Third week of March 2023

國立中興大學
NATIONAL CHUNG HSING UNIVERSITY

Experimental Log

Smart Motorcycle Online Diagnosis and Detection and Evaluation System for Driving Behavior

Department of Electrical Engineering, National Chung Hsing University
Student: Yen, Wei-Liang (William)
Supervisor: Prof. Hsiao-Ping Tsai

01

Last week's review comments

- Confirm the grouping effect
- It is better to divide the experiment into several groups
- Observe the characteristics of group data

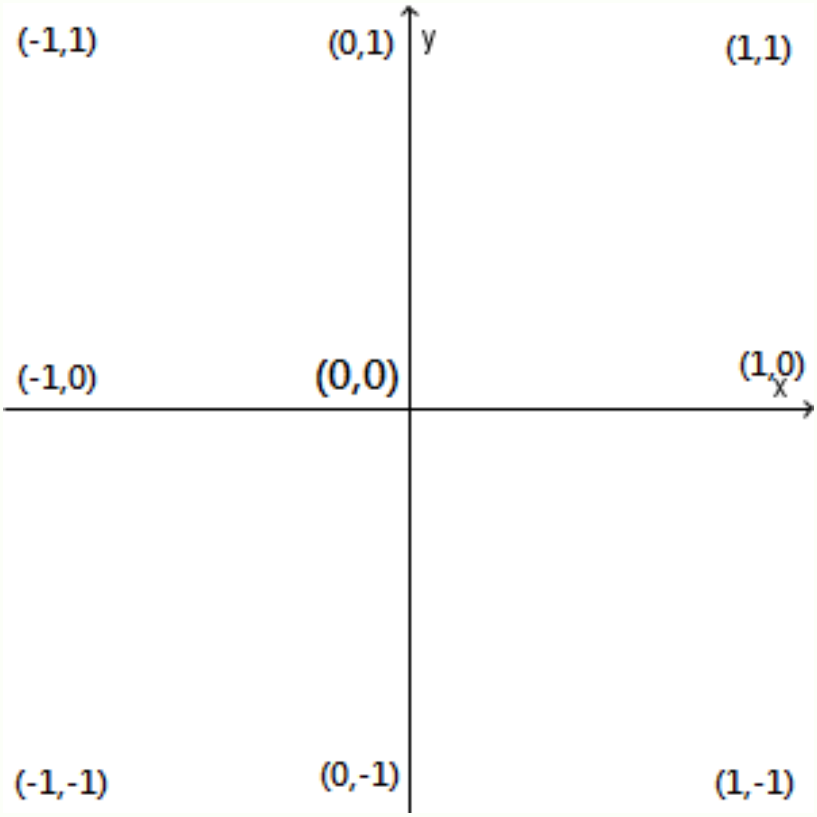
Fixes and Improvements

Q：Confirm the grouping effect:

A：Use PCA to reduce the main features to 2. Based on last week's experience, reducing the features to 3 or 2 had the same effect. The weight of each feature in the reduction process is as follows

	X-axis Acceleration	Y-axis Acceleration	Z-axis Acceleration	X-axis Angle	Y-axis Angle	Z-axis Angle
PCA1	0.469917	0.034039	0.320600	0.129014	0.045299	0.001132
PCA2	0.257958	0.199543	0.434687	0.085491	0.007442	0.014878

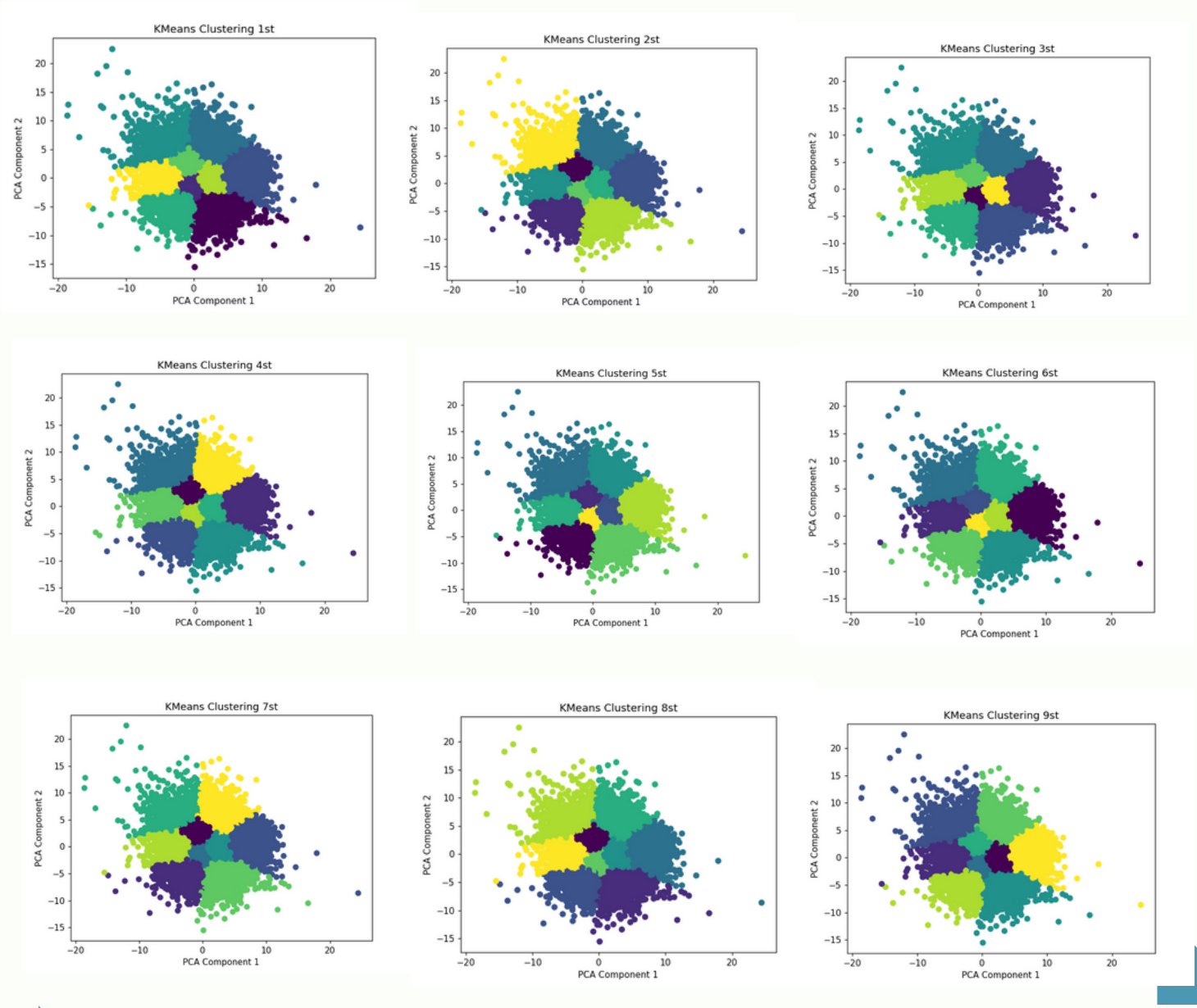
Set the values of both axes to -1, 0, and 1 so that the data can be divided into 9 groups.



