UNIVERSIDADE ESTADUAL PAULISTA
"JÚLIO DE MESQUITA FILHO"

Faculdade de Engenharia de São João da Boa Vista

Luis Alexandre Silveira Tapia

Design de Filtros Interdigitais para Tecnologia 5G via Redes Neurais Artificiais

São João da Boa Vista

2024

Luis Alexandre Silveira Tapia

Design de Filtros Interdigitais para Tecnologia 5G via Redes Neurais Artificiais

Trabalho de Conclusão de Curso apresentado ao Conselho de Curso de Engenharia Eletrônica e de Telecomunicações da Faculdade de Engenharia de São João da Boa Vista, Universidade Estadual Paulista, como requisito parcial para obtenção do diploma de Graduação em Engenharia Eletrônica e de Telecomunicações .

Orientador: Prof. Dr. Rafael Abrantes Penchel

Coorientador: Helton S. Bernardo

São João da Boa Vista

2024

T172d

Tapia, Luis Alexandre Silveira
Design de filtros interdigitais para tecnologia 5G via redes neurais artificiais / Luis Alexandre Silveira Tapia. -- São João da Boa Vista, 2024
41 p. : il., tabs.

Trabalho de conclusão de curso (Bacharelado - Engenharia de Telecomunicações) - Universidade Estadual Paulista (UNESP), Faculdade de Engenharia, São João da Boa Vista
Orientador: Rafael Abrantes Penchel
Coorientador: Helton Silva Bernardo

1. Telecomunicações. 2. Sistemas de comunicação sem fio. 3. Inteligência Artificial. I. Título.

Sistema de geração automática de fichas catalográficas da Unesp. Biblioteca da Universidade Estadual Paulista (UNESP), Faculdade de Engenharia, São João da Boa Vista. Dados fornecidos pelo autor(a).

UNIVERSIDADE ESTADUAL PAULISTA "JÚLIO DE MESQUITA FILHO"
FACULDADE DE ENGENHARIA - CÂMPUS DE SÃO JOÃO DA BOA VISTA
GRADUAÇÃO EM ENGENHARIA ELETRÔNICA E DE TELECOMUNICAÇÕES

TRABALHO DE CONCLUSÃO DE CURSO

PROJETO DE FILTROS INTERDIGITAIS PARA TECNOLOGIA 5G VIA REDES NEURAIS ARTIFICIAIS

Aluno: Luís Alexandre Silveira Tapia
Orientador: Prof. Dr. Rafael Abrantes Penchel

Banca Examinadora:

- Rafael Abrantes Penchel (Orientador)
- Luís Fernando Duarte (Examinador)
- Lucio Neri Borges (Examinador)

Os formulários de avaliação e a ata da defesa, na qual consta a aprovação do trabalho, devidamente assinados pela banca encontram-se no prontuário eletrônico do aluno.

São João da Boa Vista, 08 de julho de 2024

Agradecimentos

Agradeço primeiramente a Deus por me guiar e abençoar essa jornada. Aos meus pais, por me darem apoio, me incentivarem e acreditarem em mim, mesmo diante de incertezas. Também agradeço ao meu irmão, Daniel, pois, além do apoio aos estudos, também apoiou em outras partes da vida.

Agradeço também aos amigos que fiz na faculdade, em especial ao meu colega de casa, Felipe Franco, que compartilhou comigo momentos de crescimento, superando desafios e celebrando conquistas. Outro colega que gostaria de agradecer em especial é a Ana Luísa, que fez a pandemia se tornar mais leve.

Por fim, agradeço ao meu orientador, Prof. Dr. Rafael Abrantes Penchel, por me introduzir ao mundo das pesquisas científicas, além de apoiar meus estudos.

Resumo

Este trabalho tem como objetivo explorar a aplicação da técnica de redes neurais, especificamente a *perceptron* de múltiplas camadas, no contexto das ondas milimétricas para redes 5G. A proposta é treinar a rede neural com dados relacionados aos parâmetros de espalhamento S_{11} e S_{21}, que são indicadores críticos da performance de dispositivos de micro-ondas, para determinar as dimensões de espaçamento entre os ressonadores, ideais de um filtro interdigital de quinta ordem. Além disso, o filtro é projetado para operar eficientemente na faixa de 26 GHz a 28 GHz, uma banda crucial para a tecnologia 5G. Para alcançar este objetivo, será utilizada uma base de dados gerada através da amostragem aleatória de um filtro interdigital, implementado com técnicas de design de filtros. Os dados coletados serão então analisados e refinados para servir como entrada para o modelo de regressão baseado em rede neural. Os resultados encontrados foram uma largura de banda de 26,30 a 28,05 GHz ($\Delta = 1,75$ GHz) com frequência central de 27GHz, com $|S_{11}|$ e $|S_{12}|$ médios de -17,11 dB e -3,19 dB respectivamente, o modelo MLP demostrou uma melhoria na planicidade da largura de banda na operação do filtro com *score* avaliado em 0,9 e a perda próxima de 153 para R_{G1} e 0,98 de *score* e perda de 38,69 para R_{G2}.

