

Problem 1. (Naive Bayes) 3 points

You want to find a teammate to have a successful team project with. From your experience, the two key factors were whether your teammate had good communication skills and whether they were responsible toward deadlines. You summarized the data of your 6 past group projects below:

Success (S)	1	1	0	0	1	0
Good communication (C)	1	0	1	0	1	1
Responsible (R)	1	0	0	1	1	0

1. What is the probability of a successful project regardless of your teammate? What is the probability your teammate had good communication skills given the project was successful?

$$Prob(S = 1) = \dots\dots\dots$$

$$Prob(C = 1|S = 1) = \dots\dots\dots$$

Solution:

$$Prob(S = 1) = \frac{3}{6} = 0.5$$

$$Prob(C = 1|S = 1) = \frac{2}{3} \approx 0.67$$

2. You want to use a naive Bayes classifier to find the probability that a friend of yours, who has no communication skills and who is not too strict with deadlines, would be a good fit.

Write down the naive Bayes' classifier formula (do not plug in the numerical values)

$$Prob(S = 1|C = 0, R = 0) = \dots\dots\dots$$

$$\text{Solution: } Prob(S = 1|R = 0, C = 0) = \frac{Prob(R=0|S=1)Prob(C=0|S=1)Prob(S=1)}{Prob(R=0,C=0)}$$

3. Circle the term that is larger based on the naive Bayes classifier:

$$(a) Prob(S = 0|C = 0, R = 0) \quad (b) Prob(S = 1|C = 0, R = 0)$$

Solution:

$$Prob(S = 0|C = 0, R = 0)Prob(C = 0, R = 0) = \frac{2}{3} \frac{1}{3} \frac{3}{6} = \frac{1}{9}$$

$$Prob(S = 1|C = 0, R = 0)Prob(C = 0, R = 0) = \frac{1}{3} \frac{1}{3} \frac{3}{6} = \frac{1}{18}$$

Problem 2. (k Nearest Neighbors) 2 points

1. A company wants to detect whether its product is faulty ($y = 1$) or intact ($y = 0$) based on the product's weight (x_1) and length (x_2). The company asks you to use the k-nearest neighbor approach with $k = 1$. For a product with $x^{test} = (2, 4)$, you are given two datapoints $x^1 = (2.2, 3.8)$ with $y^1 = 1$ and $x^2 = (2.3, 4.0)$ with $y^2 = 0$. Determine the label of the test point x^{test} based on the **Manhattan** distance. (1 point)

2. Now, consider kNN for regression, where you want to predict the lifetime of the product in hours: $y \in \mathbb{R}$. Compute the label of x^{test} given that the labels of the $k = 3$ closest data points are $y^1 = 2.3$, $y^2 = 2.5$, and $y^3 = 1.8$. (1 point)

$$y^{test} = \dots\dots\dots$$

Solution:

1. Manhattan distance:

$$d(x^{test}, x^i) = |2 - 2.2| + |4 - 3.8| = 0.4$$

$$d(x^{test}, x^j) = |2 - 2.3| + |4 - 4| = 0.3$$

1-nearest neighbor using the Euclidean distance labels the product with $x^{test} = (2, 4)$ as belonging to the same class as x^2 , so to class $Y = 0$ and thus intact.

$$2. y^{test} = \frac{1}{3}(2.3 + 2.5 + 1.8) = 2.2$$

Problem 3. (CNN) 2 points

Fill in the empty blocks in the convolution of X with a 2×2 filter K and a bias $b = 0$.

1	0	0
1	1	0
1	0	1

X

0	1
1	0

K

Bias: $b = 0$

1	1
2	0

$X * K$

Problem 4. (Neural networks) 3 points

Consider a neural network $f : [-2, 2] \rightarrow \mathbb{R}$, $f(x) = W^{[1]\top} g(W^{[0]\top} x + b^{[0]}) + b^{[1]}$ with a single hidden layer and activation function $g(x) := \max(0, x)$, where $W^{[0]} = [1, 1]$, $b^{[0]} = [0, -1]^\top$, and $W^{[1]\top} = [-1, 1]$ and $b^{[1]} = 0$.

1. (1 point) Write the explicit expression for $g(W^{[0]\top} x + b^{[0]}) \in \mathbb{R}^2$ as a function of $x \in \mathbb{R}$.

2. (2 points) Write the explicit expression for the function f and draw its graph for $x \in [-2, 2]$.

Solution:

$$1. \ g(W^{[0]\top}x + b^{[0]}) = \begin{bmatrix} g(x) \\ g(x-1) \end{bmatrix}$$

2. We have

$$\begin{aligned} f(x) &= -g(x) + g(x-1) \\ &= \begin{cases} 0, & \text{if } -2 \leq x < 0, \\ -x, & \text{if } 0 \leq x < 1, \\ -1, & \text{if } 1 \leq x \leq 2. \end{cases} \end{aligned}$$