Fixes and Improvements

Q：Confirm the grouping effect:

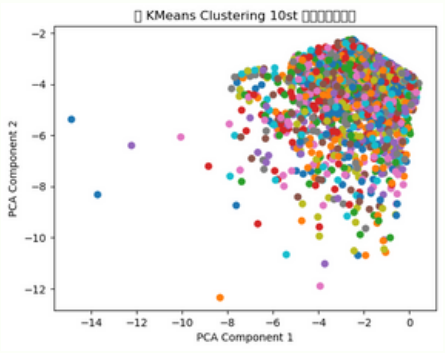
A：Since the K-means algorithm can produce different results depending on the initial points, we ran the K-means algorithm 10 times in the experiment and took the intersection of each group from the 10 experiments

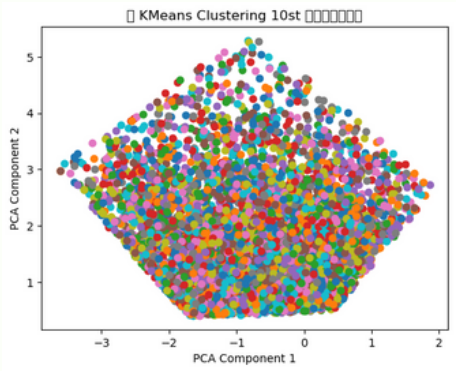


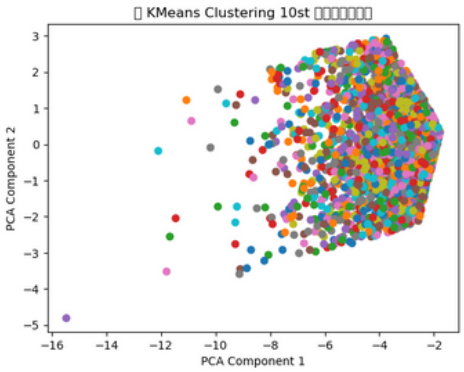
Fixes and Improvements

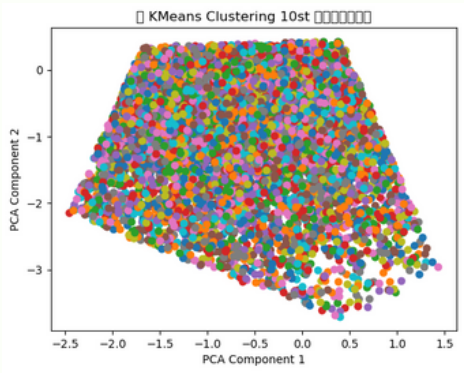
Q：Confirm the grouping effect:

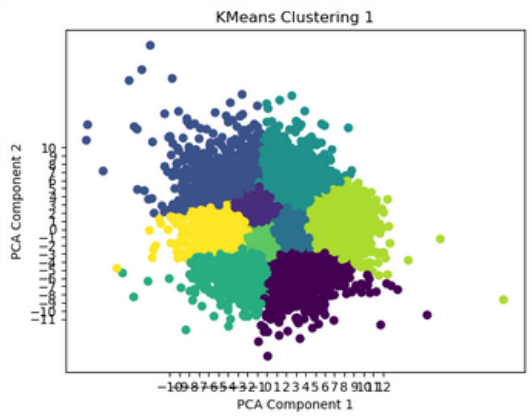
A：Since the K-means algorithm can produce different results depending on the initial points, we ran the K-means algorithm 10 times in the experiment and took the intersection of each group from the 10 experiments

