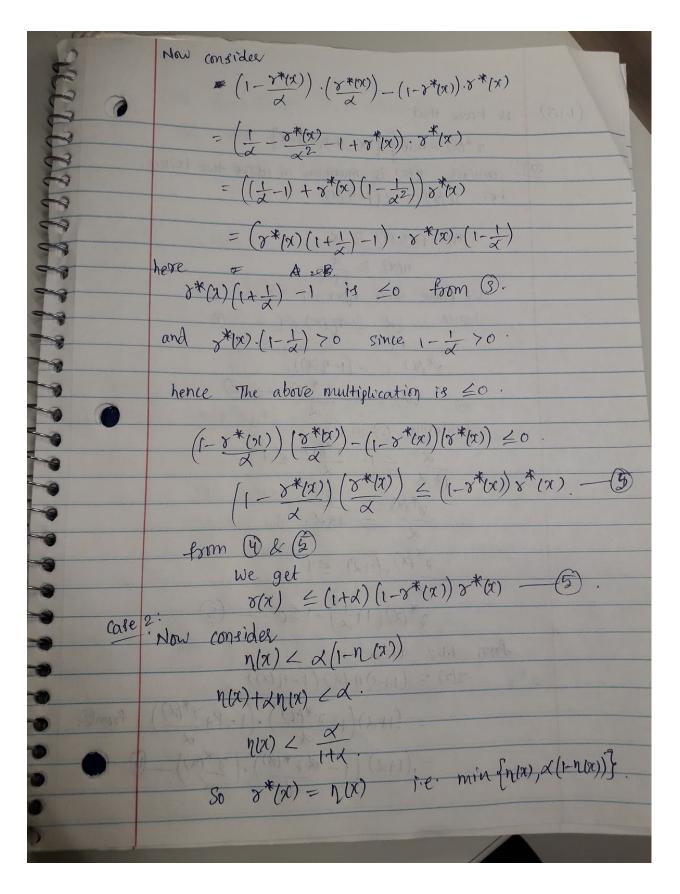
Cover page for answers.pdf CSE512 Fall 2019 - Machine Learning - Homework 3

Name	Vamshi Muthineni
Solar ID	112607946
NetID email address	vmuthineni@cs.stonybrook.edu
Names of people whom you discussed the homework with	AravindReddy Ravula, Rajireddy Annadi, Mallikarjuna, Vamshi krishna, Meghana, Fathima Cherat

1.1.1					
•	it was at a to nothing that philipping of a company				
(1.1.1)	AS We brown that the solings David Low				
	data point or is				
	data point x is $\sum_{i=1}^{M} \hat{n}_i(x) L(i, j)$				
	i li=1				
	L(i,i) is the loss value when a data point is predicted				
	as j', Be instead of actual class i'.				
	It is given for us that				
	0-positive				
	It is given for us that 0 True class.i Assigned 0 0 & 0-positive 1 0 Negative-				
	Loss matrix.				
	$L(0,1) = 1 & L(1,0) = \alpha$				
	when $j=0$ $y^*(x) = \sum_{i=0}^{J} n_i(x) \cdot L(i,j)$				
	$\gamma^*(x) = \sum_{i=0}^{\infty} n_i(x) \cdot L(i,j)$				
	$= \frac{1}{n_0(0)} \cdot L(0,0) + \frac{1}{n_1(x)} \cdot L(1,0)$				
	$\frac{\partial}{\partial x}(x)(0) + (1-1/2) \cdot x$				
53451 0	$= \hat{\eta}_0(x) \cdot (0) + (1 - \eta (x)) \cdot \alpha$ $= (1 - \eta(x)) \alpha \cdot \eta(x) \text{ is prob. that } x \text{ is positive}$				
10	So (1-1/2) is prob that				
	when j=1 1 n.(x). L(i,1) 213-re				
18.08	$=\frac{1}{100}$				
	when $j=1$ $= \frac{(1-\eta(x))}{(1-\eta(x))} \times \frac{(1-\eta(x))}{(1-$				
- ix	= 100.00+0				
	so $\sqrt{x(x)} = \min \{ n(x) + (1-n(x)) \cdot x \}.$				
0	x(0.1x1) 1x19 1x6,110 s				
	COL Motted Water and a second second				

