



UNIVERSIDADE FEDERAL DA BAHIA
ESCOLA POLITÉCNICA
PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA ELÉTRICA

CAMILLA BAHIA LARocca

DISSERTAÇÃO DE MESTRADO EM ENGENHARIA ELÉTRICA

**ENSEMBLE DE CLASSIFICADORES PARA AVALIAÇÃO
DE CHAPAS DE AÇO IF UTILIZANDO CORRENTES
PARASITAS PULSADAS**

Salvador/BA - Brasil

Outubro de 2022

CAMILLA BAHIA LARocca

**ENSEMBLE DE CLASSIFICADORES PARA AVALIAÇÃO DE
CHAPAS DE AÇO IF UTILIZANDO CORRENTES PARASITAS
PULSADAS**

DISSERTAÇÃO DE MESTRADO EM ENGENHARIA ELÉTRICA

Dissertação de Mestrado apresentada ao Programa de Pós-graduação em Engenharia Elétrica da Universidade Federal da Bahia como parte dos requisitos necessários para obtenção do título de Mestre em Engenharia Elétrica.

Orientador: Dr. Eduardo Furtado de Simas Filho
– UFBA

Coorientador: Dr. Ivan Costa da Silva – IFBA

Salvador/BA - Brasil

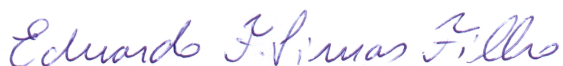
Outubro de 2022

ENSEMBLE DE CLASSIFICADORES PARA AVALIAÇÃO DE CHAPAS DE AÇO IF UTILIZANDO CORRENTES PARASITAS PULSADAS

CAMILLA BAHIA LAROCCA

Dissertação submetida ao corpo docente do programa de pós-graduação em engenharia elétrica da Universidade Federal da Bahia (PPGEE / UFBA) como parte dos requisitos necessários para obtenção do título de mestre em engenharia elétrica.

Salvador/BA - Brasil, 17 de outubro de 2022:



Eduardo Furtado de Simas Filho, D.Sc.
Orientador (UFBA)



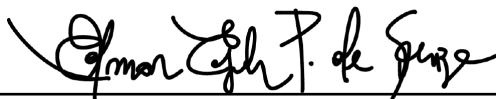
Ivan Costa da Silva, D.Sc.
Coorientador (IFBA)



Paulo Cesar Machado de Abreu Farias, D.Sc.
Avaliador Interno (UFBA)



Elineudo Pinho de Moura, D.Sc.
Avaliador Externo (UFC)



Edmar Egidio Purcino de Souza, D.Sc.
Avaliador Externo (Sorbonne Université)

Salvador/BA - Brasil
Outubro de 2022

Agradecimentos

Agradeço primeiramente ao meu orientador pela sua dedicação e paciência durante o projeto, dando todo o auxílio necessário e me ajudando em cada passo deste processo.

Aos mentores que encontrei durante a minha caminhada, que me acompanharam e me incentivaram em momentos importantes da minha vida. À Filipe Aguiar, que me ensinou que estudar não é o martírio que eu pensava. Nunca imaginei que a menina que precisava de ajuda com funções de segundo grau e vetores chegaria aqui. Ao professor Ivan C. da Silva e à professora Claudia T. T. Farias, por me ensinarem o quão instigante é a pesquisa acadêmica. Sempre dedicarei a vocês o momento no qual eu encontrei o meu caminho.

À minha mãe Maria do Carmo Vídero Bahia Larocca, meu porto seguro, que me ensinou que desistir não é uma opção e à meu pai, Angelo Grimaldi Larocca, meu exemplo e meu *nerd* favorito, que me ensinou que nunca é tarde demais para estudar um assunto novo. À minha família, por me incentivarem a cada momento, pela confiança no meu progresso e apoio emocional.

Ao meu companheiro, André Lins, por sempre acreditar em mim, mesmo quando nem eu mesma acreditava.

Agradeço também a todos meus amigos, sempre ao meu lado para dividir os momentos de alegria e angústia.

Obrigado por acreditarem em mim.

“Para os demais, podemos dizer que Muad’Dib aprendeu rápido porque primeiro lhe ensinaram como aprender. E a primeira lição de todas foi desenvolver a confiança fundamental de que ele era capaz de aprender. É surpreendente saber quantas pessoas acreditam não ser capazes de aprender e quantas outras creem que aprender é difícil. Muad’Dib sabia que toda experiência era uma lição.”

Frank Herbert, *Duna*

Resumo

ENSEMBLE DE CLASSIFICADORES PARA AVALIAÇÃO DE CHAPAS DE AÇO IF UTILIZANDO CORRENTES PARASITAS PULSADAS

Camilla Bahia Larocca

Outubro/2022

Orientadores: Eduardo Furtado de Simas Filho, D.Sc
Ivan Costa da Silva, D.Sc

Programa: Engenharia Elétrica

Os aços livres de elementos intersticiais (do inglês, *Intertitial Free* – IF) são empregados na fabricação de inúmeras peças de formatos complexos, especialmente utilizados na estamparia do chassi na indústria automobilística, em função de sua elevada ductilidade e boa resistência mecânica. O transporte do material sob a forma de bobinas, todavia, sujeita-o à cargas do próprio peso. Pode sofrer com tensões residuais geradas e, como consequência, as propriedades mecânicas do aço se tornam susceptíveis a alterações. Visto que o material é submetido a equipamentos de estampagem com parâmetros pré-estabelecidos para condições desejadas, identificar tensões residuais fora do intervalo de tolerância é um importante passo na prevenção de defeitos das peças fabricadas. Nesta perspectiva, o presente trabalho se propôs a apresentar um conjunto de técnicas inteligentes direcionadas à extração de características e classificação de sinais, objetivando a identificação de amostras de aço IF quanto a modificação de características mecânicas a partir do ensaio não destrutivo, utilizando a técnica de correntes parasitas pulsadas (PEC). Para o referido propósito, foram analisadas diferentes rotinas computacionais baseadas em técnicas de Processamento Digital de Sinais e diversos algoritmos de aprendizado de máquina: Máquina de Vetores de Suporte, K vizinhos mais próximos, Árvore de Decisão, Perceptron Multicamadas, Regressão Logística e Naive Bayes Gaussiano. Como parte adicional desta pesquisa, um ensaio de tração foi conduzido para investigar o comportamento de corpos de prova submetidos a um estresse conhecido, para geração de tensões e a sua repercussão no ensaio PEC. Para otimização da classificação, foram avaliados diferentes tipos de modelos *ensemble*, utilizando como classificadores de base os modelos previamente validados. Os resultados obtidos indicam um desempenho satisfatório na classificação dos sinais com os modelos de base e apresentam melhora com a utilização do *ensemble*.

Palavras-chave: Aprendizado de Máquina. Correntes Parasitas Pulsadas. Reconhecimento de Padrões. Ensemble. Processamento Digital de Sinais.

Abstract

ENSEMBLE OF CLASSIFIERS FOR IF STEEL PLATES EVALUATION USING PULSED EDDY CURRENTS

Camilla Bahia Larocca

October/2022

Advisors: Eduardo Furtado de Simas Filho, D.Sc
Ivan Costa da Silva, D.Sc

Department: Electrical Engineering

Interstitial Free (IF) steels are used in the manufacture of numerous components with complex shapes. They are used for drawing in the automobile industry due to their high ductility and good mechanical strength. Since it is delivered in coil form, the material is exposed to loads of its own weight and may suffer from residual stresses generated. As a consequence, the steel's mechanical properties are susceptible to changes. Since the material is submitted to equipment with pre-established parameters for desired conditions, identifying residual stresses beyond the tolerance range is an essential step in preventing defects in the manufactured parts. In this prospect, the current work proposes a set of intelligent techniques focused on feature extraction and signal classification, with the objective of identifying IF steel samples regarding the modifications of mechanical properties through non-destructive testing using the pulsed eddy current technique (PEC). For this purpose, different computational routines based on Digital Signal Processing techniques and several machine learning algorithms were analyzed: Support Vector Machine, K-nearest-neighbor, Decision Tree, Multilayer Perceptron, Logistic Regression, and Gaussian Naive Bayes. As an additional part of this research, a tensile test was conducted to investigate the behavior of specimens subjected to known stress and its repercussion on the PEC test. Different ensemble models were evaluated for classification optimization, using the previously validated models as base classifiers. The results obtained indicate a satisfactory performance in the classification of the signals with the base models and show an improvement with the use of an ensemble.

Keywords: Machine Learning. Pulsed Eddy Current. Pattern Recognition. Ensemble. Digital Signal Processing.

Lista de ilustrações

Figura 1 – Peça feita utilizando o aço IF na indústria automobilística. Fonte: Imagem retirada de (HOILE, 2000).	1
Figura 2 – Diagrama da profundidade de penetração de corrente parasita, adaptado de (CHEN; JOMDECHA; XIE, 2019).	7
Figura 3 – Diagrama do princípio do teste por corrente parasita, adaptado de (CHEN; JOMDECHA; XIE, 2019).	8
Figura 4 – Sinal de excitação, e sinal de resposta com exemplos de parâmetros adotados para avaliação dos sinais no domínio do tempo, adaptado de (HABIBALAH; SAFIZADEH, 2013).	10
Figura 5 – Etapas do processamento de sinal na classificação de sinais provenientes de END.	13
Figura 6 – Demonstração da decomposição de frequências senoidais. Retirado de (KONG; SIAUW; BAYEN, 2020).	14
Figura 7 – Exemplo dos dois primeiros componentes principais Y_1 e Y_2 do conjunto de dados distribuído entre os eixos X_1 e X_2 . Fonte: Imagem retirada de (HAN; KAMBER; PEI, 2012).	15
Figura 8 – Projeção das componentes LD1 e LD2 de uma análise de LDA. Fonte: Imagem retirada de (HO, 2017).	16
Figura 9 – Hiperplano ótimo para duas classes linearmente separáveis. Fonte: Imagem adaptada de (HAYKIN, 2009).	19
Figura 10 – Aplicação da função <i>Kernel</i> em conjuntos de dados linearmente não-separáveis. Fonte: Imagem retirada de (SHEYKHMOUSA et al., 2020).	20
Figura 11 – Exemplo do uso do modelo KNN para a definição da classe do ponto central de acordo com os k indivíduos mais próximos, para k igual a 3, 7 e 11. Fonte: Imagem retirada de (ZHANG et al., 2020).	22
Figura 12 – Exemplo de uma árvore de decisão com perguntas numéricas baseadas em duas variáveis X_1 e X_2 . Cada nó folha é definido com a região do plano $X_1 \times X_2$. Fonte: Imagem retirada de (MOHRI; ROSTAMIZADEH; TALWALKAR, 2018).	24
Figura 13 – Exemplo de uma rede perceptron multicamadas com três camadas. Fonte: Imagem adaptada de (BENTO, 2021).	26
Figura 14 – Comparação entre algoritmos convencionais de GBDT e lightGBM. Fonte: Imagem adaptada de (Microsoft Corporation, 2022).	28
Figura 15 – Matriz de Confusão.	30
Figura 16 – Método de validação cruzada denominado k-Fold para um $K=5$	32
Figura 17 – Exemplo do ganho de desempenho obtido através do uso de <i>ensembles</i> para um problema de classificação multiclasse. Fonte: Imagem retirada de (ZHANG; MA, 2012).	33

Figura 18 – Três razões fundamentais do porquê um <i>ensemble</i> de classificadores pode funcionar melhor do que um classificador individual. Fonte: Imagem retirada de (DIETTERICH, 2000).	34
Figura 19 – Estrutura de um algoritmo de agrupamento.	35
Figura 20 – Fluxograma das principais.	37
Figura 21 – Disposição dos três corpos de prova retirados da chapa para cada situação.	39
Figura 22 – Exemplo dos corpos de prova utilizados.	39
Figura 23 – Exemplo dos corpos de prova utilizados.	40
Figura 24 – Diagrama do processo experimental.	41
Figura 25 – Sinal do ensaio PEC e Sinal alisado	42
Figura 26 – Sinal PEC utilizado nas rotinas de processamento da informação.	42
Figura 27 – Processo de criação dos quatro conjuntos de dados de entrada utilizados na rotina.	43
Figura 28 – Diagrama da rotina computacional utilizada no projeto. A partir de um conjunto de dados inicial, quatro conjuntos foram criados a partir de métodos matemáticos de processamento digital de sinais e extração de atributos. Os conjuntos foram divididos em dois subconjuntos: treinamento/validação e teste. O subconjunto de treinamento e validação foi utilizado na seleção dos classificadores de base que serão considerados como candidatos para a construção dos <i>ensembles</i> . Esses serão construídos utilizando diferentes métodos de seleção e combinação de candidatos. Posteriormente, para a avaliação do melhor modelo, o subconjunto de teste será novamente inserido no processo e a partir da avaliação dos <i>ensembles</i> , análises comparativas em relação aos classificadores de base serão realizadas para a avaliação do ganho.	51
Figura 29 – Curvas de tensão <i>versus</i> deformação obtidas pelo ensaio de tração para os diferentes corpos de prova.	53
Figura 30 – Dispersão da amplitude máxima e do Tempo até o mínimo das chapas tracionadas.	55
Figura 31 – Média dos sinais para cada condição e avaliação do pico máximo e mínimo.	56
Figura 32 – Diferença no tempo de ponto mínimo e nas ondulações após o ponto de mínimo para dois sinais escolhidos aleatoriamente.	57
Figura 33 – Principais métricas utilizadas no estudo do dispersão dos sinais.	57
Figura 34 – Dispersão da amplitude máxima de cada classe.	58
Figura 35 – Dispersão da amplitude mínima de cada classe.	59
Figura 36 – Dispersão da amplitude total de cada classe.	59
Figura 37 – Dispersão do tempo até a amplitude mínima de cada classe.	60
Figura 38 – DFT do sinal de PEC.	61
Figura 39 – Distribuição da análise da LDA.	62

Lista de tabelas

Tabela 1 – Diferentes funções Kernel para SVM	21
Tabela 2 – Principais propriedades mecânicas obtidas no ensaio de tração	40
Tabela 3 – Principais propriedades mecânicas obtidas no ensaio de tração.	41
Tabela 4 – Resumo das principais características dos conjuntos de dados utilizados neste trabalho.	44
Tabela 5 – Conjuntos de parâmetros utilizados na otimização da SVM.	46
Tabela 6 – Conjuntos de parâmetros utilizados na otimização do KNN.	46
Tabela 7 – Conjuntos de parâmetros utilizados na otimização da Árvore de Decisão. . .	46
Tabela 8 – Conjuntos de parâmetros utilizados na otimização do Perceptron multicamadas.	47
Tabela 9 – Conjuntos de parâmetros utilizados na otimização do LightGBM.	47
Tabela 10 – Conjuntos de parâmetros utilizados na otimização da Regressão Logística. . .	48
Tabela 11 – Conjuntos de parâmetros utilizados na otimização do Naive Bayes Gaussiano.	48
Tabela 12 – Principais propriedades mecânicas obtidas no ensaio de tração.	52
Tabela 13 – Resultado médio do ensaio de dureza.	54
Tabela 14 – Principais propriedades extraídas dos sinais médios de PEC.	54
Tabela 15 – Principais propriedades extraídas dos sinais médios de PEC.	60
Tabela 16 – Classificadores de base que apresentaram o melhor desempenho na etapa de otimização de parâmetros para os dados PEC-FEATURES. Os algoritmos com as células hachuradas não seguiram para a próxima etapa.	64
Tabela 17 – Classificadores de base que apresentaram o melhor desempenho na etapa de otimização de parâmetros para os dados PEC-PCA. Os algoritmos com as células hachuradas não seguiram para a próxima etapa.	65
Tabela 18 – Classificadores de base que apresentaram o melhor desempenho na etapa de otimização de parâmetros para os dados DFT-PCA. Os algoritmos com as células hachuradas não seguiram para a próxima etapa.	66
Tabela 19 – Classificadores de base que apresentaram o melhor desempenho na etapa de otimização de parâmetros para os dados DFT-LDA.	67
Tabela 20 – Candidatos a classificadores de base na etapa de validação. Os modelos com as células hachuradas não seguiram para a próxima etapa.	69
Tabela 21 – Configuração final dos classificadores de base para cada método de combinação do <i>ensemble</i>	70
Tabela 22 – Resultados dos Candidatos a classificadores de base na etapa de teste.	72
Tabela 23 – Resultados finais dos <i>ensembles</i> testados.	73
Tabela 24 – Resultados com a implementação da <i>ensemble</i>	73
Tabela 25 – Matriz de confusão dos melhores resultados (em %) para o conjunto de dados PEC-FEATURES.	74
Tabela 26 – Matriz de confusão dos melhores resultados (em %) para o conjunto de dados PEC-PCA.	74

Tabela 27 – Matriz de confusão dos melhores resultados (em %) para o conjunto de dados
DFT-PCA. 75

Tabela 28 – Matriz de confusão dos melhores resultados (em %) para o conjunto de dados
DFT-LDA. 75

Lista de abreviaturas e siglas

CV	Validação Cruzada (<i>Cross Validation</i>)
DFT	Transformada Discreta de Fourier (<i>Pulsed Eddy Currents</i>)
DT	Árvore de Decisão (<i>Decision Tree</i>)
ECT	Técnica de Corrente Parasita (<i>Eddy Current Technique</i>)
EEMD	Ensemble-Decomposição em Modos Empíricos (<i>Ensemble Empirical Mode Decomposition</i>)
GBM	<i>Gradient boosting Machine</i>
HHT	Transformada de Hilbert-Huang (<i>Hilbert–Huang Transform</i>)
IF	Livre de elementos intersticiais (<i>Interstitial Free</i>)
kNN	K-Vizinhos Mais Próximos (<i>K-Nearest Neighbors</i>)
LDA	Análise Discriminante Linea (<i>Linear Discriminant Analysis</i>)
LGBM	<i>Light Gradient Boosting Machine</i>
LR	Regressão logística (<i>Logistic regression</i>)
MLP	Perceptron multicamadas (<i>Multilayer perceptron</i>)
NB	Naive Bayes Gaussiano
NDE	Ensaio não destrutivo (<i>Non Destructive Evaluation</i>)
P-VM	Ensemble por Poda com Voto Majoritário
P-VME	Ensemble por Poda com Voto Majoritário Enviesado
PCA	Análise de Componentes Principais (<i>Principal Component Analysis</i>)
PDS	Processamento Digital de Sinais
PEC	Correntes Parasitas Pulsadas (<i>Pulsed Eddy Currents</i>)
RNA	Rede Neural Artificial
SVM	Máquina de Vetores de Suporte (<i>Support Vector Machine</i>)
SC-VM	Ensemble por Seleção Construtiva com Voto Majoritário
SC-VME	Ensemble por Seleção Construtiva com Voto Majoritário Enviesado
UT	Teste de Ultrassom (<i>Ultrasonic testing</i>)

Sumário

1	INTRODUÇÃO	1
1.1	Justificativa	1
1.2	Objetivo	4
1.2.1	Objetivos Específicos	4
1.3	Contribuições	4
1.4	Organização do documento	4
2	PESQUISA BIBLIOGRÁFICA	6
2.1	Inspeção por Correntes Parasitas Pulsadas	6
2.1.1	Aplicações da técnica PEC em pesquisas relacionadas	11
2.2	Técnicas de processamento de Sinais	13
2.2.1	Transformada Discreta de Fourier	14
2.2.2	Análise de Componentes Principais	15
2.2.3	Análise do Discriminante Linear	16
2.3	Aprendizado de Máquina	17
2.3.1	Máquina de Vetores de Suporte	18
2.3.2	K-Nearest Neighbors	22
2.3.3	Árvore de Decisão	23
2.3.4	Perceptron Multicamadas	25
2.3.5	Gradiente Boosting	26
2.3.6	Regressão Logística	28
2.3.7	Naive Bayes Gaussiano	29
2.3.8	Crterios de Avaliao de Desempenho	29
2.3.9	Tcnica de Valiao Cruzada	31
2.4	Ensemble de classificadores	31
3	MÉTODO	37
3.1	Aquisio dos dados e preparao dos <i>datasets</i>	38
3.2	Classificadores de base	44
3.2.1	SVM	45
3.2.2	KNN	45
3.2.3	Árvore de Deciso	46
3.2.4	MLP	47
3.2.5	Gradient Boosting	47
3.2.6	Regresso Logística	48
3.2.7	Naive Bayes Gaussiano	48
3.2.8	Avaliao dos melhores modelos	48
3.3	Estruturação do <i>ensemble</i> de classificadores	49

4	RESULTADOS OBTIDOS	52
4.1	Ensaio de Tração	52
4.2	Avaliação dos sinais PEC	56
4.3	Classificadores de base	61
4.4	Construção dos <i>Ensembles</i>	68
4.5	Etapa de teste	71
4.6	Comparação dos Resultados	72
5	CONCLUSÃO	76
	REFERÊNCIAS	78

1 Introdução

1.1 Justificativa

Os aços IF (do inglês “*interstitial-free*”) são amplamente utilizados em peças de bicicletas, motocicletas, móveis de escritórios, componentes eletrônicos, artigos de uso pessoal, ferragens e, em especial, na indústria automobilística: são empregados na manufatura de peças da carroceria com geometrias complexas, tais como para-lama, tampa de porta-mala, assoalho do porta-malas e as partes internas das portas dianteiras e traseiras, como visto na Figura 1.

Estes aços possuem como características pequenas taxas de elementos intersticiais, ou seja, apresentam baixos teores de elementos como carbono e nitrogênio em sua composição, tipicamente $<0,0030\%$. Com o intuito de reduzir os percentuais de carbono e nitrogênio em sua solução sólida, são inseridos, por exemplo, elementos micro ligantes como o titânio e o nióbio. Nesta incorporação, o Ti e o Nb irão fixar-se aos elementos intersticiais em forma de precipitados de carbonetos e nitretos (HOILE, 2000). Tal configuração possibilita que o aço apresente características como elevada conformabilidade, boa resistência mecânica, boa soldabilidade e baixa susceptibilidade ao envelhecimento quando comparados aos aços anteriormente utilizados na indústria, conhecidos como aço baixo carbono.



Figura 1 — Peça feita utilizando o aço IF na indústria automobilística. Fonte: Imagem retirada de (HOILE, 2000).

A extensa aplicação de aços IF na indústria ocorre principalmente devido às suas características mecânicas (WANG et al., 2020). O material possibilita uma redução do peso do produto final e, conseqüentemente, uma maior economia de combustível e a diminuição de emissão de gases resultantes da sua queima. Adicionalmente, por possuir uma elevada capacidade

de deformação, é possível um aumento da segurança do usuário visto que a chaparia irá absorver parte da energia em caso de colisão (ROMANYUK et al., 2018).

O transporte e armazenamento da chapa do aço IF ocorre sob a forma de bobinas, o que implica em constante submissão às cargas resultantes do próprio peso. Como consequência, tensões residuais¹ podem surgir no material, alterando algumas propriedades mecânicas, tais quais a durabilidade da vida em fadiga², a resistência à corrosão e ao desgaste. Essas tensões possuem ainda como agravante o fato de permanecer, na maioria das vezes, incógnitas, além de influenciarem o início de uma deformação plástica nas camadas superficiais ou sub-superficiais do material (NORCINO, 2013). Visto que o material é submetido a equipamentos de modelagem com parâmetros pré-estabelecidos para condições ideais da liga, a identificação de tensões residuais fora do intervalo de tolerância é um importante passo na prevenção de falhas na estampagem.

Com o objetivo de melhorar a qualidade da produção das peças é necessário eliminar ou compensar as variações das propriedades do aço IF. Há duas formas de atingir este resultado: classificando o material como inaplicável à produção ou ajustando os parâmetros da produção em relação às propriedades do material atual. A forma convencional para identificação de tensões residuais envolve modelagem matemática ou técnicas de medição no próprio material a partir de métodos destrutivos e não destrutivos. As técnicas destrutivas baseiam-se na alteração de estado de equilíbrio das tensões residuais a partir da retirada de camadas ou usinagem da peça, enquanto as técnicas não destrutivas analisam as relações entre as propriedades físicas ou cristalográficas do material e a tensão residual. Dentre os principais métodos de medição não destrutivos e destrutivos, as técnicas de difração de raios-X e o Método do Furo são os mais utilizados, respectivamente. (NORCINO, 2013; SOUSA et al., 2012).

Dentro do universo das técnicas magnéticas na inspeção não destrutiva para a avaliação de tensões em materiais metálicos, a utilização de correntes parasitas pulsadas foi estudada por Morozov, Tian e Withers (2010b). Neste artigo, ligas de alumínio são monitoradas com sondas isotrópicas e direcionais, e submetidas a uma gama de tratamentos térmicos e não térmicos com o objetivo de aplicar tensões abaixo do limite elástico. O efeito do tratamento térmico e da deformação plástica sobre a resposta de PEC é avaliada utilizando o valor de pico do sinal diferencial. Tanto a deformação plástica quanto o tratamento térmico mostraram ter um forte efeito sobre os coeficientes de estresse, e consequentemente, sobre as propriedades magnéticas do material. O sinal de resposta do PEC indicou que a condutividade elétrica média diminuiu com o aumento da tensão aplicada.

Técnicas de inteligência computacional têm sido empregadas extensivamente em pesquisas associadas à área industrial para processamento de sinais, extração de características e aproximação de funções. A combinação da técnica não destrutiva com algoritmos de aprendizado de máquinas tem se mostrado atraente para desenvolvimento de novas técnicas e suas aplicações,

¹ De acordo com Callister Jr. e Rethwisch (2016) tensão residual é a tensão que permanece em um material que está livre de forças externas ou de variações de temperatura. Elas podem se desenvolver em peças metálicas por consequência de processos de deformação plástica e resfriamento não uniforme. Distorção e empenamento podem ocorrer caso essas tensões residuais não sejam removidas.

² De acordo com Callister Jr. e Rethwisch (2016) vida em fadiga é um parâmetro que caracteriza o comportamento de fadiga de um material e é definido pelo número de ciclos necessários para causar a falha sob um nível de tensão específico.

como algoritmos de pré-processamento e compressão de sinais, reconhecimento de padrões de análise, classificação e agrupamento de sinais (WUNDERLICH; TSCHÖPE; DUCKHORN, 2018; NICCOLAI et al., 2021). Neste contexto, Heingärtner, Born e Hora (2010) propõem um sistema de medição utilizando correntes parasitas multifrequências para a aquisição das propriedades mecânicas de uma chapa metálica em linha em produção. Posteriormente, um modelo de regressão é proposto para correlacionar as propriedades magnéticas e resistência à tração. Ainda como exemplo, o trabalho de Habibalahi et al. (2015) utiliza uma rede neural do tipo perceptron multicamadas (do inglês “*Multilayer Perceptron*” - MLP) para avaliação de tensões em chapas de alumínio a partir das informações adquiridas pelas técnicas de correntes parasitas pulsadas (do inglês “*Pulsed Eddy Current*” - PEC) e ultrassom (UT).

Adicionalmente, a utilização de técnicas auxiliares como a transformação de sinais e extração de características permitem investigar atributos ocultos, descrever padrões e analisar as informações transportadas pelos sinais. De modo complementar, estes métodos auxiliam as etapas subsequentes de aprendizagem e generalização, reduzindo a redundância de informações, os riscos de sobreajustes, e em alguns casos, melhorando a visualização dos dados e interpretações humanas. Neste cenário, o trabalho de Silva Jr. et al. (2021) utiliza as transformações discretas de Fourier para avaliação do espectro de frequência e Wavelet para a decomposição do sinal, combinadas com a compactação de informações por Análise de Componentes Principais (do inglês “*Principal Component Analysis*” - PCA) na avaliação de um tubo de aço carbono isolado termicamente utilizando a técnica de PEC. Os conjuntos formados foram avaliados em duas redes: MLP e máquina de aprendizado extremo. As redes que apresentaram os melhores desempenhos utilizaram como conjunto de entradas a combinação dos três métodos.

A combinação de parâmetros pode ser um incremento valioso em avaliações que utilizam inteligência computacional. Como exemplo, Parikh et al. (2004) propõem um algoritmo incremental de aprendizado de máquina utilizando a combinação das informações extraídas a partir de imagens resultantes das técnicas de Ultrassom e Vazamento de Fluxo Magnético. Em média, 97,2% das simulações realizadas apresentaram o modelo proposto com resultados superiores aos desempenhos de modelos das técnicas individualmente. Para um mesmo conjunto de dados, a combinação de diferentes algoritmos pode trazer benefícios. Em seu trabalho Chou, Ngo e Chong (2017) atingiu resultados promissores utilizando a combinação de rede neurais artificiais, máquinas de vetores de suporte, árvore de regressão e regressão linear para a construção de um *ensemble* na avaliação de corrosão em equipamentos metálicos localizados em ambientes marítimos.

Para a avaliação de um dos aços mais utilizados na fabricação de chassis e latarias da indústria automobilística com uma das técnicas magnéticas emergentes de inspeção não destrutiva, este trabalho busca avaliar o sinal de PEC com o auxílio de métodos de transformação de sinais e extração de características para a criação de conjuntos de dados que servirão de entradas para algoritmos de aprendizado de máquina. Posteriormente, conjuntos de classificadores serão construídos e avaliados em relação aos ganhos adquiridos quando comparados a classificadores individuais. Na seção a seguir, serão descritos os objetivos norteadores da pesquisa.

1.2 Objetivo

Classificar amostras do aço IF utilizando um modelo *ensemble* de classificadores utilizando técnicas de processamento para a análise dos sinais de Correntes Parasitas Pulsadas.

1.2.1 Objetivos Específicos

Para alcançar o objetivo geral, este trabalho apresentou os seguintes objetivos específicos:

- Estudar uma chapa de Aço IF utilizando a técnica de Ensaio Não Destrutivo de Correntes Parasitas Pulsadas.
- Estudar uma chapa de Aço IF utilizando o ensaio de tração-deformação e de microdureza de Vickers.
- Aplicar ferramentas de processamento digital de sinais para a extração de parâmetros e análise dos conjuntos de dados como a avaliação no domínio da frequência com a transformada discreta de Fourier e o estudo dos componentes principais por técnicas de redução de dimensionalidade.
- Construir classificadores de base específicos para análise das principais características extraídas dos sinais utilizando técnicas de Aprendizado de Máquinas supervisionadas.
- Construção de um algoritmo classificador multi-classe *ensemble*, alimentado pelos classificadores de base específicos.
- Classificar sinais extraídos de uma bobina de aço IF provenientes do início, meio e final do enrolamento.

1.3 Contribuições

Embora exista uma tendência crescente dos estudos da aplicação de métodos computacionais associados à técnica de PEC, não foi encontrado, até o momento, a utilização de *ensembles* para a avaliação dos aços IF utilizando correntes parasitas pulsadas. Desta maneira, esta pesquisa irá contribuir para o estado da arte a partir de um estudo comparativo dos ganhos associados à utilização de algoritmos de agrupamento em relação a classificadores individuais, utilizando como dados de entrada dos classificadores conjuntos de atributos que apresentam diferentes formas de processamento do sinal PEC para a extração de características relevantes.

1.4 Organização do documento

Esta Dissertação de Mestrado está organizada do modo descrito a seguir:

- No Capítulo 2 será realizada uma Revisão Bibliográfica, apresentando a fundamentação teórica relacionada aos conceitos fundamentais da técnica de inspeção não destrutiva utilizada na avaliação dos corpos de prova, bem como dos sistemas inteligentes de classificação utilizados.

- No Capítulo 3 serão apresentados os procedimentos de divisão e preparação das amostras, bem como a coleta dos *datasets* dos sinais PEC. Serão descritos também, as principais metodologias e técnicas propostas para a escolha dos principais classificadores de base e da *ensemble*.
- No Capítulo 4 serão avaliados e discutidos os resultados experimentais obtidos na pesquisa de Mestrado.
- No Capítulo 5 será apresentado a conclusão sobre os resultados obtidos na Dissertação de Mestrado.

2 Pesquisa Bibliográfica

2.1	Inspeção por Correntes Parasitas Pulsadas	6
2.1.1	Aplicações da técnica PEC em pesquisas relacionadas	11
2.2	Técnicas de processamento de Sinais	13
2.2.1	Transformada Discreta de Fourier	14
2.2.2	Análise de Componentes Principais	15
2.2.3	Análise do Discriminante Linear	16
2.3	Aprendizado de Máquina	17
2.3.1	Máquina de Vetores de Suporte	18
2.3.2	K-Nearest Neighbors	22
2.3.3	Árvore de Decisão	23
2.3.4	Perceptron Multicamadas	25
2.3.5	Gradiente Boosting	26
2.3.6	Regressão Logística	28
2.3.7	Naive Bayes Gaussiano	29
2.3.8	Crítérios de Avaliação de Desempenho	29
2.3.9	Técnica de Validação Cruzada	31
2.4	Ensemble de classificadores	31

Este projeto propõe o estudo da classificação de sinais extraídos de chapas de aço IF a partir da técnica de correntes parasitas pulsadas, utilizando técnicas de processamento digital de sinais aliados a algoritmos de aprendizagem de máquina. Neste capítulo são apresentadas as definições e os conceitos básicos para compreensão deste trabalho.

2.1 Inspeção por Correntes Parasitas Pulsadas

Ensaio Não Destrutivo (END) são técnicas utilizadas na avaliação de estruturas para detectar características específicas como defeitos, corrosões e modificações estruturais, determinando a sua severidade antes que falhas indesejáveis ocorram sem modificar ou alterar o objeto inspecionado em qualquer sentido (HELLIER, 2020). São aplicadas não apenas em análises qualitativas, como a detecção e caracterização de falhas, mas também em avaliações quantitativas, para mensurar características geométricas do material e examinar propriedades mecânicas e elétricas, aumentando a vida útil dos equipamentos. Os ENDs estão entre as principais ferramentas de gestão da qualidade, essenciais em vários tipos de indústrias, incluindo as indústrias de transporte, aeroespacial, automotiva, manufatura e petroquímicas (SOPHIAN; TIAN; FAN, 2017). A escolha da técnica depende do tipo de equipamento a ser inspecionado, o tipo de defeito que se espera encontrar e a precisão necessária (GUPTA et al., 2021).

A técnica de correntes parasitas (do inglês “*Eddy Current Technique*” - ECT) é um ensaio não destrutivo baseado no princípio da indução eletromagnética e é utilizada na inspeção

de materiais metálicos para caracterizar propriedades físicas e defeitos em sua estrutura. Sua premissa se baseia nas correntes de Foucault, geradas em um material condutor quando submetido a um campo magnético variável. (HELLIER, 2020) A corrente alternada, que flui através de uma bobina de teste, produz um campo magnético alternado de densidade de fluxo B . O fluxo magnético total (ϕ_p) contido dentro do corpo é o produto de B pela sua área. Ao aproximar uma bobina excitada por um sinal senoidal de um objeto metálico, o campo eletromagnético gerado irá induzir correntes circulares que fluem em planos paralelos às espiras da bobina. Estas correntes parasitas induzidas irão gerar um campo magnético opositor. Dessa forma, o fluxo magnético induzido (ϕ_s) se opõe ao fluxo magnético primário, conforme Figura 2. A interação entre os dois campos resulta em uma força eletromotriz que irá influenciar o campo magnético total da bobina, ocasionando uma mudança aparente de sua impedância (CHEN; JOMDECHA; XIE, 2019). Estas mudanças ocorrem caso haja alguma alteração na microestrutura ou falhas mecânicas no material, dado que essas características são altamente correlacionadas à condutividade elétrica, permeabilidade magnética ou disposição do campo magnético (SOPHIAN; TIAN; FAN, 2017). Os principais fatores que influenciam no sinal são descritos a seguir.

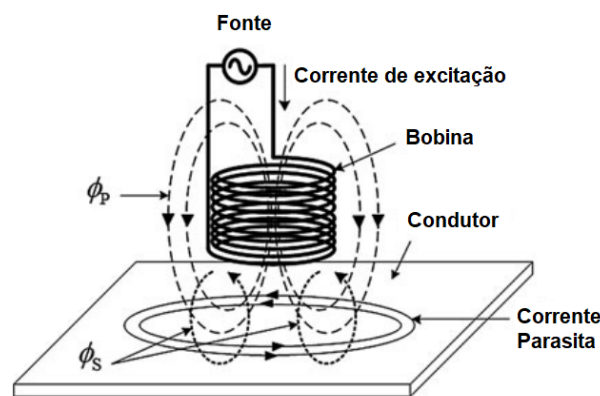


Figura 2 – Diagrama da profundidade de penetração de corrente parasita, adaptado de (CHEN; JOMDECHA; XIE, 2019).

Dentro do conjunto de variáveis que devem ser consideradas em um ensaio de ECT, a frequência de excitação é um fator importante que pode ser controlado pelo operador, visto que as propriedades do material da amostra metálica são fixas e a sonda ECT é frequentemente limitada pela geometria do material. A alteração da frequência irá modificar importantes parâmetros como as fases das correntes de Foucault induzidas, a densidade e a profundidade de penetração das correntes no corpo inspecionado.