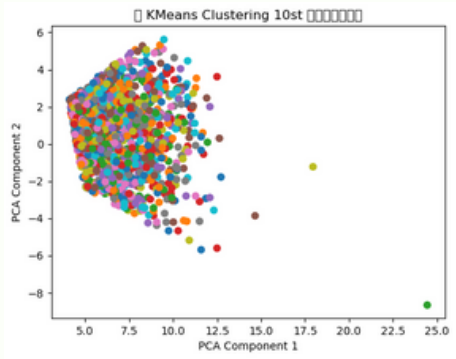


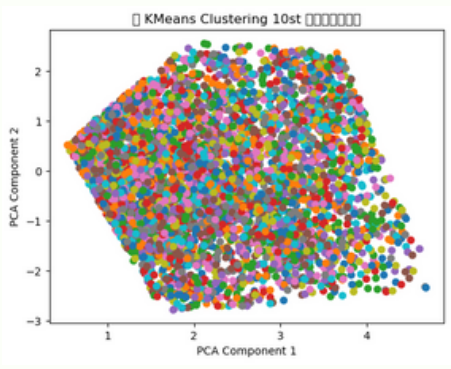


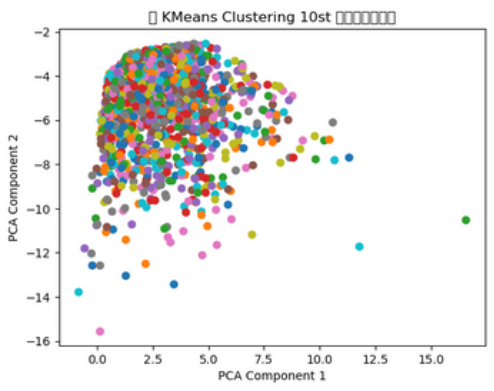


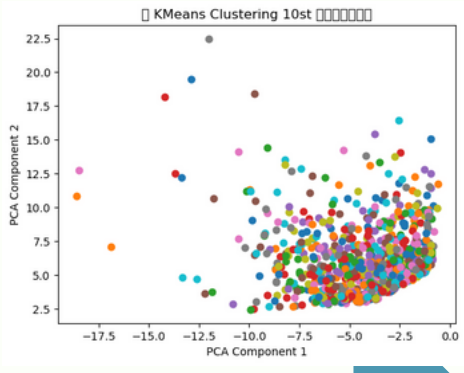










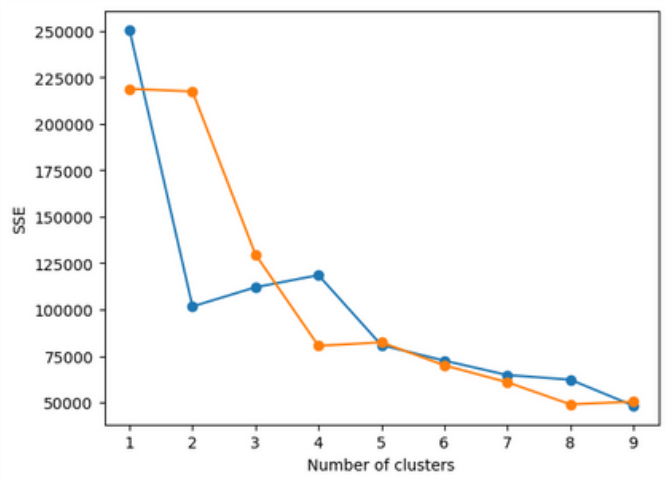


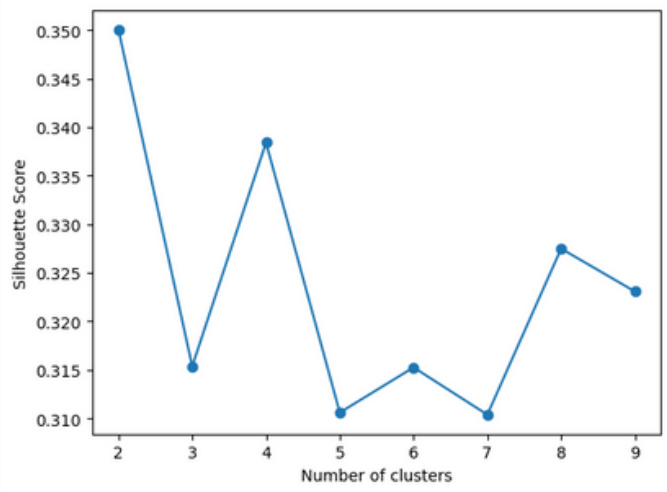
Fixes and Improvements

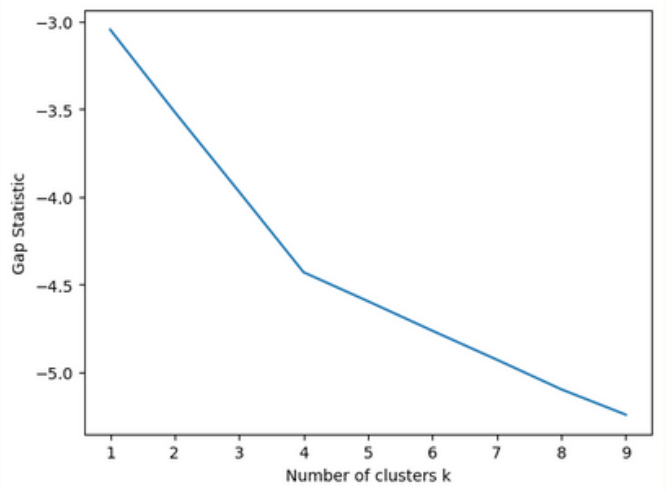
Q：It is better to divide the experiment into several groups

