

Um Método Ensemble para Classificação de Arritmias: Uma Avaliação Com Mais de 10 Mil Registros de Sinais de ECG

Rafael F. Oliveira¹, Anderson A. Ferreira¹, Gladston J. P. Moreira¹, Eduardo J. S. Luz¹

¹Departamento de Computação – Universidade Federal de Ouro Preto (UFOP)
Ouro Preto – 35.400-000 – MG – Brasil

rafael.fo@aluno.ufop.edu.br, {anderson.ferreira, gladston, eduluz}@ufop.edu.br

Abstract. *Early diagnosis of arrhythmias is of paramount importance since heart diseases cause so many deaths on the planet, according to the WHO. Therefore, automation of arrhythmia events is desirable. In this context, an automatic arrhythmia classification model is proposed based on voting ensemble, Discrete Wavelet Transform (DWT), and electrocardiogram (ECG). The model was evaluated in a large dataset with over 10,000 patients, under the inter-patient paradigm, and it achieved an average F1-score of 0.93, an efficiency increase of 2.15% over Random Forest and 1.07% over GradientBoost and XGradient Boost. Thus, the proposed model has great potential for real use due to its robustness, and generalization power.*

Resumo. *O diagnóstico precoce de arritmias é de fundamental importância, sendo uma das doenças cardíacas que causa maior número de mortes no planeta, segundo a OMS. Por isso, a automação do processo de identificação de arritmia é desejável. Nesse contexto, um modelo de classificação automática de arritmias em ECGs é proposto baseado em esquema de votação (voting ensemble) e Discrete Wavelet Transform (DWT). Avaliado em um conjunto de dados com mais de 10 mil pacientes e sob o paradigma inter-patient, o modelo proposto alcançou F1-score médio de 0,93, um aumento em eficiência de 2,15% em relação ao Random Forest e 1,07% em relação ao GradientBoost e ao XGradient Boost. Com isso, o modelo proposto apresenta grande potencial para uso real devido sua robustez e poder de generalização.*

1. Introdução

De acordo com a Organização Mundial de Saúde (OMS), doenças cardíacas são a principal causa de morte no planeta. Embora boa parte das doenças cardíacas possam ser diagnosticadas com antecedência, o número de casos continua a crescer. Estima-se que mais de 75% dos casos acontecem em países pobres e em desenvolvimento e 17,9 milhões de pessoas morreram de doenças cardíacas em 2019, representando 32% das mortes globais [WHO 2021]. Isso se deve ao fato de que medidas de prevenção, aparentemente, não estão sendo adequadas, por isso o diagnóstico precoce é de fundamental importância.

O principal exame para diagnóstico de doenças cardíacas é chamado Eletrocardiograma (ECG). Um dos problemas cardíacos detectável pelo exame de ECG são as arritmias. Arritmias são batimentos cardíacos anormais ou irregulares. Há duas categorias

de arritmias: as arritmias formadas por um único batimento cardíaco irregular, denominadas de arritmias morfológicas, e as arritmias formadas por um conjunto irregular de batimentos, denominadas arritmias rítmicas.

O processo de identificação e classificação de arritmias pode ser muito cansativo para um ser humano e ainda sujeito a erros. Portanto, a necessidade de estudos por soluções para esse problema se faz necessário. Uma solução consiste no uso de inteligência artificial para classificação automática dos padrões de arritmias. Atualmente, métodos baseados em aprendizagem de máquina (*machine learning*), tornaram-se bastante populares, obtendo resultados expressivos [Hannun et al. 2019].

O processo de classificação de arritmias pode ser dividido em três partes principais: pré-processamento, extração de atributos e classificação. Na etapa de pré-processamento, são aplicados métodos e técnicas para remoção de ruídos, como por exemplo, interferências no sinal de ECG. Após a limpeza de ruídos dos sinais, atributos são extraídos de modo que possa ocorrer uma caracterização de cada sinal de ECG. Em seguida, os atributos são utilizados para treinar e validar modelos de classificação.

Diversos métodos envolvendo o processo de classificação de arritmias têm sido explorados na literatura [Luz et al. 2016]. Para validação desses métodos, populares conjunto de dados públicos são utilizados, como *MIT Arrhythmia Database (MIT-BIH)*, *The American Heart Association (AHA)*, *The Creighton University Sustained Ventricular Arrhythmia Database (CU)* e *The European Society of Cardiology ST-T Database (EDB)*. Apesar das evidências que conjuntos de dados, mesmo que reduzidos, são úteis para área de saúde [Silveira et al. 2022], a escalabilidade de métodos de classificação deve ser investigada. Dessa forma, como garantir que os métodos vão generalizar bem para uma população maior?

