

VINÍCIUS RESENDE ROCHA

**OTIMIZAÇÃO PARAMÉTRICA DE REDES NEURAIS
ARTIFICIAIS APLICADAS EM PROCESSO DE
SOLDAGEM COM DATASETS REDUZIDOS**



UNIVERSIDADE FEDERAL DE UBERLÂNDIA
FACULDADE DE ENGENHARIA MECÂNICA

2025

VINÍCIUS RESENDE ROCHA

**OTIMIZAÇÃO PARAMÉTRICA DE REDES NEURAIIS
ARTIFICIAIS APLICADAS EM PROCESSO DE SOLDAGEM COM
DATASETS REDUZIDOS**

Tese submetida ao Programa de Pós-graduação em Engenharia Mecânica da Universidade Federal de Uberlândia, como requisito parcial para a obtenção do título de **Doutor em Engenharia Mecânica**.

Área de Concentração: Materiais e Processos de Fabricação.

Orientador: Prof. Dr. Luiz Eduardo dos Santos Paes

**UBERLÂNDIA - MG
2025**

Ficha Catalográfica Online do Sistema de Bibliotecas da UFU
com dados informados pelo(a) próprio(a) autor(a).

R672 Rocha, Vinícius Resende, 1994-
2025 Otimização Paramétrica De Redes Neurais Artificiais
Aplicadas Em Processo De Soldagem Com Datasets Reduzidos
[recurso eletrônico] / Vinícius Resende Rocha. - 2025.

Orientador: Luiz Eduardo dos Santos Paes.
Tese (Doutorado) - Universidade Federal de Uberlândia,
Pós-graduação em Engenharia Mecânica.
Modo de acesso: Internet.
Disponível em: <http://doi.org/10.14393/ufu.te.2025.92>
Inclui bibliografia.
Inclui ilustrações.

1. Engenharia mecânica. I. Paes, Luiz Eduardo dos
Santos, 1991-, (Orient.). II. Universidade Federal de
Uberlândia. Pós-graduação em Engenharia Mecânica. III.
Título.

CDU: 621

Bibliotecários responsáveis pela estrutura de acordo com o AACR2:

Gizele Cristine Nunes do Couto - CRB6/2091
Nelson Marcos Ferreira - CRB6/3074



UNIVERSIDADE FEDERAL DE UBERLÂNDIA
Coordenação do Programa de Pós-Graduação em Engenharia
Mecânica

Av. João Naves de Ávila, nº 2121, Bloco 1M, Sala 212 - Bairro Santa Mônica, Uberlândia-MG, CEP 38400-902

Telefone: (34) 3239-4282 - www.posmecanicaufu.com.br - secposmec@mecanica.ufu.br



ATA DE DEFESA - PÓS-GRADUAÇÃO

Programa de Pós-Graduação em:	Engenharia Mecânica				
Defesa de:	Tese de Doutorado Acadêmico, número 394, PPGEM				
Data:	07/02/2025	Hora de início:	14:00	Hora de encerramento:	17:00
Matrícula do Discente:	12113EMC008				
Nome do Discente:	Vinicius Resende Rocha				
Título do Trabalho:	OTIMIZAÇÃO PARAMÉTRICA DE REDES NEURAIS ARTIFICIAIS APLICADAS EM PROCESSO DE SOLDAGEM COM DATASETS REDUZIDOS				
Área de concentração:	Materiais e Processos de Fabricação				
Linha de pesquisa:	Processos de Fabricação (Usinagem e Soldagem)				
Projeto de Pesquisa de vinculação:	Não está vinculado a projeto de pesquisa				

Reuniu-se por videoconferência, a Banca Examinadora, designada pelo Colegiado do Programa de Pós-graduação em Engenharia Mecânica, assim composta: Prof. Dr. Leonardo Rosa Ribeiro da Silva - UFU; Prof. Dr. Julian Arnaldo Avila Diaz - UNESP; Prof. Dr. Fran Sergio Lobato - UFU; Prof. Dr. Hipólito Domingo Carvajal Fals - UNICAMP; Prof. Dr. Washington Martins da Silva Junior - UFU e Prof. Dr. Luiz Eduardo dos Santos Paes - UFU; orientador(a) do(a) candidato(a).

Iniciando os trabalhos o(a) presidente da mesa, Dr. Luiz Eduardo dos Santos Paes, apresentou a Comissão Examinadora e o candidato(a), agradeceu a presença do público, e concedeu ao Discente a palavra para a exposição do seu trabalho. A duração da apresentação do Discente e o tempo de arguição e resposta foram conforme as normas do Programa.

A seguir o senhor(a) presidente concedeu a palavra, pela ordem sucessivamente, aos(às) examinadores(as), que passaram a arguir o(a) candidato(a). Ultimada a arguição, que se desenvolveu dentro dos termos regimentais, a Banca, em sessão secreta, atribuiu o resultado final, considerando o(a) candidato(a):

Aprovado.

Esta defesa faz parte dos requisitos necessários à obtenção do título de Doutor.

O competente diploma será expedido após cumprimento dos demais requisitos, conforme as normas do Programa, a legislação pertinente e a regulamentação interna da UFU.

Nada mais havendo a tratar foram encerrados os trabalhos. Foi lavrada a presente ata que após lida e achada conforme foi assinada pela Banca Examinadora.



Documento assinado eletronicamente por **Leonardo Rosa Ribeiro da Silva, Professor(a) do Magistério Superior**, em 20/02/2025, às 08:32, conforme horário oficial de Brasília, com fundamento no art. 6º, § 1º, do [Decreto nº 8.539, de 8 de outubro de 2015](#).



Documento assinado eletronicamente por **Julian Arnaldo Avila Diaz, Usuário Externo**, em 20/02/2025, às 08:35, conforme horário oficial de Brasília, com fundamento no art. 6º, § 1º, do [Decreto nº 8.539, de 8 de outubro de 2015](#).



Documento assinado eletronicamente por **Fran Sergio Lobato, Professor(a) do Magistério Superior**, em 20/02/2025, às 10:42, conforme horário oficial de Brasília, com fundamento no art. 6º, § 1º, do [Decreto nº 8.539, de 8 de outubro de 2015](#).



Documento assinado eletronicamente por **Washington Martins da Silva Junior, Professor(a) do Magistério Superior**, em 20/02/2025, às 15:50, conforme horário oficial de Brasília, com fundamento no art. 6º, § 1º, do [Decreto nº 8.539, de 8 de outubro de 2015](#).



Documento assinado eletronicamente por **Hipolito Domingo Carvajal Fals, Usuário Externo**, em 20/02/2025, às 20:39, conforme horário oficial de Brasília, com fundamento no art. 6º, § 1º, do [Decreto nº 8.539, de 8 de outubro de 2015](#).



Documento assinado eletronicamente por **Luiz Eduardo dos Santos Paes, Professor(a) do Magistério Superior**, em 21/02/2025, às 10:51, conforme horário oficial de Brasília, com fundamento no art. 6º, § 1º, do [Decreto nº 8.539, de 8 de outubro de 2015](#).



A autenticidade deste documento pode ser conferida no site https://www.sei.ufu.br/sei/controlador_externo.php?acao=documento_conferir&id_orgao_acesso_externo=0, informando o código verificador **6115572** e o código CRC **F2FD57**.

Dedico meu trabalho a Deus primeiramente, aos meus pais, minha esposa, minha família e amigos, que estiveram ao meu lado durante toda a caminhada.

AGRADECIMENTOS

A realização deste trabalho não seria possível sem a contribuição inestimável de diversas pessoas e instituições às quais sou profundamente grato.

Primeiramente, gostaria de expressar minha mais sincera gratidão ao meu orientador, Dr. Luiz Eduardo dos Santos Paes, por sua orientação, paciência e apoio ao longo desta jornada. Sua expertise e dedicação foram essenciais para o desenvolvimento desta tese e para o meu crescimento como pesquisador. Sem a sua confiança e incentivo, este trabalho não seria uma realidade.

À Universidade Federal de Uberlândia (UFU), CAPES, CNPQ, Fapemig e ao Programa de Pós-Graduação em Engenharia Mecânica, meu agradecimento por fornecerem a infraestrutura, os recursos e o ambiente acadêmico necessários para a realização deste estudo.

Agradeço também a todos os professores que, ao longo dos anos, compartilharam seus conhecimentos e me inspiraram com sua paixão pelo ensino e pela pesquisa. Cada aula, conversa e orientação foram fundamentais para moldar o pesquisador que me tornei.

RESUMO

A Indústria 4.0 tem adotado, cada vez mais, soluções digitais como algoritmos de *Machine Learning* e *Deep Learning*, para aprimorar processos de fabricação, como a soldagem. Regular os parâmetros de soldagem é fundamental para alcançar geometrias de solda de alta qualidade, mas pode ser uma tarefa desafiadora para os operadores no campo. Uma abordagem promissora é utilizar técnicas de previsão de parâmetros a partir de geometrias específicas, sendo as redes neurais artificiais (RNA's) uma das mais destacadas. Essas redes têm um funcionamento inspirado no aprendizado observado em seres vivos. Para determinar a arquitetura ideal da RNA, incluindo camadas, neurônios, função de ativação e taxa de aprendizagem, muitos autores recorrem à seleção por tentativa e erro. O presente trabalho visa desenvolver uma metodologia para otimizar os parâmetros da arquitetura de uma RNA sem causar *overfitting*, além de promover comparação com técnicas de regressão. Para enfrentar a escassez de dados, são exploradas técnicas avançadas de geração de dados e validação cruzada, como *Leave One Out* (LOO), Kriging e *Synthetic Data Vault* (SDV). O objetivo da RNA é prever valores como: tensão, vazão do gás, velocidade de alimentação e stickout, com base em informações geométricas resultantes do processo de soldagem MAG no aço SAE 1020. Os dados são divididos em conjuntos de treinamento, validação e teste para garantir que a arquitetura da RNA seja definida na fase de validação e mantida constante durante o teste, evitando *overfitting*. No processo de otimização optou-se pela evolução diferencial (ED) comparando-se com a homologia simplicial (SHGO) por apresentar melhores resultados. Foi proposto um estudo do número de geração adicional de dados para as técnicas SDV e Kriging, aumentando-se o dataset padrão em 50%, 100% e 150%. Os melhores resultados foram encontrados ao aumentar os dados em 50% por SDV para a RNA através de 1 camada oculta, 2 neurônios e função de ativação logística. O aumento artificial do *dataset* melhorou a precisão preditiva da RNA, com o SDV apresentando 10,1% de melhora em relação à condição padrão com o dataset original e ED, e o LOO alcançando uma redução de erro de 55,3%. A função de ativação logística foi encontrada na maioria dos melhores resultados. Todos os resultados apresentados por técnicas de regressão se mostram inferiores aos apresentados pela RNA, entretanto, há também uma melhora do resultado a partir do aumento de dados por meio do SDV. O trabalho mostra que utilizar técnicas alternativas para melhorar o desempenho quando se trata de base de dados reduzidas melhorou os resultados de predição.

Palavras-chave: Inteligência Artificial, GMAW, Base de Dados Reduzida, Redes Neurais Artificiais, Aumento de Dados, Validação Cruzada.

ABSTRACT

Industry 4.0 has increasingly adopted digital solutions such as Machine Learning and Deep Learning algorithms to improve manufacturing processes, such as welding. Regulating welding parameters is essential to achieve high-quality weld geometries, but it can be a challenging task for operators in the field. A promising approach is to use parameter prediction techniques from specific geometries, with artificial neural networks (ANNs) being one of the most prominent. These networks have a functioning inspired by the learning observed in living beings. To determine the ideal ANN architecture, including layers, neurons, activation function and learning rate, many authors resort to trial-and-error selection. This work aims to develop a methodology to optimize the parameters of an ANN architecture without causing overfitting, in addition to promoting comparison with regression techniques. To address data limitation, advanced data generation and cross-validation techniques are explored, such as Leave One Out (LOO), Kriging and Synthetic Data Vault (SDV). The objective of the ANN is to predict values such as: voltage, gas flow, feed speed and stickout, based on geometric information resulting from the MAG welding process on SAE 1020 steel. The data are divided into training, validation and test sets to ensure that the ANN architecture is defined in the validation phase and kept constant during testing, avoiding overfitting. In the optimization process, differential evolution (DE) was chosen compared to simplicial homology (SHGO) because it presents better results. A study of the number of additional data generation for the SDV and Kriging techniques was proposed, increasing the standard dataset by 50%, 100% and 150%. The best results were found when increasing the data by 50% by SDV for the ANN through 1 hidden layer, 2 neurons and logistic activation function. Artificially augmenting the dataset improved the predictive accuracy of the ANN, with SDV showing a 10.1% improvement over the standard condition with the original dataset and ED, and LOO achieving an error reduction of 55.3%. The logistic activation function was found in most of the best results. All results presented by regression techniques are inferior to those presented by the ANN, however, there is also an improvement in the result from data augmentation through SDV. The work shows that using alternative techniques to improve performance when dealing with reduced datasets improved the prediction results.

Key-words: Artificial Intelligence, GMAW, Small Dataset, Artificial Neural Networks, Data Augmentation, Cross Validation.

LISTA DE FIGURAS

Figura 1. Desenho esquemático de soldagem GMAW.	6
Figura 2. Modos de transferência metálica no processo GMAW convencional: (a) curto-circuito; (b) globular e (c) goticular.	7
Figura 3. <i>Stickout</i> na soldagem GMAW.	12
Figura 4. Terminologia do cordão de solda.	13
Figura 5. RNA: a) Neurônio biológico b) <i>Perceptron</i> c) Sinapses biológicas d) Sinapses artificiais.	15
Figura 6. Arquitetura típica de um neurônio artificial.	16
Figura 7. Estrutura representativa da RNA.	17
Figura 8. Representação do algoritmo de <i>backpropagation</i>	19
Figura 9. Ótimos locais e globais.	24
Figura 10. Fluxograma de funcionamento da ED.	29
Figura 11. Punção de uma hiper esfera em um ponto minimizador em um espaço de busca compacto.	32
Figura 12. O processo de perfuração de uma nova hiper esfera.	33
Figura 13. Demonstração visual em superfícies com restrições não lineares.	34
Figura 14. Exemplos de correlação entre duas variáveis.	43
Figura 15. Representação gráfica da transformação da matriz de distância.	44
Figura 16. Exemplo de modelo balanceado e com <i>overfitting</i>	46
Figura 17. Fluxograma do processo de geração de dados SDV.	51
Figura 18. Representação de um semivariograma.	53
Figura 19. Modelos de semivariogramas.	54
Figura 20. Fluxograma de apresentação da metodologia.	57
Figura 21. Preparação de amostras. a) Gabarito de raiz de solda. b) Chanfro V. c) Medidas dos corpos de prova.	59
Figura 22. Parâmetros geométricos de entrada.	60
Figura 23. Separação de dados 20% para treinamento e validação.	62
Figura 24. Separação de dados para LOO.	63
Figura 25. Separação de dados com aumento de dados por Kriging ou SDV.	64
Figura 26. Proposta para a resolução do problema de soldagem.	65
Figura 27. Definição das variáveis da população perturbada e variáveis objetivo da otimização.	67
Figura 28. Parâmetros da ED.	68
Figura 29. Fluxograma do algoritmo proposto.	69
Figura 30. <i>Pair plot</i> das variáveis de entrada e saída.	71

Figura 31. Gráfico de correlação Pearson.....	72
Figura 32. Teste de Correlação de Distâncias para o <i>dataset</i> original.	73
Figura 33. Teste de Correlação de Distâncias para o <i>dataset</i> gerado com SDV.....	73
Figura 34. Teste de Correlação de Distâncias para o <i>dataset</i> gerado com Kriging.	74
Figura 35. Avaliação da RNA com teste de população na ED na validação.....	75
Figura 36. Gráfico do E_t e tempo computacional com variação das iterações da RNA..	76
Figura 37. Resultados da otimização SHGO na RNA..	78
Figura 38. Comparação das técnicas SHGO e ED.	80
Figura 39. Resultados LOO com 50 iterações da RNA.	81
Figura 40. Resultados LOO com 500 iterações da RNA.	82
Figura 41. Resultados LOO com 5000 iterações da RNA.	82
Figura 42. Gráfico com as médias das simulações LOO na validação.....	83
Figura 43. Gráfico com médias das simulações LOO no teste.....	83
Figura 44. Comparação ponto a ponto LOO na validação.	84
Figura 45. Comparação ponto a ponto LOO no teste.....	85
Figura 46. Comparação do E_t na validação do SDV e do Kriging..	87
Figura 47. Comparação de E_t na validação de SDV e Kriging.	88
Figura 48. Efeito do aumento de dados nas predições da RNA e condição padrão na validação..	90
Figura 49. Efeito do aumento de dados nas predições da RNA e condição padrão no teste.	90
Figura 50. Comparativo das técnicas na predição da RNA na validação..	91
Figura 51. Comparativo das técnicas na predição da RNA na validação..	92
Figura 52. Erros de previsão obtidos a partir da regressão linear (E_t) para todas as condições do <i>dataset</i>	94
Figura 53. Erros de previsão obtidos a partir da regressão polinomial (E_t) para todas as condições do <i>dataset</i>	96
Figura 54. Erros de previsão obtidos a partir da regressão polinomial (E_t) para todas as condições do <i>dataset</i>	98
Figura 55. Valores E_t previstos para todas as técnicas de regressão.....	100
Figura 56. Comparação de valores E_t previstos para todas as técnicas de regressão e os melhores resultados da RNA.....	101
Figura 57. Programa desenvolvido..	103
Figura 58. Programa desenvolvido para arquitetura da RNA.	104
Figura 59. Programa de soldagem desenvolvido com aumento de dados para previsão de parâmetros.....	104

LISTA DE TABELAS

Tabela 1. Estratégias comuns na ED.....	28
Tabela 2. Parâmetros experimentais de soldagem.	58
Tabela 3. Parâmetros de Entrada e Saída.....	60
Tabela 4. Incremento de dados.	61
Tabela 5. Divisão do <i>Dataset</i> em Treinamento e Testes da RNA	62
Tabela 6. Funções de Ativação Consideradas.....	65
Tabela 7. Parâmetros disponíveis na biblioteca <i>sklearn</i> , utilizada para construir a RNA.	66
Tabela 8. Notação da estratégia da ED.	68
Tabela 9. Parâmetros da arquitetura da RNA otimizados por ED.	77
Tabela 10. Parâmetros da arquitetura da RNA para o SHGO.....	79
Tabela 11. Parâmetros da arquitetura da RNA com LOO.	85
Tabela 12. Parâmetros de arquitetura da RNA para SDV e Kriging.....	87
Tabela 13. Valores de erro de previsão de regressão linear (E_t).....	94
Tabela 14. Coeficiente de determinação (R^2) para as entradas com a regressão linear.	95
Tabela 15. Erros de previsão de regressão polinomial (E_t) para todas as condições do <i>dataset</i>	96
Tabela 16. Coeficiente de determinação (R^2) para as entradas com a regressão polinomial.	97
Tabela 17. Erros de previsão de regressão de Ridge (E_t) em todas as condições do <i>dataset</i>	98
Tabela 18. Coeficiente de determinação (R^2) para as entradas com a regressão de Ridge.	99
Tabela 19. Faixa de resultados de previsão em dados inseridos no programa.	103

LISTA DE SÍMBOLOS

- $\hat{Y}_i$ - Valor previsto da variável dependente para a i-ésima observação
 $x_s(i)^q$ – Componente do vetor alvo
 $A_{j,k}$ – Distância de centro duplo de x
 $B_{j,k}$ – Distância de centro duplo de y
 E_t – Erro total somado
 H_0 – Hipótese nula
 U^{q+1} – Vetor experimental
 V^{q+1} – Vetor doador
 $X_{1r}, X_{2r}, \dots, X_{nr}$ - Variáveis independentes da regressão múltipla
 X_l - Variável independente da regressão linear
 X_p - Variável independente da regressão polinomial
 X_s^q – Custo do vetor alvo
 X_s^q – Vetor alvo
 X_s^{q+1} – Vetor alvo da próxima geração
 X_α^q, X_β^q e X_γ^q – Indivíduos aleatórios ED
 Y_i - Valor observado da variável dependente para a i-ésima observação
 Y_l - Variável dependente da regressão linear
 Y_p - Variável dependente da regressão polinomial
 Y_r - Variável dependente da regressão múltipla
 Z_p – Novos pontos de Kriging
 $a_{j,k}$ – Componente matriz de distâncias de x
 $b_{j,k}$ – Componente matriz de distâncias de y
 $\Sigma_{errototal}$ – Somatório do erro médio percentual
 $rand_i$ – Números com distribuição uniforme
 s^2 – Segundo momento central
 $u(i)^{q+1}$ – Componente do vetor experimental
 $v(i)^{q+1}$ – Componente do vetor doador
 w_i – Pesos sinápticos
 x_i – Entrada da RNA
 $\bar{y}, \bar{x}$ – Média amostral
 y_i – Saída da RNA
 α_3 – Terceiro momento central
 α_4 – Quarto momento central

β_0 - Coeficiente de interceptação
 $\beta_0, \beta_1, \dots, \beta_n$ - Coeficientes de regressão
 β_1 - Coeficiente de inclinação
 Cov – Covariância
 Cr – Taxa de cruzamento
 H – Complexo simplicial
 $MAPE$ – Erro médio percentual
 F – Fator de escala da ED
 S – Superfície abstrata
 Var - Variância
 W – Matriz de pesos da regularização L2
 Y – Ativação de neurônio
 Z – Elevação do ponto de controle
 e - erro aleatório
 f – Superfície de otimização SHGO
 $f(x)$ – Função de ativação
 n – Tamanho da amostra
 pop – População da ED
 r – Coeficiente de correlação de Pearson
 $\tanh(x)$ – Função tangente hiperbólica
 z – Fase conectada
 Ω – Termo de regularização L2
 α – Taxa de aprendizagem
 γ – Função de semivariograma
 θ – Função somatória RNA
 λ - Parâmetro de regularização
 ρ – Número de dimensões
 σ – Desvio padrão da média
 $\sigma(x)$ – Função sigmoide
 ω – Pesos das previsões por Kriging

SUMÁRIO

LISTA DE FIGURAS.....	vii
LISTA DE TABELAS.....	ix
LISTA DE SÍMBOLOS	x
1 – INTRODUÇÃO.....	1
1.1 – O problema da soldagem manual	1
1.2 – O Problema da Arquitetura de Redes Neurais	2
1.3 – O problema de Datasets Limitados em RNA's	3
1.4 – Objetivos.....	4
1.4.1 – Objetivos Gerais.....	4
1.4.2 - Objetivos Específicos.....	5
2 – REVISÃO BIBLIOGRÁFICA.....	5
2.1 - Soldagem GMAW	5
2.1.1 – Corrente de Soldagem	8
2.1.2 – Tensão de Soldagem	10
2.1.3 – Velocidade de Alimentação	11
2.1.4 – Vazão de Gás	11
2.1.5 - Stickout.....	12
2.1.5 – Geometria da Solda	13
2.2 – Redes Neurais Artificiais	14
2.2.1 – Conceitos Gerais.....	14
2.2.2 – Arquitetura da RNA	15
2.2.3 – Backpropagation	18
2.2.4 – Aplicações Industriais.....	20
2.2.5 – Otimização de Pesos Sinápticos	22
2.3 – Otimização	23
2.3.1 – Técnicas de Otimização	25
2.3.2 – Evolução Diferencial.....	27
2.3.3 - Homologia Simplicial.....	30
2.3.4 – Otimização de RNA's	34
2.4 – Regressão	35
2.4.1 – Regressão Linear	36
2.4.2 – Regressão Polinomial	37
2.4.3 – Regressão Múltipla de Ridge	37
2.4.4 – Coeficiente de Determinação	38
2.4.5 – Modelos de regressão aplicados	39

2.5 – Testes Estatísticos	41
2.5.1 – Correlação	41
2.6 – Base de Dados Reduzida.....	46
2.6.1 – Leave One Out.....	50
2.6.2 – Geração de Dados Sintéticos.....	50
2.6.3 - Kriging.....	52
2.7 – Síntese de soluções.....	55
3 – METODOLOGIA	56
3.1 – Base de Dados	57
3.2 – Escolha de Parâmetros	60
3.3 – Geração de Dados	61
3.4 – Separação de Amostra	61
3.4.1 – Caso Padrão de Otimização.....	61
3.4.2 – Caso Leave One Out (LOO).....	63
3.4.3 – Caso Geração de Dados	63
3.5 – Arquitetura da RNA	64
3.6 – Algoritmo de Otimização	67
3.7 – Otimização de Pesos Sinápticos	68
3.8 – Algoritmo Completo.....	69
3.9 – Padronização de Resultados	70
3.10 – Regressão	70
3.11 – Programa de Auxílio ao Soldador.....	70
4 – RESULTADOS E DISCUSSÃO	71
4.1 – Escolha dos Parâmetros de Soldagem	71
4.1.1 – Abordagem Estatística	71
4.2 – Evolução Diferencial	74
4.2.1 – Teste de População	75
4.2.2 - Cálculo da RNA - ED	76
4.3 – SHGO	78
4.3.1 – Comparação SHGO e ED	79
4.4 – Validação Cruzada LOO	81
4.5 - Geração de Dados.....	86
4.5.1 – Efeitos do Tamanho de Dados Sintéticos.....	89
4.6 – Comparação das Técnicas de RNA's.....	91
4.7 - Regressão	93
4.7.1 – Regressão Linear	93
4.7.2 – Regressão Polinomial	95

4.7.3 – Regressão Múltipla de Ridge	97
4.7.4 – Comparação das Regressões	99
4.8 – Comparação RNA e Regressão	101
4.9 – Programas Desenvolvidos	102
5 – CONCLUSÃO	105
5.1 – Trabalhos Futuros	106
REFERÊNCIAS	107
ANEXO A – Dataset Original	124
ANEXO B – <i>Dataset</i> usado no SDV 1.5.....	125
ANEXO C - <i>Dataset</i> usado no SDV 2	127
ANEXO D - <i>Dataset</i> usado no SDV 2.5	129

1 – INTRODUÇÃO

1.1 – O problema da soldagem manual

A soldagem é uma técnica amplamente utilizada no setor de fabricação para unir peças metálicas. Uma das técnicas mais comuns atualmente é a *Gas Metal Arc Welding* (GMAW), também conhecida como soldagem MIG/MAG (MIG – *Metal Inert Gas* e MAG – *Metal Active Gas*). Esse processo utiliza um eletrodo consumível, no formato de arame contínuo, que é derretido pelo calor gerado pelo arco elétrico estabelecido entre o eletrodo e a peça de trabalho, enquanto um gás de proteção é utilizado para evitar a contaminação da poça de fusão por impurezas atmosféricas (Wahab, 2014).

Apesar da automação na alimentação do arame, a soldagem GMAW exige uma habilidade significativa do operador, pois ele precisa configurar corretamente os parâmetros de soldagem, como a tensão, a corrente, a velocidade de alimentação do arame e a taxa de fluxo do gás. Além disso, o operador deve posicionar e mover a tocha de soldagem de maneira precisa para obter uma solda de alta qualidade, evitar defeitos e garantir a penetração adequada. Embora o arame seja alimentado automaticamente, a precisão e a qualidade da solda dependem fortemente da habilidade e do controle do operador, caracterizando a soldagem GMAW como um processo semiautomático (Wahab, 2014).

A soldagem GMAW se destaca em relação a outros processos devido à sua alta taxa de deposição, que permite uma maior velocidade de soldagem, a possibilidade de executar soldas em todas as posições e a ausência de escória (Scotti & Ponomarev, 2008).

A soldagem manual, por outro lado, é um método tradicional que depende inteiramente da habilidade e precisão do operador. Utilizando técnicas como a soldagem por arco elétrico manual (MMA), o operador controla todos os aspectos do processo, incluindo a manutenção do arco elétrico, a aplicação do material de solda e o posicionamento da tocha. Este processo é versátil, mas requer destreza e experiência para garantir a qualidade da solda e evitar defeitos.

Tanto nos processos manuais quanto nos semiautomáticos, é necessário ter habilidade e treinamento. O ajuste dos parâmetros no equipamento, como tensão, velocidade de alimentação e vazão de gás, influencia diretamente na qualidade da solda, afetando aspectos como falta de fusão, aparência, resistência e geometria da solda (Moore & Booth, 2015). O fator humano, associado a erros nos parâmetros e na execução, torna o papel do soldador essencial e crítico. Atenção constante é requerida para a correta execução do projeto designado.

1.2 – O Problema da Arquitetura de Redes Neurais

Redes neurais artificiais (RNA's) são modelos computacionais inspirados na estrutura e função de redes neurais biológicas. As RNA's são amplamente utilizadas em aplicações de aprendizado de máquina, como reconhecimento de imagem e fala, processamento de linguagem natural e veículos autônomos. A arquitetura de uma RNA é a estrutura da rede, incluindo o número de camadas, o número de neurônios em cada camada e a conectividade entre os neurônios (Schmidhuber, 2015). As camadas são estruturas organizacionais que agrupam um conjunto de neurônios, enquanto os neurônios são unidades de processamento que recebem entradas, as processam através de uma função de ativação, e produzem uma saída. Juntos, eles formam a arquitetura básica de uma rede neural artificial. A estrutura da rede desempenha um papel importante na determinação de seu desempenho e eficiência. Desempenho está relacionado ao quão bem uma tarefa é executada, enquanto eficiência está relacionada ao quão bem uma tarefa é executada em termos do uso de recursos, como tempo de processamento e consumo de energia (LeCun *et al.*, 2015). Ser eficiente, nesse caso, significa que a RNA pode realizar suas tarefas com o mínimo possível de recursos computacionais, mantendo ao mesmo tempo um alto nível de precisão e velocidade. Isso é crucial em aplicações onde os recursos são limitados ou onde a rapidez é essencial (Cheng *et al.*, 2018).

Nos últimos anos, houve um aumento significativo em pesquisas focadas no desenvolvimento de novas arquiteturas de RNA visando melhorar o desempenho de modelos de aprendizado de máquina. Vários tipos de RNA's foram empregados no contexto da soldagem. Em Wu *et al.* (2022), uma rede neural *feedforward* de múltiplas camadas foi usada para prever com precisão e rapidez a distorção de soldagem para uma sequência de soldagem arbitrária. Há também as redes neurais recorrentes, as quais foram utilizadas por Jaypuria *et al.* (2020), em um comparativo com redes neurais *feedforward* na modelagem de soldagem por feixe de elétrons. Redes convolucionais, como no trabalho de Cruz *et al.* (2021), abordam a otimização de hiper parâmetros de uma RNA para processamento de imagens de soldas por arco submerso e redes neurais profundas, que são parâmetros cujo valor é definido antes do processo de treinamento da rede e que não são atualizados durante o treinamento como taxa de aprendizagem, número de camadas e neurônios. Cada um desses tipos possui características arquitetônicas únicas que os tornam adequados para aplicações específicas (Schmidhuber, 2015). Outros trabalhos também podem ser citados, como por exemplo o de Chang *et al.* (2020), Lei *et al.* (2019), Mendonça *et al.* (2016), Rong *et al.* (2016), Ribeiro, (2018) e Moghaddam & Kolahan, (2021). Entretanto, com exceção do trabalho de Moghaddam & Kolahan, (2021), todos os outros citados utilizam parâmetros para a arquitetura da RNA definidos através de tentativa e erro.

Em geral, os parâmetros de uma RNA incluem o número de camadas, o número de neurônios em cada camada, as funções de ativação, e o algoritmo de otimização usado durante o treinamento (Srivastava et al., 2014). Assim como no artigo de Moghaddam & Kolahan, (2021), a presente tese investiga uma metodologia para definição dos parâmetros de uma RNA para predição de parâmetros de soldagem. Porém, também são investigadas as relações entre um maior número de parâmetros de entrada e saída, além de se definir um método para superar problemas de *dataset* pequenos encontrados no contexto da soldagem. Ainda, são contemplados parâmetros adicionais da arquitetura da RNA para serem otimizados através de um algoritmo de evolução diferencial. Isso torna o estudo amplo para que consiga abranger o máximo de variáveis e possibilitar seu uso real em aplicações industriais.

O uso de redes neurais para predição de parâmetros de soldagem oferece uma abordagem inovadora para prever variáveis complexas e interdependentes em processos de fabricação. A capacidade das redes neurais de aprender padrões e relações não-lineares nos dados torna-as especialmente adequadas para lidar com a complexidade dos processos de soldagem, onde múltiplos fatores influenciam no resultado final (Budu, 2014; R. Kumar et al., 2015).

Embora as redes neurais sejam uma boa solução, elas também podem ser computacionalmente intensivas e exigir grandes conjuntos de dados para treinamento. Nesse contexto, a alternativa de utilizar modelos de regressão oferece uma abordagem mais simples e interpretável. Os modelos de regressão, embora menos flexíveis em termos de captura de complexidade, são mais fáceis de interpretar e requerem menos recursos computacionais. Portanto, a escolha entre redes neurais e modelos de regressão depende das necessidades específicas do problema de predição, considerando *trade-offs* entre precisão, interpretabilidade e recursos computacionais disponíveis (Aslam et al., 2021; Najwa Mohd Rizal et al., 2022).

1.3 – O problema de Datasets Limitados em RNA's

O problema de conjuntos de dados limitados é comumente encontrado no contexto geral das ciências e também em processos de fabricação. Nesse contexto, a disponibilidade de dados limitados pode apresentar desafios adicionais para o desenvolvimento e aprimoramento de modelos de rede neural (Abdul Lateh et al., 2017). São diversos fatores que causam tais problemas como variações e heterogeneidades dos dados, qualidade dos dados (ruídos), restrições técnicas, restrições orçamentárias, e a transferência de aprendizado limitada.

Com conjuntos de dados reduzidos, pode haver uma falta de representatividade em termos de repetitividade. Isso significa que o conjunto de dados pode não capturar

adequadamente todas as possíveis variações, causando limitações na capacidade do modelo de rede neural de generalizar novas situações ou condições de operação (Y. Zhu et al., 2023).

Os dados coletados em processos de fabricação comumente podem conter ruídos e imperfeições devido a condições inadequadas de coleta (V. Kumar et al., 2022). Com conjuntos de dados pequenos, esses ruídos tendem a impactar significativamente no treinamento e no desempenho do modelo de rede neural, dificultando a identificação de padrões relevantes e prejudicando a generalização.

Em muitos casos, a coleta de dados em processos de fabricação pode ser cara, demorada ou inviável devido a restrições práticas. Isso leva a conjuntos de dados menores, o que limita a quantidade de informações disponíveis para treinar um modelo de rede neural de maneira adequada.

A transferência de aprendizado constitui uma abordagem utilizada para aproveitar modelos pré-treinados em conjuntos de dados grandes e aplicá-los a problemas relacionados. No entanto, com conjuntos de dados pequenos, há uma limitação na disponibilidade de modelos pré-treinados relevantes, dificultando a aplicação efetiva da técnica (Moreno-Barea et al., 2020).

1.4 – Objetivos

O presente trabalho se propõe a resolver as seguintes questões de pesquisa:

- Somos capazes de definir uma metodologia de parametrização otimizada de RNA aplicada em soldagem MAG sem utilizar técnicas de tentativa e erro?
- É possível prever parâmetros de configuração do processo de soldagem MAG através da geometria de solda desejada em RNA's evitando overfitting?
- É possível viabilizar o uso de datasets reduzidos em RNA's de predição em soldagem?
- É possível viabilizar o uso de um dataset aumentado artificialmente em problemas de predições em soldagem?
- Técnicas de regressão podem ser mais viáveis que RNA's em problemas de soldagem?

A partir das questões é então possível definir os objetivos do trabalho em sequência.

1.4.1 – Objetivos Gerais

O objetivo geral da presente tese é desenvolver uma metodologia inovadora de otimização das variáveis utilizadas em uma rede neural artificial, destinada à predição de parâmetros de soldagem do processo GMAW, promovendo a viabilidade de aplicação de redes neurais artificiais e modelos de regressão em *datasets* pequenos utilizando técnicas de aumento de dados e validação cruzada nos dados obtidos no processo de soldagem GMAW.

1.4.2 - Objetivos Específicos

Para a consecução deste objetivo geral são estabelecidos os seguintes objetivos específicos:

- Compreender a relação entre os parâmetros de entrada (geometria da peça e do cordão de solda) e os parâmetros de saída (tensão, corrente, velocidade de soldagem, stickout e vazão de gás).
- Analisar a influência dos parâmetros da rede neural, como função de ativação, número de camadas ocultas e número de neurônios, sobre os resultados.
- Desenvolver uma técnica de otimização evolutiva diferencial para a determinação de parâmetros de inicialização da rede neural artificial para a predição de parâmetros de soldagem GMAW em diversas condições.
- Desenvolver e aplicar um algoritmo de otimização global SHGO e comparar seus resultados com a técnica de evolução diferencial.
- Propor e comparar técnicas para melhorar o desempenho de predição em redes neurais, incluindo validação cruzada e aumento de dados.
- Validar os resultados obtidos com a rede neural artificial e compará-los com modelos de regressão linear, polinomial e múltipla (Ridge).
- Utilizar técnicas para aprimorar os resultados das redes neurais na predição de parâmetros de soldagem GMAW em datasets pequenos, evitando overfitting.

2 – REVISÃO BIBLIOGRÁFICA

2.1 - Soldagem GMAW

O *gas metal arc welding* (GMAW) é um processo de soldagem amplamente utilizado em diversos setores da indústria devido à alta velocidade de soldagem, baixo custo e facilidade de automação (Monika & Chauhan, 2017; Rocha et al., 2024). Nesse processo, um arco elétrico é formado entre um eletrodo consumível (arame) e a peça de trabalho, fundindo o metal e unindo as peças. A poça de fusão protegida por um gás, que geralmente consiste em uma mistura de argônio, hélio e dióxido de carbono que busca como resultado, a qualidade de solda (Pires et al., 2007).

A qualidade em solda GMAW refere-se à medida em que a solda atende aos requisitos e especificações estabelecidas para sua aplicação, incluindo resistência, durabilidade, aparência e conformidade com padrões industriais. A qualidade da solda é avaliada com base em vários critérios, como a penetração e uniformidade do cordão de solda, a ausência de

defeitos (porosidade, trincas, inclusões), propriedades mecânicas (resistência à tração, dureza), e a estética da solda (Cary & Scott, 2005).

A Figura 1 mostra o desenho esquemático do processo GMAW.

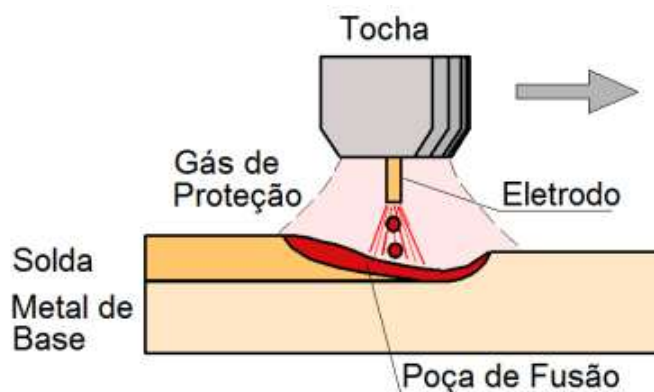


Figura 1. Desenho esquemático de soldagem GMAW. Fonte: Marques et al. (2017).

O processo GMAW foi desenvolvido pela primeira vez na década de 1940 e desde então se tornou um dos processos de soldagem mais utilizados no mundo devido à sua flexibilidade e versatilidade (Modenesi et al., 2012). Nos últimos anos, houve avanços significativos na tecnologia, incluindo melhorias no arame, gás de proteção e equipamentos de soldagem.

O gás de proteção usado na soldagem GMAW pode ser inerte (por exemplo, argônio, hélio), ativo (por exemplo, dióxido de carbono), ou uma mistura dos dois. Os gases inertes não reagem com a poça de fusão e são usados para soldar metais não ferrosos como alumínio, cobre e magnésio. Os gases ativos reagem com a poça de fusão e são usados para soldar metais ferrosos como o aço. De acordo com a *American Welding Society* (AWS), a escolha do gás de proteção depende do material a ser soldado, da posição de soldagem e das propriedades desejadas (AWS, 2004).

Os modos de transferência da GMAW referem-se à maneira como o metal de adição é transferido da tocha de soldagem para a peça. Existem três principais modos de transferência nesse processo convencional: curto-circuito, globular e goticular (spray). A Figura 2 apresenta esses modos de transferência.

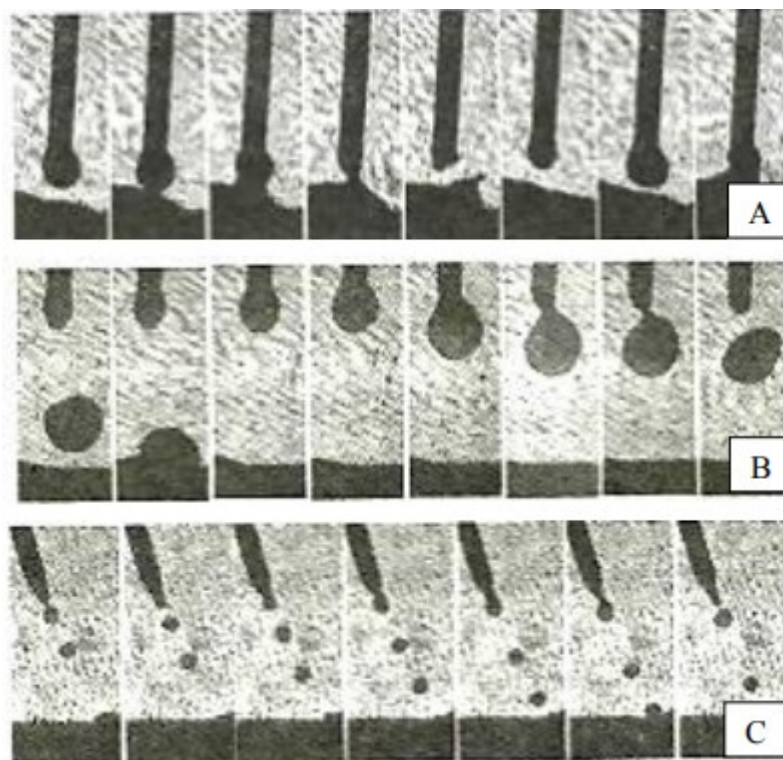


Figura 2. Modos de transferência metálica no processo GMAW convencional: (a) curto-circuito; (b) globular e (c) goticular. Fonte: (Scotti & Ponomarev, 2008).

No modo de transferência por curto-circuito, o arco elétrico é interrompido periodicamente quando o metal de adição toca o metal de base, causando um curto-circuito. Esse curto-circuito gera calor suficiente para fundir o metal de adição e criar a união. É um modo adequado para soldagem em posições verticais e sobre cabeça, com correntes relativamente baixas (Marques et al., 2017).

No modo de transferência spray, o metal de adição é pulverizado em pequenas gotas pela alta corrente e tensão aplicadas. Essas gotas são projetadas na direção da peça a ser soldada, criando uma transferência mais estável. Esse modo é usado em soldagens de alta velocidade e para espessuras maiores, permitindo uma alta taxa de deposição (Modenesi et al., 2012).

A escolha do modo de transferência depende do tipo de aplicação, espessura do material, posição de soldagem e requisitos de qualidade. Cada modo de transferência possui vantagens e limitações específicas, e o soldador deve selecionar o mais adequado para obter os melhores resultados em cada situação.

Uma das principais vantagens do GMAW é sua alta produção, que é função da taxa de deposição, e pode chegar a valores maiores que 5 kg/h (Soeiro Junior et al., 2015). No entanto, o processo também é suscetível a diversos defeitos, como mordeduras, falta de fusão, porosidade e respingos. Para minimizar esses defeitos, é importante otimizar os

parâmetros de soldagem, como corrente, tensão, velocidade de alimentação do arame e composição do gás de proteção (Ahsan et al., 2016).

Estudos recentes têm focado na melhoria do desempenho do GMAW usando tecnologias avançadas de soldagem, como GMAW pulsado, transferência de curto-circuito modificada e controle da forma de onda, que se refere à capacidade de ajustar e controlar a forma da corrente elétrica durante a soldagem. O GMAW pulsado, por exemplo, usa uma corrente de alta frequência para modular a corrente de soldagem e produzir uma série de pulsos, o que ajuda a reduzir a entrada de calor e melhorar a qualidade da solda. A transferência por curto-circuito modificada, por outro lado, usa um curto-circuito de baixa corrente que reduz respingos e melhora a estabilidade do arco (Cary, 2004).

Outro aspecto importante do GMAW é a seleção do consumível a ser depositado. O arame deve ter composição química, diâmetro adequados para garantir uma boa qualidade de solda. Existem vários tipos de arames disponíveis, como arames sólidos e arames tubulares, cada um com suas próprias vantagens e desvantagens (Nishanth & Venkatesan, 2022).

Para obter os melhores resultados durante a soldagem, é importante selecionar cuidadosamente os parâmetros, usar tecnologias avançadas, escolher o arame apropriado e se for o caso, um soldador bem treinado e com ferramentas que possibilitem mitigar erros.

Dentre as limitações desse processo, cabe destacar que não é adequado soldar materiais altamente reativos como titânio e zircônio. O processo também é sensível ao vento e correntes de ar, que podem interromper o gás de proteção e afetar a qualidade da solda. Por fim, o fato da fonte de calor e o material de adição serem dependentes, dificulta a controlabilidade da poça de fusão.

Os principais parâmetros de soldagem são aqueles que influenciam diretamente o processo de fusão e solidificação do metal, afetando a qualidade e as características da solda. Alguns desses parâmetros incluem a corrente de soldagem, a tensão de soldagem, a velocidade de alimentação do arame, o tipo e o diâmetro do eletrodo, o gás de proteção, vazão do gás, entre outros. Alguns desses parâmetros serão abordados em sequência para entendimento de ajuste adequado para cada aplicação, considerando o material a ser soldado, a espessura, a posição da soldagem e as especificações técnicas requeridas.

2.1.1 – Corrente de Soldagem

Um dos parâmetros mais críticos no GMAW é a corrente de soldagem, que determina a profundidade de penetração da solda (Hayat, 2011). Em fontes de energia com característica estática de corrente constante, a corrente é diretamente regulada. No entanto, em fontes de

tensão constante, a corrente é função da tensão e da velocidade de alimentação do arame, ou seja, é indiretamente regulada.

Um dos desafios na seleção da corrente de soldagem adequada é equilibrar a necessidade de boa penetração e alta velocidade de deslocamento com o risco de defeitos, como respingos e porosidade. A corrente de soldagem deve ser alta o suficiente para obter uma boa fusão entre os metais base, mas não tão alta que cause respingos excessivos ou distorção (J. Yu & Kim, 2018).

Vários estudos investigaram os efeitos da corrente de soldagem no processo GMAW. Alguns trabalhos como de Ibrahim et al. (2012) e Joseph et al. (2005) relatam que o aumento da corrente de soldagem resultou em maior profundidade de penetração e Zeng et al. (2017) encontram que também aumentou o risco de porosidade e respingos. Da mesma forma, Thao et al. (2008) descobriram que uma maior corrente de soldagem levou ao aumento da largura da solda, mas também aumentou o risco de fusão incompleta. Em geral, a corrente de soldagem utilizada no GMAW varia de 50 A a 500 A, dependendo da espessura da peça, do tipo de gás de proteção e do diâmetro do arame. Para obter qualidade e produtividade de soldagem é importante otimizar a corrente de soldagem com base no material a ser soldado, na espessura da peça de trabalho e nas propriedades desejadas. Por exemplo, Vijayavel et al. (2018) utilizaram um método de otimização para determinar a corrente de soldagem mais apropriada para unir chapas de aço de alta resistência. Os autores descobriram que uma corrente de soldagem mais baixa produz melhor qualidade e reduz o risco de distorção. Tecnologias avançadas de soldagem, como GMAW pulsado e o controle da forma de onda, auxiliam na redução da entrada de calor (Brumm & Bürkner, 2015).

O GMAW pulsado utiliza uma corrente de alta frequência para produzir uma série de pulsos, o que reduz a entrada de calor e melhora a estabilidade do arco (Silva et al., 2020). O controle da forma de onda, por outro lado, permite que a corrente de soldagem seja variada em tempo real para se adaptar às mudanças no ambiente de soldagem (Fan et al., 2012). A transferência por curto-circuito modificada utiliza um curto-circuito de baixa corrente, o que reduz respingos e melhora a estabilidade do arco (Singh, 2020).

Outro aspecto importante da corrente de soldagem é seu efeito nas propriedades metalúrgicas da solda. Altas correntes de soldagem podem levar a uma microestrutura mais grosseira e menor tenacidade, enquanto baixas correntes podem resultar em fusão incompleta e falta de penetração (Chacón-Fernández et al., 2022). Isso se deve à relação direta entre corrente de soldagem e taxa de resfriamento, que determina a microestrutura final. Portanto, é importante selecionar a corrente de soldagem apropriada para obter as propriedades mecânicas desejadas da solda.

Paes et al. (2022) apresentam um trabalho sobre a sensibilidade de variáveis em um processo de soldagem TIG (*Tungsten Inert Gas*) em que se avalia a influência da energia de

soldagem, corrente, e velocidade de soldagem no volume da zona termicamente afetada (ZTA). A partir de uma metodologia de superfície de resposta e um modelo numérico de volumes finitos foi possível demonstrar que a corrente afeta o volume da ZTA e a penetração com influência mais forte do que a velocidade de soldagem.

2.1.2 – Tensão de Soldagem

Um dos parâmetros mais importantes no GMAW é a tensão de soldagem, que, assim como a corrente desempenha um papel importante na determinação da entrada de calor, penetração da solda e formato do cordão. Em fontes de energia com característica estática de tensão constante, a tensão é diretamente regulada. No entanto, em fontes de corrente constante, a tensão é função da corrente e da velocidade de alimentação do arame, ou seja, é indiretamente regulada.

Vários estudos investigaram os efeitos da tensão de soldagem no processo GMAW. Por exemplo, Che Wan et al. (2011) estudaram o efeito da tensão de soldagem na microestrutura e nas propriedades mecânicas de juntas soldadas em uma liga de alumínio. Eles descobriram que o aumento da tensão de soldagem levou à maior penetração da solda, mas também resultou em uma microestrutura mais grosseira e propriedades mecânicas mais baixas, como resistência à tração e dureza. Jindal et al. (2014) investigaram a influência da tensão de soldagem no desempenho de um aço de baixa liga e alta resistência. Eles relataram que o aumento da tensão de soldagem resultou em maior taxa de deposição e penetração mais profunda, mas também levou ao aumento de respingos e cordões de solda mais largos. Zhang et al. (2019) estudaram o efeito da tensão de soldagem na penetração da solda e na geometria do cordão em uma liga de alumínio. Os autores descobriram que o aumento da tensão levou a uma penetração mais profunda da solda, maior largura do cordão e maior altura do reforço.

Sudhakar & Srinivas (2019) investigaram a influência da tensão de soldagem na geometria da solda de um aço de baixo carbono. Eles relataram que o aumento da tensão de soldagem resultou em um tamanho maior do cordão de solda e maior altura do reforço, mas também levou a um maior risco de mordedura.

Além da penetração da solda e da geometria do cordão, a tensão de soldagem também afeta a qualidade da solda. Por exemplo, Kannan et al. (2022) investigaram o efeito da tensão de soldagem na formação de porosidade no GMAW de um aço com 12 mm de espessura. Eles relataram que o aumento da tensão de soldagem levou a uma maior incidência de porosidade.

