

UNIVERSIDADE FEDERAL DA BAHIA
ESCOLA POLITÉCNICA
Programa De Pós-Graduação Em Engenharia Elétrica

VIVIANE OLIVEIRA DAS MERCÊS

**Aprendizado de Máquina Aplicada em Modelagem da
Eficiência de Acoplamento entre Guias Dielétricos.**

SALVADOR

2024

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

VIVIANE OLIVEIRA DAS MERCÊS

**Aprendizado de Máquina Aplicada em Modelagem da
Eficiência de Acoplamento entre Guias Dielétricos.**

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

Orientador: Prof. Dr. Vitaly Félix Rodríguez Esquerre.

Coorientador: Prof. Dr. Anderson Dourado Sisnando.

SALVADOR

2024

Ficha catalográfica elaborada pelo Sistema Universitário de Bibliotecas (SIBI/UFBA),
com os dados fornecidos pelo(a) autor(a).

Oliveira das Mercês, Viviane
Aprendizado de Máquina Aplicada em Modelagem da
Eficiência de Acoplamento entre Guias Dielétricos. /
Viviane Oliveira das Mercês. -- Salvador, 2024.
115 f. : il

Orientador: Vitaly Félix Rodríguez Esquerre.
Coorientador: Anderson Dourado Sisnando.
Dissertação (Mestrado - Programa de Pós-Graduação em
Engenharia Elétrica e Computação) -- Universidade
Federal da Bahia, Universidade Federal da Bahia, 2024.

1. Sistemas de comunicação óptica. 2. Redes neurais
artificiais. 3. Guias de onda ópticos. 4. Inteligência
artificial. I. Félix Rodríguez Esquerre, Vitaly. II.
Dourado Sisnando, Anderson. III. Título.



UNIVERSIDADE FEDERAL DA BAHIA
ESCOLA POLITÉCNICA



PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA ELÉTRICA

**Aprendizado de Máquina Aplicada em Modelagem da
Eficiência de Acoplamento entre Guias Dielétricos.**

DISSERTAÇÃO DE MESTRADO

Autora: Viviane Oliveira das Mercês

BANCA EXAMINADORA

Prof. Dr. Vitaly Felix Rodriguez Esquerre
(Orientador)

Prof. Dr. Anderson Dourado Sisnando
(Coorientador)

Prof. Dr. Fabrício Gerônimo Simões Silva

Prof. Dr. Eduardo Furtado de Simas Filho

DEDICATÓRIA

*Aos meus pais, irmão, familiares
e amigos , pelo incentivo, apoio,
compreensão e amor, para que
atravessar as tribulações e
lograr este êxito.*

AGRADECIMENTOS

Às forças do universo que continuamente trabalham para estamos aqui, nesta era, neste tempo-espaço.

Aos pais, amor eterno; orientadores, faróis sábios; irmão, apoio constante; sobrinho, alegria pura, e o sorriso que alegra meu viver.

Às minhas queridas famílias paterna e materna, primos, tios, aos sobrinhos e cunhada, sempre ao meu redor, apoiando-me, trazendo alegria e motivação.

Aos queridos que me acompanharam na jornada da dissertação: aos meus colegas expresso minha profunda gratidão. Amigos e familiares, luz nos dias sombrios.

A crença no divino nos guiou, enfrentamos tristezas e tribulações, fortalecendo laços.

Cada palavra escrita é fruto da união e amor compartilhados. Nossa jornada, repleta de desafios, se tornou triunfo conjunto.

Ao meu orientador Prof. Dr. Vitaly, pelo apoio, amizade, compreensão e competência dados em todo o caminhar desta pesquisa, só tenho que agradecer.

Ao Prof. Dr. Anderson Dourado pelas ricas sugestões no exame de qualificação, a minha admiração e reconhecimento.

Aos professores e alunos do Curso de Pós Graduação PPGEE, participantes da pesquisa, pela sua disposição.

À CAPES, pelos recursos financeiros investidos na pesquisa e atividades acadêmicas.

Com humildade, agradeço por cada presença e apoio nessa conquista.

O ocaso é o caos reorganizado em que para começar ciclos, outro precisou findar-se.

RESUMO

Esta dissertação tem como propósito o uso do aprendizado de máquina para a formulação e concepção de modelos miméticos dos dispositivos fotônicos do tipo *TAPER* destinados ao acoplamento de guias de onda com diferentes estruturas geométricas. O objetivo é avaliar a eficiência de acoplamento em função de variações específicas nas características geométricas na Banda C.

Além da fundamentação teórica, foi necessário preparar uma base de dados para treinar as redes de aprendizado. Esta base consiste em soluções numéricas obtidas por meio de método numérico baseado em elementos finitos, bem como informações previamente publicadas. Esses dados foram consolidados em um conjunto abrangente cujo atributos correspondem às variações na dimensão dos segmentos componentes do *taper* (o comprimento denotado como “ a ”), e a saída é a eficiência de acoplamento (representada por “ η ”).

Desenvolveu-se então uma arquitetura de rede neural com os parâmetros de entrada : comprimento de cada um dos 15 segmentos do *taper*(a), comprimento de onda(λ), índices de refração do núcleo(n_1) e do substrato(n_2) e como parâmetro de saída: a relação entre a potência de entrada(P_{in}) e a potência de saída(P_{out}) dada pela eficiência de acoplamento(η).

Para esta arquitetura explorou-se as variações nos algoritmos de treinamento e nas funções de ativação. Essas variações foram utilizadas para avaliar o desempenho dos modelos propostos , considerando critérios como acurácia, precisão, simplicidade, além dos custos computacionais envolvidos.

Como resultado , as arquiteturas desenvolvidas demonstraram desempenhos melhores que os valores delimitados pelos critérios de parada, com o erro médio quadrático menor do que 10^{-7} e taxa de regressão ou coeficiente de determinação R^2 de 100% em mais de 92% dos totais de 81 modelos avaliados com utilização reduzida de recursos computacionais.

Esse estudo visa contribuir para a melhoria da compreensão e projeto de dispositivos fotônicos por meio da aplicação sinérgica de aprendizado de máquina e técnicas tradicionais.

Palavras-chave: Aprendizado de Máquina, Redes Neurais Artificiais, Guias de Onda, *Multi-Layer Perceptron*, *Taper* Fotônico, Eficiência de Acoplamento, Inteligência Artificial.

ABSTRACT

This dissertation aims to use machine learning for the formulation and design of mimetic models of TAPER-type photonic devices intended for coupling waveguides with different geometric structures. The objective is to evaluate the coupling efficiency as a function of specific variations in geometric characteristics in the C-Band.

In addition to the theoretical foundation, it was necessary to prepare a database to train the learning networks. This database consists of numerical solutions obtained through a finite element-based numerical method, as well as previously published information. These data were consolidated into a comprehensive set whose attributes correspond to variations in the dimensions of the taper segments (the length denoted as “ a ”), and the output is the coupling efficiency (represented by “ η ”).

A neural network architecture was then developed with the input parameters: length of each of the 15 taper segments (a), wavelength (λ), refractive indices of the core (n_1) and substrate (n_2), and as the output parameter: the ratio between the input power (P_{in}) and the output power (P_{out}) given by the coupling efficiency (η).

For this architecture, variations in training algorithms and activation functions were explored. These variations were used to evaluate the performance of the proposed models, considering criteria such as accuracy, precision, simplicity, and the computational costs involved.

As a result, the developed architectures demonstrated performances better than the values defined by the stopping criteria, with a mean squared error less than 10^{-7} and a regression rate or determination coefficient R^2 of 100% in more than 92% of the total of 81 models evaluated with reduced use of computational resources.

This study aims to contribute to the improvement of the understanding and design of photonic devices through the synergistic application of machine learning and traditional techniques

Keywords: Machine Learning; Data Pre-Processing; Tapers; Feedforward; Waveguide; Back-propagation; Coupling Efficiency, Artificial Intelligence.

O homem que move montanhas começa carregando pequenas pedras."

Confúcio- Analectos [551 a.C. e 479 a.C].

SUMÁRIO

ÍNDICE DAS FIGURAS	13
LISTA DE ABREVIATURAS.....	16
PUBLICAÇÕES.....	17
INTRODUÇÃO.....	18
CAPÍTULO 1- INTRODUÇÃO.....	28
1 GUIAS DE ONDA FOTÔNICOS E SUAS APLICAÇÕES.....	29
1.1 Guia de Onda Planar (1D)	30
1.2 Guias de Onda do tipo Canal ou Fibras Ópticas (2D)	31
1.3 Modos de Propagação.....	31
1.4 Guia de onda segmentado (SWG)	33
1.5 Métodos numéricos.....	39
1.6 Análise do dispositivo.....	43
1.7 Conclusão do capítulo.....	46
CAPÍTULO 2 - METODOLOGIA DA PESQUISA.....	47
2 FUNDAMENTOS E PROCEDIMENTOS.	47
2.1 Caminhos da pesquisa.....	49
2.2 Abordagem e Natureza da Pesquisa.....	49
2.3 Procedimentos Metodológicos.....	50
2.4 Consulta à Bibliografia	52
2.4.1 Instrumentos de coleta de dados	53
2.4.2 Análise de Desempenho.....	53
2.5 Conclusão do Capítulo.....	55
CAPÍTULO 3 – APRENDIZADO DE MÁQUINA	56

3	APRENDIZADO DE MÁQUINA OU <i>MACHINE LEARNING</i>	56
3.1	Aprendizado Profundo (<i>Deep Learning</i>)	56
3.2	Redes Neurais Artificiais	57
3.2.1	Bio- Inspiração.....	58
3.2.2	Modelo Matemático do Neurônio.....	59
3.2.3	Funções de Ativação.....	60
3.2.4	Diretrizes aplicadas – Redes Neurais Artificiais	62
3.3	Classificação das Redes Neurais.....	63
3.3.1	Classificação quanto aos Tipos de Aprendizado:	63
3.3.2	Classificação quanto aos Tipos de Arquitetura.....	63
3.4	Algoritmos de Aprendizado.....	66
3.4.1	Validação Cruzada.....	67
3.4.2	Condição de Parada	67
3.4.3	Algoritmos de treinamento implementados.....	68
3.4.4	Backpropagation - Multilayer Perceptron(MLP).....	68
3.5	Conclusão do Capítulo.....	70
	 CAPÍTULO 4 – REDE NEURAL ARTIFICIAL	71
4	O MODELO RNA.....	71
4.1	Síntese do Modelo	72
4.2	Resultados encontrados:	74
4.2.1	Diagrama da RNA:	75
4.2.2	Softwares Elaborados/Códigos Utilizados.....	76
4.3	Consolidação dos resultados.....	77
4.3.1	Fase I – Variação de Algoritmos de Treinamento	77
4.3.2	Fase I – Resultados das Simulações	78
4.3.3	Fase II – Variação de Funções de Ativação por camada.	79
4.3.4	Fase II – Resultado das Simulações.....	82
4.4	Conclusão do capítulo.....	93
	 CAPÍTULO 5 - CONCLUSÕES.....	95
5	CONSIDERAÇÕES FINAIS	95

5.1 Contribuições	95
5.2 Extensões	96
5.3 SUPORTE ACADÊMICO	99
REFERÊNCIAS BIBLIOGRÁFICAS	100
APÊNDICES:	108

ÍNDICE DAS FIGURAS

Figura 1 - Configurações de guias de onda planares. Confinamento do eixo x: guias slab do tipo step index (a) e índice gradual (b). Confinamento lateral: raised strip (c), canal (d) e rib (d). (Adaptado de LIFANTE, 2003).....	31
Figura 2 – Características dos modos em uma estrutura de camada, dependendo do tamanho da componente paralela (constante de propagação) do vetor de ondas. (Adaptado de Yariv, 1989).	32
Figura 3- Padrão de raios ópticos em um guia de onda planar multimodo(SILVA, 2012).	33
Figura 4-- Diferentes formas de propagação de um raio luminoso em um guia de onda: (a) modo de radiação; (b) modo do substrato; (c) modo guiado.	33
Figura 5 - GUIA DE ONDA CONTÍNUO (Fonte: Produzido pelo autor).....	34
Figura 6 - (a) Diagrama do esquemático de dispersão na primeira zona de Brillouin do guia de onda periódico com propagação longitudinal e (b) Uma representação equivalente do 'índice efetivo do modo de Bloch (nB) do guia de onda periódico como uma função da função da razão do comprimento de onda com o período de segmentação $\lambda^- = \lambda/\Lambda$ (HALIR et al., 2015)	35
Figura 7 Configuração do modo 2-D de aumento gradual (up-taper) do afastamento dos segmentos de um guia de onda segmentado (uma visão geral).Fonte : Adaptado de (WEISSMAN; HARDY, 1992).....	37
Figura 8 - Esquemático do Taper 2D. Adaptado de(SISNANDO et al., 2020)	39
Figura 9 - Elementos discretos para definição de subdomínios em FEM. Adaptado de (SEGUNDO, 2018).....	40
Figura 10 - Região discretizada por elementos finitos. Fonte: Produzido pelo autor - adaptado de (RUBIO-MERCEDES et al., 2014).....	41
Figura 11 - Diagrama de análise do taper para obtenção dos dados consolidados. (Adaptado de (OLIVEIRA DAS MERCÊS; DOURADO SISNANDO; RODRIGUEZ ESQUERRE, 2021) e (DOURADO SISNANDO; RODRÍGUEZ ESQUERRE; RUBIO MERCEDES, 2016)	43
Figura 12 - Condições de Contorno para o Taper. Fonte: Adaptado de(SISNANDO et al., 2020).....	51
Figura 13- Tipos de Publicações e Referências por períodos . Fonte: Próprio autor.....	53
Figura 14 - Componentes basicos de um Neurônio. Fonte: Adaptado de(HAYKIN, 1999)	58
Figura 15 - Neurônio Artificial. Fonte: Adaptado de(HAYKIN, 1999)	59
Figura 16 - Representação da Função Sigmoide. Fonte: Produzido pelo autor.	60
Figura 17 - Representação da Função Tangente Hiperbólica. Fonte: Produzido pelo autor.....	61
Figura 18- Representação da Função Linear	62
Figura 19- Representação simples de uma rede Single-Layer.Fonte: Adaptado de (BOLODURINA; ZABRODINA, 2021).....	65

Figura 20-Representação simples de uma rede Multilayer.Fonte: Adaptado de (BOLODURINA; ZABRODINA, 2021).....	65
Figura 21- Arquitetura básica de uma Rede Recorrente. Fonte: Adaptado de (HAYKIN, 1999)	66
Figura 22- Arquitetura Simples RNA.....	69
Figura 23 -Fluxo de Trabalho na Arquitetura da RNA.....	72
Figura 24- (a)Taper segmentado no GID (b)Malha representativa do Taper segmentado.	73
Figura 25- Configuração das variáveis de entrada e saída para a eficiência de acoplamento do conjunto de comprimento do segmento do guia de onda.....	74
Figura 26- RNA - MLP - DIAGRAMA DE TRABALHO. Fonte: Autor.	76
Figura 27 - RNA-MLP -LM - REGRESSÃO.....	78
Figura 28-RNA-MLP -LM - PERFORMANCE.....	78
Figura 29 - RNA-MLP -BR - PERFORMANCE	79
Figura 30 - RNA-MLP -BR - REGRESSÃO	79
Figura 31- Arranjo de Funções de Ativação vs. MSE	80
Figura 32 - Arranjo de Funções de Ativação vs Números de Épocas.....	80
Figura 33 - Arranjo de Funções de Ativação vs Regressão.	81
Figura 34 - Arranjo de Funções de Ativação vs. Tempo Decorrido	81
Figura 35- Tempo de execução para arranjos da tabela 8.....	83
Figura 36-Resultados do Treinamento, Validação e Testes - Arranjo 1 da tabela 8.....	83
Figura 37 - Performance - Arranjo 1 da Tabela 8.....	84
Figura 38- Falha na Validação do Arranjo 1 da tabela 8.	84
Figura 39 -Resultados do Treinamento, Validação e Testes - Arranjo 2 da tabela 8.....	84
Figura 40- Falha na Validação do Arranjo 2 da Tabela 8.....	85
Figura 41 - Performance do Arranjo 2 da Tabela 8.	85
Figura 42-Resultados do Treinamento, Validação e Testes - Arranjo 3 da tabela 8.....	85
Figura 44-Falha na Validação do Arranjo 3 da tabela 8.	86
Figura 43- Performance do Arranjo 3 da Tabela 8	86
Figura 45-Resultados do Treinamento, Validação e Testes - Arranjo 4 da tabela 8.....	86
Figura 46-Performance do Arranjo 4 da tabela 8.....	Erro! Indicador não definido.

Figura 47-Falha na Validação do Arranjo 4 da tabela 8.	87
Figura 48- Resultados do Treinamento, Validação e Testes - Arranjo 5 da tabela 8.....	87
Figura 49 -Performance do Arranjo 5 da tabela 8.....	88
Figura 50 - Performance do Arranjo 5 da tabela 8.....	88
Figura 53 - Os 5 arranjos com menos épocas a atingir ao critério de parada.	89
Figura 54- Os 5 melhores resultados dos Arranjos quanto à MSE	89
Figura 55 - Resultados do Treinamento, Validação e Testes - Arranjo 1 da tabela 10.....	90
Figura 56 - Falha na Validação do Arranjo 1 da tabela 10	91
Figura 57- Performance do Arranjo 1da tabela 10.....	91
Figura 58-Resultados do Treinamento, Validação e Testes - Arranjo 5 da tabela 10.....	91
Figura 59 – Performance do Arranjo 5 da tabela 10.....	92
Figura 60- Falha na Validação do Arranjo 5 da Tabela 10.....	92

LISTA DE ABREVIATURAS

Adaline – Adaptive Linear Element

ANN: Artificial Neural Network

BR: Bayesian Regularization

RNA: Rede Neural Artificial

CNN: Convolutional Neural Network

CWG: Continuous Wave Guide

DNN: Deep Neural Network

FEM: Finite Elements Method

GPU: Graphics Processing Unit

LM: Lavenberg-Macquartd

MLP: Multi-Layer Perceptron

PML: Perfectly Matched Layer

PSW: Periodically Segmented Waveguides

SWG: Segmented Wave Guide

RAM :Random Access Memory

DDR3:Double Data Rate 3

Sigmoid: Sigmoid Activation Function (Função de Ativação Sigmoide)

Softmax: Softmax Activation Function (Função de Ativação Softmax)

TanH: Hyperbolic Tangent (Tangente Hiperbólica)

PUBLICAÇÕES

1. DAS MERCÊS, Viviane Oliveira; SISNANDO, Anderson Dourado; RODRIGUEZ-ESQUERRE, Vitaly F. “*Machine learning modeling of periodical subwavelength tapers coupling efficiency*”. In: Applications of Machine Learning” 2021. SPIE, 2021. p. 299-303.
2. DAS MERCÊS, Viviane Oliveira ; VIEIRA, Luana da França; SISNANDO, Anderson Dourado; RODRIGUEZ ESQUERRE, Vitaly F “*Machine learning analysis of waveguides bending*”. In: Integrated Optics: Design, Devices, Systems and Applications VII. SPIE, 2023. p. PC125750I.
3. DAS MERCÊS, V. O.; SISNANDO, A. D.; RODRIGUEZ-ESQUERRE, V. F. “Aprendizado de Máquina Aplicado na Análise da Eficiência de Tapers Periódicos Segmentados”. Em: XI Conferência Nacional em Comunicações, Redes e Segurança da Informação (ENCOM), 2021, Campina Grande-PB. Anais do XI Conferência Nacional em Comunicações, Redes e Segurança da Informação (ENCOM). Campina Grande/PB: Instituto de Estudos Avançados em Comunicações (Iecom), p. 44-45, 2021.

OUTRAS PUBLICAÇÕES DO AUTOR:

1. DOS SANTOS, Ângelo G.; DAS MERCÊS, Viviane O. “Rastreador do Ponto de Máxima Potência Baseado em Controle Preditivo para o Conversor SEPIC em Sistemas Fotovoltaicos”. In: Congresso Brasileiro de Automática-CBA. 2020.

INTRODUÇÃO

“Ele tinha um apetite natural pelas maravilhas do universo. Queria conhecer a ciência. O problema é que toda a ciência se perdera pelos filtros antes de chegar até ele. Os nossos temas culturais, o nosso sistema educacional, os nossos meios de comunicação haviam traído esse homem. O que a sociedade permitia que escoasse pelos seus canais era principalmente simulacro e confusão. Nunca lhe ensinara como distinguir a ciência verdadeira da imitação barata.”

- Carl Segan, **O mundo assombrado pelos demônios** (1995)

Dentre os problemas que permeiam o campo da fotônica e da inteligência artificial, a especificidade do conhecimento e a produção científica que realmente é disponibilizada se obstem diante da complexidade de cálculos, processamento e análise de resultados por vezes oriundos de trabalhos analíticos e de programas que requerem alta capacidade computacional para lograr algum resultado rápido e satisfatório que possam ser implementados com uma boa relação de custo x benefício. Os estudos investigativos tanto na área da eletro-ótica quanto na engenharia computacional por detrás do aprendizado de máquina demonstram clara vantagem no uso concomitante das técnicas computacionais em solucionadores autômatos de cálculos complexos. Algumas destas pesquisas têm se destacado dentre as relacionadas ao tema deste projeto. Vários autores como, por exemplo Michio Kaku que destaca os avanços na física diante das tecnologias atuais ou Francesco Musumeci e demais pesquisadores que abordam e reforçam a necessidade de uso das ferramentas computacionais no desenvolvimento da tecnologia óptica(KAKU, 2012; MUSUMECI et al., 2019).

Além das características específicas dos dispositivos que a fotônica abrange, uma das principais justificativas para este projeto são as suas dimensões cuja ordem de grandeza é bastante reduzida. E para atingir o grau de precisão necessário, altos custos são demandados, o que pode inviabilizar economicamente até os estudos e pesquisas relacionados. Este tipo de projeto irá depender portanto, de investimento de diversos setores para apoio na produção científica(ANDRADE, 2022), atrelada a investimentos em tecnologias transversais(ANDRADE, 2021), principalmente em países em

desenvolvimento que, de acordo com uma pesquisa da UNESCO , ressalta a corrida contra o tempo no que se refere a saltos tecnológicos.(UNESCO et al., 2021)

Focando na crucial necessidade de aprimoramento profissional no contexto da aplicação de inteligência artificial em fotônica na engenharia, percebemos uma urgência ainda maior na capacitação dos profissionais que atuam nessa área(FLAMINI et al., 2020). Além disso, os engenheiros enfrentam desafios relacionados à aplicação prática, garantindo que os sistemas inteligentes estejam alinhados com os objetivos de projetos específicos(MOREIRA et al., 2023).

Essas são apenas algumas das questões enfrentadas pelos engenheiros no campo da inteligência artificial aplicada à fotônica, tornando essencial uma formação inicial sólida e um compromisso contínuo com o desenvolvimento profissional ao longo de suas carreiras(HUSSAIN et al., 2023).

Nas pesquisas relacionadas ao ensino de inteligência artificial, vê-se uma vasta área interessada em trabalhar com as dificuldades que são encontradas pelos pesquisadores e desenvolvedores, tentando contribuir, através dessas investigações, para a melhoria contínua neste avanço (Aggarwal, 2021). No entanto, observamos uma notável carência no cenário das pesquisas voltadas para a implementação da inteligência artificial nos centros de pesquisa do Brasil, especialmente no que diz respeito à formação continuada dos profissionais que atuam nessa área mesmo estando sob implicações ético-jurídicas(FORNASIER; KNEBEL, 2020). Embora seja reconhecido que esses especialistas se dediquem arduamente à sua atualização nas mais diversas áreas de conhecimento, nem sempre esse esforço abrange os conhecimentos específicos essenciais para a efetiva disseminação da inteligência artificial (Gupta et al., 2023).

Este desafio se torna ainda mais premente considerando que muitos desses profissionais estão envolvidos na formação da próxima geração de cientistas e pesquisadores em inteligência artificial no país. Portanto, a necessidade de uma abordagem mais holística, que englobe tanto os avanços tecnológicos quanto a capacitação recorrente especializada, emerge como um requisito vital para a efetiva integração da inteligência artificial nos centros de pesquisa do Brasil(GT IA, 2023).

Algumas pesquisas (HAMMOND; CAMACHO, 2019) que têm como tema abordagens separadas para aplicação do aprendizado de máquina em diversas áreas, e outras para aplicações computacionais em solucionadores complexos (LIM; PSALTIS, 2022).

Entretanto, uma revisão bibliográfica recentemente realizada pela Universidade de Stanford, na Califórnia (JIANG; CHEN; FAN, 2021), apresenta um registro evolutivo dos últimos anos sobre o uso do aprendizado de máquina. Essa revisão que gerou trabalhos que destacaram a importância de diversificar a aplicabilidade tecnológica na obtenção de bons resultados.

Segundo Jiang, Chen e Fan (JIANG; CHEN; FAN, 2021), os sistemas fotônicos são analisados considerando dois problemas. O primeiro, de forma direta e de mais fácil solução, considera que, dado uma estrutura, a resposta eletromagnética pode ser encontrada usando simuladores eletromagnéticos numéricos bem estabelecidos. Porém, o inverso, que é buscar uma estrutura fotônica adequada para uma resposta eletromagnética desejada, o próprio Jiang considera um processo desafiador e complexo devido ao espaço de solução não convexo onde existem muitos pontos de possíveis soluções, com múltiplos mínimos locais sem uma única solução ótima evidente.

As redes neurais profundas (*deep ANN*), um ramo da inteligência artificial, são poderosas ferramentas que desempenham um papel fundamental na simulação e no projeto de tecnologias fotônicas. Essas redes possuem a capacidade de resolver dois tipos de problemas: os diretos, envolvendo a avaliação precisa de equações complexas, como as equações de Maxwell, elencadas na Tabela 1, e os inversos, que buscam projetar estruturas fotônicas com base em respostas eletromagnéticas desejadas (JIANG; CHEN; FAN, 2021).

Equações de Maxwell	
Lei de Gauss	
Equação 1-Lei de Gauss para Eletricidade	
Forma Integral: $\oint \vec{E} \cdot d\vec{A} = \frac{Q_{int}}{\epsilon_0} \quad (1.1)$	Forma Diferencial: $\vec{\nabla} \cdot \vec{E} = \frac{\rho}{\epsilon_0} \quad (1.2)$
Lei de Gauss para o Magnetismo	
Equação 2 -Lei de Gauss para Magnetismo	
Forma Integral: $\oint \vec{B} \cdot d\vec{A} = 0 \quad (2.1)$	Forma Diferencial: $\vec{\nabla} \cdot \vec{B} = 0 \quad (2.2)$
Lei de Ampère-Maxwell	
Equação 3 -Lei de Ampère- Maxwell	
Forma Integral: $\oint \vec{B} \cdot d\vec{l} = \mu_0 \left(i_c + \frac{\epsilon_0 d\phi_E}{dt} \right) \quad (3.1)$	Forma Diferencial: $\vec{\nabla} \cdot \vec{B} = \mu_0 \vec{j} + \mu_0 \epsilon_0 \frac{\partial \vec{E}}{\partial t} \quad (3.2)$

Lei de Faraday-Lenz	
Equação 4 - Lei de Faraday-Lenz	
Forma Integral: $\oint \vec{E} \cdot d\vec{l} = - \frac{d\phi_B}{dt} \quad (4.1)$	Forma Diferencial: $\vec{\nabla} \cdot \vec{E} = - \frac{\partial \vec{B}}{\partial t} \quad (4.2)$

Tabela 1- Equações de Maxwell nas Formas Integral e Diferencial. Fonte: Adaptado de (SADIKU, 2004)

Em 2012, Noriega colocou em perspectiva o avanço da óptica integrada, comparando-a ao guiamento de ondas eletromagnéticas. Desde o final da década de 1960, quando trabalhos pioneiros como o de Marcatili e Miller em 1969 exploraram a análise de guias de onda, sinalizando o início de uma evolução tecnológica, o conceito de óptica integrada tem sido discutido (MARCATILI; MILLER, 1969). Essa discussão marca o início de uma jornada que culminaria em avanços revolucionários no campo da comunicação e da computação (RUBIO NORIEGA, 2012).

Em seu trabalho, (RUBIO NORIEGA, 2012) contemporiza uma análise comparativa abrangente entre as interconexões ópticas e as interconexões eletrônicas, levando em consideração uma série de critérios essenciais, como largura de banda, multiplexação por divisão de comprimento de onda (WDM), perdas de acoplamento, dimensões físicas, e aspectos indiretos como peso, consumo de energia e economia, que são afetados com a redução de tamanho e da complexidade dos dispositivos, contribuindo para redução do consumo por serem mais eficientes refletindo na viabilidade econômica sistêmica.

Os resultados dessa comparação revelaram que as interconexões ópticas representavam uma promessa significativa no cenário das comunicações de alta velocidade. Novos estudos como o de (SHAHBAZ; BUTT; PIRAMIDOWICZ, 2023) corroboram tal avanço: pesquisadores fizeram progressos substanciais no aumento da sensibilidade e limite de detecção (*LoD*) dos sensores fotônicos de silício (Si). Ao aprimorar o projeto de guias de onda e ressonadores e incorporar materiais avançados, esses sensores aumentaram consideravelmente a precisão e passaram a detectar até mesmo quantidades mínimas de substâncias químicas e biológicas, tornando-os adequados para uma variedade maior de aplicações, incluindo monitoramento ambiental, cuidados de saúde e segurança.

A fotônica integrada, termo utilizado para designar a construção de guias de onda e dispositivos fotônicos em um único chip de silício, oferece inúmeras vantagens, incluindo tamanho reduzido, maior portabilidade e custo-efetividade, possibilitando

avanços em técnicas de fabricação, como nanofabricação e produção em massa de dispositivos fotônicos de silício de forma econômica, com sistemas de maior desempenho e mais eficientes tanto em aplicações diversas(TSIRIGOTIS et al., 2023).

Uma comparação com a transmissão de dados por meio de circuitos elétricos revelou que, nas interconexões ópticas, o desempenho é mais limitado pelo circuito integrado do que pelas perdas no guia óptico ou pela dispersão, especialmente em distâncias de até um metro. Com esse desafio de distância superado, a tecnologia óptica abre caminho para formas inovadoras de projetar sistemas de computação de alto desempenho, segundo(HUNSPERGER, 2009):

“A enumeração das muitas vantagens dos circuitos integrados ópticos e das interconexões de fibras ópticas sugere que a óptica integrada pode substituir completamente a eletrônica convencional, e de fato algumas pessoas sustentam essa visão. No entanto, a opinião predominante é que é mais provável que sistemas ópticos integrados complementem substancialmente, mas não substituam completamente, os sistemas eletrônicos. O custo relativamente alto do desenvolvimento da tecnologia óptica provavelmente limitará suas aplicações a situações em que o desempenho superior alcançável possa justificar os gastos.”

As interconexões ópticas superam limitações de distância em comparação com circuitos elétricos, impulsionando sistemas de computação de alto desempenho. Oferecem benefícios como redução de *crosstalk* (conversa cruzada) e interferência EMI, além de potencial para diminuir tamanho e custo de barramentos de dados de alta velocidade(DUONG et al., 2019).

Com base em pesquisas mais recentes, é importante notar que as interconexões ópticas continuam a ser uma área de grande interesse e desenvolvimento na busca por melhorias significativas no desempenho e na eficiência dos sistemas de comunicação e computação de alta velocidade(WANG et al., 2021; ZHAO; SHI; ZHANG, 2021) .

Um dos desafios principais passou ser na possibilidade da fabricação em larga escala de estruturas complexas com maior assertividade, precisão e efetividade(BHANDARI et al., 2023; O'BRIEN, 2018).

Tais desafios impulsionaram o uso de novos mecanismos que sobrepujassem modelos matemáticos e ferramentas analíticas dependentes de processos manuais e capacidade computacionais limitadas, de forma acessível e pragmática(CHAND; CHEEMA; KAUR, 2023).

A integração entre a inteligência artificial e a fotônica, manifestada no entremeio dessas tecnologias, tem inaugurado novas perspectivas e potencializado avanços

significativos no panorama da pesquisa tecnológica e é o principal motivador do desenvolvimento desta dissertação. A eficiência de ferramentas fundamentada na inteligência artificial destaca-se por fatores como mínima intervenção do usuário e a capacidade de solução de problemas complexos(GOSTIMIROVIC; YE, 2018).

No entanto, é atualmente que testemunhamos a plena maturação desse conceito, e essa transformação é impulsionada, em grande parte, pelas inovações computacionais, notadamente o uso da inteligência artificial e suas derivações, como as redes neurais artificiais, no campo do aprendizado de máquina.

As redes neurais artificiais, inspiradas pelo funcionamento do cérebro humano, provaram ser altamente eficazes na otimização de sistemas complexos, como os dispositivos ópticos utilizados em redes de comunicação(MA et al., 2021). Elas são capazes de aprender e aprimorar continuamente os processos de projeto, tornando os dispositivos ópticos mais eficientes e, ao mesmo tempo, reduzindo seu tamanho(FREIRE et al., 2023).

Porém, a aplicação de redes neurais profundas não é isenta de desafios, especialmente quando se trata de gerenciar recursos computacionais em simulações extensas. Além disso, a natureza não convexa do espaço de solução nos problemas inversos torna a resolução desses desafios complexa(JIANG; CHEN; FAN, 2021).

Apesar desses desafios, o aprendizado profundo tem o potencial de causar um impacto significativo na área de tecnologias fotônicas, representando uma abordagem promissora para a resolução de problemas complexos podendo-se afirmar que:

“Em primeiro lugar, o aprendizado profundo é um método comprovado para a captura, interpolação e otimização de fenômenos altamente complexos em diversos campos, que vão desde o controle de robôs e descoberta de medicamentos até a classificação de imagens e tradução de idiomas. Esses algoritmos estão se tornando ainda mais poderosos, especialmente devido ao recente crescimento explosivo do campo da ciência de dados.” (p.02) (JIANG; CHEN; FAN, 2021)

Neste contexto, o presente trabalho tem como objeto de investigação, a análise e síntese do uso de redes neurais, com diversas arquiteturas lógicas, diante de variações construtivas do dispositivo fotônico conhecido como *taper*(funil), na janela da banda C, em uma faixa específica de comprimento de onda centrada em torno de 1550 nm [1530~1565 nm] utilizada em comunicações ópticas.

Para avançar neste trabalho os objetivos específicos traçados foram :

- Pesquisa de Técnicas de Aprendizado de Máquina: Investigar e compilar as

abordagens de aprendizado de máquina aplicadas no projeto de dispositivos fotônicos, particularmente o guia de onda dielétrico segmentado (taper).