Neste contexto, este trabalho propõe um modelo de classificação automática de arritmias em sinais de ECGs, baseado em votação (*voting ensemble*), utilizando um conjunto de dados com muitos registros e mais representativo. Os experimentos foram validados sob o paradigma *inter-patient* no conjunto de dados Chapman University and Shaoxing People's Hospital (CUSPH) com mais de 10 mil pacientes. Para o *ensemble*, investiga-se os atributos extraídos por meio de *Discrete Wavelet Transform (DWT)* e dados coletados do próprio conjunto de dados. Quatro métodos de classificação são avaliados, sendo eles: *Extreme Gradient Boost*, *Random Forest*, *Gradient Boost* e *Voting Ensemble*.

Como contribuição, o presente trabalho permitiu investigar o desempenho de um modelo de classificação *voting ensemble* diante de um conjunto de dados com muitos registros (mais de 10 mil pacientes). O modelo proposto apresentou um aumento em eficiência de 2,15% em relação ao *Random Forest* e 1,07% em relação ao *Gradient Boost* e ao *XGradient Boost*. Vinculado a esse desempenho, a interpretabilidade dos resultados permitiu verificar que alguns atributos apresentaram maior importância para o modelo, influenciando assim no seu desempenho.

O trabalho é organizado como segue. A Seção 2 descreve uma breve revisão da literatura referente às etapas necessárias para um sistema de classificação de ECG, bem como os métodos existentes. A Seção 3 detalha o conjunto de dados utilizado para validação do modelo. A metodologia proposta para a classificação de ECG é descrita na Seção 4. Além disso, na Seção 5, a descrição dos experimentos, resultados e discussões

são realizados. Finalmente, as observações finais são reportadas na Seção 6.

2. Trabalhos relacionados

O bom desempenho de um modelo de *machine learning* normalmente depende diretamente de um grande conjunto de dados. Uma grande representatividade de dados (registros de pacientes) permite que o modelo seja mais generalista, porém isso só é possível com um grande conjunto de dados para tornar o conjunto de treinamento o mais diversificado possível. Contudo, nem sempre isso é possível, uma vez que os conjuntos de dados mais populares apresentam poucos registros, impossibilitando que métodos/técnicas sejam validados em um grande volume de dados [Hannun et al. 2019].

Com um conjunto de dados suficientemente representativo, um modelo de classificação de arritmias pode ser dividido em três etapas principais: 1) pré-processamento, 2) extração de atributos e 3) classificação. Para cada etapa, diversos métodos tem sido propostos na literatura. Alguns desses métodos são revistos nessa seção.

2.1. Pré-processamento

Antes da extração de informações úteis do sinal, algumas técnicas de processamento precisam ser aplicadas previamente. A etapa de pré-processamento é a etapa correspondente à limpeza do sinal de ECG (remoções de interferências na linha de energia do sinal, por exemplo). Na literatura, alguns métodos têm sido propostos, como filtros (passa-baixa, média móvel, passa-alta) [Garcia et al. 2017], decomposição empírica [Rakshit and Das 2018], filtros adaptativos [Poungponsri and Yu 2013] e *wavelet transform* [Wang et al. 2018]. Deve-se observar que a aplicação excessiva de filtros pode ocasionar na distorção da morfologia do sinal, tornando-o inutilizável para a classificação e detecção de arritmias [Luz et al. 2016].

Em [Dias et al. 2021], utilizou-se dois filtros de média móvel, um deles com largura de 200 ms e outro com largura de 600 ms para remoção de ruídos do sinal, ocasionado pela respiração ou pelo movimento do paciente. Em [Wang et al. 2018], os autores propuseram um método de seleção *wavelet* com limiar adaptativo para redução de ruído em ECG baseado em funções *wavelet*. O limiar foi ajustado dinamicamente utilizando heurística e o número de camadas da decomposição da função *wavelet*. Em [Wang et al. 2016], a decomposição adaptativa de Fourier foi escolhida para diminuir o ruído dos sinais de ECG.

2.2. Extração de atributos e classificação

A detecção de arritmias com base na classificação de ECG tornou-se um tópico de pesquisa muito popular na literatura nos últimos anos [Luz et al. 2016]. Os métodos tradicionais utilizam atributos extraídos do sinal, como o intervalo RR, coeficientes *wavelets*, amplitude, complexo QRS, onda P, duração, etc. Os principais pontos e intervalos, denominados de pontos fiduciais, de um sinal ECG são apresentados na Figura 1. Dentre os pontos apresentados, há o segmento QRS, que normalmente é utilizado durante o processo de segmentação de um sinal de ECG em batimentos menores.

Após a etapa de pré-processamento, podem ainda ser aplicadas outras técnicas para adequar o sinal de forma que as melhores informações possam ser extraídas, tais como normalização e amostragem. Alguns trabalhos utilizam uma etapa de segmentação antes

da extração de atributos [Dias et al. 2021, Mousavi and Afghah 2019, Wang et al. 2021], onde os sinais são divididos em batimentos individuais de acordo com um tamanho fixo de amostras. Essa etapa de segmentação requer detecção das ondas QRS e pontos fiduciais dos batimentos [Wang et al. 2021]. Neste trabalho não foi realizada a etapa de segmentação, sendo o sinal de ECG utilizado por completo.

