



Universidade de Brasília

Instituto de Ciências Exatas
Departamento de Ciência da Computação
Departamento de Engenharia Elétrica

**Diagnóstico semiautomático de arritmias cardíacas
utilizando técnicas de aprendizado de máquina e
transformadas Wavelet**

Yuri Castro do Amaral

Monografia apresentada como requisito parcial
para conclusão do Curso de Engenharia da Computação

Orientador
Prof. Dr. Pedro de Azevedo Berger

Brasília
2019

Dedicatória

Dedico este trabalho a minha família e amigos por sempre acreditarem em mim.

Agradecimentos

Primeiramente, a Deus por sempre estar presente na minha vida e guiar os meus passos. Aos meus pais e ao meu irmão por sempre estarem do meu lado. A minha tia Rita por não medir esforços para me ajudar quando preciso. Aos meus colegas de curso que se tornaram grandes amigos para toda vida. Ao meu orientador Pedro Berger pela oportunidade de elaboração deste trabalho e aos membros da banca que se dispuseram a avaliá-lo.

Resumo

Normalmente, o diagnóstico de arritmias cardíacas é realizado de forma manual pelos médicos através da análise das ondas do eletrocardiograma (ECG). Desta forma, surgiu a necessidade da implementação de sistemas que automatizem a detecção para agilizar e melhorar a precisão do processo. Este trabalho propõe uma ferramenta computacional, utilizando técnicas de aprendizado de máquina, para classificação de batimentos cardíacos dentre quatro categorias propostas pela *Association for the Advancement of Medical Instrumentation* (AAMI): batimentos normais, ectópicos ventriculares, ectópicos supraventriculares e fusão. Os classificadores utilizados (redes neurais, máquina de vetores de suporte e análise do discriminante linear) foram treinados utilizando sinais do Arrhythmia database (MIT-BIH)

Ferramentas para detecção de doenças em sinais biológicos costumam lidar com conjuntos de dados desbalanceados, pois os eventos de interesse ocorrem raramente. Este trabalho propõe a geração de dados sintéticos utilizando o algoritmo *Adaptive Synthetic Sampling Approach* (ADASYN) para equilibrar as quantidades de exemplos de treino dos modelos preditivos. Além disso, utilizou-se o algoritmo *Sequential Forward Floating Selection* (SFFS) para a seleção automática das características que mais contribuíram na construção dos mesmos. O modelo proposto obteve 89% de sensibilidade para a classe dos batimentos de fusão e 94% para os ectópicos supraventriculares, as quais são as classes com menos amostras nos conjuntos de treino e teste. A aplicação da técnica de reamostragem se mostrou eficaz e possibilita que trabalhos semelhantes a utilizem para contornar o desbalanceamento.

Palavras-chave: Aprendizado de máquina, Classificação, Arritmia cardíaca, Eletrocardiograma, AAMI, ADASYN

Abstract

Usually, the diagnosis of cardiac arrhythmias is performed manually by doctors through the analysis of electrocardiogram waves. Thus, it came up the need to develop systems that automate the detection for improve process speed and accuracy. This work proposes a computational tool, using machine learning techniques, to classify heartbeat in four categories developed by *Association for the Advancement of Medical Instrumentation* (AAMI): normal beats, ventricular ectopic beats, supra-ventricular ectopic beats and fusion beats. The classifiers used (neural networks, support vector machine and linear discriminant analysis) were trained using signals from the Arrhythmia database (MIT-BIH)

Tools for disease detection using biological signals often deal with imbalanced datasets, because events of interest are rare. This work applies synthetic data generation using the *Adaptive Synthetic Sampling Approach* (ADASYN) algorithm to balance the samples quantities used for training the predictive models. In addition, the *Sequential Forward Floating Selection* (SFFS) algorithm was used to automatically select the most important features for building those models. The proposed model achieved sensitivity of 89% for the fusion beats class and 94% for the supra-ventricular ectopic beats class, which correspond to the ones with the smallest sample quantities on the training and test sets. The application of the resampling technique is effective and allows similar works to use it for deal with imbalacing problem.

Keywords: Machine Learning, Classification, Cardiac arrhythmia, Electrocardiogram, AAMI, ADASYN

Sumário

1	Introdução	1
1.1	Descrição do problema	1
1.2	Objetivos	2
1.2.1	Objetivo geral	2
1.2.2	Objetivos específicos	2
1.3	Descrição dos capítulos	3
2	Referencial Teórico	4
2.1	Eletrocardiograma	4
2.1.1	Arritmias cardíacas	5
2.2	Aprendizado de máquina	5
2.2.1	Classificação	6
2.2.2	Métricas de desempenho	6
2.2.2.1	Validação cruzada	9
2.2.3	O processo de mineração de dados	10
2.2.3.1	Coleta de dados e preparação	10
2.2.3.2	Escolha dos modelos	10
2.2.3.3	Ajuste dos parâmetros	11
2.2.3.4	Seleção de características	11
2.2.4	Classificadores	14
2.2.4.1	Redes neurais artificiais	14
2.2.4.2	Máquina de vetores de suporte	15
2.2.4.3	Análise do discriminante linear	16
2.2.5	O problema do desbalanceamento	17
2.2.5.1	ADASYN	18
3	Classificação semiautomática de arritmias cardíacas no ECG	20
3.1	Pré-processamento dos sinais	20
3.1.1	Transformada <i>Wavelet</i>	20

3.1.2	Remoção de artefatos	21
3.2	Segmentação	22
3.2.1	detecção dos picos R Pan-Tompkins	22
3.3	Extração de características	22
3.3.1	Intervalo RR	23
3.3.2	<i>Wavelet</i>	23
3.3.3	Estatísticas temporais	24
3.3.4	Características morfológicas	24
3.4	Seleção de características	24
3.5	Classificação	24
4	Metodologia	26
4.1	Banco de dados	26
4.2	AAMI	28
4.3	Divisão dos conjuntos de dados	28
4.4	Etapas	29
4.4.1	Pré-processamento	30
4.4.2	Segmentação	32
4.4.3	Extração de características	33
4.4.4	Seleção de características e ajuste dos hiperparâmetros	34
4.4.4.1	métrica para seleção	34
4.4.4.2	Seleção de características	35
4.4.4.3	ajuste dos hiperparâmetros	36
4.4.5	Reamostragem	38
4.4.6	Classificação	39
5	Resultados	41
5.1	Métricas de desempenho	41
5.2	Experimentos	42
5.2.1	Conjunto original de dados	42
5.2.2	Conjunto com dados reamostrados	43
5.2.3	Características selecionadas	44
5.3	Comparação com trabalhos da literatura	45
6	Conclusão	50
6.1	Trabalhos futuros	50
	Referências	52

Lista de Figuras

2.1	Eletrocardiograma normal - Figura adaptada.	4
2.2	Figura adaptada - Posicionamento dos nós SA e AV em um coração simplificado (átrio direito em azul no canto superior esquerdo da figura).	6
2.3	Modelo de classificação baseado no mapeamento entre os conjuntos X e y através da função f	7
2.4	Exemplo de matriz de confusão para três classes.	7
2.5	Figura baseada em [1] - Exemplo validação cruzada.	9
2.6	Esquemático de divisão do conjunto de dados original com etapa de ajuste dos hiperparâmetros.	12
2.7	Figura inspirada em [2] - Neurônio artificial.	14
2.8	Hiperplanos possíveis que delimitam as classes.	15
2.9	Larguras das margens para dois hiperplanos construídos sobre o mesmo conjunto de dados.	16
2.10	Distribuições gaussianas para três classes e fronteiras de decisão geradas pela LDA.	17
2.11	Exemplo para o algoritmo ADASYN com $K = 5$, os quadrados azuis e os círculos vermelhos representam as classes majoritária e minoritária respectivamente. Duas amostras sintéticas foram geradas sobre o vetor de diferença entre X_1 e X_2 e uma entre X_1 e X_3	19
3.1	Decomposição DWT em dois níveis.	21
3.2	Características relacionadas ao intervalo RR para um segmento de um ECG.	23
4.1	Divisão interpaciente do banco de dados.	30
4.2	Fluxo aplicado para a ferramenta de detecção desenvolvida.	31
4.3	Resultado do pré-processamento para um trecho de um registro do MIT-BIH. Na parte superior, o sinal original contaminado com os artefatos e, na inferior, o sinal filtrado.	32
4.4	Distribuição das amostras do conjunto de validação. À esquerda para o <i>fold</i> 1 e à direita para o <i>fold</i> 2.	34

4.5	Seleção de características via SFFS com LDA. O eixo y mostra o jk_{index} para cada subconjunto de características do eixo x. A curva em azul corresponde ao desempenho médio, e a área ao redor ao desvio padrão nos <i>folds</i>	36
4.6	Mapa de calor com os resultados do <i>GridSearchCV</i> para a SVM.	37
4.7	Mapa de calor com os resultados do <i>GridSearchCV</i> para a MLP.	38
4.8	Distribuição das amostras entre as classes para o conjunto de teste e treino respectivamente.	39
4.9	Distribuição das amostras, após o balanceamento, das classes para o conjunto treino.	40
5.1	Matrizes de confusão normalizadas resultantes da avaliação da MLP, SVM e LDA em DS2, seguindo as especificações do experimento 1.	47
5.2	Matrizes de confusão normalizadas resultantes da avaliação da MLP, SVM e LDA em DS2, seguindo as especificações do experimento 2.	48
5.3	Matrizes de confusão normalizadas resultantes da avaliação da MLP, SVM e LDA em DS2, seguindo as especificações do experimento 3.	49

Lista de Tabelas

4.1	Informações dos pacientes pertencentes à base de dados MIT-BIH	27
4.2	Agrupamentos de classes proposto pela AAMI	29
4.3	Divisão dos pacientes ao longo dos conjuntos definidos pela divisão interpa- ciente	30
5.1	Métricas resultantes do experimento 1 para os três modelos classificatórios. .	42
5.2	Métricas resultantes do experimento 2 para os três modelos classificatórios. .	43
5.3	Métricas resultantes do experimentos 3 para os três modelos classificatórios. .	44
5.4	Comparação das métricas do modelo proposto neste projeto (LDA-EXP3) com alguns trabalhos da literatura relacionados.	45

Lista de Abreviaturas e Siglas

AAMI *Association for the Advancement of Medical Instrumentation.*

Acc acurácia.

ADASYN *Adaptive Synthetic Sampling Approach.*

AM aprendizado de máquina.

AV junção atrioventricular.

DWT transformada *Wavelet* discreta.

ECG eletrocardiograma.

FN Falso negativo.

FP Falso positivo.

HOS estatísticas de alta ordem.

KNN K-vizinhos mais próximos.

LDA *Linear Discriminant Analysis.*

MIT-BIH Arrhythmia database.

MLP *Multilayer perceptron.*

RNA Rede neural artificial.

SA nó sinoatrial.

SFFS *Sequential Forward Floating Selection.*

SVM *Support vector machine.*

TN Verdadeiro negativo.

TP Verdadeiro positivo.

TPR Taxa de verdadeiros positivos.

Capítulo 1

Introdução

Segundo Thaler et al. [3], as arritmias cardíacas se manifestam através de variações na frequência, regularidade ou origem dos impulsos elétricos originados pelas células do coração. Os batimentos irregulares ocorrem tanto de forma pontual quanto sustentada, podendo os últimos perdurarem durante a vida inteira de uma pessoa. Algumas categorias de arritmia não representam nenhuma ameaça a vida do paciente, entretanto, existem as graves, que em algumas situações, podem levar à morte súbita dos mesmos [3].

As mortes cardíacas súbitas e arritmias representam cerca de 15 a 20% do total global de óbitos [4], um percentual crítico que gera grande preocupação por parte dos especialistas em realizar diagnósticos rápidos e precisos para tratá-las, em cada caso, da maneira mais adequada.

Muitas arritmias cardíacas possuem manifestações clínicas como palpitações e tonturas, contudo, o ideal é que elas sejam identificadas previamente para evitar implicações na saúde do paciente que possam levar à morte [3]. Deste modo, as arritmias podem ser identificadas analisando anormalidades nas ondas do eletrocardiograma (ECG), um exame não invasivo responsável por monitorar a atividade elétrica do coração.

Geralmente, as arritmias aparecem no ECG através de variações temporais e morfológicas em sequências de batimentos do sinal. Tais eventos viabilizam a detecção das mesmas através da inspeção visual dos registros pelos médicos [5].

1.1 Descrição do problema

As arritmias ocorrem em momentos específicos do dia, desta forma, longos registros, chamados de registro de ritmo, são necessários para monitorá-las adequadamente e podem chegar a ter horas de duração [6].

Pelo longo tamanho dos registros de ritmo, o diagnóstico manual é um processo demorado e cansativo, e por estes motivos, bastante suscetível a erros. A utilização de uma

ferramenta computacional que seja capaz de identificar e classificar batimentos cardíacos arrítmicos pode auxiliar os médicos na precisão e agilidade no processo de diagnóstico. Além disso, as ferramentas podem se estender além da função de apenas distinguir entre batimentos normais e anormais para detectar categorias específicas de arritmia, pois, cada uma tem tratamentos e consequências distintas.

A construção destas ferramentas de carácter semiautomático vem sendo vastamente explorada na literatura por vários autores e já alcançou resultados bastante promissores.

1.2 Objetivos

1.2.1 Objetivo geral

Este trabalho visa desenvolver, analisar e comparar modelos computacionais de aprendizado de máquina (e.g. redes neurais, máquinas de vetores de suporte e análise do discriminante linear) que classifiquem batimentos cardíacos em quatro grupos: batimentos normais, batimentos ectópicos ventriculares, batimentos ectópicos supraventriculares e batimentos de fusão, utilizando os sinais adquiridos do *Arrhythmia database* (MIT-BIH) [7], uma base de dados comum em trabalhos neste domínio.

1.2.2 Objetivos específicos

- Elaborar uma ferramenta de detecção que siga as principais sugestões da *Association for the Advancement of Medical Instrumentation* (AAMI)¹, uma organização responsável em propor recomendações para tecnologias médicas. Estas recomendações permitem que os resultados sejam padronizados e conseqüentemente seja possível compará-los com soluções existentes que seguiram os mesmos padrões.
- Pelo fato dos batimentos ectópicos serem raros, os sinais são majoritariamente dominados pelos de característica normal. Geralmente, este desbalanceamento provoca um viés na classificação de alguns batimentos arrítmicos para a classe normal, deste modo, é preciso aplicar técnicas que contornem as consequências proporcionadas pelo desbalanceamento. Assim, outro objetivo é analisar o impacto da aplicação de uma técnica para geração de dados sintéticos (ADASYN) de modo a equilibrar a quantidade de amostras entre as classes e verificar o ganho de desempenho nas classificações.

¹<https://www.aami.org/membershipcommunity/content.aspx?ItemNumber=1292&navItemNumber=4603>

- Aplicar algoritmo para seleção automática de características a fim de melhorar o desempenho geral da classificação e a interpretabilidade dos modelos explicitando quais foram selecionadas.
- Por fim, o último objetivo é realizar uma análise completa da literatura e comparar as métricas obtidas através dos modelos gerados com trabalhos semelhantes.

1.3 Descrição dos capítulos

Este trabalho é dividido em 5 capítulos, descritos a seguir:

- Capítulo 1: O primeiro capítulo é a introdução, cujos propósitos são descrever o contexto do trabalho, explicar, de maneira geral, o problema ao qual este projeto visa solucionar e propor uma solução, explicitando pontualmente cada um dos objetivos.
- Capítulo 2: No referencial teórico são esclarecidos todos os conceitos necessários para fundamentar as técnicas e algoritmos utilizados ao longo do presente trabalho.
- Capítulo 3: No terceiro capítulo é feita uma breve descrição das etapas adotadas no processo de diagnóstico semiautomático de arritmias cardíacas no eletrocardiograma. Além disso, também é apresentada uma revisão geral da literatura apontando os procedimentos adotados por outros trabalhos relacionados.
- Capítulo 4: O capítulo de metodologia tem como finalidade apresentar as informações e características do banco de dados de sinais ECG utilizado (MIT-BIH). Elucidar, de maneira detalhada, todas as etapas utilizadas para preparação dos sinais e construção dos modelos classificatórios, especificamente as de pré-processamento, segmentação, extração de características, seleção de características, ajustes dos hiperparâmetros, reamostragem e classificação.
- Capítulo 5: Nos resultados estão descritos os experimentos realizados utilizando os conjuntos de dados elaborados nas etapas anteriores, além das métricas e matrizes de confusão associadas a cada um dos classificadores treinados. Por fim, neste capítulo também é feita uma comparação dos resultados relativos ao melhor modelo gerado com alguns trabalhos da literatura aos quais utilizaram abordagens similares.
- Capítulo 6: A conclusão é responsável por fazer uma breve análise crítica dos resultados alcançados e expor algumas propostas para trabalhos futuros.

Capítulo 2

Referencial Teórico

2.1 Eletrocardiograma

O eletrocardiograma (ECG) é uma ferramenta médica, barata e não invasiva responsável por registrar, através de eletrodos na pele, os potenciais elétricos originados do impulso cardíaco, vastamente utilizado no diagnóstico de doenças como infartos e arritmias [8].

Um ECG normal com dois batimentos pode ser visto na Figura 2.1, é constituído pelas ondas P e T e pelas ondas Q, R e S, que juntas compõem o complexo QRS. As ondas P e as pertencentes ao complexo QRS são ondas de despolarização, visto que se originam por despolarizações de células atriais e ventriculares respectivamente. Já a onda T é uma onda de repolarização por ser originada no momento da repolarização ventricular [8].