- Simulação do taper para obtenção da solução numérica: Executar simulações tapers com variações geométricas para explorar diferentes configurações e obtenção da eficiência.
- Coleta de Dados para Treinamento: Reunir dados ópticos e geométricos resultantes das simulações para treinar redes neurais artificiais.
- Desenvolvimento de Arquiteturas de Redes Neurais: Criar arquiteturas simplificadas de RNA para obtenção de eficiências preditas diante da alteração geométrica do taper.
- Análise dos Resultados dos Modelos de RNA para o taper: Utilizando as métricas estatísticas, observar a qualidade dos valores preditos pelos modelos.
- Validação e Contribuição da Pesquisa: Formular conclusões que confirmem a eficácia dos resultados obtidos e discutir a relevância do uso de aprendizado de máquina no projeto de dispositivos fotônicos.

Fundamentados nessas justificativas, pretende-se demonstrar que a análise e síntese do modelo de rede neural envolvem a compreensão profunda do funcionamento do modelo do *taper*, que se baseia em ferramentas numéricas, bem como a criação ou adaptação de um novo modelo com o uso de Inteligência Artificial para atender a um conjunto específico de requisitos ou resolver um problema particular, através dos seguintes passos:

1. ANÁLISE DAS FERRAMENTAS NUMÉRICAS

A análise de modelos de acopladores ópticos através do Método dos Elementos Finitos (do inglês *Finite Elements Method* ou FEM) é uma técnica fundamentalmente importante no campo da engenharia e da ciência aplicada. Ela desempenha um papel essencial na avaliação de estruturas complexas e sistemas, permitindo uma compreensão aprofundada de como esses sistemas se comportam sob diferentes condições e carregamentos(que se referem a forças, pressões e outras influências que atuam sobre a estrutura, como forças mecânicas, pressões de fluidos e variações de temperatura). No cerne dessa abordagem estão uma série de questões orientadoras que servem como diretrizes fundamentais para a condução de uma análise FEM eficaz(SCHAD, 1985).

É amplamente utilizada em simulações computacionais para prever e analisar se um produto foi bem dimensionado com base em seu modelo. Ele divide o domínio do problema em elementos finitos, permitindo a resolução eficiente de impasses de engenharia. A seleção criteriosa de elementos, a configuração da malha e a incorporação adequada de parâmetros de material são essenciais para obter resultados confiáveis e relevantes. Contudo, diante do avanço tecnológico, o custo computacional para o uso de ferramentas numéricas torna sua aplicação em larga escala bastante limitada. É uma desvantagem potencial o tempo computacional necessário para resolver problemas complexos com o FEM. Dependendo do tamanho da malha, da complexidade do problema e dos recursos computacionais disponíveis, as simulações com o FEM podem exigir um tempo significativo de processamento(LIU; LI; PARK, 2022).

2. SINTESE DO NOVO MODELO

O novo modelo desenvolvido pro aprendizado de máquina visa obter as eficiências de acoplamento diante às variações geométricas do taper em tempo plausível , com esforço computacional reduzido e que não exija conhecimento específicos e complexos em áreas da engenharia e da física, que o modelo atual não permite.

Com base nisso é designado uma arquitetura de rede neural adequada, com hiperparâmetros definidos , tais como taxa de aprendizado, número de épocas e validação, e o ajuste fino feito por meio de experimentação iterativa até atingir ao(s) critério(s) desejado(s) (LIAO et al., 2022)e (ARDEN; SAFITRI, 2022) .

Essas são configurações externas ao modelo e afetam diretamente o processo de aprendizagem e, portanto, o desempenho do modelo(FILIPPIS et al., 2018)

A etapa de iteração e ajuste é uma fase crítica do processo. Com base nos resultados obtidos na validação do modelo, verifica-se se é necessário realizar novas iterações.

A etapa subsequente, o teste e implantação é o momento em que o modelo é submetido a um conjunto de testes independente e avaliado o resultado logrado.

Um modelo baseado em aprendizado de máquina poderá capturar relações complexas e não lineares em dados provenientes de modelos matemáticos , levando a previsões mais precisas e detalhadas quando comparados entre si.

O processo de projeto e análise de dispositivos fotônicos irá apreciar ganhos em

tempo de simulação e recursos computacionais uma vez que o aprendizado de máquina aceleram o processo de obtenção de respostas além da alta capacidade de generalizar a partir de dados de treinamento para fazer novas previsões o que pode ser útil para lidar com variações e incertezas nos dispositivos fotônicos.(BHANDARI et al., 2023)

Já a integração de algoritmos de aprendizagem de máquina com técnicas de otimização promove descobertas automatizada de projetos otimizados , melhorando significativamente o desempenho dos dispositivos fotônicos , além da flexibilidade e adaptabilidade diante de alterações nas condições operacionais e no ambiente em comparação com modelos fenomenológicos estáticos(DA ROS et al., 2023).

Assim, são apresentados os caminhos da pesquisa nos próximos capítulos, como descritos a seguir:

O Capítulo 1 fornece uma introdução geral à guias de onda fotônicos e suas variações físicas estruturais em guias de onda *SWG* conhecidas como *TAPERS* segmentados, listando suas vantagens em relação aos componentes contínuos e de estruturas homogêneas convencionais. É fornecida uma explicação geral sobre as diferentes ferramentas computacionais utilizadas, bem como uma descrição concisa dos métodos de simulação usados no restante desta dissertação.

Uma introdução ao principal objetivo desta dissertação é apresentada, bem como o objetivo de desenvolver emuladores computacionais baseados em AI que representem eficazmente dispositivos ópticos eficientes e os diversos métodos aplicados.

O Capítulo 2 apresenta a metodologia da pesquisa e suas implicações e detalha sobre a aplicação de técnicas de aprendizado de máquina, com foco em redes neurais artificiais, para otimização de dispositivos de acoplamentos fotônicos, utilizando uma abordagem quantitativa e descritiva, e baseando-se em um levantamento bibliográfico e instrumentos de coleta de dados como simulação e modelagem estruturadas.

O Capítulo 3 aborda a aplicação do aprendizado de máquina no processo de emular os dispositivos, com fundamentações teóricas e síntese do experimento. Os principais aspectos da teoria das redes neurais artificiais, enfatizando a compreensão profunda desta ferramenta e as razões fundamentais que sustentam sua aplicação também são explanados. Além disso, o texto explora a bio-inspiração por trás dos modelos de redes neurais, destacando a estrutura do neurônio biológico e sua abstração matemática, bem como as funções de ativação utilizadas em redes neurais artificiais.

O Capítulo 4 apresenta os resultados alcançados, as previsões disponibilizadas e

análises dos dados obtidos diante da abordagem das arquiteturas das redes estudadas, os resultados são discutidos com base no referencial teórico.

O Capítulo 5 apresenta as conclusões sobre o trabalho, os principais problemas do uso da IA (Inteligência Artificial) em modelos miméticos substitutivos de dispositivos acopladores ópticos, bem como os desafios que ainda precisam ser resolvidos. Uma perspectiva futura para ANN aplicadas à óptica integrada é proposta bem como as considerações finais sobre a pesquisa, e demais complemento textual de suporte à esta dissertação.

CAPÍTULO 1- INTRODUÇÃO

"O conhecimento é uma questão de ciência e nenhuma desonestidade ou presunção é de modo algum permitida. O que é exigido é definitivamente o oposto - honestidade e modéstia."

"*On Practice*" (July 1937), Selected Works, Vol. I, p.300. (Tse-Tung, 1937)

Em fotônica, campo da ciência e tecnologia que abrange o estudo e a manipulação da luz, ocorre o desenvolvimento de componentes e sistemas que utilizam fótons para transmissão, processamento e armazenamento de informações e tem aplicações em diversas áreas .

A fotônica deriva de conhecimento interdisciplinar que envolve física, engenharia elétrica, ciência dos materiais e óptica (TANSU, 2014). Dentro desta área, a fotônica é aplicada em dispositivos, dos quais destacam-se os acopladores ópticos. Suas estruturas têm dimensões na ordem de grandeza de até nanômetros. No entanto, a minimização do dispositivo esbarra em suas próprias dimensões, nos processos de fabricação e na viabilidade de gerar elementos próximos aos ideais sem perdas ou interferências, tanto internas quanto externas (LEE et al., 2002).

Os guias de onda fotônicos têm desempenhado um papel fundamental na revolução das comunicações ópticas e na manipulação de luz em escala cada vez menor e com maneiras de direcionar e controlar a propagação da luz de forma eficiente(FLORES et al., 2016).

No entanto, foi somente nas últimas décadas que os avanços na fabricação e na compreensão teórica permitiram o desenvolvimento de guias de onda fotônicos altamente eficientes e versáteis. Esses dispositivos têm uma ampla gama de aplicações, desde telecomunicações de alta velocidade até sensores avançados e processadores quânticos (KIM; YAN, 2018). Neste breve olhar sobre a evolução dos guias de onda fotônicos, exploraremos como eles passaram de conceitos teóricos a blocos de construção essenciais para a tecnologia moderna.

Em seu livro *Photonic Crystals* (Joannopoulos et al., 2008) demonstra a capacidade de um cristal fotônico tridimensional em confinar a luz em todas as três dimensões. Este fenômeno permite que materiais cuidadosamente projetados restrinjam a luz próxima a um único ponto, criando uma cavidade óptica, a direcionem ao longo de

uma trajetória específica por meio de um guia de onda ou a mantenham presa em uma superfície bidimensional (JOANNOPOULOS et al., 2008).

Porém, é notório que a fabricação de estruturas verdadeiramente periódicas nas três dimensões permanece uma tarefa desafiadora (MOON; YANG, 2010). Nesse capítulo, é feita uma exploração do progresso alcançado ao considerar estruturas mais simples, nomeadamente guias de onda dielétricos periódicos. Também são abordadas a teoria dos guias de onda contínuos e periódicos, suas adaptações, juntamente com suas expressões aproximadas que facilitam o cálculo da eficiência de acoplamento entre eles. Esses guias de onda apresentam um padrão periódico unidimensional ao longo da direção de propagação, mas possuem uma espessura e largura finitas (LI; BURKE, 1992).

1 GUIAS DE ONDA FOTÔNICOS E SUAS APLICAÇÕES

O guia de onda óptico é o elemento fundamental que interconecta os diversos dispositivos de um circuito integrado óptico, assim como uma trilha metálica faz em um circuito elétrico integrado (HUNSPERGER, 2009).

Guias de onda ópticos são dispositivos que, analogamente às linhas de transmissão (LT), transportam ondas propagadas ao longo de seu comprimento. Eles apresentam vantagens devido à variedade de configurações de diferentes modos de propagação, como ondas transversais elétricas (TE) e transversais magnéticas (TM), e suas respostas em aplicações de alta frequência. Diferentemente das linhas de transmissão que perdem sua eficiência devido ao efeito '*skin effect*' (ou efeito pelicular), um fenômeno associado a perdas em dielétricos de guias de onda no intervalo de frequência de micro-ondas. Esses guias de onda são usados para obter maior largura de banda, bem como menor atenuação de sinal (SADIKU, 2004). O guia de onda é o elemento fundamental na tecnologia óptica integrada, definido como uma estrutura óptica que permite o confinamento da luz internamente a seus contornos graças à reflexão interna total, garantida por um meio com um índice de refração maior que o do meio que o cerca. Um tipo de classificação dos guias de onda é em decorrência das dimensões do confinamento da luz, com isso guias planares, guias canalizados ou fibras ópticas ou cristais fotônicos são guias de onda que representam as dimensões do confinamento da luz, a saber: 1D, 2D e 3D respectivamente (LIFANTE, 2003).

Em seu livro, Hunsperger descreve os modos ópticos que podem existir em um guia de onda plano em camadas, que podem ser descritos tanto por um método físico-

óptico, baseado na solução da equação das ondas de Maxwell, quanto pelo método de óptica geométrica, derivada dos princípios da óptica clássica (HUNSPERGER, 2009).

Os guias de onda ópticos podem ser encontrados em uma variedade de configurações, muitas vezes determinadas pelo material e pelo método de fabricação escolhido. No que diz respeito ao perfil do índice de refração, os guias de onda podem ser divididos em dois grupos: os uniformes, que possuem um perfil de índice em forma de degrau, e os não uniformes, que possuem um perfil de índice gradual. Em ambos os casos, o índice de refração no núcleo do guia de onda é maior do que nos meios circundantes.

Classificados quanto à sua geometria os guias podem ser subdivididos em planares ou do tipo canal (cilíndricos ou retangulares) e um dos guias ópticos que se baseiam em variações das estruturas, apresentadas no subitem 1.2, é o guia do tipo "*slab*" (bloco), que é uma estrutura ideal composta por camadas adjacentes com variações de espessura e índices de refração. Estes guias também são conhecidos como guias de onda planares. Essa estrutura possui simetria em relação ao eixo de propagação (eixo z) e em relação a um dos eixos transversais ao eixo de propagação.

Já o vocábulo "fotônico" começou a ganhar amplitude em sua aplicação para denotar guias de ondas e outros componentes ópticos correlatos durante o decênio de 1990, porém décadas antes o IEEE publicou artigos com o termo sendo associado a condicionamento da luz em algum estágio (MOLNAR, 1950) e (EDEN; COLEMAN, 1963). À medida que o campo de pesquisa na fotônica e nas telecomunicações ópticas progrediu, tornou-se patente a necessidade de um termo específico para referenciar dispositivos e sistemas ópticos que fazem uso de fótons como portadores de informação. Consequentemente, a terminologia "fotônico" foi adotada com o intuito de estabelecer uma distinção nítida entre esses dispositivos e os sistemas elétricos convencionais.

1.1 Guia de Onda Planar (1D)

Estas estruturas ópticas, comumente chamadas de guias planares, confinam a radiação óptica em uma dimensão e podem ser classificadas de acordo com a distribuição do índice de refração ou perfis de índice degrau ou gradual. Estes são ajustados a fim de manter a reflexão total interna ao elemento, através da alteração abrupta ou gradual dos índices de refração entre os meios na sua geografia planar. (LIFANTE, 2003)

1.2 Guias de Onda do tipo Canal ou Fibras Ópticas (2D)

Nestas estruturas, o confinamento da luz além de ocorrer em uma dimensão axial, possui um espalhamento em outra dimensão perpendicular, devido à difração, e para evitar este efeito e manter o confinamento, além do contraste entre os índices de refração do meio central com o que o recobre, os contornos laterais também deverão apresentar índice menor. Já as fibras são guias de onda tipo canal, com a geometria cilíndrica já bastante difundida e aperfeiçoada inclusive com propostas de uso de novos materiais e compostos (HAN et al., 2023).

No caso deste guia, o perfil de índices de refração varia apenas ao longo de uma direção, que é o eixo x . Os guias "slab" do tipo "step index" (índice degrau) convencional, ilustrados na Figura 1(a) tem como característica principal uma camada de maior índice de refração adjacente às camadas de menor índice de refração. Nos guias de índice gradual -Figura 1(b)- o índice de refração varia de forma gradual, sendo maior na região do núcleo e menor na direção externa. As estruturas slab são amplamente empregadas na modelagem e simulação do comportamento eletromagnético para o projeto de guias de onda.

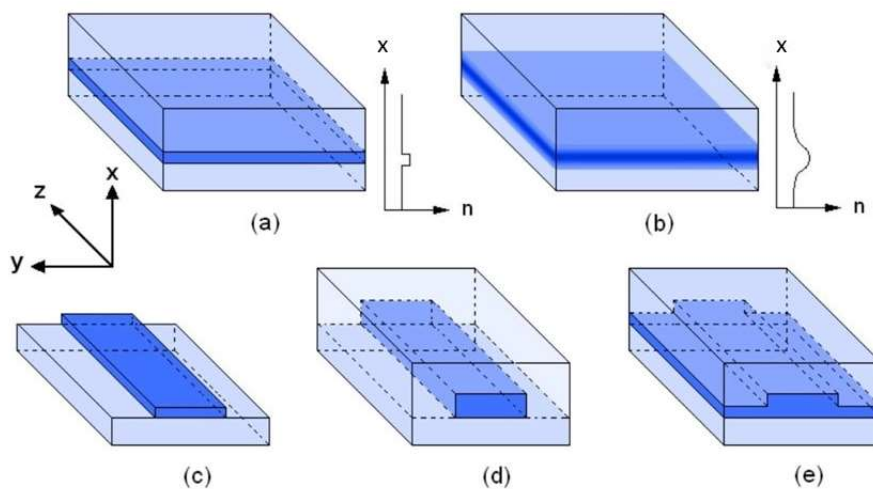


Figura 1 - Configurações de guias de onda planares. Confinamento do eixo x : guias slab do tipo step index (a) e índice gradual (b). Confinamento lateral: raised strip (c), canal (d) e rib (e). (Adaptado de LIFANTE, 2003)

1.3 Modos de Propagação

A descrição da teoria do comportamento dos modos em guias de onda planares é explicada por (HUNSPERGER, 2009), e indicada a seguir na Figura 2 que foi apresentada por (YARIV, 1989). Para o presente trabalho, um tipo de guia tridimensional é abordado, onde considera-se que a mudança da forma do modo com a constante de

propagação β , em uma situação com frequência constante ω , o número de onda k ($k=\omega/c$, c é a velocidade da luz no vácuo) e índices de refração $n_2 > n_3 > n_1$, que ocorre comumente em guias de onda, onde n_2 é maior que tanto n_1 quanto n_3 . Visualizando a propagação da onda na camada interna 2 como uma onda plana propagando-se em um certo ângulo θ em relação ao eixo horizontal nota-se os triângulos resultantes com os lados β , h (hipotenusa) e kn_2 em que esta propagação pode ser considerada formalmente como a de uma onda plana ao longo da direção de h . O modo (a) não é fisicamente possível, já os modos (b), (c) são modos bem confinados, adequados para guia de ondas. Modo (d) é menos útil, pois perde energia ao se propagar.

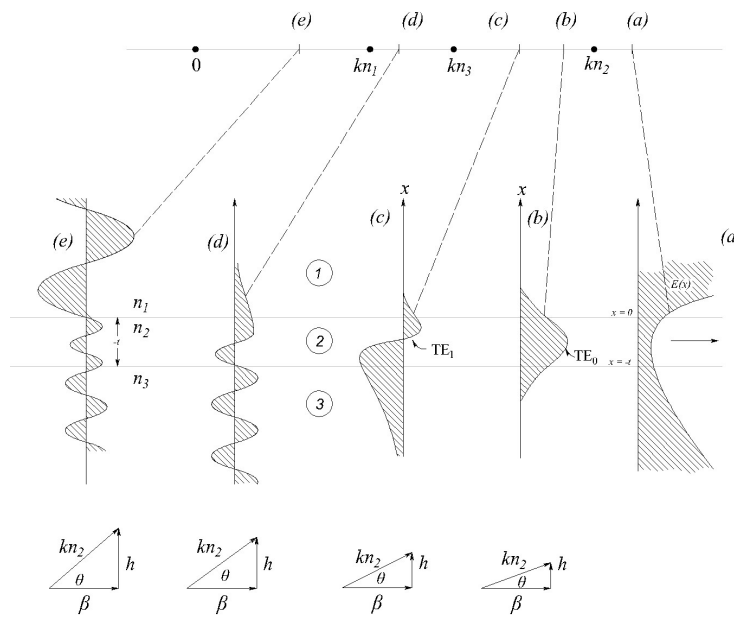


Figura 2 – Características dos modos em uma estrutura de camada, dependendo do tamanho da componente paralela (constante de propagação) do vetor de ondas. (Adaptado de Yariv, 1989).

Outra forma de compreensão dos modos guiados é através da óptica de raios ou óptica geométrica (CACHO; DUARTE, 2010). Nessa abordagem, a luz é considerada como ondas planas que viajam pelo núcleo do guia de onda em trajetórias que formam um padrão de "zig-zag" nas interfaces entre as regiões do núcleo e da cobertura. Cada modo apresenta uma velocidade de fase distinta, resultando em ângulos de reflexão diferentes para cada modo óptico. A Figura 4 apresenta 2 modos de propagação em um guia, utilizando a abordagem da óptica geométrica.

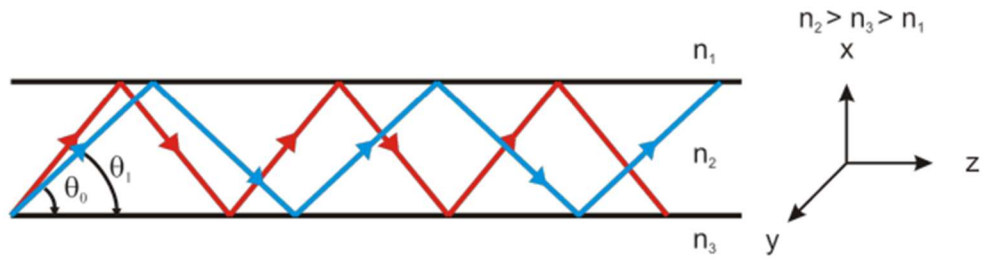


Figura 3- Padrão de raios ópticos em um guia de onda planar multimodo(SILVA, 2012).

Na Figura 4 as formas de propagação são apresentadas e quando o ângulo de incidência é pequeno, $\theta_2 < \theta_1$ e θ_3 , a luz incidente é refratada de acordo com a Lei de Snell e emerge do guia conforme ilustrado na Figura 4(a). Esse modo não confinado da luz é conhecido como "modo radiativo". Quando $\theta_2 < \theta_3 < \theta_1$, temos a situação descrita na Figura 4(b), onde a luz é refratada na interface entre n_3 - n_2 e totalmente refletida na interface n_2 - n_1 . Esse é o "modo do substrato". Por fim, quando $\theta_2 > \theta_1$ e θ_3 ocorre reflexão total em ambas as interfaces, resultando na confinância da luz no núcleo, conforme mostrado na Figura 4(c). Esse é o "modo guiado".

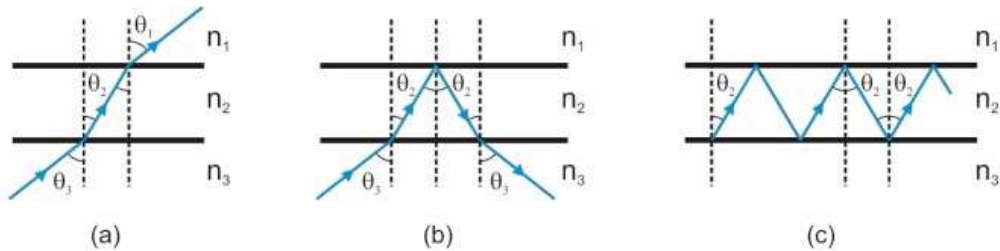


Figura 4-- Diferentes formas de propagação de um raio luminoso em um guia de onda: (a) modo de radiação; (b) modo do substrato; (c) modo guiado.

1.4 Guia de onda segmentado (SWG)

O guia de onda segmentado tem como característica principal a alternância entre elementos com índices de refração diferentes de modo que mantenha a maior proporção da luz guiada confinada na direção desejada. Apesar dessa variação, o guia ainda é capaz de manter a propriedade de transportar a onda óptica em seu interior (HALIR et al., 2015). Guias de onda ópticos desenvolvidos a partir do método de elementos finitos, descritos subitem 1.5.1, tiveram sua abordagem inicial voltada para garantir a transmissão da luz por longas distâncias, porém, com a possibilidade de miniaturização de dispositivos ópticos para ordens de grandezas que passam a tratar o fóton como o elemento fundamental na transmissão da informação, estes guias foram sendo adequados para realizar tratamento da luz tal qual processamento da informação , em dispositivos de

fotônica integrada (PICs- *Photonic Integrated Circuit*) ou em nanotecnologias (DOERR, 2015).

Neste espaço um detalhamento de guias de onda segmentados é apresentado com suas principais aplicações.

1.4.1 Características e materiais dos guias de onda.

Neste trabalho, os elementos componentes dos dispositivos fotônicos são primordialmente compostos de sílica e silício, sem interferência de elementos metálicos ou metamateriais. Neste campo um aprofundamento das características destes materiais e as características construtivas dos guias são apresentados a exemplo do guia de onda contínuo da Figura 5.

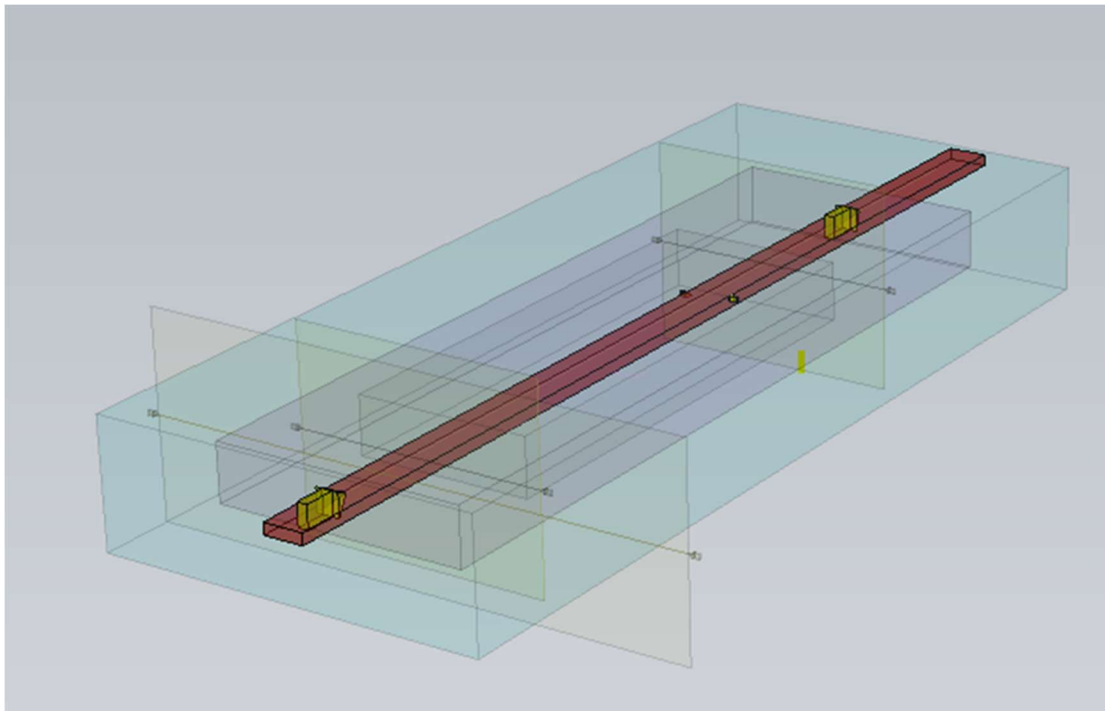


Figura 5 - GUIA DE ONDA CONTÍNUO (Fonte: Produzido pelo autor)

Os guias de onda contínuos (CWG - *Continuous Waveguide*) possuem uma simetria translacional contínua, já que o índice de refração não varia na direção de propagação. Portanto, as características de propagação e a análise modal são amplamente conhecidas (LEE et al., 2009) Por outro lado, os guias de onda periodicamente segmentados (PSW) não exibem uma simetria translacional contínua, mas sim uma periodicidade (RUBIO-NORIEGA, 2012).

1.4.2 Guia de onda segmentado do tipo Grade de Sub-comprimento de Onda (SWG)

Com a dependência do guia de onda segmentado, diante do seu período de segmentação Λ , à frequência de operação (ou para o comprimento de onda equivalente), quando Λ é igual ao comprimento de onda efetivo, irá se obter um dispositivo se comportando como uma grade de Bragg. Já quando Λ for maior do que a condição de Bragg, a luz irá ser parcialmente radiada do guia de onda. Por fim, se Λ for pequeno o suficiente, o guia de onda segmentado irá se comportar de forma parecida com um guia de onda contínuo, como visto a seguir (HALIR et al., 2015).

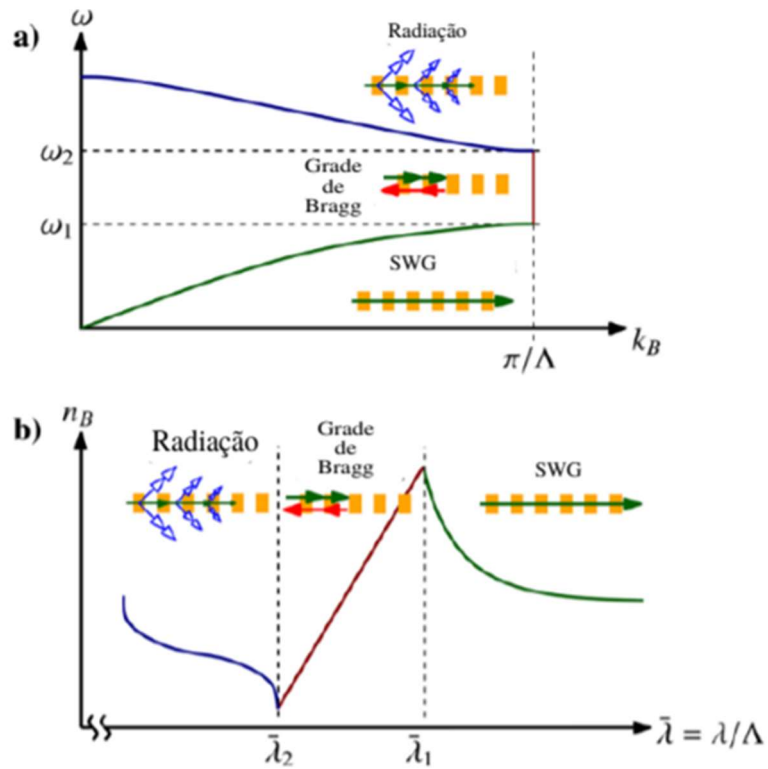


Figura 6 - (a) Diagrama do esquemático de dispersão na primeira zona de Brillouin do guia de onda periódico com propagação longitudinal e (b) Uma representação equivalente do 'índice efetivo do modo de Bloch' (n_B) do guia de onda periódico como uma função da razão do comprimento de onda com o período de segmentação $\bar{\lambda} = \lambda/\Lambda$ (HALIR et al., 2015)

O guia segmentado possibilita controle do desconfinamento da energia do modo na região de guiamento, reduzindo as perdas por espalhamento nos limites entre o revestimento e o núcleo, ao mesmo tempo que permite o controle do índice de refração efetivo sem precisar adicionar outros materiais, aumentando a versatilidade em sua aplicação (Gonçalves et al., 2016).

Eles permitem controle independente dos modos tanto vertical quanto

lateralmente (WEISSMAN et al., 1995). Isso é usado para expandir a seção transversal do modo e reduzir as perdas por dispersão. A diferença chave em relação aos guias de onda contínuos é que o formato do modo varia periodicamente com a distância. Um desafio comum é o alto contraste de índices de refração, que pode resultar em dispersão e comprometer a flexibilidade. Reduzir as paredes verticais dos guias de onda retangulares diminui as perdas de propagação, mas pode limitar o suporte a certas polarizações. A segmentação periódica do guia de onda reduz o índice de refração do núcleo e resulta em baixas perdas de propagação. Os guias de onda segmentados apresentam uma distribuição de perfil de modo semelhante ao contínuo. Os guias de onda segmentados na direção de propagação normal mantêm altas taxas de transmissão mesmo em curvas, especialmente com um raio de curvatura maior que três vezes o comprimento de onda (RUBIO-MERCEDES et al., 2014).

1.4.3 *TAPER*

Taper é uma estrutura de interface que evita a mudança abrupta das geometrias de guias, que garantem uma transição mais suave evitando assim perdas na transmissão. Convencionalmente o *taper* é projetado como uma peça afunilada com dimensões de entrada e saída similares aos dispositivos aos quais se conecta cujo comprimento, largura e período de segmentação são calculados para eficiência máxima, porém extremamente dependentes da transição adiabática para que isso ocorra (WEISSMAN et al., 1995)

Em *tapers* mais longos a perda de potência da luz transmitida chega a ser insignificante no guia de onda de saída (DA SILVA, 2003; DOURADO SISNANDO et al., 2020), porém a diminuição do comprimento do *taper* aumentará a perda de potência (MOLESKY et al., 2018) (RODRÍGUEZ ESQUERRE; ISÍDIO LIMA; DOURADO SISNANDO, 2010), ou seja, o regime adiabático de um acoplador de guia de ondas. Se diminuirmos o comprimento do afunilamento, a perda de potência aumenta (CHUGH et al., 2019b). Ou, para alcançar grande eficiência de acoplamento entre guias de onda segmentadas periódicas (PSW) o uso de um guia de ondas segmentado como um afunilamento entre eles é outra possibilidade (RODRÍGUEZ ESQUERRE; ISÍDIO LIMA; DOURADO SISNANDO, 2010).

Poder manipular o índice de refração do núcleo do guia de ondas ao longo do circuito óptico é uma grande vantagem em trabalhar com estruturas retas PSW, e pode fornecer muitas aplicações, desde estabilidade e alto confinamento da luz no guia de

ondas filtra os modos de ordens superiores (CHUGH et al., 2019b), (DA SILVA FERREIRA et al., 2018). O uso de *taper* segmentados de guias de onda foi proposto por (DA SILVA FERREIRA et al., 2018), (Rodríguez-Esquerre et al., 2010)] e (HORNIK; STINCHCOMBE; WHITE, 1989) com o objetivo de transformar as formas de acoplamento entre os guias de onda.

Um guia de onda segmentado, com modificações graduais nas larguras dos segmentos pode ser usada como um transformador de modo, cuja aplicação resolve problemas de descasamento entre guias com geometrias diferentes (OSKOOI et al., 2012), gerando estruturas que permite um controle do modo da combinação do ciclo de serviço (ou razão de preenchimento) e da largura dos segmentos.

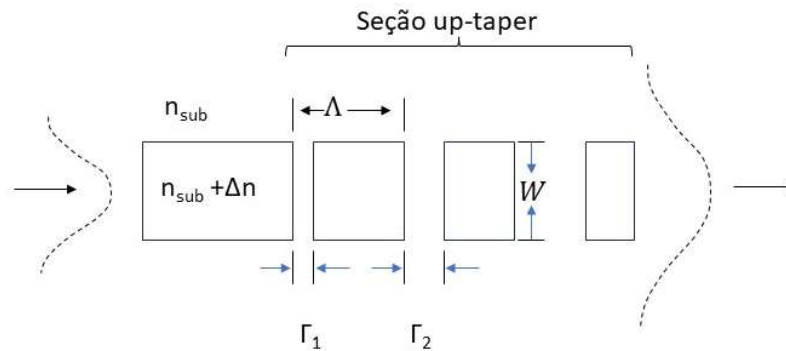


Figura 7 Configuração do modo 2-D de aumento gradual (up-taper) do afastamento dos segmentos de um guia de onda segmentado (uma visão geral). Fonte : Adaptado de (WEISSMAN; HARDY, 1992)

No modelo do *taper* proposto por (WEISSMAN; HARDY, 1992) indicada na Figura 7 o afastamento dos segmentos dos guias obedece a uma taxa de incremento que é ajustável para avaliar a potência propagada, a partir desta representação diversos estudos derivados das variações do período Λ , comprimento de segmento W em detrimento da largura do afastamento entre eles.

