

HEURIMIND Lab

Heuristic Models For Intelligent Decision Systems

DataSets repository and Results

Low quality datasets

Datasets with fuzzy values

Used in the paper: "Feature subset selection Filter-Wrapper based on low quality data"

Below you can find all the Classification datasets available. For each dataset, it is shown its name, attributes (the table details the number of Real/Nominal attributes in the data), number of instances, and classes (number of possible values of the output variable). In addition, the table shows if the corresponding dataset has missing values, interval values, fuzzy values and/or noise (the table shows the percentage of low quality values). The table allows to download each dataset (inside a ZIP file). Additionally, it is possible to obtain the dataset already partitioned, by means of a 5-folds cross validation procedure.

 Datasets with normalized values into the interval [0,1].

 Missing values in attribute 6. This attribute contains 3% of missing values. The overall percentage of missing values is 0.25%.

Name	#Attributes (R/N)	#Examp.	#Class.	Missing values	Fuzzy values	Dataset	3x5fcv	
(AUS) Australian	14 (6/8)	690	2	No	Yes (10%)			
(BCW) Wisconsin Br. Cancer	9 (9/0)	699	2	Yes (3%)*	Yes (10%)			
(GER) German credit	24 (24/0)	1000	2	No	Yes (10%)			
(GLA) Glass identification	9 (9/0)	214	7	No	Yes (10%)			
(HEA) Statlog Heart	13 (13/0)	270	2	No	Yes (10%)			
(ION) Ionosphere	34 (34/0)	351	2	No	Yes (10%)			
(IRP) Iris Plant	4 (4/0)	150	3	No	Yes (10%)			
(PIM) Pima Indian Diabetes	8 (8/0)	768	2	No	Yes (10%)			
(SON) Sonar	60 (60/0)	208	2	No	Yes (10%)			
(SPE) Spectft heart	44 (44/0)	267	2	No	Yes (10%)			
(VEH) Vehicle	18 (18/0)	846	4	No	Yes (10%)			
(WDC) Wisconsin Diag. Br. Cancer	31 (31/0)	569	2	No	Yes (10%)			
(WIN) Wine	13 (13/0)	178	3	No	Yes (10%)			
(ZOO) Zoo	16 (1/16)	101	7	No	Yes (10%)			

All datasets




[info.pdf](#)