A：

一、Elbow Method
二、Silhouette Score
三、 Gap Statistic

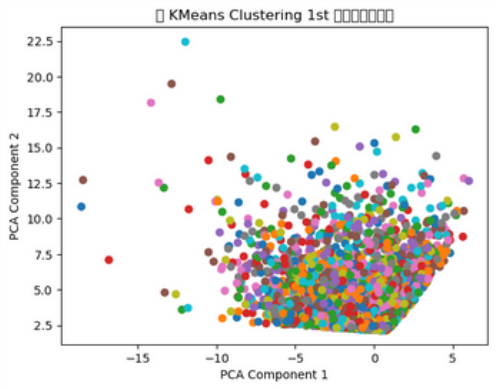


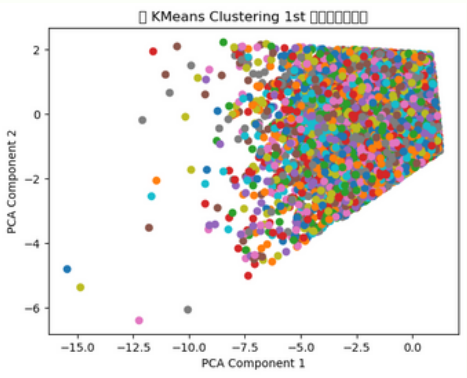


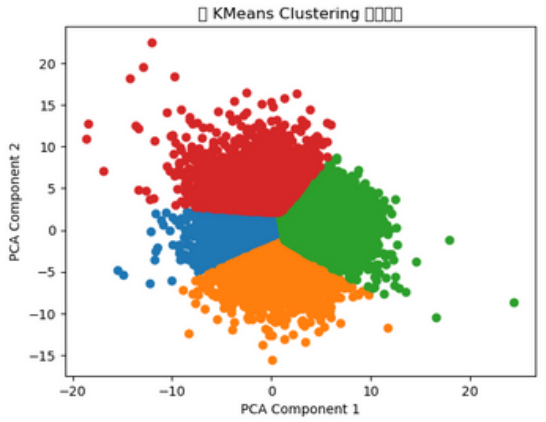


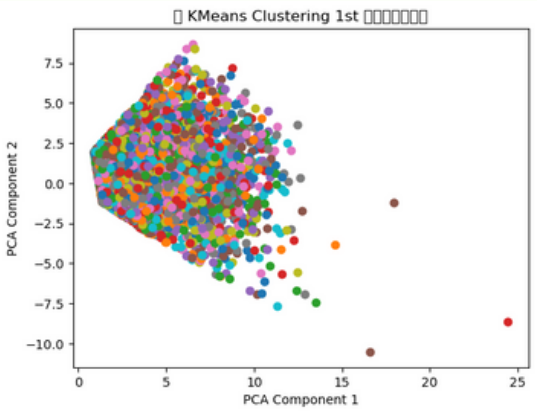
Fixes and Improvements

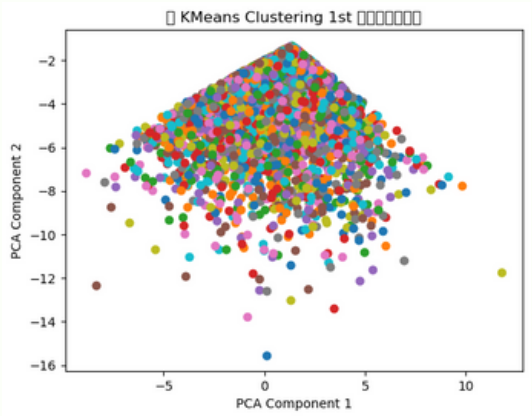
A： Into four groups





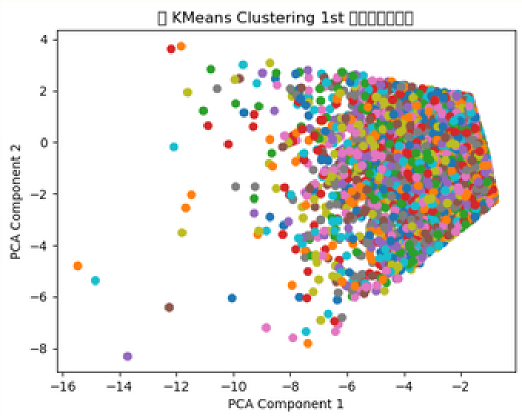


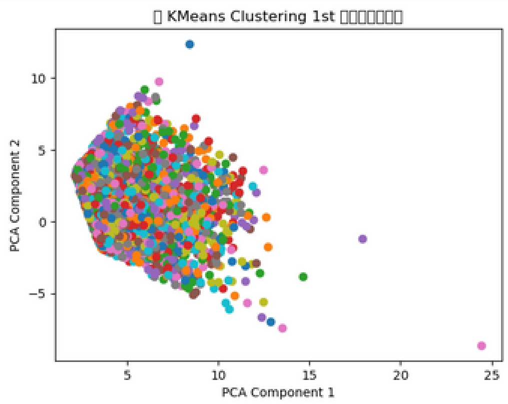


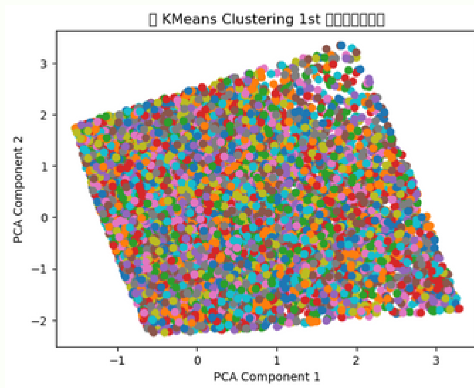


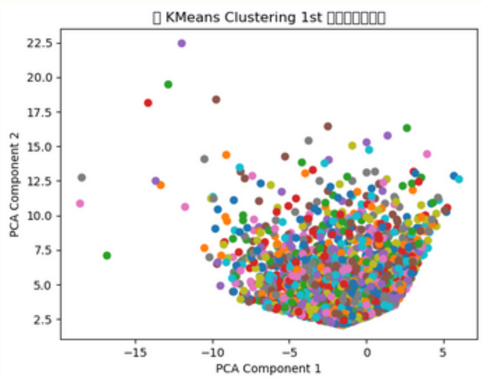
Fixes and Improvements

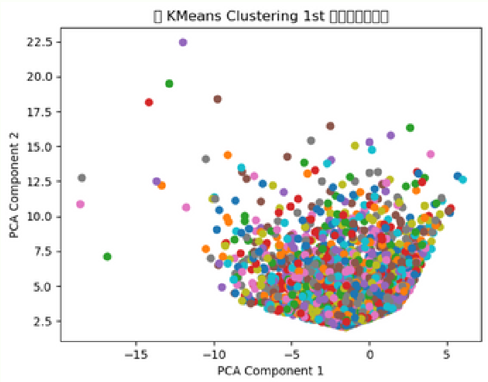
A： Into five groups

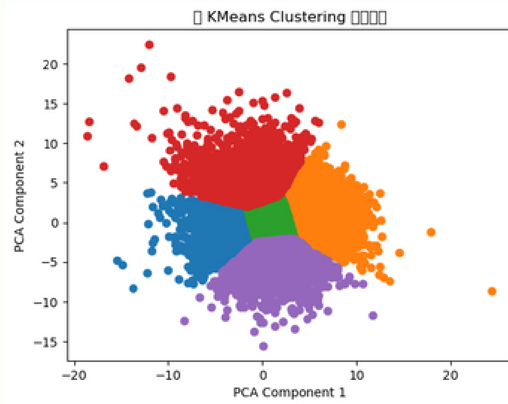












Fixes and Improvements

A:

方法一：Elbow Method

