**60.Bach Chorales Data Set**

1. 数据库网址

https://archive.ics.uci.edu/ml/datasets/Bach+Chorales

2. 数据库描述

【1.[数据集名称]数据集由[机构名或人名]采集；】The data used in our experiments were collected by Mainous and Ottman edition.Mainous, Frank D. and Robert W. Ottman, eds.【2.用于[什么实验目的]】We used to examining the prediction and generation of music using a multiple viewpoint system, a collection of independent views of the musical surface each of which models a specific type of musical phenomena. Both the general style and a particular piece are modeled using dual short-term and long-term theories, and the model is created using machine learning techniques on a corpus of musical examples.【3】Number of Instances is 100 Chorales, each with ~45 events. Number of Attributes is 6 (nominal) per event, such as :(a) start-time, measured in 16th notes from chorale beginning (time 0) ;(b) pitch, MIDI number (60 = C4, 61 = C#4, 72 = C5, etc.) ;(c) duration, measured in 16th notes;(d) key signature, number of sharps or flats, positive if key signature has sharps, negative if key signature has flats ;(e) time signature, in 16th notes per bar;(f) fermata, true or false depending on whether event is under a fermata. 【4】The database has 100 samples, respectively belong to training with 80 samples and testing with 20 samples. The categories can be as shown in Table 1.

Table 1 Category Distribution of Bach Chorales Dataset [根据数据库绘制]

|  |  |  |  |
| --- | --- | --- | --- |
| Categories | optdigits.tra | optdigits.tes | Total Number of Samples |
| Short-term | 40 | 10 | 50 |
| Long-term | 40 | 10 | 50 |
| Total number of samples in total | 80 | 20 | 100 |

3. 精简描述

The Data in our experiment were collected by Mainous and Ottman edition.Mainous, Frank D. and Robert W. Ottman, eds. The dataset includes have 100 samples, which used to to examining the prediction and generation of music using a multiple viewpoint system, a collection of independent views of the musical surface each of which models a specific type of musical phenomena. Through which, we divided the dataset into two part, training data set with 80 samples and forecasting data set with 20 samples.