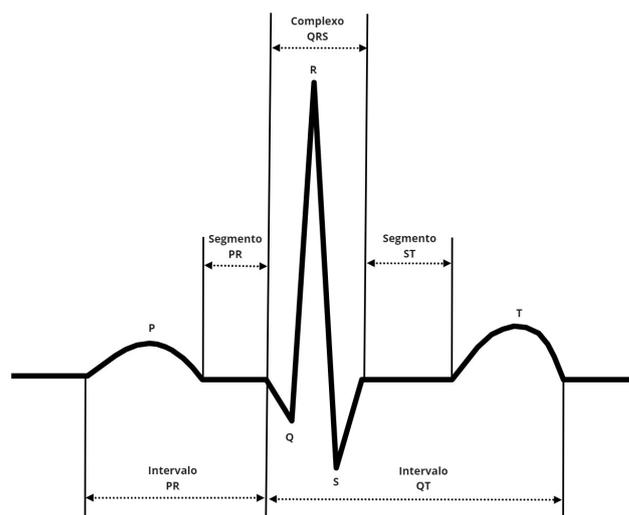


Figura 1. Pontos fiduciais de um batimento de sinal de ECG. Fonte: Adaptada de [Zheng et al. 2020].

A identificação dos sinais de ECG ocorre a partir dos atributos extraídos previamente. Com isso, diversos atributos e métodos têm sido utilizados na literatura. Em [De Chazal et al. 2004], combinou-se atributos rítmicos, duração de onda e atributos morfológicos. Em [Wang et al. 2018], os autores propuseram um método de seleção de atributos baseado em transformação *wavelet* com limiar adaptativo, visando reduzir o ruído do sinal de ECG. A *Wavelet Transform* (WT) pode ser utilizada para extração de atributos, uma vez que a mesma realiza uma decomposição adaptativa em tempo-frequência do sinal de ECG. Pela representação multirresolução é possível descrever a estrutura do sinal por apenas alguns coeficientes no domínio *Wavelet* [Sternickel 2002].

Com o objetivo de tornar o modelo em desenvolvimento mais robusto e adaptável a um cenário real, alguns trabalhos na literatura exploram dois paradigmas de validação de dados: *intra-patient* e *inter-patient* [Mousavi and Afghah 2019]. No paradigma *intra-patient*, o conjunto de treinamento e teste é formado por sinais de ECG de um mesmo paciente, enquanto que no paradigma *inter-patient*, o conjunto de treinamento e teste são obtidos de sinais de ECG de pacientes diferentes. Esse último permite uma generalização mais próxima de um cenário real.

Após a extração dos atributos do sinal de ECG, um classificador de aprendizado de máquina é treinado utilizando os atributos do sinal. A escolha do classificador depende do problema a ser tratado. Na literatura há diversos modelos, dos quais *Random Forest* [Yang et al. 2021], *Extreme Gradient Boost* [Shi et al. 2019], *Gradient Boost* [Shi et al. 2019] e *Voting Ensemble* foram utilizados nesse trabalho.

3. Conjunto de dados

O conjunto de dados utilizado neste trabalho foi a Chapman University and Shaoxing People's Hospital (CUSPH)¹ [Zheng et al. 2020]. O conjunto de dados é composto por 10.646 sinais de ECGs de pacientes diferentes, sendo 5.956 homens e 4.690 mulheres. Dentre esses pacientes, 17% apresentam arritmia normal e 83% apresentam pelo menos algum tipo de arritmia. A idade dos pacientes varia de 4 a 98 anos. Cada sinal contém informação de 12 canais (configurações das posições dos elétrodos no paciente) com duração de 10 segundos e amostrados a uma taxa 500 Hz. Neste trabalho, foi utilizada o canal II.

Tabela 1. Descrição das arritmias (classes) do conjunto de dados.

Classes	Frequência	Agrupamento	Total
SR (<i>Sinus Rhythm</i>)	1.826	SR	2.225
SI (<i>Sinus Irregularity</i>)	399		
AFIB (<i>Atrial Fibrillation</i>)	1.780	AFIB	2.225
AF (<i>Atrial Flutter</i>)	445		
ST (<i>Sinus Tachycardia</i>)	1.568		
SVT (<i>Supraventricular Tachycardia</i>)	587		
AT (<i>Atrial Tachycardia</i>)	121		
AVNRT (<i>Atrioventricular Node Reentrant Tachycardia</i>)	16	GSVT	2.307
AVRT (<i>Atrioventricular Reentrant Tachycardia</i>)	8		
SAAWR (<i>Sinus Atrium to Atrial Wandering Rhythm</i>)	7		
SB (<i>Sinus Bradycardia</i>)	3.889	SB	3.889
Total de registros			10.646

