CS 482/682 – Machine Learning

TEST 2

Chapter 4- Data Preprocessing., Ch5-Non Linear Models

Your Name: KEY

1. (3 points) Given the following dataset, provide the revised data table by using one hot encoding for the categorical data. For the numeric column Test-Ch1 use standard scalar i.e make the mean 0 and standard deviation 1. For Test-Ch8 use Min-Max scaling.

Test-Ch1	Test-Ch8	Sex
2	8	Male
8	9	Female
7	8	Female
3	5	Male
10	3	Female

SOLUTION

Mean ([2 8 7 3 10]) = 30/5 = 6
Standard deviation of (2 8 7 3 10]) = sqrt($((-4)^2 + 2^2 + 1^2 + (-3)^2 + 4^2)/5$) = sqrt($((-4)^2 + 2^2 + 1^2 + (-3)^2 + 4^2)/5$) = sqrt((16+4+1+9+16)/5) = sqrt(46/5) = sqrt(9.2) = 3.03

Min ([8 9 8 5 3]) = 3
Max ([8 9 8 5 3]) = 9
Range ([8 9 8 5 3]) = 9-3 = 6

(TestCh1-mean)std_dev	(TestCh8-Min)/Range	Sex_male	Sex_Femal	
-4/3.03 = -1.32	5/6=.83	1	0	
2/3.03 = .66	6/6=1	0	1	
1/3.03 = .33	5/6=.83	0	1	
-3/3.03=.99	2/6=.33	1	0	
4/3.03=1.32	0/6=0	0	1	

2. (2 points) An SVM model was developed for a binary classification problem in which there are three predictors. The two classes are Class 1 (target value is 1) and Class 2 (target value is -1) The values of margin points (support vectors, SV1, SV2 and SV3), the values of weights α_i for the margin points are given along with bias term

   $\beta_0 = -2$

Support Vector Number	Support Vector Value	Target Value	α_i
SV1	(2,2,1)	-1	1.00
SV2	(-1,-2,2)	1	.3
SV3	(-2,-1,3	1	.6

Classify the unknown data `u = (-2, -4, -1 )` as either Class 1 or Class 2.

$D(u) = -2 + [(2,2,1) *(-2,-4,-1)^T*(-1)*1 + (-1,-2,2)*(-2,-4,-1)^T*1*(.3) + (-2,-1,3)*(-2,-4,-1)^T (1)(.6) =$