Uma característica de interesse do ECG é o intervalo RR, a distância temporal entre dois picos R consecutivos a qual é diretamente ligada à frequência cardíaca.

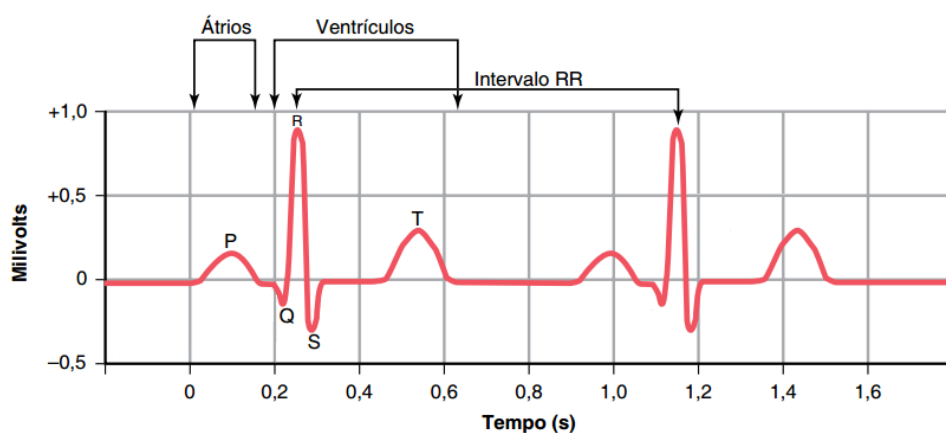


Figura 2.1: Eletrocardiograma normal - Figura adaptada (Fonte: [8]).

O coração é um órgão complexo, desta forma, várias combinações e posicionamentos de eletrodos, chamadas de derivações, são necessárias para que toda a atividade elétrica seja monitorada [3]. O ECG padrão é composto por 12 derivações, onde cada uma é mais perceptível à determinada parte do coração, característica crucial para a detecção de algumas doenças [3].

2.1.1 Arritmias cardíacas

A frequência cardíaca normal de um ser humano é de cerca de 60 a 100 batimentos por minuto, eventos fora dessa faixa, isolados ou sustentados, são considerados arritmias [3]. Nem todas as arritmias são danosas, para as que são, o diagnóstico rápido através do eletrocardiograma é de extrema importância para evitar, por exemplo, a morte súbita do paciente. Como as arritmias podem ocorrer em momentos específicos e imprevisíveis, longas medidas de batimentos de uma ou mais derivações, chamadas de registros de ritmo, são realizadas para captá-los [3].

As células cardíacas encarregadas em gerar as ondas elétricas normais são denominadas células de marca-passo. Localizado no átrio direito, o nó sinoatrial (SA) corresponde ao conjunto principal dessas células, que em situações normais, é o marca-passo cardíaco mais acelerado. Os ritmos ectópicos são aqueles gerados por células de marca-passo diferentes das do nó SA como as pertencentes à junção atrioventricular (AV) [3].

A Figura 2.2 ilustra a estrutura simplificada do coração e as posições do nó sinoatrial e junção atrioventricular. As arritmias supraventriculares são as originadas nos átrios ou no nó AV, por outro lado, as de origem abaixo do nó AV são chamadas de ventriculares. É de extrema importância distinguir as categorias de arritmia tendo em vista que o tratamento e gravidade são diferentes [3].

2.2 Aprendizado de máquina

Segundo Mohri et al. [9], aprendizado de máquina (AM), ou em inglês *Machine Learning*, é definido como modelos computacionais que se utilizam da experiência para aumentar o desempenho ou realizar previsões em outros conjuntos de dados.

Existem dois grandes grupos de aprendizado: de um lado, o supervisionado, onde rótulos das categorias de algumas instâncias de dados são fornecidos para que o modelo passe por uma etapa de treinamento que o permita realizar previsões em outros conjuntos de dados não categorizados. Além disso, existem os modelos de AM não supervisionados, os quais não rotulam os exemplos de entrada de tal modo que as associações possam ser feitas, por exemplo, por agrupamento dos dados com características próximas [9].

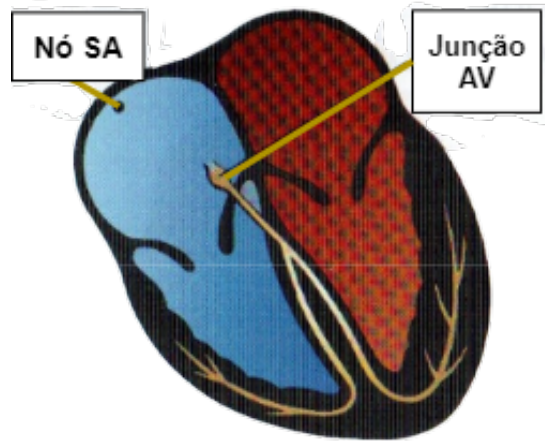


Figura 2.2: Figura adaptada - Posicionamento dos nós SA e AV em um coração simplificado (átrio direito em azul no canto superior esquerdo da figura). (Fonte: [3]).

2.2.1 Classificação

Os dados de entrada também chamados de exemplos, amostras ou *data points* são tuplas X que correspondem a um grupo de características/atributos que o descrevem. Cada tupla é associada à categoria predefinida ao qual pertence (y), também chamada de classe ou rótulo. A classificação é um dos principais problemas tratados pelo AM e pode ser dividida em duas etapas: A primeira delas é a de treinamento que consiste em construir o modelo de classificação que possa exercer um aprendizado sobre um grupo de dados de entrada denominado conjunto de treinamento. Já a segunda etapa diz respeito à categorização dos dados de outro conjunto, o de teste, utilizando o modelo construído anteriormente. Os dados de teste devem ser diferentes do da etapa de treinamento, isto é, não pode ter sido usados para treinar o classificador para que a avaliação seja fidedigna [10].

Como exemplificado na Figura 2.3 (inspirada em [11]), a classificação pode ser entendida como uma função $f()$ que mapeia um vetor de características X em um rótulo y , onde f pode ser, por exemplo, classificadores como as redes neurais [10].

2.2.2 Métricas de desempenho

Em problemas de classificação é preciso registrar as previsões do modelo de AM submetido ao conjunto de teste para medir seu desempenho. A matriz de confusão é uma tabela que apresenta a performance geral de um classificador [12]. Uma linha e uma coluna são

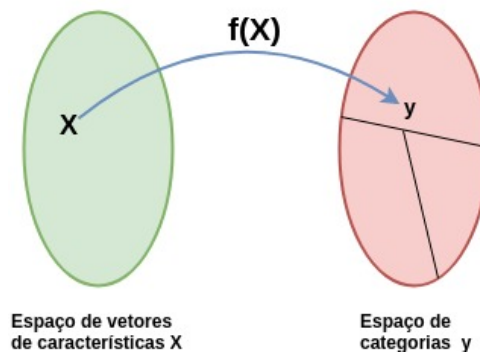


Figura 2.3: Modelo de classificação baseado no mapeamento entre os conjuntos X e y através da função f .

alocadas para cada classe do problema. As linhas representam as classes verdadeiras e as colunas as previstas, isto é, para cada elemento da tabela, o valor em questão refere-se a quantos exemplos da classe denotada pela linha foram classificados como a classe da coluna. Boas classificações possuem os valores concentrados na diagonal principal da matriz, o que significa que os dados foram previstos corretamente [13].

A Figura 2.4 apresenta um exemplo de matriz de confusão para seis amostras divididas em três classes. Uma escala de cores foi adicionada para melhorar a visualização, onde a intensidade da cor é diretamente proporcional à quantidade de dados de cada entrada da matriz.

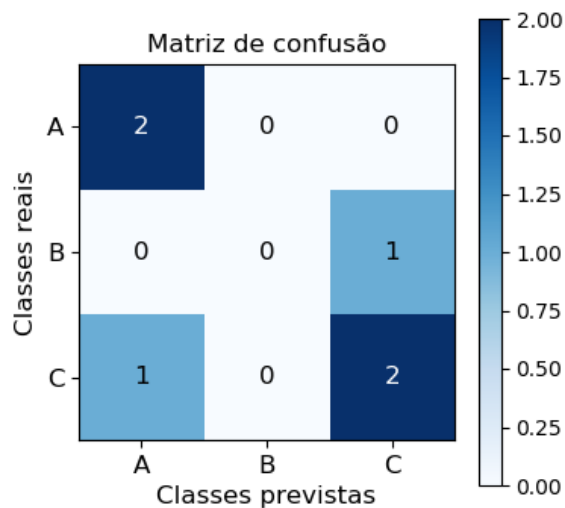


Figura 2.4: Exemplo de matriz de confusão para três classes.

A matriz de confusão registra toda informação da classificação, porém, por se tratar

de um registro geral, se torna difícil compará-la com as geradas por outros modelos. Consequentemente, é importante representá-la em valores específicos que sintetizem e deem diferentes perspectivas do desempenho obtido [12]. Estes valores mencionados correspondem às métricas de desempenho as quais podem ser calculadas a partir dos seguintes elementos da matriz:

- Verdadeiro positivo (TP): São exemplos pertencentes a uma dada classe previstos corretamente.
- Falso positivo (FP): São exemplos não pertencentes a uma dada classe, mas que lhe foram atribuídos.
- Verdadeiro negativo (TN): São exemplos não pertencentes a uma dada classe que, corretamente, não lhe foram atribuídos.
- Falso negativo (FN): São exemplos pertencentes a uma dada classe, mas que não lhe foram atribuídos.

A acurácia (Acc), vide Equação 2.1, é uma métrica comum em AM. Seu objetivo é medir globalmente a taxa de acertos em relação ao número total de predições. Contudo, não é muito recomendada para conjuntos de dados desbalanceados, pois mesmo com baixa taxa de acerto para as classes de menor quantidade, o valor da acurácia fica condicionado majoritariamente pela classe dominante [12].

$$Acurácia = \frac{TP + TN}{TP + TN + FP + FN} \quad (2.1)$$

Sensibilidade, *Recall* ou Taxa de verdadeiros positivos (TPR) é descrita pela Equação 2.2 e relaciona os TP ao número total de exemplos de uma classe, isto é, representa o percentual de acerto de uma classe associado a todos os exemplos realmente pertencentes a ela. Para uma classe de arritmia, por exemplo, é a taxa de pacientes doentes devidamente diagnosticados.

$$Sensibilidade = \frac{TP}{TP + FN} \quad (2.2)$$

A precisão, na Equação 2.3, é a fração entre os verdadeiros positivos e a soma dos verdadeiros com os falsos positivos, ou melhor, a taxa de exemplos que fazem parte de uma classe dentro todos atribuídos a ela.

$$Precisão = \frac{TP}{TP + FP} \quad (2.3)$$

*Cohen's Kappa*¹ (k) compara a concordância entre dois avaliadores, em classificação, rótulos e predições. É apresentada na Equação 2.4, onde P_o denota a taxa observada de concordância e P_e a taxa esperada de concordância ao acaso. Resumidamente, representa o grau de concordância excluindo as concordâncias casuais [14].

$$k = \frac{P_o - P_e}{1 - P_e} \quad (2.4)$$

2.2.2.1 Validação cruzada

Quando um modelo de aprendizado de máquina detém boa performance em relação aos exemplos de treinamento, porém não apresenta boa generalização aos de teste, isto é, que não foram utilizados para construir o modelo, é possível que ele tenha sofrido sobreajuste ou, em inglês, *overfitting* [1].

Em situações onde o conjunto de dados é pequeno ou se deseja verificar de maneira mais robusta a generalização do modelo, pode-se aplicar a técnica de validação cruzada *k-fold*. Esta metodologia consiste em particionar o conjunto de dados em k subconjuntos de mesmo tamanho, chamados de *folds*, e reiteradamente selecionar um deles para teste e o restante para treino até que todos tenham sido utilizados para teste exatamente uma vez [12].

Os resultados globais são obtidos pela média das métricas calculadas em cada iteração. A Figura 2.5 mostra um exemplo de funcionamento do algoritmo. Os dados foram subdivididos em cinco *folds*, para as respectivas iterações em cada linha, os quadrados verdes são as partições de teste e os cinzas as de treinamento.

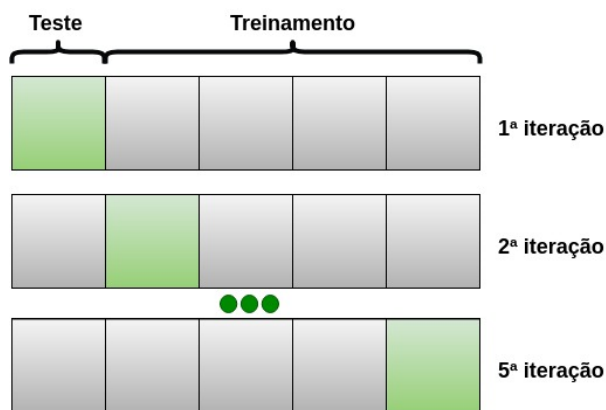


Figura 2.5: Figura baseada em [1] - Exemplo validação cruzada.

¹https://scikit-learn.org/stable/modules/generated/sklearn.metrics.cohen_kappa_score.html

2.2.3 O processo de mineração de dados

Segundo Marsland [15], o processo de mineração de dados é composto normalmente por algumas etapas, que serão descritas a seguir:

2.2.3.1 Coleta de dados e preparação

Algumas vezes o conjunto de dados não está inicialmente coletado e/ou preparado para o processo de aprendizado. Os dados de entrada podem se originar de fontes variadas como páginas *web* ou de bancos de dados governamentais ou abertos. Desta forma, os primeiros passos são carregá-los ou extraí-los, retirar informações relevantes e aplicar técnicas de processamento a fim de prepará-los para as etapas seguintes.

Em um contexto de aprendizado supervisionado é preciso que também sejam extraídas as categorias as quais pertencem os exemplos [15].

Alguns classificadores são influenciados por características não escaladas, podendo demorar a convergir no treinamento ou então priorizar as que possuem valores mais altos em detrimento das restantes. A normalização das características é uma etapa de processamento utilizada para contornar os efeitos colaterais citados anteriormente [10].

A normalização *Z-score* é mostrada na Equação 2.5, onde \mathbf{x} é uma instância, \mathbf{u} a média e \mathbf{s} o desvio padrão da distribuição dos dados de entrada.

$$z = \frac{x - u}{s} \quad (2.5)$$

2.2.3.2 Escolha dos modelos

A decisão de qual classificador usar varia dependendo do problema. Existem alguns critérios que podem ser utilizados para avaliar sua viabilidade [10]:

- Velocidade - Existem situações onde os recursos computacionais são limitados ou que o *hardware* é incapaz de treinar os modelos em tempo viável. Assim o custo computacional do classificador pode ser um fator determinante para sua escolha.
- Robustez - Problemas reais lidam com ruídos, informações indesejadas agregadas aos dados de entrada. Um modelo robusto é capaz de realizar classificações acuradas mesmos nestas situações.
- Escalabilidade - Com o advento da *internet*, grandes massas de dados são geradas por dia. A escalabilidade diz se um modelo é capaz de lidar com quantidades maiores de dados.
- Interpretabilidade - É um critério que se fundamenta na capacidade de interpretação da classificação. As árvores de decisão, por exemplo, constroem uma estrutura hie-

rárquica das características que define o fluxo de atribuição a uma classe, permitindo visualizar como a atribuição para um exemplo foi tomada.

2.2.3.3 Ajuste dos parâmetros

Geralmente os classificadores possuem parâmetros especiais que influenciam diretamente no aprendizado acerca dos dados de treinamento. Os chamados hiperparâmetros podem ser ajustados em um grupo de dados reservado chamado de validação. A Figura 2.6 ilustra uma metodologia de divisão do conjunto original de dados que inclui a etapa de ajuste de hiperparâmetros. Primeiramente, os dados são divididos em treino e teste, em seguida, o de treino subdividido em treino e validação.

Repetidamente atribui-se hiperparâmetros ao classificador e, após o treinamento, verifica-se sua performance sobre os dados de validação com o objetivo de encontrar a combinação que minimiza o erro global. Por fim, com os hiperparâmetros selecionados na etapa anterior, realiza-se a predição final sobre o conjunto de teste.

*Grid Search*² é uma solução para automatizar o ajuste dos hiperparâmetros. A partir de coleções de valores predefinidos de alguns parâmetros, treina-se o modelo combinando-os de todas as formas possíveis. Cada associação de valores é submetida à validação cruzada de forma que os hiperparâmetros selecionados sejam aqueles que obtiverem o maior valor de uma métrica específica.

A complexidade deste algoritmo é alta, tendo em vista que adicionar novos valores acarreta em várias combinações de parâmetros treinados várias vezes ao longo dos *folds*. Portanto, é importante limitar a quantidade de parâmetros, valores e *folds* para que o tempo de execução não se torne inviável.

2.2.3.4 Seleção de características

O objetivo da seleção de características é reduzir a dimensionalidade dos vetores de entrada coletados anteriormente, restringindo-o apenas às características que possuam maior contribuição para a performance da classificação, isto é, as que não contenham informações redundantes ou irrelevantes [12].

Sequential Forward Floating Selection (SFFS)³ é um algoritmo guloso para seleção de características. Seu funcionamento consiste em iterativamente avaliar vários subconjuntos de características através de um classificador e selecionar, a cada passo, o que detiver a maior desempenho.