Palavras-chave: redes neurais; filtro interdigital; 5G; ondas milimétricas.

Abstract

This work aims to explore the application of the neural network technique, specifically the multilayer perceptron, in the context of millimeter waves for 5G networks. The proposal is to train the neural network with data related to the scattering parameters S_{11} and S_{21}, which are critical indicators of the performance of microwave devices, to determine the spacing dimensions between the resonators, ideals of a fifth-order interdigital filter. Furthermore, the filter is designed to operate efficiently in the 26 GHz to 28 GHz band, a crucial band for 5G technology. To achieve this objective, a database generated through random sampling of an interdigital filter, implemented with filter design techniques, will be used. The collected data will then be analyzed and refined to serve as input to the neural network-based regression model. The results found were a bandwidth of 26.30 to 28.05 GHz ($\Delta = 1.75$ GHz) with a central frequency of 27GHz, with $|S_{11}|$ and $|S_{12}|$ averages of -17.11 dB and -3.19 dB respectively, the MLP model demonstrated an improvement in bandwidth flatness in filter operation with *score* evaluated at 0.9 and loss close to 153 for R_{G1} and 0.98 of *score* and loss of 38.69 for R_{G2}.

Keywords: neural networks; interdigital filter; 5G; millimeter waves.

Lista de ilustrações

Lista de tabelas

Lista de abreviaturas e siglas

IoT	Internet das Coisas
EB	Exabytes
AR	Realidade aumentada
URLLC	Comunicações de baixa latência e alta confiabilidade
mMTC	Comunicações massivas de tipo máquina
eMBB	Banda larga móvel aprimorada
ADS	Advanced Design System
mmWave	Ondas milimétricas
ANNs	Redes neurais artificiais
TSV	Vias através do substrato
MLP	Multicamadas de *perceptron*
MSE	Erro quadrático médio

Sumário

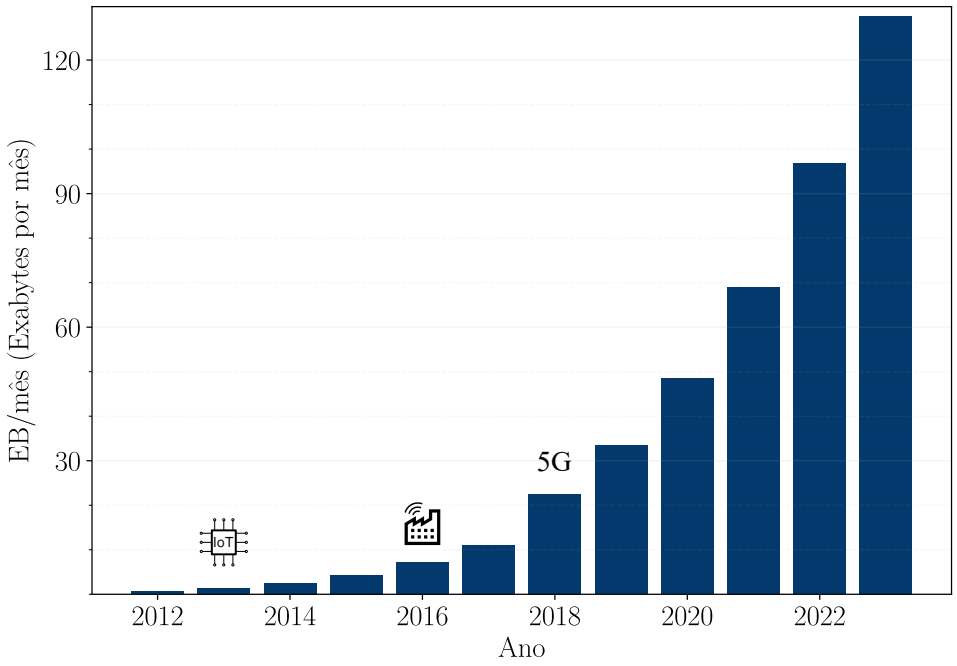
1 Introdução

De acordo com a Ericsson (2023), o tráfego de dados nas redes móveis tem crescido significativamente nos últimos anos, como ilustrado na Figura 1. Esse aumento começou a se destacar em 2013, impulsionado pela popularização dos dispositivos de Internet das Coisas (IoT), que incluem desde equipamentos para casas inteligentes até dispositivos vestíveis e equipamentos médicos (BRAUN, 2024).

