Statlog (Australian Credit Approval) Data Set

Download: Data Folder, Data Set Description

Abstract: This file concerns credit card applications. This database exists elsewhere in the repository (Credit Screening Database) in a slightly different form

Data Set Characteristics:	Multivariate	Number of Instances:	690	Area:	Financial
Attribute Characteristics:	Categorical, Integer, Real	Number of Attributes:	14	Date Donated	N/A
Associated Tasks:	Classification	Missing Values?	Yes	Number of Web Hits:	89629

Source:

(confidential)

Submitted by quinlan '@' cs.su.oz.au

Data Set Information:

This file concerns credit card applications. All attribute names and values have been changed to meaningless symbols to protect confidentiality of the data.

This dataset is interesting because there is a good mix of attributes -- continuous, nominal with small numbers of values, and nominal with larger numbers of values. There are also a few missing values.

Attribute Information:

There are 6 numerical and 8 categorical attributes. The labels have been changed for the convenience of the statistical algorithms. For example, attribute 4 originally had 3 labels p,g,gg and these have been changed to labels 1,2,3.

A1: 0,1 CATEGORICAL (formerly: a,b)
A2: continuous.
A3: continuous.
A4: 1,2,3 CATEGORICAL (formerly: p,g,gg)
A5: 1, 2,3,4,5, 6,7,8,9,10,11,12,13,14 CATEGORICAL (formerly: ff,d,i,k,j,aa,m,c,w, e, q, r,cc, x)
A6: 1, 2,3, 4,5,6,7,8,9 CATEGORICAL (formerly: ff,dd,j,bb,v,n,o,h,z)
A7: continuous.
A8: 1, 0 CATEGORICAL (formerly: t, f)
A9: 1, 0 CATEGORICAL (formerly: t, f)

A10: continuous.
A11: 1, 0 CATEGORICAL (formerly t, f)
A12: 1, 2, 3 CATEGORICAL (formerly: s, g, p)
A13: continuous.
A14: continuous.
A15: 1,2 class attribute (formerly: +,-)

Relevant Papers:

Ross Quinlan. "Simplifying decision trees", Int J Man-Machine Studies 27, Dec 1987, pp. 221-234.
[Web Link]

Ross Quinlan. "C4.5: Programs for Machine Learning", Morgan Kaufmann, Oct 1992
[Web Link]

Papers That Cite This Data Set[1]:

Jeroen Eggermont and Joost N. Kok and Walter A. Kosters. Genetic Programming for data classification: partitioning the search space. SAC. 2004. [View Context].

Bart Hamers and J. A. K Suykens. Coupled Transductive Ensemble Learning of Kernel Models. Bart De Moor. 2003. [View Context].

Xiaoming Huo. FBP: A Frontier-Based Tree-Pruning Algorithm. Seoung Bum Kim. 2002. [View Context].

Endre Boros and Peter Hammer and Toshihide Ibaraki and Alexander Kogan and Eddy Mayoraz and Ilya B. Muchnik. An Implementation of Logical Analysis of Data. IEEE Trans. Knowl. Data Eng, 12. 2000. [View Context].

Mark A. Hall. Department of Computer Science Hamilton, NewZealand Correlation-based Feature Selection for Machine Learning. Doctor of Philosophy at The University of Waikato. 1999. [View Context].

Rudy Setiono and Huan Liu. NeuroLinear: From neural networks to oblique decision rules. Neurocomputing, 17. 1997. [View Context].

Adil M. Bagirov and Alex Rubinov and A. N. Soukhojak and John Yearwood. Unsupervised and supervised data classification via nonsmooth and global optimization. School of Information Technology and Mathematical Sciences, The University of Ballarat. [View Context].

Wl/odzisl/aw Duch and Karol Grudzi nski and Grzegorz Stawski. SYMBOLIC FEATURES IN NEURAL NETWORKS. Department of Computer Methods, Nicolaus Copernicus University. [View Context].

Hussein A. Abbass. Pareto Neuro-Evolution: Constructing Ensemble of Neural Networks Using Multi-objective Optimization. Artificial Life and Adaptive Robotics (A.L.A.R.) Lab, School of Information Technology and Electrical Engineering, Australian Defence Force Academy. [View Context].

Krzysztof Grabczewski and Wl/odzisl/aw Duch. THE SEPARABILITY OF SPLIT VALUE CRITERION. Department of Computer Methods, Nicolaus Copernicus University. [View Context].

Bart Baesens and Stijn Viaene and Tony Van Gestel and J. A. K Suykens and Guido Dedene and Bart De Moor and Jan Vanthienen and Katholieke Universiteit Leuven. An Empirical Assessment of Kernel Type Performance for Least Squares Support Vector Machine Classifiers. Dept. Applied Economic Sciences. [View Context].

Citation Request:

Please refer to the Machine Learning Repository's citation policy