A frequência de excitação determina o alcance das correntes parasitas no interior do material. Em geral, componentes de alta frequência tem um menor alcance de penetração e são influenciados por defeitos superficiais. Além disso, possuem um tempo de difusão menor do que os sinais compostos por frequências mais baixas, que possuem uma maior profundidade de penetração e são influenciados tanto por defeitos na superfície como defeitos mais internos (ARJUN et al., 2015).

A profundidade de penetração da corrente parasita está relacionada à concentração das

correntes circulares em relação à sua localização no material. Devido ao efeito pelicular, também conhecido como *skin effect*, a concentração máxima de correntes está situada próximo à superfície adjacente à bobina, decrescendo exponencialmente com a penetração ao longo da parede do equipamento. A profundidade na qual a densidade de corrente se encontra em $1/e$, ou 36,8% do valor da superfície, é conhecido como a profundidade de penetração padrão (γ), dado de acordo com a Equação 2.1. Além da frequência de teste, o decaimento da distribuição das correntes ao longo da profundidade do material é influenciado também pela condutividade elétrica (σ) e pela permeabilidade magnética (μ). A equação do efeito pelicular pode ser deduzida a partir das equações de campo de Maxwell, conforme descrito em Uemura (2018).

$$\gamma = \sqrt{\frac{1}{\pi f \mu \sigma}} \quad (2.1)$$

Visto que a sensibilidade do sinal de ECT à falha depende da densidade de correntes naquela localização (CHEN; JOMDECHA; XIE, 2019), a definição da frequência de operação em função do limiar de excitação e sensibilidade do equipamento está relacionado com a profundidade de penetração padrão no material. Em 2δ encontra-se 13,5% da densidade de corrente inicial e em 3δ , também conhecido como o ponto de penetração de corrente efetiva, a corrente decai para 5% do valor inicial (UEMURA, 2018), conforme mostrado em Figura 3. O ponto 3δ é definido como o limite onde o sinal de corrente é recebido com uma relação sinal-ruído suficiente para a análise.

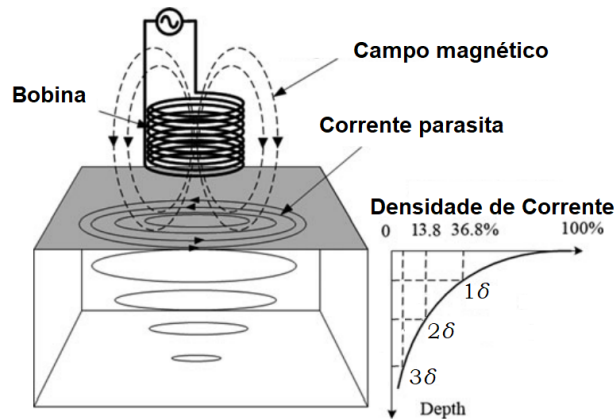


Figura 3 – Diagrama do princípio do teste por corrente parasita, adaptado de (CHEN; JOMDECHA; XIE, 2019).

Além da frequência, o *Lift-off* e o manuseio da sonda também são características controladas pelo operador. *Lift-off* é o termo usado para descrever qualquer espaço entre o espécime em teste e a sonda ECT. O aumento da distância entre a sonda e o corpo de prova diminuirá a densidade da corrente parasita no espécime, o que significa que a sensibilidade do teste é reduzida. Geralmente, a distância de *Lift-off*, o ângulo entre a sonda e o corpo e a pressão aplicada devem ser mantidas constantes durante a aplicação. Qualquer mudança no manuseio da sonda pode causar mudanças significativas nos sinais ECT. Adicionalmente, o operador da bobina deve realizar o ensaio longe das bordas do material, visto que as correntes parasitas são distorcidas quando a bobina se aproxima da extremidade do corpo de prova. Isso ocorre devido

à impossibilidade das correntes de Foucault fluírem para fora do modelo, sendo esta influência conhecida como efeito de borda. As modificações do sinal devido ao manuseio, em geral, são mais robustas e podem mascarar qualquer sinal proveniente de falhas.

Em relação às propriedades eletromagnéticas dos materiais, a permeabilidade magnética e a condutividade elétrica, como visto na Equação 2.1, afetam a profundidade de penetração e, conseqüentemente, a densidade de corrente. Normalmente, estas propriedades podem ser simplificadas como homogêneas para todo um material condutivo, contudo, caso ocorra uma pequena alteração local, estas pequenas variações têm um grande efeito sobre a impedância da bobina ECT. Tais mudanças na impedância da bobina são, em geral, maiores quando comparado àquelas causadas por falhas, podendo novamente mascarar a avaliação desejada. Já no âmbito das propriedades mecânicas, mudanças na dureza do material e a formação de tensões residuais podem afetar propriedades elétricas. Isso ocorre devido à condutividade de um metal policristalino se tornar anisotrópica sob a influência da tensão elástica devido ao efeito da piezoresistividade. Uma vez que os sensores de EC são sensíveis às mudanças de condutividade, o método de corrente parasita tem o potencial de proporcionar um indicador do estado de tensão residual em metais e ligas (MOROZOV; TIAN; WITHERS, 2010a).

A técnica de correntes parasitas possui algumas vantagens por ser um método que utiliza as propriedades elétricas e magnéticas dos materiais. Não há a necessidade de contato direto com o corpo a ser inspecionado, reduzindo custos referentes a preparação prévia da superfície e possibilitando a sua utilização em equipamentos revestidos. Outros pontos positivos da técnica são a alta sensibilidade, a alta velocidade do processamento de sinais e o baixo custo de operação e montagem do equipamento ao comparar com outras técnicas de END. O sistema pode ser embutido e inserido em rotinas de testes mecanizados para classificar uma alta demanda de materiais, para monitorar sua deterioração e de equipamentos em serviço, e para verificar a qualidade do processo. A análise do sinal, interpretação e avaliação do ECT podem ser realizadas em tempo real. Além disso, os sistemas podem ser desenvolvidos para automação de sistemas de teste, para classificação de peças de teste, para controle dos processos de fabricação e para o controle estatístico de qualidade (CHEN; JOMDECHA; XIE, 2019).

Por outro lado, existem desvantagens que incluem a limitação da inspeção de materiais metálicos com condutividade elétrica significativa, o constante reajuste do equipamento para diferentes defeitos, a necessidade do aprendizado de técnicas de interpretação do sinal e a sensibilidade da bobina a diferentes alturas de *Lift-off*, dado que irregularidades na superfície do material podem deixar o sinal ruidoso.

Diferentemente da técnica convencional de correntes parasitas, a técnica de correntes parasitas pulsadas (do inglês, “*Pulsed Eddy Current*” - PEC) tem como princípio a utilização de um espectro contínuo de frequências, utilizando repetitivos pulsos de ondas quadradas de curta duração com um determinado ciclo de trabalho (*duty cycle*) para excitar a bobina (ARJUN et al., 2015). A técnica transitória de corrente de excitação contém uma rica gama de frequências e aumenta a força do campo de excitação magnético. Esta configuração potencializa as informações sobre as condições do teste, possibilitando uma maior sensibilidade, além de uma maior profundidade de penetração das correntes, sem a necessidade da mudança da sonda ou da

frequência de operação. Inúmeras combinações entre a resposta magnética, os sinais de referência, as combinações de sensores, sondas e as metodologias distintas de processamento digital de sinais instigam pesquisas para o seu aperfeiçoamento. Nos últimos anos, o PEC tem alcançado progressos consideráveis na detecção de rachaduras e corrosão em estruturas de fuselagem de aeronaves, etc (CHEN; JOMDECHA; XIE, 2019).

A análise da resposta de PEC é feita, na grande maioria dos casos, no domínio do tempo, onde o foco é a avaliação do formato da curva e sua variação. A pesquisa de Sophian et al. (2003) utilizou a diferença entre um sinal de referência e a resposta magnética, obtendo parâmetros característicos para a detecção de variação na espessura do material. O tempo em que o sinal atinge a amplitude nula contém informações sobre a localização de falhas na estrutura em relação a sua profundidade. Quando comparadas falhas de mesmas dimensões, existe uma relação inversamente proporcional entre a profundidade em que a descontinuidade está inserida e o tempo de resposta. De forma análoga, em termos da amplitude, o ponto máximo é proporcional ao tamanho do defeito em relação à perda de metal (BIEBER et al., 1997). De forma complementar, o trabalho de Habibalahi e Safizadeh (2013) analisa as características variantes no sinal em relação à tensão para um ensaio de medida de estresse com uma sonda que utiliza duas bobinas para excitação e aquisição da informação. O sinal de excitação e o sinal resultante com os parâmetros tempo de pico, valor de pico, ponto de subida e travessia zero são apresentados na Figura 4.

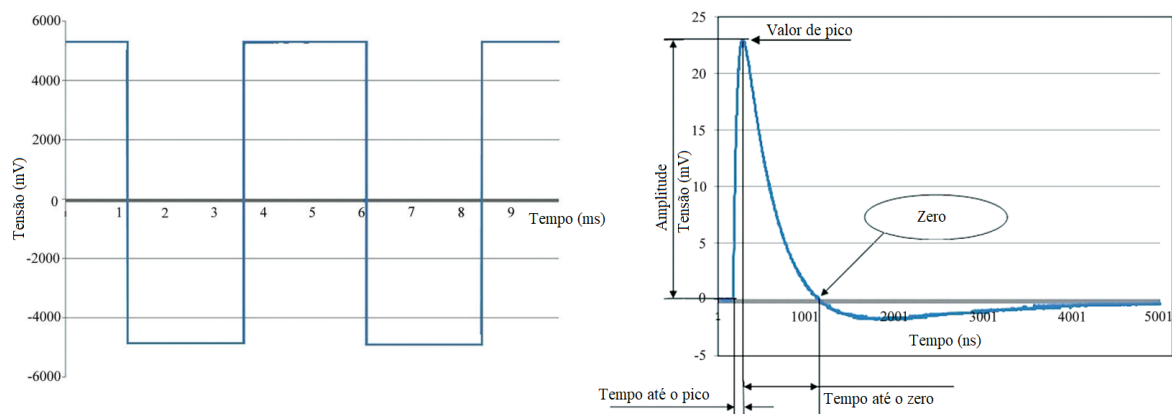


Figura 4 – Sinal de excitação, e sinal de resposta com exemplos de parâmetros adotados para avaliação dos sinais no domínio do tempo, adaptado de (HABIBALAH; SAFIZADEH, 2013).

O fator de maior importância a ser considerado na técnica de PEC é a possibilidade da utilização de diversas metodologias de processamento de sinais para a extração de informações. Inúmeras combinações entre as respostas magnéticas adquiridas, os arranjos de sensores e sondas e as metodologias distintas de tratamentos de dados instigam pesquisas para o aperfeiçoamento da técnica e em aplicações distintas. Trabalhos que relacionam PEC com processamento de sinal e aprendizado de máquinas, temas abordados nesta pesquisa, são apresentados com maiores detalhes na seção 2.1.1.

A técnica de correntes parasitas pulsadas compartilha dos pontos positivos da técnica convencional de correntes parasitas como uma alta sensibilidade e aplicabilidade. Somado a estas características, apresenta uma ampla gama de informações que podem ser combinadas e processadas, um sistema mais robusto quando consideradas as influências acometidas por

interferências e, consequentemente, aumento da relação sinal ruído do sistema.

2.1.1 Aplicações da técnica PEC em pesquisas relacionadas

Análises envolvendo a associação entre o domínio do tempo e da frequência e algoritmos de processamento de sinais para extração e definição de novos parâmetros relevantes se tornaram possíveis, em função dos incrementos da capacidade computacional. Como exemplo, [Tian et al. \(2013\)](#) estudaram a possibilidade de utilizar a normalização no espectro da frequência para remover os efeitos da variação do *lift-off* (distância entre a sonda e o corpo de prova) e as decorrências quando associadas às mudanças nas propriedades do material (condutividade elétrica e permeabilidade magnética) e nas forças de tração aplicadas às chapas de alumínio. A alteração do *lift-off* apresentou uma variação linear na magnitude espectral, podendo ser anulada com a normalização. Contudo, quando associados a variação de condutividade elétrica, permeabilidade magnética e valores de tensões associados à tração, os espectros magnéticos das diferentes distâncias apresentaram padrões não lineares e complexos, impossibilitando a sua eliminação. [Alamin et al. \(2012\)](#) aplicaram a análise de componentes principais no sinal de resposta de uma sonda PEC em placas de aço com níveis variáveis de corrosão para determinar quais são as características principais do domínio de tempo-frequência e posteriormente classificar e agrupar amostras semelhantes conforme as suas condições de superfície. Nos grupos de sinais não-normalizados o componente principal correspondeu à amplitude do sinal, ao passo que nos grupos normalizados, o referido componente está relacionado à taxa de variação da carga na forma de onda PEC. A pesquisa de [Tian et al. \(2005\)](#) utiliza a transformada de Wavelet para segregar os ruídos de alta frequência e extrair os outros componentes relevantes na decomposição do sinal. A PCA foi aplicada a este novo conjunto de dados para a classificação e quantificação dos sinais de PEC na detecção de defeitos em corpos de prova paramagnéticos e ferromagnéticos. A integração das técnicas proporcionou, além da redução de dimensionalidade, uma melhor identificação na localização de falhas.

Em trabalhos mais recentes, [Piao et al. \(2019\)](#) utilizaram como fundamento as variações das propriedades elétricas e magnéticas ocorridas nas imediações de uma descontinuidade. A pesquisa avalia a distribuição descontínua da condutividade e da permeabilidade não uniforme em corpo de prova ferromagnético contendo defeitos distintos. Para a detecção de tais intercorrências, foi analisado os padrões de distribuição dependentes da condutividade e da permeabilidade extraídos do sinal PEC. Ao se basear na natureza da difusão das correntes parasitas no material e na sensibilidade do sinal de resposta, [Ge et al. \(2020\)](#) propuseram uma característica denominada valor de pós-pico para a classificação de defeitos utilizando uma bobina absoluta. As correntes inicialmente induzidas nas camadas mais externas atingem, em primeiro, as descontinuidades superficiais, posteriormente se difundindo pelo material, onde encontram as descontinuidades sub-superficiais. Dessa forma, após o tempo de pico do sinal de resposta existe uma seção, denominada pós-pico, que apresenta uma melhor sensibilidade na detecção de defeitos sub-superficiais. Contudo, os experimentos mostraram que as características relacionadas ao ponto máximo ainda possuem um melhor desempenho na detecção de defeitos superficiais. Ainda na área das estratégias para reduzir os efeitos das alterações de *lift-off* e tornar as avaliações de espessura mais confiáveis, o trabalho de [Wen et al. \(2020\)](#) propôs um ajuste do parâmetro

conhecido como ponto de interseção do *lift-off* a partir da variação do tempo de carga do pulso de excitação. Seu trabalho mostrou que ensaios em amostras mais espessas se beneficiam ao utilizar sinais de excitação com um tempo de carga maior, e de forma complementar, recomenda-se usar sinais de excitação com um tempo de carga menor para as amostras de menor espessura.

Essas novas configurações, associadas às técnicas de aprendizagem de máquinas fornecem uma ampla gama de informações relacionadas ao espaçamento entre camadas, propriedades magnéticas, perda de material, diferenças de *lift-off* entre outros (FAN et al., 2016). Kaiyu et al. (2021) consideraram o aspecto não-estacionário do sinal e utilizaram o método da *Ensemble Decomposition* em Modos Empíricos (do inglês “*Ensemble Empirical Mode Decomposition*” - EEMD) associada à transformada de Hilbert-Huang (do inglês “*Hilbert-Huang Transform*” - HHT) no sinal PEC para extrair novas características, como o pico e a energia do espectro marginal. Para comparação, um segundo conjunto de dados composto por características mais tradicionais, como a amplitude de pico no domínio do tempo e a densidade espectral no domínio da frequência, foi criado. Os dois conjuntos serviram de entradas para uma rede neural do tipo *Backpropagation* cujo objetivo era estimar as propriedades ferromagnéticas de ligas de aço. As redes que tinham como conjunto de entrada as características tradicionais apresentaram erros entre 6,82% e 3,42%. As novas características extraídas pelo método EEMD-HHT apresentaram resultados superiores, com a faixa de erro variando de 5,70% à 2,51%. A combinação dos dois conjuntos de entrada apresentou os melhores resultados, com erros variando entre 3,85% e 2,06%. Com o objetivo de identificar duas classes de defeitos distintas em um tubo de aço carbono isolado, Silva et al. (2019) empregaram a Transformada Rápida de Fourier (do inglês “*Discrete Fourier Transform*” - DFT) e a Transformada Discreta de Wavelet para a extração de parâmetros do sinal de PEC. Posteriormente, foi utilizada PCA para selecionar as características dominantes e reduzir a dimensionalidade dos dados a serem aplicados em uma rede Perceptron multicamadas. A utilização do método de otimização por enxames de partículas impôs restrições durante o processo de classificação, resultando em redes com eficiência de até 91.2%.

Adicionalmente, a junção de dados provenientes de múltiplos sistemas no auxílio da análise de problemas e tomadas de decisões vem ganhando popularidade em aplicações militares e robóticas (GROS, 1996) e, mais recentemente, em ensaios não destrutivos. O trabalho de Habibalahi et al. (2015) propõe um método de junção de dados utilizando redes neurais para avaliação de tensões em chapas de alumínio a partir das informações adquiridas pelas técnicas de PEC e Ultrassom. O valor de pico (extraído do sinal PEC) e o tempo do transiente (extraído do sinal de UT) foram normalizados e os dados agregados, combinados e fundidos utilizando a rede Perceptron multicamadas. O erro médio da rede utilizando a fusão de dados foi de cerca de 5,5%, uma redução de cerca de 55% em comparação com o erro médio nos métodos UT e PEC. Em contrapartida, Wei et al. (2020) propuseram uma fusão de perspectiva física entre o transdutor acústico electromagnético e a técnica de PEC para detecção de defeitos a partir da criação de um sensor híbrido, podendo simultaneamente gerar ultrassom e correntes de Foucault no material.

A ampla gama de possibilidades viabiliza a utilização da técnica de correntes parasitas pulsada para a análise e investigação das propriedades mecânicas, magnéticas e estruturais de materiais metálicos em diversas situações. A combinação de distintas configurações dos sistemas físicos de geração dos sinais e técnicas computacionais ainda é estudada e motiva diversos

pesquisadores da área.

2.2 Técnicas de processamento de Sinais

Os conjuntos de dados utilizados atualmente, em especial caso sejam provenientes do mundo real, são altamente suscetíveis a ruídos e inconsistências adquiridas em vários estágios da aquisição do sinal. Ruídos e distorções devido aos circuitos eletrônicos do instrumento, acoplamentos de sensores, digitalização dos sinais e erros associados a dispositivos e circunstâncias de medições são algumas situações que podem interferir no resultado. Informações de baixa qualidade poderão induzir a resultados e análises insatisfatórias (HAN; KAMBER; PEI, 2012), gerando um estado de confusão e incerteza nas observações. Para contornar esta situação, as pesquisas e aplicações industriais dos últimos anos tem se concentrado em processos que melhoram a confiabilidade e o conteúdo das informações. Muitos destes estudos estão relacionados ao uso de conceitos de processamento de sinais, que já foram usados com sucesso em outras áreas de engenharia, como radar, sonar e geofísica (VEJDANNIK et al., 2019).

O Processamento Digital de Sinais (PDS) pode ser compreendido como um conjunto de técnicas de transformações para a conversão dos sinais em informações úteis para um computador digital. A utilização de diversas ferramentas de PDS e extração de características principais permitem o refino do conjunto de dados em diversas aplicações.

Em rotinas de ensaios não destrutivos para a classificação de sinais, o sistema de processamento de sinal consiste em três blocos básicos: seleção dos dados (ou pré-processamento), extração de características principais e redução de características, conforme mostrado na Figura 5. Filtragem, redução de ruídos e correção de inconsistências são alguns métodos empregados no pré-processamento para eliminação de dados inválidos, não coesos ou pouco informativos. O papel da etapa de extração de características tem como finalidade selecionar as características mais relevantes para o reconhecimento de padrões. Adicionalmente, as transformações de dados também podem ser aplicadas para que os sinais sejam dimensionados a outros domínios utilizando processos de manipulação matemática. Na etapa de extração de características, são gerados conjuntos reduzidos de parâmetros que podem traduzir a verdadeira natureza do conjunto inicial, eliminando recursos redundantes ou que não forneçam informações notáveis em detrimento dos recursos computacionais de processos futuros. Estas etapas podem melhorar a precisão e a eficiência dos algoritmos de classificação que se seguirão, visto que existem muitos fatores que compreendem a qualidade dos dados, incluindo precisão, integridade, consistência, confiabilidade e interpretabilidade (KUBAT, 2017).



Figura 5 – Etapas do processamento de sinal na classificação de sinais provenientes de END.

Para a transformação, extração de características e interpretação dos sinais adquiridos, são utilizadas técnicas de processamento de sinais visando produzir novas representações da informação, que podem revelar aspectos significativos para o problema. Nesta seção, serão comentados as técnicas utilizadas neste trabalho:

2.2.1 Transformada Discreta de Fourier

A Transformada Discreta de Fourier, (do inglês “*Discrete Fourier Transform*” - DFT) é a equivalente da Transformada Contínua de Fourier aplicada a sinais discretos. É uma transformada amplamente utilizada em sinais práticos com amostras espaçadas em um intervalo de tempo finito, como sinais de rádio, ondas sonoras ou medições de temperatura, para a avaliação do espectro de frequência, processamento de fala, remoção de ruído, imagem e reconhecimento de padrões.

A ferramenta DFT converte os dados amostrados no domínio do tempo em dados amostrados no domínio da frequência, decompondo qualquer sinal em uma soma de simples ondas senoidais. Para um sinal de comprimento finito $x[n]$, a DTF pode ser definida a partir da Equação 2.2, onde N são os números de amostras, n é a amostra atual, k é a frequência atual, sendo $k \in [0, N - 1]$, $x[n]$ é o valor do seno na amostra n e X_k é a DFT que inclui informações de amplitude e fase (ANTONIOU, 2006).

$$X_k = \sum_{n=0}^{N-1} x[n] \cdot e^{-2j\pi k n / N}, 0 \leq k \leq N - 1 \quad (2.2)$$

Dessa forma, é possível discernir e medir facilmente as diferentes frequências, a amplitude e a fase que compõem o espectro de frequência do sinal original. A Figura 6 mostra a ideia por trás da DFT, onde o sinal acima é, na verdade, o resultado da soma de 3 ondas sinusoidais diferentes.

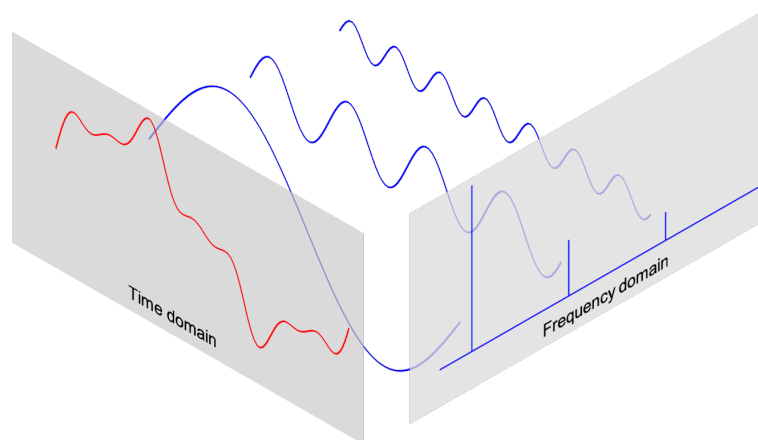


Figura 6 – Demonstração da decomposição de frequências senoidais. Retirado de (KONG; SIAUW; BAYEN, 2020).

O emprego da DFT em análises de sinais de PEC já é amplamente difundida em pesquisas da área. O uso da análise do domínio da frequência produz resultados promissores quando associamos a relação entre a frequência do sinal PEC e a profundidade de penetração do sinal (QIU et al., 2013). De forma complementar a análise do sinal de PEC no domínio do tempo, Pan et al. (2013) propuseram um método para análise do sinal PEC no domínio da frequência, extraindo características similares, tais como, frequência de pico, amplitude de frequência de pico, frequência de cruzamento zero, etc. Estas características foram usadas como entrada para um classificador baseado na análise de componentes principais, permitindo a localização de defeitos usinados em uma estrutura de várias camadas. Outros trabalhos que utilizam o espectro de potência FFT na extração de parâmetros para a classificação de defeitos são Peng et al. (2016) e Larocca et al. (2018).

2.2.2 Análise de Componentes Principais

Uma das principais questões no reconhecimento de padrões dos ensaios não destrutivos é a seleção de uma gama adequada de características, que forneçam informações independentes ou discriminantes exigidas pelo classificador para a avaliação correta dos defeitos. Muitas vezes, uma grande coleção de características potencialmente úteis são coletadas. Contudo, nem todas as características extraídas são relevantes para a tarefa de reconhecimento de padrões. Algumas são importantes apenas quando relacionadas a outras características e algumas podem ser consideradas apenas ruídos quando analisado um contexto específico (KUNCHEVA, 2004). Estas características implicam algum tipo de dúvida e ambiguidade sobre o problema, e consequentemente reduzem a capacidade de identificação (DUDA; HART; STORK, 2001).

A fim de melhorar o desempenho e aumentar a eficiência computacional, um passo fundamental no projeto de classificação de sinais é a seleção de um conjunto de características relevantes, reduzindo a dimensionalidade dos dados de entrada. A seleção das características deve cumprir um critério fundamental: as características devem reter todas e somente as informações significativas que estão incluídas nos dados, de modo a reproduzir com alta confiabilidade as informações do conjunto de dados inicial (VEJDANNIK et al., 2019).

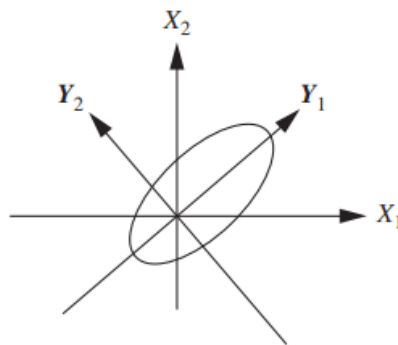


Figura 7 – Exemplo dos dois primeiros componentes principais Y_1 e Y_2 do conjunto de dados distribuído entre os eixos X_1 e X_2 . Fonte: Imagem retirada de (HAN; KAMBER; PEI, 2012).

A técnica de análise de componentes principais (do inglês “*Principal Component Analysis*” - PCA) é um método de redução de dimensionalidade não supervisionado que visa preservar a

reprodução das informações contidas nos dados originais. Dessa forma, a PCA procura por k vetores ortogonais n -dimensionais, chamados componentes principais, que melhor podem ser usados para representar o conjunto de dados originais descritos por n atributos ou dimensões, onde $k \leq n$ (HAN; KAMBER; PEI, 2012). A representação dos dois primeiros componentes principais para um determinado conjunto de dados é mostrado na Figura 7.

É uma ferramenta de análise descritiva de dados amplamente utilizada, sendo empregada para analisar as relações entre um grande número de variáveis e em encontrar novos fatores que são funções lineares daquelas presentes no conjunto de dados originais. Estas novas variáveis devem, a partir de transformações lineares, criar um conjunto de elementos não relacionadas umas com as outras que retenham a maior parte do arranjo de variabilidade e correlação da matriz de dados iniciais.

Além da redução de dimensionalidade e extração de características principais, o seu emprego também possibilita a investigação de dependência entre fatores e identificações de padrões, relações de causalidade e associações. Diversos trabalhos utilizaram PCA na etapa de extração de características e redução de dimensionalidade para a avaliação de sinais PEC (SOPHIAN et al., 2003; TIAN et al., 2005; BINFENG; FEILU; DAN, 2007; SILVA et al., 2019; NAFIAH et al., 2021; SHA et al., 2021).

2.2.3 Análise do Discriminante Linear

Enquanto o PCA não faz distinção entre diferentes classes na busca das características com a maior variabilidade, a Análise de Discriminante Linear (do inglês “*Linear Discriminant Analysis*” - LDA) procura uma combinação linear das variáveis e tenta explicitamente modelar a diferença entre as classes de dados. É uma ferramenta utilizada para, além da redução de dimensão, a classificação e visualização de dados.

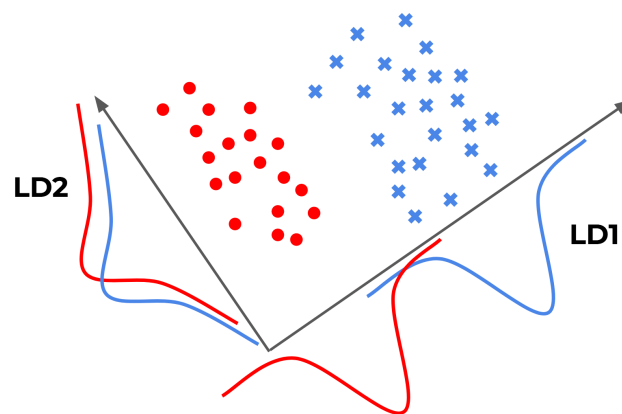


Figura 8 – Projeção das componentes LD1 e LD2 de uma análise de LDA. Fonte: Imagem retirada de (HO, 2017).

A técnica tem como princípio encontrar uma matriz de orientação $\mathbf{W}$ que reduza os vetores de característica de alta dimensão pertencentes a diferentes classes a um espaço de característica dimensional inferior (SHARMA; PALIWAL, 2014). Durante o treinamento, o algoritmo aprende os eixos mais discriminatórios entre as classes que, conseqüentemente, podem então ser usados

para definir um hiperplano onde os dados serão projetados. Dessa forma, objetivamos que os vetores de característica de uma classe neste espaço dimensional inferior estejam bem separados dos vetores de características de outras classes. Para uma fácil visualização, a Figura 8 apresenta um espaço dimensional 2D construído após análise com LDA contendo as projeções LD1 e LD2. Enquanto componente LD1 apresenta uma boa separação entre classes, a componente LD2 apresenta uma sobreposição entre classes. Para que os agrupamentos consigam achar um limite de separação ótimo, são consideradas todas as possíveis projeções e selecionadas aquelas que apresentem a melhor divisão dos conjuntos.

De forma mais detalhada, a LDA é realizada em três etapas: inicialmente é calculada a separabilidade entre diferentes classes, considerando a distância entre os centros das diferentes classes (variância entre classes); na segunda etapa, é calculada a distância entre a média e as amostras de cada classe (variância dentro da classe); já na terceira etapa, é construído o espaço dimensional inferior, que maximiza a razão da variância entre classes e minimiza a variância intra-classe, garantindo assim a máxima separabilidade (THARWAT et al., 2017).

A LDA é uma técnica promissora e uma boa escolha para a avaliação dos dados em etapas predecessoras a algoritmos de classificação (GÉRON, 2019). Embora sejam encontrados poucos trabalhos associando a utilização da Análise da Discriminante Linear à PEC (QIU et al., 2013; CHEN et al., 2014; PENG et al., 2016), existem diversas aplicações quando correlacionado as outras diversas técnicas de END (KESHARAJU et al., 2014; VEJDANNIK; SADR, 2016; HERTER et al., 2021; HWANG et al., 2022).

2.3 Aprendizado de Máquina

As técnicas de aprendizagem de máquinas fazem parte da área dos algoritmos computacionais cujo objetivo é simular a capacidade humana no aprendizado a partir dos estímulos externos. Técnicas baseadas na aprendizagem de máquinas têm sido aplicadas com sucesso em diversos campos que vão desde o reconhecimento de padrões, visão computacional, planejamento de trajetórias, finanças, entretenimento até aplicações biomédicas e médicas. (NAQA; MURPHY, 2015) A principal vantagem de usar esta abordagem computacional é desenvolver um sistema de apoio à decisão para identificar e classificar com precisão as falhas e defeitos, o que proporciona um desempenho padronizado e imparcial (VEJDANNIK et al., 2019).

Em problemas de classificação, a área de reconhecimento de padrões estuda parâmetros ou comportamentos previsíveis a partir de características associadas a uma classe, a um grupo ou a uma categoria (DUDA; HART; STORK, 2001). A construção de modelos computacionais eficazes é fundamental para a classificação de sinais. Tendo em vista este objetivo, existem diversas maneiras de um algoritmo computacional se adaptar em resposta ao aprendizado, uma vez que os dados introduzidos podem ser selecionados e ponderados para fornecer resultados mais eficientes.

Em relação ao tipo de abordagem no aprendizado de máquina, existem três vertentes populares: aprendizagem por reforço, atividade não-supervisionada e supervisionada. Na aprendizagem por reforço, o modelo realiza uma ação dentre uma série de possíveis opções em um

ambiente e recebe uma recompensa conforme o resultado dessa ação. O modelo irá observar, interagir e modificar o ambiente no treinamento e seu objetivo final é receber a maior recompensa possível (SKANSI, 2018). São algoritmos utilizados, geralmente, em controle estatístico e inteligências artificiais. Nas atividades não-supervisionadas ou *clustering*, agrupam um conjunto de dados a partir da similaridade da população, não sendo necessária a definição prévia das classes no conjunto de treinamento (NASCIMENTO, 2019). São utilizadas para agrupamentos, quantização vetorial, extração de características principais, codificação de sinal e análise de dados. Por fim temos as atividades supervisionadas, que ocorrem quando as classes dos conjuntos de treinamento são previamente definidas. Esta será a linha adotada neste projeto, dado que existem três classes pré-definidas como referência para a classificação. É amplamente utilizado para classificações, aproximações, controles estatísticos, modelagem, processamentos de sinais e otimizações (DU; SWAMY, 2014).

Existem diversos algoritmos que podem ser aplicados em atividades de classificação supervisionadas. Nesta seção, será abordado sete classificadores utilizados nesta pesquisa: Máquina de vetores de suporte, k-Nearest Neighbors, Árvore de decisão, Perceptron multicamadas, *Gradient Boosting*, Regressão Logística e Naive Bayes Gaussiano.

2.3.1 Máquina de Vetores de Suporte

Máquina de Vetores de Suporte (do inglês “*Support Vector Machine*” - SVM) é uma técnica versátil de aprendizado de máquina capaz de realizar classificações lineares e não lineares, e apresenta uma boa resposta para conjuntos de dados complexos de pequeno e médio porte (GÉRON, 2019). A técnica possui um excelente desempenho em várias aplicações de reconhecimento de padrões e existe um movimento crescente de implementação em muitos campos de pesquisa, como, por exemplo, categorização de texto (e hipertexto), classificação de imagem, bioinformática (classificação de proteínas e classificação de câncer), detecção facial e controle preditivo generalizado (CERVANTES et al., 2020).

Seu principal objetivo é mapear o padrão do espaço de entrada $\mathbb{R}^k$ em um espaço de maior dimensão $\mathbb{R}^n$, no qual o método estabelece a fronteira de decisão ótima, com margem máxima entre os pontos mais próximos de cada classe que minimize o erro de reconhecimento, (NASCIMENTO et al., 2009; MIN; HONG; CHO, 2007), através de uma função obtida a partir dos exemplos conhecidos na fase de treinamento.

A equação da fronteira de decisão, também denominada de hiperplano, é dada por:

$$\mathbf{w} \cdot \mathbf{x} + b = 0 \quad (2.3)$$

onde $\mathbf{w} = [w_1, w_2, \dots, w_n]$ são os pesos, $\mathbf{x} = [x_1, x_2, \dots, x_n]$ é o padrão de entrada mapeado para o espaço de atributos e b é um parâmetro de viés (*bias*). Sabemos que se o ponto x_i encontra-se no hiperplano, $w \cdot x + b = 0$. Caso contrário, cada x_i será associado a um valor determinado indicando se o elemento pertence a uma classe:

$$\mathbf{w} \cdot \mathbf{x} + b \neq 0 \quad (2.4)$$

Dado um conjunto de dados $D = \{(x_i, y_i) | i=1, \dots, N\} \subset \mathbb{R}^n \times \{-1, 1\}$, e sendo $f : x \rightarrow \{-1, 1\}$, iremos utilizar a informação da classe y para a determinação do melhor hiperplano. Definindo $f = y(\mathbf{w} \cdot \mathbf{x} + b)$, f será sempre positivo se o ponto for classificado corretamente e negativo caso classificado incorretamente. Dividindo f pelo comprimento do vetor $\mathbf{w}$ (representado por $\|\mathbf{w}\|$), obtemos distância algébrica de um ponto x_i até a fronteira de decisão, denominada ρ e definida por:

$$\rho = y \left(\frac{\mathbf{w}}{\|\mathbf{w}\|} \cdot \mathbf{x} + \frac{b}{\|\mathbf{w}\|} \right) \quad (2.5)$$

São calculados as distâncias para cada exemplo do treinamento. Os indivíduos de cada classe que apresentaram o menor valor de ρ , irão definir os limites da margem de separação (ρ_0). Com posse dessas informações, é possível comparar diferentes hiperplanos com distintos valores de $\mathbf{w}$ e b visando encontrar o hiperplano ótimo com o maior valor de ρ_0 . A Figura 9 apresenta a construção de um hiperplano ótimo para um espaço de entrada bidimensional.

