國 立 陽 明 交 通 大 學
NATIONAL YANG MING CHIAO TUNG UNIVERSITY

Python 與機器學習演算法 **SCMA30084**
Python and Machine Learning Algorithms

第二次段考 June 8, 2026 **Exam 2**

姓名 Name : Solution

學號 Student ID # :

Lecturer:	Jephian Lin 林晉宏
Contents:	cover page, **5 pages** of questions, score page at the end
To be answered:	on the test paper
Duration:	**180 minutes**
Total points:	**20 points** + 2 extra points

Do not open this packet until instructed to do so.

Instructions:

- Enter your **Name** and **Student ID #** before you start.
- If the answer is too long, round it to **two decimal places**. Numerical answers with a **tolerance** ± 0.05 are accepted.
- Any work necessary to arrive at an answer must be shown on the examination paper. Marks will not be given for final answers that are not supported by appropriate work.
- Clearly indicate your final answer to each question either by **underlining it or circling it**. If multiple answers are shown then no marks will be awarded.
- Please answer the problems in **English**.

1. Let

$$X = \begin{bmatrix} 0.10 & 0.10 & 0.10 & 0.10 & 0.10 & 0.10 \\ 1.51 & 1.51 & 1.51 & 1.51 & 1.51 & 1.51 \\ 0.48 & 0.86 & -1.03 & 0.48 & 0.86 & -1.03 \end{bmatrix}$$

be a data matrix whose rows are the samples.

(a) [2pt] Apply PCA to X. Find the **first two principal components** and find the corresponding **explained variances**.

model.components_ → $\vec{P}_1 = (0.32, 0.25, 0.58, 0.32, 0.25, 0.58)$

$\vec{P}_2 = (0.32, 0.5, -0.39, 0.32, 0.5, -0.39)$

model.explained_variance_ $var_1 = 4.43$

$var_2 = 0.9$

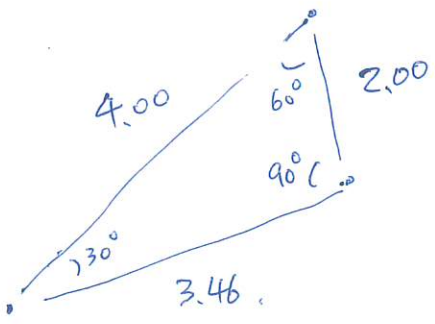
(b) [1pt] Explain why the third explained variance is almost zero.

The three points in X fall on a 2-dim affine plane.

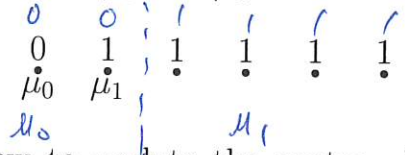
(c) [2pt] Apply MDS to X and reduce its dimension to 2. **Plot** the three points in $\mathbb{R}^2$ to form a triangle. Label **each side** with its length and each **interior angle** with its measure.

MDS preserves the pairwise distances if possible.

In the case, all distances are preserved.



2. Consider the following dataset and apply the k-means clustering algorithm with $k = 2$. Suppose that, at some iteration, the algorithm has the configuration shown below, where the number above each point represents its cluster label, and μ_0 and μ_1 are the current cluster centers.



(a) [1pt] **Describe** how to update the center. **Draw** the new cluster centers above.

new μ_j = center of all points labeled by j.

(b) [1pt] After updating the cluster centers, **describe** how to update the cluster labels. **Mark** the new cluster labels above.

For each point, find the closest center μ_j and label the point by j.

3. Apply DBSCAN to the iris dataset with `eps=0.3` and `min_samples=5`.

(a) [1pt] **How many clusters** are there in the output? For each cluster, **how many data points** does it contain?

```
y_new = model.fit_predict(X)
np.unique(y_new, return_counts=True)
```

→ (-1, 0, 1, 2)
(96, 37, 12, 5)

So 3 clusters and their sizes are 37, 12, 5.

(b) [1pt] How many points are labeled as noise?

96.

This page is supposed to have 5 points in total.
So 1 point will be granted.

