

---

---

# Detecção de Arritmias Cardíacas com WiSARD

Ygor Canalli

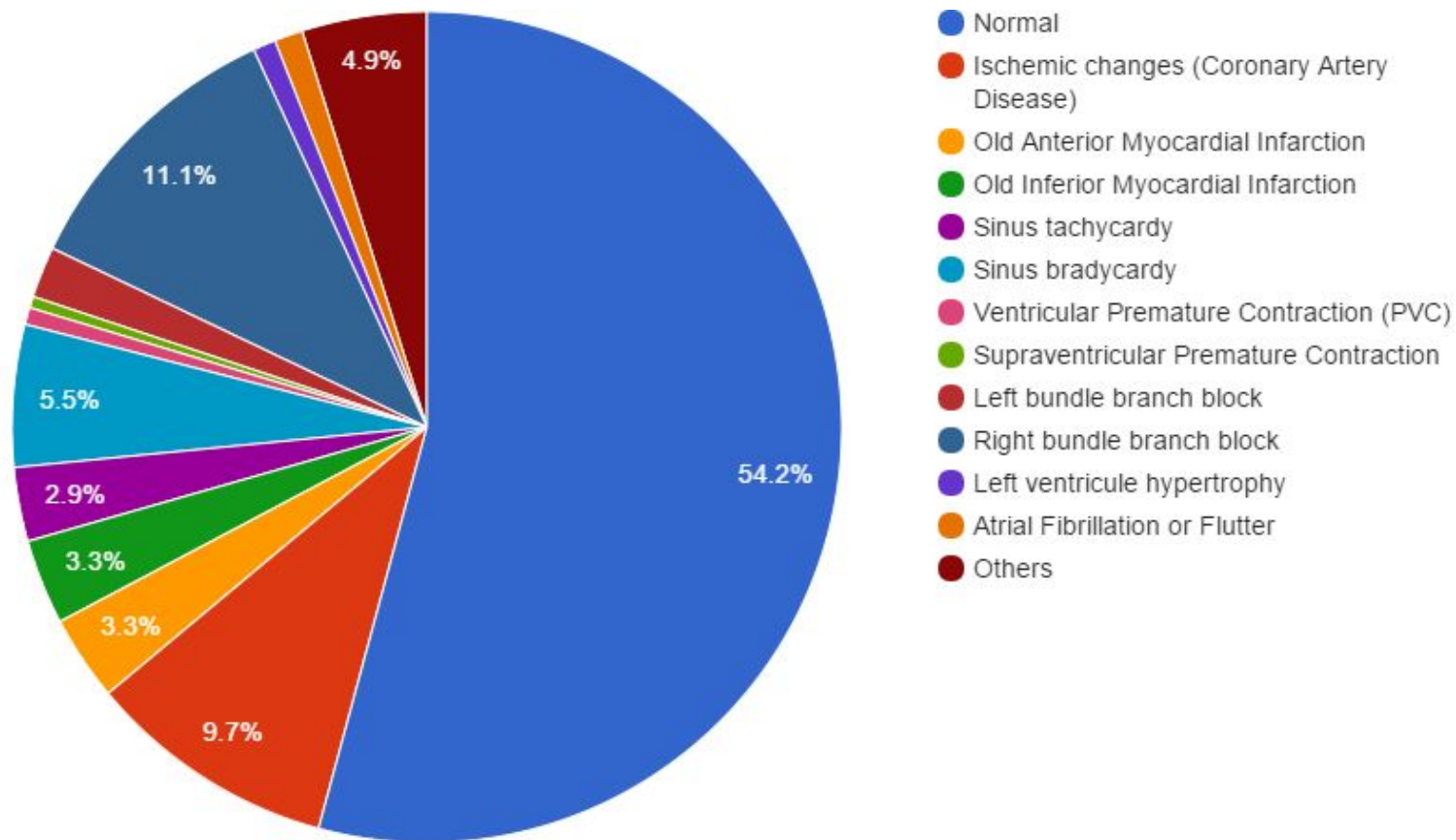
---

---

# Dataset

- *Dataset Arrhythmia* disponível no Repositório UCI Machine Learning
- Consiste num problema de classificação
  - Classe 1: Sem Arritmia
  - Classes 2-15: Tipos distintos de Arritmia (Ex.: Contração Ventricular Pré-matura, Taquicardia, Braquicardia)
  - Classe 16: Arritmia não classificada
- Possui 452 Amostras
- Utilizado num trabalho de H. Altay Guvenir et. al.
  - Publicado 1997
  - Propôs o algoritmo VF15
  - Acurácia de 62%,