2.1.3 – Velocidade de Alimentação

A velocidade de alimentação do arame também constitui um parâmetro importante, e afeta a qualidade e eficiência do processo de soldagem. Esse parâmetro determina a quantidade de deposição de metal e a forma do cordão de solda. Velocidades de arame mais altas resultam em maiores taxas de deposição de metal e cordões de solda mais largos. Velocidades de arame mais baixas resultam em taxas de deposição mais lentas e cordões de solda mais estreitos. No entanto, velocidades de arame excessivamente altas também podem causar respingos, porosidade e redução da qualidade da solda (AWS, 2004).

A velocidade do arame também afeta a penetração e o formato da solda. Velocidades de arame mais altas, para a mesma energia de soldagem, resultam em penetração de solda mais rasa e formas de cordão de solda mais planas. Velocidades de arame mais baixas, para a mesma energia de soldagem, resultam em penetração de solda mais profunda e formas de cordão de solda mais convexas (Chavda et al., 2014).

A velocidade do arame de soldagem é influenciada por vários fatores, incluindo o diâmetro do eletrodo, a tensão de soldagem e a velocidade de deslocamento. Esses fatores são interdependentes e alterações em um parâmetro podem afetar a velocidade do arame e a qualidade da solda resultante.

2.1.4 – Vazão de Gás

A vazão ou fluxo de gás é um parâmetro na soldagem GMAW que afeta a cobertura do gás de proteção, a estabilidade do arco, a profundidade de penetração, e a geometria do cordão. O gás de proteção é usado para proteger a poça de fusão e o arame do eletrodo da contaminação atmosférica e da oxidação durante a soldagem (Weman, 2012). O fluxo de gás é normalmente controlado ajustando-se as configurações do regulador e do bocal.

Alguns trabalhos investigaram os efeitos da vazão de gás no processo GMAW. Por exemplo, Karadeniz et al. (2007) estudaram o efeito da vazão de gás na profundidade de penetração e na geometria do cordão em um aço com baixo teor de carbono. Eles descobriram que o aumento no fluxo de gás resultou em melhor cobertura de gás de proteção, além de uma ZTA mais estreita, o que levou a uma penetração mais profunda e menor largura do cordão.

Da mesma forma, Liu et al. (2023) investigaram a influência da vazão de gás na formação de respingos e na qualidade da solda em um aço inoxidável. Eles relataram que um maior fluxo de gás resultou em mais respingos, porém aumentou a qualidade da solda devido à cobertura suficiente do gás de proteção.

Para otimizar o fluxo de gás no processo GMAW, alguns métodos foram propostos, como dinâmica de fluidos computacional (CFD) e o método dos elementos finitos (MEF)

(Bitharas et al., 2018). Esses métodos podem prever e visualizar a o fluxo de gás mais apropriado para um determinado conjunto de parâmetros de entrada e requisitos de soldagem. Vários outros estudos investigaram a influência da vazão de gás na geometria da solda. Por exemplo, Chuaiphan & Srijaroenpramong (2020) estudaram o efeito da vazão de gás na geometria do cordão e na profundidade de penetração no GMAW do aço inoxidável SUS 201. Eles descobriram que o aumento do fluxo de gás levou a uma largura de cordão mais estreita, profundidade de penetração aumentada e reforço de solda reduzido. Da mesma forma, Hu et al. (2020) investigaram os efeitos da vazão de gás no formato do cordão de solda e na profundidade de penetração no GMAW através de uma simulação numérica. Eles relataram que o aumento do fluxo de gás resultou em uma poça de fusão mais profunda e com maior ZTA.

O fluxo de gás ideal depende do material, dos requisitos de soldagem e de outros parâmetros do processo, e pode ser determinado usando técnicas avançadas de modelagem.

2.1.5 - *Stickout*

O *stickout* refere-se à distância entre a extremidade do bico de contato e a ponta do arame. O *stickout* incorreto pode levar a problemas como respingos excessivos, transferência instável, dificuldades na ignição do arco e falta de fusão adequada. A Figura 3 mostra esquematicamente o *stickout*.

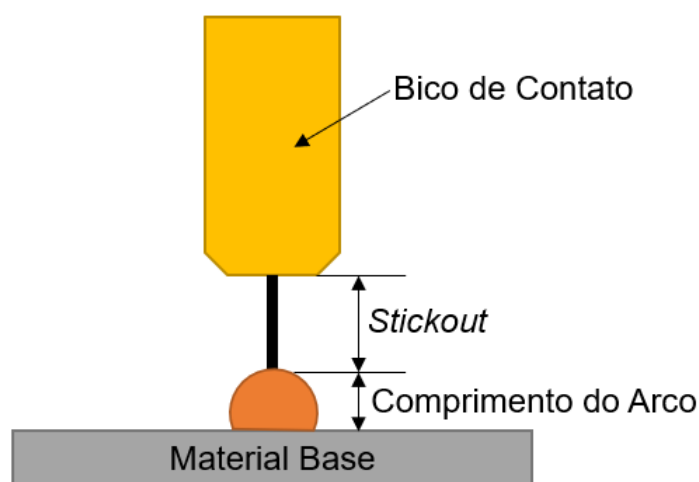


Figura 3. *Stickout* na soldagem GMAW.

Um *stickout* longo pode resultar em respingos excessivos, nos quais gotas de metal fundido são ejetadas da poça de fusão. Isso pode levar à formação de salpicos ao redor da solda e pode ser prejudicial à qualidade e à estética da junta soldada (Mandal, 2017).

A transferência de metal pode se tornar instável com um *stickout* longo e também causar interrupções no arco. Pode ocorrer uma transferência de metal em gotas maiores, levando a uma aparência irregular da solda e à formação de porosidades (Datta et al., 2006).

Em seu trabalho, Dhas et al. (2022) incluíram o *stickout* nos dados de entrada para previsão da qualidade de uma soldagem em arco submerso usando o método de Taguchi e lógica Fuzzy.

2.1.5 – Geometria da Solda

A geometria da solda (largura e a altura do cordão) constitui um fator crítico que afeta a resistência mecânica.

A largura e a altura do cordão são dimensões que refletem o tamanho e o formato do cordão de solda, resultante do metal fundido depositado ao longo da junta. Tais parâmetros podem ser influenciados por vários fatores, incluindo a corrente de soldagem, a velocidade de deslocamento, e o ângulo do eletrodo. O ângulo do eletrodo refere-se à inclinação do eletrodo em relação à superfície da peça durante o processo de soldagem.

Várias normas de soldagem fornecem orientação sobre as faixas aceitáveis para esses parâmetros. Por exemplo, a norma de soldagem estrutural AWS D1.1 da *American Welding Society* especifica aberturas de raiz mínimas e máximas e larguras de cordão para vários tipos de juntas de soldagem.

Para que melhor se compreendam os parâmetros analisados no presente trabalho, a Figura 4 exibe um desenho esquemático com a terminologia utilizada para o cordão de solda.

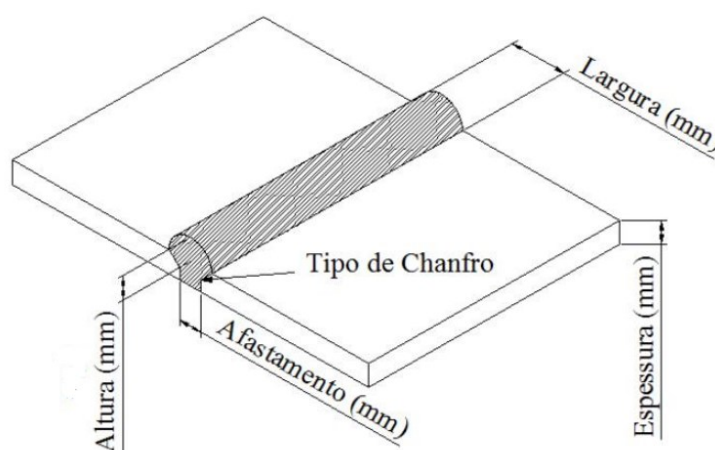


Figura 4. Terminologia do cordão de solda. Fonte: Adaptado de (Ribeiro, 2018).

2.2 – Redes Neurais Artificiais

Neste item são apresentados conceitos e estudos realizados visando a aplicação das redes neurais artificiais (RNA's).

2.2.1 – Conceitos Gerais

Para que os objetivos do trabalho sejam atingidos é necessário a utilização de alguma técnica capaz de correlacionar os parâmetros de soldagem com a geometria do cordão desejada. Existem algumas opções, e dentre essas se destacaram a regressão linear e as redes neurais artificiais (RNA's).

A regressão é uma técnica bem difundida e utilizada quando o problema possui múltiplas entradas e saída única. Poderia ser adotada caso apenas um parâmetro devesse ser encontrado, o que não é o caso deste trabalho. Além disso, a regressão tem como característica não se comportar bem com dados não lineares. Por outro lado, trata-se de um método mais simples e difundido.

As RNA's resultam da aplicação de métodos computacionais, inspirados pela forma que o sistema nervoso dos seres vivos funciona, para a resolução de problemas não-lineares. Esses sistemas são compostos por unidades de processamento, ou neurônios artificiais, que são interligados por sinapses artificiais. Vetores ou matrizes de pesos representam as sinapses (Flauzino et al., 2016).

As RNA's são um tipo de algoritmo de aprendizado de máquina que pode aprender com dados anteriores para fazer previsões ou decisões. As RNA's consistem em nós interconectados, ou neurônios, organizados em camadas, que permitem que a rede aprenda e identifique padrões nos dados, usados para fazer previsões ou decisões. Uma representação da RNA é mostrada na Figura 5.

O bloco de construção básico de uma RNA é um *perceptron* (B), que recebe os sinais de entrada, os processa e produz uma saída. As RNA's consistem em vários *perceptrons* organizados em camadas (D). A entrada é representada por x_i , a saída y_i e a função de ativação $f(x)$. A camada de entrada recebe os dados de entrada e a camada de saída produz a saída. As camadas ocultas entre as camadas de entrada e saída permitem que a rede aprenda e identifique padrões complexos nos dados.

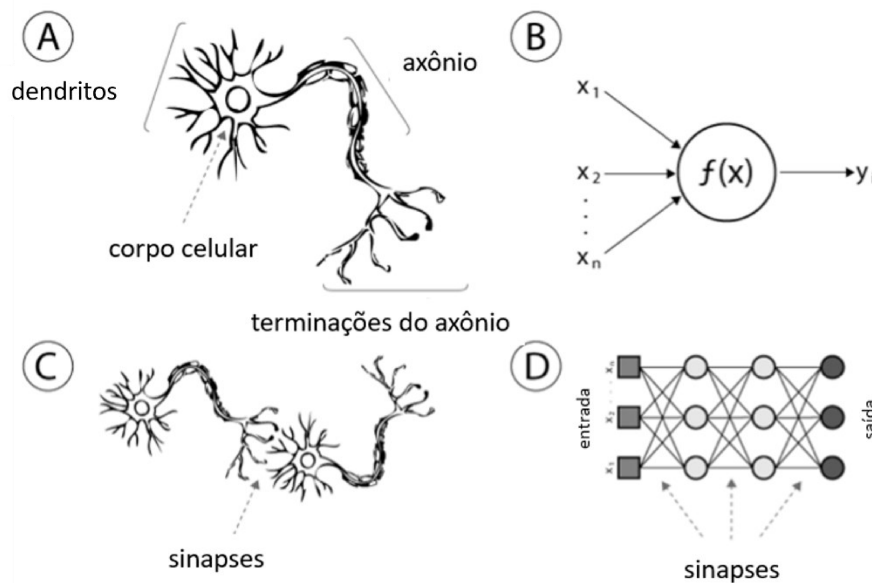


Figura 5. RNA: a) Neurônio biológico b) *Perceptron* c) Sinapses biológicas d) Sinapses artificiais. Fonte: Adaptado de datascience.eu, 2020.

Algumas das características de uma rede neural artificial (RNA) incluem (Ajit et al., 2020):

- Adaptação por experiência;
- Capacidade de aprendizado;
- Habilidade de generalização;
- Organização de dados;
- Tolerância a falhas;
- Armazenamento distribuído;
- Múltiplas entradas e saídas.

2.2.2 – Arquitetura da RNA

O neurônio ou *perceptron* é uma parte essencial para o funcionamento da RNA. O funcionamento do neurônio se dá pelo processamento dos dados de entrada que são associados com pesos sinápticos. Após a multiplicação pelos pesos é feito um somatório dos sinais de entrada e a partir disso há a ativação ou não ativação do neurônio através de uma função de ativação (Oliveira & Rocha, 2021). A ativação de um neurônio é o processo de transformar as informações de entrada em um valor de saída, que é passado para os neurônios na camada seguinte da rede. A estrutura do neurônio é apresentada na Figura 6.

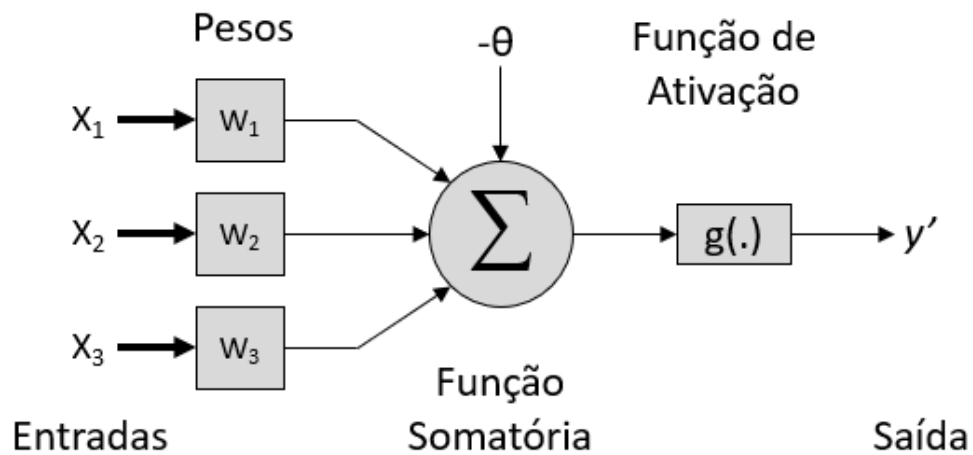


Figura 6. Arquitetura típica de um neurônio artificial.

A entrada é representada por x_i , a saída y_i , a função de ativação $f(x)$ os pesos sinápticos w_i e a função somatória θ . A função somatória em uma rede neural artificial é responsável por calcular a soma ponderada das entradas de um neurônio antes da aplicação da função de ativação. É o primeiro passo na ativação de um neurônio.

Durante construção da arquitetura básica de uma RNA é necessário determinar a quantidade de neurônios, camadas ocultas e função de ativação. A Figura 7 exibe duas camadas ocultas, com cinco neurônios em cada, além de possuir cinco entradas e quatro saídas, como no problema do presente trabalho.

Na fase de entrada, a rede neural recebe dados de entrada, que são geralmente representados por um vetor de valores numéricos. Esses dados podem ser imagens, áudio, texto, entre outros tipos. Cada valor do vetor de entrada é conectado a um neurônio de entrada na rede neural nas arquiteturas mais utilizadas.

Na fase de processamento, os neurônios da rede neural são ativados ou desativados de acordo com a força e direção das conexões sinápticas entre eles. Cada neurônio recebe sinais de entrada dos neurônios conectados a ele, multiplica esses sinais pelos pesos sinápticos associados a cada conexão e os soma para gerar um sinal de saída. O sinal de saída é então transmitido aos neurônios conectados a ele na próxima camada da rede neural (Goodfellow et al., 2016).

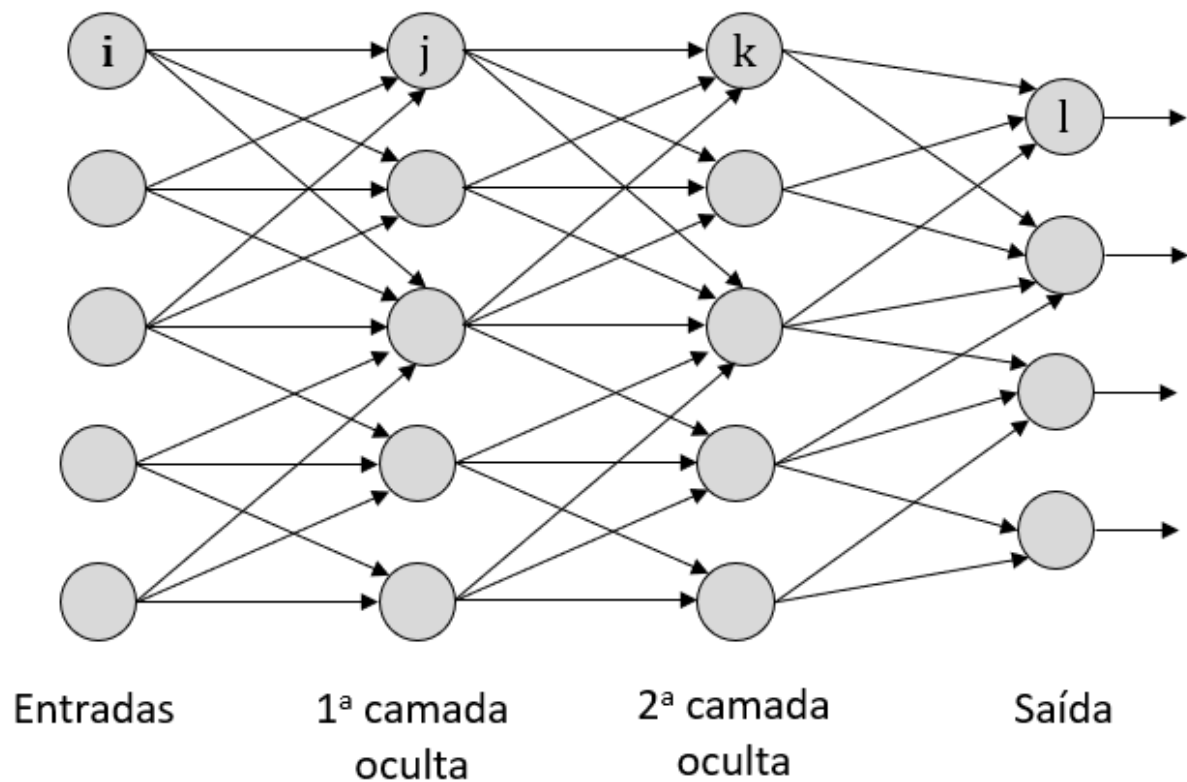


Figura 7. Estrutura representativa da RNA.

O processo de ativação e propagação do sinal é repetido em todas as camadas da rede neural até que se chegue à camada de saída, onde a rede neural gera uma resposta ou saída para a tarefa que está sendo realizada (Hecht-Nielsen, 1992). Por exemplo, em uma rede neural que realiza a classificação de imagens de animais, a saída poderia ser um vetor de probabilidades que indica a probabilidade de a imagem pertencer a cada uma das classes possíveis (cachorro, gato, pássaro, etc.).

Na fase de treinamento, a resposta da rede neural é comparada com a resposta desejada. Um algoritmo de aprendizado é usado para ajustar os pesos sinápticos de forma a minimizar o erro entre a resposta da rede neural e a resposta desejada. Esse processo de ajuste dos pesos sinápticos é repetido durante o treinamento da rede neural até que essa possa produzir saídas precisas e confiáveis para novos dados de entrada.

As funções de ativação são responsáveis por verificar se a informação de que um neurônio está recebendo é relevante e deve ser repassada, ou se a informação deve ser ignorada. A seguir se exhibe matematicamente o processamento feito por essa função.

$$Y = \text{Ativação}(\sum(\text{pesos} \times \text{entradas}) \times \text{bias}) \quad (1)$$

Os pesos sinápticos em uma rede neural artificial RNA são valores numéricos que determinam a força e direção das conexões entre os neurônios. São muito importantes, pois permitem que se aprenda a partir dos dados e gere saídas satisfatórias para cada aplicação. Já o bias refere-se a um parâmetro adicional em uma função de ativação que é usado para ajustar a saída do neurônio, independentemente das entradas recebidas. O bias permite que a função de ativação seja deslocada para cima ou para baixo, o que pode ser crucial para o aprendizado da rede.

Cada conexão entre os neurônios em uma RNA tem um peso sináptico associado a ela. O valor do peso sináptico indica a importância relativa daquela conexão para a tarefa que a RNA está realizando. Por exemplo, em uma RNA que está aprendendo a classificar imagens de animais, os pesos sinápticos associados às conexões que se relacionam com as características dos animais (como o formato das orelhas, tamanho do corpo, cor da pele, etc.) serão ajustados para que a RNA possa reconhecer essas características e fazer a classificação correta (Fleck et al., 2016).

Os pesos sinápticos são geralmente inicializados aleatoriamente e ajustados durante o treinamento da RNA através de algoritmos de aprendizado. Durante o treinamento, a RNA é apresentada a exemplos de entrada (por exemplo, imagens de animais) e sua saída é comparada com a saída desejada (por exemplo, o rótulo que indica o tipo de animal).

A arquitetura da rede neural artificial define como os neurônios estão organizados, enquanto o algoritmo de *backpropagation* permite que a rede seja treinada ajustando os pesos sinápticos, sendo abordado no item a seguir. Ambos trabalham juntos para possibilitar o aprendizado e a capacidade de generalização da rede neural, permitindo que ela seja aplicada em uma variedade de tarefas de processamento de informações

2.2.3 – *Backpropagation*

Backpropagation é um algoritmo de treinamento utilizado em RNA's que permite ajustar os pesos sinápticos da rede de forma a minimizar o erro entre a saída produzida pela RNA e a saída desejada (Werbos, 1990). O algoritmo foi proposto por Werbos em 1974 e é considerado uma das principais técnicas de treinamento.

O processo de treinamento por *backpropagation* envolve a realimentação do erro da saída para as camadas anteriores da rede, ajustando os pesos sinápticos de forma a minimizar esse erro. O algoritmo é realizado em várias etapas (Wythoff, 1993):

- Propagação para frente ou *feedforward*: A rede neural recebe uma entrada e propaga o sinal através das camadas até calcular a saída. Esse processo é realizado multiplicando os pesos sinápticos pelas entradas e somando os resultados, aplicando uma função de ativação em cada neurônio.

- **Cálculo do erro:** É calculado o erro da saída da rede neural em relação à saída exemplo. Esse erro pode ser medido usando diferentes métricas, como a média do erro quadrático ou a entropia cruzada.
- **Realimentação do erro:** O erro é propagado para trás pela rede, a partir da camada de saída até a camada de entrada, ajustando os pesos sinápticos de acordo com a contribuição de cada conexão para o erro, assim como apresentado na Figura 8. Para isso, o gradiente da função de perda em relação aos pesos sinápticos é calculado por meio da regra da cadeia, permitindo a atualização dos pesos em direção ao gradiente negativo. A função de perda, também conhecida como função de custo ou função de erro, é uma medida utilizada em redes neurais artificiais para quantificar o quão bem os resultados da rede correspondem aos valores desejados. Ela mede a discrepância entre as saídas previstas pela rede neural e os valores alvo esperados (N. Zhang et al., 2019).
- **Atualização dos pesos:** Os pesos sinápticos são atualizados em direção ao gradiente negativo, multiplicando-os por uma taxa de aprendizado que controla a velocidade de aprendizado da RNA.
- **Repetição:** O processo é repetido para um conjunto de exemplos de treinamento até que o erro da rede seja minimizado.

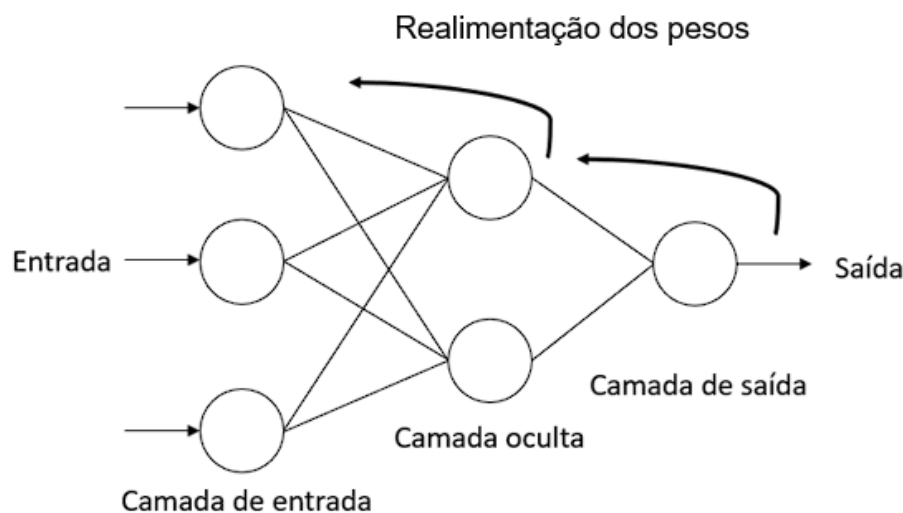


Figura 8. Representação do algoritmo de *backpropagation*.

Esse processo é chamado de *backpropagation* porque o erro é propagado da camada de saída para a camada de entrada, permitindo que os pesos da RNA sejam ajustados de forma a minimizar o erro. O algoritmo é eficiente para RNA's com múltiplas camadas e pode ser utilizado em diferentes tipos de tarefas de aprendizado supervisionado, como classificação, regressão, e sequência para sequência (Goh, 1995).

Em redes de *backpropagation* surge o problema do gradiente desvanecente (*vanishing gradient*), que pode ocorrer durante o treinamento de RNA's profundas, especialmente em arquiteturas como as recorrentes.

Quando uma RNA é treinada usando algoritmos de realimentação do gradiente, o objetivo é ajustar os pesos das conexões para minimizar a função de perda. A realimentação envolve a propagação do erro calculado na camada de saída para as camadas anteriores, ajustando os pesos ao longo do caminho para reduzir o erro gradualmente. No entanto, o problema do gradiente desvanecente ocorre quando os gradientes que são propagados de volta para as camadas anteriores se tornam muito pequenos à medida que avançam na rede (Ding et al., 2018; Rasamoelina et al., 2020).

No trabalho de Ding et al. (2018) é feito um estudo sobre o efeito e características das funções de ativação em RNA's. Os resultados para a função sigmoide – logística apresenta problemas de gradiente desvanecente em redes com mais de 5 camadas ocultas, ou seja, em problemas de *deep learning*. Para se evitar esse problema, é possível também usar funções de ativação não saturadas, como a ReLu.

No mesmo trabalho entende-se que a função de ativação tangente hiperbólica apresenta estrutura similar à função “logística”, entretanto apresenta vantagens como convergência mais rápida além de ser simétrica a partir da origem.

2.2.4 – Aplicações Industriais

As RNA's têm sido aplicadas em vários campos, incluindo reconhecimento de imagem, processamento de linguagem natural e previsão financeira. São usadas também em aplicações industriais, como otimização e controle de processos. As RNA's são adotadas na fabricação para prever a qualidade do produto, otimizar os parâmetros de produção e monitorar as variáveis do processo (Saddeeq & Abdulazeez, 2020).

Uma das principais vantagens das RNA's é a capacidade de aprender com dados anteriores e se adaptar a novas situações. Isso significa que a rede pode melhorar com o tempo, à medida que mais dados são alimentados no sistema. As RNA's também podem processar grandes quantidades de dados rapidamente, tornando-as adequadas para o controle em tempo real de processos industriais (Goodfellow et al., 2016).

Existem diferentes tipos de RNA's, incluindo redes *feedforward*, redes recorrentes e redes convolucionais. As redes *feedforward* são o tipo mais simples de RNA, onde os dados fluem da camada de entrada para a camada de saída. As redes recorrentes têm conexões de feedback entre os neurônios, o que lhes permite processar dados sequenciais. As redes convolucionais são projetadas para processar dados com uma topologia semelhante a uma grade, como imagens.

RNA's são uma ferramenta poderosa para análise de dados e tomada de decisão com capacidade de aprender e se adaptar a novas situações, o que as torna adequadas para várias aplicações, incluindo processos industriais como a soldagem, visando a previsão da qualidade da solda e controle dos parâmetros.

De forma mais específica, as RNA's têm sido usadas na soldagem para prever a penetração, o formato do cordão e defeitos como porosidade e trincas. As RNA's também são usadas para otimizar parâmetros de soldagem, como corrente, velocidade, e taxa de alimentação do eletrodo. Além disso, as RNA's são ferramentas para controlar os processos, a partir da regulação da tensão de soldagem e monitoramento da poça de fusão (Sathiya et al., 2012).

O desempenho das RNA's em aplicações de soldagem depende da qualidade e quantidade dos dados usados para treinar a rede. A qualidade dos dados é crítica, pois dados imprecisos ou incompletos podem levar a um desempenho ruim. A quantidade de dados também é importante, pois as RNA's requerem uma quantidade suficiente para generalizar e fazer previsões adequadas (Balasubramanian et al., 2010).

As RNA's mostraram resultados promissores em aplicações de soldagem, com alguns estudos relatando altas taxas de precisão na previsão da qualidade da solda e na otimização dos parâmetros (Mezaache et al., 2022). No entanto, o uso de RNA's na soldagem ainda é relativamente novo e mais pesquisas são necessárias para entender completamente seu potencial e limitações nesse campo.

No estudo de Giarollo et al. (2022) foi aplicado um design experimental Box-Behnken para analisar a relação entre variáveis do processo de soldagem e a geometria do cordão de solda em GMAW, utilizando RNA's com treinamento Bayesiano com *backpropagation* e modelos de regressão de segundo grau. A rede neural com quatro camadas e duas camadas ocultas mostrou um desempenho superior na predição da geometria do cordão, com um erro médio de 10.843, em comparação ao modelo de regressão, que apresentou um erro médio de 13.213. Apesar da superioridade da rede neural, o modelo de regressão foi mais simples de implementar e também eficaz.

A implementação de RNA's na soldagem também apresenta alguns desafios. Um desses é a necessidade de uma grande quantidade de dados para treinar a rede com precisão, o que nem sempre está disponível. Outro desafio é a complexidade dos processos de soldagem, que podem exigir várias RNA's para modelar diferentes aspectos. Além disso, as RNA's exigem uma seleção cuidadosa dos parâmetros de entrada, pois o uso de muitas ou poucas entradas pode afetar o resultado. A seleção de funções de ativação apropriadas e arquiteturas de rede também é crítica para alcançar o desempenho ideal (Balasubramanian et al., 2010).

2.2.5 – Otimização de Pesos Sinápticos

Os pesos sinápticos da RNA são importantes para se atingir bons resultados. Os valores de entrada da rede são multiplicados por tais pesos e podem definir a qualidade e o tempo de convergência. Em função disso, na própria biblioteca utilizada para implementação da RNA (*sklearn.neural_network*) existem métodos de otimização dos pesos sinápticos: BFGS (*Broyden-Fletcher-Goldfarb-Shanno*), LBFGS (*Limited-Broyden-Fletcher-Goldfarb-Shanno*), SGD (*Stochastic Gradient Descent*) e Adam (*Extended Stochastic Gradient Descent*) (K. He et al., 2015).

O Adam é um método feito a partir de uma extensão do SGD com bons resultados comparados aos do SGD, entretanto funciona melhor com uma larga base de dados, maior que 100000 dados na maioria dos casos (Khan et al., 2020).

O LBFGS é utilizado para resolver problemas de minimização com memória limitada. O algoritmo usa apenas uma estimativa da matriz Hessiana inversa, enquanto o BFGS, que é a versão não limitada, usa a matriz completa, o que gera um maior custo computacional em função do aumento da memória alocada (D. C. Liu & Nocedal, 1989). O algoritmo começa com uma estimativa inicial do valor ótimo, x_0 , e prossegue de forma iterativa para refinar essa estimativa com uma sequência de melhores estimativas $x_1, x_2, \dots$. As derivadas da função $g_k := \nabla f(x_k)$ são usadas como chave do algoritmo para identificar a direção da descida mais íngreme e também para formar uma estimativa da matriz Hessiana (segunda derivada) de $f(x)$ (Pytlak & Tarnawski, 2007). O LBFGS tem a desvantagem de custos computacionais adicionais na realização de uma atualização de nível dois para a aproximação Hessiana (inversa) em cada etapa. Embora isso seja razoavelmente rápido, começa a aumentar à medida que o espaço de busca aumenta, ou seja, em grandes *datasets* (Bottou et al., 2018).

Outro parâmetro de controle e otimização pela ED é a taxa de regularização ou taxa de aprendizagem. Esse parâmetro atua no controle do *overfitting*, que é um fenômeno que ocorre quando um modelo estatístico se ajusta muito bem ao conjunto de treinamento, mas se mostra ineficaz ao prever novos resultados.

A regularização em uma rede neural artificial é uma técnica utilizada para evitar o *overfitting*, ou seja, o ajuste excessivo do modelo aos dados de treinamento, resultando em uma baixa capacidade de generalização para novos dados. A regularização ajuda a controlar a complexidade da rede neural, evitando que ela se torne muito sensível aos dados de treinamento e reduzindo a chance de ocorrerem problemas como o excesso de ajuste (Shah et al., 2016).

Existem duas técnicas comuns de regularização em redes neurais artificiais:

- Regularização L1 e L2 (também conhecida como penalização de peso): Essas técnicas adicionam uma penalização aos termos de perda da função de custo original,

com o objetivo de reduzir os valores dos pesos sinápticos. A regularização L1 adiciona uma penalidade proporcional à soma dos valores absolutos dos pesos, enquanto a regularização L2 adiciona uma penalidade proporcional à soma dos quadrados dos pesos. Isso faz com que a rede dê preferência a pesos menores, o que ajuda a evitar a complexidade excessiva.

- *Dropout*: é uma técnica de regularização que envolve a exclusão aleatória de unidades (neurônios) durante o treinamento da rede neural. Durante o treinamento, cada unidade tem uma probabilidade definida de ser temporariamente removida. Isso ajuda a evitar que as unidades se tornem muito dependentes de outras unidades específicas, forçando a rede a aprender representações mais robustas e generalizáveis (Srivastava et al., 2014).

Durante a regularização L2, escolhida devido à sua disponibilidade na biblioteca utilizada, a função de perda da rede neural é estendida por um termo de regularização, Ω , dado por:

$$\Omega(W) = ||W||_2^2 = \sum_i \sum_j w_{ij}^2 \quad (2)$$

O termo de regularização Ω é definido como a Norma Euclidiana (ou norma L2) das matrizes de peso, ou seja, a soma de todos os valores de peso ao quadrado de uma matriz de peso. O termo de regularização é ponderado pelo α dividido por dois e adicionado à função de perda regular, que é escolhida para a tarefa atual. Em outras palavras, α rege o quanto se regulariza o modelo e consequentemente diminui-se os pesos a cada vez que o algoritmo é atualizado.

2.3 – Otimização

Otimização em engenharia é o processo de encontrar a melhor solução para um problema, maximizando ou minimizando uma função objetivo específica, enquanto satisfaz um conjunto de restrições. Essa função objetivo geralmente é uma medida do desempenho do sistema que está sendo otimizado, como minimizar custos, maximizar a eficiência ou maximizar o rendimento (Snyman & Wilke, 2018).

O processo de otimização envolve a formulação de um modelo matemático que descreve o sistema que está sendo otimizado, definindo a função objetivo e as restrições e, em seguida, usando técnicas matemáticas para encontrar a solução ideal.

Existem diferentes tipos de técnicas de otimização usadas na engenharia, incluindo programação linear, programação não linear, programação dinâmica e otimização

estocástica. Cada uma dessas técnicas tem seus próprios pontos fortes e fracos e é adequada para diferentes tipos de problemas.

A otimização é amplamente utilizada na engenharia para uma variedade de aplicações, incluindo design de produtos, processos, gerenciamento da cadeia de suprimentos e sistemas de controle. Por exemplo, no projeto do produto, a otimização pode ser usada para determinar as dimensões e materiais ideais para um componente ou reduzir seu peso, enquanto na otimização do processo, pode ser usada para determinar as condições operacionais ideais para maximizar a eficiência e reduzir o desperdício (Rocha, 2024; Venter, 2010).

No geral, a otimização é uma ferramenta importante na engenharia que ajuda a melhorar o desempenho de sistemas e processos, reduzir custos e aumentar a eficiência procurando cada vez mais chegar no ótimo global ou em algum ótimo local relevante.

Na otimização, um ótimo local é uma solução que parece ser a melhor solução em uma determinada região do espaço de busca, mas não é necessariamente a solução ótima global. Em outras palavras, um ótimo local é uma solução que não pode ser melhorada por pequenas perturbações nos valores de seus parâmetros (de Jonge & Scarf, 2020).

Ao otimizar uma função, podem haver vários ótimos locais, e encontrar o ótimo global pode ser um desafio. Isso é particularmente verdadeiro em problemas de otimização não convexos, em que a função objetivo tem vários picos e vales, como mostrado na Figura 9.

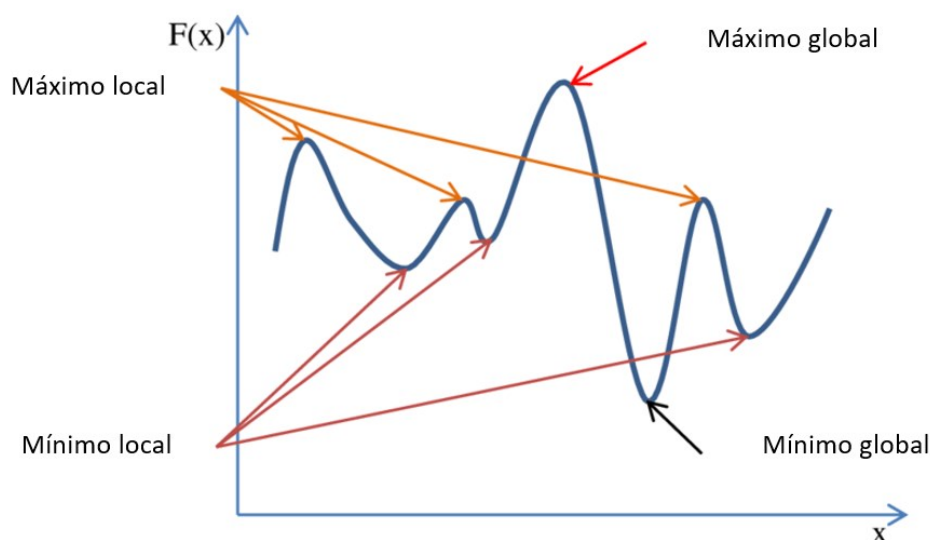


Figura 9. Ótimos locais e globais. Fonte: Adaptado de (Alves Ribeiro, 2016).

Encontrar um ótimo local geralmente é o primeiro passo em um problema de otimização. Com isso, é possível fornecer informações valiosas sobre o comportamento da função objetivo e ajudar a orientar esforços de otimização adicionais. No entanto, é importante reconhecer

que um ótimo local não é necessariamente a melhor solução, e que uma otimização adicional pode ser necessária para encontrar o ótimo global (Ribeiro, 2016).

Existem várias estratégias para lidar com ótimos locais na otimização:

- Usando um algoritmo de otimização diferente: alguns algoritmos de otimização são mais adequados para encontrar ótimos globais do que outros. Alterar o algoritmo usado pode ajudar a superar os ótimos locais.
- Partindo de múltiplos pontos iniciais: Ao iniciar o algoritmo de otimização a partir de múltiplos pontos iniciais, é possível explorar diferentes regiões do espaço de busca e aumentar a chance de encontrar o ótimo global (Peng et al., 2016).
- Incorporando aleatoriedade: métodos de otimização estocástica, que incorporam aleatoriedade no processo de otimização podem ajudar a explorar diferentes regiões do espaço de busca e evitar ficar preso em ótimos locais (Maroufpoor et al., 2020).

Em resumo, um ótimo local é uma solução que parece ser a melhor solução em uma determinada região do espaço de busca, mas não é necessariamente a solução ótima global. Reconhecer e lidar com ótimos locais é um aspecto importante da otimização.

2.3.1 – Técnicas de Otimização

A otimização desempenha um papel crítico nos problemas de soldagem, pois permite que os engenheiros identifiquem o melhor conjunto de parâmetros para uma determinada aplicação. A soldagem é um processo complexo que envolve a interação de múltiplas variáveis, como velocidade de soldagem, temperatura, entrada de calor, habilidade do operador quando manual, além de condição do material e dos eletrodos, o que torna a otimização um desafio. Nos próximos parágrafos são apresentados diversos métodos de otimização usados nos problemas de soldagem, mostrando vantagens e desvantagens de cada um.

A *Particle Swarm Optimization* (PSO) é uma técnica criada em 1995 e possui como vantagens a capacidade de resolver problemas complexos em diferentes tipos de aplicações, com rápida convergência. Entretanto, também apresenta problema para sair de mínimos locais. A regulação dos parâmetros iniciais define criticamente a qualidade dos resultados (Venter & Sobieszcanski-Sobieski, 2003).

No estudo realizado por Wang et al. (2015), uma abordagem discreta de Otimização por Enxame de Partículas (PSO) foi empregada para o planejamento do trajeto de um robô de soldagem. Inicialmente, os autores destacaram a importância de uma função de aptidão multiobjetivo, considerando o comprimento do trajeto de soldagem, a deformação da solda e o consumo de energia, esse último composto por dois fatores: a operação de soldagem, dependente dos parâmetros do processo, e o consumo energético do robô. Entretanto, no

estudo, apenas uma única função objetivo relacionada ao trajeto mais curto foi implementada. Os operadores de crossover, mutação e partição foram utilizados. Os pesquisadores compararam três abordagens distintas: PSO básico, PSO com partição (P-PSO) e PSO com mutação e partição (PM-PSO). O estudo foi conduzido considerando uma porta de carro como caso de estudo, com 115 juntas soldadas. Os resultados obtidos demonstraram que a abordagem híbrida PM-PSO apresentou o melhor desempenho.

Os algoritmos genéticos (AG) ou *genetic algorithms* (GA) foram desenvolvidos na década de 1960 e têm como característica se comportar bem com um grande número de variáveis, além de não necessitar de informações das derivadas. Porém, apresentam um longo tempo de convergência, dificuldade de ajustar parâmetros iniciais e não há garantias de se encontrar os mínimos globais nos problemas (Konak et al., 2006). O algoritmo parte de uma população inicial e essas são submetidas a um processo evolutivo de forma que os indivíduos mais adaptados continuarão se reproduzindo entre eles, e ao longo de várias gerações o mais adaptado se sobressai (Romero-Hdz et al., 2016).

Vasudevan et al. (2010) discutem o desenvolvimento de um modelo computacional usando AG para determinar os parâmetros ótimos/quase ótimos do processo GTAW para obter o perfil desejado do cordão de solda durante a soldagem automática do aço inoxidável 316LN. Utilizando os dados experimentais gerados sobre a influência das variáveis do processo na geometria do cordão de solda, foram desenvolvidos modelos de regressão que correlacionam os parâmetros do formato do cordão de solda com os parâmetros do processo, para determinar a função objetivo no algoritmo genético. Obteve-se convergência entre o perfil desejado do cordão de solda e o perfil do cordão de solda calculado pelo modelo.

Em Choudhary et al. (2020), tem-se a combinação de otimização por PSO e AG, para otimizar os parâmetros do processo a arco submerso (SAW). O foco foi dado em diferentes parâmetros diretos, como tensão, taxa de alimentação do arame, velocidade de soldagem e distância entre a bocal e a placa, e parâmetros indiretos, como condição do fluxo e espessura da placa, respectivamente. No final, o desempenho do PSO-AG híbrido foi comparado com os algoritmos PSO e GA tradicionais sugerindo que o PSO-AG híbrido é uma opção melhor do que os outros dois algoritmos, ou seja, PSO e AG, para obter os parâmetros ideais do processo SAW.

No trabalho de Yifei et al. (2018), um AG e um PSO são aplicados à otimização do caminho do robô de soldagem e a eficácia dos dois algoritmos é verificada por meio de simulação no *MATLAB*. Os resultados mostram que ambos os algoritmos possuem uma boa capacidade de convergência e otimização para o caminho de soldagem. As partículas do algoritmo PSO apresentam um recurso de “memória”, e a solução ótima ou solução quase ótima pode ser obtida em menos iterações por “auto” aprendizado e aprendizado de “grupos”. Toda a população do AG é relativamente homogênea mais próxima da solução ótima com

compartilhamento de informações. Porém, no algoritmo PSO, apenas a solução ótima global e a solução ótima individual passam informações para outras partículas em um fluxo unidirecional.

O *Gravitational Search Algorithm* (GSA) foi proposto em 2009 com convergência rápida, mas fica preso em mínimos locais com facilidade (Shareef et al., 2015). O GSA é considerado ainda desafiador porque, muitas vezes, as relações não lineares entre entradas e saídas não são explicitamente conhecidas (Song et al., 2016).

No trabalho de Zhu et al. (2023) é apresentado um método de layout de otimização de dois estágios para agarrar pontos de peças de chapa metálica durante as operações de montagem do robô. O método é baseado no modelo substituto GSA-Kriging, que ajuda a reduzir a deformação e melhorar a precisão da localização. Há também um estudo de caso de uma peça que faz parte da porta de um carro e compara os resultados com outros métodos de otimização, como função de base radial e algoritmo de otimização de enxame de partículas. O artigo conclui que o método proposto é eficaz na redução da deformação e na melhoria da precisão da montagem de chapas metálicas.

O *Firefly Algorithm* (FA) foi criado em 2008 com uma estrutura de simples implementação e com características de subdivisões automáticas. Como desvantagens apresenta resultados de ótimos locais e sem guardar em memória da melhor situação da simulação (Abd Ali et al., 2015).

Para melhorar a qualidade da soldagem TIG em uma placa de alumínio, Kumar et al. (2017) propuseram um modelo pelo qual a corrente de soldagem, a vazão do gás de proteção e a polaridade da corrente podem ser controladas durante o processo de soldagem através de uma otimização FA. Estudou-se o efeito da corrente de soldagem, da polaridade da corrente e da vazão do gás de proteção na resistência à tração da junta soldada. Com base no número de parâmetros e seus níveis, a técnica de metodologia de superfície de resposta foi selecionada como o projeto de experimento.

2.3.2 – Evolução Diferencial

No presente trabalho o enfoque foi direcionado para a Evolução Diferencial, ED (no inglês: *Differential Evolution*, DE) que é uma técnica de otimização inicialmente apresentada em 1995 por Storn e Price, com característica de fácil implementação, rápida convergência, e poucos parâmetros de ajuste para inicialização (Price et al., 2006)

A ED se inspira nos mecanismos de seleção natural e na genética de populações de seres vivos, com operadores de mutação, cruzamento e seleção para gerar novos indivíduos de uma população em busca do mais adaptado. Essa técnica tem como característica a pequena quantidade de parâmetros utilizados (Itaborahy Filho, 2018).

A ED é uma técnica de otimização populacional que utiliza parâmetros específicos para guiar sua busca por soluções ótimas. Os principais parâmetros são o tamanho da População que refere-se ao número de indivíduos presentes em cada geração. Um tamanho de população maior pode proporcionar uma exploração mais ampla do espaço de busca, mas aumenta o custo computacional. O número de gerações define o critério de parada do algoritmo, indicando quantas iterações serão realizadas antes de encerrar o processo de otimização. Já o fator de perturbação (F): Também conhecido como fator de mutação, é um valor que controla a amplificação das diferenças vetoriais durante a mutação. Geralmente, F é um valor real no intervalo $[0, 2]$. A probabilidade de cruzamento (Cr) determina a probabilidade de os componentes do vetor mutante serem incorporados no vetor de teste durante o cruzamento. Valores típicos de CR estão entre 0,5 e 1 (Das et al., 2009; Lobato & Steffen Jr, 2016).

Por fim, a estratégia refere-se à abordagem específica utilizada para gerar novos candidatos a solução. As estratégias são geralmente denotadas por notações como "DE/rand/1/bin", onde cada parte indica a escolha do vetor base, o número de vetores de diferença utilizados e o tipo de crossover aplicado. A Tabela 1 apresenta algumas estratégias tipicamente utilizadas na ED.

Tabela 1. Estratégias comuns na ED.

Estratégia	Descrição
DE/rand/1/bin	Escolhe aleatoriamente um vetor base e adiciona a diferença entre dois vetores aleatórios multiplicada por um fator de mutação. Usa crossover binomial.
DE/best/1/bin	Usa o melhor vetor da população como base, aumentando a convergência, mas podendo levar a mínimos locais.
DE/rand/2/bin	Utiliza dois pares de vetores diferenciais, aumentando a diversidade na mutação.
DE/best/2/bin	Usa o melhor vetor como base e adiciona dois pares de vetores diferenciais, equilibrando exploração e convergência.
DE/rand-to-best/1/bin	Combina um vetor aleatório e o melhor vetor da população para direcionar a busca.
DE/rand/1/exp	Igual a DE/rand/1/bin, mas usa crossover exponencial, favorecendo mudanças concentradas em partes do vetor.
DE/best/1/exp	Semelhante à DE/best/1/bin, mas com crossover exponencial, podendo acelerar a convergência para soluções locais.

Fonte: (Amritpal Singh & Sushil Kumar, 2016).

A Figura 10 apresenta um diagrama explicativo do algoritmo.

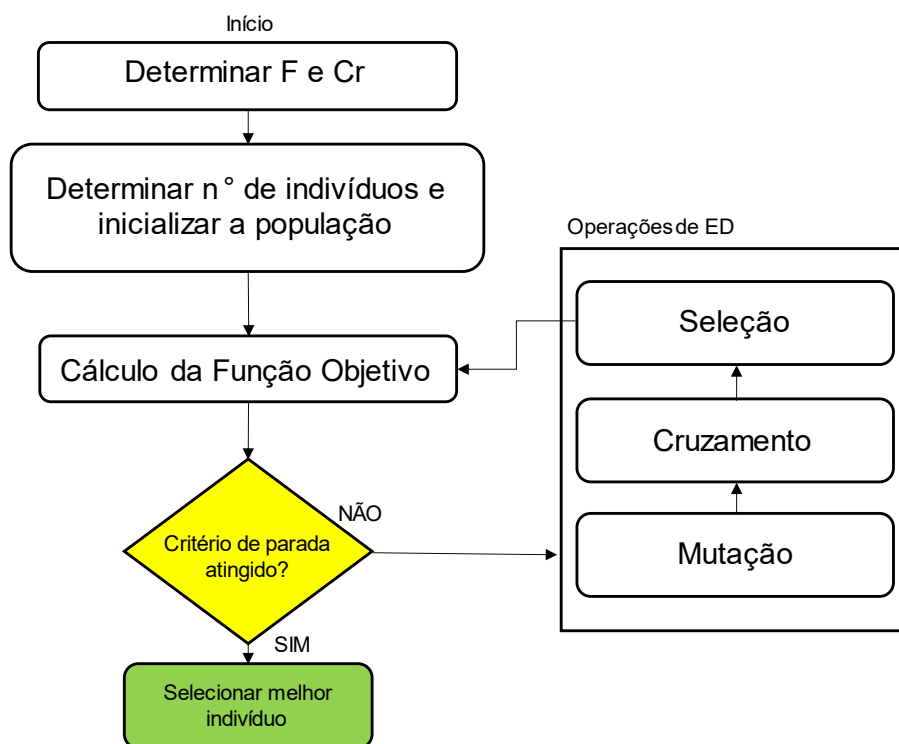


Figura 10. Fluxograma de funcionamento da ED.