是一種常見的通過 SSE（Sum of Squared Errors，誤差平方和）來判斷聚類分群效果的方法，其基本思想是：隨著分群數 k 的增加，SSE 會逐漸變小，因為更多的簇中心會更好地適應數據，但當 k 增加到一定程度時，SSE 的下降幅度會變緩，這時候再增加 k 反而會導致 SSE 變大。

方法二：Silhouette

在使用 Silhouette 分數來評估聚類結果時，一般認為分數的範圍在 -1 到 1 之間，越接近 1 表示聚類效果越好，越接近 -1 則表示聚類效果越差。通常可以使用以下規則來解釋 Silhouette 分數：

- 分數接近 1 表示聚類效果非常好
- 分數在 0.7 到 0.99 之間表示聚類效果很好
- 分數在 0.5 到 0.7 之間表示聚類效果一般，需要進一步檢查
- 分數在 0.25 到 0.5 之間表示聚類效果較差
- 分數在 -1 到 0.25 之間表示聚類效果非常差

當然，這些規則只是一個大致的參考，實際上應該根據具體的資料和分析目的來選擇最適合的分群數和分數閾值。此外，還應該結合可視化等方法來綜合評估聚類效果。

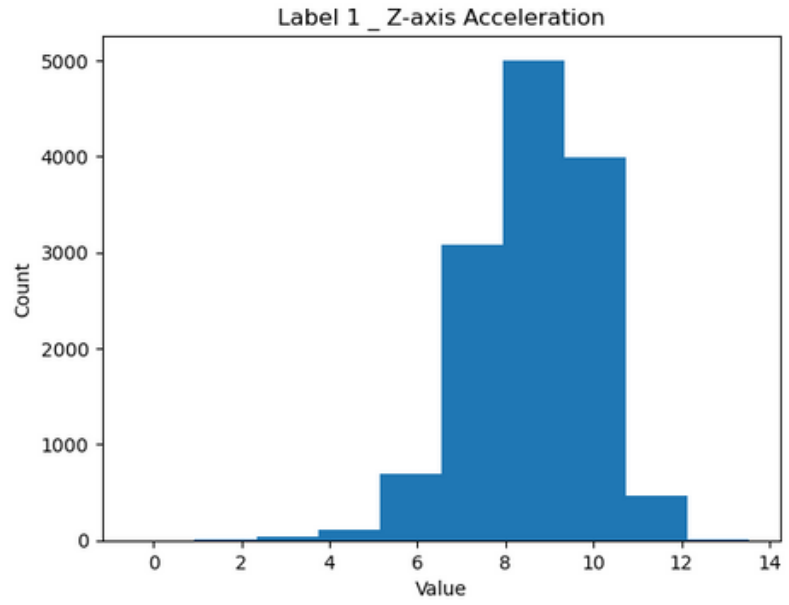
方法三：Gap Statistic

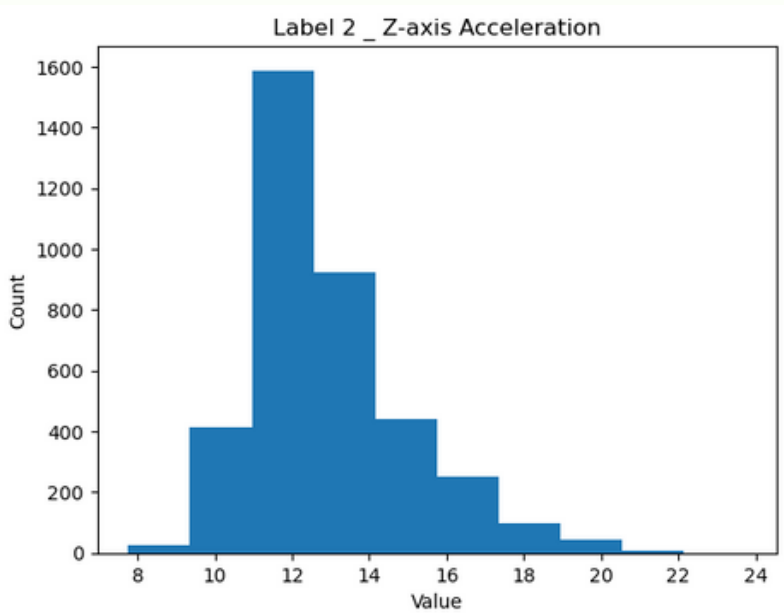
Gap Statistic 分數越大越好。因為 Gap Statistic 的值是透過比較模擬數據和實際數據的 SSE 差異來計算，而 SSE 值越小代表分群效果越好，所以 Gap Statistic 值越大代表分群效果越好。