O conjunto de dados contém 11 tipos de arritmias (classes) com diferentes tamanhos de amostras. O conjunto de dados é desbalanceado devido a presença de classes com poucas amostras (classes minoritárias). Por isso, os autores de [Zheng et al. 2020] propuseram um agrupamento por taxonomia em quatro grupos durante a composição do conjunto de dados. A Tabela 1 apresenta o detalhamento de todas as arritmias presentes no conjunto de dados, bem como o agrupamento de cada classe. O conjunto de dados ainda contém 11 atributos dos sinais, sendo eles: taxa ventricular, taxa atrial, duração QRS, intervalo QT, QT corrigido, eixo R, eixo T, contagem QRS, QOnset, QOffset e TOffset, além da idade e do gênero de cada paciente.

Como cada registro no conjunto de dados corresponde a um paciente diferente, o paradigma *inter-patient* foi aplicado, formando um conjunto de dados para treinamento diferente do conjunto de dados para teste, permitindo assim, uma generalização do modelo classificador para um cenário real.

4. Metodologia

A seguir, é apresentada a metodologia para classificação automática de arritmias em sinais de ECG. O diagrama da Figura 2 apresenta as etapas do sistema de classificação utilizado: pré-processamento, extração de atributos e classificação.

¹<https://doi.org/10.6084/m9.figshare.c.4560497.v2>

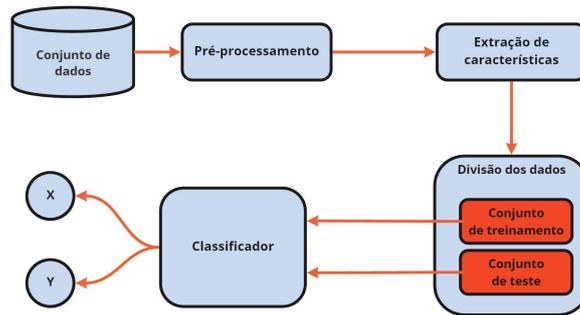


Figura 2. Diagrama do sistema de classificação de arritmia.

4.1. Pré-processamento

Nessa etapa, uma análise descritiva permite verificar como os dados estão presentes no conjunto de dados, bem como sua distribuição, presença de valores nulos, *outliers* e duplicatas, correlações de atributos, dentre outros. Neste trabalho, esta etapa consiste no agrupamento em classes de arritmias propostas pelos autores de [Zheng et al. 2020] conforme a Tabela 1, na remoção de duplicatas e na remoção de dois atributos desnecessários (*FileName* e *Beat*) para o modelo.

Quanto ao processo de remoção de ruídos, o conjunto de dados dispõe de sinais com remoção de ruídos. Os autores de [Zheng et al. 2020] utilizaram uma sequência de três filtros para tal processo, sendo um filtro passa-baixa *Butterworth* com passa-banda de 50 Hz e faixa de 60 Hz, um suavizador *LOESS* e *Non Local Means - (NLM)*. Logo, a remoção de ruídos não foi necessária durante a etapa de pré-processamento.

4.2. Extração de atributos

Nessa etapa, o objetivo é realizar o aumento de atributos dos sinais de ECG através da extração de atributos dos sinais no domínio da frequência. Logo, os sinais são convertidos para o domínio tempo-frequência através da técnica *Discrete Wavelet Transform (DWT)*. Como os sinais exibem atributos de frequência diferentes, essa diferença de comportamento é capaz de gerar atributos distintos para cada sinal, possibilitando a classificação dos mesmos.

Wavelet Transform (WT) é uma série composta por operações de escalonamento e deslocamento em base de uma função geradora chamada de *mother wavelet*, podendo ser contínuas ou discretas. WT pode representar uma informação local do sinal de ECG no domínio do tempo e frequência para ajustar as sub-bandas de alta e baixa frequência separadamente através da janelas de tempo e de frequência. WT tem sido utilizada em diferentes metodologias, como detecção de batimentos [Kadambe et al. 1999] e extração de atributos [Zhao and Zhang 2005] em sinais de ECG.

A *Continuous Wavelet Transformation (CWT)* realiza os escalonamentos e deslocamentos de forma contínua enquanto que a *Discrete Wavelet Transform (DWT)* realiza as operações de forma discreta. As fórmulas (1) e (2) [Kadambe et al. 1999] representam as funções contínuas e discretas, respectivamente:

$$CWT_x(a, b) = \frac{1}{\sqrt{|a|}} \int_{-\infty}^{\infty} x(t) \bar{\psi} \left(\frac{t-b}{a} \right) dt \quad DWT_x(b, 2^i) = \frac{1}{\sqrt{2^i}} \int_{-\infty}^{\infty} x(t) \bar{\psi} \left(\frac{t-b}{2^i} \right) dt$$

(1)
(2)

sendo $x(t)$ o sinal original, $\bar{\psi}(t)$ a função *wavelet*, $a \in \mathcal{R}$ e $a \neq 0$ fator de escalonamento, b fator de deslocamento e $i = 1, 2, \dots$

A DWT decompõe o sinal de ECG iterativamente e, em cada iteração, o sinal é decomposto a partir do valor médio da frequência do sinal de entrada. Para isso, utiliza-se Filtros de Passagem Alta (FPA) e Filtros de Passagem Baixa (FPB). Após a decomposição, a depender do tipo da função geradora utilizada, é possível reconstituir o sinal original utilizando os coeficientes *wavelets* com pouca perda de informação.