Uma estrutura que usa o modo *taper* é um acoplador do tipo guia de onda segmentado ‘afunilado’ que serve de interface entre um guia de onda contínuo e um guia de onda segmentado é utilizado para sintetizar um modelo de rede neural cuja respostas são a representação dos resultados reais.

Esta é uma estratégia viável para aumentar a eficiência de acoplamento entre guias de onda contínuos e guias de onda periodicamente segmentados (PSW) e envolve o uso de um guia de onda segmentado como um acoplador de guia de onda cônico, por meio da variação de seus parâmetros geométricos (como o tamanho do comprimento).

Uma das principais vantagens de lidar com PSWs é a capacidade de manipular o índice de refração efetivo do núcleo ao longo do circuito óptico. Além disso, o confinamento da luz em PSWs permite filtrar os modos de ordem mais elevada (BOCK et al., 2010) Publicações recentes tais como (HALIR et al., 2015) e (DOURADO SISNANDO; RODRÍGUEZ ESQUERRE; RUBIO MERCEDES, 2016; SISNANDO et al., 2020) propuseram o uso de guias de onda segmentados periodicamente para alterar o índice de guia efetivo e, assim, manipular a distribuição do campo dentro dos guias de onda.

1.4.4 *TAPER* SEGMENTADO

A Figura 8 representa o *taper* supracitado com as características construtivas indicadas como Λ , w e a que são o período, a largura e o comprimento dos segmentos, respectivamente, n é o índice de refração dos segmentos. O PSW é composto por silício e sílica com índices de refração alternados $n_1=3.476$ e $n_2=1.444$, respectivamente, e valores de largura variável de $w=270, 300$ e 330 nm . Os segmentos têm um período fixo de $\Lambda=300 \text{ nm}$. Esses parâmetros são usados como variáveis de entrada do modelo sintetizado em redes neurais artificiais e em associação com o um dos principais métodos numéricos (a saber FEM-2D, explanado na próxima subseção) para computar os campos e o acoplamento de potência das estruturas otimizadas. Devido à propriedade de simetria que a estrutura apresenta ao longo da direção y , foi possível reduzir o esforço computacional e o tempo de processamento ao discretizar apenas metade da geometria.

Essa abordagem resulta em uma análise mais eficiente e eficaz, uma vez que permite concentrar os recursos computacionais na metade da estrutura que é suficiente para capturar e representar com precisão o comportamento e as propriedades do sistema como um todo. Além disso, ao adotar essa estratégia de redução de dimensionalidade computacional, é possível obter resultados mais rápidos sem comprometer a precisão ou a representatividade do modelo analítico.

A simplificação computacional para cálculo da eficiência de acoplamento (ou coeficiente de transmissão) do afinamento otimizado do PSW (variando seus segmentos de comprimento) entre um guia de onda contínuo reto (CWG) e um PSW reto é nosso principal problema. A região de afinamento consiste em um número variável de segmentos entre 10 e 15 segmentos para obter um afinamento pequeno. Afinamentos previamente publicados considerando de 33 a 200 segmentos (CHEBEN et al.,

2006) foram observados para avaliação de resultados.

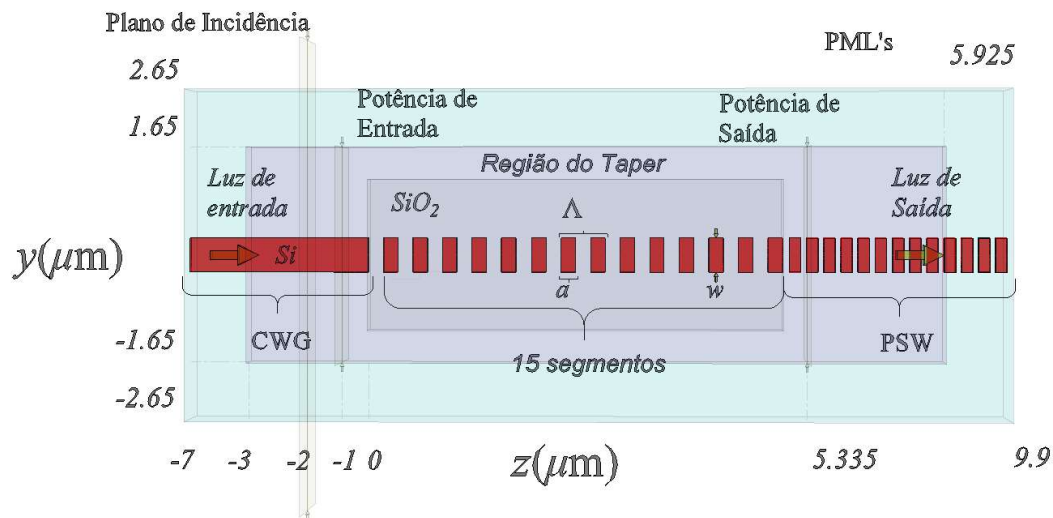


Figura 8 - Esquemático do Taper 2D. Adaptado de(SISNANDO et al., 2020)

1.5 Métodos numéricos

Os métodos numéricos representam a aplicação de algoritmos para formular e resolver problemas matemáticos, utilizando operações aritméticas de menor complexidade. Seu objetivo primordial consiste em encontrar soluções aproximadas para problemas complexos(RUGGIERO; LOPES, 1997).

Um algoritmo, por sua vez, é um conjunto finito de operações organizadas e sequenciadas que viabilizam a resolução de um determinado problema. Ele se manifesta como uma sequência de instruções ou regras definidas que, por meio de passos sucessivos, permitem uma aproximação do resultado real(FILHO, 2018).

A análise numérica se dedica ao estudo de algoritmos que buscam resultados numéricos para problemas provenientes de diversas áreas do conhecimento humano, os quais são modelados matematicamente. Geralmente, os algoritmos dos métodos numéricos podem ser classificados em diretos, recursivos e iterativos. Por exemplo, os métodos iterativos apresentam uma sequência de passos que podem ou não convergir para o valor aproximado da solução exata. A meta da análise numérica é descobrir sequências que se aproximem dos valores exatos, utilizando o menor número possível de operações elementares.

1.5.1 MÉTODOS DOS ELEMENTOS FINITOS

O Método dos Elementos Finitos (FEM) foi concebido para resolver equações diferenciais parciais em domínios finitos. Por meio desse método, é possível transformar domínios contínuos originais de problemas físicos em sistemas de equações lineares. Esse processo de transformação envolve a discretização da região de interesse em subdomínios, conhecidos como elementos finitos, seguido da derivação das equações governantes para cada elemento, do agrupamento desses elementos na região de interesse e, por fim, na resolução do sistema de equações resultante.

Na Figura 9, estão ilustrados alguns dos elementos típicos utilizados nesse método.

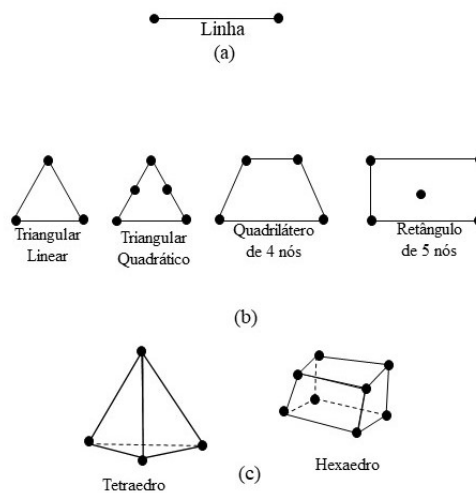


Figura 9 - Elementos discretos para definição de subdomínios em FEM. Adaptado de (SEGUNDO, 2018)

O Método dos Elementos Finitos (FEM) teve seu tratamento matemático inicial proposto por Courant em 1943 (COURANT, 1943). Antes aplicado à análise estrutural, o FEM foi posteriormente adotado no campo do eletromagnetismo em 1989 (SADIKU, 1989), e aperfeiçoadas com técnicas computacionais em anos seguintes (Sadiku, 2009). Atualmente, o FEM é amplamente utilizado em diversas áreas da engenharia elétrica, muitas vezes em conjunto com outras abordagens (DA SILVA, 2003) que incluiu a análise de propagação de feixe (BPM - *Beam Propagation Method*) e aplicado com êxito na análise de guias de onda ópticos.

A simulação de sistemas fotônicos por meio do Método dos Elementos Finitos (FEM) representa um avanço significativo na capacidade de lidar com geometrias complexas e características não homogêneas. Esse método numérico, com sua eficiência e versatilidade, supera outros métodos mais simples, como o Método dos Momentos

(MOM) e o Método das Diferenças Finitas (MDF), tornando-se a escolha preferencial para problemas que envolvem regiões com alta intensidade e variação de campos eletromagnéticos.

Essa abordagem complexa, embora desafiadora em sua implementação, tornou-se viável devido ao avanço computacional e ao desenvolvimento de programas robustos. Agora, problemas que antes seriam considerados impraticáveis podem ser analisados com sucesso, proporcionando uma compreensão mais profunda do comportamento eletromagnético de sistemas diversos.

Em seu trabalho, Sadiku (SADIKU, 1989) estabelece que a análise de um sistema por meio do Método dos Elementos Finitos (FEM) pode ser dividida em quatro etapas:

1. Discretizar a área de interesse em um conjunto limitado de elementos;
2. Estabelecer equações fundamentais específicas para cada elemento individual;
3. Combinação de todos os elementos na região de interesse;
4. Solução do sistema de equações resultante.

Na Figura 10 observa-se um dispositivo *taper*, descrito na subseção 1.4.4, já discretizado em elementos triangulares lineares, para ser submetido à análise.

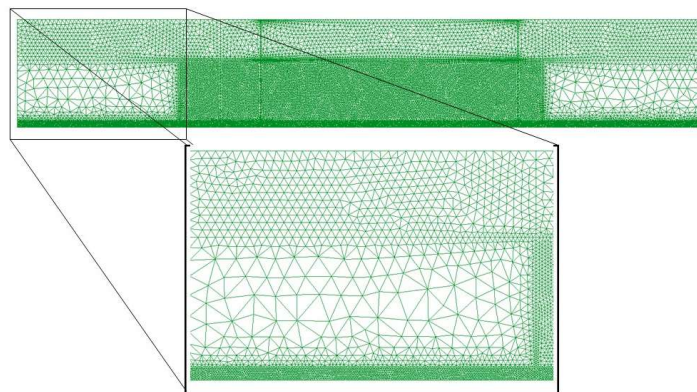


Figura 10 - Região discretizada por elementos finitos. Fonte: Produzido pelo autor - adaptado de (RUBIO-MERCEDES et al., 2014)

1.5.2 SUITE DE SOFTWARES APLICADOS AO TRABALHO:

Como mencionado anteriormente, a criação de modelos matemáticos discretos,

resultantes da aplicação do FEM por meio de programação de computador, é altamente complexa.

É evidente que a implementação do FEM requer pessoal especializado em diversas áreas, como engenharia elétrica, sistemas computacionais e matemática. Devido a isso, as empresas dedicadas ao desenvolvimento de software comercial de EF oferecem seus produtos a preços muito altos, o que exige um investimento considerável para a aquisição e aplicação.

Na área educacional da engenharia elétrica, especialmente nas universidades, o treinamento nesse tipo de metodologia de projeto torna-se difícil de incorporar nos currículos devido aos altos custos das licenças do software comercial. Isso, por sua vez, impede que os estudantes obtenham o conhecimento necessário sobre essas metodologias, o que lhes permitiria ingressar no campo de trabalho de maneira eficaz.

Neste trabalho, a incursão em inteligência artificial, especificamente em uso de redes neurais artificiais, promove um avanço no esforço computacional e de inteligência afim de tornar acessível protótipos e modelos antes de simulações em softwares dedicados proprietários de alto valor.

Para o dispositivo *taper*, descrito em 1.4.4, foram trabalhados a discretização e análise dele através de programas principais: GID , MATLAB® e uma ferramenta numérica.

GID (*Geometry and Date*) é um gerador de malhas para geometrias complexas 2D e 3D. Por outro lado, MATLAB® (MATrix LABoratory) é um resolvidor no qual formulações matemáticas do FEM podem ser implementadas de maneira simples. Tanto o MATLAB® quanto o GID são produtos da década de 80 , e ambos os programas têm sido constantemente avaliados e utilizados em diversas aplicações, tanto educacionais quanto industriais.

Já o FEMTOOL, é uma ferramenta numérica desenvolvida pelos pesquisadores Cosme Rubio e Hugo Figueroa(RUBIO-MERCEDES et al., 2014), com código proprietário desenvolvido parte em FORTRAN de interfaceamento sistêmico com o MATLAB®. Esta ferramenta cunhada de FEM-2D *Scattering Simulator*, foi ajustada para gerar os resultados de eficiência óptica diante de alterações no dispositivo *taper* citado em 1.4.4.

Para tratamento dos resultados obtidos o MATLAB® e Python 3.9 foram utilizados. O Python na versão 3.9 utilizado neste trabalho é uma linguagem orientada a

objetos interativa e de alto nível que possui uma extensa biblioteca de classes para programação, modulação e manipulação de dados.

1.6 Análise do dispositivo

O *taper* designado neste trabalho, foi desenvolvido como guias segmentados são criados configurando colunas retangulares com um índice de refração n_1 , cercadas pelo substrato com índice de refração n_2 . Para delimitar o domínio computacional, utilizam-se camadas de absorção perfeitamente casadas (PML do inglês *Perfectly Matched Layer*), que são adicionadas nas bordas do domínio da simulação para garantir condições de contorno infinitas ou semi-infinitas eficazmente.. Além disso, empregamos o simulador FEMTOOL, indicado em 1.5.2. Esse simulador é baseado no Método dos Elementos Finitos (FEM), no qual calculamos a distribuição do campo e a potência transmitida.

O diagrama da Figura 11 apresenta o processo de obtenção dos dados, através da captação dos resultados dos cálculos numéricos pelo FEM 2D integrado ao MATLAB® para simulação dos modos de propagação e obtenção do cálculo da eficiência do dispositivo mediante alterações nas características óptico-geométricas do *taper*.

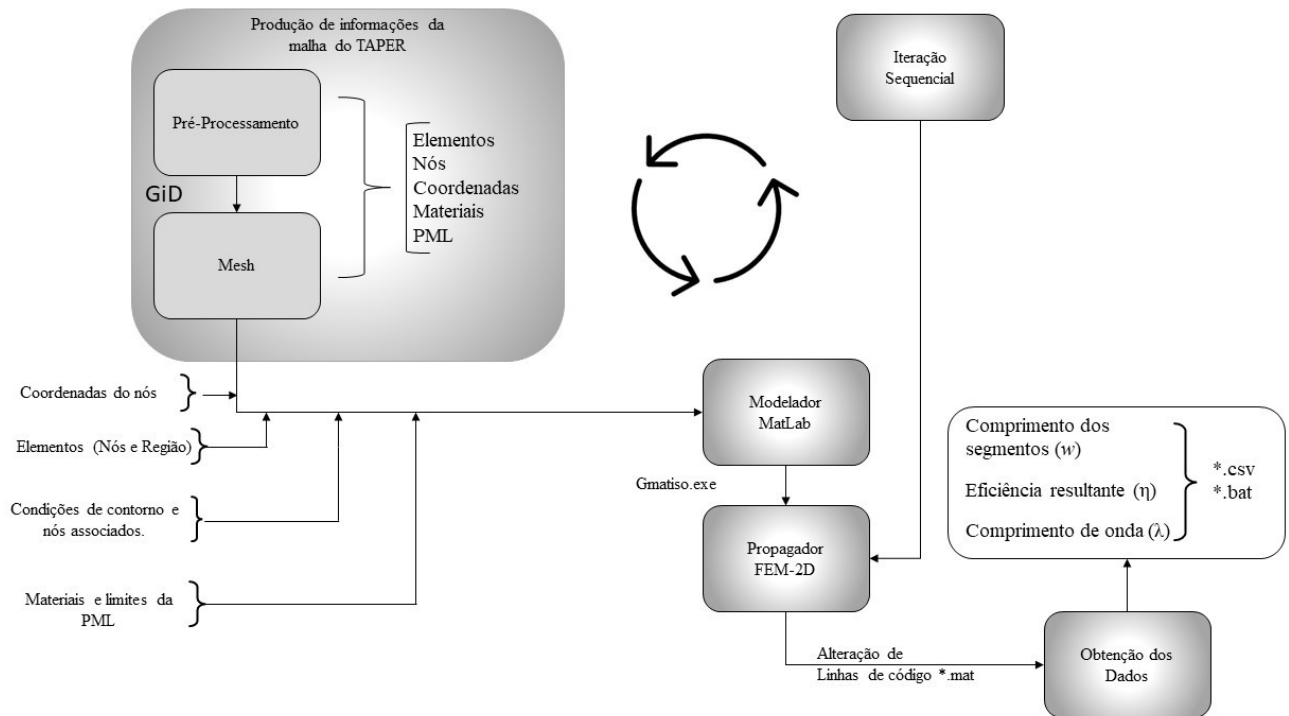


Figura 11 - Diagrama de análise do taper para obtenção dos dados consolidados. (Adaptado de (OLIVEIRA DAS MERCÊS; DOURADO SISNANDO; RODRIGUEZ ESQUERRE, 2021) e (DOURADO SISNANDO; RODRÍGUEZ ESQUERRE; RUBIO MERCEDES, 2016)

As condições de contorno e domínio computacional foram configuradas de acordo com a adaptação da publicação prévia de (DOURADO SISNANDO; RODRÍGUEZ ESQUERRE; RUBIO MERCEDES, 2016; OLIVEIRA DAS MERCÊS; DOURADO SISNANDO; RODRIGUEZ ESQUERRE, 2021)

Este trabalho empregou técnicas avançadas de modelagem, incluindo o método de elemento finito 2D (FEM), para encontrar a eficiência de acoplamento de uma estrutura de *taper* entre uma guia de onda contínua reta (CWG) e uma guia de onda segmentada periódica reta (PSW). O *taper* em questão consiste em 15 segmentos de comprimento variável, visando alcançar o PSW de forma reduzida e eficiente. A geometria da PSW serve como base para o projeto do *taper* segmentado, caracterizada pelo período dos segmentos (Λ), o comprimento do segmento com o índice de refração mais alto (n_1), e a comprimento do segmento (a).

Neste contexto, o guia de onda óptico segmentado que irá sofrer as alterações na largura do segmento, foi construída com materiais alternados de silício e sílica, possuindo índices de refração $n_1=3.476$ e $n_2=1.444$, respectivamente. A região do guia de onda é revestido com sílica, e a largura padrão dos segmentos é estabelecida em 300 nm . Para a PSW, tanto o período quanto a largura dos segmentos foram mantidos fixos em $\Lambda = w = 300 \text{ nm}$. A eficiência e o acoplamento de potência foram quantificados através de cálculos utilizando o método de elemento finito 2D (FEM-2D), obtida pela Equação 5. O comprimento de onda operacional para essa estrutura foi estabelecido em $\lambda=1.55 \mu\text{m}$ [na janela de $1.5 \sim 1.6 \text{ nm}$], por fazer parte do range prioritariamente utilizado em comunicações ópticas com menor susceptibilidade às interferências (SECTOR; ITU, [s.d.]).

$$\eta = \frac{P_{out}}{P_{in}} \quad (5)$$

Equação 5- Eficiência de Acoplamento

Na propagação em um guia periodicamente segmentado, os as formulações relevantes do método dos elementos finitos em duas dimensões para a equação de onda na forma de Helmholtz, incorporando as particularidades da PML anisotrópica, foram apresentadas no estudo dos guias de onda periódicos (BOCK et al., 2010; HALIR et al., 2015).

Em (RUBIO-NORIEGA, 2012) foi avaliada a contribuição da variação do ciclo

de serviço, a variação da largura dos segmentos em guias de onda segmentados no coeficiente de transmissão para comprimentos de onda próximos ao adotado neste trabalho. Além disso é explanado como a função do *taper* é realizado através da variação gradual da largura do guia segmentado, permitindo o acoplamento eficiente entre diferentes guias de onda ou dispositivos ópticos.

O *taper* então pode ser implementado através de diferentes métodos, como o uso de gradientes de índice de refração ou a variação da largura do guia segmentado ao longo de um determinado comprimento.

A região do *taper* objeto deste trabalho consiste em 15 segmentos, a fim de obter uma pequena junção cônica. Como comparação, relatado anteriormente, consideraram 33 a 200 segmentos (SISNANDO et al., 2020). A eficiência ou o acoplamento de energia é calculado usando o método dos elementos finitos (GONÇALVES et al. 2016), (RUBIO-MERCEDES et al., 2009), (SISNANDO et al. 2016).

Os segmentos têm largura e um período fixo $w = 0,300 \mu m$ e esses valores foram escolhidos, pois segundo (RUBIO-MERCEDES et al., 2009), para um comprimento de onda de $\lambda = 1,550 \mu m$, o período da estrutura deve ser da ordem de $300 nm$. As variáveis de entrada dos segmentos de silício $\{a1, a2, ..., a15\}$. Cada comprimento de segmento foi codificado usando 8 bits e podem assumir valores no intervalo $[30 nm, 270 nm]$ com passos de $0,9375 nm$.

Para mostrar a validade e utilidade do modelo neural do *taper* designado os resultados comparativos entre RNA-MLP e FEM são comentadas a seguir.

As soluções numéricas, utilizadas para o treinamento da RNA proposta, foram obtidas utilizando-se um método eficiente de elemento finito de domínio de frequência (RUBIO-MERCEDES, 2012) O domínio computacional foi limitado por $x \in [-7 \mu m, 9,9 \mu m]$ e $y \in [0 \mu m, -2,65 \mu m]$. O campo de incidentes é o modo fundamental do guia de onda contínuo de entrada com uma largura de $300 nm$, com um campo elétrico polarizado ao longo da direção x , colocado em $z = -2 \mu m$, as potências de entrada e saída (P_{in} e P_{out}) são computados pela integração do componente z do vetor *Poynting* (S_z) nos planos $z = -1,000 \mu m$ e $z = 5,335 \mu m$, respectivamente, e as larguras do PML são de $4,000 \mu m$ e $3,975 \mu m$ nos lados esquerdo e direito, respectivamente, e $1,000 \mu m$ na direção y , a fim de evitar reflexos indesejáveis dos limites computacionais. O *taper* começa em $z = 0$. As malhas resultantes têm cerca de 50.000 elementos e 100.000 nós e a convergência foi garantida para esses números. A eficiência do acoplamento foi obtida como a razão P_{out}

$/P_{in}$.

O objetivo principal trata então de viabilizar uma ferramenta computacional acessível, projetada para a criação, análise e desenvolvimento de dispositivos fotônicos, tal qual o *taper* supracitado, a fim de que o impacto em alterações físicas dos dispositivos seja registrado. Essa abordagem se baseia na simplicidade e em um alto nível de abstração, incorporando características importantes, como baixo custo de produção e eficiência notável na transmissão de potência óptica em largura de banda ampla. Esses atributos tornam essa ferramenta tecnicamente relevante no contexto das comunicações.

1.7 Conclusão do capítulo

O capítulo começa discutindo a importância dos guias de onda fotônicos na revolução das comunicações ópticas e na manipulação da luz em escala nanométrica. Destaca-se a necessidade de criar dispositivos compactos, levando em consideração as dimensões, processos de fabricação e a obtenção de resultados ideais sem perdas ou interferências.

É apresentada de maneira sintetizada tanto a teoria dos guias de onda contínuos e periódicos quanto suas expressões aproximadas que facilitam o cálculo da eficiência do acoplamento entre eles. Os recursos computacionais utilizados para análise numérica e modelagem do dispositivo é apresentada no CAPÍTULO 2 seguinte.

A abordagem visa a criação de uma ferramenta computacional acessível e eficaz para o desenvolvimento de dispositivos fotônicos, com foco na transmissão eficiente de potência óptica em largura de banda larga.

CAPÍTULO 2 - METODOLOGIA DA PESQUISA

O melhor instrumento de síntese, e o hífen mais natural entre cientista e filósofo é a história da ciência." (SARTON, 1916).

2 FUNDAMENTOS E PROCEDIMENTOS.

Para o estabelecimento das métricas de análise e síntese do presente trabalho, inicialmente foi realizado um estudo bibliográfico sobre as técnicas computacionais de aprendizado de máquina mais empregadas atualmente. Além disso, foram investigados os problemas e aplicações em que as redes neurais artificiais (RNA) têm sido mais utilizadas e as características conceituais de cada uma dessas técnicas.

Em seguida, como forma de analisar as características, vantagens e limitações, algumas das técnicas de RNA mais conhecidas foram implementadas em MATLAB®.

O MATLAB® suporta duas técnicas de aprendizado de máquina, o supervisionado e o não - supervisionado além das funções de ativação básicas passíveis de uso neste trabalho e juntamente com o *toolbox*(caixa de ferramentas) *nn*tools permitem a construção de modelos, com recursos para ajustes de parâmetros e geração de códigos em produção ou em posterior *deployment*(disponibilização).

As técnicas utilizadas no MATLAB® e que já possuem configurações prévias são as MLP- LM, MLP-BR submetidos a funções de ativação dos tipos TANSIG, LOGSIG, PURELIN em combinações simples, criando um bloco de arquitetura com diversificação das funções de ativação, explanada no subitem 4.3.3.

Para uma análise abrangente e comparação das arquiteturas das RNAs desenvolvidas, este estudo se concentrou na aquisição de dados numéricos, justificados por meio de ferramentas numéricas bem estabelecidas comumente utilizadas na avaliação de soluções absolutas para dados de entrada. Em seguida foi desenvolvido um código para captar as eficiências de potência da saída do *taper* estudado com base em modificações em suas características estruturais. Enfim foram aplicada as técnicas de aprendizado de máquina para correlacionar variações estruturais nos dados de entrada

com sua saída correspondente, seguida pela validação da multicolinearidade.

Essas sequencias de atividades foram delineados para encapsular a essência da pesquisa proposta, facilitando uma abordagem mais robusta e consistente para a análise experimental de várias configurações de RNA. Ao fazê-lo, este estudo buscou destacar a adequação e desempenho de diferentes arquiteturas de RNA na abordagem de tarefas complexas, contribuindo para um entendimento mais detalhado de suas aplicações práticas em cenários do mundo real.

Para todos os problemas selecionados, avaliamos o desempenho dos diferentes arranjos de Redes Neurais em termos de performance e eficácia. Isso foi feito por meio da análise de métricas como erro, tempo de execução e tempo necessário para a convergência de cada método. Os resultados experimentais foram apresentados utilizando como medida de desempenho o erro médio quadrático (MSE do inglês *mean square error*).

Dado que esses algoritmos são baseados em técnicas estocásticas, a avaliação de sua robustez após várias execuções independentes desempenha um papel fundamental na determinação de sua aplicabilidade em situações do mundo real. Portanto, realizamos múltiplas execuções e calculamos a média e o desvio padrão dos resultados obtidos.

Obtenção dos dados no pré-processamento:

Uma vez definidos os limites e domínio específicos para o dispositivo estudado, a obtenção dos dados para pré-processamento é efetuada coletando diretamente os resultados da Femtool FEM2D, através de uma interpolação no código C/Java da IDE do Matlab®, indicada no APENDICE A, armazenados em arquivos com extensões *.bat ou *.csv.

Em seguida, uma avaliação das premissas espaciais que definem as condições de contorno irá certificar a pertinência do conjunto de dados obtidos ao domínio especificado. Para isso, um código de avaliação em Python, foi desenvolvido para informar as métricas de limites mínimos e máximos das alterações dos comprimentos dos segmentos (a_1 a a_{15}) que devem estar entre as faixas absolutas descritas na Tabela 2 a seguir:

Variáveis	a_1	a_2	a_3	a_4	a_5	a_6	a_7	a_8	a_9	a_{10}	a_{11}	a_{12}	a_{13}	a_{14}	a_{15}
min	0,03	0,03469	0,03094	0,03938	0,03703	0,0525	0,03188	0,03	0,075	0,05719	0,04969	0,09563	0,03094	0,04031	0,04781
max	0,26344	0,25219	0,25219	0,25312	0,25453	0,26906	0,25594	0,2625	0,26438	0,26719	0,26156	0,2625	0,2625	0,26531	0,26531

Tabela 2-- Condições de contorno das larguras dos segmentos do taper.

E o mesmo código fornece os dados estatísticos dos resultados numéricos das eficiências encontradas limitadas na faixa de [0 a 100%]. O código está indicado no Apêndice B

O conjunto de dados obtidos para uso neste trabalho possui as seguintes características:

Total de Atributos (a_1 até a_{15} e η) – 4.787.625

Total de Atributos selecionados (a_1 até a_{15} e η) – 4.506.000

Main Dataset ou conjunto de dados principais – resultado.csv ou resultado.xlsx

Os atributos designados no conjunto de dados engloba as variáveis de entrada (colunas a_1 até a_{15}) e a saída (coluna η) que são submetidas na rede neural para treinamento, validação e testes, cuja visão geral estatística do conjunto de dados é apresentada na Tabela 3 abaixo:

Variáveis	a_1	a_2	a_3	a_4	a_5	a_6	a_7	a_8	a_9	a_{10}	a_{11}	a_{12}	a_{13}	a_{14}	a_{15}	η
contagem	281625	281625	281625	281625	281625	281625	281625	281625	281625	281625	281625	281625	281625	281625	281625	281625
média	0,04152	0,08012	0,0788	0,17431	0,1151	0,15569	0,1566	0,1596	0,19757	0,13258	0,18213	0,22083	0,1217	0,22464	0,21186	0,84291
desvio padrão	0,02714	0,0367	0,03382	0,05811	0,04022	0,05414	0,03362	0,03054	0,02702	0,03775	0,043	0,03361	0,03227	0,03296	0,03691	0,10682
min	0,03	0,03469	0,03094	0,03938	0,03703	0,0525	0,03188	0,03	0,075	0,05719	0,04969	0,09563	0,03094	0,04031	0,04781	0,43376
25%	0,03094	0,06844	0,06094	0,10312	0,10453	0,11344	0,15187	0,1575	0,195	0,12656	0,18281	0,22875	0,11531	0,19594	0,21281	0,78657
50%	0,03281	0,07219	0,07031	0,2175	0,11953	0,14344	0,15562	0,16219	0,19875	0,13406	0,19406	0,2325	0,12094	0,24094	0,22687	0,86499
75%	0,04031	0,09469	0,08156	0,22125	0,13453	0,14719	0,165	0,16875	0,20062	0,14625	0,20156	0,23344	0,12094	0,24469	0,2325	0,92414
max	0,26344	0,25219	0,25219	0,25312	0,25453	0,26906	0,25594	0,2625	0,26438	0,26719	0,26156	0,2625	0,2625	0,26531	0,26531	0,99691

Tabela 3 - Estatísticas geral do conjunto de dados principal.

2.1 Caminhos da pesquisa

Tendo em vista todos os pressupostos teóricos e as justificativas para a pesquisa, tem-se que a pesquisa deveria responder às seguintes questões:

"É possível propor modelos miméticos para dispositivos de acoplamentos fotônicos, elaborados por meio da implementação de aprendizado de máquina e aprendizado profundo na otimização de suas características espaciais, apoiados por análise bibliográfica e simulações computacionais, e compará-los diretamente com resultados obtidos por métodos numéricos na faixa de micro-ondas da banda convencional, de forma a garantir que suas configurações apresentem alta eficiência de acoplamento e possam ser fabricadas em larga escala a baixo custo?"

A metodologia foi então desenvolvida para chegar a uma resposta coerente com o trabalho.

2.2 Abordagem e Natureza da Pesquisa

Nesta pesquisa foi adotada a abordagem quantitativa que, conforme (GONÇALVES, 2014) são utilizados procedimentos de quantificação dos dados obtidos por meio de instrumentos padronizados, a partir dos quais se coletam resultados sintetizados com o posterior emprego de recursos e técnicas matemáticas e estatísticas (GONÇALVES, 2014).

Registra-se, no entanto, que a finalidade básica da pesquisa foi identificar, a partir do ponto de vista aritmético, quais são as consequências das alterações dimensionais do elemento desejadas para o satisfatório funcionamento dele.

Dentro dessa abordagem, o enfoque também se fez sobre a pesquisa descritiva, por ter sido necessário descrever a constituição geométrica possível dentro das condições de contorno previstas. De acordo com (GIL, 2002), p. 28), “As pesquisas deste tipo têm como objetivo primordial a descrição das características de determinada população ou fenômeno ou o estabelecimento de relações entre variáveis [...]”.

Em dimensão complementar registra-se a realização de comparações entre os resultados obtidos e aquelas consideradas mais relevantes no processo de síntese. Quanto aos procedimentos de investigação, foi adotada a pesquisa bibliográfica utilizada para a elaboração do referencial teórico, que de acordo com (GIL, 2002, p. 50), “ [...] é desenvolvida a partir de material já elaborado, constituído principalmente de livros e artigos científicos [...]”.