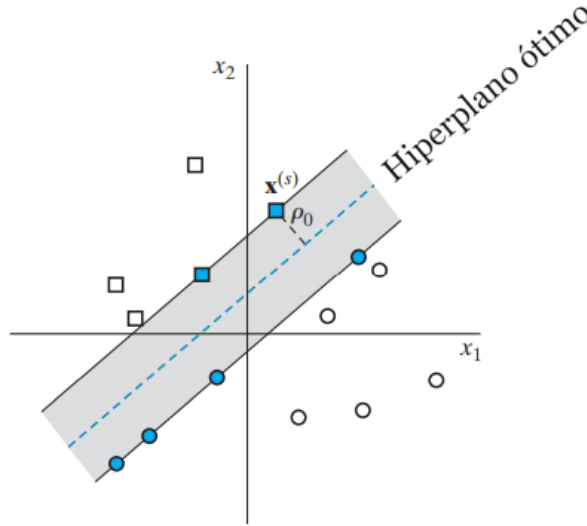


Figura 9 – Hiperplano ótimo para duas classes linearmente separáveis. Fonte: Imagem adaptada de (HAYKIN, 2009).

Em resumo, o hiperplano ideal que visamos encontrar é único no sentido que o vetor de peso ideal $\mathbf{w}$ proporciona a máxima separação possível entre exemplos de classes distintas. Esta condição ótima é obtida utilizando o método dos multiplicadores de Lagrange para solucionar um problema de otimização, onde encontraremos a função objetivo que visa maximizar os multiplicadores de Lagrange α_i , $i = 1, \dots, N$ e minimizar os argumentos $\mathbf{w}$, b , dada por (HAYKIN, 2009):

$$Q(\alpha) = \sum_{i=1}^N \alpha_i - \frac{1}{2} \sum_{i=1}^N \sum_{j=1}^N \alpha_i \alpha_j y_i y_j \mathbf{x}_i \cdot \mathbf{x}_j \quad (2.6)$$

sujeita à

$$\sum_{i=1}^N \alpha_i y_i = 0, \quad i = 1, \dots, N \quad (2.7)$$

$$0 \leq \alpha_i \leq C, \quad i = 1, \dots, N \quad (2.8)$$

onde C é um hiperparâmetro de regularização dado pelo usuário. Tendo determinado a solução ótima dos α_i , $i = 1, \dots, N$ multiplicadores de Lagrange, a função de decisão é obtida como:

$$y(\mathbf{x}) = \sum_{i=1}^N \alpha_i y_i \mathbf{x}_i^T \mathbf{x} + b \quad (2.9)$$

Tendo em vista que o algoritmo do SVM foi estruturado inicialmente para classificações binárias, o emprego de funções de núcleo, também conhecidas como funções de *Kernel* $k(\cdot, \cdot)$, se faz necessária nos casos em que não é possível estabelecer uma relação linear entre as variáveis. Isso ocorre devido ao Teorema de Cover (COVER, 1965), que afirma que um espaço não-linearmente separável pode ser transformado em um novo espaço, onde os padrões têm alta probabilidade de tornarem-se linearmente separáveis, através de geometria clássica combinatória, sob duas condições: a transformação deve ser não-linear e a dimensão do espaço de características deve ser maior em relação à dimensão do espaço de entrada. Dessa forma, a função de *Kernel* aplica uma transformação nas variáveis independentes e a SVM realiza uma mudança de dimensionalidade para um espaço de ordem mais elevada, se transformando então em um problema de classificação linear (GONCALVES, 2010), conforme mostrado na Figura 10.

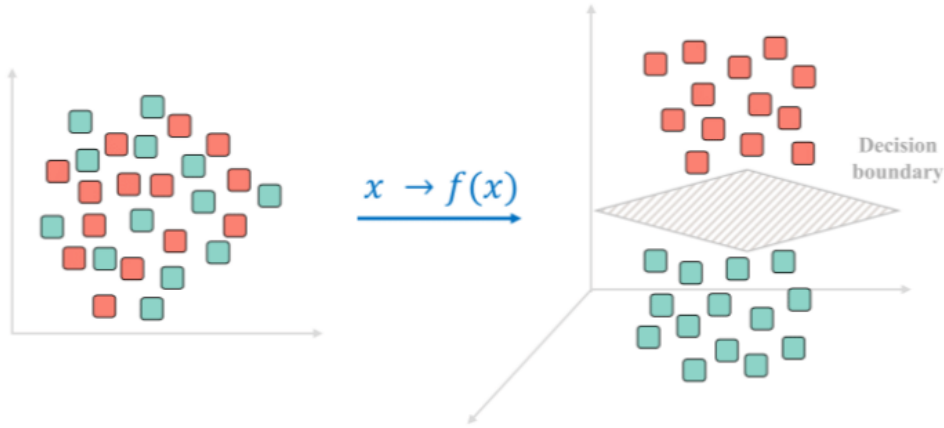


Figura 10 – Aplicação da função *Kernel* em conjuntos de dados linearmente não-separáveis. Fonte: Imagem retirada de (SHEYKHOUSA et al., 2020).

As Equações 2.6 e 2.9 podem ser reescritas da seguinte maneira quando uma função de *Kernel* é empregada no processo (DU; SWAMY, 2014):

$$Q(\alpha) = \sum_{i=1}^N \alpha_i - \frac{1}{2} \sum_{i=1}^N \sum_{j=1}^N \alpha_i \alpha_j y_i y_j K(\mathbf{x}_i \cdot \mathbf{x}_j) \quad (2.10)$$

$$y(\mathbf{x}) = \sum_{i=1}^N \alpha_i y_i K(\mathbf{x}_i, \mathbf{x}) + b \quad (2.11)$$

Funções *Kernel* que são frequentemente utilizadas são a base radial, linear, polinomial e sigmoide (KANCHERLA; BODAPATI; N, 2019). O *Kernel* Linear é aplicável em conjuntos

de dados que, ao expandir o domínio para uma nova representação, tornem-se linearmente separável ao definir um hiperplano entre eles. A função de *Kernel* polinomial possui um bom desempenho global e a dimensão da função de mapeamento cresce com o grau polinomial n , contudo a complexidade computacional também aumenta. A função de *Kernel* base Radial (em inglês, “*Radial Basis Function*” - RBF) é a forma que possui a maior generalização e é altamente utilizada devido a sua semelhança com a distribuição gaussiana. Adicionalmente, a função calcula a similaridade de dois pontos X_1 e X_2 . A função sigmoide, também conhecida como função de núcleo de tangente hiperbólica, se aproxima de uma rede neural de duas camadas e tem uma excelente precisão de classificação (REN; HU; MIAO, 2016; AL-MEJIBLI et al., 2018). As transformações e seus hiperparâmetros estão definidos na Tabela 1 abaixo.

Tabela 1 – Diferentes funções Kernel para SVM

	Formula	hiperparâmetros
Linear	$K(x, x_i) = x \cdot x_i$	-
Polinomial	$K(x, x_i) = [\gamma(x \cdot x_i) + cte]^n$	γ , cte , n
Base Radial	$K(x, x_i) = \exp(-\gamma \ x - x_i\ ^2)$	γ
Sigmoide	$K(x, x_i) = \tanh(\gamma(x \cdot x_i) + cte)$	γ , cte

O desempenho da SVM está diretamente relacionada à escolha da função de *Kernel*, do hiperparâmetro de penalidade C e do hiperparâmetro γ . O hiperparâmetro C controla o balanço entre os valores classificados incorretamente e maximização de margem ao hiperplano. Com um hiperparâmetro C maior, obtemos uma margem do hiperplano de maior valor, o que pode gerar uma alta penalidade para os pontos não separáveis ou um aumento na quantidade de vetores-suporte usados para construir a função de decisão, gerando um sobreajuste do modelo. Caso o hiperparâmetro C seja muito pequeno, pode-se gerar um modelo que desempenha subajuste. O hiperparâmetro γ está presente na função de algumas funções de *Kernels* e define a curvatura em um limite de decisão. Um gama alto significa um limite de decisão com uma curvatura mais acentuada e, consequentemente, um gama baixo significa um limite de decisão com uma curvatura menos acentuada (GÉRON, 2019).

Quanto a utilização da técnica na área de ensaios não destrutivos, a SVM vem sendo utilizada para a identificação, classificação e quantificação de defeitos. Existem contudo, apenas alguns poucos estudos relacionados à PEC. Diversos pesquisadores já utilizaram SVM no auxílio de localizações e classificações de defeitos, Chen et al. (2014) bem como em conjunto com algoritmos de extração de características para localização e classificação de defeitos em estrutura multicamadas. O trabalho de He et al. (2013a) utilizou PCA para a geração dos dados de entrada da SVM, com a finalidade de identificar e classificar automaticamente defeitos localizados em diferentes regiões entre camadas, com variações do efeito de *lift-off* em ligas de alumínio fabricadas para simular as estruturas de aeronaves. Posteriormente He et al. (2013b) propuseram um equipamento de inspeção a partir da integração de correntes parasitas e investigaram a sua utilização na classificação automatizada de defeitos.

Outra vertente da utilização de SVM associadas às correntes parasitas pulsadas é a

classificação de espessura de chapas metálicas. Ramos et al. (2014) utilizaram uma SVM multi-classe para estimar diferentes espessura de chapas de aço inoxidável e alumínio, variando de 1 mm à 7 mm. Em pesquisas mais recentes Angani et al. (2016) utilizaram as frequências de pico das componentes real e imaginária da DFT como parâmetros de entrada da SVM na análise de chapas de aço inoxidável com espessuras variando de 0.8 a 4 milímetros e *lift-off* variando de 0.5 à 2 milímetros.

Os estudos citados apresentam uma alta taxa de sucesso na utilização de máquina de vetores de suporte combinados à técnica de correntes parasitas pulsadas para caracterização de defeitos. Contudo, ainda perdura uma lacuna na sua utilização para a avaliação da estrutura metálica de materiais.

2.3.2 K-Nearest Neighbors

O classificador K vizinhos mais próximos (do inglês “*k-Nearest Neighbors*” - KNN) é um algoritmo de classificação não paramétrico e de fácil implementação. Tem como princípio a ideia de que, proposto um conjunto de dados dispostos em um espaço geométrico, aqueles indivíduos que apresentarem similaridade irão possuir distâncias entre si menores (KUNCHEVA, 2004). Dessa forma, o algoritmo inicialmente define agrupamentos com os exemplos de treinamento e, para um novo exemplo x' , busca identificar os k indivíduos mais próximos e associá-lo à classe majoritária dos seus vizinhos, conforme apresentado na Figura 11.

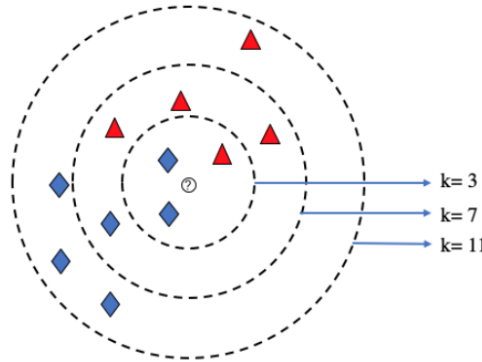


Figura 11 – Exemplo do uso do modelo KNN para a definição da classe do ponto central de acordo com os k indivíduos mais próximos, para k igual a 3, 7 e 11. Fonte: Imagem retirada de (ZHANG et al., 2020).

A teoria por trás da classificação é extremamente simples. Dado um conjunto de treinamento $D = \{(x_i, y_i) | i=1, \dots, N\}$, temos que $x_i \in \mathbf{R}^n$ são vetores de características para um determinado exemplo, n é o número de características existentes e $y_i \in \{c_1, \dots, c_m\}$, onde $\mathbf{c}$ é o vetor com todas as possíveis classes. O algoritmo da KNN encontra k indivíduos do grupo de treinamento que apresentem a menor distância para um novo ponto não categorizado $x' \in \mathbf{R}^n$, a partir de uma função de distância estabelecida. A escolha mais comum é generalização da distância de Manhattan e da distância Euclidiana, conhecida como Distância de Minkowski:

$$dist(x_i, x') = \left(\sum_{j=1}^n |x_{ij} - x'_{ij}|^p \right)^{1/p}, \quad p \geq 1 \quad (2.12)$$

onde o valor de p é escolhido por experimentação e para $p = 1$ obtemos a distância de Manhattan ou L1 e $p = 2$, a distância Euclideana ou L2. Os respectivos rótulos y_i para os k indivíduos são agregados e, por fim, y' é escolhido a partir da classe predominante no grupo avaliado.

O KNN não possui rotinas extensas de otimização e/ou ajuste de diversos parâmetros durante a etapa de treinamento. A quantidade de hiperparâmetros são reduzidos, sendo o valor da constante k a variável de maior peso na performance do algoritmo. O seu valor ótimo pode ser encontrado utilizando rotinas de tentativa e erro.

A técnica é utilizada em diversas aplicações de reconhecimento de padrões em áreas como classificação textual (SHAH et al., 2020), bioinformática (MURUGAN; NAIR; KUMAR, 2019; SHABAN et al., 2020) e detecção facial (DINO; ABDULRAZZAQ, 2019). Em pesquisas associadas a avaliações não destrutivas, a sua utilização também é bem sucedida: o algoritmo de KNN é utilizado para a detecção de áreas de corrosão de diversos tamanhos e profundidade em rebites de aeronaves com a técnica de EC (LE et al., 2022) e para a interpolação das respostas das técnicas de vazamento de fluxo magnético e ultrassom em medições de corrosão para tubulações industriais (HAMED et al., 2015). A sua utilização também foi estudada para a detecção de carburização presente em tubos de aço HP retirados de fornos de pirólise a partir da análise do espectro de frequência de sinais de ultrassom (RODRIGUES et al., 2019).

2.3.3 Árvore de Decisão

O algoritmo árvore de decisão (em inglês, “*Decision Tree*”) é categorizado como não-paramétrico e pode ser utilizado para classificação de grupos de dados ou regressão de modelos. Pertence aos mais simples e intuitivos algoritmos de aprendizagem da máquina e sua implementação é de fácil representação, possui baixo custo computacional e produz modelos de alta qualidade. É considerado uma das abordagens mais poderosas e populares na área de ciência de dados e é utilizado como base para a construção de diversos modelos mais robustos (ROKACH, 2016).

São construídas usando um algoritmo que estratifica repetidamente os dados em grupos menores baseados na etiqueta de classe (KINGSFORD; SALZBERG, 2008). A partir de uma análise do conjunto de dados, feita a escolha de uma característica que melhor divide os objetos em dois grupos a partir de uma métrica pré-estabelecida, seguindo-se da primeira divisão. Esta separação inicial é chamada de nó-raiz da árvore de decisão e suas ramificações são chamadas de nó-interno ou nó de decisão. Na etapa subsequente, um novo limite de decisão é criado para cada nó. O procedimento é conduzido recursivamente até uma divisão satisfatória, onde são gerados os nós-folha, ou nó resposta, que contém a decisão da classe referente ao caminho percorrido. O algoritmo termina quando a base da árvore consiste inteiramente de nós-folha. Geralmente, uma árvore de profundidade p pode produzir até 2^p regiões (BIEDRZYCKI, 2021). Um exemplo de árvore de decisão e sua divisão em um conjunto de dados $D \subset \mathbf{R}^2$ é mostrado na Figura 12.

As divisões são feitas com base em uma medida de impureza, onde as classificações incorretas são quantificadas em cada nó. Para um melhor entendimento, denominaremos de nós filhos as ramificações de um nó de decisão. Para determinar a variável de decisão de um nó, uma faixa de valores referente a cada característica do conjunto é testada. Para cada teste, a impureza

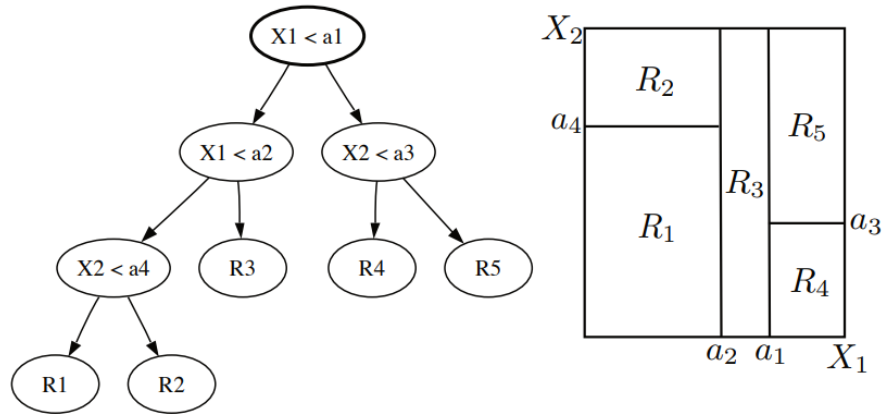


Figura 12 – Exemplo de uma árvore de decisão com perguntas numéricas baseadas em duas variáveis X_1 e X_2 . Cada nó folha é definido com a região do plano $X_1 \times X_2$. Fonte: Imagem retirada de (MOHRI; ROSTAMIZADEH; TALWALKAR, 2018).

total é calculada considerando a média ponderada da impureza dos seus nós filhos, onde aquela partição que apresentar o menor grau de impureza será a escolhida. Este processo é repetido para cada nó da árvore até que a impureza em um nó de decisão seja inferior à média ponderada da impureza em seus nós filhos. Quando isto acontece, na maioria dos casos o nó é considerado um nó folha (KINGSFORD; SALZBERG, 2008).

A qualidade da divisão é baseada no ganho de informação, que é a medida que quantifica a mudança de impureza após a divisão. Em resumo, a construção da árvore de decisão é baseada na procura do atributo que irá retornar a menor soma de impurezas de cada região, maximizando o ganho de informação. Sua fórmula para uma divisão em um determinado nó de decisão é dada por (MOHRI; ROSTAMIZADEH; TALWALKAR, 2018):

$$GI(D_p, f) = I(D_p) - \sum_{r=1}^R \frac{|D_r|}{|D_p|} I(D_r) \quad (2.13)$$

onde D_p é o conjunto de dados do nó de decisão, f é a variável de decisão, R é o número de folhas, D_r é o conjunto de dados nos nós filhos de índice r e $I(D)$ é a medida de impureza no conjunto D . Há várias medidas de impureza utilizadas e aquelas mais comumente empregadas são a Impureza de Gini e Entropia.

A impureza de Gini mede a frequência com que um elemento aleatório do conjunto de dados será incorretamente classificado. A impureza de Gini para uma região r criada a partir de uma divisão pode ser definida como:

$$I_G = \sum_{c=1}^C p_c(1 - p_c) \quad (2.14)$$

onde p_c é a probabilidade de escolha da classe c no nó e C é o número total de classes. O valor da impureza de Gini varia de 0 à 0,5, sendo 0 o valor que indica quando todos os indivíduos

pertencem a mesma classe e 0,5 quando os indivíduos estão uniformemente distribuídos entre as classes.

A entropia é uma medida similar ao Gini, variando de 0 à 1 e quantifica a falta de homogeneidade dos dados de entrada em relação a sua classificação. Sua equação é dada por:

$$I_E = - \sum_{c=1}^C p_c \log_2 p_c \quad (2.15)$$

As árvores de decisão podem ser usadas para descobrir características e extrair padrões em grandes bases de dados. Estas características, juntamente com suas interpretação intuitiva, são algumas das razões do porquê o algoritmo tem sido usado extensivamente para diversas frentes de pesquisa da bioinformática, finanças, análises preditivas e avaliações de empresas (JENA; DEHURI, 2020; CHARBUTY; ABDULAZEEZ, 2021). Em áreas associadas a avaliações de defeitos por técnicas não destrutivas, a utilização da árvore de decisão e algoritmos derivados já é uma prática bem estabelecida. Em pesquisas recentes utilizando o sinal de PEC, Liu et al. (2019) empregou o algoritmo para a detecção de defeitos inseridos em estruturas metálicas de aeronaves a partir de características selecionadas do domínio do tempo, enquanto Jin et al. (2021) avaliou estruturas tubulares multicamadas utilizadas em poços de petróleo e gás com o auxílio da PCA.

2.3.4 Perceptron Multicamadas

A rede do tipo perceptron multicamadas (do inglês, “*Multi-layer Perceptron*” - MLP) é uma das arquiteturas mais populares entre os modelos de redes neurais artificiais (KUNCHEVA, 2004). São estruturas composta por conjuntos de camadas estruturada em níveis. A rede é alimentada com o conjunto de dados a ser avaliado no primeiro nível, chamado de camada de entrada. Os níveis intermediários são chamados de camadas ocultas e podem ser compostos por uma ou mais camadas. O último nível é chamado de camada de saída, e tem uma unidade para cada valor de saída da rede ou classe avaliada. Todas as camadas, exceto pela saída, possuem pesos associados à elas e são completamente conectadas com o próximo nível.

O perceptron multicamadas se enquadra na categoria de algoritmos *feedforward*. Na camada oculta, os sinais de entrada são combinados linearmente com os pesos sinápticos $w_1, w_2, \dots, w_n$ e somados a um valor de *bias* (viés). O resultado é, então, submetido a uma função matemática de ativação $\varphi(\cdot)$ responsável associar o resultado da combinação a um certo intervalo e produzindo o valor de saída v_n . Uma vez que cada camada está conectada ao nível seguinte, o resultado é propagado para o próximo nível. Na camada de saída, as informações são avaliadas a partir de uma soma ponderada com a adição de um viés e, novamente submetidas a uma função de ativação para gerar a saída da rede, conforme mostrada na figura 13.

Com exceção da camada de entrada, todas as outras camadas são constituídas por neurônios e, portanto, apresentam uma alta capacidade computacional. São capazes de aproximar qualquer função contínua, restrita a uma região compacta do espaço de aproximação, dado um número suficiente de neurônios na camada intermediária (DU; SWAMY, 2014). Estas

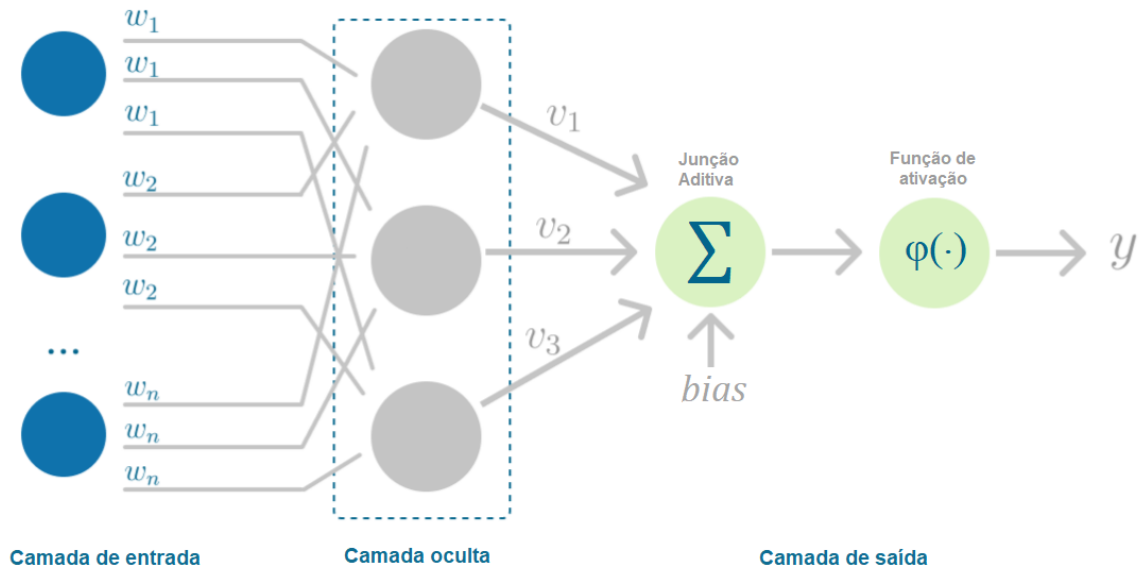


Figura 13 – Exemplo de uma rede perceptron multicamadas com três camadas. Fonte: Imagem adaptada de (BENTO, 2021).

redes apresentam resultados satisfatórios para uma grande variedade de aplicações tais como classificação de padrões, previsão e aproximação de funções.

A rede foi utilizada para a avaliação dos sinais de PEC e ECT em diversas pesquisas. Para a detecção de corrosão de um tubo de aço carbono isolado termicamente, Silva Jr. et al. (2021) utilizou um sistema eletrônico inteligente para gerar, adquirir e processar os sinais PEC. Na etapa de processamento, duas arquiteturas de classificadores foram consideradas: perceptron multicamadas e máquina de aprendizagem extrema. As rotinas com o MLP apresentaram os menores tempos de processamento e as maiores acurácias. A rede também foi utilizada para a avaliação de discontinuidades em camadas intermediárias de estruturas aeronáuticas (LIU et al., 2019), para a avaliação de estresse e tensões residuais (HABIBALAH et al., 2015) e caracterização e quantificação de defeitos (JO; LEE, 2009; GHONI et al., 2014; BUI; PHAM; CUNG, 2022).

2.3.5 Gradiente Boosting

O gradient boosting é um algoritmo de aprendizado de máquina supervisionada utilizado em tarefas de classificação e regressão, sendo incluído dentro do grupo de classificadores *ensemble*. É baseado na ideia da construção gradual de um modelo através de avaliações estatísticas e do método do gradiente (NATEKIN; KNOLL, 2013) com o intuito de minimizar uma função de erro associada ao problema estudado. A partir de um modelo inicial de previsão fraca, seu objetivo principal é acrescentar novos modelos para o conjunto, de forma sequencial, que buscam reduzir o erro do modelo anterior a partir do método do gradiente descendente para a minimizar o problema (KEPRATE; RATNAYAKE, 2017). Na construção do classificador, geralmente, são utilizadas árvores de decisão como modelo fraco de base. Estes modelos são conhecidos como *gradient boosting decision trees* (GBDT). Um resumo da rotina do algoritmo é apresentado abaixo (HASTIE; TIBSHIRANI; FRIEDMAN, 2009).

Dado um conjunto de treinamento $D = \{(x_i, y_i) | i = 1, \dots, N\}$, e uma função de erro $L(y_i, F(x_i))$, o modelo é inicializado para uma estimativa inicial F_0 que minimize a função de erro, onde y_i é referente ao i -ésimo valor observado e γ é o valor predito. Este processo irá definir a primeira árvore de decisão:

$$F_0(x) = \underset{\gamma}{\operatorname{argmin}} = \sum_{i=1}^N L(y_i, \gamma) \quad (2.16)$$

Neste ponto, o modelo é definido por $F_M(x) = F_0(x)$. Esta função no entanto, está sujeita a atualizações até que todos os classificadores da base M sejam adicionados a ela com pesos adequados. Para a construção das árvores subsequentes, é implementado um ciclo fundamentado nos erros cometidos pelos modelos anteriores. Este erro é dado pela diferença entre o valor observado e o valor predito da variável de interesse, chamado de resíduo:

$$r_{i,m} = \left[\frac{\partial L(y_i, F_{m-1}(x_i))}{\partial F_{m-1}(x_i)} \right], \text{ para } i = 1, 2, \dots, n \quad (2.17)$$

onde $F_{m-1}(x_i)$ refere-se ao valor predito pela árvore criada no passo anterior. Dessa forma, é modelada uma nova árvore para os valores $r_{i,m}$, criando as $R_{j,m}$ folhas para $j = 1, \dots, J_m$. Neste momento, o m -ésimo classificador ainda não foi adicionado.

Para cada folha da árvore $F_m(x)$, calculamos os seus valores de saída com o objetivo de encontrar o γ – *ótimo* a partir do seguinte problema de otimização:

$$\gamma_{j,m} = \underset{\gamma}{\operatorname{argmin}} = \sum_{i=1}^N L(y_i, F_{m-1}(x_i) + \gamma) \quad (2.18)$$

O ciclo utiliza esta nova árvore para atualizar a predição de cada indivíduo. Este processo é realizado através da equação:

$$F_m(x) = F_{m-1}(x) + \nu \sum_{j=1}^{J_m} \gamma_{j,m} I(x \in R_{j,m}) \quad (2.19)$$

A nova previsão $F_m(x)$ é baseada na anterior $F_{m-1}(x_i)$ e na árvore recém criada. O somatório indica que os valores de saída para todas as folhas $R_{j,m}$ sejam somados em função de ν , que é o coeficiente de aprendizado que vai escalonar a árvore. Ao fim do ciclo, $F_M(x)$ irá retornar e esta função que será utilizada para fazer as novas predições. Após M iterações, obtemos o modelo final $F_M(x)$.

A partir do modelo principal, algumas rotinas derivadas foram criadas e já são amplamente utilizadas em aplicações distintas. Entre eles, o *light gradient boosting machine* ou apenas *lightGBM*, utilizado nesta pesquisa como um dos modelos de classificadores de base, é uma ferramenta originalmente desenvolvida pela *Microsoft* que utiliza *GBDT*. Devido à sua eficiência, precisão e interpretabilidade, o *GBDT* alcança alto desempenho em muitas tarefas de aprendizado de máquina, como classificação multiclasse e predição (KE et al., 2017).

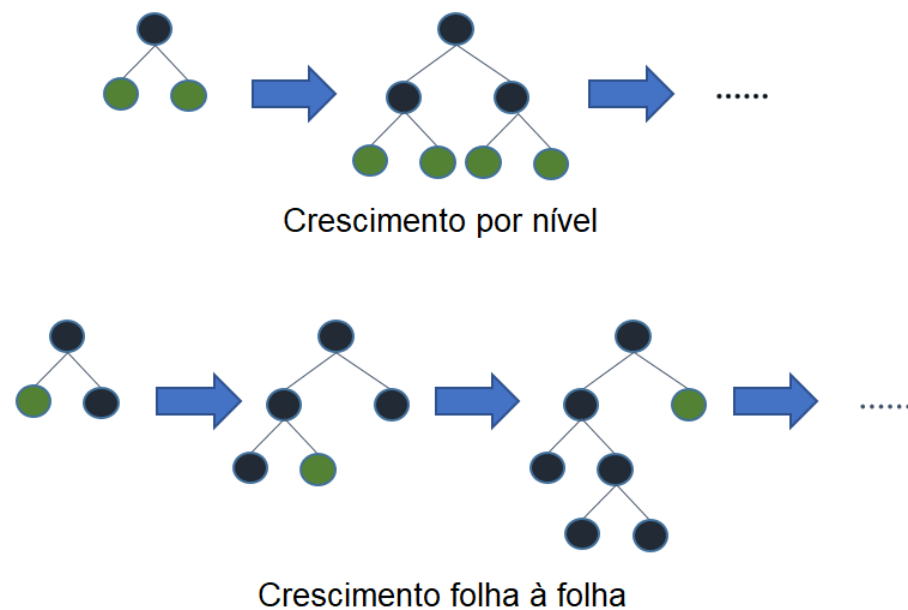


Figura 14 – Comparação entre algoritmos convencionais de GBDT e lightGBM. Fonte: Imagem adaptada de (Microsoft Corporation, 2022).

O *lightGBM* divide a árvore folha por folha, ao contrário de outros algoritmos de GBDT que desenvolvem a árvore por nível (Microsoft Corporation, 2022), conforme mostrado na Figura 14. Ele seleciona a folha que irá minimizar a função de erro para concentrar o seu crescimento. Dessa forma, a estrutura continua a crescer com os ramos e folhas mais promissoras. O seu desenvolvimento visa melhorar a escalabilidade do processo de *gradient boosting* para dados altamente dimensionais. Com ele é possível acelerar o treinamento, reduzir o consumo de memória e combinar diferentes modelos para maximizar o aprendizado paralelo (WANG; ZHANG; ZHAO, 2017).

2.3.6 Regressão Logística

A regressão logística (do inglês, “*logistic regression*” - LR) é um dos mais populares modelos de Aprendizagem de Máquina, que se enquadra na técnica de aprendizagem supervisionada (DREISEITL; OHNO-MACHADO, 2002). É um algoritmo de estimativa de classe resultante da aplicação da função logística (também chamada de sigmoide) sobre a regressão linear para a construção de um modelo que consiste em prever o valor de um atributo desejado, chamado variável dependente, com base nos valores de outros atributos, chamados variáveis independentes.

Sua principal diferença em relação à regressão linear está na previsão de uma variável dependente categórica, ou seja, pode assumir apenas um número dentro de um conjunto limitado de possibilidades. Dessa forma, é utilizado para a análise e classificação de conjuntos de dados de resposta binária e proporcional. Pode também fornecer probabilidades e se estender a problemas de classificação multiclasse (MAALOUF, 2011).

Segundo Mesquita (2014), embora a regressão logística fosse inicialmente utilizada na área médica para classificação de pacientes saudáveis e doentes, a eficiência desta técnica viabilizou

sua implementação nas mais diversas áreas do conhecimento, avaliações geopolíticas, estudos em bioinformática, pesquisa de mercado e, inclusive, em pesquisas associadas à áreas preditivas. Em áreas associadas à aeronáutica, [Stott et al. \(2015\)](#) e [Stott \(2014\)](#) utilizam a regressão logística para a avaliação do modelo aplicado na detecção de falhas em peças de alumínio multicamadas utilizando PEC e [Brence e Brown \(2002\)](#) utiliza o modelo de aprendizado de máquina LR para a avaliação de corrosão utilizando ECT.

2.3.7 Naive Bayes Gaussiano

O modelo de Naive Bayes Gaussiano é um algoritmo de classificação estatística do tipo supervisionado. Tendo como principal premissa a desconexão entre as variáveis do problema, o modelo realiza uma classificação probabilística de observações.

Dado um vetor de características $X = (x_1, x_2, \dots, x_n, \dots, x_N)$ e uma classe y , o Teorema de Bayes é expresso por:

$$P(y|X) = \frac{P(X|y)P(y)}{P(X)} \quad (2.20)$$

onde $P(y|X)$ é a probabilidade *a posteriori* de y condicionada a X , $P(X|y)$ é a probabilidade *a posteriori* de X condicionada a y e $P(X)$ e $P(y)$ são as probabilidades *a priori* de X e y . Dessa forma, para cada novo indivíduo não rotulado, o algoritmo prevê uma classe $\hat{y}$ com base na probabilidade *a posteriori* da classe y :

$$\hat{y} = \underset{y \in Y}{\operatorname{argmax}} P(y) \prod_{n=1}^N (x_n|y) \quad (2.21)$$

Uma vantagem do algoritmo de Naive Bayes é que ele requer apenas um pequeno número de dados de treinamento para estimar os parâmetros necessários em classificações. Adicionalmente, o algoritmo mostra uma alta previsão de desempenho e obtém resultados comparáveis a outras técnicas de classificação, tais como árvores de decisão e redes neurais ([ONAN; KORUKOĞLU; BULUT, 2016](#)). Esta suposição foi confirmada no trabalho de [Qiu et al. \(2014\)](#), onde estes dois algoritmos foram comparados ao Naive Bayes na avaliação da impedância da bobina de ECT para a classificação de defeitos em estruturas de aeronaves. Em pesquisas relacionadas à utilização de correntes parasitas pulsadas, [Qiu et al. \(2013\)](#) utilizou o algoritmo associado à PCA para a identificação de rachaduras e cavidades inseridas em um corpo de prova de forma superficial e subsuperficial, apresentando resultados promissores.

