

**UNIVERSIDADE FEDERAL DE JUIZ DE FORA
FACULDADE DE ENGENHARIA & INSTITUTO DE CIÊNCIAS EXATAS
PROGRAMA DE PÓS-GRADUAÇÃO EM MODELAGEM
COMPUTACIONAL**

Tales Humberto de Aquino Boratto

**Aprendizado de Máquina para a Classificação Automática de Pratos de
Bateria Conforme a Proporção de Estanho Presente em suas Ligas de Bronze**

Juiz de Fora

2022

Tales Humberto de Aquino Boratto

Aprendizado de Máquina para a Classificação Automática de Pratos de Bateria Conforme a Proporção de Estanho Presente em suas Ligas de Bronze

Dissertação apresentada ao Programa de Pós-Graduação em Modelagem Computacional da Faculdade de Engenharia & Instituto de Ciências Exatas da Universidade Federal de Juiz de Fora como requisito parcial à obtenção do título de Mestre em Modelagem Computacional. Área de concentração: Modelagem Computacional.

Orientador: Prof. Dr. Leonardo Goliatt da Fonseca

Coorientador: Prof. Dr. Alexandre Abrahão Cury

Juiz de Fora

2022

Ficha catalográfica elaborada através do programa de geração automática da Biblioteca Universitária da UFJF, com os dados fornecidos pelo(a) autor(a)

Boratto, Tales Humberto de Aquino.

Aprendizado de Máquina para a Classificação Automática de Pratos de Bateria Conforme a Proporção de Estanho Presente em suas Ligas de Bronze / Tales Humberto de Aquino Boratto. -- 2022. 104 f. : il.

Orientador: Leonardo Goliatt da Fonseca

Coorientador: Alexandre Abrahão Cury

Dissertação (mestrado acadêmico) - Universidade Federal de Juiz de Fora, ICE/Engenharia. Programa de Pós-Graduação em Modelagem Computacional, 2022.

1. Aprendizado de Máquina. 2. Caracterização de Materiais. 3. Computação Musical. 4. Ligas de Bronze. 5. Pratos de Bateria. I. Fonseca, Leonardo Goliatt da, orient. II. Cury, Alexandre Abrahão, coorient. III. Título.

Tales Humberto de Aquino Boratto

Aprendizado de Máquina para a Classificação Automática de Pratos de Bateria Conforme a Proporção de Estanho Presente em suas Ligas de Bronze

Dissertação apresentada ao Programa de Pós-graduação em Modelagem Computacional da Universidade Federal de Juiz de Fora como requisito parcial à obtenção do título de Mestre em Modelagem Computacional. Área de concentração: Modelagem Computacional.

Aprovada em 18 de novembro de 2022.

BANCA EXAMINADORA

Prof(a) Dr(a). Leonardo Goliatt da Fonseca - Orientador

Universidade Federal de Juiz de Fora

Prof(a) Dr(a). Alexandre Abrahão Cury - Coorientador

Universidade Federal de Juiz de Fora

Prof(a) Dr(a). Eduardo Pestana de Aguiar

Universidade Federal de Juiz de Fora

Prof(a) Dr(a). Flávio Luiz Schiavoni

Universidade Federal de São João Del Rei

Prof(a) Dr(a). Elineudo Pinho de Moura

Universidade Federal do Ceará

Juiz de Fora, 31/10/2022.



Documento assinado eletronicamente por **Eduardo Pestana de Aguiar, Professor(a)**, em 18/11/2022, às 17:02, conforme horário oficial de Brasília, com fundamento no § 3º do art. 4º do [Decreto nº 10.543, de 13 de novembro de 2020](#).



Documento assinado eletronicamente por **Leonardo Goliatt da Fonseca, Professor(a)**, em 18/11/2022, às 17:03, conforme horário oficial de Brasília, com fundamento no § 3º do art. 4º do [Decreto nº 10.543, de 13 de novembro de 2020](#).



Documento assinado eletronicamente por **Alexandre Abrahao Cury, Professor(a)**, em 18/11/2022, às 17:03, conforme horário oficial de Brasília, com fundamento no § 3º do art. 4º do [Decreto nº 10.543, de 13 de novembro de 2020](#).



Documento assinado eletronicamente por **Elineudo Pinho de Moura, Usuário Externo**, em 18/11/2022, às 17:08, conforme horário oficial de Brasília, com fundamento no § 3º do art. 4º do [Decreto nº 10.543, de 13 de novembro de 2020](#).



Documento assinado eletronicamente por **Flávio Luiz Schiavoni, Usuário Externo**, em 29/11/2022, às 15:00, conforme horário oficial de Brasília, com fundamento no § 3º do art. 4º do [Decreto nº 10.543, de 13 de novembro de 2020](#).



A autenticidade deste documento pode ser conferida no Portal do SEI-Uffj (www2.uffj.br/SEI) através do ícone Conferência de Documentos, informando o código verificador **1019267** e o código CRC **CC2D887C**.

AGRADECIMENTOS

Primeiramente, agradeço a Deus por sempre iluminar os meus caminhos e me fornecer saúde e força para continuar caminhando.

Agradeço também à minha família. Aos meus pais, pelo amor, apoio incondicional e pela minha formação como pessoa. E, aos meus irmãos, por todo suporte durante essa trajetória.

Aos meus orientadores, professores e amigos Leonardo Goliatt e Alexandre Cury por acreditarem na minha proposta e me guiarem nessa área totalmente nova e por me darem todo o suporte necessário. Sem vocês, este trabalho não existiria.

Ao companheiro de música e amigo Sérgio Leite e à toda a equipe do estúdio Argus, por todo suporte e atenção e por disponibilizar a estrutura e os equipamentos necessários para a aquisição dos dados deste trabalho.

Agradeço ao meu mentor musical e amigo Bráulio Mayrink e à toda equipe da escola de música LIVE pela minha formação musical e por estarem sempre me apoiando em meus projetos. Ao amigo Francisco Domene por me inserir no mundo dos pratos, me ensinar, incentivar e dar todo o suporte possível em minhas pesquisas. Ao amigo Bruno Schell por também me incentivar e me ensinar bastante sobre pratos.

Agradeço ao professor Eduardo Aguiar pela paciência e ensinamentos passados no decorrer do curso. Ao professor Flávio Luiz Schiavoni, por me apresentar e abrir as portas da área de computação musical. E também não poderia deixar de agradecer ao professor Elineudo Moura pela colaboração e disponibilidade. É um prazer tê-los como membros da banca. Muito obrigado!

À minha namorada e a todos os meus amigos que estiveram presentes e que conheci ao longo desta trajetória. São tantos, mas cada um sabe o valor e a contribuição de vocês para com este trabalho! Deixo aqui o meu muito obrigado a todos vocês!

Por fim, não poderia deixar de agradecer à Universidade Federal de Juiz de Fora e ao Programa de Pós Graduação em Modelagem Computacional pela infraestrutura e pela formação acadêmica, e também à Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Código de Financiamento 001 - pelo suporte financeiro.

O medo mora perto das ideias loucas.

(Dudu Falcão)

RESUMO

Os diversos fatores que influenciam no som de um prato de bateria, combinados com os aspectos externos que modificam a sua percepção sonora, tornam o estudo da acústica destes instrumentos mais desafiadora. Pesquisas que focam no estudo de pratos de bateria são encontradas na literatura com diversas abordagens e métodos de investigação. Entretanto, ao considerar o contexto de aprendizado de máquina, a aplicação deste tipo de técnica para o estudo de pratos de bateria ainda é pouco explorada e, normalmente, foca na identificação das categorias de pratos e o modo como foram tocados. Dessa forma, o uso de ferramentas computacionais baseadas na aprendizagem de máquina para investigar os elementos acústicos destes instrumentos torna-se um tópico de estudo com potencial para ser desenvolvido. Inspirado nisso, este trabalho propõe a classificação de pratos de bateria conforme seus materiais constituintes, uma vez que a liga metálica assume uma parcela de responsabilidade significativa em sua acústica. Para tanto, um procedimento para a captura padronizada dos sons destes instrumentos foi desenvolvido e possibilitou construir uma base de dados composta por 276 áudios, referente a 4 pratos de bateria. Diferentes conjuntos de atributos, técnicas de seleção e classificadores foram avaliados, com a finalidade de propor um modelo computacional eficiente para a resolver tal problema. Como resultado, uma acurácia média de 99,75% foi alcançada ao aplicar a regressão logística em um subconjunto de 30 atributos formados por 17 características temporais e 13 Coeficientes Mel-Cepstrais de Frequência (MFCC), selecionados por uma técnica baseada na análise do limiar de correlação entre as variáveis. Por outro lado, ao priorizar a dimensionalidade do subconjunto selecionado, a utilização de uma técnica de seleção baseada na regressão LASSO, seguida da análise de correlação, proporciona identificar um subconjunto de 5 atributos que promovem uma acurácia média de 99,37%, também utilizando a regressão logística como classificador.

Palavras-chave: Aprendizado de Máquina. Computação Musical. Ligas de Bronze. Pratos de Bateria. Reconhecimento de Materiais.

ABSTRACT

The many factors that influence the sound of a cymbal, combined with the external aspects that modify its sound perception, make the study of the acoustics of these instruments more challenging. Research that focuses on the study of drum cymbals is found in the literature, with various approaches and methods of investigation. However, considering the context of machine learning, the application of this type of technique to the study of drum cymbals is still little explored and usually focus on identifying the categories of cymbals and the way they were played. Thus, the use of machine learning-based computational tools to investigate the acoustic elements of these instruments becomes a topic of study with potential for further development. Inspired by this, this paper proposes classifying drum cymbals according to their constituent materials, since the metal alloy assumes a significant share of responsibility in their acoustics. To this end, a procedure for the standardized capture of the sounds of these instruments was developed and allowed to build a database consisting of 276 audios, referring to 4 cymbals. Different sets of attributes, selection techniques and classifiers were evaluated in order to propose an efficient computational model to solve this problem. As a result, an average accuracy of 99.75% was achieved when applying logistic regression on a subset of 30 attributes formed by 17 temporal features and 13 Mel-Frequency Cepstral Coefficients, selected by a technique based on the analysis of the correlation threshold between variables. On the other hand, when prioritizing the dimensionality of the selected subset, the use of a selection technique based on LASSO regression followed by correlation analysis, provides to identify a subset of 5 attributes that promote an average accuracy of 99.37%, also using logistic regression as a classifier.

Keywords: Machine Learning. Music Computing. Bronze Alloys. Cymbals. Material Recognition.

LISTA DE ILUSTRAÇÕES

Figura 1 – Detalhamento das regiões de um prato de bateria.	23
Figura 2 – Imagem ilustrativa da categoria de prato conhecida por <i>China</i> . .	24
Figura 3 – Identificação de algumas categorias de pratos.	25
Figura 4 – Diagramas de fase cobre-estanho (Cu-Sn) na região rica em cobre.	33
Figura 5 – Fluxograma da Fabricação Artesanal.	35
Figura 6 – Metal líquido sendo vertido nas lingoteiras côncavas durante o processo de fundição.	36
Figura 7 – Vista lateral da laminação de planos.	37
Figura 8 – Ilustração em vista superior e lateral de um <i>blank</i> de bronze B20 sem cúpula.	38
Figura 9 – Ilustração dos processos de martelamento manual e usinagem. .	39
Figura 10 – Fluxograma da Fabricação Não Artesanal Brasileira.	40
Figura 11 – Ilustração de <i>blanks</i> de Bronze B8.	41
Figura 12 – Ilustração do Processo de Repuxo.	42
Figura 13 – Fluxograma da Fabricação Híbrida.	42
Figura 14 – Influência do arco de curvatura no som de um prato.	45
Figura 15 – Ilustração da transformação do som em sinal de áudio.	46
Figura 16 – Ilustração da resolução de um sinal segundo o número de bits de sua representação.	47
Figura 17 – Comportamento da correlação entre duas variáveis numéricas. .	49
Figura 18 – Representação das etapas do teste ANOVA.	50
Figura 19 – Relação entre variância intra-grupo e inter-grupos.	51
Figura 20 – Fluxograma das etapas do trabalho.	53
Figura 21 – Curvaturas medidas dos pratos utilizados.	54
Figura 22 – Pratos utilizados para a aquisição dos áudios.	55
Figura 23 – Dano na borda do prato.	55
Figura 24 – Resposta em frequência e padrão polar do microfone AKG C414 XL II.	56
Figura 25 – Resposta em frequência e padrão polar do microfone Shure PG81.	56
Figura 26 – Configuração do procedimento para a aquisição de dados.	57
Figura 27 – Pedal de bumbo utilizado para a aquisição dos sons.	58
Figura 28 – Ilustração de um arquivo <i>.rpp</i> gerado no processo de gravação. .	59
Figura 29 – Etapas para o cálculo dos MFCCs.	63
Figura 30 – Formato da Janela de Hamming.	64
Figura 31 – Método de construção do banco de filtros.	64
Figura 32 – Resposta em frequência do banco de filtros.	65
Figura 33 – Ilustração das funções de pertinência.	69

Figura 34 – Fluxograma ilustrativo da sequência de operações realizadas nos experimentos computacionais.	73
Figura 35 – Gráfico de barras com as melhores métricas de desempenho obtidas por cada técnica de seleção, considerando o conjunto C3.	75
Figura 36 – Matrizes de confusão média considerando os melhores modelos em cada abordagem de seleção de características aplicadas à base de dados C3.	77
Figura 37 – Comportamento da acurácia média em relação aos limiares de correlação referentes ao conjunto de dados C3.	78
Figura 38 – Comportamento das acurácias em relação ao limiar de correlação para o conjunto de dados C1.	93
Figura 39 – Comportamento das acurácias em relação ao limiar de correlação para o conjunto de dados C2.	93
Figura 40 – Comportamento das acurácias em relação ao limiar de correlação para o conjunto de dados C3.	93
Figura 41 – Comportamento das acurácias em relação ao limiar de correlação para o conjunto de dados C4.	94
Figura 42 – Comportamento das acurácias em relação ao limiar de correlação para o conjunto de dados C5.	94
Figura 43 – Gráfico de caixas ilustrando a distribuição dos valores dos atributos temporais X1 a X6.	95
Figura 44 – Gráfico de caixas ilustrando a distribuição dos valores dos atributos temporais X7 a X12.	96
Figura 45 – Gráfico de caixas ilustrando a distribuição dos valores dos atributos temporais X13 a X18.	97
Figura 46 – Gráfico de caixas ilustrando a distribuição dos valores dos MFCCs 1 a 6.	98
Figura 47 – Gráfico de caixas ilustrando a distribuição dos valores dos MFCCs 7 a 12.	99
Figura 48 – Gráfico de caixas ilustrando a distribuição dos valores dos MFCCs 13 a 18.	100
Figura 49 – Gráfico de caixas ilustrando a distribuição dos valores dos MFCCs 19 a 24.	101
Figura 50 – Gráfico de caixas ilustrando a distribuição dos valores dos MFCCs 25 a 30.	102
Figura 51 – Gráfico de caixas ilustrando a distribuição dos valores dos MFCCs 31 a 36.	103
Figura 52 – Gráfico de caixas ilustrando a distribuição dos valores dos MFCCs 37 a 40.	104

LISTA DE TABELAS

Tabela 1	– Exemplos de nomenclatura de ligas de bronze utilizadas em pratos de bateria.	23
Tabela 2	– Abordagem convencional para a interpretação dos coeficientes de correlação.	49
Tabela 3	– Detalhamento da identificação dos arquivos de áudios.	59
Tabela 4	– Detalhamento das bases de dados utilizadas nos experimentos computacionais.	66
Tabela 5	– Melhores desempenhos médios percentuais obtidos para cada conjunto de atributos, bem como o modelo de classificação e a técnica de seleção associados.	74
Tabela 6	– Resultado do teste ANOVA ao comparar a distribuição dos desempenhos obtidos pelos melhores casos dos conjuntos de atributos C3, C4 e C5. O valor P obtido é significativo para $P < 0,05$	75
Tabela 7	– Classificadores associados aos melhores casos de cada abordagem de seleção, dimensão do subconjunto ótimo e listagem dos atributos selecionados, correspondentes ao conjunto de dados C3.	76
Tabela 8	– Desempenhos médios obtidos para o conjunto de dados C1, considerando todas as técnicas de seleção e os dois classificadores implementados.	91
Tabela 9	– Desempenhos médios obtidos para o conjunto de dados C2, considerando todas as técnicas de seleção e os dois classificadores implementados.	91
Tabela 10	– Desempenhos médios obtidos para o conjunto de dados C3, considerando todas as técnicas de seleção e os dois classificadores implementados.	92
Tabela 11	– Desempenhos médios obtidos para o conjunto de dados C4, considerando todas as técnicas de seleção e os dois classificadores implementados.	92
Tabela 12	– Desempenhos médios obtidos para o conjunto de dados C5, considerando todas as técnicas de seleção e os dois classificadores implementados.	92

LISTA DE ABREVIATURAS E SIGLAS

PPGMC	Programa de Pós-Graduação em Modelagem Computacional
UFJF	Universidade Federal de Juiz de Fora
ANOVA	Análise de Variância (<i>Analisis Of Variance</i>)
B8	Bronze com 8% de Estanho em sua composição química teórica
B10	Bronze com 10% de Estanho em sua composição química teórica
B20	Bronze com 20% de Estanho em sua composição química teórica
CCC	Cúbica de Corpo Centrado
CDA	Análise Discriminante Canônica (<i>Canonical Discriminant Anaysis</i>)
CF	Frequências Centrais (<i>Center Frequencies</i>)
CFC	Cúbica de Face Centrada
DCT	Transformada Discreta de Cosseno (<i>Discrete Cossine Transform</i>)
FFT	Transformada Rápida de Fourier (<i>Fast Fourier Transform</i>)
FN	Falso Negativo
FP	Falso Positivo
KNN	K-Vizinhos Mais Próximos (<i>K-Nearest Neighbors</i>)
LFC	Cepstrum de Frequência Linear (<i>Linear Frequency Cepstrum</i>)
LSF	Frequências Espectrais Lineares (<i>Linear Spectral Frequencies</i>)
MFCC	Coefficientes Mel-Cepstrais de Frequência (<i>Mel Frequency Cepstral Coef-ficients</i>)
MIR	Recuperação de Informação Musical (<i>Musical Information Retrieval</i>)
OVR	Um contra o Resto (<i>One Versus Rest</i>)
OVA	Um contra Todos (<i>One Versus All</i>)
PP	Positivo Previsto
RL	Regressão Logística
SVM	Máquina de Vetor Suporte (<i>Support Vector Machine</i>)
TP	Positivo Verdadeiro (<i>True Positive</i>)
TSFEL	Biblioteca para a Extração de Atributos de Séries Temporais (<i>Time Series Feature Extraction Library</i>)
ZCR	Taxa de cruzamento do zero (<i>Zero Crossing Rate</i>)

LISTA DE SÍMBOLOS

α	Letra grega Alfa
β	Letra grega Beta
σ	Letra grega Sigma
θ	Letra grega Theta
Cu	Representação do elemento químico Cobre
Sn	Representação do elemento químico Estanho
\in	Símbolo matemático para Pertencimento
\mathbb{R}	Representação do conjunto de números Reais

SUMÁRIO

1	INTRODUÇÃO	23
1.1	OBJETIVOS	29
1.1.1	Objetivo Geral	29
1.1.2	Objetivos Específicos	29
1.2	RELEVÂNCIA E CONTRIBUIÇÕES	29
1.2.1	Motivação Pessoal	30
1.3	ESTRUTURA DO TRABALHO	31
2	FUNDAMENTAÇÃO TEÓRICA	33
2.1	LIGAS DE BRONZE	33
2.2	PROCESSOS DE FABRICAÇÃO	34
2.2.1	Método Artesanal Clássico	35
2.2.2	Método Não-Artesanal Brasileiro	38
2.2.3	Outros Métodos	40
2.3	ACÚSTICA	42
2.4	SINAIS DE ÁUDIO	45
2.5	CORRELAÇÃO	48
2.6	ANÁLISE DE VARIÂNCIA (ANOVA)	49
3	MATERIAL E MÉTODOS	53
3.1	AQUISIÇÃO DOS ÁUDIOS	53
3.1.1	Processamento dos Sinais	58
3.2	EXTRAÇÃO DE ATRIBUTOS	60
3.2.1	Atributos Temporais	60
3.2.2	Coefficientes Mel-Cepstrais de Frequência	61
3.3	BASES DE DADOS	64
3.4	MODELOS DE CLASSIFICAÇÃO	66
3.4.1	Regressão Logística	66
3.4.2	Classificador Fuzzy	68
3.5	SELEÇÃO DE ATRIBUTOS	70
3.6	INDICADORES DE DESEMPENHO	71
4	RESULTADOS E DISCUSSÃO	73
4.1	EXPERIMENTOS COMPUTACIONAIS	73
4.2	RESULTADOS E ANÁLISES	74
4.3	VANTAGENS E LIMITAÇÕES	79
5	CONCLUSÃO	81
5.1	TRABALHOS FUTUROS	82
5.2	PARTICIPAÇÃO E COLABORAÇÃO EM EVENTOS	82
5.3	PRODUÇÕES ACADÊMICAS	83

REFERÊNCIAS	85
APÊNDICE A – DESEMPENHOS ALCANÇADOS EM CADA CASO CONSIDERADO	91
APÊNDICE B – DISTRIBUIÇÃO DOS VALORES DE CADA ATRIBUTO EXTRAÍDO	95

1 INTRODUÇÃO

Os pratos de bateria são instrumentos musicais percussivos classificados como idiofônicos¹. Desde os seus primeiros exemplares, datados de cerca de 5000 anos (PINKSTERBOER; MATTINGLY, 1992), estes instrumentos musicais foram utilizados com diferentes finalidades e de diversas maneiras até se tornarem o que conhecemos hoje como componentes fundamentais na configuração dos instrumentos de qualquer baterista da atualidade. A extensa variedade de modelos e suas aplicações mostram o quanto estes instrumentos têm se tornado cada vez mais relevantes no contexto musical. Dentre os possíveis materiais com os quais podem ser fabricados, os bronzes, tratados como sendo basicamente ligas metálicas de cobre (*Cu*) com adição de estanho (*Sn*), destacam-se por serem os mais utilizados. No entanto, a depender da concentração de estanho, diferentes ligas de bronzes podem ser obtidas, proporcionando materiais com propriedades mecânicas e acústicas distintas. Por conta disso, grande parte das empresas adota uma nomenclatura de seus materiais baseada em suas composições químicas, conforme os exemplos mostrados na Tabela 1. Nela, a letra *B* refere-se a *Bronze* e a numeração subsequente indica a porcentagem de estanho em peso presente na liga.

Tabela 1 – Exemplos de nomenclatura de ligas de bronze utilizadas em pratos de bateria.

Liga de Bronze	Sn (% peso)	Cu (% peso)
B8	8	92
B10	10	90
B20	20	80

De modo geral, os pratos de bateria apresentam uma geometria axissimétrica simples que consiste em uma superfície côncava dotada de um ressalto localizado em sua região central, como mostrado pela Figura 1.

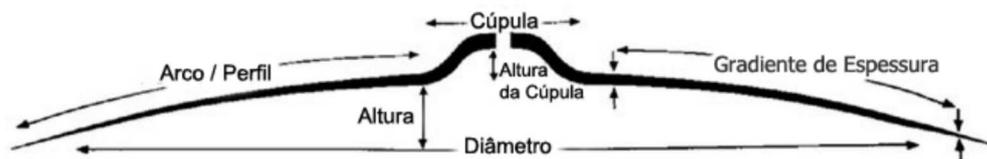
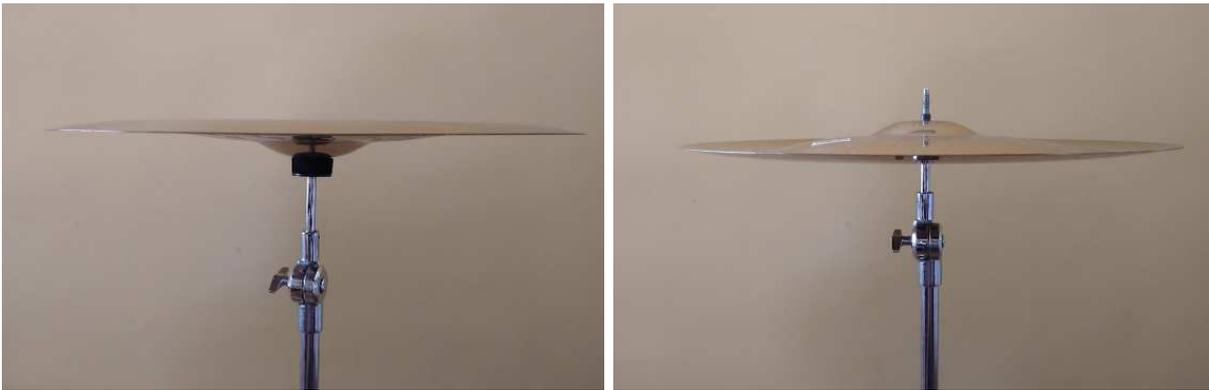


Figura 1 – Detalhamento das regiões de um prato de bateria.

Fonte: (PINKSTERBOER; MATTINGLY, 1992)

É importante ressaltar que os parâmetros descritos na Figura 1 variam para cada prato, conforme a característica acústica desejada para tal instrumento. Indo além, determinados tipos de pratos podem apresentar uma estrutura geométrica bastante diferente, como é o caso dos *Chinas* (Figura 2) que apresenta a borda curvada para o lado oposto ao convencional mostrado na Figura 1. Ainda nesse sentido, os pratos de bateria

¹ Aqueles cujo som é produzido pela vibração da sua própria estrutura.



(a) Posicionado com a cúpula voltada para baixo (b) Posicionado com a cúpula voltada para cima. (maneira mais usual de ser utilizado).

Figura 2 – Imagem ilustrativa da categoria de prato conhecida por *China*.

Fonte: Acervo do Autor.

podem ser divididos por categorias quanto à sua aplicação ou à maneira com que são utilizados. Algumas destas categorias são descritas abaixo e ilustradas em destaque na Figura 3.

1. *Hi-Hat*: Também conhecidos como “Chimbal”, esta categoria de prato é, na verdade, uma combinação de dois pratos que trabalham em conjunto, sendo montados em um suporte apropriado em que o baterista consegue regular a distância entre eles com o pé, através de um pedal. Geralmente, o chimbal é utilizado para a condução de uma música, podendo apresentar uma variedade sonora devido à sua regulagem de abertura (distância entre os pratos). Adicionalmente, esta categoria de prato também pode ser utilizada com o pé, principalmente para marcação rítmica.
2. *Crash*: Também são conhecidos como “Pratos de Ataque”, pois a sua principal utilização é em momentos de destaque ou acentuações das músicas. Sua geometria é detalhada na Figura 1.
3. *Splash*: São pratos de geometria semelhante aos *Crashes*, porém com diâmetros menores, apresentando sons de duração extremamente curta. Sendo assim, são mais utilizados em acentuações pontuais e rápidas, dando um efeito diferente na música.
4. *Ride*: Também conhecidos por “Prato de Condução” ou apenas “Condução”, esta categoria de pratos é geometricamente semelhante aos *Crashes*, embora sejam, de maneira geral, mais pesados e com diâmetros maiores. Além disso, possuem características acústicas diferentes, uma vez que a sua utilização principal é voltada para a condução de uma música.
5. *China*: Esta categoria de prato se diferencia das demais por apresentar uma geometria diferenciada, como mencionado anteriormente e ilustrado na Figura 2. Os *Chinas*

apresentam um som bastante característico e são normalmente utilizados também em acentuações musicais, embora com menor frequência do que os *Crashes*, ou para dar destaque em algum momento da música.

6. *Stack*: Como a tradução do nome sugere, esta categoria de prato consiste do empilhamento de 2 ou mais pratos. Apresentam uma gama de possibilidades sonoras, a depender da combinação dos pratos empilhados. No entanto, de maneira geral, apresentam um som característico e são utilizados para dar efeitos nas músicas, devido aos seus sons diferenciados.



Figura 3 – Identificação de algumas categorias de pratos.

Fonte: (STEEN, 2022)

Apesar desta simplicidade geométrica, a acústica de um prato é definida por inúmeros fatores que contribuem de diferentes maneiras na construção do som destes instrumentos, como será abordado com mais detalhes na Subseção 2.3. SCHROEDER et al. (SCHROEDER et al., 2007) afirmam que o som de um prato possui formas de ondas que não se repetem, sendo compostas por um grande número de frequências não relacionadas. Portanto, sem um senso de tom definido. Além disso, é dito naquele mesmo trabalho que a taxa de propagação das ondas em um prato é inversamente proporcional às dimensões da flexão da superfície na região solicitada durante o impacto. O pulso inicial, assim como as suas reflexões causadas pela limitação da borda do prato, propagam até a região central do prato, onde são novamente refletidas em direção a borda e se espalham por todo o corpo do instrumento, fazendo-o vibrar. Estes aspectos são reafirmados por HARRISON; HILL (HARRISON; HILL, 2013), que utilizam dos argumentos de SCHROEDER et al.

(SCHROEDER et al., 2007) para procurar a melhor região de captação do som de um prato em atuações ao vivo.

Uma vez que a complexidade acústica dos pratos de bateria é consequência dos diversos aspectos que contribuem para a construção e a percepção de seus sons, este é um tema que aguça a curiosidade de pesquisadores que se propõem a investigar alguns destes elementos. Estudar os padrões de vibração talvez seja a maneira mais comum de buscar compreender o comportamento acústico dos pratos, como é o caso de (GIULIO et al., 2001) e (PERRIN et al., 2008). Ainda nesse sentido, um modelo detalhado em elementos finitos foi desenvolvido por KASELOURIS et al. (KASELOURIS et al., 2022) para estudar o comportamento vibroacústico de um prato de ataque de 19 polegadas e de um *splash* de 10 polegadas de diâmetro. JAVH; SLAVIČ; BOLTEŽAR (JAVH; SLAVIČ; BOLTEŽAR, 2017) analisaram os potenciais de medições ópticas com câmaras de alta velocidade no campo da análise modal, utilizando um método simplificado de fluxo óptico baseado em gradiente otimizado para deslocamentos harmônicos de subpixel. Já AIHARA; ITO (AIHARA; ITO, 2022) centraram suas atenções na investigação do mecanismo de geração de som destes instrumentos, avaliando e comparando as características vibratórias e acústicas de pratos criados por diferentes fabricantes e métodos de processamento.

A influência dos elementos que contribuem para a formação da acústica dos pratos de bateria também tem sido estudada. OGAWA et al. (OGAWA et al., 2020) investigaram o efeito que o tamanho da cúpula provoca nas características acústicas de pratos de bateria. Para isso, os sons e as vibrações foram medidas, a frequência natural e a forma do modo de vibração foram obtidas através de elementos finitos, calculando, para cada modo, a eficiência da radiação sonora. OSAMURA et al. (OSAMURA et al., 2016) utilizaram a radiação síncrotron para medir a distribuição das tensões/deformações residuais em dois tipos de pratos de bateria — após a sua formação via *spinforming* e após o martelamento — relacionando quantitativamente com as características de frequência de seus sons. Similarmente, KURATANI et al. (KURATANI et al., 2016) procuraram investigar o efeito da etapa de martelamento nos pratos de bateria, utilizando a análise de elementos finitos como ferramenta. Mais recentemente, NGUYEN; TOUZÉ (NGUYEN; TOUZÉ, 2019) consideraram uma variação de espessura ao longo do raio (*taper*) no modelo utilizado para investigar as vibrações não lineares de pratos de bateria.

Ademais, o estudo da influência do material na acústica de instrumentos musicais é um assunto frequentemente trabalhado, seja no desenvolvimento de novos materiais ou na investigação dos que já são utilizados. SLAMET et al. (SLAMET et al., 2021) investigaram o efeito da microestrutura, densidade e propriedades mecânicas em relação às propriedades acústicas em ligas de bronzes (Cu-Sn) com proporções de estanho que variam de 20% a 24% em peso. SUGITA; PRIAMBADI (SUGITA; PRIAMBADI, 2017) tiveram como objetivo estudar o efeito do espaçamento dos braços dendríticos sobre as características acústicas da liga de bronze Cu-20Sn. Como resultados, o tamanho das

formas de dendríticas afetaram tanto as propriedades mecânicas quanto acústicas deste material. Já SUGITA; SOEKRISNO; SUYITNO (SUGITA; SOEKRISNO; SUYITNO, 2011) estavam interessados em investigar o efeito da composição química de bronzes de silício (Cu-Si) sobre as suas propriedades mecânicas e acústicas, com o objetivo de encontrar uma nova alternativa de material a ser utilizado em instrumentos musicais.