2.3 Procedimentos Metodológicos

Para o treinamento da rede neural artificial RNA-MLP, adotamos uma abordagem detalhada no Capítulo 3.2 para obtenção das informações das entradas e saídas, aplicando as camadas PML definidas para o dispositivo *taper* descrito em 1.4.4 e elencado na Tabela 4 e representado na Figura 12 :

Configuração inicial do <i>taper</i>	15 segmentos (<i>w</i>)			
Limites de comprimento/ <i>w</i>	[30 nm~270 nm]			
PML $\perp$ <i>y</i> (<i>largura</i>)	dy1=	1.000	dy2=	1.000
PML <i>y</i> (<i>inicio</i>)	yb1=	1.650	yb2=	-1.650
PML $\perp$ <i>z</i> (<i>largura</i>)	dz1=	4.000	dz2=	3.975
PML <i>z</i> (<i>inicio</i>)	zb1=	-3.000	zb2=	5.925

Tabela 4- Condições de Contorno do Taper

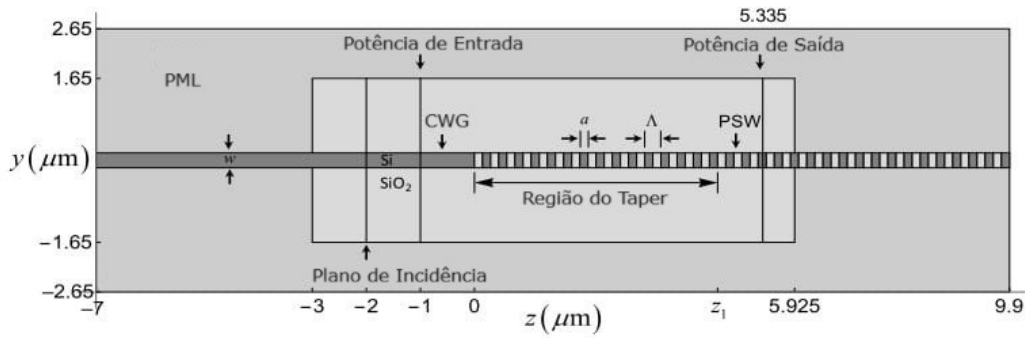


Figura 12 - Condições de Contorno para o Taper. Fonte: Adaptado de(SISNANDO et al., 2020)

O dispositivo *taper* acima possui uma estrutura que permite o controle das distribuições dos modos, é simétrico em relação ao eixo central (y a partir de 0) e tem o domínio de computação para a obtenção dos resultados numéricos vistos em estudos prévios tal como (DOURADO SISNANDO et al., 2020; DOURADO SISNANDO; RODRÍGUEZ ESQUERRE; RUBIO MERCEDES, 2016; RODRÍGUEZ ESQUERRE; ISÍDIO LIMA; DOURADO SISNANDO, 2010) , que já consideraram este ganho no esforço computacional.

A coleta e organização dos dados foram realizados em C e em Python, conforme indicado nos APÊNDICES: A e B.

A implementação da rede RNA-MLP foi realizada em duas fases , a primeira utilizando o software *Matrix Laboratory* (MATLAB®, 2012), e o treinamento foi executado um computador com processador núcleo i7-9750H, relógio de 2,60 GHz e RAM DDR3 de 16 GB, 1.333 MHz, sistema operacional Windows 10 Home Single. e a segunda fase ,também usando o Matlab® executado em um computador notebook equipado com um processador Intel (R) Core (TM) i7-9750H CPU @ 2.60GHz 2.59 GHz e 16 GB de RAM, além de uma GPU GeForce RTX 2060.@1200MHz.

As RNA desenvolvidas possuem 3 camadas ocultas e máximo de 5 camadas ao todo. Uma variação nas funções de ativação das camadas ocultas foi feita para observar as métricas do tempo de processamento e de esforço computacional, de acordo com a Tabela 11, no APÊNDICE C:

Em todas as simulações efetuadas, o comprimento de onda permanece na banda C em $\lambda = [1,5 \text{ a } 1,6] \text{ nm}$, com uma discretização de $0,1 \text{ nm}$. Os valores dos índices de refração dos materiais estrato, substrato e superstrato foram estabelecidos como 3,476,

1,444 e 1,000 , respectivamente. O número de neurônios componentes da RNA foram definido como 15 neurônios para captação das entradas (1 neurônio para cada largura do segmento a_n , onde n é a ordem do segmento a partir de $z=0$) , 3 camadas ocultas com a seguinte distribuição [18-27-28] e uma camada de saída (que representa a eficiência do dispositivo), cada uma com funções de ativação designadas em sequência combinatória designadas no subitem 4.3.3 , sendo alterados conforme avançou as simulações.

O primeiro modelo apresentado na fase I, descrito no subitem 4.3.1, foram elaboradas duas redes neurais submetidas aos dois principais algoritmos de treinamento já bastante difundidos em aprendizado de máquina: O *Levenberg- Marquardt* e a Regularização Bayesiana. Neste modelo não houve variação nas funções de ativação, apenas nos algoritmos de treinamento.

2.4 Consulta à Bibliografia

Como parte inicial e fundamental, foi feito um levantamento bibliográfico tendo como tema: fotônica e dispositivos fotônicos, redes neurais artificiais. Houve consultas principalmente a livros e artigos de periódicos. Além disso, foram consultados os resumos relacionados à temática, disponíveis no banco de dados da Capes, toda referência está relacionada no subitem REFERÊNCIAS BIBLIOGRÁFICAS deste documento. Esse levantamento foi o norte da pesquisa, indicando os caminhos e a maneira como caminhar nessa trajetória, como também refinando o olhar para o objeto de pesquisa.

A consulta levou em consideração temas relacionados à aprendizado de máquina e dispositivos fotônicos, as quais foram refinadas com base em:

- 1 – Palavras chaves
- 2 – Relevância e Reputação das revistas e editoriais.
- 3 – Tempo das Publicações

Ao todo foram referenciados 148 documentos entre publicações, programas de computadores, sítios da internet , livros, artigos e relatórios separados por ano de publicação ou revisão, conforme os gráficos da Figura 13 a seguir.

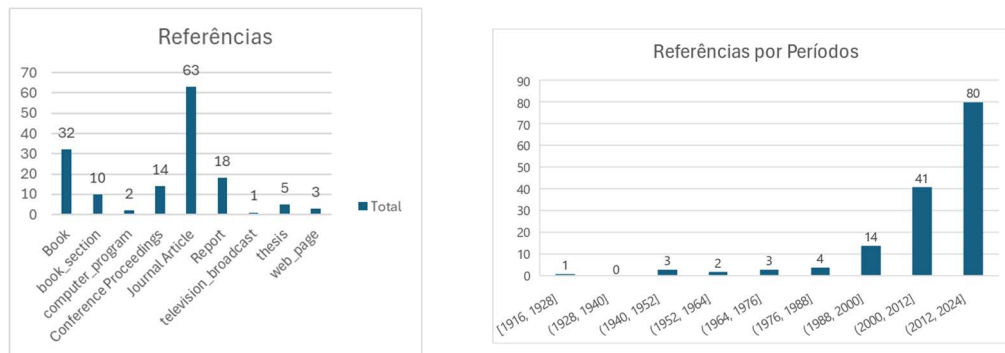


Figura 13- Tipos de Publicações e Referências por períodos . Fonte: Próprio autor.

2.4.1 Instrumentos de coleta de dados

Para o desenvolvimento da pesquisa, foram escolhidos dois instrumentos de coleta de dados: simulação e modelagem estruturadas.

Esse referencial teórico é considerado importante, com vistas à compreensão do que está por trás dos efeitos do modelo proposto.

A Tabela 5 a seguir ilustra os objetivos e caminhos para o seu alcance:

Objetivos	Metodologia	Instrumento	Métricas
ANÁLISE DOCUMENTAL	Comparação dos avanços tecnológicos na área de fotônica e de Inteligência Artificial, com a vanguarda da tecnologia.	Catálogo Eletrônico Bibliotecas Bancos de Dados Bibliográficos Conferências	Relevância Acessibilidade Impacto socioeducacional Confiabilidade
Simulação Computacional	Produção da base de dados para treinamento, validação e testes..	Interpelação de modelos matemáticos e simulações numéricas(FEM2D) para obtenção dos dados com os frameworks Python.	Obtenção da Eficiência Pré-processamento dos dados
Aprendizado de Máquina	Desenvolver e treinar modelos de aprendizado de máquina.	Implementação de algoritmos de aprendizado de máquina utilizando Matlab®.	Precisão do modelo, Latência, Custos Computacionais

Tabela 5- Objetivos e Caminhos Metodológicos

2.4.2 Análise de Desempenho

Nesta pesquisa, os dados para avaliação do desempenho da RNA foram

previamente preparados para submeter ao treinamento de várias configurações da RNA usando conjunto de dados obtidos através do FEM2D em períodos de duração entre 1 e 10 horas . Para treinar a RNA com estes conjuntos e alcançar o total de configurações estabelecidas no capítulo 4, foi observado que cada modelo submetido a estes dados para treinamento, validação e testes não ultrapassou o período de 10 horas.

Para este trabalho considerou-se que os modelos neurais específicos tiveram suas capacidades preditivas avaliadas até 50% do período para o qual foi treinado.

Os passos individuais foram proporcionais ao tempo de aprendizado do modelo. Também se realizaram estudos com intervalos fixos, de menos de 5 minutos para partes do conjunto de dados. Tais avaliações foram superadas pelas simulações abrangentes que garantiram a consolidação das etapas de treinamento, validação e testes simulatórios.

A análise das habilidades preditivas usou como medida e indicador o erro quadrático médio e o indicador de regressão linear a seguir :

- 1) Erro Quadrático Médio (EQM) / *Mean Square Error* (MSE) , dado pela Equação 6 :

$$MSE = \frac{1}{n} \sum_{i=1}^n (y_i^s - y_i^n)^2 \quad (6)$$

Onde

y_i^n = refere-se ao valor da i-ésima quantidade de saída do modelo e,

y_i^s = representa o valor da i-ésima quantidade de saída do sistema

- 2) Indicador de Regressão ou coeficiente de determinação , dado pela Equação 7:

$$R^2 = \frac{\sum_{i=1}^n (y_i^n - \bar{y}^n)^2}{\sum_{i=1}^n (y_i^s - \bar{y}^s)^2} \quad (7)$$

Onde

$\bar{y}^n$ = refere-se ao valor médio do valor de saída do modelo e,

$\bar{y}^s$ = representa o valor médio de saída do sistema.

Tais métricas acima oferecem insights quantitativos sobre a eficácia e a precisão

dos modelos em relação aos dados de teste. O EQM(ou MSE) é mensurado para validar o ajuste do modelo da RNA aos resultados obtidos através das soluções numéricas, e valores de MSE acima de 10^{-7} (valor tipicamente aceito) foram negligenciados . O R^2 é uma métrica de avaliação do desempenho do modelo de aprendizado de máquina, uma medida comum de ajuste do modelo aos dados observados e varia de 0 a 1, onde 1 indica um ajuste perfeito. A utilização do R^2 permite avaliar a eficácia do modelo de aprendizado de máquina na geração de projetos otimizados de dispositivos fotônicos.

2.5 Conclusão do Capítulo

O documento apresenta um estudo completo sobre o uso de redes neurais artificiais (RNAs) em um dispositivo de acoplamentos fotônicos, com foco nas técnicas computacionais de aprendizagem de máquina utilizadas.

Para isso, foi realizada revisão bibliográfica, implementação em MATLAB® e análise abrangente das arquiteturas de RNAs, além da avaliação do desempenho dos diferentes arranjos em termos de eficácia e rendimento. Diversas execuções foram realizadas e métricas, como MAE e coeficiente de determinação, foram calculadas para avaliar a performance dos algoritmos baseados em técnicas estocásticas. Além disso, foram estabelecidos objetivos estratégicos para a pesquisa.

Neste sentido, priorizou-se a obtenção de dados no pré-processamento, a definição precisa dos limites e domínios para o dispositivo estudado, e a proposição de modelos miméticos para acoplamentos fotônicos.

Com o intuito de complementar esta investigação, foi realizado um extenso levantamento bibliográfico sobre fotônica, dispositivos fotônicos no capítulo 1 e redes neurais artificiais no subitem 3.2 do capítulo a seguir.

Além disso, foram empregados instrumentos como simulações e modelagens estruturadas para coletar dados, implementar a rede neural artificial e validar a sua utilização.

Essa metodologia permitiu a aplicação de técnicas de *Machine Learning* na criação de um modelo de RNA específico para obter uma resposta a alteração geométrica do *taper* com resultados satisfatórios e eficientes e compatíveis com as técnicas empregadas na literatura, além de contribuir para a redução de custos e a otimização do processo de fabricação desses dispositivos.

CAPÍTULO 3 – APRENDIZADO DE MÁQUINA

*“O que todos devemos fazer é nos certificar que estamos usando a inteligência artificial de uma maneira que beneficie a humanidade, e não que a deteriore”
(Tom Cook, CEO da Apple)*

Neste capítulo, são abordados os principais aspectos da teoria do aprendizado de máquina e suas derivações enfatizando a compreensão profunda desta ferramenta e as razões fundamentais que sustentam sua aplicação no presente trabalho.

3 APRENDIZADO DE MÁQUINA OU *MACHINE LEARNING*

Aprendizado de Máquina, ou *Machine Learning* em inglês, é o termo associado a uma das formas da inteligência artificial. Ela se caracteriza por possibilitar que sistemas computacionais sejam capazes de aprender ou reagir a dados que lhes são apresentados. É um método de análise de dados que automatiza o processo de construção de modelos analíticos e que é capaz de identificar e reconhecer padrões, e a partir daí, tomem decisões com o mínimo de intervenção humana, sem a necessidade de programação explícita.

Aprendizado de máquina também pode ser definido como a maneira de fornecer conhecimento aos computadores por meio de dados, observações e interação externa para resolver um raciocínio lógico com base em armazenamento, comparação e reatividade aos estímulos externos. Esse conhecimento adquirido permite que os computadores generalizem seus processos corretamente para novas configurações (FAGGELLA; BENGIO, 2019)

3.1 Aprendizado Profundo (*Deep Learning*)

O aprendizado profundo é uma das vertentes do aprendizado de máquina em que redes neurais artificiais, algoritmos inspirados no cérebro humano, aprendem com grandes quantidades de dados, chamados *BigData*. As tarefas de treinamento são ajustadas reiteradamente para melhorar o resultado, agindo diretamente nas camadas intermediárias da rede neural, o que explica a designação de “aprendizagem profunda.

O aprendizado profundo possibilita soluções de problemas complexos, inclusive a partir de um conjunto de dados muito diversificado, não estruturado e interconectado.

E por ser uma técnica de busca evolutiva para resolver os problemas inversos , elimina a necessidade de uma outra rede e evita a constrição desnecessária do espaço de projeto, e alocação indevida de memória; e que através de uma abordagem híbrida usando DNN (*Deep Neural Network*) pode ser usado como substituto numa solução para o problema de otimização(PRESS, 2019).

3.2 Redes Neurais Artificiais

Redes Neurais são modelos matemáticos que se inspiram na estrutura neural de organismos inteligentes e têm a capacidade de adquirir conhecimento por meio da experiência, visando solucionar problemas relacionados à predição, reconhecimento de padrões e classificação (CHAND; CHEEMA; KAUR, 2023) , uma forma avançada de tecnologia computacional que tem aplicações amplas em várias áreas da ciência e da vida humana(BRAGA; LUDERMIR; CARVALHO, 2000; HAYKIN, 1999) . A combinação de IA (Inteligência Artificial) e aprendizado de máquina resulta na concepção rede neural artificial (RNA). As redes neurais artificiais podem ser usadas como componentes baseados em hardware ou software. Diferentes topologias e algoritmos de aprendizado são utilizados nas redes neurais artificiais(GURJAR; PATEL, 2021).

Entre a gama de atividades aplicáveis destaca-se simulação de sistemas não lineares, controle de operações industriais, reconhecimento de padrões, modelagem de problemas eletromagnéticos, processamento e análise de sinais, automação de processos, robótica, classificação de dados, análise de imagens, análise vocal e predição comportamental e estudo meteorológicos(HAYKIN, 1999).

O princípio central da teoria das redes neurais reside no fato de que, ao prover exemplos do relacionamento entre variáveis de entrada (x) e um alvo (t), a rede neural é capaz de capturar essa relação, permitindo a generalização dessas informações para novas possibilidades. Os primeiros modelos matemáticos inspirados em atividades no sistema neural foram descritos pelo neurofisiologista Warren McCulloch e o matemático Walter Pitts(MCCULLOCH; PITTS, 1943) .

Uma característica distintiva e crucial destes modelos é sua capacidade de realizar aproximações universais , ou seja, realizar mapeamento não linear entre entradas e saídas. Além disso, as redes neurais têm a notável habilidade de aprender e se adaptar ao ambiente em que estão operando e possuem a capacidade de generalização conforme discutido (LEE, 1986) . Essas redes utilizam um modelo matemático fundamentado na

estrutura neural dos organismos inteligentes. Na natureza esta estrutura é composta por uma intrincada rede formada por milhões de neurônios interconectados por sinapses. Essas sinapses não apenas transmitem informações entre os neurônios, mas também possuem pesos que modificam as informações à medida que passam pela entrada.

Assim, uma informação pode ser ajustada com incrementos ou decrementos, ou não sofrer alterações quando é transmitida de um neurônio para outro. Pesquisas científicas anteriores tais como demonstraram que a memória da estrutura neural reside nas conexões sinápticas, e não nos próprios neurônios. Essas conexões são modificadas ou mantidas dependendo da quantidade de informação recebida ou durante o processo de aprendizagem.

Em (HAYKIN, 1999) uma definição prática e funcional de uma rede neural é apresentada, e consiste em descrevê-la como uma estrutura interconectada de elementos de processamento simples, conhecidos como unidades ou nós, cuja operação se assemelha de maneira geral ao funcionamento dos neurônios animais. A capacidade de processamento da rede é armazenada nas forças de conexão entre essas unidades, denominadas pesos, que são ajustados por meio de um processo de adaptação durante o treinamento com um conjunto de padrões. Nas subseções a seguir, abordar-se-á conceitos e explicações atreladas à modelagem de um sistema neural para efeito computacional.

3.2.1 Bio- Inspiração

Conforme citado, os modelos de redes neurais são fundamentados no funcionamento do cérebro humano, cuja unidade básica é o neurônio, uma célula especializada na transmissão e armazenamento de informações. Um neurônio típico consiste em três partes distintas: os dendritos (I), o corpo celular ou soma (II) e o axônio (III) de acordo com a Figura 14 a seguir.

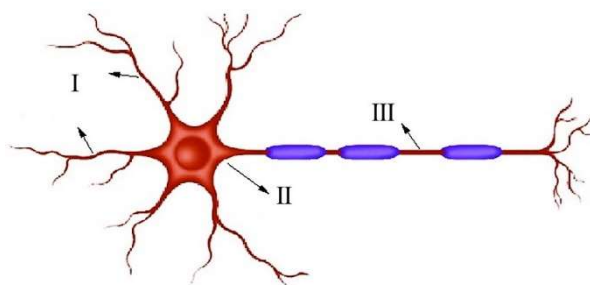


Figura 14 - Componentes básicos de um Neurônio. Fonte: Adaptado de (HAYKIN, 1999)

Os dendritos recebem os estímulos e os encaminham ao corpo celular. No corpo

celular, os estímulos são processados e posteriormente enviados aos axônios, que transmitem os impulsos ou estímulos para outras células. A interface entre os dendritos e as extremidades do axônio é denominada sinapse, responsável por regular o fluxo dos estímulos e impulsos. Quando esses impulsos atingem um limiar específico, o estímulo é transmitido; caso contrário, não é. A interação coletiva dos neurônios regula nossas atividades, desde as mais básicas até as mais complexas. O conjunto interconectado dessas células, conhecido como redes neurais é responsável por essa regulação(BRAGA; LUDERMIR; CARVALHO, 2000) .

3.2.2 Modelo Matemático do Neurônio

Nesta subseção, é apresentada uma visão macro do que se entende por neurônio artificial. Uma abordagem mais completa nesse contexto pode ser encontrada tanto no trabalho de (GURNEY, 2017) quanto nos livros de (HAYKIN, 1999) e (Bishop, 1995).

A Figura 15 ilustra um diagrama simplificado da representação matemática do comportamento do neurônio biológico:

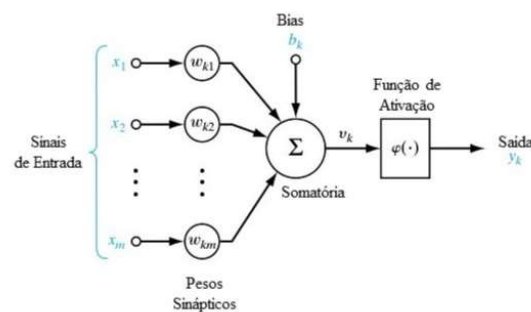


Figura 15 - Neurônio Artificial. Fonte: Adaptado de(HAYKIN, 1999)

As entradas são equiparadas aos impulsos ou informações que chegam aos dendritos (vide I da Figura 14). Os pesos refletem o grau de excitação do impulso, enquanto a função aditiva ou somador se assemelha ao corpo celular (vide II Figura 14), onde a informação é processada. O viés corresponde a uma tendência sistemática do estimador de fornecer uma resposta incorreta. A função de ativação desempenha o papel da sinapse, regulando a saída da informação. Por fim, os axônios representam a saída (vide III da Figura 14).

De acordo com(BRAGA; LUDERMIR; CARVALHO, 2000) um modelo básico de RNA possui diferentes componentes, incluindo:

- **Conjunto de sinapses:** conexões entre os neurônios da RNA. Cada uma delas

possui um peso sináptico;

- **Integrador:** realiza a soma dos sinais de entrada da RNA, ponderados pelos pesos sinápticos;

- **Função de ativação:** restringe a amplitude do valor de saída de um neurônio;

- **Bias:** valor aplicado externamente a cada neurônio e tem o efeito de aumentar ou diminuir a entrada líquida da função de ativação.

3.2.3 Funções de Ativação

Segundo (MACKAY, 1996) as saídas de uma rede y são funções das entradas x , parametrizadas pelos pesos w , ou seja $y = f(x; w)$. Essas funções são conhecidas como funções de ativação.

Em redes neurais artificiais, diversas funções de ativação podem ser aplicadas, desde que sejam contínuas, monotônicas crescente e definidas em $x \in \mathbb{R}$ e $y(x; w) \in [-1, 1]$.

Neste trabalho, 3 funções de ativação foram utilizadas para calibração da arquitetura da rede, e uma específica foi mantida para ajuste da saída, são elas:

a) Função Sigmoid/Logistic:

É definida como uma função crescente com um equilíbrio adequado entre comportamento linear e não linear, e assume um intervalo de variação entre 0 e 1 e demonstrada pela Equação 8 a seguir. O termo *sigmoid* significa ‘em forma de S’ e dentre as funções de ativação conhecidas, esta é a mais frequentemente utilizada, representada na Figura 16 a seguir.

$$f(x) = \frac{1}{1 + e^{-x}} \quad (10)$$

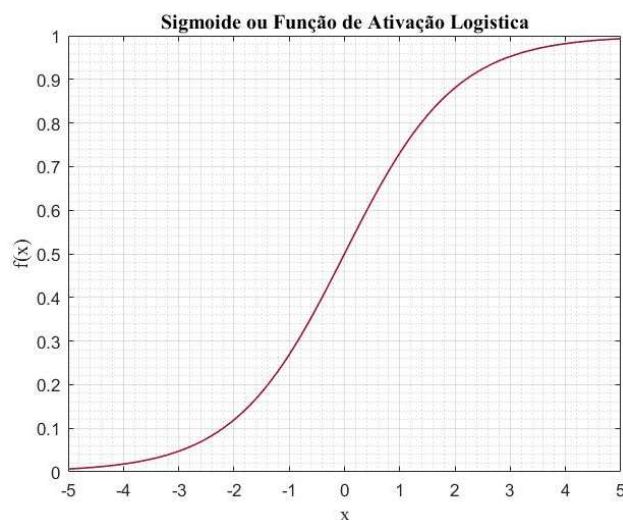


Figura 16 - Representação da Função Sigmoid. Fonte: Produzido pelo autor.

Essa função aceita qualquer valor real como entrada e quanto maior a entrada (mais positiva), mais próximo o valor de saída é de 1.0, por outro lado que quanto menor a entrada (mais negativa), mais próximo o valor de saída é de 0.0. Tanto em (HAYKIN, 1999) quanto em (BISHOP, 1995) destacam a importância da função de ativação sigmoide em redes neurais e suas implicações. O primeiro ressalta como a função de ativação sigmoide logística é fundamental para interpretar as saídas do discriminante como probabilidades a posteriori, indo além de uma simples decisão de classificação. Além disso, ele menciona que a interpretação das saídas das redes em termos de probabilidades é um tópico importante em redes neurais artificiais (HAYKIN, 1999). Isso implica que a utilização da função sigmoide não apenas facilita a classificação de dados, mas também permite uma compreensão mais refinada das saídas em termos de confiança nas previsões.

Já em seu trabalho Bishop complementa essa discussão ao explicar o significado do termo "sigmoide" e como a função sigmoide logística especificamente mapeia intervalos de valores para uma escala de probabilidades entre 0 e 1. Além disso, destaca que, em certas circunstâncias, a função sigmoide logística pode ser aproximada por uma função linear, sugerindo uma relação entre redes com funções de ativação sigmoide e redes lineares. Isso enfatiza a versatilidade da função sigmoide em adaptar-se a diferentes condições de entrada e a sua capacidade de suavizar a transição entre comportamento linear e não linear (BISHOP, 1995).

b) Função Tan-Sigmoid / Tanh

Como mencionado anteriormente, a função *logsig* na prática, esta função é amplamente empregada, sendo uma função diferenciável onde podemos encontrar a inclinação da curva sigmoide entre quaisquer dois pontos. Outra função diferenciável que também é sigmoide é a tangente hiperbólica, com representação gráfica e matemática na

Figura 17

$$f(x) = \frac{2}{1 - e^{-2x}} - 1 \quad (11)$$

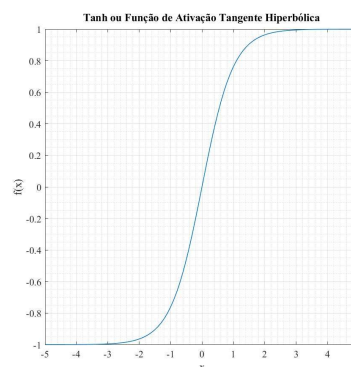


Figura 17 - Representação da Função Tangente Hiperbólica. Fonte: Produzido pelo autor.

c) Função Purelin

A referida função reflete um tipo de função de transferência linear que calcula a saída da camada diante de apenas uma entrada e neurônios com este tipo de função de ativação são utilizados em redes do tipo *single-layer* (ou única camada) tais como redes ADALINE(HAGAN et al., 1996).

Como representação matemática e gráfica temos a figura a seguir:

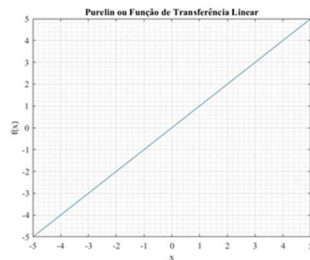


Figura 18- Representação da Função Linear .
Fonte: Produzido pelo autor.

$$f(x) = \text{purelin}(n)=n$$

$$n = \sum_{i=1}^N X_i * W_{ip} + b_p \quad (14)$$

Onde n irá representar a resposta da relação dos pesos e do bias de cada neurônio na saída única desta rede singular.

3.2.4 Diretrizes aplicadas – Redes Neurais

Artificiais

Até o final da década de 1960, as redes neurais receberam grande atenção científica. No entanto, após a publicação de "*Perceptrons*" em 1969 por Marvin Minsky e Seymour Papert, o interesse nessa área diminuiu. A análise de Minsky e Papert revelou limites fundamentais para a rede do tipo de *perceptrons* de camada única e apontou que algoritmos de aprendizado não garantiam a convergência para redes *perceptron* com mais de uma camada(BRAGA; LUDERMIR; CARVALHO, 2000). Este cenário mudou nos anos 1980, em parte graças ao impacto do artigo de Hopfield em 1982 e ao livro de dois volumes de Rumelhart e McClelland em 1986, que tiveram um papel significativo no ressurgimento do interesse em redes neurais e influenciaram o desenvolvimento subsequente da área(HAYKIN, 1999).

O *perceptron*, desenvolvido por Rosenblatt em 1957, foi um dos primeiros exemplos de redes neurais de camada única, aplicado principalmente a problemas de classificação de imagens binárias simples. Seu desempenho foi aprimorado por Rosenblatt por meio da introdução de uma camada de elementos de processamento fixos, que transformam dados brutos de entrada em um discriminante linear generalizado, conceito semelhante ao das *adalines* propostas por Widrow e outros pesquisadores(BISHOP, 1995).

O retorno das pesquisas em redes neurais teve um impacto significativo no avanço dos sistemas de reconhecimento de padrões. Embora as redes neurais artificiais tenham uma gama de funcionalidades além do reconhecimento de padrões, é inegável que os

avanços nessa área dependem cada vez mais dessas redes (JIANG; CHEN; FAN, 2021). Nas próximas subseções, são delineados os princípios fundamentais das redes neurais, apresentadas abordagens diferentes dessas redes nas topologias mais utilizadas, e examinados estudos que desenvolveram modelos eficazes para dispositivo fotônico objeto deste estudo e suas variações.

3.3 Classificação das Redes Neurais

As características pelas quais as redes neurais artificiais podem ser classificadas incluem o tipo de aprendizado e o tipo de arquitetura.

3.3.1 Classificação quanto aos Tipos de Aprendizado:

O aprendizado em redes neurais ocorre por meio de modificações significativas nas sinapses neuronais, que são afetadas pela ativação neuronal. Quando algumas conexões são utilizadas com mais frequência, elas são fortalecidas, enquanto outras são enfraquecidas. Isso explica por que é necessário um período de treinamento ao implementar uma rede neural artificial para uma aplicação específica (HAYKIN, 1999).

São três os tipos básicos de aprendizado de rede neural abordados:

- **Supervisionado:** Nesse tipo, a rede neural aprende a partir de exemplos pré-definidos de entradas e saídas, ajustando os pesos das conexões entre os neurônios até que a diferença entre as saídas produzidas pela rede e as saídas esperadas seja mínima;
- **Não-supervisionado:** Nesse tipo, a rede neural analisa os dados sem ter um objetivo específico, buscando identificar padrões ou características nos dados. O aprendizado é baseado nessas descobertas;
- **Híbrido:** Nesse tipo, a rede neural combina os tipos supervisionado e não-supervisionado. Assim, uma parte da rede pode usar um tipo de aprendizado e outra parte pode usar outro tipo.

3.3.2 Classificação quanto aos Tipos de Arquitetura

Existem diferentes formas de configurar uma rede neural artificial, dependendo do tipo de problema a ser solucionado, como por exemplo, usando configurações *feedforward*, recorrentes ou convolucionais. (JIANG; CHEN; FAN, 2021)

a) *Redes Feedforward*

Redes *feedforward* são um tipo comum de arquitetura neural em que a informação flui apenas em uma direção, da camada de entrada para a camada de saída, sem ciclos ou

retroalimentação.

O processamento de retropropagação em uma rede neural se inicia com toda a rede em um estado inativo. Um padrão externo, composto por um conjunto de sinais, é aplicado à camada de entrada, onde cada sinal estimula um dos nós. Cada nó na camada de entrada gera um sinal de saída simples, cuja magnitude depende do estímulo total recebido pela unidade. As saídas dos nós em uma camada são então repassadas como padrão de entrada para os nós das camadas seguintes. Esse ciclo se repete até que os nós na camada de saída produzam uma resposta para o vetor de entrada atual.

As redes com apenas uma camada de pesos adaptáveis em termos da diversidade de funções são limitantes. Para lidar com mapeamentos mais amplos, considera-se transformações sequenciais correspondentes a redes com várias camadas de pesos adaptáveis. Assim, redes com duas camadas de pesos têm a capacidade de aproximar qualquer mapeamento funcional contínuo. Além disso, é observado que, de maneira mais geral, é possível considerar diagramas de rede arbitrários, desde que não contenham *loops* de *feedback*, garantindo que as saídas da rede possam ser calculadas como funções explícitas das entradas e dos pesos. Redes com funções de ativação diferenciáveis utilizam a retropropagação de erro para encontrar as derivadas em relação aos pesos e vieses. Este método é crucial para muitos algoritmos de treinamento em redes multicamadas. A discussão inclui técnicas para avaliar e aproximar as segundas derivadas de uma função de erro, que compõem a matriz de Hessiana, essencial no contexto de redes neurais (BISHOP, 1995). Quanto à quantidade de camadas tem-se:

i. Single Layer (Uma única camada)

Uma forma simples de organizar uma rede neural artificial é em camadas, onde a primeira camada recebe as entradas e as repassa para a última camada, que produz as saídas. Os neurônios da primeira camada são apenas repetidores, ou seja, eles não modificam o sinal de entrada. Já os neurônios da última camada são processadores, que realizam alguma operação sobre o sinal, como mostrado na Figura 19. Nessa rede, o fluxo de sinais é apenas em uma direção (*feedforward*): os sinais vão da entrada para a saída, sem retorno.

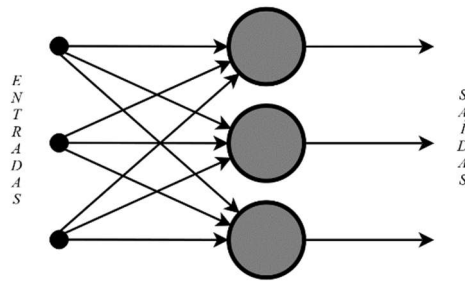


Figura 19- Representação simples de uma rede Single-Layer.Fonte: Adaptado de (BOLODURINA; ZABRODINA, 2021)

ii. Multilayers (Múltiplas Camadas)

Uma outra forma de organizar uma rede neural artificial é em camadas intermediárias ou escondidas, que ficam entre a camada de entrada e a camada de saída. Essas camadas têm neurônios que fazem algum cálculo sobre o sinal, mas não são responsáveis pela saída final(HURWITZ; KIRSCH, 2018). Ao usar uma ou mais camadas intermediárias, a rede ganha mais capacidade de processar e armazenar informações não-lineares. As saídas de cada camada da rede servem como entradas para a próxima camada. As redes *feedforward* com várias camadas são normalmente treinadas com o algoritmo de *retro-propagação* do erro (*error backpropagation*), mas há outros algoritmos possíveis. Esse algoritmo consiste em enviar o sinal de entrada pela rede (*feedforward*) e depois voltar o sinal de erro pela rede (*backpropagation*), como mostrado na Figura 20.

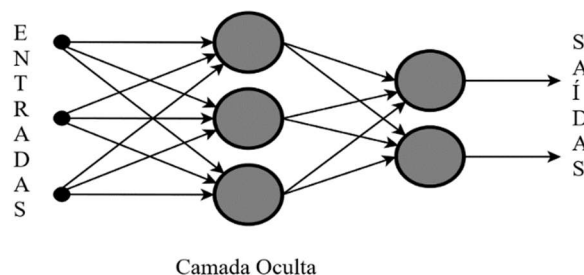


Figura 20-Representação simples de uma rede Multilayer.Fonte: Adaptado de (BOLODURINA; ZABRODINA, 2021)

As topologias Single-Layer (de única camada) ou Multilayer (de múltiplas camadas) são duas topologias dos tipos básicos de RNA com a propagação para frente (*feedforward*).

b) Redes Feedback (Redes Recorrentes)

Uma forma diferente de organizar uma rede neural artificial é em redes

recorrentes, que têm pelo menos um ciclo (*loop*) de *feedback*, vide a Figura 21 seguir. O ciclo de *feedback* afeta muito a habilidade de aprender e o desempenho da rede. Além disso, esse ciclo usa ramos especiais formados por unidades de atraso (z^{-1}), que fazem a rede ter um comportamento dinâmico não-linear, se a rede tiver elementos não-lineares.(HAYKIN, 1999).

