Homework 1

Habin Lee

September 29, 2024

1 Confusion matrix

The results of a binary classification model are shown below:

	Actual Positive	Actual Negative	Total
Predicted Positive	50	10	60
Predicted Negative	5	35	40
Toatal	55	45	100

Table 1: Table 1: Confusion Matrix

True position(TP) = 50 True Negative(TN) = 35 False Positive(FP) = 10 False Negative(FN) = 5

1. Calculate the accuracy.
Accuracy = $\frac{\text{TP+TN}}{\text{Total}} = \frac{50+35}{100} = 0.85$

2. Calculate the precision.

precision = $\frac{\text{TP}}{\text{TP+FP}} = \frac{50}{50+10} = 0.8333$

3. Calculate the recall.

recall = $\frac{\text{TP}}{\text{TP+FN}} = \frac{50}{50+5} = 0.9091$

4. Calculate the F1 score.

F1score = $2 \times \frac{\text{precision} \times \text{recall}}{\text{precision+recall}} = \frac{0.8333 \times 0.9091}{0.8333+0.9091} = 0.8696$

2 Linear Regression

MSE(mean squared error) Data Sample $D = (x_1, y_1), (x_2, y_2), \ldots (x_m, y_m)$ y_i is answer data of

$$E(f; D) = \frac{1}{m} \sum_{i=1}^{m} (f(x_i) - y_i)^2$$

$$E(f; D) = \frac{1}{m} \sum_{i=1}^{m} \text{II}(f(x_u) \neq y_i)$$

$$acc(f; D) = \frac{1}{m} \sum_{i=1}^{m} \text{II}(f(x_u) = y_i)$$
$$= 1 - E(f; D)$$

Linear Regression Solution

article amsmath

Linear Regression Solution

Given the dataset:

x	y
1	2
2	3
3	5
4	4
5	6

We need to derive the closed-form solutions for the parameters w (slope) and b (intercept) using mean square error (MSE).

Detailed Calculation Steps

(a) Derive the closed-form solutions: The linear regression model is given by:

$$y = wx + b$$

The formulas for w and b are:

$$w = \frac{n(\sum xy) - (\sum x)(\sum y)}{n(\sum x^2) - (\sum x)^2}$$

$$b = \frac{\sum y - w(\sum x)}{n}$$

Where: - $n = 5$ - $\sum x = 1 + 2 + 3 + 4 + 5 = 15$ - $\sum y = 2 + 3 + 5 + 4 + 6 = 20$ - $\sum xy = (1)(2) + (2)(3) + (3)(5) + (4)(4) + (5)(6) = 69$ - $\sum x^2 = 1^2 + 2^2 + 3^2 + 4^2 + 5^2 = 55$

(b) Use the solution derived above:

1. Calculate w:

$$w = \frac{5(69) - (15)(20)}{5(55) - (15)^2} = \frac{345 - 300}{275 - 225} = \frac{45}{50} = 0.9$$

2. Calculate b:

$$b = \frac{20 - 0.9(15)}{5} = \frac{20 - 13.5}{5} = \frac{6.5}{5} = 1.3$$

Result Verification

(a) Clearly state the final values of w and b: - Slope (w): 0.9 - Intercept (b): 1.3

Thus, the linear regression model is:

$$y = 0.9x + 1.3$$