No contexto de computação musical, a aplicação de técnicas de aprendizado de máquina tem ganhado cada vez mais espaço na área, seja em problemas de composição musical, separação de fontes, de reconhecimento de instrumentos musicais, dentre outros. Dando destaque para esta última aplicação, existem estudos que buscam a reconhecer se um instrumento é percussivo ou não percussivo, como é o caso de (KRZYWICKI, 2019), classificar diferentes instrumentos musicais (NAGAWADE; RATNAPARKHE, 2017; RATHIKARANI; DHANALAKSHMI; S., 2020; VIEIRA et al., 2021), reconhecer as diferentes partes da bateria (CHHABRA et al., 2020; HERRERA; YETERIAN; GOUYON, 2002) e também estudos que se utilizam destas ferramentas de aprendizado de máquina para a classificação de pratos de bateria como é o caso de (CAVACO; ALMEIDA, 2012; SOUZA; BATISTA; SOUZA-FILHO, 2015) e também do já mencionado (HERRERA; YETERIAN; GOUYON, 2002).

CAVACO; ALMEIDA (CAVACO; ALMEIDA, 2012) recorreram à matriz de fatoração não negativa (*Non-Negative Matrix Factorization*) e ao modelo K-vizinhos mais próximos (*K-Nearest Neighbors - KNN*) (com $K=1$) para a classificação de diferentes pratos de bateria, considerando 7 combinações de categorias de pratos:

1. *Splash* e *China*,
2. *Crashes* de 14" e 16",
3. *Splash* e *Crash* de 16",
4. *China* e *Crash* 16",
5. *hi-hat* fechado e *Ride*, ambos golpeados na borda,
6. *China* e *Crashes* de 14" e 16",
7. *Splash* e *Crashes* de 14" e 16".

Dessa forma, os autores conseguiram atingir uma taxa de classificação de 95% para os grupos que consideram duas categorias de pratos e 86% para os grupos que consideram 3 tipos.

SOUZA; BATISTA; SOUZA-FILHO (SOUZA; BATISTA; SOUZA-FILHO, 2015) aprofundaram as investigações e buscaram não só classificar pratos de bateria conforme os seus tipos, no caso *China*, *Crash*, *Hi-hat*, *Ride* e *Splash*, como também a maneira como se

deu o toque, o que constitui um problema com 12 rótulos para a classificação. Para isso, os autores exploraram diversos descritores de áudios, tais como Frequências Espectrais Lineares (*Linear Spectral Frequencies - LSF*), Coeficientes Mel Cepstrais de Frequência (*Mel-Frequency Cepstral Coefficients - MFCC*), Cepstrum de Frequência Linear (*Linear Frequency Cepstrum - LFC*), Temporais e Espectrais, e também diferentes classificadores: Naive Bayes, C4.5, Florestas Aleatórias (*Random Forest*), KNN e Máquina de Vetor Suporte (*Support Vector Machine - SVM*) com Otimização Sequencial Mínima. Por fim, para a classificação por categorias de pratos, os autores conseguiram atingir uma acurácia de 96,59%. Já para a classificação quanto à maneira do toque, atingiram uma acurácia de 91,54%.

No caso de HERRERA; YETERIAN; GOUYON (HERRERA; YETERIAN; GOUYON, 2002), 50 atributos, 3 métodos de seleção de características e 5 classificadores foram utilizados para encontrar as melhores variáveis em três problemas de classificação que envolvem os tambores e os pratos de baterias. A primeira categoria de problema consistiu em identificar se o som referia-se a um tambor ou a um prato. Neste caso, uma acurácia de 99,3% foi atingida ao utilizar a Análise Discriminante Canônica (*Canonical Discriminant Analysis - CDA*) tanto para a seleção de atributos quanto como classificador. A segunda categoria consistiu em classificar quanto aos grupos de tambores e pratos (Bumbo, Caixa, Tom, Chimbal e Pratos em geral), sendo possível atingir um desempenho de 97,4% ao utilizar a técnica de *ReliefF* para a seleção das características e o K^* (CLEARY; TRIGG, 1995) como classificador. Por fim, a última categoria de problema visava classificar cada instrumento em suas sub-categorias: Bumbo, Caixa, Tom Grave, Tom Médio, Tom Agudo, Chimbal Aberto, Chimbal Fechado, Prato de Condução (*Ride*) e Prato de Ataque (*Crash*). Utilizando o CDA como seletor de atributos e o K^* como classificador, os autores conseguiram atingir um desempenho de 90,7% de classificação.

Em particular, estes tipos de modelos computacionais podem ajudar significativamente na análise dos fatores que compõem a acústica destes instrumentos musicais, tais como o material a partir do qual eles são fabricados. Uma abordagem que recupera informações sobre o material constituinte de um prato através do seu som pode ser interessante para as empresas manterem um certo padrão de produção na fundição de ligas de bronze, uma vez que uma análise da composição química do material é um procedimento relativamente caro. Especificamente, para pequenos fabricantes, que reciclam a própria sucata de fabricações anteriores e até mesmo sucata de outros fabricantes, esta questão torna-se ainda mais relevante, uma vez que o investimento financeiro para estimar o material obtido após o processo de fundição se torna mais acessível. Estudos semelhantes, tais como os de (FU et al., 2018; GUO; JIANG; GAO, 2022), utilizam o som de impacto para reconhecimento de materiais ou de propriedades. FU et al. (FU et al., 2018) utilizaram os sons de toque e deslizamento para classificar os materiais conforme as 9 categorias de madeira selecionadas. Para isso, os autores recorreram a diferentes descritores de alto nível

e a Máquina de Aprendizagem Extrema (*Extreme Learning Machine - ELM*). Já GUO; JIANG; GAO (GUO; JIANG; GAO, 2022) investigaram o uso do som de impacto para a predição do grau de compactação de solos. A correlação entre os níveis de compactação e os centroides de uma determinada sub-banda de seus espectros de frequência foram traçados para diferentes intensidades de impacto.

Inspirados nestes trabalhos, este estudo propõe a classificação de pratos de bateria conforme suas respectivas ligas de bronze via captação de seus sons, cujos objetivos são descritos na Seção 1.1.

1.1 OBJETIVOS

1.1.1 Objetivo Geral

O objetivo principal desta dissertação é propor um modelo computacional baseado em aprendizado de máquina para a classificação de pratos de bateria conforme três diferentes proporções teóricas de estanho presente nas ligas de bronze.

1.1.2 Objetivos Específicos

Como objetivos específicos necessários para atingir o objetivo geral, destacam-se:

1. Construir uma base de dados através da captação do som de pratos de bateria de diferentes ligas de bronze;
2. Explorar diferentes descritores de áudio e determinar o conjunto mais adequado a ser utilizado;
3. Avaliar diferentes abordagens de seleção de atributos;
4. Comparar o desempenho de três classificadores: Regressão Logística, Fuzzy com função de pertinência triangular e Fuzzy com função de pertinência Gaussiana.

1.2 RELEVÂNCIA E CONTRIBUIÇÕES

Como mencionado anteriormente, estudos que investigam a influência do material em um instrumento musical têm sido realizados, ainda que com pouca adesão de técnicas de aprendizado de máquina ou inteligência artificial como ferramentas. Ainda assim, trabalhos que utilizam tais ferramentas para o reconhecimento de instrumentos musicais ou até mesmo para identificar aspectos da acústica destes instrumentos através de seus sons, podem ser encontrados na literatura.

Por outro lado, existem pesquisas similares que focam no reconhecimento de materiais a partir da coleta dos sons de contato e que recorrem a modelos computacionais

para auxiliar na identificação ou classificação das amostras. Entretanto, no melhor conhecimento do autor deste trabalho, pesquisas que se baseiam no uso do som de pratos de bateria para classificá-los conforme suas ligas de bronze, utilizando técnicas de aprendizado de máquina, ainda não foram realizadas.

Sendo assim, este é um trabalho interdisciplinar que contribui para diversas áreas do conhecimento, dado que propõe um modelo computacional para solucionar um problema ainda não investigado e que pode ser útil nas áreas da engenharia mecânica ou metalúrgica, computação musical e até mesmo na área musical. O estudo de problemas deste gênero com enfoque na engenharia mecânica ou metalúrgica pode ajudar na área de controle de qualidade dos materiais e processos de fundição e até mesmo auxiliar na identificação de fraudes, isto é, quando um produto é vendido como sendo de um determinado material, mas que, na realidade, não é.

Quando o enfoque é dado para a área de computação musical, esta categoria de problema pode servir como motivação para a realização de estudos similares aplicados a diferentes instrumentos musicais ou até mesmo auxiliar, por exemplo, no ajuste dos modelos em problemas que focam na separação de fontes. Em se tratando do meio musical, este estudo pode servir como uma ferramenta que auxilie na determinação dos parâmetros de equalização, mixagem e demais operações realizadas pelo engenheiro de som de um estúdio de gravação.

1.2.1 Motivação Pessoal

A minha relação com a música vem desde cedo. Ainda quando criança comecei minhas aulas de bateria, me apaixonei pelo instrumento e desde então nunca mais parei de tocar. Sempre fui muito curioso e busquei acompanhar os músicos profissionais da cidade para conhecer cada vez mais um pouco desse instrumento maravilhoso. No entanto, costumo atribuir o início da minha aproximação com os pratos lá para meados de 2013, em uma apresentação do Dudu Lima Trio. Nessa ocasião, o baterista do grupo utilizou um prato de condução em que era possível vê-lo vibrando a cada batida. Aquilo me fascinou e foi a partir daí que eu comecei a querer conhecer e saber mais sobre pratos.

Logo que comecei a minha graduação em 2014, eu já pensava em estudar ou até mesmo realizar o meu TCC sobre pratos de bateria. Eu só precisava conhecer mais a fundo aquele instrumento que havia me “enfeitado”. Basicamente foi assim que em 2019 eu comecei a minha jornada acadêmica, tendo como tema central os pratos de bateria. E mesmo após ter defendido o meu TCC, percebi que a cada passo que eu caminhava, ainda tinha muito o que aprender e a oferecer para a comunidade científica, já que, por ser um tema um tanto quanto incomum, não é algo tão explorado cientificamente. Isso me motivou a dar continuidade a esses estudos em um mestrado e assim conseguir retribuir de uma maneira melhor para a comunidade musical e científica.

Espero que com esse trabalho eu consiga, de certa maneira, dividir um pouco da minha alegria e emoção em poder pesquisar sobre esse tema especial para mim.

1.3 ESTRUTURA DO TRABALHO

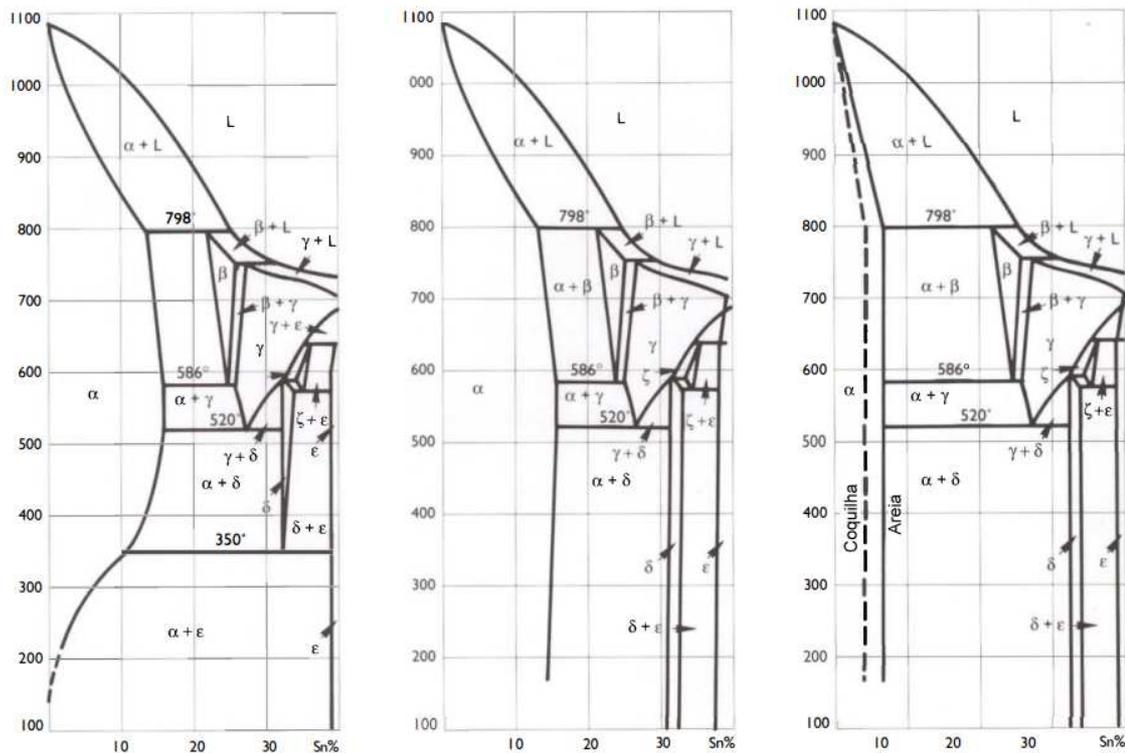
Este trabalho está dividido em cinco capítulos: Introdução, Fundamentação Teórica, Material e Métodos, Resultados e Discussão e, por fim, Conclusão. No Capítulo 1 é feita uma apresentação do tema deste trabalho, trazendo uma contextualização do problema abordado, a justificativa, as contribuições e os objetivos deste estudo. O Capítulo 2 é reservado para introduzir os principais conceitos envolvidos para o melhor entendimento deste estudo. Sendo assim, serão abordadas discussões a respeito de ligas de bronze, processos de fabricação, acústica, áudios digitais, descritores de áudio, correlação e análise de variância. Na sequência, o Capítulo 3 descreve o procedimento desenvolvido para a captação dos sons; os atributos extraídos dos sinais; as bases de dados construídas; os modelos de classificação utilizados; as técnicas de seleção de atributos e, por fim, os indicadores de desempenho utilizados no trabalho. Os resultados dos experimentos computacionais, bem como suas análises e discussões, são mostrados no Capítulo 4. Para finalizar, o Capítulo 5 conclui este trabalho e aborda as principais contribuições deste estudo, além de trazer sugestões para trabalhos futuros e as produções acadêmicas relacionadas.

2 FUNDAMENTAÇÃO TEÓRICA

O presente capítulo tem por objetivo apresentar os conceitos necessários para o melhor entendimento deste trabalho.

2.1 LIGAS DE BRONZE

Dentre os diversos fatores relacionados à acústica dos pratos de bateria, os materiais e o modo como são processados apresentam uma grande parcela do resultado acústico final destes instrumentos (WHITE; MACMILLAN, 2002). As propriedades mecânicas e acústicas das ligas de cobre tornam-nas preferíveis para este tipo de aplicação, dando destaque para os bronzes, que podem ser tratados basicamente como ligas metálicas de cobre (Cu) com a adição de estanho (Sn). No entanto, como pode ser constatado na Figura 4, diferentes ligas de bronze Cu-Sn podem ser obtidas, sendo que a variação na concentração de estanho acarreta alterações de propriedades (até mesmo microestrutural) do material (WHITE; MACMILLAN, 2002).



(a) Diagrama de equilíbrio. (b) [Diagrama para tratamentos térmicos. (c) Diagrama para fundição.

Figura 4 – Diagramas de fase cobre-estanho (Cu-Sn) na região rica em cobre. As linhas sólidas indicam o equilíbrio diagrama de fases. As letras gregas α , β , γ , δ , ζ e ϵ representam diferentes fases (compostos metálicos) sólidas. A letra L indica a fase líquida do material.

Fonte: Adaptado de (SCOTT, 1992).

Ao observar o diagrama de fases Cu-Sn em equilíbrio disposto na Figura 4(a), é possível perceber que o estanho é solúvel no cobre até cerca de 16% em peso, o que acontece na faixa de temperatura entre 520 °C e 586 °C. Na prática, diferentes versões do diagrama são construídas para representar a variação dos limites das soluções sólidas em condições de um tratamento térmico de recozimento (Figura 4(b)) ou normais de fundição (Figura 4(c)) (SCOTT, 1992). Ao restringir as possibilidades de materiais para as três ligas de interesse deste trabalho (B8, B10 e B20) - descritas na Tabela 1, tem-se que, após serem devidamente tratados, os bronzes B8 (com 8% em peso de estanho em sua composição química) e B10 (10% em peso de estanho) são formados unicamente pela fase α , isto é, são encontrados como sendo um material homogêneo. Por outro lado, os bronzes B20 utilizados em pratos de bateria são materiais heterogêneos que apresentam as fases α e β em sua microestrutura (WHITE; MACMILLAN, 2002).

A fase α presente nestes materiais é de estrutura cristalina Cúbica de Face Centrada (CFC), ao passo que o precipitado β apresenta estrutura cristalina Cúbica de Corpo Centrado (CCC) (WHITE; MACMILLAN, 2002; MORTAZAVI; OUDBASHI, 2010). Sendo assim, e sabendo que a estrutura CFC é, em geral, mais dúctil² que as demais estruturas cristalinas (GROOVER, 2010), os bronzes homogêneos apresentam uma capacidade de deformação bastante superior aos heterogêneos, que apresentam uma maior dificuldade de fabricação devido à sua maior fragilidade, estando, por isso, associados ao processo de fabricação artesanal detalhado na Subseção 2.2.1.

2.2 PROCESSOS DE FABRICAÇÃO

Delimitar a fronteira da classificação dos processos de fabricação de pratos de bateria é uma questão discutível. É comum atribuir o título de fabricação artesanal a qualquer método de fabricação que se utilize de martelamento manual e da etapa de torneamento, independente da liga de bronze, uma vez que são processos realizados manualmente e que dependem intrinsecamente da habilidade de operador para atingir o resultado esperado. No entanto, este trabalho traz uma abordagem diferente. A fabricação artesanal, apresentada na subseção 2.2.1, é considerada desde o processo de fundição do material, dado que, embora se tenha um certo rigor e padrão para a realização do processo, geralmente não há uma verificação da composição química do fundido. A fabricação não-artesanal — subseção 2.2.2 — é considerada como sendo aquela que compra o bronze de empresas especializadas e que utiliza maquinários para as etapas de conformação que diminuem a influência da habilidade do operador no resultado esperado. Por fim, as fabricações que compram a matéria-prima de empresas especializadas em sua produção e que se utilizam das etapas de martelamento manual e torneamento são classificadas, neste trabalho, como

² A ductilidade é a propriedade mecânica que mede o grau de deformação plástica (permanente) suportado pelo material até a ruptura.

sendo um método híbrido de fabricação, apresentado na subseção 2.2.3.

De toda forma, entender os processos de fabricação de um prato de bateria ajuda a compreender melhor como é dada a construção da acústica destes instrumentos. Cada modificação realizada no material bruto tem uma consequência acústica, podendo ser perceptível ou não. Dessa forma, ainda que dois pratos tenham sido identicamente fabricados, as variáveis que assumem um caráter aleatório durante a fabricação, podem causar uma variação acústica perceptível entre estes instrumentos (BORATTO et al., 2021).

2.2.1 Método Artesanal Clássico

Esta abordagem de fabricação está normalmente associada a ligas de bronze heterogêneas, com destaque para os bronzes B20, visto que são historicamente mais tradicionais para a utilização em pratos de bateria. Dessa forma, por preservar grande parte dos aspectos principais da fabricação tradicional turca utilizada há centenas de anos, o método artesanal clássico talvez seja o tipo de fabricação de pratos de bateria mais conhecido no mundo, podendo ser resumido seis etapas como mostrado pelo fluxograma da Figura 5.

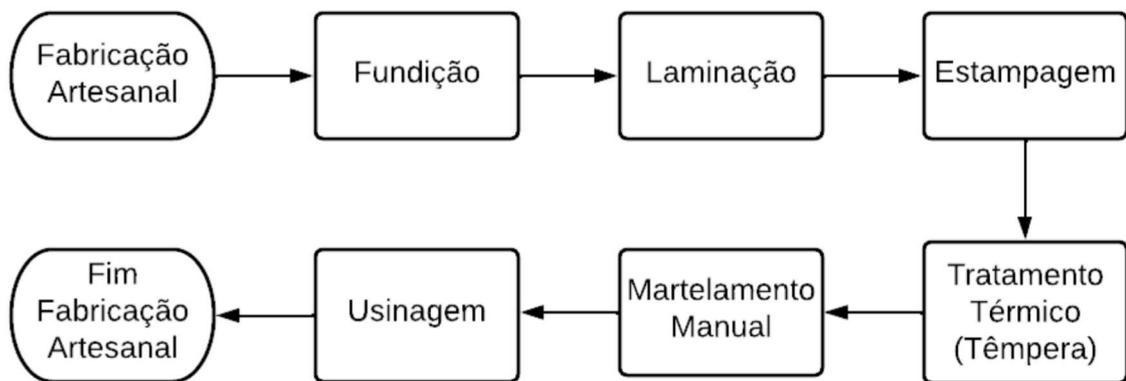


Figura 5 – Fluxograma da Fabricação Artesanal.

Fonte: Elaborado pelo Autor.

Por conta de esta abordagem trabalhar com materiais que não são comumente encontrados no mercado para outras aplicações, este tipo de fabricação inicia-se com a produção da liga de bronze em um processo conhecido por fundição por gravidade em molde permanente. Neste caso, este processo consiste em inicialmente derreter uma determinada quantidade de cobre e, posteriormente, após este metal atingir o estado líquido, adicionar uma quantidade de estanho na proporção desejada (em geral, em torno

de 20% do peso total dos metais). Por fim, a mistura líquida destes metais é vertida em um molde permanente chamado lingoteira que, após a solidificação, dará origem a um lingote, como mostrado na Figura 6.



Figura 6 – Metal líquido sendo vertido nas lingoteiras côncavas durante o processo de fundição.

Fonte: Acervo Domene Cymbals

Os lingotes obtidos por este processo apresentam um formato côncavo de seção transversal circular e dimensões na ordem de 35 mm de espessura e 100 mm de raio³. Dessa forma, estes produtos são submetidos a um processo de laminação a quente visando reduzir a espessura e aumentar a área superficial dos lingotes até que atinjam as dimensões aproximadas à que se deseja nos pratos em produção. Este processo de fabricação caracteriza-se por submeter o material a um esforço de compressão realizado por dois cilindros fixos que giram em sentidos opostos com a mesma velocidade superficial e separados por uma distância menor que a espessura do material (GROOVER, 2010; CHIAVERINI, 1986), como mostrado na Figura 7. Já o termo *laminação a quente* é utilizado para designar o processamento mecânico do material (laminação) em uma temperatura superior à temperatura de recristalização⁴ (temperatura na qual a recristalização termina em exatamente 1 hora). Sendo assim, esta etapa da produção é caracterizada pelo aquecimento do lingote até cerca de 600 °C antes de cada passe de laminação.

Após atingir a espessura e área superficial adequadas, esta chapa é novamente aquecida e então submetida ao processo de estampagem, cuja finalidade é introduzir o ressalto central observado no prato, popularmente conhecido como cúpula ou *bell*.

Até então, o material em suas condições normais apresenta uma estrutura muito quebradiça. O aumento da temperatura nos processos de conformação a quente leva o

³ Informação fornecida por Francisco Domene (Domene Cymbals) durante conversa informal.

⁴ A recristalização é o processo de formação de um novo conjunto de grão livres de deformação e que são equiaxiais.

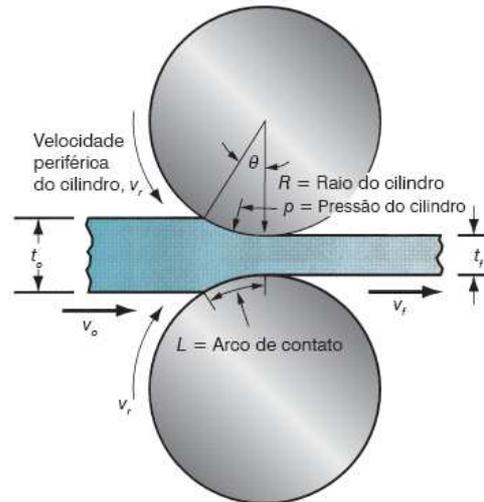


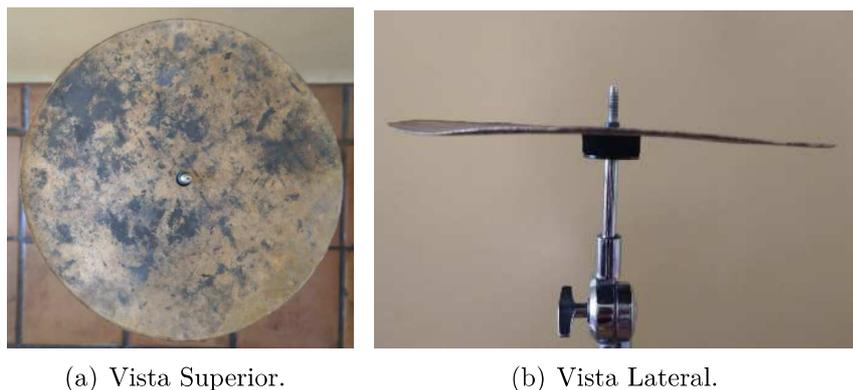
Figura 7 – Vista lateral da laminação de planos, indicando as espessuras inicial (t_0) e final (t_f), velocidades de entrada (v_0) e de saída (v_f) da peça, o ângulo de contato com os cilindros (θ) e outras características.

Fonte: (GROOVER, 2010)

material para uma composição de fases α e β . Entretanto, a lenta taxa de resfriamento ao ambiente decompõe essas fases em α e δ , sendo esta última a responsável pela fragilização do material (MORTAZAVI; OUDBASHI, 2010). Para minimizar esse efeito, realiza-se um tratamento térmico de têmpera cujo objetivo é evitar a formação da fase δ , preservando a estrutura do material na fase aquecida $\alpha + \beta$ e deixando o precipitado β em formato de agulhas (SCOTT, 1992; MORTAZAVI; OUDBASHI, 2010). Este tratamento é realizado aquecendo o material até cerca de $600\text{ }^\circ\text{C}$ e, em seguida, resfriando-o bruscamente em água à temperatura ambiente.

Ao final destas etapas, tem-se uma chapa com a superfície bastante deformada, exceto pela cúpula estampada em sua região central (caso tenha), e com um perímetro irregular. Portanto, é necessário realizar um recorte ao longo de seu contorno, com a finalidade de se prover um formato circular com diâmetro ligeiramente superior ao do prato desejado. Dessa forma, chega-se a um subproduto da fabricação conhecido como *blank* — Figura 8 — que é o material bruto que pode dar origem a qualquer prato com as determinadas especificações de espessura, diâmetro e tamanho de cúpula.

A partir desta etapa, o material todo é trabalhado a frio e em temperatura ambiente. O primeiro processo ao qual o *blank* é submetido é o de conformação por forjamento livre manual, também conhecido simplesmente por martelamento manual. O objetivo principal desta operação é atingir o formato côncavo do prato através de uma imensa quantidade de impactos do martelo sobre as superfícies superior e inferior da chapa. Além disso, como mencionado anteriormente, toda a superfície do material bruto é irregular



(a) Vista Superior.

(b) Vista Lateral.

Figura 8 – Ilustração em vista superior e lateral de um *blank* de Bronze B20 com 10 polegadas de diâmetro e sem cúpula.

Fonte: Acervo do Autor.

(não plana), como pode ser visto pela Figura 8(b). Portanto, para se atingir o equilíbrio na distribuição de tensões do prato e, conseqüentemente, o resultado geométrico final desejado, determinadas localidades da superfície necessitarão de uma quantidade maior ou menor de martelamento, provocando uma certa aleatoriedade na distribuição das marcas do martelo pela extensão da área do material. Por fim, já com seu formato, o prato é levado até um torno adaptado para esta aplicação, onde será realizada a etapa da usinagem. Primeiramente, o operador faz o ajuste mais preciso do diâmetro externo, através de um procedimento de sangria próximo à borda, para então realizar as operações de desbaste e acabamento superficial. A Figura 9 ilustra as operações de martelamento e usinagem.

2.2.2 Método Não-Artesanal Brasileiro

O método de fabricação não-artesanal está associado a ligas de bronze homogêneas, dado que estes materiais facilitam os processos de conformação, conseguindo um aumento da produtividade, redução de custo de fabricação e maior padronização de um mesmo instrumento, como será visto a seguir. O fluxograma apresentado na Figura 10 ilustra resumidamente as etapas utilizadas por esse método em empresas brasileiras. Embora aparente ser um processo mais longo, as etapas são mais rápidas, o que possibilita a maior produtividade deste método de fabricação.

Por possuírem outras finalidades na indústria, algumas das ligas de bronze homogêneas são comercializadas por empresas especializadas, como é o caso das ligas C52100 e C52400, popularmente conhecidas entre os bateristas como B8 e B10, respectivamente. Sendo assim, a primeira etapa deste método de fabricação consiste na compra de chapas ou bobinas de bronze já com a espessura determinada, conforme a especificação dos pratos que se desejam produzir. Para o caso de bobinas, uma etapa para retificá-las deve ser realizada visando transformá-las em chapas. Na sequência, cortes circulares com



(a) Martelamento manual.

(b) Usinagem.

Figura 9 – Ilustração dos processos de martelamento manual e usinagem.

Fonte: Acervo de Domene Cymbals

diâmetros aproximados aos pratos em fabricação são realizados dando origem ao *blank* que, diferentemente do método anterior, apresenta superfície plana e um perímetro bastante regular, conforme pode ser constatado na Figura 11. Para isso, os meios utilizados neste processo podem ser diversos. Duas alternativas são descritas, a saber:

1. Inicialmente, a chapa é recortada em porções quadradas cuja dimensão das arestas é ligeiramente superior ao do prato em fabricação. Furos centrais são realizados nestes cortes e, por fim, é feita a apara de suas arestas visando atingir o formato circular.
2. As chapas são submetidas a um processo de corte a laser⁵. Esta alternativa consiste apenas desta etapa, cabendo ao operador somente definir o tamanho diametral e a disposição dos cortes circulares para otimizar a utilização da chapa.

A partir do *blank*, o formato do prato, em sua totalidade, é garantido por meio da aplicação de um único processo de fabricação, ilustrado na Figura 12: o repuxo. De acordo com (GROOVER, 2010), este processo consiste em conformar gradualmente uma peça axissimétrica a partir da aplicação de uma força de compressão localizada, utilizando-se uma ferramenta arredondada. Para tanto, o *blank*, ou a chapa plana metálica em forma de disco, é adaptado em um eixo rotativo entre o mandril e o grampo. Na sequência, a

⁵ Informação fornecida em conversa particular com Bruno Schell, sócio-proprietário da empresa Redfoot Cymbals.

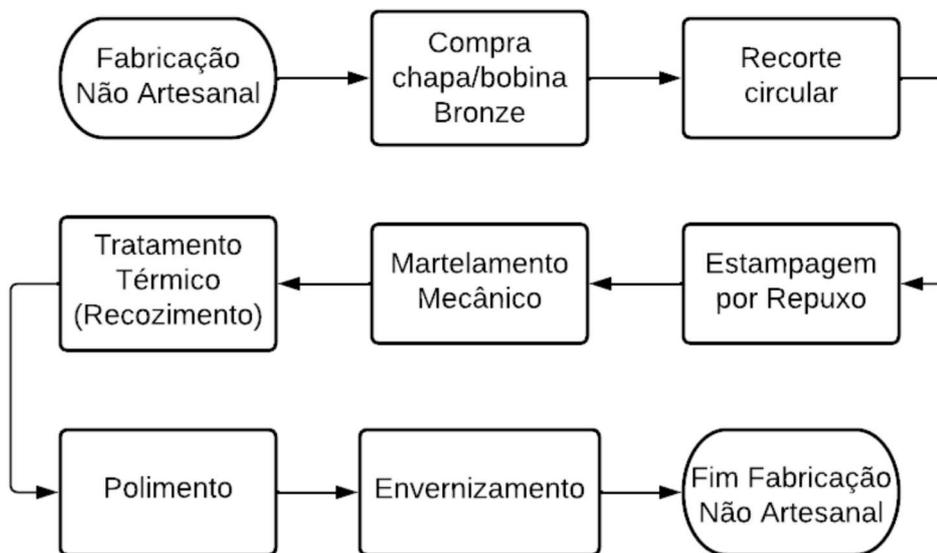


Figura 10 – Fluxograma da Fabricação Não Artesanal Brasileira.

ferramenta arredondada (ou rolete) pressiona esta chapa contra o molde e, dessa forma, a chapa é deformada gradualmente e assume a geometria do molde que, neste caso, apresenta a forma final do prato desejado.