No início da implementação do problema é necessário definir os valores de alguns parâmetros como o peso diferencial (F) e a taxa de cruzamento (Cr). Então, é necessário inicializar a população com indivíduos gerados aleatoriamente para a determinação do tamanho da população. Aumentar a população gera um custo computacional maior ao problema.

O cálculo da função objetivo ou também cálculo de *fitness* em uma ED é uma etapa fundamental para avaliar o desempenho e a qualidade das soluções. Nesse contexto, o *fitness* refere-se à medida de quão bem uma solução candidata atende aos critérios de otimização definidos pelo problema em questão. Este cálculo envolve a avaliação do desempenho ou adequação de cada solução candidata em relação aos objetivos específicos do problema. Isso geralmente é feito por meio de uma função de avaliação, também conhecida como função objetivo ou função de aptidão, que mapeia as soluções candidatas em um valor numérico que representa sua qualidade ou adequação (Price et al., 2006; Xia et al., 2021).

Na fase de operações da ED são realizadas a mutação, o cruzamento e a seleção. Essas operações (seleção, recombinação e mutação) são repetidas por várias gerações até que um critério de parada seja atingido, como o número máximo de iterações ou a convergência para uma solução satisfatória. Ao longo do processo evolutivo, espera-se que a população evolua e se adapte, com soluções candidatas de maior *fitness* emergindo gradualmente.

As definições utilizadas nos procedimentos a seguir são baseadas nos trabalhos de Oliveira (2006) e Price et al. (2006).

A Equação 3 rege o processo de mutação e é apresentada logo abaixo.

$$V^{q+1} = X_{\alpha}^q + F(X_{\beta}^q - X_{\gamma}^q) \quad (3)$$

Em que V^{q+1} é o vetor doador e X_{α}^q , X_{β}^q e X_{γ}^q são indivíduos mutualmente distintos e aleatórios escolhidos da população. F é uma constante real chamada de fator de escala $\in [0, 2]$ que determina o tamanho do passo a ser dado na direção definida pelo vetor diferença $X_{\beta}^q - X_{\gamma}^q$.

No cruzamento, há a combinação dos vetores doadores com componentes do vetor alvo, ou seja, de outro vetor aleatoriamente escolhido. A Equação 4 representa essa etapa.

$$u(i)^{q+1} = \begin{cases} v(i)^{q+1} & \text{se } rand_i \leq CR \\ x_s(i)^q & \text{se } rand_i > CR \end{cases} \quad i = 1, 2, \dots, n \quad (4)$$

Sendo X_s^q o vetor alvo e $u(i)^{q+1}$, $x_s(i)^q$ e $v(i)^{q+1}$ são os componentes dos vetores experimental, alvo e doador, respectivamente. $rand_i$ são números com distribuição uniforme no intervalo $[0, 1]$, gerados aleatoriamente.

Compara-se o valor de custo do vetor experimental e do vetor alvo, sendo o de menor custo o escolhido para passar para a geração seguinte. Essa relação é mostrada na Equação 5.

$$\begin{cases} \text{Se } f(U^{q+1}) \leq f(X_s^q) \text{ então } X_s^{q+1} = U^{q+1} \\ \text{Se } f(U^{q+1}) > f(X_s^q) \text{ então } X_s^{q+1} = X_s^q \end{cases} \quad (5)$$

X_s^{q+1} é o vetor alvo da próxima geração, U^{q+1} o vetor experimental e X_s^q o custo do vetor alvo. Torna-se necessário definir um critério de parada, nesse caso, o número de iterações consecutivas.

2.3.3 - Homologia Simplicial

O *Simplicial Homology Global Optimization* (SHGO) é uma técnica de otimização global que combina conceitos da teoria de homologia simplicial com métodos de busca local para encontrar o mínimo global de uma função em um espaço multidimensional.

A abordagem do SHGO é baseada na construção de complexos simpliciais no espaço de busca, que são estruturas formadas por combinações de pontos e suas interconexões. Esses complexos capturam informações sobre a topologia do espaço de busca e permitem explorar sua estrutura de forma eficiente (Ren et al., 2022; Tereshin et al., 2020).

O SHGO divide o espaço de busca em regiões menores e aplica métodos de busca local em cada região para encontrar mínimos locais. Em seguida, os mínimos locais são comparados utilizando critérios de homologia simplicial para determinar o mínimo global. A homologia simplicial é uma técnica matemática que permite analisar a conectividade e a estrutura topológica dos complexos simpliciais (Tereshin et al., 2020; H. Zhao et al., 2022).

O SHGO é capaz de lidar com funções objetivo não suaves ou com ruído. É uma abordagem versátil que pode ser aplicada a problemas de otimização com restrições e com múltiplas variáveis. No entanto, assim como outras técnicas de otimização global, a eficácia do SHGO depende da natureza do problema em questão e da definição adequada dos parâmetros de entrada (Endres et al., 2018).

A seguir é utilizada do trabalho de Endres et al. (2018), para melhor entendimento do algoritmo.

O algoritmo utiliza conceitos da teoria de homologia integral combinatória para encontrar subdomínios que são, aproximadamente, localmente convexos e fornece caracterizações da função objetivo à medida que o algoritmo progride, sendo feito em várias etapas.

Inicialmente a construção de um complexo simplicial H feito a partir dos pontos amostrais mapeados através de f como vértices seguindo as construções descritas. Em seguida, um homomorfismo é encontrado entre H e K . Outro complexo simplicial que existe em uma superfície construída abstrata S .

A variedade n -dimensional S é uma soma g conexa de g tori (figura geométrica gerada pela revolução de um círculo) $S := S_0 \# S_1 \# \dots \# S_{g-1}$. Em um caso 2D, as Figuras 11 e 12 demonstram essa construção geométrica.

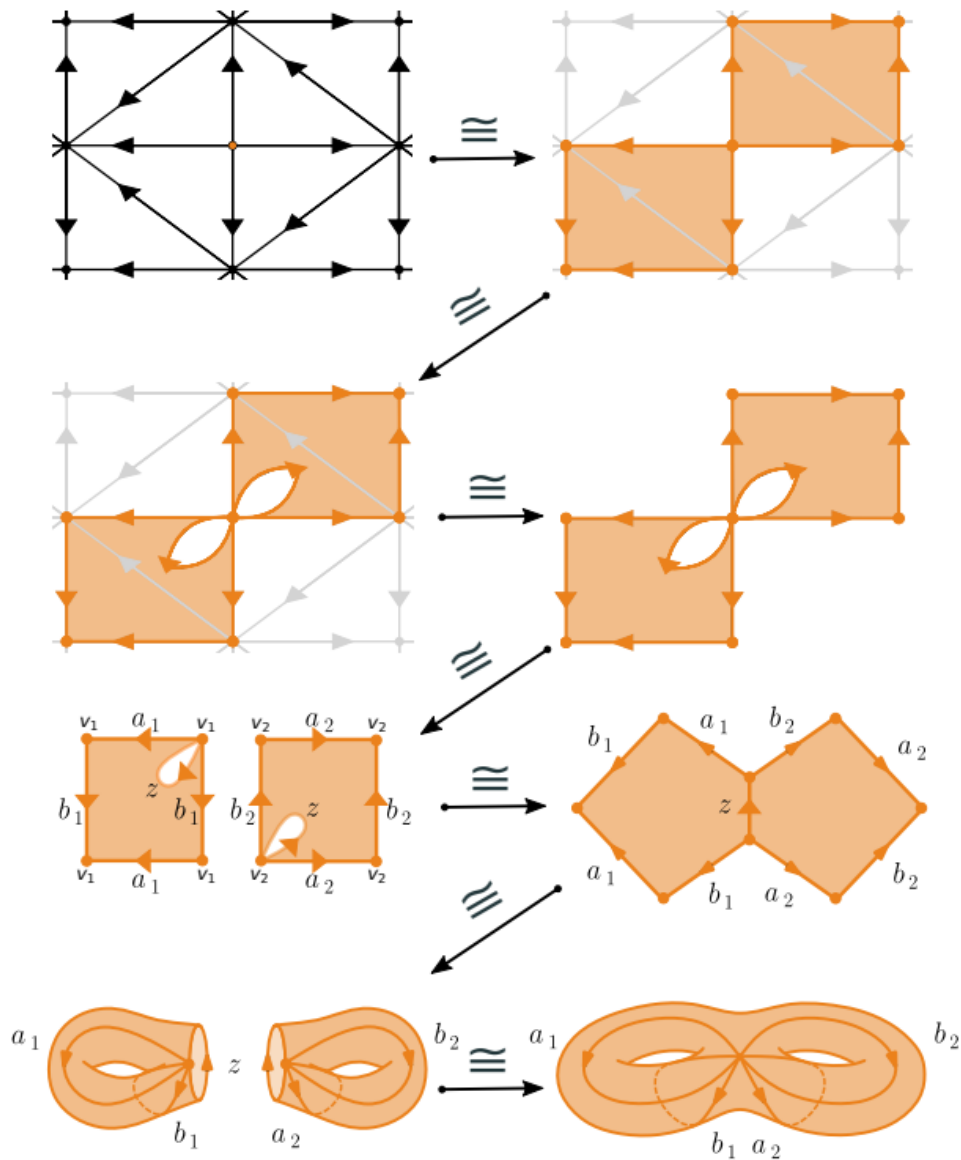


Figura 11. Punção de uma hiper esfera em um ponto minimizador em um espaço de busca compacto. Fonte: (Endres et al., 2018).

Inicia-se a compreensão da figura 11 identificando um ponto minimizador no gráfico $H^1(\cong K^1)$. O complexo inicial existe na (hiper)superfície de um toro S_0 n -dimensional, de modo que o restante de K^1 seja conectado e compacto. É puncionada uma hiper esfera no ponto minimizador e identificadas as arestas resultantes (ou $(n - 1)$) para problemas de dimensão superior. Em seguida, é feita a escolha do restante do complexo simplicial para as faces e vértices do modelo. Fazendo as identificações para S_0 e cola-se a face identificada e conectada z $(n - 1)$ resultante da punção da hiper esfera. As outras faces (ou seja, $(n - 1)$) são conectadas da maneira usual para construções do tori.

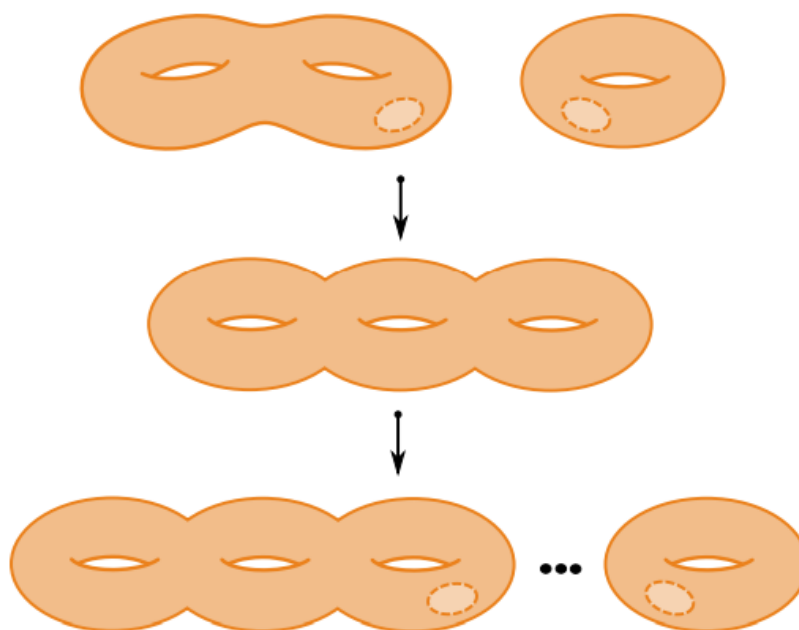


Figura 12. O processo de perfuração de uma nova hiper esfera. Fonte: (Endres et al., 2018).

Usando uma extensão do teorema do ponto fixo de Brouwer adaptado para lidar com restrições não lineares, é provado que cada um dos pontos minimizadores na figura 12 corresponde a um subdomínio contendo um mínimo local único quando o problema é adequadamente amostrado (Henle, 1994). Através do Teorema da Invariância e dos Axiomas de Eilenberg & Steenrod, (2015) desenha-se outro homomorfismo entre as superfícies de f e S .

A partir das propriedades conhecidas de S se deduz propriedades da função desconhecida f . A propriedade correspondente mais importante são os grupos de homologia de S denotados como $H_i(S)$. O posto de um dos grupos $H_1(S)$ corresponde ao número de mínimos locais em f . À medida que a amostragem aumenta e mais mínimos locais são encontrados, também aumenta a classificação de $H_1(S)$.

Ao usar amostragem uniforme, isso fornece uma indicação de sua multimodalidade e a dispersão das soluções. Além disso, foi provado em Endres et al., 2018 que o posto de $H_1(S)$ não pode aumentar além do número real de mínimos locais em f após amostragem adequada. Finalmente, usando as propriedades abelianas dos grupos de homologia, é possível entender todas as propriedades anteriores comprovadas para manter as descontinuidades não lineares, conforme demonstrado geometricamente na Figura 13.

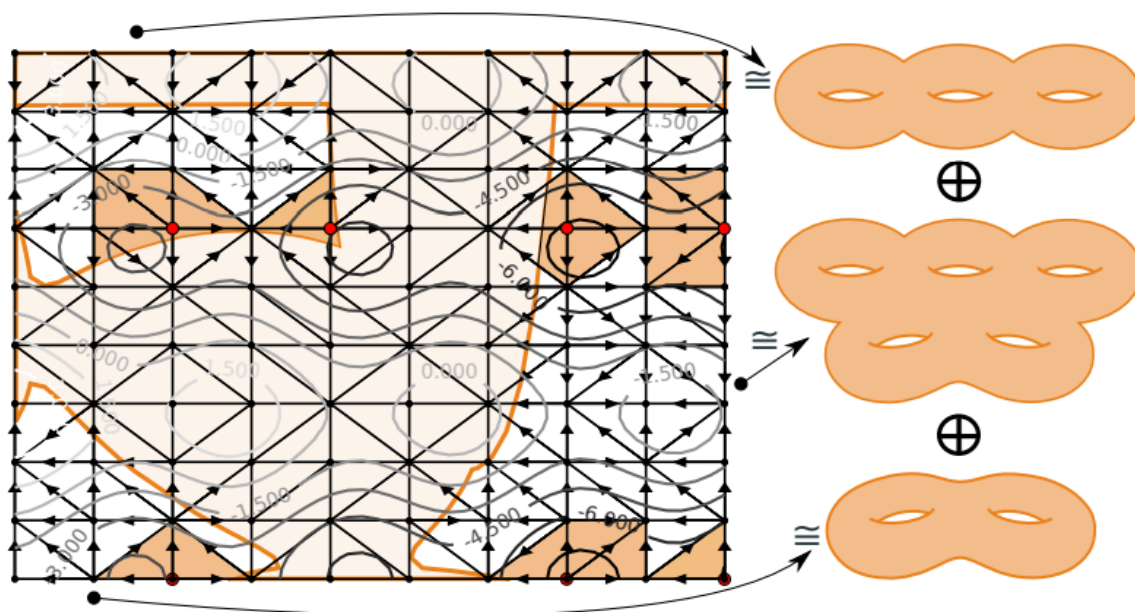


Figura 13. Demonstração visual em superfícies com restrições não lineares. Fonte: (Endres et al., 2018).

Tem-se que a região sombreada da figura 13 é inviável. Os vértices dos pontos mapeados ao infinito possuem arestas não direcionadas, portanto não formam complexos simpliciais na homologia integral.

2.3.4 – Otimização de RNA's

No trabalho de Abdolrasol et al. (2021) é possível identificar tendências da indústria e da academia sobre a aplicação de métodos de otimização em RNA's. São usadas metodologias de otimização de arquitetura, pesos, iterações, *overfitting*, critérios de parada, dentre outras em numerosas áreas de estudo. As técnicas de otimização vão desde algoritmos genéticos (AG), à criação de novos algoritmos como extensões dos já amplamente difundidos. Dentre os trabalhos com foco na otimização da arquitetura de redes neurais é apresentada em sequência um resumo de alguns desses trabalhos.

No trabalho de Ukaoha & Igodan (2020) há um esforço para determinar parâmetros otimizados de uma rede neural profunda através de um algoritmo de AG numa aplicação de classificação binária para admissão de estudantes em uma universidade. Usando-se também de um algoritmo AG, Cicek & Ozturk (2021) realizaram um estudo para otimização para número de neurônios, número de camadas ocultas e pesos para previsão de séries temporais para dados da vida real. Otimizando os pesos e limites de uma RNA do tipo *backpropagation* Q. Jiang et al. (2019) conseguiram prever problemas de risco de investimento na rede elétrica através do AG.

O *Backtracking Search Optimization Algorithm* (BSA) é usado em diversos trabalhos em aplicações contemporâneas, por exemplo, para melhorar o controle de carga de baterias de íon-lítio, através da otimização de uma rede neural de retropropagação afim de se obter taxas de aprendizado ótimas, além do valor ideal dos neurônios da camada oculta (Hannan et al., 2018).

As *Feedforward Neural Networks* (FNNs) são usadas em Jian-wei (2004) e se baseiam na otimização de um cardume peixes artificiais, ou *Artificial Fish Swarm Optimization* (AFSA) para obter o número de camadas ocultas treinadas com o objetivo de substituir o *backpropagation* em RNA's.

No estudo de Sheng et al. (2017), é utilizado um algoritmo memético adaptativo com uma mutação baseada em classificação, *Adaptive Memetic Algorithm With A Rank-Based Mutation* (AMARM) para otimizar arquiteturas de RNA por meio de ajuste fino nos pesos de conexão entre neurônios, assim como nas camadas ocultas.

Por meio de pesquisas sobre a aplicação da Evolução Diferencial (ED) na otimização de soldagem, vários estudos relevantes foram identificados. Panda et al. (2014) propuseram um método que combina redes neurais com ED para otimizar parâmetros de soldagem, melhorando com sucesso a resistência da junta e reduzindo o consumo de energia ao modelar a relação entre fatores de soldagem, como corrente, força, tempo e espessura da chapa. Da mesma forma, Alkayem et al. (2019) usaram redes neurais artificiais (RNA) para prever a qualidade da solda com base em vários parâmetros e aplicaram algoritmos evolutivos como DE para otimizar saídas como resistência à tração e dureza, com o algoritmo Neuro-DEMO fornecendo resultados superiores. Expandindo isso, Alkayem et al. (2017) exploraram algoritmos neuroevolutivos para soldagem por fricção e agitação (FSW), onde modelos de RNA foram acoplados com DE multiobjetivo e otimização de enxame de partículas (PSO) para melhorar as qualidades da solda, identificando PSO como o mais eficiente. A principal distinção deste trabalho está em seu foco na otimização de parâmetros de soldagem especificamente para soldagem a arco elétrico a gás (GMAW), visando não apenas a resistência da junta, mas também o desenvolvimento de uma aplicação prática para auxiliar soldadores em campo, o que não foi explorado em estudos anteriores. A otimização de funções de ativação também não é comum como o número de camadas, neurônios e pesos além da taxa de aprendizagem, que nos trabalhos abordados possuem valores num intervalo $[0 - 0,8]$ (Abdolrasol et al., 2021).

2.4 – Regressão

Neste tópico são abordadas técnicas de regressão usadas para predição de parâmetros do *dataset* de soldagem para comparação com outras técnicas do trabalho.

2.4.1 – Regressão Linear

A regressão linear é uma das técnicas mais antigas e amplamente utilizadas em estatística e análise de dados. Ela visa modelar a relação entre uma variável dependente (ou resposta) e uma ou mais variáveis independentes (ou preditoras) através de uma equação linear. O modelo de regressão linear simples é expresso como:

$$Y_l = \beta_0 + \beta_1 X_l + e \quad (6)$$

onde, Y_l é a variável dependente, X_l é a variável independente, β_0 é o coeficiente de interceptação, β_1 é o coeficiente de inclinação e e é o erro aleatório.

As características lineares dos modelos de regressão tornam-nos particularmente adequados para modelar relações lineares entre variáveis, que são comuns em muitos aspectos dos processos de soldagem, por exemplo. Um dos principais benefícios da regressão linear é sua interpretabilidade. Os coeficientes de regressão fornecem informações diretas sobre a relação entre as variáveis independentes e dependentes, permitindo uma compreensão clara dos efeitos de cada variável no resultado. Além disso, a validação dos modelos de regressão pode ser realizada usando técnicas como validação cruzada, *bootstrap* ou divisão de dados em conjuntos de treinamento e teste (Maulud & Abdulazeez, 2020).

Geralmente, a análise de regressão linear começa com a exploração e visualização dos dados, buscando identificar relações potenciais entre as variáveis. A análise de correlação é frequentemente empregada para avaliar a força e a direção das relações entre as variáveis, fornecendo insights iniciais sobre quais variáveis podem ser incluídas no modelo de regressão. Em seguida, os modelos de regressão linear são ajustados aos dados, estimando-se os coeficientes de regressão (Acharya et al., 2019; Maulud & Abdulazeez, 2020).

Embora a regressão linear seja uma técnica poderosa e amplamente aplicável, ela também tem suas limitações. A principal limitação está na suposição de linearidade entre as variáveis, o que pode não ser válida em todos os casos. Em processos de soldagem, por exemplo, as relações entre os parâmetros de soldagem e as características da junta soldada podem ser não-lineares ou complexas. Nesses casos, técnicas de regressão não linear ou modelos mais flexíveis podem ser mais apropriados. No entanto, apesar de suas limitações, a regressão linear continua a ser uma ferramenta valiosa na análise e previsão de parâmetros em processos de soldagem, fornecendo insights úteis para otimização e controle de processo (T. Lei et al., 2018; Vedrtnam et al., 2018).

2.4.2 – Regressão Polinomial

A regressão polinomial é uma extensão da regressão linear, onde a relação entre a variável dependente e as variáveis independentes é modelada por meio de uma função polinomial de grau n . A forma geral de um modelo de regressão polinomial de grau n é expressa como:

$$Y_p = \beta_0 + \beta_1 X_p + \beta_2 X_p^2 + \dots + \beta_n X_p^n + e \quad (7)$$

onde, Y_p é a variável dependente, X_p é a variável independente, $\beta_0, \beta_1, \dots, \beta_n$ são os coeficientes de regressão e e é o erro aleatório.

A regressão polinomial é frequentemente utilizada quando a relação entre as variáveis não pode ser adequadamente modelada por um modelo linear simples. Em processos de soldagem, onde a relação entre os parâmetros de soldagem e as características da junta soldada pode ser não-linear ou complexa, a regressão polinomial pode oferecer uma alternativa viável para modelar essas relações. Ao permitir que os modelos se ajustem a padrões não-lineares nos dados, a regressão polinomial pode capturar melhor a complexidade dos processos de soldagem e fornecer previsões mais precisas dos parâmetros de interesse (Avery, 2010; Budu, 2014).

Um dos desafios ao usar a regressão polinomial é determinar o grau do polinômio apropriado para o modelo. Um polinômio de grau muito alto pode levar ao *overfitting* dos dados, resultando em um modelo que se ajusta muito bem aos dados de treinamento, mas não generaliza bem para novos dados. Por outro lado, um polinômio de grau muito baixo pode não capturar adequadamente a estrutura dos dados, resultando em *underfitting*. A seleção adequada do grau do polinômio é crucial para obter um modelo que balanceie o viés e a variância. Técnicas como validação cruzada e critérios de seleção de modelo, como o critério de informação de Akaike ou o critério de informação bayesiana, podem ser utilizadas para determinar o grau ótimo do polinômio (Avery, 2010; Kleijnen, 2017).

Apesar dos desafios associados à regressão polinomial, sua capacidade de modelar relações não-lineares faz dela uma ferramenta valiosa na análise e previsão de parâmetros em processos de soldagem.

2.4.3 – Regressão Múltipla de Ridge

A regressão múltipla de ridge é uma técnica de modelagem estatística que visa lidar com o problema de multicolinearidade em modelos de regressão múltipla. A multicolinearidade ocorre quando duas ou mais variáveis independentes estão altamente correlacionadas entre si, o que pode levar a coeficientes de regressão instáveis ou mal

condicionados. A regressão de ridge introduz um termo de regularização na função objetivo do modelo, penalizando coeficientes de regressão grandes, o que ajuda a estabilizar os estimadores e reduzir a variância dos coeficientes (McDonald, 2009).

A forma geral do modelo de regressão múltipla de ridge é expressa como:

$$Y_r = \beta_0 + \beta_1 X_{1r} + \beta_2 X_{2r} + \dots + \beta_p X_{pr} + e \quad (8)$$

onde, Y_r é a variável dependente, $X_{1r}, X_{2r}, \dots, X_{pr}$ são as variáveis independentes, $\beta_0, \beta_1, \dots, \beta_p$ são os coeficientes de regressão e e é o erro aleatório.

Na regressão de Ridge, a função objetivo a ser minimizada é ajustada adicionando uma penalidade proporcional ao quadrado dos coeficientes de regressão:

$$\text{Minimizar } \sum_{i=1}^n (Y_i - \hat{Y}_i)^2 + \lambda \sum_{j=1}^p \beta_j^2 \quad (9)$$

onde, n é o número de observações, Y_i é o valor observado da variável dependente para a i -ésima observação, $\hat{Y}_i$ é o valor previsto da variável dependente para a i -ésima observação, λ é o parâmetro de regularização, que controla a força da penalidade.

A regressão de Ridge é especialmente útil quando as variáveis independentes estão altamente correlacionadas entre si, o que pode levar a estimativas instáveis ou de alta variância nos coeficientes de regressão. Ao introduzir a penalidade de regularização, a regressão de Ridge ajuda a reduzir a variância dos estimadores, melhorando a estabilidade e a generalização do modelo (McDonald, 2009; Vovk, 2013).

A seleção do parâmetro de regularização λ é um aspecto importante em sua aplicação. Um valor muito baixo de λ pode resultar em estimativas semelhantes às obtidas pela regressão linear padrão, enquanto um valor muito alto de λ pode levar a estimativas muito pequenas dos coeficientes de regressão. Técnicas como validação cruzada ou critérios de seleção de modelo podem ser utilizadas para determinar o valor ótimo de λ que equilibra o viés e a variância do modelo (Dorugade, 2014; McDonald, 2009).

2.4.4 – Coeficiente de Determinação

O coeficiente de determinação, ou R^2 , desempenha um papel importante na avaliação da qualidade e adequação dos modelos de regressão. Ao fornecer uma medida da proporção da variabilidade na variável dependente que é explicada pelo modelo, o R^2 permite aos pesquisadores e profissionais determinarem o quão bem o modelo se ajusta aos dados observados (Su et al., 2012).

Matematicamente, o R^2 é calculado como:

$$R^2 = 1 - \frac{SSE}{SST} \quad (10)$$

onde, SSE (Soma dos Quadrados dos Erros) é a soma dos quadrados das diferenças entre os valores observados e os valores previstos pelo modelo, SST (Soma Total dos Quadrados) é a soma dos quadrados das diferenças entre os valores observados e a média dos valores observados.

A importância do R^2 reside em sua capacidade de indicar a eficácia do modelo em capturar e explicar a relação entre as variáveis independentes e dependentes. Um valor elevado de R^2 próximo de 1 sugere que o modelo é capaz de explicar uma grande parte da variabilidade na variável dependente, o que é desejável em muitas aplicações. Por outro lado, um valor baixo de R^2 próximo de 0 indica que o modelo não consegue explicar adequadamente a variação na variável dependente, sugerindo a necessidade de reavaliar ou aprimorar o modelo (Su et al., 2012).

Na prática, o R^2 é amplamente utilizado em diversas áreas, como engenharias, ciências sociais, econômicas, médicas e científicas. É frequentemente empregado para comparar diferentes modelos de regressão e determinar qual deles melhor se ajusta aos dados. Além disso, o R^2 pode ser usado para comunicar os resultados do modelo para stakeholders e tomadores de decisão, fornecendo uma medida quantitativa da sua eficácia (Engström et al., 2010; Su et al., 2012).

2.4.5 – Modelos de regressão aplicados

A análise de regressão serve como eixo na busca pela compreensão e otimização dos parâmetros de soldagem. Ao estabelecer relações matemáticas entre parâmetros de entrada e resultados de soldagem, os modelos de regressão fornecem uma estrutura sistemática para prever e controlar (Hengl et al., 2007). A seguir são apresentados trabalhos de predição de problemas de soldagem com técnicas de regressão e machine learning (ML), que é um campo da inteligência artificial que desenvolve algoritmos e modelos que permitem a sistemas aprender e melhorar automaticamente a partir de dados sem intervenção explícita. Ele se baseia em métodos estatísticos e computacionais para identificar padrões e fazer previsões ou decisões com base em novos dados.

Saeheaw (2023) fez uma comparação de múltiplas técnicas de aprendizado profundo e aprendizado de máquina com regressão na previsão da geometria do cordão com as entradas de velocidade do arame, tensão e *stickout*. A precisão dos resultados é de 88-91%, mostrando a eficácia das técnicas de ML. R. Yu et al. (2020) aplicaram um método de previsão em tempo

real para o modo e profundidade de penetração da solda durante o processo GMAW. Utilizando-se características visuais bidimensionais da poça de fusão, um sistema de detecção de visão monocular é estabelecido. As principais características, como área, comprimento e largura da poça de fusão são extraídas. Modelos de regressão de vetores de suporte e redes neurais de retropropagação são então desenvolvidos para prever a profundidade da soldagem em tempo real. Resultados experimentais demonstram a eficácia dos modelos propostos com precisão de 95,7%.

Costa et al. (2022) usam uma regressão vetorial de suporte e programação quadrática sequencial para prever e otimizar parâmetros de soldagem subaquática molhada a arco com eletrodo tubular (FCAW). Os experimentos revelaram configurações ideais (tensão = 24 V, velocidade de alimentação do arame = 8.890 mm/min, velocidade de soldagem = 3,4 mm/s) para porosidade superficial mínima em soldas de aço ASTM A36. A abordagem computacional aproxima efetivamente os níveis de porosidade superficial, demonstrando seu potencial no aumento da precisão. Mallaradhya et al. (2022) examinam em aço inoxidável parâmetros de soldagem por resistência a ponto (força do eletrodo, tensão e corrente de soldagem). As descobertas destacam a influência primordial da força do eletrodo na qualidade da solda. Os modelos de regressão e de rede neural apresentam bom desempenho, com a regressão ligeiramente se aproximando na previsão de respostas ideais. O treinamento contínuo da rede neural melhora o desempenho ao longo do tempo.

As previsões da qualidade da soldagem usando regressão são frequentemente usadas como em D. Zhao et al. (2020) que comparam modelos de regressão e de redes neurais artificiais para prever o diâmetro da lentilha em juntas soldadas por pontos. A liga de titânio TC2 foi soldada por ponto, e variações distintas do sinal de potência foram observadas para diferentes condições de soldagem. Cinco características foram extraídas para classificar as soldas. Foram utilizados modelos de regressão *stepwise* e de rede neural de retropropagação, com diferenças de desempenho vinculadas às suas capacidades de mapeamento de dados.

Lee e Rhee (2000) criaram uma equação de regressão preditiva para parâmetros de processo de soldagem em GMAW com folga de ranhura para soldagem de topo. O modelo de regressão demonstra precisão com um erro máximo de previsão inferior a 9,5% no processo direto, e inferior a 6,5% no processo inverso. Isso afirma a viabilidade de obtenção de uma equação de regressão para o sistema de predição de *back-bead* através de ambos os processos. T. Lei et al. (2018) utilizam o método Taguchi e análise de regressão para investigar a relação entre os parâmetros de soldagem (potência do laser, velocidade de soldagem e distância focal) e a morfologia da poça fundida. A potência do laser tem um impacto maior do que a velocidade de soldagem e a distância focal. A análise estatística estabelece quatro equações de regressão não linear precisas (erros relativos médios < 5%) relacionando os parâmetros do processo às características da poça fundida. Shevchenko et

al. (2018) exploram a soldagem por eletro contato de trilhos ferroviários usando refluxo pulsante em uma máquina K1000, com testes de flexão transversal estática de três pontos. Através de análise de regressão múltipla, identifica os principais parâmetros de soldagem que influenciam a força de flexão e a deflexão. Um modelo matemático é criado para prever a qualidade da costura soldada e controlar parâmetros tecnológicos no processo de soldagem por contato de trilhos.

Vedrtnam et al. (2018) em uma soldagem por arco submerso conduziram soldas *cord-on-plate* em placas de aço inoxidável com uma máquina automatizada, o estudo empregou metodologia de superfície de resposta) e análise de regressão para estabelecer relações de entrada-saída. Os resultados mostram que os parâmetros de soldagem influenciam a geometria do cordão e as propriedades mecânicas da junta. Sevvel et al. (2020) abordam uma alternativa aos métodos tradicionais para seleção de parâmetros de processo de soldagem e introduz um modelo de regressão quadrática através da metodologia de superfície de resposta para soldagem por fricção (FSW). O modelo estabelece relações entre parâmetros de entrada (velocidade de rotação da ferramenta, força axial, velocidade de deslocamento da ferramenta) e cinco respostas de saída. Os resultados mostram um ajuste próximo entre os resultados previstos e experimentais, com discrepâncias de $\pm 1\%$. O estudo também identifica a combinação ideal de parâmetros de entrada para alcançar os resultados de saída desejados.

O uso de técnicas de regressão com aumento de dados pode levar a problemas de *overfitting*, causados principalmente pela multicolinearidade, que é uma condição que ocorre quando algumas variáveis preditoras no modelo estão correlacionadas com outras variáveis preditoras. A multicolinearidade forte é problemática porque pode aumentar a variância dos coeficientes de regressão, como observado no trabalho de Kroll & Song (2013), onde a presença de multicolinearidade afetou a regressão. Aviral Gupta et al. (2017) também estudaram que a multicolinearidade e a complexidade do conjunto de dados podem causar *overfitting* e *underfitting*, mas o risco foi reduzido com o uso da regressão Ridge.

2.5 – Testes Estatísticos

Neste item são apresentados temas estatísticos para validação da utilização das variáveis de entrada e saída da RNA.

2.5.1 – Correlação

Os testes de correlação são técnicas estatísticas utilizadas para avaliar a relação entre duas ou mais variáveis. Esses testes são amplamente utilizados em diversas áreas, incluindo ciências sociais, biologia, psicologia, economia, entre outras. O objetivo é identificar a

presença de uma relação entre as variáveis, bem como avaliar a força e a direção dessa relação. Existem diferentes tipos de testes de correlação, como o teste de correlação de Pearson, o teste de correlação de Spearman e o teste de correlação de Kendall, cada um com suas próprias suposições e aplicações. A escolha do teste de correlação adequado depende das características das variáveis em estudo e dos objetivos da pesquisa.

2.5.1.1 – Teste de Correlação de Pearson

O teste de correlação de Pearson é uma técnica estatística utilizada para avaliar a relação linear entre duas variáveis quantitativas desenvolvido por Karl Pearson e é amplamente utilizado em diversas áreas.

O coeficiente de correlação de Pearson, representado por r , é calculado a partir das médias, desvios padrão e covariância das duas variáveis. O coeficiente de correlação varia de -1 a 1, onde -1 indica uma correlação perfeitamente negativa (ou inversa), 0 indica ausência de correlação e 1 indica uma correlação perfeitamente positiva. Em (11) tem-se a covariância amostral entre duas variáveis Y_1 e Y_2 e (12) calcula o coeficiente de correlação de Pearson (Temizel et al., 2011).

$$Cov(y_1, y_2) = \frac{1}{n-1} \sum_{i=1}^n (y_{1i} - \bar{y}_1)(y_{2i} - \bar{y}_2) \quad (11)$$

$$r = \frac{Cov(y_1, y_2)}{\sqrt{\sum_{i=1}^n (y_{1i} - \bar{y}_1)^2} \cdot \sqrt{\sum_{i=1}^n (y_{2i} - \bar{y}_2)^2}} \quad (12)$$

Uma correlação perfeitamente negativa significa que uma variável aumenta à medida que a outra diminui, enquanto uma correlação perfeitamente positiva indica que ambas as variáveis aumentam ou diminuem juntas. A ausência de correlação significa que as duas variáveis não estão relacionadas linearmente. Um exemplo de correlações entre variáveis é mostrado na Figura 14.

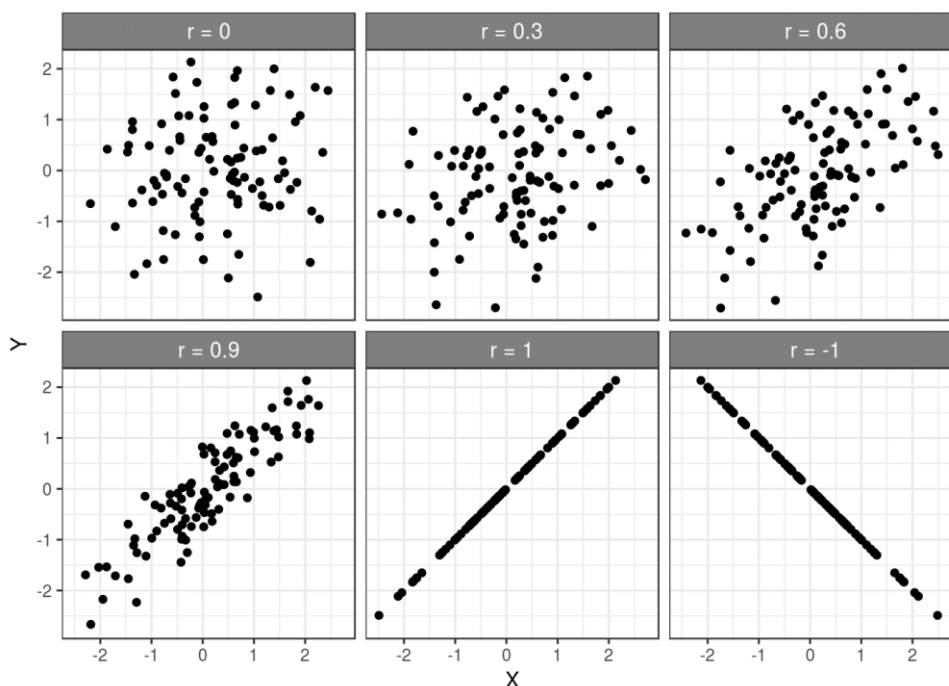


Figura 14. Exemplos de correlação entre duas variáveis.

O teste de correlação de Pearson é usado para avaliar se o coeficiente de correlação entre duas variáveis é significativamente diferente de zero. O teste de hipótese é baseado na estatística *t* de Student, que é calculada a partir do coeficiente de correlação, do tamanho da amostra e do grau de liberdade.

O valor-*p* do teste de correlação de Pearson indica a probabilidade de obter uma correlação tão forte quanto a observada na amostra, assumindo que a correlação verdadeira na população é igual a zero. Se o valor-*p* for menor que o nível de significância (geralmente 0,05), a hipótese nula de que a correlação é igual a zero é rejeitada, o que significa que há evidências estatísticas para concluir que há uma relação linear entre as duas variáveis. Se o valor-*p* for maior que o nível de significância, a hipótese nula não é rejeitada, o que significa que não há evidências suficientes para concluir que as duas variáveis estão linearmente relacionadas (Figueiredo & Silva Junior, 2010).

É importante notar que o teste de correlação de Pearson pressupõe que as duas variáveis sejam distribuídas normalmente e que essa relação seja linear, sendo necessário um teste de normalidade prévio para determinar seu uso. Se essas suposições não forem atendidas, outros tipos de correlação, como a correlação de Spearman ou de Kendall, podem ser mais apropriados (Asuero et al., 2006).

2.5.2.2 – Correlação de Distância

O teste de correlação de distância (*distance correlation*) é uma medida estatística que avalia a relação entre duas variáveis, considerando tanto a associação linear quanto a não linear entre elas. Ao contrário dos métodos de correlação tradicionais, como o coeficiente de correlação de Pearson, o teste de correlação de distância é capaz de capturar relações não lineares entre as variáveis (Székely & Rizzo, 2009).

A distância correlação mede a dependência entre duas variáveis em termos de similaridade de padrões em seus valores, independentemente da forma funcional dessa relação. Ela é calculada em duas etapas principais, abordadas a seguir.

Cálculo da matriz de distância: A primeira etapa envolve o cálculo das matrizes de distância para as duas variáveis em estudo. A matriz de distância captura a dissimilaridade entre pares de observações em cada variável.

Seja $(X_k, Y_k), n = 1, 2, \dots, n$ uma amostra estatística de um par de duas variáveis aleatórias, X e Y . Assim, são calculadas as matrizes de distância n por n com $a_{j,k}$ e $b_{j,k}$ contendo todas as distâncias pareadas dados por:

$$\begin{aligned} a_{j,k} &= |X_j - X_k|, j, k = 1, 2, \dots, n \\ b_{j,k} &= |Y_j - Y_k|, j, k = 1, 2, \dots, n \end{aligned} \quad (13)$$

Em seguida, são tomadas as distâncias de centro duplo:

$$\begin{aligned} A_{j,k} &= a_{j,k} - \bar{a}_{j.} - \bar{a}_{.k} + \bar{a}_{..} \\ B_{j,k} &= b_{j,k} - \bar{b}_{j.} - \bar{b}_{.k} + \bar{b}_{..} \end{aligned} \quad (14)$$

A Figura 15 mostra a apresentação realizada.

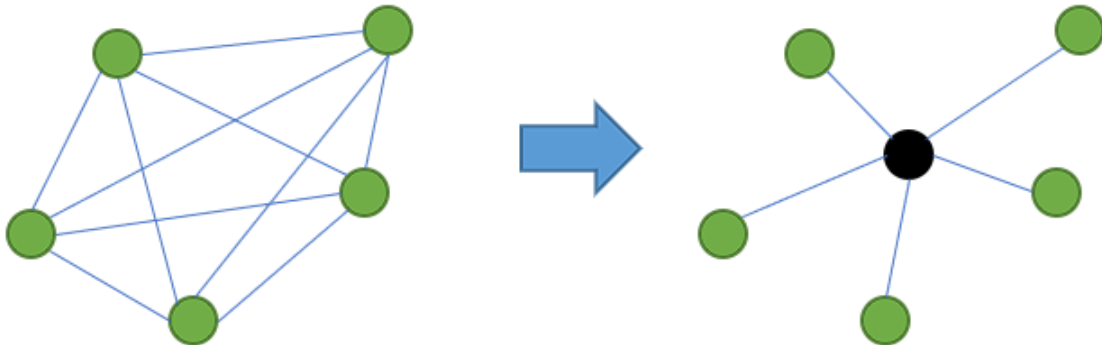


Figura 15. Representação gráfica da transformação da matriz de distância.

A razão para esses cálculos é que qualquer tipo de covariância é o produto cruzado de momentos. Momentos são medidas estatísticas usadas para descrever a forma de uma distribuição. O primeiro momento é a média, o segundo momento é a variância, e assim por diante. Como as distâncias não são momentos, temos que calculá-las em momentos. Para calcular esses momentos, é necessário calcular primeiro os desvios da média, que é o que a centralização dupla alcança.

Cálculo da distância correlação: A distância correlação é calculada como a medida de associação entre as matrizes de distância das duas variáveis. Ela varia entre 0 e 2 na maioria dos softwares e algoritmos encontrados, sendo 1 indicativo de ausência de correlação, 2 indicativo de correlação positiva perfeita, e 0 correlação negativa perfeita.

Por fim, é calculada a média aritmética dos produtos A e B para obter a covariância da distância da amostra ao quadrado:

$$dCov_n^2(X, Y) = \frac{1}{n^2} \sum_{j=1}^n \sum_{k=1}^n A_{j,k} B_{j,k} \quad (15)$$

A variância da distância é simplesmente a covariância da distância de duas variáveis idênticas. É a raiz quadrada do seguinte:

$$dVar_n^2(X) = dCov_n^2(X, X) = \frac{1}{n^2} \sum_{k,l} A_{k,l}^2 \quad (16)$$

O teste de correlação de distância permite uma avaliação mais ampla das relações entre as variáveis, especialmente quando as relações não são lineares ou quando há presença de *outliers*. No entanto, é importante ressaltar que o teste de correlação de distância não fornece informações sobre a direção ou a causalidade da relação entre as variáveis (Lyons, 2013).

A principal vantagem do teste de correlação de distância é sua capacidade de capturar relações não lineares, tornando-o uma ferramenta útil em áreas como genética, bioinformática e análise de séries temporais, onde as relações entre variáveis muitas vezes são complexas e não lineares (Kosorok, 2009; Székely & Rizzo, 2012).

É importante observar que o teste de correlação de distância não é tão amplamente conhecido ou utilizado quanto os métodos tradicionais de correlação, mas tem sido objeto de estudos e pesquisa em estatística e análise de dados. No trabalho de Li et al. (2023), o método é utilizado para filtrar e selecionar parâmetros para uso em uma RNA a fim de prever o consumo de energia para um *datacenter*. Através da correlação de distâncias foi possível eliminar parâmetros redundantes e atingir um resultado menor de erro em aproximadamente 5%.

2.6 – Base de Dados Reduzida

Um problema comum no treinamento de RNA's é a disponibilidade de conjuntos de dados limitados. Esses conjuntos de dados contêm uma quantidade limitada de exemplos para treinar a rede neural, o que pode levar a vários desafios e dificuldades durante o processo de treinamento. Aqui estão alguns problemas associados aos pequenos conjuntos de dados:

- **Generalização inadequada:** As redes neurais aprendem com os exemplos fornecidos durante o treinamento e procuram generalizar esses exemplos para novos dados. Com pequenos conjuntos de dados, há menos exemplos para a rede neural aprender e, portanto, existe o risco de que a rede se torne excessivamente ajustada aos dados de treinamento, não sendo capaz de generalizar bem para novos dados. Isso é conhecido como *overfitting*, mostrado na comparação com uma boa adaptação, na Figura 16.

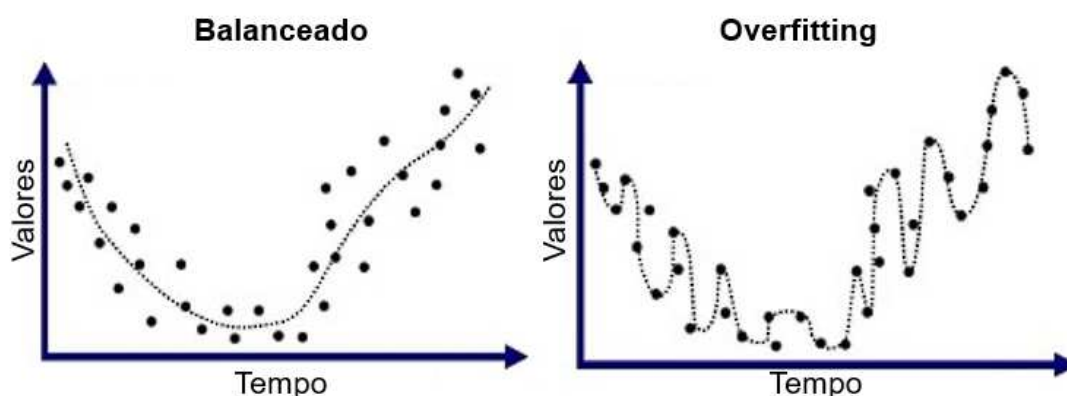


Figura 16. Exemplo de modelo balanceado e com *overfitting*. Fonte: Adaptado de <https://abracd.org/overfitting-e-underfitting-em-machine-learning/>.

- **Instabilidade do treinamento:** Com pequenos conjuntos de dados, a variabilidade nos exemplos pode ser menor, o que pode tornar o processo de treinamento instável. Pequenas mudanças nos dados de treinamento podem levar a grandes variações nos resultados da rede neural treinada. Isso ocorre porque a rede pode se concentrar em padrões ou ruídos irrelevantes presentes no conjunto de dados pequeno (Haykin, 1998). (Haykin, 1998)
- **Dificuldade em aprender representações complexas:** As redes neurais geralmente exigem um grande número de exemplos para aprender representações complexas. Com pequenos conjuntos de dados, a capacidade da rede neural de aprender representações mais ricas e abstratas pode ser limitada. Isso pode ser problemático em tarefas onde a extração de características é crucial (Schmidhuber, 2015).

- Dificuldade em ajustar hiper parâmetros: As redes neurais possuem vários hiper parâmetros, como taxa de aprendizado, tamanho do lote e número de camadas ocultas. A escolha adequada desses hiper parâmetros é essencial para obter um bom desempenho da rede. No entanto, com pequenos conjuntos de dados, pode ser difícil ajustar esses hiper parâmetros devido à falta de exemplos suficientes para avaliar seu impacto de forma confiável (L. Yang & Shami, 2020).

Existem algumas estratégias que podem ajudar a mitigar os desafios associados aos pequenos conjuntos de dados:

- Aumento de dados (*data augmentation*): Essa técnica envolve a criação de novos exemplos de treinamento sinteticamente a partir dos exemplos existentes, aplicando transformações como rotações, inversões, zooms, entre outros (Shorten & Khoshgoftaar, 2019; Taylor & Nitschke, 2018). Isso aumenta a quantidade de dados disponíveis para treinamento e pode ajudar a melhorar a generalização da rede neural, dentre essas técnicas tem-se o Kriging e Geração Sintética, por exemplo.
- Transferência de aprendizado (*transfer learning*): Nessa abordagem, uma rede neural pré-treinada em um conjunto de dados grande e relacionado é usada como ponto de partida para treinar uma nova rede neural no conjunto de dados menor. Isso permite que a nova rede neural aproveite os conhecimentos prévios da rede pré-treinada, melhorando assim seu desempenho com menos dados (Shin et al., 2016).
- Regularização: O uso de técnicas de regularização, como a adição de termos de penalidade aos pesos da rede neural (por exemplo, L1 ou L2), pode ajudar a evitar o *overfitting*, mesmo com pequenos conjuntos de dados.
- Validação Cruzada: A técnica *Leave One Out* (LOO) é uma dessas estratégias, em que cada exemplo de treinamento é usado como conjunto de dados de treinamento para treinar a rede neural, enquanto o exemplo restante é usado como conjunto de dados de teste para avaliar o desempenho da rede (Berrar, 2019).

Em muitos casos envolvendo processos de fabricação, a coleta de dados pode ser cara, demorada ou inviável devido a restrições práticas. Isso leva a conjuntos de dados menores, o que limita a quantidade de informações disponíveis para treinar um modelo de rede neural de maneira adequada. Embora haja áreas como a detecção de microestruturas que possuem ampla base de dados, isso não é verdade para a maioria dos processos (Feng et al., 2019).

No trabalho de (Feng et al., 2019) considerou-se um pequeno *dataset* (447 pontos) para identificar defeitos metalúrgicos através de uma RNA. Foi proposto um pré treinamento

dos dados e isso gerou uma acurácia melhor, ou seja, os dados de teste apresentaram previsões mais próximas dos valores verdadeiros.

Pequenos *datasets* são problemas comuns a quase todas as áreas da ciência em problemas específicos. A partir do problema foram desenvolvidas técnicas para aprimorar as RNA's e obter melhores previsões. A seguir são apresentados alguns trabalhos gerais e também na área da soldagem.