²https://scikit-learn.org/stable/modules/generated/sklearn.model_selection.GridSearchCV.html#sklearn.model_selection.GridSearchCV

³http://rasbt.github.io/mlxtend/user_guide/feature_selection/SequentialFeatureSelector/

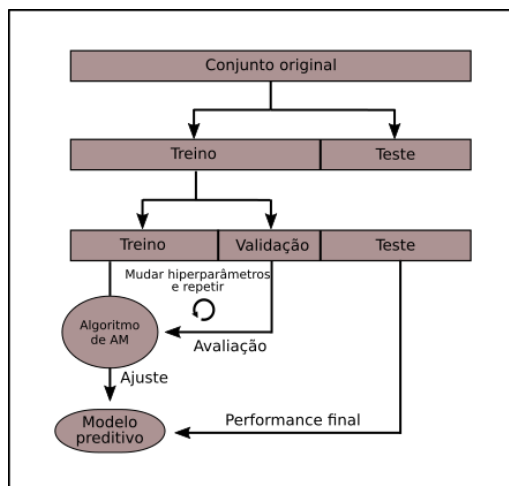


Figura 2.6: Esquemático de divisão do conjunto de dados original com etapa de ajuste dos hiperparâmetros (Fonte: [1]).

Considerando m como a dimensão do vetor de entrada (X), n a dimensão almejada do vetor de saída (S) e k a dimensão do vetor em cada iteração, o algoritmo é executado até que o vetor de saída S tenha exatamente n características ($k = n$), dado que n seja menor que m .

O algoritmo é implementado através dos seguintes quatro passos:

1. Inicia-se o subconjunto de saída S com nenhuma característica (conjunto nulo) e k igual à 0.
2. De forma individual e alternada, cada característica de X que não esta contida em S é adicionada a este subconjunto. O subconjunto atual é utilizado para treinar um classificador que, em seguida, avalia dados de validação utilizando alguma função como critério, como por exemplo a precisão (este processo pode ser feito através de validação cruzada). A característica adicionada a qual permitiu que o subconjunto S obtivesse o maior valor do critério citado anteriormente, é selecionada para fazer parte de S e incrementa-se o valor de k em 1.
3. Em seguida, por ser *floating*, a etapa de exclusão condicional é executada. De forma análoga, se k maior que 2, cada característica é removida alternadamente do conjunto S e este subconjunto é reavaliado utilizando o critério. A característica removida a qual permitiu que o subconjunto S obtivesse o maior valor do critério é removida do conjunto S e k é decrementado em 1, caso a remoção não seja refletida no aumento de valor do critério, esta etapa é ignorada.
4. A execução volta para o passo 2 e o algoritmo continua até que o valor de k seja igual ao de n .

A título de exemplificação, consideremos um ambiente em que o espaço de entrada X possua quatro características ($m = 4$) e que desejamos selecionar três delas ($n = 3$). O conjunto D corresponde às características ainda disponíveis para serem avaliadas, logo:

$$X = \{a, b, c, d\} \quad S = \{\emptyset\}, k = 0 \quad D = X - S = \{a, b, c, d\} \quad (2.6)$$

Na primeira iteração, cada característica de D é alternadamente adicionada à S e avaliada pelo critério J :

$$S' = \{a\}, J' = 0.2 \quad S'' = \{b\}, J'' = 0.21 \quad S''' = \{c\}, J''' = 0.1 \quad S'''' = \{d\}, J'''' = 0.15 \quad (2.7)$$

Podemos observar que J é maximizado quando adicionamos b à S , logo, b foi selecionado para compor S e k passa de 0 para 1:

$$X = \{a, b, c, d\} \quad S = \{b\}, k = 1 \quad D = X - S = \{a, c, d\} \quad (2.8)$$

Na implementação do algoritmo, se o k for menor ou igual a 2 a etapa de exclusão condicional não é executada. Assim, na segunda interação, novamente adicionamos cada característica de D à S e as avaliamos:

$$S' = \{b, a\}, J' = 0.31 \quad S'' = \{b, c\}, J'' = 0.29 \quad S''' = \{b, d\}, J''' = 0.25 \quad (2.9)$$

Portanto, a característica a é escolhida para compor S e k se torna 2:

$$X = \{a, b, c, d\} \quad S = \{a, b\}, k = 2 \quad D = X - S = \{c, d\} \quad (2.10)$$

Na terceira iteração, avalia-se novamente as características de D :

$$S' = \{a, b, c\}, J' = 0.33 \quad S'' = \{a, b, d\}, J'' = 0.32 \quad (2.11)$$

A característica c é escolhida para compor S e k se torna 3.

$$X = \{a, b, c, d\} \quad S = \{a, b, c\}, k = 3 \quad D = X - S = \{d\} \quad (2.12)$$

Como $k = 3$, a etapa de exclusão condicional é executada, cada característica é removida individualmente e temos os seguintes valores de critério:

$$S' = \{a, c\}, J' = 0.4 \quad S'' = \{b, c\}, J'' = 0.29 \quad S''' = \{a, b\}, J''' = 0.31 \quad (2.13)$$

Como o conjunto $S' = \{a, c\}$ obteve o J maior que o do S atual ($S = \{a, b, c\}$), S se torna S' e o k é decrementado:

$$X = \{a, b, c, d\} \quad S = \{a, c\}, k = 2 \quad D = X - S = \{b, d\} \quad (2.14)$$

O processo segue até que k seja igual à n .

2.2.4 Classificadores

2.2.4.1 Redes neurais artificiais

Rede neural artificial (RNA) é um modelo de aprendizado de máquina inspirado nas redes neurais biológicas. É composta por associações ponderadas de componentes chamados neurônios artificiais, mostrado na Figura 2.7. Cada neurônio possui os seguintes elementos: [2]

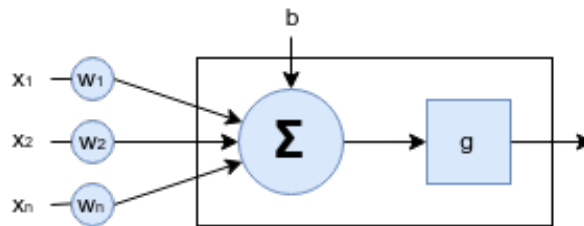


Figura 2.7: Figura inspirada em [2] - Neurônio artificial.

- Conexões de entrada: conjunto de entradas x_i com seus respectivos pesos w_i , sendo i um número inteiro de 1 a n que representa a quantidade de entradas. Uma das entradas possui valor fixo e é denominada bias (b).
- Função de entrada: função agregadora dos sinais através da combinação linear entre as conexões de entrada e seus respectivos pesos ($f = x_1.w_1 + x_2.w_2 + \dots + x_n.w_n$)
- Função de ativação: função g responsável por calcular os limites de ativação do resultado da função de entrada.
- Função de saída: função de saída do neurônio, normalmente é o próprio resultado da função de ativação.

Perceptron multicamadas, em inglês *Multilayer perceptron* (MLP), é uma RNA organizada em várias camadas de neurônios: uma de entrada, uma ou mais ocultas e uma de saída. Em uma estrutura tradicional, os neurônios de cada camada são ligados aos da camada seguinte com pesos em cada associação. A quantidade de neurônios na camada de entrada é igual à dimensão do vetor de características e na de saída ao número de classes (em problemas multiclasse). Já a quantidade de camadas ocultas e neurônios em cada uma delas são hiperparâmetros do modelo cuja seleção deve ser feita na etapa de ajuste dos mesmos [12].

O processo de aprendizado da MLP pode ser dado em duas etapas. A primeira consiste em submeter um vetor de características na camada de entrada, obter o resultado propagado pelas ligações entre os neurônios e, com essa informação, determinar a diferença entre a saída esperada e obtida utilizando alguma função de erro. Na segunda, o gradiente do erro calculado é propagado da saída para as camadas anteriores, atualizando os pesos das conexões com o propósito de minimizar o erro global. O processo da segunda etapa é conhecido como algoritmo de retropropagação do erro [2].

2.2.4.2 Máquina de vetores de suporte

Support vector machine (SVM), em português máquina de vetores de suporte, é um classificador ao qual objetivo é construir um hiperplano (fronteira de decisão) que separe duas classes em um espaço n-dimensional. A Figura 2.8 apresenta um plano com dados pertencentes à duas classes (círculo branco e círculo preto), as linhas tracejadas simbolizam alguns hiperplanos possíveis [10].

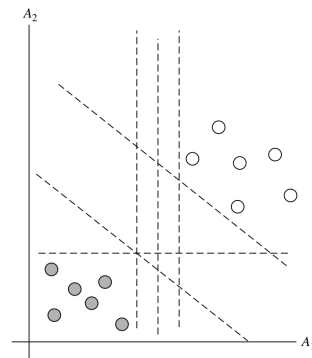


Figura 2.8: Hiperplanos possíveis que delimitam as classes (Fonte: [10]).

O objetivo da SVM é encontrar o hiperplano com máxima margem, a maior distância entre as projeções paralelas do hiperplano em questão, em ambas as direções, que tocam as amostras mais próximas de cada classe. A Figura 2.9 mostra dois hiperplanos e suas

respectivas margens. Devido a sua margem maior, foi escolhido o da direita, garantindo maior capacidade de generalização para o classificador correspondente [10].

Nem sempre os dados são linearmente separáveis, para lidar com essas situações, funções denominadas *kernel* são utilizadas para mapear o espaço de entrada em um de maior dimensão com o intuito de permitir a construção dos hiperplanos. As funções *kernel* podem ser lineares, sigmoidais, de base radial, polinomiais, etc.

Os conjuntos de dados podem conter valores irregulares ou atípicos, como por exemplo, uma amostra de uma classe (círculo preto) próxima à outra (círculos brancos) na Figura 2.8. Estes eventos acarretam em uma fronteira de decisão com margem inferior à ideal. Como solução, o SVM possui um fator C , que permite a algumas amostras serem classificadas erroneamente para aumentar o tamanho da margem, este fator é inversamente proporcional ao nível de classificação errônea.

Originalmente, o SVM é projetado apenas para classificações binárias. Já que muitos problemas do mundo real são multiclasse, existem algumas técnicas para adaptá-lo à estes contextos. Sendo K o número de classes, a metodologia um-contra-resto treina K classificadores binários, um para cada classe $k_i \in K$. Em cada um dos classificadores, considera-se as amostras pertencentes a k_i como a classe positiva, e todas as outras como a classe negativa, e utiliza-se os modelos treinados para decidir a classe final [12].

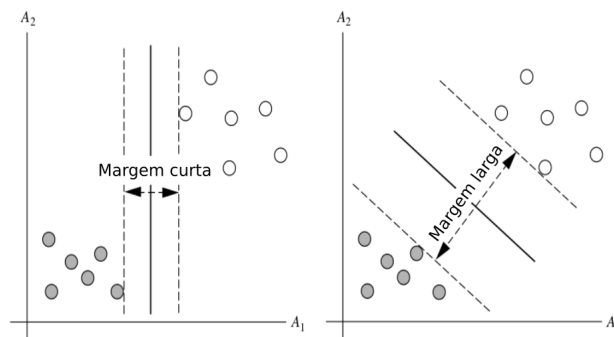


Figura 2.9: Larguras das margens para dois hiperplanos construídos sobre o mesmo conjunto de dados (Fonte: [10]).

2.2.4.3 Análise do discriminante linear

A Análise do Discriminante Linear, em inglês *Linear Discriminant Analysis* (LDA), é um modelo simples e não paramétrico projetado para redução de dimensionalidade e classificação que se utiliza de métodos probabilísticos para linearmente separar as amostras [16].

As fronteiras de decisão são definidas através da aplicação do teorema de *Bayes*. Para isso, a LDA parte do pressuposto de que as amostras de cada classe seguem distribuições gaussianas que compartilham a mesma matriz de covariância. Os parâmetros destas distribuições como as médias, matriz de covariância e probabilidades a *priori* de cada classe são calculadas a partir das amostras de treinamento [17].

A Figura 2.10 apresenta um problema classificatório com três classes modelado com a LDA. Nela, as distribuições gaussianas são retratadas pelas elipses, as quais possuem a mesma matriz de covariância e médias (centróides) distintas. As linhas correspondem às fronteiras de decisão originada pela regra de *Bayes*, e a partir delas é possível verificar a qual classe pertence uma amostra de teste.

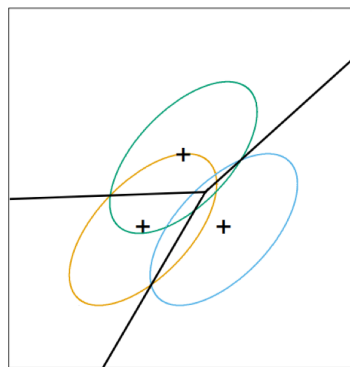


Figura 2.10: Distribuições gaussianas para três classes e fronteiras de decisão geradas pela LDA (Fonte: [17]).

2.2.5 O problema do desbalanceamento

É comum que problemas reais de AM lidem com quantidades desbalanceadas de amostras entre as classes. Quando a quantidade de exemplos de uma classe é predominante em relação às das outras, ela é chamada de majoritária, por outro lado, quando possui menos amostras é dita minoritária. Alguns modelos de AM podem ter perda de desempenho quando submetidos a dados com essa característica, sendo importante explorar técnicas que procurem resolver tal problema [18].

O desbalanceamento pode aparecer de dois modos, quando a classe de interesse é minoritária ou quando as instâncias são raras [19].

O diagnóstico de batimentos ectópicos, por exemplo, é um problema desbalanceado que apresenta as duas propriedades supracitadas. Por grande parte dos batimentos serem normais, as arritmias são pouco representadas, o que significa que as instâncias são raras, concomitantemente, são a classe de interesse para que o diagnóstico seja correto. Caso o

desbalanceamento não seja minimizado, os dados podem ficar sobreajustados em relação aos batimentos normais que, na prática, implica em não diagnóstico de pacientes doentes.

Existem algumas metodologias que podem ser aplicadas como tentativa de sanar este problema, por exemplo:

- Sub/Sobreamostragem aleatória - Alterar a quantidade de amostras a fim de equilibrar a representatividade das classes. Pode ser feita para as classes majoritárias através da seleção de parte dos dados (subamostragem) ou para as classes minoritárias com a replicação dos mesmos (sobreamostragem) [18].
- Geração de dados artificiais - Criar dados sintéticos utilizando técnicas sobre os dados reais, por exemplo, através de interpolações [18].

2.2.5.1 ADASYN

Adaptive Synthetic Sampling Approach (ADASYN) é uma técnica de sobreamostragem para geração de dados sintéticos. O objetivo deste algoritmo é gerar mais dados para as amostras que são mais difíceis de aprender. Primeiramente, calcula-se a diferença entre as quantidades das classes do problema para se definir, globalmente, quantas amostras precisam ser criadas. Em seguida, para cada amostra da classe minoritária busca-se os K-vizinhos mais próximos verificando quantos deles pertencem à majoritária. Desta forma, é possível construir uma distribuição de densidade que permite calcular quantos dados sintéticos precisam ser criados para uma amostra específica até que se atinja a quantidade global citada anteriormente, dando ênfase a àquelas que possuem vizinhanças predominantemente majoritárias [19].

Os dados sintéticos são criados sobre o vetor de diferença entre duas amostras minoritárias dentro da mesma vizinhança, multiplicado por um número aleatório que varia de zero a um, a Figura 2.11 exemplifica o funcionamento da interpolação [19].

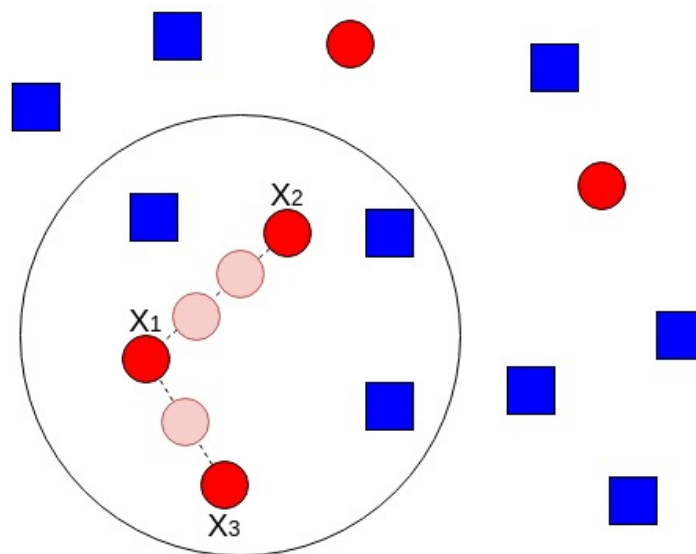


Figura 2.11: Exemplo para o algoritmo ADASYN com $K = 5$, os quadrados azuis e os círculos vermelhos representam as classes majoritária e minoritária respectivamente. Duas amostras sintéticas foram geradas sobre o vetor de diferença entre X_1 e X_2 e uma entre X_1 e X_3 .

Capítulo 3

Classificação semiautomática de arritmias cardíacas no ECG

A aplicação de técnicas de aprendizado de máquina em sinais ECG para a distinção automática de batimentos normais e anormais vem sendo vastamente explorada na literatura. No âmbito de classificação de batimentos ectópicos, os estudiosos seguem normalmente o seguinte fluxo: pré-processamento dos sinais, segmentação, extração de características, seleção de características e classificação. As subseções 2.3.1 à 2.3.4 irão discorrer sobre cada etapa do processo.

3.1 Pré-processamento dos sinais

A primeira etapa de classificação dos batimentos deve ser processar os sinais para remover propriedades de não interesse, com o intuito de melhorar o desempenho dos classificadores.

3.1.1 Transformada *Wavelet*

A Transformada *Wavelet* é um recurso matemático utilizado em sinais para remoção de ruídos, compressão, extração de características, etc. Pode ser vista como uma extensão da transformada de Fourier, cuja diferença é a representação do sinal em tempo-escala (múltiplas resoluções) ao invés de apenas frequência [6].

