

# Classificação de arritmias em sinais ECG

Eduardo Luz    Rensso Mora

Universidade Federal de Ouro Preto

06/10/2010

# Summary

- 1 Introdução
- 2 Teste não tendencioso
- 3 Método para análise
- 4 Conclusão

# Introdução

- Impacto dos registros escolhidos para treino.
  - Dados tendenciosos
  - Dados não tendenciosos

# Resultados encontrados na literatura

Tabela: Precisão dos métodos

Método	Classificador	Número de classes	Precisão
Método proposto por Mehmet Engin (2004)	Neuro-Fuzzy	4	98%
Método proposto por Guler e Ubeyli (2005)	combined NN	4	96.94%
Método proposto por Song et al (2005)	SVM	6	99.35%
Método proposto por Ye, Coimbra e Kumar (2010)	SVM	13	99.91%

# Base das dados não tendenciosa

- Teste mais 'justo' - Não usar mesmos pacientes para Treino e Teste!.

# Dados

- Banco de dados de arritmia do MIT-BIH
- Seis classes : batimentos tipo **N**, **R**, **A**, **V**, **F** e **f**.

## BD MIT-BIH

Registros

101

102

103

104

105

106

107

108

109

111

112

113

114

115

.

.

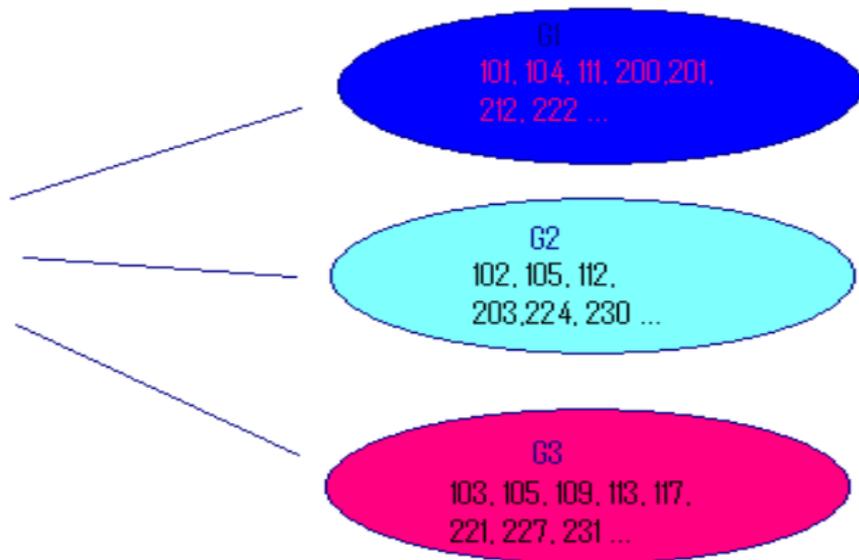


Figura: Grupos de registros