Hong-Wei et al. (2015), através de uma RNA, promoveu um pré treinamento e ajuste fino de parâmetros da rede para melhorar as previsões de emoções humanas por meio de um banco de dados pequeno de imagens. Foi possível atingir melhores resultados ajustando os parâmetros da rede, sendo aproximadamente 20% melhor que os dados de predição usados para comparação.

Na pandemia global de COVID-19, o uso de RNA's para analisar imagens de radiografia de tórax (CXR) para diagnóstico de COVID-19 e triagem de pacientes se torna relevante no trabalho de Oh et al. (2020). A base de dados usada no início da pandemia era pequena. Para contornar o problema, utilizou-se de uma abordagem de rede neural convolucional baseada em *patches* com um número relativamente pequeno de parâmetros treináveis para o diagnóstico de COVID-19. O método proposto é inspirado em análise estatística dos potenciais biomarcadores de imagem das radiografias CXR. Resultados experimentais mostram que o método atinge desempenho bom e fornece mapas com biomarcadores clinicamente interpretáveis.

Tripicchio et al. (2020) exploram uma RNA para um sistema de monitoramento capaz de realizar verificações de qualidade precisas para detectar defeitos de soldagem em injetores de consumíveis durante a fase de produção. Uma estratégia de ponderação na matriz de confusão foi usada e permitiu obter 96,30% de acurácia no desempenho de detecção de defeitos mesmo no caso de conjuntos de dados pequenos e desbalanceados na fase de teste, comparando-se com resultados reais.

Pan et al. (2020) propuseram um novo modelo de aprendizado de transferência baseado em *MobileNet* como um extrator de características de defeitos de soldagem. Usando o conjunto de dados *ImageNet* (dados de defeitos sem soldagem) para pré-treinar um modelo *MobileNet*, e então foi possível enviar modelo *MobileNet* para o campo de classificação de defeitos de soldagem. Nos resultados da metodologia proposta em um conjunto de dados pequeno de defeitos de soldagem, obteve-se precisão da previsão de 97,69%.

Aumentar dados tem se mostrado uma metodologia eficaz para abordar o problema de dados desequilibrados em aplicações de manufatura. Dados desequilibrados referem-se a conjuntos de dados onde o número de instâncias de uma classe supera significativamente o número de instâncias de outra classe, o que muitas vezes leva a um desempenho subótimo dos modelos de aprendizado de máquina. Y. Kim et al. (2023) propuseram um método para

detectar anomalias usando aprendizado de representação auto supervisionado e modelos generativos profundos, demonstrando melhorias significativas em precisão, recall e *F1-score* após aplicar técnicas de aumento de dados com StyleGAN. De forma semelhante, T. Wang et al. (2023) desenvolveram um modelo de previsão de qualidade em sistemas *ciber* físicos humanos usando a Árvore de Decisão com Aumento de Dados e *Gradient Boosting* (DAGBDT), que mostrou redução no erro médio de previsão, facilitando decisões de otimização de produtos em sistemas de manufatura inteligente. Ruediger-Flore et al. (2023) demonstraram que o uso de dados sintéticos gerados a partir de CAD para treinar modelos de classificação de peças pode melhorar significativamente o desempenho de modelos de aprendizado profundo por meio de aprendizado por transferência, especialmente em cenários envolvendo produção de pequenos lotes ou produtos personalizados. Além disso, Dasari et al. (2022) abordaram o problema de desequilíbrio de classes em dados de processos de manufatura aditiva propondo um método de aumento de dados adaptativo baseado em clustering (CADA), que superou métodos tradicionais de *oversampling* como SMOTE e melhorou o desempenho do modelo na classe minoritária.

No contexto de soldagem, onde conjuntos de dados reais podem ser limitados, o aumento de dados visa melhorar a generalização do modelo ao introduzir cenários e variações diversas no conjunto de treinamento. Aumentar o conjunto de dados envolve sintetizar instâncias adicionais por meio de técnicas como jittering, escalonamento e rotação (Kim et al., 2023; Ruediger-Flore et al., 2023; T. Wang et al., 2023). O estudo de R. Yu et al. (2020) aborda o desafio de prever a penetração da junta de solda desenvolvendo um modelo baseado na superfície do poço de solda. Para lidar com o problema de conjuntos de dados pequenos, os autores propõem o uso de regressão por vetores de suporte (SVR) e a comparam com uma rede neural MLP (*Multi-Layer Perceptron*). Os resultados mostram que o modelo SVR otimizado oferece previsões mais precisas para a largura do cordão traseiro com uma quantidade relativamente pequena de dados de treinamento, melhorando a capacidade de prever a penetração da junta de solda. L. Li et al. (2024) propuseram um método para aumentar imagens de alta resolução de defeitos de soldagem ao fundir imagens reais e geradas, resultando em uma melhoria na média de precisão (mAP) para os modelos YOLO, proporcionando uma solução eficaz para a geração de amostras de dados em inspeções industriais. Jain et al. (2022) demonstraram que o uso de GANs para gerar imagens sintéticas de defeitos de superfície pode melhorar significativamente o desempenho de redes neurais convolucionais (CNNs) na classificação de defeitos, mostrando sensibilidade e especificidade superiores em comparação com métodos clássicos de aumento de dados. W. Yang et al. (2023) propuseram um algoritmo de atenção auto-supervisionada com gerador de rede adversarial auxiliar (ASSL-ACGAN), que superou métodos tradicionais como WGAN e

ACGAN em termos de qualidade de amostra e precisão na identificação de defeitos, sugerindo um esquema viável para identificação de defeitos em conjuntos de dados pequenos.

2.6.1 – *Leave One Out*

Quando não se pensa em técnicas de aumento de dados, há também a possibilidade de utilizar a validação cruzada para problemas de pequenos *datasets*. Dentre essas técnicas destaca-se para este trabalho o *Leave One Out* (LOO), sendo esse método caracterizado por se utilizar menos dados para teste e mais para treinamento no mesmo *dataset*. Embora a técnica apresente muitas vantagens e simplicidade de implementação em diversos tipos de problema, o trabalho de Gronau & Wagenmakers (2019) investiga através de três exemplos que o LOO não é uma solução para qualquer seleção de modelo, pois apresenta um suporte estatístico limitado para os modelos simples, ou seja, problemas para fornecer evidências e garantir a validade dos resultados.

A validação cruzada LOO é assim chamada porque, para cada iteração, ela deixa um único exemplo (observação) do conjunto de dados como conjunto de teste, e utiliza os demais exemplos como conjunto de treinamento. Em outras palavras, para um conjunto de dados com "N" exemplos, ela criará "N" divisões, cada uma com "N-1" exemplos no conjunto de treinamento e um único exemplo no conjunto de teste.

A validação cruzada LOO é uma técnica útil para avaliar o desempenho de modelos de aprendizado de máquina em conjuntos de dados pequenos, embora possa ser computacionalmente custosa e produzir resultados mais instáveis em comparação com outras técnicas de validação cruzada, como a validação cruzada *k-fold* (Wong, 2015).

2.6.2 – *Geração de Dados Sintéticos*

A técnica de geração de dados sintéticos, conhecida como *Synthetic Data Vault* (SDV), é uma abordagem que visa criar conjuntos de dados sintéticos com características semelhantes aos dados reais, preservando sua privacidade e confidencialidade, ou seja, os dados sintéticos são gerados para imitar as propriedades estatísticas dos dados reais sem revelar informações específicas sobre indivíduos ou entidades reais, por exemplo.

O SDV utiliza algoritmos e modelos estatísticos para aprender a estrutura dos dados reais e gerar novos dados que sigam essa mesma estrutura. Em vez de usar diretamente os dados reais, o SDV cria um *vault* (ou cofre) de dados sintéticos que podem ser usados para análises, testes, desenvolvimento de modelos e outras atividades sem expor as informações sensíveis dos dados originais (Patki et al., 2016).

O processo de geração de dados sintéticos pelo SDV envolve algumas etapas apresentadas na Figura 17.

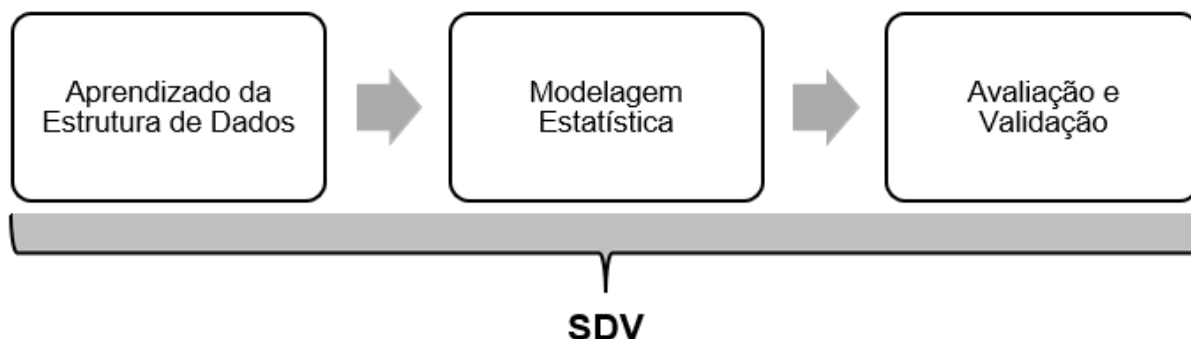


Figura 17. Fluxograma do processo de geração de dados SDV.

Na fase de aprendizado da estrutura dos dados o SDV analisa os dados reais para identificar padrões, distribuições, correlações e outras características importantes.

Com base na estrutura aprendida, a modelagem estatística do SDV utiliza modelos estatísticos, como redes generativas adversariais ou modelos de mistura gaussiana, para gerar novos dados que sigam a mesma estrutura.

Na avaliação e validação os dados sintéticos gerados são avaliados e validados para garantir que sejam estatisticamente similares aos dados reais e preservem sua qualidade e características importantes.

A geração de dados sintéticos pelo SDV apresenta algumas vantagens. Primeiro, ela permite que organizações compartilhem dados para fins de pesquisa, colaboração ou análise, sem expor informações confidenciais ou pessoais dos indivíduos envolvidos. Além disso, a geração de dados sintéticos pode ajudar a contornar problemas de escassez de dados reais ou problemas de acesso a dados sensíveis (Hittmeir et al., 2019).

No entanto, é importante observar que os dados sintéticos podem não capturar todas as nuances e detalhes dos dados reais, apresentando limitações em certos contextos. Portanto, a utilização de dados sintéticos deve ser cuidadosamente avaliada de acordo com a finalidade e contexto em que serão aplicados (Hittmeir et al., 2019; Patki et al., 2016).

A seguir são apresentados trabalhos usando o SDV para a geração de dados. Por ser uma técnica relativamente nova (2016) não há muitos estudos nas áreas de processos de fabricação.

K. Zhang et al. (2022) utilizaram o SDV para gerar dados sintéticos e comparar sua qualidade com um modelo baseado em rede contraditória generativa não sequencial existente chamado CTGAN. Para comparar os dados sintéticos sequenciais com sua contraparte real, criou-se uma nova métrica chamada *Multi-Sequence Aggregate Similarity* (MSAS). O modelo

SDV apresentou padrões de aprendizado de nível mais alto do que os modelos não sequenciais.

Hittmeir et al. (2019) produz um trabalho atestando a viabilidade do uso do SDV em dados públicos e privados, em que a privacidade do indivíduo deve ser preservada. O SDV produz dados sintéticos similares e condizentes aos dados reais, mas sem violar o sigilo para consultas públicas.

Kannan & Nandwana (2023) apresentaram uma abordagem genérica para acelerar a descoberta de ligas metálicas combinando cálculos CALPHAD (simulações de diagramas de fase) de alto rendimento, geração de dados sintéticos e mineração de dados. Os resultados sugeriram o projeto de um aço superbainítico que formam bainita a 200 °C em tempos de transformação mais baixos.

He et al. (2022) conduziram um estudo do efeito de geração sintética de dados em um caso de fadiga multiaxial com 5 *datasets*. Os resultados demonstraram que a geração de dados sintéticos melhorou o aprendizado nas técnicas de *machine learning* utilizadas para predição da vida do material.

Gregores Coto et al. (2023) apresentaram um estudo do uso de técnicas para auxílio na criação de novas ligas. A técnica SDV também é usada para criação de dados e comparada com dados reais de diagramas de fase usando a simulação CALPHAD. Os resultados demonstram que a técnica é promissora e permite a criação de inúmeras novas ligas a partir de ligas já desenvolvidas.

2.6.3 - Kriging

A técnica de *kriging*, também conhecida como interpolação *kriging* ou krigagem, é um método estatístico utilizado para estimar ou prever valores em locais não amostrados dentro de uma área geográfica. Foi criada por Daniel G. Krige e pelo matemático francês Georges Matheron. É frequentemente aplicado em problemas de interpolação espacial, onde os dados estão distribuídos de forma irregular no espaço e também é chamado de Processo Gaussiano de Regressão (Cressie, 1990).

O *kriging* baseia-se em modelos de variograma, que descrevem a variabilidade espacial dos dados. O variograma é calculado a partir das diferenças entre os valores observados em diferentes locais e fornece informações sobre a estrutura de dependência espacial dos dados (Oliver & Webster, 1990).

Existem diferentes métodos de *kriging*, como o *kriging* ordinário, *kriging* universal, *kriging* simples e *kriging* de bloco, cada um com suas próprias suposições e abordagens. Em geral, o *kriging* estima os valores desconhecidos ponderando as observações próximas aos locais de interesse, levando em consideração a correlação espacial medida pelo variograma.

A principal vantagem do *kriging* é sua capacidade de fornecer estimativas espacialmente suaves e coerentes, incorporando informações espaciais e evitando grandes variações abruptas nos valores estimados. Além disso, o *kriging* também fornece estimativas das incertezas associadas às previsões, o que é útil para avaliar a confiabilidade dos resultados (Oliver & Webster, 1990; Van Beers & Kleijnen, 2004).

No entanto, o *kriging* também apresenta algumas limitações. Ele pressupõe que os dados sigam um modelo estatístico específico (estacionariedade), e a qualidade das estimativas depende da precisão do modelo de variograma selecionado. Além disso, o *kriging* pode ser computacionalmente intensivo, especialmente quando aplicado a grandes conjuntos de dados (Hengl et al., 2007).

Os passos para realizar um *kriging* ordinário são a seguir apresentados baseados nos trabalhos de (Cressie, 1990; Hengl et al., 2007; Oliver & Webster, 1990; Van Beers & Kleijnen, 2004):

1. Inicialmente é criada uma função de semivariograma (uma representação visual da covariância exibida entre cada par de pontos nos dados amostrados) com base em pontos conhecidos para determinar a dependência espacial do modelo de autocorrelação. O semivariograma é matematicamente expresso por:

$$\gamma(d) = \frac{1}{2n} \sum_{i=1}^n \{Z(x_i) - Z(x_i + d)\}^2 \quad (17)$$

x_i é a localização do i -ésimo ponto, Z é o valor da elevação do ponto de controle, d é a distância entre os pontos de controle e n é o número de pares de pontos cuja distância é d .

A Figura 18 tem um variograma construído de acordo com as diferentes distâncias, d .

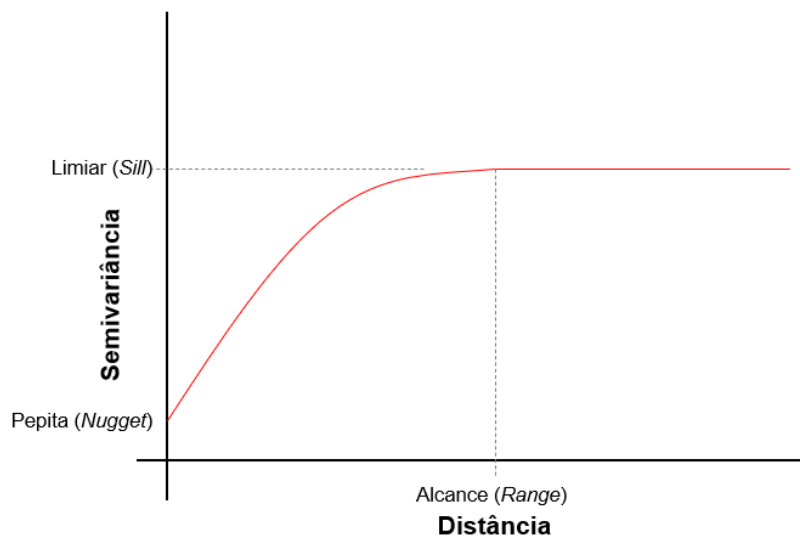


Figura 18. Representação de um semivariograma.

Os semivariogramas quantificam a autocorrelação porque representam graficamente a variância de todos os pares de dados de acordo com a distância. Isso significa que as chances são de que os pontos mais próximos estejam mais relacionados e tenham uma pequena semivariância, enquanto os pontos distantes estão menos relacionados e têm uma alta semivariância.

Porém, a uma certa distância alcance (*range*), a autocorrelação se torna independente. Onde essa variação se nivela, é chamada limiar (*sill*). Isso significa que não há mais nenhuma autocorrelação espacial ou relação entre a proximidade de seus pontos de dados. A pepita (*nugget*) é a interceptação y do variograma. Em termos práticos, o nugget representa a variabilidade em pequena escala dos dados.

2. O melhor modelo de variação teórica é selecionado de acordo com a figura do semivariograma. Os modelos teóricos comumente usados são o modelo esférico, o circular, o linear, o modelo exponencial e o modelo gaussiano, apresentados na Figura 19.

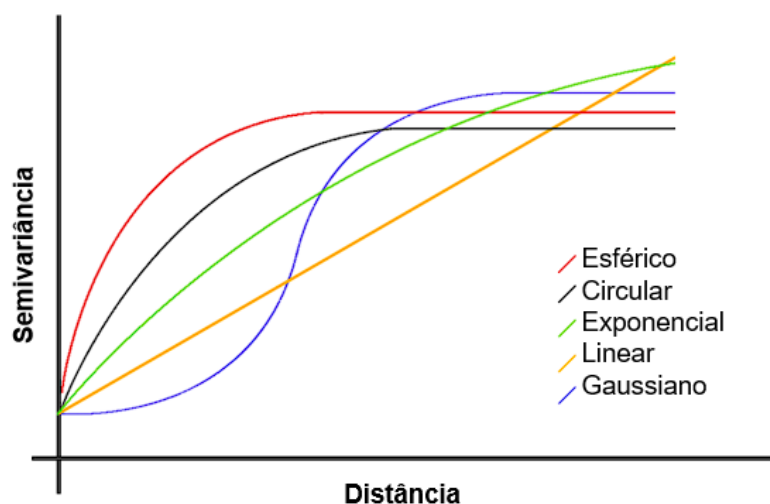


Figura 19. Modelos de semivariogramas. Fonte: Adaptado de <https://gisgeography.com/kriging-interpolation-prediction/>

3. O modelo de ajuste é usado para prever o valor de um ponto desconhecido da seguinte forma:

$$Z_p = \sum_{i=1}^n \omega_i Z_i \quad (18)$$

Z_i representa o valor do ponto i , ω_i os pesos do ponto i e Z_p o valor previsto. A maioria dos trabalhos envolvendo Kriging usa o método para predição de valores ao invés de geração de mais dados. Alguns desses trabalhos são apresentados a seguir.

P. Jiang et al. (2016) promoveram a otimização de parâmetros do processo de soldagem a laser combinando Kriging com algoritmo de otimização de enxame de partículas. A técnica foi usada para estabelecer a relação entre os parâmetros da poça de fusão em um aço galvanizado de 0,8 mm de espessura. Os resultados do modelo híbrido apresentaram efetividade melhor do que o padrão sem Kriging.

Gao et al. (2016) propuseram a otimização dos parâmetros da soldagem híbrida laser-arco, de topo, em aço inoxidável 316L usando o modelo Kriging e GA. O modelo Kriging foi usado para interpolar a relação de parâmetros da geometria da solda. Experimentos de verificação demonstram que os valores ótimos obtidos estão de acordo com os resultados experimentais.

Y. Yang et al. (2018) realizaram uma integração de dados por Kriging a partir de modelos de soldagem a laser de alta penetração de baixa e alta fidelidade obtidos com o método dos elementos finitos. O modelo previu a relação de parâmetros da poça de fusão e a distorção angular. Os resultados se mostraram promissores para identificar as relações propostas pelo Kriging.

No trabalho de Ruan et al. (2018), experimentos de soldagem a laser foram conduzidos em aço inoxidável 2205 em pontos de amostragem. As relações entre os parâmetros de soldagem e a largura da solda são construídas usando o modelo de krigagem estocástica. Outra técnica de validação cruzada foi usada para comparação (*k-fold*). A krigagem apresentou um erro quadrático médio menor em quatro dos cinco grupos dos dados testados.

Z. Liu & Wang (2022) propuseram um modelo de Kriging e um algoritmo evolutivo assistido para aumento de dados de vetores guiados em uma otimização multiobjetivo. O método evolutivo se baseava em passos anteriores da simulação para fornecer novos dados para uso na krigagem. Os resultados mostraram que as técnicas foram válidas para os problemas de otimização dinâmica multiobjetivo.

2.7 – Síntese de soluções

A partir da revisão bibliográfica feita este trabalho propõe uma metodologia inovadora para a determinação dos parâmetros de arquitetura de uma RNA utilizando um algoritmo de ED. A função objetivo dessa otimização é definida pelo erro percentual das variáveis de saída de um processo de soldagem MAG, garantindo que a RNA seja ajustada para fornecer previsões mais precisas.

Dada a limitação do conjunto de dados, a técnica SDV é empregada para a criação de dados artificiais, permitindo expandir o dataset e melhorar a capacidade preditiva do modelo. Além disso, a validação cruzada LOO é utilizada para avaliar o desempenho da RNA e evitar viés na análise dos resultados. Para garantir uma comparação abrangente, a otimização dos

parâmetros da rede também é realizada com o algoritmo SHGO e, na geração de dados artificiais, a técnica de Kriging é explorada como alternativa ao SDV.

A fim de verificar o desempenho preditivo da RNA, os resultados são comparados com modelos tradicionais de regressão, incluindo regressão linear, Ridge e polinomial. Para entender as relações entre os parâmetros de entrada e saída do processo de soldagem MAG, são realizados testes de correlação de Pearson e de distâncias, permitindo identificar quais variáveis possuem maior influência no processo.

Por fim, para mitigar o *overfitting*, os parâmetros da RNA são determinados durante a fase de validação e posteriormente utilizados na fase de testes, garantindo que o modelo generalize melhor para novos dados. Essa abordagem integrada possibilita uma análise robusta e eficaz para otimizar a predição de parâmetros no processo de soldagem MAG.

3 – METODOLOGIA

A metodologia proposta visa a determinação otimizada de parâmetros de uma RNA para prever variáveis do processo de soldagem MAG. Inicialmente, são coletados dados experimentais do processo de soldagem, incluindo tensão, corrente, velocidade de alimentação do arame, stickout e vazão de gás.

Devido à limitação do *dataset* original, aplicam-se técnicas de aumento de dados, como SDV e Kriging, para gerar novos pontos estatisticamente coerentes com a distribuição original e estimar valores intermediários. Após essa etapa, o conjunto de dados expandido é separado em três subconjuntos: treinamento, para ajuste dos pesos da RNA; validação, para seleção dos melhores hiperparâmetros; e teste, para avaliar a performance final do modelo e evitar *overfitting*.

A otimização da arquitetura da RNA é realizada utilizando duas técnicas distintas. A primeira é a Evolução Diferencial (ED) e a segunda é o SHGO (*Simplicial Homology Global Optimization*). Os hiperparâmetros ajustados incluem o número de camadas ocultas e neurônios, função de ativação e taxa de aprendizado. Com a RNA otimizada, o modelo é treinado e aplicado na predição dos parâmetros de soldagem, sendo seus resultados comparados com técnicas de regressão linear, polinomial e Ridge. Para garantir a generalização dos modelos, utiliza-se a validação cruzada *Leave-One-Out* (LOO), reduzindo a variabilidade estatística causada por conjuntos de dados pequenos.

A Figura 20 apresenta um fluxograma com a metodologia geral deste trabalho.

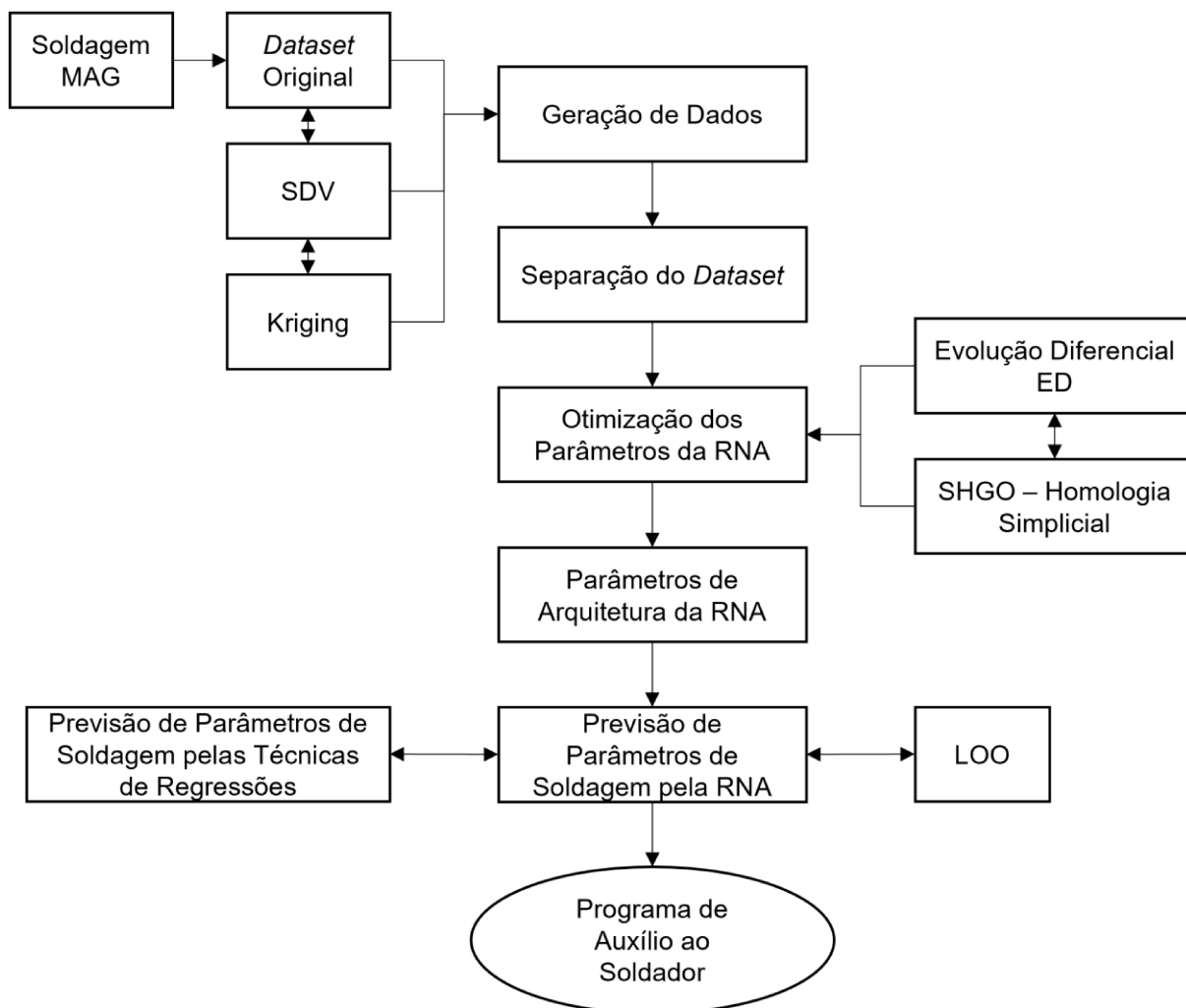


Figura 20. Fluxograma de apresentação da metodologia.

Por fim, os modelos treinados são integrados em um programa de auxílio ao soldador, permitindo que operadores e engenheiros obtenham recomendações automatizadas para ajuste dos parâmetros de soldagem com base nas condições de entrada do processo.

3.1 – Base de Dados

O problema de pesquisa da presente tese consiste em encontrar os parâmetros de inicialização de uma rede neural artificial que forneçam a melhor estimativa entre determinadas “entradas” e “saídas” relacionadas ao processo de soldagem MAG. Conforme mencionado, a literatura existente não aplica uma metodologia coerente para a utilização desses parâmetros, ou se utiliza da tentativa e erro.

A Tabela 2 mostra os dados e características do processo experimental realizado para produção dos corpos de prova que foi extraído do trabalho de Ribeiro (2018).

Tabela 2. Parâmetros experimentais de soldagem.

Parâmetros de Soldagem	
Processo de Soldagem	MAG
Material	SAE 1020
Espessura	3.18 - 9.52 mm (1/8" - 3/8")
Arame	AWS ER70S6
Chanfro	Reto - V (60°)
Equipamento	ESAB Smashweld 316
Gás de Proteção	CO ₂
Abertura de raiz	1 - 4.5 mm
Tipo de junta	Topo
Ângulo de soldagem	45°
Tipo de Corrente	Corrente direta com Polaridade invertida
Direção	Puxando

O SAE 1020 utilizado é um aço de baixo carbono com boa soldabilidade e ampla gama de aplicações na indústria. O arame AWS ER70S6 foi projetado para soldagem de aços macios e de baixa liga, sendo adequado para o SAE 1020. As chapas de aço foram cortadas em guilhotina, sendo o tipo de chanfro usinado. A Figura 21 mostra a fixação das placas, além das dimensões.

Em (a) é mostrado o gabarito de fixação e espaçamento entre as chapas de aço. Em (b) é mostrado o chanfro em V, enquanto em (c) estão as dimensões. A abertura da raiz, o comprimento da chapa e a espessura variam na faixa exibida na Figura 20.

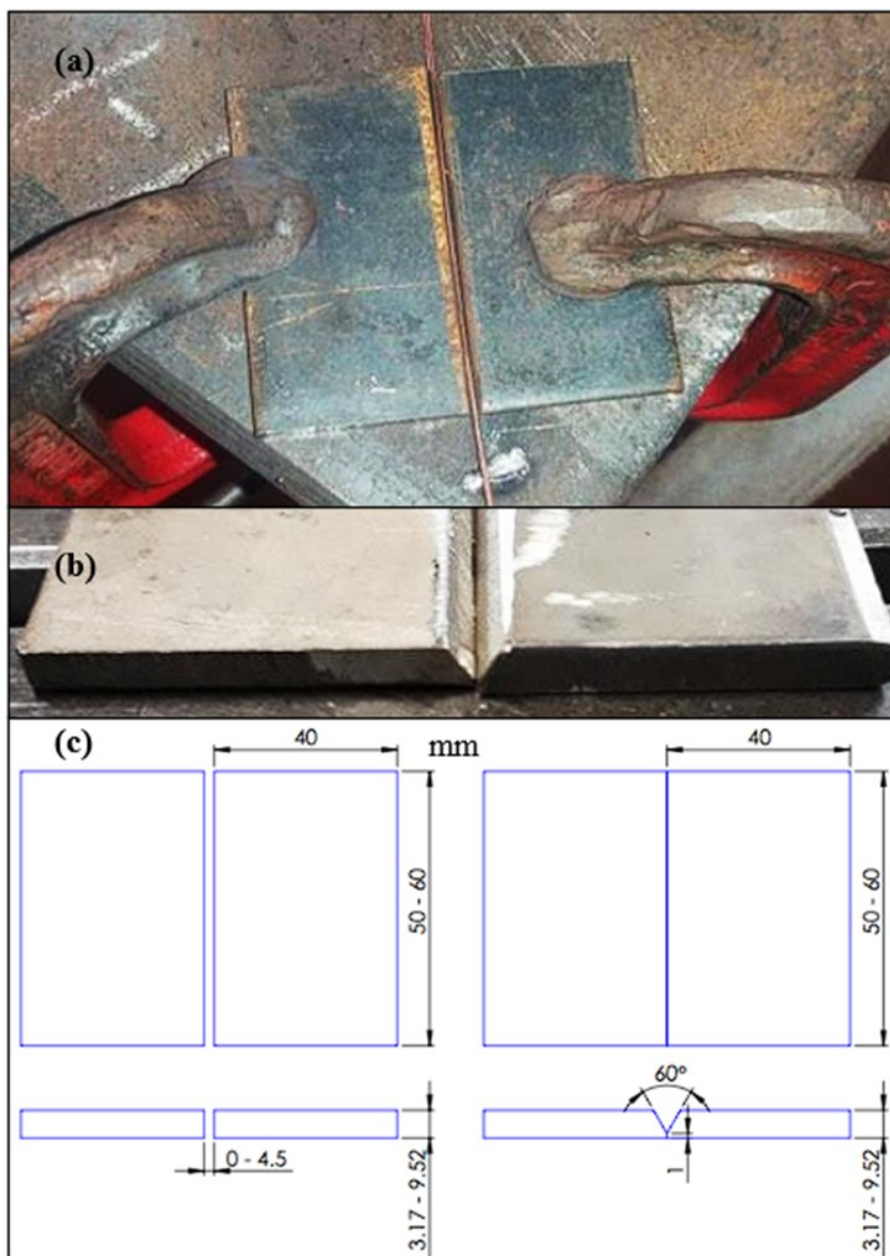


Figura 21. Preparação de amostras. a) Gabarito de raiz de solda. b) Chanfro V. c) Medidas dos corpos de prova. Fonte: Adaptado de Ribeiro (2018)

A Tabela 3 apresenta os dados de entrada e saída considerados. As entradas dizem respeito à condição da junta (espessura da chapa, tipo de chanfro, afastamento entre peças, altura do cordão e largura do cordão), enquanto as saídas são os parâmetros do processo (vazão do gás, tensão, velocidade de alimentação do arame e *stickout*). Logo, almeja-se que o soldador, ao especificar as características da junta, seja capaz de selecionar os melhores parâmetros, com auxílio de uma RNA. Para o desenvolvimento do trabalho foi utilizada a série de dados de Ribeiro (2018).

Tabela 3. Parâmetros de Entrada e Saída.

Parâmetros de Entrada	Parâmetros de Saída
Espessura da Chapa (mm)	Vazão do Gás (L/min)
Tipo de Chanfro	Tensão (V)
Afastamento entre Peças (mm)	Velocidade de Alimentação do
	Arame (m/min)
Altura do Cordão (mm)	<i>Stickout</i> (mm)
Largura do Cordão (mm)	

O processo de soldagem GMAW possui uma grande variedade de parâmetros. Os parâmetros de saída mencionados foram selecionados com base no seu grau de influência sobre os parâmetros de entrada, sendo descartado apenas a corrente por sua relação direta com a tensão.

Os parâmetros geométricos de entrada são mostrados na Figura 22.

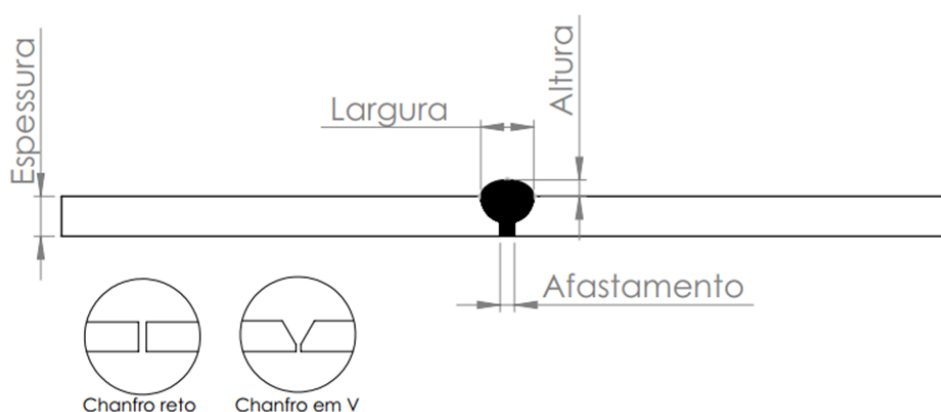


Figura 22. Parâmetros geométricos de entrada.

3.2 – Escolha de Parâmetros

Com os recursos do software PAST, uma ferramenta estatística com ampla gama de recursos estatísticos e gráficos para explorar e interpretar dados, sendo também gratuito e de código aberto, foram realizadas diversas análises sobre as amostras da base de dados para melhor seleção de parâmetros.

Inicialmente tem-se a plotagem de todas variáveis de entrada e saída no estilo *pair plot* para verificação visual de alguma possível relação. Para mensurar as possíveis relações não lineares foi conduzido um teste de correlação de distâncias entre parâmetros de entrada e saída. Pelo meio da biblioteca *scipy* e a função *spatial.distance.correlation* os parâmetros de

correlação foram determinados e abordados em forma tabular tanto para o conjunto de dados original e os aumentados.

3.3 – Geração de Dados

Pela implementação da técnica de Kriging e SDV na linguagem *Python* foram criados 50%, 100% e 150% mais dados para o *dataset* original. A partir dos testes realizados na RNA (abordados adiante, os erros foram comparados e apresentados, e foi definida a técnica com melhor desempenho para avaliar o comportamento dos erros de predição através do aumento da geração de dados.

A biblioteca *sklearn* e módulo *GaussianProcessRegressor* foram utilizados para geração de dados por Kriging. A biblioteca *sdv* e o módulo *GaussianCopulaSynthesizer* foram utilizados no caso de geração de dados sintéticos SDV.

A Tabela 4 mostra a divisão e aumento do *dataset* para as técnicas mencionadas.

Tabela 4. Incremento de dados.

Tamanho Dataset	Técnica	Kriging			SDV		
	Condição	1,5	2,0	2,5	1,5	2,0	2,5
	Original	44	44	44	44	44	44
	Incremento	22	44	66	22	44	66
	Total	66	88	110	66	88	110

Este trabalho foca em aumentar o conjunto de dados em até 2,5 vezes o seu tamanho original. Embora aumentar o conjunto de dados possa melhorar o desempenho e a robustez do modelo, aumentá-lo além desse limite pode também introduzir um risco maior de overfitting e criar diferentes relações entre os parâmetros. O aumento excessivo de dados pode levar o modelo a aprender com ruídos e padrões irrelevantes, reduzindo assim sua capacidade de se generalizar bem para novos dados não vistos. Além disso, a multicolinearidade dos parâmetros pode levar a erros maiores na regressão (Aviral Gupta et al., 2017; Kroll & Song, 2013).

3.4 – Separação de Amostra

3.4.1 – Caso Padrão de Otimização

A partir dos dados fornecidos pelo trabalho de Ribeiro (2018) foi possível desenvolver um *dataset* em formato *csv* para ser implementado na linguagem *Python*. Com o auxílio da

ferramenta *Google Colab*, os dados foram separados em entradas e saídas para a rede neural. Os dados foram separados em teste e validação (20%) e treinamento (80%) para a RNA. A seleção dos dados foi realizada de forma aleatória pela biblioteca do *Python sklearn* e o recurso *train_test_split*. A Tabela 5 exibe a divisão do conjunto de dados utilizados no treinamento e nos testes.

Tabela 5. Divisão do *Dataset* em Treinamento e Testes da RNA

Quantidade de Conjunto de Dados	Amostras para Treinamento	Amostras para Validação	Amostras para Testes
44	35	4	5

Quando um modelo de aprendizado de máquina é treinado, é essencial avaliar seu desempenho em um conjunto de dados de teste separado que não foi usado durante a fase de treinamento. Isso ajuda a garantir que o modelo se generalize bem para dados não vistos, proporcionando uma medida realista de seu desempenho e prevenindo problemas como o sobreajuste. A fase de teste serve como um ponto de referência objetivo para verificar que a metodologia é sólida e que as previsões do modelo são confiáveis quando aplicadas a novos dados (Bishop & Nasrabadi, 2006; Hastie et al., 2009).

Mediante a separação dos dados em treinamento, validação e teste é possível padronizar quais conjuntos de dados sempre serão usados para o algoritmo desenvolvido no trabalho e ainda prevenir o *overfitting* encontrando os parâmetros da RNA na fase de validação e fixando esses parâmetros na fase de testes. Dessa forma, elimina-se a possibilidade de grande variação dos resultados devido a testes em amostras diferentes. A Figura 23 apresenta uma representação da separação de dados dessa fase.

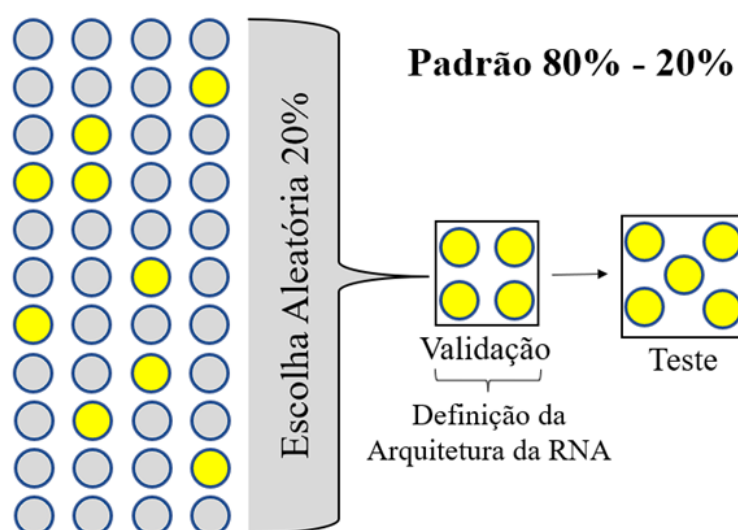


Figura 23. Separação de dados 20% para treinamento e validação.

3.4.2 – Caso *Leave One Out* (LOO)

A separação dos dados para a técnica de validação cruzada foi adotada separando-se 9 amostras individuais do *dataset* para teste de cada uma e posterior inferência estatística sobre os resultados. Essas amostras foram as mesmas adotadas na etapa padrão, para gerar comparabilidade. A separação para a LOO é representada na Figura 24.

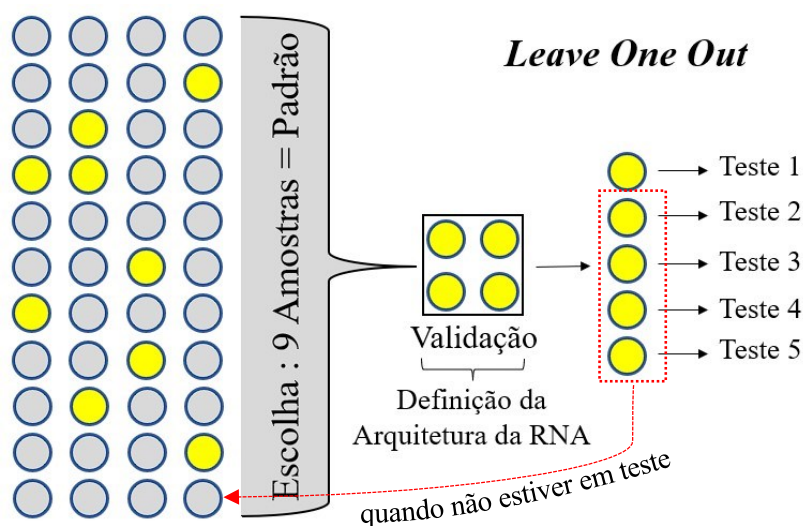


Figura 24. Separação de dados para LOO.

Na fase de validação são feitas 5 determinações de parâmetros de arquitetura, pois enquanto um teste será feito na fase posterior, os testes restantes integram o treinamento, até totalizar todas as opções possíveis. Cada amostra é testada no algoritmo desenvolvido individualmente para posterior cálculo da média.

3.4.3 – Caso *Geração de Dados*

Com a geração de mais dados através do Kriging e SDV foram mantidas as mesmas amostras para testes e validação, para que a comparação dos métodos seja coerente. A Figura 25 mostra a separação de dados para os casos de aumento de dados.

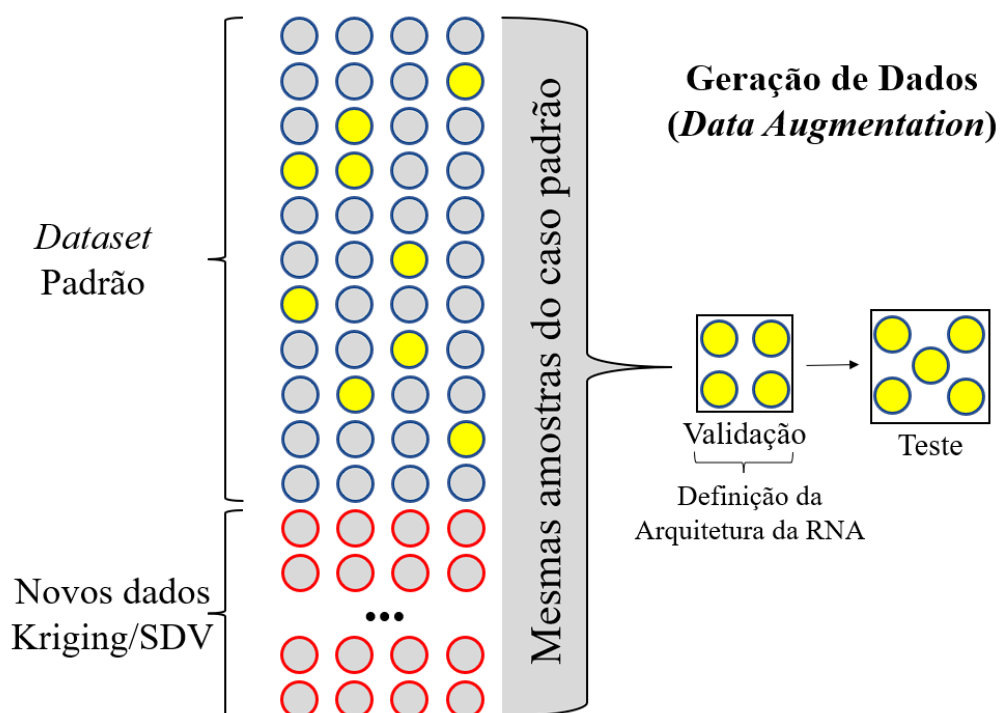


Figura 25. Separação de dados com aumento de dados por Kriging ou SDV.

3.5 – Arquitetura da RNA

No trabalho de Ribeiro (2018), observou-se que a estrutura utilizada para a arquitetura da RNA era composta por duas camadas ocultas com 10 ou 20 neurônios em cada e funções de ativação “sigmoide” e “linear”. Já no trabalho de Moghaddam & Kolahan (2021), até três camadas ocultas foram utilizadas, com até cinco neurônios, porém a função de ativação não foi mencionada.

Com base nos trabalhos citados acima, verificou-se a necessidade de se expandir o estudo das camadas, neurônios e função de ativação no presente trabalho. A Figura 26 apresenta, de forma esquemática, a RNA proposta para a resolução do problema de soldagem com todos os parâmetros definidos com base na possibilidade de controle usando-se a biblioteca *scikit-learn* e o módulo *MLPRegressor* do *Python*.

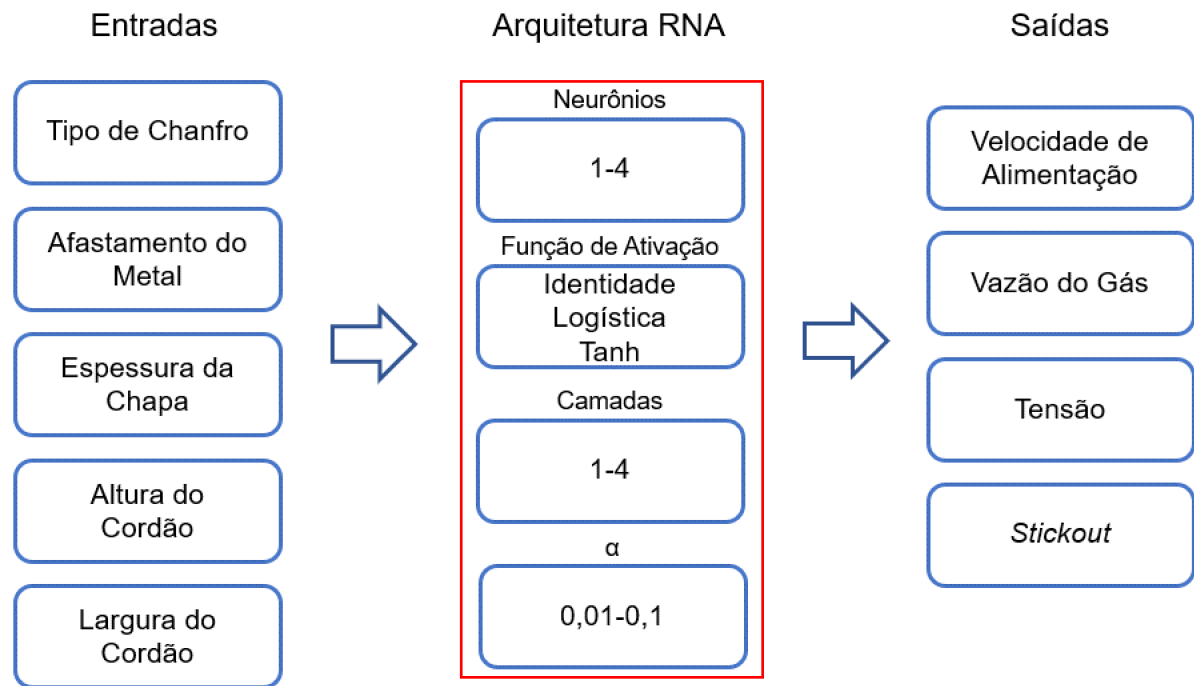
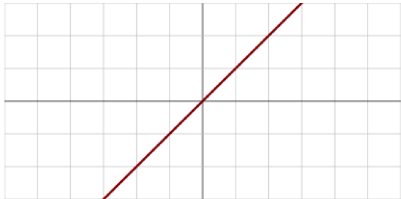
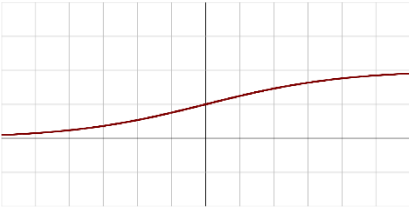
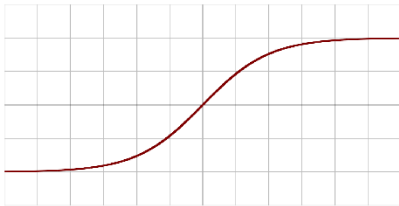


Figura 26. Proposta para a resolução do problema de soldagem.

As funções matemáticas que representam as funções de ativação escolhidas, assim como a forma e faixa de atuação são mostradas na Tabela 6.

Tabela 6. Funções de Ativação Consideradas.

Nome	Forma	Função, $g(x)$	Faixa
Identidade, Linear		x	$(-\infty, +\infty)$
Logística, sigmoid		$\sigma(x) \doteq \frac{1}{1 + e^{-x}}$	$(0,1)$
Tanh, Tansig		$\tanh(x) \doteq \frac{e^x - e^{-x}}{e^x + e^{-x}}$	$(-1,1)$

As funções de ativação foram definidas também de acordo com a biblioteca da linguagem Python utilizada no trabalho, e será abordada a seguir.

A biblioteca selecionada para o algoritmo da RNA possui a denominação *sklearn*. Ela possui o módulo *neuralnetworks*, que engloba alguns algoritmos para uso em predições de parâmetros.

