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 cardio log’s and the programs classification. Taking the cardio log’s as a gold standard we aim to minimize this difference by means of machine learning tools.

***Number of Instances: 452***

***Number of Attributes: 279*** *(206 of which are linear valued and the rest are nominal)*

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **S.No.** | **Feature Name** | **Feature Unit** | **Feature Type** | **Feature Desc** | **Action Taken** |
| 1 | Age | Years | Continuous | Age of subject | MinMax scaler |
| 2 | Gender | Boolean  0=M and 1=F | Nominal | Gender of subject | One hot encoder |
| 3 | Height | Centimeters | Continuous | Height of subject | Fix two outliers then MinMax scaler |
| 4 | Weight | Kilograms | Continuous | Weight of subject | Fix outliers then MinMax Scaler |
| 5 | QRS Duration | Milliseconds | Continuous | Average duration of QRS | Fix outliers then MinMax Scaler |
| 6 | P-R interval | Milliseconds | Continuous | Average duration between onset of P and Q waves | Fix outliers then MinMax Scaler |
| 7 | Q-T interval | Milliseconds | Continuous | Average duration between onset of Q and offset of T waves | Fix outliers then MinMax Scaler |
| 8 | T interval | Milliseconds | Continuous | Average duration of T wave | Fix outliers then MinMax Scaler |
| 9 | P interval | Milliseconds | Continuous | Average duration of P wave | Fix outliers then MinMax Scaler |
| 10 | QRS vector angle | Degrees | Discrete | Vector angle on front plane | Fix outliers then MinMax Scaler |
| 11 | T vector angle | Degrees | Discrete | Vector angle on front plane | Contains Missing values |
| 12 | P vector angle | Degrees | Discrete | Vector angle on front plane | Contains Missing values |
| 13 | QRST vector angle | Degrees | Discrete | Vector angle on front plane | Contains Missing values |
| 14 | J vector angle | Degrees | Discrete | Vector angle on front plane | Contains Missing values |
| 15 | Heart Rate | BPM | Discrete | Number of heart beats per minute | Contains Missing values |
| 16 | DI Q wave width | Milliseconds | Discrete | Average Q wave width DI Channel |  |
| 17 | DI R wave width | Milliseconds | Discrete | Average R wave width DI Channel |  |
| 18 | DI S wave width | Milliseconds | Discrete | Average S wave width DI Channel |  |
| 19 | DI R1 wave width | Milliseconds | Discrete | DI channel -- Average width of small wave after R | All records are 0 after fixing. |
| 20 | DI S1 wave width | Milliseconds | Discrete | DI channel -- Average width of small wave after S | All records are 0 after fixing. |
| 21 | DI Number of intrinsic deflections |  | Discrete | Number of intrinsic deflections |  |
| 22 | DI Ragged R wave | Boolean | Nominal | Existence of ragged R wave | For now, using OHE,  however, in the later phase will implement the Target Class Level Mean Encoding. |
| 23 | DI R wave diphasic derivation | Boolean | Nominal | Existence of diphasic derivation of R wave |
| 24 | DI Ragged P wave | Boolean | Nominal | Existence of ragged P wave |
| 25 | DI P Wave diphasic derivation | Boolean | Nominal | Existence of diphasic derivation of P wave |
| 26 | DI Ragged T wave | Boolean | Nominal | Existence of ragged T wave |
| 27 | DI T wave diphasic derivation | Boolean | Nominal | Existence of diphasic derivation of T wave |
|  | 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 |  |  |  |  |
|  | DI Amplitude JJ Wave |  | Continuous | \* 0.1 millivolt |  |
|  | DI Amplitude Q Wave |  | Continuous | \* 0.1 millivolt |  |
|  | DI Amplitude R Wave |  | Continuous | \* 0.1 millivolt |  |
|  | DI Amplitude S Wave |  | Continuous | \* 0.1 millivolt |  |
|  | DI Amplitude R1 Wave |  | Continuous | \* 0.1 millivolt |  |
|  | DI Amplitude S1 Wave |  | Continuous | \* 0.1 millivolt |  |
|  | DI Amplitude P Wave |  | Continuous | \* 0.1 millivolt |  |
|  | DI Amplitude T Wave |  | Continuous | \* 0.1 millivolt |  |
|  | DI\_QRSA |  | Continuous | Sum of areas of all segments divided by 10  (Area= width \* height / 2) |  |
|  | DI\_QRSTA |  | Continuous | QRSA + 0.5 \* width of T wave \* 0.1 \* height of T wave.  (If T is diphasic then the bigger segment is considered) |  |
|  | 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 |  |  |  |  |
|  | Diagnosis Class |  | Nominal | 16 classes |  |