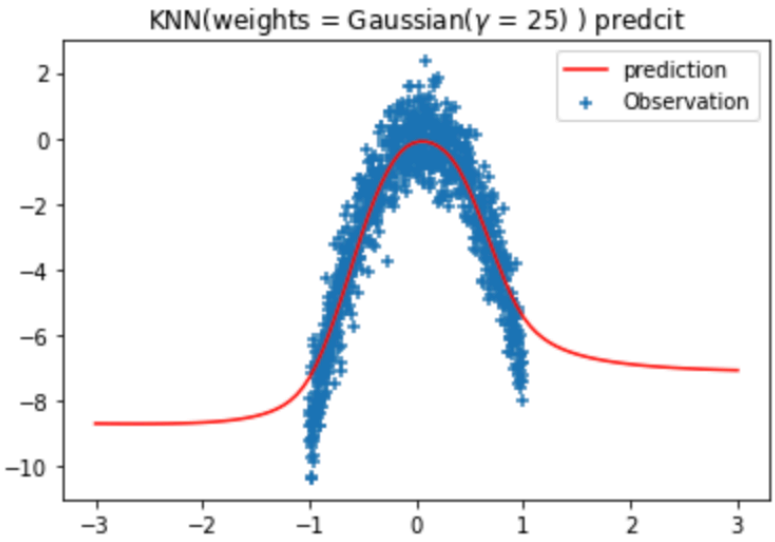
If the data of this assignment is used for prediction, a weighted KNN model is established, where the nearest neighbor is the sample number of the data set. When the gamma value is equal to 0, KNN's regression prediction for all samples is the same value, which is the average expected output of all samples. With the increase of gamma, KNN regression gets better and better.





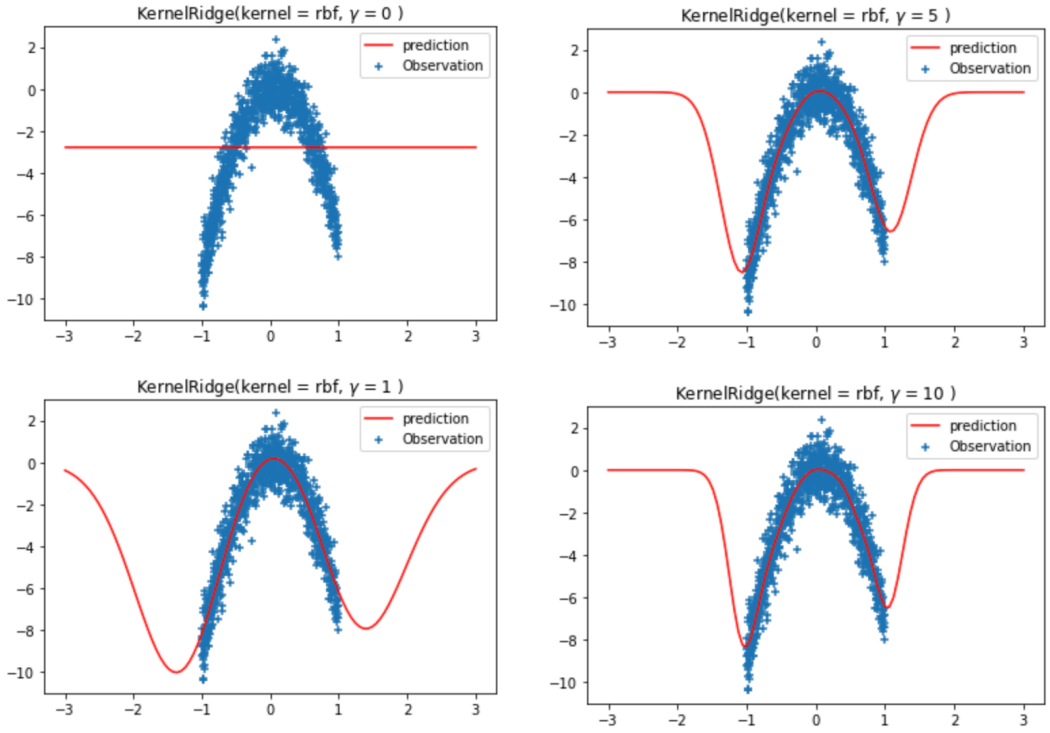


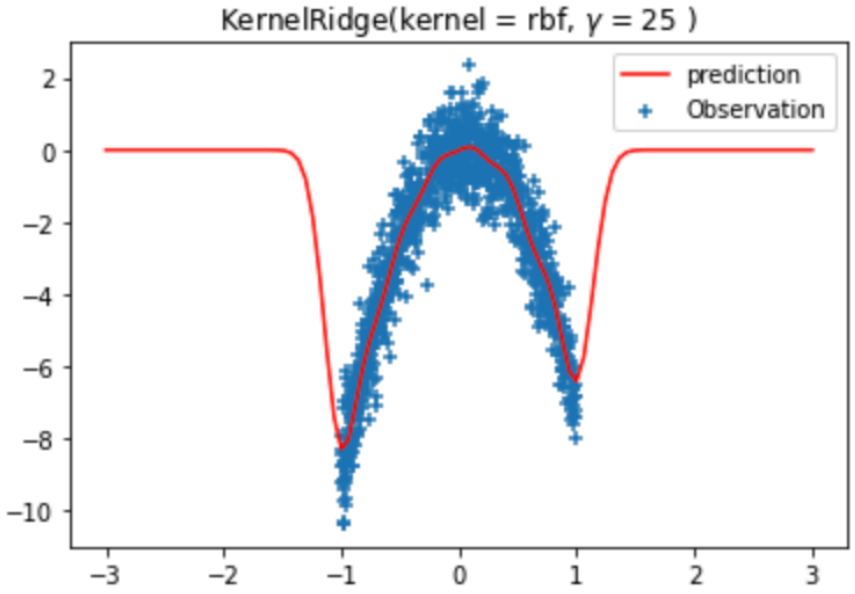




(b)

The results are similar for the KernelRidge (RBF kernel) regression model, although the KernelRidge prediction appears to be a little more complex than the KNN graph.





(c)

In order to find the KNN regression and the best parameters of kernelRidge, the 5-fold cross validation and grid search algorithm are adopted to find the best KNN model and kernelRidge. By comparison, the weight parameter in the optimal KNN model is the Gaussian weight function (Gamma =25). The optimal kernelRidge parameter is C=100, Gamma =1, kernel=' RBF '.

The optimal KNN model and kernelRidge model are used to predict the data again. Through comparison, it can be found that the prediction accuracy of kernelRidge model is a little higher. The MSE of KNN model is 0.8104. KernelRidge had 0.6153.