(1.1.2)	Asymptotic disk
	y(x) (an be given by
athlewas-(r	Toronto (Contradició Face (Con A Conc.
See James - ($Y(x) = P(Y=1) \cdot P(Y_{\text{neighbor}} = 0) + \alpha \cdot P(Y=0) \cdot P(Y_{\text{neighbor}} = 1)$
	$= n(\alpha) \cdot \left[1 - n(\alpha)\right] + \alpha \cdot \left(1 - n(\alpha)\right) \cdot n(\alpha)$
	hence (2) = 10(21) - 10(21) + 12 + 12 (21)
	$\sigma(x) = n(x) \cdot \left[1 - n(x) + \lambda - \alpha \cdot n(x) \right]$
	$= n(x) \cdot \left[(1-n(x)) \cdot (1+d) \right]$
	$\gamma(x) = \eta(x) \cdot (1-\eta(x)) \cdot (1+x)$
	(O) 2011 - 10 - 10 - 10 - 10 - 10 - 10 - 10
(1 =	the disport of the state of the
	knot atab at a qual a nava
(4,14)0 -1	to be be an attendary of the Colon of Charles
(4 MP 4 =	() - when the part of the particular of the same same of the particular of the same same same of the same same same of the same same same same same same of the same same same same same same same sam
F31 34 50	
A TOTAL TO	ar-1)+(a(1)6-1)-(x)1, -(1)x
	0

1.1.3
Ballone of the state of the sta
(1.1.3) We know that
$\gamma^*(x) = \min \left\{ \eta(x), (\mu(x)) \cdot x \right\}$
case: 1 meilos pro
case: consider n(x) is maximum in above two terms.
$(A) \subset X (1-(X))$
$n(x) + \alpha n(x) \ge \alpha$. $n(x) \ge \alpha$ We know that $n(x) \le 1$ hence $\alpha \le n(x) \le 1$.
$M(n) + dn(n) \geq \alpha$.
$n(x) \geq \frac{\alpha}{1+\alpha}$ We know that $n(x) \leq 1$ hence
We know that Ita.
hance
$\frac{1}{2}$
hence $\angle 1 + 2 + 1 = 0$
$\delta^{\pi}(x) = \lambda(1-\eta(x)).$
$n(x)$, $\gamma^*(x)$ — (2)
(V) = (
$N(x) = (-\frac{y^*(x)}{x}) - (2)$ Substituting (2) in (1)
x 21-8*(x) 21
1+d 2
A CONTRACTOR OF
$\frac{8^{*}(x)}{\alpha} \leq \frac{1}{1+\alpha}.$
d
$y^*(x) \cdot (1+x) \leq 1$
マ*(ス)(1+よ)-1=0・一多・
8*(x) (1+ x) -1 =0 · - 3
from 1.1.2 we've
$\gamma(x) = (1+\alpha) \gamma(x) (1-\gamma(x))$
(~ */\sigma) (r-V+8*(\sigma)) from (\sigma)
= (1+d)(1-2++)
$= (1+\alpha) \left(1 - \alpha \gamma^*(\alpha)\right) \cdot \left(3^*(\alpha)\right) - \alpha$
$from 1.1.2 we've$ $r(x) = (1+d)r(x)(1-r(x))$ $= (1+d)\left(1-\frac{x^*(x)}{a}\right)\cdot\left(x-x+\frac{x^*(x)}{a}\right) from \mathfrak{D},$ $= (1+d)\left(1-\frac{x^*(x)}{a}\right)\cdot\left(\frac{x^*(x)}{a}\right)$



substituting in 1.1.2. we get $\gamma(x) = (+x) \eta(x) (1 - \eta(x))$ $= (1+\lambda) \gamma^*(x) (1-\gamma^*(x)) - 6.$ from (5) & 6 it is clear that $\gamma(x) \leq (1+\lambda)\gamma^*(x)(1-\gamma^*(x))$ 1 XXXXXXXXX

(1.1.4)	I NN charifier
	Let R be the asymptotic risk of the 1-NN classified
	E & be the bayes oisk.
	Since $R(x) = E(x(x))$
TO CALL	$48001 \cdot 1.1.3$ $8(x) \leq (1+2) 8^{*}(x)(1-8^{*}(x))$
	$E(\gamma(x)) \leq E\left[(1+\lambda)\gamma^*(x)(1-\gamma^*(x))\right]$
	E (1+2) E [8*(x) - 8*(x)2]
	< (1+x) [= (x*(x)) - E(x*(x)2)]
	$\leq (1+x) \left[\left\{ \left\{ \left\{ \left\{ \left\{ \right\}^{*}(x) \right\} - \left[\left\{ \left\{ \left\{ \left\{ \right\}^{*}(x) \right\} \right\} \right]^{2} \right] \right\} \right] \right] \right]$
	Since $E(x^2) = Vad(x) + [E(x)]^2$
	$\leq (1+\alpha) \left[E(s^*(\alpha)(1-E(s^*(\alpha))) - Val(s^*(\alpha)) \right]$
	€ (1+x) [2*00 (1-2*) - Var (8*(x))]
	$\leq (1+2) R^*(1-R^*) - (1+2) var(r^*(x))$
	hence. $R \leq (1+d) R^{*}(1-R^{*})$ since $1+d \approx 0$ val($5*(21)$) ≥ 0
	2