A partir de 2016, a evolução para a Indústria 4.0 também contribuiu para esse crescimento, ao lado da crescente adoção de *smartphones* e de *softwares* como serviços (QUEIROZ et al., 2022). Estes fatores combinados resultaram em uma taxa média de crescimento notável nos anos de 2016 e 2017.

Entretanto, foi entre 2017 e 2018, com a introdução e rápida adoção da tecnologia de quinta geração (5G), que o aumento no tráfego de dados móveis se tornou ainda mais pronunciado (INTELLIGENCE, 2024), (ERICSSON, 2023). Isso ocorreu pois, a chegada do 5G proporcionou velocidades de envio e recebimento muito maiores e menor latência do sinal, facilitando ainda mais o uso intensivo de dados em uma variedade de aplicações e serviços.

Figura 1. Média de tráfego de dados por mês ao longo dos anos.



Fonte: Elaborado pelo autor

Além disso, a ascensão do 5G trouxe uma maior geração de receita, ainda mais quando a receita vem de serviços ofertados ao consumidor (ERICSSON, 2020). Conforme

relata Ericsson (2020), em 2017 esse mercado já representava cerca de 56% da receita total da empresa, onde essa porção tende a aumentar graças a adoção da tecnologia 5G. De acordo com Ericsson (2023), essa tendência ascendente no tráfego de dados móveis deve alcançar 325 exabytes (EB) por mês até 2028, marcando um aumento de quatro vezes em relação aos níveis de 2022 (ERICSSON, 2023).

Esse cenário é reflexo dos benefícios que a tecnologia do 5G tende a oferecer, ou seja, uma tecnologia capaz de suportar aplicações com alto uso dados, como a realidade aumentada (AR) ou realidade virtual ou veículos autônomos, ao mesmo tempo que oferece maior segurança e menor latência, faz com que os consumidores sejam mais receptivos ao 5G, mesmo que ele seja mais caro (NOKIA, 2023).

Em uma definição mais formal, a rede móvel 5G é estruturada em três principais serviços: as comunicações de baixa latência e alta confiabilidade (URLLC), comunicações massivas de tipo máquina (mMTC) e banda larga móvel aprimorada (eMBB). O URLLC atende a aplicações que requerem latência mínima e alta confiabilidade, como a condução autônoma e sistemas de controle industrial. O mMTC suporta um grande número de dispositivos IoT de baixa potência que transferem pequenos volumes de dados. O eMBB, projetado para aplicações de alta largura de banda, pode proporcionar taxas de dados de até 20 Gbps em *downlink* e 10 Gbps em *uplink* (POPOVSKI et al., 2018).

No aspecto de banda suportada, o 5G é capaz de operar em um espectro amplo de frequências, fazendo uso de frequências baixas (<1 GHz), para ampla região de cobertura, frequências médias (1-6 GHz) para um equilíbrio entre região de cobertura e capacidade de transmissão de dados e frequências elevadas (> 24 GHz) para fornecer altas taxas de dados (ATEYA et al., 2018).

Conceitualmente, o espectro de ondas milimétricas é estabelecido dentro da faixa de 30 GHz a 300 GHz (Lé et al., 2022), (DEJEN et al., 2022). Entretanto, no contexto do 5G, muitas empresas referem-se a essa faixa como tendo início em frequências a partir de 24 GHz. Isso ocorre devido à sua capacidade de fornecer altas taxas de dados e aliviar o congestionamento da rede. Apesar disso, as ondas milimétricas oferecem oferecem alta capacidade de transmissão de dados, mas ao mesmo tempo estão propensas a alta perda de caminho (*path loss*), o que restringe o seu uso a distâncias curtas (podendo ser utilizado para melhorar a segurança e a privacidade dos dados). Nessas altas frequências, os componentes do circuito devem ser significativamente menores que o comprimento de onda, necessitando do uso de uma abordagem de elementos distribuídos, como linhas de transmissão, para mimetizar o comportamento dos elementos concentrados tradicionais (resistores, capacitores e indutores)(Lé et al., 2022).

O desenvolvimento de componentes em ondas milimétricas (*mmWave*), como filtros ou antenas, exige o uso de softwares especializados como o *Advanced Design System (ADS)* e o *Ansys Eletromagnetics*. Através do ADS, os resultados são obtidos rapidamente usando

os métodos aproximados, tornando-os ideais para análises preliminares de parâmetros S, padrões de radiação e características de componentes. Por outro lado, o Ansys EM realiza análises eletromagnéticas de onda completa, proporcionando dados mais precisos, mas com maior tempo de processamento. Esses desafios de tempo de simulação em componentes mmWave frequentemente levam à exploração de técnicas alternativas, como a aprendizagem de máquina.