A função geradora *Daubechies (db14)* foi utilizada no processo de extração de atributos de cada sinal de ECG. Essa função gerou 08 sub-bandas diferentes para cada sinal. Em cada sub-banda foram extraídos 14 atributos, apresentado na Tabela 2, totalizando 112 novos atributos para cada sinal. *Zero crossing* e *mean crossing* significam quantas vezes o sinal de ECG cortou em ($y = 0$) e $y = \text{media}(y)$, respectivamente. *Skewness* mede a assimetria da distribuição dos dados e *Kurtosis* descreve o grau em que os dados se agrupam nas caudas ou no pico de sua distribuição em relação à distribuição normal. Após a extração, os novos atributos gerados foram concatenados com os atributos originais do conjunto de dados.

Tabela 2. Conjunto de atributos para cada sinal de ECG.

Atributos				
• 5-percentil	• 25-percentil	• desvio padrão	• <i>skewness</i>	• <i>zero crossing</i>
• <i>kurtosis</i>	• entropia	• média quadrática	• variância	• <i>mean crossing</i>
• 75-percentil	• 95-percentil	• mediana	• média	

4.3. Classificação

Em relação a modelos de classificação de arritmia, os algoritmos mais populares para essa tarefa, segundo [Luz et al. 2016], são: *SVM*, redes neurais, discriminantes lineares e computação de reservatório com regressão logística (*Reservoir Computing With Logistic Regression - RC*).

SVM [Vapnik 1999] é um dos classificadores mais populares presente na literatura para classificação de arritmias, porém este modelo apresenta comportamento desfavorável para dados desbalanceados [Luz et al. 2016]. Discriminantes lineares é um método estatístico baseado em funções discriminantes. As funções são estimadas a partir do conjunto de treinamento e são lineares em relação ao vetor de atributos. Este método é vantajoso em dados desbalanceados e requer baixo tempo de treinamento [Luz et al. 2016]. Segundo [Rodan and Tino 2010], os modelos *RC* são modelos dinâmicos que visam processar um sinal de série temporal através de um reservatório dinâmico não adaptável e uma leitura dinâmica do reservatório. Essa técnica é robusta ao desequilíbrio de classes.

Redes neurais artificiais, por sua vez, tem apresentado muito interesse na literatura, sendo as estruturas *Multilayer Perceptron (MLP)* e *Probabilistic Neural Networks*

(*PNN*) as mais populares. Estes tipos de rede apresentam problemas com generalização e alto tempo computacional de treinamento [Luz et al. 2016]. Segundo [Osowski et al. 2008], a combinação de modelos *ensemble* tem sido uma proposta interessante para este problema, principalmente na redução da incidência de falso negativos. Por isso, neste contexto, este trabalho investiga as performances de quatro modelos *ensemble*: *Extreme Gradient Boost (XGradient Boost)*, *Random Forest*, *Gradient Boost* e *Voting Ensemble*.

Algoritmos de aprendizagem de máquina combinados (*Ensembles*) são modelos compostos que combinam uma série de classificadores de baixo desempenho com o objetivo de criar um classificador aprimorado. Os modelos são ajustados utilizando função de perda diferenciável (*loss function*) e algoritmo de otimização de gradiente descendente (*gradient descent*).

Voting Ensemble é um modelo que combina previsões de múltiplos modelos. Há duas metodologias para as previsões deste modelo: *hard mode* (soma dos votos dos modelos), sendo esta a utilizada neste trabalho e, *soft mode* (soma das probabilidades dos modelos). Este tipo de modelo pode oferecer uma baixa variância nas previsões em relação aos modelos individuais, diminuindo o *overfitting*. Neste trabalho, utilizou-se a combinação dos modelos *XGradient Boost*, *Random Forest* e *Gradient Boost* para a criação do modelo *voting ensemble* com o objetivo de melhorar a performance de classificação.

5. Resultados

Nesta seção, são apresentados as configurações do experimentos, os resultados e suas discussões.

5.1. Configurações e métricas

Os experimentos foram conduzidos em um ambiente Windows 11 com ambiente virtual Anaconda Python 3.9 com Intel®i5-1035G1 com 04 cores (threads) de 1,2GHz, 8GB RAM DDR4 e GPU GeForce MX350 de 2GB de RAM.