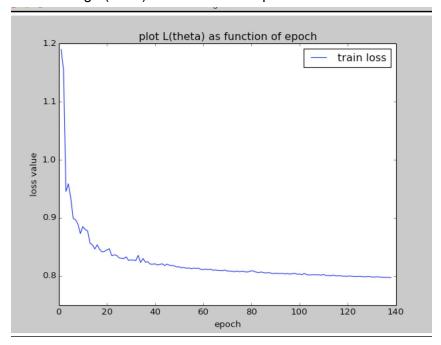
(1.2.1)	As we know that Asymptotic risk in k-NN classifies $ \gamma(x) = P(y=1) \cdot P(y_{\text{majority-neighbors}}^{(y)} + P(y=0) \cdot P(y_{\text{majority-neighbors}}^{(y)} = 1) $
	given = from in the data that $P(Y=1) = n(x), P(Y_{\text{majority-neighbors}} = 1) = g(n,k)$ $P(Y=0) = 1 - n(x), P(Y_{\text{majority-neighbors}} = 0) = 1 - g(n,k)$ Substituting in the above eqn. $S(x) = n(x) \cdot (1 - g(n,k)) + (1 - n(x)) \cdot g(n,k)$

	mate as not soon as the contract of the soon of the so
(1.2.2)	As we know
	$8*(x) = \min\{n(x), 1-n(x)\}$
	from the above derivation we've
	$g(x) = \eta(x) \cdot (1 - g(n, k)) + (1 - \eta(x)) \cdot g(n, k) - 0$
	So when $n(x) \leq \frac{1}{2}$
	$g^*(x) = \min \left\{ n(x), 1 - n(x) \right\}$
	n(x) + n(x)
-	i.e. $\eta(x) = \gamma^*(x)$
	Substituting in (1)
	$\eta(x) = \delta^*(x) (1-g(n,k)) + (1-\delta^*(x))g(n,k)$
	= (2/1) - o(2) - g(1) +) =
	$= \gamma^{*}(x) - \gamma^{*}(x) - g(n, k) + (1 - \beta^{*}(x)) \cdot g(n, k)$
	$8(\pi) = 8^*(\pi) + (1-28^*(\pi)) \cdot 9(8, 1) - 2$
	when $p^*(x)$ $n(n) \ge \frac{1}{2}$
	$\gamma^*(\alpha) = (-\eta(\alpha) \Rightarrow \eta(\alpha) = (-\gamma^*(\alpha)).$
	substituting in ()
	$s(x) = (1 - x^{*}(x)) \cdot (1 - g(n, k)) + (x - x + x^{*}(x)) \cdot g(n, k)$
	$= (1-r^*(x) + [r^*(x)-1+r^*(x)] \cdot g(n, k)$
	=1-8*(x)+[2x*(x)-1].9(1-8*(x),K)
	$= 1 - 8^{*}(x) + (28^{*}(x) - 1) \cdot [1 - 9(8^{*}(x), k)]$
	* · · · · · · · · · · · · · · · · · · ·
	$\sigma(x) = \gamma^{+}(x) + (1-2\gamma^{+}(x)) \cdot g(\gamma^{+}(x), \kappa) - g(\gamma^{+}(x), \kappa)$

1.2.0
So from @, B, it is clear for us that
, , , , , , , , , , , , , , , , , , , ,
in any case
So from (a), (b), it is clear for us that in any case
A Participal and att
and the language of the same of the same to save
(1.2.3) Hoeffding's inequality pro is
6 - (5)4, -1,000
$P(H(n)) \geq (P+E)n = exp(2e-2E^2n)$.
H(n) -> no. of points that are positive
P > probability that a point is positive
n = total points.
A LOUIS AND A LOUIS AND A CONTRACT OF THE PARTY OF THE PA
$9/3\sqrt[4]x), k) = 2/(H(n) = 2)$ When n is ever
$g(x^*x), k) = \begin{cases} P(H(n) \ge \frac{k}{2}) & \text{when n is even} \\ P(H(n) \ge \frac{k}{2}) & \text{when n is odd} \end{cases}$
(20) (10) 6-1) 1-(1, 19) (1) 2-(1) 1-(1) (20)
SO P(HZ K) = P(H E(8*(X)+0.5-8*(X))K)
observing with Hoeffdings inequality
it's clear that $p = x^*(x)$
$\varepsilon = 0.5 - \gamma^*(x)$
E = 0.5 - 0 (1)
n = K. milatiladus
Control of the Man of Control of the Man of
So we'll get
$9(8*(x), k) = p(H \ge \frac{k}{2}) \le exp(-2(0.5 - 8*(x))k)$
Make in the state of the state
1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
(1-08/12)+(x)-0-1 =
(1) (N, CN*) P. (CN* XX -1) + (M* x = (M) C

