Homework2

Course: Machine Learning

Name: Xiang Lei (雷翔)

Student ID: 2053932

Date: June 2024

Q1. Calculate the gradient of the following multivariate function:

(1) $u = xy + y^2 + 5$

(2) $u = \ln \sqrt{(x^2 + y^2 + z^2)},\ at\ (1, 2, -2)$

Ans.

$$\frac{\partial u}{\partial x} = y$$

$$\frac{\partial u}{\partial y} = x + 2y$$

Therefore, the gradient ∇u is:

$$\nabla u = (y, x + 2y)$$

(2)

$$u = \ln \sqrt{x^2 + y^2 + z^2} = \frac{1}{2} \ln(x^2 + y^2 + z^2)$$

$$\frac{\partial u}{\partial x} = \frac{\partial}{\partial x} \left(\frac{1}{2} \ln(x^2 + y^2 + z^2) \right) = \frac{1}{2} \cdot \frac{2x}{x^2 + y^2 + z^2} = \frac{x}{x^2 + y^2 + z^2}$$

$$\frac{\partial u}{\partial y} = \frac{\partial}{\partial y} \left(\frac{1}{2} \ln(x^2 + y^2 + z^2) \right) = \frac{1}{2} \cdot \frac{2y}{x^2 + y^2 + z^2} = \frac{y}{x^2 + y^2 + z^2}$$

$$\frac{\partial u}{\partial z} = \frac{\partial}{\partial z} \left(\frac{1}{2} \ln(x^2 + y^2 + z^2) \right) = \frac{1}{2} \cdot \frac{2z}{x^2 + y^2 + z^2} = \frac{z}{x^2 + y^2 + z^2}$$

Therefore, the gradient ∇u is:

$$\nabla u = \left(\frac{x}{x^2 + y^2 + z^2}, \frac{y}{x^2 + y^2 + z^2}, \frac{z}{x^2 + y^2 + z^2} \right)$$

At the point $(1, 2, -2)$:

$$x^2 + y^2 + z^2 = 1^2 + 2^2 + (-2)^2 = 1 + 4 + 4 = 9$$

So, the gradient at the point $(1, 2, -2)$ is:

$$\nabla u = \left(\frac{1}{9}, \frac{2}{9}, -\frac{2}{9} \right)$$

Q2. As we all know, whether to sleep in is a complex question that depends on multiple variables. The following is a random selection of student A's 12-day data on sleeping in. Please build a decision tree based on this data, and use the information gain to divide the attributes. An illustration of the calculation process and the final decision tree is required. Hint: For some nodes, you may not need to calculate conditional entropy, but directly make decision by observing the data.

Ans.

Season	After 8:00	Wind	Sleep in
spring	no	breeze	yes
winter	no	no wind	yes
autumn	yes	breeze	yes
winter	no	no wind	yes
summer	no	breeze	yes
winter	yes	breeze	yes
winter	no	gale	yes
winter	no	no wind	yes
spring	yes	no wind	no
summer	yes	gale	no
summer	no	gale	no
autumn	yes	breeze	yes

$$Ent(D) = \sum_c -P(D=c)\log_2 P(D=c) = -\frac{9}{12}\times\log_2\frac{9}{12} - \frac{3}{12}\times\log_2\frac{3}{12} = 0.8113$$

For Season:

spring: 1 yes and 1 no

summer: 1 yes and 2 no

autumn: 2 yes and 0 no

winter: 5 yes and 0 no

$$Ent(Season = spring) = -\frac{1}{2}\times\log_2\frac{1}{2} - \frac{1}{2}\times\log_2\frac{1}{2} = 1$$

$$Ent(Season = summer) = -\frac{1}{3}\times\log_2\frac{1}{3} - \frac{2}{3}\times\log_2\frac{2}{3} = 0.9183$$

$$Ent(Season = autumn) = -\frac{2}{2}\times\log_2\frac{2}{2} = 0$$

$$Ent(Season = winter) = -\frac{5}{5} \times \log_2 \frac{5}{5} = 0$$

$$Gain(D, Season) = Ent(D) - \sum_v \frac{D^v}{D} Ent(D^v) = 0.8113 - 1 \times \frac{2}{12} - 0.9182 \times \frac{3}{12} = 0.4151$$