Posteriormente, os pratos, já com seus formatos, são submetidos ao processo de martelamento que, diferente do método de fabricação artesanal, é realizado por máquinas. Esta etapa, no entanto, possui como objetivo principal trabalhar a acústica dos instrumentos, conferindo-lhes deformações distribuídas de maneira padronizada (não aleatória) sobre a superfície do prato e com formas específicas. Na sequência, é realizado um tratamento térmico de recozimento para alívio das tensões internas do material, cuja temperatura de operação pode variar conforme o modelo do prato e o fabricante. Por fim, as etapas de polimento e aplicação de verniz podem ser tratadas como acabamentos. O ato de polir o prato consegue eliminar eventuais riscos presentes na chapa ou bobina de bronze adquirida. No entanto, muitas das vezes essa etapa é executada com o objetivo de se produzir um efeito acústico e estético, favorecendo a propagação de frequências mais altas e deixando o prato com um aspecto visual mais brilhante. Por outro lado, a aplicação do verniz pode ser vista como uma maneira de evitar a oxidação do material, embora também produza um efeito acústico nestes instrumentos.

2.2.3 Outros Métodos

O surgimento de um precipitado nas ligas heterogêneas de bronze provoca uma maior fragilidade do material e, conseqüentemente, dificuldade de processamento. No entanto, a fabricação artesanal apresentada consegue lidar com estas limitações, de maneira

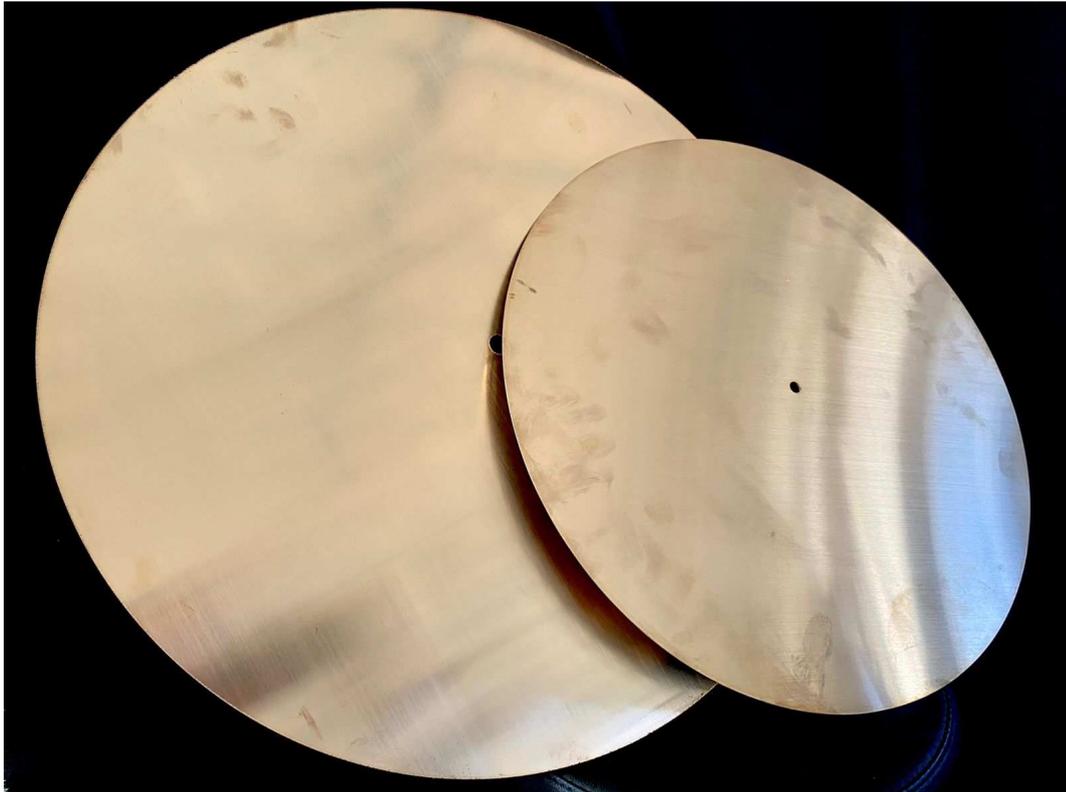


Figura 11 – Ilustração de *blanks* de Bronze B8.

Fonte: Imagem cedida por Bruno Schell.

que é possível associar este tipo de material à tal método de fabricação no Brasil. Por outro lado, as ligas homogêneas de bronze contam com uma boa capacidade de serem processadas. Portanto, dois caminhos podem surgir para a fabricação: utilizar-se do método não artesanal já apresentado ou combinar os métodos e buscar uma fabricação híbrida.

Neste caso, a fabricação híbrida consiste em valer-se das etapas iniciais do método não artesanal e as etapas finais do método artesanal. Desta forma, esta abordagem de produção acarreta um aumento da produtividade da fabricação, visto que otimiza os processos iniciais até a obtenção do *blank*, ao mesmo tempo que consegue um melhor trabalho acústico do material, proveniente das etapas finais do método artesanal. A Figura 13 mostra um fluxograma esquemático de uma abordagem híbrida de fabricação.

Em um cenário internacional, existem algumas outras formas de se fabricar pratos de bateria. No entanto, a maioria são variações dos métodos artesanal, não artesanal e até mesmo do método híbrido. Como exceção, existe um método para fabricação que se baseia em um processo denominado *Rotocasting*. Neste processo, o bronze líquido fundido é vertido em um molde rotativo (com velocidade de rotação próxima da 1000 rpm) e, dessa forma, o processo de solidificação é dada de forma diferente (PINKSTERBOER;

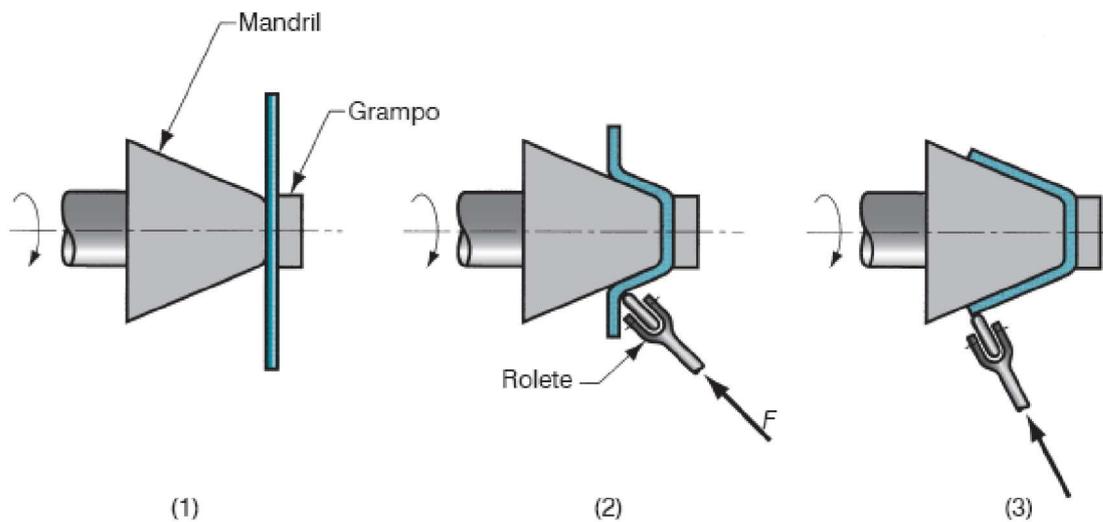


Figura 12 – Ilustração do processo de repuxo convencional: (1) montagem inicial; (2) durante a operação; e (3) final do processo.

Fonte: (GROOVER, 2010)

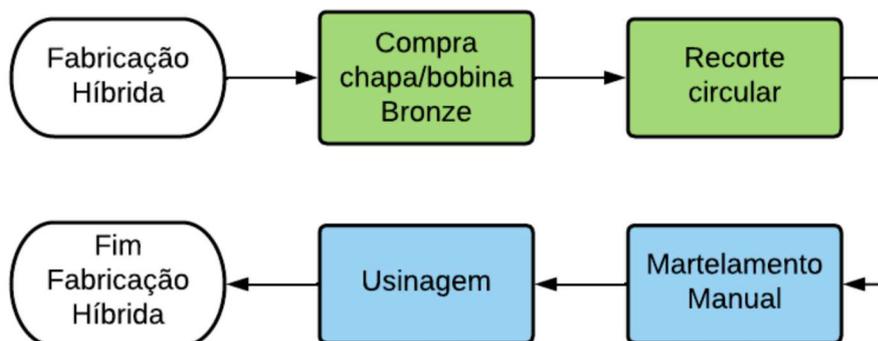


Figura 13 – Fluxograma da Fabricação Híbrida. Em verde estão representados os processos provenientes da fabricação não artesanal e em azul, os provenientes da fabricação artesanal clássica.

MATTINGLY, 1992).

2.3 ACÚSTICA

Como já mencionado anteriormente, cada aspecto do prato implica em uma consequência acústica. Conhecer estes fatores, bem como as suas influências no som do prato, é de suma importância para se projetar um novo instrumento e até mesmo para realizar alguma alteração acústica no mesmo pois, dessa forma, garante-se que o resultado acústico atingido seja próximo ao desejado.

É possível agrupar e elencar alguns destes fatores visando obter uma abordagem mais geral e didática. No entanto, deve-se ter em mente que diferentes fatores podem estar diretamente relacionados. Portanto, esta abordagem pode se tornar redundantes em alguns aspectos. Além disso, vale ressaltar que o efeito acústico de um fator pode compensar o de outro.

1. Aspectos Químicos

- a) Composição Química do Material: É o aspecto em estudo neste trabalho. A princípio, sabe-se que a velocidade de propagação (V) de uma onda em um meio sólido está relacionada ao módulo de elasticidade (E) e à densidade (ρ) do material ($V = \sqrt{\frac{E}{\rho}}$). Sendo assim, as diferentes combinações dos elementos químicos que compõem um material provocam alterações em suas propriedades mecânicas e, conseqüentemente, modificações nas características acústicas do instrumento.
- b) Camada de Oxidação: Está normalmente associada aos pratos advindos da fabricação artesanal. Esta camada de oxidação, proveniente do próprio processo fabril, dificulta a vibração do prato, acarretando um som mais “compacto” e deixando os toques de baquetas mais definidos.
- c) Presença de Verniz: A aplicação de verniz, embora tenha o objetivo de evitar oxidações, também proporciona um ligeiro “abafamento” do som dos pratos, devido à adição desta película de material sobre toda a superfície destes instrumentos.

2. Aspectos Físicos

- a) Peso: Este fator é resultado da combinação de seu material, diâmetro e espessura. Assim, de maneira geral, o peso de um prato está associado à quantidade de material que deve ser colocada para vibrar. Dessa forma, uma maior quantidade de massa provoca um aumento na sustentação e projeção do som, reposta mais lenta ao toque, além de deixar o instrumento com uma tendência de som mais agudo (PINKSTERBOER; MATTINGLY, 1992).
- b) Marcas de Martelo: O processo de martelamento cria regiões de compressão localizada (marcas do martelo). Estes pontos de impacto tornam-se mais densos, rígidos e resilientes⁶. O aumento da tensão no material provocada por esse processo acarreta uma maior sustentação do som, tendência de som mais agudo e reforça a gama de frequências emitidas pelo instrumento (PINKSTERBOER; MATTINGLY, 1992).
- c) Tensionamento: O tensionamento da superfície do material tem relação com a concavidade ou curvatura do prato, que será abordado na sequência. Por

⁶ Refere-se à habilidade de um material absorver energia no regime elástico.

outro lado, o tensionamento da borda do prato também apresenta-se como um fator extremamente importante na acústica destes instrumentos. Bordas mais tensionadas vibram mais rapidamente e, portanto, com menor amplitude, ao passo que bordas com menor tensionamento produzem uma vibração mais lenta e com maior amplitude (SCHROEDER et al., 2007). Sendo assim, por serem instrumento idiofônicos, vibrações mais rápidas e de menor amplitude estão associadas a sons mais agudos (ondas de maior frequência), ao contrário dos instrumentos com bordas menos tensionadas que tendem a produzir sons mais graves, isto é, ondas de menor frequência (BORATTO et al., 2021).

- d) Acabamento Superficial: Está relacionado à dimensão dos sulcos tonais da(s) superfície(s) dos pratos, gerados durante o processo de usinagem (fabricação artesanal) ou processo de repuxo (fabricação não artesanal). Acabamentos mais refinados, isto é, sulcos mais curtos e rasos ou até mesmo superfícies polidas, favorecem a propagação de frequências mais altas, tornando o som mais agudo (PINKSTERBOER; MATTINGLY, 1992; BORATTO et al., 2021). Desta maneira, existe uma extensa possibilidade de trabalho acústico na etapa de usinagem, dado que é possível combinar diferentes acabamentos superficiais em um mesmo instrumento, ou seja, combinar diferentes sulcos tonais e até mesmo não realizar o processo de desbaste⁷ em alguma região do prato (BORATTO et al., 2021).

3. Aspectos Geométricos

- a) Tamanho da Cúpula: Tem relação com a emissão de harmônicos e sensação de volume sonoro. Quanto maior a cúpula, mais intensos ficam estes aspectos acústicos (PINKSTERBOER; MATTINGLY, 1992).
- b) Diâmetro: Além das influências relacionadas ao peso do prato, como mencionado anteriormente, uma maior área de superfície favorece a propagação de frequências mais baixas⁸, além de conseguir movimentar uma coluna de ar maior. Portanto, conseguir mais projeção de volume e sustentação do som (PINKSTERBOER; MATTINGLY, 1992).
- c) Espessura: Além de ter influência no peso do prato, este fator necessita de uma atenção especial quando se refere à existência de um gradiente de espessura ao longo do raio. Esta variação gradual é denominada *taper* e sua função é facilitar a movimentação inicial do prato, já que a borda passa a conter uma quantidade inferior de material com relação à região central. Em termos acústicos, a presença do *taper* provoca uma resposta mais rápida ao toque, menor sustentação

⁷ Termo que designa a retirada de material durante o processo de usinagem.

⁸ A frequência tem relação inversa com o comprimento de onda sonora. Quanto maior a frequência, menor o comprimento da onda.

do som e tende a deixar o som do instrumento mais grave (PINKSTERBOER; MATTINGLY, 1992), isto é, favorecer a emissão de frequências mais baixas.

- d) Curvatura do Prato: Este aspecto geométrico tem relação com a tensão interna da superfície do prato. O aumento deste tensionamento provoca o aumento do arco de curvatura do instrumento. Sendo assim, pratos mais côncavos vibram mais rapidamente. Portanto, estão associados a sons mais agudos (PINKSTERBOER; MATTINGLY, 1992; BORATTO et al., 2021), como pode ser observado na Figura 14.



Figura 14 – Influência do arco de curvatura no som de um prato.

Fonte: (PAISTE, 2022)

2.4 SINAIS DE ÁUDIO

Antes de introduzir o conceito e as características de sinais de áudios, é necessário compreender o que é o som. Basicamente, o *som* pode ser entendido como sendo resultado de uma perturbação do meio. Um instrumento musical quando tocado, uma voz ou qualquer outro tipo de fonte sonora provoca uma oscilação da pressão em regiões de compressão e de rarefação, caracterizando-se, portanto, como uma onda mecânica longitudinal⁹ (MÜLLER, 2015). No caso de pratos de bateria, por serem instrumentos idiofônicos, a própria movimentação do corpo é quem causa o deslocamento de ar responsável por criar estas oscilações de pressão.

Quando esta oscilação da pressão é captada por um microfone, surge o conceito *sinais de áudio* (MÜLLER, 2015). O áudio é a captação do som por um microfone, que atua como um transdutor que converte a energia mecânica do som em um sinal elétrico. Sendo assim, tal conceito pode ser entendido como uma forma de representar sons eletricamente. A Figura 15 ilustra todo esse processo de transformação de um som em sinal de áudio.

⁹ Ondas longitudinais são aquelas que possuem a mesma direção de vibração e propagação.

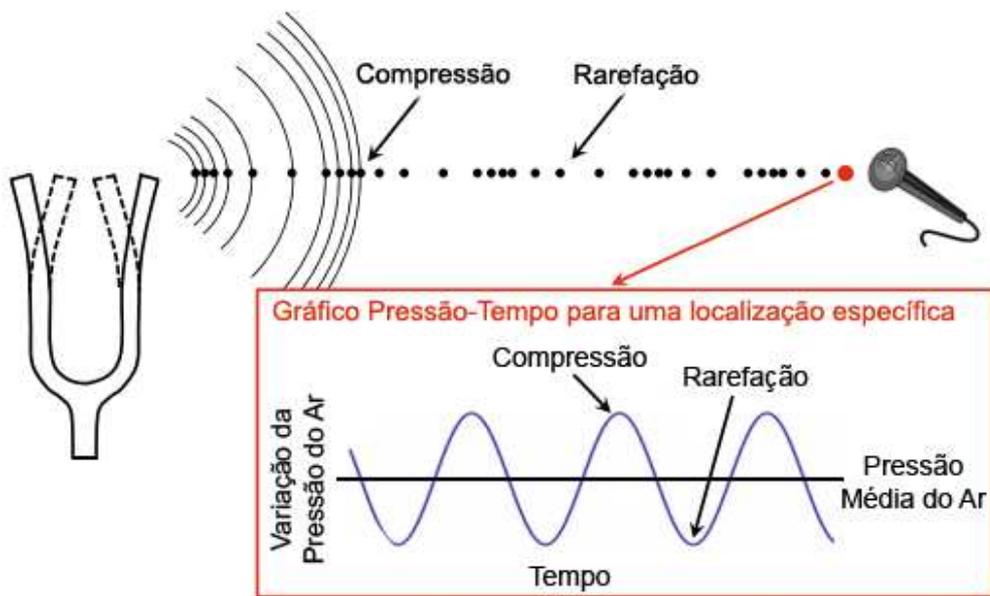


Figura 15 – Ilustração da transformação do som em sinal de áudio. A perturbação do meio causada por um instrumento musical gera ondas mecânicas longitudinais (som). Estas perturbações são registradas por um microfone e convertidas em ondas digitais transversais (sinal de áudio).

Fonte: (MÜLLER, 2015)

Ao representar um som em um domínio digital, o sinal deixa de ser contínuo e passa a ser discreto, de forma que a resolução desta representação seja determinada por dois parâmetros: taxa ou frequência de amostragem (*sampling rate*) e o número ou profundidade de bits (*bit depth*). A taxa de amostragem pode ser entendida como sendo a quantidade registros de informação (amostras) que é feita a cada segundo. Para evitar possíveis problemas e a perda de mais informação do som, o valor da frequência de amostragem não pode ser escolhido arbitrariamente, de maneira que para a sua determinação é necessário utilizar como base o Teorema da Amostragem (Teorema 1).

Teorema 1 (Teorema da Amostragem). (OPPENHEIM; SCHAFER, 2009; DINIZ; SILVA; NETTO, 2010) *Se um sinal contínuo no tempo $x_c(t)$ for banda limitada, isto é, a sua Transformada de Fourier é tal que $X_c(j\omega) = 0$, para $|\Omega| > |\Omega_N|$, então $x_c(t)$ pode ser completamente recuperado do sinal de tempo discreto $x(n) = x_c(nT)$, com $n = 0, \pm 1, \pm 2, \dots$, se a frequência de amostragem Ω_s satisfizer:*

$$\Omega_s = \frac{2\pi}{T} \geq 2\Omega_N$$

A Frequência Ω_N é conhecida por Frequência de Nyquist e o termo $2\Omega_N$ é refere-se à Taxa de Nyquist-Shannon.

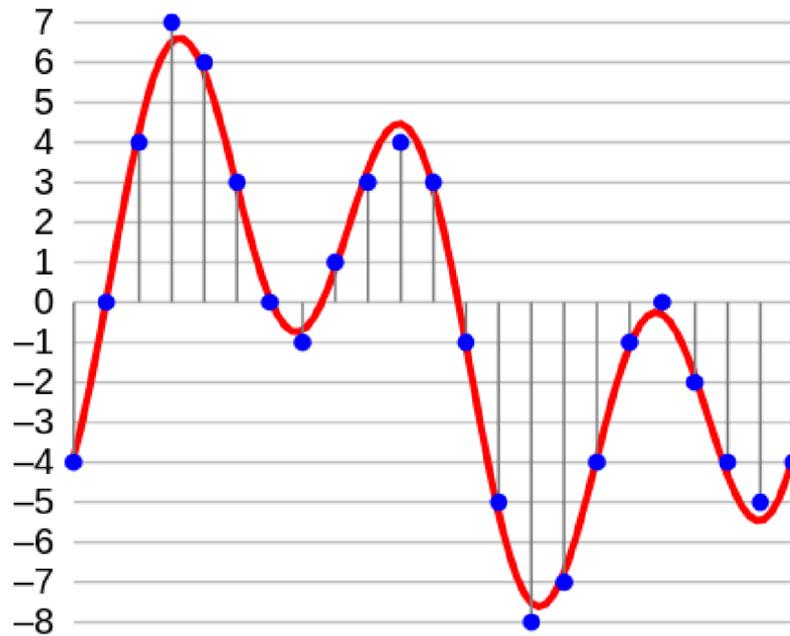


Figura 16 – Ilustração da resolução de um sinal segundo o número de bits de sua representação. Em vermelho, tem-se o sinal analógico e em azul a sua representação digitalizada com 4 bits de profundidade (16 níveis de amplitude). Cada ponto azul indica o valor da amplitude assumido pelo sinal digital nos instantes do tempo.

Fonte: (KEELEY, 2019)

A Frequência de *Nyquist* é normalmente tratada como a frequência máxima representada por determinada taxa de amostragem (Ω_s). Sendo assim, é comum que no contexto musical sejam utilizadas frequências de amostragem (Ω_s) acima de 44,1 kHz para os sinais de áudio, dado que a percepção auditiva humana é limitada superiormente em aproximadamente 22 kHz (Ω_N).

Por outro lado, a profundidade de bits está relacionada à resolução vertical do sinal, determinando os possíveis valores de amplitude que podem ser assumidos e, portanto, influenciando a relação sinal-ruído (*Signal to Noise Ratio* - *SNR*), dado que a representação de um som no domínio digital é feita de maneira discreta, isto é, não contínua. Sendo assim, quanto maior a profundidade de bits, maior a resolução do sinal de áudio. Em termos quantitativos, uma quantidade N de bits é capaz de representar 2^N valores. Portanto, um sinal com 16 bits de profundidade ($N = 16$) é capaz de representar $2^{16} = 65536$ valores e, da mesma maneira, um sinal com 24 bits ($N = 24$) consegue representar $2^{24} = 16777216$ valores. A Figura 16 ilustra este processo, mostrando a comparação entre um sinal analógico ilustrativo e a sua representação digitalizada com 4 bits de profundidade, totalizando 16 níveis de amplitude, representados em bit como: 0000, 0001, 0010, 0100, 0011, 0101, 0110, 0111, 1000, 1001, 1010, 1100, 1011, 1101, 1110, 1111.

Então, de maneira resumida, a taxa de amostragem está relacionada à resolução

horizontal do sinal, enquanto o número de bits está relacionado à resolução vertical.

2.5 CORRELAÇÃO

Quando se deseja medir a relação entre duas variáveis, isto é, determinar o grau de associação entre elas, geralmente utiliza-se uma métrica simples conhecida como coeficiente de correlação. No entanto, é importante ressaltar que a correlação entre duas variáveis não necessariamente implica em uma relação causal em que a mudança de uma variável seja uma consequência direta e correspondente da mudança da outra (BOSLAUGH, 2008).

Para a determinação da relação entre as variáveis é possível utilizar diversos métodos, tais como o coeficiente de correlação de Pearson, o coeficiente de ordem de classificação de Spearman, o coeficiente de correlação ponto-biserial e o coeficiente de correlação phi. Neste trabalho, optou-se por utilizar o método de Correlação de Pearson, por ser o mais utilizado para variáveis numéricas. Portanto, esta subseção restringirá o detalhamento unicamente para esta abordagem de determinação de correlação. Os demais métodos anteriormente mencionados podem ser encontrados em (BOSLAUGH, 2008).

O método de Pearson utiliza a fórmula apresentada pela Equação 2.1 para determinar a correlação entre duas variáveis e assume que os valores para seus coeficientes estão dispostos no intervalo $[-1, 1]$. Sendo assim, $Cor(X, Y) = -1$ indica uma correlação total negativa, $Cor(X, Y) = 0$ significa não haver correlação entre as variáveis e $Cor(X, Y) = 1$ designa uma correlação total positiva.

$$Cor(X, Y) = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2} \sqrt{\sum_{i=1}^n (y_i - \bar{y})^2}} \quad (2.1)$$

onde x_i e y_i são os valores medidos para as variáveis X e Y, respectivamente, e \bar{x} e \bar{y} são as médias aritméticas de ambas as variáveis, ou seja, $\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i$ e $\bar{y} = \frac{1}{n} \sum_{i=1}^n y_i$.

Conforme ilustrado na Figura 17, uma correlação positiva é dita quando o comportamento crescente de uma variável é também observado em outra, de forma que quanto maior o limiar da correlação entre elas, mais forte é essa relação de proporcionalidade direta entre tais variáveis. Uma correlação negativa, por outro lado, significa que a relação entre as variáveis é inversamente proporcional, isto é: enquanto uma variável aumenta, a outra decresce. Adicionalmente, o grau de associação negativa fica mais intenso quando o limiar se aproxima de -1 (NETTLETON, 2014).

A interpretação para os valores destes coeficientes auxilia no entendimento e nas análises realizadas. Diferentes interpretações podem ser encontradas na literatura como a de (SCHOBER; BOER; SCHWARTE, 2018), apresentada pela Tabela 2.

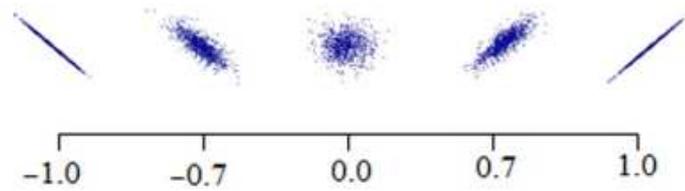


Figura 17 – Comportamento da correlação entre duas variáveis numéricas.

Fonte: (NETTLETON, 2014)

Magnitude Absoluta do Coeficiente de Correlação	Interpretação
0,00 - 0,10	Correlação Negligenciável
0,10 - 0,39	Correlação Fraca
0,40 - 0,69	Correlação Moderada
0,70 - 0,89	Correlação Forte
0,90 - 1,00	Correlação Muito Forte

Tabela 2 – Abordagem convencional para a interpretação dos coeficientes de correlação.

Fonte: (SCHOBER; BOER; SCHWARTE, 2018)

2.6 ANÁLISE DE VARIÂNCIA (ANOVA)

A Análise de Variância (do inglês *ANalysis Of VAriance* — ANOVA) é uma ferramenta estatística utilizada para detectar diferenças significativas entre as médias de vários grupos de distribuição (SAWYER, 2009). O teste ANOVA consiste da comparação quantitativa entre um aspecto interno e um externo dos grupos de distribuições em análises. O primeiro deles é referente à dispersão das observações de cada grupo, enquanto o aspecto externo consiste em avaliar a variação entre as médias dos grupos de distribuições. No entanto, existem alguns pressupostos que devem ser atendidos para que a análise de variância seja corretamente utilizada (OSTERTAGOVA; OSTERTAG, 2013):

1. As amostras devem ser independentes uma das outras.
2. As distribuições dos grupos devem seguir um padrão Gaussiano.
3. Homoscedasticidade ou Homogeneidade da Variância, isto é, a variância em cada grupo de distribuição é a mesma.

Sendo assim, o primeiro passo a ser realizado é tomar a média geral (\bar{y}_{total}) dos dados — Figura 18(a) — conforme a Equação 2.2. Na sequência, os grupos são separados e as médias para cada um deles (\bar{y}_j) são calculadas — Figura 18(b) — conforme a Equação 2.3. Com a média global e de cada grupo devidamente calculadas, é possível mensurar a distância entre a média de cada conjunto com relação à média geral ($\bar{y}_j - \bar{y}_{total}$) e a distância

entre a média de cada agrupamento (subíndice j) com relação aos seus respectivos valores internos ($y_{ij} - \bar{y}_j$) — Figura 18(c). Com isso, calcula-se a variância entre os conjuntos de distribuições ($\hat{\sigma}_{externa}^2$) e a variância interna dos conjuntos ($\hat{\sigma}_{interna}^2$), segundo as Equações 2.4 e 2.5, respectivamente (KIM, 2017).

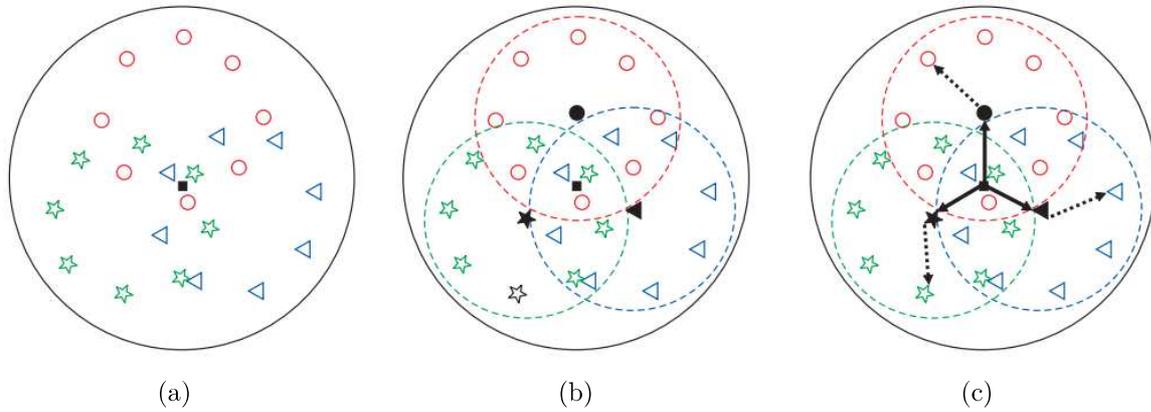


Figura 18 – Representação das etapas do teste ANOVA. (a) O quadrado sólido é sugerido como o valor da média dos dados globais. (b) Os dados são divididos em três grupos, que serão representados pelas suas médias (figuras sólidas). (c) A divisão dos grupos é avaliada a partir da comparação entre as distâncias da média de cada grupo com relação à média global (setas sólidas) e as distâncias das médias dos grupos com relação aos seus dados internos (setas pontilhadas). A primeira métrica de distância, ilustrada pelas setas sólidas, representa a variância inter-grupo. Já a segunda, ilustrada pelas setas pontilhadas, representa a variância intra-grupo.

Fonte: (KIM, 2017)

$$\bar{y}_{total} = \frac{\sum_j^k \sum_i^n y_{ij}}{nk} \quad (2.2)$$

$$\bar{y}_j = \frac{\sum_i^n y_i}{n} \quad (2.3)$$

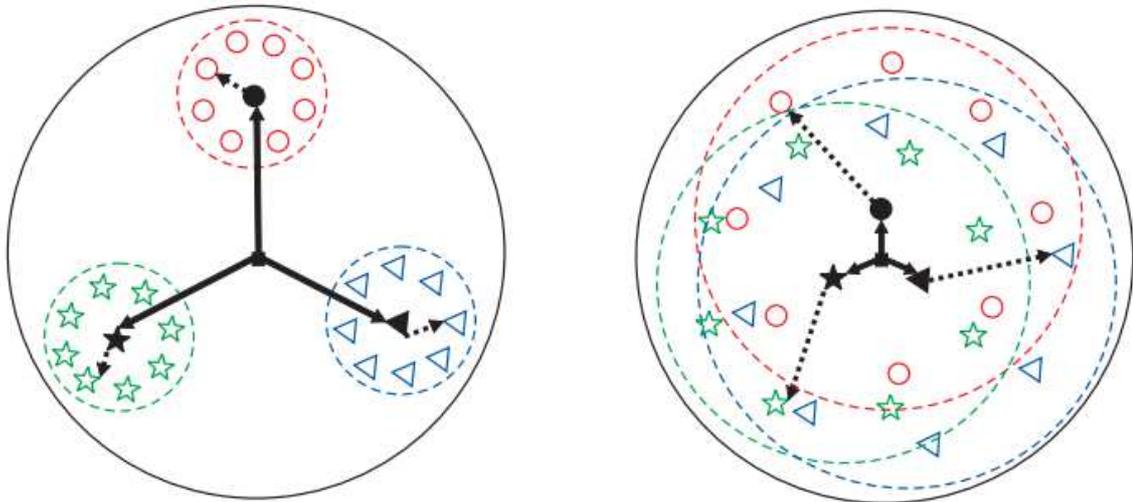
$$\hat{\sigma}_{externa}^2 = \frac{n \sum_j^k (\bar{y}_j - \bar{y}_{total})^2}{k - 1} \quad (2.4)$$

$$\hat{\sigma}_{interna}^2 = \frac{\sum_j^k \sum_i^n (y_{ij} - \bar{y}_j)^2}{N - k} \quad (2.5)$$

onde N representa a quantidade total de amostras e k a quantidade de agrupamento de dados.