1.2.4	
er.	
6	
	(1.2.4)
	We need to prove
2	We need to prove $7 \times 1 \times 1 + \sqrt{2} \times 1 \times 1 = 1$
7	
T	from 1.2.2, we know that
	$\delta(x) = \gamma^*(x) + (1-2\gamma^*(x)) g(\gamma^*(x), K)$
	rewriting the exuation as follows
	$8(x) = r^*(x) + \frac{1}{\sqrt{2}k} \left[\sqrt{2k} (1-2r^*(x)) g(r^*(x), k) \right] - 9$
70	0(x) = 8 (x) + 2x (1-20 07) J(0 0.7)]
7	from 1.2.3
	$g(x^*(x), k) \leq \exp(-2(0.5 - x^*(x))^2 k)$
	substitute & in G.
-	120 1 *120 1 [-1. 00*44) mal also attack!
	$-8(x) \leq x^*(x) + \frac{1}{\sqrt{2}k} \left[\sqrt{2k} \left(1 - 2x^*(x) \right) \cdot \exp\left(-2(0.5 - x^*(x)) k \right) \right] - \frac{1}{\sqrt{2}k} \left[\sqrt{2k} \left(1 - 2x^*(x) \right) \cdot \exp\left(-2(0.5 - x^*(x)) k \right) \right]$
	217
	$= 8 \times (x) + \frac{1}{\sqrt{2} \times (2 \times (0.5 - 8 \times 1x))} \cdot \exp(2(0.5 - 8 \times 1x)^2 \times)$
9	Jek Line
9	# Let z = 2/2× (0.5- x*(x)
	H Q Z = 2 2 x (0.3 - 0 (1)
	the the above ean-will be of form
	$\leq x^{*}(x) + \frac{1}{\sqrt{2}k} \left[2z \exp(-z^{2}) \right]$
•	= 0 (1) V2K []
10	The values for 22 exp(-2)) lies between [-0.858,0.858]
9	which is < 1.
	which is 51.
	1 2 (2) (2*/2)+ -
	hence $\gamma(x) \leq \gamma^{*}(x) + \frac{1}{\sqrt{2K}}$
	· · · · · · · · · · · · · · · · · · ·

- a. Number of epochs before exiting = 138
- b. Plot showing L(theta) as a function of epoch

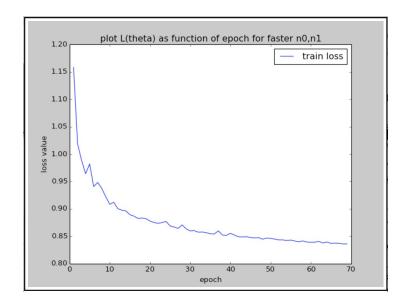


c. Final value of L(theta) after the optimization = 0.8

2.3.2.

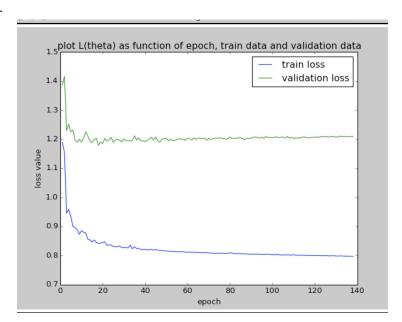
a. The pair of n0,n1 that leads to faster convergence is 0.1, 5.
 The no of epochs taken is 70.
 Final value of L(theta) = 0.85

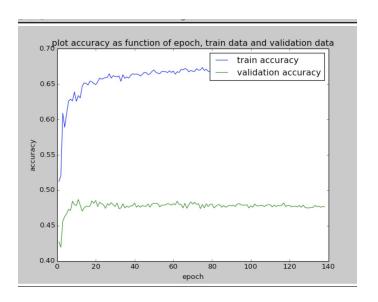
b.



2.3.3

a.





2.3.4. Confusion matrix

On Train Data

Predicted Values

		1	2	3	4
	1	536	185	28	34
Actual	2	139	911	114	105
Values	3	40	231	353	192
	4	29	113	93	897

On Validation Data

		Predi	Predicted Values				
		1 2 3 4					
	1	169	94	26	27		
Actual	2	133	359	77	91		
Values	3	39	197	77	125		
	4	25	116	90	355		

2.4

1. Best accuracy obtained on test data on kaggle - 46%