O algoritmo utilizado é o *MLPRegressor*, uma rede neural que utiliza o *perceptron* multicamadas juntamente com a otimização de pesos disponível. Esse algoritmo ainda possui a opção de utilizar diferentes funções de ativação, taxa de aprendizagem constante, adaptativa e decrescente. Também permite o controle do número máximo de iterações, tolerância numérica para evolução, e critério de parada nas iterações previamente definidas.

Outros parâmetros podem ser utilizados na parametrização da RNA, em virtude da combinação do método de otimização dos pesos usados e da função de ativação. A Tabela 7 apresenta as opções disponíveis e os valores considerados padrões para a biblioteca.

Tabela 7. Parâmetros disponíveis na biblioteca *sklearn*, utilizada para construir a RNA.

Atributo	Padrão	Uso	Atributo	Padrão	Uso
hidden_layer_sizes	(-100)	Controlado	Tol	0,0001	Controlado
activation	relu	Controlado	verbose	False	-
solver	adam	Controlado	warm_start	False	-
alpha	0,0001	Controlado	momentum	0,9	X
batch_size	auto	X	early_stopping	False	X
learning_rate	constant	X	validation_fraction	0,1	X
learning_rate_init	0,001	X	beta_1	0,9	X
power_t	0,5	X	beta_2	0,999	X
max_iter	200	Controlado	epsilon	1E-08	X
shuffle	True	X	n_iter_no_change	10	X
random_state	None	Controlado	max_fun	15000	-

Acima, o termo “controlado” significa que o parâmetro será controlado pelo algoritmo desenvolvido no trabalho, “X” presume-se que a variável não será utilizada no momento e “-” significa que o valor padrão é escolhido.

Com essa implementação de código é possível variar os parâmetros da arquitetura da rede neural conforme proposto. Os resultados gerados utilizam dos dados separados para teste, e a partir das predições realizadas, calcula um erro com base no valor real (referência) de cada variável. Para que todos os erros sejam comparáveis é utilizada a função do erro absoluto médio percentual (*mean absolute percentagem error – MAPE*), dado por:

$$MAPE(y, \hat{y}) = \frac{1}{n_{amostras}} \sum_{i=0}^{n_{amostras}-1} \frac{|y_i - \hat{y}_i|}{\max(\epsilon, |y_i|)} \quad (14)$$

É então calculado o erro percentual para cada saída, assim como o erro total, ou seja, a soma de todos os erros. O algoritmo de evolução diferencial tem como objetivo a minimização desse erro.

3.6 – Algoritmo de Otimização

Para a otimização dos parâmetros da arquitetura da RNA foi utilizado um algoritmo de Evolução Diferencial (ED) em que se gerou uma população para os parâmetros nas etapas de mutação, cruzamento e seleção. A função objetivo da otimização foi definida para minimizar a somatória dos erros associados às previsões feitas para cada parâmetro de saída da RNA.

A razão de se utilizar um algoritmo de otimização está relacionada com a busca de valores ótimos para os parâmetros da RNA através da minimização do erro total somado dos parâmetros de saída, ou E_t . A Figura 27 exhibe o procedimento de forma esquemática.

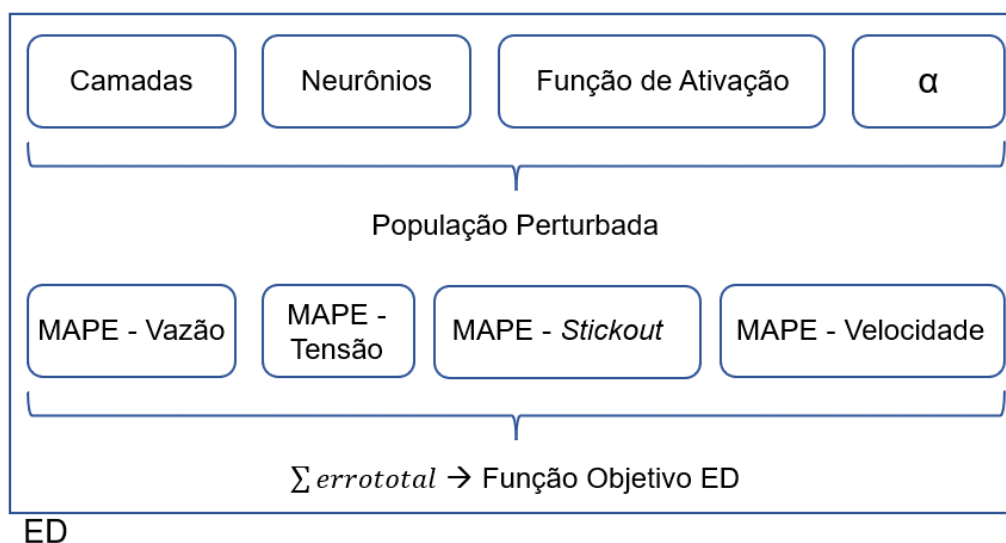


Figura 27. Definição das variáveis da população perturbada e variáveis objetivo da otimização.

As soluções candidatas, representadas pelos indivíduos, foram geradas com base em uma semente comum para repetitividade e comparabilidade dos resultados. A implementação da ED segue passos comuns as técnicas de otimizações baseadas em algoritmos genéticos que são apresentados no fluxograma da Figura 28.

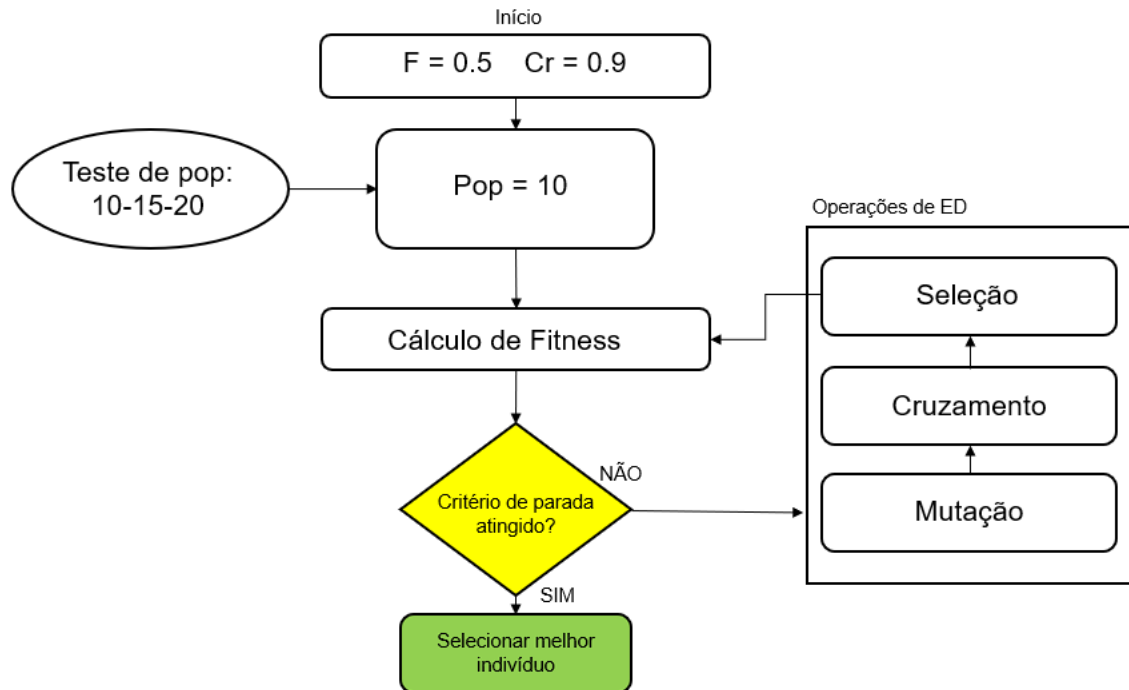


Figura 28. Par metros da ED.

O teste de popula  o representado na figura 28   uma fase anterior para defini  o experimental de par metros para a ED. S o realizados testes com o *dataset* padr o em 50 itera  es com popula  es iniciais da ED em 10, 15 e 20. A partir dos resultados do erro e tempo computacional gasto, uma pondera  o do resultado foi realizada utilizando-se a popula  o de 10 para as demais etapas. Como estrat gia adotada na ED utiliza-se da mostrada na Tabela 8.

Tabela 8. Nota  o da estrat gia da ED.

Muta��o	Nota��o
$V^{q+1} = X_{best}^q + F(X_{\beta}^q - X_{\gamma}^q)$	ED/best/1/bin

Em que *best* representa o vetor que ser  perturbado, ou seja, o vetor de menor custo da popula  o, 1   a quantidade de diferen as ponderadas para a perturba  o de *best*, e *bin* representa o cruzamento do tipo binomial.

3.7 – Otimiza  o de Pesos Sin pticos

A taxa de aprendizagem, ou α ,   objeto de estudo e implementa  es desde a cria  o de RNA's. Entretanto, a maioria dos trabalhos n o possuem uma metodologia para determina  o desse valor, ficando geralmente entre 0 e 1 (Kanti & Rao, 2008).

Sendo α um parâmetro contínuo, se justifica o uso da otimização por ED e não simplesmente uma análise combinatória. A biblioteca em *Python* utiliza um valor padrão para α de 0,001. Entretanto, após testes prévios no algoritmo definiu-se um intervalo [0,01 - 0,1] para que a ED atue. Para fins de comparação da ED se adota também um algoritmo SHGO com 50 iterações e mesmos limites da ED.

3.8 – Algoritmo Completo

Após estudo dos métodos necessários para otimização dos parâmetros da arquitetura da rede neural e dos pesos sinápticos, é necessário acoplá-los em um mesmo algoritmo para geração de resultados. O fluxograma da Figura 29 apresenta a estruturação dos algoritmos de otimização e RNA desenvolvidos no presente trabalho.

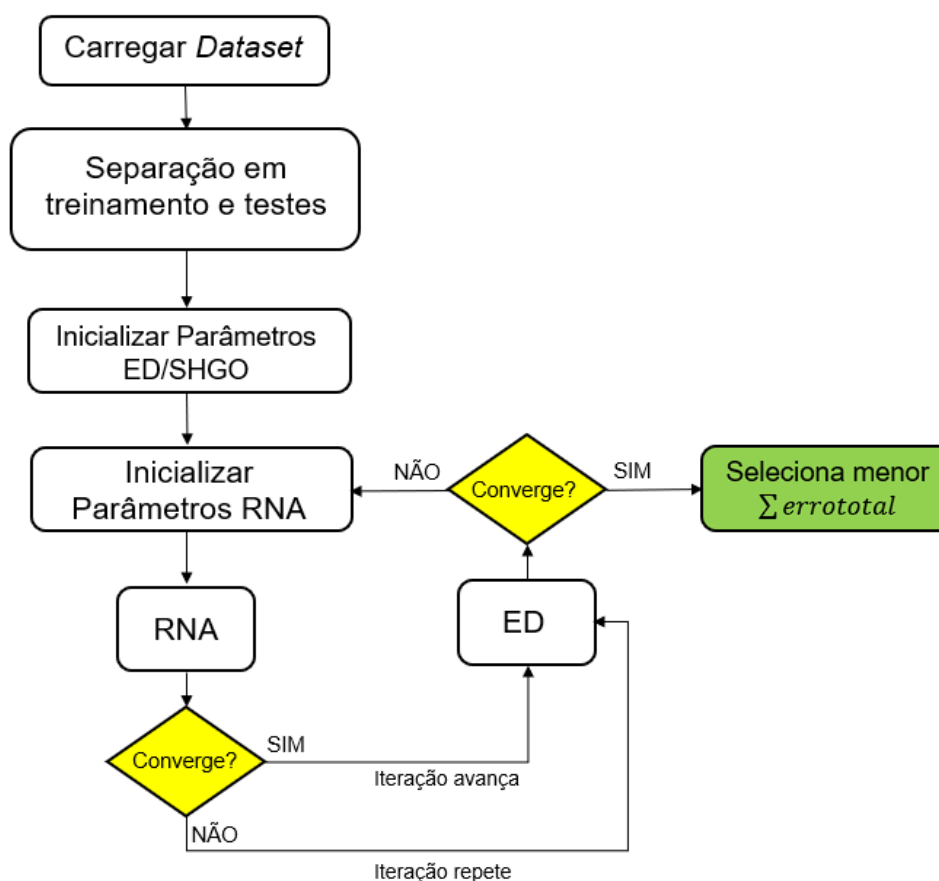


Figura 29. Fluxograma do algoritmo proposto.

A partir do fluxograma apresentado é possível identificar a relação estabelecida entre a RNA, ED (ou SHGO) e otimização de pesos. Após a etapa de carregamento e separação dos dados, feita com o recurso *train_test_split*, iniciam-se os parâmetros da ED como apresentados na figura 13, e a partir disso se inicia a RNA com os parâmetros mostrados na

tabela 7, em que o número de iterações é variado para análise de convergência. Sendo assim, um laço para integração dos algoritmos é iniciado entre RNA e ED, até que ocorra a convergência dos resultados da minimização do erro total pela ED.

Quando o resultado da RNA não converge, é feita uma repetição da iteração definida pela ED. Caso ocorra a convergência, a iteração avança para a próxima. Assim, como o melhor resultado do E_t é gravado no fim da simulação realizada, há também a coleta de outras variáveis como: *MAPE* (Vazão, Tensão Velocidade de Alimentação e *Stickout*), α , Camadas, Neurônios e Função de Ativação. O próximo tópico mostra os resultados obtidos utilizando-se análises gráficas do algoritmo para o problema de soldagem em questão.

3.9 – Padronização de Resultados

Para gerar comparabilidade de resultados, é definido o seguinte:

- Os mesmos pontos de teste foram definidos para todos os casos;
- Todos os dados de validação de testes são os mesmos para cada condição testada;
- A semente foi fixada em todos os casos para a RNA e para a ED;
- O tempo computacional é calculado com o mesmo gasto de memória e processamento pela ferramenta *Google Colab* em 1,2 Gb.

3.10 – Regressão

Todas as condições de *dataset* apresentadas na Tabela 4 são então testadas em código *Python* utilizando técnicas de regressão: linear, múltipla de Ridge e polinomial. A biblioteca utilizada foi *sklearn.preprocessing* e *sklearn.linear_model*. O coeficiente de determinação, R^2 , é então determinado para indicar a proporção da variância no parâmetro dependente que é previsível a partir da variável independente.

3.11 – Programa de Auxílio ao Soldador

Utilizando a biblioteca *tkinter* do *Python* é então desenvolvido e apresentado um programa que como entrada são inseridos os dados geométricos da solda e como saída se apresenta os parâmetros previstos com a melhor arquitetura de RNA encontrada no *dataset* original. Da mesma forma, apresenta-se um programa que aumenta os dados sinteticamente para calcular as mesmas saídas.

Por fim, é desenvolvido um programa para apresentar a melhor arquitetura encontrada para uso de parâmetros na RNA apresentada.

4 – RESULTADOS E DISCUSSÃO

4.1 – Escolha dos Parâmetros de Soldagem

A partir de uma abordagem estatística e teórica são apresentados os critérios de escolha dos parâmetros de entrada e saída da RNA.

4.1.1 – Abordagem Estatística

Antes de realizar a análise de correlação, a Figura 30 mostra visualmente a relação entre as variáveis de entrada e saída através de um *pair plot*.

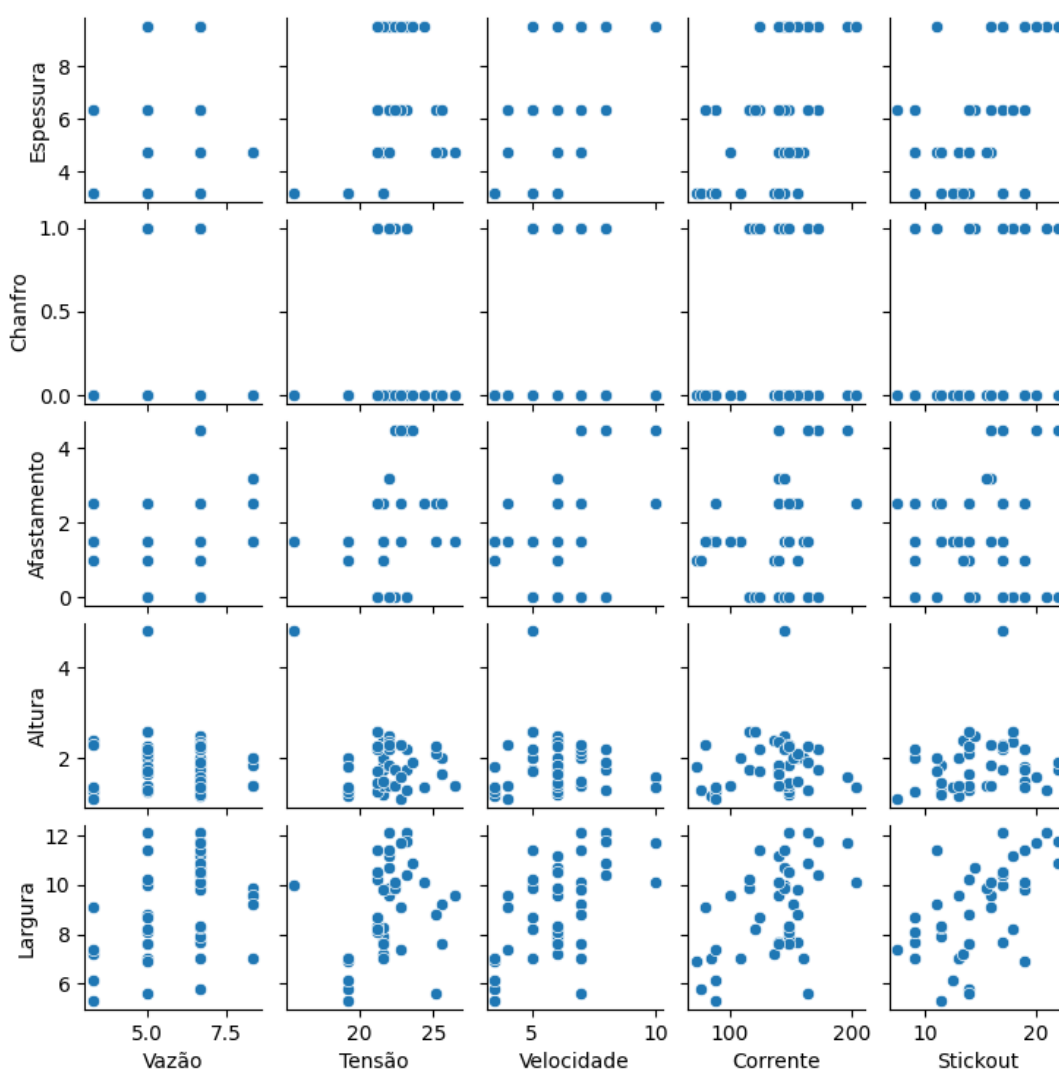


Figura 30. *Pair plot* das variáveis de entrada e saída.

A partir da análise visual da relação entre as variáveis não é possível identificar uma possível correlação, sendo então necessário recorrer aos testes estatísticos para verificar. Com o software PAST foi possível plotar um gráfico visual que representa o nível das correlações entre todas as variáveis de entrada e saída. Esses resultados são apresentados na Figura 31. Quanto maior o tamanho da elipse, maior é o valor de correlação. A cor azul indica uma correlação positiva, enquanto a cor vermelha, negativa, respectivamente.

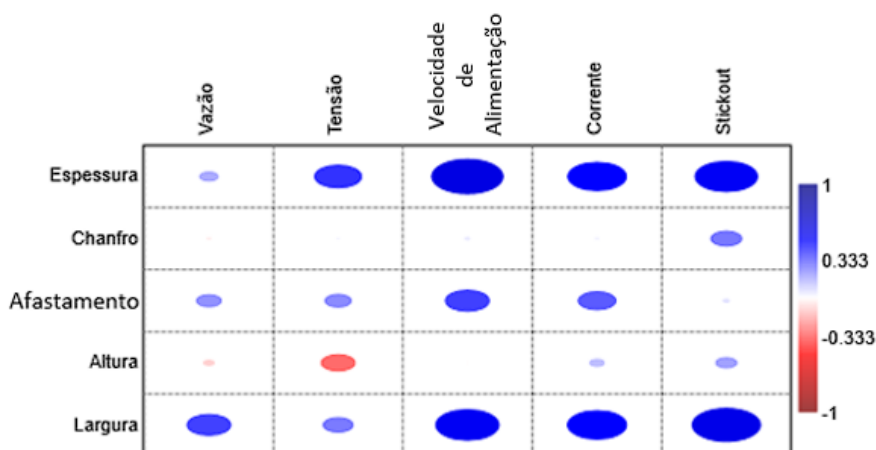


Figura 31. Gráfico de correlação Pearson.

As maiores correlações positivas apresentadas entre as variáveis são apresentadas por: Velocidade de Alimentação x Espessura, Corrente x Espessura, Tensão x Espessura, Corrente x Largura, *Stickout* x Espessura e *Stickout* x Largura.

A tensão e a corrente se relacionam diretamente através do ajuste da tensão na máquina, em que uma depende diretamente da outra, sendo um fator justificativo para utilização da tensão por ser o parâmetro passível de ajuste pelo soldador (Modenesi *et al.*, 2005).

Como não é possível apenas prever que a relações entre os parâmetros analisados sejam apenas lineares torna-se necessária a aplicação de outro teste que consiga avaliar possíveis correlações não lineares também.

A análise de correlação de distâncias é um recurso recente e que pode mensurar relações lineares e não lineares em amostras. A Figura 32 apresenta os resultados dos testes de correlação realizados na linguagem *Python* e transferidos para uma forma numérica e visual para melhor entendimento do conjunto de dados original.

Entrada/Saída	Vazão	Tensão	Velocidade	Corrente	Stickout
Espessura	0,841	0,599	0,391	0,501	0,467
Chanfro	1,029	0,984	0,962	0,978	0,737
Afastamento	0,787	0,774	0,626	0,679	0,948
Altura	1,087	1,289	1,004	0,877	0,819
Largura	0,628	0,743	0,460	0,496	0,416



Correlação Positiva
Correlação Negativa

Figura 32. Teste de Correlação de Distâncias para o *dataset* original.

A partir do teste de distâncias realizado é possível identificar que o parâmetro com menor correlação apresentado é o chanfro. Isso mostra o efeito do tipo de variável (qualitativa) em relação às demais variáveis. É encontrada uma maior correlação entre o chanfro e o *stickout*. Dentre as demais variáveis, é possível verificar correlações possíveis na maioria delas.

Comparando o teste de Pearson e o teste de distâncias não se identificam diferenças nas correlações entre os testes. Isso sugere que as relações entre as amostras podem ser consideradas lineares e que o teste de correlação de Pearson bastaria como justificativa para a base de dados de soldagem utilizada.

Para garantir a qualidade dos dados aumentados, o teste de correlação de distância também é realizado para o caso SDV e Kriging. A Figura 33 apresenta essa análise para SDV e a Figura 34 para Kriging.

Entrada/Saída	Vazão	Tensão	Velocidade	Corrente	Stickout
Espessura	1,026	0,609	0,448	0,541	0,440
Chanfro	1,084	1,007	0,881	1,021	0,744
Afastamento	0,700	0,853	0,869	0,701	1,027
Altura	1,107	1,257	0,877	0,913	0,752
Largura	0,752	0,736	0,516	0,521	0,340



Correlação Positiva
Correlação Negativa

Figura 33. Teste de Correlação de Distâncias para o *dataset* gerado com SDV.

As correlações da Figura 33 apresentam valores similares em comparação ao conjunto de dados original. Isso indica que as relações foram mantidas no novo conjunto de dados, dando uma boa condição para evitar *overfitting* nos dados aumentados.

Entrada/Saída	Vazão	Tensão	Velocidade	Corrente	Stickout
Espessura	0,977	1,102	0,828	0,822	0,863
Chanfro	1,327	1,413	1,364	1,217	1,127
Afastamento	0,718	0,673	0,504	0,559	0,725
Altura	1,001	0,743	0,767	0,711	0,895
Largura	0,548	0,716	0,426	0,509	0,430



Correlação Positiva
Correlação Negativa

Figura 34. Teste de Correlação de Distâncias para o *dataset* gerado com Kriging.

Na Figura 34 é possível ver que algumas correlações foram alteradas do conjunto de dados original. Esse efeito é observado principalmente no tipo de chanfro e espessura. As outras correlações são semelhantes aos dados originais. A mudança observada pode indicar desafios para essa técnica em garantir as propriedades dos dados originais e pode levar a possíveis resultados diferentes nas previsões.

4.2 – Evolução Diferencial

Na literatura não há um consenso sobre a definição do fator de escala ou fator de perturbação, F , sendo que de acordo com o trabalho de Storn & Price (1996) recomenda iniciar o parâmetro com o valor de 0,5. Sendo assim, as simulações foram iniciadas com o mesmo valor.

A taxa de cruzamento, Cr , possui limites entre zero e um, e representa a probabilidade do vetor experimental herdar os valores das componentes do vetor doador. Adota-se o valor de 0,9 por meio de simulações prévias realizadas no algoritmo, em que valores menores que 0,6 apresentaram os piores resultados e os maiores que 0,7 os melhores.

O tamanho da população (n), refere-se à quantidade de indivíduos aleatoriamente gerados para os vetores. Um valor pequeno gera poucas possibilidades de cruzamento, enquanto valores grandes implica em mais possibilidades, mas apresenta custo computacional maior e nem sempre chega em resultados melhores. O tamanho da população foi testado em 10, 15 e 20, com resultados apresentados na próxima sessão.

Como critério de parada foi adotado o número de iterações em 50, sendo esse parâmetro definido através de testes no algoritmo.

A estratégia “ED/best/1/bin” foi definida devido à sua simplicidade e eficiência comprovada por outros trabalhos como otimização paramétrica de imagens de soldagem em Cruz et al. (2021) e investigação de parâmetros de soldagem por atrito nos trabalhos de Srichok et al. (2020) e Luesak et al. (2023).

4.2.1 – Teste de População

Com a variação da população da ED em 10, 15 e 20 foram feitas 5000 iterações da RNA. A Figura 35 apresenta os resultados com o $MAPE$ ($\sum erro_{total}$ ou E_t) e o tempo computacional gasto em cada condição.

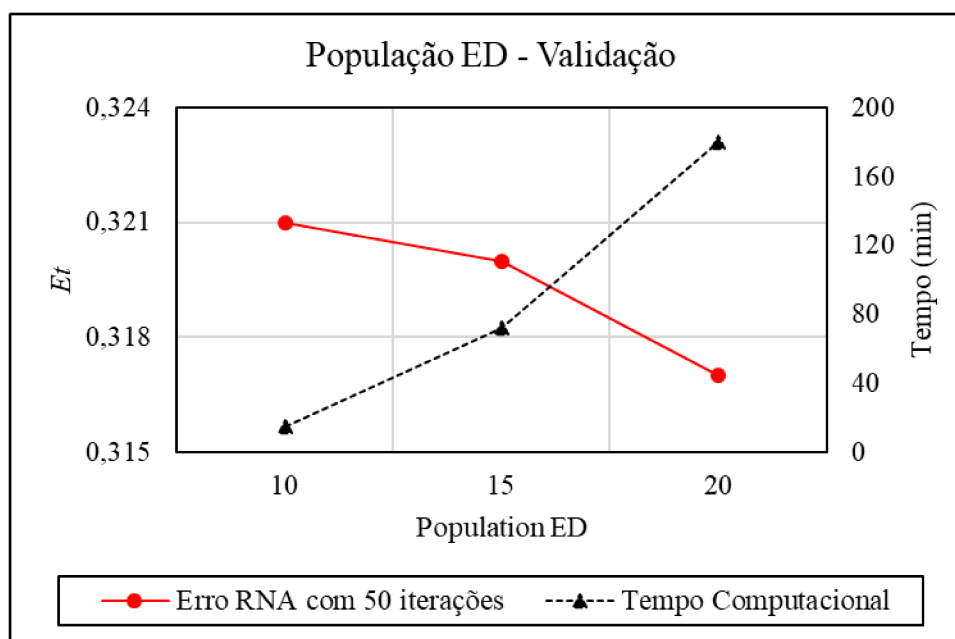


Figura 35. Avaliação da RNA com teste de população na ED na validação. A população mais adequada é circulada (10).

A diferença do E_t mais baixo para o mais alto é de apenas 0,004 enquanto a diferença do melhor tempo (15 minutos) para o mais alto (180 minutos) é de 165 minutos. A partir de uma ponderação de custo/benefício do valor do erro e tempo computacional gasto é então escolhido o valor da população de 10 para os demais testes na sequência do trabalho.

É interessante notar que houve convergência dos resultados tanto no valor do erro da RNA quanto no tempo computacional. Isso demonstra o que foi constatado pelo método proposto por Storn & Price (1996) em que um número maior de população gera também um gasto computacional maior. Pelo trabalho de Mallipeddi & Suganthan (2008) também se entende que uma população menor apresenta convergência mais rápida, entretanto pode gerar uma estagnação nos resultados que pode justificar o pior resultado da RNA. O contrário é válido para populações maiores, ou seja, apresenta convergência mais lenta e menor chance de estagnação.

4.2.2 - Cálculo da RNA - ED

Com a população da ED definida em 10, é possível analisar o efeito do aumento de iterações da RNA nos resultados de predição do presente *dataset* de soldagem. A Figura 36 mostra o gráfico da evolução do somatório dos erros individuais de cada parâmetro de saída, ou seja, o *MAPE* (E_t). O tempo computacional para processamento também é disposto no gráfico.

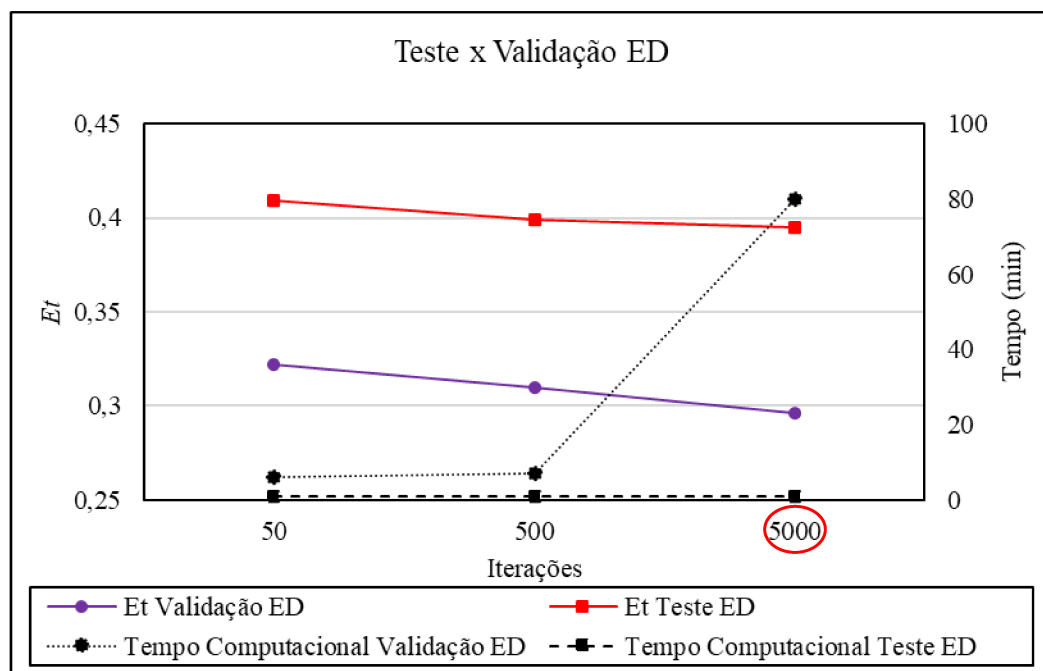


Figura 36. Gráfico do E_t e tempo computacional com variação das iterações da RNA. A condição de iterações mais precisa nos testes é circulada (5000).

O melhor resultado na fase de definição da arquitetura da RNA, ou seja, na validação, apresenta um valor de 0,296, com um custo computacional de 80 minutos, sendo referentes a 5000 iterações da RNA. O pior resultado foi obtido com 50 iterações da RNA e valores de E_t de 0,321, e tempo computacional de 6 minutos. A simulação com 500 iterações apresenta os valores intermediários para o valor do erro e também para o tempo gasto.

Nota-se uma pequena diferença entre os resultados de erro para os três casos analisados, sugerindo uma sutil evolução nos valores à medida em que se aumenta o número de iterações da RNA. Isso indica que também em casos em que a simulação deve ser feita em tempo real, é necessário analisar a relação custo-benefício do aumento das iterações. Um tempo computacional aproximadamente doze vezes menor produziu um erro 0,025 menor.

Em termos de convergência de resultados é possível notar a coerência, em que um número maior de iterações gera um resultado de predição melhor e com custo computacional maior, para o caso de validação. Como essa é uma característica já esperada pelas próprias

definições das RNA's, isso já era esperado (Schmidhuber, 2015). Entretanto, um número maior de iterações pode gerar *overfitting* como no caso reportado em Leshno & Spector (1996). Porém, justamente para evitar o *overfitting* se utiliza das etapas de validação e teste.

Fixando-se os parâmetros encontrados na fase de validação, parte-se então para os testes em outro conjunto de dados separados do treinamento e validação. Isso garante que a ED não gerará um *overfitting* nos resultados de predição na fase de testes, pois os parâmetros já são fixados em uma etapa anterior.

O melhor resultado do erro para fase de teste apresenta valor de 0,395, referente a 5000 iterações da RNA. O pior resultado foi obtido com 50 iterações da RNA, e valores de E_t de 0,408. A simulação com 500 iterações apresenta os valores intermediários para o erro. O tempo referente à fase de testes é de apenas um minuto, pois todos os parâmetros já são fixados.

A Tabela 9 apresenta os parâmetros da arquitetura da RNA otimizados pela ED a partir do E_t já apresentado. Como as sementes de geração inicial para a ED e para a RNA foram fixados, não é necessário conduzir diversas simulações para se encontrar uma convergência dos resultados.

Tabela 9. Parâmetros da arquitetura da RNA otimizados por ED.

Iterações	Camadas	Neurônios	Ativação	α
50	1	4	tanh	0,097
500	1	2	logística	0,022
5000	3	3	logística	0,049

Como pode ser observado na Tabela 9, há convergência dos resultados para 500 e 5000 iterações apenas para a função de ativação (logística). O valor de α não foi semelhante em nenhuma condição. O melhor resultado apresentou então 3 camadas, 3 neurônios, função de ativação logística, $\alpha = 0,049$, e número de iterações da RNA de 5000. Como α representa uma medida de quão rápido o algoritmo atualiza os pesos de aprendizagem, isso mostra que uma taxa de aprendizagem mais rápida apresentou os melhores resultados para a técnica ED.

Através dos estudos de Ding et al. (2018) entende-se que a rede neural com função de ativação logística do presente trabalho não apresenta o gradiente evanescente do problema, pois não possui mais que 5 camadas ocultas. Essa característica também é descrita por Rasamoelina et al. (2020). Em suas pesquisas, o melhor desempenho para função logística acontece em redes com poucas camadas. RNA's com muitas camadas frequentemente apresentam problemas de otimização após um certo ponto com essa função de ativação. Como no presente trabalho são utilizadas poucas camadas, o resultado da ED é consistente.

A função tangente hiperbólica aparece em 50 iterações e geralmente é mais utilizada do que a logística nas RNA's por sua maior velocidade de convergência, entretanto, apresenta o problema de *vanishing gradient* também (Ding et al., 2018; Rasamoelina et al., 2020).

Como se avalia o erro somado das quatro variáveis de saída (vazão, tensão, velocidade e *stickout*), os valores apresentados são maiores que os vistos em trabalhos em que se avaliam apenas os erros individuais em parâmetros de soldagem. Nesses, os resultados de erro são menores que 0,1 como em Ribeiro (2018) e em Dhas et al. (2022).

Para melhor entendimento e comparabilidade da otimização de RNA's, o próximo tópico apresenta uma comparação entre as técnicas SHGO e ED para escolha do método com melhor desempenho para prosseguimento da pesquisa.

4.3 – SHGO

Alterando-se o algoritmo de otimização do trabalho para o SHGO, torna-se necessário comparar as técnicas de otimização para a RNA. A Figura 37 mostra o gráfico relativo à influência do incremento das iterações da RNA no E_t e no tempo computacional.

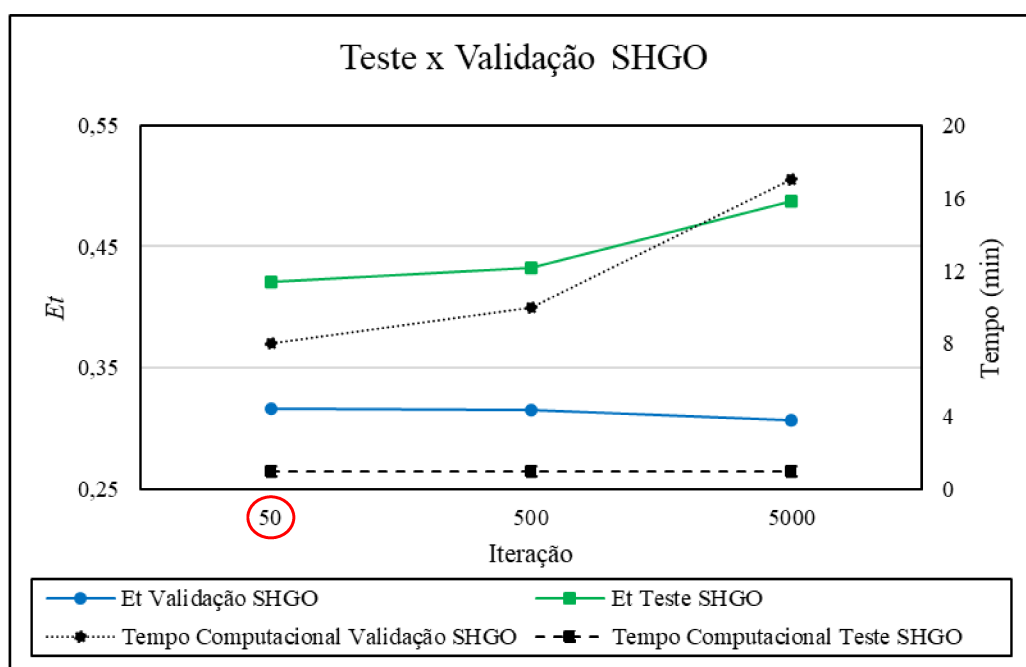


Figura 37. Resultados da otimização SHGO na RNA. A condição de iterações mais precisa nos testes é circulada (50).

A partir da figura 36 verifica-se a convergência dos resultados tanto para o E_t quanto para o tempo gasto nas simulações na fase de validação. Os melhores resultados são encontrados com 500 e 5000 iterações, com valor de erro de 0,315 e 0,307, respectivamente.

É importante ressaltar que não há garantias de melhora nos resultados com o aumento de iterações, apenas se garante que o resultado não piora (Nah et al., 2019). Os tempos gastos variam de 8 a 17 minutos para 50 e 5000 iterações, respectivamente.

Os melhores resultados da fase de teste são encontrados com 50 iterações, com E_t de 0,421. Os parâmetros definidos na fase anterior são usados na fase de teste. Os parâmetros encontrados são mostrados na Tabela 9.

Tabela 10. Parâmetros da arquitetura da RNA para o SHGO.

Iterações	Camadas	Neurônios	Ativação	α	E_t teste
50	1	4	tanh	0.099	0,421
500	3	3	logística	0.021	0,432
5000	3	3	logística	0.024	0,487

Assim como na análise feita para a tabela 9 da arquitetura resultante do ED, a função de ativação apresentou os mesmos resultados para o número de iterações testadas, com a “logística” apresentando os melhores resultados. Apresentando um nível de convergência entre 500 e 5.000 iterações, o único valor alterado foi o α , ainda uma pequena variação de 0,003.

É importante ressaltar que a convergência dos resultados de E_t só é esperada na fase de validação (4 conjuntos), onde o algoritmo de otimização atua para encontrar a melhor arquitetura, minimizando E_t . Na fase de testes (5 conjuntos diferentes dos de validação) os parâmetros são fixados nos mesmos encontrados para a fase de validação, o que não garante que para o conjunto de dados testados serão os melhores parâmetros da arquitetura da RNA. Afim de promover a escolha da melhor técnica para o caso proposto apresenta-se o próximo tópico de comparação das técnicas de otimização.

4.3.1 – Comparação SHGO e ED

A Figura 38 apresenta os resultados encontrados nas técnicas de ED e SHGO à medida em que se aumenta o número de iterações na RNA.

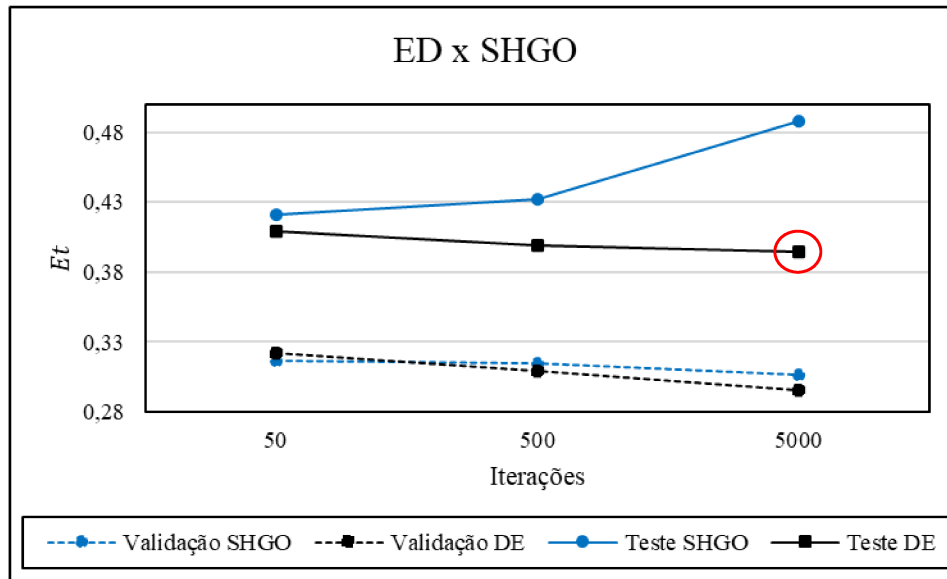


Figura 38. Comparação das técnicas SHGO e ED. O melhor resultado da técnica mais adequada é circulada (ED – 5000 iterações).

Comparando-se as duas técnicas de otimização entende-se que a ED apresenta os melhores valores para os erros na fase de testes. Entretanto, comparando-se os tempos computacionais das figuras 36 e 37 nota-se que o algoritmo SHGO possui convergência mais rápida, principalmente em 5.000.

A ED é uma técnica de otimização populacional baseada em algoritmos genéticos utilizando uma população de soluções candidatas e operadores genéticos para gerar novas soluções. Uma característica marcante é a rápida capacidade de convergência, como identificado no trabalho de Z. Hu et al. (2013), que apresenta um estudo de convergência com o algoritmo de ED e como resultado mostra rápida convergência com 25 iterações em seu estudo.

Na fase de validação, os resultados são muito semelhantes, com E_t do SHGO melhor que ED com apenas 50 iterações, enquanto nas demais condições há ED apresentando melhores resultados. Por outro lado, apenas analisando-se a fase de teste, a SHGO emprega uma estratégia de divisão do espaço de busca em regiões discretas chamadas de complexos simpliciais, que são estruturas geométricas formadas por vértices, arestas, faces e assim por diante. O menor tempo computacional associado ao algoritmo SHGO é coerente com o que foi visto na comparação entre as técnicas ED e SHGO feita no trabalho de Endres et al. (2018). As condições testadas eram em problemas multimodais com restrições de intervalo.

Como os resultados apontam que o método com menores erros é a ED, segue-se a pesquisa com essa técnica. A seguir são apresentados os resultados das técnicas para aprimorar valores dos erros em *datasets* pequenos.

4.4 – Validação Cruzada LOO

Para melhorar os resultados de Redes Neurais Artificiais (RNA) enfrentando as limitações impostas por um conjunto de dados pequeno, foi adotada uma abordagem baseada na pesquisa de Y. Zhu et al. (2023). O conjunto de dados utilizado nesta pesquisa, contendo apenas 200 amostras, é ainda menor do que o considerado pequeno no estudo mencionado. Para mitigar os desafios associados ao uso de conjuntos de dados tão limitados, foram exploradas técnicas específicas como aumento de dados e validação cruzada, buscando otimizar o desempenho da RNA apesar das restrições significativas de dados.

Para solucionar o problema é então proposta a técnica LOO para a validação cruzada dos dados como proposto nos trabalhos de Wong (2015) e Berrar (2019). A técnica é utilizada nesse tipo de problema pois tem a capacidade de aumentar as amostras para treinamento da RNA, usando apenas uma amostra por vez no teste.

Inicialmente, são realizadas as análises do efeito do aumento das iterações na RNA no Et e tempo de processamento em todos os 9 pontos (4 para validação e 5 para teste) que foram usados como parcela de 20% utilizadas nas etapas anteriores. A Figura 39 apresenta os resultados referentes a 50 iterações.

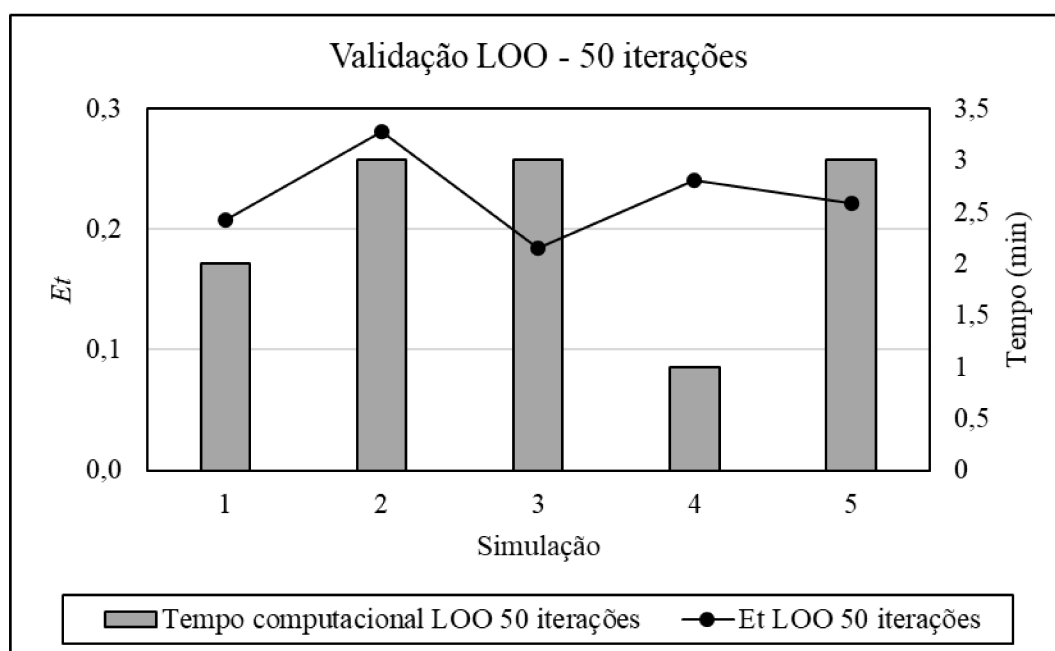


Figura 39. Resultados LOO com 50 iterações da RNA.

Analisando-se os erros da validação LOO é possível ver que são valores em sua maioria menores que os apresentados na técnica padrão da RNA. Os valores dos tempos computacionais também são variáveis, assim como os erros, e isso já era esperado pois como são pontos diferentes, o algoritmo também se adapta de forma diferente.

A Figura 40 mostra essa análise para 500 iterações da RNA.

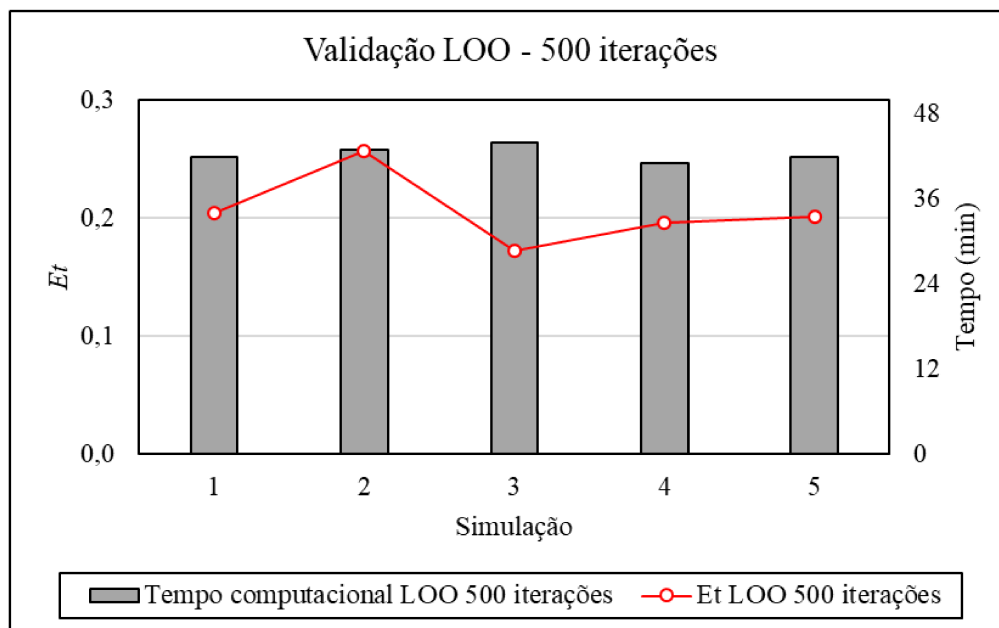


Figura 40. Resultados LOO com 500 iterações da RNA.

Os valores apresentados também mostram variações similares aos apresentados pela Figura 39. Porém, é importante notar que os tempos são maiores, resultado da maior complexidade envolvida no algoritmo com mais iterações. A Figura 41 dispõe dos resultados da técnica LOO para 5000 iterações nos pontos testados individualmente.

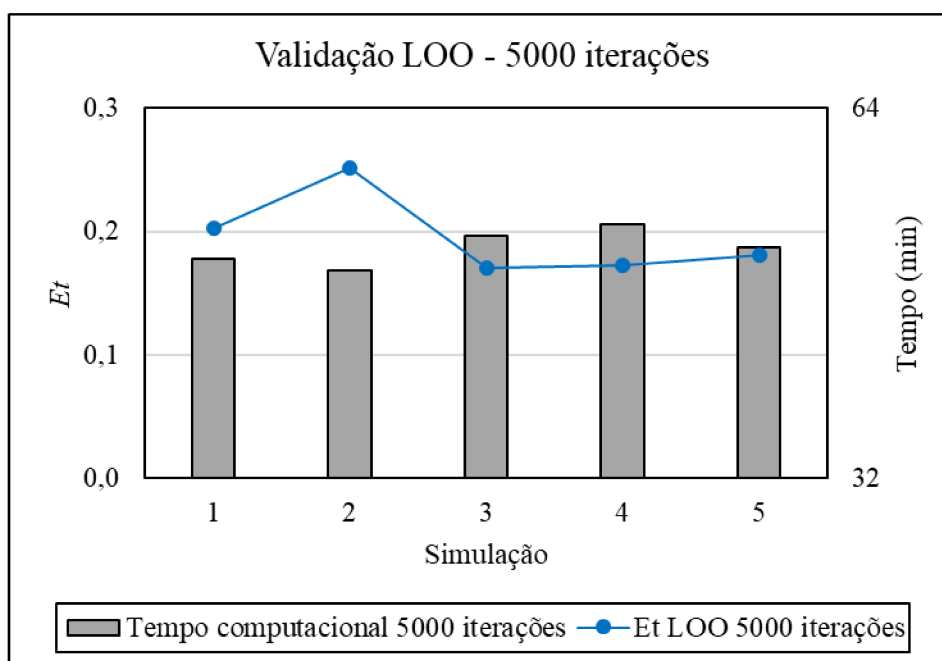


Figura 41. Resultados LOO com 5000 iterações da RNA.