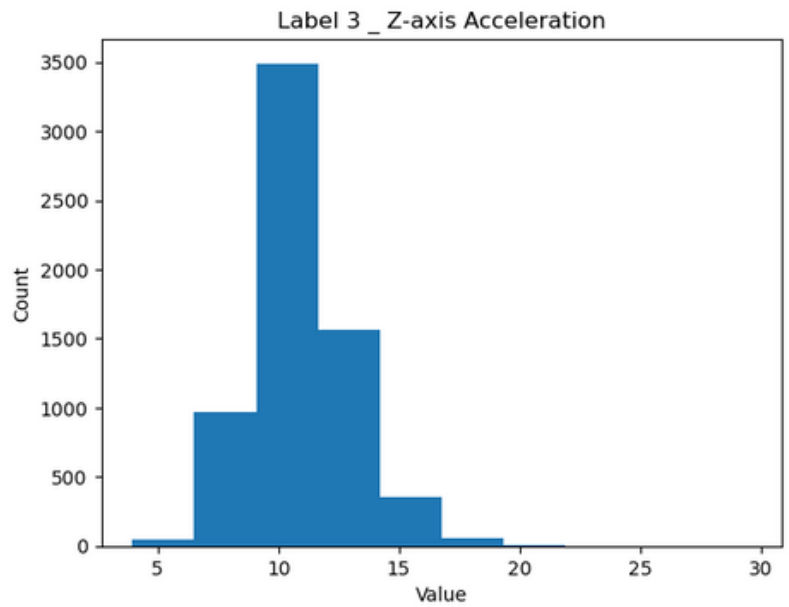
Fixes and Improvements

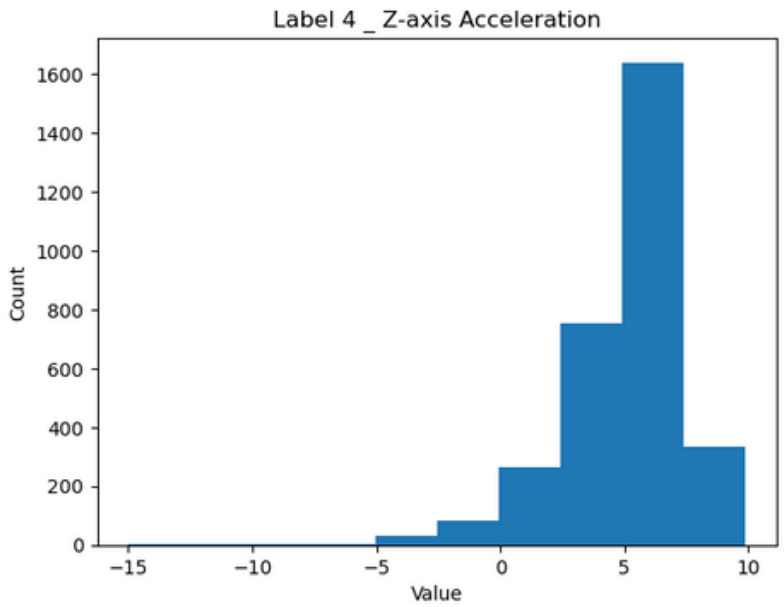
Q: Divided into four groups, the value relationship diagram for each group and Z-axis Acceleration

A:



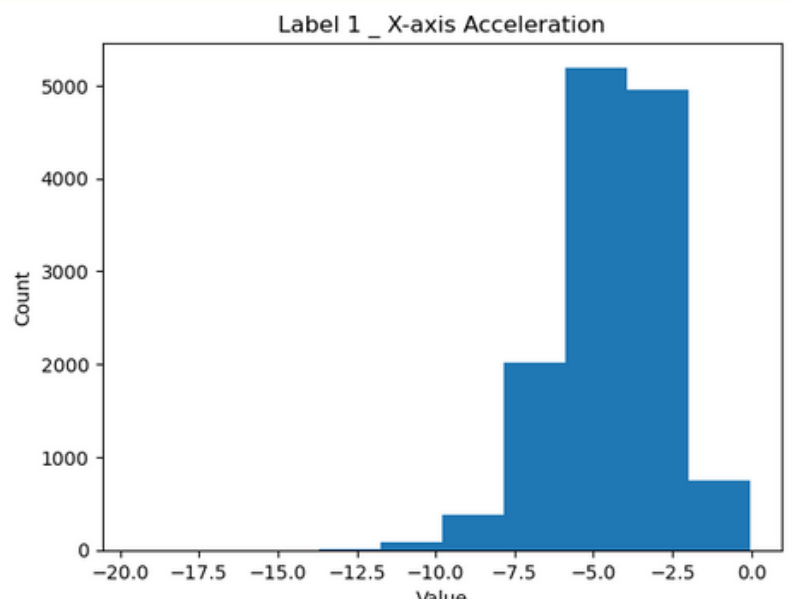


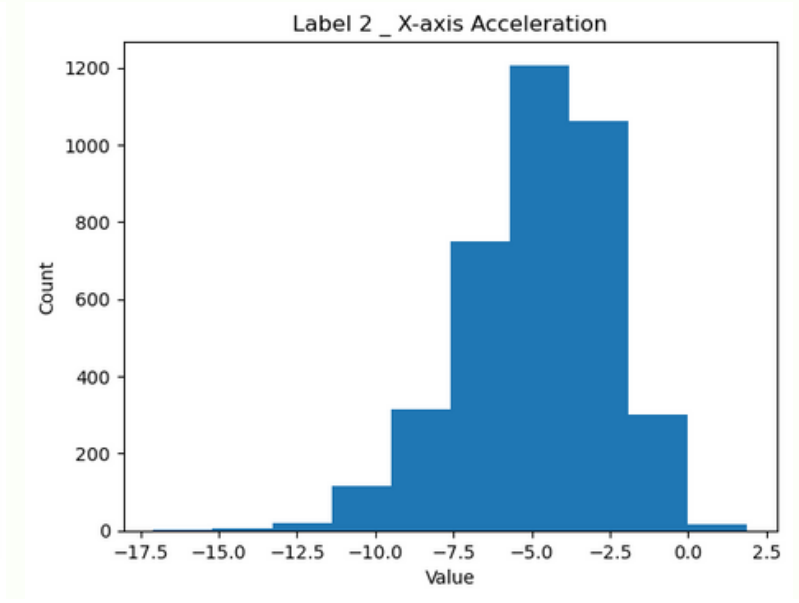


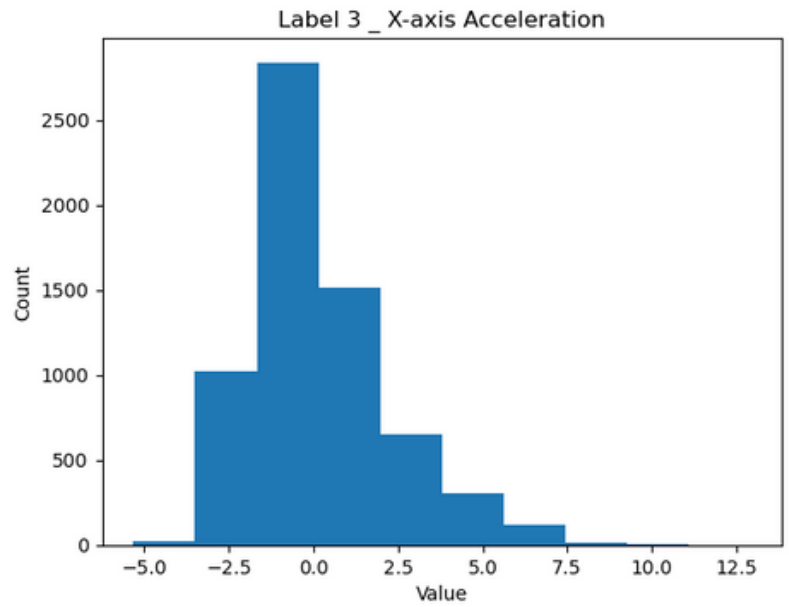


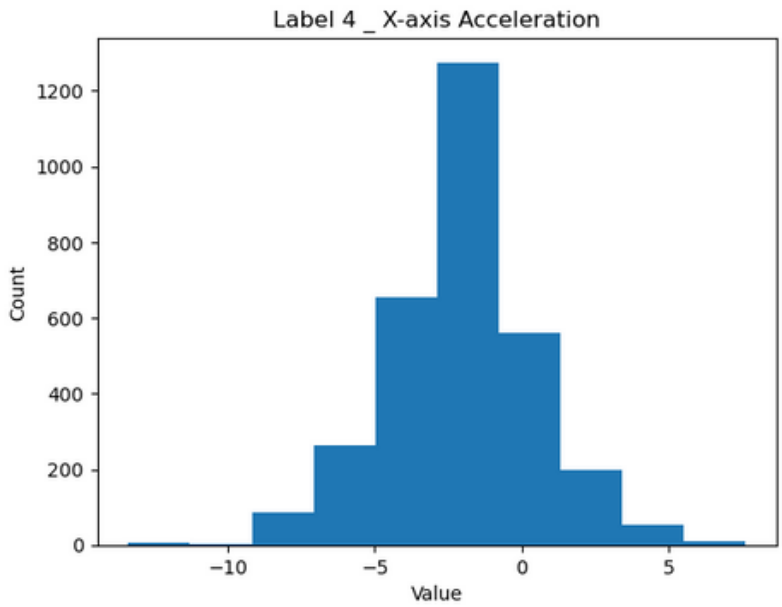
Fixes and Improvements

Q：Q：Divided into four groups, the value relationship diagram for each group and X-axis Acceleration