Para avaliar o desempenho do modelo de classificação proposto, três métricas frequentemente mais utilizadas na literatura são [Luz et al. 2016]: predição positiva ($+P^k$), sensibilidade (Se^k) e F1-score (F_s^k), sendo $k \in \{\text{AFIB, GSVT, SB, SR}\}$, caracterizando assim, uma análise multi-classe. A métrica F_s será considerada a escolha do melhor modelo, devido ser a média ponderada entre a $+P$ e Se .

As métricas são baseadas nas medidas de verdadeiros positivos (TP), falsos negativos (FN), verdadeiros negativos (TN) e falsos positivos (FP). As Fórmulas 4, 5 e 6 correspondem às métricas ($+P^k$), (Se^k) e (F_s^k), respectivamente:

$$+P^k = \frac{TP}{TP + FP} \times 100$$

(4)

$$Se^k = \frac{TP}{TP + FN} \times 100$$

(5)

$$F_s^k = 2 * \frac{Se^k * (+P^k)}{Se^k + (+P^k)}$$

(6)

5.2. Resultados e discussões

A validação de cada modelo foi realizada utilizando *10-fold cross-validation* por meio do método *StratifiedKFold* da biblioteca *Scikit-learn* [Pedregosa et al. 2011]. Este

método permite que a divisão de *folds* ocorra preservando o percentual de amostras de cada classe. O resultado final corresponde à média das métricas obtidas em cada iteração.

O modelo *voting ensemble* obteve a média ponderada mais alta de F_S de 0,93, um aumento em eficiência de 2,15% em relação ao *Random Forest* e 1,07% em relação ao *GradientBoost* e ao *XGradient Boost*. A matriz de confusão da Tabela 3 apresenta a performance do modelo *voting ensemble* em cada classe. Percebe-se que o modelo *voting ensemble* apresentou um bom desempenho de predição nas respectivas classes. Os resultados das métricas analisadas são detalhados na Tabela 4. O modelo de *voting ensemble* alcançou os melhores F_S para as classes **AFIB**, **GSVT**, **SB** e **SR** de 0,85, 0,91, 0,99 e 0,94, respectivamente. O modelo *XGradient Boost* apresentou a melhor Se de 0,85 para a classe **AFIB** e $+P$ de 0,93 para a classe **SR**.

Tabela 3. Matriz de confusão do modelo *Voting Ensemble*.

		Predição				TOTAL
		AFIB	GSVT	SB	SR	
Verdade	AFIB	1799	239	47	140	2225
	GSVT	146	2125	8	28	2307
	SB	17	0	3861	11	3889
	SR	64	8	19	2134	2225
TOTAL		2026	2372	3935	2313	

Com a finalidade de aumentar a interpretabilidade dos resultados, a técnica de importância de atributos foi utilizada para o modelo *voting ensemble*. Essa técnica permite classificar e selecionar os atributos mais importantes para o modelo. A Tabela 5 apresenta os vinte atributos mais importantes para o modelo *voting ensemble* através dos coeficientes dos modelos². Vale ressaltar que nessa análise é considerado tanto os atributos originais quanto os atributos extraídos pela DWT.

A importância de atributos apresenta quais deles mais contribuem para o melhor desempenho do modelo. Observa-se que além dos atributos originais do conjunto de dados, alguns atributos extraídos pela DWT (atributos com prefixo *Dwt_*) contribuíram para o bom desempenho do modelo. Isso demonstra que os atributos extraídos pela DWT contribuem para uma boa caracterização da morfologia do sinal de ECG e complementam os atributos originais do conjunto de dados.

Ainda, percebe-se que atributos originais relacionados à frequência do sinais de ECG (*ventricular rate*, *atrial rate*, *QRSCount*) demonstram grande importância para o modelo. Isso se dá devido às classes de arritmias, como a taquicardia e bradicardia, que podem ser detectadas apenas pela variação da frequência. Todavia, algumas arritmias alteram a morfologia do sinal e demandam atributos que permitam a detecção da variação na forma do sinal, como intervalo QT, duração QRS e coeficientes *wavelets*. Ressalta-se ainda, que a idade dos pacientes apresentou grande importância para o modelo. Sabe-se que taquicardia provoca o aumento da frequência cardíaca do indivíduo. No entanto, indivíduos em idade infantil também apresentam frequência cardíaca elevada. Nesse caso, diagnóstico em crianças, o atributo idade pode ser de fundamental importância para a correta detecção de arritmias.

²Obtido através da função `feature_importances_` implementado na biblioteca `scikit-learn`.

Os resultados apresentam que o modelo baseado em votação pode melhorar a performance de modelos de classificação. Essa é uma abordagem relevante pois sua metodologia utiliza o desempenho coletivo de múltiplos modelos, além de apresentar robustez e poder de generalização, visto ser vantajoso em grandes conjuntos de dados como a utilizada neste trabalho. Vale ressaltar que o modelo *voting ensemble* faz sentido quando os modelos envolvidos apresentam desempenho similares. Caso contrário, o desempenho será comprometido.