Assim como nos resultados para 50 e 500 iterações, a Figura 41 possui variação similar, mas com escala de tempos maiores que os previamente testados. A Figura 42 mostra as médias calculadas para as simulações em forma gráfica.

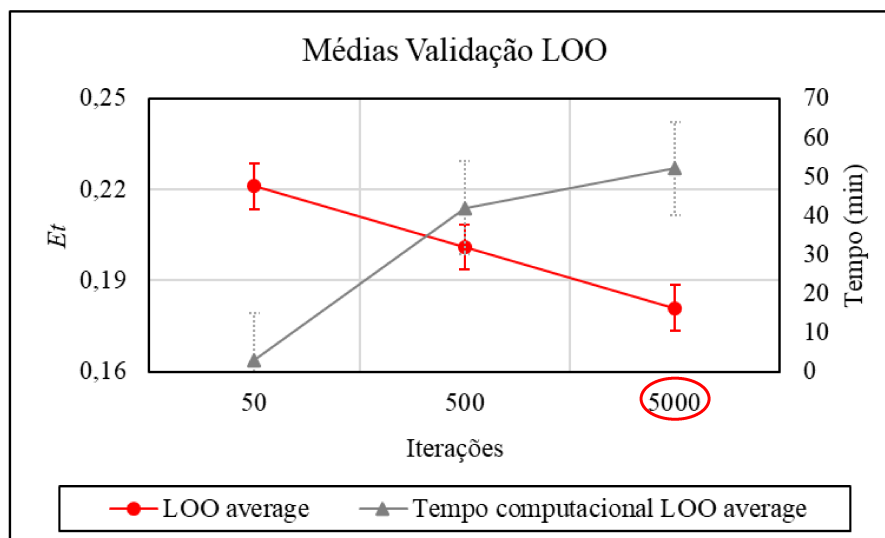


Figura 42. Gráfico com as médias das simulações LOO na validação. A condição de iterações com melhor resultado para a validação com LOO é circulada (5000).

Os valores médios de erro de validação variam de 0,221 em 50 iterações a 0,181 em 5.000 iterações (redução de 18,1%). O tempo aumenta de 3 para 52 minutos (aumento de 1733%), inversamente proporcional à redução do erro. A Figura 43 mostra a mesma análise para a fase de testes.

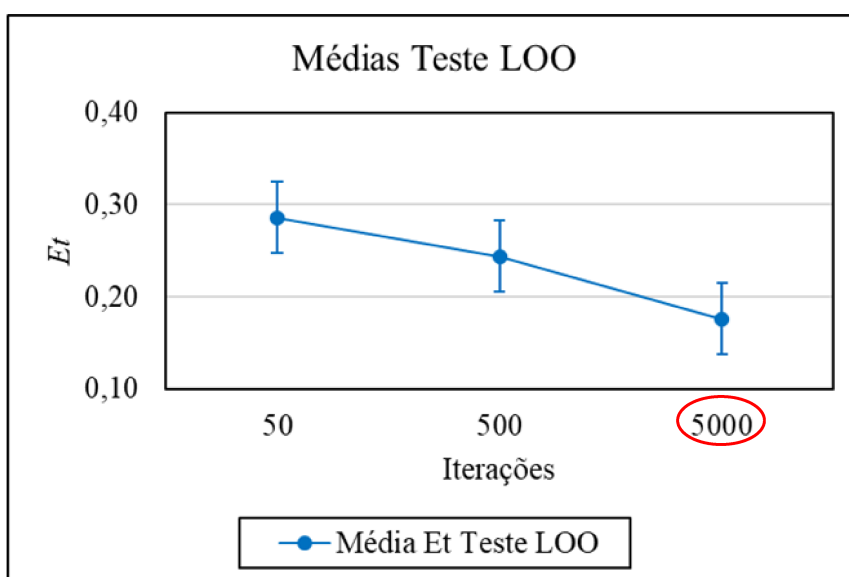


Figura 43. Gráfico com médias das simulações LOO no teste. A condição de iterações com melhor resultado para o teste com LOO é circulada (5000).

Com 5000 apresenta-se o melhor valor médio de E_t (0,176), enquanto o pior resultado é observado em 50 iterações e E_t de 0,286. É interessante notar que os valores são aproximadamente metade dos apresentados sem a técnica de validação cruzada. Portanto, é uma técnica bem útil para problemas de poucas amostras em um *dataset* de soldagem GMAW.

No trabalho de Wong (2015) há também uma melhora nos resultados usando o algoritmo de validação cruzada LOO, mas deve ser salientado que o trabalho não envolve problemas de soldagem. A Figura 44 apresenta a análise ponto a ponto da LOO nas três condições testadas na validação.

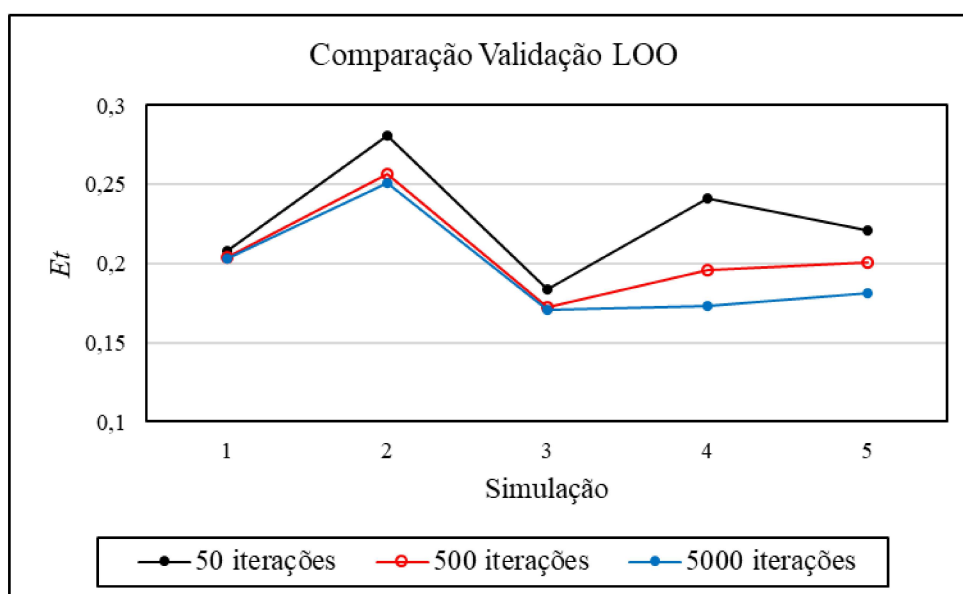


Figura 44. Comparação ponto a ponto LOO na validação.

Assim como nos valores apresentados pelas médias, é possível ver que o maior número de iterações promove o melhor resultado em todos os pontos, exceto no ponto 3 que possui o mesmo valor de E_t , comprovando o que foi dito em Nah et al. (2019), que um número maior de iterações não garante melhores resultados. A mesma comparação anterior é apresentada na Figura 45 para a fase de teste.

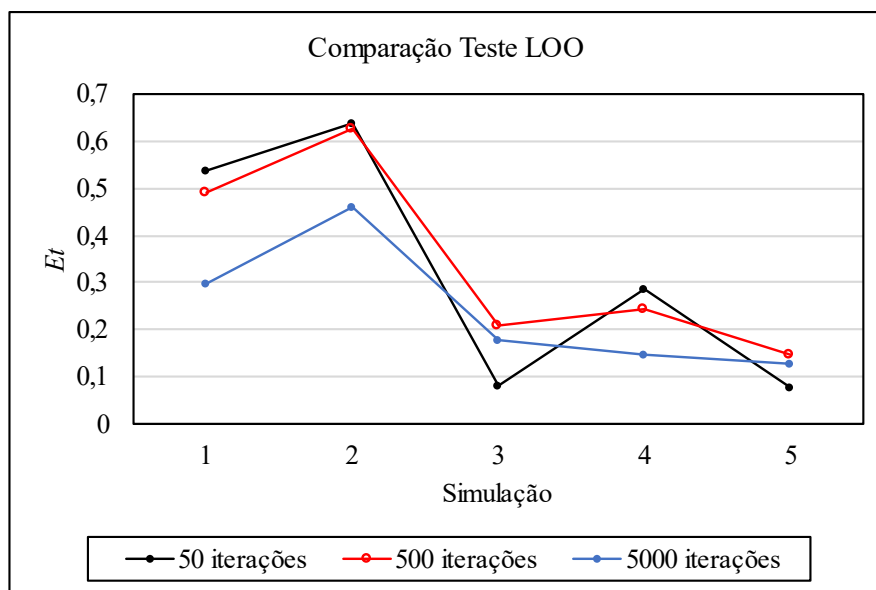


Figura 45. Comparação ponto a ponto LOO no teste.

Como já apresentado anteriormente, não há garantias que o aumento de iterações melhore o resultado na fase de teste, pois os parâmetros da arquitetura da RNA foram definidos para outro conjunto de dados na fase de validação. Os resultados da fase de teste apresentam E_t maiores que os apresentados na validação em todos os pontos testados, entretanto, como já apresentado pela Figura 43, os valores médios dos 5 pontos são menores que as outras técnicas testadas na mesma fase. Os parâmetros da arquitetura da RNA com a LOO são apresentados na Tabela 11.

Tabela 11. Parâmetros da arquitetura da RNA com LOO.

Iterações	E_t Teste LOO	Camadas	Neurônios	Ativação	α
50	0,538	1	4	logística	0,020
	0,640	2	2	logística	0,012
	0,082	4	4	identidade	0,097
	0,286	2	2	logística	0,053
	0,078	2	1	identidade	0,081
500	0,491	4	4	logística	0,021
	0,626	1	2	logística	0,072
	0,208	3	3	logística	0,059
	0,244	1	4	logística	0,072
	0,147	2	2	logística	0,021
5000	0,299	1	4	logística	0,080
	0,462	1	4	logística	0,098
	0,176	1	4	logística	0,014
	0,147	1	4	logística	0,053
	0,127	3	3	logística	0,078

Há maior convergência nos resultados para função de ativação, camadas e neurônios para 5.000 iterações. Isso pode mostrar o maior poder de convergência com esse número de iterações na RNA. Apenas dois pontos tiveram sua função de ativação alterada de “logística”, apresentando a função identidade com 50 iterações, o que indicou que na maioria dos casos a convergência para a função de ativação para esses parâmetros da arquitetura já havia sido alcançada com 500 iterações.

A função de ativação de identidade é apresentada apenas para 50 iterações. Por ser uma função mais simples que as demais e exigir menor gasto computacional para sua utilização, é, portanto, mais fácil de ser encontrada em simulações com menos iterações. Porém, mesmo sendo uma função mais simples, não é sinônimo de resultados de baixa qualidade, como mostra Wanto et al. (2017), que encontrou uma precisão de 94% (a mais alta entre as testadas) em uma previsão de densidade populacional com RNA usando uma função de ativação linear.

O valor de α é o valor que mais muda à medida em que se aumenta nas iterações. Isso indica que cada conjunto de dados possui uma taxa ótima de aprendizado diferente, gerando assim os melhores resultados para o caso de estudo.

Determinar os melhores valores de α não é uma tarefa trivial, pois depende do conjunto de dados, da arquitetura da rede e do problema específico a ser abordado. Atualmente, vem utilizando técnicas de controle adaptativo da taxa de aprendizagem, conforme proposto pela ED no presente trabalho. A maioria dos trabalhos utiliza uma taxa de aprendizagem adaptativa que decai durante o treinamento, porém algumas propostas como a presente em Takase et al. (2018), que utilizam uma taxa que pode aumentar ou diminuir com base nos resultados de predição da RNA.

Com a avaliação a validação cruzada sendo considerada uma prática com bons resultados, é necessário entender o efeito de técnicas de aumento de dados, que serão abordadas nos próximos itens.

4.5 - Geração de Dados

A seguir são apresentados os resultados da técnica Kriging e sua comparação com o SDV. A Figura 46 mostra o erro da fase de validação com incremento de iterações, e a Tabela 12 apresenta os parâmetros da RNA com o tempo de processamento computacional da fase de validação para ambos os métodos de aumento de dados com todas as condições utilizadas.

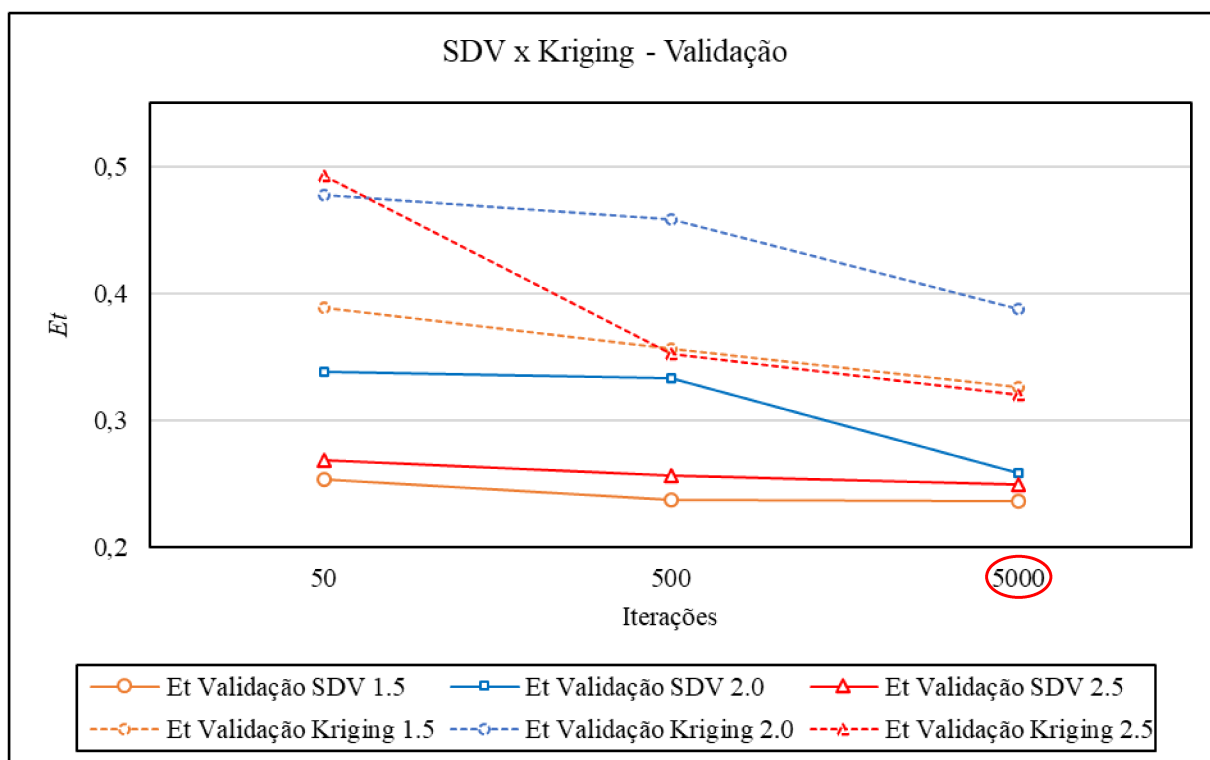


Figura 46. Comparação do E_t na validação do SDV e do Kriging. A condição de iterações com melhor resultado para a validação com aumento de dados é circulado (5000).

Tabela 12. Parâmetros de arquitetura da RNA para SDV e Kriging.

Iterações	Método	Condição	E_t Teste	Camadas	Neurônios	Ativação	α
50	SDV	1.5	0,355	1	2	logística	0,063
		2	0,385	1	4	identidade	0,074
		2.5	0,366	4	4	identidade	0,084
	Kriging	1.5	0,535	3	3	identidade	0,024
		2	0,415	1	2	identidade	0,056
		2.5	0,408	2	2	logística	0,058
500	SDV	1.5	0,418	4	4	tanh	0,069
		2	0,418	2	2	identidade	0,044
		2.5	0,409	4	3	tanh	0,053
	Kriging	1.5	0,642	1	2	logística	0,029
		2	0,541	1	3	logística	0,073
		2.5	0,708	1	4	logística	0,057
5000	SDV	1.5	0,403	3	3	logística	0,073
		2	0,608	3	3	logística	0,061
		2.5	0,450	1	4	tanh	0,055
	Kriging	1.5	0,707	1	4	tanh	0,028
		2	0,804	1	3	tanh	0,079
		2.5	0,464	4	4	logística	0,099

Para o SDV, o tempo varia de 2 a 77 minutos (3850%), para 50 (condição 1.5) e 5000 (condição 2.5) iterações, respectivamente. O erro apresenta convergência em todas as condições. Isso também é válido para o tempo gasto, que aumenta com o número de iterações. Os dados analisados com a técnica SDV apresentaram maior sensibilidade à mudança em 5.000 iterações da RNA apenas na condição 2, gerando um resultado visivelmente melhor do que aqueles observados com menos iterações. A condição 1.5 apresentou os melhores resultados na técnica SDV em fase de validação.

O Kriging é geralmente utilizado em técnicas de predição, mas foi adaptada para gerar mais dados no presente trabalho. Os critérios de convergência avaliados nas condições anteriores, como diminuição dos erros e aumento do tempo computacional para medir o aumento das iterações, também são observados para a krigagem. 5.000 iterações foram encontradas como o melhor resultado para o E_t na validação. Os tempos computacionais variam de 1 a 100 minutos (10.000%), enquanto o erro melhora com o incremento das iterações em todas as condições testadas.

A Figura 47 mostra a evolução do erro da fase de teste nas simulações a partir do aumento no número de iterações com a técnica SDV e krigagem com 50%, 100% e 150% a mais de dados gerados no conjunto de simulações (condição 1.5, 2 e 2.5).

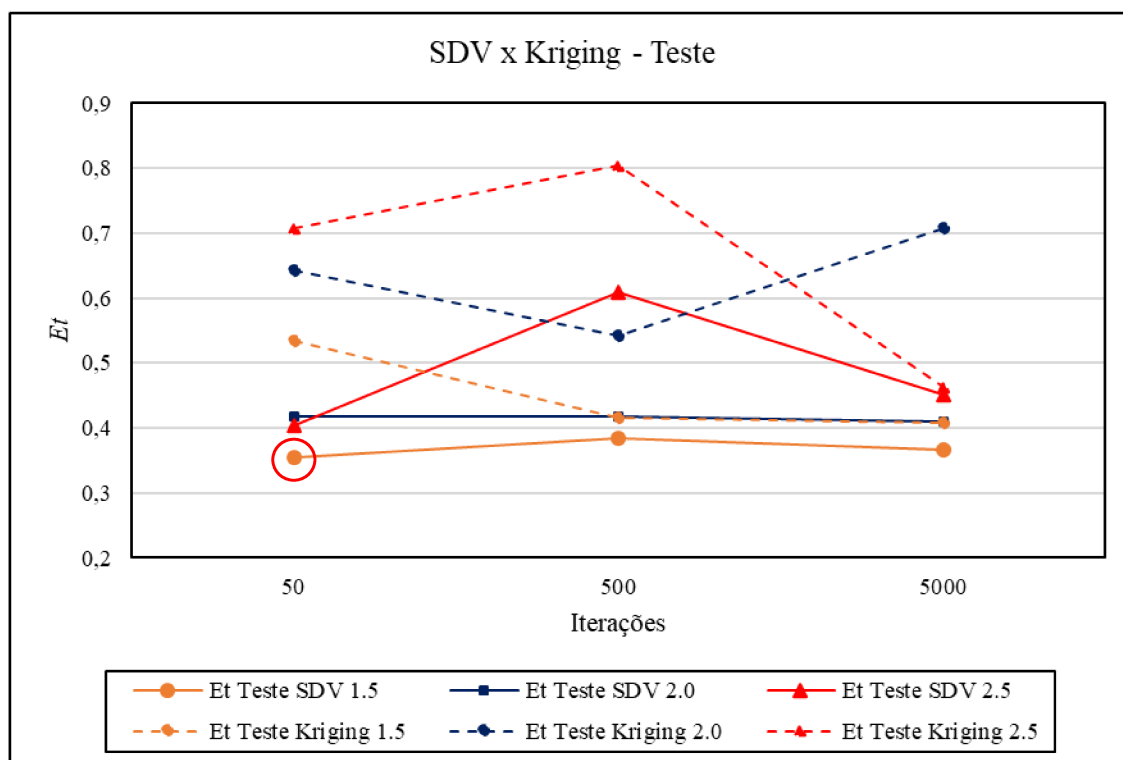


Figura 47. Comparação de E_t na validação de SDV e Kriging. Melhor resultado nos testes com aumento de dados é circulado (SDV 1.5).

O E_t vai de 0,355 com 50 iterações para 0,385 (8,9%) com 500 iterações em SDV com a condição de melhor resultado (1.5). Na melhor condição de Kriging (1.5), o melhor resultado é 0,408 com 5000 iterações, e o pior 0,535 com 50 iterações.

Das técnicas de geração de dados, os melhores resultados são obtidos com a técnica SDV, ainda melhores que os dados apresentados sem dados sintéticos na condição 1.5. Essa melhoria também é identificada em Moreno-Barea et al. (2020) que utiliza a geração de mais dados em um pequeno conjunto de dados e avalia a melhoria na precisão das previsões.

A técnica de Kriging apresentou resultados de erros superiores aos observados sem aumento de dados, o que difere do observado em Liu & Wang (2022), onde o resultado é melhorado. Porém, é necessário ressaltar que nesse trabalho é utilizado um algoritmo evolutivo em conjunto com o Kriging para geração de dados. No presente trabalho, a ED é utilizada para determinar os parâmetros da arquitetura da RNA.

Os tempos computacionais para a técnica SDV são em sua maioria os mais altos, mas também apresentam os melhores resultados para o E_t . Como a técnica SDV é muito recente, ainda não existem avaliações na área de soldagem, portanto a comparação não é possível.

4.5.1 – Efeitos do Tamanho de Dados Sintéticos

Para avaliação de um melhor valor de geração de dados por SDV e Kriging é proposta a avaliação de 3 condições de aumento de amostras, em 50% (1.5), em 100% (2) e 150% (2.5). Para comparação com o valor original obtido com o *dataset* (ED padrão), esse valor também é disposto na análise da Figura 48 para a validação, e na Figura 49 para a fase de teste. A condição padrão usada em ambos os casos mostra apenas o melhor resultado encontrado, independentemente do número de iterações.

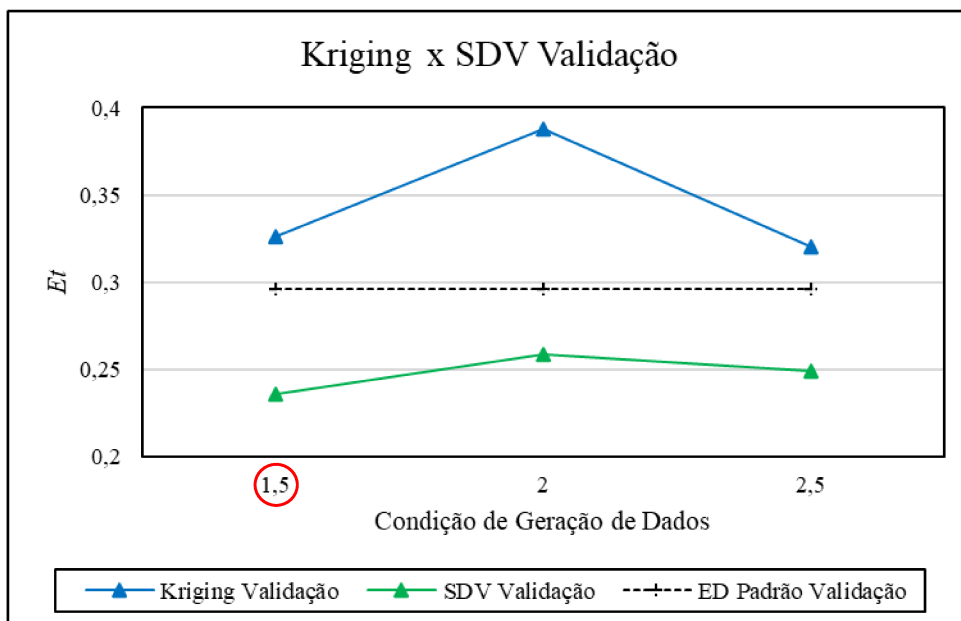


Figura 48. Efeito do aumento de dados nas previsões da RNA e condição padrão na validação. A condição de aumento de dados de melhor resultado na validação (SDV 1.5) é circulada.

Nota-se que a técnica SDV possui os melhores resultados comparando-se inclusive com a condição da ED padrão em todas as condições de aumento analisadas. Já a técnica Kriging não melhora os resultados para a fase de validação em comparação ao *dataset* original.

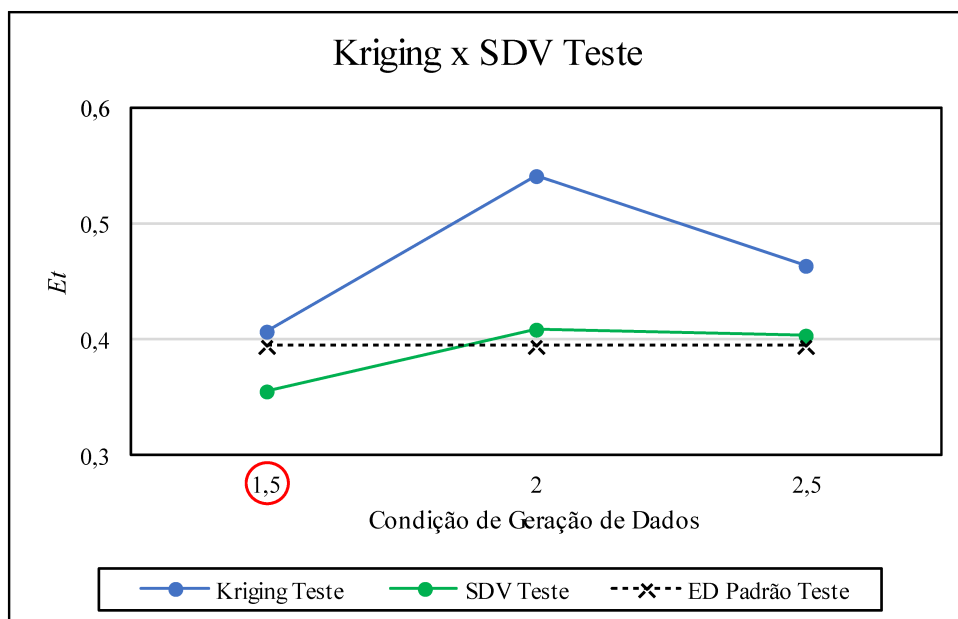


Figura 49. Efeito do aumento de dados nas previsões da RNA e condição padrão no teste. A condição de aumento de dados de melhor resultado no teste (SDV 1.5) é circulada.

Inicialmente, é importante ressaltar que os resultados foram apresentados com a melhor condição de iteração alcançada para cada técnica. Na Figura 48, os valores na condição 1.5 de SDV resultaram em melhores valores de E_t em comparação com a condição de teste da ED padrão. Isso mostra a viabilidade da técnica SDV aumentando-se o tamanho das amostras iniciais. Já na condição 2, os valores começaram a ficar maiores em relação à condição de teste da ED.

Os tempos computacionais aumentam progressivamente à medida que o conjunto de dados aumenta (Tabela 12), mostrando que a complexidade do problema também aumenta com o tamanho da amostra. O melhor resultado do SDV apresentado foi o aumento de 50% nos dados de 44 para 66 amostras, saindo do erro de 0,394 na condição original para 0,355 (10,98%) na condição em questão. Isso mostra a eficiência da técnica SDV no caso de soldagem analisado, sendo assim uma grande possibilidade para futuros estudos em outras áreas de processos de fabricação.

No Kriging, nenhuma condição de aumento de dados teve melhor desempenho do que o conjunto de dados original, o que indica a inviabilidade da técnica.

4.6 – Comparação das Técnicas de RNA's

Neste tópico são analisadas todas as técnicas em um mesmo gráfico, Figura 50, para avaliação da fase de validação.

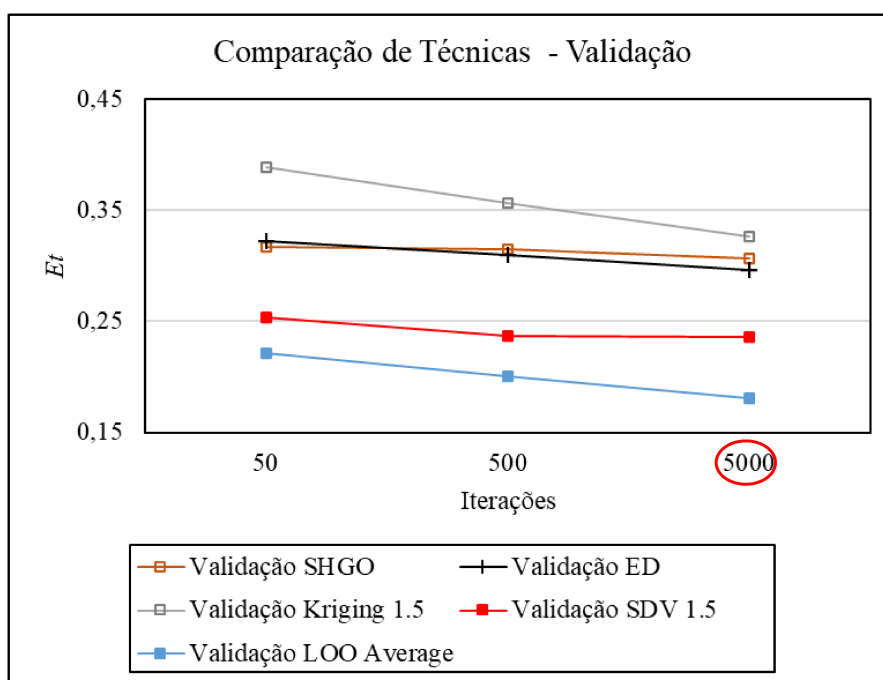


Figura 50. Comparativo das técnicas na predição da RNA na validação. O melhor resultado usando a RNA na validação é circulado (LOO – 5000 iterações).

Do ponto de vista da validação, a classificação de desempenho pode ser feita da seguinte forma:

1. Validação Cruzada: LOO.
2. Cofre de dados sintéticos: SDV.
3. Otimização ED com o conjunto de dados original.
4. Otimização SHGO com o conjunto de dados original.
5. Aumento de dados por krigagem.

É importante destacar que a validação é uma fase anterior à simulação, onde os resultados podem ser afetados pelo *overfitting*, uma vez que o algoritmo de otimização está atingindo o menor erro do conjunto de dados utilizado nessa etapa. Por esse motivo, a fase de teste é realizada com outros conjuntos de dados para evitar *overfitting*, e é o melhor resultado a ser considerado.

Por fim, na Figura 51 é realizada a análise dos melhores resultados de predições encontrados para cada técnica na condição de teste.

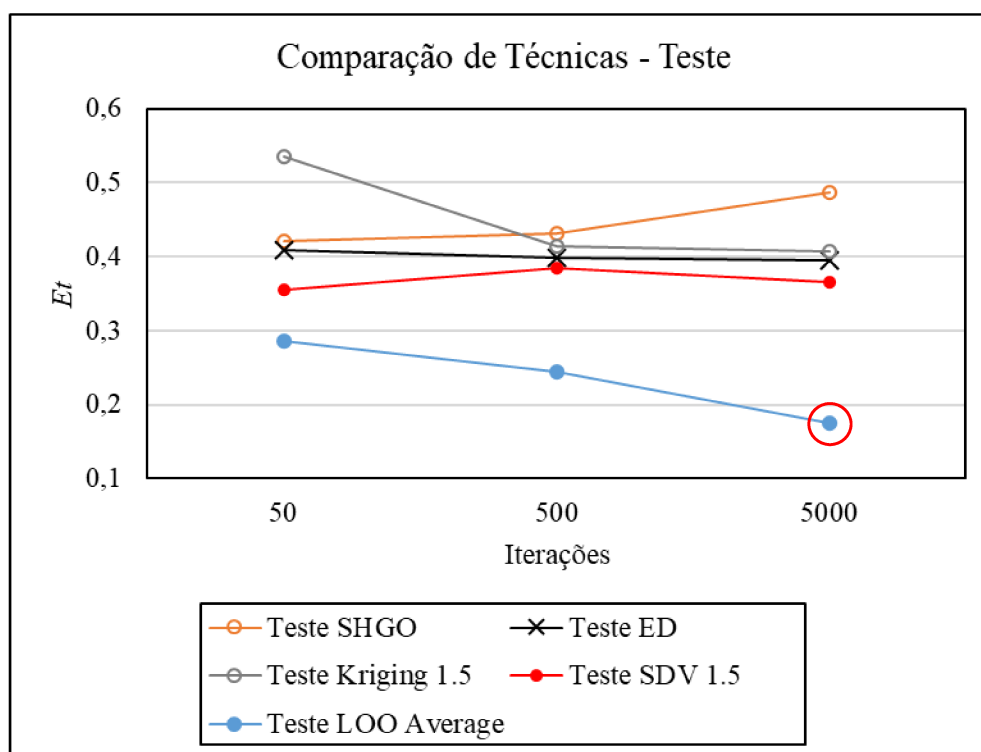


Figura 51. Comparativo das técnicas na predição da RNA na validação. Melhor resultado usando RNA no teste é circulado (LOO – 5000 iterações).

A classificação em termos de desempenho de erros é mostrada abaixo:

1. Validação Cruzada: LOO.
2. Cofre de dados sintéticos: SDV.
3. Otimização ED com o conjunto de dados original.
4. Aumento de dados por krigagem.
5. Otimização SHGO com o conjunto de dados original.

Cada uma das técnicas foi comparada com outras no mesmo propósito para definir a melhor viabilidade para o caso de soldagem MAG em questão. Destaca-se a validação cruzada sendo uma alternativa bem interessante para o problema gerado por pequenos *datasets*. A técnica de SDV é uma alternativa ao mesmo problema, tendo resultados da RNA melhorados em relação à condição original.

Também se comprova a validade de cada uma das técnicas utilizadas, com maiores ou menores erros associados, conseguindo encontrar valores ótimos de parâmetros de arquitetura sem a utilização de métodos de tentativa e erro, como reportado por Moghaddam & Kolahan (2021) para um problema complexo de múltiplas entradas e saídas em uma base de dados reduzida de soldagem.

Por fim, é necessário realizar uma comparação com uma técnica mais simples de ML como a regressão para comprovar a efetividade das RNA aplicadas em problemas de soldagem como o proposto.

4.7 - Regressão

Os resultados de regressão são divididos nas predições de cada técnica utilizada e na comparação com a soma de todos os erros de saída individuais das técnicas de RNA .

4.7.1 – Regressão Linear

Após realizar a regressão em cada saída (velocidade do arame, tensão, vazão do gás e *stickout*) é necessário avaliar os resultados da previsão. A Figura 52 apresenta o MAPE obtido com regressão linear.

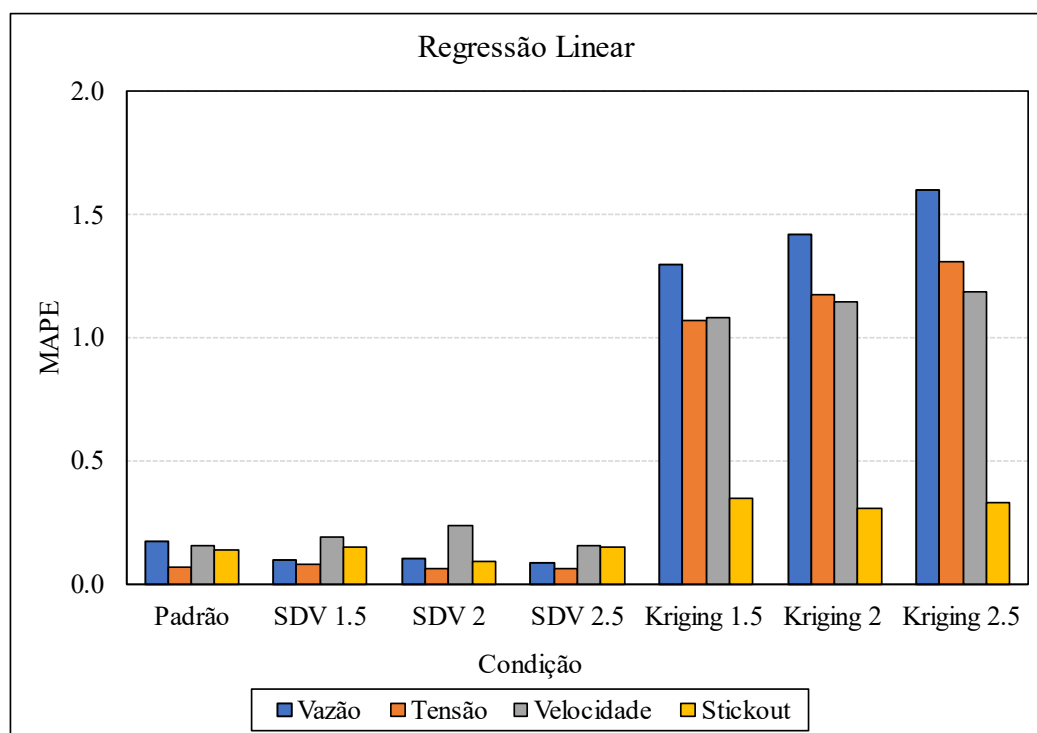


Figura 52. Erros de previsão obtidos a partir da regressão linear (E_t) para todas as condições do *dataset*.

Pela Figura 50 é possível perceber que o aumento dos dados da krigagem não gerou melhores resultados de predição do que os dos casos padrão e por SDV. Vale ressaltar também que o SDV produziu resultados semelhantes ou melhores que a condição padrão. Para uma análise mais aprofundada, é então necessário analisar os resultados numéricos das previsões de regressão linear disponíveis na Tabela 13. Esse resultado mostra o MAPE individual para cada output e o E_t somado. As cores utilizadas são uma forma de diferenciar os melhores dos piores resultados.

Tabela 13. Valores de erro de previsão de regressão linear (E_t).

Condição	Vazão	Tensão	Velocidade	Stickout	E_t Regressão
Padrão	0,173	0,068	0,157	0,139	0,537
SDV 1.5	0,096	0,080	0,191	0,148	0,515
SDV 2	0,103	0,062	0,236	0,090	0,491
SDV 2.5	0,082	0,063	0,152	0,146	0,443
Kriging 1.5	1,293	1,066	1,079	0,346	3,784
Kriging 2	1,420	1,172	1,146	0,305	4,043
Kriging 2.5	1,601	1,309	1,183	0,332	4,425

*As cores mais quentes representam erros mais elevados, enquanto as cores mais frias apresentam erros mais baixos.

A partir da Tabela 12 é possível inferir que a condição SDV apresentou as melhores previsões, tanto individual quanto coletivamente, para as saídas analisadas. Os melhores resultados de vazão de gás foram 0,082, velocidade 0,152 e E_t 0,443, todos na condição SDV 2,5, mostrando que o aumento dos dados é promissor. Na condição SDV 2.0, os melhores valores de tensão são 0,062 e para *stickout* 0,09. Comparando com a condição padrão testada, todo o aumento de dados realizado pelo SDV melhorou os resultados de predição, enquanto na krigagem não houve benefícios em nenhum parâmetro previsto.

A Tabela 14 apresenta os valores calculados para R^2 em cada saída e condição testada na regressão linear.

Tabela 14. Coeficiente de determinação (R^2) para as entradas com a regressão linear.

Condição	Vazão	Tensão	Velocidade	Stickout
Padrão	0,464	0,210	0,112	0,374
SDV 1.5	0,639	0,001	0,252	0,630
SDV 2	0,331	0,326	0,052	0,734
SDV 2.5	0,541	0,004	0,190	0,500
Kriging 1.5	0,034	0,189	0,595	0,008
Kriging 2	0,023	0,131	0,614	0,016
Kriging 2.5	0,014	0,099	0,642	0,024

*Apenas valores > 0,5 são destacados.

A partir da Tabela 14 tem-se que apenas os valores destacados apresentam resultados mais próximos de 1, o que pode indicar um bom ajuste do modelo aos dados, onde a variável independente explica a maior parte da variabilidade na variável dependente. Destaca-se o valor de 0,734 para a saída Stickout e condição SDV 2 como o melhor ajuste. Para as demais, não destacadas, tem-se uma indicação que o modelo de regressão não é capaz de explicar muito da variabilidade na variável dependente.

4.7.2 – Regressão Polinomial

Para a regressão polinomial, o mesmo tipo de análise é feita e apresentada na Figura 53. O MAPE é analisado para cada condição testada em cada saída.

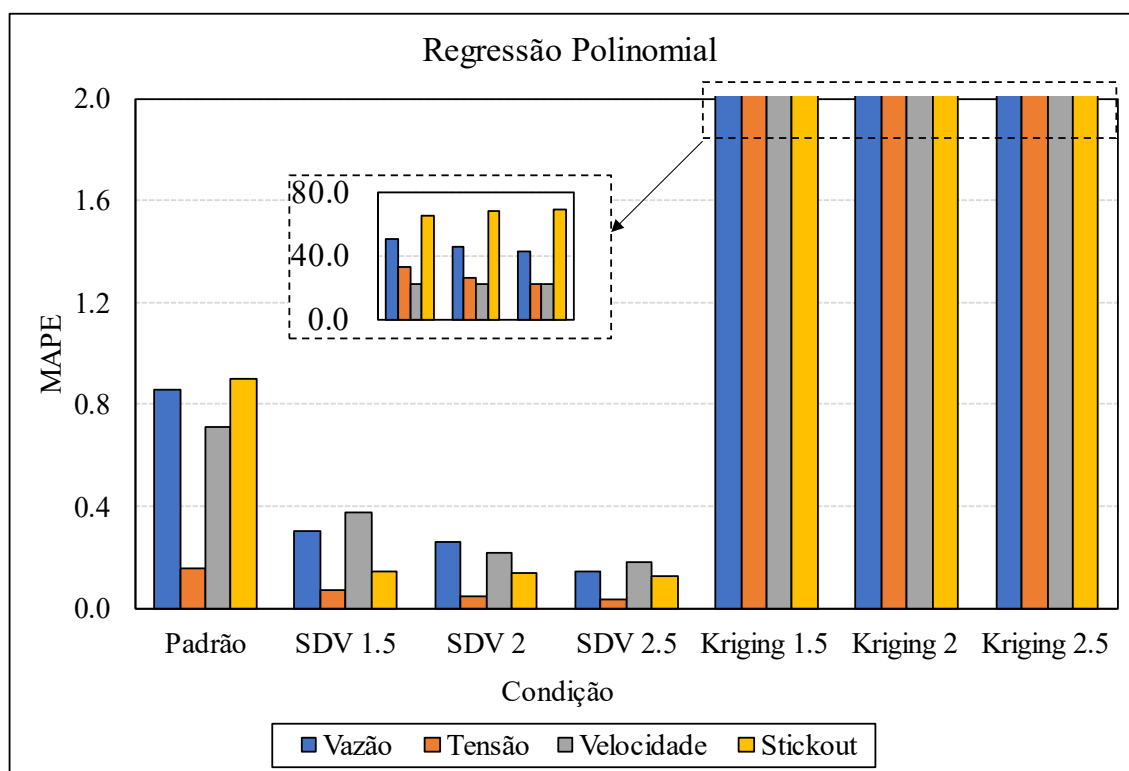


Figura 53. Erros de previsão obtidos a partir da regressão polinomial (E_t) para todas as condições do *dataset*. A área destacada é mostrada em uma escala diferente para visualização adequada dos resultados.

Observa-se que os resultados da previsão por regressão polinomial apresentam desempenho notavelmente inferior em comparação à regressão linear em condições padrão e de krigagem. Para o SDV há maior comparabilidade e é necessário analisar os valores numéricos da Tabela 15.

Tabela 15. Erros de previsão de regressão polinomial (E_t) para todas as condições do *dataset*.

Condição	Vazão	Tensão	Velocidade	Stickout	E_t Regressão
Padrão	0,858	0,159	0,714	0,898	2,629
SDV 1.5	0,305	0,074	0,375	0,144	0,898
SDV 2	0,261	0,045	0,219	0,140	0,665
SDV 2.5	0,142	0,037	0,179	0,126	0,484
Kriging 1.5	50,944	32,892	22,787	65,903	172,526
Kriging 2	45,727	26,308	22,838	68,012	162,885
Kriging 2.5	43,371	22,399	22,546	69,009	157,325

*As cores mais quentes representam erros mais elevados, enquanto as cores mais frias apresentam erros mais baixos.

Para a regressão polinomial, a Tabela 15 mostra o que já havia sido identificado pela regressão linear, ou seja, o aumento de dados por SDV apresenta os melhores resultados de predição dos parâmetros, enquanto a krigagem apresenta os piores. Neste caso analisado, a condição SDV 2.5 apresentou os melhores resultados para todas as saídas e também para E_t . Isso mostra uma tendência de melhorar as previsões à medida que os dados aumentam, ou seja, quanto mais dados melhor para a técnica nesse problema de soldagem. É interessante notar que a krigagem apresentou valores muito elevados para as previsões, mostrando um problema para sua utilização no aumento de dados para o caso analisado neste trabalho. O melhor resultado para E_t , de 0,484, na condição SDV 2.5, é semelhante aos encontrados na regressão linear, enquanto as demais condições apresentam valores muito superiores aos encontrados anteriormente. Os coeficientes de determinação da regressão polinomial são apresentados na Tabela 16.

Tabela 16. Coeficiente de determinação (R^2) para as entradas com a regressão polinomial.

Condição	Vazão	Tensão	Velocidade	Stickout
Padrão	0,336	0,000	0,444	0,052
SDV 1.5	0,033	0,021	0,068	0,983
SDV 2	0,140	0,578	0,002	0,652
SDV 2.5	0,228	0,014	0,095	0,912
Kriging 1.5	0,048	0,012	0,368	0,660
Kriging 2	0,043	0,009	0,378	0,665
Kriging 2.5	0,042	0,007	0,385	0,668

*Apenas valores > 0,5 são destacados.

Os resultados destacados da Tabela 16 sugerem que o modelo de regressão se ajusta bem aos dados, com a maioria das relações encontradas no *stickout* com bom ajuste. O melhor ajuste é encontrado na condição SDV 1.5 com R^2 de 0,983. Já para as demais, com exceção de uma condição na tensão (SDV 2), os valores de R^2 são baixos, o que não indica um bom ajuste da regressão aos dados.

4.7.3 – Regressão Múltipla de Ridge

Por fim, para a regressão múltipla de Ridge, a Figura 54 mostra graficamente as previsões para cada saída do problema de soldagem analisado.

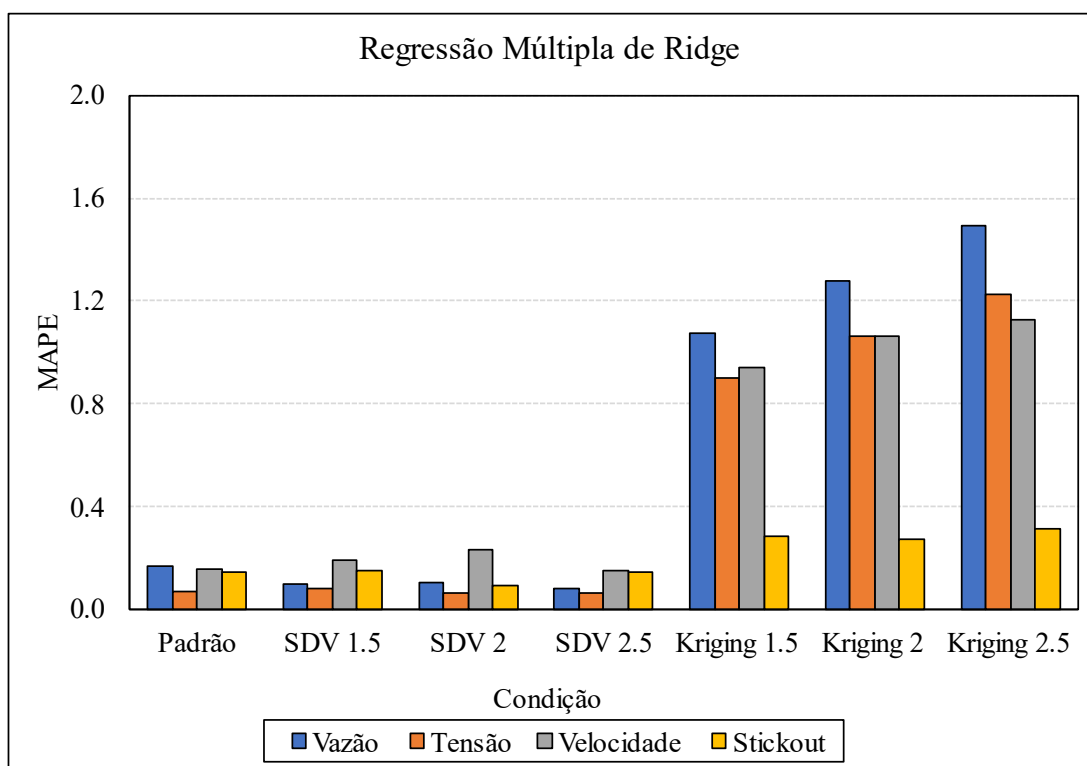


Figura 54. Erros de previsão obtidos a partir da regressão polinomial (E_t) para todas as condições do *dataset*. A área destacada é mostrada em uma escala diferente para visualização adequada dos resultados.

É interessante analisar que a técnica Ridge em regressão múltipla apresenta os melhores resultados de predição para a técnica de Kriging. Porém, assim como nas demais regressões, a krigagem apresenta os piores resultados em comparação ao caso padrão e ao SDV. Nos casos padrão e SDV também há bom desempenho, com seus valores numéricos apresentados na Tabela 17.

Tabela 17. Erros de previsão de regressão de Ridge (E_t) em todas as condições do *dataset*.

Condição	Vazão	Tensão	Velocidade	Stickout	E_t Regressão
Padrão	0,166	0,066	0,155	0,143	0,530
SDV 1.5	0,096	0,079	0,190	0,148	0,513
SDV 2	0,103	0,060	0,233	0,090	0,486
SDV 2.5	0,081	0,062	0,152	0,146	0,441
Kriging 1.5	1,076	0,898	0,943	0,283	3,200
Kriging 2	1,279	1,061	1,061	0,274	3,675
Kriging 2.5	1,490	1,225	1,124	0,311	4,150

*As cores mais quentes representam erros mais elevados, enquanto as cores mais frias apresentam erros mais baixos.