O sinal resultante da transformada é gerado através de operações de dilatação e translação sobre ondas específicas chamadas *Wavelets* mãe. Existem várias famílias dessas ondas, como a *Daubechies* e *Symlets*, cuja decisão de uso depende do contexto de utilização.

A transformada *Wavelet* discreta (DWT) pode ser implementada com uma associação de filtros passa-alta $g(n)$, passa-baixa $h(n)$ e subamostradores que selecionam metade das

amostras. As subbandas geradas pelos filtros passa-baixas são chamadas de aproximações e pelos passa-altas de detalhes. O processo consiste em decompor as aproximações de cada nível repetidamente em um número de níveis pré definidos [20].

A Figura 3.1 mostra a decomposição de um sinal em dois níveis utilizando a DWT, no primeiro, o sinal $x(n)$ é decomposto em aproximações (A_1) e detalhes (D_1) e subamostrado pela metade (círculos). Já no segundo, A_1 é decomposto em A_2 e D_2 e subamostrado novamente. As reticências na figura representam a extensão da DWT para mais níveis de decomposição.

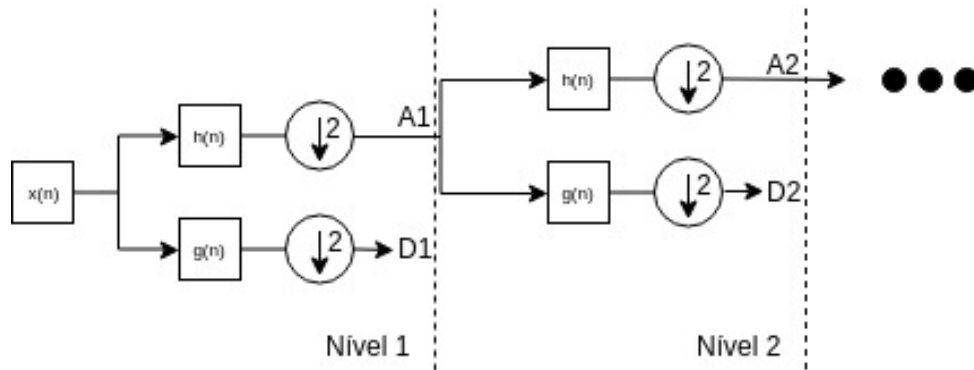


Figura 3.1: Decomposição DWT em dois níveis.

3.1.2 Remoção de artefatos

Naturalmente, sinais ECG possuem pequena amplitude [21]. Assim, podem ser contaminados por artefatos, sinais que influenciam negativamente a qualidade do processo de automação do diagnóstico.

Os artefatos relativos ao ECG podem se originar de várias fontes como a atividade bioelétrica de outras células que não sejam o coração ou até mesmo por particularidades dos instrumentos de medida como interferências eletromagnéticas, qualidade dos eletrodos utilizados ou sinais espúrios de equipamentos próximos [21].

Dois artefatos comuns em sinais ECG são a *Baseline Wander* e a interferência na linha de energia. O primeiro manifesta-se nas baixas frequências. É originado principalmente devido a perturbações entre a pele e os eletrodos, ocasionadas por exemplo, pela respiração do paciente. Já o segundo é um artefato bastante comum que se manifesta nas altas frequências [22].

Várias técnicas de remoção de artefatos são utilizadas na literatura como filtros de mediana [23, 22, 24], filtro morfológico [25] e baseados na transformada *Wavelet* [26, 27] sendo o último utilizado neste trabalho.

O funcionamento geral da remoção de artefatos utilizando a transformada *Wavelet* consiste em decompor o sinal em subbandas, diminuir ou zerar aquelas cuja frequência são dominadas pelos artefatos e, por fim, realizar a transformada inversa. Para este tipo de filtragem é necessário ter conhecimento das bandas de frequência onde se encontram os artefatos, a fim de se evitar a perda de informações pertencentes àquele sinal, como é o caso de componentes de frequência que diferem os batimentos normais dos ectópicos.

3.2 Segmentação

Com os sinais filtrados, a próxima etapa é a de segmentação. A classificação de arritmias é feita a nível dos batimentos, logo, os sinais ECG devem ser segmentados em trechos que contenham toda informação sobre um ciclo cardíaco. Um problema a parte da literatura é a delimitação, que consiste projetar algoritmos de detecção automática de pontos específicos do ECG como as ondas P, Q, R, S e T.

Quando a frequência de amostragem do sinal é conhecida, considera-se o pico R como o ponto central e seleciona-se um número fixo de amostras à direita e à esquerda para delimitar um segmento. A segmentação automática é de suma importância para a aplicação dos sistemas de classificação na prática, porém, os erros deste passo se propagam para as etapas posteriores.

3.2.1 detecção dos picos R Pan-Tompkins

Pan e *Tompkins* [28] é um algoritmo de detecção das ondas R que se fundamenta em quatro estágios:

Primeiramente, o sinal ECG é filtrado para atenuar artefatos e outros componentes do sinal à medida do possível. O filtro projetado seleciona, aproximadamente, a banda de 5 à 12 Hz, cuja energia é predominantemente pertencente ao complexo QRS. Em seguida, um filtro que aproxima uma derivada é aplicado com o objetivo de extrair informações de inclinação do ECG. Após isso, cada amostra do sinal é elevada ao quadrado para eliminar valores negativos e aumentar a contribuição das altas frequências. Por fim, uma janela deslizante de integração é utilizada para obter um sinal que contenha informações à respeito do complexo QRS. Com este sinal é possível definir a posição dos picos R [28].

3.3 Extração de características

A extração de características envolve calcular valores sobre cada batimento que representem bem as peculiaridades de cada classe. Além disso, nesta etapa também extraem-se as

classes para cada batimento. Na literatura, as classes são coletadas dos bancos de dados que fornecem os sinais através de anotações, elaboradas por profissionais da área médica, em amostras específicas informando os eventos que ocorreram naquele instante. Nos tópicos a seguir serão apresentadas algumas características comuns utilizadas por autores da área.

3.3.1 Intervalo RR

O intervalo RR (Figura 2.1) é o tempo entre dois batimentos subsequentes. Variações deste intervalo implicam em alterações morfológicas do ECG que podem ter sido ocasionadas por alguns tipos de arritmia [29]. Características relativas ao intervalo RR são muito populares na literatura [30, 23, 31].

A Figura 3.2 apresenta um trecho de aproximadamente 12 segundos para um ECG de 30 minutos. O círculo vermelho referencia o pico R de um segmento do sinal. Para o batimento em questão, a característica RR-anterior, em vermelho, diz respeito à distância temporal entre o pico R atual e o anterior, a RR-posterior, em verde, à distância entre o pico atual e o seguinte, a RR-local, em laranja, à média dos dez últimos intervalos RR e por fim, em roxo, a RR-global que é a média dos intervalos para todo o sinal.

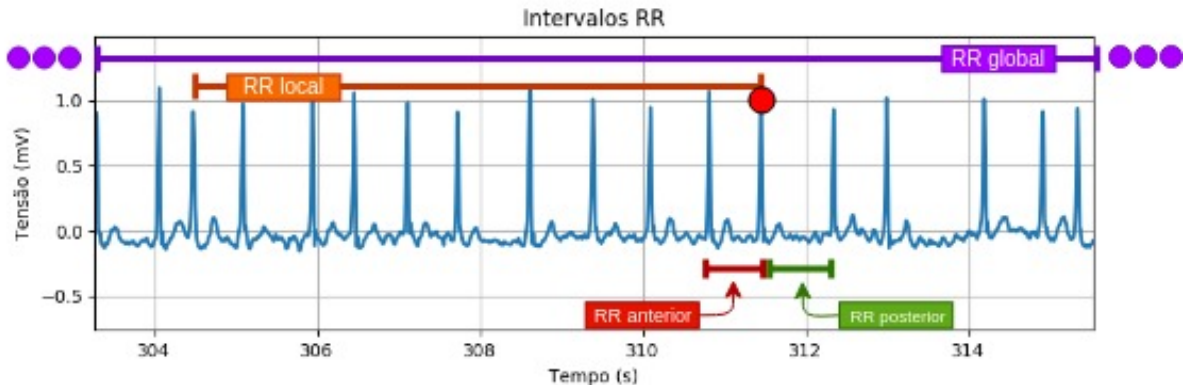


Figura 3.2: Características relacionadas ao intervalo RR para um segmento de um ECG.

3.3.2 Wavelet

Como citado previamente, uma das aplicações da DWT em sinais ECG é a extração dos coeficientes gerados em cada nível, que representam bem a característica do sinal em uma resolução multi-escala [31]. Na literatura, os autores utilizam a DWT para extrair características do ECG [30, 32, 26, 31]. Além disso, é comum calcular estatísticas sobre as subbandas para reduzir a dimensão do vetor de características [22, 6].

3.3.3 Estatísticas temporais

Estatísticas temporais também já foram utilizadas na literatura [23]. Especificamente, as estatísticas de alta ordem (HOS) são características representativas no domínio temporal dos sinais. Variância, *Skewness* e *Kurtosis* são as *HOS* de segunda, terceira e quarta ordem, respectivamente, representam a variabilidade, simetria e achatamento da distribuição de dados. Mostraram-se efetivas em trabalhos da literatura [23, 33, 22, 31].

3.3.4 Características morfológicas

A morfologia dos batimentos ECG detém grande capacidade de distinção entre as classes de arritmia. Pontos de máximo/mínimo, duração e algumas amostras (interpoladas para ter tamanho fixo) das ondas P,T e do complexo QRS são alguns exemplos de características desta categoria. Para extrair tais informações é preciso utilizar algoritmos de delimitação para detectar pontos de início e fim para cada componente do ECG (pontos fiduciais). Igualmente, vem sendo explorada na literatura [25, 23, 5].

3.4 Seleção de características

A seleção de características é uma etapa que nem sempre é aplicada pelos trabalhos da área. Em alguns deles, a seleção é manual pela submissão de várias combinações de características ao modelos e escolhendo as que resultam nas melhores métricas [5, 31]. Por outro lado, já foram aplicados algoritmos de seleção automática que propõe resultados mais robustos através da seleção baseada em algoritmos genéticos [22] e SSFS [23].

3.5 Classificação

A classificação é a última etapa do processo de automatização do diagnóstico, consiste em distinguir os batimentos em algumas categorias. As classes utilizadas pelos autores são variadas e podem ser: normais e anormais [32], tipos de arritmia (e.g. fibrilação atrial, taquicardia ventricular e etc) [34] e agrupamentos de batimentos ectópicos (e.g. supraventriculares, ventriculares e fusão) [26], sendo o último utilizado neste trabalho.

Grandes variedades de modelos classificatórios como k-vizinhos mais próximos [22], redes neurais convolucionais [34, 35, 36] e redes neurais *fuzzy* [33] foram utilizados na literatura. Luz et al. [29] realizou um trabalho de pesquisa dos artigos relacionados à detecção automática de arritmias e percebeu que os modelos SVM [6, 31], MLP [26, 20] e LDA [23, 5] estão entre os mais utilizados para tal tarefa.

Particularmente, LDA se mostrou bastante rápido e teve bom desempenho nos trabalhos que utilizam a metodologia de divisão interpaciente da base de dados, a qual será explicada com detalhes no capítulo 4.

Capítulo 4

Metodologia

4.1 Banco de dados

O banco de dados utilizado para treino, validação e teste dos modelos foi o *Arrhythmia Database* (MIT-BIH) [37], a primeira base aberta com conteúdo disponível para uso em sistemas de detecção de arritmia, cujas medidas ECG foram obtidas de pacientes do *Beth Israel Hospital Arrhythmia Laboratory* [37].

Segundo a descrição do banco, o MIT-BIH é composto por 48 registros, com duração de aproximadamente trinta minutos, obtidos de 47 pacientes distintos. As 23 medidas cujo número de referência são iniciados pelo dígito 1 são de pacientes aleatoriamente selecionados, enquanto as 25 restantes pertencem a pacientes cujos ECGs contêm algumas categorias de arritmia.

Os sinais analógicos foram filtrados por um passa-banda de 0.1 a 100 Hz e digitalizados a uma taxa de amostragem de 360 Hz [37]. A Tabela 4.1 apresenta as informações relativas a cada registro do MIT-BIH, são elas: o número de referência, idade, nome das derivações para cada um dos dois canais e sexo do paciente. As medidas foram coletadas a partir de pacientes com idades na faixa de 23 a 89 anos de ambos sexos, à exceção dos registros 103 e 219, em que as idades não foram especificadas.

Além disto, existem dois canais obrigatoriamente para cada medida, porém eles podem variar, dependendo do paciente, devido a implicações específicas no momento da medição [37]. Para 45 delas, a primeira derivação é a MLII (modificação da derivação II com os elétrodos posicionados sobre o tórax) e para as três restantes a V5. Já a segunda derivação é a V1, V2, V5, V4 e MLII para 40, 4, 2, 1 e 1 pacientes respectivamente.

Para viabilizar trabalhos supervisionados de detecção semiautomática, é ideal que anotações com os rótulos de cada classe, feitas por profissionais da área médica, sejam fornecidas pelo provedor do banco de dados. Na concepção do MIT-BIH, as anotações foram formuladas sobre todos os registros da seguinte maneira: a princípio foram gera-

das marcações automáticas através de um detector QRS, em seguida, dois cardiologistas analisaram cada registro para marcar os batimentos de arritmia, reportar comentários e verificar/corrigir inconsistências. Por fim, nos casos de anotações divergentes, os médicos entraram em consenso para definir a decisão final [37].

Após a execução do processo supracitado, foram realizados ao total aproximadamente 109000 rótulos. Os raros erros presentes nas anotações vêm sendo corrigidos para aumentar a confiabilidade do banco de dados [37].

É importante ressaltar que alguns destes registros contém batimentos de pacientes com marca-passo, neste banco, este tipo de sinal não é seguramente refletido no ECG devido a sua frequência pertencer a uma banda distante das medidas analógicas [7].

Num.	Idade	Canal 1	Canal 2	Sexo	Num.	Idade	Canal 1	Canal 2	Sexo
100	69	MLII	V5	M	⋮	⋮	⋮	⋮	⋮
101	75	MLII	V1	F	201	68	MLII	V1	M
102	84	V5	V2	F	202	68	MLII	V1	M
103	-	MLII	V2	M	203	43	MLII	V1	M
104	66	V5	V2	F	205	59	MLII	V1	M
105	73	MLII	V1	F	207	89	MLII	V1	F
106	24	MLII	V1	F	208	23	MLII	V1	F
107	63	MLII	V1	M	209	62	MLII	V1	M
108	87	MLII	V1	F	210	89	MLII	V1	M
109	64	MLII	V1	M	212	32	MLII	V1	F
111	47	MLII	V1	F	213	61	MLII	V1	M
112	54	MLII	V1	M	214	53	MLII	V1	M
113	24	MLII	V1	F	215	81	MLII	V1	M
114	72	V5	MLII	F	217	65	MLII	V1	M
115	39	MLII	V1	F	219	-	MLII	V1	M
116	68	MLII	V1	M	220	87	MLII	V1	F
117	69	MLII	V2	M	221	83	MLII	V1	M
118	69	MLII	V1	M	222	84	MLII	V1	F
119	51	MLII	V1	F	223	73	MLII	V1	M
121	83	MLII	V1	F	228	80	MLII	V1	F
122	51	MLII	V1	M	230	32	MLII	V1	M
123	63	MLII	V5	F	231	72	MLII	V1	F
124	77	MLII	V4	M	232	76	MLII	V1	F
200	64	MLII	V1	M	233	57	MLII	V1	M
⋮	⋮	⋮	⋮	⋮	234	56	MLII	V1	F

Tabela 4.1: Informações dos pacientes pertencentes à base de dados MIT-BIH

4.2 AAMI

A *Association for the Advancement of Medical Instrumentation* (AAMI) é uma organização que tem como um dos objetivos propor especificações para o desenvolvimento de tecnologias médicas. Este trabalho foi fundamentado em algumas recomendações da AAMI [38] desenvolvidas para padronizar e viabilizar a reprodução de trabalhos relacionados à arritmia.

Sua primeira recomendação é referente à escolha do banco de dados. Por ser aberto, bem descrito e consistente, a AAMI avaliou que o MIT-BIH é qualificado para ser aplicado em sistemas de arritmia, além de ser muito comum em projetos deste âmbito o que motivou sua utilização neste trabalho.

Diversos tipos de batimentos estão presentes nas anotações do MIT-BIH. A AAMI sugere que eles sejam agrupados em cinco categorias principais a fim de melhorar a representatividade dos que ocorrem poucas vezes e padronizar as classes dos trabalhos da literatura. A Tabela 4.2 apresenta as cinco classes (N, S, V, F e Q) definidas pela AAMI e as categorias de batimentos associadas a cada uma delas.

Tendo em vista que muitos autores descartam a classe Q devido a sua baixa ocorrência e grau de relevância em ambientes reais [25, 39, 31, 23], somente as classes N, S, V e F foram consideradas.

A AAMI também recomenda que os pacientes com marca-passo sejam excluídos do processo, pois no MIT-BIH os batimentos desta categoria não possuem qualidade suficiente para o propósito do trabalho. Logo, os registros 102, 104, 107 e 217 foram removidos do conjunto de entrada.

4.3 Divisão dos conjuntos de dados

Comumente, a divisão dos conjuntos de treino e teste eram feitas ao nível de batimentos, por exemplo, por validação cruzada. Porém, Chazal et al. [5] concluiu que este tipo de metodologia (intrapaciente) deve ser evitada, pois batimentos de um mesmo paciente podem aparecer tanto no treino quanto no teste. Este acontecimento acarreta em modelos com métricas excelentes, mas que não refletem os ambientes reais onde são treinados com batimentos de alguns pacientes e avaliados em outros.

