AI/ML - CS 337 Lab - Lab Assignment 3

B.Nikhil 170050099

September 15, 2019

Problem 1: Logistic Regression

Accuracy on test data is 89.94. If we observe the data we can see the number of subscribers to term deposit is 5289 among 45212 people which is nearly 88.30. Therefore if we just predict *no* for everyone then we get 88.30 accuracy. This implies accuracy is not the good performance metric.

Problem 2: Kernel Perceptron

Accuracy comes to be 100 for all the three kernels(linear, Polynomial with degree 3 and Gaussian with sigma 4). Accuracy for Gaussian with sigma 5 is 79.125(bad). Values in alpha are all positive integers, by bounding each alpha value with 1 we get 86.12 accuracy for Gaussian, 100 for linear and 0 for polynomial. Similarly by bounding alpha with 2 we get 0.0275 accuracy for Gaussian, 100 for linear and 0 again for polynomial. We observe a drastic change in case of polynomial less change in case of Gaussian and no change in case of linear.

Problem 3: Kernel Logistic Regression

In main, data is read from *dataset1.txt* and the function *logisticKernel* is called with Gaussian kernels of different sigma. Following Plot shows number of mistakes for each given sigma.

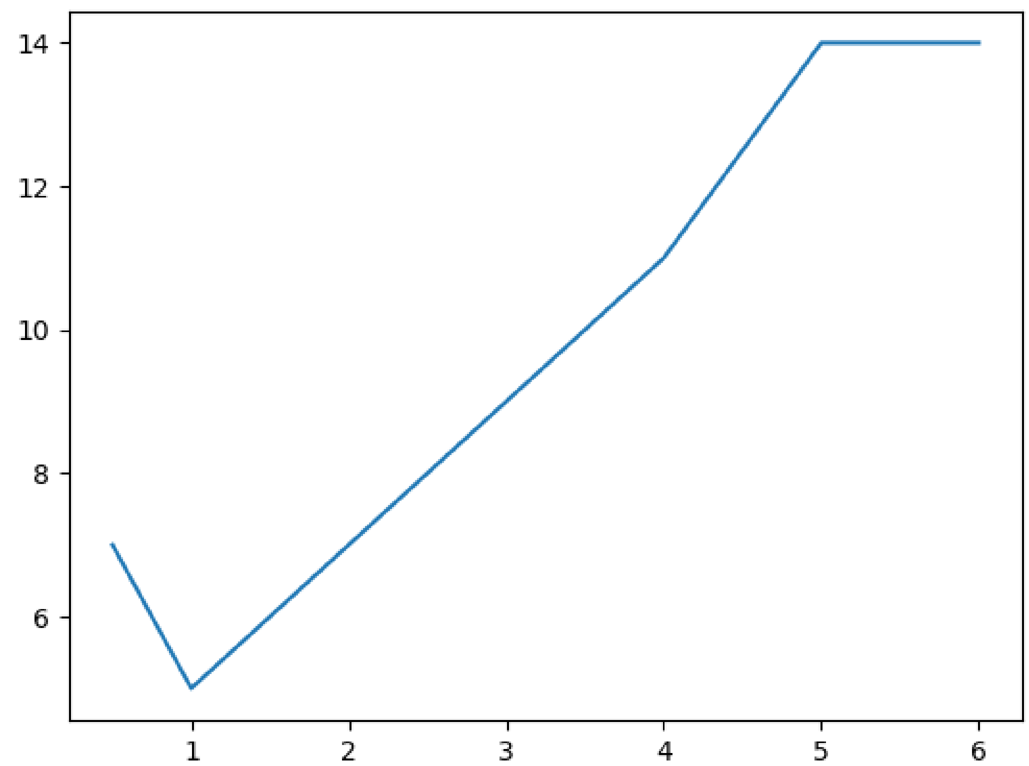

Figure 1: Number of mistakes for each given sigma

Problem 4: Kernel Ridge Regression

Variation of Regression Function with sigma

Gaussian Kernel is given by

$$k(\boldsymbol{x}, \boldsymbol{y}) = \exp \frac{-\|\boldsymbol{x} - \boldsymbol{y}\|}{2\sigma^2}$$

We know that the kernel function determines the similarity between query point and the data point. Decreasing sigma will increase kernel function value and the curve tends to fit/move towards data point.

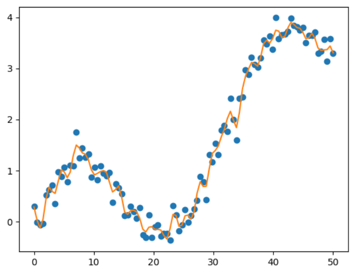

Figure 2: Regression function alongside Training examples for sigma 1 and lambda 0.01.

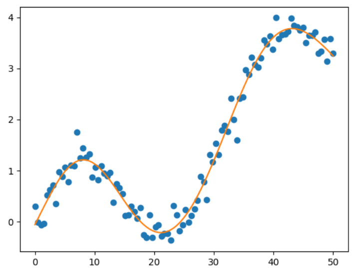

Figure 3: Regression function alongside Training examples for sigma 10 and lambda 0.01.

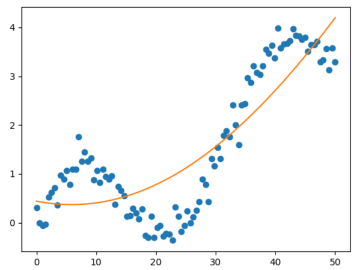

Figure 4: Regression function alongside Training examples for sigma 100 and lambda 0.01.

