1

Based on whether the training data has labeled information, learning tasks can be broadly divided into two categories: "supervised learning" and "unsupervised learning".

Supervised learning is a type of machine learning where the model is trained on labeled data, meaning the input data is paired with corresponding output labels. The goal is to learn a mapping function that can predict the output labels for new, unseen input data.

Unsupervised learning, on the other hand, is a type of machine learning where the model is trained on unlabeled data. The goal is to discover patterns, structures, or relationships within the data without any predefined output labels.

Two representative algorithms for supervised learning are:

- Linear Regression: It is used for predicting a continuous output variable based on one or more input features.
- Support Vector Machines (SVM): It is a powerful algorithm used for classification and regression tasks. SVM finds the best hyperplane that separates the data into different classes.

Two representative algorithms for unsupervised learning are:

- K-means Clustering: It is used to partition data into distinct groups or clusters based on their similarity.
- Principal Component Analysis (PCA): It is used to reduce the dimensionality of high-dimensional data while preserving most of the important information.

2

No.	X	Y
0	0	2
1	2	2
2	3	1

1. $No.2$ is the test sample.
   Let $No.0$,$No.1$ fit the linear regression model,we have $w = 0, b = 2$.
   We test $No.2$, $output = 2$
2. $No.0$ is the test sample.
   Let $No.1$,$No.2$ fit the linear regression model,we have $w = -1, b = 4$.
   We test $No.0$, $output = 4$
3. $No.1$ is the test sample.
   Let $No.0$,$No.2$ fit the linear regression model,we have $w = -1/3, b = 2$.
   We test $No.1$, $output = 4/3$

$$MSE = \frac{1}{3}\sum_{i=1}^{m}(f(x_i) - y_i)^2 = \frac{1}{3}((2-1)^2 + (4-2)^2 + (\frac{4}{3}-2)^2)$$
$$= \frac{1}{3}(1 + 4 + \frac{4}{9}) = \frac{1}{3} \times \frac{49}{9} = \frac{49}{27} \approxeq 1.815$$

3

$$MAE = \frac{1}{m}\sum_{i=1}^{m}|f(x_i) - y_i| = \frac{22}{12} = 1.833$$

$$MSE = \frac{1}{m}\sum_{i=1}^{m}(f(x_i) - y_i)^2 = \frac{56}{12} = 4.667$$

4

(1)

A:

	Predicted A	Predicted not A
Actual A	40	30
Actual not A	35	155

$$Precision = \frac{TP}{TP+FP} = \frac{40}{75} = 0.533$$

$$Recall = \frac{TP}{TP+FN} = \frac{40}{70} = 0.571$$

B:

	Predicted B	Predicted not B
Actual B	85	75
Actual not B	30	70

$$Precision = \frac{TP}{TP+FP} = \frac{85}{115} = 0.739$$

$$Recall = \frac{TP}{TP+FN} = \frac{85}{160} = 0.531$$

C:

	Predicted C	Predicted not C
Actual C	20	10
Actual not C	50	180

$$Precision = \frac{TP}{TP+FP} = \frac{20}{70} = 0.286$$

$$Recall = \frac{TP}{TP+FN} = \frac{20}{30} = 0.667$$

(2)

Macro-average:

$$P_{macro} = \frac{1}{N}\sum_{i=1}^{N} P_i = 0.5194$$

$$R_{macro} = \frac{1}{N}\sum_{i=1}^{N} R_i = 0.5898$$

Weighted-average:

$$w = [\frac{70}{260}, \frac{160}{260}, \frac{30}{260}]$$

$$P_{weighted} = \sum_{i=1}^{N} w_i \cdot P_i = 0.6314$$

$$R_{weighted} = \sum_{i=1}^{N} w_i \cdot R_i = 0.5577$$