Chazal et al. especificou a divisão do MIT-BIH em dois grupos de pacientes distintos e específicos para treino e teste (50% dos pacientes em cada um dos conjuntos), chamados respectivamente de DS1 e DS2, aos quais vem sendo igualmente utilizados por outros autores [23, 31]. Com a necessidade de selecionar as características e ajustar os hiper-

Classe de batimento - AAMI	Anotações do MIT-BIH
Normal (N)	Batimento normal (N) Bloqueio de ramo esquerdo (L) Bloqueio de ramo direito (R) Batimento de escape atrial (e) Batimento nodal (juncional) de escape (j)
Ectópico supraventricular (S)	Batimento atrial prematuro (A) Batimento atrial prematuro aberrante (a) Batimento nodal (juncional) prematuro (J) Batimento supraventricular prematuro (S)
Ectópico ventricular (V)	Contração ventricular prematura (V) Batimento de escape ventricular (E)
Fusão (F)	Fusão dos batimentos normal e ectópico ventricular (F)
Desconhecido (Q)	Batimento ritmado (/) Fusão dos batimentos ritmado e normal (f) Batimento não classificado (U)

Tabela 4.2: Agrupamentos de classes proposto pela AAMI

parâmetros em um conjunto diferente ao de teste, Mar et al. [23] dividiu o conjunto de treino citado anteriormente em dois subconjuntos para validação.

A Figura 4.1 apresenta, de maneira mais clara como foi feita a divisão dos conjuntos de dados neste projeto, utilizando a metodologia interpaciente da mesma forma que Mar et al. Em cinza, 44 ECGs oriundos de pacientes sem marca-passo foram divididos em dois grupos: DS1 (azul) para treino dos classificadores e DS2 (vermelho) para teste final, ambos com 22 ECGs. o primeiro é subdividido em dois grupos, com 11 pacientes distintos em cada um, que são submetidos à *2-fold* validação cruzada para ajustar os hiperparâmetros e selecionar as características antes do teste final.

Como mostra a Tabela 4.3, foram alocados os mesmos pacientes apresentados por Mar et al. para cada conjunto, decisão que permite a comparação dos resultados com os trabalhos que seguem a mesma abordagem.

4.4 Etapas

Usualmente, as ferramentas de detecção semiautomática de arritmias são compostas por algumas etapas bem definidas: pré-processamento, segmentação, extração de características, seleção de características e classificação. Neste trabalho todas elas foram aplicadas, inclusive a de seleção de características que nem sempre é incluída no processo. Além disso,

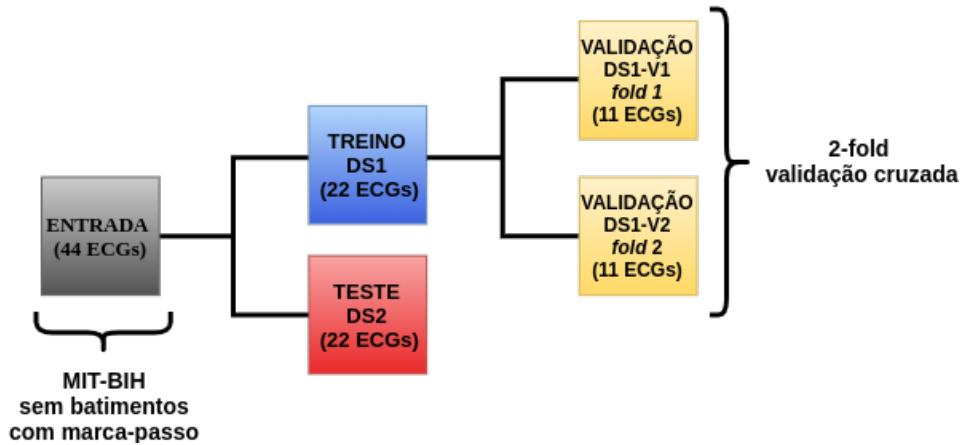


Figura 4.1: Divisão interpaciente do banco de dados.

	Pacientes
Treino DS1	101, 106, 108, 109, 112, 114, 115, 116, 118, 119, 122, 124, 201, 203, 205, 207, 208, 209, 215, 220, 223 e 230
Teste DS2	100, 103, 105, 111, 113, 117, 121, 123, 200, 202, 210, 212, 213, 214, 219, 221, 222, 228, 231, 232, 233 e 234
Validação DS1-V1	109, 114, 118, 119, 124, 201, 203, 205, 215, 220 e 223
Validação DS1-V2	101, 106, 108, 112, 115, 116, 122, 207, 208, 209 e 230

Tabela 4.3: Divisão dos pacientes ao longo dos conjuntos definidos pela divisão interpaciente

também adicionou-se a etapa de reamostragem para contornar problemas de desbalanceamento.

A Figura 4.2 ilustra um diagrama cronológico das etapas adotadas. Do pré-processamento ao ajuste dos hiperparâmetros o fluxo é linear, em seguida, três caminhos são tomados: um em roxo (1), um em azul (2) e um em verde (3). Finalmente, todos eles se encontram na classificação, gerando suas respectivas métricas. As subseções 4.4.1 a 4.4.6 irão discorrer detalhadamente sobre cada uma das etapas.

4.4.1 Pré-processamento

A primeira etapa se constitui em filtrar alguns artefatos dos sinais pertencentes ao MIT-BIH como o *Baseline Wandering* e os ruídos de alta frequência. A remoção foi feita através da decomposição do sinal utilizando a transformada *Wavelet* discreta.

Primeiramente foi necessário decidir qual família de *Wavelets* seria aplicada. Singh et al. [21] realizou um estudo comparativo de várias famílias e concluiu que a *Daubechies* é a mais indicada para a remoção de ruídos do ECG. O pré-processamento aplicado foi

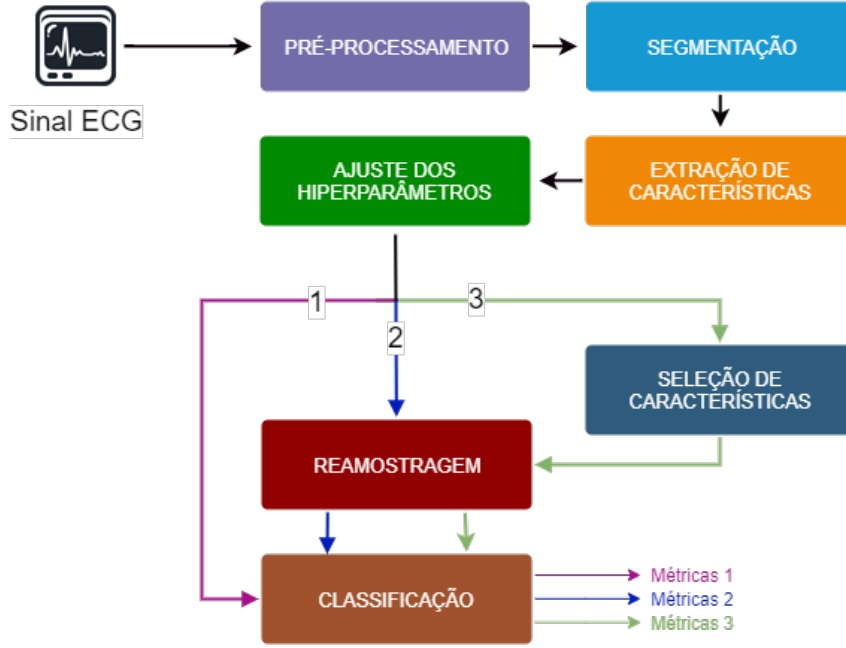


Figura 4.2: Fluxo aplicado para a ferramenta de detecção desenvolvida.

inspirado em Roshan Joy et al. [26], onde as subbandas ruidosas foram zeradas antes da reconstrução do sinal pela transformada inversa. Porém, zerar todos os coeficientes pode acarretar na perda de informação, assim, as subbandas que detêm os artefatos de alta frequência não foram zeradas, mas sim diminuídas aplicando técnicas de limiarização, como relatado por Lin et al [40]. Já as subbandas que contém os de baixas frequência foram zeradas pois são dominadas majoritariamente por estes artefatos [40]. O procedimento adotado é descrito a seguir:

Primeiramente, os sinais de cada registro do banco foram submetidos a 9 níveis de decomposição da DWT com a *Wavelet* mãe db6, o que resultou em 9 detalhes (D_1 a D_9) e uma aproximação A_9 . Os coeficientes A_9 têm largura de banda de 0 a 0.3516 Hz, que em grande maioria são a *Baseline Wander*, logo, todos eles foram zerados. Já os coeficientes D_1 e D_2 juntos englobam as frequências do intervalo de 45 a 180 Hz e foram diminuídos através do limiar suave, em inglês *soft threshold*, apresentado na Equação 4.1. A Equação 4.2 (t) apresenta o *threshold* (t), proposto por Donoho et al. [41], associado a cada subbanda do sinal, onde N é o tamanho do sinal original em amostras, σ o desvio padrão do ruído e D_j os coeficientes do detalhe para o nível j .

$$D_j[n] = \begin{cases} \text{sin}(D_j[n]) (|D_j[n]| - t), & \text{para } |D_j[n]| \geq t \\ 0, & \text{para } |D_j[n]| < t \end{cases} \quad (4.1)$$

$$t = \sigma \sqrt{2 \log N} \quad | \quad \sigma = \frac{\text{mediana}(|D_j|)}{0.6745} \quad (4.2)$$

Finalmente, com as subbandas processadas, o sinal foi reconstruído aplicando a DWT inversa para se obter sua versão filtrada.

A Figura 4.3 mostra o resultado do pré-processamento para um intervalo de aproximadamente 5 segundos do paciente 100. Na parte superior, o trecho original do sinal e na inferior, o sinal filtrado. Visualmente, as flutuações geradas pelo *Baseline Wander* diminuíram e o sinal foi suavizado sem alterar sua morfologia.

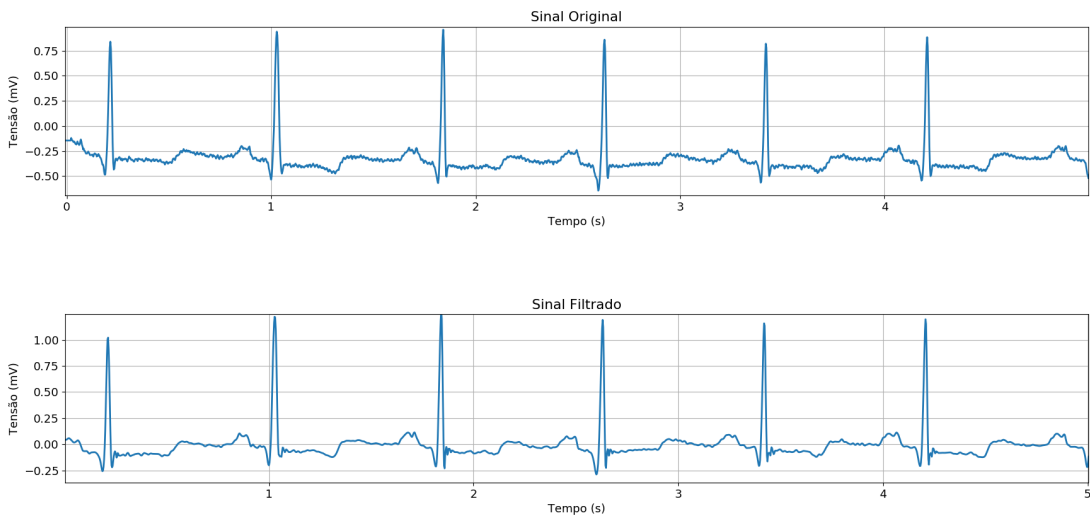


Figura 4.3: Resultado do pré-processamento para um trecho de um registro do MIT-BIH. Na parte superior, o sinal original contaminado com os artefatos e, na inferior, o sinal filtrado.

4.4.2 Segmentação

Com o sinal pré-processado, a etapa seguinte consiste em segmentá-lo por batimentos para que seja possível extrair informações úteis de cada um deles.

*Ecgpuwave*¹ é um algoritmo de detecção de pontos fiduciais no ECG fundamentado no algoritmo *Pan-Tompkins*. A partir dele, gerou-se anotações com os pontos fiduciais (picos R) relativos a cada sinal ECG. As anotações foram lidas e cada segmento foi definido com 99 amostras (275 ms) à esquerda e 144 (400 ms) à direita de cada pico R. A decisão de uma quantidade maior de amostras à direita se baseia em Afkham et al. [24], ao qual

¹<https://physionet.org/physiotools/wag/ecgpuw-1.htm>

relata que tal duração permite capturar todo comprimento da onda T para a maioria das situações.

Além disso, foram descartados os primeiros e últimos segmentos de cada medida que não possuíam amostras suficientes em qualquer um dos lados para padronizar esta etapa.

Como a onda R é de curta duração, alguns picos ficaram levemente deslocados em relação à posição em que deveriam realmente estar, por isso foi aplicado um algoritmo de correção de picos fornecido pela biblioteca de leitura do banco de dados.

A correção funciona da seguinte forma: para o segmento em questão, obtém-se a versão suavizada do sinal através de um filtro de média, sendo a nova posição do pico R a amostra do sinal original que possui a maior distância, em módulo, até a versão suavizada, isto é, o máximo ou mínimo local.

4.4.3 Extração de características

Com os segmentos definidos, foi possível extrair as características e as classes associadas a cada um através das anotações. As características foram computadas apenas sobre os sinais derivados do canal MLII com o propósito de simplificar o sistema. As quatro categorias de características extraídas são apresentadas a seguir:

- Intervalo RR: Os 4 descritores temporais relacionados à esta categoria foram calculados: RR-anterior, RR-posterior, RR-local e RR-global.
- Estatísticas temporais: Foram extraídas três HOS do segmento (variância, *skewness* e *kurtosis*) além do desvio padrão e média, totalizando 5 características.
- Morfológicas temporais: Calcular a duração das ondas P e T demanda que o algoritmo de segmentação encontre os pontos de início e fim de cada uma, processo que pode ser suscetível a erros. Assim, foram calculadas apenas características morfológicas associadas ao segmento como um todo: energia, amplitude máxima, mínima e pico a pico.
- Wavelet: Roshan Joy et al. [26] verificou que o quarto nível de decomposição através da Wavelet mãe *FIR approximation of Meyer wavelet* (*dmey*) é capaz de discriminar bem as classes de arritmia. As subbandas A_4 e D_4 foram usadas, porém a dimensão de seus coeficientes é maior que 250. Assim, calculou-se as seguintes características morfológicas e estatísticas sobre os coeficientes: máximo, mínimo, pico a pico, HOS, média e desvio padrão resultando em 16 características.

Ao final, a extração resultou em 100381 *data points* com 29 características cada, além da classe e número de referência do registro de onde foi retirado o segmento.

4.4.4 Seleção de características e ajuste dos hiperparâmetros

Após a extração das características de cada segmento, o conjunto de entrada estava pronto para as etapas subsequentes, assim, primeiramente aplicou-se a divisão interpaciente explicada anteriormente e utilizou-se os conjuntos de validação tanto para ajustar os hiperparâmetros dos modelos escolhidos, quanto para selecionar as características que mais contribuem para a distinção das classes.

A Figura 4.4 apresenta as quantidades das amostras nos *fold*s do conjunto de validação. Acima de cada barra estão as porcentagens de amostras pertencentes a cada classe do *fold* em questão. Como esperado, a classe N é majoritária e todas as outras são classes minoritárias. A classe F, por exemplo, detêm apenas 0.1% da quantidade total do *fold* 1 e 1.5% no *fold* 2.

Apesar dos dois *fold*s serem extremamente desbalanceados, optou-se por não realizar nenhum tipo de intervenção nos dados, pois o conjunto de validação tem como propósito fornecer uma estimativa genérica dos hiperparâmetros e características. Assim, estes tipos de experimento foram feitos apenas sobre os dados na etapa de treino final.

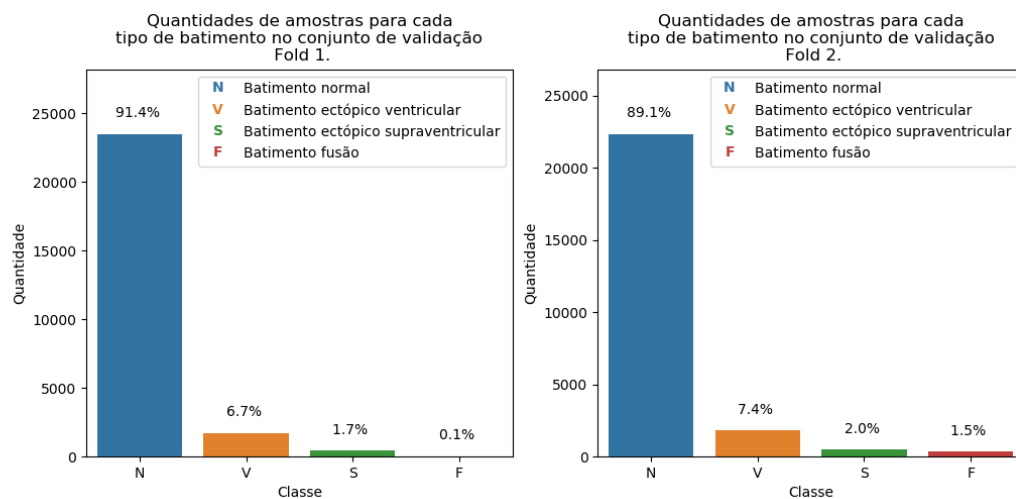


Figura 4.4: Distribuição das amostras do conjunto de validação. À esquerda para o *fold* 1 e à direita para o *fold* 2.