2.3.8 Critérios de Avaliação de Desempenho

Com a finalidade de realizar análises dos resultados experimentais obtidos, algumas métricas são utilizadas para avaliar quantitativamente a capacidade de erro e acerto dos modelos. A forma mais natural para o cálculo do erro em modelos classificadores é encontrando a proporção dos objetos classificados incorretamente ([KUNCHEVA, 2004](#)). Para um classificador D , de tamanho $N \times n$ com vetor de saída $\hat{y}$ e resposta real y , o erro(Err) é definido por:

$$Err(D) = \frac{1}{N} \sum_{i=1}^N I(y_i, \hat{y}_i) \quad (2.22)$$

em que $I(y_i, \hat{y}_i)$ é um indicador de variável, sendo igual a 1 em caso de $y_i = \hat{y}_i$ e igual a zero caso $y_i \neq \hat{y}_i$. De forma complementar, a Acurácia (Acc) é a avaliação do percentual de vezes que os modelos apresentaram a saída correta, independente da classe, e é calculado por $1 - Err(D)$. Para modelos multi-classe, onde o conjunto de entrada possui N amostras e K classes, e o vetor de saída $\hat{y}_{i,k}$ e resposta real $y_{i,k}$ são as observações i da classe k, a acurácia pode ser definida como:

$$Acc(D) = \frac{1}{N} \sum_{i=1}^N \sum_{k=1}^K I(y_{i,k}, \hat{y}_{i,k}) \quad (2.23)$$

Para facilitar o entendimento e a visualização dos acertos e erros da saída do classificador, a matriz de confusão é utilizada para relacionar as contagens em que cada classe é classificada de forma correta e incorreta. Ela é dividida em 4 parâmetros principais: O parâmetro Verdadeiros Positivos (VP) referem-se às instâncias positivas que foram corretamente classificadas. O parâmetro Falso Positivo (FP) representam as instâncias negativas que foram incorretamente rotuladas como positivas. O parâmetro Verdadeiro Negativo (VN) referem-se às instâncias negativas que foram corretamente previstas pelo classificador e o Falso Negativo (FN), por sua vez, instâncias positivas que foram incorretamente rotuladas como negativas. A sua distribuição está mostrada na Figura 15. As colunas da matriz de confusão representam as instâncias de uma classe prevista, enquanto as linhas demonstram as classes reais.

		CLASSE PREDITA	
		P	N
CLASSE REAL	P	VP	FN
	N	FP	VN

Figura 15 – Matriz de Confusão.

De posse dessas informações é possível se obter outras métricas para avaliações mais apuradas sobre possíveis erros de classificação. Precisão é a medida que quantifica, de todos os dados classificados como positivos, quantos são realmente positivos. Em contrapartida, o *recall*, também chamado de sensibilidade ou taxa de verdadeiro positivo, avalia a porcentagem de dados classificados como positivos quando analisado a população real de positivos que existem no conjunto de dados avaliado. O *F1-score* é a média harmônica entre precisão e a sensibilidade (DENG et al., 2016) e é uma forma de avaliar as duas métricas simultaneamente. Suas fórmulas são dadas por:

$$Precisão = \frac{VP}{VP + FP} \quad (2.24)$$

$$Recall = \frac{VP}{VP + FN} \quad (2.25)$$

$$F1 \text{ score} = 2 \times \frac{Precisão \times Recall}{Precisão + Recall} \quad (2.26)$$

2.3.9 Técnica de Validação Cruzada

Em algoritmos de aprendizado supervisionado, o ajuste de hiperparâmetros é norteado pelo erro medido entre a saída da máquina e a saída desejada. O objetivo destas avaliações é encontrar um classificador que apresente um erro mínimo a partir de um processo de aperfeiçoamento com análises apuradas e um grande número de testes para cada tipo de *datasets* trabalhados (PRATES, 2020). Contudo, é importante considerar que o modelo não deve aprender a exata representação do conjunto utilizado para os testes, mas sim construir um modelo estatístico que seja capaz de generalizar uma solução. Dessa forma, além de minimizar o erro, a escolha do classificador e do melhor conjunto de hiperparâmetros deve lidar com situações não apresentadas durante seu desenvolvimento.

Para que o modelo escolhido atenda estas duas particularidades, é comumente utilizado a técnica de validação cruzada (do inglês, *Cross-Validation* – CV) para otimização de hiperparâmetros. Este procedimento utiliza a reamostragem para estimar o erro de teste associado a um modelo de aprendizagem de máquina para uma dada combinação de hiperparâmetros. Dessa forma, é possível utilizá-lo para encontrar os melhores valores de hiperparâmetros que produzem um desempenho de generalização satisfatório.

O método de validação cruzada denominado *k-Fold* divide o conjunto de treinamento em k subconjuntos (ou *Folds*) mutuamente exclusivos do mesmo tamanho. Inicialmente, o primeiro subconjunto é separado para ser utilizado como um conjunto de validação e os $k - 1$ subconjuntos restantes são utilizados para treinamento dos modelos escolhidos. Finalizada a análise, o conjunto de validação passa a ser o conjunto adjacente e uma nova rodada de treinamento é iniciada. O processo de validação cruzada é repetido k vezes, de modo que a cada repetição, um dos k subconjuntos sejam utilizados como dado de teste para validação do modelo exatamente uma vez. Esse processo resultará em k estimativas do erro de teste. A Figura 16 apresenta essa rotina para $k = 5$ (DU; SWAMY, 2014). O resultado final desse processo é o desempenho médio do classificador nos k testes. O objetivo de repetir os testes múltiplas vezes é aumentar a confiabilidade da estimativa da precisão do classificador, produzindo subconjuntos de teste e validação distintos para um conjunto inicial limitado (PEDREGOSA et al., 2011).

2.4 Ensemble de classificadores

Visando a contínua melhoria de métodos de aprendizado de máquina, os métodos *ensemble* tem sido amplamente estudado nos últimos anos como uma alternativa para aumentar a eficiência e precisão de rotinas (SAGI; ROKACH, 2018). São compostos por conjuntos de classificadores

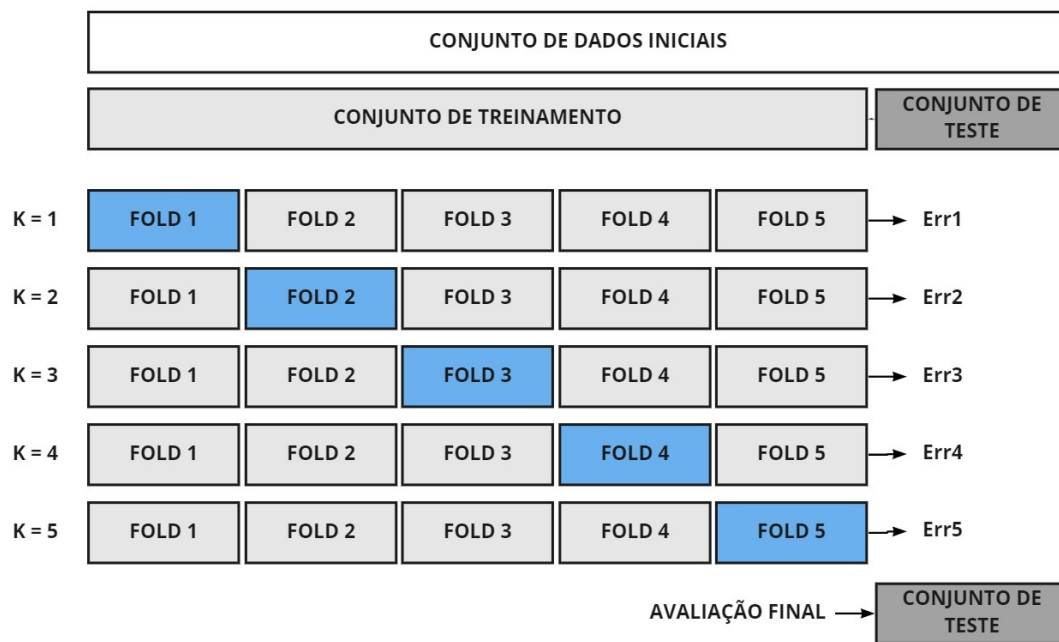


Figura 16 – Método de validação cruzada denominado k-Fold para um K=5.

menores e/ou específicos, que através de combinações sequenciais ou independentes, criam um classificador mais robusto. Tal algoritmo decorrente da fusão de modelos possibilita uma maior compreensão e confiabilidade à análise de dados do que uma decisão que de outra forma seria baseada somente em uma das fontes (PARIKH et al., 2004). Esta metodologia flerta com a ideia apresentada por Surowiecki (2004) no seu livro “Sabedoria das Multidões” (do inglês “*The Wisdom of the Crowds*”): a agregação de informação em grupos resulta em decisões que, são quase sempre melhores do que as que poderiam ser feitas por qualquer membro do grupo, quando estes membros apresentam diversidade de opiniões, independência, descentralização do conhecimento e uma agregação adequada das decisões individuais.

A diversidade dos classificadores de base permite que diferentes limites de decisão sejam gerados pelo uso de parâmetros de treinamento diversos, podendo considerar aspectos distintos e, em simultâneo, relevantes para a solução de um dado problema. A ideia por trás da combinação é que cada classificador cometerá diferentes padrões de erro, e através da combinação estratégica desses modelos a variação e o viés de classificação podem ser reduzidos e a dependência dos resultados em características de um único conjunto de treinamento eliminado (OPITZ; MACLIN, 1999; KUNCHEVA, 2004). Dessa forma, ao invés de buscar o melhor conjunto de características e o melhor classificador, buscamos agora o melhor conjunto de classificadores e então o melhor método de combinação.

Um exemplo de um modelo de *ensemble* para a classificação de um conjunto com três classes é mostrado na Figura 17, onde podemos visualizar o ganho de desempenho que pode ser obtido com a aplicação deste algoritmo. As fronteiras de decisão de três classificadores são apresentadas nos modelos 1, 2 e 3. Cada classificador apresenta uma proposta de fronteira distinta, apresentando erros de classificação em diferentes indivíduos. No entanto, a combinação desses três classificadores em um *ensemble*, neste exemplo representado a através de voto majoritário, é

capaz de gerar uma nova fronteira de decisão capaz de separar corretamente todas as amostras do conjunto de dados em questão.

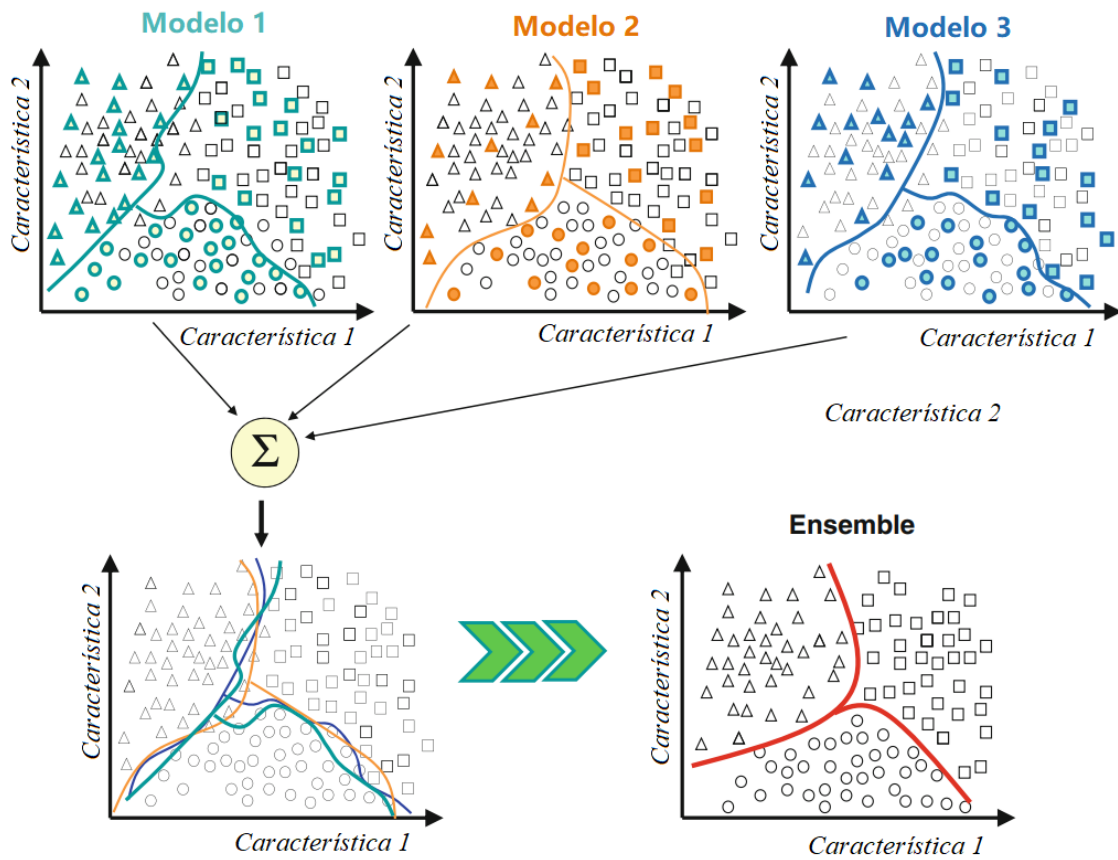


Figura 17 – Exemplo do ganho de desempenho obtido através do uso de *ensembles* para um problema de classificação multiclasse. Fonte: Imagem retirada de (ZHANG; MA, 2012).

Existem três razões majoritárias para explicar o porquê de um *ensemble* de classificadores apresentar, geralmente, um resultado melhor do que um classificador individual, apresentadas por Dietterich (2000): a primeira, chamada de estatística, surge quando a quantidade de dados de treinamento é muito pequena e um espaço de busca H que contém um elevado número de classificadores com precisões semelhantes e hipóteses diferentes. Ao invés de escolher um classificador, usando métodos de conjunto, o algoritmo pode calcular a média das hipóteses e reduzir o risco de escolher um classificador inadequado. Adicionalmente, pode encontrar uma boa aproximação para a hipótese verdadeira f . O segundo motivo, chamado de computacional, surge quando os dados de treinamento são suficientes, mas o algoritmo de aprendizado ainda está preso na ótica local. Em alguns algoritmos o processo de otimização é realizado a partir de inicializações aleatórias e, a depender do ponto escolhido, pode ser muito difícil computacionalmente para o modelo encontrar a melhor hipótese. Um conjunto no qual a busca local será executada a partir de diferentes pontos de partida pode fornecer uma melhor aproximação da hipótese f desconhecida do que qualquer um dos classificadores individuais. Por fim, a razão representacional surge quando a função da hipótese f não pode ser representada por nenhuma das hipóteses em H . Contudo, ao realizar uma soma ponderada das funções de h_n , pode ser possível expandir o espaço de funções representáveis H e englobar a hipótese f . Isso ocorre visto que uma população diversa

de funções permite uma exploração maior do espaço de busca e, conseqüentemente, uma maior eficiência na localização de soluções melhores (COELHO, 2006). Uma representação gráfica das três razões é apresentada na Figura 18.

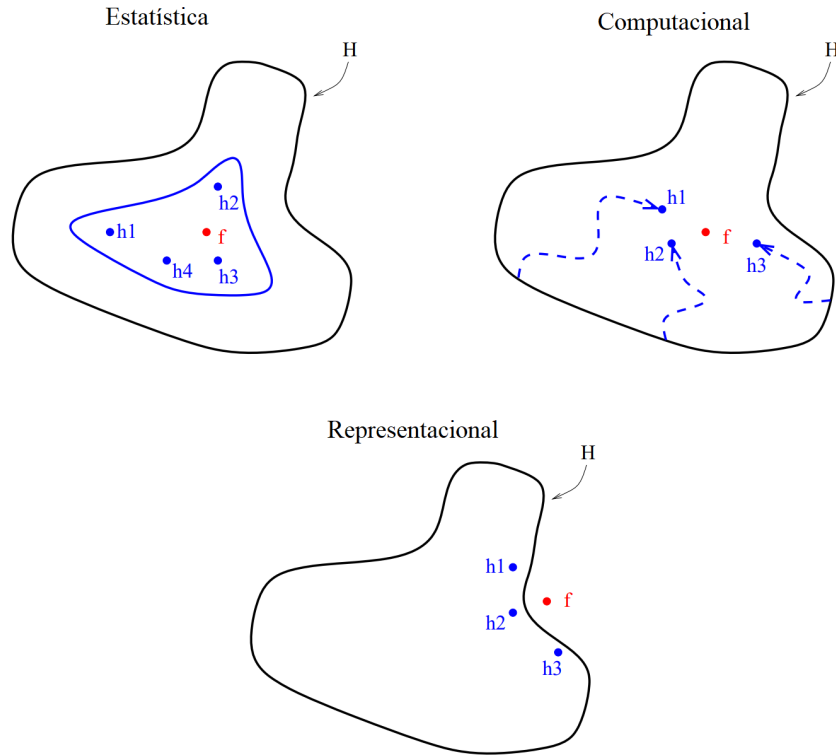


Figura 18 – Três razões fundamentais do porquê um *ensemble* de classificadores pode funcionar melhor do que um classificador individual. Fonte: Imagem retirada de (DIETTERICH, 2000).

A construção de um *ensemble* efetivo, segundo Zhang e Ma (2012), está baseada nas estratégias utilizadas em três pilares fundamentais:

- (i) Escolha de um conjunto de elementos candidatos a fazerem parte do *ensemble*;
- (ii) Seleção do algoritmo para combinação dos candidatos que contribuem positivamente em sua resposta;
- (iii) Escolha do método utilizado para a combinação das saídas geradas pelos classificadores base.

O primeiro pilar, representado pelo bloco (I) na Figura 19 objetiva encontrar candidatos promissores de forma a aumentar a diversidade de membros do conjunto. A diversidade dos classificadores é de suma importância em sistemas baseados em agrupamentos, visto que se todos os candidatos fornecem o mesmo resultado, não há ganhos com a sua combinação. Dois classificadores são chamados de diversos caso apresentem classificações incorretas e diferentes entre si para um novo conjunto de dados, não visto previamente (DU; SWAMY, 2014). Os modelos de um conjunto podem ser mais dependentes se eles compartilharem o mesmo algoritmo e/ou conjunto de dados usado no treinamento. Dessa forma, estratégias em diferentes níveis da construção do modelo são empregadas para que a diversidade nos candidatos seja alcançada.

Entre as mais comuns estão os métodos que atuam no nível dos dados de treinamento e os

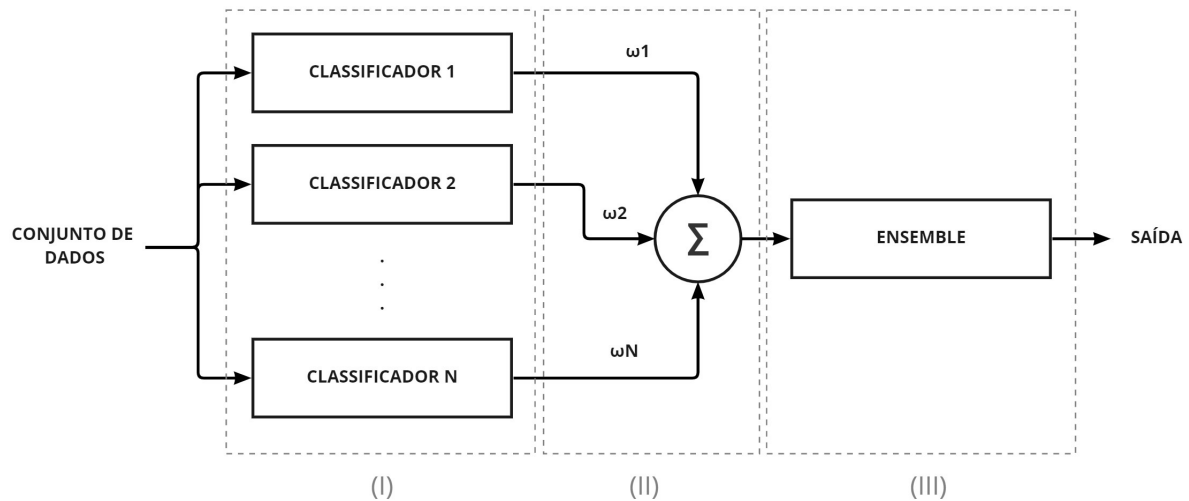


Figura 19 – Estrutura de um algoritmo de agrupamento.

métodos que atuam no nível dos classificadores de base. Para a diversificação do conjunto inicial, podem ser empregadas estratégias de amostragem e a utilização de diferentes subconjuntos de características, visando gerar componentes para a produção de mapeamentos diferentes. Para a diversificação dos classificadores de base existem métodos que buscam manipular as configurações de hiperparâmetros ou eleger modelos heterogêneos de forma que diferentes conjuntos de hipóteses sejam selecionados. Existem ainda, neste grupo, métodos que atuam sobre a forma de exploração do espaço de hipóteses, alterando a forma de exploração do espaço para a convergência em hipóteses heterogêneas, mesmo se apresentarem um ponto de partida único (ZHANG; MA, 2012; KUNCHEVA, 2004).

O *ensemble* é um algoritmo de meta-aprendizagem de máquinas que procura um melhor desempenho combinando as previsões de múltiplos modelos. Dessa forma, o segundo pilar, representado pelo bloco (II), refere-se a técnica de combinação dos classificadores de base. Embora haja um número ilimitado de combinações de agrupamento, três métodos dominam o campo da aprendizagem por *ensemble*: *bagging*, o *boosting* e o *stacking*.

Bootstrap aggregation, ou apenas *Bagging*, é um tipo de agrupamento homogêneo que busca a diversidade a partir da variação dos dados de entrada por amostragem *Bootstrap* (EFRON; TIBSHIRANI, 1993), utilizando subconjuntos distintos derivados de um mesmo conjunto de dados. Neste método geralmente utiliza-se apenas um algoritmo de aprendizado de máquina e cada modelo, independente em relação ao outro, é treinado com um subconjunto único retirado do conjunto original de dados. Seu principal algoritmo é a Floresta Aleatória (do inglês “*Random Forest*”). No *Boosting*, ou aprendizado por reforço, utiliza-se um conjunto único de dados de entrada e tem como objetivo criar modelos sequenciais e adaptativos com forte correlação, onde os pesos dos vetores do conjunto de dados de treinamento são modificados a cada nível de forma que o nível subsequente concentre-se em instâncias mal classificadas por modelos gerados anteriormente. A ideia geral por trás desta rotina é a correção de erros de classificação e diversos algoritmos amplamente utilizados são desenvolvidos a partir desta estrutura, como o *AdaBoost* e o *Gradient boosting* (GÉRON, 2019). Por fim, o *Stacked Generalization*, ou apenas *stacking*, também

chamado de ensacamento, visa explorar a independência dos classificadores e combiná-los, para minimizar o erro, em um mesmo nível. A rotina envolve o ajuste de muitos tipos diferentes de modelos para um mesmo conjunto de dados de entrada, chamados de classificadores de primeiro nível (DU; SWAMY, 2014; BROWNEE, 2020). O *stacking* é a técnica utilizada neste projeto.

Concluindo, a combinação das saídas dos classificadores de base para as definições finais das classes na saída do modelo *ensemble*, representado pelo bloco (III), pode ser realizada de diferentes maneiras: através do voto majoritário, onde um mesmo conjunto é avaliado pelos classificadores de base e a classe que obtiver o maior número de ocorrências será aquela a ser atribuída ao conjunto; através do voto ponderado dos resultados dos classificadores de base, considerando que cada classificador recebe pesos de acordo com seu desempenho de classificação; e por fim, através da meta-aprendizagem, quando os resultados de um grupo de classificadores são utilizados como base de treino para um novo classificador. Seu objetivo é aprender a melhor forma de combinar as previsões provenientes do primeiro nível. A combinação das previsões é uma parte central do método de agrupamento e dependem muito dos tipos de modelos que compõem o *ensemble* (BROWNEE, 2020). As três combinações serão utilizadas neste estudo e serão explicadas em maiores detalhes na seção seguinte.

A utilização de *ensemble* de classificadores vem sendo estudada na resolução de diversos problemas. Questões relacionadas à classificação, como reconhecimento facial, reconhecimento de caracteres, análise de imagens, engenharia de software, diagnóstico médico e relacionadas a regressão, como aproximação de funções e predição de séries temporais comprovam a capacidade do sistema (COELHO, 2006).

Embora um número considerável de pesquisas tenham sido dedicadas ao tema, menos atenção tem sido dada à utilização da técnica com sinais de PEC. Na área de testes não destrutivos, Parikh et al. (2004) combinaram dois *ensemble* com classificadores base do tipo perceptron multicamadas para avaliar diferentes tipos de defeitos em chapas de aço inoxidável: um alimentado pelas imagens do ensaio de ultrassom e outro pelas imagens da técnica de vazamento de fluxo magnético. A pesquisa de Chou, Ngo e Chong (2017) utilizou quatro máquinas de aprendizagem bem conhecidas, incluindo redes neurais artificiais, máquinas de vetores de suporte, árvore de regressão e regressão linear para a construção de diferentes modelos, visando encontrar aquele que apresentasse o menor erro na predição de risco de corrosão em concreto armado e no aço carbono em ambientes marítimos. Em pesquisas mais recentes, Marino, Virupakshappa e Oruklu (2020) utilizaram uma *ensemble* construída com uma rede neural de Wavelet, uma rede neural convolucional e uma rede autoencoder variacional como classificadores de base para a avaliação de sinais A-scan reais e simulados de ensaios com ultrassom que continham uma ampla gama de defeitos de diferentes naturezas. Adicionalmente, o trabalho de Ben Seghier, Höche e Zheludkevich (2022) comparou a utilização dos modelos bem estabelecidos na literatura como floresta aleatória, árvore de regressão, gradient boosting, adaBoost, e gradient para a previsão da taxa de corrosão interna em oleodutos e gasodutos.

3 Método

3.1	Aquisição dos dados e preparação dos <i>datasets</i>	38
3.2	Classificadores de base	44
3.2.1	SVM	45
3.2.2	KNN	45
3.2.3	Árvore de Decisão	46
3.2.4	MLP	47
3.2.5	Gradient Boosting	47
3.2.6	Regressão Logística	48
3.2.7	Naive Bayes Gaussiano	48
3.2.8	Avaliação dos melhores modelos	48
3.3	Estruturação do <i>ensemble</i> de classificadores	49

Para compreensão do objeto de estudo, foi inicialmente explorada a teoria da aplicação da técnica de inspeção não destrutiva de Correntes Parasitas Pulsadas para a avaliação de estruturas e materiais metálicos, em especial para observação de tensões residuais. Outro ponto considerado neste estudo foi a pesquisa sobre técnicas de pré-processamento de sinais e algoritmos de redes classificadoras já difundidos na comunidade acadêmica na avaliação de materiais metálicos. A Figura 20 ilustra as principais atividades conduzidas neste trabalho, dividido em três blocos de atividades: levantamento teórico, seguido por abordagem experimental e posteriormente, análise e discussão dos resultados.

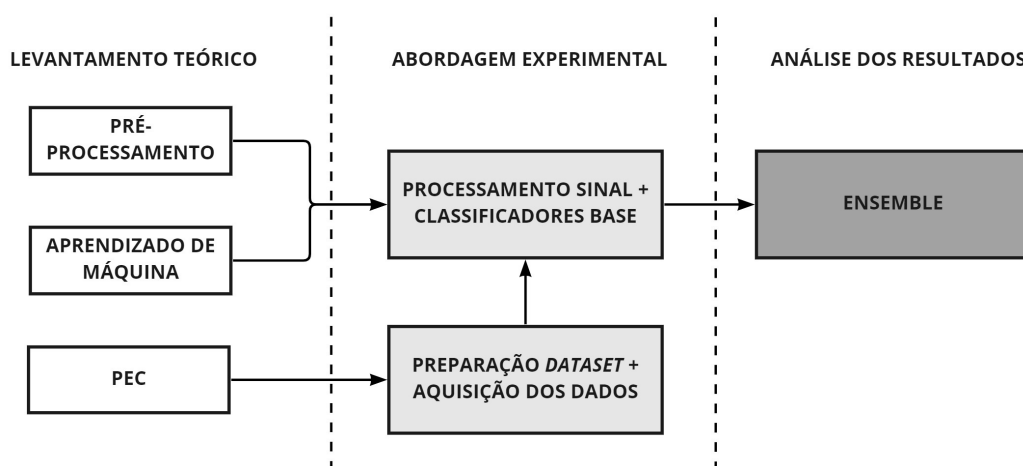


Figura 20 – Fluxograma das principais.

O bloco do levantamento teórico proporciona uma base sólida para as principais temáticas trabalhadas nesta pesquisa. De forma simplificada, pode-se enumerar os temas mais significativos para se atingir os objetivos propostos neste estudo:

- Técnicas clássicas de aprendizado de máquina;

- *Ensemble*;
- Técnica de Correntes Parasitas Pulsadas;
- Técnicas de pré-processamento de sinais temporais.

O Capítulo 02 presente nesta dissertação apresentou o resultado do levantamento teórico realizado ao longo do projeto. No que diz respeito ao bloco de atividades envolvidas pela abordagem experimental, a seção 3.1 irá descrever o desenvolvimento dos processos ocorridas na preparação das amostras, aquisição de sinais e preparo do *dataset*. Já a seção 3.2 discorrerá sobre as rotinas utilizadas nos processamentos dos sinais e na escolha dos classificadores de base. Por fim, a seção 3.3 irá detalhar a construção do *ensemble* utilizado.

3.1 Aquisição dos dados e preparação dos *datasets*

Para a realização do projeto foram utilizados como corpos de prova iniciais nove amostras cedidas por uma montadora automobilística. Essas amostras foram retiradas da bobina de aço utilizada na fabricação das peças na linha de produção e exemplificam as condições do material como recebido para estampagem. Três chapas com dimensões de 1530×1175 mm em diferentes posições foram utilizadas para a produção dos corpos de prova: uma retirada da parte mais interna da bobina, uma retirada do meio e uma chapa retirada da parte mais externa do objeto. Em cada chapa, três amostras de aproximadamente 300×300 mm foram extraídas conforme disposição mostrada na Figura 21.

A Figura 22a exemplifica as amostras do aço IF recebidas em laboratório. As chapas foram, então, seccionadas conforme a indicação na Figura 22b e divididas em dois conjuntos de corpos de prova. O primeiro conjunto, formado pelas duas seções retiradas das chapas iniciais, totaliza dezoito amostras e foi utilizado no estudo de tração. A nomenclatura ocorreu da seguinte maneira: a partir da divisão da amostra 01 do Início, as duas seções menores foram denominadas Início 1A e Início 1B e assim sucessivamente. As seções restantes de área maior compõem o segundo conjunto, totalizando nove amostras e foram utilizadas para aquisição de sinais PEC e alimentação de sistemas inteligentes de classificação.

As amostras do primeiro conjunto foram utilizadas para uma avaliação comportamental quando submetidas a estresse. Para realizar o ensaio de tração, as chapas foram divididas em grupos: 1A, 1B, 2A, 2B, 3A e 3B, conforme exemplo na Figura 23a abaixo.

A amostra 1A de cada caso foi separada para referência. Ensaio de tração foram conduzidos em uma máquina de ensaios EMIC DL30000N, com capacidade de carga de até 300 kN nos quinze corpos de prova restantes. A amostra 1B foi tracionada até a ruptura para determinar a curva tensão deformação de cada caso (início, meio e final). A Figura 23b mostra a chapa Final-1B após ensaio de tração, em comparação com a chapa Final-1A. A partir da delimitação do limite elástico da curva, foi determinado a força de tração que seria aplicada em cada amostra. As dimensões dos corpos de prova, seu papel no ensaio e as características do teste estão sintetizados na Tabela 2.

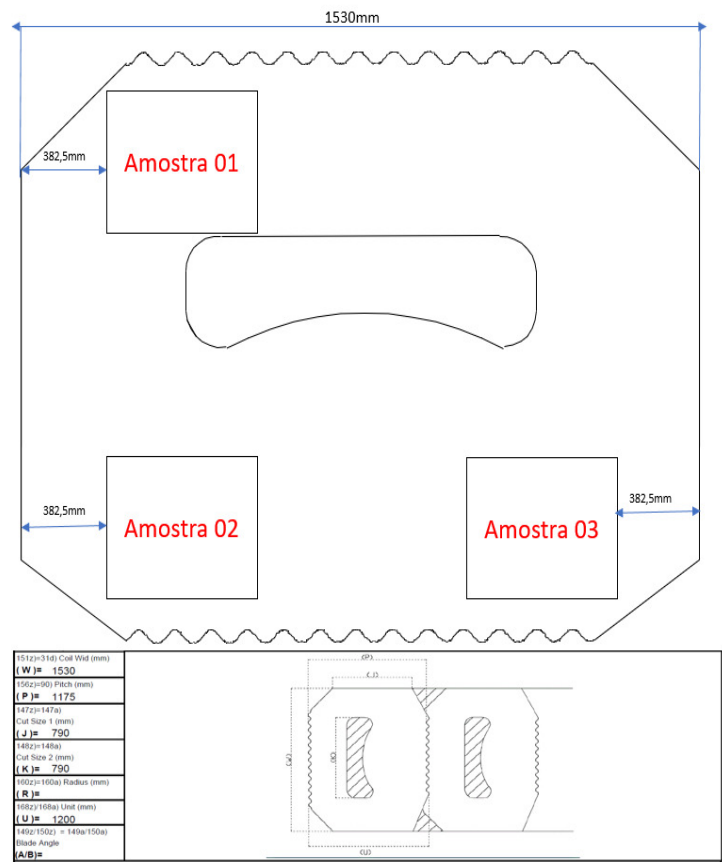
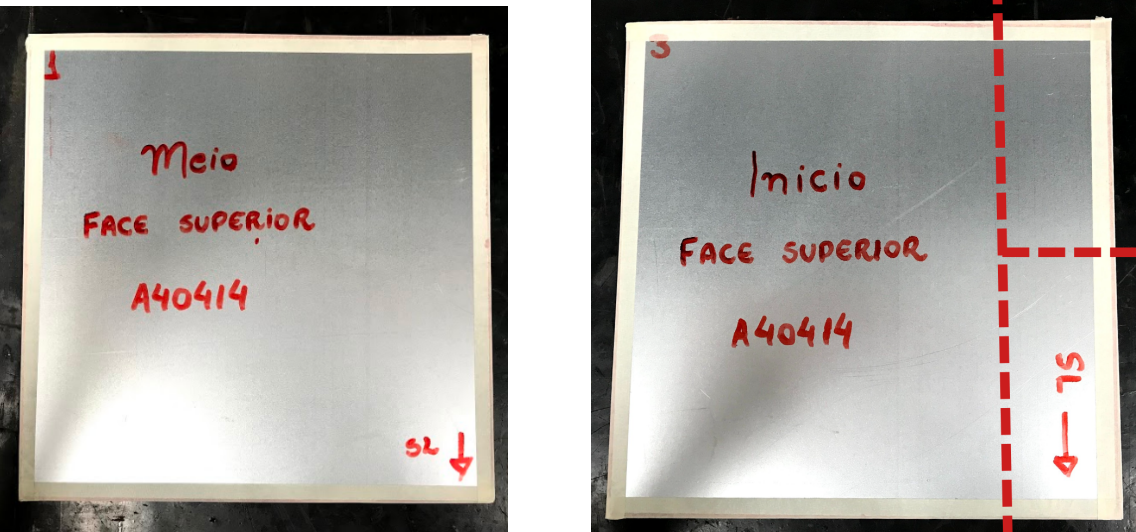
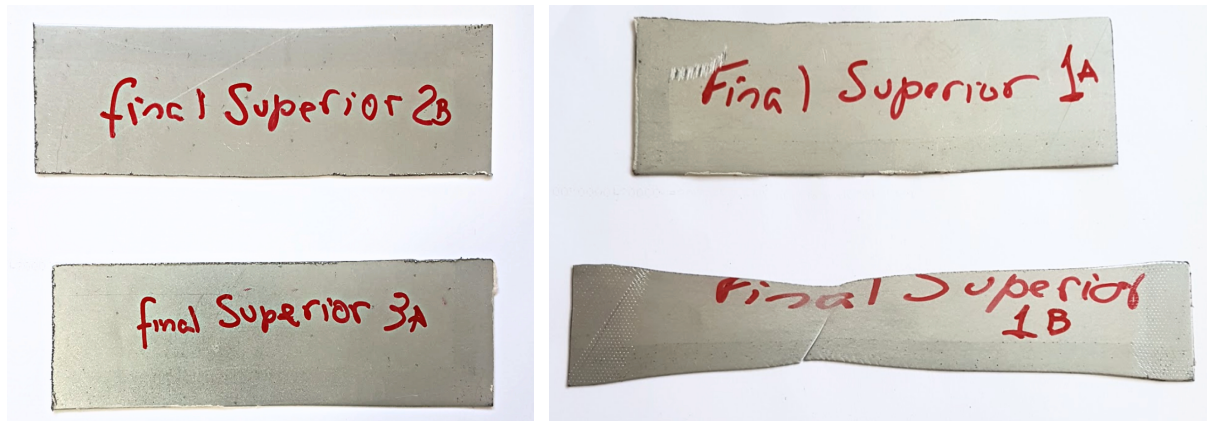


Figura 21 – Disposição dos três corpos de prova retirados da chapa para cada situação.



(a) Amostra da chapa inicial recebida. (b) Local do corte para a divisão das amostras.

Figura 22 – Exemplo dos corpos de prova utilizados.



(a) Amostras utilizadas para o ensaio de tração. (b) Comparação da amostra de referência com a amostra 1B utilizada para determinar a curva tensão deformação.

Figura 23 – Exemplo dos corpos de prova utilizados.