# Teste não tendencioso

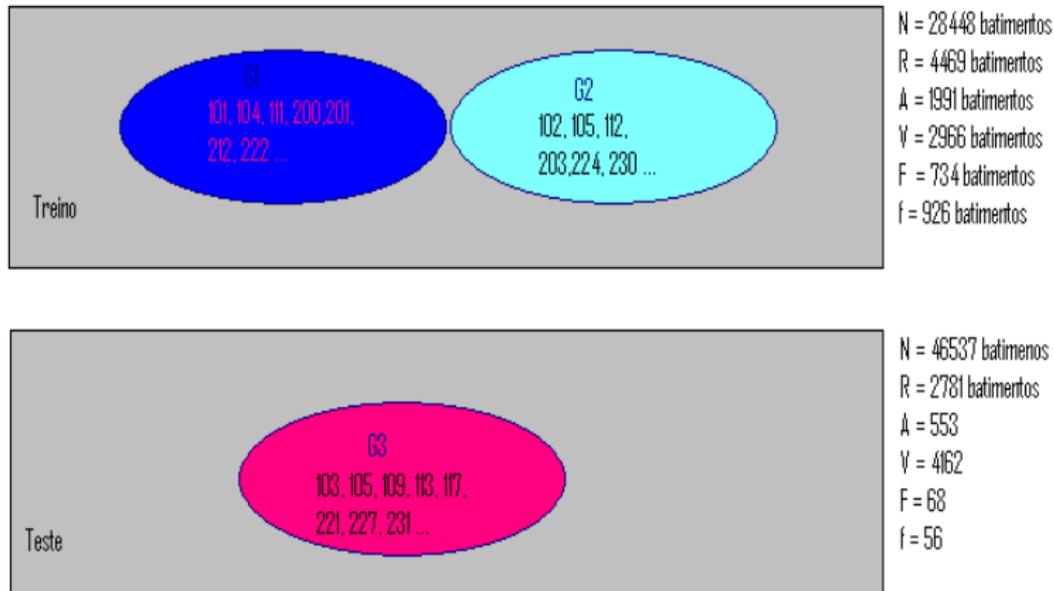


Figura: Grupos para treino e teste

## Características : método proposto por Song et al, 2005

- 19 características extraídas da transformada wavelet (coeficientes do sinal de detalhe dos níveis 4 a 7).
- 2 características extraídas de informação do ritmo cardíaco (intervalo RR).

$$valor = \begin{cases} 1 & \frac{K}{RR_i} \\ 2 & \frac{K}{RR_{i+1}} \end{cases} \quad (1)$$

# Primeiro teste : treino com MLP

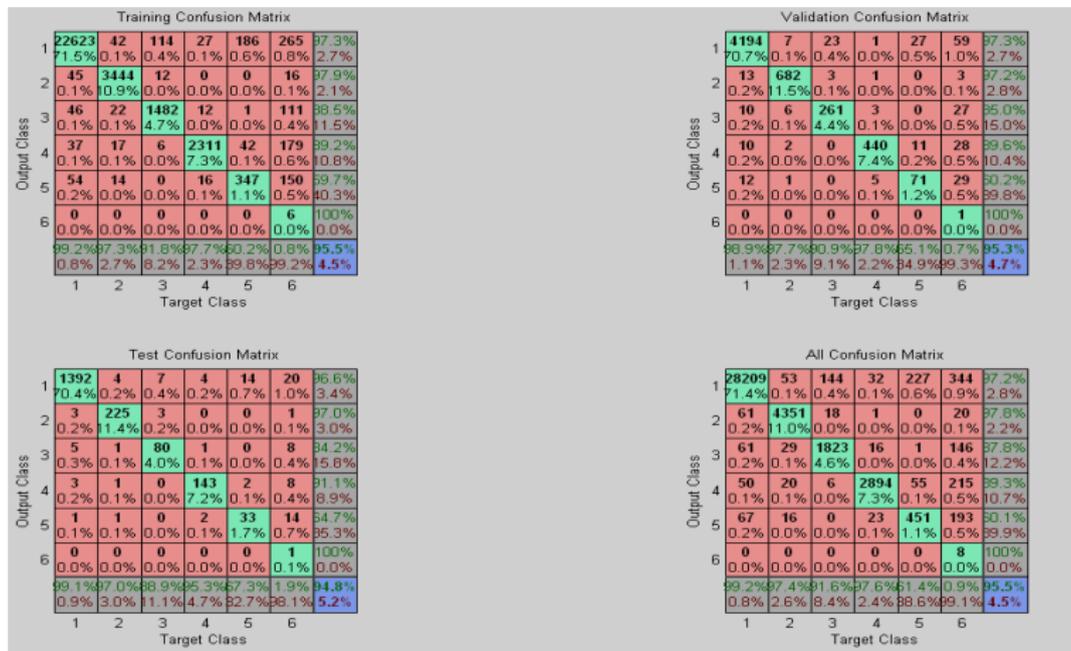


Figura: Matriz de confusão

# Primeiro teste : teste com MLP

Confusion Matrix

1	37392 69.0%	1594 2.9%	366 0.7%	240 0.4%	23 0.0%	21 0.0%	94.3% 5.7%
2	170 0.3%	1097 2.0%	5 0.0%	65 0.1%	0 0.0%	0 0.0%	82.0% 18.0%
3	1054 1.9%	1 0.0%	124 0.2%	423 0.8%	4 0.0%	2 0.0%	7.7% 92.3%
4	5414 10.0%	3 0.0%	37 0.1%	3360 6.2%	27 0.0%	19 0.0%	37.9% 62.1%
5	2507 4.6%	86 0.2%	21 0.0%	74 0.1%	14 0.0%	14 0.0%	0.5% 99.5%
6	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	NaN% NaN%
	80.3% 19.7%	39.4% 60.6%	22.4% 77.6%	80.7% 19.3%	20.6% 79.4%	0.0% 100%	77.5% 22.5%
	1	2	3	4	5	6	