A aprendizagem de máquina é uma área em expansão da ciência da computação que pode ser aplicada em diversas áreas, incluindo telecomunicações. Diferente da programação tradicional, onde as regras são explicitamente codificadas, a aprendizagem de máquina adapta-se através de iterações para melhorar o desempenho (NAQA; MURPHY, 2015), (ALDAYA et al., 2022). As redes neurais artificiais (ANNs), um ramo da aprendizagem de máquina, funcionam através da mimetização do processo de aprendizado biológico. Uma ANN é composta por neurônios, onde cada neurônio processa entradas com base em pesos ajustáveis, e o ajuste contínuo desses pesos melhora a precisão do modelo (AGGARWAL, 2018). No contexto do design de filtros interdigital para 5G ondas milimétricas, as ANNs podem ser usadas para determinar as dimensões ideais do filtro com base na resposta dos parâmetros S, tornando o processo de design mais eficaz.

1.1 Objetivo

Este trabalho tem como objetivo modelar os principais parâmetros de um filtro interdigital passa-faixa para aplicações 5G de ondas milimétricas, operando na faixa de 26 GHz a 28 GHz, utilizando uma rede neural artificial para aumentar a largura de banda. As dimensões do filtro interdigital serão definidas com base no substrato ROGERS RT/Duroid 6006. Inicialmente, o design do filtro será baseado em equações de design tradicionais. Posteriormente, para aprimorar a eficiência e a precisão do design, uma rede neural artificial será implementada para determinar as melhores dimensões do filtro com base na resposta dos parâmetros S.

1.2 Organização do trabalho

O trabalho está organizado em 4 capítulos:

Capítulo 1. É feita uma abordagem inicial sobre o tema, com introdução e levantamento de alguns artigos relacionados ao assunto, junto com os objetivos e justificativas para a realização deste trabalho.

Capítulo 2. Foca na parte teórica e nos cálculos do modelo apresentado, sendo este o interdigital. Neste capitulo são abordados conceitos sobre o funcionamento dos filtros, os tipos

existentes de filtros, método de perda de inserção e design de filtros interdigitais.

Capítulo 3. É mostrado o conceito sobre ANN, além do o processo de modelagem, de uma rede neural artificial de regressão.

Capítulo 4. É a conclusão com as considerações finais sobre os modelos, suas aplicações com uma análise sobre o projeto e os possíveis temas a serem aprofundados em caso de continuidade desse trabalho futuramente.

2 Conceitos sobre filtro interdigital

Neste capítulo serão abordados temas que permitam a análise desempenho, definição e *design* de filtros interdigitais. Por meio da matriz de parâmetros de espalhamento S é possível avaliar como o filtro está se comportando dentro da faixa de frequência desejada. Além disso, é importante elucidar os conceitos da modelagem de filtros em radiofrequência, assim como suas características básicas. Por fim, é importante definir um método inicial de design para poder fabricar um filtro que seja capaz de operar na faixa de ondas milimétricas.

2.1 Matriz de parâmetros de espalhamento S

Segundo Pozar (2012) e Hong (2011), em um sistema de radiofrequência é comum fazer uso de técnicas que avaliam as suas ondas de tensão, ondas de corrente e impedância de entrada. Uma dessas técnicas avalia a relação entre o sinal incidente e o refletido nas portas do sistema. Tal técnica é denominada Matriz de Espalhamento e estabelece uma relação dada pela equação (2.1).

$$S_{ij} = \left.\frac{V_i^-}{V_j^+}\right|_{V_k^+=0} \quad \text{Para } k \neq j, \tag{2.1}$$

onde S_{ij} é um elemento da Matriz de Espalhamento S e é obtido através a relação entre as amplitudes do sinal de tensão incidente V_j^+ e do sinal refletido V_i^-. Com isso, é estabelecido um elo entre o sinal que entra na porta j e o sinal que sai na porta i. Vale salientar que para essa condição ocorrer, todos os sinais de todas as outras portas devem ser nulos, garantindo também que as impedâncias das demais portas sejam casadas para evitar possíveis reflexões. Ao garantir essa relação, tem-se para S_{ii} o coeficiente de reflexão e para S_{ij} o coeficiente de transmissão da porta j para a i (POZAR, 2012). Pode-se generalizar a matriz apresentada na Equação (2.2):

$$\begin{bmatrix} V_1^- \\ V_2^- \\ \vdots \\ V_N^- \end{bmatrix} = \begin{bmatrix} S_{11} & S_{12} & \cdots & S_{1N} \\ S_{21} & & & \\ \vdots & & \ddots & \\ S_{N1} & \cdots & \cdots & S_{NN} \end{bmatrix} \begin{bmatrix} V_1^+ \\ V_2^+ \\ \vdots \\ V_N^+ \end{bmatrix}. \tag{2.2}$$

