Machine Learning Homework 1

Question

1. Explain what is supervised and unsupervised learning, and list two representative algorithms respectively.

2. For *K*-fold cross-validation, if K = the number of samples, it is called leave-one-out method. Suppose you have the following data: both input and output have only one variable. Use a linear regression model ($y = wx + b$) to fit the data. So what is the mean squared error (MSE) obtained using Leave-One Out cross-validation?

X(independent variable)	Y(dependent variable)
0	2
2	2
3	1

3. Let us take an example of actual demand and forecasted demand for a brand of ice creams in a shop in a year. In the table, M = Month, AD = Actual Demand, FD = Forecasted Demand.

M	1	2	3	4	5	6	7	8	9	10	11	12
AD	42	45	49	55	57	60	62	58	54	50	44	40
FD	44	46	48	50	55	60	64	60	53	48	42	38

Please calculate the MAE(mean absolute error) and MSE(mean square error).

4. Consider a triple classification problem that requires the recognition of three classes, A, B and C.

	Predicted		
	A	**B**	**C**
A	40	20	10
B	35	85	40
C	0	10	20

Given the confusion matrix, please complete the following calculations.

(1) Calculate the precision and recall for each class respectively.

(2) Use both Macro-average and Weighted-average to calculate the precision and recall of the whole classifier. If you are not sure about these concepts, use a search engine. Retain 4 decimal places (only for (2)).

5. Try to implement an agent with practical application value on Tongji AI. Provide an introduction to its functions and screenshots of its application.

Submission

1. **Format**: Please submit a pdf/doc/docx file and name it in this format: **HW1+Student_ID+Name**. Example: HW1+1234567+张三.pdf

2. **Deadline**: 2025/4/25 23:59. You have 2 weeks.

3. Please submit your homework to **Canvas**.

4. **Late policy**: **7** free late days

    a) Use up to 4 late days per assignment.

    b) Afterwards, 25% off per day late.