Uma diferença significativa entre os grupos de distribuições é constatada quando existir uma grande variação entre as médias dos grupos em comparação com a variação das observações de cada conjunto - Figura 19(a). Por outro lado, caso a variação do

aspecto interno seja superior às do aspecto externo, pode-se concluir que as diferenças entre as médias dos conjuntos de distribuições são decorrência do acaso, portanto, as médias provavelmente não são significativamente diferentes - Figura 19(b) (KIM, 2017).



(a) A variância entre os grupos é consideravelmente maior que as variâncias internas.

(b) A variância entre os grupos é consideravelmente menor que as variâncias internas.

Figura 19 – Relação entre variância intra-grupo e inter-grupos. As setas pontilhadas representam as variâncias intra-grupos e as setas sólidas, as variâncias inter-grupos.

Fonte: (KIM, 2017)

Para quantificar a relação entre os aspectos interno e externo e, conseqüentemente, obter a significância das diferenças entre as médias das distribuições, um teste de hipótese usando a distribuição F é aplicado. Portanto, supondo que a hipótese nula seja verdadeira ($H_0 : \bar{y}_1 = \bar{y}_2 = \dots = \bar{y}_k$), isto é, supondo que não exista uma diferença estatisticamente significativa entre as médias das distribuições, então o Fator F, definido pela Equação 2.6 segue uma distribuição F com (k-1) e (N-k) graus de liberdade (GILLARD, 2020).

$$F = \frac{\hat{\sigma}_{externa}^2}{\hat{\sigma}_{interna}^2} \quad (2.6)$$

Para finalizar a análise, é necessário estabelecer um grau de significância (α) para o Fator-F. De maneira geral, trabalha-se com um nível de significância de 5% para a hipótese nula ser rejeitada. Portanto, se o Fator F excede o percentil de 95%, a hipótese nula é rejeitada. Generalizando esta análise para um nível de significância α qualquer, o intervalo (I) de valores para os quais o Fator F não rejeita a hipótese nula é dada pela Equação 2.7 (GILLARD, 2020). Na prática, é comum que essa análise de significância seja baseada em um parâmetro conhecido por *p-value* ou valor-p, que indica a probabilidade de rejeição da hipótese nela. Neste caso, quando $p < \alpha$ pode-se dizer que a hipótese nula é rejeitada com uma probabilidade maior que $(1 - \alpha) \cdot 100\%$ (OSTERTAGOVA; OSTERTAG, 2013).

$$I = \{F : 0 \leq F \leq F_{1-\alpha}(k-1, N-k)\} \quad (2.7)$$

3 MATERIAL E MÉTODOS

O fluxograma apresentado na Figura 20 representa as macroetapas realizadas neste trabalho. A primeira etapa consistiu da aquisição dos áudios, o que envolve a seleção dos pratos e equipamentos, o procedimento desenvolvido para a captação de seus sons e o pós-processamento realizado nos arquivos obtidos no processo de gravação. Em seguida, atributos temporais e coeficientes mel-cepstrais de frequência foram extraídos dos sinais, possibilitando dar origem a cinco bases de dados. Três técnicas de seleção de atributos foram utilizadas para avaliar os conjuntos de características anteriormente definidos. Adicionalmente, três classificadores também foram avaliados, com o objetivo final de encontrar um modelo computacional e um subconjunto ótimo de atributos que alcancem os melhores desempenhos.

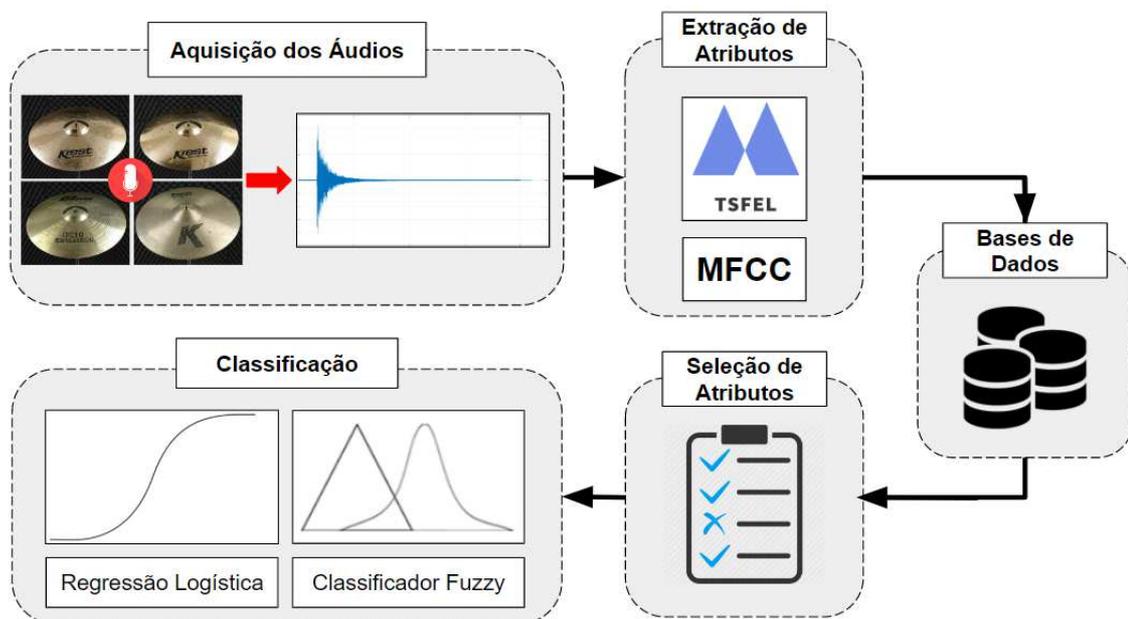


Figura 20 – Fluxograma das etapas do trabalho.

Fonte: Elaborado pelo Autor (2022).

3.1 AQUISIÇÃO DOS ÁUDIOS

Por se tratar de uma abordagem que, no melhor do conhecimento do autor, ainda não foi explorada, evitou-se, inicialmente, gerar esforços para selecionar uma grande variedade e quantidade de pratos. Portanto, foram selecionados apenas quatro pratos de bateria do tipo *Crash* com 18 polegadas de diâmetro, distribuídos entre três classes referentes às suas ligas de bronze (B8, B10, e B20) e de geometrias aproximadas, conforme mostrado pela Figura 21. Representando a classe de bronze B20, um prato do modelo *K*

Dark, fabricado pela empresa *Zildjian*, foi selecionado — Figura 22(a). Para a classe B10, utilizou-se um prato da marca *N Gate*, modelo *K10 Expression* - Figura 22(b). Por fim, duas versões de um prato da *Krest* modelo *Fusion* foram selecionadas para representar a classe de bronze B8 - Figuras 22(c) e 22(d). A primeira versão consiste em um equipamento novo e a outra de um equipamento danificado com um amassado na borda que se estende por 25,4 mm, 10 mm na direção radial e flexão máxima de 1 mm, conforme é mostrado pela Figura 23.

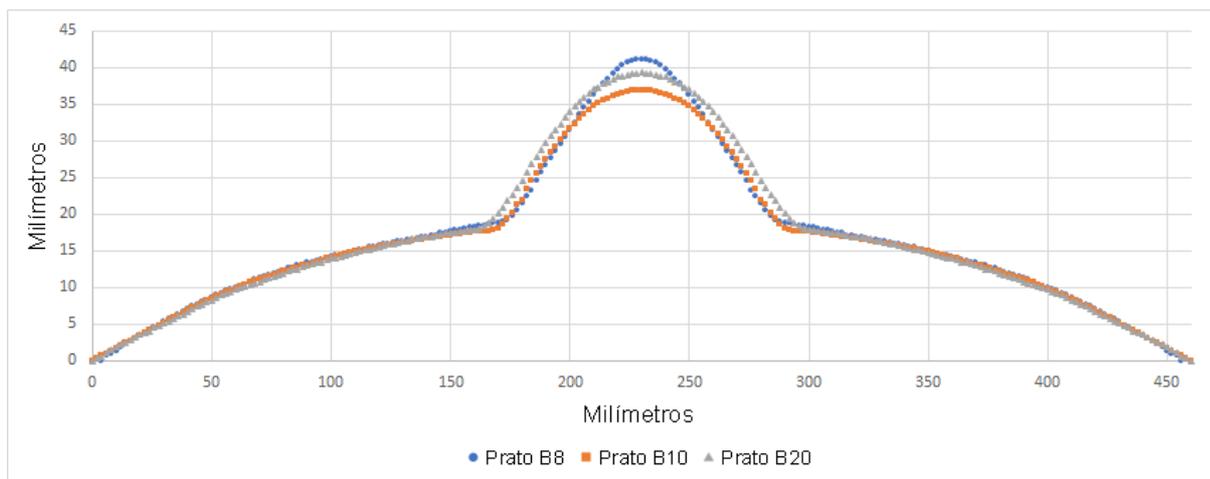


Figura 21 – Curvaturas medidas dos pratos utilizados.

Fonte: (BORATTO, 2019)

Na busca de uma representação do som mais fiel possível à realidade, a determinação dos microfones a serem utilizados no procedimento de captação se baseou na análise dos gráficos de resposta em frequência dos equipamentos disponíveis. Sendo assim, optou-se por selecionar aqueles que apresentaram a curva mais próxima de 0 dB, isto é, com menor ganho nas frequências. Os demais equipamentos de gravação foram utilizados conforme a disponibilidade da estrutura do estúdio de gravação e também ajustados com ganhos nulos. Portanto, a aquisição dos áudios foi conduzida na sala de gravação do estúdio ARGUS, equipada com um microfone *AKG C414 XL II*, ajustado como cardioide, e um *Shure PG81*, também cardioide, cujas especificações podem ser vistas, respectivamente, nas Figuras 24 e 25, interface de áudio *Behringer Xr18* (24 bits e 48 kHz) e o programa *REAPER* como *software* para a gravação.

Após a definição dos equipamentos a serem utilizados, a montagem do procedimento para a captação dos sons foi realizada com o auxílio de uma estante (ou suporte) apropriada para pratos, um pedal de bumbo de acionamento direto (*Direct Drive*)⁹ e um banco utilizado

⁹ Este termo refere-se ao mecanismo de transferência de movimento da base para o batedor do pedal feito através de uma peça única maciça e não por corrente, deixando o equipamento mais sensível e silencioso.



(a) Zildjian K Dark - B20.



(b) NGate K10 Expression - B10.



(c) Krest Fusion Cymbal - B8 (Novo).



(d) Krest Fusion Cymbal - B8 (Danificado).

Figura 22 – Pratos utilizados para a aquisição dos áudios.

Fonte: Autor (2021).



Figura 23 – Dano na borda do prato.

Fonte: Autor (2021).

apenas para elevar a altura do pedal, como ilustrado na Figura 26. O suporte utilizado para apoiar os pratos foi ajustado com uma leve inclinação e mínima pressão do feltro superior. O primeiro ajuste foi realizado para permitir que o batedor do pedal pudesse atingir a

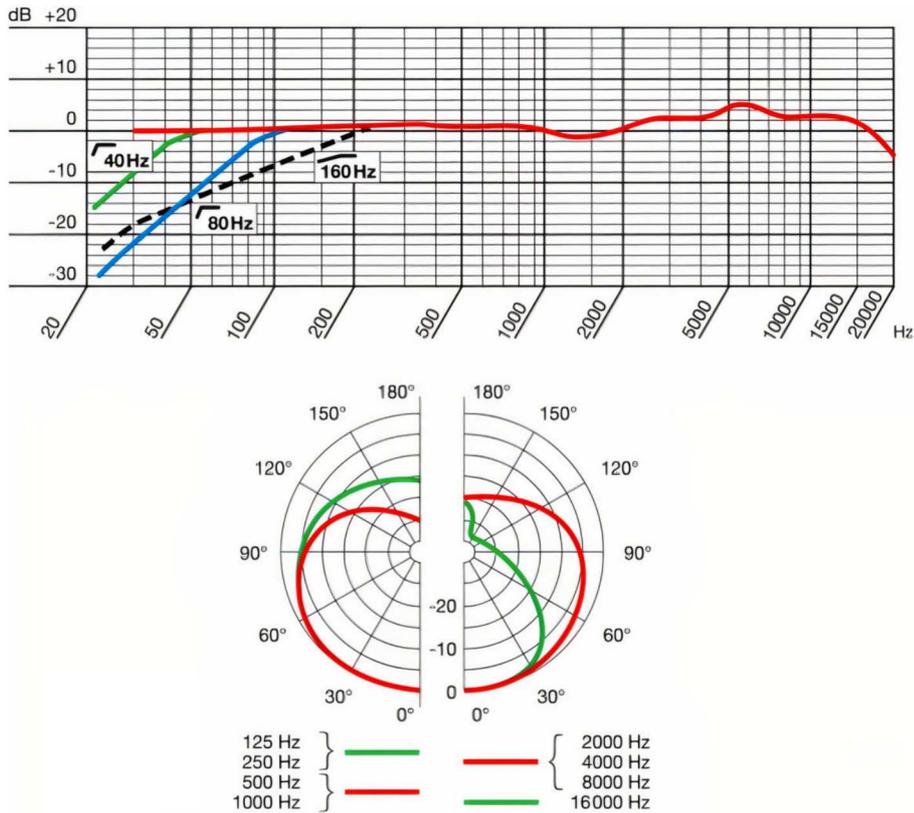


Figura 24 – Resposta em frequência e padrão polar do microfone AKG C414 XL II.

Fonte: (AKG, 2021)

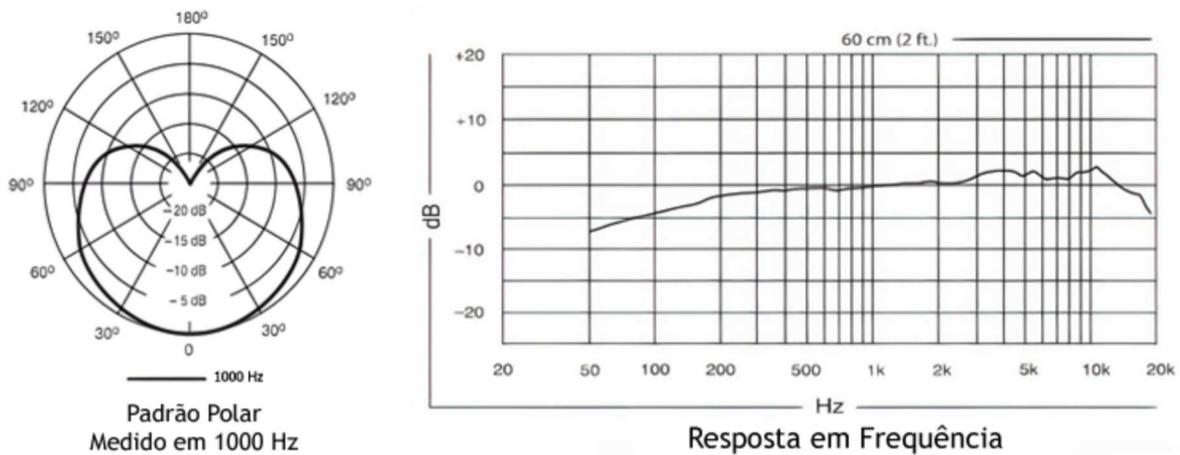


Figura 25 – Padrão polar e resposta em frequência do microfone Shure PG81.

Fonte: (SHURE, 2021)

borda do instrumento nos impactos. Já o segundo ajuste cumpriu apenas a função de restringir o movimento vertical do prato para fora da estante, de maneira que interferência na mobilidade de vibração fosse mínima. O pedal de bumbo necessitou ser adaptado em

uma superfície mais elevada e também ajustado com uma determinada inclinação para que, em conjunto com a angulação do prato, o batedor conseguisse atingir a borda do prato com toda a superfície de contato de feltro. Dessa forma, este equipamento foi fixado em uma banquetta, como mostra a Figura 26. Por fim, a montagem dos preparativos é finalizada com o posicionamento dos microfones, dispostos na direção radial dos pratos a uma distância de cerca de 250 mm da borda. Estes equipamentos foram conectados via cabos do tipo XLR à canais separados da interface de áudio, sendo possível, portanto, captar o som ao mesmo tempo pelos dois microfones, mas de forma independente.



Figura 26 – Configuração do procedimento para a aquisição de dados. 1) Estante do Prato. 2) Prato. 3) Microfone Shure PG 81. 4) Microfone AKG C414 XL II. 5) Pedal de bumbo. 6) Batedor do Pedal.

Fonte: Autor (2021).

Uma escala adesivada ao pedal de bumbo, como mostrado na Figura 27, permitiu que o arco do batedor pudesse ser ajustado com amplitudes semelhantes em cada impacto. Dessa forma, foi possível obter uma certa padronização das batidas, visto que conseguiu-se manter aproximadamente a maneira e a força com que se davam os impactos. Além

disso, este método também se mostrou vantajoso por ter um tempo de contato suficiente, durante as batidas, para não abafar o som do instrumento, uma vez que a mola do pedal é responsável por restaurar o batedor para a sua posição de equilíbrio.



(a) Ilustração do pedal separadamente. (b) Ilustração do pedal montado e ajustado para o procedimento de aquisição dos áudios.

Figura 27 – Pedal de bumbo utilizado para a aquisição dos sons. 1) Mecanismo *Direct-drive*. 2) Escala Adesivada. 3) Mola.

Fonte: Autor (2022).

O procedimento de captura dos sons foi conduzido em dois ambientes diferentes. Inicialmente, optou-se por cercar a área de captação por colchões, visando minimizar o reflexo das ondas sonoras no ambiente. Na sequência, tais colchões foram retirados e o processo foi conduzido no ambiente típico da sala de gravação. Além disso, foram aplicadas no mínimo 15 batidas em cada prato, cabendo ao responsável pelo procedimento decidir, por critérios subjetivos, a necessidade de mais batidas.

3.1.1 Processamento dos Sinais

O procedimento de aquisição dos áudios, previamente descrito, gerou 8 arquivos em formato *.rpp*, provenientes do *REAPER*. Cada arquivo contém a captação dos sons de todos os impactos provenientes de um determinado modelo de prato e em um dos ambientes, considerando os dois microfones utilizados, conforme é ilustrado pela Figura 28 e detalhado pela Tabela 3. Sendo assim, o particionamento e a renderização dos áudios foram realizados manualmente por meio do mesmo *software* de gravação utilizado, tomando como base uma duração de 21 segundos e frequência de amostragem de 44,1 kHz, conforme também ilustrado pela Figura 28, em que cada retângulo azul representa um particionamento realizado que deu origem a um arquivo de áudio. Dessa forma, foi

possível obter 276 arquivos de áudios em formato *.wav* (*Waveform Audio File Format*), sem compressão. Subdividindo-os por classes, tem-se 134 áudios para o material B8, 76 para o B10, e 66 para o B20. Uma descrição mais detalhada pode ser constatada na Tabela 3.

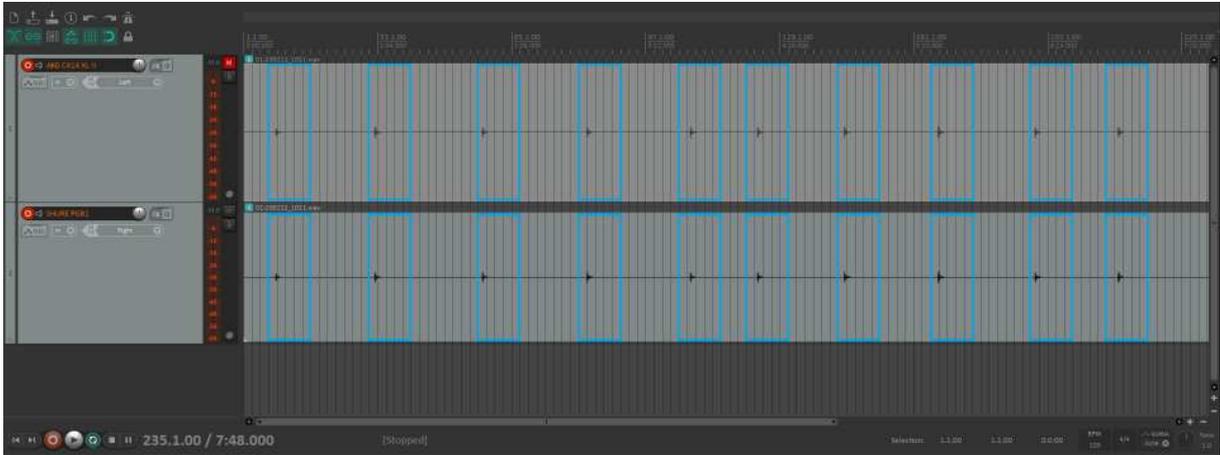


Figura 28 – Ilustração de um arquivo *.rpp* gerado no processo de gravação. Os registros dos microfones estão isolados: a primeira linha refere-se ao *AKG C414 XL II* e a segunda ao *Shure PG81*. Cada retângulo azul representa um particionamento realizado que deu origem a um arquivo de áudio *.wav*.

Fonte: Autor (2021).

Tabela 3 – Detalhamento da identificação dos arquivos de áudios.

Fabricante	Prato		Condição	Classe	Ambiente	# de Áudios (# Impactos)
	Série					
Krest	Fusion	Sem Dano	B8	com colchão	30 (15)	
				sem colchão	40 (20)	
Krest	Fusion	Danificado	B8	com colchão	32 (16)	
					sem colchão	32 (16)
NGate	K10 Expression	Sem Dano	B10	com colchão	30 (15)	
					sem colchão	46 (23)
Zildjian	K Dark	Sem Dano	B20	com colchão	30 (15)	
					sem colchão	36 (18)

Durante a análise exploratória dos dados, foi percebida a existência de arquivos renderizados com durações inferiores e superiores aos 21 segundos padronizados. Para contornar o problema dos arquivos de áudio com duração superior aos 21 segundos, optou-se por eliminar uma parcela “extra” do final do sinal, com a garantia de que a porção significativa do áudio não fosse afetada. Para os arquivos com duração inferior, a estratégia

utilizada foi a de aplicar o procedimento de *zero padding*, isto é, preencher o final do sinal com zeros até que o áudio complete a duração padrão de 21 segundos (ou os 926100 pontos do sinal¹⁰). Além disso, também foi verificada a presença de dois áudios ruidosos com falas. Estes, no entanto, não sofreram nenhum tratamento, sendo mantidos na base de dados com a finalidade de “estressar” os modelos computacionais.

3.2 EXTRAÇÃO DE ATRIBUTOS

Conforme mencionado anteriormente na Seção 1.1, um dos objetivos específicos deste trabalho é explorar atributos dos sinais. Dessa forma, atributos do domínio do tempo e coeficientes mel-cepstrais de frequência foram extraídos dos sinais, sendo detalhados nas Subseções 3.2.1 e 3.2.2, respectivamente.

3.2.1 Atributos Temporais

Tomando como base as informações disponibilizadas pela biblioteca para a extração de atributos de séries temporais TSFEL (*Time Series Feature Extraction Library*) (BARANDAS et al., 2020) utilizada, os 18 atributos temporais extraídos de cada sinal (x) são descritos a seguir, considerando sinais compostos de N pontos.

1. X_1 - **Energia Absoluta:** Calcula a energia absoluta do sinal, definida como a soma do quadrado de cada valor absoluto da série temporal: $\sum_{i=1}^N |x_i|^2$
2. X_2 - **Área sob a curva:** Calcula a área sob a curva do sinal através da regra dos trapézios: $\sum_{i=1}^N (t_i - t_{i-1}) \cdot \frac{(x_i + x_{i-1})}{2}$
3. X_3 - **Autocorrelação:** Calcula a autocorrelação do sinal, ou seja, a correlação cruzada da mesma sequência. \bar{x} é o complexo conjugado do sinal x e l é o atraso (*lag*): $\sum_{n \in Z} x(n) \bar{x}(n - l)$
4. X_4 - **Centroide ao longo do eixo do tempo:** $\frac{\sum_{i=0}^N t_i \cdot x_i^2}{\sum_{i=0}^N x_i^2}$
5. X_5 - **Entropia:** Calcula a entropia do sinal utilizando a Entropia Shannon (CRUTCHFIELD; FELDMAN, 2003). $-\sum_{k \in x} P(k) \cdot \log_2 P(k)$
6. X_6 - **Média do sinal derivado absoluto:** $media(|\Delta x|)$
7. X_7 - **Média do sinal derivado:** $media(\Delta x)$
8. X_8 - **Mediana do sinal derivado absoluto:** $mediana(|\Delta x|)$
9. X_9 - **Mediana do sinal derivado:** $mediana(\Delta x)$

¹⁰ A quantidade de pontos é calculada pelo produto da duração em segundos pela frequência de amostragem.

10. X_{10} - **Número de pontos de viragem negativos:** Calcula a quantidade de pontos de mínimo, ou vales, do sinal. O cálculo é feito a partir da contagem do número de vezes em que a derivada do sinal de áudio (Δx) deixa de ser negativa ($\Delta x[i] < 0$) e torna-se positiva ($\Delta x[i + 1] > 0$).
11. X_{11} - **Número de picos a partir de uma vizinhança definida:** Calcula o número de picos suporte em um determinado intervalo. Um pico de suporte S é definido como uma subsequência do sinal onde ocorre um valor que é maior que seus S vizinhos à esquerda e à direita (TIANO; BONIFATI; NG, 2021). Por padrão, o algoritmo utilizado configura $S = 10$, ou seja, a vizinhança é definida como 10 valores para a esquerda e 10 valores para a direita.
12. X_{12} - **Amplitude de pico a pico:** Calcula a distância entre os pontos globais de máximo e de mínimo do sinal. $|max(x) - min(x)|$
13. X_{13} - **Número de pontos de viragem positivos:** Calcula a quantidade de pontos de máximo, ou picos, do sinal. O cálculo é feito a partir da contagem do número de vezes em que a derivada do sinal de áudio (Δx) deixa de ser positiva ($\Delta x[i] > 0$) e torna-se negativa ($\Delta x[i + 1] < 0$).
14. X_{14} - **Distância percorrida pelo sinal:** $\sum_{i=0}^{N-1} \sqrt{1 + \Delta x_i^2}$
15. X_{15} - **Inclinação:** Calcula a inclinação do sinal ou, em outras palavras, o coeficiente angular (α) de uma reta ajustada aos pontos da série temporal, conforme a equação: $x = \alpha t + \beta$.
16. X_{16} - **Soma dos incrementos absolutos do sinal:** $\sum_{i=0}^{N-1} |\Delta x_i|$
17. X_{17} - **Energia total do sinal:** $\frac{\sum_{i=0}^N x_i^2}{t_N - t_0}$
18. X_{18} - **Taxa de cruzamento do zero (Zero Crossing Rate - ZCR):** Calcula a taxa com que o sinal de entrada muda de valores positivos para negativos e vice-versa, conforme mostrado pelo Algoritmo 1.

3.2.2 Coeficientes Mel-Cepstrais de Frequência

De acordo com Oppenheim e Schaffer em (OPPENHEIM; SCHAFER, 2004), *Cepstrum* é o termo utilizado para se referir ao espectro do log do espectro de uma forma de onda no tempo. Já o *Mel Cepstrum de Frequência* é derivado do logaritmo do espectro obtido de um banco de filtros com frequências centrais e larguras de banda definidas por um intervalo de frequência constante.

Os coeficientes mel-cepstrais de frequência (*Mel Frequency Cepstral Coefficients* - *MFCC*) são descritores de áudio que relacionam a frequência real e o tom percebido,

Algoritmo 1: Cálculo da taxa com que o sinal de entrada cruza o zero

Entrada: x

início

$signal =$ lista de zeros de tamanho N

$zc =$ lista de zeros de tamanho N

para $i \in [0, N]$ **faça**

se $x[i] > 0$ **então**

$signal[i] = 1$

fim

se $x[i] < 0$ **então**

$signal[i] = 0$

fim

fim

$\Delta signal = signal[i + 1] - signal[i]$

para $i \in [0, N]$ **faça**

se $\Delta signal[i] > 0$ **ou** $\Delta signal[i] < 0$ **então**

$zc[i] = 1$

fim

se $\Delta signal[i] = 0$ **então**

$zc[i] = 0$

fim

fim

$ZCR = \sum_i^N zc[i]$

fim

Saída: ZCR

mapeando as frequências para corresponder aproximadamente à capacidade de percepção auditiva humana. Sendo assim, tal mapeamento é aproximadamente linear até 1 kHz e logarítmico para frequências superiores (LOGAN, 2000; MUDA; BEGAM; ELAMVAZUTHI, 2010). Por conta disso, os MFCCs têm sido uma boa alternativa e amplamente utilizados em problemas de Recuperação de Informação Musical (*Music Information Retrieval - MIR*), o que inclui o reconhecimento de instrumentos musicais, como é o caso dos trabalhos (BHALKE; RAO; BORMANE, 2016; LOUGHRAN et al., 2008; LOUGHRAN et al., 2018; MALÍK; ORJEŠEK, 2016; NAGAWADE; RATNAPARKHE, 2017; SOUZA; BATISTA; SOUZA-FILHO, 2015).

A lógica utilizada para implementar a função de obtenção destes coeficientes em Python foi baseada no Relatório Técnico de Malcolm Slaney de 1998 (SLANEY, 1998) e é ilustrada na Figura 29. Sendo assim, ainda de acordo com SLANEY (SLANEY, 1998), a sequência de processamentos utilizados para a extração dos coeficientes cepstrais inclui para cada bloco de dados:

1. Janelar os dados com a Janela de Hamming;
2. Obter a Transformada Rápida de Fourier (*Fast Fourier Transform — FFT*), bem

como a sua magnitude;

3. Converter os dados da frequência (FFT) em saídas de banco de filtro;
4. Encontrar o log na base 10 (retificação não-linear);
5. Encontrar a transformação discreta de cosseno (*Discrete Cosine Transform - DCT*) para reduzir a dimensionalidade.

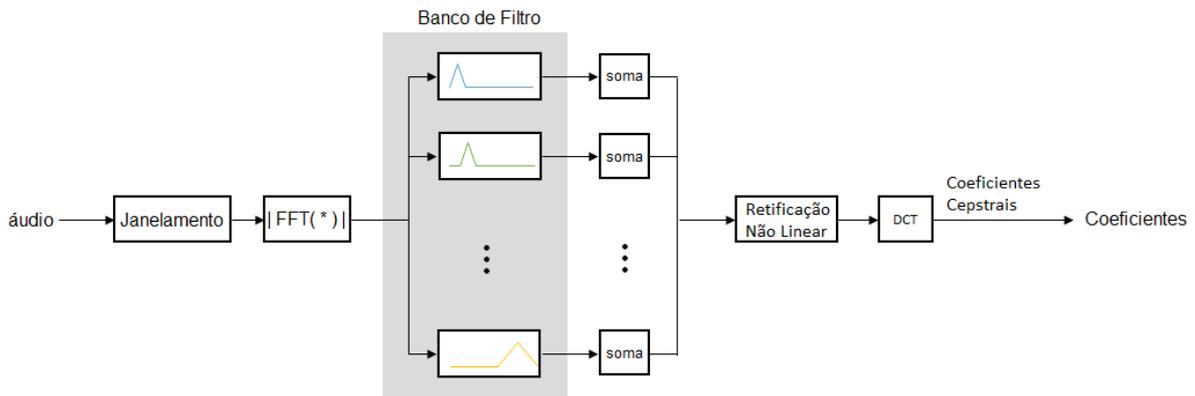


Figura 29 – Etapas para o cálculo dos MFCCs.