Tabela 4. Performance dos modelos de classificação.

Grupos	Gradient Boost			XGradient Boost			Random Forest			Voting Ensemble		
	+P	Se	F _s	+P	Se	F _s	+P	Se	F _s	+P	Se	F _s
AFIB	0,88	0,80	0,84	0,89	0,82	0,85	0,85	0,71	0,78	0,89	0,81	0,85
GSVT	0,89	0,92	0,91	0,90	0,92	0,91	0,86	0,92	0,89	0,90	0,92	0,91
SB	0,98	0,99	0,99	0,98	0,99	0,99	0,97	0,99	0,98	0,98	0,99	0,99
SR	0,92	0,96	0,94	0,93	0,96	0,94	0,89	0,94	0,91	0,92	0,96	0,94
Acc			0,92			0,92			0,91			0,93
Média Macro	0,91	0,91	0,91	0,92	0,92	0,92	0,89	0,89	0,89	0,92	0,92	0,92
Média Ponderada	0,92	0,92	0,92	0,92	0,92	0,92	0,91	0,91	0,91	0,93	0,93	0,93

Tabela 5. Os 20 atributos mais importantes para o modelo *Voting Ensemble*.

Atributos	Importância	Atributos	Importância	Atributos	Importância	Atributos	Importância
1. Ventricular Rate	1,2362	6. Dwt_15	0,0446	11. Gender_Female	0,0216	16. Dwt_6	0,0148
2. Atrial Rate	0,4580	7. QTInterval	0,0402	12. Dwt_1	0,0212	17. Dwt_9	0,0145
3. Patient Age	0,2098	8. TOffset	0,0389	13. Dwt_2	0,0205	18. Dwt_5	0,0144
4. QRSCount	0,1024	9. RAxis	0,0244	14. Dwt_3	0,0181	19. Dwt_11	0,0133
5. TAxis	0,0648	10. QTCorrected	0,0239	15. QRS Duration	0,0169	20. Dwt_7	0,0130

6. Conclusões

Neste trabalho, foi proposto um modelo de classificação automática de arritmia baseado em modelo *voting ensemble* e *Discrete Wavelet Transform - DWT* para extração de atributos em um conjunto de dados com 10.646 pacientes. Inicialmente, os sinais sem ruídos disponibilizados pelo conjunto de dados são decompostos em sub-bandas de frequências, de acordo com a função geradora *wavelet* db14, gerando assim 112 novos atributos. Os mesmos foram combinados com os 15 atributos originais do conjunto de dados, totalizando 127 atributos por sinal. Por fim, quatro classificadores (*XGradient Boost*, *Random Forest*, *GradientBoost* e *Voting Ensemble*) foram utilizados e comparados de acordo com suas performances. Como simulação de aplicabilidade em cenários reais, os modelos foram treinados de acordo com o paradigma *inter-patient*.

Os resultados apresentam uma performance considerável de classificação no modelo *voting ensemble*. Quando comparado ao desempenho individuais dos modelos, o modelo *voting ensemble* apresentou F_s médio de 0,93, um aumento em eficiência de 2,15% em relação ao *Random Forest* e 1,07% em relação ao *GradientBoost* e ao *XGradient Boost*. Isso torna claro que as técnicas que envolvem a combinação de modelos de aprendizado de máquina (*ensemble*) podem ajudar na melhoria de performance de modelos de predição.

O modelo proposto permite ainda uma escalabilidade das técnicas de aprendizado de máquina em grandes conjuntos de dados devido sua robustez e poder de generalização.

Com a interpretabilidade manual dos atributos através da importância de atributos foi possível detectar os que mais influenciaram no desempenho do modelo. Foi possível identificar que atributos relacionados à frequência e à morfologia do sinal de ECG (*ventricular rate*, *atrial rate*, *QRSCount*) apresentaram grande importância para a identificação das arritmias, uma vez que arritmias podem ser identificadas por diferentes maneiras. A idade do paciente também apresentou grande importância para o modelo. Por fim, a interpretabilidade dos atributos ainda permite uma melhor compreensão e dimensionalidade dos atributos contidos em grande conjuntos de dados.

Agradecimentos

Os autores agradecem a *Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Financiamento código 001, Fundação de Amparo à Pesquisa do Estado de Minas Gerais (FAPEMIG, APQ-01518-21), Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) e Universidade Federal de Ouro Preto (UFOP)* pelo suporte no desenvolvimento do presente trabalho.

