COMS4030 Assignment 1 - report

Wesley Earl Stander - 1056114 - Coms Hons

May 21, 2020

1 Question 1

1.1 Question 1.1

Record #	Distance from the record	Label of the record
11	1.752027397046062	0
Prediction		0

1.2 Question 1.2

Record #	Distance from the record	Label of the record
28	1.6488177582740906	1
25	1.7365195075207192	1
27	2.5044160996128424	1
Prediction		1

1.3 Question 1.3

Record #	Distance from the record	Label of the record
36	1.0435516278555652	1
39	1.542335890783846	1
29	2.808006410249093	1
30	3.9457065273534977	1
4	4.538788384580184	0
Prediction		1

K-NN	Training Error Rate	Test Error Rate
1-NN	0.0	0.025
3-NN	0.0	0.025
5-NN	0.05	0.05
7-NN	0.05	0.05
9-NN	0.05	0.05
11-NN	0.05	0.1
13-NN	0.075	0.025

1.4 Question 1.4

2 Question 2

2.1 Question 2.1

$h\theta(x)$ is in the range of 0 to 1 and if β is negative or 0 then the classifier cannot learn from an arbitrary data set. This is because if β is 0 then the classifier will always return the same label and if β is negative then the label will be inverted.

2.2 Question 2.2

$$\begin{aligned}
\frac{\delta}{\delta\theta_k}(h_\theta(x^{(n)})) &= \frac{\delta}{\delta z}(g(z)) \\
&= (-1)(1+e^{-\beta z})^{-2}(-\beta e^{-\beta z}) \\
&= \frac{\beta e^{-\beta z}}{(1+e^{-\beta z})^2} = \frac{e^{-\beta z}}{1+e^{-\beta z}}\beta g(z) \\
&= (1-\frac{1}{1+e^{-\beta z}})\beta g(z) = (1-g(z))\beta g(z) \\
&= \beta(1-h_\theta(x^{(n)}))(h_\theta(x^{(n)}))
\end{aligned} \tag{1}$$

$$\begin{aligned}
\frac{\delta J(\theta)}{\delta\theta_k} &= \frac{\delta}{\delta\theta_k}[\frac{1}{4N}\sum_{n=1}^{N}(h_\theta(x^{(n)})-y^{(n)})^4] \\
&= \frac{1}{N}\sum_{n=1}^{N}(h_\theta(x^{(n)}-y^{(n)})^3\frac{\delta}{\delta\theta_k}(h_\theta(x^{(n)})-y^{(n)}) \\
&= \frac{1}{N}\sum_{n=1}^{N}(h_\theta(x^{(n)})-y^{(n)})^3\beta(1-h_\theta(x^{(n)}))(h_\theta(x^{(n)}))x_k^{(n)} \\
\theta_k &\leftarrow \theta_k - \alpha[\frac{1}{N}\sum_{n=1}^{N}(h_\theta(x^{(n)})-y^{(n)})^3\beta(1-h_\theta(x^{(n)}))(h_\theta(x^{(n)}))x_k^{(n)}]
\end{aligned} \tag{2}$$

2.3 Question 2.3

Testing record #	Classifier Output $h_\theta(x^{(n)})$	Final Output
5	0.19	0
10	0.03	0
15	0.05	0
20	0.12	0
25	0.85	1
30	0.9	1
35	0.86	1
40	0.94	1

3 Question 3

3.1 Question 3.1

$$
\begin{aligned}
g_1(z) &= \frac{1}{1+e^{-(\theta_0 x_0+\theta_1 x_1+\theta_2 x_2+\theta_3 x_3+\theta_4 x_4)}} \\
g_2(z) &= \theta_6 g_1(z) + \theta_5 x_0 \\
g_3(z) &= \theta_7 g_2(z) + \theta_8 x_0 \\
h_\theta(x) &= g_3(z) \\
&= \theta_7 g_2(z) + \theta_8 x_0 \\
&= \theta_7(\theta_6 g_1(z) + \theta_5 x_0) + \theta_8 x_0 \qquad (3) \\
&= \theta_7(\theta_6(\frac{1}{1+e^{-(\theta_0 x_0+\theta_1 x_1+\theta_2 x_2+\theta_3 x_3+\theta_4 x_4)}}) + \theta_5 x_0) + \theta_8 x_0 \qquad (4)
\end{aligned}
$$

3.2 Question 3.2

$$\theta_8 \leftarrow \theta_8 - \alpha[\frac{1}{N}\sum_{n=1}^{N}(h_\theta(x^{(n)}) - y^{(n)})(x_0)] \tag{5}$$

$$\theta_7 \leftarrow \theta_7 - \alpha[\frac{1}{N}\sum_{n=1}^{N}(h_\theta(x^{(n)}) - y^{(n)})(g_2(z))] \tag{6}$$

$$\theta_6 \leftarrow \theta_6 - \alpha[\frac{1}{N}\sum_{n=1}^{N}(h_\theta(x^{(n)}) - y^{(n)})(\theta_7 g_1(z))] \tag{7}$$

$$\theta_5 \leftarrow \theta_5 - \alpha[\frac{1}{N}\sum_{n=1}^{N}(h_\theta(x^{(n)}) - y^{(n)})(\theta_7 x_0)] \tag{8}$$