**109.Energy efficiency**

1. 数据库网址

http://archive.ics.uci.edu/ml/datasets/Energy+efficiency

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 E. Alpaydin, C. Kaynak, from Department of Computer Engineering,Bogazici University at July,1998.【2.用于[什么实验目的]】We used preprocessing programs made available by NIST to extract normalized bitmaps of handwritten digits from a preprinted form.【3】

【4】The database has 5620 samples, respectively belong to optdigits.tra with 3823 samples and optidigits.tes with 1797 samples. The categories of network system include seven categories, as shown in Table 1.

Table 1 Category Distribution of Network System [根据数据库绘制]

|  |  |  |  |
| --- | --- | --- | --- |
| Invasion Categories | optdigits.tra | optdigits.tes | Total Number of Samples |
|  |  |  |  |
|  |  |  |  |
| Total number of samples in total |  |  | 768 |

|  |  |
| --- | --- |
| **Abstract**: This study looked into assessing the heating load and cooling load requirements of buildings (that is, energy efficiency) as a function of building parameters. |  |

**Source:**

The dataset was created by Angeliki Xifara (angxifara **'@'** gmail.com, Civil/Structural Engineer) and was processed by Athanasios Tsanas (tsanasthanasis **'@'** gmail.com, Oxford Centre for Industrial and Applied Mathematics, University of Oxford, UK).

**Data Set Information:**

We perform energy analysis using 12 different building shapes simulated in Ecotect. The buildings differ with respect to the glazing area, the glazing area distribution, and the orientation, amongst other parameters. We simulate various settings as functions of the afore-mentioned characteristics to obtain 768 building shapes. The dataset comprises 768 samples and 8 features, aiming to predict two real valued responses. It can also be used as a multi-class classification problem if the response is rounded to the nearest integer.

**Attribute Information:**

The dataset contains eight attributes (or features, denoted by X1...X8) and two responses (or outcomes, denoted by y1 and y2). The aim is to use the eight features to predict each of the two responses.   
  
Specifically:   
X1 Relative Compactness   
X2 Surface Area   
X3 Wall Area   
X4 Roof Area   
X5 Overall Height   
X6 Orientation   
X7 Glazing Area   
X8 Glazing Area Distribution   
y1 Heating Load   
y2 Cooling Load

**Relevant Papers:**

A. Tsanas, A. Xifara: 'Accurate quantitative estimation of energy performance of residential buildings using statistical machine learning tools', Energy and Buildings, Vol. 49, pp. 560-567, 2012