$= -2 + (-4-8-1)*(-1) + (2+8-2)*(.3) + (4+4-3)*(.6) = -2 + 13 +2.4 + 3.0 =$

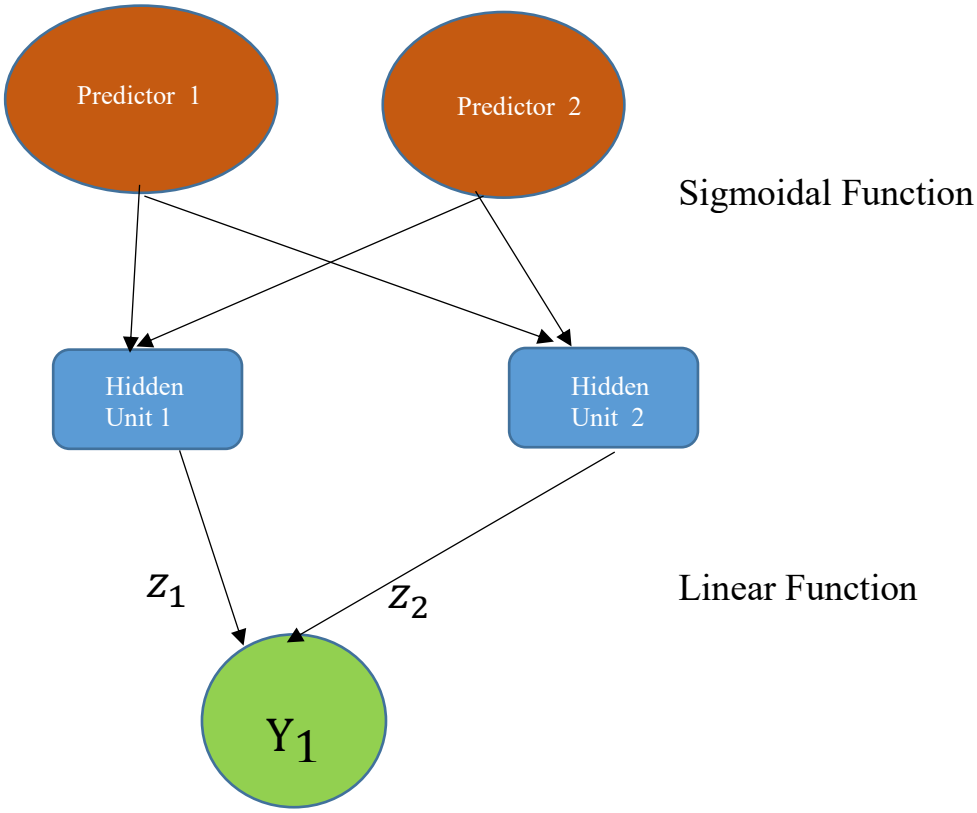
$11+2.4+3.0 = 11+5.4 = 16.4 > 0$

Therefore the unknown point belongs to Class 1

3. (4 +1 = 5 points) A neural network with 2 hidden units is shown in the figure below for a regression problem, Assume the following values for hidden units 1 and 2 and the outcome unit Y_1.

$$\alpha_1 = (\alpha_{01}, \alpha_{11}, \alpha_{21}) = (1,1,0) \quad \alpha_2 = (\alpha_{02}, \alpha_{12}, \alpha_{22}) = (1,0,1)$$
$$\beta_1 = (\beta_{01}, \beta_{11}, \beta_{21}) = (1,1,1)$$

a) Find the prediction Y_1 made by the outcome unit for the input data with predictor values [2 3].
b) If the correct outcome is 5, calculate the absolute value of the error in the prediction.



a) $z_1 = \sigma(u_1)$ where $u_1 = \alpha_{01} + x_{11} * \alpha_{11} + x_{12}\alpha_{21} = 1 + 1*2 + 0*3 = 3$
$z_2 = \sigma(u_2)$ where $u_2 = \alpha_{02} + x_{11} * \alpha_{12} + x_{12}\alpha_{22} = 1 + 0*2 + 1*3 = 4$

$\sigma(4) = 1/(1+e^4) = 1/(1+54.59) = 1/55.59 = 0.0183$

$Y_1 = \beta_{01} + z_1 * \beta_{11} + z_2 \beta_{21} = 1 + 1 * .0474 + 1 * .0183 = 1.0657$

b) Absolute value of error: $|5 - 1.0657| = 3.93$

4. (5 points) (CS-682 Students ONLY) : Given the learning rate is .4 use the equations below to find the update of α's and β's in the problem above after finding the predicted value for [2 3] and the correct outcome being 5. Note that x' stands for transpose of x in the equations below.

$$\frac{\partial R_i}{\partial \beta_{km}} = -2(y_{ik} - f_k(x_i))g_k'(\beta_k^T z_i)z_{mi},$$

$$\frac{\partial R_i}{\partial \alpha_{m\ell}} = -\sum_{k=1}^{K} 2(y_{ik} - f_k(x_i))g_k'(\beta_k^T z_i)\beta_{km}\sigma'(\alpha_m^T x_i)x_{i\ell}.$$

$$\beta_{km}^{(r+1)} = \beta_{km}^{(r)} - \gamma_r \sum_{i=1}^{N} \frac{\partial R_i}{\partial \beta_{km}^{(r)}},$$

$$\alpha_{m\ell}^{(r+1)} = \alpha_{m\ell}^{(r)} - \gamma_r \sum_{i=1}^{N} \frac{\partial R_i}{\partial \alpha_{m\ell}^{(r)}},$$

$$g_k(T) = \frac{e^{T_k}}{\sum_{\ell=1}^{K} e^{T_\ell}}.$$

$$\sigma(u) = \frac{1}{1 + e^u}$$