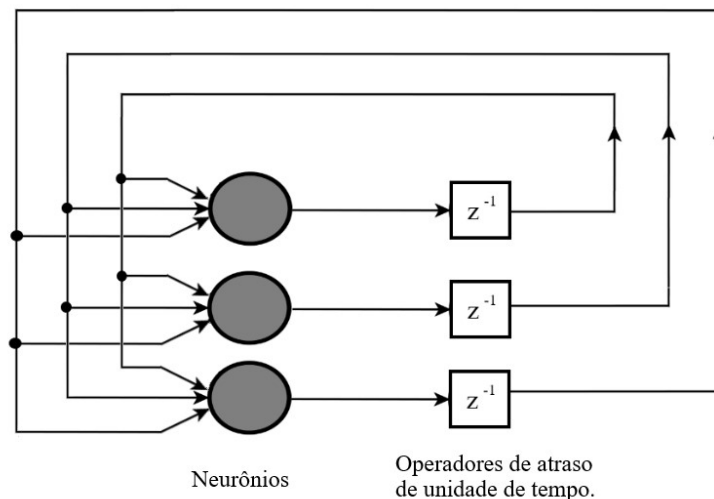


Figura 21- Arquitetura básica de uma Rede Recorrente. Fonte: Adaptado de (HAYKIN, 1999)

3.4 Algoritmos de Aprendizado

Na categorização das redes neurais os tipos básicos de aprendizados introduzidos no subitem 3.3.1, os do tipo aprendizado supervisionado, adotado neste trabalho propicia que a rede receba uma série de padrões ou vetores de entrada com sua saída ou resposta desejada e através da comparação entre os alvos (*targets*) e valores encontrados na saída (*output*) possibilite a alteração dos parâmetros internos (*weight* ou pesos sinápticos) de forma a aproximar a sua resposta à saída desejada(ROSENBLATT; PAPERT, 1969). Este processo é repetido até que a diferença entre a resposta gerada pela rede e a resposta esperada para cada modo de apresentação seja menor que um limite pré-definido. Treinar uma rede neural, ou aprendizagem supervisionada, envolve essencialmente minimizar o erro entre o alvo e a saída.

Tais algoritmos que trabalham neste tipo de aprendizado resumam os seguintes pontos:

- a) A RNA é estimulada a partir de uma abstração inicial com a inserção dos padrões ou atributos em sua camada de entrada.
- b) A estrutura interna da RNA é alterada em função destes padrões e do alvo.
- c) A resposta da RNA ao estímulo é diferente da anterior devido a estas

alterações da estrutura interna.

O algoritmo de retro propagação, explanado no próximo tópico, conta com protocolos e arranjos essenciais no processamento da rede: a validação cruzada e a condição de parada.

3.4.1 Validação Cruzada

O objetivo do algoritmo de retropropagação de erro é ajustar os pesos sinápticos da rede neural para mapear os vetores de entrada e saída. Portanto, a rede pode aprender com esse relacionamento e descobrir outros relacionamentos válidos para os vetores de entrada que ela não conhece. A rede precisa então da capacidade de generalizar o que aprendeu e utilizá-lo com sucesso em novas situações.

A validação cruzada é um método estatístico antigo (KRZANOWSKI, 1987) que ajuda a determinar se uma rede neural pode generalizar bem o que aprendeu. Os dados para treinamento são divididos em duas partes:

1. Conjunto de treinamento (estimativa): usado para treinar a rede neural.
2. Conjunto de validação: usado para verificar ou confirmar se a rede está aprendendo bem.

Depois que a rede neural tiver sido treinada por um certo número de épocas (uma época que usa todos os dados de treinamento), o treinamento é interrompido e a rede é testada usando dados de validação. Dependendo do problema que queremos resolver, esse processo continua até que a rede apresente bons resultados via dados de validação.

A razão para esse particionamento é testar o modelo usando dados diferentes dos dados usados para variar os pesos sinápticos. Isso também evita que a rede neural seja sobre treinada (*overtraining*). Uma rede '*overtraining*' grava os dados de treinamento e não generaliza bem o que aprendeu.

3.4.2 Condição de Parada

O algoritmo de treinamento precisa de um critério ou condição para cessar o ciclo de comparação do *backpropagation*. Como condição ideal encontrar um valor mínimo global como critério de encerramento, para evitar que só seja encontrado um valor mínimo local no erro, outros critérios de parada usados concomitantemente(KRZANOWSKI, 1987):

1. Treinar a rede neural por um número fixo de vezes ou épocas.

2. Parar quando o erro médio quadrático (MSE) chegar em um valor designado.
3. Parar quando a taxa de diminuição do erro médio quadrático entre os ciclos atingir a um valor mínimo.
4. Parar quando o erro da validação cruzada aumenta.

3.4.3 Algoritmos de treinamento implementados

Neste trabalho, os softwares utilizados foram elencados no subitem 1.5.2 juntamente com a linguagem de programação para redes neurais junto com uma planilha EXCEL™. Entre as alternativas de configuração de rede existentes, optou-se pelas que apresentavam a maior responsividade nos seus algoritmos em duas fases: A primeira variando os algoritmos de treinamento mantendo todos os parâmetros a fim de identificar qual dos algoritmos apresenta a melhor resposta e esta fase foi consolidada através das publicações indicadas neste documento e detalhadas nos próximos tópicos. E a segunda etapa mantendo o algoritmo inalterado e variando as funções de ativação, a fim de cobrir o maior número de possibilidades possível, possibilitando que a análise dos modelos fosse realizada em uma ampla variedade de aspectos. Dessa forma, os tipos de rede escolhidos para o estudo foram os seguintes:

- *Perceptron* Multicamadas (ou *Multilayer Perceptron*-MLP; rede “feed-forward”) com algoritmo de Regularização Bayesiana (MLP- BR), onde BR é do inglês *Bayesian Regularization*, na fase I.
- *Perceptron* Multicamadas (ou MLP; rede “feed-forward”) com algoritmo de *Levenberg-Marquardt* (MLP – LM), na fase I e fase II.

Na fase II, a rede neural definida foi submetida ao seguinte escalonamento das funções de ativação, mantendo a camada de saída definida para a função PURELIN apenas, conforme indicada na tabela no APÊNDICE .

3.4.4 Backpropagation - Multilayer Perceptron(MLP)

As redes neurais artificiais multicamadas (RNAs) para este trabalho receberam diferentes valores de segmentos de comprimento para cada segmento que compõe o *taper* referenciado na Figura 8 como entradas e, em seguida, produz a estimativa da eficiência de acoplamento de energia por meio da função de ativação. Algoritmos de retropropagação (ou *backpropagation*) Levenberg-Marquardt é então operado com a função de desempenho, que é uma função da estimativa baseada em RNA.

As variáveis de peso e tendência são ajustadas de acordo com o método dos algoritmos selecionados (neste caso Levenberg-Marquardt), e o algoritmo de retropropagação é usado para calcular a matriz Jacobiana da função de desempenho em relação às variáveis de peso e viés (ROSENBLATT; PAPERT, 1969). Com pesos e vieses atualizados, RNA ainda estima o acoplamento de eficiência de energia na próxima etapa. A arquitetura da RNA mostrada na Figura 22. Demonstra a interação entre as saídas das camadas precedentes e entrada das próximas no intermeio entre as camadas de entrada e saída da rede neural.

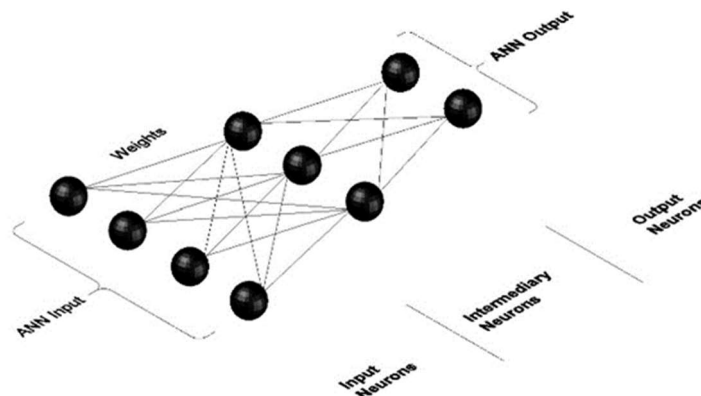


Figura 22- Arquitetura Simples RNA

O modelo *Perceptron* de camada única é ampliado pelo Perceptron Multicamadas, que tem uma camada de unidades de entrada ligada a uma ou mais camadas de unidades intermediárias, chamadas de camadas ocultas, e uma camada de unidades de saída (ROSENBLATT; PAPERT, 1969). Essa rede aprende de forma supervisionada, usando principalmente o algoritmo “*back-propagation*”.

Algumas características importantes são:

- as unidades da rede têm uma função de ativação não linear (normalmente a função sigmoide);
- a rede tem uma ou mais camadas ocultas, que lhe dão a capacidade de resolver problemas complexos, captando as características mais relevantes dos padrões de entrada;
- a rede tem alto nível de conectividade, que possibilita interação entre as unidades.

O *Perceptron* de camada única se diferencia então, desse modelo *multilayer* que apresenta a(s) camada(s) oculta(s), que lhe conferem maior poder computacional, mas também tornam o algoritmo de treinamento mais difícil (KUSSL et al., 2001).

A arquitetura da RNA é concebida considerando o número de entradas (a),

número de saídas (η), tamanho do banco de dados e as funções de ativação para o correto equilíbrio das camadas ocultas, representada no CAPÍTULO 4.

Os critérios de parada da RNA definidos para este trabalho são: 2.000 épocas ou erro quadrático médio (MSE) de 10^{-7} .

3.5 Conclusão do Capítulo

O capítulo apresenta uma visão geral do Aprendizado de Máquina (*Machine Learning*) e do Aprendizado Profundo (*Deep Learning*). Ele destaca a importância dessas tecnologias na análise de dados e na tomada de decisões automatizadas.

O capítulo discute os diferentes tipos de aprendizado em redes neurais, incluindo o aprendizado supervisionado, que é um método que permite que a rede receba uma série de padrões ou vetores de entrada com sua saída ou resposta desejada. Além disso, o capítulo aborda os algoritmos de aprendizado, com ênfase no algoritmo de retropropagação.

O capítulo também discute a arquitetura de redes neurais e os critérios de parada para o treinamento fornecendo uma visão geral revisada do Aprendizado de Máquina e do Aprendizado Profundo utilizados neste trabalho.

O Capítulo 3 apresenta a teoria das redes neurais artificiais, enfatizando sua importância e aplicabilidade em diversos campos. Primeiro, é apresentada a base biológica dos neurônios, descrevendo como os dendritos recebem estímulos, que são processados dentro do corpo celular e transmitidos através das sinapses pelos axônios. A seguir, é discutida a abstração matemática dos neurônios artificiais, com foco nos componentes e modelos matemáticos que representam seu comportamento.

Este capítulo também explora funções de ativação, enfatizando a relação entre entradas e saídas da rede neural parametrizadas por pesos. Além disso, são propostos modelos matemáticos inspirados na atividade do sistema neural, enfatizando a capacidade das redes neurais de realizar aproximações universais e mapear entradas para saídas não lineares, bem como sua capacidade de aprender e se adaptar ao seu ambiente.

A importância das redes neurais artificiais hoje tem sido enfatizada, destacando sua capacidade de aquisição de conhecimento através da experiência e sua ampla aplicação em simulação de sistemas não lineares, controle de operação industrial, reconhecimento de padrões, análise de imagens, automação de processos etc. Por fim, este capítulo destaca a necessidade de uma compreensão mais profunda desta ferramenta para explorar o seu potencial.

CAPÍTULO 4 – REDE NEURAL ARTIFICIAL

“As formas primitivas de inteligência artificial que já temos provaram ser muito úteis, mas eu acredito que se o desenvolvimento de uma inteligência artificial completa acontecer, poderá significar o fim da raça humana. Uma vez que os humanos desenvolvam a inteligência artificial, ela seguirá seu próprio caminho e se redesenhará a uma taxa cada vez maior. Os humanos, limitados pela lenta evolução biológica, não poderiam competir e seriam substituídos”
(Stephen Hawking , 2014)

4 O MODELO RNA.

Para estabelecer um modelo representativo de um sistema, todas as entradas / saídas que irão compor os atributos para a criação de um conjunto de dados específico para o *taper* foram obtidas através de diversas alterações analíticas das larguras dos 15 segmentos deste *taper* e a eficiência resultante de cada uma destas alterações.

Para isso, um programa computacional externo (Modelo Híbrido Físico-ML) foi adaptado e submetido a várias entradas de valores entre limites específicos, para obtenção de resultados matemáticos absolutos(OLIVEIRA DAS MERCÊS; DOURADO SISNANDO; RODRIGUEZ ESQUERRE, 2021).

Um grande bloco de dados foi gerado com o tempo de processamento que são indicados numa tabela (Etapa 1). Este conjunto de dados é composto 15 entradas ou variáveis independentes e uma saída (ou variável dependente), onde cada linha no banco de dados representa 1 modelo de treinamento com 15 larguras(1 por coluna) e uma única saída na última coluna. Constituindo assim um uma relação 15x1 para cada conjunto iterativo a ser aplicado a RNA.

Em seguida, dividindo todo o bloco em 2 conjuntos no tamanho de proporção de 70 e 30% para treinamento e validação, respectivamente, para se obter um equilíbrio entre ter dados suficientes para treinamento e teste conforme subitem 3.4.1, e um bloco gerado em separado de ser usado como conjunto de teste (Etapa2).

Após uma abstração inicial, uma possível arquitetura de uma rede neural artificial é proposta considerando o número de entradas, número de saídas, tamanho do banco de dados bem como as funções de ativação para equilíbrio correto das camadas ocultas (Etapa 3), submetendo a rede neural artificial para um treinamento considerando: número de épocas, erro quadrático médio objetivo e desempenho de regressão como critérios de parada (Etapa 4).

Após sucessivos envios de treinamentos, os objetivos são avaliados com o conjunto de validação (Etapa 5) e, caso não sejam atendidos, o modelo recebe uma nova configuração e é novamente submetido a treinamento e validação (Etapa 6). Uma vez atingidos os objetivos, os testes são realizados utilizando um conjunto de dados diferente dos anteriores para verificar a qualidade dos resultados e sua precisão (Etapa 7).

Seguindo essas etapas, uma RNA foi construída com características de entrada, camadas ocultas e saídas, seguindo o fluxo de trabalho conforme Figura 23 a seguir:

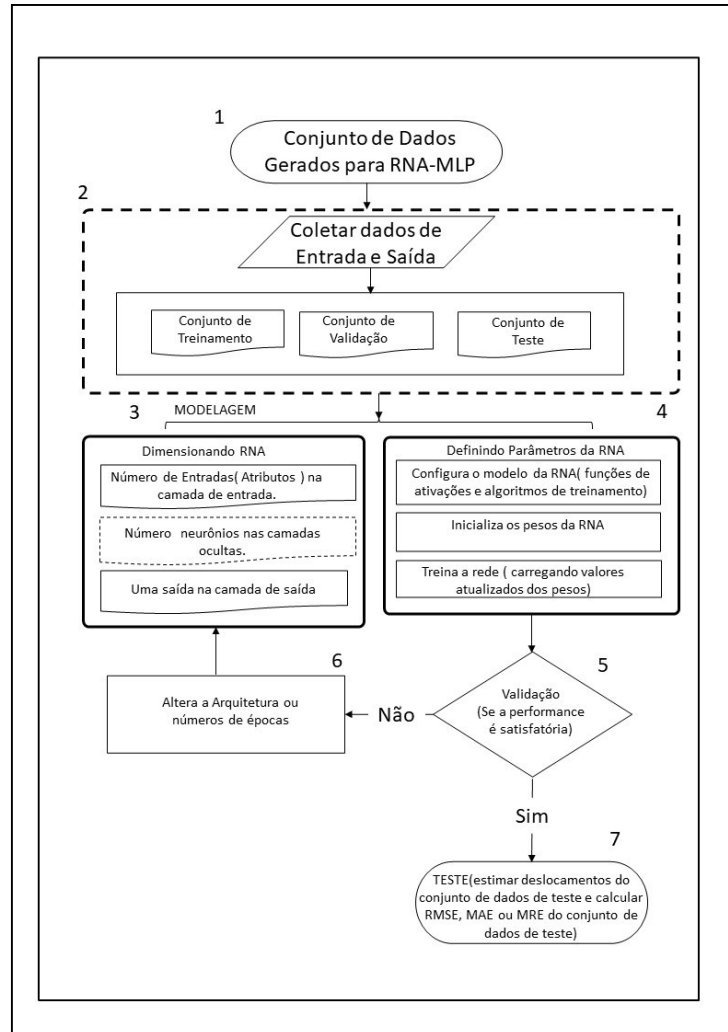


Figura 23 -Fluxo de Trabalho na Arquitetura da RNA

4.1 Síntese do Modelo

A partir da realização da pesquisa bibliográfica a respeito dos dispositivos passíveis de modelagem e confrontar resultados por outros métodos, foram analisadas estruturas existentes e seus respectivos resultados, os quais são comparados ao longo da Banda C, conforme mencionado no subitem 2.3 e para diferentes respostas eficientes de

acoplamento das estruturas avaliadas. Em seguida, são submetidas a um processo de otimização dos pesos atuadores nos neurônios com base nas respostas dos algoritmos de otimização bioinspirados gerando novas modelagens e transportados para os simuladores e elaborada novas análises da eficiência de cada estrutura modelada.

A primeira etapa neste método consiste em discretizar a geometria do dispositivo em análise. Para isto, é utilizado o programa computacional GiD. Esta etapa faz parte dos conjuntos de ações do pré-processamento.

A Figura 24 mostra um exemplo de um acoplador óptico segmentado do tipo *taper* montado no GiD, onde se observa a malha discretizada com indicação de nós e elemento geométrico fundamental, que neste caso é triangular. Nestas imagens verificam-se as posições do PML, contornos e planos de incidência e do limiar captação da potência óptica.

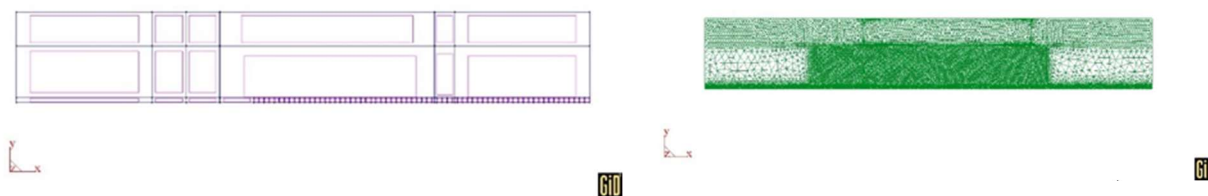


Figura 24- (a)Taper segmentado no GiD (b)Malha representativa do Taper segmentado.

O processamento consiste em resolver a equação de onda escalar para o campo magnético no domínio da frequência, as condições de contorno e campo magnético inicial, com o equacionamento final da etapa de processamento resulta em um sistema linear de equações, através do programa computacional FEM2D(RUBIO-MERCEDES et al., 2009) .Um banco de dados robusto é gerado conforme metodologia indicada no subitem 2.

Uma vez resolvido este sistema, inicia-se a etapa de pós-processamento. Para isto, utilizam-se programas computacionais comerciais como o MATLAB® e o WOLFRAM MATHEMATICA ©, indicados no subitem 1.5.2.

O próximo passo é emular uma Rede Neural Artificial designada através de softwares com capacidade computacional robusta, submeter a treinamento, validação e teste, sendo recompensado com um processo de otimização atuante nos kernels das camadas ocultas até observar os critérios de paradas.

A seguir, com as informações estruturadas do novo modelo desejado, um modelo de RNA é determinado e comparado a um método numérico elaborado a partir do valor encontrado como saída da rede neural, cuja configuração é apresentada na Figura 25. A

arquitetura da RNA é concebida considerando o número de entradas (a), número de saídas (η), tamanho do banco de dados, bem como as funções de ativação para o equilíbrio correto das camadas ocultas.

Após sucessivos ciclos de treinamento, os valores previstos são avaliados com o conjunto de validação e em caso de divergências o modelo é alterado para uma nova configuração e ele é treinado e validado. Uma vez alcançado o critério de parada, definido no 3.4.4, testes são realizados usando um conjunto de dados diferente dos anteriores para verificar a qualidade dos resultados e sua precisão

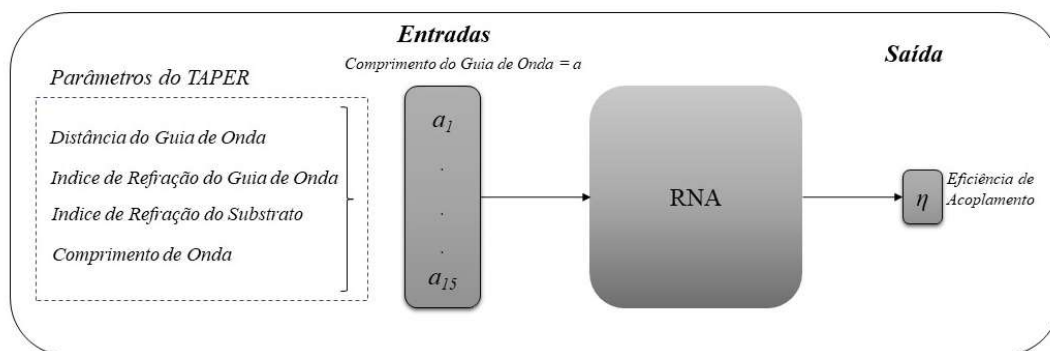


Figura 25- Configuração das variáveis de entrada e saída para a eficiência de acoplamento do conjunto de comprimento do segmento do guia de onda.

Os resultados obtidos apresentam uma boa concordância quando comparados com os dados experimentais.

4.2 Resultados encontrados:

O dispositivo modelado foi mostrado na Figura 8 do subitem 2.3 que apresenta seus respectivos parâmetros ópticos e geométricos que são as variáveis de entrada, como indicada no subitem 1.4.4, e com o uso dos algoritmos desenvolvidos em associação com o FEM-2D, calculou-se os campos e a potência das estruturas resultantes. Assim para este problema calculamos a eficiência de acoplamento ou coeficiente de transmissão entre um CWG direto 2D e um PSW direto 2D através de um cone PSW com comprimento variável segmentos, como a variável dependente, considerada como alvo.

Conforme descrito no subitem 1.4.4, a geometria e o domínio computacional do PSW usado como base para o afinilamento segmentado apresenta Λ como o período dos segmentos, a é o comprimento do segmento com maior índice de refração (n_1) e w é a largura do segmento e dos guias de onda contínuo. Foi então considerado um guia de onda óptico segmentado composto de silício e sílica, com índices de refração alternados: $\bar{n}_A =$

3,476 e $n_2 = 1,444$, respectivamente, cuja condições de contorno foram descritas no subitem 1.6.

Na próxima seção, o RNA-MLP (LM e BR) foi elaborado, e suas aplicações são apresentadas. Posteriormente, descreve-se a estratégia de modelagem usando ML para prever a eficiência de acoplamento entre os guias de onda PSW e CWG, com os atributos de entrada escolhidos. Por fim, os resultados numéricos obtidos pelas redes neurais MLP são mostrados em comparação com os valores encontrados analiticamente pelo FEM, juntamente com as principais conclusões.

4.2.1 Diagrama da RNA:

A RNA MLP é composta por camadas de neurônios: uma camada de entrada, três camadas ocultas e uma camada de saída. A estrutura apresentada foi definida com base na quantidade de entradas e saída desejadas e as camadas ocultas definidas de acordo com os resultados obtidos a partir da abstração inicial, chegando a uma arquitetura suficientemente simples e com poucos neurônios capazes de processar o conjunto de dados obtidos de acordo com a sua natureza não-linear e complexa.

1. Camada de entrada (15 neurônios): Isso corresponde às características de entrada do seu modelo. Cada neurônio na camada de entrada representa uma entrada única do seu conjunto de dados e na configuração do RNA as 15 entradas são os valores absolutos do comprimento em *nm* de cada um dos 15 segmentos componentes do *taper*.

2. Primeira camada oculta (18 neurônios): Cada um dos 15 neurônios da camada de entrada se conecta a todos os 18 neurônios desta camada. Portanto, haverá $15 \times 18 = 270$ conexões vindas da camada de entrada para a primeira camada oculta.

3. Segunda camada oculta (27 neurônios): Cada neurônio na primeira camada oculta se conecta a todos os 27 neurônios na segunda camada oculta. Isso cria $18 \times 27 = 486$ conexões.

4. Terceira camada oculta (28 neurônios): Semelhantemente, cada um dos 27 neurônios da segunda camada oculta se conectará a todos os 28 neurônios da terceira camada oculta, resultando em um total de $27 \times 28 = 756$ conexões.

5. Camada de saída (1 neurônio): Todos os 28 neurônios da terceira camada oculta então se conectam ao único neurônio na camada de saída, que fornece a saída do modelo, totalizando 28 conexões.

Ao todo, 89 nós e 1540 conexões representadas a seguir:

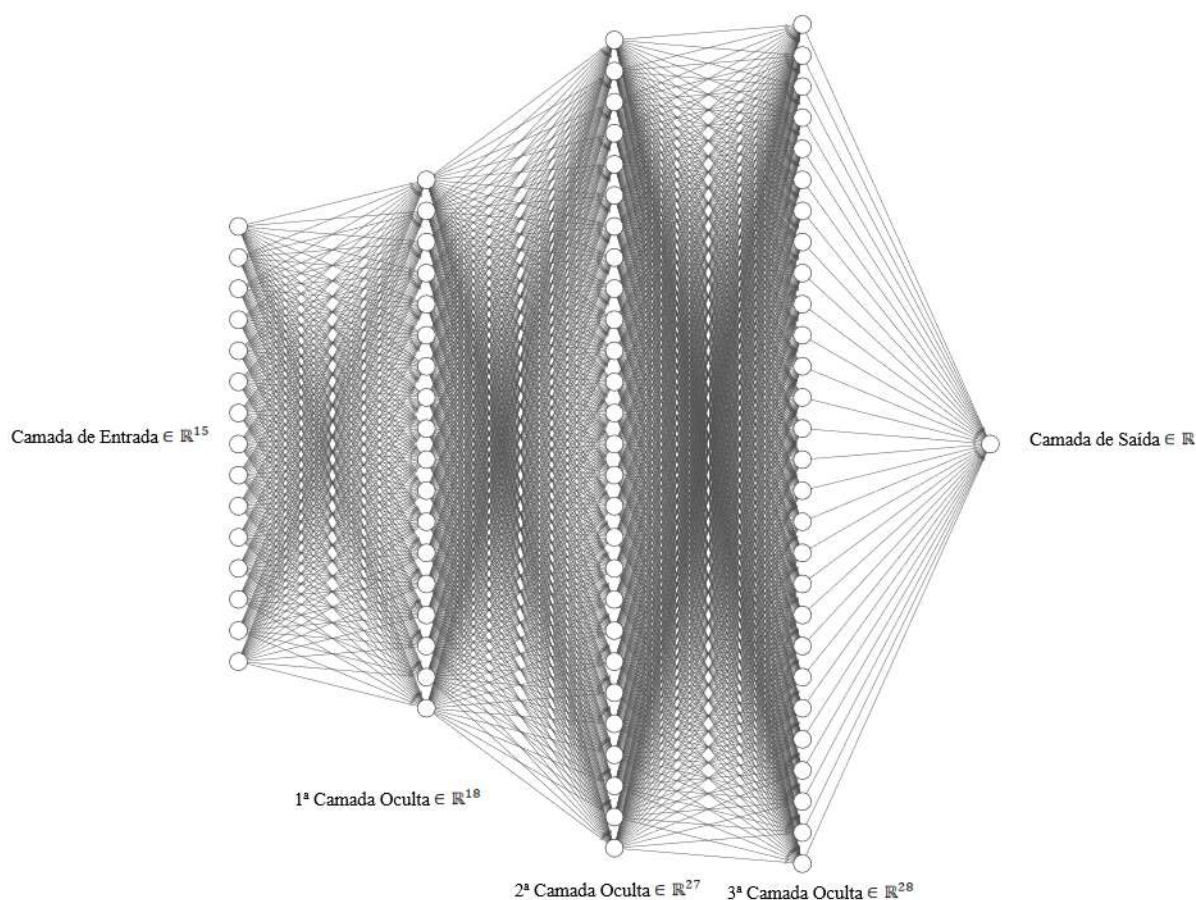


Figura 26- RNA - MLP - DIAGRAMA DE TRABALHO. Fonte: Autor.

4.2.2 Softwares Elaborados/Códigos Utilizados

Ao longo deste trabalho alguns códigos já foram elencados e este subitem trata especificamente sobre a RNA desenvolvida.

As etapas a seguir ilustram como foram elaboradas as sequências de comandos para desenvolvimento do código final *MLP4C.m* descrito no APENDICE D:

O Matlab® sendo um ambiente de programação que dentre outras aplicações, possui uma caixa de ferramentas de redes neurais (*Neural Network Toolbox*), que permite criar, treinar, testar e simular redes neurais.

É possível pode usar linhas de comando, m-files ou uma interface gráfica (*nntool*) para trabalhar com as redes neurais no Matlab®. O código da rede neural artificial (RNA) foi desenvolvido através dos seguintes passos:

- Foram definidos os padrões de entrada e saída, que são os dados que são

usados para treinar e testar a rede, armazenados em matrizes, vetores ou arquivos de texto.

- A definição da arquitetura da rede, especificando o número de camadas, o número de neurônios, as funções de ativação, o algoritmo de treinamento e outros parâmetros, foi executada com a função *feedforwardnet* criando uma rede *feedforward*,
- A seguir, definiu-se os parâmetros de treinamento, que são os valores que controlam o processo de aprendizado da rede, como o número de épocas, o erro final desejado e falhas na validação, com a estrutura *net.trainParam*.
- O treinamento da rede foi feito com função *train*.
- Para testar a rede a função *sim* foi usada para simular a rede com os dados de teste e comparar os resultados com os valores esperados.

4.3 Consolidação dos resultados

Para o *taper* designado no item 1.4.4, o modelo de Rede Neural foi submetido a uma série de configurações as quais podemos separar em duas fases. Na primeira fase (explanada no item 3.4.3) houve variação entre algoritmos de treinamento, na fase II, a variância é acerca das funções de ativação por camada conforme indicado na **Erro! Fonte de referência não encontrada..**

4.3.1 Fase I – Variação de Algoritmos de Treinamento

Para prever a resposta da eficiência de acoplamento do *taper*, a RNA multicamadas é inicialmente construída com os dados de entrada referentes a possíveis valores de comprimento de cada segmento dentro da faixa descrita. Foram obtidos modelos de rede neural para o *taper* mostrado na figura 7, que possui uma solução analítica utilizada para gerar os dados de treinamento. A expressão que permite a computação do desempenho do acoplamento é dada por $\eta = P_{out} / P_{in}$ (OLIVEIRA DAS MERCÊS; DOURADO SISNANDO; RODRIGUEZ ESQUERRE, 2021; SISNANDO et al., 2020). Os detalhes da arquitetura do sistema de alto nível e da estrutura de componentes descritos na subseção 4.1 apresenta a arquitetura de RNA utilizada nesta fase indicada na Figura 25.

O modelo geral da RNA-MLP foi submetido a diversos algoritmos de treinamento baseados na retropropagação, porém, entre as métricas utilizadas, 7ps

algoritmos LM e BR mostraram o melhor desempenho entre outros, aproveitando a precisão de desempenho e o tempo de retorno computacional observado.

4.3.2 Fase I – Resultados das Simulações

Os dados construtivos do dispositivo foram definidos no item 1.4.4 e as condições de contorno designadas na Tabela 2 e os dados de treinamento utilizados para cada variável contida no intervalo: a $[0,0300\mu\text{m}, 0,27\mu\text{m}]$; Inicialmente, para encontrar a eficiência para o conjunto de entradas, foi criado um conjunto de 282.000 amostras, 70% foram utilizadas como treinamento e 30% amostras foram utilizadas para validação e teste da rede. A rede MLP foi configurada para executar 2000 iterações, com as funções de ativação *tan-sigmoid*, *log-sigmoid* e lineares e os algoritmos de treinamento foram o *Levenberg – Marquardt* e a Regularização Bayesiana.

Os erros médios quadráticos apresentados pelas redes MLP-LM e MLP-BR, respectivamente. Como observado, o MLP-LM apresenta, nas mesmas condições, resultados muito mais rápidos que o MLP-BR, que permite emular uma forma avançada de métodos analíticos para recuperar os mesmos resultados.

A comparação entre os resultados obtidos pelas redes neurais e as soluções analíticas são apresentadas a seguir:

Nos ciclos de treinamento com *Lavenberg-Macquardt*, a melhor performance é demonstrada na Figura 28 com o índice de regressão igual 1 demonstrada na Figura 27

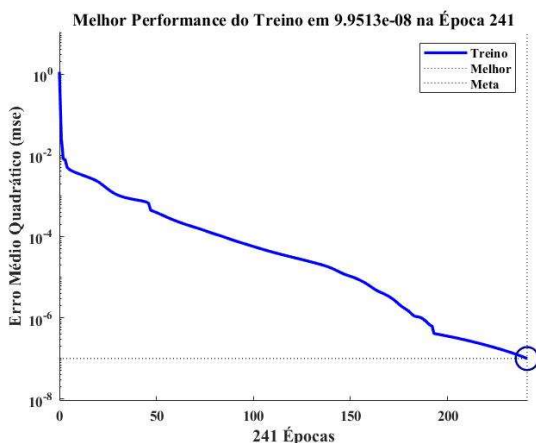


Figura 28-RNA-MLP -LM - PERFORMANCE

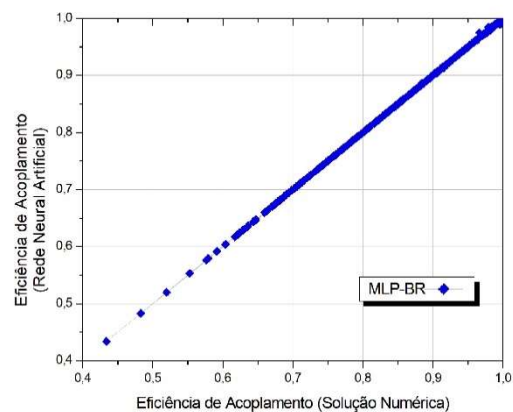


Figura 27 - RNA-MLP -LM - REGRESSÃO

Já nos ciclos de treinamento com Regularização Bayesiana, a melhor performance é demonstrada na Figura 29 com índice de regressão igual a 1 demonstrada na Figura 30.