4.4.4.1 métrica para seleção

Com a aplicação de técnicas para a seleção de características, surgiu a necessidade de formular uma métrica que represente, em um único valor, quão bem a classificação das arritmias desempenhou, pois, normalmente, as métricas usuais são dominadas pela classe N devido a sua quantidade de amostras ser superior em relação às demais. Logo, Mar

et al. [23] propôs o jk_{index} , apresentado na Equação 4.3, ao qual já foi empregado como métrica de performance final em outros trabalhos [31].

A métrica é obtida pela combinação linear do *Cohen's Kappa* (k) e o j_{index} de modo que o valor final fique no intervalo de 0 a 1. O k é adotado como métrica global da matriz de confusão por ser menos influenciado pelo desbalanceamento quando comparado à outras métricas. Já o jk_{index} , apresentado na Equação 4.4, é adotado por avaliar a classificação das arritmias mais representativas, as S e V.

$$jk_{index} = w_1 \cdot k + w_2 \cdot j_{index} \quad | \quad w_1 = \frac{1}{2}, \quad w_2 = \frac{1}{8} \quad (4.3)$$

$$j_{index} = Sensibilidade_S + Sensibilidade_V + Precisão_S + Precisão_V \quad (4.4)$$

4.4.4.2 Seleção de características

A seleção de características foi feita utilizando os conjuntos de validação através da técnica *SSFS* com o jk_{index} para avaliação de performance das características. O classificador selecionado para avaliação das características foi o LDA por ter pouco custo computacional além de ser um bom classificador para o propósito deste trabalho.

O algoritmo foi executado para selecionar a mesma quantidade que a dimensão do espaço original, isto é, 29 características. A Figura 4.5 mostra os resultados do jk_{index} nos *folds* para cada passo do *SSFS*. É notório que o aumento excessivo da dimensão acarreta em queda da performance geral do classificador. Deste modo, o subconjunto com 14 características foi selecionado por obter o maior desempenho médio.

As características selecionadas foram as seguintes: RR-global, RR-anterior, média, *kurtosis*, máximo D_4 , mínimo D_4 , pico à pico D_4 , *skewness* D_4 , *kurtosis* D_4 , mínimo A_4 , desvio padrão A_4 , média A_4 , variância A_4 e *kurtosis* A_4 .

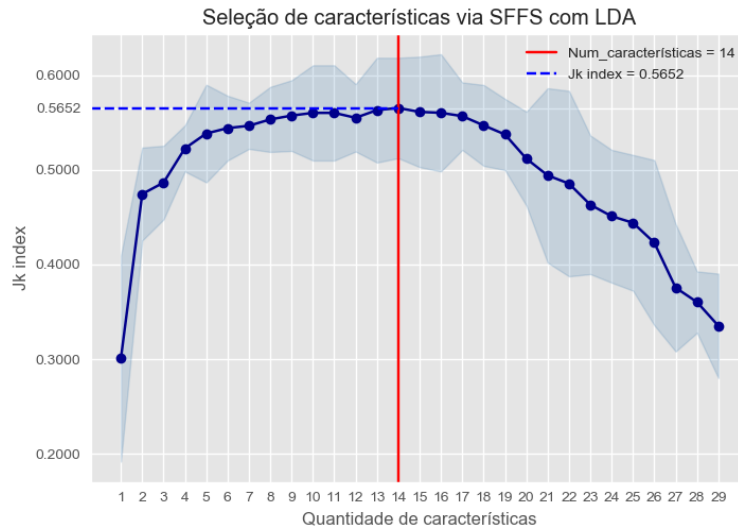


Figura 4.5: Seleção de características via SFFS com LDA. O eixo y mostra o jk_{index} para cada subconjunto de características do eixo x. A curva em azul corresponde ao desempenho médio, e a área ao redor ao desvio padrão nos *folds*.

4.4.4.3 ajuste dos hiperparâmetros

Os classificadores utilizados neste trabalho foram: SVM, MLP e LDA. Para cada um deles os hiperparâmetros foram ajustados via *GridSearchCV* nos dois *folds* do conjunto de validação, também utilizando o jk_{index} para ranquear a melhor combinação dos mesmos.

Para o SVM foram fornecidas as quatro funções *kernel* disponíveis na biblioteca (*linear*, *sigmoid*, *rbf* e *poly*) além de vários valores do parâmetro de classificação incorreta (C).

Por conceito, o LDA é um classificador não paramétrico, e por isso, utilizou-se a configuração padrão da biblioteca.

Por fim, para a MLP foram estimados o número de neurônios da camada oculta, a função de ativação e a taxa de aprendizado. Fixou-se o número de épocas em um valor muito alto (5000) e, para garantir a convergência, utilizou-se o valor padrão de tolerância na variação da função de custo ($tol = 10^{-4}$), isto é, caso a função de custo não variasse mais que *tol* durante dez épocas consecutivas, o treinamento seria interrompido assumindo a convergência da MLP.

Os resultados do *GridSearchCV* para a SVM são apresentados na Figura 4.6 em forma de mapa de calor, de modo que o eixo x exhibe as funções *kernel* e o eixo y, os valores de C. Cada célula contém a média do jk_{index} nos dois *folds* de validação. Para auxiliar a visualização, adicionou-se uma escala de cor à direita cuja intensidade do azul é associada a valores maiores. É possível perceber que o *kernel* linear com valores pequenos de

C obtiveram no geral os melhores resultados, portanto os hiperparâmetros selecionados foram o *kernel* linear com C igual à 0.1.

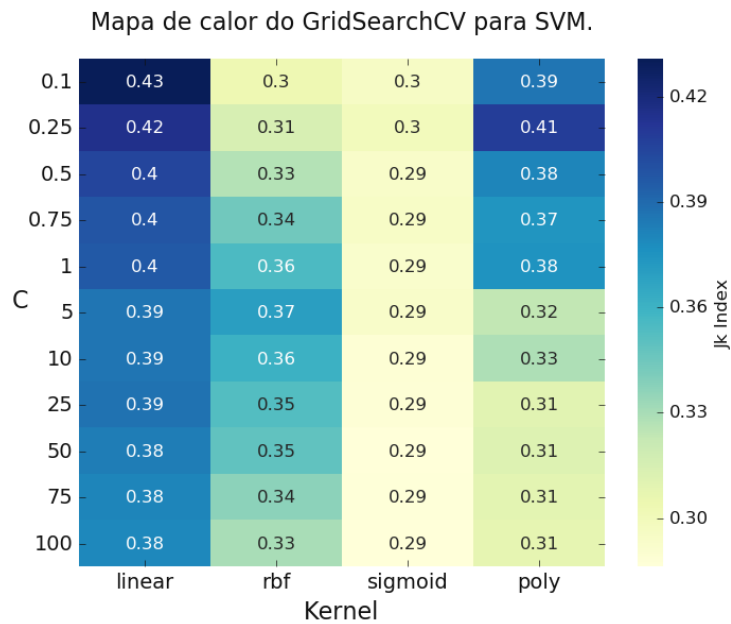


Figura 4.6: Mapa de calor com os resultados do *GridSearchCV* para a SVM.

Analogamente, a Figura 4.7 exibe os valores do *GridSearchCV* para a MLP. Porém como foram submetidos três categorias de hiperparâmetros, o gráfico possui três dimensões, além da escala de cor para denotar o desempenho. O eixo x apresenta as quantidades de neurônios, o y as taxas de aprendizado e o z as funções de ativação. O eixo z é composto por variáveis categóricas, isto é, não possuem valores numéricos que permitam ordená-las no eixo. Desta forma, foram posicionadas em ordem arbitrária de modo que ficassem igualmente distribuídas no eixo. O melhor resultado foi obtido pela combinação de hiperparâmetros circulada em vermelho cujos valores são 14 neurônios na camada oculta, tangente hiperbólica (*tanh*) como função de ativação e taxa de aprendizado de 0.001.

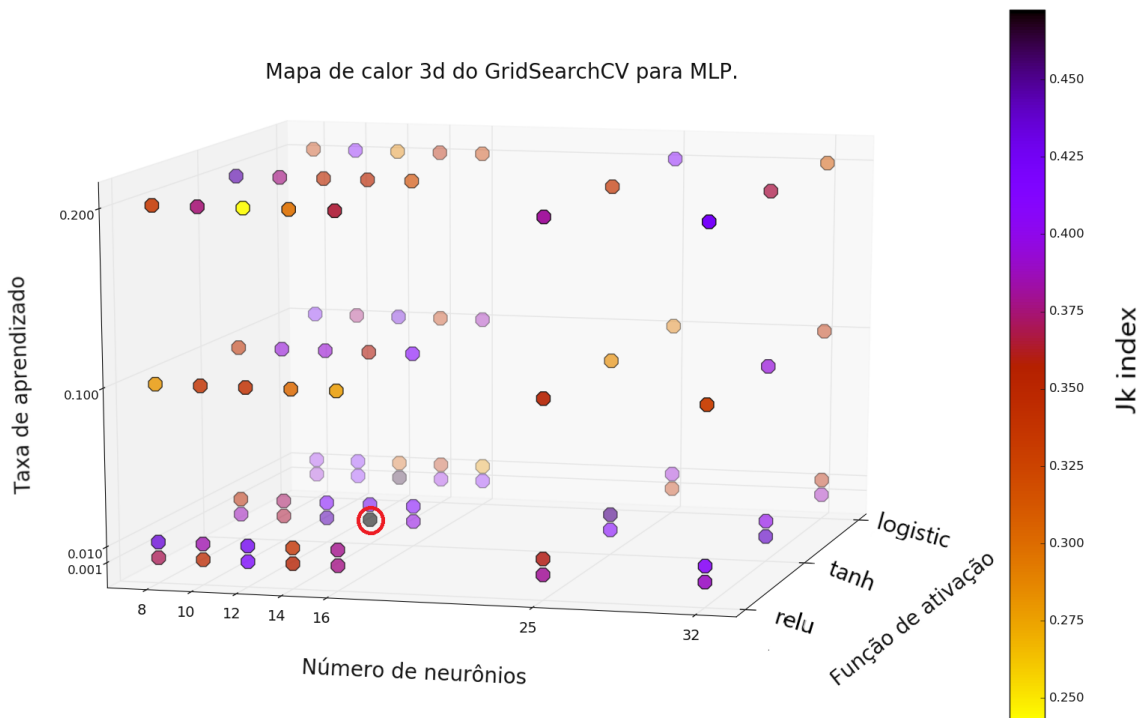


Figura 4.7: Mapa de calor com os resultados do *GridSearchCV* para a MLP.

4.4.5 Reamostragem

Na literatura normalmente os trabalhos lidam com o desbalanceamento sob a perspectiva dos classificadores, por exemplo, aplicando pesos no treinamento. Como relatado na pesquisa sobre trabalhos da área feita por Luz et al. [29], métodos para tratar o desbalanceamento sob a perspectiva dos dados não são muito explorados pelos autores.

Uma grande contribuição deste trabalho é justamente equilibrar os dados de treino aplicando uma técnica robusta de geração de dados sintéticos para as classes minoritárias e verificando seu reflexo na performance da classificação.

A Figura 4.8 mostra, à esquerda, as quantidades de amostras para o conjunto de treino (DS1) e , à direita, para o conjunto de teste (DS2). É evidente o domínio da classe N em detrimento das restantes em ambos os conjuntos.

O algoritmo de reamostragem escolhido foi o ADASYN por gerar os dados sintéticos de uma forma mais inteligente que seu antecessor (SMOTE), criando mais amostras para as quais são mais difíceis de aprender. A vizinhança utilizada foi a padrão da biblioteca ($k = 5$).

Somente as quantidades de amostras no conjunto de treino foram modificadas, pois é

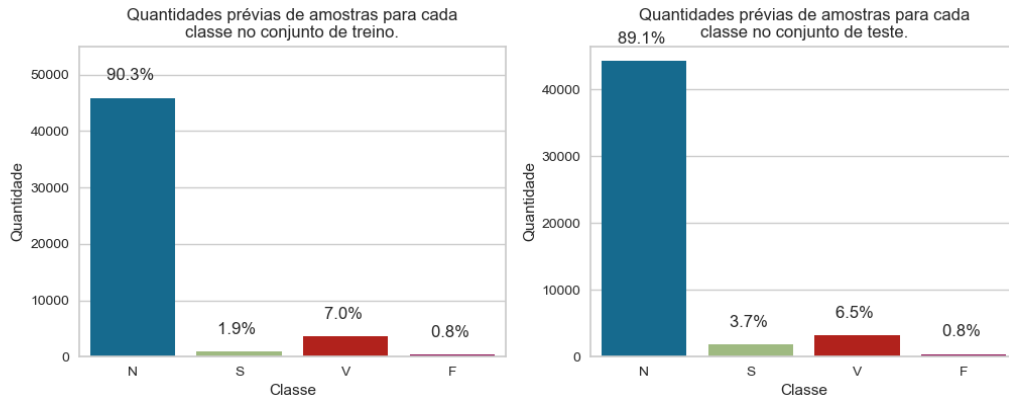


Figura 4.8: Distribuição das amostras entre as classes para o conjunto de teste e treino respectivamente.

preciso que o conjunto de teste reflita os ambientes reais, onde os dados são naturalmente desbalanceados. O objetivo da reamostragem é apenas eliminar o viés para as classes majoritárias no treinamento dos modelos.

Primeiramente, a classe N foi aleatoriamente subamostrada de forma que sua quantidade final ficasse igual à soma das quantidades de amostras das classes S, V e F. Isso foi feito para que a proporção entre batimentos normais e ectópicos fossem iguais. Em seguida, para as classes V, S e F, foram geradas as amostras sintéticas com o ADASYN para que cada uma detivesse ao final aproximadamente a mesma quantidade de amostras que a classe N.

O resultado da reamostragem é apresentado na Figura 4.9, onde cada classe tem aproximadamente 25% do conjunto final de treino. As pequenas variações nas quantidades se dão justamente pelo efeito citado anteriormente de geração de mais dados para as amostras mais difíceis de aprender.

4.4.6 Classificação

Como foi dito anteriormente, a última etapa consiste na classificação dos batimentos entre as quatro classes da AAMI treinando com apenas os dados de DS1 e avaliando em DS2. Tal tarefa foi realizada utilizando cada um dos ambientes de treino originados por cada um dos fluxos (1, 2 e 3) da Figura 4.2. Os ambientes em questão foram classificados utilizando os três modelos preditivos com seus melhores hiperparâmetros selecionados na etapa de ajuste.

O Capítulo 5 discutirá detalhadamente sobre cada um dos experimentos aplicados, suas motivações e respectivas métricas de performance para todos os modelos gerados.

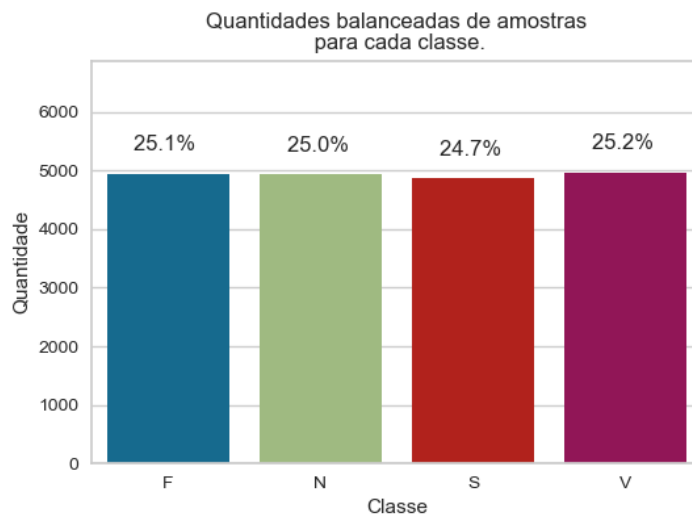


Figura 4.9: Distribuição das amostras, após o balanceamento, das classes para o conjunto treino.

Capítulo 5

Resultados

Neste capítulo são apresentados os resultados dos três experimentos classificatórios realizados, suas respectivas análises e a comparação dos modelos de maiores métricas com alguns trabalhos da literatura que utilizaram a metodologia interpaciente e seguiram as recomendações da AAMI.

5.1 Métricas de desempenho

Os três experimentos foram desenvolvidos sequencialmente com o intuito de aprimorar as métricas selecionadas. Os resultados foram obtidos da classificação de DS2 pelos três classificadores treinados com as amostras dos conjuntos de entrada associados à cada um dos experimentos.

Para avaliação de desempenho, as métricas utilizadas foram a matriz de confusão normalizada, acurácia e as precisões e sensibilidades relacionadas a cada uma das classes. Estes conjuntos de valores foram registrados em tabelas para permitir uma análise da performance de cada experimento bem como sua comparação com os restantes. Além das métricas supracitadas também estimou-se o jk_{index} , apresentado anteriormente na Equação 4.3, para os modelos selecionados para comparação com a literatura.

Como recomendação da AAMI [38], as classificações de batimentos de fusão como ventriculares não devem ser penalizadas ou recompensadas pelos classificadores. Em outras palavras, os FP de V devem incluir apenas as amostras de N e S erroneamente atribuídas à V. Tal recomendação reflete, algumas vezes, na mudança dos valores de algumas métricas utilizadas, como por exemplo, aumento da precisão para esta classe e do jk_{index} .

A seção 5.2 discorrerá detalhadamente sobre os experimentos efetuados em ordem cronológica.