Tabela 2 – Principais propriedades mecânicas obtidas no ensaio de tração

Amostra	Comprimento (mm)	Largura (mm)	Altura (mm)	Tipo de amostra	Tensão aplicada (N)
Início-1A	153,0	36,8	0,8	Referência	-
Meio-1A	150,0	35,3	0,8	Referência	-
Final-1A	153,0	49,0	0,8	Referência	-
Início-1B	157,0	36,8	0,8	Tracionada até a falha	7019
Meio-1B	149,0	36,1	0,8	Tracionada até a falha	6.647
Final-1B	149,0	37,5	0,8	Tracionada até a falha	7.139
Início-2A	151,0	36,9	0,8	Força pré-definida	1.700
Meio-2A	156,0	37,3	0,8	Força pré-definida	1.700
Final-2A	148,0	38,2	0,8	Força pré-definida	1.700
Início-2B	154,0	36,0	0,8	Força pré-definida	2.500
Meio-2B	150,0	37,6	0,8	Força pré-definida	2.500
Final-2B	148,0	37,6	0,8	Força pré-definida	2.500
Início-3A	149,0	37,0	0,8	Força pré-definida	3.400
Meio-3A	158,0	37,5	0,8	Força pré-definida	3.400
Final-3A	149,0	38,0	0,8	Força pré-definida	3.400
Início-3B	154,0	36,4	0,8	Força pré-definida	3.750
Meio-3B	145,0	36,6	0,8	Força pré-definida	3.750
Final-3B	156,0	38,6	0,8	Força pré-definida	3.817

Após o ensaio de tração, foram coletados 56 sinais das amostras 1A, 2A, 2B, 3A e 3B utilizando a técnica PEC em posições aleatórias. No total, 840 sinais serviram como o conjunto de dados para a primeira parte deste trabalho. De forma complementar, o ensaio de dureza Vickers foi realizado para a avaliação das amostras 1A, 2A, 2B, 3A e 3B. A dureza Vickers se baseia na resistência que o material oferece à penetração de uma pirâmide de diamante submetida a determinada carga e seu valor é o quociente da carga aplicada pela área de impressão deixada no corpo ensaiado. Foram adquiridos cinco valores de dureza para cada um dos corpos ensaiados, totalizando 75 valores.

O sistema experimental utilizado na aquisição dos sinais consiste em um sistema PEC, uma fonte de alimentação e um gerador de função. A coleta de dados experimentais foi realizada utilizando uma sonda PEC composta por uma bobina geradora de campo magnético e uma bobina para a leitura do campo secundário ao redor de um único núcleo de ferrite. As propriedades ferromagnéticas da ferrite contribuem para a ampliação do fluxo magnético produzido pela sonda. A bobina geradora foi alimentada por um gerador de função, ajustado para produzir um pulso de onda quadrada de 5 Vpp, com uma frequência de 100 Hz e *Duty cycle* de 50%. A resposta magnética, lida pela bobina foi digitalizada com uma frequência de amostragem de 200 kHz e 12 bits de quantização. Ao final, cada assinatura PEC foi gravada com 2.500 amostras temporais. O diagrama do processo experimental é mostrado na Figura 24.

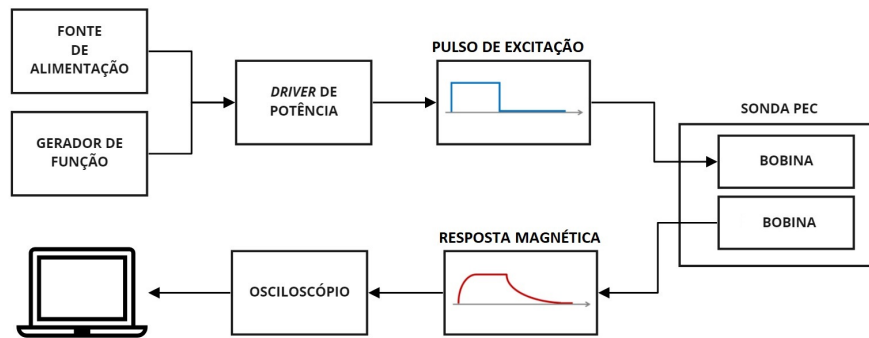


Figura 24 – Diagrama do processo experimental.

As amostras que compõem o segundo conjunto, formado pela seção maior do corte previamente realizado, foram utilizadas para a classificação das condições em que o material é recebido para estampagem e foram utilizadas, posteriormente, para o treinamento e teste dos classificadores. As dimensões dos corpos de prova individuais estão indicados na Tabela 3. A coleta dos sinais foi realizada com o mesmo sistema experimental. Foram coletados 300 sinais das amostras 1, 2 e 3 de cada condição (início, meio e final) utilizando a técnica PEC em posições aleatórias. No total, 2.700 sinais compõem o conjunto de dados para a segunda parte deste trabalho.

Tabela 3 – Principais propriedades mecânicas obtidas no ensaio de tração.

Amostra	Comprimento (mm)	Largura (mm)	Altura (mm)
Início-1	263,2	310,0	0,8
Meio-1	264,3	399,0	0,8
Final-1	251,0	302,0	0,8
Início-2	263,6	305,0	0,8
Meio-2	262,4	306,0	0,8
Final-2	262,1	296,0	0,8
Início-3	263,3	303,0	0,8
Meio-3	262,0	303,0	0,8
Final-3	261,7	305,0	0,8

Na etapa de pré-processamento, foi realizado um estudo preliminar do conjunto de dados

disponível. Inicialmente foi executado um alisamento utilizando a regressão local LOESS³ a cada dez pontos temporais para a suavização do ruído no sinal, cujo objetivo principal era manter, sem alteração, suas características temporais. Uma amostra dos sinais extraídos para o projeto e a curva alisada são mostrados na Figura 25. O sinal consiste em uma curva com 2500 amostras temporais com informações de amplitude. O tempo total do sinal é de 10ms e o intervalo de amostragem do sinal é de 4 μ s.

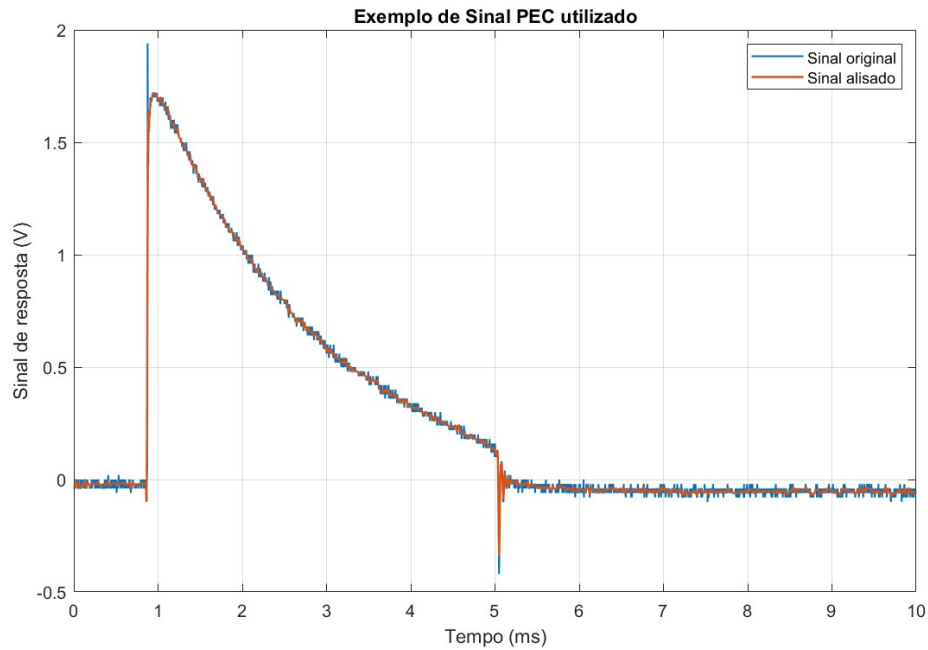


Figura 25 – Sinal do ensaio PEC e Sinal alisado

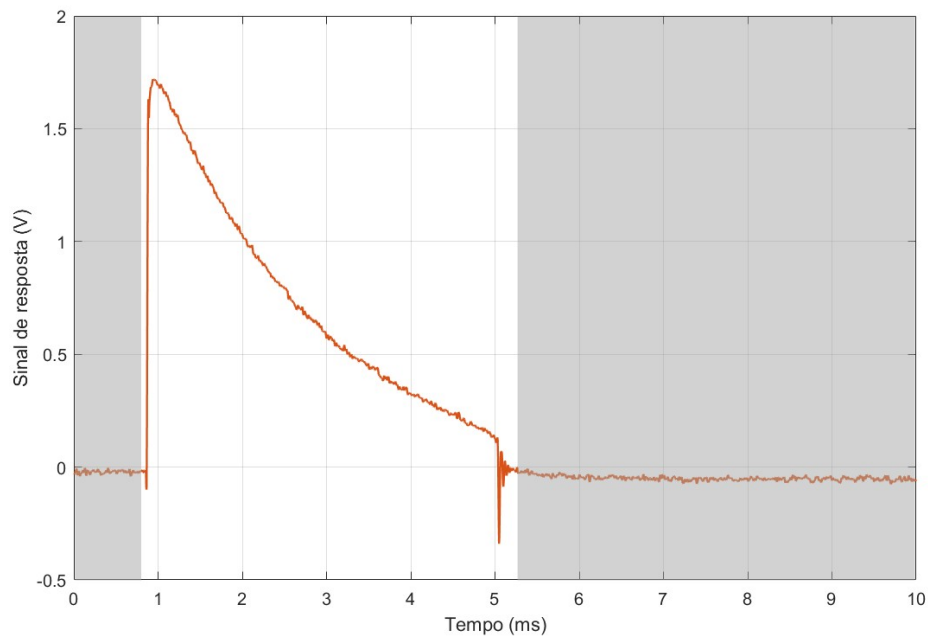


Figura 26 – Sinal PEC utilizado nas rotinas de processamento da informação.

³ Função de regressão linear com ponderação local para suavização dos dados.

Para reduzir o volume dos dados de entrada, foram excluídas as seções que não apresentaram variância no sinal, representado pela área hachurada na Figura 26. Dessa forma, o sinal que apresentava 2.500 pontos, apresentou após a redução, 1.101 pontos.

Visto que o sinal ainda apresentava um número elevado de características para serem processados por algoritmos de aprendizado de máquina, foram aplicadas algumas técnicas de redução de dimensionalidade para a criação de novos conjuntos de dados. Em um primeiro momento, foi aplicada a técnica de PCA no conjunto de dados dos sinais PEC, mantendo 250 componentes principais que retêm 95.12% da variância inicial, apresentando uma redução de 77,29%. O segundo método aplicado foi a escolha de atributos do sinal PEC. As métricas utilizadas são: mediana, média, desvio padrão, variação e erro quadrático médio em relação à amplitude média, amplitude máxima do sinal, amplitude mínima do sinal, variação de amplitude, tempo de pico, tempo de ponto mínimo e tempo entre amplitude máxima e mínima. Esse conjunto possui onze atributos, apresentando uma redução de 99,00%.

Com o objetivo de melhorar o desempenho dos sistemas de classificação, a Transformada Discreta de Fourier também foi aplicada ao conjunto de entrada dos sinais PEC para avaliação do domínio da frequência, resultando em um terceiro conjunto de dados contendo os 1025 coeficientes da DFT. Posteriormente, foram aplicadas a este conjunto a PCA e a LDA. Para a PCA foram escolhidos 390 componentes principais que retêm 95.03% da variância inicial e representam uma redução de 64,58%. A escolha do número de componentes do LDA está intimamente ligado ao número de classes do problema. Usualmente, é escolhido o número imediatamente inferior à quantidade de classes. Dessa forma, para a LDA foram escolhidos dois componentes principais. O diagrama do processo é mostrado na Figura 27.

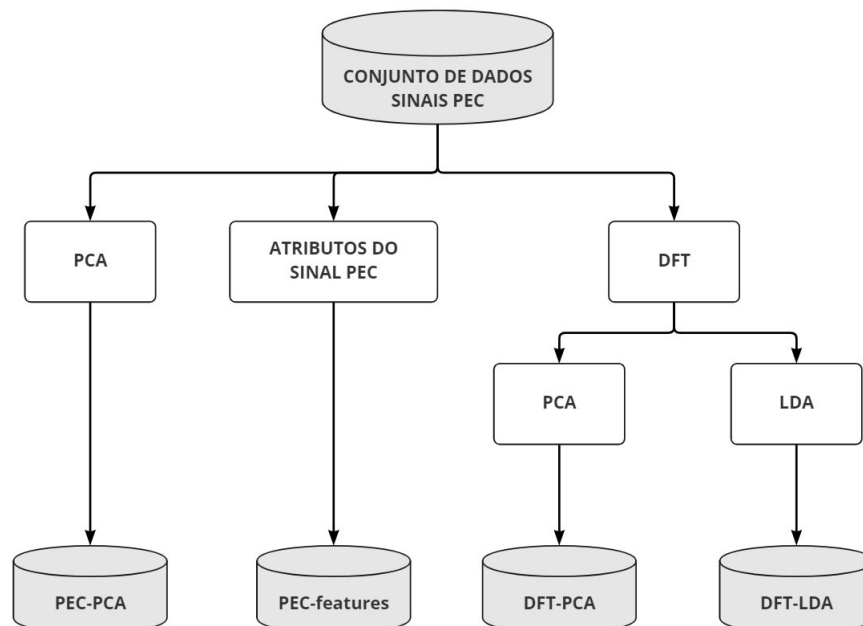


Figura 27 – Processo de criação dos quatro conjuntos de dados de entrada utilizados na rotina.

Ao final, os quatro conjuntos de dados utilizados neste trabalho foram *PEC-features*, *PEC-PCA*, *DFT-PCA* e *DFT-LDA*. O conjunto *PEC-features* possui onze atributos extraídos do sinal PEC, o conjunto *PEC-PCA* possui 250 atributos, o conjunto *DFT-PCA* possui 390

atributos e o conjunto DFT-LDA possui 2 atributos. Todos os conjuntos possuem 2.700 amostras divididas igualmente em três classes - início, meio e final. A Tabela 4 apresenta um resumo das principais características dos conjuntos de dados utilizados nos experimentos deste trabalho.

Tabela 4 – Resumo das principais características dos conjuntos de dados utilizados neste trabalho.

Conjunto de dados	Numero de classes	Numero de atributos
PEC-features	3	11
PEC-PCA	3	250
DFT-PCA	3	390
DFT-LDA	3	2

Na divisão dos conjuntos de dados para a utilização nos algoritmos, foi adotado o particionamento em três subconjuntos: treinamento, validação e teste. A partir do conjunto original, 20% dos sinais foram reservados para teste. Este subconjunto foi separado no início do processo de forma aleatória e esses dados não foram introduzidos até a definição final dos modelos, quando foi realizada a etapa de teste. Os sinais restantes foram novamente particionados, considerando 20% para a validação e 80% para o treinamento. Ao final, o subconjunto de treinamento possui 1.728 amostras, o subconjunto validação possui 432 amostras e o subconjunto teste possui 540 amostras.

O subconjunto de treinamento foi utilizado na etapa treinamento dos classificadores, o subconjunto de validação nas etapas de seleção e combinação de componentes e o subconjunto de teste foi utilizado para verificar o desempenho do *ensemble* e dos classificadores de base individualmente.

Para que o conjunto de teste e validação apresentasse uma heterogeneidade maior nas etapas de treinamento, o método da validação cruzada foi aplicado. Sua utilização irá explorar a flutuação estatística dos conjuntos de dados e auxiliar a escolha do melhor modelo não enviesado. A sua utilização em cada modelo será comentada em detalhes nas próximas seções.

3.2 Classificadores de base

Como mencionado na seção referente à pesquisa bibliográfica, a construção de um *ensemble* pode ser dividida em três etapas: geração dos candidatos, seleção e combinação de componentes de base. Para a primeira etapa anteriormente mencionada, neste trabalho foram escolhidos para compor o conjunto de candidatos sete algoritmos de aprendizado de máquina vistos constantemente na literatura. São estes: Máquina de Vetores de Suporte (SVM), K vizinhos mais próximos (KNN), Árvore de Decisão (DT), Perceptron Multicamadas (MLP), *Gradient Boosting* (GBM), Regressão Logística (LR) e Naive Bayes Gaussiano (NB). Para a escolha de uma configuração de modelo de base satisfatório, foi realizada uma etapa de otimização de parâmetros.

A otimização de parâmetros é uma fase inicial implementada no treinamento dos modelos individuais a fim de encontrar aquele que apresente o melhor resultado dentro do universo estu-

dado. Foram escolhidos alguns hiperparâmetros-chave de cada algoritmo que controlam o processo de aprendizagem e, para cada característica variada, um conjunto de valores foi eleito de forma que representasse uma grande faixa dentro do universo de possibilidades. Estes parâmetros foram avaliados através da execução do método de tentativa e erro até encontrada a melhor combinação. Dessa forma, foi implantada uma rotina utilizando a biblioteca *sklearn.model_selection.GridSearchCV()* (PEDREGOSA et al., 2011) que avaliaria $n = (m_1)(m_2) \dots (m_p)$ modelos, sendo p a quantidade de parâmetros avaliados, m_1 o número de valores avaliados no primeiro parâmetro, m_2 o número de valores avaliados no segundo parâmetro, e assim sucessivamente.

Para avaliação do melhor modelo dentro do conjunto de n possibilidades, foi utilizado uma validação cruzada K-fold com o conjunto de teste. Este conjunto foi dividido em cinco subconjuntos de tamanhos iguais ($K=5$), onde um subconjunto é escolhido para a validação e os quatro restantes para treinamento, mantendo a proporção 80%-Teste e 20%-validação. Para cada inicialização do modelo, a escolha do subconjunto varia, formando assim cinco possibilidades de *datasets* diferentes. Dessa forma, podemos garantir que cada combinação de parâmetros seja avaliada de uma forma mais robusta, onde o resultado médio de cada modelo seja uma estimativa mais precisa do desempenho médio para cinco conjuntos de dados diferentes. Por se tratar de uma avaliação primária dos modelos, cada combinação parâmetros-*dataset* foi inicializada apenas uma vez. O parâmetro de randomização para a divisão dos conjuntos nas validações cruzadas foi ajustado de forma que todos os treinamentos utilizassem os mesmos subconjuntos treinamento-validação, possibilitando a comparação e reprodução dos modelos. Os resultados preliminares dos testes realizados serão apresentados no próximo capítulo.

O conjunto de parâmetros escolhidos para cada técnica individual e seus respectivos parâmetros padrões são mostrados nas próximas sub-seções. A descrição detalhada de cada parâmetro e sua influência na criação dos limites de decisão pode ser encontrada em Pedregosa et al. (2011) e Ke et al. (2017).

3.2.1 SVM

Para a configuração dos modelos das Máquinas de Vetores de Suporte, foi utilizado a biblioteca específica para classificação *sklearn.svm.SVC()* (PEDREGOSA et al., 2011). O conjunto de parâmetros inicializados para implementar a otimização é apresentados na Tabela 5 abaixo. Com os parâmetros escolhidos, foi possível avaliar 720 modelos distintos.

3.2.2 KNN

Para a configuração dos modelos do K-ésimo Vizinho mais Próximo, foi utilizado a biblioteca específica para classificação *sklearn.neighbors.KNeighborsClassifier()* (PEDREGOSA et al., 2011). O conjunto de parâmetros inicializados para implementar a otimização é apresentados na Tabela 6 abaixo. Com os parâmetros escolhidos, foi possível avaliar 240 modelos distintos.

Tabela 5 – Conjuntos de parâmetros utilizados na otimização da SVM.

Parâmetro	Parâmetros utilizados para a avaliação
C	$1e^{-2}$ $1e^{-1}$ 1, 5, 10, 20, 50, 100, 1.000
Kernel	base radial, linear, sigmoid, polinomial
Grau (<i>degree</i>)	3, 5
γ (<i>gamma</i>)	$1e^{-4}$, $1e^{-3}$, $1e^{-2}$, $1e^{-1}$, 1
Função de Decisão (<i>decision_function_shape</i>)	One- <i>vs</i> -rest, One- <i>vs</i> -one

Tabela 6 – Conjuntos de parâmetros utilizados na otimização do KNN.

Parâmetro	Parâmetros utilizados para a avaliação
Vizinhos (<i>n_neighbors</i>)	1, 3, 5, 7, 10, 20, 30, 40, 50, 100
p	1, 2
Tamanho da folha (<i>leaf_size</i>)	1, 5, 10, 15, 30, 50
Pesos (<i>weights</i>)	uniform, distance

3.2.3 Árvore de Decisão

Para a configuração dos modelos das Árvores de decisão, foi utilizado a biblioteca *sklearn.tree.DecisionTreeClassifier()* (PEDREGOSA et al., 2011). O conjunto de parâmetros inicializados para implementar a otimização é apresentados na Tabela 7 abaixo. Com os parâmetros escolhidos, foi possível avaliar 216 modelos distintos

Tabela 7 – Conjuntos de parâmetros utilizados na otimização da Árvore de Decisão.

Parâmetro	Parâmetros utilizados para a avaliação
Máx. n. <i>Features</i> (<i>max_features</i>)	auto, sqrt, log2
Profundidade máxima (<i>max_depth</i>)	5, 7, 10
Custo da complexidade de poda (<i>ccp_alpha</i>)	$1e^{-3}$, $1e^{-2}$, $1e^{-1}$, 0
Função Critério (<i>criterion</i>)	gini, entropy
Min. n. amostras para divisão (<i>min_samples_split</i>)	2, 3, 4

3.2.4 MLP

Para a configuração dos modelos do Perceptron multicamadas, foi utilizado a biblioteca *sklearn.neural_network.MLPClassifier()* (PEDREGOSA et al., 2011). O conjunto de parâmetros inicializados para implementar a otimização é apresentados na Tabela 8 abaixo. Com os parâmetros escolhidos, foi possível avaliar 360 modelos distintos

Tabela 8 – Conjuntos de parâmetros utilizados na otimização do Perceptron multicamadas.

Parâmetro	Parâmetros utilizados para a avaliação
Tamanho das camadas ocultas (<i>hidden_layer_sizes</i>)	(10), (20), (50), (100), (200) (10,10),(10,20),(10,50),(10,100),(10,200) (50,10),(50,20),(50,50),(50,100),(50,200)
Função de ativação (<i>activation</i>)	logistic, tanh, relu
Solver para otimização (<i>solver</i>)	sgd, adam
α (<i>alpha</i>)	$1e^{-4}$, $5e^{-2}$
Taxa de aprendizado (<i>learning_rate</i>)	constant, adaptive
N. máx. iterações (<i>max_iter</i>)	1.000

3.2.5 Gradient Boosting

Para a utilização do algoritmo de *Gradient Boosting* nesta pesquisa, foi escolhida uma estrutura desenvolvida pela Microsoft chamada de *Light Gradient Boosting Machine - LightGBM*, disponível através da biblioteca *lightgbm.LGBMClassifier()*. O funcionamento da estrutura do LightGBM é explicada em maiores detalhes na seção 2.3.5. O conjunto de parâmetros inicializados para implementar a otimização é apresentados na Tabela 9 abaixo. Com os parâmetros escolhidos, foi possível avaliar 360 modelos distintos

Tabela 9 – Conjuntos de parâmetros utilizados na otimização do LightGBM.

Parâmetro	Parâmetros utilizados para a avaliação
N. folhas (<i>num_leaves</i>)	20, 40, 50, 60, 80, 100, 200
subamostras (<i>subsample</i>)	$1e^{-1}$, $5e^{-1}$, $7e^{-1}$
N. mín. amostras em uma folha (<i>min_child_sample</i>)	5, 10, 20, 50
Taxa de aprendizado (<i>learning_rate</i>)	$5e^{-2}$, $1e^{-1}$, $2e^{-1}$

3.2.6 Regressão Logística

Para a configuração dos modelos de Regressão Logística, foi utilizado a biblioteca `sklearn.linear_model.LogisticRegression()` (PEDREGOSA et al., 2011). O conjunto de parâmetros inicializados para implementar a otimização é apresentados na Tabela 10 abaixo. Com os parâmetros escolhidos, foi possível avaliar 216 modelos distintos

Tabela 10 – Conjuntos de parâmetros utilizados na otimização da Regressão Logística.

Parâmetro	Parâmetros utilizados para a avaliação
Penalidade (<i>penalty</i>)	11, 12
C	$1e^{-3}$, $1e^{-2}$, $1e^{-1}$, 1, 5, 10, 20, 50, 100, 1.000
Algoritmo de otimização (<i>solver</i>)	newton-cg, sag, saga, lbfgs
Multiclasse (<i>multi_class</i>)	One-vs-rest, auto, multinomial

3.2.7 Naive Bayes Gaussiano

Para a configuração dos modelos de Naive Bayes Gaussiano, foi utilizado a biblioteca `sklearn.naive_bayes.GaussianNB()` (PEDREGOSA et al., 2011). O conjunto de parâmetros inicializados para implementar a otimização é apresentados na Tabela 11 abaixo. Com os parâmetros escolhidos, foi possível avaliar 150 modelos distintos

Tabela 11 – Conjuntos de parâmetros utilizados na otimização do Naive Bayes Gaussiano.

Parâmetro	Parâmetros utilizados para a avaliação
Variância (<i>var_smoothing</i>)	150 valores distribuídos igualmente no espaço $[1e^{-9}, 1e^{-15}]$

3.2.8 Avaliação dos melhores modelos

O objetivo da busca realizada é personalizar um conjunto eficiente de parâmetros que possa atuar bem com o conjunto de dados utilizado. Inicialmente, a diferença da acurácia de validação e treinamento foi utilizada como primeiro filtro para a avaliação de subajuste ou sobreajuste do modelo. Um limite de 5,0% foi estipulado a partir das observações dos exemplos de treinamento. Apenas aqueles com uma diferença absoluta menor do que 5,0% foram considerados aptos para seguir com a avaliação. Para a escolha dos melhores modelos foi utilizada como medida de avaliação de desempenho a média da acurácia de validação. Caso mais de um modelo apresentasse a mesma acurácia, aquele que exibisse o menor tempo de execução no treinamento seria o escolhido. Assim, os resultados das diversas inicializações foram analisados e o modelo de cada técnica que apresentou o melhor desempenho seguiu para uma segunda etapa de avaliação.

Para a etapa de seleção dos classificadores de base que irão compor os *ensembles*, foram novamente avaliados os modelos eleitos para cada *dataset* utilizando a validação cruzada com

$K=5$. Dessa vez, para cada configuração dos subconjuntos treinamento-validação avaliada, 60 reinicializações foram realizadas, totalizando 300 iniciações para cada modelo. A escolha dos classificadores que vieram a formar uma população de candidatos para os *ensembles* foi baseada na ideia dos classificadores fracos, ou seja, os modelos devem apresentar um resultado ligeiramente melhor do que modelos aleatórios. Dessa forma, aqueles modelos que não atingiram uma acurácia, mínima de 50%, não foram selecionados para a próxima fase. Os candidatos restantes foram considerados aptos para a construção do modelo de agrupamento.

3.3 Estruturação do *ensemble* de classificadores

Vencida a etapa de geração de uma população de classificadores candidatos para os *ensembles*, é necessário então encontrar aqueles candidatos que irão contribuir para aumentar o desempenho de classificação, visto que a inclusão de todos os candidatos pode degradar o seu desempenho. Foram aplicadas três técnicas de seleção de componentes: Construtiva com Exploração, Poda com Exploração e o algoritmo de agrupamento conhecido como *Stacking*. A partir das situações em que as técnicas de seleção de componentes não retornavam resultados satisfatórios, uma escolha arbitrária era feita para montar os agrupamentos, através do método de tentativa e erro até encontrar uma solução que fosse compatível. Estes casos serão comentados no capítulo 5.

As rotinas de Seleção Construtiva e Poda foram utilizadas para reduzir a complexidade computacional de um *ensemble* que apresentasse um conjunto de classificadores base com pouco ou nenhum efeito sobre o seu desempenho, encontrando a combinação de membros que resultou em uma melhor resposta do que o uso cego de todos modelos candidatos.

A rotina de Seleção Construtiva começa com um conjunto vazio e foi adicionado a ele o modelo que apresenta o melhor desempenho. Os outros candidatos foram, então, adicionados um a um se, e somente se, resultassem em um aumento de desempenho sobre o conjunto anterior. De funcionamento semelhante, a rotina de Poda começou com um conjunto contendo todos os candidatos, sendo retirados, um a um, os modelos que apresentaram uma redução de desempenho sobre o conjunto anterior. A cada adição ou retirada, os modelos de referência foram atualizados. A combinação dos componentes foi feita por voto majoritário dos percentuais de cada classe, e o resultado do *ensemble* foi avaliado utilizando uma validação cruzada de $K=5$ com 10 repetições para cada combinação de subconjuntos treinamento-validação, totalizando 50 repetições para cada modelo *ensemble* testado. Os modelos escolhidos pelas rotinas seguiram para a fase de teste.

O algoritmo de *Stacking* é baseado no conceito de meta-aprendizado, onde existe um nível zero de classificadores base e suas saídas são avaliadas por uma máquina de nível um. Para a construção do algoritmo, foi realizado uma etapa de otimização para avaliar a combinação dos classificadores de nível zero e o melhor candidato ao nível um. Considerando k como o número de modelos disponíveis, nesta avaliação, foram combinados todas as possibilidades de combinações de (k) candidatos do nível zero, tomados p a p e multiplicados por k candidatos do nível um.

$$n = k \left[1 + \sum_{p=2}^{(k-1)} C_p^k \right] \quad (3.1)$$

onde

$$C_p^k = \frac{k!}{(k-p)!p!} \quad (3.2)$$

A validação cruzada com $K=5$ foi utilizada para avaliação do melhor modelo. Dado o número elevado de inicializações, cada combinação foi avaliada apenas uma vez. O melhor modelo foi selecionado para seguir na etapa de teste. Neste ponto, três modelos de *ensemble* haviam sido construídos para cada *dataset*.

Para uma maior diversidade dos modelos de agrupamento, foram criados mais dois modelos derivados da Seleção Construtiva e Poda: Seleção Construtiva com Voto Majoritário Enviesado e Poda com Voto Majoritário Enviesado. Diferentemente do Voto Majoritário que assume que todos os classificadores base são igualmente habilidosos na criação dos limites de decisão, o Voto Majoritário enviesado atribui um peso fixo a partir da previsão feita na etapa de validação para a combinação dos componentes. Dessa forma aqueles classificadores que apresentarem um maior índice de acertos irão ter mais peso na decisão final da hipótese criada pelo *ensemble*.

Finalmente, os modelos construídos e avaliados na etapa de teste são:

- Máquina de Vetores de Suporte (SVM);
- K vizinhos mais próximos (KNN);
- Árvore de Decisão (DT);
- Perceptron Multicamadas (MLP);
- *Gradient Boosting - Light Gradient Boosting Machine* (LGBM);
- Regressão Logística (LR);
- Naive Bayes Gaussiano (NB);
- *Ensemble* por Seleção Construtiva com Voto Majoritário (SC-VM);
- *Ensemble* por Seleção Construtiva com Voto Majoritário Enviesado (SC-VME);
- *Ensemble* por Poda com Voto Majoritário (P-VM);
- *Ensemble* por Poda com Voto Majoritário Enviesado (P-VME);
- *Ensemble* por *Stacking*;

É importante enfatizar que nem todos os modelos listados foram testados para todos os conjuntos de entrada. As avaliações preliminares nas etapas de treinamento que não apresentaram resultados significantes para avançar para a próxima etapa, menores que 50% de acurácia, serão comentadas no capítulo 4 (Resultados).

Por fim, foi utilizado o conjunto de dados previamente separado para o teste dos classificadores de base e dos *ensembles* selecionados. Para a avaliação dos classificadores, foram realizadas dez inicializações para cada modelo. Para fins comparativos dos ganhos associados à utilização dos agrupamentos, os melhores resultados dos classificadores de base e *ensembles* para cada conjunto de dados foram separados e analisados comparativamente.

Um diagrama do procedimento de experimentos computacionais implementado para a realização desta pesquisa é mostrado na Figura 28.

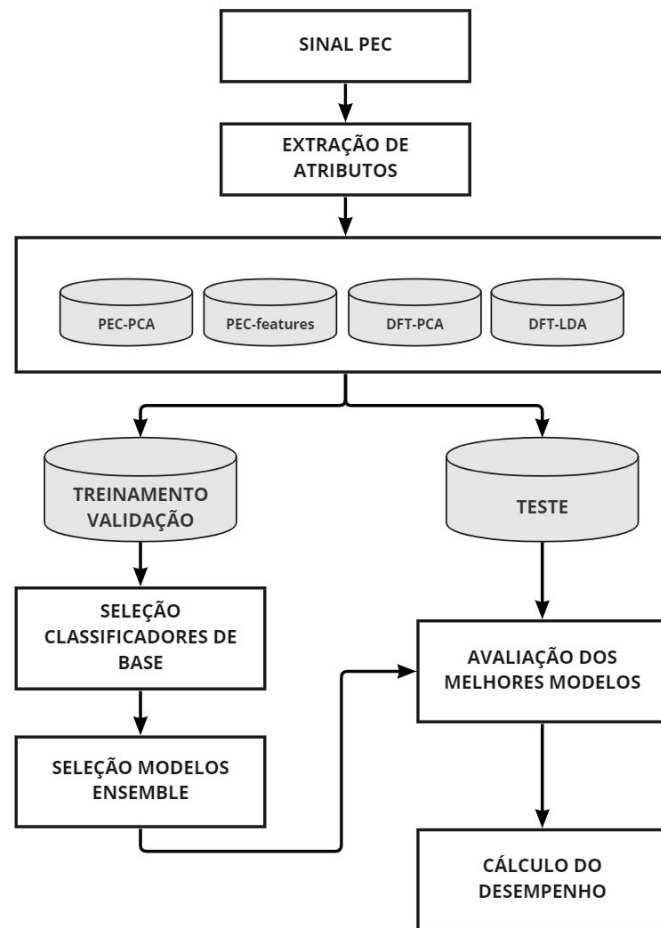


Figura 28 – Diagrama da rotina computacional utilizada no projeto. A partir de um conjunto de dados inicial, quatro conjuntos foram criados a partir de métodos matemáticos de processamento digital de sinais e extração de atributos. Os conjuntos foram divididos em dois subconjuntos: treinamento/validação e teste. O subconjunto de treinamento e validação foi utilizado na seleção dos classificadores de base que serão considerados como candidatos para a construção dos *ensembles*. Esses serão construídos utilizando diferentes métodos de seleção e combinação de candidatos. Posteriormente, para a avaliação do melhor modelo, o subconjunto de teste será novamente inserido no processo e a partir da avaliação dos *ensembles*, análises comparativas em relação aos classificadores de base serão realizadas para a avaliação do ganho.

4 Resultados Obtidos

4.1	Ensaio de Tração	52
4.2	Avaliação dos sinais PEC	56
4.3	Classificadores de base	61
4.4	Construção dos <i>Ensembles</i>	68
4.5	Etapas de teste	71
4.6	Comparação dos Resultados	72

Neste capítulo são apresentadas os principais resultados experimentais obtidos através da metodologia proposta nesta Dissertação de Mestrado. Na seção 4.1 são apresentados os resultados do ensaio de tração e discutido como o sinal PEC se comporta quando o aço IF é submetido a diferentes forças. Nas seções subsequentes, são apresentados os resultados dos classificadores de base para os conjuntos de dados descritos na seção 3.1. As referidas metodologias são, então, comparadas entre si. Posteriormente, são apresentados e discutidos os resultados do *ensemble* de classificadores.

4.1 Ensaio de Tração

Dado que não foram apresentados maiores informações em relação às tensões residuais contidas em cada corpo de prova, para um maior entendimento do conjunto estudado, foi realizado um ensaio de tração. Seis amostras foram seccionadas de cada conjunto (Início, Meio e Final), nomeadamente: 1A, 1B, 2A, 2B, 3A e 3B. As amostras 1A foram separadas para referência e as amostras 1B foram tracionadas até a ruptura para gerar as curvas de tração-deformação. Os corpos 2A, 2B, 3A e 3B foram, então, tracionados com diferentes valores e suas consequências sobre os corpos foram avaliadas.

Após o levantamento das curvas tensão-deformação, foi realizada a interpretação de cada uma delas. Foram definidas algumas propriedades mecânicas básicas do ensaio: tensão limite de escoamento, tensão limite de resistência à tração e alongamento. Os valores das principais propriedades mecânicas obtidas nos ensaios são apresentados na Tabela 12.

Tabela 12 – Principais propriedades mecânicas obtidas no ensaio de tração.

Amostra	Tensão limite de escoamento (MPa)	Tensão Limite de resistência à tração (MPa)	Deformação
Início-1B	0,653	1,215	0,316
Meio-1B	0,623	1,236	0,320
Final-1B	0,683	1,278	0,310

Dentre as três amostras, a Final-1B é a que apresenta o maior limite de escoamento, a maior tensão limite de resistência à tração e o menor alongamento percentual. A amostra Meio-1B apresenta o menor valor de tensão limite de escoamento e o maior alongamento percentual. Visto

que a dureza de um material é a medida da sua resistência à deformação plástica localizada, podemos associar a partir dos dados apresentados que a amostra Final-1B é a mais dura, seguida pelo Início-1B e Meio-1B.

