**93.Entree Chicago Recommendation（芝加哥餐厅）**

1. 数据库网址

http://archive.ics.uci.edu/ml/datasets/Entree+Chicago+Recommendation+Data

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 |  |  | 50672 |

|  |  |
| --- | --- |
| **Abstract**: This data contains a record of user interactions with the Entree Chicago restaurant recommendation system. |  |

**Source:**

Original Owner and Donor:   
  
Robin Burke   
University of California, Irvine   
Department of Information and Computer Science   
Irvine, CA 92697   
  
Now here:   
<http://josquin.cti.depaul.edu/~rburke/>

**Data Set Information:**

This data records interactions with Entree Chicago restaurant recommendation system (originally [[Web Link]](http://infolab.cs.uchicago.edu/entree)) from September, 1996 to April, 1999. The data is organized into files roughly spanning a quarter year -- with Q3 1996 and Q2 1999 each only containing one month.   
  
Each line in a session file represents a session of user interaction with the system. The (tab-separated) fields are as follows:   
  
Date, IP, Entry point, Rated restaurant1, ..., Rated restaurantN, End point   
  
1. Entry point:   
  
Users can use a restaurant from any city as a entry point, but they always get recommendations for Chicago restaurants. The entry point therefore draws from a larger universe of restaurants than the rest of the data.   
Entry points have the form nnnX, where nnn is a numeric restaurant ID and X is a character A-H that encodes the city.   
  
A = Atlanta   
B = Boston   
C = Chicago   
D = Los Angeles   
E = New Orleans   
F = New York   
G = San Francisco   
H = Washington DC   
  
2. Rated Restaurant:   
  
These are all Chicago restaurants.   
These entries have the form nnnX, where nnn is a numeric restaurant ID and X is a character L-T that encodes the navigation operation.   
  
L = browse (move from one restaurant in a list of recommendations to another)   
M = cheaper (search for a restaurant like this one, but cheaper)   
N = nicer ( " " , but nicer)   
O = closer (unused in the production version of the system)   
P = more traditional (search for a restaurant like this, but serving more traditional cuisine)   
Q = more creative (search for a restaurant serving more creative cuisine)   
R = more lively (search for a restaurant with a livelier atmosphere)   
S = quieter (search for a restaurant with a quieter atmosphere)   
T = change cuisine (search for a restaurant like this, but serving a different kind of food) Note that with this tweak, we would ideally like to know what cuisine the user wanted to change to, but this information was not recorded.   
  
3. End point:   
  
Just the numeric id for the (Chicago) restaurant that the user saw last. In our experiments, we are assuming that this was a good suggestion, but it is also possible that user just gives up.   
Some potentially useful data is missing. In many cases, we don't know the starting point because the user input a set of selection criteria (such as "inexpensive traditional Mexican") using a form submission, rather than starting from a known restaurant. These queries were not recorded. This is denoted by a 0 in the entry point field. Some sessions do not have a known end point. This is marked by -1 in the end point field.   
  
  
In addition to the user's interactions, there is also data linking the restaurant ID with its name and features such as "fabulous wine lists", "good for younger kids", and "Ethopian" cuisine. This data is stored by city (e.g. Atlanta, Boston, etc.) and is in the following format:   
  
restaurant id [tab] restaurant name [tab] restaurant features (3 digits ids separated by spaces)