5.2 Experimentos

5.2.1 Conjunto original de dados

O primeiro experimento (fluxo 1 da Figura 4.2) foi realizado treinando os modelos sem qualquer manipulação na quantidade de amostras pertencentes à DS1 e com todas as características calculadas na etapa de extração apresentada previamente na Seção 4.4.3.

A Figura 5.1 exibe as matrizes de confusão normalizadas resultantes das três classificações deste experimento. Elas revelam que a maior parte das predições corretas ficaram restritas às classes V e N, ao passo que houve uma tendência na predição da classes minoritárias restantes (S, F) como N, isto é, uma baixa quantia de TP para estas classes. Em relação a F, por exemplo, a taxa de TP (sensibilidade) foi de aproximadamente 20% para a MLP, 30% para a LDA e apenas 8% para a SVM, já para S, foram de 22%, 11% e 12% respectivamente.

A Tabela 5.1 apresenta a precisão e sensibilidade dos três classificadores para cada classe bem como a acurácia associada. As métricas em questão foram selecionadas por serem utilizadas na análise dos resultados em trabalhos com propósitos e metodologias similares [23, 31].

	N		S		V		F		*
	Sensibilidade	Precisão	Sensibilidade	Precisão	Sensibilidade	Precisão	Sensibilidade	Precisão	Acurácia
MLP	0.911	0.963	0.216	0.249	0.943	0.799	0.205	0.037	0.882
SVM	0.994	0.952	0.119	0.695	0.866	0.928	0.080	0.756	0.946
LDA	0.973	0.945	0.108	0.169	0.738	0.915	0.295	0.318	0.920

Tabela 5.1: Métricas resultantes do experimento 1 para os três modelos classificatórios.

Como esperado, o valor da acurácia se mostrou inadequado para avaliar a performance em conjuntos de dados desbalanceados. Neste caso, a acurácia foi maior que 88% para os três classificadores enquanto as sensibilidades para S foram inferiores à 22% e para F 30%, ou seja, apesar de um acerto global de 88% menos de 30% dos batimentos ectópicos pertencentes à cada uma dessas classes de arritmia foram corretamente previstos.

As precisões da SVM para as classes S e F foram, na devida ordem, 70% e 76%, podendo ser consideradas relativamente altas quando comparadas às dos outros modelos deste experimento que obtiveram menos de 25% e 32% desta métrica para as mesmas classes.

Para uma classe de arritmia \mathbf{X} , os FP remetem aos pacientes saudáveis ou portadores de outras arritmias classificados como \mathbf{X} . Já os FN de \mathbf{X} dizem respeito aos portadores desta arritmia que não estão sendo corretamente diagnosticados. No atual contexto, para as classes de arritmia, os FP são menos críticos que os FN, pois ao se tratar de uma ferramenta médica semiautomática todas as amostras preditas positivas, isto é, classificadas como batimentos ectópicos, serão analisadas posteriormente por um profissional da área

da saúde para formular a decisão final, de forma que os FP terão a oportunidade de serem corrigidos.

Assim, o objetivo traçado foi diminuir a quantia de FN e aumentar os TP das classes de arritmia, sendo tais valores diretamente refletidos no aumento das sensibilidades dessas classes nos modelos. Como tentativa de aumentar estas métricas, foi realizado um experimento com os dados gerados pela etapa de reamostragem da Seção 4.4.5.

5.2.2 Conjunto com dados reamostrados

Neste experimento, os dados utilizados para treino foram os de DS1 submetidos ao procedimento de reamostragem explicitado na Seção 4.4.5, também com as 29 características extraídas na Seção 4.4.3, cuja origem é representada pelo fluxo de 2 da Figura 4.2.

A Figura 5.2 apresenta as três matrizes de confusão resultantes dos testes dos modelos treinados com o novo conjunto de dados de entrada. Comparadas às anteriores, vale ressaltar que houve uma melhoria nas predições corretas para a classe S, principalmente para a LDA e SVM. Em contrapartida, as classificações para F ainda se mantiveram dispersas, principalmente entre N, V e a própria F.

A Tabela 5.2 apresenta as métricas resultantes da classificação e, ao analisá-las, percebe-se uma redução na sensibilidade para a classe N em relação ao primeiro experimento. Para a MLP a queda foi de aproximadamente 19%, para a SVM 17% e para a LDA 14%. Em compensação, as precisões para a mesma classe obtiveram um aumento de até 5%.

	N		S		V		F		*
	Sensibilidade	Precisão	Sensibilidade	Precisão	Sensibilidade	Precisão	Sensibilidade	Precisão	
MLP	0.719	0.985	0.506	0.263	0.915	0.592	0.306	0.013	0.720
SVM	0.821	0.994	0.878	0.233	0.893	0.688	0.166	0.033	0.822
LDA	0.837	0.995	0.881	0.253	0.865	0.746	0.386	0.067	0.837

Tabela 5.2: Métricas resultantes do experimento 2 para os três modelos classificatórios.

Em relação à classe S, a reamostragem proporcionou um ganho de sensibilidade para a LDA e a SVM, que obtiveram 88% desta métrica. Por outro lado, a sensibilidade para a MLP foi de 51%, que apesar do aumento, continua sendo um valor insatisfatório para classificar um arritmia tão significativa.

Em todos os casos a maior parte das predições de F correspondem aos FP distribuídos para as classes N e V. Tal evento é coerente com a realidade tendo em vista que os batimentos de fusão correspondem, como o próprio nome diz, à fusão de batimentos normais e ectópicos ventriculares, logo, é de se esperar que estes possuam características próximas. Devido à esta circunstância, as sensibilidades de F foram 31%, 17% e 39% para

a MLP, SVM e LDA respectivamente. Já os valores das precisões da classe F foram: 1%, 3% e 7% para os mesmos modelos.

As acurácias globais reduziram até 16% em relação ao experimento anterior, principalmente condicionados pela queda de TP da classe N, que representam 89.1% das amostras do conjunto de teste.

5.2.3 Características selecionadas

Como visto anteriormente, a aplicação da técnica de reamostragem promoveu, no geral, um aumento significativo na sensibilidade da classe S. Porém, muitas características não agregam muita informação no processo de classificação, sendo elas bastante correlacionadas ou trazendo ruídos para o espaço dos dados de entrada. Normalmente, a alta dimensão deixa os modelos mais complexos e diminuem sua performance. Logo, aplicou-se um experimento utilizando os dados de DS1 reamostrados com apenas as 14 características originadas da etapa de seleção de características elucidada na Seção 4.4.4.2 (fluxo 3 da Figura 4.2).

Na Figura 5.3 estão as três matrizes de confusão resultantes dos testes dos modelos treinados para este experimento. Visualmente percebe-se uma maior concentração das predições nas diagonais principais, principalmente para a LDA e para a SVM. A Tabela 5.3 apresenta as métricas para o experimento em questão, que comparadas com as do anterior, mostram uma redução na sensibilidade de N em todos os casos.

	N		S		V		F		*
	Sensibilidade	Precisão	Sensibilidade	Precisão	Sensibilidade	Precisão	Sensibilidade	Precisão	Acurácia
MLP	0.664	0.991	0.908	0.284	0.942	0.627	0.290	0.012	0.687
SVM	0.741	0.997	0.943	0.253	0.758	0.633	0.907	0.058	0.751
LDA	0.768	0.997	0.938	0.268	0.793	0.705	0.894	0.062	0.776

Tabela 5.3: Métricas resultantes do experimentos 3 para os três modelos classificatórios.

Para a MLP, a sensibilidade foi de 67% para N e 29% para F. Por tais valores, este modelo mostra-se menos competitivo quando comparado aos restantes deste experimento. Entretanto, apesar desta redução no valor das sensibilidades, as precisões de N foram em torno de 100% nos três modelos. Tais valores indicam que aproximadamente todas as amostras classificadas como N pertencem realmente a esta classe.

Já a LDA e a SVM obtiveram sensibilidades de 77% e 74% para a classe de batimentos normais, valores inferiores aos do experimento passado associados aos mesmos classificadores. Em contrapartida, as sensibilidades de F aumentaram para 89% e 91%, mesmo esta classe representando apenas 0.8% do conjunto de teste. Para os batimentos pertencentes a S, as sensibilidades também foram superiores às dos experimentos anteriores: 91% para MLP e 94% para LDA e SVM.

5.3 Comparação com trabalhos da literatura

Por ser uma tema bastante explorado na literatura é de suma importância comparar os resultados obtidos com outros trabalhos para avaliar a contribuição das técnicas e características utilizadas no processo de automação. Logo, os modelos de cada experimento foram analisados e o considerado com melhores métricas, foi selecionado para ser comparado.

Nenhum modelo foi selecionado no experimento 1 devido as sensibilidades associadas às classes S e F estarem simultaneamente abaixo de 22% e 30% respectivamente, as quais são inferiores aos modelos dos experimento 2 e 3. Para o experimento 2, apesar das sensibilidades das classes N, V e S, nenhum modelo foi selecionado devido ao fato de apresentarem sensibilidade para F de no máximo 39%, valor inferior a alguns do experimento 3, que obtiveram, em média, 90% para esta métrica. Em relação ao experimento 3, o LDA (LDA-EXP3) foi selecionado devido as métricas avaliadas na Tabela 5.3 serem, no geral, maiores para o classificador em questão em relação aos outros do mesmo experimento.

A Tabela 5.4 mostra alguns trabalhos que utilizaram a metodologia interpaciente bem como as recomendações da AAMI para formular trabalhos de detecção semiautomática de arritmias cardíacas.

Autores	N		S		V		F		*	*
	Sensibilidade	Precisão	Sensibilidade	Precisão	Sensibilidade	Precisão	Sensibilidade	Precisão	Acurácia	jk index
Mar et al. [23]	0.896	0.991	0.832	0.335	0.868	0.859	0.611	0.166	0.890	0.602
Mondéjar et al. [31]	0.959	0.982	0.781	0.497	0.947	0.939	0.124	0.236	0.945	0.773
Chazal et al. [5]	0.871	0.992	0.760	0.385	0.803	0.866	0.894	0.086	0.862	0.612
LDA-exp3	0.768	0.997	0.938	0.268	0.793	0.705	0.894	0.062	0.776	0.542

Tabela 5.4: Comparação das métricas do modelo proposto neste projeto (LDA-EXP3) com alguns trabalhos da literatura relacionados.

O modelo proposto por Mar et al. [23] apresentou sensibilidade 13% maior que o LDA-EXP3 para a classe N, entretanto, para a classe S, houve troca entre sensibilidade e precisão de modo que a sensibilidade para o LDA-EXP3 fosse maior. Por fim, houve um ganho de 28% de sensibilidade para a classe F.

Já Mondéjar et al. [31] obteve as maiores sensibilidades de N e V, maiores precisões para S, V e F, além da maior acurácia e jk_{index} . Em contrapartida, a sensibilidade para F foi baixa, apenas 16.6%. Para aplicação das ferramentas em ambientes reais é preciso que a detecção de todas as arritmias sejam representativas, pois elas são a classe de interesse do problema. O LDA-EXP3, apesar de algumas métricas inferiores, obteve altas sensibilidades para as três classes de arritmia.

Por fim, quando comparado ao modelo proposto por Chazal et al. [5], o LDA-EXP3 obteve sensibilidades semelhantes para as classes F e V. Por outro lado, houve uma queda de sensibilidade para a classe N, porém um ganho para a classe S.

Comparado aos três trabalhos, o LDA-EXP3 obteve sensibilidade inferior para a classe normal, porém a precisão desta mesma classe foi de aproximadamente 100%. Desta forma, nem todos os batimentos normais são classificados corretamente, mas quase a totalidade das amostras que são atribuídas a esta classe são realmente pertencentes a ela. Além disto, podemos perceber também as maiores sensibilidades para a classe S e F dentre os trabalhos selecionados para comparação. Desta forma, o LDA-EXP3 é um modelo confiável, que apesar de menor precisão para as classes de arritmia, consegue classificar grande parte delas corretamente além de distingui-las.

Em relação ao jk_{index} , o modelo proposto obteve 54%, enquanto Mar et al. obteve 60%, Chazal et al. 61% e Mondéjar et al. 77%. Apesar do último apresentar um valor superior aos demais, a sensibilidade para a classe F foi de apenas 12%, valor bastante inferior aos dos outros trabalhos.

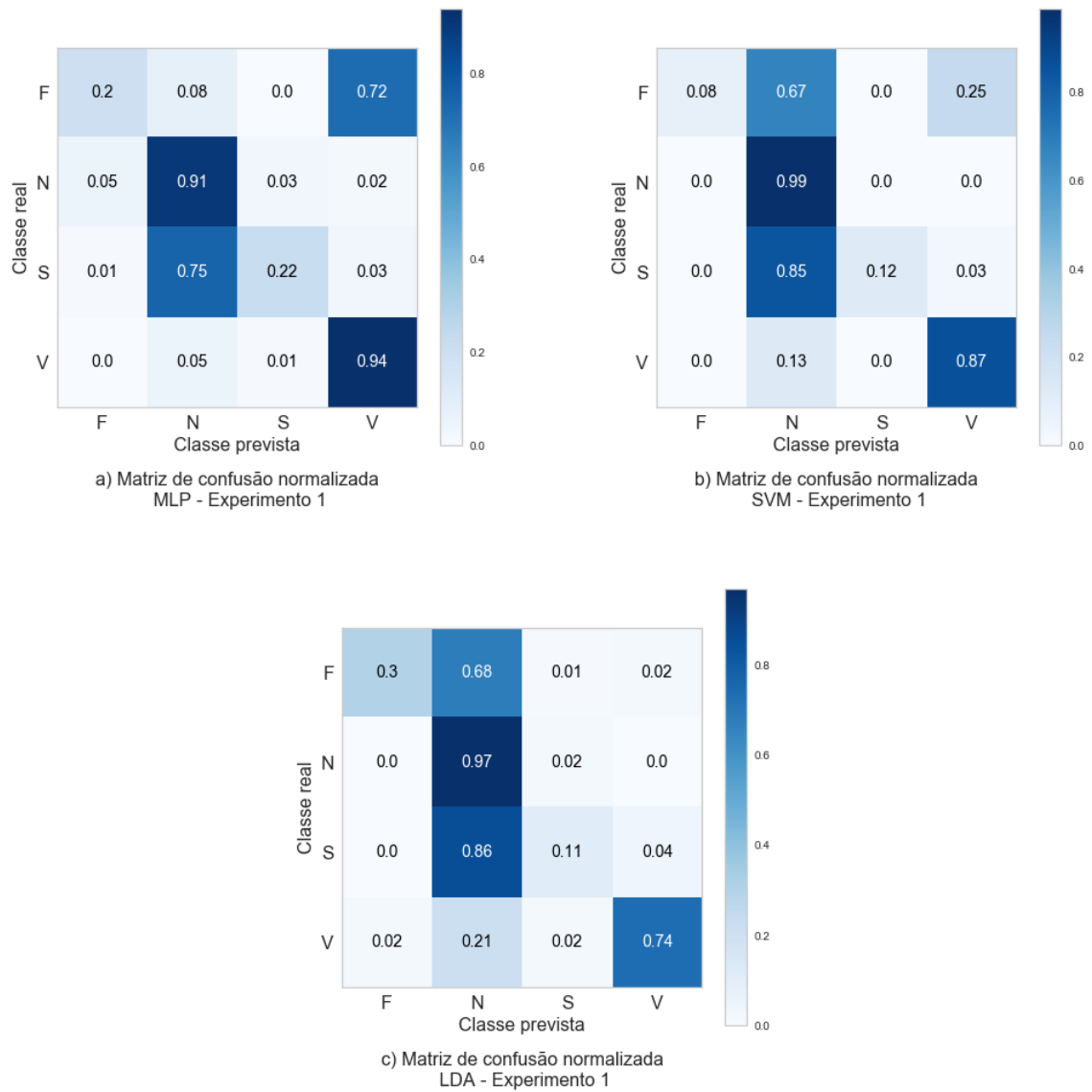


Figura 5.1: Matrizes de confusão normalizadas resultantes da avaliação da MLP, SVM e LDA em DS2, seguindo as especificações do experimento 1.

