COL774: Machine Learning	**Date: February 12, 2019**

Assignment 1 Report

Entry Number: 2016CS10363	**Name: Manish Tanwar**

1 A

2 Locally Weighted Linear Regression:

(a) Linear Regression (unweighted):

$$\theta = \begin{bmatrix}\theta_0 \\ \theta_1\end{bmatrix} = \begin{bmatrix}0.327680 \\ 0.175316\end{bmatrix}$$

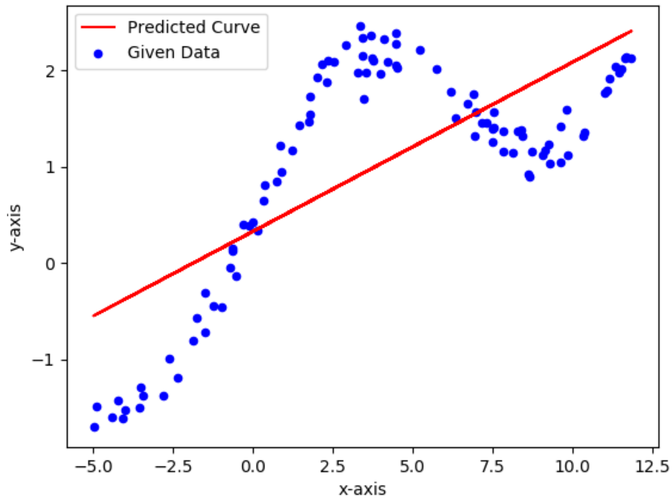


Figure 1: Linear Regressin Plot (Underfitting)

Linear Regression(unweighted) is not a good fit for the data as shown in the Figure 1(underfitting).

(b) Locally Weighted Linear Regression:

Weights : $$w^{(i)} = \exp\left(-\frac{(x - x^{(i)})^2}{2\tau^2}\right) \qquad \text{(where } \tau = \text{Bandwidth Parameter)}$$

Error Function: $$J(\theta) = \frac{1}{2m}(X\theta - Y)^T W (X\theta - Y) \qquad \text{(where } W = diag(w^{(i)})\text{)}$$

Minima: $$\nabla_\theta J(\theta) = 0 \quad \Rightarrow \quad \theta = (X^T W X)^{-1} X^T W Y$$

Plots:

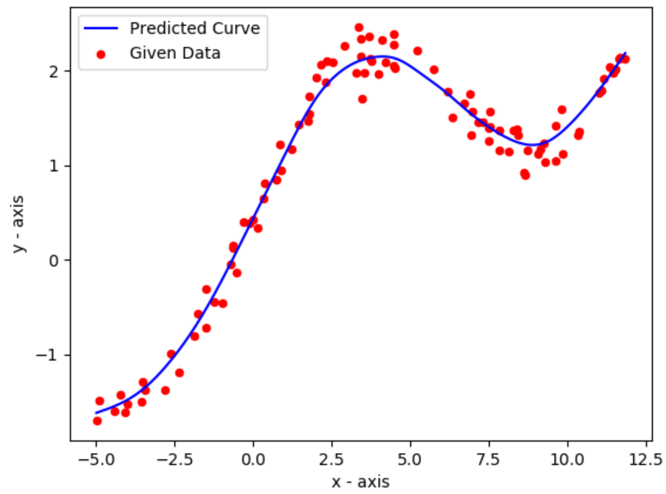


Figure 2: Locally Weighted Linear Regressin Plot ($\tau = 0.8$)

(c) Plots on Varying Bandwidth Parameter(τ)

