HW4

Started: Aug 15 at 7:05pm

Quiz Instructions

Question 1   1 pts

We consider a 3-state (a, b, c) Markov Chain. We observe the following sequence of data: b a a b a b c c b c a b c c a b c c a c b c a b c c b b b a b c. Please use MLE and Laplace Smoothing with k = 2 to learn the following transition probabilities: P(a|a), P(b|a), P(c|a), P(a|b), P(b|b), P(c|b), P(a|c), P(b|c), P(c|c).

Question 2   2 pts

The following is a tabular dataset on the outdoor conditions for playing an outdoor game.

outlook	temperature	humidity	windy	play
overcast	cool	normal	true	Yes

overcast	mild	high	true	Yes
overcast	hot	high	false	Yes
overcast	hot	normal	false	Yes
rainy	mild	high	false	Yes
rainy	cool	normal	false	Yes
rainy	mild	normal	false	Yes
sunny	cool	normal	false	Yes
sunny	mild	normal	true	Yes
rainy	cool	normal	true	No
rainy	mild	high	true	No
sunny	hot	high	true	No
sunny	mild	high	false	No
sunny	hot	high	false	No
sunny	cool	high	false	?

Please use the Naive Bayesian Network to classify the play for the last row. We also assume that Laplace-smoothing with k = 2 is used.

Question 3   2 pts

We apply decision tree algorithm to the above question to classify the play for the last row. Please show the decision

tree as well.

Question 4

1 pts

We consider the following dataset on house price and house size:

X: price (1000 $)	Y: size (1000 square feet)
145	1.25
65	1.1
60	0.75
75	1.05
110	1.15
165	1.75

Please use the linear regression formula derived in class to figure out the linear equation.

Question 5
2 pts

We consider the following dataset on the T-shirt size vs height and weight:

Height (cm)	Weight (kg)	T shirt size
158	58	M
158	59	M
158	63	M
160	59	M
160	60	M
163	60	M
163	61	M
160	64	L
163	64	L
165	61	L
165	62	L
165	65	L
168	62	L
168	63	L
168	66	L
170	63	L
170	64	L
170	68	L

We use KNN with K = 3 to predict a T-shirt size for the following persons. Please use Manhattan distance.

1. A person of height 161cm and of weight 61kg
2. A person of height 164cm and of weight 59kg

Question 6   **2 pts**

We have the following eight 2D data points:

X	Y
24.5	3.5
26.5	7.0
21.5	7.5
25.0	15.5
2.5	20.5
9.5	23.0
11.0	20.0
21.5	4.0

Two initial centroids (12.0, 12.5), (15.0, 15.5). Please find the next two centroids after one iteration using K-means with K = 2 and Euclidean distance.

Question 7   2 pts

We have the following six 1D data points: 1.6, 2.1, 1.8, 5.3, 6.7, 4.7. We consider EM with two components with the following initial setup: $\pi_1 = 0.4$, $\mu_1 = 1$, $\sigma_1^2 = 2$, $\pi_2 = 0.6$, $\mu_2 = 2$, $\sigma_2^2 = 2$. Please find the updated values of these 6 variables after one iteration using EM.

Question 8   **2 pts**

We consider function $f(W,X) = \sigma(w_0x_0 + w_1x_1 + b)$, where $X = (x_0, x_1)$ is the input, $W = (w_0, w_1)$ is weight vector, b is the bias, and $\sigma(x)$ is the sigmoid function. Use Backpropagation to draw a work flow as shown on the lecture slides to obtain the partial derivatives of $\partial f/\partial w_0$, $\partial f/\partial w_1$ and $\partial f/\partial b$ at $W = (1, -2)$ and $b = -3$ as $X = (-5, -3)$.

Not saved   Submit Quiz