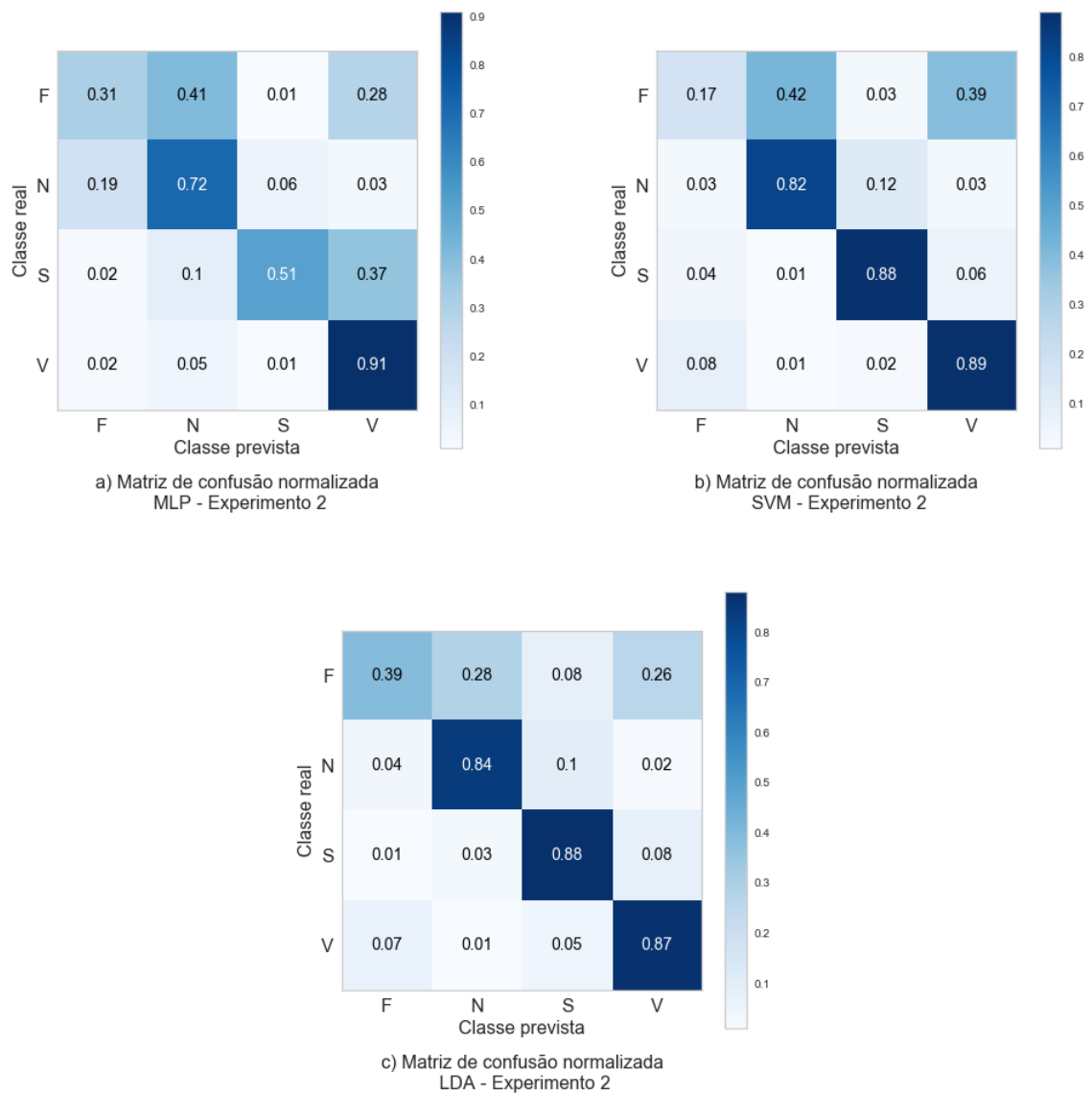


Figura 5.2: Matrizes de confusão normalizadas resultantes da avaliação da MLP, SVM e LDA em DS2, seguindo as especificações do experimento 2.

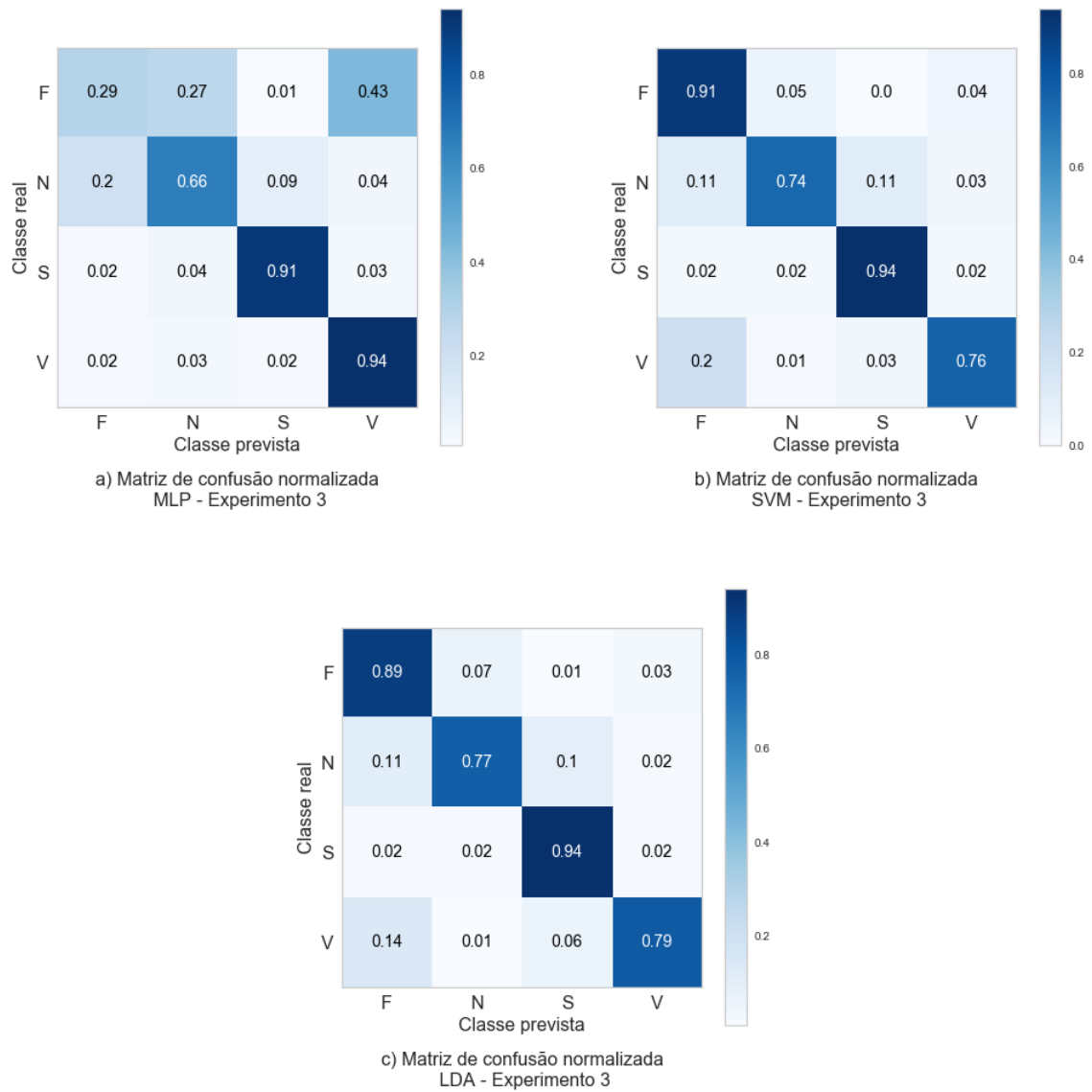


Figura 5.3: Matrizes de confusão normalizadas resultantes da avaliação da MLP, SVM e LDA em DS2, seguindo as especificações do experimento 3.

Capítulo 6

Conclusão

O presente trabalho propôs modelos eficientes para a classificação de arritmias cardíacas através de técnicas de aprendizado de máquina utilizando características obtidas a partir de sinais ECG do MIT-BIH. Além disso, o desenvolvimento foi embasado na metodologia interpaciente para construção dos conjuntos de dados e nas recomendações da AAMI.

A proposta de solucionar o desbalanceamento através da subamostragem aleatória das classe N e da geração de dados sintéticos utilizando a técnica ADASYN no conjunto de dados de entrada proporcionou melhoria significativa nas performances gerais dos classificadores. Além disso, a redução da dimensão do espaço de entrada também refletiu na melhoria de algumas métricas, principalmente as sensibilidades das classes S e F.

Comparados com a literatura, os classificadores propostos conseguem distinguir bem entre as classes de arritmia ao custo de uma queda na sensibilidade para a classe N. Em outras palavras, há uma quantia maior de FN para N, porém, para S, F e V, a maioria dos batimentos pertencentes a estas classes são corretamente classificados.

Assim como Mondéjar et al. [31], as características foram extraídas de apenas uma derivação para que os modelos demandassem de apenas uma medida para serem construídos. Apesar da simplificação associada à esta decisão, a utilização de múltiplas derivações poderia vir a agregar mais informação acerca das arritmias e melhorar ainda mais a classificação das mesmas.

6.1 Trabalhos futuros

Normalmente, os trabalhos da literatura propõem seus detectores semiautomáticos utilizando os registros do MIT-BIH. Tal ação é de extrema importância para padronizar os resultados obtidos e viabilizar a comparação entre trabalhos distintos, entretanto, é interessante que também sejam avaliados em outras bases de dados para verificar a capacidade generalização e conseqüentemente proporcionar uma melhor estimativa da aplicabilidade

dos mesmos em ambientes médicos. Logo, como trabalho futuro propõe-se a utilização de outros bancos para este propósito.

A obtenção dos sinais do ECG necessita que seja feita uma integração com os aparelhos de monitoramento cardíacos responsáveis por captá-los. A construção de uma aplicação integrada com modelos baseados em aprendizado profundo que classificam fotos de batimentos permitiria uma solução mais prática, a qual possibilitaria com que os médicos tirassem fotos dos trechos aos quais hajam dúvidas acerca do tipo de batimento.

Pelo bom reflexo da metodologia aplicada, outra possibilidade seria testa-la com sinais de outros tipos de doenças cardíacas, como infartos do miocárdio, verificando o impacto da geração dos dados sintéticos para a classificação de batimentos desta categoria.

Referências

- [1] Raschka, Sebastian e Vahid Mirjalili: *Python Machine Learning*. Packt Publishing Ltd, setembro 2017. ix, 9, 12
- [2] Kasabov, Nikola K.: *Foundations of Neural Networks, Fuzzy Systems, and Knowledge Engineering*. Marcel Alencar, 1996. ix, 14, 15
- [3] Thaler, Malcolm S.: *ECG Essencial - 7ed.* Artmed Editora, julho 2016. 1, 5, 6
- [4] Srinivasan, Neil e Richard Schilling: *Sudden cardiac death and arrhythmias*. Arrhythmia Electrophysiology Review, 7, 2018. 1
- [5] Philip de Chazal, M. O'Dwyer e R. B. Reilly: *Automatic classification of heartbeats using ECG morphology and heartbeat interval features*. IEEE Transactions on Biomedical Engineering, 51(7), julho 2004. 1, 24, 28, 45
- [6] Übeyli, Elif Derya: *ECG beats classification using multiclass support vector machines with error correcting output codes*. Digital Signal Processing, 17(3), maio 2007. 1, 20, 23, 24
- [7] Moody, G. B. e R. G. Mark: *The impact of the MIT-BIH Arrhythmia Database*. IEEE Engineering in Medicine and Biology Magazine, 20(3), maio 2001. 2, 27
- [8] Hall, John E. e Arthur C. GUYTON: *Guyton E Hall Tratado De Fisiologia Médica*. Elsevier Brasil, dezembro 2011. 4
- [9] Mohri, Mehryar, Afshin Rostamizadeh e Ameet Talwalkar: *Foundations of Machine Learning*. MIT Press, dezembro 2018. 5
- [10] Han, Jiawei, Jian Pei e Micheline Kamber: *Data Mining: Concepts and Techniques*. Elsevier, junho 2011. 6, 10, 15, 16
- [11] Haykin, Simon: *Redes Neurais: Princípios e Prática*. Bookman Editora, 2007. 6
- [12] Tan, Pang Ning, Michael Steinbach e Vipin Kumar: *Introduction to Data Mining*. Pearson Addison Wesley, 2006. 6, 8, 9, 11, 15, 16
- [13] Witten, Ian H. e Eibe Frank: *Data Mining: Practical Machine Learning Tools and Techniques, Second Edition*. Elsevier, julho 2005. 7
- [14] Cohen, Jacob: *A Coefficient of Agreement for Nominal Scales*. Educational and Psychological Measurement, 20(1), abril 1960. 9

- [15] Marsland, Stephen: *Machine Learning: An Algorithmic Perspective, Second Edition*. CRC Press, setembro 2015. 10
- [16] 1.2. linear and quadratic discriminant analysis — scikit-learn 0.21.3 documentation. https://scikit-learn.org/stable/modules/lda_qda.html#id4, acesso em 2019-09-15. 16
- [17] Hastie, Trevor, Robert Tibshirani e Jerome H. Friedman: *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*. Springer, 2009. 17
- [18] Gu, Q., Z. Cai, L. Zhu e B. Huang: *Data Mining on Imbalanced Data Sets*. Em *2008 International Conference on Advanced Computer Theory and Engineering*, dezembro 2008. 17, 18
- [19] Haibo He, Yang Bai, E. A. Garcia e Shutao Li: *ADASYN: Adaptive synthetic sampling approach for imbalanced learning*. Em *2008 IEEE International Joint Conference on Neural Networks (IEEE World Congress on Computational Intelligence)*, junho 2008. 17, 18
- [20] Inan, O. T., L. Giovangrandi e G. T. A. Kovacs: *Robust Neural-Network-Based Classification of Premature Ventricular Contractions Using Wavelet Transform and Timing Interval Features*. *IEEE Transactions on Biomedical Engineering*, 53(12), dezembro 2006. 21, 24
- [21] Singh, Brij N. e Arvind K. Tiwari: *Optimal selection of wavelet basis function applied to ECG signal denoising*. *Digital Signal Processing*, 16(3), maio 2006. 21, 30
- [22] Kutlu, Yakup e Damla Kuntalp: *A multi-stage automatic arrhythmia recognition and classification system*. *Computers in Biology and Medicine*, 41(1), janeiro 2011. 21, 23, 24
- [23] Mar, T., S. Zaunseder, J. P. Martínez, M. Llamedo e R. Poll: *Optimization of ECG Classification by Means of Feature Selection*. *IEEE Transactions on Biomedical Engineering*, 58(8), agosto 2011. 21, 23, 24, 28, 29, 35, 42, 45
- [24] Ghorbani Afkhami, Rashid, Ghanbar Azarnia e Mohammad Ali Tinati: *Cardiac arrhythmia classification using statistical and mixture modeling features of ECG signals*. *Pattern Recognition Letters*, 70, janeiro 2016. 21, 32
- [25] zhu, w, X. Chen, Y. Wang e L. Wang: *Arrhythmia Recognition and Classification Using ECG Morphology and Segment Feature Analysis*. *IEEE/ACM Transactions on Computational Biology and Bioinformatics*, 16(1), janeiro 2019. 21, 24, 28
- [26] Martis, Roshan Joy, U. Rajendra Acharya e Lim Choo Min: *ECG beat classification using PCA, LDA, ICA and Discrete Wavelet Transform*. *Biomedical Signal Processing and Control*, 8(5), setembro 2013. 21, 23, 24, 31, 33
- [27] Desai, U., R. J. Martis, C. G. Nayak e and G. Seshikala: *Machine intelligent diagnosis of ECG for arrhythmia classification using DWT, ICA and SVM techniques*. Em *2015 Annual IEEE India Conference (INDICON)*, dezembro 2015. 21

- [28] Pan, J. e W. J. Tompkins: *A Real-Time QRS Detection Algorithm*. IEEE Transactions on Biomedical Engineering, BME-32(3), março 1985. 22
- [29] Luz, Eduardo José da S., William Robson Schwartz, Guillermo Cámara-Chávez e David Menotti: *ECG-based heartbeat classification for arrhythmia detection: A survey*. Computer Methods and Programs in Biomedicine, 127, abril 2016. 23, 24, 38
- [30] Llamedo, M. e J. P. Martinez: *Heartbeat Classification Using Feature Selection Driven by Database Generalization Criteria*. IEEE Transactions on Biomedical Engineering, 58(3), março 2011. 23
- [31] Mondéjar-Guerra, V., J. Novo, J. Rouco, M.G. Penedo e M. Ortega: *Heartbeat classification fusing temporal and morphological information of ecgs via ensemble of classifiers*, 2019. 23, 24, 28, 35, 42, 45, 50
- [32] Rai, Hari Mohan, Anurag Trivedi e Shailja Shukla: *ECG signal processing for abnormalities detection using multi-resolution wavelet transform and Artificial Neural Network classifier*. Measurement, 46(9), novembro 2013. 23, 24
- [33] and, S. Osowski: *ECG beat recognition using fuzzy hybrid neural network*. IEEE Transactions on Biomedical Engineering, 48(11), novembro 2001. 24
- [34] Yıldırım, Özal, Paweł Pławiak, Ru San Tan e U. Rajendra Acharya: *Arrhythmia detection using deep convolutional neural network with long duration ECG signals*. Computers in Biology and Medicine, 102:411–420, novembro 2018. 24
- [35] Acharya, U. Rajendra, Shu Lih Oh, Yuki Hagiwara, Jen Hong Tan, Muhammad Adam, Arkadiusz Gertych e Ru San Tan: *A deep convolutional neural network model to classify heartbeats*. Computers in Biology and Medicine, 89, outubro 2017. 24
- [36] Xia, Yong, Naren Wulan, Kuanquan Wang e Henggui Zhang: *Detecting atrial fibrillation by deep convolutional neural networks*. Computers in Biology and Medicine, 93:84–92, fevereiro 2018. 24
- [37] *Mit-bih arrhythmia database directory*. <https://archive.physionet.org/physiobank/database/html/mitdbdir/intro.htm>, acesso em 19 Agosto 2019. 26, 27
- [38] Association for the Advancement of Medical Instrumentation e American National Standards Institute: *Testing and reporting performance results of cardiac rhythm and ST-segment measurement algorithms*. The Association, Arlington, VA, 1999. OCLC: 56564461. 28, 41
- [39] Zhang, Zhancheng, Jun Dong, Xiaoqing Luo, Kup Sze Choi e Xiaojun Wu: *Heart-beat classification using disease-specific feature selection*. Computers in Biology and Medicine, 46, março 2014. 28
- [40] Lin, H., Sz-Ying Liang, Y. Ho, Y. Lin e H. Ma: *Discrete-wavelet-transform-based noise reduction and R wave detection for ECG signals*. outubro 2013. 31

- [41] Donoho, David L e Iain M Johnstone: *Ideal spatial adaptation by wavelet shrinkage*. Biometrika, 81(3), setembro 1994. <https://doi.org/10.1093/biomet/81.3.425>.
31