Sendo N correspondente ao número de portas do circuito. Pode-se ainda estabelecer uma relação com o número de parâmetros que a matriz de espalhamento S terá e o número de portas do sistema. No caso a relação é dada por 2^n. Em outras palavras, se o sistema em estudo possui uma porta, ele terá apenas um parâmetro S, o S_{11}; se ele possui duas

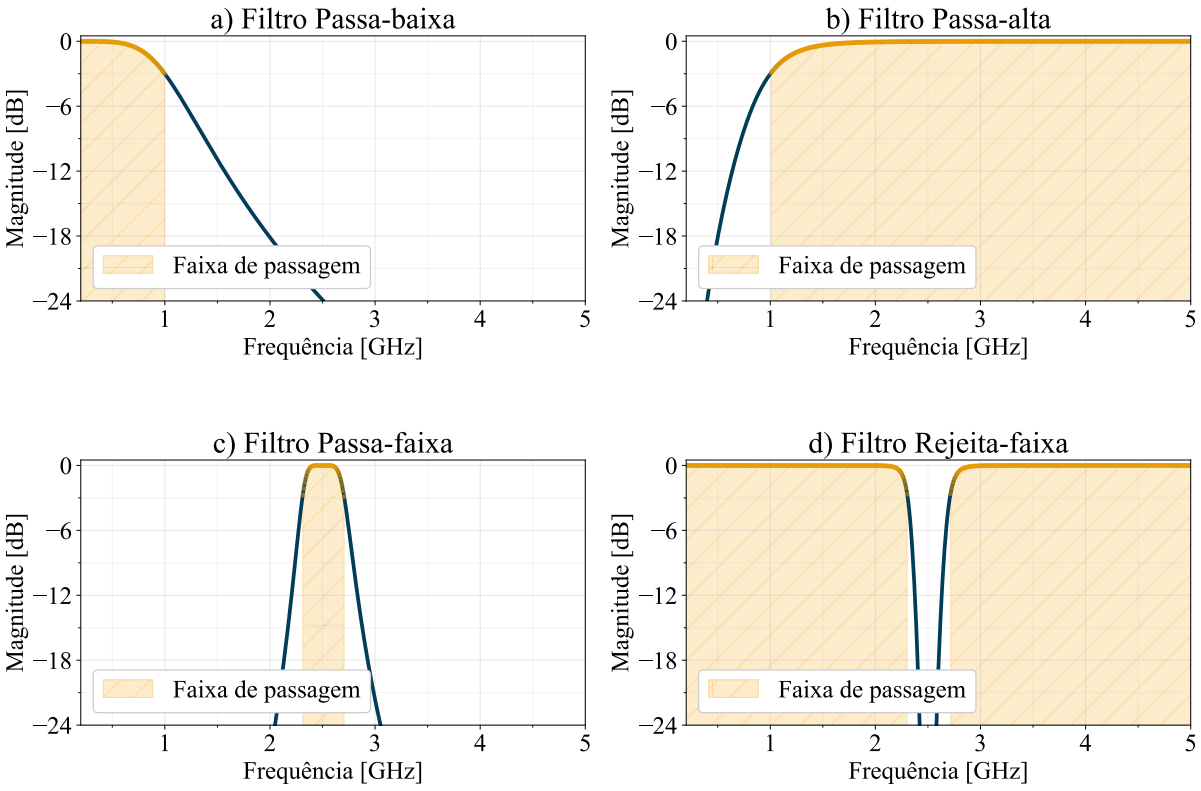
portas, a matriz irá possuir quatro elementos, o S_{11}, S_{21}, S_{12} e S_{22} e assim por diante (PENCHEL, 2024).

2.2 Caracterização de filtros

Os filtros em aplicações de radiofrequência podem ser projetados usando parâmetros concentrados, com componentes como resistores, capacitores e indutores, ou parâmetros distribuídos, utilizando elementos como linhas de transmissão e *stubs* (POZAR, 2012). Independentemente do método, os filtros compartilham características comuns: eles estabelecem a relação entre o sinal de entrada e o de saída com base na frequência. Possuem uma faixa de passagem delimitada por frequências de corte, onde a atenuação é tipicamente de 3 dB a mais que na faixa de passagem. Filtros também podem ser categorizados como passa-baixa, passa-alta, passa-faixa ou rejeita-faixa (HONG, 2011).