É possível perceber que a regressão múltipla de Ridge apresenta os melhores resultados entre as demais regressões testadas. Há um comportamento semelhante dos resultados em comparação à regressão linear. O melhor E_t , 0,441, bem como a melhor previsão de fluxo de gás, 0,081, e velocidade, 0,152 são encontrados na condição SDV 2.5. A tensão de 0,062 e o *stickout* de 0,090 são obtidos no SDV 2.0. O resultado do SDV melhorou com o aumento na base de dados criada sinteticamente. Quanto à krigagem, embora não tenha sido melhor que as demais técnicas, houve um desempenho melhor em comparação às demais regressões analisadas anteriormente. Na Tabela 18 encontram-se os coeficientes R^2 para cada saída nas predições realizadas pela regressão múltipla de Ridge.

Tabela 18. Coeficiente de determinação (R^2) para as entradas com a regressão de Ridge.

Condição	Vazão	Tensão	Velocidade	Stickout
Padrão	0,441	0,178	0,116	0,387
SDV 1.5	0,634	0,001	0,243	0,626
SDV 2	0,348	0,314	0,048	0,729
SDV 2.5	0,542	0,003	0,186	0,499
Kriging 1.5	0,067	0,239	0,526	0,008
Kriging 2	0,035	0,150	0,593	0,027
Kriging 2.5	0,021	0,108	0,639	0,030

*Apenas valores > 0,5 são destacados.

Os resultados de R^2 mostram aproximadamente os mesmos resultados encontrados para a regressão polinomial da Tabela 16, indicando que a maioria dos valores (exceto a maioria de *stickout* e uma condição de tensão) de R^2 se aproximam de 0, indicando que o modelo de regressão não consegue explicar uma grande parte da variação na variável dependente, sugerindo um ajuste insatisfatório do modelo aos dados. É importante notar que o R^2 não indica a causalidade entre as variáveis, apenas a força da relação linear entre elas.

4.7.4 – Comparação das Regressões

O erro total somado, E_t , foi escolhido estrategicamente para permitir a comparabilidade com os valores dos produtos. A partir do somatório do MAPE, utilizado neste trabalho, é apresentada a Figura 55, que contém uma comparação para todas as condições testadas para os métodos de regressão.

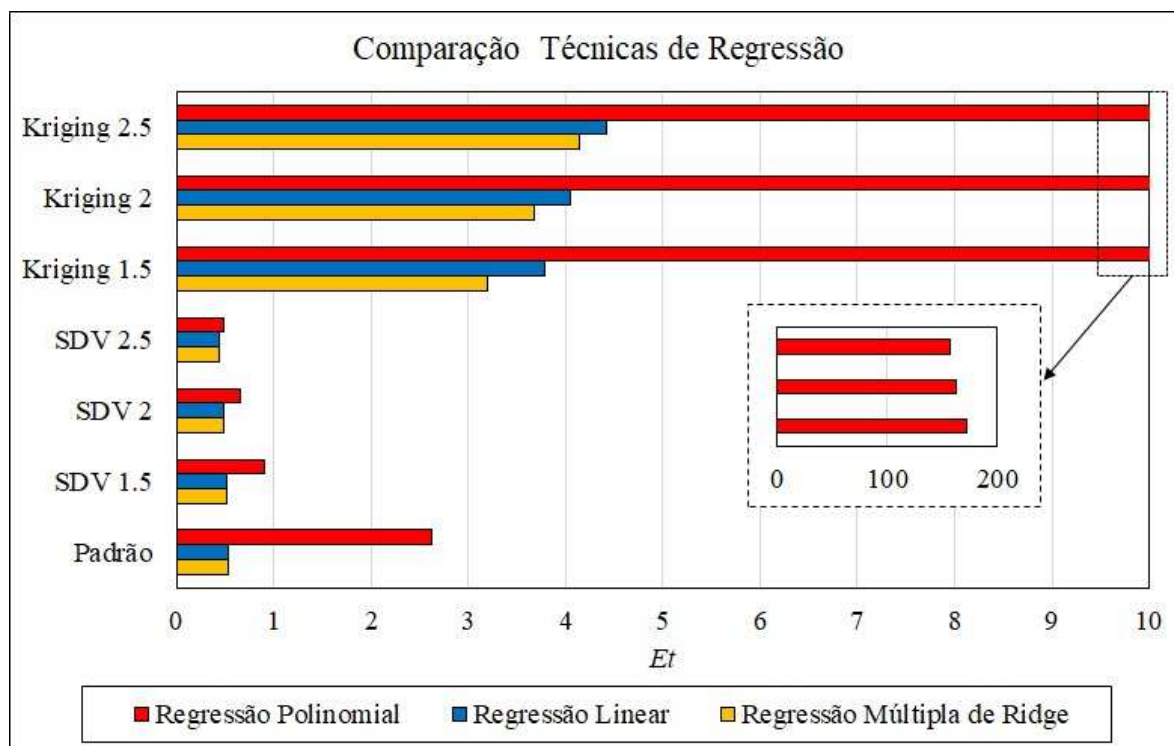


Figura 55. Valores E_t previstos para todas as técnicas de regressão. A área destacada é mostrada em uma escala diferente para visualização adequada dos resultados.

Ao analisar todas as técnicas e condições do conjunto de dados, é possível classificar as seguintes técnicas, do maior para o menor desempenho: regressão múltipla de Ridge; Regressão linear e regressão polinomial, respectivamente. Como o problema de soldagem é considerado complexo e não linear, esperava-se que uma regressão não linear tivesse melhor desempenho que as técnicas de regressão linear. Analisando-se o efeito do aumento dos dados por SDV e krigagem, é possível determinar a seguinte classificação de desempenho de predição: SDV 2.5, SDV 2.0, SDV 1.5, Padrão, krigagem 1.5, krigagem 2.0 e krigagem 2.5, respectivamente.

O SDV foi executado com os melhores valores de predição, melhor que a condição padrão. Os melhores resultados foram encontrados para a condição SDV 2.5. Isso pode indicar o destaque e a viabilidade de aumentar os dados sintéticos para aumentar a precisão dos modelos preditivos nesse caso de soldagem. A técnica de Kriging ainda precisa de desenvolvimento para ser utilizada para o aumento de dados, como comprovado a partir dos piores resultados preditivos apresentados.

Existem algumas razões que podem induzir as previsões com Kriging aos piores resultados. Usar a técnica de Kriging para aumentar artificialmente os dados em uma regressão pode piorar os resultados devido à introdução de ruído adicional e suposições inapropriadas sobre a estrutura espacial dos dados. Extrapolação irrealista e alta dimensionalidade podem produzir valores não representativos, enquanto o *overfitting* pode

reduzir a generalização do modelo. Além disso, a aplicação incorreta de Kriging e informações redundantes podem prejudicar a precisão das previsões. Finalmente, a análise de correlações da Figura 34 mostrou uma divergência do conjunto de dados original, indicando possíveis desafios no aumento de dados por Kriging.

4.8 – Comparação RNA e Regressão

A fim de se comparar as técnicas de regressão e RNA utilizadas é apresentada a Figura 56, que possui os resultados de E_t para todas as condições de aumento de dados e também todas as diferentes técnicas de regressão, além de incluir os melhores resultados obtidos pela RNA e ED para cada condição de *dataset*.

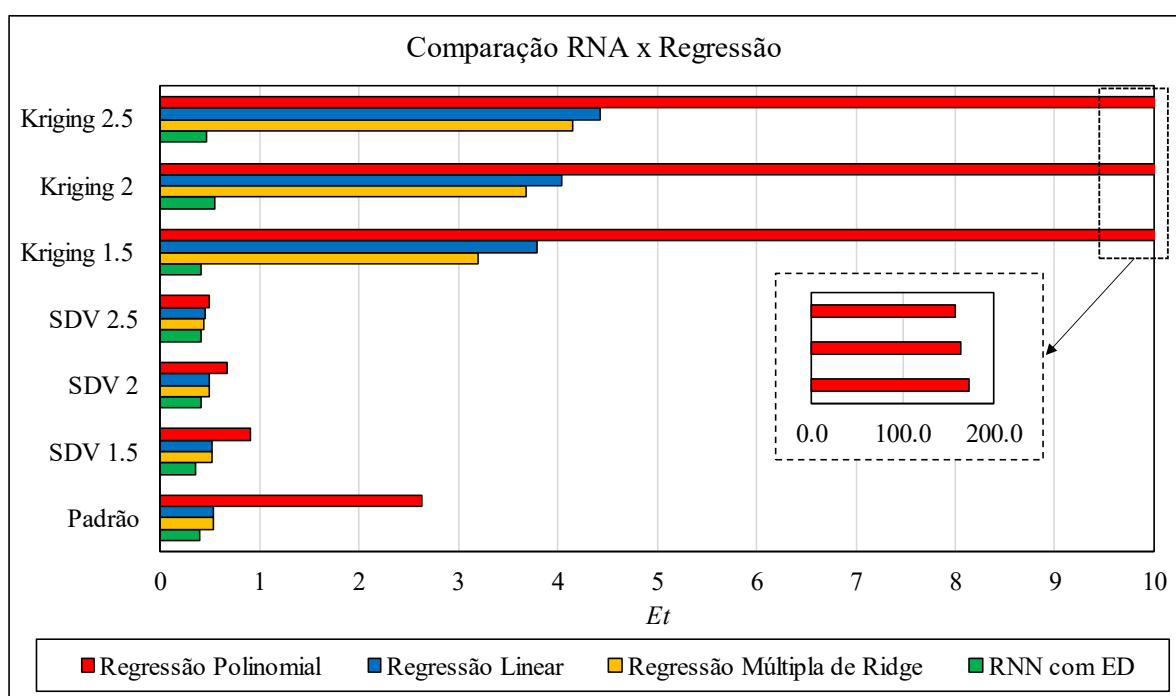


Figura 56. Comparação de valores E_t previstos para todas as técnicas de regressão e os melhores resultados da RNA. A área destacada é mostrada em uma escala diferente para visualização adequada dos resultados.

Como o problema de soldagem é considerado complexo e não linear, esperava-se que uma RNA tivesse um desempenho melhor que as técnicas de regressão. Esse comportamento é melhor observado na técnica de krigagem, na qual as regressões utilizadas não obtiveram resultados precisos, mostrando grande adaptabilidade a problemas não lineares pela RNA.

Os resultados para a RNA com ED da Figura 56 são os melhores encontrados pela técnica, ou seja, o número de iterações para cada condição pode variar, assim como apresentado pela Figura 51.

A partir de todos os resultados analisados é possível então verificar que a RNA possui um desempenho melhor para o *dataset* utilizado, inclusive para todas as condições de aumento de dados para SDV e especialmente para o Kriging. O melhor resultado E_t encontrado foi de 0,355 para a RNA com ED na condição de SDV 1.5, sendo que o melhor para a técnica de regressão foi encontrado na técnica linear com E_t de 0,443 na condição de SDV 2.5. Em ambas as metodologias o aumento de dados por SDV se provou mais preciso que a condição padrão.

Embora as técnicas de regressões sejam mais simples de implantar comparando-se com as RNA's, assim como indicado por Giarollo et al. (2022), os resultados superiores da RNA viabilizam seu uso.

4.9 – Programas Desenvolvidos

Um dos principais resultados desta pesquisa foi o desenvolvimento de um programa de auxílio ao soldador, permitindo a recomendação automatizada dos parâmetros de soldagem com base nas características geométricas da solda. Esse programa utiliza a RNA otimizada por ED para prever os valores ideais de tensão, velocidade de alimentação, *stickout* (ou DBCP – Distância Bico Contato Peça) e vazão de gás, garantindo que a soldagem atinja a qualidade esperada.

O funcionamento do programa é simples e intuitivo: o usuário insere como entrada os valores desejados de largura e altura, espessura de chapa e abertura de raiz e o sistema retorna os parâmetros exatos para configurar a máquina de soldagem MAG. O modelo utiliza os melhores parâmetros encontrados para a condição de dados original.

A Figura 57 apresenta diferentes etapas do programa desenvolvido para auxílio na configuração de parâmetros de soldagem. Em a) há a interface inicial do programa, onde o usuário é recebido e orientado sobre o funcionamento da ferramenta, b) possui inserção de dados fictícios, permitindo ao programa realizar as previsões dos parâmetros ideais para configuração da máquina de soldagem, em c) há o aviso de sucesso nas condições, ou seja, caso atendam aos critérios do modelo, um aviso de sucesso é exibido, indicando que as previsões podem ser geradas e, por fim, em d) são mostrados os parâmetros previstos. Esses valores são baseados no modelo de RNA otimizado e podem ser utilizados diretamente pelo operador.

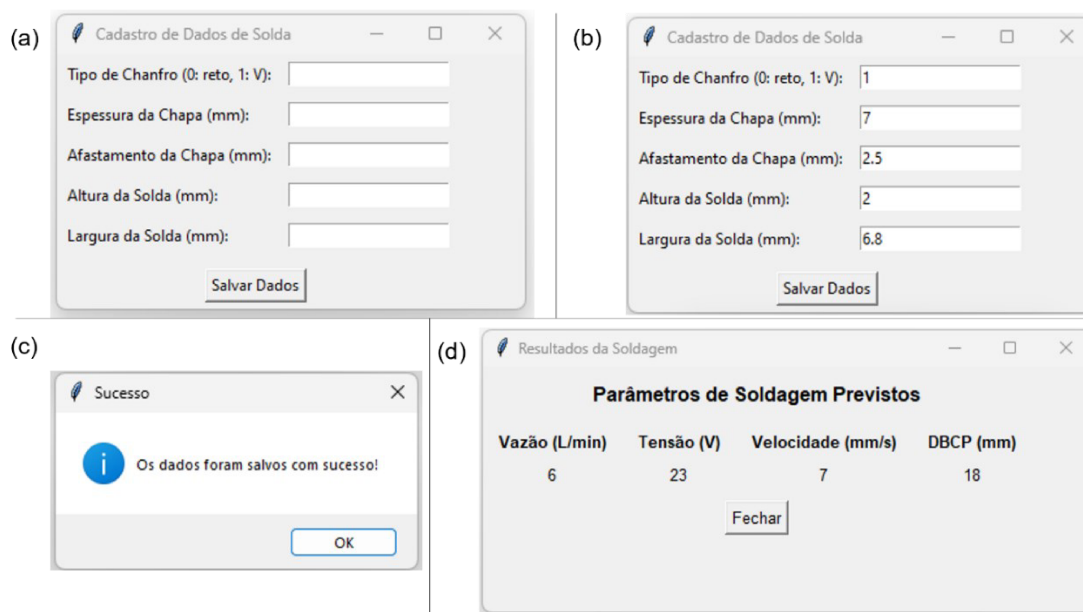


Figura 57. Programa desenvolvido. (a) Apresentação inicial. (b) Inserção de dados fictícios. (c) Aviso de sucesso nas condições. (d) Parâmetros previstos.

Além disso, um dos diferenciais desse programa é sua baixa exigência computacional. Ele pode ser executado em qualquer dispositivo que possua Python instalado, sem necessidade de hardware especializado. O acesso à internet é necessário apenas na primeira execução, para a instalação das dependências e carregamento do modelo otimizado. Após essa etapa, o programa pode rodar *offline*, garantindo acessibilidade e flexibilidade para soldadores em ambientes industriais.

A Tabela 19 apresenta a faixa de valores obtidos nos erros de previsão para o conjunto de dados de testes (5) não vistos pelo programa. O melhor valor apresenta 0,6% de erro enquanto o pior, 25,6%.

Tabela 19. Faixa de resultados de previsão em dados inseridos no programa.

Parâmetros	Vazão (%)	Tensão (%)	Velocidade (%)	DBCP (%)
Faixa de Erros	11,6 - 18,4	0,6 - 9,1	1,3 - 12,8	3,5 - 25,6

Esse resultado representa um avanço significativo na automatização do ajuste de parâmetros de soldagem, tornando o processo mais eficiente, reduzindo desperdícios de material e minimizando a dependência de tentativa e erro para definir configurações ideais.

Dois outros programas foram desenvolvidos. O primeiro exibe os parâmetros otimizados da RNA, permitindo sua extrapolação para diversas áreas que envolvam a previsão de múltiplas entradas e saídas. Já o segundo solicita ao usuário um percentual de aumento do *dataset* pelo SDV, mantendo as mesmas entradas e saídas previamente apresentadas.

A Figura 58 exibe o programa responsável pela determinação dos parâmetros de arquitetura da RNA via ED, onde qualquer ajuste deve ser feito diretamente no código-fonte.

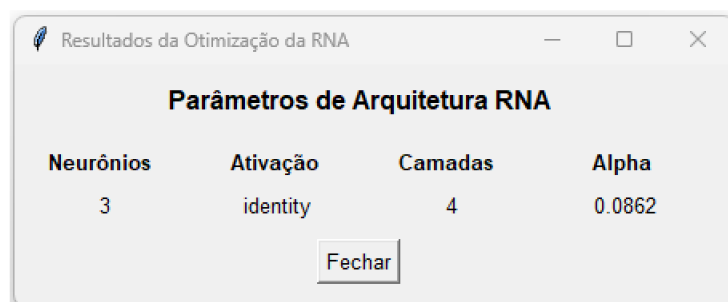


Figura 58. Programa desenvolvido para arquitetura da RNA.

A Figura 59 apresenta o *layout* do programa de previsão de parâmetros de soldagem com aumento de dados pelo SDV. Nas telas a), b) e c), o usuário insere o percentual de aumento desejado e os dados geométricos da solda, enquanto em d) são exibidos os resultados da previsão.

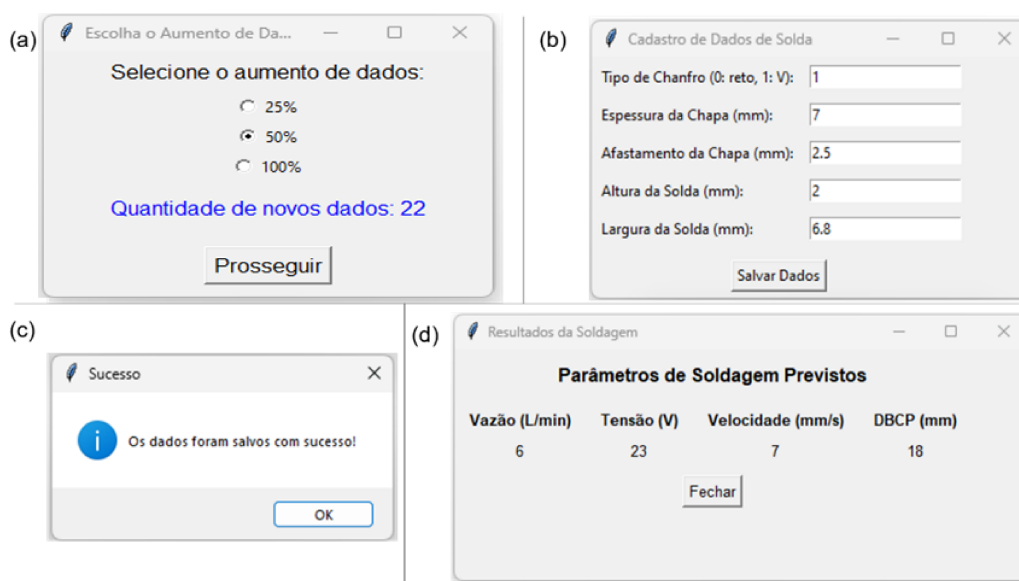


Figura 59. Programa de soldagem desenvolvido com aumento de dados para previsão de parâmetros. (a) Seleção de aumento de dados. (b) Inserção de dados fictícios. (c) Aviso de sucesso nas condições. (d) Parâmetros previstos.

Assim, os programas se mostram ferramentas valiosas para soldadores e engenheiros, auxiliando na padronização e otimização do processo de soldagem MAG, além de possibilitar a aplicação da metodologia em outras áreas de estudo.

5 – CONCLUSÃO

Neste estudo foi proposto o desenvolvimento de uma metodologia inovadora para otimizar variáveis em uma RNA destinada à predição de parâmetros de soldagem no processo GMAW através do uso de ED para otimização dos parâmetros da rede. Utilizou-se de comparações de múltiplas técnicas como regressões e técnicas de validação cruzada que possibilitaram validar o desempenho da predição da RNA. Essas técnicas foram projetadas para incorporar parâmetros de soldagem derivados de variáveis geométricas na soldagem GMAW e a partir das análises realizadas foi possível uma série de avanços descritos a seguir:

- Compreendeu-se a relação física entre os parâmetros de entrada (geometria da peça e geometria do cordão de solda) e os parâmetros de saída (tensão, corrente, velocidade de soldagem, *stickout* e vazão de gás) vendo através de análises de correlações lineares e não lineares que todas as variáveis indicaram alguma correlação.
- A análise de correlações nos dados aumentados serviu como forma de garantir que as mesmas propriedades do *dataset* original fossem mantidas. Apenas os dados gerados pela técnica de Kriging mostraram alguma diferença dos dados originais.
- Analisou-se a influência dos parâmetros da RNA sobre os resultados. A função de ativação logística foi encontrada na maioria dos melhores resultados enquanto o número de neurônios e camadas ocultas variou de 1 a 2 nos melhores resultados.
- Implementou-se um algoritmo de otimização utilizando o método de evolução diferencial (ED), que mostrou eficácia superior, mas convergência mais lenta comparada ao algoritmo SHGO.
- A análise de iteração revelou melhores erros de predição para variáveis de saída, com valores de 0,394 para ED e 0,408 para SHGO.
- A viabilidade de aplicação de RNA's e modelos de regressão em *datasets* pequenos foi comprovada utilizando técnicas de aumento de dados e validação cruzada, melhorando o desempenho preditivo.
- Aumentar o tamanho do conjunto de dados por meios artificiais melhora a precisão preditiva da RNA em dados anteriormente não vistos. O melhor resultado do SDV apresentou 10,1% e o LOO 55,3% de melhora em relação à condição ED padrão.
- Os melhores resultados foram obtidos ao aumentar os dados em 50% utilizando SDV para a RNA com uma camada oculta, dois neurônios e função de ativação logística.
- A técnica SDV superou o Kriging em todas as condições de aumento de dados, com erros de predição aprimorados de 0,355 para SDV e 0,408 para o Kriging. Enquanto

a técnica LOO apresentou a menor taxa de erro registrada em 0,176 na média dos erros das variáveis de saída testadas.

- Resultados da RNA foram validados e comparados com modelos de regressão linear, polinomial e múltipla (Ridge), mostrando superioridade da RNA.
- O aumento de dados artificiais por meio do SDV também melhorou os resultados das predições por regressão linear, mas resultados inferiores aos da RNA.
- A metodologia de definir os parâmetros da arquitetura da RNA antes do teste significou usar uma fase de validação e teste com diferentes conjuntos de dados sendo estas técnicas eficazes para aprimorar os resultados das RNA's sem causar *overfitting*.

Ao avaliar sistematicamente várias técnicas, otimizar arquiteturas de RNA, evitar *overfitting*, comparar com técnicas mais simples e investigar a influência de parâmetros-chave, este estudo contribuiu com a melhoria da previsão de parâmetros de soldagem. As tendências e recomendações identificadas abrem caminhos para o avanço na área de processos de fabricação e modelagem preditiva.

5.1 – Trabalhos Futuros

Como avanços sugeridos para este estudo, indica-se investigar o efeito das técnicas de geração de dados sintéticos em novas áreas referentes a processos de fabricação e avaliação do efeito do aumento de dados em técnicas de previsão. Sugere-se também utilizar o *dataset* disponibilizado nos anexos para implementação de outras técnicas de inteligência artificial ou otimização para aprimorar o desempenho das previsões encontradas.

Além disso, sugere-se empregar a mesma metodologia para determinação de parâmetros ótimos de uma RNA em quaisquer áreas que demandem um sistema de processamento de entradas e saídas. Para isso, há também os programas desenvolvidos que podem auxiliar em novos estudos.

Sugere-se também focar no aspecto de qualidade de solda e microestrutura a partir da mesma metodologia apresentada.

É também sugerido comparar os resultados obtidos com a RNA a partir das técnicas de Kriging para avaliar sua viabilidade em predição direta dos parâmetros.

REFERÊNCIAS

- Abd Ali, J., Hannan, M., & Mohamed, A. (2015). A Novel Quantum-Behaved Lightning Search Algorithm Approach to Improve the Fuzzy Logic Speed Controller for an Induction Motor Drive. *Energies*, 8(11), 13112–13136. <https://doi.org/10.3390/en81112358>
- Abdolrasol, M. G. M., Hussain, S. M. S., Ustun, T. S., Sarker, M. R., Hannan, M. A., Mohamed, R., Ali, J. A., Mekhilef, S., & Milad, A. (2021). Artificial Neural Networks Based Optimization Techniques: A Review. *Electronics*, 10(21), 2689. <https://doi.org/10.3390/electronics10212689>
- Abdul Lateh, M., Kamilah Muda, A., Izzah Mohd Yusof, Z., Azilah Muda, N., & Sanusi Azmi, M. (2017). Handling a Small Dataset Problem in Prediction Model by employ Artificial Data Generation Approach: A Review. *Journal of Physics: Conference Series*, 892, 012016. <https://doi.org/10.1088/1742-6596/892/1/012016>
- Acharya, M. S., Armaan, A., & Antony, A. S. (2019). A Comparison of Regression Models for Prediction of Graduate Admissions. *2019 International Conference on Computational Intelligence in Data Science (ICCIDS)*, 1–5. <https://doi.org/10.1109/ICCIDS.2019.8862140>
- Ahsan, Md. R. U., Kim, Y. R., Kim, C. H., Kim, J. W., Ashiri, R., & Park, Y. D. (2016). Porosity formation mechanisms in cold metal transfer (CMT) gas metal arc welding (GMAW) of zinc coated steels. *Science and Technology of Welding and Joining*, 21, 209–215.
- Ajit, A., Acharya, K., & Samanta, A. (2020). A Review of Convolutional Neural Networks. *2020 International Conference on Emerging Trends in Information Technology and Engineering (Ic-ETITE)*, 1–5. <https://doi.org/10.1109/ic-ETITE47903.2020.049>
- Alkayem, N. F., Parida, B., & Pal, S. (2017). Optimization of friction stir welding process parameters using soft computing techniques. *Soft Computing*, 21(23), 7083–7098. <https://doi.org/10.1007/s00500-016-2251-6>
- Alkayem, N. F., Parida, B., & Pal, S. (2019). Optimization of friction stir welding process using NSGA-II and DEMO. *Neural Computing and Applications*, 31(S2), 947–956. <https://doi.org/10.1007/s00521-017-3059-8>
- Alves Ribeiro, V. (2016). *Soft Computing Techniques for UAVs Classifiers using Micro-Doppler Signatures*.
- Amritpal Singh, & Sushil Kumar. (2016). *Differential Evolution: An Overview* (pp. 209–217). https://doi.org/10.1007/978-981-10-0448-3_17
- Aslam, J., Latif, W., Wasif, M., Hussain, I., & Javaid, S. (2021). Comparison of Regression and Neural Network Model for Short Term Load Forecasting: A Case Study. *The 1st International Conference on Energy, Power and Environment*, 29. <https://doi.org/10.3390/engproc2021012029>

- Asuero, A. G., Sayago, A., & González, A. G. (2006). The Correlation Coefficient: An Overview. *Critical Reviews in Analytical Chemistry*, 36(1), 41–59. <https://doi.org/10.1080/10408340500526766>
- Avery, M. (2010). *Literature Review for Local Polynomial Regression*. <https://api.semanticscholar.org/CorpusID:41160687>
- Aviral Gupta, Akshay Sharma, & Dr. Amita Goel. (2017). Review of Regression Analysis Models. *International Journal of Engineering Research And*, V6(08). <https://doi.org/10.17577/IJERTV6IS080060>
- Azadi Moghaddam, M., & Kolahan, F. (2021). Modeling and optimization of A-GTAW process using back propagation neural network and heuristic algorithms. *International Journal of Pressure Vessels and Piping*, 194, 104531. <https://doi.org/10.1016/j.ijpvp.2021.104531>
- Balasubramanian, K. R., Buvanashakaran, G., & Sankaranarayanan, K. (2010). Modeling of laser beam welding of stainless steel sheet butt joint using neural networks. *CIRP Journal of Manufacturing Science and Technology*, 3(1), 80–84. <https://doi.org/10.1016/j.cirpj.2010.07.001>
- Berrar, D. (2019). Cross-Validation. In *Encyclopedia of Bioinformatics and Computational Biology* (pp. 542–545). Elsevier. <https://doi.org/10.1016/B978-0-12-809633-8.20349-X>
- Bishop, C. M., & Nasrabadi, N. M. (2006). *Pattern recognition and machine learning* (Vol. 4, Issue 4). Springer.
- Bitharas, I., McPherson, N. A., McGhie, W., Roy, D., & Moore, A. J. (2018). Visualisation and optimisation of shielding gas coverage during gas metal arc welding. *Journal of Materials Processing Technology*, 255, 451–462. <https://doi.org/10.1016/j.jmatprotec.2017.11.048>
- Bottou, L., Curtis, F. E., & Nocedal, J. (2018). *Optimization Methods for Large-Scale Machine Learning*.
- Brumm, S., & Bürkner, G. (2015). Gas Metal arc Pulse Welding with Alternating Current for Lightweight Materials. *Materials Today: Proceedings*, 2, S179–S187. <https://doi.org/10.1016/j.matpr.2015.05.008>
- Budu, K. (2014). Comparison of Wavelet-Based ANN and Regression Models for Reservoir Inflow Forecasting. *Journal of Hydrologic Engineering*, 19(7), 1385–1400. [https://doi.org/10.1061/\(ASCE\)HE.1943-5584.0000892](https://doi.org/10.1061/(ASCE)HE.1943-5584.0000892)
- Cary, H. B. (2004). Modern Welding Technology 5/e. *Industrial Robot: An International Journal*, 31(4), 376–376. <https://doi.org/10.1108/ir.2004.31.4.376.3>
- Chacón-Fernández, S., Portolés García, A., & Romaní Labanda, G. (2022). Analysis of the Influence of GMAW Process Parameters on the Properties and Microstructure of S32001 Steel. *Materials*, 15(18), 6498. <https://doi.org/10.3390/ma15186498>

- Chang, Y., Yue, J., Guo, R., Liu, W., & Li, L. (2020). Penetration quality prediction of asymmetrical fillet root welding based on optimized BP neural network. *Journal of Manufacturing Processes*, 50, 247–254. <https://doi.org/10.1016/j.jmapro.2019.12.022>
- Chavda, S. P., Desai, J. V., & Patel, T. M. (2014). A review on optimization of MIG Welding parameters using Taguchi's DOE Method. *International Journal of Engineering and Management Research (IJEMR)*, 4(1), 16–21.
- Che Wan, M. N., Samo, K., Nurazilla, Musa, A., & Muzathik, A. (2011, April). *The Effect of Arc Voltage and Welding Current on Mechanical and Microstructure Properties of 5083-Aluminium Alloy Joints used in Marine Applications*.
- Cheng, J., Wang, P., Li, G., Hu, Q., & Lu, H. (2018). Recent advances in efficient computation of deep convolutional neural networks. *Frontiers of Information Technology & Electronic Engineering*, 19(1), 64–77. <https://doi.org/10.1631/FITEE.1700789>
- Choudhary, A., Kumar, M., Gupta, M. K., Unune, D. K., & Mia, M. (2020). Mathematical modeling and intelligent optimization of submerged arc welding process parameters using hybrid PSO-GA evolutionary algorithms. *Neural Computing and Applications*, 32(10), 5761–5774. <https://doi.org/10.1007/s00521-019-04404-5>
- Chuaiphan, W., & Srijaroenpramong, L. (2020). Effect of hydrogen in argon shielding gas for welding stainless steel grade SUS 201 by GTA welding process. *Journal of Advanced Joining Processes*, 1, 100016. <https://doi.org/10.1016/j.jajp.2020.100016>
- Costa, P. S., Altamirano-Guerrero, G., Ochoa-Palacios, R. M., Reséndiz-Flores, E. O., Guía-Hernández, L. A., & Ramírez-Luna, L. E. (2022). Optimization of welding parameters in underwater wet FCAW on a structural steel using support vector regression and sequential quadratic programming. *The International Journal of Advanced Manufacturing Technology*, 121(5–6), 4225–4236. <https://doi.org/10.1007/s00170-022-09584-4>
- Cressie, N. (1990). The origins of kriging. *Mathematical Geology*, 22(3), 239–252. <https://doi.org/10.1007/BF00889887>
- Cruz, Y. J., Rivas, M., Quiza, R., Villalonga, A., Haber, R. E., & Beruvides, G. (2021). Ensemble of convolutional neural networks based on an evolutionary algorithm applied to an industrial welding process. *Computers in Industry*, 133, 103530. <https://doi.org/10.1016/j.compind.2021.103530>
- Das, S., Abraham, A., & Konar, A. (2009). Differential Evolution Algorithm: Foundations and Perspectives. In *Metaheuristic Clustering* (pp. 63–110). Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-540-93964-1_2
- Dasari, S. K., Cheddad, A., Palmquist, J., & Lundberg, L. (2022). Clustering-based adaptive data augmentation for class-imbalance in machine learning (CADA): additive manufacturing use case. *Neural Computing and Applications*. <https://doi.org/10.1007/s00521-022-07347-6>

- Datta, S., Sundar, M., Bandyopadhyay, A., Nandi, G., Pal, P.K., & Roy, S.C. (2006). Effect of Process Parameters on Bead Geometry, Quality and Performance of Submerged Arc Weldment - Experimental and Statistical Analyses. *Journal of the Mechanical Behavior of Materials*, 17(3), 149–192. <https://doi.org/10.1515/JMBM.2006.17.3.149>
- de Jonge, B., & Scarf, P. A. (2020). A review on maintenance optimization. *European Journal of Operational Research*, 285(3), 805–824. <https://doi.org/10.1016/j.ejor.2019.09.047>
- de Oliveira, G., & Resende Rocha, V. (2021). Use Of Artificial Neural Networks In Fault Detection And Diagnosis In Three-Phase Induction Motors. *Proceedings of the 26th International Congress of Mechanical Engineering*. <https://doi.org/10.26678/ABCM.COBEM2021.COB2021-1608>
- Ding, B., Qian, H., & Zhou, J. (2018). Activation functions and their characteristics in deep neural networks. *2018 Chinese Control And Decision Conference (CCDC)*, 1836–1841. <https://doi.org/10.1109/CCDC.2018.8407425>
- Dorugade, A. V. (2014). New ridge parameters for ridge regression. *Journal of the Association of Arab Universities for Basic and Applied Sciences*, 15(1), 94–99. <https://doi.org/10.1016/j.jaubas.2013.03.005>
- dos Santos Paes, L. E., Andrade, J. R., Lobato, F. S., dos Santos Magalhães, E., Ponomarov, V., de Souza, F. J., & Vilarinho, L. O. (2022). Sensitivity analysis and multi-objective optimization of tungsten inert gas (TIG) welding based on numerical simulation. *The International Journal of Advanced Manufacturing Technology*, 122(2), 783–797. <https://doi.org/10.1007/s00170-022-09934-2>
- e José Alexandre Silva Junior, D. F. F. (2010). Desvendando os Mistérios do Coeficiente de Correlação de Pearson (r). *Revista Política Hoje*, 18(1). <https://periodicos.ufpe.br/revistas/politica hoje/article/view/3852>
- Edwin Raja Dhas, J., Anton Savio Lewise, K., & Laxmi, G. (2022). Submerged arc welding process parameter prediction using predictive modeling techniques. *Materials Today: Proceedings*, 64, 402–409. <https://doi.org/10.1016/j.matpr.2022.04.757>
- Eilenberg, S., & Steenrod, N. (2015). *Foundations of algebraic topology* (Vol. 2193). Princeton University Press.
- Endres, S. C., Sandrock, C., & Focke, W. W. (2018). A simplicial homology algorithm for Lipschitz optimisation. *Journal of Global Optimization*, 72(2), 181–217. <https://doi.org/10.1007/s10898-018-0645-y>
- Engström, E., Runeson, P., & Skoglund, M. (2010). A systematic review on regression test selection techniques. *Information and Software Technology*, 52(1), 14–30. <https://doi.org/10.1016/j.infsof.2009.07.001>

- Erzurum Cicek, Z. I., & Kamisli Ozturk, Z. (2021). Optimizing the artificial neural network parameters using a biased random key genetic algorithm for time series forecasting. *Applied Soft Computing*, 102, 107091. <https://doi.org/10.1016/j.asoc.2021.107091>
- Fan, Y. Y., Yang, C. L., Lin, S. B., Fan, C. L., & Liu, W. G. (2012). Ultrasonic wave assisted GMAW. *Weld J*, 91(3), 91s–99s.
- Feng, S., Zhou, H., & Dong, H. (2019). Using deep neural network with small dataset to predict material defects. *Materials & Design*, 162, 300–310. <https://doi.org/10.1016/j.matdes.2018.11.060>
- Flauzino, R. A., Silva, I. N., & Spatti, D. H. (2016). *Redes Neurais Artificiais para Engenharia e Ciências Aplicadas: Fundamentos Teóricos e Aspectos Práticos*. Sao Paulo: Artliber Editora Ltda.
- Fleck, L., Tavares, M. H. F., Eyng, E., Helmann, A. C., & Andrade, M. A. de M. (2016). Redes neurais artificiais: Princípios básicos. *Revista Eletrônica Científica Inovação e Tecnologia*, 1(13), 47–57.
- Gao, Z., Shao, X., Jiang, P., Cao, L., Zhou, Q., Yue, C., Liu, Y., & Wang, C. (2016). Parameters optimization of hybrid fiber laser-arc butt welding on 316L stainless steel using Kriging model and GA. *Optics & Laser Technology*, 83, 153–162. <https://doi.org/10.1016/j.optlastec.2016.04.001>
- Giarollo, D. F., Hackenhaar, W., Mazzaferro, C. C. P., & Mazzaferro, J. A. E. (2022). Bead Geometry Prediction in Pulsed GMAW Welding: A Comparative Study on the Performance of Artificial Neural Network and Regression Models. *Soldagem & Inspeção*, 27. <https://doi.org/10.1590/0104-9224/si27.22>
- Goh, A. T. C. (1995). Back-propagation neural networks for modeling complex systems. *Artificial Intelligence in Engineering*, 9(3), 143–151. [https://doi.org/10.1016/0954-1810\(94\)00011-S](https://doi.org/10.1016/0954-1810(94)00011-S)
- Goodfellow, I., Bengio, Y., & Courville, A. (2016). *Deep learning*. MIT press.
- Gregores Coto, A., Precker, C. E., Andersson, T., Laukkanen, A., Suhonen, T., Rodriguez, P. R., & Munos-Landon, S. (2023). The use of generative models to speed up the discovery of materials. *Computer Methods in Material Science*, 23(1). <https://doi.org/10.7494/cmms.2023.1.0802>
- Gronau, Q. F., & Wagenmakers, E.-J. (2019). Limitations of Bayesian Leave-One-Out Cross-Validation for Model Selection. *Computational Brain & Behavior*, 2(1), 1–11. <https://doi.org/10.1007/s42113-018-0011-7>
- Hannan, M. A., Lipu, M. S. H., Hussain, A., Saad, M. H., & Ayob, A. (2018). Neural Network Approach for Estimating State of Charge of Lithium-Ion Battery Using Backtracking Search Algorithm. *IEEE Access*, 6, 10069–10079. <https://doi.org/10.1109/ACCESS.2018.2797976>

- Hastie, T., Tibshirani, R., Friedman, J. H., & Friedman, J. H. (2009). *The elements of statistical learning: data mining, inference, and prediction* (Vol. 2). Springer.
- Hayat, F. (2011). The effects of the welding current on heat input, nugget geometry, and the mechanical and fractural properties of resistance spot welding on Mg/Al dissimilar materials. *Materials & Design*, 32(4), 2476–2484. <https://doi.org/10.1016/j.matdes.2010.11.015>
- Haykin, S. (1998). *Neural Networks: A Comprehensive Foundation* (2nd ed.). Prentice Hall PTR.
- He, G., Zhao, Y., & Yan, C. (2022). Application of tabular data synthesis using generative adversarial networks on machine learning-based multiaxial fatigue life prediction. *International Journal of Pressure Vessels and Piping*, 199, 104779. <https://doi.org/10.1016/j.ijpvp.2022.104779>
- He, K., Zhang, X., Ren, S., & Sun, J. (2015). *Delving Deep into Rectifiers: Surpassing Human-Level Performance on ImageNet Classification*. <https://arxiv.org/abs/1502.01852>
- HECHT-NIELSEN, R. (1992). Theory of the Backpropagation Neural Network**Based on “nonindent” by Robert Hecht-Nielsen, which appeared in Proceedings of the International Joint Conference on Neural Networks 1, 593–611, June 1989. © 1989 IEEE. In *Neural Networks for Perception* (pp. 65–93). Elsevier. <https://doi.org/10.1016/B978-0-12-741252-8.50010-8>
- Hengl, T., Heuvelink, G. B. M., & Rossiter, D. G. (2007). About regression-kriging: From equations to case studies. *Computers & Geosciences*, 33(10), 1301–1315. <https://doi.org/10.1016/j.cageo.2007.05.001>
- Henle, M. (1994). *A Combinatorial Introduction to Topology* (Vol. 1). Courier Corporation.
- Hittmeir, M., Ekelhart, A., & Mayer, R. (2019). On the Utility of Synthetic Data. *Proceedings of the 14th International Conference on Availability, Reliability and Security*, 1–6. <https://doi.org/10.1145/3339252.3339281>
- Hong-Wei, N., Nguyen, V. D., Vonikakis, V., & Winkler, S. (2015). Deep Learning for Emotion Recognition on Small Datasets using Transfer Learning. *Proceedings of the 2015 ACM on International Conference on Multimodal Interaction*, 443–449. <https://doi.org/10.1145/2818346.2830593>
- Howard B. Cary, & Scott C. Helzer. (2005). Modern Welding Technology. In *Pearson/Prentice Hall* (Vol. 6).
- Hu, Q., Zhang, L., Pu, J., & Zhu, C. (2020). Effect of shielding gas flow rate on arc behaviors in GMAW with single cable-typed wire. *International Journal of Modern Physics B*, 34(01n03), 2040059. <https://doi.org/10.1142/S0217979220400597>

- Hu, Z., Xiong, S., Su, Q., & Zhang, X. (2013). Sufficient Conditions for Global Convergence of Differential Evolution Algorithm. *Journal of Applied Mathematics*, 2013, 1–14. <https://doi.org/10.1155/2013/193196>
- Ibrahim, I. A., Mohamat, S. A., Amir, A., & Ghalib, A. (2012). The Effect of Gas Metal Arc Welding (GMAW) Processes on Different Welding Parameters. *Procedia Engineering*, 41, 1502–1506. <https://doi.org/10.1016/j.proeng.2012.07.342>
- Itaborahy Filho, M. A. (2018). *Análise de algoritmos genéticos e evolução diferencial para otimização de funções não-lineares multimodais*. Universidade Tecnológica Federal do Paraná.
- Jain, S., Seth, G., Paruthi, A., Soni, U., & Kumar, G. (2022). Synthetic data augmentation for surface defect detection and classification using deep learning. *Journal of Intelligent Manufacturing*, 33(4), 1007–1020. <https://doi.org/10.1007/s10845-020-01710-x>
- Jaypuria, S., Gupta, S. K., & Pratihari, D. K. (2020). *Comparative Study of Feed-Forward and Recurrent Neural Networks in Modeling of Electron Beam Welding* (pp. 521–531). https://doi.org/10.1007/978-981-32-9433-2_45
- Jiang, P., Cao, L., Zhou, Q., Gao, Z., Rong, Y., & Shao, X. (2016). Optimization of welding process parameters by combining Kriging surrogate with particle swarm optimization algorithm. *The International Journal of Advanced Manufacturing Technology*, 86(9–12), 2473–2483. <https://doi.org/10.1007/s00170-016-8382-1>
- Jiang, Q., Huang, R., Huang, Y., Chen, S., He, Y., Lan, L., & Liu, C. (2019). Application of BP Neural Network Based on Genetic Algorithm Optimization in Evaluation of Power Grid Investment Risk. *IEEE Access*, 7, 154827–154835. <https://doi.org/10.1109/ACCESS.2019.2944609>
- Jian-wei, M. (2004). Optimization of feed-forward neural networks based on artificial fish-swarm algorithm. *Computer Engineering*.
- Jindal, S., Chhibber, R., & Mehta, N. P. (2014). Effect of welding parameters on bead profile, microhardness and H₂ content in submerged arc welding of high-strength low-alloy steel. *Proceedings of the Institution of Mechanical Engineers, Part B: Journal of Engineering Manufacture*, 228(1), 82–94. <https://doi.org/10.1177/0954405413495846>
- Joseph, A., Farson, D., Harwig, D., & Richardson, R. (2005). Influence of GMAW-P current waveforms on heat input and weld bead shape. *Science and Technology of Welding and Joining*, 10(3), 311–318. <https://doi.org/10.1179/174329305X40624>
- Kannan, R., & Nandwana, P. (2023). Accelerated alloy discovery using synthetic data generation and data mining. *Scripta Materialia*, 228, 115335. <https://doi.org/10.1016/j.scriptamat.2023.115335>

- Karadeniz, E., Ozsarac, U., & Yildiz, C. (2007). The effect of process parameters on penetration in gas metal arc welding processes. *Materials & Design*, 28(2), 649–656. <https://doi.org/10.1016/j.matdes.2005.07.014>
- Khan, A. H., Cao, X., Li, S., Katsikis, V. N., & Liao, L. (2020). BAS-ADAM: an ADAM based approach to improve the performance of beetle antennae search optimizer. *IEEE/CAA Journal of Automatica Sinica*, 7(2), 461–471. <https://doi.org/10.1109/JAS.2020.1003048>
- Kim, Y., Lee, T., Hyun, Y., Coatanea, E., Mika, S., Mo, J., & Yoo, Y. (2023). Self-supervised representation learning anomaly detection methodology based on boosting algorithms enhanced by data augmentation using StyleGAN for manufacturing imbalanced data. *Computers in Industry*, 153, 104024. <https://doi.org/10.1016/j.compind.2023.104024>
- Kleijnen, J. P. C. (2017). Regression and Kriging metamodels with their experimental designs in simulation: A review. *European Journal of Operational Research*, 256(1), 1–16. <https://doi.org/10.1016/j.ejor.2016.06.041>
- Konak, A., Coit, D. W., & Smith, A. E. (2006). Multi-objective optimization using genetic algorithms: A tutorial. *Reliability Engineering & System Safety*, 91(9), 992–1007. <https://doi.org/10.1016/j.ress.2005.11.018>
- Kosorok, M. R. (2009). Discussion of: Brownian distance covariance. *The Annals of Applied Statistics*, 3(4). <https://doi.org/10.1214/09-AOAS312B>
- Kroll, C. N., & Song, P. (2013). Impact of multicollinearity on small sample hydrologic regression models. *Water Resources Research*, 49(6), 3756–3769. <https://doi.org/10.1002/wrcr.20315>
- Kumar, R., Aggarwal, R. K., & Sharma, J. D. (2015). Comparison of regression and artificial neural network models for estimation of global solar radiations. *Renewable and Sustainable Energy Reviews*, 52, 1294–1299. <https://doi.org/10.1016/j.rser.2015.08.021>
- Kumar, R., Ramesh Mevada, N., Rathore, S., Agarwal, N., Rajput, V., & Barad, A. S. (2017). Experimental Investigation and Optimization of TIG Welding Parameters on Aluminum 6061 Alloy Using Firefly Algorithm. *IOP Conference Series: Materials Science and Engineering*, 225, 012153. <https://doi.org/10.1088/1757-899X/225/1/012153>
- Kumar, V., Parida, M. K., & Albert, S. K. (2022). The state-of-the-art methodologies for quality analysis of arc welding process using weld data acquisition and analysis techniques. *International Journal of System Assurance Engineering and Management*, 13(1), 34–56. <https://doi.org/10.1007/s13198-021-01282-w>
- Lee, J. I., & Rhee, S. (2000). Prediction of process parameters for gas metal arc welding by multiple regression analysis. *Proceedings of the Institution of Mechanical Engineers, Part B: Journal of Engineering Manufacture*, 214(6), 443–449. <https://doi.org/10.1243/0954405001517757>

- Lei, T., Rong, Y., Xu, J., & Huang, Y. (2018). Experiment study and regression analysis of molten pool in laser welding. *Optics & Laser Technology*, 108, 534–541. <https://doi.org/10.1016/j.optlastec.2018.07.053>
- Lei, Z., Shen, J., Wang, Q., & Chen, Y. (2019). Real-time weld geometry prediction based on multi-information using neural network optimized by PCA and GA during thin-plate laser welding. *Journal of Manufacturing Processes*, 43, 207–217. <https://doi.org/10.1016/j.jmapro.2019.05.013>
- Leshno, M., & Spector, Y. (1996). Neural network prediction analysis: The bankruptcy case. *Neurocomputing*, 10(2), 125–147. [https://doi.org/10.1016/0925-2312\(94\)00060-3](https://doi.org/10.1016/0925-2312(94)00060-3)
- Li, C., Zhu, D., Hu, C., Li, X., Nan, S., & Huang, H. (2023). ECDX: Energy consumption prediction model based on distance correlation and XGBoost for edge data center. *Information Sciences*, 643, 119218. <https://doi.org/10.1016/j.ins.2023.119218>
- Li, L., Wang, P., Ren, J., Lü, Z., Li, X., Gao, H., & Di, R. (2024). Synthetic data augmentation for high-resolution X-ray welding defect detection and classification based on a small number of real samples. *Engineering Applications of Artificial Intelligence*, 133, 108379. <https://doi.org/10.1016/j.engappai.2024.108379>
- Liu, D. C., & Nocedal, J. (1989). On the limited memory BFGS method for large scale optimization. *Mathematical Programming*, 45(1–3), 503–528. <https://doi.org/10.1007/BF01589116>
- Liu, Z., & Wang, H. (2022). A data augmentation based Kriging-assisted reference vector guided evolutionary algorithm for expensive dynamic multi-objective optimization. *Swarm and Evolutionary Computation*, 75, 101173. <https://doi.org/10.1016/j.swevo.2022.101173>
- Liu, Z., Yang, Y., Han, C., Zhou, H., Zhou, H., Wang, M., Liu, L., Wang, H., Bai, Y., & Wang, D. (2023). Effects of gas flow parameters on droplet spatter features and dynamics during large-scale laser powder bed fusion. *Materials & Design*, 225, 111534. <https://doi.org/10.1016/j.matdes.2022.111534>
- Lobato, F. S., & Steffen Jr, V. (2016). Desenvolvimento de um Algoritmo de Otimização Auto-Adaptativo para a Determinação de um Protocolo Otimizado para a Administração de Drogas no Tratamento de Tumores. *TEMA (São Carlos)*, 17(2), 225. <https://doi.org/10.5540/tema.2016.017.02.0225>
- Luesak, P., Pitakaso, R., Sethanan, K., Golinska-Dawson, P., Srichok, T., & Chokanat, P. (2023). Multi-Objective Modified Differential Evolution Methods for the Optimal Parameters of Aluminum Friction Stir Welding Processes of AA6061-T6 and AA5083-H112. *Metals*, 13(2), 252. <https://doi.org/10.3390/met13020252>
- Lyons, R. (2013). Distance covariance in metric spaces. *The Annals of Probability*, 41(5). <https://doi.org/10.1214/12-AOP803>

