Student ID________________________ Student Name________________________

Machine Learning

Logistic regression and SVM

1. Given historical data as below, if current weights **W** for logistic regression model (in which $P(y=1|\mathbf{X})=\sigma(\mathbf{W}^T\mathbf{X})$, $\mathbf{X}=[x1,x2]^T$) is $[0,0]^T$, update **W** one step according to the gradient ascend. The learning rate η=0.1

y	x1	x2
1	2	1
0	1	2
0	3	3

$$\mathbf{w}^{\text{new}} = \mathbf{w}^{\text{old}} + \eta \sum_{t} \left\{ y_n - \sigma\left(\left(\mathbf{w}^{\text{old}} \right)^{\top} \mathbf{x}_n \right) \right\} \mathbf{x}_n.$$

$=[0\ 0]^T+0.1*\{ (1- \sigma([0\ 0]*[2\ 1]^T))*[2\ 1]^T + (0- \sigma([0\ 0]*[1\ 2]^T))* [1\ 2]^T + (0- \sigma([0\ 0]*[3\ 3]^T))* [3\ 3]^T \}$

$=[0\ 0]^T+0.1*\{ (1- \sigma(0))*[2\ 1]^T + (0- \sigma(0))*[1\ 2]^T + (0- \sigma(0))*[3\ 3]^T \}$

$=[0\ 0]^T+0.1*\{ 0.5*[2\ 1]^T - 0.5*[1\ 2]^T - 0.5*[3\ 3]^T \}$

$=[-0.1, -0.2]^T$

2. Given a dataset as below, what is the solution of SVM model, where y=sign($\mathbf{W}^T\mathbf{X}$+b), $\mathbf{X}$=[x1,x2]^T
That is, solve the **W** and b. (Hint, solve the dual problem)

y	x1	x2
1	2	1
-1	1	2
-1	3	3

Dual problem

$$\max_{\boldsymbol{\alpha}} \quad \mathcal{G}(\boldsymbol{\alpha}) = -\frac{1}{2}\sum_{i=1}^{N}\sum_{j=1}^{N}\alpha_i\alpha_j y_i y_j \boldsymbol{x}_i^{\top}\boldsymbol{x}_j + \sum_{i=1}^{N}\alpha_i,$$

$$\text{subject to} \quad \sum_{i=1}^{N}\alpha_i y_i = 0,$$

$$\alpha_i \geq 0, \quad i = 1, \ldots, N.$$

Max G(a1, a2, a3)
= -0.5*{(a1*a1*1*1*[2 1]*[2 1]^T) + (a1*a2*1*(-1)*[2 1]*[1 2]^T)+ (a1*a3*1*(-1)*[2 1]*[3 3]^T)
+(a2*a1*(-1)*1*[1 2]*[2 1]^T) + (a2*a2*(-1)*(-1)*[1 2]*[1 2]^T)+ (a2*a3*(-1)*(-1)*[1 2]*[3 3]^T)
+(a3*a1*(-1)*1*[3 3]*[2 1]^T) + (a3*a2*(-1)*(-1)*[3 3]*[1 2]^T)+ (a3*a3*(-1)*(-1)*[3 3]*[3 3]^T)}
+a1+a2+a3
Subject to a1*1+a2*(-1)+a3*(-1)=0, a1>=0, a2>=0. a3>=0

Max G(a1, a2, a3)
= -0.5*{5*a1*a1 -4a1*a2 -9*a1*a3
-4*a2*a1 + 5*a2*a2+ 9*a2*a3*
-9*a3*a1 + 9*a3*a2*+ 18*a3*a3}
+a1+a2+a3
Subject to a1=a2*+a3*, a1>=0, a2>=0. a3>=0

Replacing a1=a2+a3 into G
Max G(a2, a3)
= -0.5*{5*(a2+a3)*(a2+a3) -4(a2+a3)*a2 -9*(a2+a3)*a3
-4*a2*(a2+a3) + 5*a2*a2+ 9*a2*a3*
-9*a3*(a2+a3) + 9*a3*a2*+ 18*a3*a3}
+a2+a3+a2+a3
Subject to a2=>0. a3>=0
That is
Max G(a2,a3)=-a2*a2-2.5a3*a3-a2*a3+2*a2+2*a3 (this is one degree 2 polynomial)
Subject to a2>=0 a3>=0

dG/da2=0 =>-2*a2-a3+2=0 dG/da3=0 => -5*a3-a2+2=0
⇨ a2=8/9 a3=2/9 => a1=a2+a3=10/9
⇨ W=a1*(1)*[2 1]^T + a2*(-1)*[1 2]^T + a3*(-1)*[3 3]^T =[2/3 -4/3]^T
a1>0, so X=[2 1]^T is the support vector, for this X, we have y($\mathbf{W}^T\mathbf{X}$+b)=1 or 1([2/3 -4/3]*[2 1]^T +b) =1
=>b=1