Conforme ilustra a Figura 2(a), no filtro passa-baixa o nível de sinal que transita pelo filtro não sofre atenuação para as frequências que estão abaixo da frequência de corte ($f_c = 1$ GHz). No filtro passa-alta (Figura 2(b)) o sinal atenuado é aquele que está abaixo da frequência de corte ($f_c = 1\ GHz$) e no filtro passa-faixa (Figura 2(c)), só há transmissão de sinal entre as frequências de corte inferior ($f_1 = 2,3$ GHz) e superior ($f_2 = 2,7$ GHz). Por fim, no filtro rejeita-faixa há um comportamento oposto ao filtro passa-faixa.

Figura 2. Tipos de filtros



Fonte: Elaborado pelo autor

Dependendo da relação gerada pela entrada e saída, a forma de atenuação do sinal também muda, em outras palavras, conforme Ogata (2011) define, pode-se definir essa relação pela Equação (2.3):

$$H(s) = \frac{Y(s)}{X(s)}, \tag{2.3}$$

onde: $H(s)$ é a função transferência;
$Y(s)$ é a transformada de Laplace do sinal de saída;
$X(s)$ é a transformada de Laplace do sinal de entrada;
s é a frequência complexa definida por $s = \sigma + j\omega$, em que:
σ é a frequência real e,
ω é a frequência angular;

O fato da função de transferência estar escrita em função da variável complexa s, permite encontrar características da função através dos polos do polinômio do divisor e dos zeros do polinômio do numerador. Um polo é um valor da variável complexa s que faz o polinômio denominador ser nulo, em outras palavras, faz a função tender ao infinito, sendo importante para averiguar a estabilidade do sistema e o regime transiente. Já o zero, conforme o próprio nome diz, faz o polinômio numerador ser nulo, influenciando na resposta do sistema em relação a frequência e, também, na estabilidade dele (OGATA, 2011).

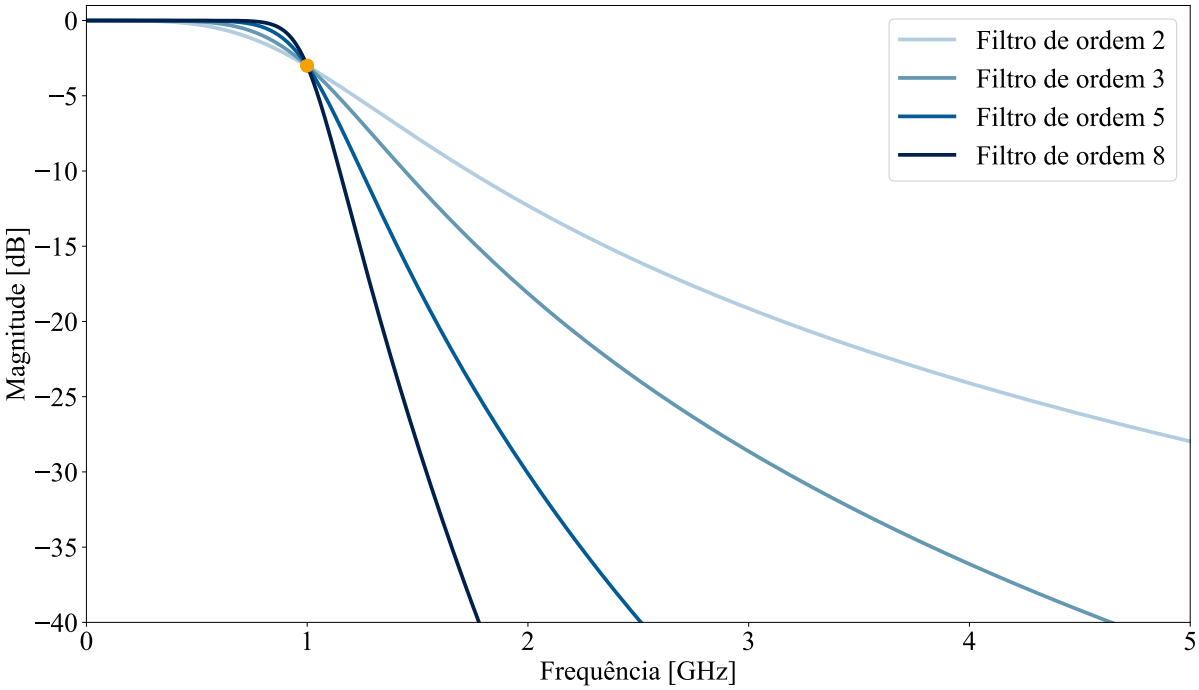
Segundo Hong (2011), um exemplo de função transferência para o filtro é a função *Butterworth* ou totalmente plana (em tradução livre para o termo *maximally flat*). A magnitude da função de transferência *Butterworth* para um filtro passa baixa pode ser dada pela Equação (2.4):