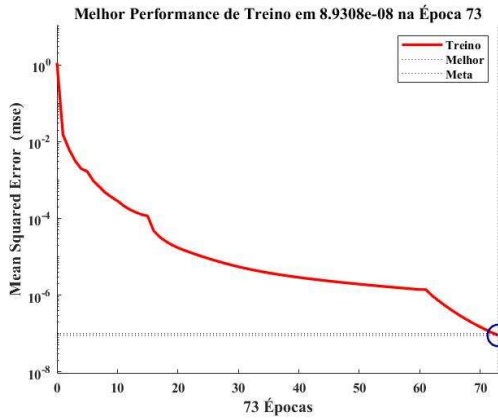


Figura 29 - RNA-MLP -BR - PERFORMANCE

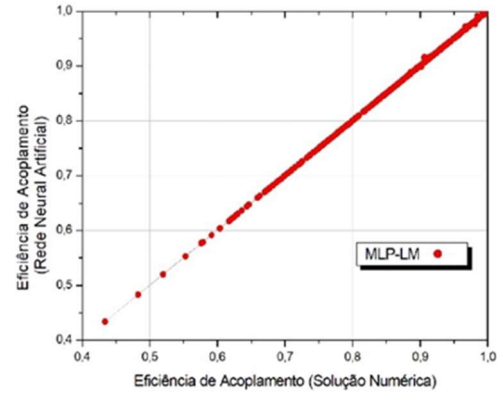


Figura 30 - RNA-MLP -BR - REGRESSÃO

Todas as simulações foram realizadas usando um computador com processador núcleo i7-9750H, relógio de 2,60 GHz e RAM DDR3 de 16 GB, 1,333 MHz, sistema operacional Windows 10 Home Single. As redes foram implementadas na Matlab®. A tabela 1 mostra um resumo da resposta computacional que mostra os resultados desta comparação, sob mesmas condições.

Algoritmo	Época	MSE	TIME (hh:mm:ss)	Regressão
Levenberg-Marquardt	73	$8,9 \times 10^{-7}$	00:24:30	1
Regularização bayesiana	471	$9,75 \times 10^{-8}$	02:21:11	1

Tabela 6- Comparação Única LM X BR treinamento

4.3.3 Fase II – Variação de Funções de Ativação por camada.

Nesta fase do trabalho optou-se por efetuar simulações categorizadas por análise combinatória com um arranjo com repetição (A_r) de forma que Z é o conjunto do número de camadas da RNA (r) exceto a camada de saída que se mantém inalterada, conforme explanado no item 3.4.3, e que assumem os possíveis valores distintos (n) de funções de ativação : *purelin*, *logsig* e *tansig*. :

$$Z = \{L1, L2, L3, L4\}$$

$$n = 3$$

$$r = 4$$

Obtendo-se :

$$A_r = (n, r) = n^r \text{ que resulta : } A(3,4) = 3^4 = 81$$

As 81 possibilidades estão relacionadas no APÊNDICE C:.

Em relação à métrica de comparação do erro em relação aos valores alvos, o gráfico da **Erro! Fonte de referência não encontrada.** a seguir denota que, à exceção do arranjo puramente linear, a média quadrática do erro é melhor em 98,76% das simulações efetuadas com valores menores que 10^{-7} .

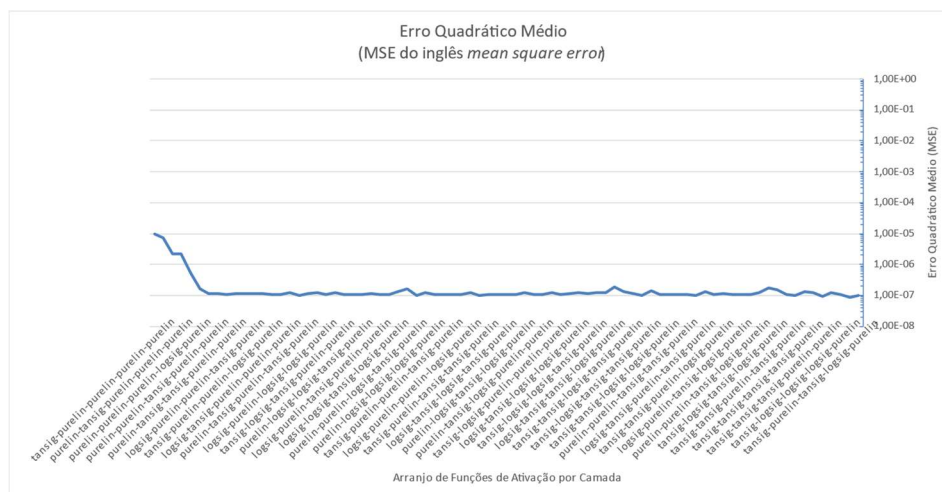


Figura 31- Arranjo de Funções de Ativação vs. MSE

Quanto à performance desempenhada em função da condição de parada referente aos ciclos de treinamento ou época, 92,59% alcançaram a Regressão de 100% sem atingir o número máximo de épocas estipuladas, 90,12% atingiram $R=1$ em menos de 50% do limite do critério de parada definidos para épocas, e 81,48% performaram em menos de 500 épocas, como observado no gráfico da Figura 32 a seguir:

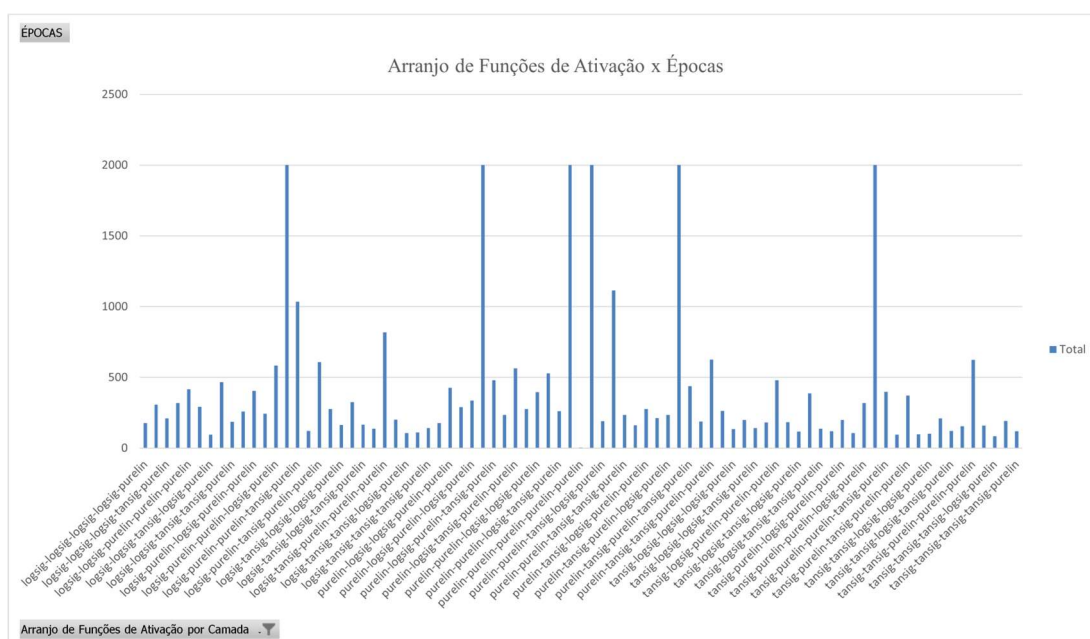


Figura 32 - Arranjo de Funções de Ativação vs Números de Épocas

Ainda sobre a capacidade de generalização do modelo trabalhado, a regressão de 100% foi atingida em 98,76%, e representando graficamente na Figura 33, verifica-se a sutil diferença apresentada no modelo onde a função de ativação puramente linear foi aplicada a regressão é observada em 75,9%.

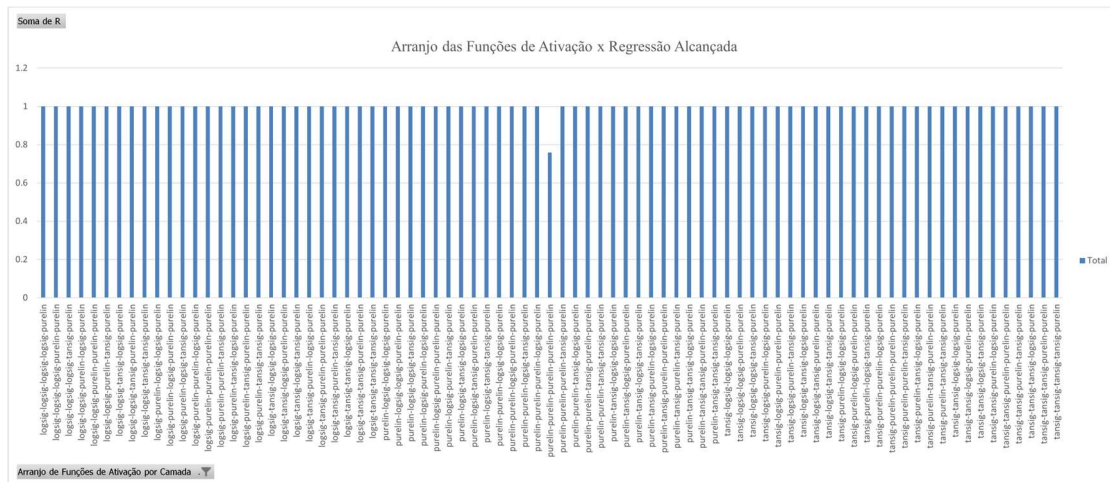


Figura 33 - Arranjo de Funções de Ativação vs Regressão.

O tempo de produção dos resultados para cada modelo ajustado pode ser visualizado na Figura 34 a seguir. O tempo de processamento manteve-se dentro do limiar encontrado na FASE I em 56,79% , e 1,24% não performa dentro dos critérios de parada designados, como demonstrado.

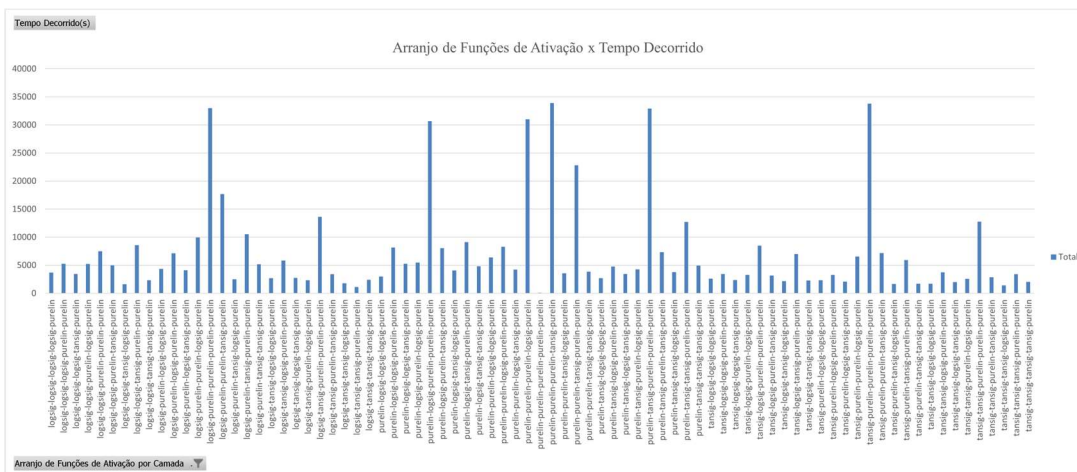


Figura 34 - Arranjo de Funções de Ativação vs. Tempo Decorrido

Dentre este rol de possibilidades, foram escolhidos para considerações sobre os resultados alcançados os 5 arranjos melhores classificados quanto ao tempo decorrido, número de épocas e regressão abaixo de 100% para demonstrar que o modelo possibilitou o uso de uma grande variação de funções de ativação atingindo a mesma qualidade e a escolha da configuração mais rápida que justifica a substituição dos cálculos numéricos

diretamente pelo modelo escolhido além da integração em softwares de simulação com menor custo operacional.

Para todos os processos o conjunto de treinamento, validação e teste, foram definidos em 2 conjuntos com tamanhos de 70% e 30% para o treinamento e validação, respectivamente, outro conjunto separado para ser usado como um conjunto de testes.

Dentre os principais resultados destacaram-se os que obtiveram melhor desempenho na busca pelo índice de regressão de 100%.

4.3.4 Fase II – Resultado das Simulações

Com os dados construtivos do dispositivo foram definidos no item 1.4.4 e as condições de contorno designadas na tabela 3 e os dados de treinamento utilizados para cada variável contida no intervalo: a $[0,0300\mu\text{m}, 0,27\mu\text{m}]$.

Para encontrar a eficiência para o conjunto de insumos, foi criado um conjunto de 282.000 amostras, 70% foram utilizadas como treinamento e 30% amostras foram utilizadas para validação e teste da rede. Outro conjunto de dados foram produzidos para avaliar a capacidade de generalização do modelo.

Dos destaques anteriormente mencionados, consolidamos nas tabelas 8,9,10 e 11 a seguir, os arranjos, configurações e parâmetros sob vieses de tempo de resposta, ciclos de iteração e melhor resultado de erro quadrático médio.

a. Quanto ao Tempo decorrido:

Os principais arranjos da RNA com o menor tempo de treinamento e simulação decorridos retrata a robustez do modelo de 4 camadas com uma alta capacidade de permissibilidade em combinações, mesmo com uma abstração inicial que ignora uma relação ordinária, após algumas substituições a agilidade do modelo seria verificada. O gráfico da Figura 35 mostra a relação entre os arranjos destacados e o tempo de execução.

Neurônios	Arranjo	15	18	27	28	1	Tempo Decorrido
Camadas		Entrada	Oculto 1	Oculto 2	Oculto 3	Saída	
Funções de Ativação	1	<i>tansig</i>	<i>purelin</i>	<i>tansig</i>	<i>logsig</i>	<i>purelin</i>	0h14min9,6s
	2	<i>tansig</i>	<i>tansig</i>	<i>tansig</i>	<i>logsig</i>	<i>purelin</i>	0h15min0s
	3	<i>logsig</i>	<i>logsig</i>	<i>tansig</i>	<i>logsig</i>	<i>purelin</i>	0h15min31,8s
	4	<i>tansig</i>	<i>tansig</i>	<i>logsig</i>	<i>logsig</i>	<i>purelin</i>	0h26min39,6s
	5	<i>tansig</i>	<i>purelin</i>	<i>tansig</i>	<i>tansig</i>	<i>purelin</i>	0h31min28s

Tabela 7 - Seleção de Arranjos quanto ao tempo decorrido.



Figura 35- Tempo de execução para arranjos da tabela 8

No arranjo 1 da Tabela 7, as camadas de entrada[15] e ocultas [18-27-28] são ativadas através das funções de ativação indicadas e apresenta uma regressão de 100% conforme Figura 36 a seguir, tanto para treino, validação e testes. A melhor performance é obtida na época indicada na Figura 37 e sem falhas de verificação durante a validação, conforme Figura 38.

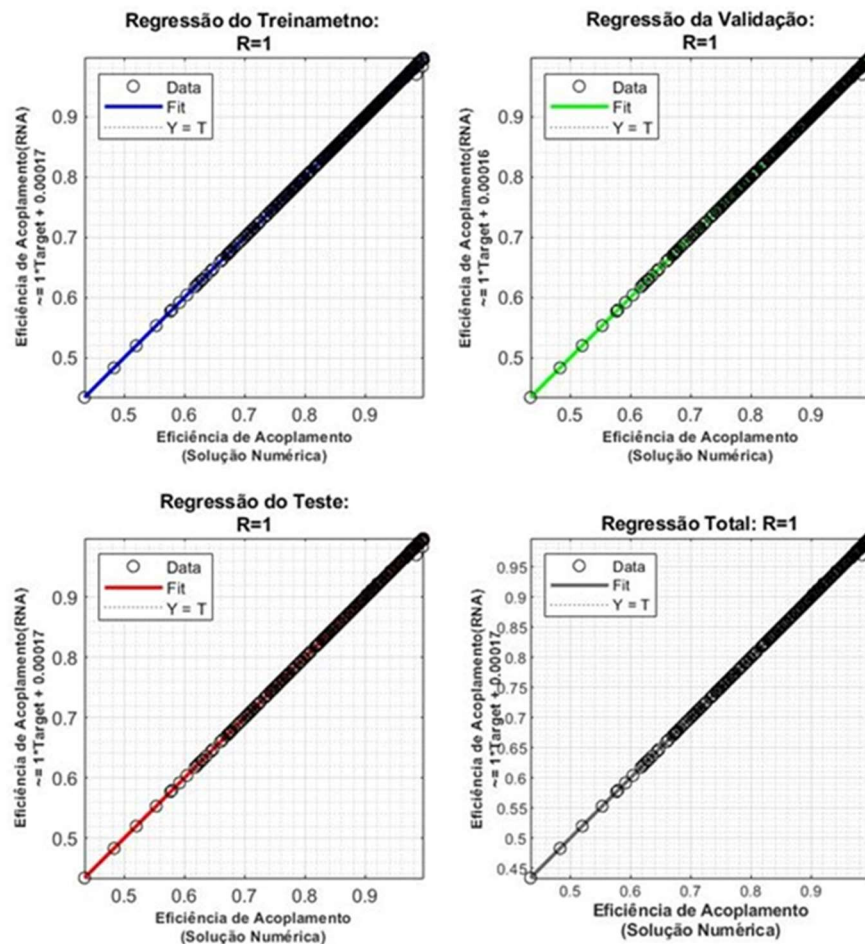


Figura 36-Resultados do Treinamento, Validação e Testes - Arranjo 1 da tabela 8.

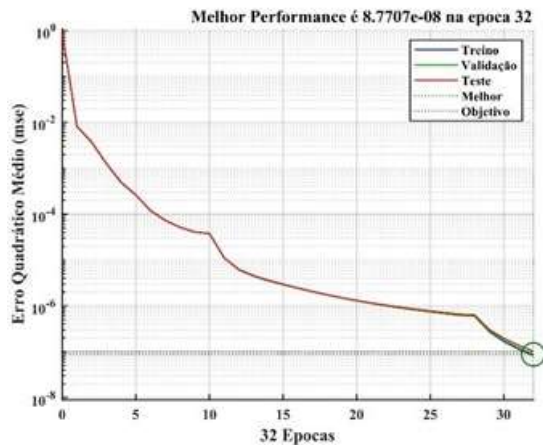


Figura 37 - Performance - Arranjo 1 da Tabela 8

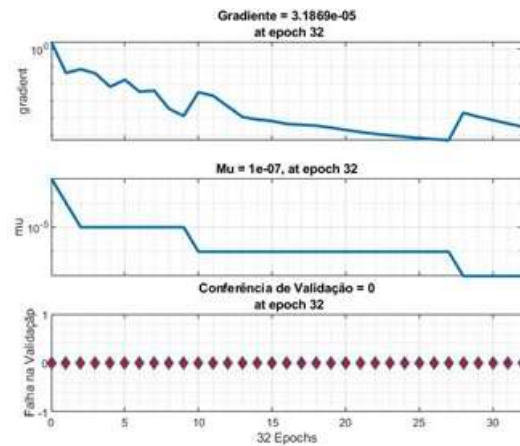


Figura 38- Falha na Validação do Arranjo 1 da tabela 8.

No arranjo 2 da Tabela 7, as camadas de entrada e as ocultas [18-27-28] são ativadas através das funções de ativação indicadas e apresenta uma regressão de 100% conforme Figura 39 a seguir, tanto para treino, validação e testes. A melhor performance é obtida na época indicada na Figura 41 e sem falhas na validação, conforme Figura 40

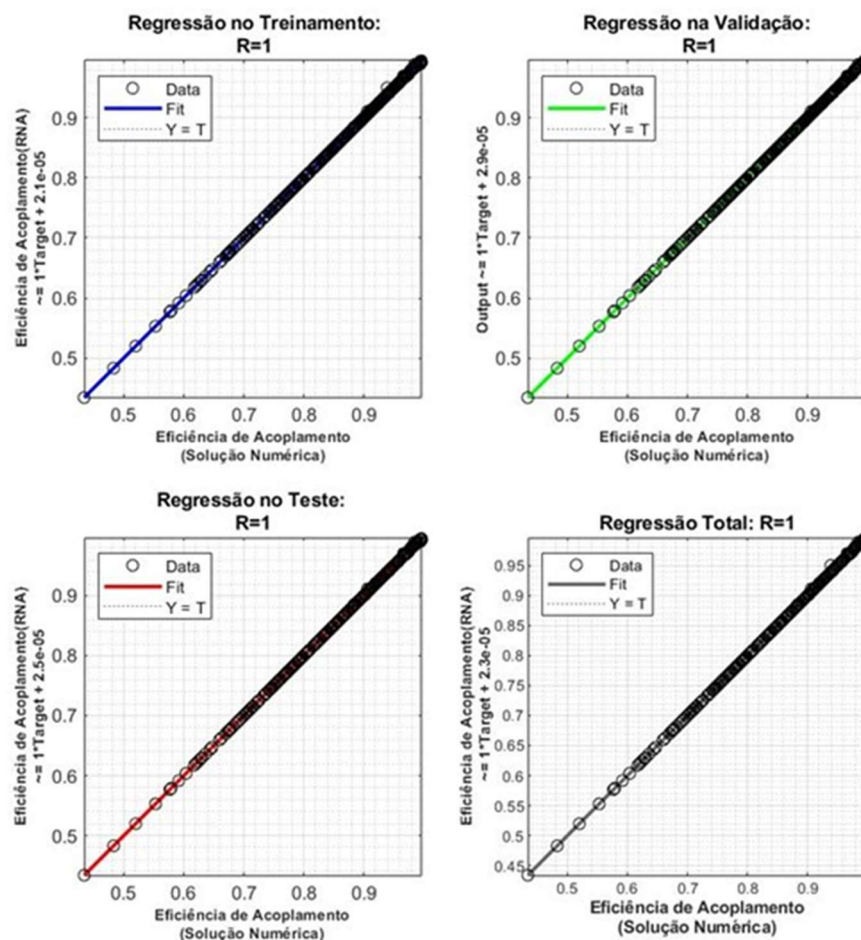


Figura 39 - Resultados do Treinamento, Validação e Testes - Arranjo 2 da tabela 8.

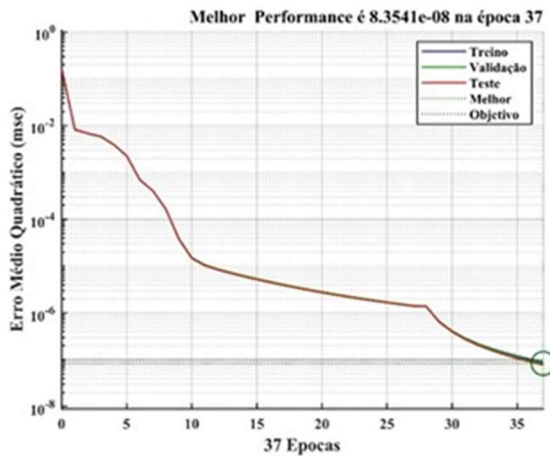


Figura 41 - Performance do Arranjo 2 da Tabela 8.

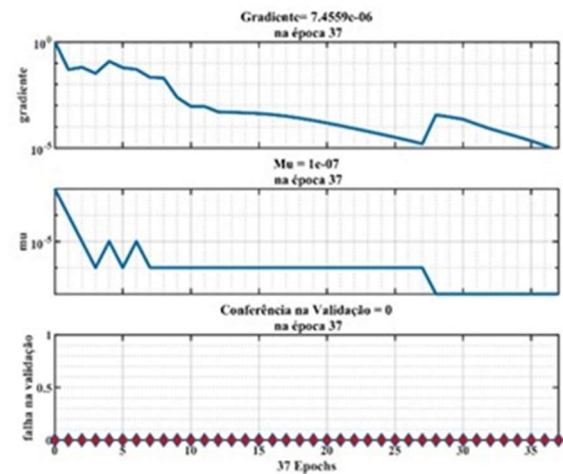


Figura 40- Falha na Validação do Arranjo 2 da Tabela 8

Já no arranjo 3 da Tabela 7, as camadas de entrada e as ocultas são ativadas através das funções de ativação indicadas e apresenta uma regressão de 100% no treinamento e no teste, vide Figura 42, porém na validação foi observada uma intercorrência na época 50, dentro do limiar estabelecido na condição de parada no conforme a seguir, tanto para treino, validação e testes. Haverá uma discrepância entre a melhor performance do treino contra validação obtida na época indicada na Figura 43 e a falha de validação indicada na Figura 44.

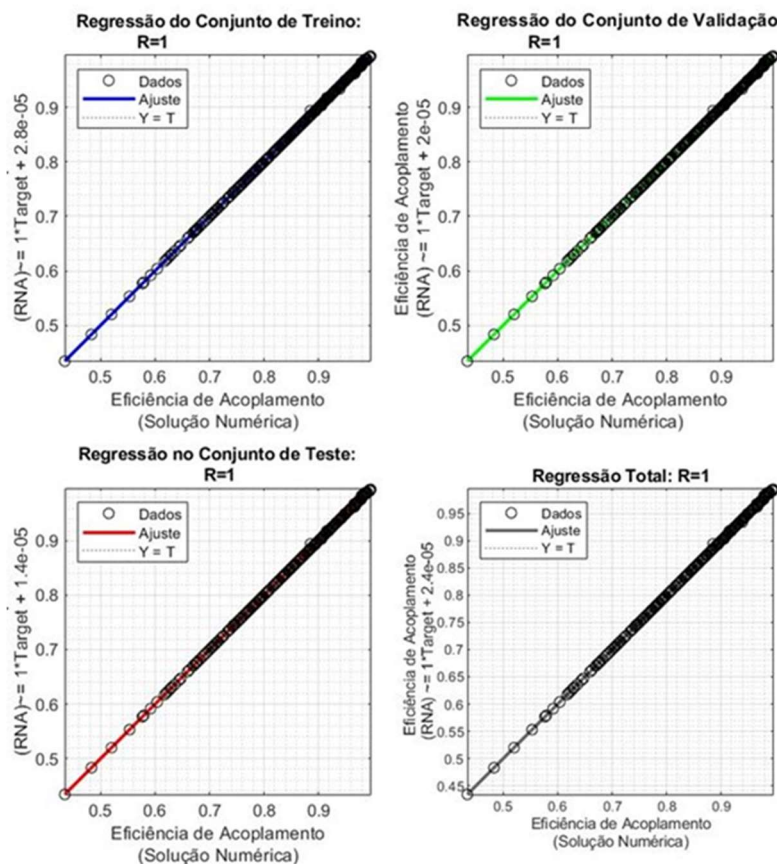


Figura 42-Resultados do Treinamento, Validação e Testes - Arranjo 3 da tabela 8.

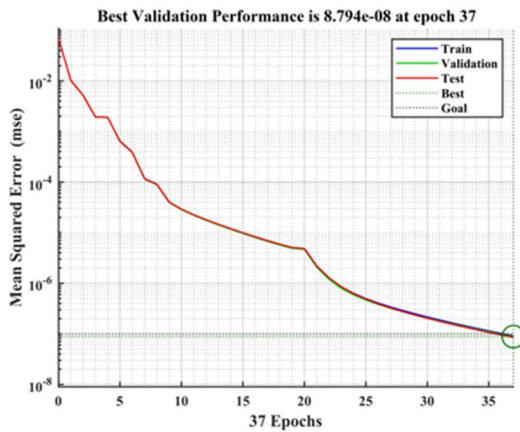


Figura 43- Performance do Arranjo 3 da Tabela 8

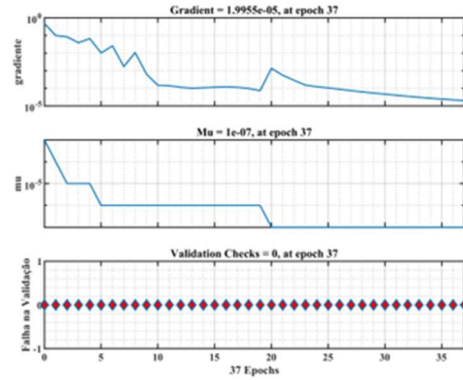


Figura 44-Falha na Validação do Arranjo 3 da tabela 8.

No arranjo 4 da Tabela 7, as camadas de entrada e as intermediárias são ativadas através das funções de ativação indicadas e apresenta uma regressão de 100% conforme Figura 45 a seguir, tanto para treino, validação e testes. A melhor performance é obtida na época indicada na **Erro! Fonte de referência não encontrada.** e sem falhas de verificação durante a validação, conforme Figura 47.

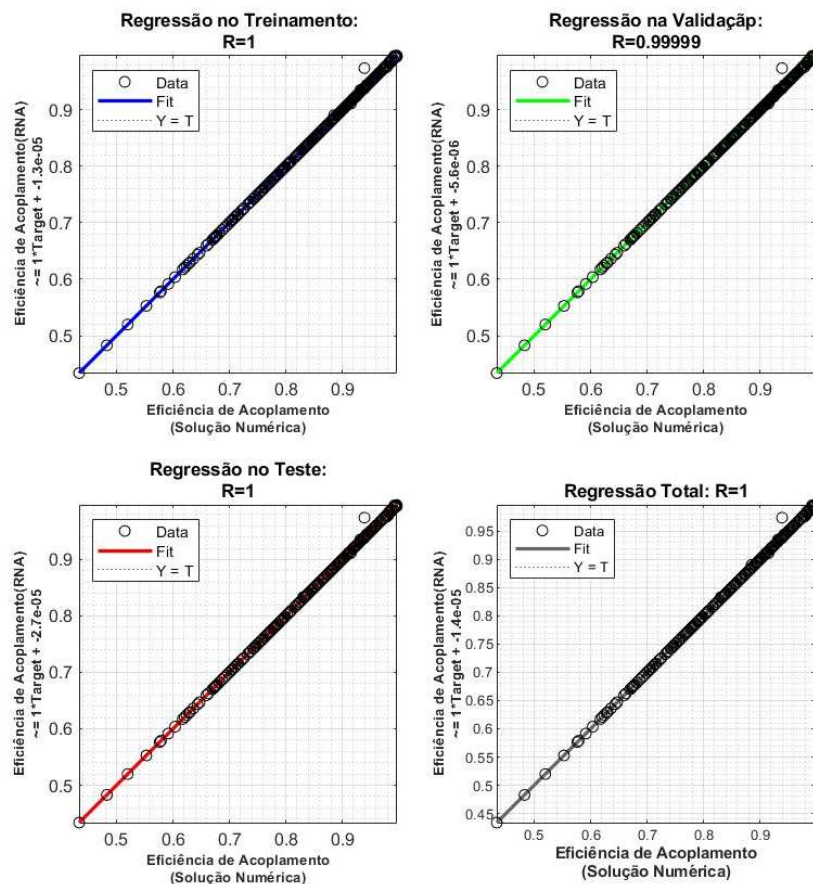


Figura 45-Resultados do Treinamento, Validação e Testes - Arranjo 4 da tabela 8.

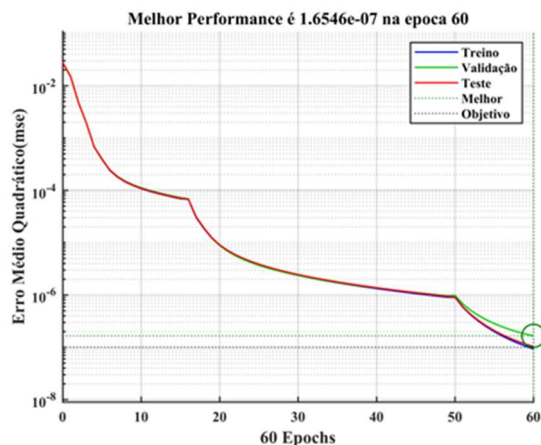


Figura 46 -Performance do Arranjo 4 da tabela 8.

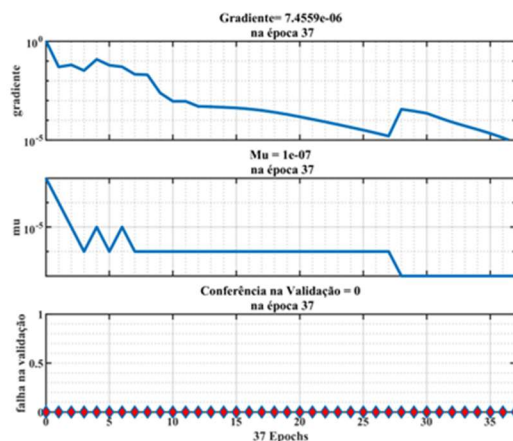


Figura 47-Falha na Validação do Arranjo 4 da tabela 8.

Para o arranjo 5 da Tabela 7, as camadas de entrada e as ocultas são ativadas através das funções de ativação indicadas e apresenta uma regressão de 100% conforme Figura 48 a seguir, tanto para treino, validação e testes. A melhor performance é obtida na época indicada na Figura 49 e sem falhas de verificação durante a validação, conforme Figura 50.

