**86.Housing**

1. 数据库网址

https://archive.ics.uci.edu/ml/datasets/Housing

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 |  |  | 506 |

**Source:**

Origin:   
  
This dataset was taken from the StatLib library which is maintained at Carnegie Mellon University.   
  
Creator:   
  
Harrison, D. and Rubinfeld, D.L.   
'Hedonic prices and the demand for clean air', J. Environ. Economics & Management, vol.5, 81-102, 1978.

**Data Set Information:**

Concerns housing values in suburbs of Boston.

**Attribute Information:**

1. CRIM: per capita crime rate by town   
2. ZN: proportion of residential land zoned for lots over 25,000 sq.ft.   
3. INDUS: proportion of non-retail business acres per town   
4. CHAS: Charles River dummy variable (= 1 if tract bounds river; 0 otherwise)   
5. NOX: nitric oxides concentration (parts per 10 million)   
6. RM: average number of rooms per dwelling   
7. AGE: proportion of owner-occupied units built prior to 1940   
8. DIS: weighted distances to five Boston employment centres   
9. RAD: index of accessibility to radial highways   
10. TAX: full-value property-tax rate per $10,000   
11. PTRATIO: pupil-teacher ratio by town   
12. B: 1000(Bk - 0.63)^2 where Bk is the proportion of blacks by town   
13. LSTAT: % lower status of the population   
14. MEDV: Median value of owner-occupied homes in $1000's

2. Sources:

(a) Origin: This dataset was taken from the StatLib library which is

maintained at Carnegie Mellon University.

(b) Creator: Harrison, D. and Rubinfeld, D.L. 'Hedonic prices and the

demand for clean air', J. Environ. Economics & Management,

vol.5, 81-102, 1978.

(c) Date: July 7, 1993

4. Relevant Information:

Concerns housing values in suburbs of Boston.

5. Number of Instances: 506

6. Number of Attributes: 13 continuous attributes (including "class"

attribute "MEDV"), 1 binary-valued attribute.

7. Attribute Information:

1. CRIM per capita crime rate by town

2. ZN proportion of residential land zoned for lots over

25,000 sq.ft.

3. INDUS proportion of non-retail business acres per town

4. CHAS Charles River dummy variable (= 1 if tract bounds

river; 0 otherwise)

5. NOX nitric oxides concentration (parts per 10 million)

6. RM average number of rooms per dwelling

7. AGE proportion of owner-occupied units built prior to 1940

8. DIS weighted distances to five Boston employment centres

9. RAD index of accessibility to radial highways

10. TAX full-value property-tax rate per $10,000

11. PTRATIO pupil-teacher ratio by town

12. B 1000(Bk - 0.63)^2 where Bk is the proportion of blacks

by town

13. LSTAT % lower status of the population

14. MEDV Median value of owner-occupied homes in $1000's

8. Missing Attribute Values: None.