Readme

**1. Introduction**

This document outlines the structure and usage of our dataset, which is comprised of two types of cross-validation datasets: a feature dataset and an EPC (Electronic Product Code) dataset. Both datasets are designed to evaluate classification performance under different cross-validation methodologies.

**2. Dataset Overview**

The dataset includes two types of cross-validation datasets:

A **Feature Dataset** containing various features for classification tasks.

An **EPC Dataset** which may include data related to Electronic Product Codes, used for specific classification or tracking purposes in supply chain or inventory management.

**3. Cross-Validation**

**File Naming Convention**

Filenames without the prefix CVB are associated with the first cross-validation method.

Filenames with the prefix CVB are associated with the second cross-validation method.

**4. Details of Cross-Validation Methods**

1. First Cross-Validation Method

The first cross-validation method focuses on distinguishing the classification performance across different label categories. This method requires classification to be performed under the same distance, comparing different category labels without the presence of a single true label among many false ones.

1. Second Cross-Validation Method

The second cross-validation method is tailored for scenarios where a specific label category contains only one true label, with the rest being false. For this method, the classification task involves distinguishing between the feature data of the second category at a certain distance d and the first 29 rows of the feature data of the first category at the same distance d.

Filenames containing specific codes correspond to the following label categories:

* WYUAN7017: Guangzhou Wang yuan Electronics Alien7017
* WYUAN9662: Guangzhou Wang yuan Electronics Alien 9662
* WYUAN9640: Guangzhou Wang yuan Electronics Alien 9640
* QB9662: Shenzhen Qibao Technology Alien 9662
* QB9654: Shenzhen Qibao Technology Alien 9654
* LJY9662: Nanjing Lejay Technology Alien 9662
* LJY9654: Nanjing Lejay Technology Alien 9654

**4. Feature Count**

Data files prefixed with the following indicate the number of features they contain:

* SevenFeats: 7 features
* TwentyeightFeats: 28 features
* OnehundredandfourFeats: 104 features

**5. Usage Guidelines**

* Before utilizing this dataset, ensure you understand the two cross-validation methods and their representation within the dataset.
* The label category codes in the filenames allow for quick identification of the category to which the data belongs.
* The prefixes indicating the number of features help you understand the dimensionality of features included in each data file.
* For the first method, analyze the classification performance by comparing datasets with the same distance but different category labels.
* For the second method, focus on the classification task where datasets from identical category at the same distance are compared, with emphasis on the first 29 rows of the second category's feature data.

**6. Notes**

* Please ensure compliance with the relevant data usage protocols and copyright statements before using the dataset.
* Should you encounter any issues during the use of the dataset, please contact us promptly.
* Contact the corresponding author for any inquiries or issues related to the dataset.