Na Figura 29 são apresentadas as curvas de tração para as amostras ensaiadas, onde podemos verificar visualmente estas diferenças.

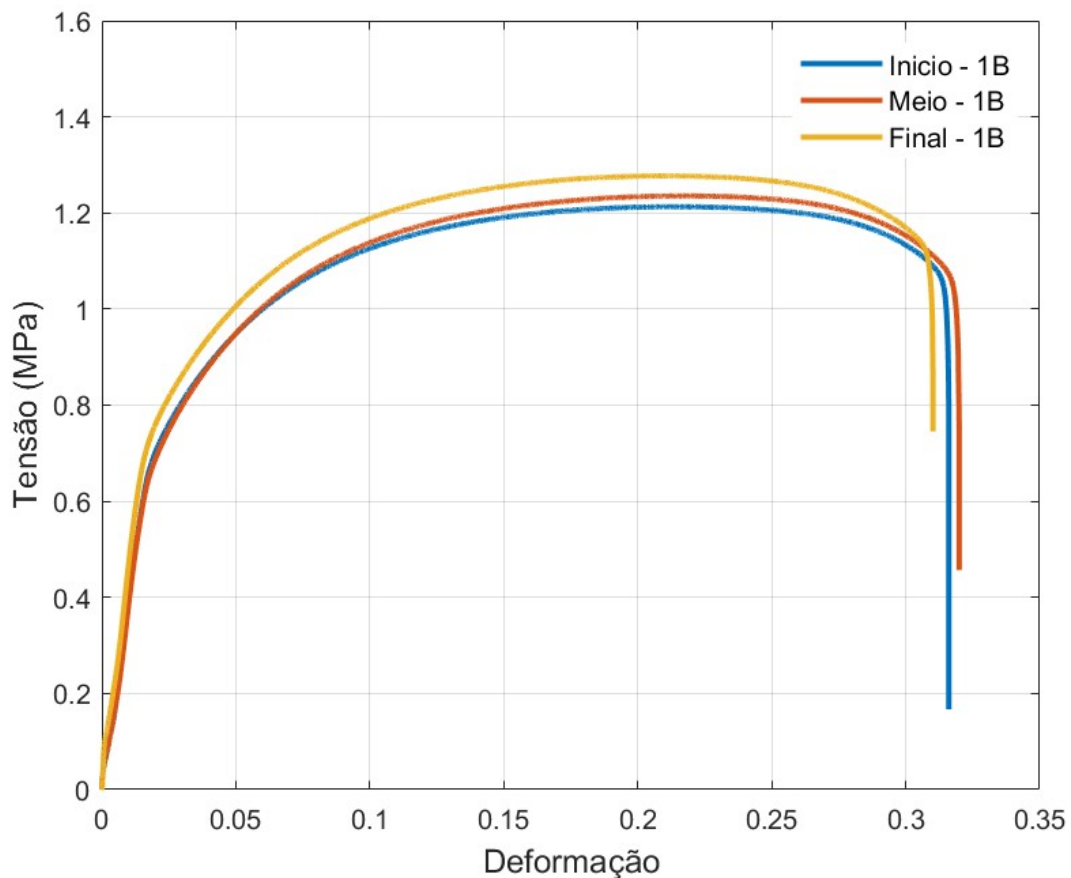


Figura 29 – Curvas de tensão *versus* deformação obtidas pelo ensaio de tração para os diferentes corpos de prova.

Cinquenta e seis sinais de PEC de cada amostra tracionadas foram adquiridos para uma posterior avaliação e, sequencialmente, um ensaio de dureza foi realizado. Os valores médios de cinco aquisições são apresentados na Tabela 13. Podemos perceber uma relação diretamente proporcional entre o carregamento aplicado às amostras e sua medida de dureza. Geralmente, a dureza deve aumentar com a tensão residual presente na superfície do material. Desta forma, é possível estabelecer uma relação onde as amostras mais duras foram aquelas as que sofreram com uma maior incidência de cargas e, conseqüentemente, possuem um maior nível de tensões residuais. Contudo, é importante ter em mente que a tensão residual assume apenas uma parte da responsabilidade do total da medida de tensão, sendo esta apenas uma avaliação primária.

Uma hipótese para que os corpos de prova retirados do final e início da bobina apresentem uma dureza maior é que estas localidades sofrem um maior grau de esforço mecânico durante o transporte. A parte inicial, por estar localizada internamente, apresenta uma circunferência de enrolamento menor e, conseqüentemente, uma maior torção. De forma complementar, o material

Tabela 13 – Resultado médio do ensaio de dureza.

Amostra	Tensão Aplicada (N)	Classe		
		Início	Meio	Final
1A	Referência	85,83	84,48	88,14
2A	1.700	82,86	81,58	80,88
2B	2.500	84,58	82,35	81,87
3A	3.400	87,28	85,45	82,62
3B	3.750/3.817	91,70	88,64	112,52

da parte final, por estar localizado externamente, apoia todo o peso da bobina sobre ele.

As consequências do ensaio de tração no sinal de PEC são verificadas na Tabela 14 a seguir, comparando algumas principais características. É possível perceber uma tendência de redução da amplitude máxima e do tempo até o ponto mínimo, a medida que o corpo sofre com uma maior tração. Contudo, não é possível classificar o sinal de forma satisfatória utilizando como referência apenas as medidas apresentadas. Na Figura 30 é possível visualizar o comportamento decadente da amplitude máxima e do Tempo até o ponto mínimo.

Tabela 14 – Principais propriedades extraídas dos sinais médios de PEC.

Classe	Amplitude máxima (V)	Amplitude mínima (V)	Variação do sinal (V)	Tempo até o mínimo (ms)
Início - 2A	$1,874 \pm 0,007$	$-0,379 \pm 0,007$	$2,253 \pm 0,010$	$4,847 \pm 0,011$
Meio - 2A	$1,884 \pm 0,008$	$-0,391 \pm 0,09$	$2,274 \pm 0,010$	$4,887 \pm 0,016$
Final - 2A	$1,880 \pm 0,010$	$-0,382 \pm 0,010$	$2,262 \pm 0,010$	$4,883 \pm 0,023$
Início - 2B	$1,867 \pm 0,008$	$-0,386 \pm 0,009$	$2,253 \pm 0,011$	$4,797 \pm 0,012$
Meio - 2B	$1,870 \pm 0,008$	$-0,383 \pm 0,011$	$2,253 \pm 0,014$	$4,819 \pm 0,013$
Final - 2B	$1,863 \pm 0,007$	$-0,388 \pm 0,010$	$2,251 \pm 0,011$	$4,814 \pm 0,015$
Início - 3A	$1,861 \pm 0,008$	$-0,394 \pm 0,010$	$2,255 \pm 0,011$	$4,835 \pm 0,023$
Meio - 3A	$1,861 \pm 0,009$	$-0,390 \pm 0,009$	$2,250 \pm 0,010$	$4,826 \pm 0,020$
Final - 3A	$1,872 \pm 0,010$	$-0,383 \pm 0,009$	$2,256 \pm 0,011$	$4,838 \pm 0,026$
Início - 3B	$1,854 \pm 0,008$	$-0,379 \pm 0,008$	$2,233 \pm 0,008$	$4,738 \pm 0,018$
Meio - 3B	$1,861 \pm 0,007$	$-0,383 \pm 0,011$	$2,244 \pm 0,010$	$4,756 \pm 0,019$
Final - 3B	$1,866 \pm 0,009$	$-0,391 \pm 0,010$	$2,257 \pm 0,014$	$4,789 \pm 0,016$

É importante salientar que não é recomendado comparar os sinais de PEC das amostras de tração com os sinais adquiridos para a avaliação dos classificadores. Enquanto os sinais dos corpos tracionados são afetados pelo efeito de borda, visto que os corpos tracionados apresentam uma largura um pouco maior do que a sonda, os outros sinais foram adquiridos em chapas muito maiores, onde este efeito não é presente. Contudo, para efeitos de comparação entre condições semelhantes de aquisição dos sinais, entende-se que não há grandes interferências.

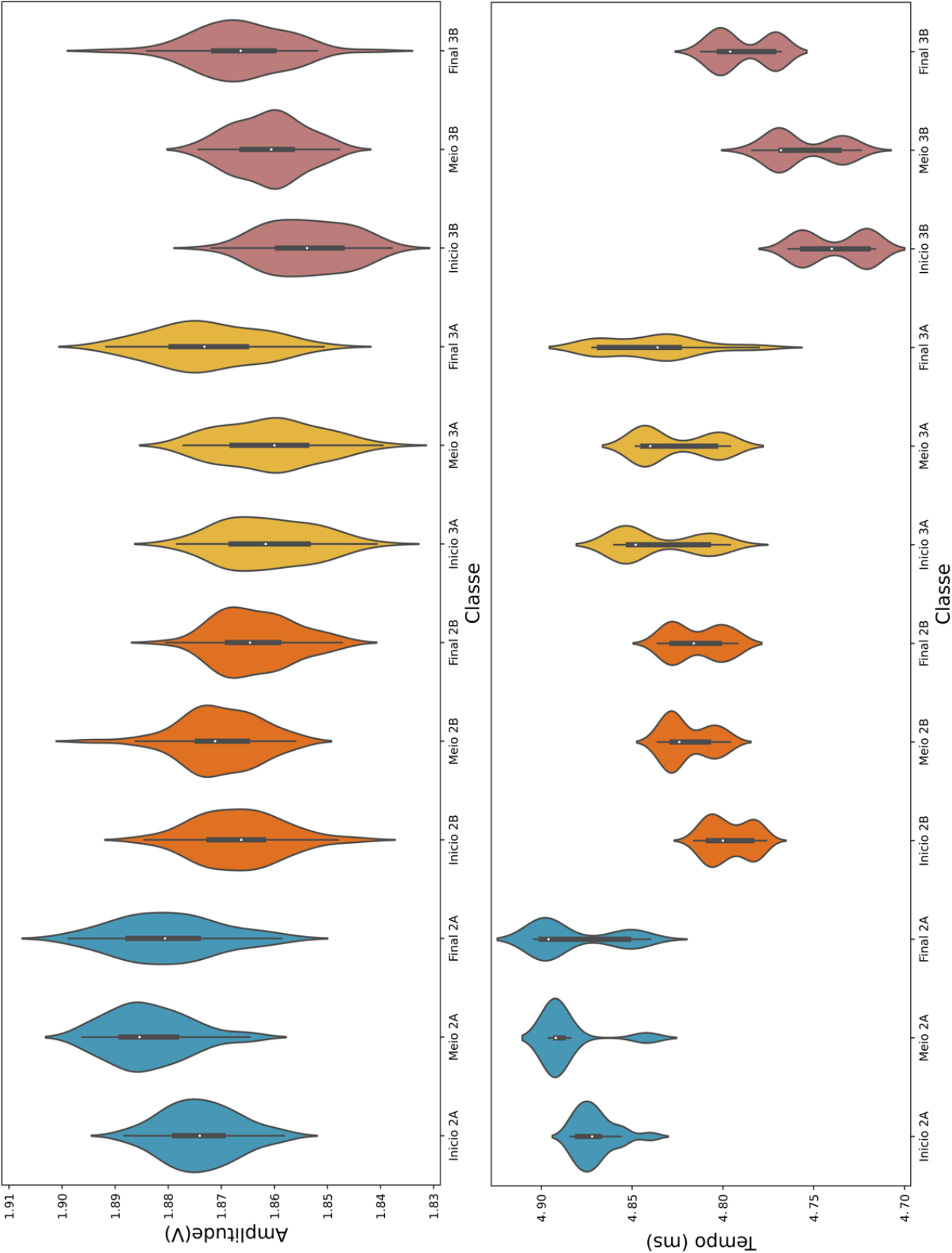


Figura 30 – Dispersão da amplitude máxima e do Tempo até o mínimo das chapas tracionadas.

4.2 Avaliação dos sinais PEC

Foi realizado um estudo preliminar para o entendimento dos sinais extraídos das chapas de aço IF. Visualmente, são sinais coincidentes na sua forma e variação, e os pontos de divergência se encontram na amplitude de pico e amplitude mínima. Para auxiliar em uma visualização e análise geral dos três conjuntos estudados, foram criados três sinais que correspondem à média das 900 medições realizadas em cada condição. A comparação das médias é mostrada na Figura 31.

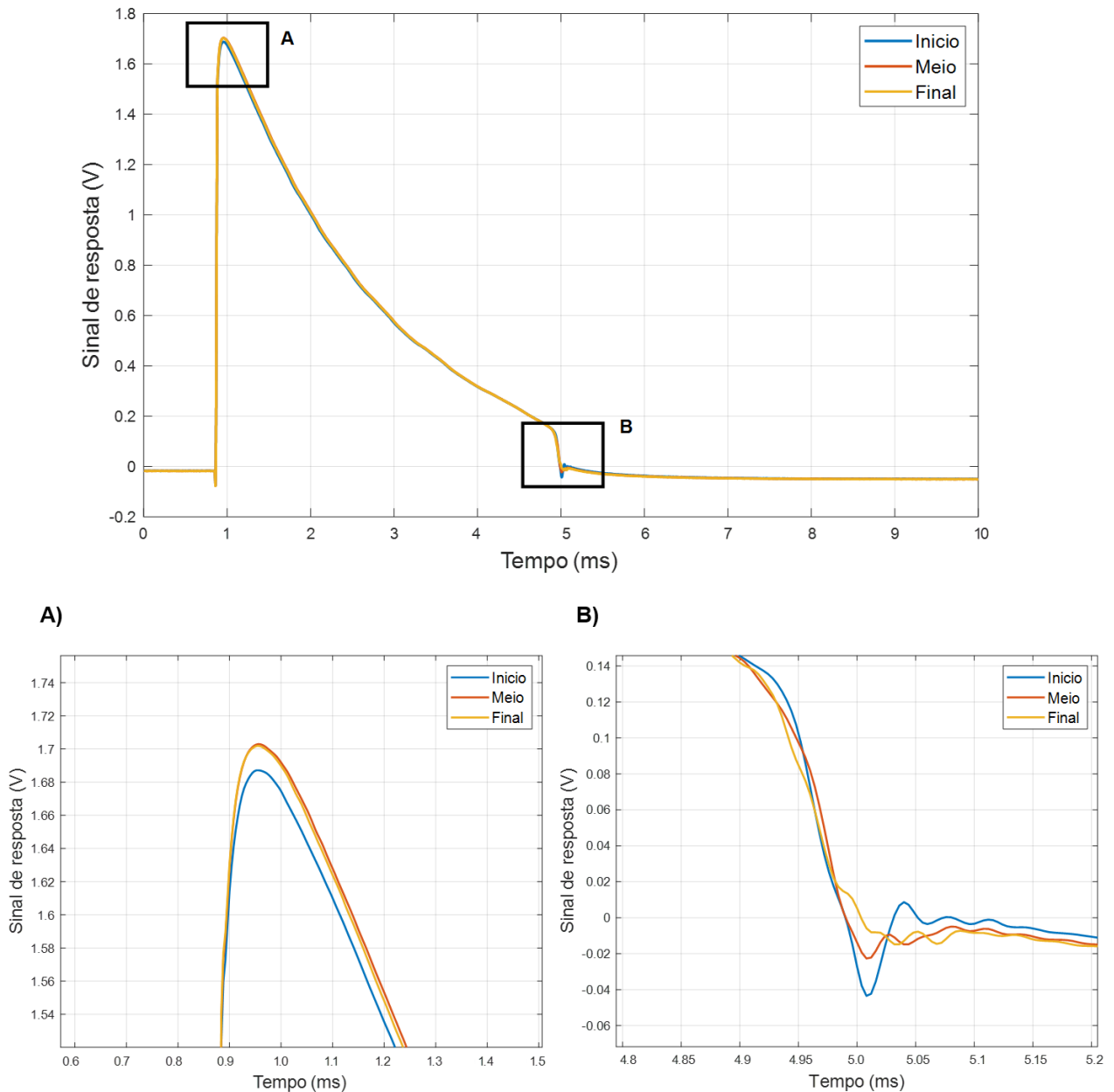


Figura 31 – Média dos sinais para cada condição e avaliação do pico máximo e mínimo.

Em relação à amplitude máxima, o sinal médio do início apresenta uma amplitude de pico de pelo menos 14,9 mV menor quando comparado as outras condições. Já a diferença entre o meio e o final encontra-se entre 1,0 mV. Deste modo, é possível separar visualmente o intervalo do sinal de início, contudo, não é possível separar o meio do final, visto que estas se sobrepõem. Na curva

referente ao decaimento do sinal, as três médias apresentam uma alta similaridade. No entanto, no que se refere, as medidas de amplitude mínima, os sinais médios foram altamente influenciadas pelas ondulações presente no mesmo após o ponto de mínimo, conforme mostrado em detalhes na Figura 32 para um exemplo de dois sinais escolhidos aleatoriamente. Adicionalmente, existe uma diferença no tempo de valor mínimo que, quando realizado a média dos sinais, uma interferência destrutiva nas oscilações pode gerar distorções na região. Dessa forma, não é factível a avaliação dos pontos mínimos, amplitude total e tempo de mínimo a partir da média dos sinais. A Figura 33 mostra o que cada um destes parâmetros representa em um sinal de PEC .

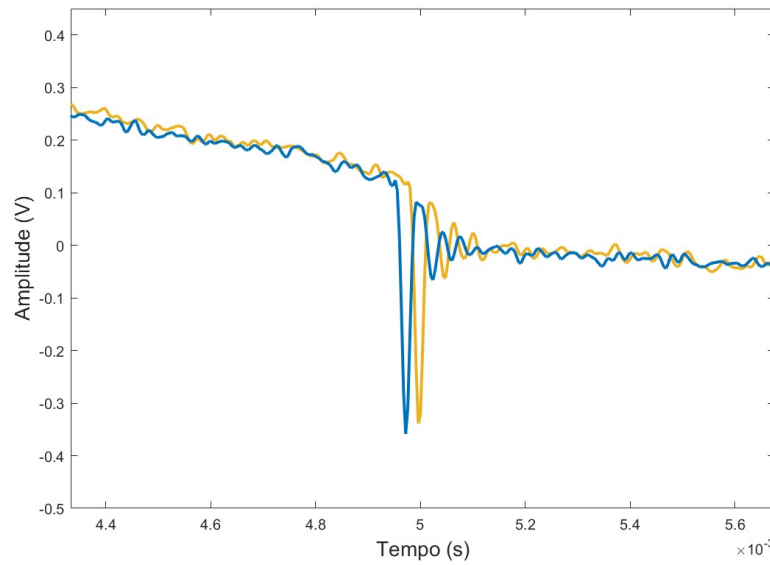


Figura 32 – Diferença no tempo de ponto mínimo e nas ondulações após o ponto de mínimo para dois sinais escolhidos aleatoriamente.

Para complementar a avaliação das médias e entender como uma classe poderia interferir em outra na utilização do conjunto total a ser trabalhado, foi realizado o estudo da distribuição

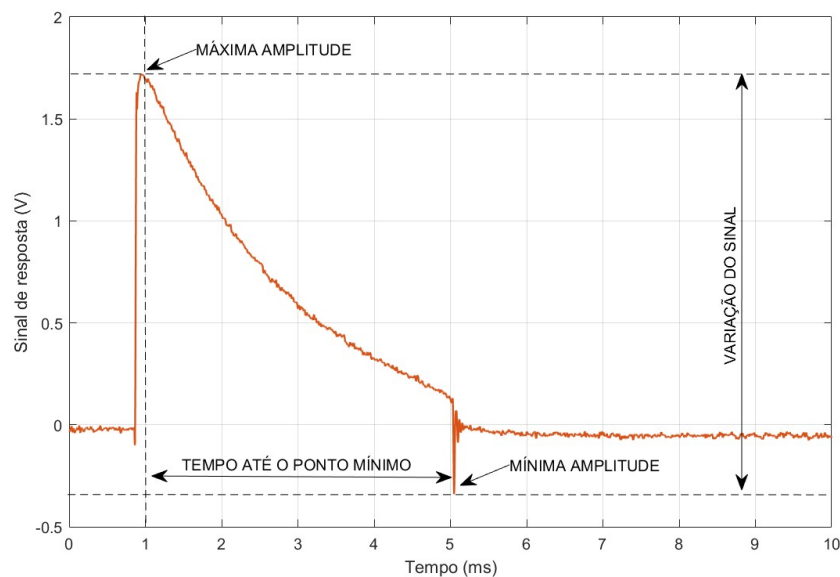


Figura 33 – Principais métricas utilizadas no estudo do dispersão dos sinais.

com os 900 sinais de cada classe para os valores de máxima amplitude, amplitude mínima, variação do sinal e tempo até o ponto mínimo, mostrados a seguir.

A distribuição das amplitudes máximas são mostradas na Figura 34 e ocorrem da seguinte maneira: para a classe início, o valor mínimo se encontra em 1,650 V e o valor máximo em 1,821 V. O segundo e o terceiro quartil estão localizados entre 1,686 V e 1,705 V e a mediana está localizada em 1,697 V. Para a classe meio, o valor mínimo se encontra também em 1,650 V e o valor máximo em 1,858 V. O segundo e o terceiro quartil estão localizados entre 1,700 V e 1,723 V e a mediana está localizada em 1,711 V. Finalmente, a classe final apresenta o valor mínimo em 1,653 V e o valor máximo em 1,816 V. De forma semelhante com a classe meio, o segundo e o terceiro quartil estão localizados entre 1,700 V e 1,720 V e a mediana está localizada em 1,708 V.

Em acordo ao estudo das médias dos sinais, as amplitudes máximas da classe início são, em geral, concentradas em valores menores do que as classes meio e final. A classe meio apresenta a maior dispersão, contudo as três classes possuem uma alta sobreposição.

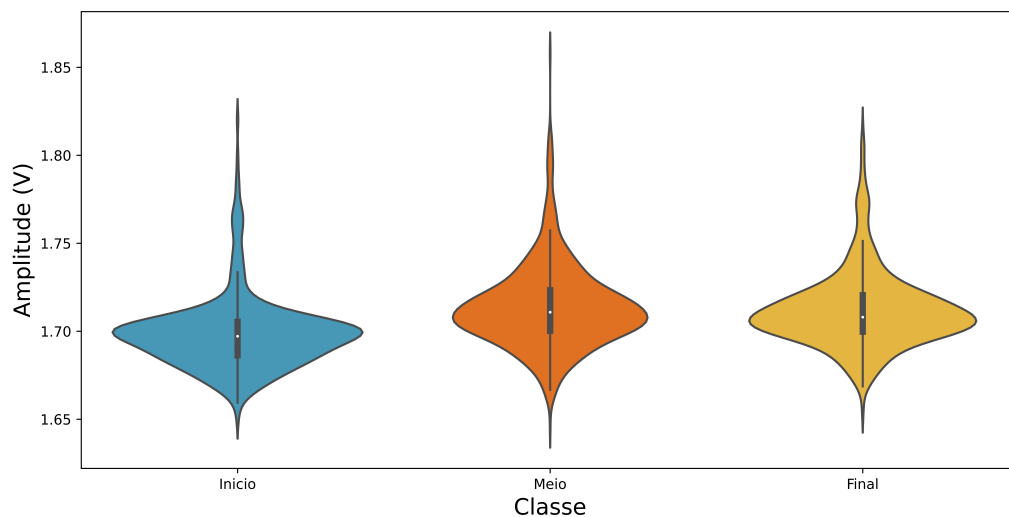


Figura 34 – Dispersão da amplitude máxima de cada classe.

A distribuição das amplitudes mínimas são mostradas na Figura 35 e ocorrem da seguinte maneira: para a classe início, o valor mínimo se encontra em -0,370 V e o valor máximo em -0,301 V. O segundo e o terceiro quartil estão localizados entre -0,340 e -0,326 V e a mediana está localizada em -0,333 V. Para a classe meio, o valor mínimo se encontra em -0,380 V e o valor máximo em -0,311 V. O segundo e o terceiro quartil estão localizados entre -0,353 V e -0,336 V e a mediana está localizada em -0,345 V. Finalmente, a classe final apresenta o valor mínimo em -0,391 V e o valor máximo em 0,317 V. O segundo e o terceiro quartil estão localizados entre -0,360 V e -0,343 V e a mediana está localizada em -0,351 V.

Conforme comentado anteriormente, as amplitudes mínimas apresentadas na média dos sinais expressa um valor em desacordo quando analisado a dispersão dos valores dos sinais individuais. As amplitudes mínimas da classe início são, em geral, concentradas em valores menores, seguida pelas classes meio e final. Apesar de sobrepostos, é possível perceber uma

tendência nas regiões de maior frequência de cada classe, de um aumento da amplitude mínima.

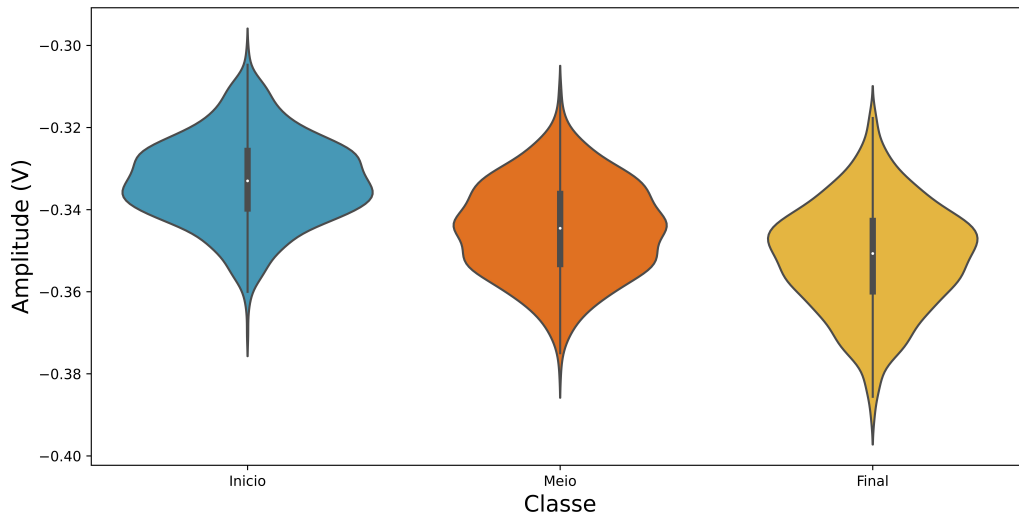


Figura 35 – Dispersão da amplitude mínima de cada classe.

A distribuição das variações de amplitude são mostradas na Figura 36 e ocorrem da seguinte maneira: para a classe início, o valor mínimo se encontra em 1,976 V e o valor máximo em 2,174 V. O segundo e o terceiro quartil estão localizados entre 2,017 e 2,040 V e a mediana está localizada em 2,029 V. Para a classe meio, o valor mínimo se encontra em 1,981 V e o valor máximo em 2,229 V. O segundo e o terceiro quartil estão localizados entre 2,040 V e 2,072 V e a mediana está localizada em 2,056 V. Finalmente, a classe final apresenta o valor mínimo em 2,005 V e o valor máximo em 2,189 V. O segundo e o terceiro quartil estão localizados entre 2,045 V e 2,078 V e a mediana está localizada em 2,060 V.

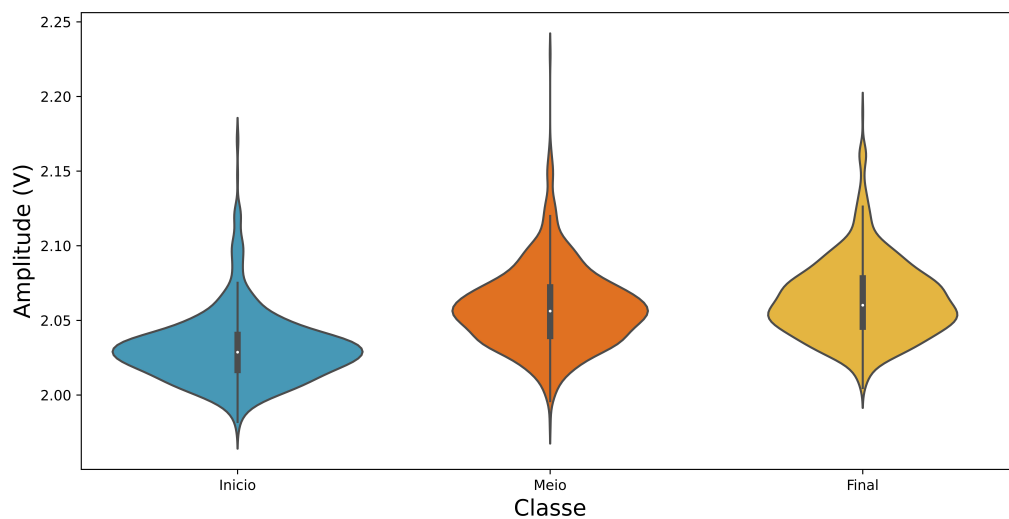


Figura 36 – Dispersão da amplitude total de cada classe.

As variações de amplitude dos sinais da classe início apresentam, em geral, maior

frequência em valores menores do que as classes meio e final. De forma equivalente à amplitude máxima, a classe meio apresenta a maior dispersão e uma grande semelhança com a classe final. As três classes possuem uma alta sobreposição.

Por fim, a distribuição dos tempos até o ponto mínimo são mostradas na Figura 36 e ocorre da seguinte maneira: para a classe início, o valor mínimo se encontra em 4,900 ms e o valor máximo em 5,096 ms. O segundo e o terceiro quartil estão localizados entre 4,976 ms e 5,12 ms e a mediana está localizada em 5,096 ms. Para a classe meio, o valor mínimo se encontra em 4,884 ms e o valor máximo em 5,100 ms. O segundo e o terceiro quartil estão localizados entre 4,980 ms e 5,020 ms e a mediana está localizada em 5,000 ms. Finalmente, a classe final apresenta o valor mínimo em 4,900 ms e o valor máximo em 5,104 ms. O segundo e o terceiro quartil estão localizados entre 4,972 ms e 5,028 ms e a mediana está localizada em 5,000 ms.

As três classes apresentam uma grande sobreposição e semelhança na concentração das maiores frequências.

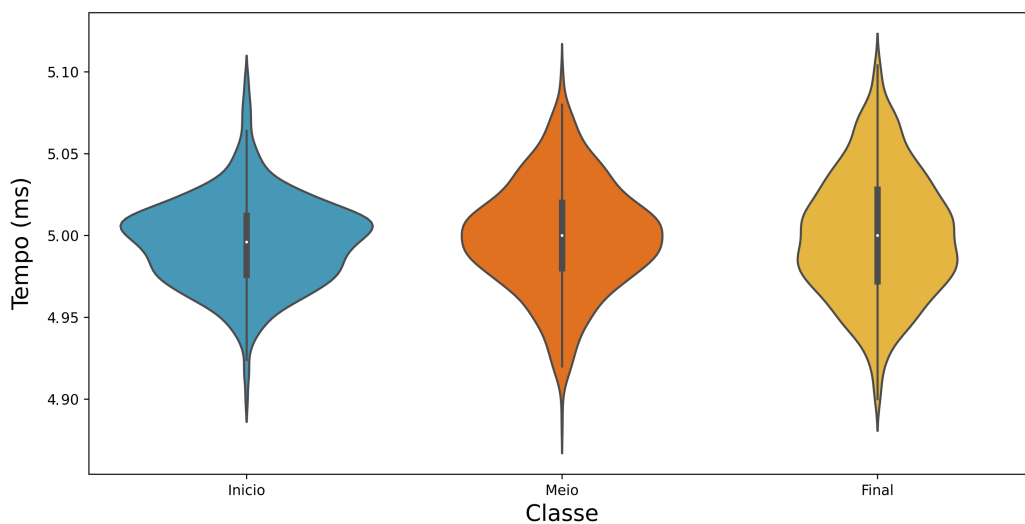


Figura 37 – Dispersão do tempo até a amplitude mínima de cada classe.

Um resumo dos valores das médias da máxima amplitude, amplitude mínima, variação do sinal e tempo até o ponto mínimo e seus desvios padrões é mostrado na Tabela 15.

Tabela 15 – Principais propriedades extraídas dos sinais médios de PEC.

Classe	Amplitude máxima (V)	Amplitude mínima (V)	Variação do sinal (V)	Tempo até o mínimo (ms)
Início	$1,698 \pm 0,021$	$-0,333 \pm 0,012$	$2,031 \pm 0,023$	$4,996 \pm 0,027$
Meio	$1,714 \pm 0,024$	$-0,345 \pm 0,011$	$2,058 \pm 0,027$	$4,999 \pm 0,033$
Final	$1,712 \pm 0,022$	$-0,351 \pm 0,013$	$2,064 \pm 0,026$	$5,000 \pm 0,038$

Como comentado na subseção 3.1, foram realizadas algumas transformações nos conjuntos de dados para gerar novos grupos. A Transformada Discreta de Fourier foi aplicada ao conjunto de entrada dos sinais PEC para avaliação do domínio da frequência. Um sinal resultante de cada classe e o detalhamento das frequências de maior densidade espectral são mostrados na Figura 38.

O eixo horizontal representa a frequência e o eixo vertical representa a densidade espectral, ou a amplitude da DFT. As respostas de frequência dos sinais, novamente, são visualmente muito semelhantes entre si, não representando uma clara distinção entre as classes.

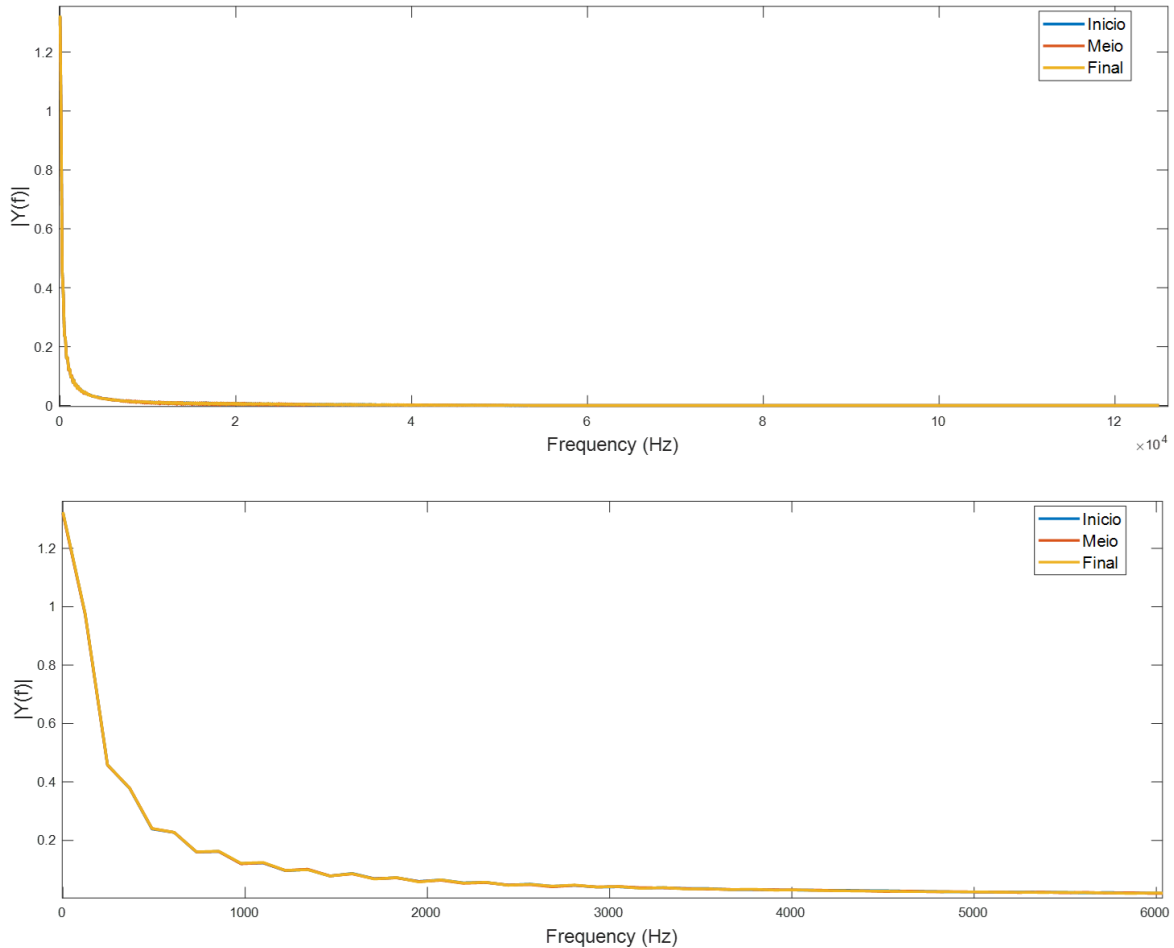


Figura 38 – DFT do sinal de PEC.