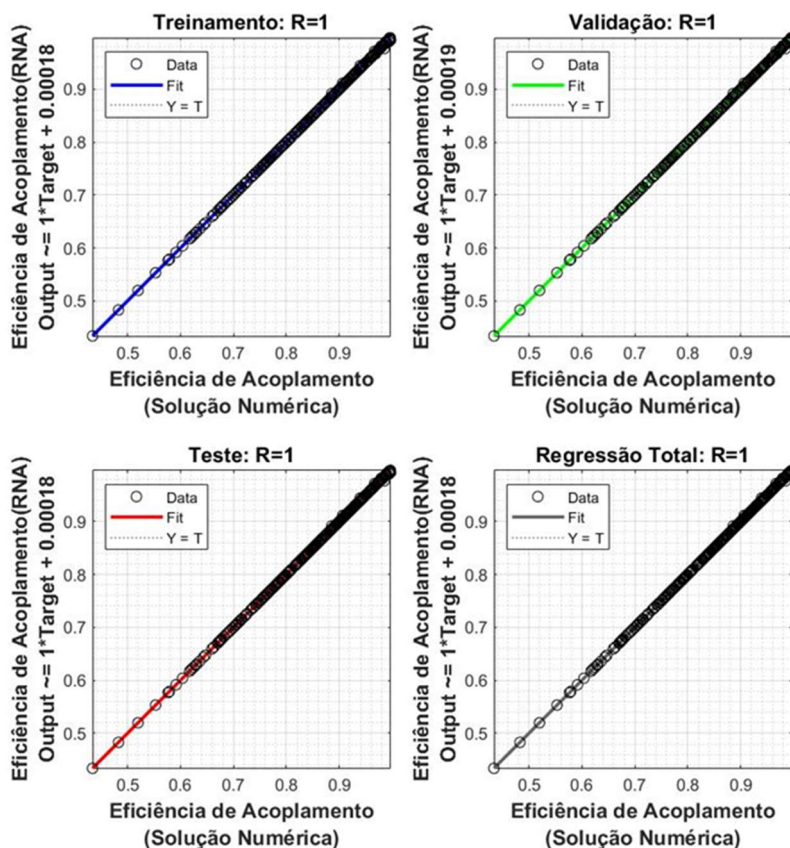


Figura 48- Resultados do Treinamento, Validação e Testes - Arranjo 5 da tabela 8.

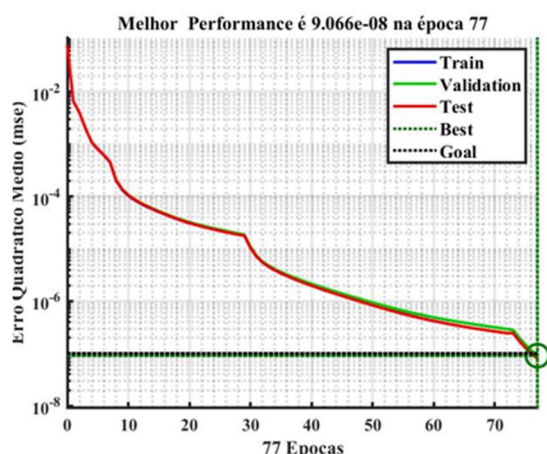


Figura 49 - Performance do Arranjo 5 da tabela 8

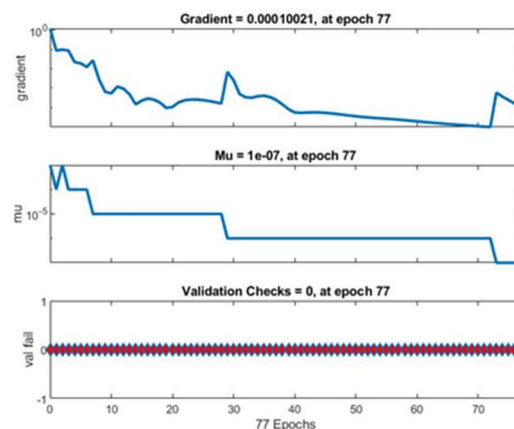


Figura 50 - Performance do Arranjo 5 da tabela 8

b. Quanto às Épocas alcançadas:

As iterações cíclicas são dependentes direta e condicionalmente do recurso computacional além das características intrínsecas a uma rede neural artificial tais como número de neurônios, camadas ocultas e tipo de aprendizados.

Neste trabalho pôde-se relacionar o tempo com as épocas e em alguns casos o menor tempo de processamento decorrido equivale ao menor número de épocas.

Observa-se através da Tabela 8 a equivalência com a Tabela 7, em que os arranjos com menor tempo decorrido são também os arranjos com menos épocas dispensadas.

Na Figura 51, estão dispostos os arranjos cujo resultados foram os melhores em relação ao critério de parada que foi atingir o erro quadrático médio inferior a 10^{-7} .

Neurônios	Arranjo	15	18	27	28	1	Épocas	R
Camadas		Entrada	Oculto 1	Oculto 2	Oculto 3	Saída		
Funções de Ativação	1	<i>tansig</i>	<i>purelin</i>	<i>tansig</i>	<i>logsig</i>	<i>purelin</i>	32	1
	2	<i>logsig</i>	<i>logsig</i>	<i>tansig</i>	<i>logsig</i>	<i>purelin</i>	37	1
	3	<i>tansig</i>	<i>tansig</i>	<i>tansig</i>	<i>logsig</i>	<i>purelin</i>	37	1
	4	<i>tansig</i>	<i>tansig</i>	<i>logsig</i>	<i>logsig</i>	<i>purelin</i>	60	1
	5	<i>tansig</i>	<i>purelin</i>	<i>tansig</i>	<i>tansig</i>	<i>purelin</i>	77	1

Tabela 8 - Seleção de Arranjos quanto ao número de épocas.

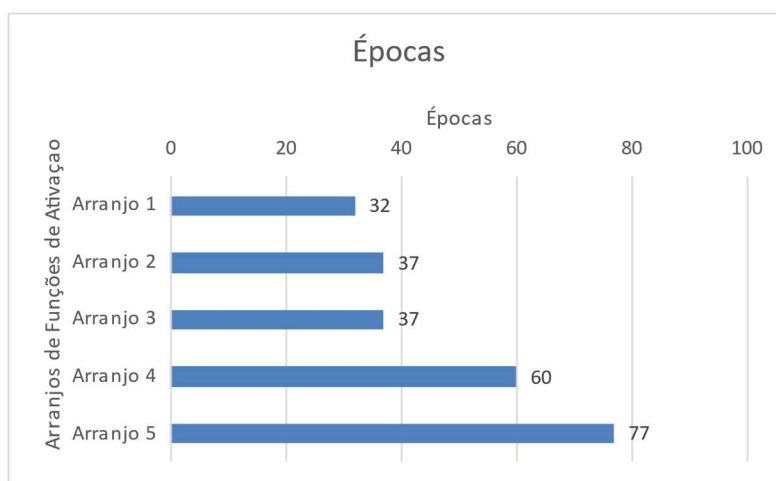


Figura 51 - Os 5 arranjos com menos épocas a atingir ao critério de parada.

c. Quanto ao menor erro médio quadrático.

Outra forma de verificar a robustez do modelo é verificar o quão ele se separa dos valores reais e o quanto ele consegue restringir a sua previsibilidade de erro para que este não impacte na predição final.

Uma das métricas, o *mse*, foi menor nos 5 arranjos a seguir, o arranjo de número 5, conforme Figura 52.

Neurônios	Arranjo	15	18	27	28	1	MSE	R	tempo decorrido
Camadas		Entrada	Oculto 1	Oculto 2	Oculto 3	Saída			
Funções de Ativação	1	<i>tansig</i>	<i>logsig</i>	<i>logsig</i>	<i>logsig</i>	<i>purelin</i>	8,33E-08	1	0h43min49,2s
	2	<i>tansig</i>	<i>tansig</i>	<i>tansig</i>	<i>logsig</i>	<i>purelin</i>	8,35E-08	1	0h15min0s
	3	<i>logsig</i>	<i>logsig</i>	<i>tansig</i>	<i>logsig</i>	<i>purelin</i>	8,49E-08	1	0h15min31,8s
	4	<i>tansig</i>	<i>purelin</i>	<i>tansig</i>	<i>logsig</i>	<i>purelin</i>	8,77E-08	1	0h14min9,6s
	5	<i>logsig</i>	<i>tansig</i>	<i>tansig</i>	<i>purelin</i>	<i>purelin</i>	9,82E-08	1	0h57min36s

Tabela 9 - Seleção de Arranjo quanto ao menor MSE calculado.

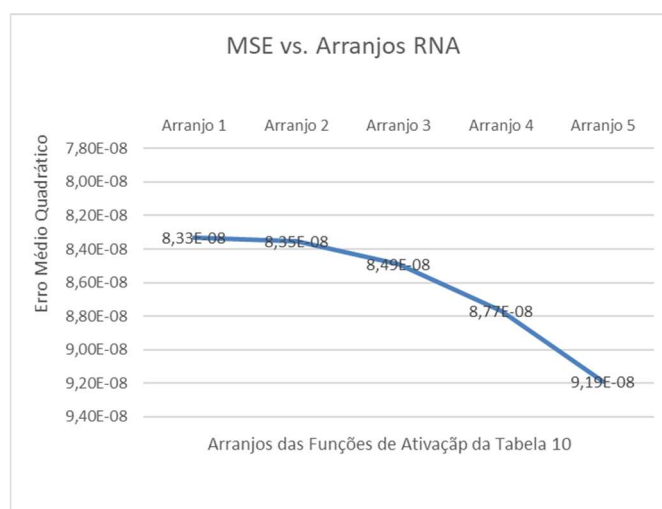


Figura 52- Os 5 melhores resultados dos Arranjos quanto à MSE

Também observa-se as equivalências dos arranjos entre a Tabela 7 e a Tabela 8 e a Tabela 9. Porém os arranjos 1 e 5 da Tabela 9 não aparecem nas tabelas anteriores os quais se distanciam dos demais pelo tempo de execução decorrido. Estes 2 arranjos apresentaram os resultados satisfatórios a seguir.

No arranjo 1 da Tabela 9, as camadas de entrada e as ocultas [15-18-27-28] são ativadas através das funções de ativação indicadas e apresenta uma regressão de 100% conforme a Figura 53

seguir, tanto para treino, validação e testes. A melhor performance é obtida na época indicada na Figura 54 e sem falhas de verificação durante a validação, conforme Figura 55.

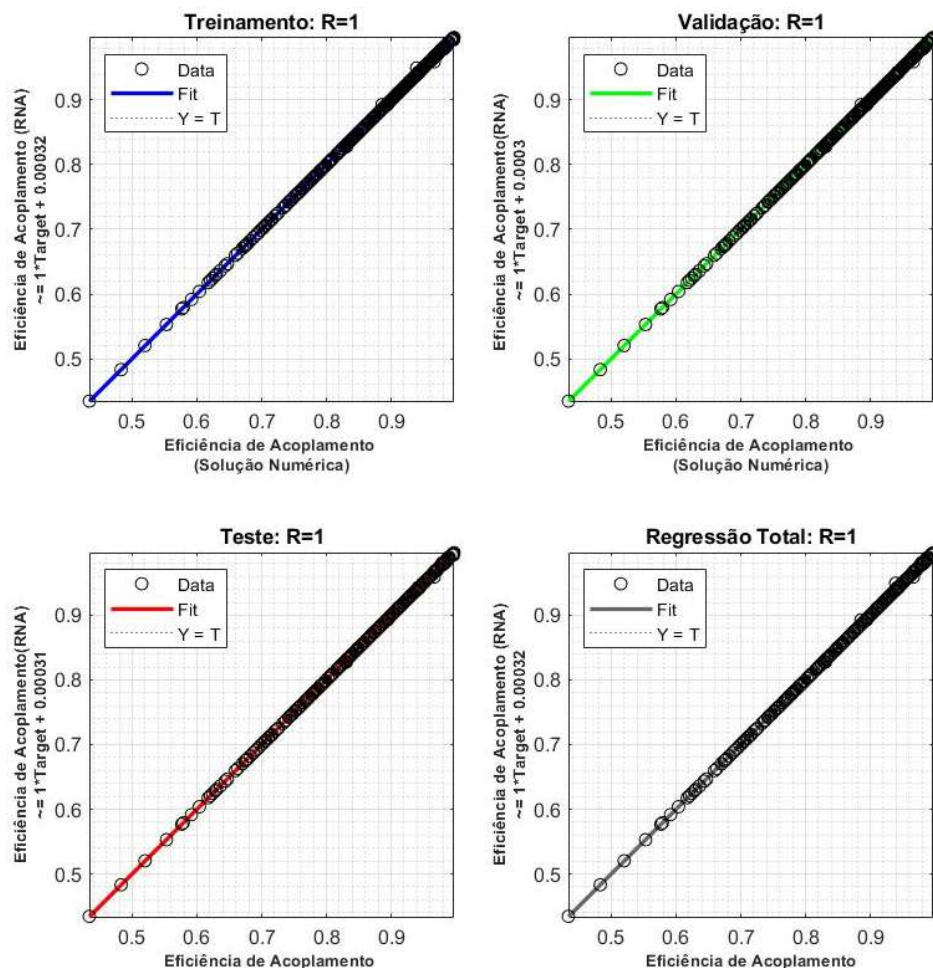


Figura 53 - Resultados do Treinamento, Validação e Testes - Arranjo 1 da tabela 10.

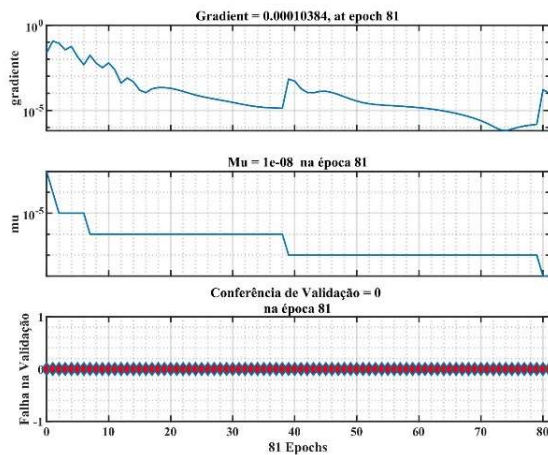


Figura 54- Performance do Arranjo 1 da tabela 10

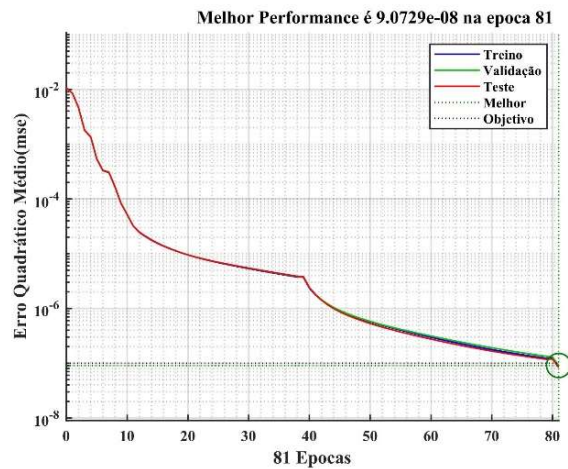


Figura 55- Falha na Validação do Arranjo 1 da tabela 10

Por último, arranjo 5 da Tabela 9, as camadas de entrada e as intermediárias são ativadas através das combinação de 3 funções indicadas e apresenta regressão de 100% tanto no treinamento quanto na validação, mas foi garantida a menor taxa de erro entre todas as arquiteturas conforme Figura 56 a seguir. A melhor performance é obtida na época indicada na Figura 58 e sem falhas de verificação durante a validação, conforme Figura 57.

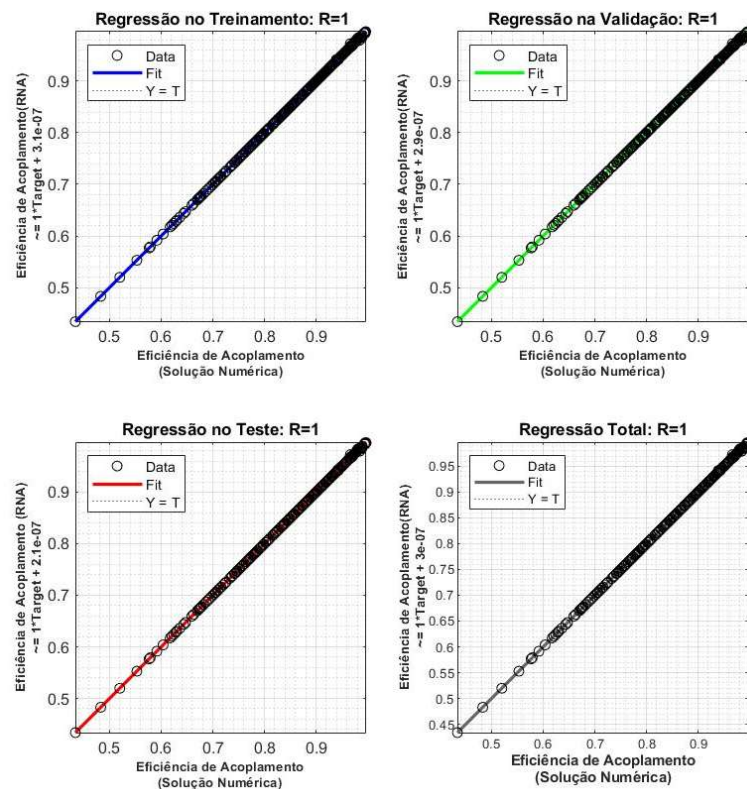


Figura 56-Resultados do Treinamento, Validação e Testes - Arranjo 5 da tabela 10.

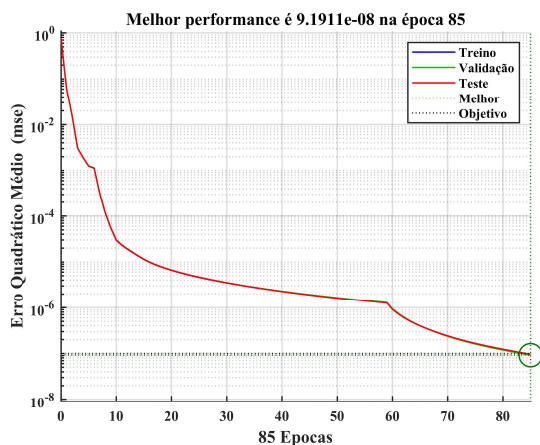


Figura 58 – Performance do Arranjo 5 da tabela 10

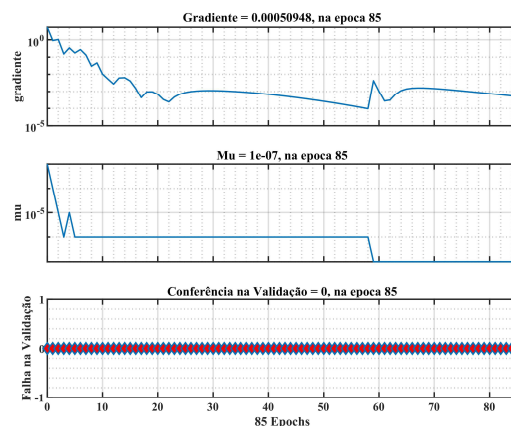


Figura 57- Falha na Validação do Arranjo 5 da Tabela 10

d. Quanto Regressão abaixo de 100%:

Na seleção dos arranjos na Tabela 10 observa-se os critérios de parada foram acionados antes de que a rede convergisse para um modelo cuja regressão atingisse o desejado, porém a tendência faria atender caso as condições de parada fossem readequadas.

Neurônios	15	18	27	28	1	MSE	Época	R	TEMPO TOTAL
Camadas	Entrada	Oculto 1	Oculto 2	Oculto 3	Saída				
Funções de Ativação	<i>purelin</i>	<i>purelin</i>	<i>purelin</i>	<i>logsig</i>	<i>purelin</i>	1.62E-07	2000	0.99999	8h36.73min
	<i>purelin</i>	<i>purelin</i>	<i>purelin</i>	<i>tansig</i>	<i>purelin</i>	5.46E-07	2000	0.99998	9h24.92min
	<i>purelin</i>	<i>tansig</i>	<i>purelin</i>	<i>purelin</i>	<i>purelin</i>	2.13E-06	2000	0.99991	9h8.41min
	<i>tansig</i>	<i>purelin</i>	<i>purelin</i>	<i>purelin</i>	<i>purelin</i>	7.29E-06	2000	0.99968	9h22.95min
	<i>logsig</i>	<i>purelin</i>	<i>purelin</i>	<i>purelin</i>	<i>purelin</i>	9.38E-06	2000	0.99959	9h10.26min

Tabela 10 - Seleção de Arranjo quanto à Regressão abaixo de 100%

Como resultado da pesquisa, vários resultados significativos de teste foram obtidos e resumidos para os valores médios de erro MSE (Erro Quadrático Médio), referente aos modelos de RNA atualizados com as variações de funções de ativação, destacando se os arranjos quanto ao tempo decorrido, quanto às épocas alcançadas e quanto ao menor média de erro quadrático obtido.

4.4 Conclusão do capítulo

Algoritmos baseados em redes MLP-LM e MLP-BR foram implementados para modelar *tapers* segmentados subdivididos em 15 segmentos (variáveis de entrada) para obter a máxima eficiência de acoplamento. Foi feita uma comparação entre os resultados obtidos com as redes neurais MLP (LM e BR) e os dados numéricos da eficiência de acoplamento dos *tapers*, apresentando excelentes resultados com erros menores que 10^{-7} , com esforço computacional reduzido e muito mais simples do que o métodos numéricos comumente utilizados (como o FEM), mostrando conformidade com os resultados obtidos com outros Métodos.

As principais vantagens de um modelo baseado em RNA são simplicidade, redução de tempo e esforço computacional, e sua aplicação em problemas de síntese.

Pode-se concluir que para a análise do *taper* segmentado dividido em 15 partes, ambas as redes, MLP-LM e O MLP-BR pode ser aplicado como uma ferramenta eficiente. Os resultados obtidos pelo MLP-LM foram ligeiramente melhores do que aqueles obtido pela rede MLP-BR e com redução do esforço computacional.

Os dispositivos utilizados em acoplamento de guias de onda já possuem habilidades e conhecimentos avançados sobre manufatura e comercialização projetados através de softwares simuladores (TAYLOR; TANEV, 2007) que requisitam conhecimentos complexos para cálculos de matemática aplicada e conhecimento robusto em áreas específicas da Física, Engenharia e Computação.

Com o avanço das técnicas de processamento de informações baseadas em comportamentos da natureza e do aprofundamento do estudo neuromórfico, as redes neurais artificiais promovem um modo de desenvolvimento de soluções complexas com base no aprendizado de máquina que trata os processos computacionais e de conhecimento como parte integrante de um processador capaz de sintetizar todo o sistema num modelo com uma boa resposta aos estímulos apreciados. (MONROE, 2014).

Como atualmente, estudos trouxeram modelos de RNA que representam com sucesso dispositivos ópticos voltados para o guiamento da luz (RUBIO-MERCEDES et al., 2009), (MERCÊS et al. 2021), e análises das modificações estruturais para apresentar respostas desejadas também são demonstradas através da simplificação mimetizada por heurísticas sinápticas e aprimorada com novas técnicas que balizarão o comportamento dos neurónios ou dos filtros associados (kernels) no momento das escolhas de qual solução melhor corresponde ao modelo referenciado (CHUGH et al., 2019) e

(KUDYSHEV et al., 2020) , o tema estudado além de grande interesse, está em equidade temporal com pesquisas desenvolvidas com abrangência global, tem potencial para apresentar resultados originais para divulgação em revistas especializadas, podendo contribuir para o estudo e trabalhos colaborativos com outras entidades de pesquisa.

O capítulo versa sobre a utilização de redes neurais artificiais (RNA) para otimizar parâmetros de resposta única ou multi-resposta em nanoestruturas de dispositivos ópticos. O estudo utiliza modelos baseados em MLP (*Perceptron Multicamadas*) e algoritmos de treinamento como *Levenberg-Marquardt* e *Bayesian Regularization*.

O processo de treinamento da RNA consiste em dividir um grande bloco de dados em dois conjuntos, um para treinamento e validação e outro para teste. A RNA é submetida a sucessivos treinamentos até que os objetivos sejam atingidos e os testes sejam realizados para verificar a qualidade dos resultados e sua precisão.

Os resultados obtidos com a utilização da RNA em diferentes arranjos e configurações é demonstrado. Além disso, o capítulo destaca a importância da abstração inicial na construção da RNA e a necessidade de avaliar a capacidade de generalização do modelo.

CAPÍTULO 5 - CONCLUSÕES

5 CONSIDERAÇÕES FINAIS

Este trabalho de vanguarda desdobra-se em uma gama diversificada de possibilidades, incrementando os limites convencionais no campo dos dispositivos fotônicos. Essa abordagem simplificada em linguagem de alto nível permite que até mesmo indivíduos com conhecimento básico possam se envolver no processo de projeção, análise e desenvolvimento de dispositivos fotônicos, impulsionando assim a inovação e a criatividade em diversas esferas tecnológicas.

Por fim, apresento as contribuições deste trabalho e proponho algumas possibilidades para projetos futuros que possam estender ou complementar esta dissertação.

5.1 Contribuições

A pesquisa em redes neurais artificiais, fundamentada em princípios de teoria de aproximação de funções e análise numérica, leva a descobertas importantes de análise e síntese que são essenciais para a aplicação dessas estruturas conexionistas a uma variedade de problemas não-lineares, aqui destacado no campo da fotônica em dispositivos dielétricos.

A arquitetura de rede neural conhecida como *perceptron* de múltiplas camadas (MLP, do inglês *multilayer perceptron*), usada para treinamento supervisionado, foi examinada e aplicada neste trabalho de maneira aprofundada, resultando em:

- em estratégias voltadas ao treinamento das arquiteturas MLP- adequando as arquiteturas experimentalmente às modificações das funções de ativação mais comuns;
- no estudo dos fundamentos no projeto de algoritmos de treinamento eficientes para redes do tipo MLP;
- na compreensão dos méritos dos métodos computacionais convencionais aplicados em um guia dielétrico diante das possibilidades de simplificação permitida pelo

uso das redes neurais artificiais. ;

- na implementação de uma grande variedade de métodos de treinamento das RNAs na análise do custo computacional frente aos métodos numéricos convencionais
- no estudo da capacidade de generalização dos modelos de redes neurais resultantes;
- na comparação de desempenho de diferentes algoritmos de treinamento quando aplicados a vários problemas clássicos e de natureza distinta;
- no estudo de estratégias de determinação da dimensão das arquiteturas empregadas;

Este trabalho foi finalizado com estudos sobre as respostas obtidas diante das variantes das arquiteturas neurais e funções de ativações típicas. O método de distribuição combinatório das funções entre as camadas resultaram em modelos adequados capazes de generalização dos resultados independente do arranjo , e como previsto, redes com funções de ativações puramente lineares ou sequencialmente similares não se adequam ao aprendizado para a natureza dos dados trabalhados.

Os resultados deste estudo se destacam por sua fundamentação teórica e por sua conexão com outros campos de pesquisa e prática científica já consolidados. Isso é particularmente verdadeiro em relação aos aspectos ligados às ferramentas mais avançadas e flexíveis utilizadas nesses campos. Tais características são justificadas pela natureza interdisciplinar das redes neurais artificiais, pela sofisticação que define todas as contribuições relacionadas ao campo e por sua propensão inerente para lidar com problemas de alta complexidade.

5.2 Extensões

Há uma variedade de tópicos ligados ao tema desta pesquisa que podem ser explorados com base nos resultados apresentados nesta dissertação. Alguns deles incluem:

- Pesquisas sobre a habilidade de generalização das redes, utilizando a teoria de regularização e comparando com os resultados obtidos através de procedimentos de validação cruzada;
- Promover a inclusão de novos parâmetros como variáveis de entrada que alterem características notáveis dos dispositivos e conseqüentemente suas propriedades e

eficiência.

- Comparação de desempenho (incluindo a complexidade computacional do processo de busca) de diferentes métodos de aprendizado.
- Uso de estratégias que não empregam diferenciação, como algoritmos genéticos (AG) e *simulated annealing* (SA), na determinação de arquiteturas e pesos ideais da rede;
- Avaliação de desempenho de outros tipos de função de ativação para os neurônios das redes do tipo MLP, por exemplo, extraídos de um conjunto de funções candidatas;
- Análise mais aprofundada do efeito bias/variância;
- Estudo de técnicas eficazes para a divisão do conjunto amostral, em treinamento, validação e teste, assegurando uma maior capacidade de generalização da rede treinada;
- Implementação computacional em arquiteturas paralelas dos métodos de segunda ordem;
- Estudo de métodos construtivos de modelos em simuladores realistas voltados à manufatura e produção;
- Desenvolver uma ferramenta computacional acessível e de fácil utilização, a fim de democratizar o acesso a uma tecnologia anteriormente restrita a especialistas, através de um software de código aberto e com processamento em nuvem.

Além disso, os modelos desenvolvidos aqui podem ser usados para o projeto de acopladores de guia de onda considerando diferentes materiais, número de seções, variação do período e simulações 3D.

E ao destacar o baixo custo de produção e a eficiência excepcional na transmissão de potência óptica em bandas largas, esse avanço tecnológico promete revolucionar o panorama das telecomunicações, abrindo caminho para novas aplicações e soluções em conectividade.

A capacidade desses dispositivos de operar em banda larga os torna especialmente adequados para enfrentar os desafios emergentes de comunicação e transmissão de dados em alta velocidade, contribuindo assim para a construção de uma infraestrutura tecnológica mais robusta e resiliente.

Ainda sob este aspecto, a integração de técnicas de modelagem baseadas em Inteligência Artificial marca um salto significativo em direção à automação e otimização dos processos computacionais. Ao simplificar a concepção e a implementação de

dispositivos ópticos fotônicos tridimensionais, essas técnicas abrem caminho para a criação de soluções personalizadas e adaptáveis, perfeitamente alinhadas com as demandas específicas de cada contexto de aplicação. A capacidade de integrar-se com softwares de código aberto fortalece ainda mais a escalabilidade e a flexibilidade dessas soluções, estimulando a colaboração e a participação da comunidade global de desenvolvedores e inovadores.

5.3 SUPORTE ACADÊMICO

A pesquisa deste trabalho foi realizada na infraestrutura do Laboratório de Telecomunicações (LABTEL) da escola Politécnica da Universidade Federal da Bahia – UFBA, onde estão disponibilizados softwares usados para as simulações incluindo: Matlab®, e suítes de edição da Microsoft e *OpenSource* tais como o LaTeX e *Jupyter*(EXECUTABLE BOOKS COMMUNITY, 2021), bem como toda a estrutura de computadores.

Além destes, o recurso computacional mais utilizado foi o notebook pessoal marca Dell modelo G5 i7 que possui um processador Intel Core i7 de nona geração, com seis núcleos, 12 *threads* e frequência de até 4.5 GHz, para processamento e multitarefa. Possui também 16 GB de memória RAM DDR4, que permite rodar vários aplicativos simultaneamente sem perda de desempenho, e um SSD de 512 GB NVMe, que oferece uma alta velocidade de leitura e gravação de dados.

O resultado deste trabalho respalda o desenvolvimento de estruturas fotônicas que apresentem boas características de eficiência com custo operacional reduzido e que tenham possibilidade de aplicação em larga escala industrial, abrindo novas perspectivas para o avanço da fotônica e suas tecnologias associadas.

REFERÊNCIAS BIBLIOGRÁFICAS

AGGARWAL, C. C. An Introduction to Artificial Intelligence. Em: **Artificial Intelligence**. Cham: Springer International Publishing, 2021. p. 1–34.

ANDRADE, R. DE O. **Alicerce para Inovar**. Disponível em: <<https://revistapesquisa.fapesp.br/alicerce-para-inovar/>>. Acesso em: 1 set. 2023.

ANDRADE, R. DE O. **Forças Armadas dos Estados Unidos apoiam projetos de pesquisa básica no Brasil**. Disponível em: <<https://revistapesquisa.fapesp.br/forcas-armadas-dos-estados-unidos-apoiam-projetos-de-pesquisa-basica-no-brasil/>>. Acesso em: 1 set. 2023.

ARDEN, F.; SAFITRI, C. **Hyperparameter Tuning Algorithm Comparison with Machine Learning Algorithms**. 2022 6th International Conference on Information Technology, Information Systems and Electrical Engineering (ICITISEE). **Anais...**2022.

BHANDARI, J. K. et al. **Development of Electronic-Photonic 3D System in Package: Architecture, Integration, and Scaling**. 2023 IEEE 73rd Electronic Components and Technology Conference (ECTC). **Anais...**2023.

BISHOP, C. M. **Neural Networks for Pattern Recognition**. CLARENDON PRESS ed. Oxford: Oxford University Press, 1995.

BOCK, P. J. et al. Subwavelength grating periodic structures in silicon-on-insulator: a new type of microphotonic waveguide. **Optics Express**, v. 18, n. 19, p. 20251, 13 set. 2010.

BOLODURINA, I.; ZABRODINA, L. *Investigation of Optimization Algorithms for Neural Network Solutions of Optimal Control Problems with Mixed Constraints*. **Machines**, v. 9, n. 5, p. 102, 17 maio 2021.

BRAGA, A. DE P.; LUDERMIR, T. B.; CARVALHO, A. C. P. **Redes neurais artificiais: teoria e aplicações**. Rio de Janeiro: LTC, 2000.

CACHO, D.; DUARTE, V. **Produção e caracterização de guias de onda de telureto e germanato para aplicações em optoeletrônica**. Tese de doutorado-São Paulo: Universidade de São Paulo, 31 ago. 2010.

CHAND, L.; CHEEMA, S. S.; KAUR, M. *Understanding neural networks*. Em: DEEP SINGH, C.; KAUR, H. (Eds.). **Factories of the future**. Chichester: John Wiley & Sons, 2023. p. 83–102.

CHEBEN, P. et al. *Subwavelength waveguide grating for mode conversion and light coupling in integrated optics*. **Optics Express**, v. 14, p. 4695–4702, 29 maio 2006.

CHUGH, S. et al. *Machine learning approach for computing optical properties of a photonic crystal fiber*. **Optics Express**, v. 27, n. 25, p. 36414, 9 dez. 2019a.

CHUGH, S. et al. *Machine learning regression approach to the nanophotonic waveguide analyses*. **Journal of Lightwave Technology**, v. 37, n. 24, p. 6080–6089, 15 dez. 2019b.

COURANT, R. *Variational methods for the solution of problems of equilibrium and vibrations*. **Bulletin of the American Mathematical Society**, v. 49, n. 1, p. 1–23, 1943.

DA ROS, F. et al. *Machine learning for photonics: from computing to communication*. *Proceedings of 2023 IEEE Photonics Society Summer Topicals Meeting Series*. **Anais...Giardini-Naxos : IEEE**, 19 jul. 2023.

DA SILVA FERREIRA, A. et al. *Towards an integrated evolutionary strategy and artificial neural network computational tool for designing photonic coupler devices*. **Applied Soft Computing Journal**, v. 65, p. 1–11, abr. 2018.

DA SILVA, J. P. *Simulação por elementos finitos da propagação de feixes ópticos em estruturas fotônicas*. Tese de doutorado-Campinas: Universidade Estadual de Campinas, 2003.

DOERR, C. R. *Silicon photonic integration in telecommunications*. **Frontiers in Physics**, v. 3, p. 1–16, ago. 2015.