# Distribuição das classes

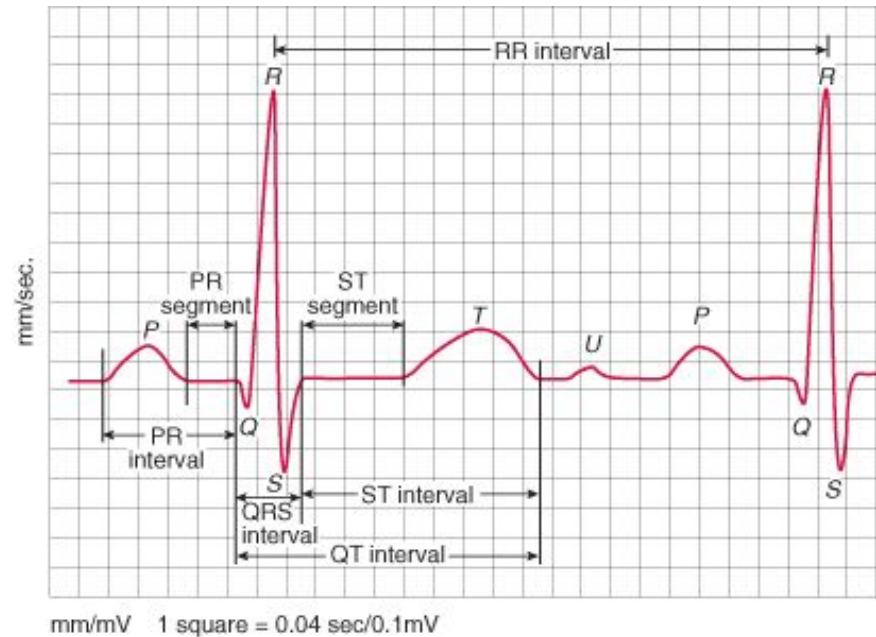


# Features

- 279 atributos
  - 206 contínuos
  - 73 categóricos
- Dados faltosos
  - 0,43% dos contínuos
  - 0% dos categóricos
- Dados Clínicos. Ex.:
  - Idade
  - Sexo
  - Peso

# Features

- Dados de Eletrocardiograma
  - Canais DII - DIII, AVL, AVR, AVF, V1 - V6
  - Informações extraídas das ondas. Ex.:
    - Frequência cardíaca
    - Largura da Onda R
    - Amplitude da Onda QQ
    - Intervalo P-R
    - Existencia de Derivação difásica da Onda R



Fonte: Life in The Fastlane

# Codificação dos Dados

- Categóricos: *One-Hot-Encoding*
  - Ex.: Classes A,B,C
  - A = 001
  - B = 010
  - C = 100
- Contínuos: Codificação termômetro
  - Ex.: Termômetro de tamanho 5 para intervalo de 0 a 100
  - $0 \leq x \leq 20$  = 00001
  - $20 < x \leq 40$  = 00011
  - $40 < x \leq 60$  = 00111
  - $60 < x \leq 80$  = 01111
  - $80 < x \leq 100$  = 11111

# Técnicas utilizadas

- WiSARD convencional
- *Bleaching*
  - Limite máximo de 15
  - Alguns casos prosseguiam indefinidamente empatados
- Separação por sexo
  - Treinamento de um WiSARD para cada sexo
  - Melhora considerável nos resultados
- Tratamento de dados faltosos
  - Contínuos: preencher com a média
  - Categóricos: preencher com a moda

# Configurações avaliadas

- Uso dos atributos
  - Apenas contínuos
  - Apenas categóricos
  - Todos
- Tamanho do termômetro
  - Entre 4 e 15 *bins*
- Tamanho de endereçamento
  - Baseado na quantidade de colunas da matriz codificada
  - 5 primeiros divisores (caso existam)
  - Ex.: 3090 *features*
    - [2, 3, 5, 6, 10]



# Resultados

- Experimentação utilizando *10-fold cross-validation*
  - Divide-se aleatoriamente o *dataset* em 10 *folds* (fatias)
  - Para cada *fold*
    - Treinar com os 9 demais
    - Utilizar a atual como teste
  - Calcular a acurácia média das 10 rodadas
- Resultados ao redor de 59% (em geral)
- Melhor configuração
  - Uso das *features*: apenas contínuas
  - Tamanho do termômetro: 14 *bins*
  - Tamanho do endereço: 4 bits
  - Acurácia: 0,624227

# Referências

- ❑ Arrhythmia Dataset. UCI Machine Learning Repository. Disponível em <https://archive.ics.uci.edu/ml/datasets/Arrhythmia>
- ❑ Güvenir, H., Burak Acar, and Gülsen Demiröz. "A supervised machine learning algorithm for arrhythmia analysis." *Computers in Cardiology 1997*. IEEE, 1997.
- ❑ The P Wave. Life in The Fastlane. Disponível em <http://lifeinthefastlane.com/ecg-library/basics/p-wave/>