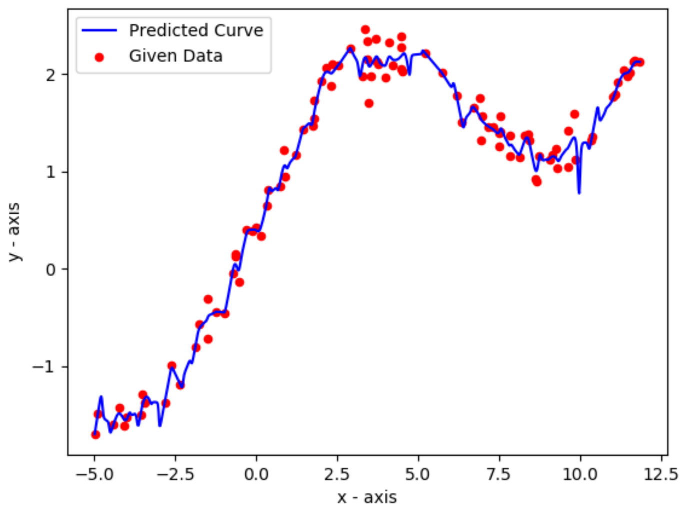


Figure 3: $\tau = 0.1$

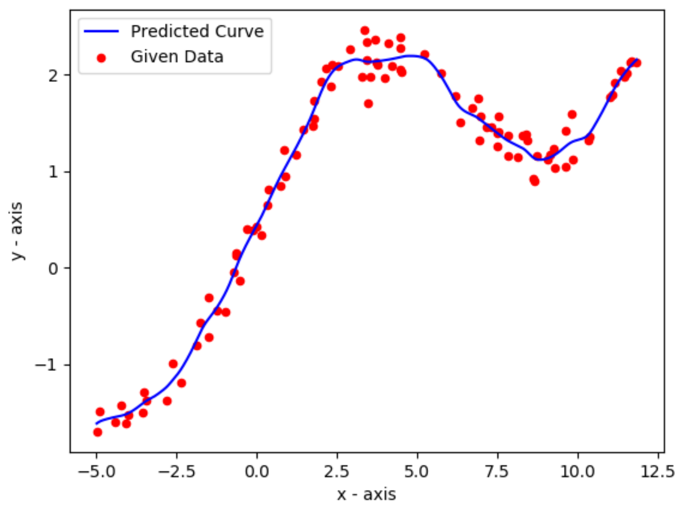


Figure 4: $\tau = 0.3$

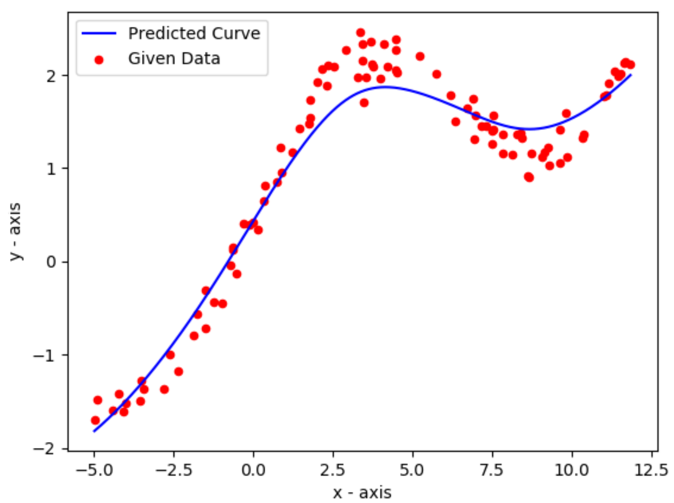


Figure 5: $\tau = 2$

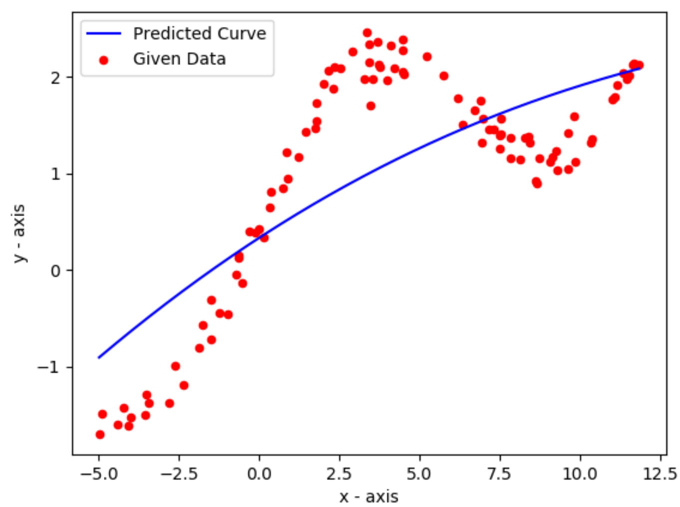


Figure 6: $\tau = 10$

Analysis:

- When Bandwidth Parameter(τ) is:
  - **Too Small:** the model **overfits** the training data.
  - **Too Large:** The model tries to **underfits** the training data.
- $\tau = 0.8$ works the best.

3 Logistic Regression:

Log Likelihood: $$LL(\theta) = \sum_{i=1}^{m} y^{(i)} log(h_\theta(x^{(i)})) + (1 - y^{(i)}) log(1 - h_\theta(x^{(i)}))$$

$$\nabla_\theta LL(\theta) = X^T(Y - g(X\theta)) \qquad \text{(where } g(x) = \frac{1}{1+\exp(-x)}\text{)}$$

Hessian Matrix: $$H = \nabla^2_\theta LL(\theta) = -X^T D X$$

$$\text{(where } D = diag(\, g(x^{(i)T}\theta)(1 - g(x^{(i)T}\theta))\,)\text{)}$$

Newton's Method:

$$\theta^{(t+1)} = \theta^{(t)} - H^{-1} \nabla_\theta LL(\theta)\big|_{\theta_t}$$

Convergence Condition:

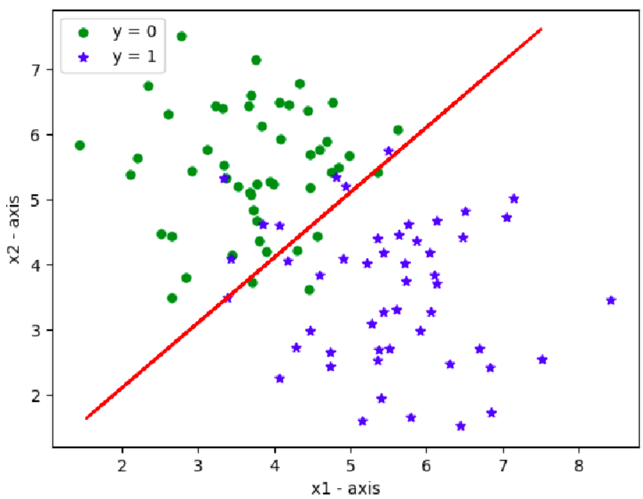
$$\left|\theta_j^{(t+1)} - \theta_j^{(t)}\right| < \epsilon \qquad \text{(for a sufficiently small } \epsilon\text{)}$$

Resulting Parameters:

$$\theta = \begin{bmatrix} \theta_0 \\ \theta_1 \\ \theta_2 \end{bmatrix} = \begin{bmatrix} 0.223295 \\ 1.962616 \\ -1.964861 \end{bmatrix}$$

Decision Boundary is the straight line boundary separating the region where $h_\theta(x) \geq 0.5$ (class $y = 1$) from where $h_\theta(x) \leq 0.5$ (class $y = 0$).

Plot:



4 Gaussian Discrmimant Analysis:

a