Variation of Regression Function with lambda

Increasing Lambda(regularizing parameter) reduces over-fitting by decreasing the curvature of the fit. As we can see from the following plots, curvature of the fit decreases from left to right.

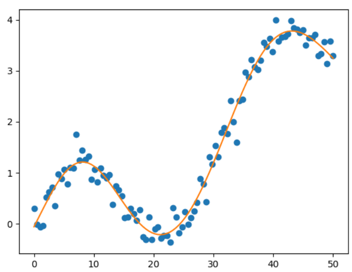

Figure 5: Regression function alongside Training examples for sigma 10 and lambda 0.01.

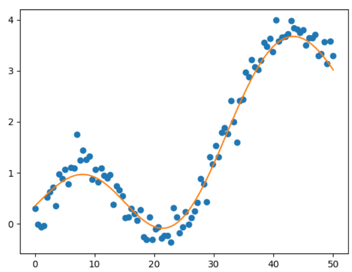

Figure 6: Regression function alongside Training examples for sigma 10 and lambda 1.

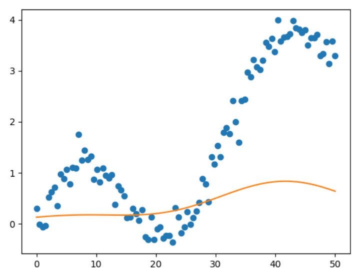

Figure 7: Regression function alongside Training examples for sigma 10 and lambda 100.

$\sigma \to \infty$ and $\sigma \to 0$

As sigma tends to zero, kernel values will be 0 if the norm distance between x and y is more. Therefore,for the points in between data points, curve goes to zero. But for the points closer to data points, curve tries to fit to data points.

As sigma tends to infinity, kernel values tends to 1 and irrespective of norm distance between x and y. Therefore, the regression function would become constant. Refer figure 8 and 9 for sigma tending to zero and sigma tending to infinity.

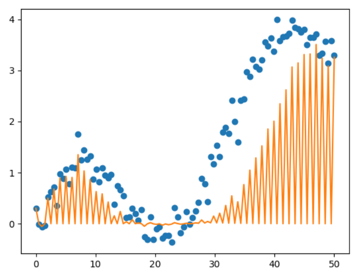

Figure 8: Regression function alongside Training examples for sigma 0.0001 and lambda 0.01.

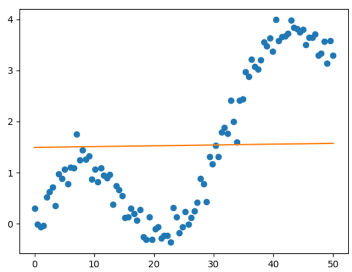

Figure 9: Regression function alongside Training examples for sigma 10000 and lambda 0.01.

My kernel to estimate total fuel consumption

Trying out different possibilities, kernel obtained by doing product of polynomial and Gaussian kernels gives good fitting.

$$k(\boldsymbol{x}, \boldsymbol{y}) = (1 + x^T y)^3 \exp \frac{-\|\boldsymbol{x} - \boldsymbol{y}\|}{2\sigma^2}$$

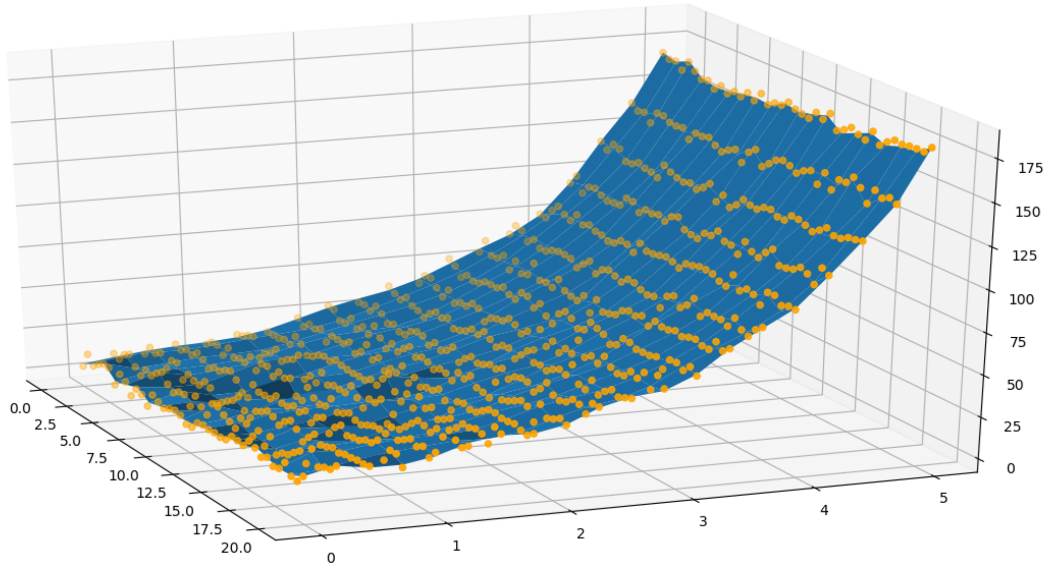

Figure 10: Estimate of fuel by using my kernel

Closed Form solution for f

Yes It is possible to get closed form solution for f, In fact, we use closed form solution to get regression function. Regression Function f is given by,

$$f = \sum_{i=1}^{m} \alpha_i k(\boldsymbol{x}, \boldsymbol{x_i})$$

where

$$\alpha = \Phi^T (\Phi\Phi^T + \lambda I) y$$