Output Class

Target Class

Figura: Matriz de confusão

## Segundo teste : teste com SVM

**Tabela:** Matriz de confusão SVM :  $C = 10$  e  $\text{gamma} = 1/(2 * \text{num features})$

						(%)
36935	125	326	5196	952	3003	79.36
1812	935	0	9	1	24	33.62
412	1	59	38	5	38	10.66
253	192	293	3133	227	64	75.27
24	0	0	32	6	6	8.82
12	0	0	11	0	33	58.92
93.62	74.62	8.70	37.21	0.5	1.04	75.89

## Chazal, O'Dwyer, Reilly - 2004

- Padrão da Association for the Advanced of Medical Instrumentaion(AAMI)
- Poucos utilizam o padrão, tornando os resultados tendenciosos.
  - 2 trabalho em 2004, de acordo com Chazal, O'Dwyer, Reilly (2004)
  - 3 trabalho em 2009, de acordo com Ince, Kiranyaz e Gabbouj (2009)

## Análise para 3 configurações

- Conf. 1 : Garante que os registros de um pacientes vão estar presentes em ambos os grupos (teino e teste)
- Conf. 2 : Escolha aleatória dos batimentos para treino e teste.
- Conf. 3 : Garante que os registros de um paciente vão pertencer apenas a um grupo (treino ou teste).

# Configuração 1

A divisão é feita por paciente, de acordo com o trabalho de Ye, Coimbra e Kumar (2010):

- 13% dos batimentos do tipo N para treino, separados aleatoriamente.
- 40% dos batimentos do tipo R para treino, separados aleatoriamente.
- 40% dos batimentos do tipo A para treino, separados aleatoriamente.
- 40% dos batimentos do tipo V para treino, separados aleatoriamente.
- 50% dos batimentos do tipo F para treino, separados aleatoriamente.
- 50% dos batimentos do tipo f para treino, separados aleatoriamente.

## Configuração 2

Os batimentos de toda a base são lidos e organizados aleatoriamente.

- 13% dos batimentos do tipo N para treino
- 40% dos batimentos do tipo R para treino
- 40% dos batimentos do tipo A para treino
- 40% dos batimentos do tipo V para treino
- 50% dos batimentos do tipo F para treino
- 50% dos batimentos do tipo f para treino

## Configuração 3

Os registros dos pacientes são divididos em 3 grupos, sendo que cada pacientes petence apenas a 1 grupo.

# Resultados - Configuração 1 usando MLP

Todo o processo foi executado 5 vezes.

- Precisão total média = 95.78%, desvio padrão = 0.7463
- Sensibilidade classe N = 98.44, desvio padrão = 0.3578
- Sensitividade classe N = 97.20, desvio padrão = 0.5788
- Sensibilidade classe R = 94.22, desvio padrão = 1.5770
- Sensitividade classe R = 98.54, desvio padrão = 0.6044
- Sensibilidade classe A = 62.20, desvio padrão = 6.3577
- Sensitividade classe A = 64.96, desvio padrão = 8.7440
- Sensibilidade classe V = 79.57, desvio padrão = 6.0116
- Sensitividade classe V = 94.66, desvio padrão = 0.6229
- Sensibilidade classe F = 45.88, desvio padrão = 8.7236
- Sensitividade classe F = 30.10, desvio padrão = 8.4380
- Sensibilidade classe f = 59.14, desvio padrão = 6.8347
- Sensitividade classe f = 39.94, desvio padrão = 11.0428

# Resultados - Configuração 1 usando SVM

Todo o processo foi executado 5 vezes. Teste com SVM - kernel RBF,  $C = 10$  e  $\text{gamma} = 0.023810 = (1/2 * \text{num features})$