$$|H(j\omega)| = \frac{1}{\sqrt{1 + \left(\frac{\omega}{\omega_c}\right)^{2N}}}, \tag{2.4}$$

onde N é a ordem do filtro e ω_c é a frequência angular de corte, dada por $\omega_c = 2\pi f_c$ do filtro. Vale salientar, que para um filtro, a função transferência dele é o equivalente ao parâmetro $S_{21}(j\omega)$, que representa a relação entre o sinal de entrada na porta 1 e o sinal de saída na porta 2, por isso, pode-se ler $|S_{21}|^2 = |H(j\omega)|^2$, conforme relata Hong (2011). Além disso, a ordem de um filtro é um parâmetro notório, pois ela é responsável pela complexidade do filtro, tornando-o maior conforme o seu valor aumenta. Indo além, à medida que a ordem aumenta, mais abrupta é a queda da faixa de transição de um filtro, conforme ilustra a Figura 3. Isso ocorre pois, a ordem diz respeito à quantidade mínima de elementos reativos necessários para a resposta desejada do filtro, em outras palavras, se um filtro é de terceira ordem, ele irá precisar de três elementos reativos, seja um capacitor

e dois indutores ou um indutor ou dois capacitores, dependendo da topologia adotada (FIORE, 2021), (UNIVERSITY, [s.d.]).

Figura 3. Ordem de um filtro *Butterworth*



Fonte: Elaborado pelo autor

Na Figura 3 são apresentados diversos filtros passa-baixa e com frequência de corte em 1 GHz ($f_c = 1$ GHz), que foram modelados com parâmetros concentrados, cuja a função de transferência é a *Butterworth*. No caso, há um filtro de segunda, terceira, quinta e oitava ordem e conforme a ordem do filtro aumenta, mais intensa é a atenuação do sinal à medida que se afasta da faixa de passagem. Além disso, para $N \to \infty$, o filtro torna-se ideal. Outro fator importante ilustrado é o ponto destacado em amarelo, que independente da ordem do filtro possui a mesma posição (frequência de 1 GHz para uma magnitude de -3 dB). Isso ocorre pois o ponto indica que quando a frequência é igual a frequência de corte ($f = f_c$), tem-se que a intensidade do sinal de entrada cai pela metade, em termos de frequência angular, $|H(j0)|^2 = \frac{1}{2}$ para $\omega = \omega_c$(UNIVERSITY, [s.d.]).

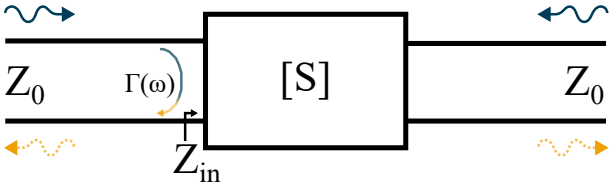
Além dessa função de transferência, também existem outras que mudam a forma com que o filtro responde a variação de frequência, como por exemplo, *Chebyshev I*, *Chebyshev II*, *Bessel*, entre outras mais. A melhor função é aquela que melhor se adapta ao projeto, levando em consideração a seletividade, complexidade do filtro e a banda de passagem.

2.2.1 Modelo de parâmetros concentrados

De acordo com Pozar (2012) e Analysis (2024), o método da perda de inserção possui como vantagem um alto controle sobre a banda de passagem e rejeição. Isso ocorre pois permite ao usuário definir as frequências de corte, atenuação e comportamento da banda de passagem dentro do escopo do método.

Dessa forma, considere o sistema apresentado na Figura 4, onde há a presença de uma rede de micro-ondas com duas portas, que estão casadas com uma impedância Z_0. O sistema possui uma impedância de entrada Z_{in} e um coeficiente de reflexão $\Gamma(\omega)$.

Figura 4. Sistema a ser analisado no método de inserção.



Fonte: Elaborado pelo autor.

Ao definir a potência incidente como P_0, pode-se também concluir que a potência recebida pela carga é a potência incidente menos a potência que é refletida, em outras palavras, $P_{load} = P_0(1 - |\Gamma|^2)$. À luz disso, de acordo com Penchel (2024), a razão da potência perdida pode ser definida pela relação entre esses dois fatores, conforme é apresentado na Equação (2.5):

$$P_{LR} = \frac{P_0}{P_{load}} = \frac{1}{1 - |\Gamma|^2}. \tag{2.5}$$

Desta forma, de acordo com Pozar (2012), P_{LR} é equivalente a $|S_{21}|^2$ se a fonte e a carga estiverem casadas. Indo além, a perda de inserção (IL), em dB é dada pela Equação (2.6):

