




[About](#) [Citation Policy](#) [Donate a Data Set](#)
[Contact](#)

☒ Repository ☐ Web 

[View ALL Data Sets](#)

Arrhythmia Data Set

Download: [Data Folder](#), [Data Set Description](#)



Abstract: Distinguish between the presence and absence of cardiac arrhythmia and classify it in one of the 16 groups.

Data Set Characteristics:	Multivariate	Number of Instances:	452	Area:	Life
Attribute Characteristics:	Categorical, Integer, Real	Number of Attributes:	279	Date Donated	1998-01-01
Associated Tasks:	Classification	Missing Values?	Yes	Number of Web Hits:	299230

Source:

Original Owners of Database:

1. H. Altay Guvenir, PhD.,
Bilkent University,
Department of Computer Engineering and Information Science,
06533 Ankara, Turkey
Phone: +90 (312) 266 4133
Email: guvenir '@' cs.bilkent.edu.tr

2. Burak Acar, M.S.,
Bilkent University,
EE Eng. Dept.
06533 Ankara, Turkey
Email: buraka '@' ee.bilkent.edu.tr

3. Haldun Muderrisoglu, M.D., Ph.D.,
Baskent University,
School of Medicine
Ankara, Turkey

Donor:

H. Altay Guvenir
Bilkent University,
Department of Computer Engineering and Information Science,
06533 Ankara, Turkey
Phone: +90 (312) 266 4133
Email: guvenir '@' cs.bilkent.edu.tr

Data Set Information:

This database contains 279 attributes, 206 of which are linear valued and the rest are nominal.

Concerning the study of H. Altay Guvenir: "The aim is to distinguish between the presence and absence of cardiac arrhythmia and to classify it in one of the 16 groups. Class 01 refers to 'normal' ECG classes 02 to 15 refers to different classes of arrhythmia and class 16 refers to the rest of unclassified ones. For the time being, there exists a computer program that makes such a classification. However there are differences between the cardiologist's and the programs classification. Taking the cardiologist's as a gold standard we aim to minimise this difference by means of machine learning tools."

The names and id numbers of the patients were recently removed from the database.

Attribute Information:

-- Complete attribute documentation:

- 1 Age: Age in years , linear
- 2 Sex: Sex (0 = male; 1 = female) , nominal
- 3 Height: Height in centimeters , linear
- 4 Weight: Weight in kilograms , linear
- 5 QRS duration: Average of QRS duration in msec., linear
- 6 P-R interval: Average duration between onset of P and Q waves in msec., linear
- 7 Q-T interval: Average duration between onset of Q and offset of T waves in msec., linear
- 8 T interval: Average duration of T wave in msec., linear
- 9 P interval: Average duration of P wave in msec., linear
- Vector angles in degrees on front plane of:, linear
- 10 QRS
- 11 T
- 12 P
- 13 QRST
- 14 J

- 15 Heart rate: Number of heart beats per minute ,linear

- Of channel DI:
- Average width, in msec., of: linear
- 16 Q wave
- 17 R wave
- 18 S wave
- 19 R' wave, small peak just after R
- 20 S' wave

- 21 Number of intrinsic deflections, linear

- 22 Existence of ragged R wave, nominal
- 23 Existence of diphasic derivation of R wave, nominal
- 24 Existence of ragged P wave, nominal
- 25 Existence of diphasic derivation of P wave, nominal
- 26 Existence of ragged T wave, nominal
- 27 Existence of diphasic derivation of T wave, nominal

- Of channel DII:
- 28 .. 39 (similar to 16 .. 27 of channel DI)
- Of channels DIII:
- 40 .. 51
- Of channel AVR:
- 52 .. 63

Of channel AVL:

64 .. 75

Of channel AVF:

76 .. 87

Of channel V1:

88 .. 99

Of channel V2:

100 .. 111

Of channel V3:

112 .. 123

Of channel V4:

124 .. 135

Of channel V5:

136 .. 147

Of channel V6:

148 .. 159

Of channel DI:

Amplitude , * 0.1 milivolt, of

160 JJ wave, linear

161 Q wave, linear

162 R wave, linear

163 S wave, linear

164 R' wave, linear

165 S' wave, linear

166 P wave, linear

167 T wave, linear

168 QRSA , Sum of areas of all segments divided by 10, ($\text{Area} = \text{width} * \text{height} / 2$), linear

169 QRSTA = QRSA + 0.5 * width of T wave * 0.1 * height of T wave. (If T is diphasic then the bigger segment is considered), linear

Of channel DII:

170 .. 179

Of channel DIII:

180 .. 189

Of channel AVR:

190 .. 199

Of channel AVL:

200 .. 209

Of channel AVF:

210 .. 219

Of channel V1:

220 .. 229

Of channel V2:

230 .. 239

Of channel V3:

240 .. 249

Of channel V4:

250 .. 259

Of channel V5:

260 .. 269

Of channel V6:

270 .. 279

Relevant Papers:

H. Altay Guvenir, Burak Acar, Gulsen Demiroz, Ayhan Cekin "A Supervised Machine Learning Algorithm for

Arrhythmia Analysis." Proceedings of the Computers in Cardiology Conference, Lund, Sweden, 1997.
[\[Web Link\]](#)

Papers That Cite This Data Set¹:



Krista Lagus and Esa Alhoniemi and Jeremias Seppa and Antti Honkela and Arno Wagner. [INDEPENDENT VARIABLE GROUP ANALYSIS IN LEARNING COMPACT REPRESENTATIONS FOR DATA](#). Neural Networks Research Centre, Helsinki University of Technology. [\[View Context\]](#).

Gisele L. Pappa and Alex Alves Freitas and Celso A A Kaestner. [AMultiobjective Genetic Algorithm for Attribute Selection](#). Computing Laboratory Pontificia Universidade Catolica do Parana University of Kent at Canterbury. [\[View Context\]](#).

Shay Cohen and Eytan Ruppin and Gideon Dror. [Feature Selection Based on the Shapley Value](#). School of Computer Sciences Tel-Aviv University. [\[View Context\]](#).

Citation Request:

Please refer to the Machine Learning Repository's [citation policy](#)

[1] Papers were automatically harvested and associated with this data set, in collaboration with [Rexa.info](#)

Supported By:



In Collaboration With:



[About](#) || [Citation Policy](#) || [Donation Policy](#) || [Contact](#) || [CML](#)