Machine Learning: Homework #2

Professor Shuang Liang

2252709
Xuanhe Yang

May, 2025

Problem 1

Calculate the gradient of the following multivariate functions:
(1) $u = xy + y^2 + 5$
The gradient is:

$$\nabla u = \left(\frac{\partial u}{\partial x}, \frac{\partial u}{\partial y} \right) \tag{1}$$
$$\frac{\partial u}{\partial x} = y \tag{2}$$
$$\frac{\partial u}{\partial y} = x + 2y \tag{3}$$
$$\therefore \nabla u = (y, x + 2y) \tag{4}$$

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

$$u = \ln \sqrt{x^2 + y^2 + z^2} = \ln(x^2 + y^2 + z^2)^{1/2} = \frac{1}{2} \ln(x^2 + y^2 + z^2) \tag{5}$$

Calculate the partial derivatives:

$$\frac{\partial u}{\partial x} = \frac{1}{2} \cdot \frac{1}{x^2 + y^2 + z^2} \cdot 2x = \frac{x}{x^2 + y^2 + z^2} \tag{6}$$
$$\frac{\partial u}{\partial y} = \frac{1}{2} \cdot \frac{1}{x^2 + y^2 + z^2} \cdot 2y = \frac{y}{x^2 + y^2 + z^2} \tag{7}$$
$$\frac{\partial u}{\partial z} = \frac{1}{2} \cdot \frac{1}{x^2 + y^2 + z^2} \cdot 2z = \frac{z}{x^2 + y^2 + z^2} \tag{8}$$

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

$$x^2 + y^2 + z^2 = 1^2 + 2^2 + (-2)^2 = 1 + 4 + 4 = 9 \tag{9}$$

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

$$\nabla u \bigg|_{(1,2,-2)} = \left(\frac{1}{9}, \frac{2}{9}, \frac{-2}{9} \right) = \frac{1}{9}(1, 2, -2) \tag{10}$$

Problem 2

Building a Decision Tree for Sleep-in Prediction
Given the 12-day dataset, we'll build a decision tree using information gain.
Step 1: Calculate initial entropy
Total samples: 12 (9 yes, 3 no)

$$H(Sleep) = -\frac{9}{12} \log_2 \frac{9}{12} - \frac{3}{12} \log_2 \frac{3}{12} = 0.811 \tag{11}$$

Step 2: Calculate information gain for each attribute
For Season:

- Spring (2): 1 yes, 1 no $\rightarrow H = 1.0$
- Summer (3): 1 yes, 2 no $\rightarrow H = 0.918$
- Autumn (2): 2 yes, 0 no $\rightarrow H = 0$

Problem 2 continued on next page...

- Winter (5): 5 yes, 0 no $\to H = 0$

$$H(Sleep|Season) = \frac{2}{12}(1.0) + \frac{3}{12}(0.918) + \frac{2}{12}(0) + \frac{5}{12}(0) = 0.396 \tag{12}$$
$$IG(Season) = 0.811 - 0.396 = 0.415 \tag{13}$$

<u>For After 8:00:</u>

- Yes (5): 3 yes, 2 no $\to H = 0.971$
- No (7): 6 yes, 1 no $\to H = 0.592$

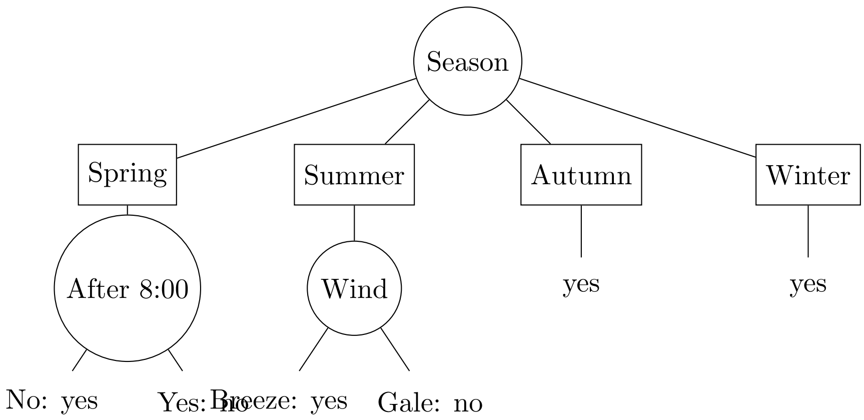
$$H(Sleep|After8) = \frac{5}{12}(0.971) + \frac{7}{12}(0.592) = 0.764 \tag{14}$$
$$IG(After8) = 0.811 - 0.764 = 0.047 \tag{15}$$

<u>For Wind:</u>

- Breeze (5): 5 yes, 0 no $\to H = 0$
- No wind (4): 3 yes, 1 no $\to H = 0.811$
- Gale (3): 1 yes, 2 no $\to H = 0.918$

$$H(Sleep|Wind) = \frac{5}{12}(0) + \frac{4}{12}(0.811) + \frac{3}{12}(0.918) = 0.500 \tag{16}$$
$$IG(Wind) = 0.811 - 0.500 = 0.311 \tag{17}$$

Step 3: Build the decision tree
Since $IG(Season) = 0.415$ is highest, we choose Season as root node.



Note: For Autumn and Winter branches, all instances have Sleep in = yes, so no further splitting is needed. For Spring and Summer, we need to split further based on the instances in each branch.

Problem 3

Given the data where $\mathbf{x}$ is a 2D vector with $x^{(1)} \in \{1, 2, 3\}$, $x^{(2)} \in \{S, M, L\}$, and $y \in \{-1, 1\}$. For new data $\mathbf{x} = (2, S)$, we apply Naive Bayes classification.

Problem 3 continued on next page. . .

Step 1: Calculate prior probabilities

$$P(y = 1) = \frac{10}{15} = \frac{2}{3} \tag{18}$$

$$P(y = -1) = \frac{5}{15} = \frac{1}{3} \tag{19}$$

Step 2: Calculate likelihoods
For $y = 1$ (10 instances):

$$P(x^{(1)} = 2 \mid y = 1) = \frac{4}{10} \tag{20}$$

$$P(x^{(2)} = S \mid y = 1) = \frac{1}{10} \tag{21}$$

For $y = -1$ (5 instances):

$$P(x^{(1)} = 2 \mid y = -1) = \frac{1}{5} \tag{22}$$

$$P(x^{(2)} = S \mid y = -1) = \frac{3}{5} \tag{23}$$

Step 3: Apply Naive Bayes
For $\mathbf{x} = (2, S)$:

$$P(y = 1 \mid \mathbf{x}) \propto P(y = 1) \cdot P(x^{(1)} = 2 \mid y = 1) \cdot P(x^{(2)} = S \mid y = 1) \tag{24}$$

$$= \frac{2}{3} \times \frac{4}{10} \times \frac{1}{10} = \frac{8}{300} = 0.0267 \tag{25}$$

$$P(y = -1 \mid \mathbf{x}) \propto P(y = -1) \cdot P(x^{(1)} = 2 \mid y = -1) \cdot P(x^{(2)} = S \mid y = -1) \tag{26}$$

$$= \frac{1}{3} \times \frac{1}{5} \times \frac{3}{5} = \frac{3}{75} = 0.04 \tag{27}$$

Since $P(y = -1 \mid \mathbf{x}) = 0.04 > P(y = 1 \mid \mathbf{x}) = 0.0267$, the predicted value is $\boxed{y = -1}$.