De forma complementar, a relação das duas componentes resultantes da análise da Discriminante Linear foi avaliada para análise do conjunto DFT-LDA. Conforme mostrado na Figura 39, a partir da análise com a LDA, é possível estabelecer visualmente limites entre as classes. A classe início é a que apresenta a maior separação. Entretanto, a sobreposição das classes é uma questão importante a ser considerada. No intervalo $[-1,1]$ do componente No. 2 existe uma sobreposição entre a classe Meio e Final. Embora em menor quantidade, o intervalo $[0,2]$ do componente No. 1 também apresenta uma certa confusão entre as três classes.

4.3 Classificadores de base

Para as simulações computacionais dos classificadores de base e a identificação dos melhores parâmetros, foi realizada a abordagem experimental apresentada na subseção 3.3. Uma rotina de otimização de parâmetros foi aplicada para sete algoritmos inteligentes empregados em sistemas de classificação, nomeadamente: Máquina de Vetores de Suporte (SVM), K vizinhos mais próximos (KNN), Árvore de Decisão (DT), Perceptron Multicamadas (MLP), *Gradient*

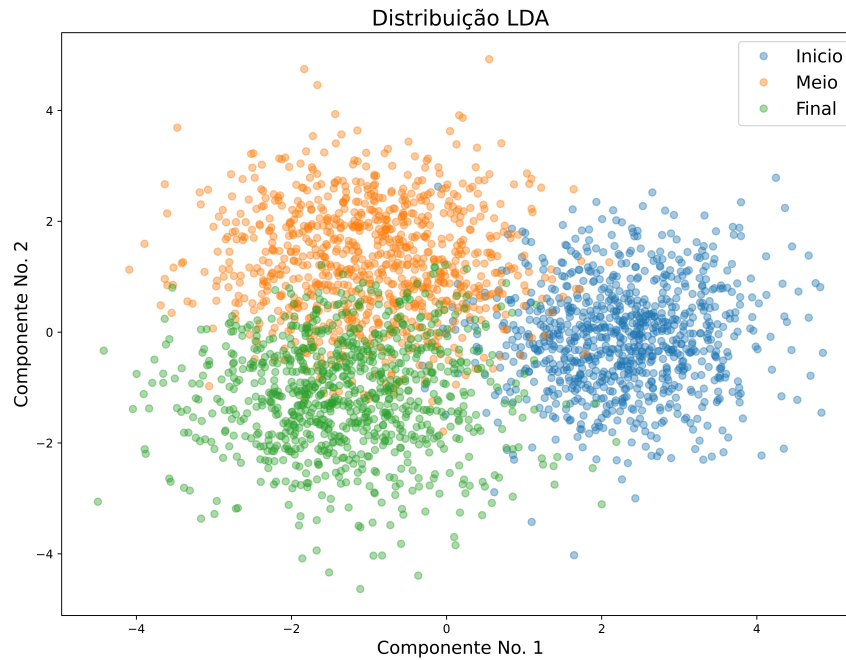


Figura 39 – Distribuição da análise da LDA.

Boosting (LGBM), Regressão Logística (LR) e Naive Bayes Gaussiano (NB). Para os referidos algoritmos de aprendizado supervisionados, foram avaliados os desempenhos de classificação a partir da acurácia de teste, validação e tempo médio de treinamento.

Considerando os dados de entrada empregados nas simulações computacionais, foram utilizados quatro *datasets* de informações extraídas dos sinais de PEC: o primeiro contendo 11 métricas do sinal de PEC (PEC-FEATURES), outro contendo a análise dos principais componentes do sinal de PEC (PEC-PCA), um terceiro contendo a análise dos principais componentes da DFT do sinal de PEC (DFT-PCA) e por fim, o último contendo os discriminantes lineares da DFT sinal de PEC (DFT-LDA), conforme especificado no capítulo 3 (Metodologia). As Tabelas 16, 17, 18 e 19, contêm as acurácias médias, seus respectivos desvios padrões, o tempo médio de treinamento e conjunto de parâmetros do modelo mais eficiente para cada classificador base retornado a partir da rotina de otimização dos parâmetros. As linhas hachuradas representam os modelos que não seguiram para a próxima etapa.

Os classificadores selecionados na análise do conjunto PEC-FEATURES apresentam uma acurácia dentro do intervalo de 50,50% a 65,69%. Aqueles que apresentaram as melhores acurácias foram a LGBM (65,69%), SVM-polinomial (63,43%) e DT (61,81%). Entretanto, o modelo da LGBM será desconsiderado na próxima fase por apresentar uma diferença entre a acurácia de teste e de validação de 8,16%, maior do que o limite de 5,00% estabelecido. O algoritmo que apresentou a menor acurácia foi o MLP, com resultados em torno de 50,51%. Em geral, os modelos apresentam um tempo de treinamento baixo, menor do que 0,3 segundos, excluído deste conjunto apenas a rede MLP, com 4,6 segundos. Os classificadores que passaram para a próxima etapa para o conjunto PEC-FEATURES foram: SVM-rbf, SVM-linear, SVM-sigmoid,

SVM-polinomial, KNN, MLP, DT, NB e LR.

Os classificadores selecionados na análise do conjunto PEC-PCA apresentaram uma alta variação da acurácia, dentro do intervalo de 45,93% a 66,76%. Aqueles que apresentaram as melhores acurácias foram a LGBM (66,76%), SVM-rbf (64,54%) e LR (64,44%). Os modelos da LGBM, SVM-linear e MLP serão desconsiderados para próxima fase por apresentarem uma diferença entre a acurácia de teste e de validação de 33,08%, 15,78% e 6,94%, respectivamente. Os modelos que apresentaram as menores acurácias foram a árvore de decisão, com 45,93% e a SVM-Polinomial, com 47,55%. Nesta etapa, os modelos que apresentaram uma acurácia abaixo de 50,00% ainda não foram desconsiderados. Os classificadores apresentam uma grande variância do tempo de treinamento, desde modelos abaixo de 0,6 segundos até modelos com 23,5 segundos. Em resumo, os classificadores que passaram para a próxima etapa para o conjunto PEC-PCA foram: SVM-rbf, SVM-sigmoid, SVM-polinomial, KNN, DT, NB e LR.

Os classificadores selecionados na análise do conjunto DFT-PCA apresentam a maior variação de acurácia, com um intervalo de 35,23% a 56,81%. Aqueles que apresentaram os melhores resultados foram a SVM-linear (56,81%), a LGBM (54,26%), e LR (54,167%). Os modelos LGBM e SVM-linear serão desconsiderados para próxima fase por apresentarem uma diferença entre a acurácia de teste e de validação de 45,60%, 19,04%. O algoritmo que apresentou a menor acurácia foi da SVM-polinomial, com 35,23%. No entanto, os modelos KNN, SVM-rbf, MLP, DT e NB também apresentaram resultados abaixo de 50,00%. Os classificadores apresentam, novamente, uma grande variância do tempo de treinamento, desde modelos abaixo de 0,05 segundos, outros em torno de 0,6 segundos, até modelos com 11,5 segundos. De modo geral, este conjunto de entrada apresentou os modelos mais fracos. Em resumo, os classificadores que passaram para a próxima etapa para o conjunto PEC-PCA foram: SVM-rbf, SVM-sigmoid, SVM-polinomial, KNN, MLP, DT, NB e LR.

Por fim, os classificadores selecionados na análise do conjunto DFT-LDA apresentam a menor variação de acurácia, com um intervalo de 87,82% a 89,12%. Aqueles que apresentaram os melhores resultados foram a MLP (89,12%), a LR (88,98%), e NB (88,84%). Todos os modelos apresentaram uma diferença da acurácia teste-validação abaixo de 5,0%. O algoritmo que apresentou a menor acurácia foi a SVM-polinomial, com 87,82%. Os candidatos apresentaram os menores tempos de treinamento dentre todos os conjuntos, com máxima em 0,63 segundos. De modo geral, este conjunto de entrada apresentou os modelos mais bem sucedidos. Todas os classificadores passaram para a próxima etapa.

Em uma análise geral, os modelos que mais apresentaram incidência de sobreajuste foram o LGBM, apesar da alta taxa de acurácia quando comparado com os outros candidatos, e o SVM-linear. Do ponto de vista dos conjuntos de dados de entrada, o PEC-PCA e o DFT-PCA foram aqueles que apresentaram os modelos com menor acurácia e/ou maior dificuldade de achar uma hipótese que apresentasse generalização.

Tabela 16 – Classificadores de base que apresentaram o melhor desempenho na etapa de otimização de parâmetros para os dados PEC-FEATURES. Os algoritmos com as células hachuradas não seguiram para a próxima etapa.

Modelo	Acurácia Validação (%)	Acurácia Treinamento (%)	Tempo médio de Treinamento (s)	Parâmetros
SVM - rbf	61,481 $\pm$ 1,497	61,400 $\pm$ 0,529	0,181 $\pm$ 0,005	$C = 1.000$ $decision_function_shape = \text{ovo}$ $gamma = 1$ $kernel = \text{rbf}$
SVM - linear	58,519 $\pm$ 1,109	58,738 $\pm$ 0,587	0,118 $\pm$ 0,006	$C = 1.000$ $decision_function_shape = \text{ovo}$ $kernel = \text{linear}$
SVM - poly	63,426 $\pm$ 1,412	63,750 $\pm$ 0,346	0,121 $\pm$ 0,006	$C = 1.000$ $decision_function_shape = \text{ovo}$ $degree = 5$ $gamma = 1$ $kernel = \text{poly}$
SVM - sigmoid	55,880 $\pm$ 0,404	55,984 $\pm$ 0,620	0,289 $\pm$ 0,005	$C = 1.000$ $decision_function_shape = \text{ovr}$ $gamma = 1$ $kernel = \text{sigmoid}$
KNN	60,880 $\pm$ 0,463	64,074 $\pm$ 0,531	0,007 $\pm$ 0,001	$leaf_size = 30$ $n_neighbors = 40$ $p = 1$ $weights = \text{uniform}$
MLP	50,509 $\pm$ 1,691	50,660 $\pm$ 0,678	4,587 $\pm$ 0,254	$activation = \text{tanh}$ $alpha = 0,0001$ $hidden_layer_sizes = (50,50)$ $learning_rate = \text{adaptive}$ $solver = \text{adam}$
LGBM	65,694 $\pm$ 2,385	73,854 $\pm$ 0,509	0,135 $\pm$ 0,002	$learning_rate = 0,05$ $min_child_samples = 50$ $num_leaves = 20$ $subsample = 0,1$
LR	58,750 $\pm$ 1,193	58,981 $\pm$ 0,487	0,118 $\pm$ 0,006	$C = 1.000$ $multi_class = \text{multinomial}$ $penalty = \text{l2}$ $solver = \text{lbfgs}$
DT	61,806 $\pm$ 2,240	64,387 $\pm$ 0,987	0,012 $\pm$ 0,000	$ccp_alpha = 0,01$ $criterion = \text{entropy}$ $max_depth = 5$ $max_features = \text{auto}$ $min_samples_split = 4$
NB	60,139 $\pm$ 0,807	60,856 $\pm$ 0,486	0,007 $\pm$ 0,001	$var_smoothing = 1,472e^{-05}$

Tabela 17 – Classificadores de base que apresentaram o melhor desempenho na etapa de otimização de parâmetros para os dados PEC-PCA. Os algoritmos com as células hachuradas não seguiram para a próxima etapa.

Modelo	Acurácia Validação (%)	Acurácia Treinamento (%)	Tempo médio de Treinamento (s)	Parâmetros
SVM - rbf	64,537 ± 2,919	67,083 ± 0,585	0,491 ± 0,006	$C = 1$ $decision_function_shape = \text{ovo}$ $gamma = 0,0001$ $kernel = \text{rbf}$
SVM - linear	64,028 ± 2,530	79,803 ± 0,581	0,419 ± 0,025	$C = 0,01$ $decision_function_shape = \text{ovo}$ $kernel = \text{linear}$
SVM - poly	47,546 ± 2,363	48,750 ± 0,967	0,489 ± 0,005	$C = 20$ $decision_function_shape = \text{ovo}$ $degree = 3$ $gamma = 0,0001$ $kernel = \text{poly}$
SVM - sigmoid	62,407 ± 2,897	64,120 ± 0,778	0,533 ± 0,010	$C = 1$ $decision_function_shape = \text{ovo}$ $gamma = 0,0001$ $kernel = \text{sigmoid}$
KNN	63,843 ± 3,557	66,319 ± 0,942	0,004 ± 0,000	$leaf_size = 50$ $n_neighbors = 40$ $p = 2$ $weights = \text{uniform}$
MLP	63,241 ± 1,993	70,185 ± 1,135	23,548 ± 0,378	$activation = \text{logistic}$ $alpha = 0,05$ $hidden_layer_sizes = (10,100)$ $learning_rate = \text{constant}$ $solver = \text{sgd}$
LGBM	66,759 ± 2,807	99,838 ± 0,043	9,935 ± 0,053	$learning_rate = 0,05$ $min_child_samples = 50$ $num_leaves = 40$ $subsample = 0,1$
LR	64,444 ± 2,826	65,579 ± 1,099	1,340 ± 0,108	$C = 0,01$ $multi_class = \text{auto}$ $penalty = \text{l1}$ $solver = \text{saga}$
DT	45,926 ± 6,169	46,458 ± 5,907	0,038 ± 0,004	$ccp_alpha = 0,01$ $criterion = \text{entropy}$ $max_depth = 10$ $max_features = \text{log2}$ $min_samples_split = 3$
NB	55,880 ± 2,399	58,449 ± 1,100	0,011 ± 0,000	$var_smoothing = 0,0122$

Tabela 18 – Classificadores de base que apresentaram o melhor desempenho na etapa de otimização de parâmetros para os dados DFT-PCA. Os algoritmos com as células hachuradas não seguiram para a próxima etapa.

Modelo	Acurácia Validação (%)	Acurácia Treinamento (%)	Tempo médio de Treinamento (s)	Parâmetros
SVM - rbf	44,722 ± 1,184	45,208 ± 0,512	0,800 ± 0,015	$C = 0,1$ $decision_function_shape = ovr$ $gamma = 0,0001$ $kernel = rbf$
SVM - linear	56,806 ± 2,706	75,845 ± 0,955	0,640 ± 0,023	$C = 0,01$ $decision_function_shape = ovo$ $kernel = linear$
SVM - poly	35,231 ± 0,628	36,331 ± 0,247	0,689 ± 0,016	$C = 20$ $decision_function_shape = ovo$ $degree = 3$ $gamma = 0,0001$ $kernel = poly$
SVM - sigmoid	51,528 ± 1,820	53,912 ± 0,558	0,745 ± 0,014	$C = 1$ $decision_function_shape = ovr$ $gamma = 0,0001$ $kernel = sigmoid$
KNN	49,259 ± 3,641	51,354 ± 0,870	0,005 ± 0,000	$leaf_size = 10$ $n_neighbors = 50$ $p = 1$ $weights = uniform$
MLP	43,380 ± 9,354	45,486 ± 12,276	8,247 ± 9,605	$activation = logistic$ $alpha = 0,05$ $hidden_layer_sizes = (10,50)$ $learning_rate = constant$ $solver = sgd$
LGBM	54,259 ± 1,468	99,861 ± 0,069	11,555 ± 0,126	$learning_rate = 0,05$ $min_child_samples = 50$ $num_leaves = 20$ $subsample = 0,1$
LR	54,167 ± 1,486	55,093 ± 0,554	1,644 ± 0,070	$C = 0,01$ $multi_class = ovr$ $penalty = l1$ $solver = saga$
DT	41,296 ± 1,855	42,002 ± 2,993	0,052 ± 0,003	$ccp_alpha = 0,01$ $criterion = entropy$ $max_depth = 5$ $max_features = sqrt$ $min_samples_split = 3$
NB	45,185 ± 1,753	46,319 ± 0,885	0,013 ± 0,001	$var_smoothing = 0,3138$

Tabela 19 – Classificadores de base que apresentaram o melhor desempenho na etapa de otimização de parâmetros para os dados DFT-LDA.

Modelo	Acurácia Validação (%)	Acurácia Treinamento (%)	Tempo médio de Treinamento (s)	Parâmetros
SVM - rbf	88,843 $\pm$ 1,789	88,854 $\pm$ 0,436	0,050 $\pm$ 0,001	$C = 50$ $decision_function_shape = \text{ovo}$ $gamma = 0,0001$ $kernel = \text{rbf}$
SVM - linear	88,843 $\pm$ 1,789	88,854 $\pm$ 0,436	0,028 $\pm$ 0,002	$C = 0,01$ $decision_function_shape = \text{ovr}$ $kernel = \text{linear}$
SVM - poly	87,824 $\pm$ 1,791	88,113 $\pm$ 0,706	0,081 $\pm$ 0,009	$C = 1$ $decision_function_shape = \text{ovr}$ $degree = 5$ $gamma = 1$ $kernel = \text{poly}$
SVM - sigmoid	88,843 $\pm$ 1,789	88,854 $\pm$ 0,436	0,054 $\pm$ 0,001	$C = 10$ $decision_function_shape = \text{ovr}$ $gamma = 0,001$ $kernel = \text{sigmoid}$
KNN	88,611 $\pm$ 1,825	88,692 $\pm$ 0,308	0,003 $\pm$ 0,000	$leaf_size = 50$ $n_neighbors = 40$ $p = 2$ $weights = \text{uniform}$
MLP	89,120 $\pm$ 1,875	88,796 $\pm$ 0,405	0,632 $\pm$ 0,044	$activation = \text{relu}$ $alpha = 0,0001$ $hidden_layer_sizes = 20$ $learning_rate = \text{constant}$ $solver = \text{adam}$
LGBM	87,870 $\pm$ 1,820	90,914 $\pm$ 0,248	0,162 $\pm$ 0,005	$learning_rate = 0,05$ $min_child_samples = 50$ $num_leaves = 40$ $subsample = 0,5$
LR	88,981 $\pm$ 1,844	88,935 $\pm$ 0,424	0,012 $\pm$ 0,000	$C = 1$ $multi_class = \text{multinomial}$ $penalty = \text{l2}$ $solver = \text{saga}$
DT	88,148 $\pm$ 1,848	88,634 $\pm$ 0,596	0,007 $\pm$ 0,001	$ccp_alpha = 0,01$ $criterion = \text{entropy}$ $max_depth = 5$ $max_features = \text{auto}$ $min_samples_split = 2$
NB	88,848 $\pm$ 1,729	88,762 $\pm$ 0,458	0,003 $\pm$ 0,000	$var_smoothing = 1,3207e^{-12}$

Na etapa seguinte, devido à relativa variabilidade dos resultados a cada execução pela utilização da validação cruzada, testou-se cada candidato apto 300 vezes. A acurácia e o F1-score médios de classificação para todos os modelos considerados é apresentada na Tabela 20. Nesta etapa, serão selecionados os candidatos aptos que irão fazer parte das rotinas de criação do *ensemble*. Dessa forma, aqueles modelos com uma acurácia abaixo de 50,00%, indicados pelas linhas hachuradas, não serão considerados.

Os modelos do conjunto PEC-FEATURES apresentam uma diferença entre seus modelos de até 17,6 pontos percentuais na acurácia. Em concordância com a análise anterior, o MLP foi o candidato que apresentou o pior resultado, com 45,73% e um elevado desvio padrão de 7,16%, sendo desclassificado para a próxima fase. Os modelos SVM-polinomial, SVM-rbf e NB foram os que apresentaram os melhores resultados, com 63,33%, 61,19% e 60,50%, respectivamente. O conjunto PEC-PCA é o que apresenta a maior variação das acurácias médias entre os modelos, com um intervalo de 23,46%. Os modelos SVM-polinomial e DT apresentaram uma acurácia abaixo de 50%, respectivamente, 46,95% e 40,87%. Os modelos que apresentaram as maiores acurácias foram SVM-rbf com 64,33%, LR com 64,31% e KNN com 63,34%.

O conjunto DFT-PCA foi o que apresentou o pior resultado dentre os *datasets* testados. Os modelos SVM-rbf, SVM-polinomial, KNN, MLP, DT e NB não seguiram para a próxima etapa por apresentarem acurácias abaixo do limite mínimo. Os candidatos SVM-sigmoid e LR atingiram resultados de 51,11% e 53,30%, respectivamente, próximo à fronteira de 50,00%. Em contrapartida, o conjunto DFT-LDA é o que apresenta a menor variância dos modelos, com 3,43 pontos percentuais, e os melhores resultados. Todos os modelos seguiram para a próxima etapa. É interessante observar que não há uma diferença expressiva entre os valores da acurácia e do F1-score, visto que o número de instâncias de cada classe são muito próximos.

4.4 Construção dos *Ensembles*

Com o conjunto de candidatos definidos, foi realizado uma exploração para a construção do melhor modelo de *ensemble*. As rotinas de Seleção Construtiva, Poda e Stacking foram aplicadas conforme comentado na seção 3.4. Na Tabela 21, pode-se conferir os classificadores de base que compõem cada *ensemble* criado a partir de determinado método de combinação e sua acurácia a partir da validação do conjunto de treinamento. A partir dos candidatos escolhidos para os agrupamentos, é possível verificar que a seleção dos algoritmos não considera apenas a acurácia individual, avaliada nas etapas anteriores. Existe uma mescla entre os classificadores de maior acurácia e de acurácia intermediária dentro do sub-conjunto dos candidatos disponíveis.

Os *ensembles* do conjunto PEC-FEATURES, DFT-PCA e DFT-LDA não apresentaram resultados promissores neste primeiro momento com o conjunto destinado à teste e validação. Os modelos resultantes do conjunto PEC-FEATURES apresentaram acurácias menores do que o classificador individual SVM-polinomial. Para o conjunto DFT-PCA, como só haviam dois candidatos disponíveis, as rotinas de SC e Poda não foram utilizadas. Dessa forma, foram agrupados os dois modelos para voto majoritário de forma arbitrária. Nenhum dos modelos apresentou resultados superiores aos classificador LR. Os agrupamentos resultantes do conjunto DFT-LDA apresentaram resultados preliminares menores do que os apresentados pelos classifica-

Tabela 20 – Candidatos a classificadores de base na etapa de validação. Os modelos com as células hachuradas não seguiram para a próxima etapa.

Conjunto Dados	Classificadores	Acurácia de Validação (%)	F1-score
PEC-FEATURES	SVM-rbf	61,189 $\pm$ 1,872	60,167 $\pm$ 1,971
	SVM-linear	58,519 $\pm$ 1,866	57,212 $\pm$ 1,938
	SVM-poly	63,326 $\pm$ 1,906	62,452 $\pm$ 2,000
	SVM-sigmoid	55,716 $\pm$ 1,937	53,974 $\pm$ 1,993
	KNN	56,711 $\pm$ 1,808	54,363 $\pm$ 1,923
	MLP	45,728 $\pm$ 7,163	39,666 $\pm$ 12,955
	LR	58,694 $\pm$ 1,930	57,667 $\pm$ 2,000
	DT	59,620 $\pm$ 2,391	58,737 $\pm$ 2,530
	NB	60,498 $\pm$ 2,087	59,653 $\pm$ 2,179
PEC-PCA	SVM-rbf	64,332 $\pm$ 1,997	63,133 $\pm$ 2,162
	SVM-poly	46,951 $\pm$ 2,563	40,481 $\pm$ 3,109
	SVM-sigmoid	62,162 $\pm$ 2,084	60,961 $\pm$ 2,227
	KNN	63,344 $\pm$ 1,968	62,257 $\pm$ 2,113
	LR	64,305 $\pm$ 2,061	63,571 $\pm$ 2,147
	DT	40,873 $\pm$ 5,862	34,070 $\pm$ 11,261
	NB	56,003 $\pm$ 2,032	51,071 $\pm$ 2,961
DFT-PCA	SVM-rbf	44,350 $\pm$ 1,662	35,567 $\pm$ 1,462
	SVM-poly	35,042 $\pm$ 0,807	21,576 $\pm$ 1,220
	SVM-sigmoid	51,114 $\pm$ 2,143	49,591 $\pm$ 2,200
	KNN	48,144 $\pm$ 1,873	44,692 $\pm$ 2,142
	MLP	40,533 $\pm$ 8,974	35,512 $\pm$ 10,987
	LR	53,350 $\pm$ 2,031	51,678 $\pm$ 2,144
	DT	38,705 $\pm$ 3,585	,011 $\pm$ 7,697
	NB	44,857 $\pm$ 1,951	39,052 $\pm$ 2,243
DFT-LDA	SVM-rbf	88,743 $\pm$ 1,344	88,740 $\pm$ 1,342
	SVM-linear	88,743 $\pm$ 1,344	88,740 $\pm$ 1,342
	SVM-poly	85,386 $\pm$ 1,844	85,530 $\pm$ 1,839
	SVM-sigmoid	88,743 $\pm$ 1,344	88,740 $\pm$ 1,342
	KNN	88,302 $\pm$ 1,401	88,293 $\pm$ 1,401
	MLP	88,642 $\pm$ 1,372	88,635 $\pm$ 1,370
	LGBM	87,942 $\pm$ 1,402	87,935 $\pm$ 1,399
	LR	88,818 $\pm$ 1,322	88,822 $\pm$ 1,320
	DT	85,750 $\pm$ 4,203	85,568 $\pm$ 4,729
	NB	88,767 $\pm$ 1,305	88,766 $\pm$ 1,304

dores LR e NB individualmente. Um ponto importante a ser comentado para este conjunto foi a dificuldade em criar agrupamentos utilizando o método de Seleção Construtiva, visto que a partir do classificador de base de maior acurácia (LR), a adição de novos classificadores não trazia benefícios. Dessa forma, foi novamente escolhido arbitrariamente um conjunto de classificadores para compor um terceiro conjunto para este *dataset*.

Tabela 21 – Configuração final dos classificadores de base para cada método de combinação do *ensemble*.

Conjunto de Dados	Método de Combinação	Acurácia (%)	Classificadores de base
PEC-FEATURES	Seleção Construtiva (SC)	$63,054 \pm 2,108$	SVC-poly KNN DT
	Poda (P)	$63,255 \pm 5,058$	SVC-poly SVC-rbf DT
	Stacking	$62,522 \pm 2,137$	Nível 0: SVC-poly KNN DT Nível 1: NB
PEC-PCA	Seleção Construtiva (SC)	$66,389 \pm 2,030$	KNN LR
	Poda (P)	$66,389 \pm 2,030$	KNN LR
	Stacking	$63,394 \pm 1,955$	Nível 0: KNN DT Nível 1: SVC-sigmoid
DFT-PCA	Escolha Arbitrária (EA)	$52,555 \pm 2,039$	KNN LR
	Stacking	$49,063 \pm 1,875$	Nível 0: SVC-sigmoid LR Nível 1: LR
DFT - LDA	Escolha Arbitrária (EA)	$88,512 \pm 1,370$	SVC-poly SVC-rbf MLP KNN DT LR
	Poda (P)	$88,752 \pm 1,334$	SVC-poly SVC-rbf SVC-linear SVC-sigmoid LGBM MLP LR NB
	Stacking	$88,519 \pm 1,377$	Nível 0: SVC-poly SVC-rbf SVC-sigmoid MLP KNN DT LR Nível 1: LR

Em contrapartida, os *ensembles* resultantes da Seleção Construtiva e Poda do conjunto PEC-PCA apresentaram resultados superiores aos classificadores individuais. O agrupamento proveniente da combinação por *Stacking* resultou em acurácias menores do que aquelas apresentadas pelos modelos SVM-rbf e LR.

4.5 Etapa de teste

Na etapa de teste, foi utilizado o conjunto previamente separado no início do projeto. Os subconjuntos treinamento e validação foram agrupados e os modelos foram, então, treinados com o subconjunto total. Todos os modelos de *ensemble* e candidatos a classificadores de base foram testados com dez reinicializações para uma avaliação comparativa.

A acurácia para todos os modelos classificadores de base considerados é apresentada na Tabela 22. Para o conjunto PEC-FEATURES, o classificador SVM-polinomial obteve o melhor desempenho, com 63,52% de exatidão, superando o segundo melhor classificador (SVM-rbf) em quase 3 pontos percentuais. A acurácia de teste ficou próxima do resultado visto anteriormente no treinamento e validação, com alguns modelos apresentando uma melhora, como SVM-linear e DT. Em contrapartida, o conjunto PEC-PCA apresentou uma boa generalização, uma vez que a acurácia de teste supera a acurácia de treinamento e validação em todos os modelos. O classificador que apresentou a melhor acurácia foi o SVM-rbf, com 67,04%, seguido pelo KNN com uma defasagem de 1,11 ponto percentual.

O conjunto DFT-PCA apresenta resultados inferiores aos vistos anteriormente com o conjunto de teste. Seus dois modelos ficam próximos do limite de decisão do considerado classificador fraco. Por fim, o conjunto DFT-LDA também apresenta uma boa generalização, com a maioria dos seus candidatos apresentando um aumento de pelo menos 2,5 pontos percentuais na acurácia. Os classificadores SVM-rbf, SVM-linear e SVM-sigmoid apresentam os melhores resultados, com 91,67% de exatidão. Estes mesmos classificadores, apesar de possuírem hiperparâmetros diferentes, apresentam os mesmos resultados. Uma teoria para este comportamento é a convergência para uma mesma hipótese.

O resultado da rodada de teste para os *ensembles* é apresentado na Tabela 23. Nesta etapa, para os métodos de Seleção Construtiva (SC) e Poda (P), foram testadas as combinações por voto majoritário (VM) e voto majoritário enviesado (VME). Para o método de combinação *Stacking*, o nível 0 são os classificadores de base e o nível 1 é o algoritmo de meta-aprendizado.

Os conjuntos de PEC-FEATURE, PEC-PCA e DFT-LDA apresentaram uma boa generalização, com uma melhora nos resultados de teste. De modo geral, o método *Stacking* apresentou o pior desempenho dentro dos métodos analisados e a utilização do Voto Majoritário Enviesado entregou uma melhora na acurácia. Adicionalmente, os melhores resultados médios destes conjuntos (P+VME com 64,593% para PEC-FEATURES, P/SC+VM com 69,630% para PEC-PCA e EA+VME com 91,778% para DFT-LDA) superaram os resultados médios dos classificadores individuais previamente apresentados (SVM-poly com 63,519% para PEC-FEATURES, SVM-rbf com 67,037% para PEC-PCA e SVM-linear com 91,667% para DFT-LDA).

O conjunto DFT-PCA, entretanto, ficou com a acurácia média abaixo da acurácia

Tabela 22 – Resultados dos Candidatos a classificadores de base na etapa de teste.

Conjunto Dados	Classificadores	Acurácia de Teste (%)	F1-score
PEC-FEATURES	SVM-rbf	60,553	59,630
	SVM-linear	56,296	55,079
	SVM-poly	63,519	62,718
	SVM-sigmoid	52,593	50,765
	KNN	53,148	50,248
	LR	57,963	56,972
	DT	60,148	59,705
	NB	58,704	56,972
PEC-PCA	SVM-rbf	67,037	66,092
	SVM-sigmoid	64,815	63,597
	KNN	65,926	64,872
	LR	64,630	63,829
	NB	57,407	51,008
DFT-PCA	SVM-sigmoid	50,000	49,028
	LR	51,667	49,936
DFT-LDA	SVM-rbf	91,667	91,625
	SVM-linear	91,667	91,625
	SVM-poly	87,963	88,042
	SVM-sigmoid	91,667	91,625
	KNN	91,481	91,429
	MLP	91,463	91,423
	LGBM	90,741	90,669
	LR	91,296	91,252
	DT	89,704	89,624
	NB	91,481	91,437

do classificador individual LR. Adicionalmente, o algoritmo de *Stacking* não apresentou bons resultados de generalização, com uma redução de 17,40 pontos percentuais em relação ao treinamento.

4.6 Comparação dos Resultados

Para uma avaliação dos ganhos com a aplicação da metodologia do *ensemble*, a Tabela 24 compara a melhor inicialização dos melhores classificadores individuais e dos métodos de agrupamento para cada conjunto de dados.

O *Dataset* PEC-FEATURES foi o conjunto em que o *ensemble* apresentou a melhora mais significativa, aumentando 2,59 pontos percentuais para a acurácia e 2,86 pontos percentuais na métrica F1 em relação ao classificador individual SVM-polinomial. O conjunto PEC-PCA apresentou uma melhora semelhante, com uma melhora de 2,59 pontos percentuais para a acurácia e 2,66 pontos percentuais na métrica F1 em relação a SVM-rbf. O conjunto DFT-PCA,

Tabela 23 – Resultados finais dos *ensembles* testados.

Melhor modelo dados	Método de Seleção + Combinação	Acurácia Validação (%)	F1-score
PEC-FEATURE	SC + VM	63,611 $\pm$ 0,850	62,925 $\pm$ 0,880
	SC + VME	64,130 $\pm$ 0,764	63,374 $\pm$ 0,878
	P + VM	64,278 $\pm$ 1,175	63,663 $\pm$ 1,183
	P + VME	64,593 $\pm$ 0,872	63,873 $\pm$ 0,859
	Stacking	63,259 $\pm$ 0,857	62,537 $\pm$ 1,000
PEC-PCA	SC + VM	69,630 $\pm$ 0,000	68,770 $\pm$ 0,000
	SC + VME	69,444 $\pm$ 0,000	68,558 $\pm$ 0,000
	P + VM	69,630 $\pm$ 0,000	68,770 $\pm$ 0,000
	P + VME	69,444 $\pm$ 0,000	68,558 $\pm$ 0,000
	Stacking	66,852 $\pm$ 0,976	66,435 $\pm$ 68,330
DFT-PCA	EA + VM	51,574 $\pm$ 0,354	50,444 $\pm$ 0,295
	EA + VME	51,630 $\pm$ 0,308	50,481 $\pm$ 0,273
	Stacking	31,667 $\pm$ 0,000	16,034 $\pm$ 0,000
DFT-LDA	EA + VM	91,759 $\pm$ 0,387	91,722 $\pm$ 0,388
	EA + VME	91,778 $\pm$ 0,222	91,741 $\pm$ 0,225
	P + VM	91,278 $\pm$ 0,130	91,236 $\pm$ 0,132
	P + VME	91,296 $\pm$ 0,083	91,255 $\pm$ 0,084
	Stacking	91,444 $\pm$ 0,216	91,397 $\pm$ 0,216

entretanto, apresentou a menor diferença de acurácia entre os melhores algoritmos, 0,37 pontos percentuais, e 0,93 pontos percentuais na métrica F1. Por fim, o conjunto DFT-LDA apresentou uma melhora de 0,55 pontos percentuais para a acurácia e 0,56 pontos percentuais na métrica F1 quando comparado à SVM-linear.

Tabela 24 – Resultados com a implementação da *ensemble*.

Conjunto Dados	Melhores modelos	Acurácia de Teste (%)	F1-score
PEC-FEATURE	SVM-poly	63,519	62,718
	P + VME	66,111	65,582
PEC-PCA	SVM-rbf	67,037	66,092
	P/SC + VM	69,630	68,770
DFT-PCA	LR	51,667	49,936
	EA + VME	52,037	50,867
DFT-LDA	SVM - linear	91,667	91,625
	EA + VME	92,220	92,188

As Tabelas 25, 26, 27 e 28 apresentam uma comparação complementar a apresentada na Tabela 24, onde é possível visualizar a matriz de confusão dos erros e acertos de cada classe. Uma vez que existe um pequeno desbalanceamento das classes no conjunto de teste, a normalização em relação à classe verdadeira é apresentado à direita para facilitar a análise.

Em relação à comparação dos conjuntos PEC-FEATURES, com a utilização do agrupamento, um aumento de 9,19 pontos percentuais ocorre na classificação correta da classe “Meio”, contudo a classe “Início” reduz seu acerto em 1,63 ponto percentual. Com o classificador individual, a classe que apresentava a maior dificuldade para a classificação correta era a classe “Meio”. Após o agrupamento, a classe com menor percentual passou a ser o “Final”.

Tabela 25 – Matriz de confusão dos melhores resultados (em %) para o conjunto de dados PEC-FEATURES.

		Classe Predita					
		Início	Meio	Final	Início	Meio	Final
SVM-poly Acerto = 63,519%	Início	153	5	26	83,15	02,72	14,13
	Meio	37	92	56	20,00	49,73	30,27
	Final	37	36	98	21,64	21,05	57,31
P + VM Acerto = 66,111%	Início	150	10	24	81,52	05,44	13,04
	Meio	31	109	45	16,76	58,92	24,32
	Final	33	40	98	19,30	23,39	57,31

De forma análoga, em relação à comparação do conjunto PEC-PCA, um aumento expressivo de 14,05 pontos percentuais ocorre na classificação correta da classe “Meio” e as classes “Início” e “Final” reduzem seu acerto em 0,55 e 6,43 pontos percentuais. Novamente, a classe que apresentava a maior dificuldade para a classificação correta era a classe “Meio” com o classificador individual e, após o agrupamento, passou a ser o “Final”.

