Assignment 3

Meriam Elabor, 1076589

March 22, 2018

Question 1.1

1-KNN		
Record	Distance from the record	Label of the record
7	1.29007933	-1
Prediction		-1

Question 1.2

3-KNN		
Record	Distance from the record	Label of the record
27	1.36993215	1
22	1.65639809	1
40	2.09803532	1
Prediction		1

Question 1.3

5-KNN		
Record	Distance from the record	Label of the record
30	3.23194266	1
9	5.03028237	-1
37	5.2278505	1
38	5.23353427	1
40	5.61435592	1
Prediction		1

Question 1.4

Testing Error Rates		
1-NN Test Error Rate	3-NN Test Error Rate	5-NN Test Error Rate
0.025	0.05	0.075

Question 2.1

$h_\Theta(x) \in (1, -1)$

Question 2.2

let $J(\Theta) = \frac{1}{N}\sum_{n=1}^{N}(h_\Theta(x^n) - y^n)^2, \quad h_\Theta(x^n) = \frac{2}{1+e^{\Theta^T.x}} - 1, \quad g_\Theta(x^n) = \frac{1}{1+e^{\Theta^T.x}},$
then $\forall k = \{0, ..., 4\}$

$\frac{\partial J(\Theta)}{\partial \theta_k} = \frac{\partial J(\Theta)}{\partial h_\Theta(x^n)} \cdot \frac{\partial h_\Theta(x^n)}{\partial \theta_k}$
$\frac{\partial J(\Theta)}{\partial h_\Theta(x^n)} = \frac{2}{N}\sum_{n=1}^{N}(h_\Theta(x^n) - y^n)$
$\frac{\partial h_\Theta(x^n)}{\partial \theta_k} = \frac{2x_k e^{\Theta^T.x}}{(1+e^{\Theta^T.x})^2} = 2x_k g_\Theta(x^n)(1 - g_\Theta(x^n))$
thus $\frac{\partial J(\Theta)}{\partial \theta_k} = 4x_k \frac{1}{N}\sum_{n=1}^{N}(h_\Theta(x^n) - y^n) g_\Theta(x^n)(1 - g_\Theta(x^n))$

and $\theta_k \to \theta_k - \alpha 4 x_k \frac{1}{N}\sum_{n=1}^{N}(h_\Theta(x^n) - y^n) g_\Theta(x^n)(1 - g_\Theta(x^n)), \forall k = \{0, ..., 4\}$

Question 2.3

Logistic Regression Classifier		
Testing record	Classifier Output $h_\Theta(x^{(n)})$	Final Output
5	-0.949479	-1
10	-0.991887	-1
15	-0.996641	-1
20	-0.992353	-1
25	0.893525	1
30	0.968980	1
35	0.991332	1
40	0.983889	1

Question 3.1

let $f(X) = \frac{1}{1+e^X}, \quad X \in \mathbb{R}$ then
$h_\Theta(x) = g_2 = f(g_1\theta_6 + x_0\theta_5)$, $h_\Theta(x) \in (0,1)$
$g_1 = f(x_0\theta_0 + x_1\theta_1 + x_2\theta_2 + x_3\theta_3 + x_4\theta_4)$

Question 3.2

let $f(X) = \frac{1}{1+e^X}, \quad X \in \mathbb{R}$ and
$J(\Theta) = (h_\Theta(x) - y)^2$ then
$\frac{d}{dX} f(X) = f(X)(1 - f(X))$
and $\quad \forall k = \{5, 6\}$

$\frac{\partial J(\Theta)}{\partial \theta_k} = \frac{\partial J(\Theta)}{\partial h_\Theta(x)} . \frac{\partial h_\Theta(x)}{\partial \theta_k}$
$\frac{\partial J(\Theta)}{\partial h_\Theta(x)} = 2(h_\Theta(x) - y)$
$\frac{\partial h_\Theta(x)}{\partial \theta_k} = 2x_k h_\Theta(x)(1 - h_\Theta(x))$
thus $\frac{\partial J(\Theta)}{\partial \theta_k} = 4x_k(h_\Theta(x) - y)h_\Theta(x)(1 - g_\Theta(x))$

and $\theta_k \to \theta_k - \alpha 4x_k(h_\Theta(x) - y)h_\Theta(x)(1 - h_\Theta(x))$, $\forall k = \{5, 6\}$

-·-
$\forall k = \{0, ..., 4\}$

$\frac{\partial J(\Theta)}{\partial \theta_k} = \frac{\partial J(\Theta)}{\partial h_\Theta(x)} . \frac{\partial h_\Theta(x)}{\partial g1} . \frac{\partial g1}{\partial \theta_k}$
$\frac{\partial J(\Theta)}{\partial h_\Theta(x)} = 2(h_\Theta(x) - y)$
$\frac{\partial h_\Theta(x)}{\partial g1} = 2\theta_6 h_\Theta(x)(1 - h_\Theta(x))$
$\frac{\partial g1}{\partial \theta_k} = x_k g1(1 - g1)$
thus $\frac{\partial J(\Theta)}{\partial \theta_k} = 4x_k\theta_6(h_\Theta(x) - y)h_\Theta(x)(1 - h_\Theta(x))g1(1 - g1)$

and $\theta_k \to \theta_k - \alpha 4x_k\theta_6(h_\Theta(x) - y)h_\Theta(x)(1 - h_\Theta(x))g1(1 - g1)$, $\forall k = \{0, ..., 4\}$