UCI Machine Learning Repository
Center for Machine Learning and Intelligent Systems

About Citation Policy Donate a Data Set Contact

Search
Repository Web
Google™

View ALL Data Sets

Blood Transfusion Service Center Data Set

Download: Data Folder, Data Set Description

Abstract: Data taken from the Blood Transfusion Service Center in Hsin-Chu City in Taiwan -- this is a classification problem.

Data Set Characteristics:	Multivariate	**Number of Instances:**	748	**Area:**	Business
Attribute Characteristics:	Real	**Number of Attributes:**	5	**Date Donated**	2008-10-03
Associated Tasks:	Classification	**Missing Values?**	N/A	**Number of Web Hits:**	356327

Source:

Original Owner and Donor
Prof. I-Cheng Yeh
Department of Information Management
Chung-Hua University,
Hsin Chu, Taiwan 30067, R.O.C.
e-mail:icyeh '@' chu.edu.tw
TEL:886-3-5186511

Date Donated: October 3, 2008

Data Set Information:

To demonstrate the RFMTC marketing model (a modified version of RFM), this study adopted the donor database of Blood Transfusion Service Center in Hsin-Chu City in Taiwan. The center passes their blood transfusion service bus to one university in Hsin-Chu City to gather blood donated about every three months. To build a FRMTC model, we selected 748 donors at random from the donor database. These 748 donor data, each one included R (Recency - months since last donation), F (Frequency - total number of donation), M (Monetary - total blood donated in c.c.), T (Time - months since first donation), and a binary variable representing whether he/she donated blood in March 2007 (1 stand for donating blood; 0 stands for not donating blood).

Attribute Information:

Given is the variable name, variable type, the measurement unit and a brief description. The "Blood Transfusion Service Center" is a classification problem. The order of this listing corresponds to the order of numerals along the rows of the database.

R (Recency - months since last donation),
F (Frequency - total number of donation),
M (Monetary - total blood donated in c.c.),

T (Time - months since first donation), and
a binary variable representing whether he/she donated blood in March 2007 (1 stand for donating blood; 0 stands for not donating blood).

Table 1 shows the descriptive statistics of the data. We selected 500 data at random as the training set, and the rest 248 as the testing set.

Table 1. Descriptive statistics of the data

Variable	Data Type	Measurement	Description	min	max	mean	std
Recency	quantitative	Months	Input	0.03	74.4	9.74	8.07
Frequency	quantitative	Times	Input	1	50	5.51	5.84
Monetary	quantitative	c.c. blood	Input	250	12500	1378.68	1459.83
Time	quantitative	Months	Input	2.27	98.3	34.42	24.32
Whether he/she donated blood in March 2007	binary	1=yes 0=no	Output	0	1	1 (24%)	0 (76%)

Relevant Papers:

Yeh, I-Cheng, Yang, King-Jang, and Ting, Tao-Ming, "Knowledge discovery on RFM model using Bernoulli sequence," Expert Systems with Applications, 2008, [Web Link]

Citation Request:

NOTE: Reuse of this database is unlimited with retention of copyright notice for Prof. I-Cheng Yeh and the following published paper:

Yeh, I-Cheng, Yang, King-Jang, and Ting, Tao-Ming, "Knowledge discovery on RFM model using Bernoulli sequence, "Expert Systems with Applications, 2008,
[Web Link]

Supported By: Rexa.info In Collaboration With: Rexa.info

About || Citation Policy || Donation Policy || Contact || CML