- Precisão total = 98.3566% std = 0.0293
- Sensibilidade classe N = 98.8173 % std = 0.0005
- Sensitividade classe N = 99.4503 % std = 0.0001
- Sensibilidade classe R = 99.4527 % std = 0.0006
- Sensitividade classe R = 99.0512 % std = 0.0015
- Sensibilidade classe A = 87.8322 % std = 0.0024
- Sensitividade classe A = 82.8371 % std = 0.0103
- Sensibilidade classe V = 97.2675 % std = 0.0016
- Sensitividade classe V = 91.4065 % std = 0.0008
- Sensibilidade classe F = 76.8911 % std = 0.0126
- Sensitividade classe F = 74.3314 % std = 0.0047
- Sensibilidade classe f = 87.3144 % std = 0.0113
- Sensitividade classe f = 84.5755 % std = 0.0150

## Resultados - Configuração 2 usando MLP

Todo o processo foi executado 5 vezes.

- Precisão total média = 95.72%, desvio padrão = 0.6058
- Sensibilidade classe N = 98.32, desvio padrão = 0.7328
- Sensitividade classe N = 97.30, desvio padrão = 0.3391
- Sensibilidade classe R = 94.38, desvio padrão = 0.8258
- Sensitividade classe R = 98.70, desvio padrão = 0.3674
- Sensibilidade classe A = 47.5800, desvio padrão = 26.6601
- Sensitividade classe A = 55.10, desvio padrão = 31.6217
- Sensibilidade classe V = 80.88, desvio padrão = 3.9366
- Sensitividade classe V = 94.76, desvio padrão = 0.8414
- Sensibilidade classe F = 36.32, desvio padrão = 12.6804
- Sensitividade classe F = 33.22, desvio padrão = 12.6646
- Sensibilidade classe f = 56.48, desvio padrão = 7.4419
- Sensitividade classe f = 39.76, desvio padrão = 12.5071

## Resultados - Configuração 2 usando SVM

Todo o processo foi executado 5 vezes. Teste com SVM - kernel RBF, C = 10 e  $\gamma = 0.023810 = (1/2 * \text{num features})$

- Precisão total = 98.3932%
- Sensibilidade classe N = 98.9110 % std = 0.0006
- Sensibilidade classe R = 99.4055 % std = 0.0004
- Sensibilidade classe R = 99.3080 % std = 0.0006
- Sensibilidade classe R = 99.0199 % std = 0.0016
- Sensibilidade classe A = 86.8951 % std = 0.0039
- Sensibilidade classe A = 84.4749 % std = 0.0124
- Sensibilidade classe V = 97.1976 % std = 0.0041
- Sensibilidade classe V = 91.5067 % std = 0.0044
- Sensibilidade classe F = 75.5666 % std = 0.0103
- Sensibilidade classe F = 75.4468 % std = 0.0209
- Sensibilidade classe f = 86.3821 % std = 0.0127
- Sensibilidade classe f = 86.2747 % std = 0.0066

# Resultados - Configuração 3 - MLP - G1 e G2 para treino e G3 para teste

- Precisão total = 77.5%
- Sensibilidade classe N = 94.30%
- Sensitividade classe N = 80.3%
- Sensibilidade classe R = 82.00%
- Sensitividade classe R = 39.04%
- Sensibilidade classe A = 7.70%
- Sensitividade classe A = 22.4%
- Sensibilidade classe V = 37.90%
- Sensitividade classe V = 80.70%
- Sensibilidade classe F = 0.50%
- Sensitividade classe F = 20.60%
- Sensibilidade classe f = 0.00%
- Sensitividade classe f = 0.00%

# Resultados - Configuração 3 - SVM - G1 e G2 para treino e G3 para teste

Teste com SVM - kernel RBF,  $C = 10$  e  $\text{gamma} = 0.023810 = (1/2 * \text{num features})$

- Precisão total = 75.89%
- Sensibilidade classe N = 79.36%
- Sensitividade classe N = 93.62%
- Sensibilidade classe R = 33.62%
- Sensitividade classe R = 74.62%
- Sensibilidade classe A = 10.66%
- Sensitividade classe A = 8.70%
- Sensibilidade classe V = 75.27%
- Sensitividade classe V = 37.21%
- Sensibilidade classe F = 8.82%
- Sensitividade classe F = 00.50%
- Sensibilidade classe f = 58.92%
- Sensitividade classe f = 1.04%