4. [2pt] Let

```
from sklearn.datasets import load_digits
digits = load_digits()
X = digits.data
y = digits.target
X_train = X[:100]
y_train = y[:100]
```

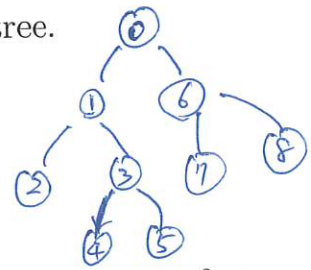
and train a k-nearest neighbor classifier with $k = 5$, `X_train`, and `y_train`. Let `p = X[1008]`. For the k nearest neighbors of `p`, determine their **indices and labels**. Then determine the **predicition of the label** of `p` under this setting.

The 5 nearest nbrs are indexed by 12, 37, 74, 84, 22
and their labels are 2, 9, 5, 2, 2.
By majority rule, the prediction is 2.

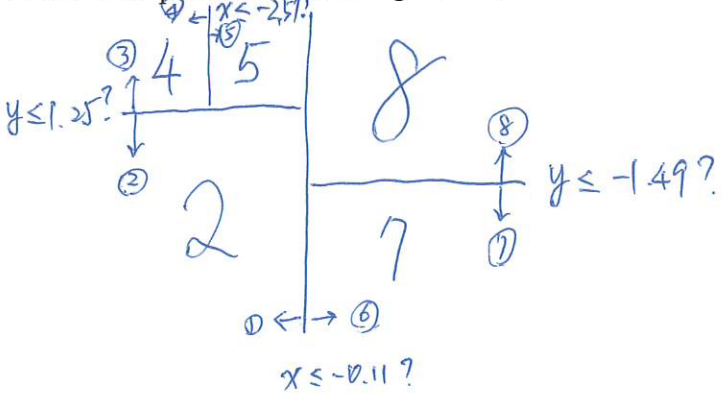
5. [3pt] Let `t` be a tree trained by the decision tree algorithm with the following information.

index	0	1	2	3	4	5	6	7	8
t.children_left	1	2	-1	4	-1	-1	7	-1	-1
t.children_right	6	3	-1	5	-1	-1	8	-1	-1
t.feature	0	1	-2	0	-2	-2	1	-2	-2
t.threshold	-0.11	1.25	-2	-2.57	-2	-2	-1.49	-2	-2
t.n_node_samples	200	86	76	10	2	8	114	22	92
t.impurity	0.5	0.17	0	0.32	0	0	0.31	0	0

(a) [1pt] Draw the tree.



(b) [2pt] Draw the partition of $\mathbb{R}^2$ given by the tree.



6. [5pt] Mathematical essay: Write a few paragraphs to introduce the *gradient descent algorithm.*

   Your score will be based on the following criteria.

   - The definition is clear.
   - Some sentences are added to explain the definition.
   - Examples or pictures are included to help understanding.
   - The sentences are complete.

7. [extra 2pt] Let A be an $m \times n$ real matrix and $\mathbf{x} \in \mathbb{R}^n$. Consider the function

$$f(A, \mathbf{x}) = \|A\mathbf{x}\|^2.$$

Let $a_{i,j}$ be the i,j-entry of A. Determine the gradient of f with resepect to A:

$$\nabla_A(f) = \left[\frac{\partial f}{\partial a_{i,j}}\right]$$

Compute the directional derivative:

$$\lim_{t\to 0} \frac{f(A+tB, \vec{x}) - f(A, \vec{x})}{t} = \lim_{t\to 0} \frac{\langle (A+tB)\vec{x}, (A+tB)\vec{x}\rangle - \langle A\vec{x}, A\vec{x}\rangle}{t}$$

$$= \lim_{t\to 0} \frac{\langle A\vec{x}+tB\vec{x}, A\vec{x}+tB\vec{x}\rangle - \langle A\vec{x}, A\vec{x}\rangle}{t}$$

$$= 2\langle A\vec{x}, B\vec{x}\rangle$$

$$= \langle 2Axx^T, B\rangle$$

So $\nabla_A(f) = 2Axx^T$

[END]

Page	Points	Score
1	5	
2	5	
3	5	
4	5	
5	2	
Total	20 (+2)	