DOURADO SISNANDO, A. et al. *Artificial neural networks approach for the analysis and synthesis of metamaterial waveguides*. In: (N. Engheta, M. A. Noginov, N. I. Zheludev, Eds.) *Metamaterials, Metadevices, and Metasystems 2020*. **Anais...California: SPIE-The International Society for Optical Engineering**, ago. 2020.

DOURADO SISNANDO, A.; RODRÍGUEZ ESQUERRE, V. F.; RUBIO MERCEDES, C. E. *Power Coupling Optimization in Periodical Segmented Waveguides by Bio-Inspired Algorithms*. *Conference on Lasers and Electro-Optics*. **Anais...Optica Publishing Group**, 5 jun. 2016. Disponível em: <https://opg.optica.org/abstract.cfm?URI=CLEO_SI-2016-JW2A.82>. Acesso em: 6 out. 2024

DUONG, L. H. K. et al. *Crosstalk noise reduction through adaptive power control in inter/intra-chip optical networks*. **IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems**, v. 38, n. 1, p. 43–56, jan. 2019.

EDEN, R. C.; COLEMAN, P. D. Proposal for microwave modulation of light employing the shift of optical absorption edge with applied electric field. **Proceedings of the IEEE**, v. 51, n. 12, p. 1776–1777, 1963.

EXECUTABLE BOOKS COMMUNITY. *Jupyter Book* (Executable Books Community, Ed.). [s.l.] Zenodo, 12 fev. 2021. Disponível em: <<https://doi.org/10.5281/zenodo.4539666>>.

FAGGELLA, D.; BENGIO, Y. *The rise of neural networks and deep learning in our everyday lives: a conversation with Yoshua Bengio*. Disponível em: <<https://emerj.com/ai-podcast-interviews/the-rise-of-neural-networks-and-deep-learning-in-our-everyday-lives-a-conversation-with-yoshua-bengio/>>. Acesso em: 8 nov. 2023.

FILHO, F. F. C. **Algoritmos numéricos: uma abordagem moderna de cálculo numérico**. 3ª ed. Belo Horizonte: LTC - Livros Técnicos e Científicos, 2018.

FILIPPIS, L. A. C. DE et al. ANN Modelling to Optimize Manufacturing Process. Em: EL-SHAHAT, A. (Ed.). *Advanced Applications for Artificial Neural Networks*. London: InTechOpen, 2018. p. 1–296.

FLAMINI, F. et al. Photonic architecture for reinforcement learning. *New Journal of Physics*, v. 22, n. 4, p. 1–12, 1 abr. 2020.

FLORES, Z. C. et al. Photonics Technology Will Transform the Software and Hardware of Telecommunications. *Journal of Industrial and Intelligent Information*, v. 4.2, p. 178–185, mar. 2016.

FORNASIER, M. DE O.; KNEBEL, N. M. P. Inteligência Artificial: Desafios e riscos ético-jurídicos. *Revista Jurídica Direito & Paz*, v. 43, p. 207–228, 21 ago. 2020.

FREIRE, P. et al. *Artificial neural networks for photonic applications-from algorithms to implementation: tutorial*. *Advances in Optics and Photonics*, v. 15, n. 3, p. 739–834, set. 2023.

GIL, A. C. *Como elaborar projetos de pesquisa*. 4ª ed. São Paulo: Editora Atlas S.A., 2002.

GONÇALVES, H. DE A. *Manual de metodologia da pesquisa científica*. 2ª ed. São Paulo: Avercamp, 2014.

GONÇALVES, M. S.; ISAYAMA, Y. H.; HERNÁNDEZ-FIGUEROA, H. E. *A novel three dimensional vector finite element method for periodic photonic devices*. *Microwave and Optical Technology Letters*, v. 58, n. 11, p. 2665–2668, nov. 2016.

GOSTIMIROVIC, D.; YE, W. N. *Automating Photonic Design with Machine Learning*. 2018 IEEE 15th International Conference on Group IV Photonics (GFP). *Anais...*2018.

GT IA. **Recomendações para o avanço da inteligência artificial no Brasil**. *GT-IA da Academia Brasileira de Ciências* Rio de Janeiro Academia Brasileira de Ciências, dez. 2023. Disponível em: <<https://www.ipea.gov.br/cts/pt/central-de-conteudo/artigos/artigos/398-abc-faz-recomendacoes-para-avanco-da-inteligencia-artificial-no-brasil>>. Acesso em: 31 mar. 2024

GUPTA, S. et al. Operationalizing digitainability: encouraging mindfulness to harness the power of digitalization for sustainable development. *Sustainability (Switzerland)*, v. 15, n. 8, 1 abr. 2023.

GURJAR, A. P.; PATEL, S. B. Fundamental categories of artificial neural networks. Em: INFORMATION RESOURCES MANAGEMENT ASSOCIATION (Ed.). *Research Anthology on Artificial Neural Network Applications*. Hershey: IGI Global, 2021. p. 1–30.

GURNEY, K. *An introduction to neural networks*. 1st. ed. London: CRC Press, 2017.

HAGAN, M. T.; DEMUTH, H. B.; BEALE, M. H. *Neural network design*. 2nd. ed. Boston: Publishing Co., 1996.

HALIR, R. et al. **Waveguide sub-wavelength structures: A review of principles and applications**. *Laser and Photonics Reviews*, Weinheim, Wiley-VCH Verlag, 1 jan. 2015.

HAMMOND, A. M.; CAMACHO, R. M. Designing integrated photonic devices using artificial neural networks. *Optics Express*, v. 27, n. 21, p. 29620, 14 out. 2019.

HAN, X. et al. Potential new material for optical fiber: Preparation and characterization of transparent fiber based on natural cellulosic fiber and epoxy. *International Journal of Biological Macromolecules*, v. 224, p. 1236–1243, jan. 2023.

HAWKING, S.; CELLAN-JONES, R. **AI could spell end of the human race**. UKBBC News, 2 dez. 2014. Disponível em: <<https://www.bbc.com/news/av/technology-30299992>>. Acesso em: 1 nov. 2023

HAYKIN, S. *Neural Networks - A Comprehensive Foundation*. Upper Saddle River: Prentice Hall, 1999.

HORNIK, K.; STINCHCOMBE, M.; WHITE, H. Multilayer feedforward networks are universal approximators. *Neural Networks*, v. 2, n. 5, p. 359–366, jan. 1989.

HUNSPERGER, R. G. Theory of Optical Waveguides. Em: **Integrated Optics**. New York, NY: Springer New York, 2009. p. 33–52.

HURWITZ, J.; KIRSCH, D. *Machine Learning IBM Limited Edition*. Hoboken: John Wiley & Sons, Inc., 2018.

HUSSAIN, M. S. et al. Examining the role of artificial intelligence in determining sustainable competitive advantage: Evidence from the pharmaceutical sector of Karachi Pakistan. *Journal of Future Sustainability*, v. 3, n. 1, p. 23–34, 2023.

JIANG, J.; CHEN, M.; FAN, J. A. **Deep neural networks for the evaluation and design of photonic devices**. *Nature Reviews Materials*, Nature Research, 1 ago. 2021.

JOANNOPOULOS, J. D. et al. *Photonic Crystals*. Princeton: Princeton University Press, 2008.

KAKU, M. *Physics of the Future: How Science Will Shape Human Destiny and Our Daily Lives by the Year 2100*. Rio de Janeiro: Rocco, 2012. v. 1

KIM, S.; YAN, R. Recent developments in photonic, plasmonic and hybrid nanowire waveguides. *Journal of Materials Chemistry C*, v. 6, p. 11795–11816, 2018.

KRZANOWSKI, W. J. Cross-validation in principal component analysis. *Biometrics*, p. 575–584, 1987.

- KUDYSHEV, Z. A. et al. Machine learning–assisted global optimization of photonic devices. *Nanophotonics*, v. 10, p. 371–383, 2020.
- KUSSUL, E. et al. **Rosenblatt perceptrons for handwritten digit recognition.** *IJCNN'01. International Joint Conference on Neural Networks. Proceedings (Cat. No. 01CH37222)*. **Anais...IEEE**, 2001.
- LEE, D. L. *Electromagnetic principles of integrated optics*. 1st. ed. Chichester: John Wiley & Sons, 1986.
- LEE, E.-H. et al. **Micro/nano-scale integrated photonic devices and circuits: issues and challenges.** *SPIE/OSA/IEEE Asia Communications and Photonics*. **Anais...2002**. Disponível em: <<https://api.semanticscholar.org/CorpusID:120918307>>
- LEE, J.-M. et al. Low bending loss metal waveguide embedded in a free-standing multilayered polymer film. *Optics Express*, v. 17, n. 1, p. 228, 5 jan. 2009.
- LI, L.; BURKE, J. J. Linear propagation characteristics of periodically segmented waveguides. *Journal of the Optical Society of America (Optics Letters)*, v. 17, n. 17, 1992.
- LIAO, L. et al. An Empirical Study of the Impact of Hyperparameter Tuning and Model Optimization on the Performance Properties of Deep Neural Networks. *ACM Transactions on Software Engineering and Methodology*, v. 31, n. 3, 1 jul. 2022.
- LIFANTE, G. *Integrated photonics : fundamentals*. Chichester: John Wiley & Sons, 2003.
- LIM, J.; PSALTIS, D. MaxwellNet: Physics-driven deep neural network training based on Maxwell's equations. *Appl Photonics*, v. 7, n. 1, p. 1–10, 31 jul. 2022.
- LIU, W. K.; LI, S.; PARK, H. S. **Eighty Years of the Finite Element Method: Birth, Evolution, and Future.** *Archives of Computational Methods in Engineering*, Springer Science and Business Media B.V., 1 out. 2022.
- MA, W. et al. **Deep learning for the design of photonic structures.** *Nature Photonics*, Nature Research, 1 fev. 2021.
- MACKAY, D. J. C. Bayesian Non-Linear Modeling for the Prediction Competition. Em: *Maximum Entropy and Bayesian Methods*. Dordrecht: Springer Netherlands, 1996. p. 221–234.
- MARCATILI, E. A. J.; MILLER, S. E. Improved Relations Describing Directional Control in Electromagnetic Wave Guidance. *Institute of Electrical and Electronics Engineers (IEEE)*, p. 1–28, 22 jan. 1969.
- MATLAB, S. *Matlab*. Natick, MA: **The MathWorks**, 2012.
- MCCULLOCH, W. S.; PITTS, W. A logical calculus of the ideas immanent in nervous activity. *The Bulletin of Mathematical Biophysics*, v. 5, n. 4, p. 115–133, dez. 1943.

- MOLESKY, S. et al. Inverse design in nanophotonics. *Nature Photonics*, v. 12, n. 11, p. 659–670, 26 nov. 2018.
- MOLNAR, J. P. Conduction phenomena in gases. *Electrical Engineering*, v. 69, n. 12, p. 1071–1076, dez. 1950.
- MONROE, D. Neuromorphic computing gets ready for the (really) big time. *Communications of the ACM*, v. 57, p. 13–15, 2014.
- MOON, J. H.; YANG, S. Chemical Aspects of Three-Dimensional Photonic Crystals. *Chemical Reviews*, v. 110, n. 1, p. 547–574, 13 jan. 2010.
- MOREIRA, P. A. et al. Artificial Intelligence and Industry 4.0? Validation of Challenges Considering the Context of an Emerging Economy Country Using Cronbach's Alpha and the Lawshe Method. *Engineering*, v. 4, n. 3, p. 2336–2351, 1 set. 2023.
- MUSUMECI, F. et al. An Overview on Application of Machine Learning Techniques in Optical Networks. *IEEE Communications Surveys and Tutorials*, v. 21, n. 2, p. 1383–1408, 1 abr. 2019.
- O'BRIEN, P. **Scaling of Integrated Photonic Packaging for Volume Manufacture.** 2018 *International Conference on Optical MEMS and Nanophotonics (OMN)*. *Anais...*2018.
- OLIVEIRA DAS MERCÊS, V.; DOURADO SISNANDO, A.; RODRIGUEZ ESQUERRE, V. F. **Machine Learning Modeling of Periodical Subwavelength Tapers Coupling Efficiency.** Proceedings of SPIE - The International Society for Optical Engineering. *Anais...*2021.
- OSKOOI, A. et al. Robust optimization of adiabatic tapers for coupling to slow-light photonic-crystal waveguides. *Optics Express*, v. 20, n. 19, p. 21558, 10 set. 2012.
- PRESS, G. **Deep Learning Pioneer Yoshua Bengio Says AI Is Not Magic And Intel AI Experts Explain Why And How.** *Forbes*, 2019. Disponível em: <<https://www.forbes.com/sites/gilpress/2019/09/20/deep-learning-pioneer-yoshua-bengio-says-ai-is-not-magicand-intel-ai-experts-explain-why-and-how/>>. Acesso em: 2 set. 2023
- RODRÍGUEZ ESQUERRE, V. F.; ISÍDIO LIMA, J. J.; DOURADO SISNANDO, A. **Efficient Neural Network Modeling of Photonic Crystal Fiber Chromatic Dispersion.** *Latin America Optics and Photonics Conference*. *Anais...*Washington, D.C.: OSA, 2010.
- ROSENBLATT, F.; PAPERT, S. *Perceptron*. Cambridge: MIT Press, 1969. v. 9
- RUBIO NORIEGA, R. E. **Estudo de guias periodicamente segmentados usando o método dos elementos finitos.** Dissertação de Mestrado-Campinas: UNICAMP, 2012.
- RUBIO-MERCEDES, C. E. et al. Novel FEM Approach for the Analysis of Cylindrically Symmetric Photonic Devices. *Journal of Lightwave Technology*, v. 27, n. 21, p. 4717–4721, nov. 2009.

RUBIO-MERCEDES, C. E. et al. Analysis of Straight Periodic Segmented Waveguide Using the 2-D Finite Element Method. *Journal of Lightwave Technology*, v. 32, n. 11, p. 2163–2169, jun. 2014.

RUBIO-NORIEGA, R. E. **Estudo de guias periodicamente segmentados usando o método dos elementos finitos**. Dissertação (Mestrado)-Campinas, SP: Universidade Estadual de Campinas, 22 ago. 2012.

RUGGIERO, M. A. G.; LOPES, V. L. DA R. **Cálculo numérico: aspectos teóricos e computacionais**. São Paulo: Makron Books do Brasil, 1997.

SADIKU, M. N. O. A simple introduction to finite element analysis of electromagnetic problems. *IEEE Transactions on Education*, v. 32, n. 2, p. 85–93, maio 1989.

SADIKU, M. N. O. *Elementos de eletromagnetismo*. São Paulo: Bookman, 2004.

SADIKU, M. N. O. *Numerical Techniques in ELECTROMAGNETICS with MATLAB*. 3rd. ed. São Paulo: CRC Press, 2009.

SARTON, G. The History of Science. *Monist*, v. 26, n. 3, p. 321–365, 1916.

SCHAD, H. Computing costs for FEM analysis of foundation engineering problems and possible ways of increasing efficiency. *International Journal for Numerical and Analytical Methods in Geomechanics*, v. 9, n. 3, p. 261–275, 7 maio 1985.

SECTOR, S.; ITU, O. F. ITU-T Series G. [s.d.].

SEGUNDO, M. T. A. B. **Análise de guias de onda ópticos segmentados com sequência quase periódica**. [Tese de Doutorado]-Natal: UFRN, 2018.

SHAHBAZ, M.; BUTT, M. A.; PIRAMIDOWICZ, R. Breakthrough in Silicon Photonics Technology in Telecommunications, Biosensing, and Gas Sensing. *Micromachines*, v. 14, n. 8, p. 1637, 19 ago. 2023.

SILVA, D. M. DA. **Fabricação e caracterização de guias de onda baseadas em filmes finos e vidros de óxido de metal pesado dopados com Er³⁺ e Yb³⁺ e contendo nanopartículas metálicas para aplicações em dispositivos fotônicos**. [Tese de Doutorado]-São Paulo: Universidade de São Paulo, 26 nov. 2012.

SISNANDO, A. D. et al. Inverse design of tapers by bio-inspired algorithms. *Journal of Microwaves, Optoelectronics and Electromagnetic Applications*, v. 19, n. 1, p. 39–49, 2020.

TANSU, N. Photonics-Advances in Fundamental Sciences and Engineering Technologies of Light. *Photonics*, v. 1, p. 1–8, 2014.

TAYLOR, J. M.; TANEV, S. **Photonic simulation software tools for education**. In: *International Topical Meeting on Education and Training in Optics and Photonics*. Anais...2007.

TSE-TUNG, M. **ON PRACTICE**. Pequim: Foreign Languages Press, 1967.

TSIRIGOTIS, A. et al. Unconventional Integrated Photonic Accelerators for High-Throughput Convolutional Neural Networks. *Intelligent Computing*, v. 2, jan. 2023.

UNESCO et al. *UNESCO science report the race against time for smarter development*. Paris: UNESCO, 2021. v. 1

WANG, Z. et al. Reduce Loss and Crosstalk in Integrated Silicon-Photonic Multistage Switching Fabrics through Multichip Partition. *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, v. 40, n. 1, p. 101–114, 1 jan. 2021.

WEISSMAN, Z. et al. Asymmetric *Y*-junction wavelength demultiplexer based on segmented waveguides. *Applied Physics Letters*, v. 67, n. 3, p. 302–304, 17 jul. 1995.

WEISSMAN, Z.; HARDY, A. 2-D mode tapering via tapered channel waveguide segmentation. *Electronics Letters*, v. 28, n. 16, p. 1514, 1992.

YARIV, A. *Quantum Electronics*. Chichester: John Wiley & Sons, 1989. v. 3rd

ZHAO, L.; SHI, P.; ZHANG, H. Bi-Directional Benes With Large Port-Counts and Low Waveguide Crossings for Optical Network-on-Chip. *IEEE Access*, v. 9, p. 115788–115800, 2021.

APÊNDICES:

APÊNDICE A:

CÓDIGOS

Intercepção em C – Matlab®

```
valFitnessFfinal(iffitness3)=Eff;
Eff_aux = [Eff_aux; Eff];

global qpw %declarando a variavel global
qpw = [qpw qp]; %obtenção das variáveis de largura

%%aqui II
% exportar w para um arquivo unico

dlmwrite('ws.bat',qp_aux,'-append');

% exportar eff2 para um arquivo unico

dlmwrite('eff2.bat',Eff_aux,'-append');

% concatenar w e eff2 para um arquivo unico

dlmwrite('bdwe.bat',[qp_aux Eff_aux],'-append');

disp("PARA ANALISE =")
jL
iffitness3
```

APÊNDICE B:

CÓDIGOS

Estatísticas dos atributos – Python

```
In [7]: import pandas as pd
import numpy as np
import glob
from sklearn.model_selection import train_test_split

In [106]: df = pd.read_csv(r'C:\TAPERSgeradores\BDAIS\1.55\bdwspais.bat', header = None)
#df

In [105]: XY = df
#XY

In [16]: #Definido inputs e outputs X, são as colunas de 0 a 14 e y a coluna 16
X = df.drop([14],axis=1)
y = df[16]

In [17]: # Dividindo o conjunto em 60% treinamento e 40% teste
X_train, X_test, y_train, y_test = train_test_split(X,y,test_size=0.4, random_state = 42)

In [18]: #Dividindo o segundo conjunto em 50% validação e 50 % teste
X_val, X_test, y_val, y_test = train_test_split(X_test, y_test, test_size=0.5, random_state = 42)

In [19]: #confirmando que o somatório é 100% - AQUI 60% TREINO, 20% VALIDAÇÃO E 20% TESTE
for dataset in [y_train, y_val, y_test]:
    print(round(len(df) / len(y), 2))

1.0
1.0
1.0

In [21]: X_train.to_csv(r'C:\TAPER\SPLIT\train_input.csv', index=False)

In [104]: #X_train

In [103]: X_val.to_csv(r'C:\TAPER\SPLIT\val_input.csv', index=False)
#X_val

In [102]: X_test.to_csv(r'C:\TAPER\SPLIT\test_input.csv', index=False)
#X_test

In [101]: y_train.to_csv(r'C:\TAPER\SPLIT\train_output.csv', index=False)
#y_train

In [100]: y_val.to_csv(r'C:\TAPER\SPLIT\val_output.csv', index=False)
#y_val

In [99]: y_test.to_csv(r'C:\TAPER\SPLIT\test_output.csv', index=False)
#y_test

In [34]: minimotrein = X_train.min(axis=0)

In [98]: #minimotrein

In [97]: maximotrein = X_train.max(axis=0)
#maximotrein

In [96]: MINMAXTREINO = minimotrein.compare(maximotrein)
#MINMAXTREINO

In [95]: #X_val

In [94]: maximoval = X_val.max(axis=0)
#maximoval

In [93]: minimoval = X_val.min(axis=0)
#minimoval
```

```

In [92]: MINMAXVAL= minimoval.compare(maximoval)
          #MINMAXVAL

In [107]: #X_test

In [91]: maximotest = X_test.max(axis=0)
          #maximotest

In [90]: minimotest= X_test.min(axis=0)
          #minimotest

In [89]: MINMAXTESTE = minimotest.compare(maximotest)
          #MINMAXTESTE

In [88]: dftreino = X_train.merge(y_train,how = 'left', left_index = True, right_index = True)
          #dftreino

In [87]: dfval = X_val.merge(y_val,how = 'left', left_index = True, right_index = True)
          #dfval

In [86]: dftest = X_test.merge(y_test,how = 'left', left_index = True, right_index = True)
          #dftest

In [85]: dfcompara1 = MINMAXTREINO.merge(MINMAXVAL,how = 'left', left_index = True, right_index = True)
          #dfcompara1

In [79]: dfcompara2 = dfcompara1.merge(MINMAXTESTE,how = 'left', left_index = True, right_index = True)

In [80]: dfcompara2.rename(columns={'self_x':'Treinomin','other_x':'Treinomax','self_y':'Valmin','other_y':'Valma

In [53]: dftreino.to_csv(r'C:\TAPER\SPLIT\treino.csv', index=False,header = None)
          dfval.to_csv(r'C:\TAPER\SPLIT\val.csv', index=False,header = None)
          dftest.to_csv(r'C:\TAPER\SPLIT\test.csv', index=False,header = None)
          dfcompara2.to_csv(r'C:\TAPER\SPLIT\MINMAX.csv', index=False)

In [81]: #XY

In [82]: #X

In [83]: #y

In [73]: # salvar resultado com .csv retirando o cabeçalho e o índice
          XY.to_csv(r'C:\TAPER\resultado.csv', index = False, header = False)

In [84]: XYstats = XY.describe()

In [76]: XYstats.to_csv(r'C:\TAPER\resultadostats.csv', index = True, header = True)

In [77]: XYstats.to_excel(r'C:\TAPER\resultadostats.xlsx', index=True)

In [ ]:

```

APÊNDICE C:

TABELA DE PARÂMETROS DE TREINAMENTO DA RNA - Funções De Ativação Por Camada – Parâmetros Alcançados.

Camada de Entrada	Camada Oculta 1	Camada Oculta 2	Camada Oculta 3	Camada de Saída	mse	epoch	R	tempo decorrido
logsig	logsig	logsig	logsig	purelin	1.09E-07	177	1	1h1min52.2s
logsig	logsig	logsig	tansig	purelin	1.67E-07	210	1	0h57min42.6s
logsig	logsig	logsig	purelin	purelin	1.05E-07	308	1	1h28min10.2s
logsig	logsig	tansig	logsig	purelin	1.36E-07	96	1	0h27min12s
logsig	logsig	tansig	tansig	purelin	1.08E-07	186	1	0h39min12s
logsig	logsig	tansig	purelin	purelin	1.09E-07	466	1	2h23min19.8s
logsig	logsig	purelin	logsig	purelin	1.13E-07	318	1	1h26min54s
logsig	logsig	purelin	tansig	purelin	1.3E-07	292	1	1h22min42s
logsig	logsig	purelin	purelin	purelin	1.07E-07	417	1	2h4min47.4s
logsig	tansig	logsig	logsig	purelin	1.24E-07	163	1	0h44min58.2s
logsig	tansig	logsig	tansig	purelin	1.09E-07	167	1	0h46min3s
logsig	tansig	logsig	purelin	purelin	1.05E-07	325	1	1h37min33.6s
logsig	tansig	tansig	logsig	purelin	1.83E-07	106	1	0h30min11.4s
logsig	tansig	tansig	tansig	purelin	1.29E-07	142	1	0h40min24s
logsig	tansig	tansig	purelin	purelin	1.07E-07	110	1	0h19min7.2s
logsig	tansig	purelin	logsig	purelin	9.24E-08	137	1	0h38min57s
logsig	tansig	purelin	tansig	purelin	1.29E-07	201	1	0h56min46.8s
logsig	tansig	purelin	purelin	purelin	1.25E-07	818	1	3h47min4.2s
logsig	purelin	logsig	logsig	purelin	1.08E-07	258	1	1h12min48.6s
logsig	purelin	logsig	tansig	purelin	1.16E-07	244	1	1h8min36.6s
logsig	purelin	logsig	purelin	purelin	1.08E-07	404	1	1h58min54s
logsig	purelin	tansig	logsig	purelin	1.03E-07	121	1	0h41min24.6s
logsig	purelin	tansig	tansig	purelin	1.01E-07	276	1	1h26min34.8s
logsig	purelin	tansig	purelin	purelin	1.06E-07	609	1	2h55min40.2s
logsig	purelin	purelin	logsig	purelin	1.19E-07	583	1	2h45min52.2s
logsig	purelin	purelin	tansig	purelin	1.15E-07	1035	1	4h54min58.2s

Tabela 11-TABELA DE PARÂMETROS DE TREINAMENTO DA RNA

tansig	logsig	logsig	logsig	purelin	8.33E-08	135	1	0h43min49.2s
tansig	logsig	logsig	tansig	purelin	1.2E-07	142	1	0h39min23.4s
tansig	logsig	logsig	purelin	purelin	1.13E-07	199	1	0h57min10.2s
tansig	logsig	tansig	logsig	purelin	1.5E-07	118	1	0h35min54s
tansig	logsig	tansig	tansig	purelin	1.44E-07	137	1	0h38min21s
tansig	logsig	tansig	purelin	purelin	1.07E-07	388	1	1h56min39s
tansig	logsig	purelin	logsig	purelin	1.25E-07	182	1	0h54min42s
tansig	logsig	purelin	tansig	purelin	1.09E-07	184	1	0h52min50.4s
tansig	logsig	purelin	purelin	purelin	1.03E-07	481	1	2h21min36s
tansig	tansig	logsig	logsig	purelin	1.1E-07	101	1	0h28min36s
tansig	tansig	logsig	tansig	purelin	1.15E-07	121	1	0h33min38.4s
tansig	tansig	logsig	purelin	purelin	1.19E-07	211	1	1h2min34.8s
tansig	tansig	tansig	logsig	purelin	1.79E-07	84	1	0h23min55.2s
tansig	tansig	tansig	tansig	purelin	1.23E-07	119	1	0h33min51s
tansig	tansig	tansig	purelin	purelin	1.26E-07	192	1	0h56min36s
tansig	tansig	purelin	logsig	purelin	1.13E-07	155	1	0h43min6.6s
tansig	tansig	purelin	tansig	purelin	1.02E-07	160	1	0h47min39.6s
tansig	tansig	purelin	purelin	purelin	1.06E-07	624	1	3h33min1.2s
tansig	purelin	logsig	logsig	purelin	1.06E-07	120	1	0h38min39.6s
tansig	purelin	logsig	tansig	purelin	1.09E-07	107	1	0h34min38.4s
tansig	purelin	logsig	purelin	purelin	1.11E-07	200	1	0h54min27s
tansig	purelin	tansig	logsig	purelin	9.68E-08	95	1	0h27min54.6s
tansig	purelin	tansig	tansig	purelin	1.08E-07	98	1	0h28min21s
tansig	purelin	tansig	purelin	purelin	1.02E-07	372	1	1h38min46.8s
tansig	purelin	purelin	logsig	purelin	1.1E-07	319	1	1h49min6.6s
tansig	purelin	purelin	tansig	purelin	1.06E-07	398	1	1h59min24.6s
tansig	purelin	purelin	purelin	purelin	7.39E-06	2000	0.99968	9h22min57s
purelin	logsig	logsig	logsig	purelin	1.07E-07	177	1	0h49min44.4s
purelin	logsig	logsig	tansig	purelin	1.07E-07	291	1	1h28min4.2s

purelin	logsig	logsig	purelin	purelin	1.08E-07	428	1	2h16min13.2s
purelin	logsig	tansig	logsig	purelin	1.18E-07	234	1	1h8min3.6s
purelin	logsig	tansig	tansig	purelin	1.11E-07	276	1	1h20min28.2s
purelin	logsig	tansig	purelin	purelin	1.2E-07	564	1	2h31min54.6s
purelin	logsig	purelin	logsig	purelin	1.03E-07	337	1	1h31min22.2s
purelin	logsig	purelin	tansig	purelin	1.1E-07	480	1	2h14min11.4s
purelin	logsig	purelin	purelin	purelin	2.25E-06	2000	1	8h31min15.6s
purelin	tansig	logsig	logsig	purelin	1.06E-07	161	1	0h44min49.8s
purelin	tansig	logsig	tansig	purelin	1.08E-07	212	1	0h57min48.6s
purelin	tansig	logsig	purelin	purelin	1.08E-07	276	1	1h19min39.6s
purelin	tansig	tansig	logsig	purelin	1.01E-07	189	1	1h3min16.2s
purelin	tansig	tansig	tansig	purelin	1.22E-07	263	1	1h22min1.8s
purelin	tansig	tansig	purelin	purelin	1.12E-07	626	1	3h32min3.6s
purelin	tansig	purelin	logsig	purelin	1.14E-07	235	1	1h10min57s
purelin	tansig	purelin	tansig	purelin	1.17E-07	437	1	2h2min27s
purelin	tansig	purelin	purelin	purelin	2.15E-06	2000	0.99991	9h8min24.6s
purelin	purelin	logsig	logsig	purelin	1.19E-07	395	1	1h46min16.2s
purelin	purelin	logsig	tansig	purelin	1E-07	262	1	1h10min43.8s
purelin	purelin	logsig	purelin	purelin	9.92E-08	528	1	2h18min10.8s
purelin	purelin	tansig	logsig	purelin	1.22E-07	191	1	0h59min50.4s
purelin	purelin	tansig	tansig	purelin	1.1E-07	235	1	1h4min37.8s
purelin	purelin	tansig	purelin	purelin	1.11E-07	1115	1	6h20min18.6s
purelin	purelin	purelin	logsig	purelin	1.62E-07	2000	0.99999	8h36min43.8s
purelin	purelin	purelin	tansig	purelin	5.48E-07	2000	0.99998	9h24min55.2s
purelin	purelin	purelin	purelin	purelin	0.004829	4	0.75918	0h1min49.8s

APENDICE D:

```
%
clc %Limpar a Tela
clear all %Liberar Memória e Evitar Conflitos de variáveis
tic % Inicia o Cronômetro

% Definir os dados de entrada e saída para a rede neural
% addpath \ (inserir o caminho da árvore de diretórios - Windows)
% addpath / (inserir o caminho da árvore de diretórios - Linux)

%carrega o arquivo que contem o conjunto de dados/atributos de entrada e
%saída
load dataset.bat -ASCII ();%(.txt,.dat,.csv)

dataset = nomedoarquivo ;%generaliza o dataset para qualquer arquivo carregado

X=dataset(:,(1:15)) ;% colunas dos segmentos do TAPER
Y=dataset(:,17);% saída obtida por método numérico e otimizado por algoritmos miméticos

% Criar a rede neural com N camadas ocultas com a seguinte organização de
% neurônios[N1 N2 N3 N(...)]
% hiddenLayerSize = [N1 N2 N3 ];
hiddenLayerSize = [1 1 1 1];%EXEMPLO DE 4 CAMADAS OCULTAS

%Estrutura da Rede
net = feedforwardnet(hiddenLayerSize,'trainlm');%(trainbr)
%Atribuindo Funções de Ativação ( purelin,logsig,tansig)

net.layers{1}.transferFcn = 'purelin';
net.layers{2}.transferFcn = 'purelin';
net.layers{3}.transferFcn = 'purelin';
net.layers{4}.transferFcn = 'purelin';

% Definir os parâmetros de treinamento da rede neural
net.trainParam.epochs = 2000;
net.trainParam.lr = 0.0001;
net.trainParam.goal = 1e-7;
net.trainParam.val = 10;

% Dividir os dados em conjuntos de treinamento, validação e teste
net.divideParam.trainRatio = 0.7;
net.divideParam.valRatio = 0.15;
net.divideParam.testRatio = 0.15;

% Treinar a rede neural com os dados de entrada e saída
[net, tr] = train(net, X', Y');

% Testar a rede neural com novos dados de entrada
load dataset2.bat -ASCII ();%(.txt,.dat,.csv)
l = size(dataset2,1)
testx = dataset2(:,1:15);%valores de entrada)
testy = dataset2(:,17);% valores de saída alvo ou target

TESTEACCURACY=sim(net,testx');
Y=TESTEACCURACY'
comparar = [Y testy], resultado = comparar;
```

```

varia = [testy - Y];
tab = tabulate(varia);

% Exibir a performance da rede neural
mse_train = tr.best_perf;
mse_val = tr.best_vperf;
mse_test = tr.best_tperf;
% mse_test1 = mean((Y_test - net(X_test')).^2);
fprintf('MSE (treinamento): %f\n', mse_train);
fprintf('MSE (validação): %f\n', mse_val);
fprintf('MSE (teste): %f\n', mse_test);

% Plotar os resultados da rede neural
% plotregression(Y', net(X'));
net.plotFcns = {'plotperform', 'plottrainstate', 'ploterrhist', ...
    'plotconfusion', 'plotroc', 'plotregression'};

%mape
abs_percentage_error = abs((testy - Y) ./ testy) * 100;
mape_value = mean(abs_percentage_error);
disp(['O MAPE é: ', num2str(mape_value), '%']);

%mae
absolute_error = abs(testy - Y);
mae_value = mean(absolute_error);

disp(['O MAE é: ', num2str(mae_value)]);

%regressão de valores testados
mdl=fitlm(testy,Y,'linear', 'RobustOpts','on')
tbl=anova(mdl)
r = mdl.Residuals.Raw
aic = mdl.ModelCriterion.AIC
bic = mdl.ModelCriterion.BIC
caic = mdl.ModelCriterion.CAIC
r2 = mdl.Rsquared.Adjusted
disp(['O r2 é: ', num2str(r2)]);
%Finaliza o Cronômetro
toc

```