ELE 364: Assignment #4 Solutions

1. (10 pts)

   (a) The prediction is as follows:

ID	**d**[1]	**d**[2]	**d**[3]	Target	Predict
1	1	2	3	27	14
2	2	3	4	39	20
3	3	4	5	51	26
4	4	5	6	63	32

   As a result: $L_2 = 1/2 \times (169 + 361 + 625 + 961) = 1058$.

   (b) The prediction is as follows (ED = ErrorDelta):

ID	**d**[1]	**d**[2]	**d**[3]	Target	Predict	Error	ED **w**[0]	ED **w**[1]	ED **w**[2]	ED **w**[3]
1	1	2	3	27	14	13	13	13	26	39
2	2	3	4	39	20	19	19	38	57	76
3	3	4	5	51	26	25	25	75	100	125
4	4	5	6	64	32	31	31	124	155	186
						Sum	88	250	338	426

   Thus, the new values after one iteration are:

   - $\mathbf{w}[0] = 0.0088$.
   - $\mathbf{w}[1] = 1.0250$.
   - $\mathbf{w}[2] = 2.0338$.
   - $\mathbf{w}[3] = 3.0426$.

2. (10 pts) The first step is to handle the missing values. The mean for the Age feature is 32.7 and the mode for the socio-economic band feature is a. Hence, the question marks in the query instances are changed to reflect these values. Each continuous feature should then be range-normalized to $[-1, 1]$. These normalized values are shown in the following table.

ID	Age	Soc.-econ. band	Shop value	Shop freq.
1	−0.347	b	0.769	−0.565
2	0.956	a	−0.065	−0.750

   With this information, we can calculate the prediction by plugging these values into the regression equation as follows.

   - Logistic$(0.668 + (-0.580 \times -0.347) + (-0.198 \times 1) + 3.409 \times 0.769 + (2.050 \times -0.565))$ $= 0.894$. Hence, yes.
   - Logistic$(0.668 + (-0.580 \times 0.956) + (-0.160 \times 1) + (3.409 \times -0.065) + (2.050 \times -0.750))$ $= 0.141$ Hence, no.

3. (10 pts) The output for the SVM model is given by

$$(-1 \times 1.6811 \times (\langle -0.4549, 0.0095, 0.2203\rangle \cdot \langle -0.8945, -0.3459, 0.5520\rangle) - 0.0216)$$
$$+(-1 \times 0.2384 \times (\langle -0.2843, -0.5253, 0.3668\rangle \cdot \langle -0.8945, -0.3459, 0.5520\rangle) - 0.0216)$$
$$+(-1 \times 0.2055 \times (\langle 0.3729, 0.0904, -1.0836\rangle \cdot \langle -0.8945, -0.3459, 0.5520\rangle) - 0.0216)$$
$$+(-1 \times 1.7139 \times (\langle 0.5580, 0.2217, 0.2115\rangle \cdot \langle -0.8945, -0.3459, 0.5520\rangle) - 0.0216)$$
$$= -0.9045 - 0.1737 - 0.2195 - 0.8083 = -2.1061$$

Because the output is less than -1, the model predicts the negative level: low risk.

4. (10 pts)

   (a) $\phi(v_1) = \langle 0.25,\ 0.64,\ 0.566,\ 0.707,\ 1.131,\ 1\rangle$ and $\phi(v_2) = \langle 1,\ 0.49,\ 0.99,\ 1.414,\ 0.99,\ 1\rangle$.

   (b) $\phi(v_1) \cdot \phi(v_2) = 0.25 \times 1 + 0.64 \times 0.49 + 0.566 \times 0.99 + 0.707 \times 1.414 + 1.131 \times 0.99 + 1 \times 1 = 4.244$.

   (c) $kernel(v_1, v_2) = (0.5 \times 1 + 0.8 \times 0.7 + 1)^2 = 4.244$. This matches the result in (b).

   (d) $kernel(v_1, v_2) = 4.244$, $kernel(v_0, v_2) = (0.9 \times 1 + 0.7 \times 0.6 + 1)^2 = 5.382$.
   $(-1 \times 0.71 \times kernel(v_0, v_2) + 0.21) + (+1 \times 0.98 \times kernel(v_1, v_2) + 0.21)$
   $= (-1 \times 0.71 \times 5.382 + 0.21) + (+1 \times 0.98 \times 4.244 + 0.21) = -3.611 + 4.369 = 0.758$
   Even though this value is less than 1 and, hence, within the margin, since it is greater than 0, the class of v_2 is predicted as 1.