# Resultados - Configuração 3 - MLP - G1 e G3 para treino e G2 para teste

- Precisão total = 76.9%
- Sensibilidade classe N = 76.0%
- Sensitividade classe N = 100.0%
- Sensibilidade classe R = 0.00%
- Sensitividade classe R = 0.00%
- Sensibilidade classe A = 0.00%
- Sensitividade classe A = 0.00%
- Sensibilidade classe V = 92.60%
- Sensitividade classe V = 73.60%
- Sensibilidade classe F = 0.00%
- Sensitividade classe F = 0.00%
- Sensibilidade classe f = 100.00%
- Sensitividade classe f = 5.00%

# Resultados - Configuração 3 - SVM - G1 e G3 para treino e G2 para teste

Teste com SVM - kernel RBF,  $C = 10$  e  $\text{gamma} = 0.023810 = (1/2 * \text{num features})$

- Precisão total = 78.38%
- Sensibilidade classe N = 99.98%
- Sensitividade classe N = 77.76%
- Sensibilidade classe R = 00.00%
- Sensitividade classe R = 00.00%
- Sensibilidade classe A = 0.49%
- Sensitividade classe A = 22.58%
- Sensibilidade classe V = 87.21%
- Sensitividade classe V = 88.85%
- Sensibilidade classe F = 0.27%
- Sensitividade classe F = 3.03%
- Sensibilidade classe f = 43.07%
- Sensitividade classe f = 100.00%

# Resultados - Configuração 3 - MLP - G2 e G3 para treino e G1 para teste

- Precisão total = 72.6%
- Sensibilidade classe N = 83.20%
- Sensitividade classe N = 99.40%
- Sensibilidade classe R = 27.60%
- Sensitividade classe R = 1.80%
- Sensibilidade classe A = 0.00%
- Sensitividade classe A = 0.00%
- Sensibilidade classe V = 5.80%
- Sensitividade classe V = 9.20%
- Sensibilidade classe F = 0.00%
- Sensitividade classe F = 0.00%
- Sensibilidade classe f = 0.00%
- Sensitividade classe f = 0.00%

# Resultados - Configuração 3 - SVM - G2 e G3 para treino e G1 para teste

Teste com SVM - kernel RBF,  $C = 10$  e  $\text{gamma} = 0.023810 = (1/2 * \text{num features})$

- Precisão total = 77.60%
- Sensibilidade classe N = 95.78%
- Sensitividade classe N = 91.67%
- Sensibilidade classe R = 00.00%
- Sensitividade classe R = 00.00%
- Sensibilidade classe A = 22.56%
- Sensitividade classe A = 55.46%
- Sensibilidade classe V = 89.40%
- Sensitividade classe V = 35.49%
- Sensibilidade classe F = 0.26%
- Sensitividade classe F = 0.23%
- Sensibilidade classe f = 16.51%
- Sensitividade classe f = 94.82%

## Resultados - Visão geral MLP

Tipo	Sensibilidade (%)				
	Conf. 1 (std)	Conf. 2 (std)	Conf. 3 G1	Conf. 3 G2	Conf. 3 G3
N	98.44 (0.35)	98.32 (0.72)	83.20	76.0	94.3
R	94.22 (1.57)	94.38 (0.82)	27.60	0.00	82.00
A	62.20 (6.35)	47.58 (26.66)	0.00	0.00	7.70
V	79.57 (6.01)	80.88 (3.93)	5.80	92.60	37.90
F	45.88 (8.72)	36.32 (12.68)	0.00	0	0.50
f	59.14 (6.83)	56.48 (7.44)	0.00	100	0.0
P. Total	95.78 (0.74)	95.72 (0.60)	72.6	76.9	77.5

## Conclusão e Trabalhos futuros

- Usar dados de forma tendenciosa favorece os resultados. Esta prática deve ser evitada.
- Com a utilização do padrão recomendado pela AAMI, os métodos podem ser efetivamente comparados.