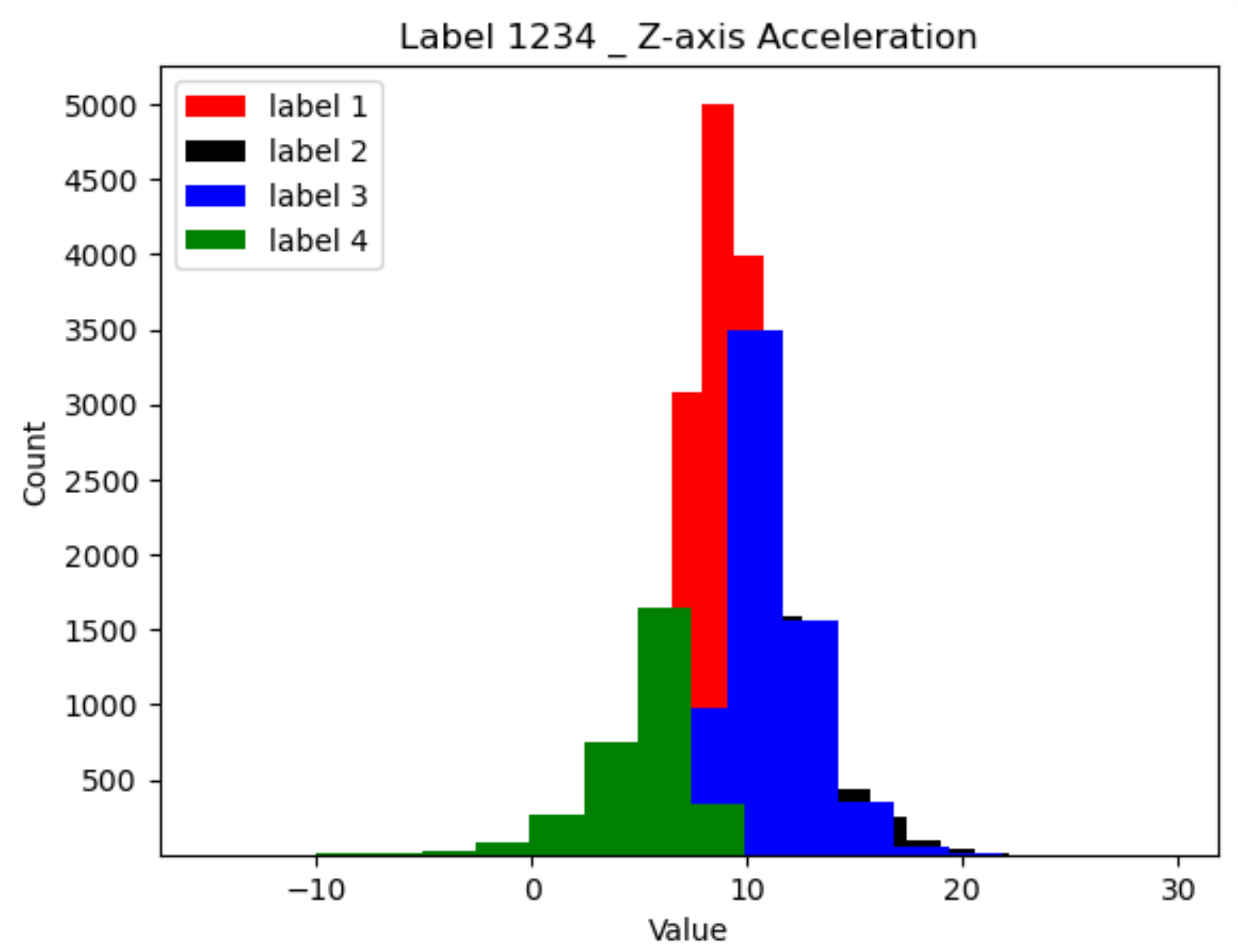






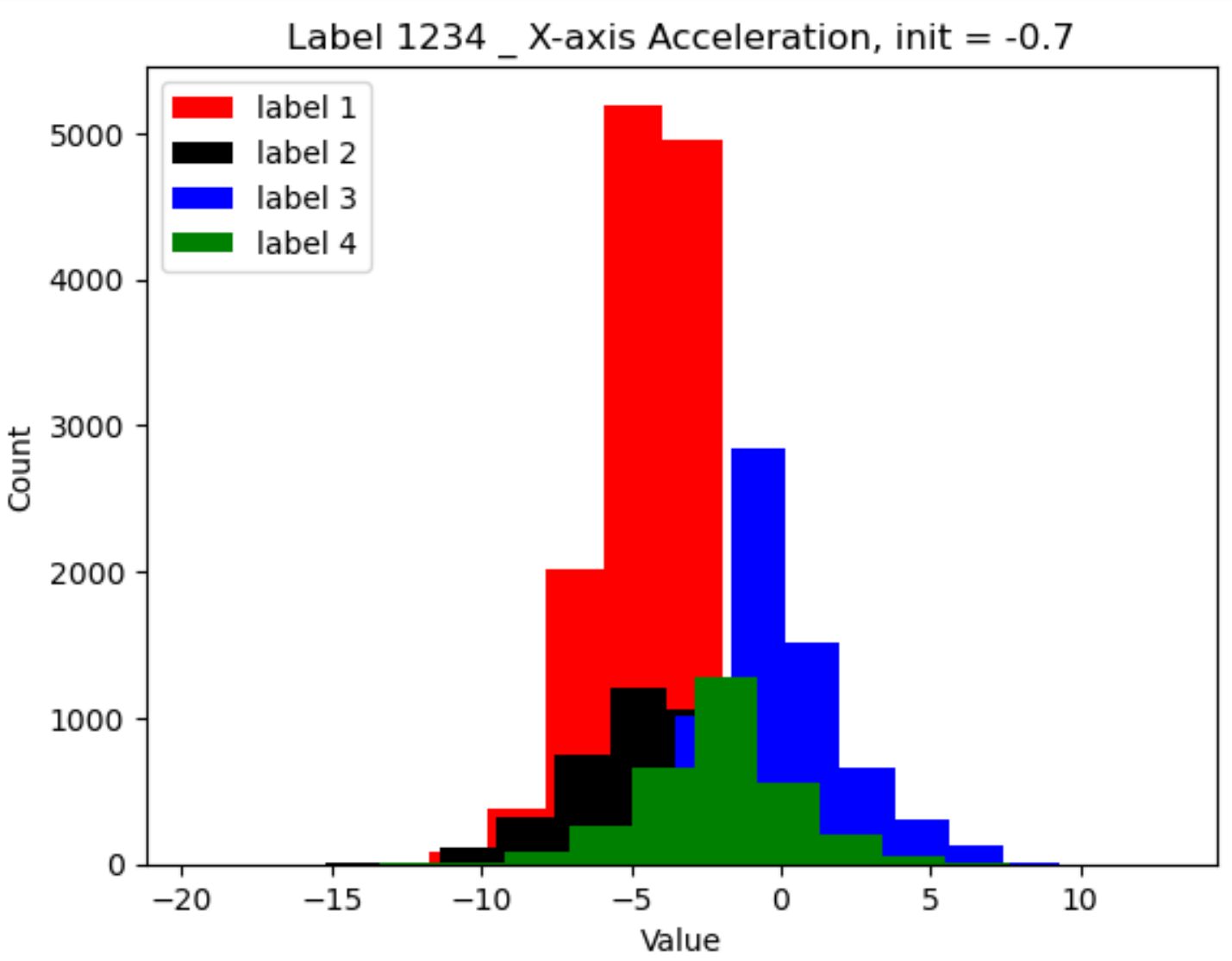
Fixes and Improvements

Q：Q：Divided into four groups, the value relationship diagram for each group and Z-axis Acceleration



Fixes and Improvements

Q：Q：Divided into four groups, the value relationship diagram for each group and X-axis Acceleration



Progress this week

- 測試分群效果
- 需要加怠速嗎?