- Mallaradhya, H. M., Kumar, M. V., & Chandra, M. S. V. (2022). Optimization of Parameters and Prediction of Response Values Using Regression and ANN Model in Resistance Spot Welding of 17-4 Precipitation Hardened Stainless Steel. *Journal of Advanced Manufacturing Systems*, 21(02), 275–291. <https://doi.org/10.1142/S0219686722500111>
- Mallipeddi, R., & Suganthan, P. N. (2008). Empirical study on the effect of population size on Differential evolution Algorithm. *2008 IEEE Congress on Evolutionary Computation (IEEE World Congress on Computational Intelligence)*, 3663–3670. <https://doi.org/10.1109/CEC.2008.4631294>
- Mandal, N. R. (2017). *Welding Parameters* (pp. 159–174). https://doi.org/10.1007/978-981-10-2955-4_14
- Maroufpoor, S., Azadnia, R., & Bozorg-Haddad, O. (2020). Stochastic optimization. In *Handbook of Probabilistic Models* (pp. 437–448). Elsevier. <https://doi.org/10.1016/B978-0-12-816514-0.00017-5>
- MARQUES, P. V., MODENESI, P. J., & BRACARENSE, A. Q. (2017). Soldagem fundamentos e tecnologia. atual. *Belo Horizonte: Editora UFMG*.
- Maulud, D., & Abdulazeez, A. M. (2020). A Review on Linear Regression Comprehensive in Machine Learning. *Journal of Applied Science and Technology Trends*, 1(4), 140–147. <https://doi.org/10.38094/jastt1457>
- McDonald, G. C. (2009). Ridge regression. *WIREs Computational Statistics*, 1(1), 93–100. <https://doi.org/10.1002/wics.14>
- Mendonça, M., Fionocchio, M. A. F., Gusmão, R. V., & Chrun, I. R. (2016). ARTIFICIAL NEURAL NETWORKS APPLIED IN IMAGE EXTRACTION FOR AUTONOMOUS CLASSIFICATION OF WELD BEADS. *Revista Científica On-Line Tecnologia – Gestão – Humanismo - ISSN: 2238-5819*, 6(2).
- Mezaache, M., Babes, B., & Chaouch, S. (2022). Optimization of Welding Input Parameters Using PSO Technique for Minimizing HAZ Width in GMAW. *Periodica Polytechnica Mechanical Engineering*, 66(2), 99–108. <https://doi.org/10.3311/PPme.14127>
- Modenesi, P. J., Marques, P. V., & Santos, D. B. (2012). *Introdução à Metalurgia da Soldagem* (Vol. 1). UFMG.
- Monika, K., & Chauhan, J. (2017). Monika, & Chauhan, J. (2017). A REVIEW PAPER ON GAS METAL ARC WELDING (GMAW) OF MILD STEEL 1018 BY USING TAGUCHI TECHNIQUE. *INTERNATIONAL JOURNAL OF CURRENT ENGINEERING AND SCIENTIFIC RESEARCH (IJCESR)*, 4(7).
- Moore, P., & Booth, G. (2015). Woodhead Publishing Series in Welding and Other Joining Technologies. In *The Welding Engineers Guide to Fracture and Fatigue* (pp. xi–xv). Elsevier. <https://doi.org/10.1016/B978-1-78242-370-6.50017-4>

- Moreno-Barea, F. J., Jerez, J. M., & Franco, L. (2020). Improving classification accuracy using data augmentation on small data sets. *Expert Systems with Applications*, 161, 113696. <https://doi.org/10.1016/j.eswa.2020.113696>
- Nah, S., Son, S., & Lee, K. M. (2019, June). Recurrent Neural Networks With Intra-Frame Iterations for Video Deblurring. *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*.
- Najwa Mohd Rizal, N., Hayder, G., Mnzool, M., Elnaïm, B. M. E., Mohammed, A. O. Y., & Khayyat, M. M. (2022). Comparison between Regression Models, Support Vector Machine (SVM), and Artificial Neural Network (ANN) in River Water Quality Prediction. *Processes*, 10(8), 1652. <https://doi.org/10.3390/pr10081652>
- Natesh Nishanth T, & K. Venkatesan. (2022). A Literature Review on Gas Metal Arc Welding (GMAW). *INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH IN TECHNOLOGY*, 9(7), 370–375.
- Oh, Y., Park, S., & Ye, J. C. (2020). Deep Learning COVID-19 Features on CXR Using Limited Training Data Sets. *IEEE Transactions on Medical Imaging*, 39(8), 2688–2700. <https://doi.org/10.1109/TMI.2020.2993291>
- Oliveira, G. T. da S., & others. (2006). *Estudo e aplicações da evolução diferencial*. UFU.
- OLIVER, M. A., & WEBSTER, R. (1990). Kriging: a method of interpolation for geographical information systems. *International Journal of Geographical Information Systems*, 4(3), 313–332. <https://doi.org/10.1080/026937990008941549>
- Pan, H., Pang, Z., Wang, Y., Wang, Y., & Chen, L. (2020). A New Image Recognition and Classification Method Combining Transfer Learning Algorithm and MobileNet Model for Welding Defects. *IEEE Access*, 8, 119951–119960. <https://doi.org/10.1109/ACCESS.2020.3005450>
- Panda, B. N., Bahubalendruni, M. V. A. R., & Biswal, B. B. (2014). Optimization of resistance spot welding parameters using differential evolution algorithm and GRNN. *2014 IEEE 8th International Conference on Intelligent Systems and Control (ISCO)*, 50–55. <https://doi.org/10.1109/ISCO.2014.7103917>
- Patki, N., Wedge, R., & Veeramachaneni, K. (2016). The Synthetic Data Vault. *2016 IEEE International Conference on Data Science and Advanced Analytics (DSAA)*, 399–410. <https://doi.org/10.1109/DSAA.2016.49>
- Peng, B., Yang, F., Yan, C., Zeng, X., & Zhou, D. (2016). Efficient multiple starting point optimization for automated analog circuit optimization via recycling simulation data. *2016 Design, Automation & Test in Europe Conference & Exhibition (DATE)*, 1417–1422.
- Pires, I., Quintino, L., & Miranda, R. M. (2007). Analysis of the influence of shielding gas mixtures on the gas metal arc welding metal transfer modes and fume formation rate. *Materials & Design*, 28(5), 1623–1631. <https://doi.org/10.1016/j.matdes.2006.02.012>

- Price, K., Storn, R. M., & Lampinen, J. A. (2006). *Differential evolution: a practical approach to global optimization*. Springer Science & Business Media.
- Pytlak, R., & Tarnawski, T. (2007). Limited memory quasi-Newton methods for problems with box constraints. *2007 American Control Conference*, 1257–1262. <https://doi.org/10.1109/ACC.2007.4282546>
- Rasamoelina, A. D., Adjailia, F., & Sincak, P. (2020). A Review of Activation Function for Artificial Neural Network. *2020 IEEE 18th World Symposium on Applied Machine Intelligence and Informatics (SAMI)*, 281–286. <https://doi.org/10.1109/SAMI48414.2020.9108717>
- Ren, J., Zhao, H., Zhang, L., Zhao, Z., Xu, Y., Cheng, Y., Wang, M., Chen, J., & Wang, J. (2022). Design optimization of cement grouting material based on adaptive boosting algorithm and simplicial homology global optimization. *Journal of Building Engineering*, 49, 104049. <https://doi.org/10.1016/j.jobbe.2022.104049>
- Ribeiro, C. R. (2018). *Uso de Redes Neurais Artificiais como Ferramenta Computacional em Processos de Soldagem*. Unifei.
- Rocha, V. R. (2024, January 8). *Topological Optimization of a Formula Student Monocoque Chassis with Fixed Torsional Stiffness and Element Size Influence in Ansys*. <https://doi.org/10.4271/2023-36-0103>
- Rocha, V. R., Rosa, D. M., & dos Santos Paes, L. E. (2024). Effect of previous heating treatments conditions and gas metal arc welding (GMAW) on mechanical and microstructural properties of HSLA steel DOMEX 460MC. *Journal of the Brazilian Society of Mechanical Sciences and Engineering*, 46(8), 475. <https://doi.org/10.1007/s40430-024-05010-w>
- Romero-Hdz, J., Saha, B., Toledo-Ramirez, G., & Beltran-Bqz, D. (2016). Welding Sequence Optimization using Artificial Intelligence Techniques, an Overview. *International Journal of Computer Science and Engineering*, 3(11), 90–95. <https://doi.org/10.14445/23488387/IJCSE-V3I11P115>
- Rong, Y., Zhang, G., Chang, Y., & Huang, Y. (2016). Integrated optimization model of laser brazing by extreme learning machine and genetic algorithm. *The International Journal of Advanced Manufacturing Technology*, 87(9–12), 2943–2950. <https://doi.org/10.1007/s00170-016-8649-6>
- Ruan, X., Zhou, Q., Shu, L., Hu, J., & Cao, L. (2018). Accurate Prediction of the Weld Bead Characteristic in Laser Keyhole Welding Based on the Stochastic Kriging Model. *Metals*, 8(7), 486. <https://doi.org/10.3390/met8070486>
- Ruediger-Flore, P., Glatt, M., Hussong, M., & Aurich, J. C. (2023). CAD-based data augmentation and transfer learning empowers part classification in manufacturing. *The*

- International Journal of Advanced Manufacturing Technology*, 125(11–12), 5605–5618.
<https://doi.org/10.1007/s00170-023-10973-6>
- Sadeeq, M. A. M., & Abdulazeez, A. M. (2020). Neural Networks Architectures Design, and Applications: A Review. *2020 International Conference on Advanced Science and Engineering (ICOASE)*, 199–204. <https://doi.org/10.1109/ICOASE51841.2020.9436582>
- Saeheaw, T. (2023). Comparison of different supervised machine learning algorithms for bead geometry prediction in GMAW process. *Engineering Solid Mechanics*, 11(2), 175–190. <https://doi.org/10.5267/j.esm.2022.12.003>
- Sathiya, P., Panneerselvam, K., & Soundararajan, R. (2012). Optimal design for laser beam butt welding process parameter using artificial neural networks and genetic algorithm for super austenitic stainless steel. *Optics & Laser Technology*, 44(6), 1905–1914. <https://doi.org/10.1016/j.optlastec.2012.01.025>
- Schmidhuber, J. (2015). Deep learning in neural networks: an overview. *Neural Networks : The Official Journal of the International Neural Network Society*, 61, 85–117. <https://doi.org/10.1016/j.neunet.2014.09.003>
- Scotti, A., & Ponomarev, V. (2008). *Soldagem MIG/MAG: melhor entendimento, melhor desempenho* (2nd ed.). Artliber Editora Ltda.
- Sevvel, P., Satheesh, C., & Senthil Kumar, R. (2020). Generation of regression models and multi-response optimization of friction stir welding technique parameters during the fabrication of AZ80A Mg alloy joints. *Transactions of the Canadian Society for Mechanical Engineering*, 44(2), 311–324. <https://doi.org/10.1139/tcsme-2019-0162>
- Shah, P., Khankhoje, U. K., & Moghaddam, M. (2016). Inverse Scattering Using a Joint Norm-Based Regularization. *IEEE Transactions on Antennas and Propagation*, 64(4), 1373–1384. <https://doi.org/10.1109/TAP.2016.2529641>
- Shareef, H., Ibrahim, A. A., & Mutlag, A. H. (2015). Lightning search algorithm. *Applied Soft Computing*, 36, 315–333. <https://doi.org/10.1016/j.asoc.2015.07.028>
- Sheng, W., Shan, P., Mao, J., Zheng, Y., Chen, S., & Wang, Z. (2017). An Adaptive Memetic Algorithm With Rank-Based Mutation for Artificial Neural Network Architecture Optimization. *IEEE Access*, 5, 18895–18908. <https://doi.org/10.1109/ACCESS.2017.2752901>
- Shevchenko, R. A., Kozyrev, N. A., Kryukov, R. E., Patrushev, A. O., & Usoltsev, A. A. (2018). Multi-factor regression analysis of the process of rails contact welding on K1000 machine. *IOP Conference Series: Materials Science and Engineering*, 411, 012088. <https://doi.org/10.1088/1757-899X/411/1/012088>
- Shin, H.-C., Roth, H. R., Gao, M., Lu, L., Xu, Z., Nogues, I., Yao, J., Mollura, D., & Summers, R. M. (2016). Deep Convolutional Neural Networks for Computer-Aided Detection: CNN

- Architectures, Dataset Characteristics and Transfer Learning. *IEEE Transactions on Medical Imaging*, 35(5), 1285–1298. <https://doi.org/10.1109/TMI.2016.2528162>
- Shorten, C., & Khoshgoftaar, T. M. (2019). A survey on Image Data Augmentation for Deep Learning. *Journal of Big Data*, 6(1), 60. <https://doi.org/10.1186/s40537-019-0197-0>
- Silva, R. H. G. e, Nogueira, R. R., Silva, F. L. dos S. e, Santos, A. G. M., & Brito, U. H. de O. de. (2020). Metallic Droplet, Weld Pool Temperatures Evaluation and Pulsed Wave Formulation in the Pulsed GMAW. *Materials Research*, 23(4). <https://doi.org/10.1590/1980-5373-mr-2020-0074>
- Singh, R. (2020). Welding and joining processes. In *Applied Welding Engineering* (pp. 157–186). Elsevier. <https://doi.org/10.1016/B978-0-12-821348-3.00015-X>
- Snyman, J. A., & Wilke, D. N. (2018). *Practical Mathematical Optimization*. 133. Springer International Publishing.
- Soeiro Junior, J. C., Luz, M. A. da, & Brandi, S. D. (2015). Comparação da Taxa e Eficiência de Deposição entre os Consumíveis ER70S-6 e E71T-1C. *Soldagem & Inspeção*, 20(1), 2–15. <https://doi.org/10.1590/0104-9224/SI2001.02>
- Song, X., Sun, G., & Li, Q. (2016). Sensitivity analysis and reliability based design optimization for high-strength steel tailor welded thin-walled structures under crashworthiness. *Thin-Walled Structures*, 109, 132–142. <https://doi.org/10.1016/j.tws.2016.09.003>
- Srichok, T., Pitakaso, R., Sethanan, K., Sirirak, W., & Kwangmuang, P. (2020). Combined Response Surface Method and Modified Differential Evolution for Parameter Optimization of Friction Stir Welding. *Processes*, 8(9), 1080. <https://doi.org/10.3390/pr8091080>
- Srivastava, N., Hinton, G., Krizhevsky, A., Sutskever, I., & Salakhutdinov, R. (2014). Dropout: A Simple Way to Prevent Neural Networks from Overfitting. *Journal of Machine Learning Research*, 15(56), 1929–1958. <http://jmlr.org/papers/v15/srivastava14a.html>
- Storn, R., & Price, K. (1996). Minimizing the real functions of the ICEC'96 contest by differential evolution. *Proceedings of IEEE International Conference on Evolutionary Computation*, 842–844. <https://doi.org/10.1109/ICEC.1996.542711>
- Su, X., Yan, X., & Tsai, C. (2012). Linear regression. *WIREs Computational Statistics*, 4(3), 275–294. <https://doi.org/10.1002/wics.1198>
- Sudhakar, U., & Srinivas, J. (2019). Mechanical characteristics and corrosion behavior of friction stir AA5251-AA6063 butt welds. *Materials Today: Proceedings*, 15, 132–137. <https://doi.org/10.1016/j.matpr.2019.05.035>
- Székel, G. J., & Rizzo, M. L. (2009). Brownian distance covariance. *The Annals of Applied Statistics*, 3(4). <https://doi.org/10.1214/09-AOAS312>
- Székel, G. J., & Rizzo, M. L. (2012). On the uniqueness of distance covariance. *Statistics & Probability Letters*, 82(12), 2278–2282. <https://doi.org/10.1016/j.spl.2012.08.007>

- Takase, T., Oyama, S., & Kurihara, M. (2018). Effective neural network training with adaptive learning rate based on training loss. *Neural Networks*, 101, 68–78. <https://doi.org/10.1016/j.neunet.2018.01.016>
- Taylor, L., & Nitschke, G. (2018). Improving Deep Learning with Generic Data Augmentation. *2018 IEEE Symposium Series on Computational Intelligence (SSCI)*, 1542–1547. <https://doi.org/10.1109/SSCI.2018.8628742>
- Temizel, A., Halici, T., Logoglu, B., Temizel, T. T., Omruuzun, F., & Karaman, E. (2011). Experiences on Image and Video Processing with CUDA and OpenCL. In *GPU Computing Gems Emerald Edition* (pp. 547–567). Elsevier. <https://doi.org/10.1016/B978-0-12-384988-5.00034-6>
- Tereshin, N. A., Padokhin, A. M., Andreeva, E. S., & Kozlovtseva, E. A. (2020). Simplicial Homology Global Optimisation in the Problem of Point-to-Point Ionospheric Ray Tracing. *2020 XXXIIIrd General Assembly and Scientific Symposium of the International Union of Radio Science*, 1–4. <https://doi.org/10.23919/URSIGASS49373.2020.9232337>
- Thao, D. T., Jeong, J. W., Kim, I. S., & Kim, J. W. H. J. (2008). Predicting Lap-Joint bead geometry in GMA welding process. *International Scientific Journal Published Monthly by the World Academy of Materials and Manufacturing Engineering*, 32(2), 121–124.
- Tripicchio, P., Camacho-Gonzalez, G., & D'Avella, S. (2020). Welding defect detection: coping with artifacts in the production line. *The International Journal of Advanced Manufacturing Technology*, 111(5–6), 1659–1669. <https://doi.org/10.1007/s00170-020-06146-4>
- Ukaoha, K., & Igodan, E. (2020). Architecture Optimization Model for the Deep Neural Network For Binary Classification Problems. *International Journal of Intelligent Computing and Information Sciences*, 0(0), 0–0. <https://doi.org/10.21608/ijicis.2020.18509.1008>
- Van Beers, W. C. M., & Kleijnen, J. P. C. (n.d.). Kriging Interpolation in Simulation: A Survey. *Proceedings of the 2004 Winter Simulation Conference, 2004.*, 107–115. <https://doi.org/10.1109/WSC.2004.1371308>
- Vasudevan, M., Kuppaswamy, M. V., & Bhaduri, A. K. (2010). Optimising process parameters for gas tungsten arc welding of an austenitic stainless steel using genetic algorithm. *Transactions of the Indian Institute of Metals*, 63(1), 1–10. <https://doi.org/10.1007/s12666-010-0001-5>
- Vedrtam, A., Singh, G., & Kumar, A. (2018). Optimizing submerged arc welding using response surface methodology, regression analysis, and genetic algorithm. *Defence Technology*, 14(3), 204–212. <https://doi.org/10.1016/j.dt.2018.01.008>
- Venkatesh Kannan, M., Arivazhagan, N., Nageswara Rao, M., Madhusudhan Reddy, G., Phani Prabhakar, K. V., & Gadhe, P. (2022). Studies on Microstructure and Mechanical Properties of Weldments Produced in 12 mm Thick Naval Grade High Strength Low Alloy Steel for Sub-Zero Application by Single and Double Pass Hybrid Laser Arc Welding.

- Journal of Materials Engineering and Performance*, 31(4), 3234–3248.
<https://doi.org/10.1007/s11665-021-06403-x>
- Venter, G. (2010). Review of Optimization Techniques. In *Encyclopedia of Aerospace Engineering*. John Wiley & Sons, Ltd. <https://doi.org/10.1002/9780470686652.eae495>
- Venter, G., & Sobieszczanski-Sobieski, J. (2003). Particle Swarm Optimization. *AIAA Journal*, 41(8), 1583–1589. <https://doi.org/10.2514/2.2111>
- Vijayavel, P., Mahesh, D., Venkatswamy, M., & Chakilam, S. (2018). OPTIMIZATION OF MIG WELDING PARAMETERS FOR IMPROVING STRENGTH OF WELDED JOINTS. 2394–2697. <https://doi.org/10.21276/ijcesr.2018.5.1.20>
- Vovk, V. (2013). Kernel Ridge Regression. In *Empirical Inference* (pp. 105–116). Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-642-41136-6_11
- Wahab, M. A. (2014). Manual Metal Arc Welding and Gas Metal Arc Welding. In *Comprehensive Materials Processing* (pp. 49–76). Elsevier. <https://doi.org/10.1016/B978-0-08-096532-1.00610-5>
- Wang, T., Hu, B., Feng, Y., Gao, X., Yang, C., & Tan, J. (2023). Data Augmentation-Based Manufacturing Quality Prediction Approach in Human Cyber-Physical Systems. *Journal of Manufacturing Science and Engineering*, 145(12). <https://doi.org/10.1115/1.4063269>
- Wang, X.-W., Shi, Y.-P., Gu, X.-S., & Ding, D.-Y. (2015). Partition Mutation PSO for Welding Robot Path Optimization (pp. 77–86). https://doi.org/10.1007/978-3-319-18997-0_6
- Wanto, A., Windarto, A. P., Hartama, D., & Parlina, I. (2017). Use of Binary Sigmoid Function And Linear Identity In Artificial Neural Networks For Forecasting Population Density. *IJISTECH (International Journal Of Information System & Technology)*, 1(1), 43. <https://doi.org/10.30645/ijistech.v1i1.6>
- Weman, K. (2012). MIG/MAG welding. In *Welding Processes Handbook* (pp. 75–97). Elsevier. <https://doi.org/10.1533/9780857095183.75>
- Werbos, P. J. (1990). Backpropagation through time: what it does and how to do it. *Proceedings of the IEEE*, 78(10), 1550–1560. <https://doi.org/10.1109/5.58337>
- Wong, T.-T. (2015a). Performance evaluation of classification algorithms by k-fold and leave-one-out cross validation. *Pattern Recognition*, 48(9), 2839–2846. <https://doi.org/10.1016/j.patcog.2015.03.009>
- Wong, T.-T. (2015b). Performance evaluation of classification algorithms by k-fold and leave-one-out cross validation. *Pattern Recognition*, 48(9), 2839–2846. <https://doi.org/10.1016/j.patcog.2015.03.009>
- Wu, C., Wang, C., & Kim, J.-W. (2022). Welding sequence optimization to reduce welding distortion based on coupled artificial neural network and swarm intelligence algorithm. *Engineering Applications of Artificial Intelligence*, 114, 105142. <https://doi.org/10.1016/j.engappai.2022.105142>

- Wythoff, B. J. (1993). Backpropagation neural networks. *Chemometrics and Intelligent Laboratory Systems*, 18(2), 115–155. [https://doi.org/10.1016/0169-7439\(93\)80052-J](https://doi.org/10.1016/0169-7439(93)80052-J)
- Xia, X., Gui, L., Zhang, Y., Xu, X., Yu, F., Wu, H., Wei, B., He, G., Li, Y., & Li, K. (2021). A fitness-based adaptive differential evolution algorithm. *Information Sciences*, 549, 116–141. <https://doi.org/10.1016/j.ins.2020.11.015>
- Yang, L., & Shami, A. (2020). On hyperparameter optimization of machine learning algorithms: Theory and practice. *Neurocomputing*, 415, 295–316. <https://doi.org/10.1016/j.neucom.2020.07.061>
- Yang, W., Xiao, Y., Shen, H., & Wang, Z. (2023). An effective data enhancement method of deep learning for small weld data defect identification. *Measurement*, 206, 112245. <https://doi.org/10.1016/j.measurement.2022.112245>
- Yang, Y., Gao, Z., & Cao, L. (2018). Identifying optimal process parameters in deep penetration laser welding by adopting Hierarchical-Kriging model. *Infrared Physics & Technology*, 92, 443–453. <https://doi.org/10.1016/j.infrared.2018.07.006>
- Yifei, T., Meng, Z., Jingwei, L., Dongbo, L., & Yulin, W. (2018). Research on Intelligent Welding Robot Path Optimization Based on GA and PSO Algorithms. *IEEE Access*, 6, 65397–65404. <https://doi.org/10.1109/ACCESS.2018.2878615>
- Yu, J., & Kim, D. (2018). Effects of welding current and torch position parameters on minimizing the weld porosity of zinc-coated steel. *The International Journal of Advanced Manufacturing Technology*, 95(1–4), 551–567. <https://doi.org/10.1007/s00170-017-1180-6>
- Yu, R., Han, J., Zhao, Z., & Bai, L. (2020). Real-Time Prediction of Welding Penetration Mode and Depth Based on Visual Characteristics of Weld Pool in GMAW Process. *IEEE Access*, 8, 81564–81573. <https://doi.org/10.1109/ACCESS.2020.2990902>
- Zeng, M., Huang, J., Zhang, Y., & Hu, P. (2017). Modeling for GMAW process with a current waveform control method. *Journal of Materials Processing Technology*, 240, 404–413. <https://doi.org/10.1016/j.jmatprotec.2016.10.018>
- ZHANG, D., ZHAO, Y., DONG, M., WANG, G., WU, A., SHAN, J., MENG, D., LIU, X., SONG, J., & ZHANG, Z. (2019). Effects of weld penetration on tensile properties of 2219 aluminum alloy TIG-welded joints. *Transactions of Nonferrous Metals Society of China*, 29(6), 1161–1168. [https://doi.org/10.1016/S1003-6326\(19\)65024-7](https://doi.org/10.1016/S1003-6326(19)65024-7)
- Zhang, K., Patki, N., & Veeramachaneni, K. (2022). *Sequential Models in the Synthetic Data Vault*.
- Zhang, N., Shen, S.-L., Zhou, A., & Xu, Y.-S. (2019). Investigation on Performance of Neural Networks Using Quadratic Relative Error Cost Function. *IEEE Access*, 7, 106642–106652. <https://doi.org/10.1109/ACCESS.2019.2930520>

- Zhao, D., Wang, Y., Liang, D., & Ivanov, M. (2020). Performances of regression model and artificial neural network in monitoring welding quality based on power signal. *Journal of Materials Research and Technology*, 9(2), 1231–1240. <https://doi.org/10.1016/j.jmrt.2019.11.050>
- Zhao, H., Wang, M., & Chang, X. (2022). Data-driven reduced order model and simplicial homology global optimization for reliability analysis and application. *Heliyon*, 8(10), e11036. <https://doi.org/10.1016/j.heliyon.2022.e11036>
- Zhu, C., Wan, X., & Zhou, Z. (2023). Two-stage optimization layout of grasping points for sheet metal part based on GSA-Kriging model. *The International Journal of Advanced Manufacturing Technology*.
- Zhu, Y., Yuan, Z., Khonsari, M. M., Zhao, S., & Yang, H. (2023). Small-Dataset Machine Learning for Wear Prediction of Laser Powder Bed Fusion Fabricated Steel. *Journal of Tribology*, 145(9). <https://doi.org/10.1115/1.4062368>

ANEXO A – Dataset Original

Entradas					Saídas				
Essespura(mm)	Chanfro: reto (0) - V (1)	Afastamento (mm)	Altura (mm)	Largura (mm)	Vazão (L/min)	Tensão (V)	Velocidade (m/min)	Corrente (A)	Stickout (mm)
3,1750	0	1,5	2,00	7,00	5,00	19,2	5,0	108	13,0
3,1750	0	1,5	4,80	10,00	5,00	15,6	5,0	144	17,0
3,1750	0	1,0	1,80	6,90	5,00	19,2	3,5	72	19,0
3,1750	0	1,0	1,30	5,80	6,70	19,2	3,5	76	14,0
3,1750	0	1,5	1,15	7,00	6,70	19,2	3,5	84	13,0
3,1750	0	1,5	1,25	5,30	3,30	19,2	3,5	88	11,5
3,1750	0	1,5	1,35	6,15	3,30	19,2	3,5	88	12,5
3,1750	0	1,0	2,40	7,20	3,30	21,6	6,0	136	13,5
3,1750	0	1,0	1,75	7,70	6,70	21,6	6,0	140	17,0
3,1750	0	1,0	2,00	7,65	6,70	21,6	6,0	156	9,0
4,7625	0	1,5	1,85	7,90	6,70	21,6	6,0	148	11,5
4,7625	0	1,5	1,20	8,25	6,70	21,6	6,0	148	11,5
4,7625	0	2,5	1,25	8,10	5,00	21,2	6,0	148	9,0
4,7625	0	2,5	1,40	7,60	5,00	21,6	7,0	148	14,0
4,7625	0	1,5	2,00	7,00	8,35	21,6	7,0	160	9,0

4,7625	0	3,2	1,85	9,60	8,35	22,0	6,0	140	16,0
4,7625	0	3,2	1,40	9,90	8,35	22,0	6,0	144	15,5
4,7625	0	1,5	1,40	9,60	8,35	26,4	4,0	100	13,0
4,7625	0	2,5	2,00	9,20	8,35	25,6	7,0	152	11,0
4,7625	0	2,5	2,10	8,80	5,00	25,2	7,0	156	14,0
4,7625	0	2,5	1,45	8,30	6,70	21,2	6,0	148	11,5
6,3500	0	2,5	2,25	10,50	6,70	21,2	6,0	148	17,0
6,3500	0	4,5	2,20	10,40	6,70	23,2	8,0	172	17,0
6,3500	0	2,5	1,10	7,40	3,30	22,8	4,0	88	7,5
6,3500	0	1,5	2,30	9,10	3,30	22,8	4,0	80	16,0
6,3500	0	1,5	2,25	5,60	5,00	25,2	7,0	164	14,0
6,3500	0	2,5	1,65	7,60	5,00	25,6	6,0	140	14,0
6,3500	1	0,0	2,50	10,70	6,70	22,0	6,0	144	14,5
6,3500	1	0,0	2,35	11,20	6,70	22,0	6,0	140	18,0
6,3500	1	0,0	1,75	9,90	6,70	22,4	5,0	116	19,0
6,3500	1	0,0	2,60	10,20	5,00	21,2	5,0	116	14,0
6,3500	1	0,0	2,20	8,70	5,00	21,2	5,0	124	9,0
6,3500	1	0,0	2,60	8,20	5,00	21,2	5,0	120	18,0
9,5250	1	0,0	1,70	11,40	5,00	21,2	5,0	124	11,0
9,5250	1	0,0	1,30	12,10	5,00	23,2	8,0	164	21,0
9,5250	1	0,0	1,75	11,80	6,70	23,2	8,0	172	22,0
9,5250	1	0,0	2,20	11,40	6,70	22,0	7,0	144	19,0
9,5250	1	0,0	2,30	12,10	6,70	22,0	7,0	148	17,0
9,5250	0	4,5	1,40	10,10	6,70	22,4	7,0	140	16,0
9,5250	0	4,5	1,90	10,90	6,70	23,6	8,0	164	22,0
9,5250	0	4,5	1,60	11,70	6,70	22,8	10,0	196	20,0
9,5250	0	2,5	1,50	9,80	6,70	21,6	7,0	156	19,0
9,5250	0	2,5	1,35	10,10	6,70	24,4	10,0	204	19,0
9,5250	0	2,5	2,25	10,50	6,70	21,2	6,0	148	17,0

ANEXO B – Dataset usado no SDV 1.5

chanfro	espessura	afastamento	altura	largura	vazão	tensão	velocidade	stickout
0	3.175	1.5	2.0	7.0	5.0	19.2	5.0	13.0
0	3.175	1.5	4.8	10.0	5.0	15.6	5.0	17.0
0	3.175	1.0	1.8	6.9	5.0	19.2	3.5	19.0
0	3.175	1.0	1.3	5.8	6.7	19.2	3.5	14.0
0	3.175	1.5	1.15	7.0	6.7	19.2	3.5	13.0
0	3.175	1.5	1.25	5.3	3.3	19.2	3.5	11.5
0	3.175	1.5	1.35	6.15	3.3	19.2	3.5	12.5

0	3.175	1.0	2.4	7.2	3.3	21.6	6.0	13.5
0	3.175	1.0	1.75	7.7	6.7	21.6	6.0	17.0
0	3.175	1.0	2.0	7.65	6.7	21.6	6.0	9.0
0	4.7625	1.5	1.85	7.9	6.7	21.6	6.0	11.5
0	4.7625	1.5	1.2	8.25	6.7	21.6	6.0	11.5
0	4.7625	2.5	1.25	8.1	5.0	21.2	6.0	9.0
0	4.7625	2.5	1.4	7.6	5.0	21.6	7.0	14.0
0	4.7625	1.5	2.0	7.0	8.35	21.6	7.0	9.0
0	4.7625	3.2	1.85	9.6	8.35	22.0	6.0	16.0
0	4.7625	3.2	1.4	9.9	8.35	22.0	6.0	15.5
0	4.7625	1.5	1.4	9.6	8.35	26.4	4.0	13.0
0	4.7625	2.5	2.0	9.2	8.35	25.6	7.0	11.0
0	4.7625	2.5	2.1	8.8	5.0	25.2	7.0	14.0
0	4.7625	2.5	1.45	8.3	6.7	21.2	6.0	11.5
0	6.35	2.5	2.25	10.5	6.7	21.2	6.0	17.0
0	6.35	4.5	2.2	10.4	6.7	23.2	8.0	17.0
0	6.35	2.5	1.1	7.4	3.3	22.8	4.0	7.5
0	6.35	1.5	2.3	9.1	3.3	22.8	4.0	16.0
0	6.35	1.5	2.25	5.6	5.0	25.2	7.0	14.0
0	6.35	2.5	1.65	7.6	5.0	25.6	6.0	14.0
1	6.35	0.0	2.5	10.7	6.7	22.0	6.0	14.5
1	6.35	0.0	2.35	11.2	6.7	22.0	6.0	18.0
1	6.35	0.0	1.75	9.9	6.7	22.4	5.0	19.0
1	6.35	0.0	2.6	10.2	5.0	21.2	5.0	14.0
1	6.35	0.0	2.2	8.7	5.0	21.2	5.0	9.0
1	6.35	0.0	2.6	8.2	5.0	21.2	5.0	18.0
1	9.525	0.0	1.7	11.4	5.0	21.2	5.0	11.0
1	9.525	0.0	1.3	12.1	5.0	23.2	8.0	21.0
1	9.525	0.0	1.75	11.8	6.7	23.2	8.0	22.0
1	9.525	0.0	2.2	11.4	6.7	22.0	7.0	19.0
1	9.525	0.0	2.3	12.1	6.7	22.0	7.0	17.0
0	9.525	4.5	1.4	10.1	6.7	22.4	7.0	16.0
0	9.525	4.5	1.9	10.9	6.7	23.6	8.0	22.0
0	9.525	4.5	1.6	11.7	6.7	22.8	10.0	20.0
0	9.525	2.5	1.5	9.8	6.7	21.6	7.0	19.0
0	9.525	2.5	1.35	10.1	6.7	24.4	10.0	19.0
0	9.525	2.5	2.25	10.5	6.7	21.2	6.0	17.0
0	8.5811	0.0	2.62	10.15	7.06	24.3	8.3	16.3
0	8.3209	1.4	1.21	7.25	3.53	22.7	4.4	17.5
0	9.0852	0.0	1.49	9.15	7.41	20.0	3.9	19.6
0	5.1582	0.1	1.67	7.38	7.9	19.8	5.8	10.2
0	5.4376	1.7	2.36	10.13	5.79	19.0	4.4	10.5
1	6.8574	0.0	1.7	9.04	5.95	20.4	4.8	13.6
0	6.3352	4.5	1.4	7.09	7.15	23.1	3.5	17.4
1	9.2633	0.0	2.73	11.89	7.85	25.5	9.3	17.7

0	7.9632	4.5	1.41	6.42	3.63	22.6	5.6	10.5
1	5.1473	0.0	1.78	6.43	7.51	18.8	3.8	9.9
0	7.2973	1.0	1.92	5.8	6.34	24.8	6.0	16.4
0	5.9049	0.1	1.21	9.84	7.89	26.0	7.0	10.8
0	9.3932	0.4	1.14	9.88	6.08	26.4	7.5	11.3
0	4.8399	4.5	2.42	5.95	7.82	24.0	5.9	9.5
1	5.4246	0.0	1.82	10.08	6.43	22.6	4.6	8.0
0	7.1745	4.5	1.25	7.21	6.91	20.7	5.4	13.2
0	5.8519	0.0	2.62	9.18	8.17	20.9	7.5	11.7
0	5.1425	4.5	1.17	5.31	7.39	25.5	4.6	12.1
0	5.2328	0.1	1.94	5.86	7.64	22.7	4.1	8.7
1	9.3854	0.0	2.06	11.81	6.75	24.5	7.6	12.3

ANEXO C - Dataset usado no SDV 2

espessura	chanfro	afastamento	altura	largura	vazão	tensão	velocidade	stickout
3.175	0	1.5	2	7	5	19.2	5	13
3.175	0	1.5	4.8	10	5	15.6	5	17
3.175	0	1	1.8	6.9	5	19.2	3.5	19
3.175	0	1	1.3	5.8	6.7	19.2	3.5	14
3.175	0	1.5	1.15	7	6.7	19.2	3.5	13
3.175	0	1.5	1.25	5.3	3.3	19.2	3.5	11.5
3.175	0	1.5	1.35	6.15	3.3	19.2	3.5	12.5
3.175	0	1	2.4	7.2	3.3	21.6	6	13.5
3.175	0	1	1.75	7.7	6.7	21.6	6	17
3.175	0	1	2	7.65	6.7	21.6	6	9
4.7625	0	1.5	1.85	7.9	6.7	21.6	6	11.5
4.7625	0	1.5	1.2	8.25	6.7	21.6	6	11.5
4.7625	0	2.5	1.25	8.1	5	21.2	6	9
4.7625	0	2.5	1.4	7.6	5	21.6	7	14
4.7625	0	1.5	2	7	8.35	21.6	7	9
4.7625	0	3.2	1.85	9.6	8.35	22	6	16
4.7625	0	3.2	1.4	9.9	8.35	22	6	15.5
4.7625	0	1.5	1.4	9.6	8.35	26.4	4	13
4.7625	0	2.5	2	9.2	8.35	25.6	7	11
4.7625	0	2.5	2.1	8.8	5	25.2	7	14
4.7625	0	2.5	1.45	8.3	6.7	21.2	6	11.5
6.35	0	2.5	2.25	10.5	6.7	21.2	6	17

6.35	0	4.5	2.2	10.4	6.7	23.2	8	17
6.35	0	2.5	1.1	7.4	3.3	22.8	4	7.5
6.35	0	1.5	2.3	9.1	3.3	22.8	4	16
6.35	0	1.5	2.25	5.6	5	25.2	7	14
6.35	0	2.5	1.65	7.6	5	25.6	6	14
6.35	1	0	2.5	10.7	6.7	22	6	14.5
6.35	1	0	2.35	11.2	6.7	22	6	18
6.35	1	0	1.75	9.9	6.7	22.4	5	19
6.35	1	0	2.6	10.2	5	21.2	5	14
6.35	1	0	2.2	8.7	5	21.2	5	9
6.35	1	0	2.6	8.2	5	21.2	5	18
9.525	1	0	1.7	11.4	5	21.2	5	11
9.525	1	0	1.3	12.1	5	23.2	8	21
9.525	1	0	1.75	11.8	6.7	23.2	8	22
9.525	1	0	2.2	11.4	6.7	22	7	19
9.525	1	0	2.3	12.1	6.7	22	7	17
9.525	0	4.5	1.4	10.1	6.7	22.4	7	16
9.525	0	4.5	1.9	10.9	6.7	23.6	8	22
9.525	0	4.5	1.6	11.7	6.7	22.8	10	20
9.525	0	2.5	1.5	9.8	6.7	21.6	7	19
9.525	0	2.5	1.35	10.1	6.7	24.4	10	19
9.525	0	2.5	2.25	10.5	6.7	21.2	6	17
9.4338	1	0.0	3.14	12.09	7.8	22.7	9.4	20.9
5.7256	1	0.0	2.38	7.49	7.64	21.2	4.8	9.0
4.7741	0	4.5	2.68	5.73	8.16	17.3	3.9	8.0
8.522	1	0.0	1.95	10.66	8.27	25.3	7.3	16.1
9.1673	0	0.6	1.53	11.08	6.09	21.6	6.5	18.0
6.0876	0	4.5	1.59	9.58	8.21	23.7	9.4	12.7
4.8715	0	0.1	1.49	8.3	4.39	15.6	3.5	10.2
7.1943	0	0.0	1.58	8.44	3.3	23.1	5.4	11.6
8.5042	1	0.0	1.8	11.32	7.73	22.9	4.7	13.9
8.9055	1	0.0	1.64	11.61	3.92	21.9	8.9	21.2
5.5219	0	0.7	2.26	5.4	3.82	19.3	3.7	9.8
5.5731	0	3.2	2.39	9.7	8.23	21.3	5.9	14.9
7.5158	1	0.0	2.64	11.21	7.43	22.3	9.3	13.6
8.0945	0	0.2	2.43	7.64	7.12	23.7	7.7	18.8
6.3375	0	0.1	1.7	8.83	6.08	21.3	5.7	13.7
6.4381	1	0.0	2.44	10.36	8.16	23.1	7.0	16.7
4.7771	0	2.7	2.33	5.78	8.01	17.2	3.5	14.7
4.7813	0	1.7	1.54	7.51	8.12	20.9	4.1	12.0
6.8151	0	0.8	1.52	11.11	6.54	23.8	6.7	11.6
9.525	1	0.0	2.69	11.62	7.52	26.4	8.1	21.4
5.0801	0	4.5	2.06	7.47	7.84	21.7	7.1	9.6
7.0534	0	0.1	1.74	9.84	6.23	21.6	4.2	9.6
9.1733	0	0.0	1.62	10.43	3.3	20.5	5.3	16.6

8.7411	0	0.0	1.63	9.6	6.34	20.4	4.7	15.8
9.3391	1	0.0	2.64	9.19	6.86	25.5	5.5	21.2
6.6674	0	1.3	1.51	6.72	7.39	20.4	3.6	17.0
5.9102	0	0.6	1.4	10.3	5.08	19.4	6.4	15.7
8.8574	0	4.1	1.63	6.17	5.28	18.3	3.6	15.7
9.525	1	0.0	4.8	11.8	4.64	22.2	10.0	20.2
8.4912	1	0.0	1.88	6.62	6.56	24.9	10.0	11.4
7.7654	0	0.2	1.52	6.94	5.91	22.0	4.2	19.8
5.9698	0	1.7	1.41	12.07	8.35	23.5	10.0	9.1
6.0657	0	0.3	1.87	5.79	5.52	22.0	4.1	9.2
7.0297	0	0.0	1.2	5.83	4.71	22.7	3.7	10.6
9.4202	1	0.0	1.63	9.93	3.3	21.3	6.5	15.6
9.3468	1	0.0	2.42	11.63	4.43	19.9	8.2	18.8
4.7645	0	4.5	1.88	6.42	8.21	22.4	3.7	11.7
7.7438	0	0.0	2.24	11.89	6.61	22.9	8.6	16.0
5.3769	0	4.5	1.28	6.52	5.87	22.0	6.0	15.6
8.6745	0	0.0	2.18	8.25	3.3	19.0	4.4	15.4
9.3727	1	0.0	2.42	9.99	3.3	22.0	6.5	13.3
5.2189	0	3.6	1.32	8.54	6.17	20.6	4.1	13.1
4.8754	0	0.0	2.57	6.15	7.63	21.2	4.5	11.3

ANEXO D - Dataset usado no SDV 2.5

espessura	chanfro	afastamento	altura	largura	vazão	tensão	velocidade	stickout
3.175	0	1.5	2	7	5	19.2	5	13
3.175	0	1.5	4.8	10	5	15.6	5	17
3.175	0	1	1.8	6.9	5	19.2	3.5	19
3.175	0	1	1.3	5.8	6.7	19.2	3.5	14
3.175	0	1.5	1.15	7	6.7	19.2	3.5	13
3.175	0	1.5	1.25	5.3	3.3	19.2	3.5	11.5
3.175	0	1.5	1.35	6.15	3.3	19.2	3.5	12.5
3.175	0	1	2.4	7.2	3.3	21.6	6	13.5
3.175	0	1	1.75	7.7	6.7	21.6	6	17
3.175	0	1	2	7.65	6.7	21.6	6	9
4.7625	0	1.5	1.85	7.9	6.7	21.6	6	11.5
4.7625	0	1.5	1.2	8.25	6.7	21.6	6	11.5
4.7625	0	2.5	1.25	8.1	5	21.2	6	9
4.7625	0	2.5	1.4	7.6	5	21.6	7	14
4.7625	0	1.5	2	7	8.35	21.6	7	9
4.7625	0	3.2	1.85	9.6	8.35	22	6	16
4.7625	0	3.2	1.4	9.9	8.35	22	6	15.5
4.7625	0	1.5	1.4	9.6	8.35	26.4	4	13
4.7625	0	2.5	2	9.2	8.35	25.6	7	11
4.7625	0	2.5	2.1	8.8	5	25.2	7	14
4.7625	0	2.5	1.45	8.3	6.7	21.2	6	11.5

6.35	0	2.5	2.25	10.5	6.7	21.2	6	17
6.35	0	4.5	2.2	10.4	6.7	23.2	8	17
6.35	0	2.5	1.1	7.4	3.3	22.8	4	7.5
6.35	0	1.5	2.3	9.1	3.3	22.8	4	16
6.35	0	1.5	2.25	5.6	5	25.2	7	14
6.35	0	2.5	1.65	7.6	5	25.6	6	14
6.35	1	0	2.5	10.7	6.7	22	6	14.5
6.35	1	0	2.35	11.2	6.7	22	6	18
6.35	1	0	1.75	9.9	6.7	22.4	5	19
6.35	1	0	2.6	10.2	5	21.2	5	14
6.35	1	0	2.2	8.7	5	21.2	5	9
6.35	1	0	2.6	8.2	5	21.2	5	18
9.525	1	0	1.7	11.4	5	21.2	5	11
9.525	1	0	1.3	12.1	5	23.2	8	21
9.525	1	0	1.75	11.8	6.7	23.2	8	22
9.525	1	0	2.2	11.4	6.7	22	7	19
9.525	1	0	2.3	12.1	6.7	22	7	17
9.525	0	4.5	1.4	10.1	6.7	22.4	7	16
9.525	0	4.5	1.9	10.9	6.7	23.6	8	22
9.525	0	4.5	1.6	11.7	6.7	22.8	10	20
9.525	0	2.5	1.5	9.8	6.7	21.6	7	19
9.525	0	2.5	1.35	10.1	6.7	24.4	10	19
9.525	0	2.5	2.25	10.5	6.7	21.2	6	17
5.6227	0	0.6	2.35	11.77	8.07	21.2	6.4	17.5
6.6107	0	3.8	1.21	6.79	6.44	21.6	6.0	14.2
4.9431	0	0.4	1.88	6.12	8.33	22.6	6.0	10.7
6.7797	1	0.0	2.73	7.44	3.78	23.2	4.0	12.0
5.229	0	0.0	1.54	8.84	6.89	20.8	4.9	9.9
7.7262	0	4.5	1.29	10.89	8.3	23.8	7.8	18.9
9.525	1	0.0	1.66	10.96	4.7	26.4	10.0	20.0
5.2986	1	0.0	1.88	11.26	8.07	21.3	4.3	19.2
5.7258	0	0.6	1.25	6.07	7.03	21.7	3.9	10.0
8.4184	0	1.7	1.29	9.69	8.05	23.6	7.7	15.0
7.6717	1	0.0	1.6	8.77	7.29	22.6	5.9	9.9
8.7959	1	0.0	2.22	11.78	7.4	22.7	7.9	15.5
4.9765	1	0.0	1.82	7.8	7.32	21.5	4.0	11.7
6.3719	0	0.5	2.19	7.16	7.22	21.8	6.7	16.7
6.3639	0	4.5	1.19	11.35	7.17	26.4	8.7	19.6
6.5795	0	4.5	2.34	10.36	7.44	20.4	6.1	15.1
9.029	0	0.0	3.1	8.15	4.67	24.4	9.1	19.7
6.8389	0	1.5	1.99	11.2	7.94	21.9	4.0	19.0
8.2236	1	0.0	2.9	7.44	6.11	19.6	4.1	15.1
9.4446	1	0.0	1.49	12.09	6.54	25.5	9.0	20.3
5.8932	0	0.0	1.54	7.3	5.46	21.6	5.7	7.5
6.8163	1	0.0	1.86	12.09	7.16	20.6	4.5	19.5

5.0605	1	0.0	2.66	9.46	6.72	22.9	6.2	11.1
8.4601	0	1.4	2.79	11.73	7.71	24.8	8.7	17.4
8.6205	1	0.0	1.81	12.03	7.94	21.5	4.8	20.8
8.4681	0	0.0	1.95	11.81	4.83	21.6	6.2	17.8
6.6263	0	0.1	1.43	6.1	4.27	24.9	4.1	8.2
9.4258	1	0.0	2.56	11.42	6.03	20.1	9.4	20.1
8.7086	0	0.4	1.42	11.97	4.78	21.8	6.5	18.0
5.814	0	0.6	1.49	5.95	6.79	20.2	3.8	12.7
9.2844	1	0.0	1.89	11.7	3.3	24.3	8.0	19.0
8.9586	0	0.0	1.44	11.11	3.88	20.2	3.5	20.4
8.7801	1	0.0	2.21	12.07	6.51	23.2	9.7	21.4
5.5132	0	0.0	1.73	5.81	3.69	18.5	3.7	15.7
5.5444	1	0.0	1.99	7.26	4.81	18.7	4.9	8.3
9.3914	1	0.0	2.65	11.87	6.53	23.4	10.0	17.2
9.2136	0	0.1	1.35	5.87	5.47	26.4	9.1	14.8
9.1295	0	0.0	2.75	11.29	4.32	21.5	6.7	20.9
8.6607	0	1.0	1.4	11.13	7.04	24.3	7.5	11.8
4.971	0	0.0	2.66	6.54	6.07	17.1	3.7	7.9
8.6433	0	1.2	1.33	10.23	7.66	22.4	4.8	20.5
6.9651	0	0.4	1.22	7.65	3.51	22.7	4.5	12.4
6.0507	0	1.2	1.14	7.75	8.1	24.4	3.8	10.8
4.925	0	3.5	1.29	8.39	8.25	22.5	3.8	8.9
7.0615	0	0.3	1.36	9.18	6.46	21.3	3.7	18.4
8.1331	0	0.0	1.86	7.6	4.67	19.9	7.5	19.7
5.3788	0	3.5	1.25	8.66	6.34	21.5	4.1	15.5
7.7228	0	0.2	1.59	10.88	7.81	22.8	4.4	16.6
8.535	0	0.1	2.23	11.17	7.48	23.3	6.4	15.8
8.2536	1	0.0	3.8	11.05	5.55	18.4	4.5	20.5
7.9665	0	4.5	2.25	9.91	7.29	22.8	8.7	19.4
5.0573	0	4.5	1.25	6.29	7.64	22.7	3.7	11.9
7.5055	0	0.4	1.74	9.51	6.28	21.7	7.6	17.5
4.9635	0	0.0	1.54	7.61	3.3	18.3	3.9	11.8
7.8572	1	0.0	2.23	10.48	4.67	23.3	8.5	19.4
8.5114	1	0.0	2.47	10.89	4.86	21.1	6.2	21.2
9.1318	1	0.0	2.22	12.08	6.13	24.0	6.8	20.7
9.0941	1	0.0	2.55	9.46	3.3	21.2	4.2	17.0
7.5204	0	0.6	2.3	11.92	8.29	26.2	8.5	19.2
9.3978	0	0.0	1.45	9.83	3.32	25.4	6.7	10.7