Fonte: Adaptado de (MATLAB, 2021).

Como mencionado, o janelamento utilizado por SLANEY (SLANEY, 1998), consequentemente também utilizado na implementação realizada pelo autor, é do tipo *Hamming*. A formulação matemática para esta categoria de janelamento é apresentada pela Equação 3.1 e está graficamente representada na Figura 30.

$$w = 0,54 - 0,46 \cdot \cos\left(\frac{2\pi n}{N}\right) \quad 0 \leq n \leq N \quad (3.1)$$

onde $N + 1$ corresponde ao comprimento da janela em número de pontos (OPPENHEIM; SCHAFER, 2009). Neste caso, optou-se por realizar um único janelamento do sinal, isto é, considerar $N + 1$ como sendo o tamanho do sinal de áudio. Sendo assim, como os sinais utilizados neste trabalho apresentam a mesma duração de 21 segundos e frequência de amostragem de 44,1 kHz, a quantidade de pontos é $N + 1 = 21 \cdot 44100 = 926100$.

A FFT (OPPENHEIM; SCHAFER, 2009) é aplicada ao sinal obtido após a operação de janelamento e o valor absoluto deste resultado será utilizado como entrada na etapa subsequente.

SLANEY (SLANEY, 1998) também mostra como o banco de filtro (indicado no item 3 acima) deve ser construído: são utilizados 13 filtros linearmente espaçados de 133,33 Hz entre as frequências centrais (*center frequencies - CF*), seguido de 27 filtros

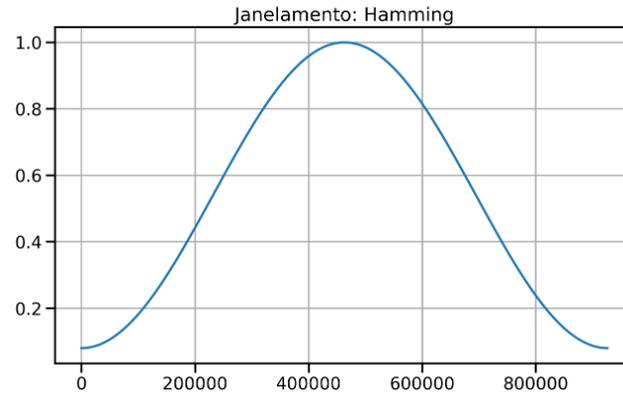


Figura 30 – Formato da janela única de Hamming utilizada.

Fonte: Autor.

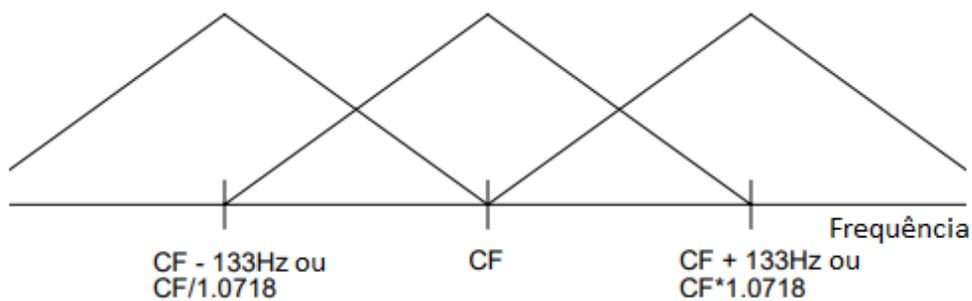


Figura 31 – Método de construção do banco de filtros proposto por Malcolm Slaney.

Fonte: (SLANEY, 1998)

logaritmicamente espaçados, separados por um fator de 1,0711703 em frequência, como mostrado pela Figura 31. Cada filtro é construído através da combinação da amplitude dos pontos da FFT, portanto a resposta em frequência dos 40 filtros é mostrada pela Figura 32.

Com relação às demais etapas, a Retificação Não Linear consiste no cálculo do logaritmo na base 10 das saídas do banco de filtro e a DCT (OPPENHEIM; SCHAFER, 2009) é utilizada para a redução da dimensionalidade antes de retornar os coeficientes calculados.

3.3 BASES DE DADOS

Baseando-se na extração de atributos temporais e de coeficientes mel-cepstrais de frequência, foi possível construir 5 bases de dados, visando explorar e avaliar as informações recuperadas dos áudios. Os conjuntos determinados estão descritos a seguir.

1. **C1:** Este conjunto de dados considera apenas os 18 Atributos Temporais extraídos dos sinais (vide Subseção 3.2.1).

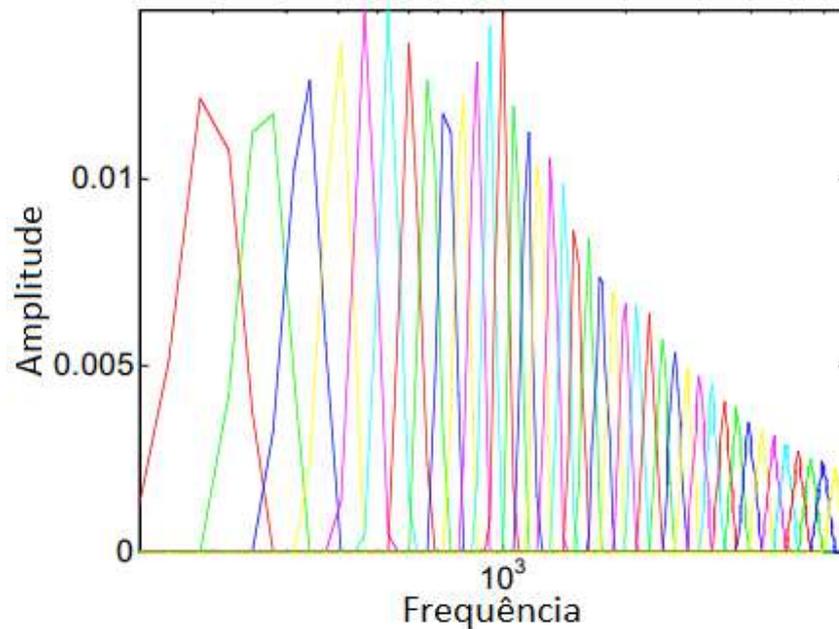


Figura 32 – Resposta em frequência do banco de filtros proposto por Malcolm Slaney.

Fonte: (SLANEY, 1998)

2. **C2:** Este conjunto de dados é formado pelos 13 primeiros Coeficientes Mel-Cepstrais de Frequência ($MFCC_{1 \rightarrow 13}$) extraídos dos sinais de áudio. A quantidade de coeficientes escolhida está relacionada à construção do banco de filtro, dado que são utilizados 13 filtros linearmente espaçados, seguidos de 27 logaritmicamente espaçados (vide Subseção 3.2.2).
3. **C3:** Este conjunto de dados é formado pela união dos conjuntos C1 e C2 anteriores. Portanto, é constituído por 18 Atributos Temporais e 13 Coeficientes Mel-Cepstrais de Frequência.
4. **C4:** Neste caso, foram considerados 40 Coeficientes Mel-Cepstrais de Frequência ($MFCC_{1 \rightarrow 40}$) extraídos dos sinais de áudio. Optou-se por incluir os demais coeficientes mel-cepstrais de frequência, relacionados ao mapeamento não linear das frequências, pois foi visto que um prato de bateria pode não apresentar um intervalo específico de frequências características.
5. **C5:** Este conjunto de dados é formado pela união dos conjuntos C1 e C4 anteriores. Portanto, é constituído por 18 Atributos Temporais e 40 Coeficientes Mel-Cepstrais de Frequência.

No entanto, como nem todos os atributos extraídos conseguem recuperar informações pertinentes dos sinais e/ou são adequados como entradas para os classificadores, optou-se por filtrar o conjunto de atributos de diferentes maneiras conforme o classificador

em operação. Sendo assim, ao considerar a Regressão Logística nos experimentos computacionais, um filtro de variância nula (Filtro 1) foi aplicado ao conjunto de atributos inicial, visando eliminar características com informações não relevantes. Já quando o classificador *Fuzzy* foi considerado, foi necessário aplicar um filtro para eliminar os atributos com médias absolutas inferiores a $1 e - 10$ com o objetivo de desconsiderar características com valores praticamente residuais, reduzindo, portanto, a quantidade de recursos computacionais a serem utilizados, além de evitar problemas de divisão por zero durante o cálculo das pertinências, como será visto a seguir na Subseção 3.4.2.

De maneira resumida, a Tabela 4 detalha as bases de dados utilizadas nos experimentos computacionais, mostrando as modificações causadas pela aplicação destes filtros.

Base de Dados	Dimensão Inicial	Variáveis Descartadas	Dimensão Final
Filtro 1			
C1	18	X_9	17
C2	13	-	13
C3	31	X_9	30
C4	40	-	40
C5	58	X_9	57
Filtro 2			
C1	18	X_7, X_9, X_{15}	15
C2	13	-	13
C3	31	X_7, X_9, X_{15}	28
C4	40	-	40
C5	58	X_7, X_9, X_{15}	55

Tabela 4 – Detalhamento das bases de dados utilizadas nos experimentos computacionais.

3.4 MODELOS DE CLASSIFICAÇÃO

3.4.1 Regressão Logística

A Regressão Logística (RL) é um classificador binário baseado em probabilidade que é amplamente utilizado em problemas de aprendizagem de máquinas. O seu modelo é uma transformação de regressão linear que estima a probabilidade de uma instância pertencer a uma determinada classe dentro de um conjunto binário de classes. E visto que a regressão logística é um modelo binário e probabilístico, a decisão de prever os rótulos torna-se simples. Se a probabilidade (\hat{p}) de uma instância \mathbf{x} for maior ou igual a 0,5, então é classificada como 1; caso contrário, é classificada como 0 (GERON, 2017), como mostrado nas Equações 3.2 e 3.3.

$$\hat{p}(y = 1|x; \theta) = \frac{1}{1 + \exp(-\theta^T x)} \quad (3.2)$$

$$\hat{y} = \begin{cases} 0, & \text{if } \hat{p} < 0,5 \\ 1, & \text{if } \hat{p} \geq 0,5 \end{cases} \quad (3.3)$$

onde $\theta \in \mathbb{R}^N$ são os parâmetros do modelo.

Embora a Regressão Logística seja concebida para a classificação binária, ela pode ser expandida para problemas multiclasse empregando uma estratégia heurística conhecida como *One-vs-Rest (OVR)* ou *One-vs-All (OVA)*. Este método encaixa um classificador binário para cada classe, que separará os dados que pertencem a essa classe em particular (saída 1) dos dados que não pertencem (saída 0) (LIU et al., 2021; PEDREGOSA et al., 2011).

A regressão logística não-regularizada, de acordo com (LEE et al., 2006), é um problema de otimização convexa sem restrições. No entanto, a regularização pode ajudar a evitar o sobreajustamento em algumas situações, por exemplo, com uma quantidade muito pequena de amostras no conjunto de treino ou uma com demasiados parâmetros para serem aprendidos (LEE et al., 2006). Em particular, a regularização *LASSO (Least Absolute Shrinkage and Selection Operator)* ou, em português, operador de menor retração absoluta e seleção) é também frequentemente utilizada para fins de seleção de características, dado que, utilizando a norma L_1 como termo de penalização, pode levar a coeficientes nulos. Assim, em problemas envolvendo múltiplos atributos que mostram correlação, a penalização de L_1 permite selecionar um atributo entre os correlacionados, associando os outros com coeficientes nulos. Portanto, este método é também conhecido como regularização L_1 .

Vale a pena lembrar que a norma 1 ($\|\cdot\|_1$) é definida como o máximo da soma dos valores absolutos para as entradas de cada linha. Portanto, dada uma matriz $A^{m \times n}$ esta norma pode ser representada matematicamente pela Equação 3.4.

$$\|A\|_1 = \max_{i \leq m} \sum_{j=1}^n |a_{ij}| \quad (3.4)$$

De acordo com (NG, 2004), é possível descrever uma regressão logística regularizada a partir da descoberta dos parâmetros θ que resolvem o problema de otimização mostrado na Equação 3.5. Note que o termo de regularização é representado por $R(\theta)$. Assim, se $R(\theta) \equiv 0$, então o modelo não é regularizado, mas se $R(\theta) = \|\theta\|_1$, o modelo torna-se L_1 regularizado e o problema de otimização a resolver torna-se representado pela Equação 3.6.

$$\min_{\theta} \sum_{i=1}^m -\log \text{Pr}(y^{(i)}|x^{(i)}; \theta) + \alpha R(\theta) \quad (3.5)$$

$$\min_{\theta} \|\theta\|_1 + C \sum_{i=1}^m \log \text{Pr}(y^{(i)}|x^{(i)}; \theta) \quad (3.6)$$

Por ser um modelo simples, a regressão logística é proposta como um modelo de classificação através do qual será possível avaliar alguns subconjuntos de características, determinados pelas técnicas de seleção de atributos descritas na Seção 3.5. Dessa forma, a função de regressão logística, disponível na estrutura do *scikit-learn*, foi configurada com os seguintes hiperparâmetros: limitação de 1000 iterações, norma L_1 como termo de penalidade e *liblinear solver* (FAN et al., 2008).

3.4.2 Classificador Fuzzy

O conceito da lógica fuzzy proposto por ZADEH em (ZADEH, 1965) manipula conceitos booleanos tradicionais, atribuindo graus de veracidade (CHERKASSKY, 1998) às informações que variam entre o completamente falso e completamente verdadeiro, passando a existir um grau de pertinência. Sendo assim, a estrutura da lógica fuzzy permite facilitar a tomada de decisões em um ambiente de imprecisão causado pela ausência de critérios bem definidos de pertencimento às classes (ZADEH, 1965).

Uma função de pertinência $f_A(x)$ para um conjunto difuso A no universo do discurso X é definida como $f_A(x) : X \rightarrow [0, 1]$, onde cada elemento de X é mapeado para um valor real entre 0 e 1. Assim, tomando um ponto qualquer $x \in X$, o valor obtido ao aplicar a função de pertinência neste ponto ($f_A(x)$) quantifica o grau de pertinência do elemento x ao conjunto difuso A , sendo, portanto, denominado valor ou grau de pertinência (ZADEH, 1965).

Para este trabalho, foram utilizadas as funções de pertinência triangular e Gaussiana, conforme (JANG CHUEN-TSAI SUN, 1997), representadas respectivamente nas Equações 3.7 e 3.8.

$$f_A(x) = \begin{cases} 0 & x \leq a \\ \frac{x-a}{m-a} & a < x \leq m \\ \frac{b-x}{b-m} & m < x < b \\ 0 & x \geq b \end{cases} \quad (3.7)$$

$$f_A(x) = \exp\left\{-\frac{(x-m)^2}{2\sigma^2}\right\} \quad (3.8)$$

Os valores de a , m e b da Equação 3.7 referem-se, respectivamente, aos vértices esquerdo, superior e direito de um triângulo acutângulo qualquer, conforme ilustrado na Figura 33. Com relação à Equação 3.8, os termos m e σ são os parâmetros utilizados para descrever uma distribuição Gaussiana e referem-se ao valor médio e ao grau de dispersão dos dados (desvio-padrão) dos valores que um atributo do conjunto de treino assume, conforme ilustrado na Figura 33.

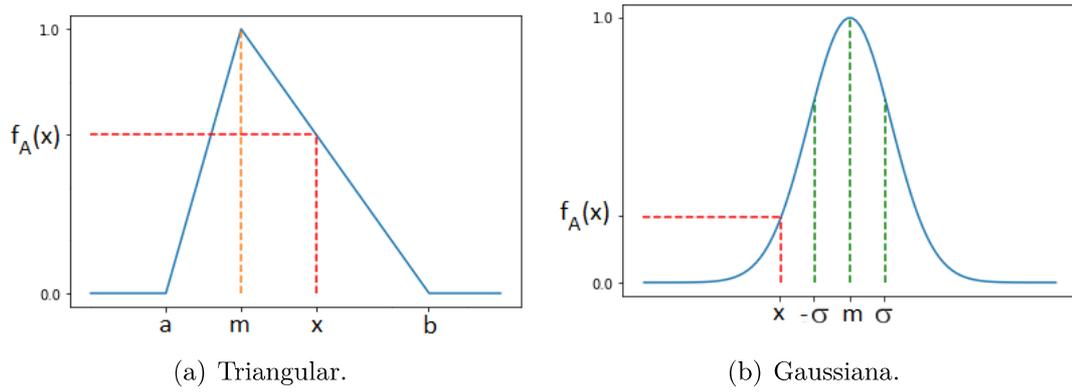


Figura 33 – Ilustração das funções de pertinência.

Fonte: Autor (2022).

Para a construção dos conjuntos de pertinência para cada classe, inicialmente foi necessário dividir a base de dados em grupos de treinamento e teste, selecionando aleatoriamente 70% das amostras para treinar o modelo e 30% para avaliá-lo. A identificação das classes dos materiais foi feita atribuindo-se os seguintes valores discretos: $B8 = 1$, $B10 = 2$ e $B20 = 3$. Desta maneira, após o grupo de dados para treinamento ser definido, um particionamento conforme os rótulos foi realizado. Como resultado, foi possível obter um conjunto de treinamento para cada material. Assim, tomando a matriz de atributos de um conjunto de treinamento (X) de uma determinada classe, os conjuntos de pertinência triangular são construídos para cada atributo e definidos pelo menor valor (parâmetro a - Equação (3.9)), pela média (parâmetro m - Equação (3.10)) e pelo maior valor (parâmetro b - Equação (3.11)) presente na coluna referente a este atributo na matrix X . Já a construção dos conjuntos de pertinência Gaussiana são definidos pela média (parâmetro m - Equação (3.10)) e desvio-padrão (parâmetro σ - Equação (3.12)) dos valores que um determinado atributo apresenta na matrix X . Sendo assim, para cada classe de material, os parâmetros a , m , b e σ das Equações (3.7) e (3.8) são calculados para cada atributo (subíndice j) segundo as Equações (3.9), (3.10), (3.11) e (3.12):

$$a_j = \min(X_j) \quad (3.9)$$

$$m_j = \frac{1}{N} \sum_{i=1}^N x_{ij} \quad (3.10)$$

$$b_j = \max(X_j) \quad (3.11)$$

$$\sigma_j = \sqrt{\frac{\sum_{i=1}^N (x_{ij} - \bar{x}_j)^2}{N}} \quad (3.12)$$

onde $i = 1, \dots, N$ representa o índice das amostras do conjunto de treinamento de cada material e j se refere aos atributos, cujo valor máximo é definido pelo número de colunas do conjunto de características em análise: C_1, C_2, C_3, C_4 ou C_5 .

3.5 SELEÇÃO DE ATRIBUTOS

Ao contrário do que se pode pensar, o acréscimo de dimensão (atributos) em um problema nem sempre significa uma melhoria no desempenho de um classificador, uma vez que as observações no espaço de alta dimensão se tornam muito esparsas. Assim, a redução da dimensionalidade pode trazer benefícios tais como a melhoria da eficiência computacional dos algoritmos e a simplificação do modelo gerado, facilitando a interpretação e visualização dos dados.

Uma forma de reduzir a dimensionalidade de um problema é selecionando atributos, procurando por um conjunto de características sub-ótimas, ou eliminando aquelas que são redundantes ou irrelevantes para o problema. A seleção de características normalmente é um procedimento integrado com o processo de aprendizagem dos métodos de aprendizado de máquina, tendo como objetivo a redução do número de características para aumentar o desempenho de um classificador.

Diversas técnicas para a escolha dos atributos podem ser encontradas na literatura, desde abordagens mais simples, como as que as utilizadas neste trabalho, passando por algoritmos baseados em meta-heurísticas, tal como a Evolução Diferencial, e até mesmo opções baseadas em redes neurais.

A análise de correlação é um método que pode ser bastante útil para identificar atributos com informações redundantes num conjunto de dados. A decisão de qual característica será selecionada depende unicamente do limite aceitável de correlação entre as variáveis. A utilização da regularização L_1 é outra forma de selecionar características, podendo estimar aquelas que são mais relevantes, especialmente em problemas que envolvem múltiplos atributos que apresentam correlação. Sendo assim, três técnicas de seleção de atributos foram consideradas neste trabalho:

1. **Técnica 1 (T1) - Análise dos Coeficientes da Regressão Logística L_1 Regularizada:** Conforme dito anteriormente, a regularização L_1 pode identificar e descartar os atributos irrelevantes associando-os a um coeficiente nulo. Desta forma, um total de 100 execuções independentes da regressão logística L_1 regularizada foram executadas, considerando conjuntos de treinamento constituídos por 70% do banco de dados. Em cada execução do algoritmo, o conjunto de treinamento era atualizado e os coeficientes com valores maiores a 10^{-6} eram contabilizados. Assim, no final de todas as execuções, as características associadas aos coeficientes contabilizados mais

de 50 vezes em pelo menos uma das três classes foram selecionadas como as mais relevantes.

2. **Técnica 2 (T2) - Análise do Limiar de Correlação:** O coeficiente de correlação é uma métrica usada para determinar o grau de associação entre duas variáveis. O método de Pearson usa a Equação 2.1 para calcular tais coeficientes, que podem assumir valores na faixa $[-1, 1]$. Assim, esta abordagem de seleção de atributos focou em avaliar o melhor limite de correlação aceitável entre as variáveis, variando o limiar de -1,0 até 1,0 em incrementos de 0,01. Para cada valor limiar, um subconjunto de atributos era formado ao selecionar apenas as características com índice de correlação inferior ao permitido (limiar). Para este caso, a avaliação do desempenho dos subconjuntos também considerou 100 execuções independentes, considerando a mesma proporção de 30% dos dados para avaliação, redefinidos a cada execução.
3. **Técnica 3 (T3) - Análise dos Coeficientes da Regressão Logística L_1 Regularizada seguida da análise do Limiar de Correlação:** A ideia principal desta técnica é primeiro eliminar os atributos irrelevantes aplicando a Técnica 1 e, em seguida, descartar as variáveis com um certo nível de redundância de informações aplicando a Técnica 2.

3.6 INDICADORES DE DESEMPENHO

Foram utilizados quatro indicadores de desempenho: Acurácia, Precisão, *Recall* e *F1*. No entanto, optou-se por tomar a primeira delas como base para a avaliação do limiar de correlação, devido a sua maior simplicidade de compreensão, já que a acurácia pode ser entendida como a porcentagem de classes previstas corretamente quando comparadas com a classificação verdadeira. Sendo assim, ao tomar a classe prevista $f(x_k)$ de uma amostra de teste e a sua classe verdadeira c_k , a acurácia pode ser calculada através da Equação 3.13, onde $I(verdadeiro) = 1$ e $I(falso) = 0$.

$$Acuracia = \frac{1}{N} \sum_{k=1}^N I(f(x_k) = c_k) \quad (3.13)$$

A precisão, mostrada na Equação 3.14, representa a taxa de casos corretamente classificados ou Verdadeiramente Positivos (*True Positive - TP*) com relação aos casos Previstos como Positivos (*Predicted Positive - PP*). Sendo assim, este indicador mostra o desempenho para cada classe individualmente.

$$Precisao = \frac{TP_k}{PP_k} \quad (3.14)$$

Já o *Recall* é uma medida que mostra a porcentagem de amostras positivas reais classificadas como positivas. Sua representação matemática é mostrada na Equação 3.15,

onde TP_k é o número de amostras que foram corretamente classificadas e FN_k é o número de amostras mal classificadas para a classe c_k (Falso Negativo - FN). Novamente, este indicador mostra o desempenho para cada classe individualmente.

$$Recall = \frac{TP_k}{TP_k + FN_k} \quad (3.15)$$

Por fim, a $F1$ também é uma medida que indica o desempenho individual para cada classe, sendo entendida como a média ponderada da Precisão e o $Recall$, atingindo o seu melhor valor em 1 e a pior pontuação em 0. Sua representação é mostrada na Equação 3.16, onde TP_k é o número de amostras que foram corretamente classificadas, FP_k é o número total de amostras mal classificadas (Falso Positivo - FP) e FN_k é o número de amostras mal classificadas para a classe c_k (Falso Negativo).

$$F1 = \frac{2TP_k}{2TP_k + FP_k + FN_k} \quad (3.16)$$

Visando facilitar as análises e a visualização dos desempenhos, os indicadores que mostram o acerto para classe individualmente (Precisão, $Recall$ e $F1$) tiveram as suas médias calculadas para chegar a um valor de referência de desempenho final, conforme mostrado pela Equação 3.17.

$$desempenho_final = \frac{classe_1 + classe_2 + classe_3}{3} \quad (3.17)$$

onde $classe_1$, $classe_2$ e $classe_3$ são os desempenhos médios das execuções independentes para cada classe (B8, B10 ou B20) individualmente.

4 RESULTADOS E DISCUSSÃO

4.1 EXPERIMENTOS COMPUTACIONAIS

Os experimentos computacionais foram conduzidos em um computador com as seguintes especificações: *CPU Intel(R) Core(TM) i5-1135G7*, 8 GB de memória RAM e sistema operacional *Windows 10 Home* versão *21H2*.

A base de dados de áudio construída (detalhada na Seção 3.1) está disponibilizada gratuitamente no repositório *Mendeley Data* (BORATTO; CURY; GOLIATT, 2022b)¹¹. Os códigos utilizados neste trabalho foram desenvolvidos em linguagem Python e são baseados nos seguintes pacotes: *pandas* (MCKINNEY, 2010), *numpy* (HARRIS et al., 2020), *scikit-learn framework* (PEDREGOSA et al., 2011), *scipy* (VIRTANEN et al., 2020) e *TSFEL* (BARANDAS et al., 2020), *librosa* (MCFEE et al., 2015), *seaborn* (WASKOM, 2021) e *matplotlib* (HUNTER, 2007).

Ao todo, foram considerados 60 casos. Para cada conjunto de atributos extraídos dos sinais (C1, C2, C3, C4 e C5), quatro abordagens de seleção foram consideradas (sem seleção, T1, T2, T3), formando subconjuntos de atributos utilizados como entrada para três modelos de aprendizado de máquina: Regressão Logística, *Fuzzy* com função de pertinência triangular e *Fuzzy* com função de pertinência Gaussiana. A Figura 34 ilustra a sequência de operações realizadas nos experimentos computacionais, que deram origem aos casos para cada conjunto de atributos.

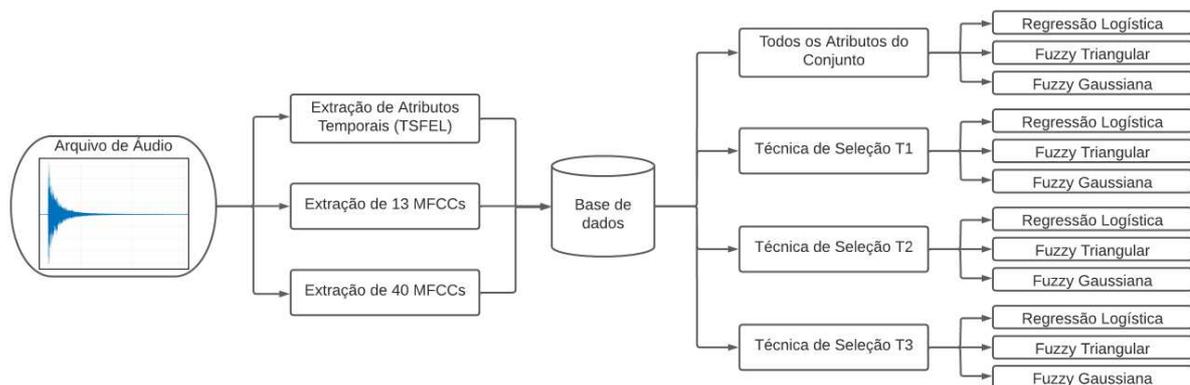


Figura 34 – Fluxograma ilustrativo da sequência de operações realizadas nos experimentos computacionais.

Para cada caso, foram consideradas 100 execuções independentes do algoritmo. A cada execução, as instâncias do conjunto de dados foram divididas aleatoriamente considerando uma proporção de 70% das entradas para treinamento e 30% para teste. Dessa forma, o resultado final obtido para cada métrica de desempenho é calculado a partir da média das 100 execuções independentes do algoritmo em cada caso sob análise.

¹¹ Link de acesso: <<https://data.mendeley.com/datasets/9tytvdx24/1>>

De posse dos resultados obtidos, buscou-se, inicialmente, identificar o conjunto de atributos mais adequado para o problema, considerando os desempenhos de classificação e a dimensionalidade inicial do conjunto. Na sequência, as técnicas de seleção foram avaliadas também com a finalidade de detectar a mais adequada. Com isso, foi possível conhecer o subconjunto ótimo de atributos selecionados responsáveis por atingir o melhor desempenho de classificação, bem como o classificador associado a este caso.

4.2 RESULTADOS E ANÁLISES

Os resultados para cada um dos 60 casos considerados neste trabalho estão dispostos no Apêndice A. De maneira resumida, a Tabela 5 traz os melhores desempenhos médios obtidos ao considerar cada um dos cinco conjuntos de dados, bem como os classificadores associados a esses casos.

Tabela 5 – Melhores desempenhos médios percentuais obtidos para cada conjunto de atributos, bem como o modelo de classificação e a técnica de seleção associados.

	Conjuntos de Atributos				
	C1	C2	C3	C4	C5
Modelo	Regressão Logística	Regressão Logística	Regressão Logística	Fuzzy Triangular	Regressão Logística
Técnica de Seleção	T2	T3	T2	T2	T2
Acurácia	97,16 ± 1,57	98,55 ± 1,19	99,75 ± 0,73	99,61 ± 0,70	99,76 ± 0,72
Precisão	97,37 ± 1,57	99,02 ± 0,83	99,69 ± 0,90	99,74 ± 0,47	99,70 ± 0,90
Recall	96,96 ± 1,73	98,02 ± 1,60	99,66 ± 0,99	99,48 ± 0,96	99,67 ± 0,98
F1	97,08 ± 2,72	98,46 ± 1,27	99,66 ± 0,99	99,60 ± 0,74	99,67 ± 0,99

A análise da Tabela 5 revela algumas informações importantes. Inicialmente, percebe-se que a Regressão Logística se sobressaiu como classificador, tendo sido associada a 4 dos 5 casos descritos na tabela. Avaliando os desempenhos médios, percebe-se que uma certa superioridade dos MFCCs em relação aos atributos temporais, uma vez que os conjuntos C2 e C4 alcançam, isoladamente, desempenhos médios superiores a C1, além de promoverem um aumento de desempenho quando associados aos atributos temporais (conjuntos C3 e C5). Isso, no entanto, já era esperado, uma vez que são descritores de áudio de alto nível.

Além disso, é possível verificar que, de maneira determinística, o conjunto de atributos C5 consegue atingir os maiores desempenhos em todas as métricas. No entanto, os desempenhos obtidos por C3 são ligeiramente inferiores (0,01%) em relação a C5 e os resultados atingidos por C4 são próximos aos de C3 e C5, com uma diferença média em torno de 0,16%. Por conta disso, um teste ANOVA é aplicado nestes 3 casos para avaliar se as diferenças existentes entre as médias dos desempenhos são estatisticamente significativas. Os resultados deste teste são apresentados na Tabela 6.

Tabela 6 – Resultado do teste ANOVA ao comparar a distribuição dos desempenhos obtidos pelos melhores casos dos conjuntos de atributos C3, C4 e C5. O valor P obtido é significativo para $P < 0,05$.

Comparativo	Métrica	Acurácia	Precisão	Recall	F1
C3 vs C4	Valor F	1,7028	0,1934	1,6986	0,2607
	Valor P	0,1934	0,6606	0,1940	0,6102
C3 vs C5	Valor F	0,0137	0,0040	0,0098	0,0064
	Valor P	0,9072	0,9496	0,9212	0,9364
C4 vs C5	Valor F	2,0436	0,1302	1,9810	0,3625
	Valor P	0,1544	0,7186	0,1608	0,5478

A diferença estatisticamente não significativa entre os melhores desempenhos alcançados pelos conjuntos de atributos C3, C4 e C5, conforme constatado pela Tabela 6, implica na necessidade de recorrer a outra métrica de avaliação para determinar o conjunto de atributos mais eficaz para esta investigação. Sendo assim, considerando que o conjunto C3 apresenta uma dimensão inicial após a filtragem (30 atributos) consideravelmente inferior em comparação a C4 (40 atributos) e C5 (57 atributos), este conjunto foi considerado como o mais relevante e adequado, devido ao menor esforço computacional inicial necessário.