For After 8:00

- After 8:00 = yes: 3 yes and 2 no

- After 8:00 = no: 6 yes and 1 no

$$Ent(After\ 8:00 = yes) = -\frac{3}{5} \times \log_2 \frac{3}{5} - \frac{2}{5} \times \log_2 \frac{2}{5} = 0.9710$$

$$Ent(After\ 8:00 = no) = -\frac{6}{7} \times \log_2 \frac{6}{7} - \frac{1}{7} \times \log_2 \frac{1}{7} = 0.5917$$

$$Gain(D, After\ 8:00) = Ent(D) - \sum_v \frac{D^v}{D} Ent(D^v) = 0.8113 - 0.9710 \times \frac{5}{12} - 0.5917 \times \frac{7}{12} = 0.0616$$

For Wind:

breeze: 5 yes and 0 no

no wind: 3 yes and 1 no

gale: 1 yes and 2 no

$$Ent(Wind = breeze) = -\frac{5}{5} \times \log_2 \frac{5}{5} = 0$$

$$Ent(Wind = nowind) = -\frac{3}{4} \times \log_2 \frac{3}{4} - \frac{1}{4} \times \log_2 \frac{1}{4} = 0.8113$$

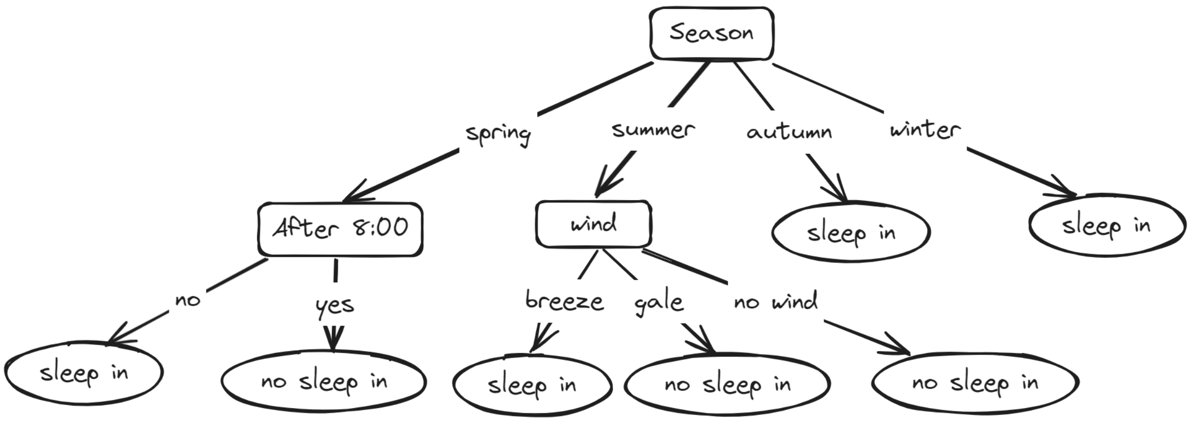
$$Ent(Wind = gale) = -\frac{1}{3} \times \log_2 \frac{1}{3} - \frac{2}{3} \times \log_2 \frac{2}{3} = 0.9183$$

$$Gain(D, Wind) = Ent(D) - \sum_v \frac{D^v}{D} Ent(D^v) = 0.8113 - 0.8113 \times \frac{4}{12} - 0.9183 \times \frac{3}{12} = 0.3113$$

So Season is chosen as root.

After getting the root, the other nodes can be decided by observing the data.

The result is the following:



Q3. Given the following data:

where x is a 2D vector, the first dimension takes values in (1, 2, 3), the second dimension takes values in (S, M, L), and y takes values in (-1, 1). Given new data x = (2, S), try the Naive Bayes method to predict the value of y at this time.

Ans.

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
x (1)	1	1	1	1	1	2	2	2	2	2	3	3	3	3	3
x (2)	S	M	M	S	S	S	M	M	L	L	L	M	M	L	L
y	-1	-1	1	1	-1	-1	1	1	1	1	1	1	1	1	-1

$$P(y=-1|x_1=2,x_2=S)=\frac{P(x_1=2,x_2=S|y=-1)P(y=-1)}{P(x_1=2,x_2=S)}$$

$$P(y=1|x_1=2,x_2=S)=\frac{P(x_1=2,x_2=S|y=1)P(y=1)}{P(x_1=2,x_2=S)}$$

$$P(x_1=2,x_2=S|y=-1)P(y=-1)=\frac{1}{5}\times\frac{3}{5}\times\frac{5}{15}=\frac{1}{25}$$

$$P(x_1=2,x_2=S|y=1)P(y=1)=\frac{4}{10}\times\frac{1}{10}\times\frac{10}{15}=\frac{2}{75}$$

So, $P(x_1=2,x_2=S|y=-1)P(y=-1)>P(x_1=2,x_2=S|y=1)P(y=1)$

We can predict the value of y is -1.