Tabela 26 – Matriz de confusão dos melhores resultados (em %) para o conjunto de dados PEC-PCA.

		Classe Predita					
		Início	Meio	Final	Início	Meio	Final
SVM-rbf Acerto = 67,037%	Início	163	4	17	88,59	02,17	09,24
	Meio	32	94	59	17,30	50,81	31,89
	Final	35	31	105	20,47	18,13	61,40
P/SC + VME Acerto = 69,630 %	Início	162	7	15	88,04	03,81	08,15
	Meio	33	120	32	17,84	64,86	17,30
	Final	34	43	94	19,88	25,15	54,97

Para o conjunto DFT-PCA, as classes “Meio” e “Final” aumentaram seu acerto em 1,63 e 2,93 pontos percentuais, enquanto a classe “Início” reduziu em 3,25 pontos percentuais. Neste conjunto, a classe que apresenta a maior dificuldade, em ambos os casos, é o “Meio”.

Finalmente, para as matrizes de confusão do conjunto DFT-LDA, as classes “Início” e

Tabela 27 – Matriz de confusão dos melhores resultados (em %) para o conjunto de dados DFT-PCA.

		Classe Predita					
		Início	Meio	Final	Início	Meio	Final
LR Acerto = 51,667%	Início	145	17	22	78,80	09,24	11,96
	Meio	48	65	72	25,95	35,13	38,92
	Final	42	60	69	24,56	35,09	40,35
SC + VME Acerto = 52,037%	Início	139	18	27	75,55	09,78	14,67
	Meio	46	68	71	24,86	36,76	38,38
	Final	40	57	74	23,39	33,33	43,28

“Final” apresentaram um aumento em 0,54 e 1,17 pontos percentuais e a classe “Meio” se manteve constante. Novamente, a classe que apresenta a maior dificuldade, em ambos os casos, é o “Meio”.

Tabela 28 – Matriz de confusão dos melhores resultados (em %) para o conjunto de dados DFT-LDA.

		Classe Predita					
		Início	Meio	Final	Início	Meio	Final
SVM-linear Acerto = 91,667%	Início	175	5	4	95,11	02,71	02,18
	Meio	4	166	15	02,16	89,73	08,11
	Final	4	13	154	02,34	07,60	90,06
EA + VME Acerto = 92,220%	Início	176	5	3	95,65	02,72	01,63
	Meio	4	166	15	02,16	89,73	08,11
	Final	3	12	156	01,75	07,02	91,23

Em todas as matrizes, a classe “Início” apresenta o maior índice de acerto para as duas situações. Dessa forma, as matrizes de confusão acima apresentaram informações concordantes com uma das hipóteses iniciais deste trabalho, onde os sinais provenientes do “Início” seriam aqueles que originariam menos erros, visto que aparentavam uma separação visual. Em conjuntos que possuem um processamento de dados menos complexo, a classe “Meio” se beneficia com a utilização do agrupamento, gerando uma melhora na classificação dos sinais Verdadeiros Positivos. De modo geral, os incidentes de Falso Positivo das classes “Início” e “Meio” estão enviesados para a classe “Final”.

5 Conclusão

Esta pesquisa apresentou um conjunto de técnicas inteligentes direcionadas à classificação de corpos de provas retirados de uma bobina de aço IF que exemplificam as condições do material como recebido para estampagem em indústrias automobilísticas. A utilização de diversos algoritmos de aprendizado de máquina associados a métodos de processamento digital de sinais foram avaliados e comparados para apoio à decisão no diagnóstico PEC. Como abordagem adicional para atender a demandas desta pesquisa, foram realizados ensaios de tração e medições de dureza para uma análise do conjunto avaliado.

Os principais objetivos pretendidos com as rotinas implantadas nesta pesquisa residiram na avaliação do ganho com a utilização de algoritmos de agrupamento. A estrutura para a construção de *ensembles* é composta por três pilares centrais: a geração de um conjunto de classificadores de base candidatos à composição do *ensemble*; a seleção dos candidatos que contribuem positivamente para a capacidade de generalização do sistema; e a combinação dos classificadores de base selecionados para se obter a saída do *ensemble*.

Para a geração dos componentes aptos a participarem da etapa de seleção, uma rotina de otimização utilizada para encontrar a melhor combinação de parâmetros foi implementada. Os algoritmos de Máquina de Vetores de Suporte (SVM), K vizinhos mais próximos (KNN), Árvore de Decisão (DT), Perceptron Multicamadas (MLP), *Gradient Boosting* (GBM), Regressão Logística (LR) e Naive Bayes Gaussiano (NB) foram utilizados como máquinas base para a escolha de modelos satisfatórios. Gerada essa população de classificadores candidatos, sub-conjuntos desses indivíduos foram testados e aqueles que apresentaram os melhores agrupamentos selecionados para a próxima etapa. Foram adotadas aqui três técnicas de seleção distintas (seleção construtiva, poda e *stacking*) e mais duas combinações de componentes (voto majoritário e voto majoritário enviesado). No final, o melhor classificador individual e o melhor *ensemble* para cada conjunto de dados distinto foi comparado para análise das melhorias encontradas.

Dentro do escopo desta pesquisa, objetivou-se também um estudo do comportamento do aço IF quando submetido a cargas distintas. Um ensaio de tração-deformação e de microdureza de Vickers foi realizado para análise dos resultados. Uma relação proporcional foi estabelecida entre carga aplicada e dureza e, consequentemente, tensão residual. Visto que mudanças nas propriedades mecânicas dos materiais desencadeiam mudanças nas propriedades magnéticas, a técnica de correntes parasitas pulsadas é sensível ao acúmulo de tensões residuais no aço IF.

Considerando o desenvolvimento dos algoritmos de classificação dos sinais de PEC, os resultados obtidos nesta pesquisa indicaram que o modelo proposto foi capaz de discriminar de forma satisfatória as diferentes condições do material, atingindo uma acurácia de até 92,22%.

No âmbito da aprendizagem de máquina, os resultados deste estudo mostram que o maior ganho da acurácia está associado à técnica de processamento de sinal utilizada para a extração das informações que irão alimentar a rede. Avaliando os melhores classificadores individuais para cada conjunto de dados, melhorias de até 40 pontos percentuais são obtidas ao utilizar técnicas

mais refinadas, ou através da combinação de rotinas. Comparativamente, a utilização de técnicas de agrupamento apresentou menores ganhos, com um aumento de até 2,59 pontos percentuais. De modo geral, sua utilização apresentou maiores ganhos quando utilizados com conjuntos de dados de acurácia intermediária, e que não apresentem limites nítidos de definição de classes.

A partir das análises apresentadas, foi possível concluir que na utilização de algoritmos de aprendizado de máquinas para classificação dos sinais, uma das etapas de maior importância é o processo de extração de informações para a construção de um conjunto de dados adequado, visto que ele está intimamente relacionado ao resultado máximo alcançado. A utilização do *ensemble* é um bom recurso para a realização de um ajuste fino da acurácia pretendida, na tentativa de otimizar o resultado em detrimento do custo computacional.

Os resultados obtidos nesta pesquisa motivam a realização de trabalhos futuros envolvendo demandas correlatas a esta investigação. Planeja-se continuar a explorar novas metodologias para a extração de informações do sinal de correntes parasitas pulsadas, o que poderia tornar ainda mais relevante a busca pelo melhor algoritmo de aprendizado de máquina. A exploração do perfil dos sinais que apresentam confusões entre classes pode revelar características que, a partir de tratamentos dedicados por técnicas de processamento digital de sinais, incrementem o desempenho de classificação. Adicionalmente, a implementação futura do método proposto em um sistema embarcado dedicado para auxílio à inspeção em campo é uma linha de pesquisa promissora quando analisado o panorama geral do estado da arte. Pretende-se, como uma outra extensão deste trabalho, realizar pesquisas interdisciplinares para uma investigação mais detalhada do aço IF e o que as modificações de suas propriedades mecânicas implicam no ensaio PEC realizado. Com um maior entendimento do material, seria possível diferenciar referências de anomalias. Outra frente que poderia ser avaliada é a questão da utilização de outra técnica não destrutiva e, com o auxílio de um *ensemble*, somá-las de forma que a classificação se torne mais robusta devido à maior diversidade de informação.

Referências

- AL-MEJIBLI, I. S. et al. Performance evaluation of kernels in support vector machine. In: *2018 1st Annual International Conference on Information and Sciences (AiCIS)*. [S.l.: s.n.], 2018. p. 96–101. Cited on page 21.
- ALAMIN, M. et al. Principal component analysis of pulsed eddy current response from corrosion in mild steel. *IEEE SENSORS JOURNAL*, v. 12, n. 8, p. 2548–2553, 2012. Cited on page 11.
- ANGANI, C. S. et al. SVM classification of thickness and lift-off using transient eddy current oscillation method. *2016 IEEE International Instrumentation and Measurement Technology Conference Proceedings*, p. 1–6, 2016. Cited on page 22.
- ANTONIOU, A. *Digital Signal Processing. Signals, systems and filters*. [S.l.]: McGraw-Hill, 2006. Cited on page 14.
- ARJUN, V. et al. Optimisation of pulsed eddy current probe for detection of sub-surface defects in stainless steel plates. *Sensors and Actuators A: Physical*, v. 226, p. 69–75, 2015. Cited 2 times on pages 7 e 9.
- Ben Seghier, M. E. A.; HöCHE, D.; ZHELUDKEVICH, M. Prediction of the internal corrosion rate for oil and gas pipeline: Implementation of ensemble learning techniques. *Journal of Natural Gas Science and Engineering*, v. 99, p. 104425, 2022. ISSN 1875-5100. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S1875510022000191>>. Cited on page 36.
- BENTO, C. *Multilayer Perceptron Explained with a Real-Life Example and Python Code: Sentiment Analysis*. [S.l.]: Medium, 2021. Cited 2 times on pages 7 e 26.
- BIEBER, J. A. et al. Time-gating of pulsed eddy current signals for defect characterization and discrimination in aircraft lap-joints. *Review of Progress in Quantitative Nondestructive Evaluation*, v. 16, p. 1915–1921, 1997. Cited on page 10.
- BIEDRZYCKI, J. *Integration of decision trees in geometric space*. Dissertação (Tese de doutorado) — Wrocław University of Science and Technology, 2021. Cited on page 23.
- BINFENG, Y.; FEILU, L.; DAN, H. Research on edge identification of a defect using pulsed eddy current based on principal component analysis. *NDTE International*, v. 40, p. 294—299, 2007. Cited on page 16.
- BRENCE, J. R.; BROWN, D. E. Data mining corrosion from eddy current non-destructive tests. *Computers Industrial Engineering*, v. 43, n. 4, p. 821–840, 2002. ISSN 0360-8352. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S0360835202001420>>. Cited on page 29.
- BROWNLEE, J. *Ensemble Learning Algorithms With Python*. [S.l.: s.n.], 2020. Cited on page 36.
- BUI, T.-D.; PHAM, V.-D.; CUNG, T.-L. Multilayer perceptron neural network and eddy current technique for estimation of the crack depth on massive metal structures. *Journal of Military Science and Technology*, n. 77, p. 3–12, 2022. Disponível em: <<https://online.jmst.info/index.php/jmst/article/view/291>>. Cited on page 26.
- CALLISTER Jr., W. D.; RETHWISCH, D. G. *Ciência e engenharia de materiais : uma introdução*. [S.l.]: LTC - Livros Técnicos e Científicos Editora Ltda, 2016. Cited on page 2.

- CERVANTES, J. et al. A comprehensive survey on support vector machine classification: Applications, challenges and trends. *Neurocomputing*, v. 408, p. 189–215, 2020. ISSN 0925-2312. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S0925231220307153>>. Cited on page 18.
- CHARBUTY, B.; ABDULAZEEZ, A. Classification based on decision tree algorithm for machine learning. *Journal of Applied Science and Technology Trends*, v. 2, n. 01, p. 20 – 28, 2021. Disponível em: <<https://www.jastt.org/index.php/jasttpath/article/view/65>>. Cited on page 25.
- CHEN, X. et al. Study on defect classification in multi-layer structures based on fisher linear discriminate analysis by using pulsed eddy current technique. *NDTE International*, v. 67, p. 46–54, 2014. Cited 2 times on pages 17 e 21.
- CHEN, Z.; JOMDECHA, C.; XIE, S. *Eddy Current Testing*. [S.l.]: Springer International Publishing, 2019. Cited 4 times on pages 7, 8, 9 e 10.
- CHOU, J.-S.; NGO, N.-T.; CHONG, W. K. The use of artificial intelligence combiners for modeling steel pitting risk and corrosion rate. *Engineering Applications of Artificial Intelligence*, v. 65, p. 471–483, 2017. ISSN 0952-1976. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S0952197616301737>>. Cited 2 times on pages 3 e 36.
- COELHO, G. P. *Geração, Seleção e Combinação de Componentes para Ensembles de Redes Neurais Aplicadas a Problemas de Classificação*. Dissertação (Dissertação de Mestrado) — Universidade Estadual de Campinas, 2006. Cited 2 times on pages 34 e 36.
- COVER, T. M. Geometrical and statistical properties of systems of linear inequalities with applications in pattern recognition. *IEEE Transactions on Electronic Computers*, n. 14, p. 326–334, 1965. Cited on page 20.
- DENG, X. et al. An improved method to construct basic probability assignment based on the confusion matrix for classification problem. *Information Sciences*, v. 340-341, p. 250–261, 2016. ISSN 0020-0255. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S002002551600044X>>. Cited on page 30.
- DIETTERICH, T. G. Ensemble methods in machine learning. In: *Multiple Classifier Systems*. Berlin, Heidelberg: Springer Berlin Heidelberg, 2000. p. 1–15. Cited 3 times on pages 8, 33 e 34.
- DINO, H. I.; ABDULRAZZAQ, M. B. Facial expression classification based on svm, knn and mlp classifiers. In: *2019 International Conference on Advanced Science and Engineering (ICOASE)*. [S.l.: s.n.], 2019. p. 70–75. Cited on page 23.
- DREISEITL, S.; OHNO-MACHADO, L. Logistic regression and artificial neural network classification models: a methodology review. *Journal of Biomedical Informatics*, v. 35, n. 5, p. 352–359, 2002. ISSN 1532-0464. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S1532046403000340>>. Cited on page 28.
- DU, K.-L.; SWAMY, M. N. S. *Neural Network and Statistical Learning*. [S.l.]: Springer, 2014. Cited 6 times on pages 18, 20, 25, 31, 34 e 36.
- DUDA, R. O.; HART, P. E.; STORK, D. G. *Pattern Classification*. [S.l.]: John Wiley Sons, 2001. Cited 2 times on pages 15 e 17.
- EFRON, B.; TIBSHIRANI, R. *An Introduction to the Bootstrap*. [S.l.: s.n.], 1993. Cited on page 35.

- FAN, M. et al. Frequency optimization for enhancement of surface defect classification using the eddy current technique. *Sensors*, v. 649, n. 16, 2016. Cited on page 12.
- GE, J. et al. Defect classification using postpeak value for pulsed eddy-current technique. *Sensors*, v. 20, n. 12, 2020. Cited on page 11.
- GÉRON, A. *Hands-On Machine Learning with Scikit-Learn, Keras, and TensorFlow: Concepts, Tools, and Techniques to Build Intelligent Systems*. [S.l.]: O'reilly, 2019. Cited 4 times on pages 17, 18, 21 e 35.
- GHONI, R. et al. Defect characterization based on eddy current technique: Technical review. *Advances in Mechanical Engineering*, v. 6, p. 182496, 2014. Cited on page 26.
- GONCALVES, A. R. Máquina de vetores suporte. *Universidade de Campinas*, p. 1–18, 2010. Cited on page 20.
- GROS, X. E. *NDTData Fusion*. 1. ed. USA: Halsted Press, 1996. Acesso em: 8 fev 2022. Cited on page 12.
- GUPTA, M. et al. Advances in applications of non-destructive testing (ndt): A review. *Advances in Materials and Processing Technologies*, p. 1–22, 2021. Cited on page 6.
- HABIBALAH, A. et al. Improving pulse eddy current and ultrasonic testing stress measurement accuracy using neural network data fusion. *IET Science, Measurement Technology*, v. 9, n. 4, p. 514 – 521, 2015. Cited 3 times on pages 3, 12 e 26.
- HABIBALAH, A.; SAFIZADEH, M. S. Forward to residual stress measurement by using pulsed eddy current technique. *Insight*, v. 55, p. 492–497, 2013. Cited 2 times on pages 7 e 10.
- HAMED, Y. et al. An application of k-nearest neighbor interpolation on calibrating corrosion measurements collected by two non-destructive techniques. In: *2015 IEEE 3rd International Conference on Smart Instrumentation, Measurement and Applications (ICSIMA)*. [S.l.: s.n.], 2015. p. 1–5. Cited on page 23.
- HAN, J.; KAMBER, M.; PEI, J. *Data Mining: Concepts and Techniques*. [S.l.]: Morgan Kaufmann Series in Data Management Systems, 2012. Cited 4 times on pages 7, 13, 15 e 16.
- HASTIE, T.; TIBSHIRANI, R.; FRIEDMAN, J. *The Elements of Statistical Learning*. [S.l.: s.n.], 2009. Cited on page 26.
- HAYKIN, S. *Neural Networks and Learning Machines*. [S.l.]: Pearson Education, Inc., 2009. Cited 2 times on pages 7 e 19.
- HE, Y. et al. Pec defect automated classification in aircraft multi-ply structures with interlayer gaps and lift-offs. *NDTE International*, v. 53, p. 39–46, 2013. Cited on page 21.
- HE, Y. et al. Support vector machine and optimised feature extraction in integrated eddy current instrument. *Measurement*, v. 46, p. 764–774, 2013. Cited on page 21.
- HEINGÄRTNER, J.; BORN, M.; HORA, P. Online acquisition of mechanical material properties of sheet metal for the prediction of product quality by eddy current. *10th European Conference on NDT*, v. 1, p. 363–369, 2010. Cited on page 3.
- HELLIER, C. *Handbook of Nondestructive Evaluation, Third Edition*. [S.l.]: McGraw-Hill Professional, 2020. Cited 2 times on pages 6 e 7.

- HERTER, S. et al. Machine learning based preprocessing to ensure validity of cross-correlated ultrasound signals for time-of-flight measurements. *Journal of Nondestructive Evaluation*, v. 40, 2021. ISSN 0963-8695. Cited on page 17.
- HO, G. *Linear Discriminant Analysis for Starters*. 2017. Disponível em: <<https://www.georgeho.org/lda/>>. Cited 2 times on pages 7 e 16.
- HOILE, S. Processing and properties of mild interstitial free steels. *Materials Science and Technology*, v. 16, n. 10, p. 1079–1093, 2000. Cited 2 times on pages 7 e 1.
- HWANG, Y.-I. et al. Detection and classification of artificial defects on stainless steel plate for a liquefied hydrogen storage vessel using short-time fourier transform of ultrasonic guided waves and linear discriminant analysis. *Applied Sciences*, v. 12, 2022. Cited on page 17.
- JENA, M.; DEHURI, S. Decision tree for classification and regression: A state-of-the art review. *Informatica*, v. 44, p. 405–420, 2020. Disponível em: <<https://www.informatica.si/index.php/informatica/article/view/3023>>. Cited on page 25.
- JIN, H. et al. Defect recognition and classification techniques for multi-layer tubular structures in oil and gas wells by using pulsed eddy current testing. In: GELMAN, L. et al. (Ed.). *Advances in Condition Monitoring and Structural Health Monitoring*. Singapore: Springer Singapore, 2021. p. 499–509. Cited on page 25.
- JO, N. H.; LEE, H.-B. A novel feature extraction for eddy current testing of steam generator tubes. *NDT E International*, v. 42, n. 7, p. 658–663, 2009. ISSN 0963-8695. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S0963869509000942>>. Cited on page 26.
- KAIYU, L. et al. Estimation method of mechanical properties of ferromagnetic materials based on pulsed eddy current. *Russian Journal of Nondestructive Testing*, v. 57, n. 2, p. 170–179, 2021. Cited on page 12.
- KANCHERLA, D.; BODAPATI, J. D.; N, V. Effect of different kernels on the performance of an SVM based classification. *International Journal of Recent Technology and Engineering (IJRTE)*, v. 7, p. 1–6, 2019. Cited on page 20.
- KE, G. et al. Lightgbm: A highly efficient gradient boosting decision tree. *Advances in neural information processing systems*, v. 30, p. 3146–3154, 2017. Cited 2 times on pages 27 e 45.
- KEPRATE, A.; RATNAYAKE, R. M. C. Using gradient boosting regressor to predict stress intensity factor of a crack propagating in small bore piping. In: *2017 IEEE International Conference on Industrial Engineering and Engineering Management (IEEM)*. [S.l.: s.n.], 2017. p. 1331–1336. Cited on page 26.
- KESHARAJU, M. et al. Ultrasonic sensor based defect detection and characterisation of ceramics. *Ultrasonics*, v. 54, n. 1, p. 312–317, 2014. ISSN 0041-624X. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S0041624X13002217>>. Cited on page 17.
- KINGSFORD, C.; SALZBERG, S. L. What are decision trees? *Nature Biotechnology*, v. 26, p. 1101–1103, 2008. Disponível em: <<https://www.nature.com/articles/nbt0908-1011#citeas>>. Cited 2 times on pages 23 e 24.
- KONG, Q.; SIAUW, T.; BAYEN, A. *Python Programming and Numerical Methods. A Guide for Engineers and Scientists*. [S.l.]: Academic Press, 2020. Cited 2 times on pages 7 e 14.
- KUBAT, M. *An Introduction to Machine Learnings*. [S.l.]: Springer International Publishing, 2017. Cited on page 13.

- KUNCHEVA, L. I. *Combining Patterns Classifiers*. [S.l.]: Wiley-interscience, 2004. Cited 6 times on pages 15, 22, 25, 29, 32 e 35.
- LAROCCA, C. B. et al. Wall thinning characterization of composite reinforced steel tube using frequency-domain pec technique and neural networks. *Journal of Nondestructive Evaluation*, v. 78, p. 37–51, 2018. Cited on page 15.
- LE, M. et al. Auto-detection of hidden corrosion in an aircraft structure by electromagnetic testing: A machine-learning approach. *Applied Sciences*, v. 12, n. 10, 2022. ISSN 2076-3417. Disponível em: <<https://www.mdpi.com/2076-3417/12/10/5175>>. Cited on page 23.
- LIU, Y. et al. Pulsed eddy current data analysis for the characterization of the second-layer discontinuities. *Journal of Nondestructive Evaluation*, v. 38, 2019. Disponível em: <<https://link.springer.com/article/10.1007/s10921-018-0545-6#citeass>>. Cited 2 times on pages 25 e 26.
- MAALOUF, M. Logistic regression in data analysis: an overview. *International Journal of Data Analysis Techniques and Strategies*, v. 3, n. 3, p. 281–299, 2011. Cited on page 28.
- MARINO, M.; VIRUPAKSHAPPA, K.; ORUKLU, E. A stacked ensemble neural network classifier for ultrasonic non-destructive evaluation applications. In: *2020 IEEE International Ultrasonics Symposium (IUS)*. [S.l.: s.n.], 2020. p. 1–4. Cited on page 36.
- MESQUITA, P. S. B. *Um modelo de regressão logística para avaliação dos programas de pós-graduação no Brasil*. Dissertação (Dissertação de Mestrado) — Universidade Estadual Do Norte Fluminense, 2014. Cited on page 28.
- Microsoft Corporation. *Features*. Microsoft Corporation, 2022. Disponível em: <<https://lightgbm.readthedocs.io/en/latest/Features.html>>. Cited 2 times on pages 7 e 28.
- MIN, J.-K.; HONG, J.-H.; CHO, S.-B. Ensemble approaches of support vector machines for multiclass classification. *Ghosh A., De R.K., Pal S.K. (eds) Pattern Recognition and Machine Intelligence. PReMI 2007. Lecture Notes in Computer Science*, n. 4815, 2007. Cited on page 18.
- MOHRI, M.; ROSTAMIZADEH, A.; TALWALKAR, A. *Foundations of Machine Learning*. [S.l.]: The MIT Press, 2018. Cited 2 times on pages 7 e 24.
- MOROZOV, M.; TIAN, G. Y.; WITHERS, P. J. Noncontact evaluation of the dependency of electrical conductivity on stress for various al alloys as a function of plastic deformation and annealing. *Journal of Applied Physics*, v. 108, 2010. Cited on page 9.
- MOROZOV, M.; TIAN, G. Y.; WITHERS, P. J. The pulsed eddy current response to applied loading of various aluminium alloys. *NDTE International*, v. 43, p. 493–500, 2010. Cited on page 2.
- MURUGAN, A.; NAIR, S. H.; KUMAR, K. P. S. Detection of skin cancer using svm, random forest and knn classifiers. *Journal of Medical Systems volume*, v. 43, 2019. Cited on page 23.
- NAFIAH, F. et al. Decoupling the influence of wall thinning and cladding thickness variations in pulsed eddy current using principal component analysis. *IEEE Sensors Journal*, v. 21, n. 19, p. 22011–22018, 2021. Cited on page 16.
- NAQA, I. E.; MURPHY, M. J. What Is Machine Learning?. In: El Naqa I., Li R., Murphy M. (eds) *Machine Learning in Radiation Oncology*. [S.l.]: Cham Springer, 2015. Cited on page 17.
- NASCIMENTO, P. de A. *Aplicando Ensemble para classificação de textos curtos em português do Brasil*. Dissertação (Dissertação de Mestrado) — Universidade Federal de Pernambuco, 2019. Cited on page 18.

- NASCIMENTO, R. F. F. et al. O algoritmo support vector machines (SVM): avaliação da separação ótima de classes em imagens CCD-CBERS-2. *Anais XIV Simpósio Brasileiro de Sensoriamento Remoto*, p. 2079–2086, 2009. Cited on page 18.
- NATEKIN, A.; KNOLL, A. Gradient boosting machines, a tutorial. *Neurorobot*, v. 7:21, 2013. ISSN 0957-4174. Disponível em: <<https://www.frontiersin.org/articles/10.3389/fnbot.2013.00021/full>>. Cited on page 26.
- NICCOLAI, A. et al. Machine learning-based detection technique for NDT in industrial manufacturing. *Mathematics*, v. 1251, n. 9, p. 1–16, 2021. Cited on page 3.
- NORCINO, A. B. *Uma contribuição ao estudo das tensões residuais no fresamento*. Dissertação (Dissertação de Mestrado) — Universidade Estadual Paulista. Faculdade de Engenharia de Ilha Solteira, 2013. Cited on page 2.
- ONAN, A.; KORUKOĞLU, S.; BULUT, H. A multiobjective weighted voting ensemble classifier based on differential evolution algorithm for text sentiment classification. *Expert Systems with Applications*, v. 62, p. 1–16, 2016. ISSN 0957-4174. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S095741741630286X>>. Cited on page 29.
- OPITZ, D.; MACLIN, R. Popular ensemble methods: An empirical study. *Journal of Artificial Intelligence Research*, v. 11, p. 169–198, 1999. Cited on page 32.
- PAN, M. et al. Pec frequency band selection for locating defects in two-layer aircraft structures with air gap variations. *IEEE Transactions on Instrumentation and Measurement*, v. 62, p. 2849 – 2856, 2013. Cited on page 15.
- PARIKH, D. et al. Ensemble of classifiers approach for NDT data fusion. *2004 IEEE International Ultrasonics, Ferroelectrics, and Frequency Control Joint 50th Anniversary Conference*, p. 1062–1065, 2004. Cited 3 times on pages 3, 32 e 36.
- PEDREGOSA, F. et al. Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research*, v. 12, p. 2825–2830, 2011. Cited 5 times on pages 31, 45, 46, 47 e 48.
- PENG, Y. et al. Defect classification using pec responses based on power spectral density analysis combined with emd and eemd. *NDT E International*, v. 78, p. 37–51, 2016. Cited 2 times on pages 15 e 17.
- PIAO, G. et al. A novel pulsed eddy current method for high-speed pipeline inline inspection. *Sensors and Actuators A: Physical*, v. 295, p. 244 – 258, 2019. Cited on page 11.
- PRATES, R. M. *Deteção e classificação de isoladores de redes de distribuição utilizando inteligência computacional e geração artificial de imagens*. Dissertação (Tese de doutorado) — Universidade Federal da Bahia, 2020. Cited on page 31.
- QIU, X. et al. Defect classification by pulsed eddy current technique in con-casting slabs based on spectrum analysis and wavelet decomposition. *Sensors and Actuators A: Physical*, v. 203, p. 272–281, 2013. Cited 3 times on pages 15, 17 e 29.
- QIU, X. B. et al. Defect classification by pulsed eddy-current technique based on power spectral density analysis combined with wavelet transform. *IEEE Transactions on Magnetics*, v. 50, n. 9, p. 1–8, 2014. Cited on page 29.
- RAMOS, H. G. et al. An SVM approach with electromagnetic methods to assess metal plate thickness. *Measurement*, v. 54, p. 201–206, 2014. Cited on page 22.

- REN, Y.; HU, F.; MIAO, H. The optimization of kernel function and its parameters for SVM in well-logging. *13th International Conference on Service Systems and Service Management (ICSSSM)*, p. 1–5, 2016. Cited on page 21.
- RODRIGUES, L. F. et al. Carburization level identification in industrial hp pipes using ultrasonic evaluation and machine learning. *Ultrasonics*, v. 94, p. 145–151, 2019. ISSN 0041-624X. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S0041624X18302889>>. Cited on page 23.
- ROKACH, L. Decision forest: Twenty years of research. *Information Fusion*, v. 27, p. 111–125, 2016. ISSN 1566-2535. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S1566253515000561>>. Cited on page 23.
- ROMANYUK, M. et al. Correlation between structural aspects and mechanical properties of an interstitial free steel for automotive application. *Advanced Materials Proceedings*, International Association of Advanced Materials, v. 3, n. 6, p. 408–413, 2018. ISSN 2002-4428. Disponível em: <https://amp.iaaonline.org/article_16152.html>. Cited on page 2.
- SAGI, O.; ROKACH, L. Ensemble learning: A survey. *WIREs Data Mining and Knowledge Discovery*, v. 8, n. 4, p. e1249, 2018. Disponível em: <<https://wires.onlinelibrary.wiley.com/doi/abs/10.1002/widm.1249>>. Cited on page 31.
- SHA, J. et al. Noncontact and nondestructive evaluation of heat-treated bearing rings using pulsed eddy current testing. *Journal of Magnetism and Magnetic Materials*, v. 521, p. 167516, 2021. ISSN 0304-8853. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S0304885320324835>>. Cited on page 16.
- SHABAN, W. M. et al. A new covid-19 patients detection strategy (cpds) based on hybrid feature selection and enhanced knn classifier. *Knowledge-Based Systems*, v. 205, p. 106270, 2020. ISSN 0950-7051. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S0950705120304573>>. Cited on page 23.
- SHAH, K. et al. A comparative analysis of logistic regression, random forest and knn models for the text classification. *Augmented Human Research*, v. 5, 2020. Cited on page 23.
- SHARMA, A.; PALIWAL, K. K. Linear discriminant analysis for the small sample size problem: an overview. *International Journal of Machine Learning and Cybernetics*, v. 6, p. 443–454, 2014. Cited on page 16.
- SHEYKHOUSA, M. et al. Support vector machine versus random forest for remote sensing image classification: A meta-analysis and systematic review. *EEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, v. 13, p. 6308–6325, 2020. Cited 2 times on pages 7 e 20.
- SILVA, G. et al. A constrained neural classifier for pulsed eddy current based flaw detection in industrial pipes. In: FERNANDES, B. J. T.; {Pereira Júnior}, A. (Ed.). *Anais do 14 Congresso Brasileiro de Inteligência Computacional*. Curitiba, PR: ABRICOM, 2019. p. 1–6. ISBN 978-856997201-3. Cited 2 times on pages 12 e 16.
- SILVA Jr., M. M. et al. Intelligent embedded system for decision support in pulsed eddy current corrosion detection using extreme learning machine. *Measurement*, v. 185, p. 1–11, 2021. Cited 2 times on pages 3 e 26.
- SKANSI, S. *Introduction to Deep Learning*. [S.l.]: Springer International Publishing AG, 2018. Cited on page 18.

- SOPHIAN, A.; TIAN, G.; FAN, M. Pulsed eddy current non-destructive testing and evaluation: A review. *Chinese Journal of Mechanical Engineering*, v. 30, p. 500–514, 2017. Cited 2 times on pages 6 e 7.
- SOPHIAN, A. et al. A feature extraction technique based on principal component analysis for pulsed eddy current NDT. *NDT International*, v. 36, p. 37–41, 2003. Cited 2 times on pages 10 e 16.
- SOUZA, D. A. de et al. *Determinação de tensores residuais em materiais metálicos utilizando micro-indentação*. Dissertação — VII Congresso Nacional de Engenharia Mecânica, 2012. Cited on page 2.
- STOTT, C. A. *Pulsed Eddy Current Inspection of Second Layer Wing Structure*. Dissertação (Dissertação de Mestrado) — Royal Military College of Canada, 2014. Cited on page 29.
- STOTT, C. A. et al. Pulsed eddy current detection of cracks in multilayer aluminum lap joints. *IEEE Sensors Journal*, v. 15, n. 2, p. 956–962, 2015. Cited on page 29.
- SUROWIECKI, J. *The Wisdom of Crowds: Why the Many Are Smarter Than the Few and How Collective Wisdom Shapes Business, Economies, Societies and Nations*. [S.l.]: Doubleday, 2004. Cited on page 32.
- THARWAT, A. et al. Linear discriminant analysis: A detailed tutorial. *AI Communications*, v. 30, p. 169–190, 2017. Cited on page 17.
- TIAN, G. et al. Wavelet-based PCA defect classification and quantification for pulsed eddy current NDT. *IEE Proceedings*, v. 154, n. 4, p. 141–148, 2005. Cited 2 times on pages 11 e 16.
- TIAN, G. Y. et al. Research on spectral response of pulsed eddy current and NDE applications. *Sensors and Actuators A: Physical*, v. 189, p. 313–320, 2013. Cited on page 11.
- UEMURA, C. O. *Probability of detection for Pulsed Eddy Current inspection of second layer cracks in aircraft lap-joint structures*. Dissertação (Dissertação de Mestrado) — Royal Military College of Canada, 2018. Cited on page 8.
- VEJDANNIK, M.; SADR, A. Application of statistical signal processing techniques to ultrasound signals for automatic microstructural characterization and classification. In: *Proceedings of the 3rd Iranian International NDT Conference*. Tehran, Iran: [s.n.], 2016. p. 1–6. Cited on page 17.
- VEJDANNIK, M. et al. *Signal Processing for NDE*. [S.l.]: Springer International Publishing, 2019. Cited 3 times on pages 13, 15 e 17.
- WANG, D.; ZHANG, Y.; ZHAO, Y. Lightgbm: An effective mirna classification method in breast cancer patients. In: *ICCB 2017*. [S.l.: s.n.], 2017. Cited on page 28.
- WANG, H. et al. Effect of rare earth ce on deep stamping properties of high-strength interstitial-free steel containing phosphorus. *Materials*, v. 13, n. 6, 2020. ISSN 1996-1944. Disponível em: <<https://www.mdpi.com/1996-1944/13/6/1473>>. Cited on page 1.
- WEI, G. et al. Physic perspective fusion of electromagnetic acoustic transducer and pulsed eddy current testing in non-destructive testing system. *Philosophical Transactions of the Royal Society A*, v. 378, p. 1–17, 2020. Cited on page 12.
- WEN, D. et al. Adjusting loi for enhancement of pulsed eddy current thickness measurement. *IEEE Transactions on Instrumentation and Measurement*, v. 69, n. 2, p. 521–527, 2020. Cited on page 11.

WUNDERLICH, C.; TSCHÖPE, C.; DUCKHORN, F. Advanced methods in NDE using machine learning approaches. *44th Annual Review of Progress in Quantitative Nondestructive Evaluation, AIP Conference Proceedings 1949, 020022*, v. 37, 2018. Cited on page 3.

ZHANG, C.; MA, Y. *Ensemble Machine Learning. Methods and Applications*. [S.l.]: Springer New York Dordrecht Heidelberg London, 2012. Cited 4 times on pages 7, 33, 34 e 35.

ZHANG, W. et al. A distributed storage and computation k-nearest neighbor algorithm based cloud-edge computing for cyber-physical-social systems. *IEEE Access*, v. 8, p. 50118–50130, 2020. Cited 2 times on pages 7 e 22.