Tomando o conjunto C3 como base para as análises, a Figura 35 detalha melhor o comportamento deste conjunto de atributos em termo dos melhores desempenhos médios alcançados por cada técnica de seleção. A Tabela 7 complementa este detalhamento informando os classificadores associados em cada caso, bem como a dimensão dos subconjuntos ótimos selecionados e a listagem dos atributos escolhidos.

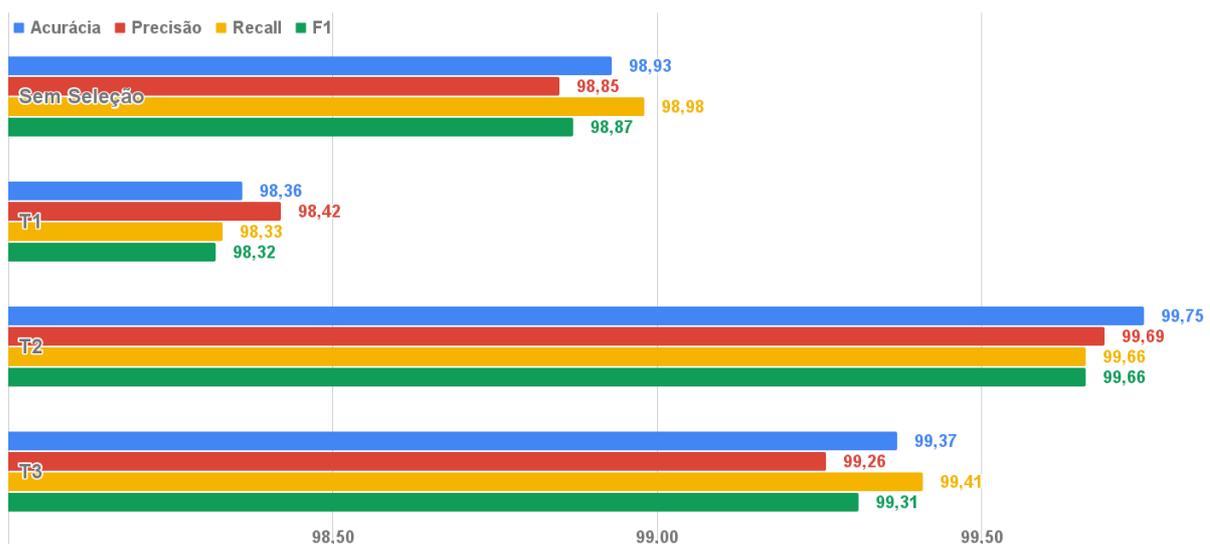


Figura 35 – Gráfico de barras com as melhores métricas de desempenho obtidas por cada técnica de seleção, considerando o conjunto C3.

Tabela 7 – Classificadores associados aos melhores casos de cada abordagem de seleção, dimensão do subconjunto ótimo e listagem dos atributos selecionados, correspondentes ao conjunto de dados C3.

Técnica de Seleção	Classificador	Dimensão do Subconjunto	Atributos Selecionados
Sem Seleção	Regressão Logística	30	$X_1, \dots, X_8, X_{10}, \dots, X_{18}$ $MFCC_1, \dots, MFCC_{13}$
T1	Regressão Logística	9	$X_{10}, X_{11}, X_{13}, X_{14}, X_{16}, X_{18},$ $MFCC_1, MFCC_{11}, MFCC_{12}$
T2	Regressão Logística	24	$X_1, X_4, X_5, X_6, X_7,$ $X_8, X_{10}, X_{11}, X_{12}, X_{15}, X_{18},$ $MFCC_1 \dots MFCC_{13}$
T3	Regressão Logística	5	$X_{10}, X_{11}, X_{14}, X_{18},$ $MFCC_{11}$

Entretanto, ainda que de maneira determinística a técnica de seleção T2 tenha apresentado superioridade nos resultados de desempenho apresentados pela Figura 35, a dimensão do subconjunto ótimo determinado (visto na Tabela 7) é consideravelmente alta quando comparada à dimensão dos subconjuntos obtidos pelas técnicas T1 e T3. Indo além, pode-se verificar que a diferença entre as métricas de desempenho obtidas por T2 e T3 são menores que 1%, estando em torno de 0,35%. Sendo assim, mesmo que um teste ANOVA aponte uma superioridade estatística do desempenho obtido por T2, é importante analisar o equilíbrio entre dimensionalidade e desempenho. Portanto, optou-se por selecionar as técnicas T2 e T3 para serem melhor detalhadas, visto que T2 atinge maiores desempenhos e T3 conseguiu resultados muito próximos com uma dimensionalidade consideravelmente inferior.

Em se tratando da relevância dos atributos, uma análise da distribuição de seus valores para cada classe, conforme é mostrado pelo Apêndice B, indica uma possível superioridade dos atributos MFCC 11, X18 (Taxa de cruzamento do zero) e MFCC 6, dado que ao menos a distribuição de uma das classes não apresenta sobreposição de intervalo com relação à distribuição dos valores das outras. Isso é mais evidente para o atributo MFCC 11, já que não há sobreposição de intervalo com nenhuma das outras classes.

Adicionalmente, a Figura 36 ilustra a matriz de confusão para cada um destes casos. Ao analisá-las, é possível perceber que o menor erro de classificação é, de fato, referente a técnica de seleção T2 e seguida por T3, considerando o conjunto de dados C3. Era de se esperar que as maiores taxas de confusão estivessem relacionadas entre os materiais B8 e B10 devido a sua proximidade em composição de estanho, como observado nas matrizes de confusão das Figuras 36(a) e 36(b). A classe de bronze B20, por outro lado, apresenta

uma maior dissemelhança com as outras ligas, uma vez que o aumento da concentração de estanho em sua composição química provocou alterações significativas neste material. No entanto, ao considerar as técnicas de seleção T2 e T3, cujas matrizes de confusão são ilustradas respectivamente pelas Figuras 36(c) e 36(d), observa-se que a maior proporção de erro (0,010) está associada à classe B20. A avaliação de tal erro revela que suas origens estão relacionadas à presença dos áudios ruidosos, de forma que ao retirá-los da base de dados, esta taxa de confusão desaparece em todos os 4 casos.

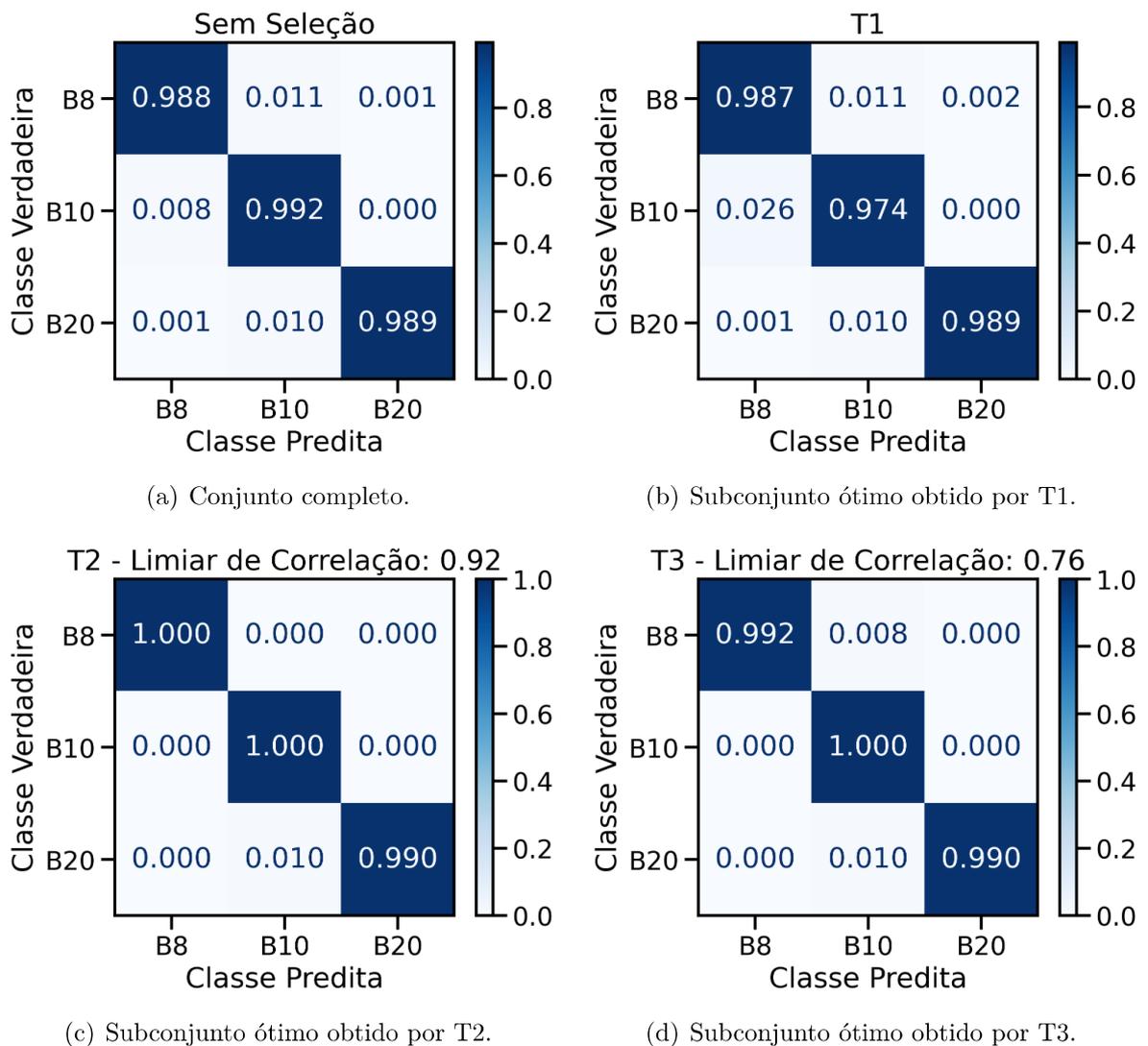
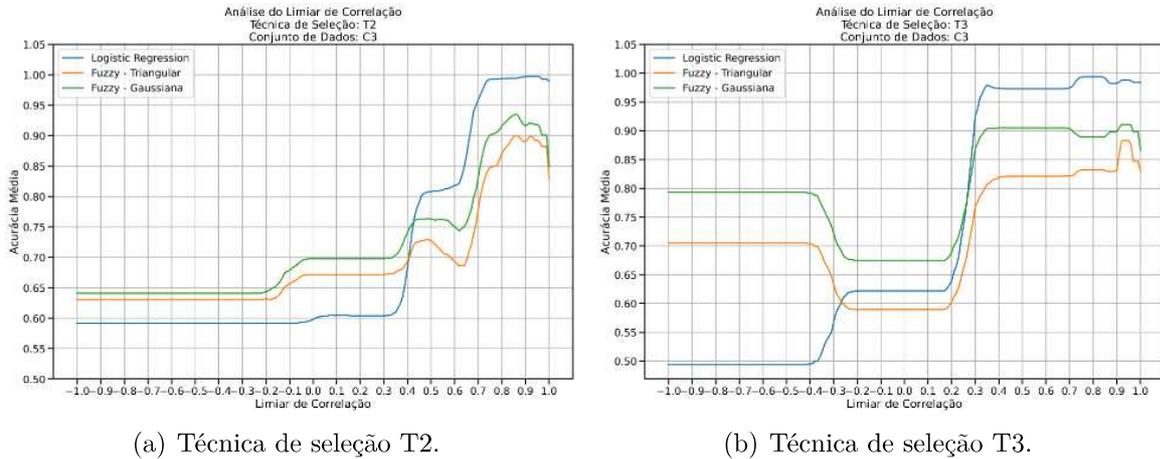


Figura 36 – Matrizes de confusão média considerando os melhores modelos em cada abordagem de seleção de características aplicadas à base de dados C3.

Portanto, para uma melhor compreensão destas técnicas, uma avaliação mais detalhada é realizada, inicialmente, através da análise do comportamento das acurácias médias em relação aos limiares de correlação, ilustrado na Figura 37. A avaliação destes gráficos mostra que as acurácias médias tendem a aumentar significativamente conforme os atributos com alto grau de correlação são descartados. É possível perceber também

que, especificamente para este conjunto de dados, a função de pertinência Gaussiana do classificador *Fuzzy* apresentou uma melhor adaptação do que a triangular, visto os maiores desempenhos médios para todos os valores limiares de correlação. Além disso, ao descartar atributos com correlações moderadas, ficou notória a superioridade da Regressão Logística.



(a) Técnica de seleção T2.

(b) Técnica de seleção T3.

Figura 37 – Comportamento da acurácia média em relação aos limiares de correlação referentes ao conjunto de dados C3 e considerando os 3 classificadores.

Esse tipo de abordagem para eliminar variáveis correlacionadas foi responsável por descartar 6 atributos do conjunto inicial C3 (Sem Seleção) ao considerar um valor de limiar de correlação de 0,92, e por descartar 4 atributos no subconjunto ótimo obtido por T1 ao considerar um valor de limiar de correlação de 0,76. Dessa forma, e dado os valores de desempenhos obtidos, pode-se inferir que a abordagem de seleção baseada na análise dos coeficientes da regressão logística (T1) conseguiu identificar grande parte dos atributos com informações pouco relevantes para o problema, ao passo que a análise do limiar de correlação (T2) foi capaz de detectar variáveis com alto grau de informações redundantes. A aplicação subsequente destas técnicas (T3), portanto, conseguiu descartar grande parte das informações não essenciais e em redundância presentes neste problema.

A análise da Figura 35 e da Tabela 7 revelam alguns pontos interessantes para serem discutidos. Um comparativo entre os desempenhos médios obtidos pelo conjunto C3 completo (sem seleção) e pelo subconjunto obtido por T1, considerando a regressão logística como classificador, mostra uma queda de desempenho após a seleção dos atributos. Em contrapartida, quando este subconjunto é submetido a uma nova técnica de seleção, o desempenho final obtido é alavancado e supera os alcançados por C3 sem seleção. Esse comportamento é uma evidência de que os atributos com níveis de correlação significativos remanescentes no conjunto, após a eliminação das variáveis não essenciais via T1, tornaram-se ainda mais relevantes, acarretando a queda de desempenho. Ao descartar variáveis via análise do grau de correlação, foi possível diminuir a redundância de informação e, conseqüentemente, o nível de importância das variáveis foi corrigido. Por outro lado,

quando a análise de seleção é feita considerando apenas o grau de correlação (T2), as variáveis de pouca relevância para o problema não causam tanto impacto, considerando a regressão logística regularizada como classificador, pois os próprios coeficientes do modelo conseguem corrigir esse problema ao atribuir pesos pequenos à essas variáveis.

4.3 VANTAGENS E LIMITAÇÕES

Tendo em vista que a base de dados construída para este trabalho é restrita em termos de quantidade e variedade de pratos e que os áudios foram captados em um ambiente controlado e com toques relativamente padronizados, os desempenhos máximos superiores a 99% podem ser consequência desses fatores, mesmo que a aquisição dos áudios tenha considerado diferentes perspectivas de microfone e ambiente. Por outro lado, ainda que o procedimento experimental desenvolvido para padronizar a coleta dos sons permita ser otimizado e não represente uma situação de um problema real, que podem ter presença de ruídos indesejados, os resultados alcançados demonstram a sua eficiência.

Apesar dos resultados promissores obtidos neste estudo, é necessário enfatizar a necessidade de expansão da base de dados quanto a número de amostras e variedade de pratos. Como resultado, avaliações mais abrangentes dos modelos de aprendizado de máquina e da influência do material sobre as características acústicas destes instrumentos podem ser realizadas. Além disso, destaca-se também a necessidade de aumentar o número de modelos de aprendizagem de máquinas, incluindo modelos de aprendizagem profunda que atuam diretamente sobre os sinais brutos. Neste caso, o uso de redes neurais, por exemplo, permite que funções mais complexas e não lineares sejam mapeadas, dado que esta categoria de modelo utiliza um número mais elevado de camadas de processamento (SHRESTHA; MAHMOOD, 2019). Dessa forma, as redes neurais podem trabalhar com os dados brutos, extraindo outras informações que não puderam ser representadas pelos descritores utilizados neste estudo.

A estratégia utilizada para avaliar os atributos, baseada no particionamento em conjuntos com uma posterior seleção, pode não ser o melhor caminho em uma aplicação real ou em trabalhos futuros. Este trabalho seguiu esta abordagem, pois se optou por investigar os mais diversos aspectos dos sinais, uma vez que este tipo de problema foi pouco ou até mesmo, no melhor dos conhecimentos do autor, não tenha sido explorado. Ainda nesse sentido, há o incentivo da aplicação de outras opções de técnicas de seleção de atributos, principalmente sobre o conjunto de atributos C5, por atingir de maneira determinística os melhores desempenhos médios. Abordagens meta-heurísticas, tais como a Evolução Diferencial podem ser alternativas interessantes para a escolha dos atributos. Técnicas baseadas em redes neurais também podem ser interessantes, visto que a rede de neurônios pode conseguir identificar e aprender relações entre as variáveis, tal como é apresentado em (NOOR, 2021), em que utilizam uma combinação de decodificadores convolucionais para

redução de ruído (*convolutional denoising autoencoder*) e uma rede neural convolucional para o aprendizado automático e não supervisionado de atributos.

Mesmo que esta base de dados limitada não permita a generalização dos resultados obtidos neste estudo, o modelo e as variáveis destacadas servem como um ponto de partida para os trabalhos futuros nesta linha. Nesse sentido, os elevados desempenhos alcançados encorajam o aprofundamento da investigação deste problema e o desenvolvimento de modelos *online* que podem ser incluídos em um dispositivo móvel de baixo custo para auxiliar fabricantes de pratos no controle da qualidade do material bruto utilizado, isto é fornecer uma certa garantia de que o material produzido é próximo ao esperado, uma vez que um teste de composição química é relativamente caro.

5 CONCLUSÃO

Grande parte dos trabalhos encontrados na literatura que se utilizam de técnicas convencionais de aprendizado de máquinas para a classificação de pratos de bateria focam em identificá-los por tipo ou até mesmo a maneira como foram tocados. Diferentemente destes estudos, este trabalho deu uma atenção especial ao material e propôs um problema para classificá-los conforme suas respectivas ligas de bronze.

Por ser um tipo de investigação incomum, inicialmente foi necessário desenvolver um procedimento para a aquisição dos áudios de maneira padronizada, o que possibilitou a captação de sons de quatro pratos de bateria, considerando a perspectiva de dois microfones e dois ambientes distintos. Em seguida, optou-se por explorar uma variedade de atributos e três diferentes técnicas de seleção, usando três diferentes classificadores como forma de avaliação. Sendo assim, as principais conclusões deste trabalho podem ser resumidas nos seguintes tópicos:

1. A combinação entre o conjunto de dados C3, a técnica de seleção T2 e a regressão logística levou ao melhor desempenho de classificação e com os menores erros.
2. Embora a técnica de seleção T2, combinada com a regressão logística, conduza ao melhor desempenho médio, a técnica T3 consegue um desempenho bastante próximo considerando aproximadamente apenas um quinto ($\approx 20\%$) da quantidade de atributos selecionados por T2.
3. Mesmo que o procedimento de aquisição dos áudios tenha sido bem realizado e considerado diferentes perspectivas de captura, a base de dados construída ainda é limitada em quantidade de amostras e variedades de pratos, não permitindo a generalização dos resultados aqui encontrados.

Os resultados promissores atingidos encorajam o aprofundamento e o desenvolvimento de novas pesquisas que busquem examinar os elementos que contribuem para a formação da acústica de pratos de bateria. Portanto, pode-se dizer que as principais contribuições deste trabalho se resumem a:

1. Promover uma inicialização do uso de aprendizado de máquina para o estudo dos elementos acústicos em pratos de bateria;
2. Possibilitar e estimular o desenvolvimento de aplicativos para auxiliar fabricantes de pratos no controle da qualidade do material fundido;
3. Promover o conhecimento sobre a influência do material na acústica de pratos de bateria;

4. Estimular o desenvolvimento de ferramentas didáticas para o ensino de acústica de pratos de bateria.
5. Contribuir para o desenvolvimento da área de reconhecimento de materiais via som por promover uma distinção entre materiais semelhantes.

5.1 TRABALHOS FUTUROS

Por ser um tipo de investigação incomum, existe uma extensa possibilidade de trabalhos futuros que podem ser desenvolvidos, tal como usar técnicas de aprendizado de máquina para estudar os demais aspectos que apresentam influência na acústica de pratos de bateria.

No entanto, grande parte das oportunidades podem ser encontradas nas limitações deste trabalho, descritas na Seção 4.3. Portanto, algumas sugestões para investigações futuras são listadas a seguir:

1. Reproduzir este trabalho em uma base de dados mais completa, considerando uma maior quantidade de amostras de material e variações de pratos em tamanho, marca, modelo, etc.
2. Explorar e avaliar outras abordagens de seleção de atributos e modelos de classificação, incluindo modelos de aprendizagem profunda. Além disso, também sugere-se utilizar imagens de espectrograma ou cocleograma como entradas de classificadores baseados em reconhecimento de imagens, tal como uma rede neural convolucional - CNN.
3. Desenvolver modelos *online* para aplicar em situações reais, até mesmo em um contexto fabril.

5.2 PARTICIPAÇÃO E COLABORAÇÃO EM EVENTOS

1. Participação

- a) XI Congresso Brasileiro de Engenharia de Fabricação (COBEF 2021).
- b) XIV Encontro Acadêmico de Modelagem Computacional (EAMC 2020).
- c) XII Congresso Iberoamericano de Acústica (FIA2022).

2. Colaboração

- a) XVIII Simpósio Brasileiro de Computação Musical (SBCM 2021) - Suporte Técnico.
- b) XV Encontro Acadêmico de Modelagem Computacional (EAMC 2022) - Suporte Técnico.

- c) XXIV Semana da Computação UFJF - Comissão Organizadora de Alunos.
- d) XXIV Semana da Computação UFJF - Ministrante do minicurso “Introdução a Recuperação de Informação de Áudio”.

5.3 PRODUÇÕES ACADÊMICAS

1. Análise dos efeitos de dois diferentes métodos de fabricação em pratos de bateria, publicado no XI Congresso Brasileiro de Engenharia de Fabricação (COBEF) - 2021 (BORATTO et al., 2021).
2. *A Fuzzy Approach to Drum Cymbals Classification*, publicado na revista *IEEE Latin America Transactions* (2022) (BORATTO; CURY; GOLIATT, 2022a).
3. Base de dados *Crash Cymbal Sounds* disponibilizada no repositório *Mendeley Data* (2022). (BORATTO; CURY; GOLIATT, 2022b).
4. *Dynamic analysis applied to cymbals with different bronze compositions*, publicado na modalidade Pôster no XII Congresso Iberoamericano de Acústica (FIA) - 2022.
5. A Recuperação de Informação Musical como ferramenta para análise sonora: Uma Inicialização, publicado no XII Congresso Iberoamericano de Acústica (FIA) - 2022.
6. *Data-driven cymbal bronze alloy identification via evolutionary machine learning with automatic feature selection*, publicado na revista *Journal of Intelligent Manufacturing* (2022) (BORATTO et al., 2022).
7. *Machine learning-based classification of bronze alloy cymbals from microphone captured data enhanced with feature selection approaches*, publicado na revista *Expert Systems With Applications* (2023) (BORATTO; CURY; GOLIATT, 2023).

REFERÊNCIAS

- AIHARA, T.; ITO, K. Relationship between chaotic vibrations and acoustic properties of percussion cymbals. *Results in Engineering*, v. 14, p. 100419, 2022. ISSN 2590-1230.
- AKG. *C414 XLII - Reference multipattern condenser microphone*. 2021. Acesso em: 28 de Mar. de 2021. Disponível em: <<https://www.akg.com/Microphones/Condenser/%20Microphones/C414+XLII.html>>.
- BARANDAS, M. et al. Tsfel: Time series feature extraction library. *SoftwareX*, Elsevier, v. 11, p. 100456, 2020.
- BHALKE, D.; RAO, C. R.; BORMANE, D. Hybridisation of mel frequency cepstral coefficient and higher order spectral features for musical instruments classification. *Archives of Acoustics*, v. 41, n. 3, 2016. ISSN 2300-262X.
- BORATTO, T.; CURY, A.; GOLIATT, L. A fuzzy approach to drum cymbals classification. *IEEE Latin America Transactions*, v. 20, n. 9, p. 2172–2180, 2022.
- BORATTO, T. H.; CURY, A.; GOLIATT, L. *Crash Cymbal Sounds*. 2022. <<https://data.mendeley.com/datasets/9tytvdx24/1>>.
- BORATTO, T. H.; CURY, A. A.; GOLIATT, L. Machine learning-based classification of bronze alloy cymbals from microphone captured data enhanced with feature selection approaches. *Expert Systems with Applications*, v. 215, p. 119378, 2023. ISSN 0957-4174.
- BORATTO, T. H. A. *Análise dinâmica aplicada à pratos de percussão de diferentes ligas de bronze*. 45 p. Monografia (Graduação em Engenharia Mecânica.) — Universidade Federal de Juiz de Fora, Juiz de Fora, 2019.
- BORATTO, T. H. A. et al. Análise dos efeitos de dois diferentes métodos de fabricação em pratos de bateria. In: *XI Congresso Brasileiro de Engenharia de Fabricação (COBEF)*. [S.l.: s.n.], 2021.
- BORATTO, T. H. A. et al. Data-driven cymbal bronze alloy identification via evolutionary machine learning with automatic feature selection. *Journal of Intelligent Manufacturing*, Springer Science and Business Media LLC, nov. 2022. Disponível em: <<https://doi.org/10.1007/s10845-022-02047-3>>.
- BOSLAUGH, P. A. W. S. *Statistics in a Nutshell: A Desktop Quick Reference*. [S.l.]: O'Reilly Media, 2008. (In a Nutshell). ISBN 0596510497; 9780596510497.
- CAVACO, S.; ALMEIDA, H. Automatic cymbal classification using non-negative matrix factorization. In: *2012 19th International Conference on Systems, Signals and Image Processing (IWSSIP)*. [S.l.: s.n.], 2012. p. 468–471.
- CHERKASSKY, V. Fuzzy inference systems: A critical review. In: KAYNAK, O. et al. (Ed.). *Computational Intelligence: Soft Computing and Fuzzy-Neuro Integration with Applications*. Berlin, Heidelberg: Springer Berlin Heidelberg, 1998. p. 177–197. ISBN 978-3-642-58930-0.
- CHHABRA, A. et al. Drum instrument classification using machine learning. In: *2020 2nd International Conference on Advances in Computing, Communication Control and Networking (ICACCCN)*. [S.l.: s.n.], 2020. p. 217–221.

CHIAVERINI, V. *Tecnologia Mecânica. Processos de Fabricação e Tratamento - Volume 2*. 2a. ed. [S.l.]: McGraw-Hill, 1986.

CLEARY, J. G.; TRIGG, L. E. K*: An instance-based learner using an entropic distance measure. In: PRIEDITIS, A.; RUSSELL, S. (Ed.). *Machine Learning Proceedings 1995*. San Francisco (CA): Morgan Kaufmann, 1995. p. 108–114. ISBN 978-1-55860-377-6.

CRUTCHFIELD, J. P.; FELDMAN, D. P. Regularities unseen, randomness observed: Levels of entropy convergence. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, v. 13, n. 1, p. 25–54, 2003.

DINIZ, P. S. R.; SILVA, E. da; NETTO, S. L. *Digital Signal Processing*. 2. ed. Cambridge, England: Cambridge University Press, 2010.

FAN, R.-E. et al. Liblinear: A library for large linear classification. *J. Mach. Learn. Res.*, JMLR.org, v. 9, p. 1871–1874, jun. 2008. ISSN 1532-4435.

FU, H. et al. Wood material recognition for industrial applications. *Systems Science & Control Engineering*, Taylor & Francis, v. 6, n. 3, p. 346–358, 2018.

GERON, A. *Hands-on machine learning with Scikit-Learn and TensorFlow : concepts, tools, and techniques to build intelligent systems*. Sebastopol, CA: O'Reilly Media, 2017. ISBN 978-1491962299.

GILLARD, J. One-way analysis of variance (anova). In: _____. *A First Course in Statistical Inference*. Cham: Springer International Publishing, 2020. p. 91–101. ISBN 978-3-030-39561-2.

GIULIO, G. et al. Experimental vibrational analysis of drum cymbals. v. 1, p. 724–730, 01 2001.

GROOVER, M. P. *Fundamentals of Modern Manufacturing*. [S.l.: s.n.], 2010. ISBN 9780470467008.

GUO, H.; JIANG, H.; GAO, M. Research on sound source material recognition technology in indoor geotechnical inspection. *Geofluids*, v. 2022, p. 1–12, 06 2022.

HARRIS, C. R. et al. Array programming with NumPy. *Nature*, Springer Science and Business Media LLC, v. 585, n. 7825, p. 357–362, set. 2020.

HARRISON, J. J.; HILL, A. J. A scientific approach to microphone placement for cymbals in live sound. In: . [S.l.: s.n.], 2013.

HERRERA, P.; YETERIAN, A.; GOUYON, F. Automatic classification of drum sounds: A comparison of feature selection methods and classification techniques. In: ANAGNOSTOPOULOU, C.; FERRAND, M.; SMAILL, A. (Ed.). *Music and Artificial Intelligence*. Berlin, Heidelberg: Springer Berlin Heidelberg, 2002. p. 69–80. ISBN 978-3-540-45722-0.

HUNTER, J. D. Matplotlib: A 2d graphics environment. *Computing in Science & Engineering*, IEEE COMPUTER SOC, v. 9, n. 3, p. 90–95, 2007.

JANG CHUEN-TSAI SUN, E. M. J.-S. R. *Neuro-Fuzzy and Soft Computing: A Computational Approach to Learning and Machine Intelligence*. 1. ed. [S.l.]: Prentice Hall, 1997. ISBN 0132610663; 9780132610667.

JAVH, J.; SLAVIČ, J.; BOLTEŽAR, M. The subpixel resolution of optical-flow-based modal analysis. *Mechanical Systems and Signal Processing*, v. 88, p. 89–99, 2017. ISSN 0888-3270.

KASELOURIS, E. et al. A detailed fem study on the vibro-acoustic behaviour of crash and splash musical cymbals. *International Journal of Circuits, Systems and Signal Processing*, v. 16, p. 948–955, 03 2022.

KEELEY, E. *Bit depth: What is it and how does it affect your work?* 2019. Disponível em: <<https://www.masteringbox.com/bit-depth/>>.

KIM, T. K. Understanding one-way ANOVA using conceptual figures. *Korean Journal of Anesthesiology*, The Korean Society of Anesthesiologists, v. 70, n. 1, p. 22, 2017.

KRZYWICKI, T. Application of meta-learning methods in the recognition of drums and cymbals on the basis of short sound samples. In: . [S.l.: s.n.], 2019.

KURATANI, F. et al. Understanding the effect of hammering process on the vibration characteristics of cymbals. *Journal of Physics: Conference Series*, IOP Publishing, v. 744, p. 012110, sep 2016.

LEE, S. et al. Efficient L1 regularized logistic regression. In: *Proceedings, The Twenty-First National Conference on Artificial Intelligence and the Eighteenth Innovative Applications of Artificial Intelligence Conference, July 16-20, 2006, Boston, Massachusetts, USA*. [S.l.]: AAAI Press, 2006. p. 401–408.

LIU, J. et al. Interclass interference suppression in multi-class problems. *Applied Sciences*, v. 11, n. 1, 2021. ISSN 2076-3417.

LOGAN, B. Mel frequency cepstral coefficients for music modeling. *Proc. 1st Int. Symposium Music Information Retrieval*, 11 2000.

LOUGHRAN, R. et al. Comparison of features in musical instrument identification using artificial neural networks. In: *5th International Symposium on Computer Music Modeling and Retrieval*. [S.l.]: Springer, 2008. p. 19–33.

LOUGHRAN, R. et al. *The use of mel-frequency cepstral coefficients in musical instrument identification*. [S.l.]: International Computer Music Association, 2018.

MALÍK, M.; ORJEŠEK, R. The comparison of selected audio features and classification techniques in the task of the musical instrument. In: . [S.l.: s.n.], 2016.

MATLAB. *Extract cepstral features from audio segment*. 2021. Acesso em: 28 de Abr. de 2021. Disponível em: <<https://www.mathworks.com/help/audio/ref/cepstralfeatureextractor-system-object.html>>.