Referências

- De Chazal, P., O'Dwyer, M., and Reilly, R. B. (2004). Automatic classification of heartbeats using ecg morphology and heartbeat interval features. *IEEE transactions on biomedical engineering*, 51(7):1196–1206.
- Dias, F. M., Monteiro, H. L., Cabral, T. W., Naji, R., Kuehni, M., and Luz, E. J. d. S. (2021). Arrhythmia classification from single-lead ecg signals using the inter-patient paradigm. *Computer Methods and Programs in Biomedicine*, 202:105948.
- Garcia, G., Moreira, G., Menotti, D., and Luz, E. (2017). Inter-patient ecg heartbeat classification with temporal vcg optimized by pso. *Scientific reports*, 7(1):1–11.
- Hannun, A. Y., Rajpurkar, P., Haghpanahi, M., Tison, G. H., Bourn, C., Turakhia, M. P., and Ng, A. Y. (2019). Cardiologist-level arrhythmia detection and classification in ambulatory electrocardiograms using a deep neural network. *Nature medicine*, 25(1):65.
- Kadambe, S., Murray, R., and Boudreaux-Bartels, G. F. (1999). Wavelet transform-based qrs complex detector. *IEEE Transactions on biomedical Engineering*, 46(7):838–848.
- Luz, E. J. d. S., Schwartz, W. R., Cámara-Chávez, G., and Menotti, D. (2016). Ecg-based heartbeat classification for arrhythmia detection: A survey. *Computer methods and programs in biomedicine*, 127:144–164.
- Mousavi, S. and Afghah, F. (2019). Inter-and intra-patient ecg heartbeat classification for arrhythmia detection: a sequence to sequence deep learning approach. In *ICASSP 2019-2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pages 1308–1312. IEEE.
- Osowski, S., Markiewicz, T., and Hoai, L. T. (2008). Recognition and classification system of arrhythmia using ensemble of neural networks. *Measurement*, 41(6):610–617.
- Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., and Duchesnay, E. (2011). Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research*, 12:2825–2830.

- Poungponsri, S. and Yu, X.-H. (2013). An adaptive filtering approach for electrocardiogram (ecg) signal noise reduction using neural networks. *Neurocomputing*, 117:206–213.
- Rakshit, M. and Das, S. (2018). An efficient ecg denoising methodology using empirical mode decomposition and adaptive switching mean filter. *Biomedical signal processing and control*, 40:140–148.
- Rodan, A. and Tino, P. (2010). Minimum complexity echo state network. *IEEE transactions on neural networks*, 22(1):131–144.
- Shi, H., Wang, H., Huang, Y., Zhao, L., Qin, C., and Liu, C. (2019). A hierarchical method based on weighted extreme gradient boosting in ecg heartbeat classification. *Computer methods and programs in biomedicine*, 171:1–10.
- Silveira, A. C. d., Sobrinho, Á., Silva, L. D. d., Costa, E. d. B., Pinheiro, M. E., and Perkusich, A. (2022). Exploring early prediction of chronic kidney disease using machine learning algorithms for small and imbalanced datasets. *Applied Sciences*, 12(7):3673.
- Sternickel, K. (2002). Automatic pattern recognition in ecg time series. *Computer methods and programs in biomedicine*, 68(2):109–115.
- Vapnik, V. (1999). *The nature of statistical learning theory*. Springer science & business media, 2nd edition.
- Wang, L., Sun, W., Chen, Y., Li, P., and Zhao, L. (2018). Wavelet transform based ecg denoising using adaptive thresholding. In *Proceedings of the 2018 7th International Conference on Bioinformatics and Biomedical Science*, pages 35–40.
- Wang, T., Lu, C., Sun, Y., Yang, M., Liu, C., and Ou, C. (2021). Automatic ecg classification using continuous wavelet transform and convolutional neural network. *Entropy*, 23(1):119.
- Wang, Z., Wan, F., Wong, C. M., and Zhang, L. (2016). Adaptive fourier decomposition based ecg denoising. *Computers in Biology and Medicine*, 77:195–205.
- WHO (2021). Cardiovascular diseases (cvds). Disponível em: [https://www.who.int/en/news-room/fact-sheets/detail/cardiovascular-diseases-\(cvds\)](https://www.who.int/en/news-room/fact-sheets/detail/cardiovascular-diseases-(cvds)).
- Yang, P., Wang, D., Zhao, W.-B., Fu, L.-H., Du, J.-L., and Su, H. (2021). Ensemble of kernel extreme learning machine based random forest classifiers for automatic heartbeat classification. *Biomedical Signal Processing and Control*, 63:102138.
- Zhao, Q. and Zhang, L. (2005). Ecg feature extraction and classification using wavelet transform and support vector machines. In *2005 International Conference on Neural Networks and Brain*, volume 2, pages 1089–1092. IEEE.
- Zheng, J., Zhang, J., Danioko, S., Yao, H., Guo, H., and Rakovski, C. (2020). A 12-lead electrocardiogram database for arrhythmia research covering more than 10,000 patients. *Scientific Data*, 7(1):1–8.