MCFEE, B. et al. *librosa: Audio and music signal analysis in python*. In: *Proceedings of the 14th python in science conference*. [S.l.: s.n.], 2015. v. 8.

MCKINNEY Wes. Data Structures for Statistical Computing in Python. In: WALT Stéfan van der; MILLMAN Jarrod (Ed.). *Proceedings of the 9th Python in Science Conference*. [S.l.: s.n.], 2010. p. 56 – 61.

MORTAZAVI, M.; OUDBASHI, O. Influence of heat treatment in formation of martensitic structure in historical bronzes. In: . [S.l.: s.n.], 2010.

MUDA, L.; BEGAM, M.; ELAMVAZUTHI, I. Voice recognition algorithms using mel frequency cepstral coefficient (MFCC) and dynamic time warping (DTW) techniques. *CoRR*, abs/1003.4083, 2010.

MÜLLER, M. *Fundamentals of Music Processing*. [S.l.]: Springer International Publishing, 2015.

NAGAWADE, M. S.; RATNAPARKHE, V. R. Musical instrument identification using mfcc. In: *2017 2nd IEEE International Conference on Recent Trends in Electronics, Information Communication Technology (RTEICT)*. [S.l.: s.n.], 2017. p. 2198–2202.

NETTLETON, D. Chapter 6 - selection of variables and factor derivation. In: NETTLETON, D. (Ed.). *Commercial Data Mining*. Boston: Morgan Kaufmann, 2014. p. 79–104. ISBN 978-0-12-416602-8.

NG, A. Y. Feature selection, l1 vs. l2 regularization, and rotational invariance. In: *Proceedings of the Twenty-First International Conference on Machine Learning*. New York, NY, USA: Association for Computing Machinery, 2004. (ICML '04), p. 78. ISBN 1581138385.

NGUYEN, Q.; TOUZÉ, C. Nonlinear vibrations of thin plates with variable thickness: Application to sound synthesis of cymbals. *The Journal of the Acoustical Society of America*, v. 145, p. 977–988, 02 2019.

NOOR, M. H. M. Feature learning using convolutional denoising autoencoder for activity recognition. *Neural Computing and Applications*, v. 33, p. 1–14, 09 2021.

OGAWA, W. et al. Effect of bell size on sound characteristics of cymbals. *Transactions of the JSME (in Japanese)*, v. 86, n. 881, p. 19–00237–19–00237, 2020.

OPPENHEIM, A.; SCHAFER, R. From frequency to quefrequency: a history of the cepstrum. *IEEE Signal Processing Magazine*, v. 21, n. 5, p. 95–106, 2004.

OPPENHEIM, A. V.; SCHAFER, R. W. *Discrete-Time Signal Processing*. 3rd. ed. USA: Prentice Hall Press, 2009. ISBN 0131988425.

OSAMURA, K. et al. The correlation between the percussive sound and the residual stress/strain distributions in a cymbal. *Journal of Materials Engineering and Performance*, v. 25, 10 2016.

OSTERTAGOVA, E.; OSTERTAG, O. Methodology and application of one-way anova. *American Journal of Mechanical Engineering*, v. 1, p. 256–261, 11 2013.

PAISTE. *Cymbal anatomy*. 2022. Acesso em: 16 de fev. de 2022. Disponível em: <<https://www.paiste.com/en/about/everything-cymbals/cymbal-anatomy>>.

PEDREGOSA, F. et al. Scikit-learn: Machine learning in python. *The Journal of Machine Learning Research*, JMLR. org, v. 12, p. 2825–2830, 2011.

PERRIN, R. et al. Normal modes of cymbals. *Proceedings of the Institute of Acoustics*, 2008.

- PINKSTERBOER, H.; MATTINGLY, R. *The Cymbal Book*. [S.l.]: Hal Leonard Publishing Corporation, 1992. (Musical Instruments Series). ISBN 9780793519200.
- RATHIKARANI, V.; DHANALAKSHMI, P.; S., P. Classification of musical instruments using svm and knn. p. 2278–3075, 05 2020.
- SAWYER, S. Analysis of variance: The fundamental concepts. *Journal of Manual & Manipulative Therapy*, v. 17, p. 27E–38E, 04 2009.
- SCHOBBER, P.; BOER, C.; SCHWARTE, L. A. Correlation coefficients: appropriate use and interpretation. *Anesthesia & Analgesia*, Wolters Kluwer, v. 126, n. 5, p. 1763–1768, 2018.
- SCHROEDER, M. et al. *Springer handbook of acoustics*. 1. ed. [S.l.]: Springer, 2007. ISBN 0387304460; 9780387304465.
- SCOTT, D. A. Metallography and microstructure of ancient and historic metals. *Studies in Conservation*, v. 37, p. 282, 1992.
- SHRESTHA, A.; MAHMOOD, A. Review of deep learning algorithms and architectures. *IEEE access*, IEEE, v. 7, p. 53040–53065, 2019.
- SHURE. *PGA81 - Microfone condensador cardioide para instrumento*. 2021. Acesso em: 02 de Abr. de 2021. Disponível em: <<https://www.shure.com/pt-BR/produtos/microfones/pga81>>.
- SLAMET, S. et al. Effect of high-tin bronze composition on physical, mechanical, and acoustic properties of gamelan materials. 08 2021.
- SLANEY, M. *Auditory Toolbox Version 2*. [S.l.], 1998.
- SOUZA, V. M. A.; BATISTA, G. E. A. P. A.; SOUZA-FILHO, N. E. Automatic classification of drum sounds with indefinite pitch. In: *Proceedings of International Joint Conference on Neural Networks (IJCNN)*. [S.l.: s.n.], 2015.
- STEEN, A. *11 types of cymbals: Everything you need to know*. 2022. Disponível em: <<https://primesound.org/types-of-cymbals/>>.
- SUGITA, I. K. G.; PRIAMBADI, I. G. N. The study of dendrite arm spacing (das) on acoustical of tin bronze 20sn alloy as gamelan bali materials. In: *Engineering and Innovative Materials V*. [S.l.]: Trans Tech Publications Ltd, 2017. (Materials Science Forum, v. 889), p. 133–137.
- SUGITA, I. K. G.; SOEKRISNO, R.; SUYITNO, S. Mechanical and damping properties of silicon bronze alloys for music applications. *Int J Eng & Technol IJET-IJENS*, v. 11, p. 98–105, 01 2011.
- TIANO, D.; BONIFATI, A.; NG, R. Featts: Feature-based time series clustering. In: _____. *Proceedings of the 2021 International Conference on Management of Data*. New York, NY, USA: Association for Computing Machinery, 2021. p. 2784–2788. ISBN 9781450383431.
- VIEIRA, R. et al. Automatic classification of instruments from supervised methods of machine learning. In: *Anais do XVIII Simpósio Brasileiro de Computação Musical*. Porto Alegre, RS, Brasil: SBC, 2021. p. 1–7. ISSN 0000-0000.

VIRTANEN, P. et al. SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python. *Nature Methods*, v. 17, p. 261–272, 2020.

WASKOM, M. L. seaborn: statistical data visualization. *Journal of Open Source Software*, The Open Journal, v. 6, n. 60, p. 3021, 2021.

WHITE, M. A.; MACMILLAN, P. The cymbal as an instructional device for materials education. *MRS Proceedings*, Cambridge University Press, v. 760, p. JJ1.6, 2002.

ZADEH, L. Fuzzy sets. *Information and Control*, v. 8, n. 3, p. 338–353, 1965. ISSN 0019-9958.

**APÊNDICE A – DESEMPENHOS ALCANÇADOS EM CADA CASO
CONSIDERADO**

Tabela 8 – Desempenhos médios obtidos para o conjunto de dados C1, considerando todas as técnicas de seleção e os dois classificadores implementados.

Seleção	Classificador	Acurácia	Precisão	Recall	F1
Sem Seleção	Regressão Logística	0.9470	0.9523	0.9442	0.9468
	Fuzzy - Triangular	0.7590	0.7499	0.7189	0.7100
	Fuzzy - Gaussiana	0.7711	0.7429	0.7358	0.7321
T1	Regressão Logística	0.9465	0.9516	0.9437	0.9462
	Fuzzy - Triangular	0.7628	0.7945	0.7214	0.7164
	Fuzzy - Gaussiana	0.8098	0.8027	0.7822	0.7803
T2	Regressão Logística	0.9716	0.9737	0.9696	0.9708
	Fuzzy - Triangular	0.7937	0.8031	0.7672	0.7593
	Fuzzy - Gaussiana	0.8730	0.8762	0.8613	0.8644
T3	Regressão Logística	0.9708	0.9731	0.9688	0.9701
	Fuzzy - Triangular	0.7724	0.8103	0.7454	0.7404
	Fuzzy - Gaussiana	0.8246	0.8184	0.8081	0.8048

Tabela 9 – Desempenhos médios obtidos para o conjunto de dados C2, considerando todas as técnicas de seleção e os dois classificadores implementados.

Seleção	Classificador	Acurácia	Precisão	Recall	F1
Sem Seleção	Regressão Logística	0.9852	0.9898	0.9797	0.9841
	Fuzzy - Triangular	0.9404	0.9499	0.9324	0.9387
	Fuzzy - Gaussiana	0.9472	0.9613	0.9371	0.9466
T1	Regressão Logística	0.9847	0.9898	0.9785	0.9835
	Fuzzy - Triangular	0.8577	0.8580	0.8476	0.8377
	Fuzzy - Gaussiana	0.8708	0.8631	0.8607	0.8564
T2	Regressão Logística	0.9852	0.9898	0.9797	0.9841
	Fuzzy - Triangular	0.9404	0.9499	0.9324	0.9387
	Fuzzy - Gaussiana	0.9472	0.9613	0.9371	0.9466
T3	Regressão Logística	0.9855	0.9902	0.9802	0.9846
	Fuzzy - Triangular	0.8953	0.8864	0.8851	0.8761
	Fuzzy - Gaussiana	0.8727	0.8726	0.8529	0.8434

Tabela 10 – Desempenhos médios obtidos para o conjunto de dados C3, considerando todas as técnicas de seleção e os dois classificadores implementados.

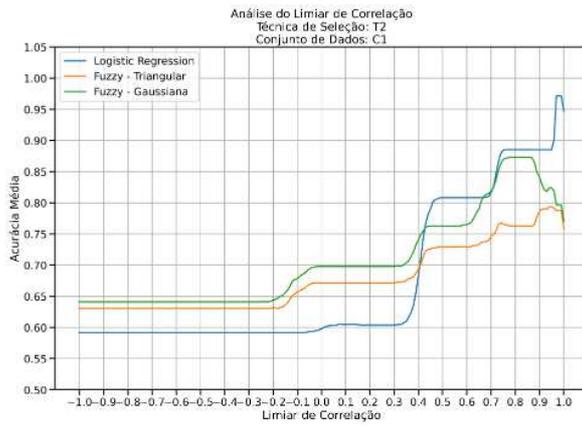
Seleção	Classificador	Acurácia	Precisão	Recall	F1
Sem Seleção	Regressão Logística	0.9893	0.9885	0.9898	0.9887
	Fuzzy - Triangular	0.8290	0.8420	0.8019	0.7991
	Fuzzy - Gaussiana	0.8492	0.8488	0.8272	0.8263
T1	Regressão Logística	0.9836	0.9842	0.9833	0.9832
	Fuzzy - Triangular	0.8287	0.8531	0.7966	0.7970
	Fuzzy - Gaussiana	0.8661	0.8633	0.8460	0.8444
T2	Regressão Logística	0.9975	0.9969	0.9966	0.9966
	Fuzzy - Triangular	0.9028	0.9168	0.8921	0.8991
	Fuzzy - Gaussiana	0.9359	0.9517	0.9248	0.9350
T3	Regressão Logística	0.9937	0.9926	0.9941	0.9931
	Fuzzy - Triangular	0.8830	0.8988	0.8644	0.8682
	Fuzzy - Gaussiana	0.9108	0.9133	0.8962	0.8987

Tabela 11 – Desempenhos médios obtidos para o conjunto de dados C4, considerando todas as técnicas de seleção e os dois classificadores implementados.

Seleção	Classificador	Acurácia	Precisão	Recall	F1
Sem Seleção	Regressão Logística	0.9917	0.9944	0.9885	0.9912
	Fuzzy - Triangular	0.9961	0.9974	0.9948	0.9960
	Fuzzy - Gaussiana	0.9929	0.9952	0.9902	0.9925
T1	Regressão Logística	0.9911	0.9940	0.9877	0.9905
	Fuzzy - Triangular	0.8860	0.8842	0.8638	0.8557
	Fuzzy - Gaussiana	0.8580	0.8704	0.8264	0.8116
T2	Regressão Logística	0.9923	0.9948	0.9894	0.9918
	Fuzzy - Triangular	0.9961	0.9974	0.9948	0.9960
	Fuzzy - Gaussiana	0.9929	0.9952	0.9902	0.9925
T3	Regressão Logística	0.9911	0.9940	0.9877	0.9905
	Fuzzy - Triangular	0.8953	0.8864	0.8851	0.8761
	Fuzzy - Gaussiana	0.8727	0.8726	0.8529	0.8434

Tabela 12 – Desempenhos médios obtidos para o conjunto de dados C5, considerando todas as técnicas de seleção e os dois classificadores implementados.

Seleção	Classificador	Acurácia	Precisão	Recall	F1
Sem Seleção	Regressão Logística	0.9925	0.9913	0.9931	0.9919
	Fuzzy - Triangular	0.9639	0.9757	0.9552	0.9636
	Fuzzy - Gaussiana	0.9742	0.9830	0.9671	0.9740
T1	Regressão Logística	0.9935	0.9924	0.9938	0.9928
	Fuzzy - Triangular	0.8061	0.8327	0.7702	0.7606
	Fuzzy - Gaussiana	0.8613	0.8639	0.8376	0.8345
T2	Regressão Logística	0.9976	0.9970	0.9967	0.9967
	Fuzzy - Triangular	0.9916	0.9944	0.9886	0.9912
	Fuzzy - Gaussiana	0.9914	0.9942	0.9885	0.9911
T3	Regressão Logística	0.9937	0.9926	0.9941	0.9931
	Fuzzy - Triangular	0.8517	0.8718	0.8263	0.8226
	Fuzzy - Gaussiana	0.9049	0.8984	0.8917	0.8888

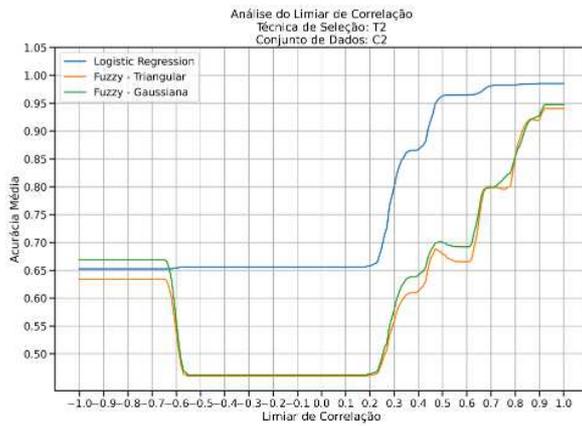


(a) Técnica de seleção T2.

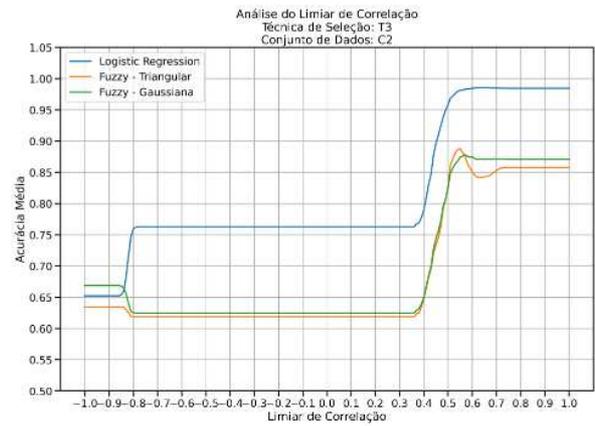


(b) Técnica de seleção T3.

Figura 38 – Comportamento das acurácias em relação ao limiar de correlação para o conjunto de dados C1.

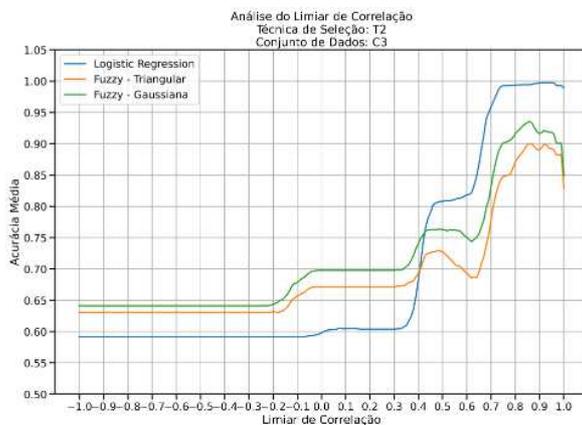


(a) Técnica de seleção T2.

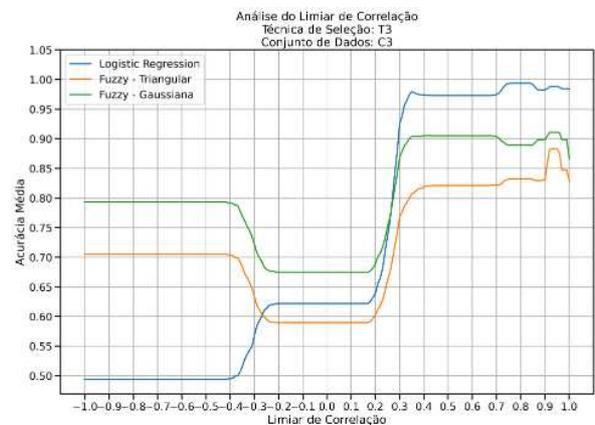


(b) Técnica de seleção T3.

Figura 39 – Comportamento das acurácias em relação ao limiar de correlação para o conjunto de dados C2.

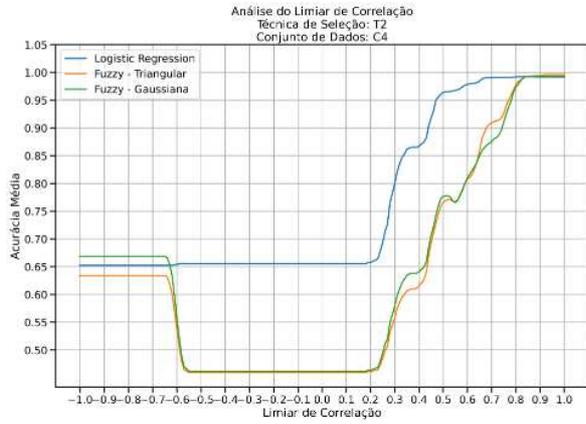


(a) Técnica de seleção T2.

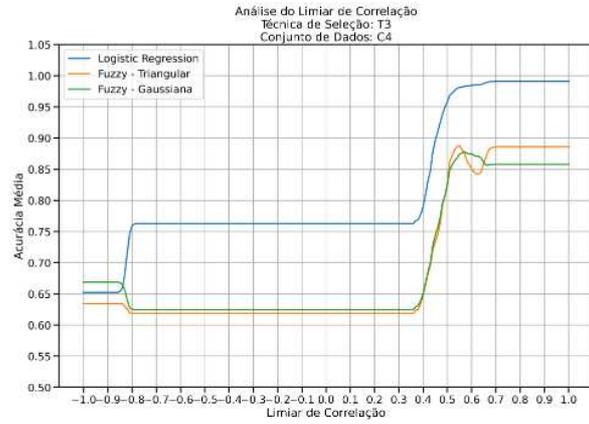


(b) Técnica de seleção T3.

Figura 40 – Comportamento das acurácias em relação ao limiar de correlação para o conjunto de dados C3.

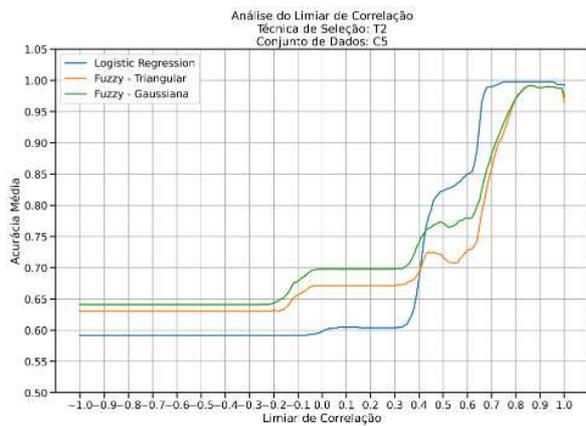


(a) Técnica de seleção T2.

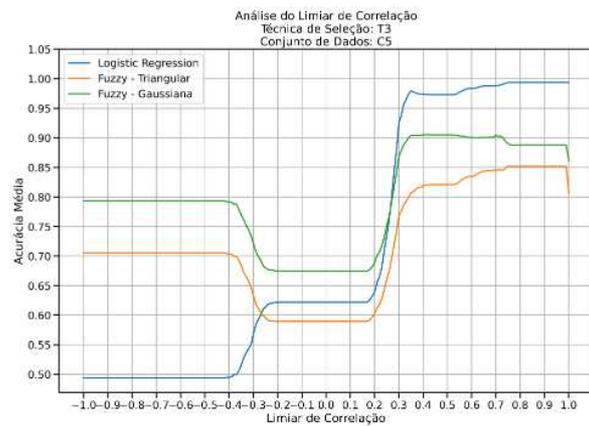


(b) Técnica de seleção T3.

Figura 41 – Comportamento das acurácias em relação ao limiar de correlação para o conjunto de dados C4.



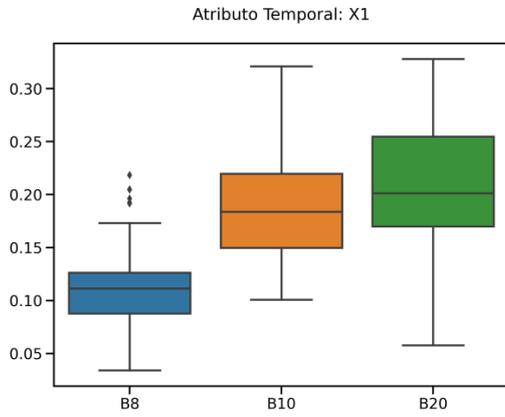
(a) Técnica de seleção T2.



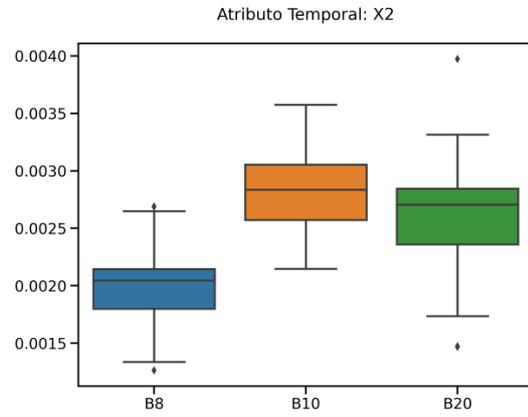
(b) Técnica de seleção T3.

Figura 42 – Comportamento das acurácias médias em relação ao limiar de correlação para o conjunto de dados C5.

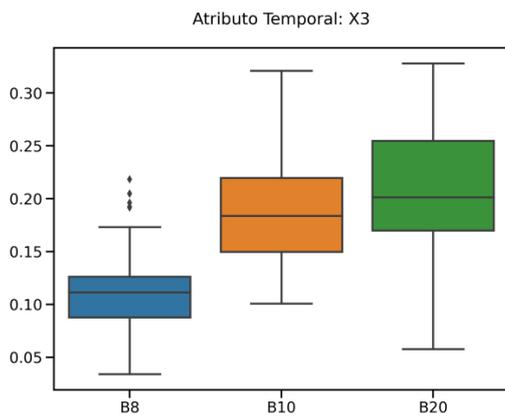
APÊNDICE B – DISTRIBUIÇÃO DOS VALORES DE CADA ATRIBUTO EXTRAÍDO



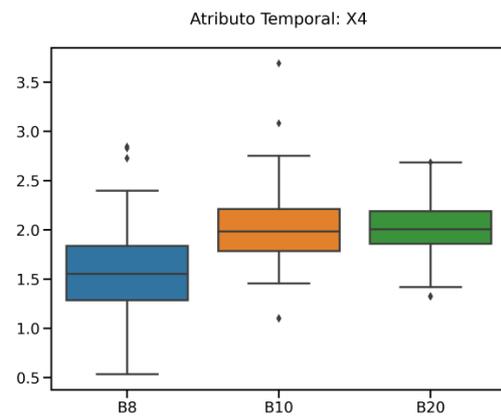
(a) Energia absoluta.



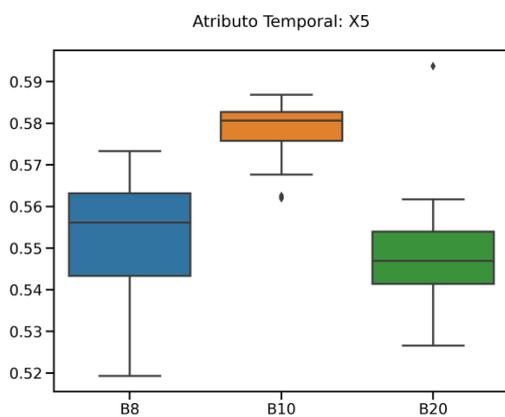
(b) Área sob a curva.



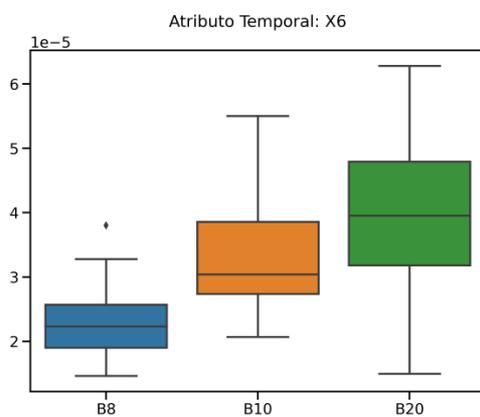
(c) Autocorrelação.



(d) Centroide ao longo do eixo do tempo.

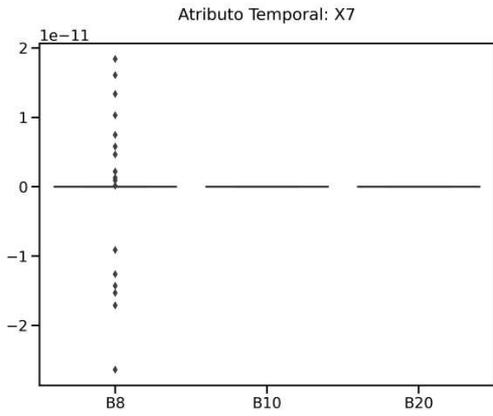


(e) Entropia.

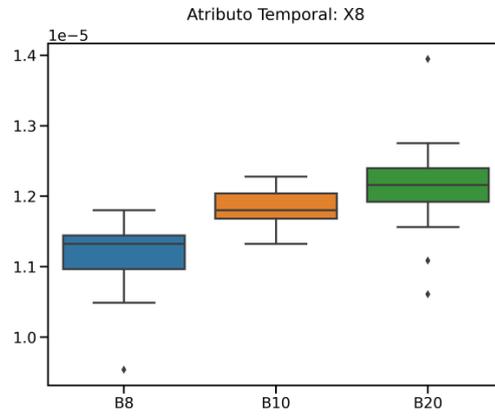


(f) Média do sinal derivado absoluto.

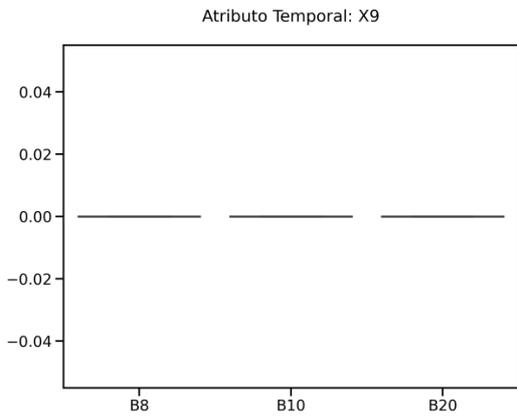
Figura 43 – Gráfico de caixas ilustrando a distribuição dos valores dos atributos temporais X1 a X6.



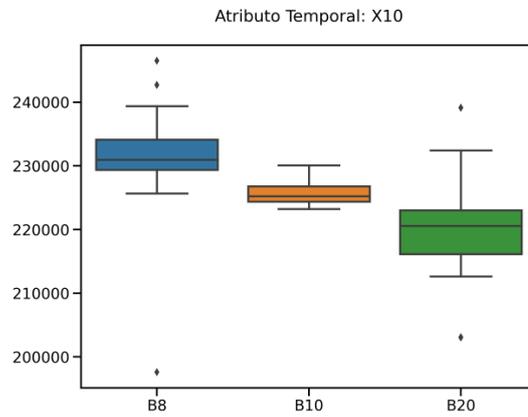
(a) Média do sinal derivado.



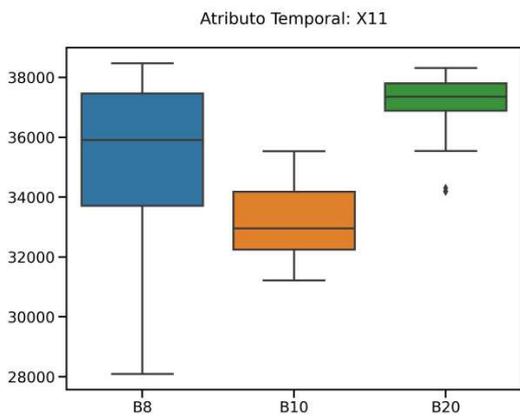
(b) Mediana do sinal derivado absoluto.



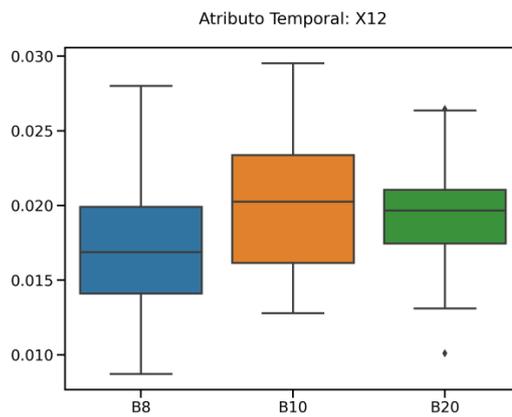
(c) Mediana do sinal derivado.



(d) Número de pontos de viragem negativos.

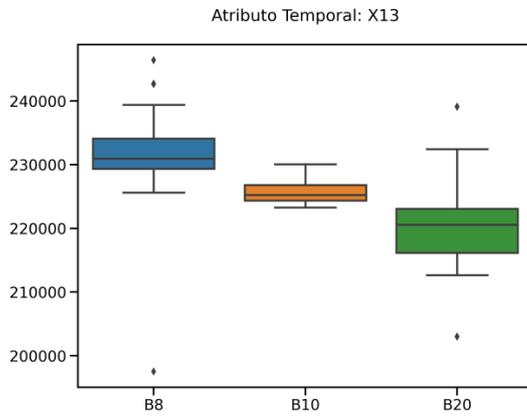


(e) Número de picos a partir de uma vizinhança definida.

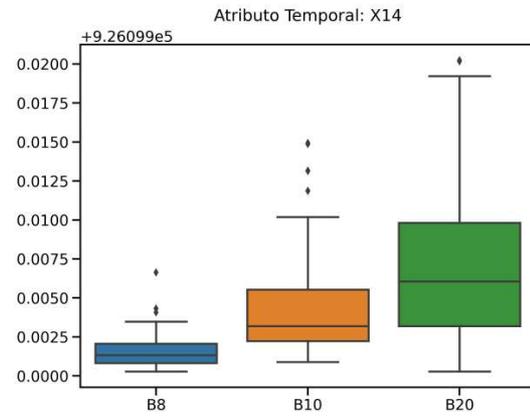


(f) Amplitude pico a pico.

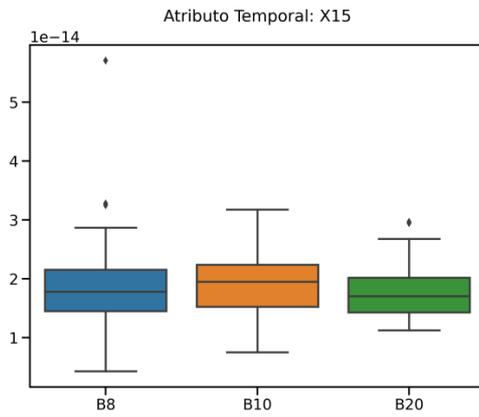
Figura 44 – Gráfico de caixas ilustrando a distribuição dos valores dos atributos temporais X7 a X12.



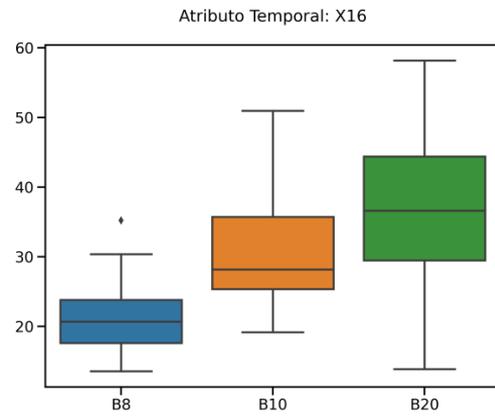
(a) Número de pontos de viragem positivos.



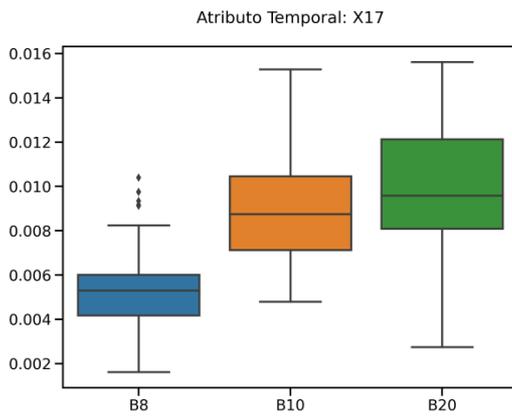
(b) Distância percorrida pelo sinal.



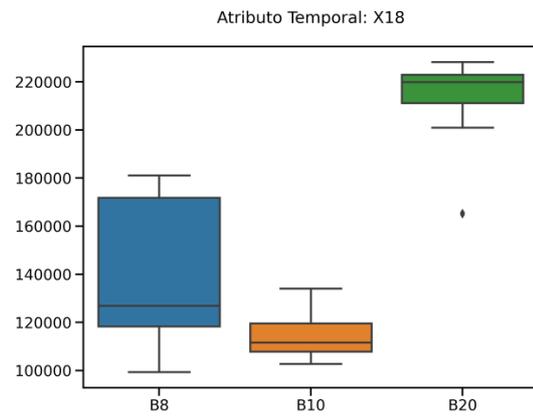
(c) Inclinação.



(d) Soma dos incrementos absolutos do sinal.

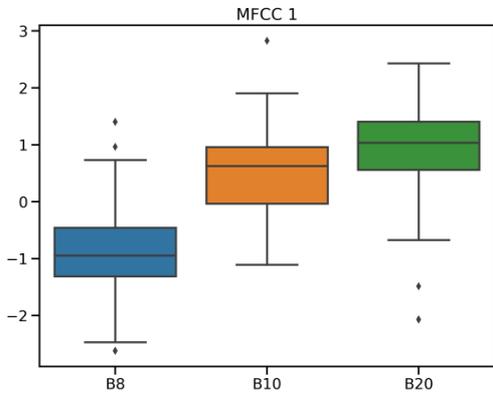


(e) Energia total do sinal.

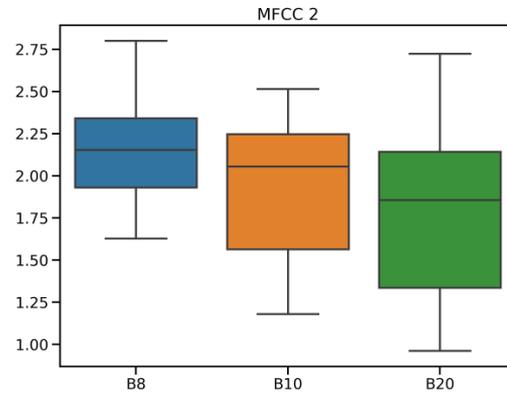


(f) Taxa de cruzamento do zero.

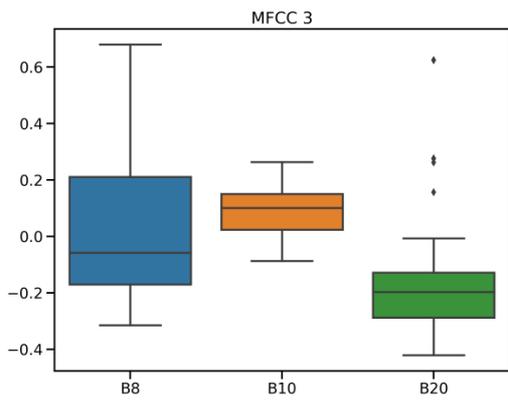
Figura 45 – Gráfico de caixas ilustrando a distribuição dos valores dos atributos temporais X13 a X18.



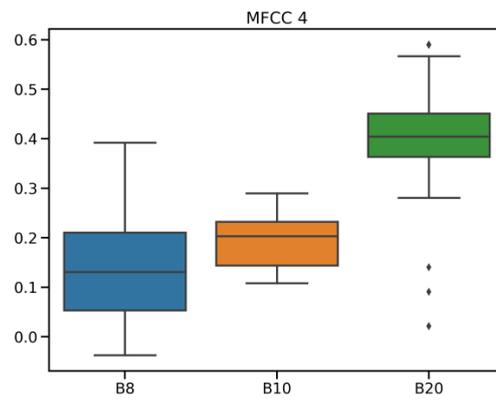
(a) MFCC 1.



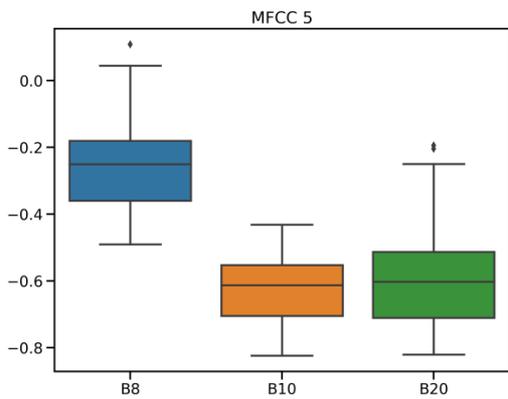
(b) MFCC 2.



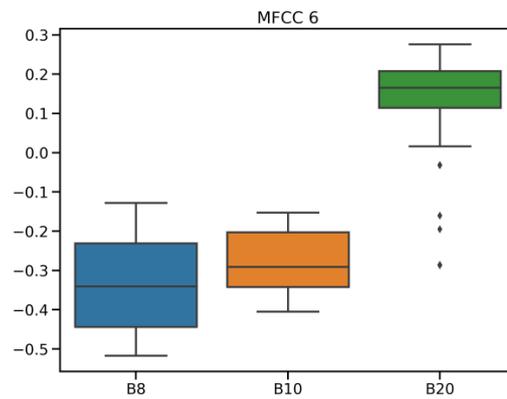
(c) MFCC 3.



(d) MFCC 4.

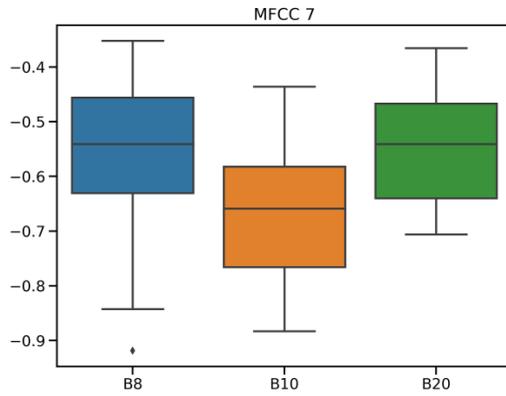


(e) MFCC 5.

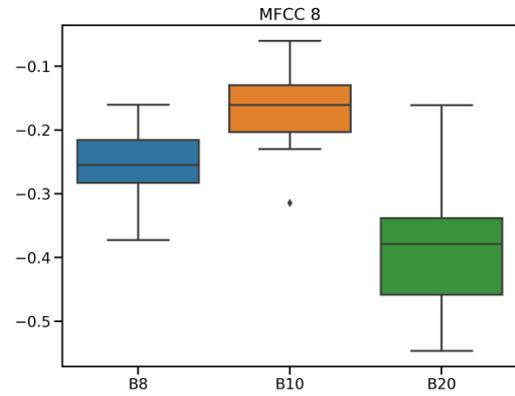


(f) MFCC 6.

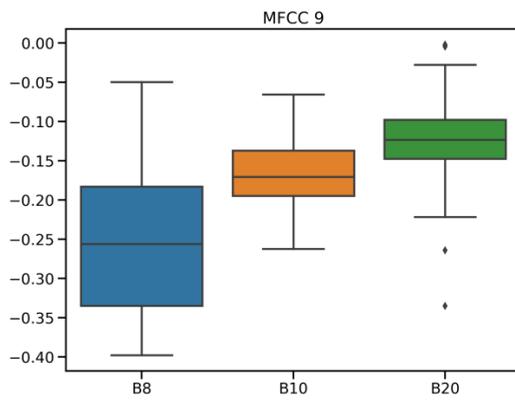
Figura 46 – Gráfico de caixas ilustrando a distribuição dos valores dos MFCCs 1 a 6.



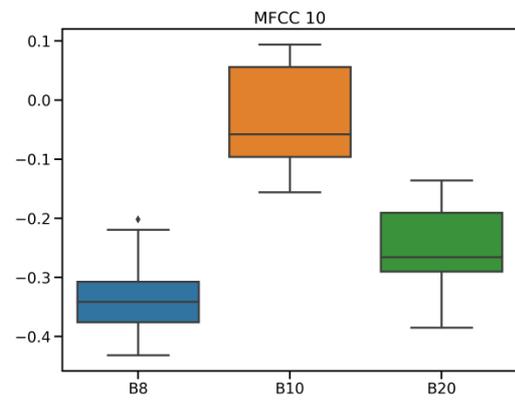
(a) MFCC 7.



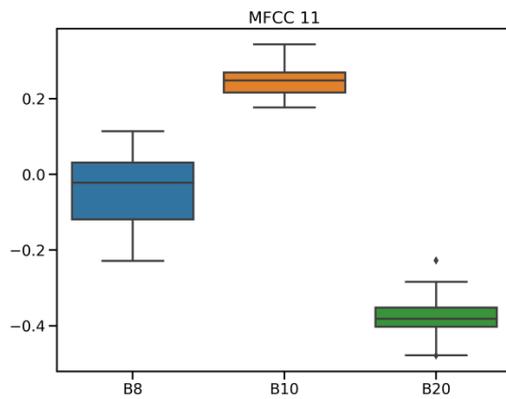
(b) MFCC 8.



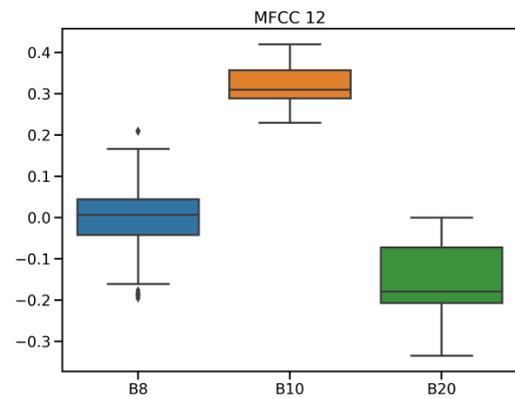
(c) MFCC 9.



(d) MFCC 10.

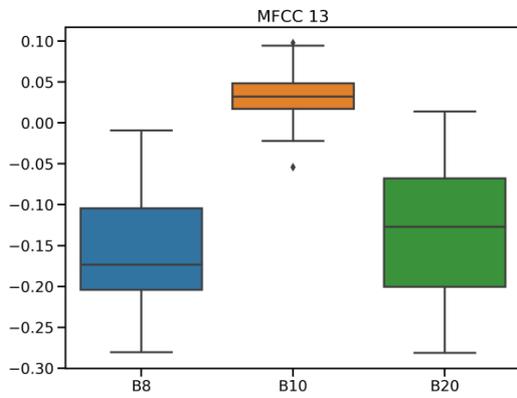


(e) MFCC 11.

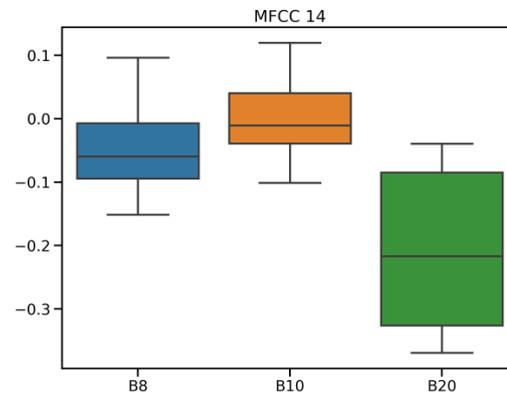


(f) MFCC 12.

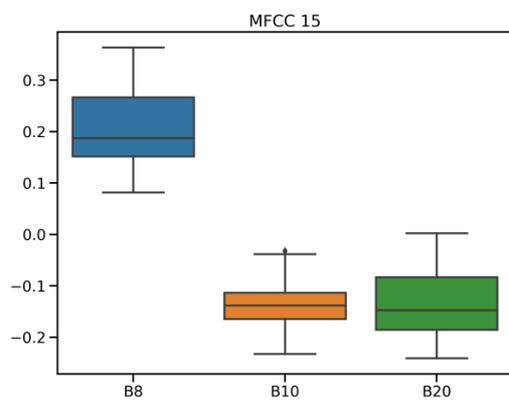
Figura 47 – Gráfico de caixas ilustrando a distribuição dos valores dos MFCCs 7 a 12.



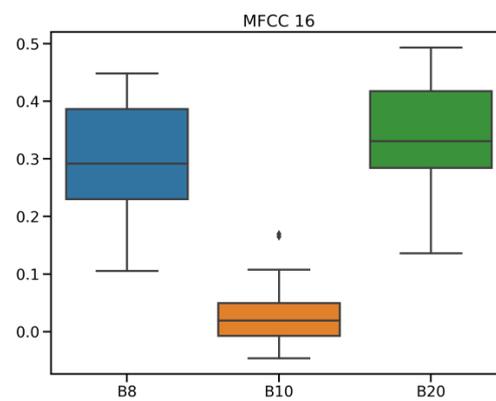
(a) MFCC 13.



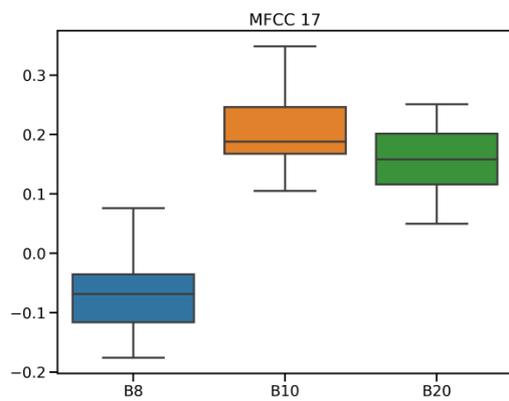
(b) MFCC 14.



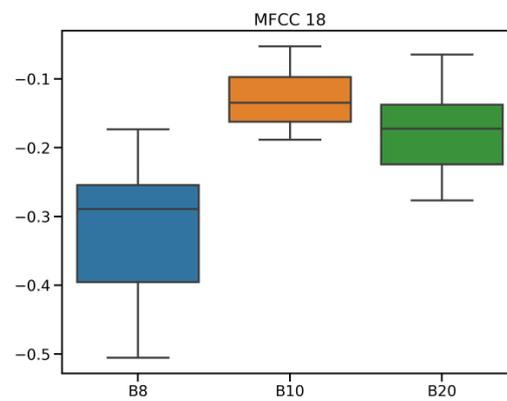
(c) MFCC 15.



(d) MFCC 16.

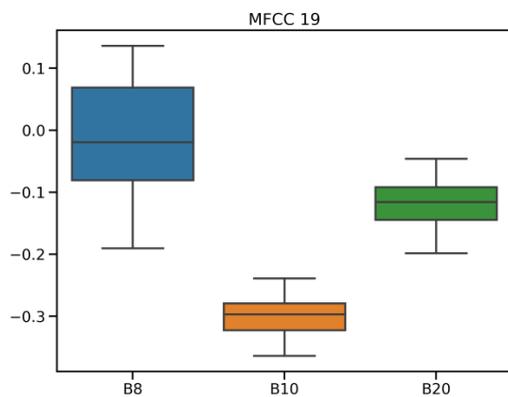


(e) MFCC 17.

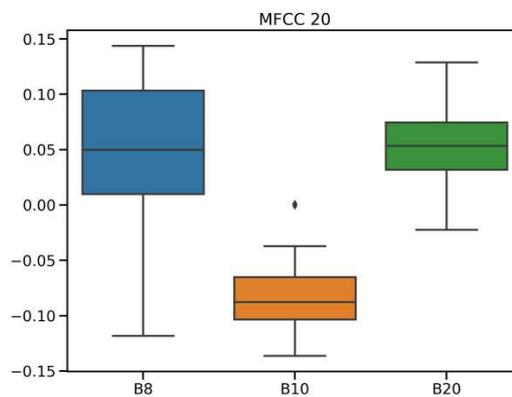


(f) MFCC 18.

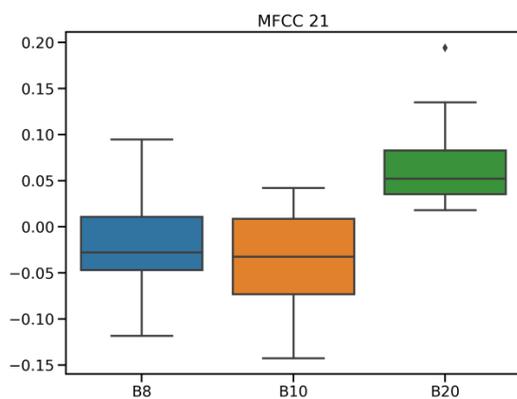
Figura 48 – Gráfico de caixas ilustrando a distribuição dos valores dos MFCCs 13 a 18.



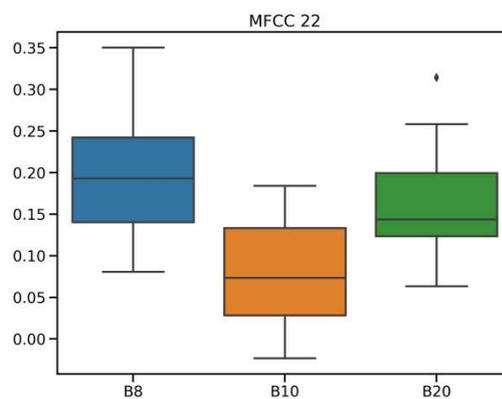
(a) MFCC 19.



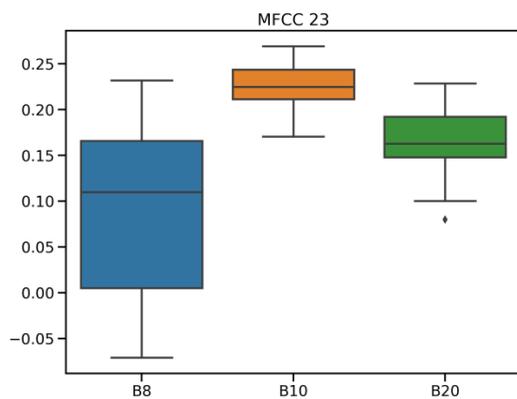
(b) MFCC 20.



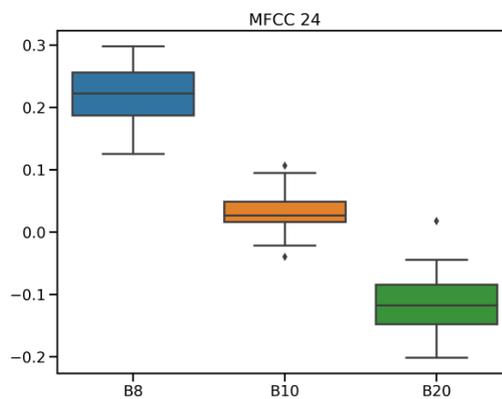
(c) MFCC 21.



(d) MFCC 22.

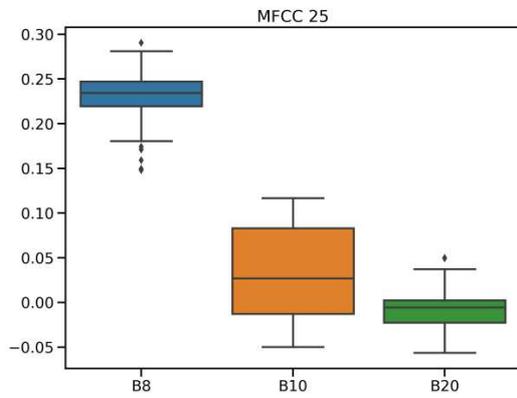


(e) MFCC 23.

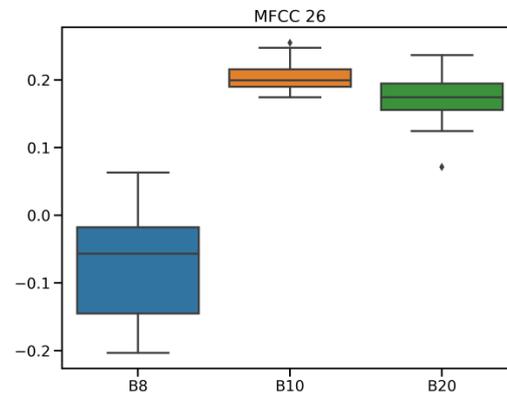


(f) MFCC 24.

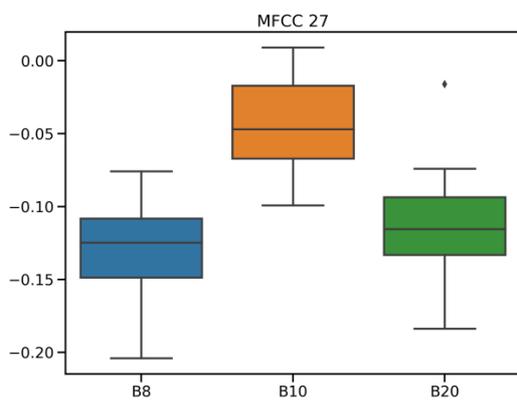
Figura 49 – Gráfico de caixas ilustrando a distribuição dos valores dos MFCCs 19 a 24.



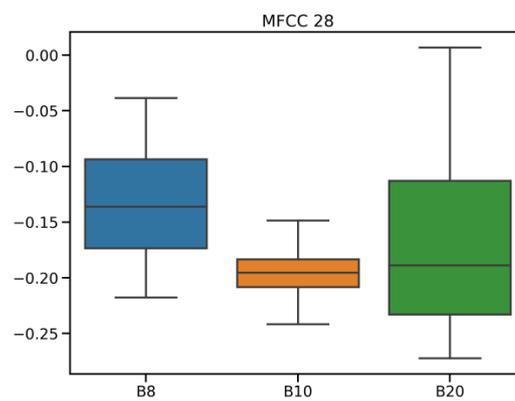
(a) MFCC 25.



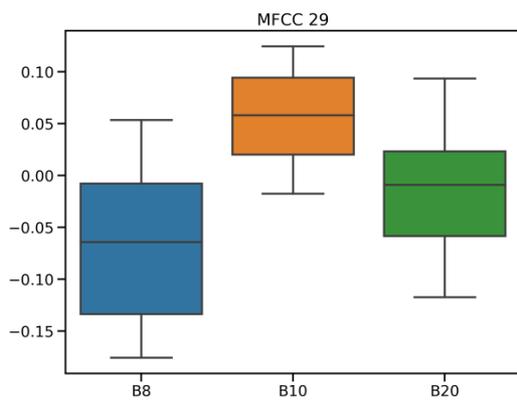
(b) MFCC 26.



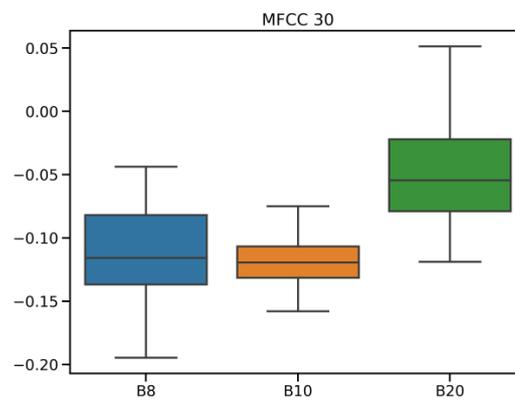
(c) MFCC 27.



(d) MFCC 28.

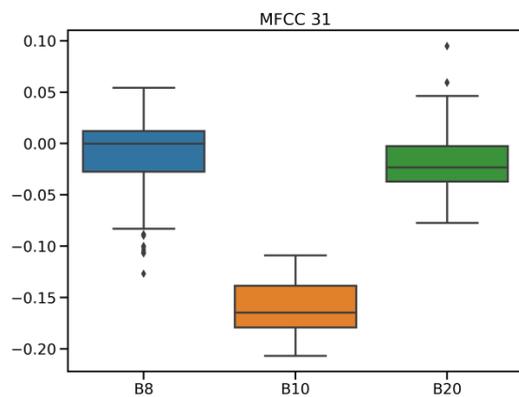


(e) MFCC 29.

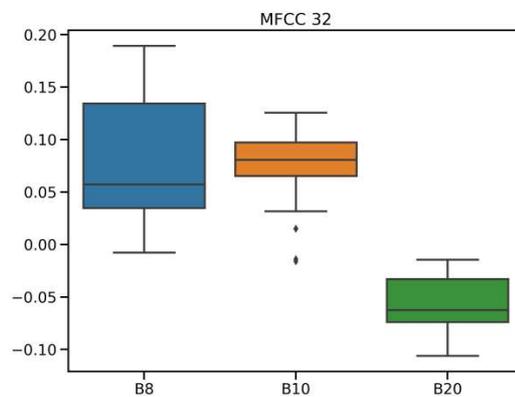


(f) MFCC 30.

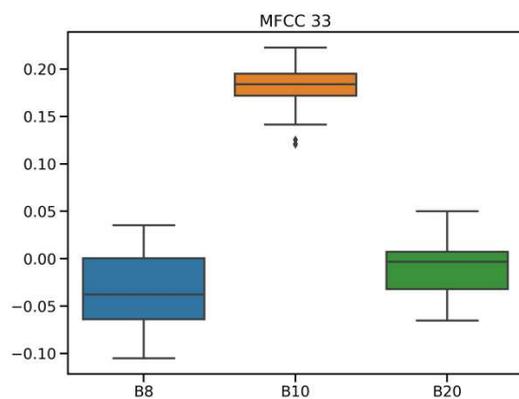
Figura 50 – Gráfico de caixas ilustrando a distribuição dos valores dos MFCCs 25 a 30.



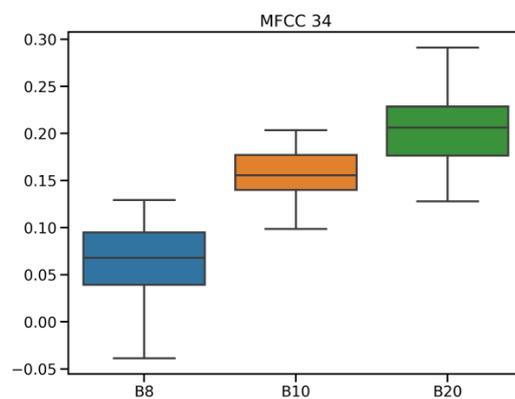
(a) MFCC 31.



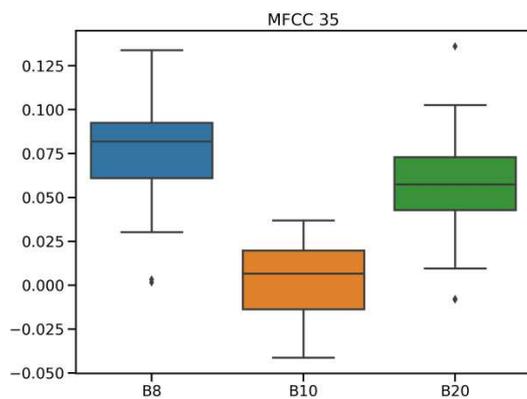
(b) MFCC 32.



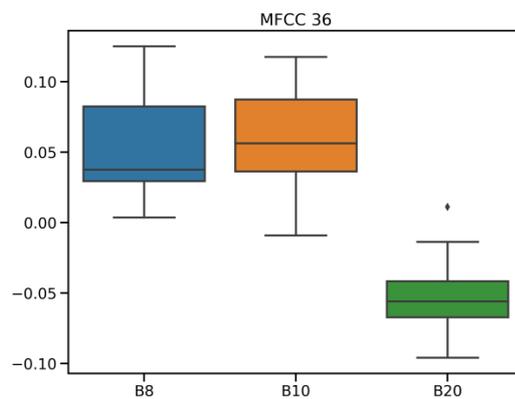
(c) MFCC 33.



(d) MFCC 34.

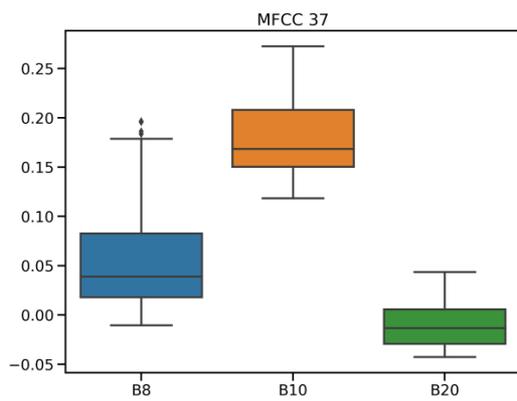


(e) MFCC 35.

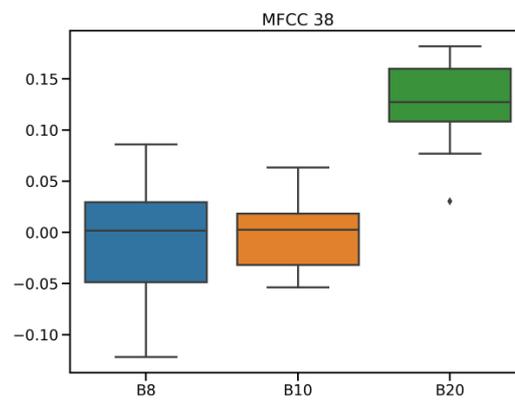


(f) MFCC 36.

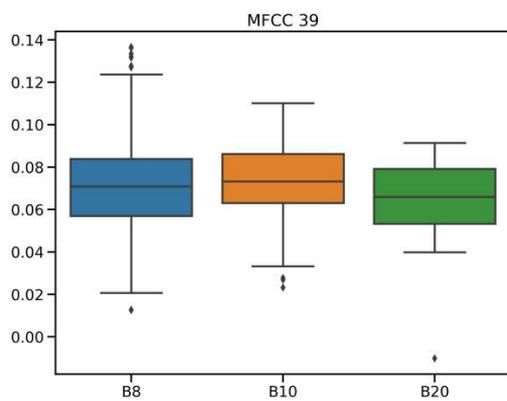
Figura 51 – Gráfico de caixas ilustrando a distribuição dos valores dos MFCCs 31 a 36.



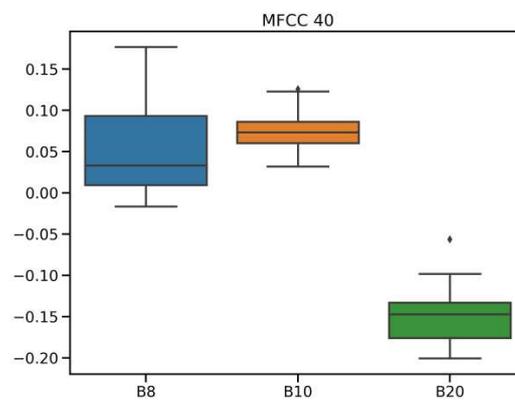
(a) MFCC 37.



(b) MFCC 38.



(c) MFCC 39.



(d) MFCC 40.

Figura 52 – Gráfico de caixas ilustrando a distribuição dos valores dos MFCCs 37 a 40.