$$IL = 10\log P_{LR}, \tag{2.6}$$

demonstrando que a IL pode ser definida com o coeficiente de reflexão dependente da frequência, para uma rede passiva em que $|\Gamma| \leq 1$. De acordo com Penchel (2024), $|\Gamma(\omega)|^2$ é uma função par e como conveniência pode ser expandido para uma série de potências pares $a + b\omega^2 + c\omega^4 + ...$ tanto no denominador, quanto no numerador e como consequência $|\Gamma(\omega)|^2$ pode ser reescrito como funções em função de ω^2:

$$|\Gamma(\omega)|^2 = \frac{M(\omega^2)}{M(\omega^2) + N(\omega^2)}, \tag{2.7}$$

onde $M(\omega^2)$ e $N(\omega^2)$ são funções polinomiais em função de ω^2. Logo, substituindo na Equação (2.5), é possível chegar na Equação (2.8):

$$P_{LR} = \frac{1}{1 - |\Gamma(\omega)|^2} = 1 + \frac{M(\omega^2)}{N(\omega^2)}. \tag{2.8}$$

À luz disso, existem inúmeras expansões que são capazes de respeitar as condições impostas. Conforme visto, existem as respostas de *Butterworth*, *Chebyshev I* e *Chebyshev II*, onde é necessário adaptar a função transferência para respeitar a equação 2.8. Tendo essas definições, o primeiro passo do método de perda de inserção é modelar um filtro passa-baixa em um circuito elétrico equivalente, para posteriormente realizar conversões para passa-alta, passa-faixa ou rejeita-faixa.

Dessa forma, de acordo com Penchel (2024), para um filtro *Butterworth* a equação da perda de inserção é dada por:

$$P_{LR} = 1 + k^2 \left(\frac{\omega}{\omega_c}\right)^{2N}, \tag{2.9}$$

onde ω_c é a frequência angular de corte do filtro, k é uma constante utilizada para a perda de inserção no início da banda de transição (para uma perda de 3 dB, a constante k assume o valor 1), N é a ordem. Dessa forma, a partir de uma rede de circuito elétrico equivalente pi ou T, é possível obter a impedância de entrada e com sua função, obter o coeficiente de reflexão para poder modelar a função da perda de inserção. Entretanto, conforme relata Pozar (2012), é possível generalizar esse procedimento ao assumir parâmetros normalizados para os circuitos mostrados na Figura 5. Com isso, ao assumir que as impedâncias sejam normalizadas em relação a impedância da fonte ($g_0 = 1$) e a frequência angular de corte também seja equivalente a 1, podemos utilizar os valores da Tabela 1 para obter os valores de indutância e capacitância necessários para modelar um filtro.

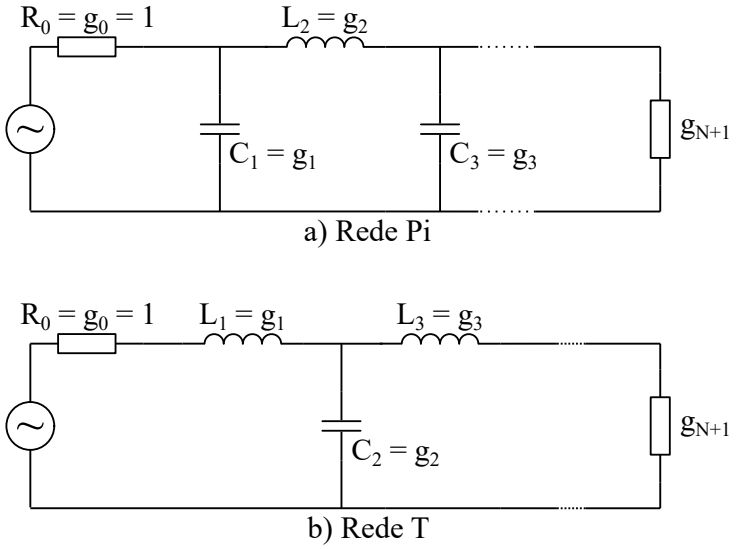
Tabela 1. Parâmetros normalizados para um filtro com resposta Butterworth.

N	g_1	g_2	g_3	g_4	g_5	g_6
1	2,00	1,00				
2	1,41	1,41	1,00			
3	1,00	2,00	1,00	1,00		
4	0,77	1,85	1,85	0,77	1,00	
5	0,62	1,62	2,00	1,62	0,62	1,00

Fonte: Adaptado de Pozar (2012).

Na Figura 5 são mostrados os circuitos equivalentes de rede pi (5.a) e T (5.b) utilizados para o design de filtros e na Tabela 1 são ilustrados os parâmetros normalizados para um filtro com resposta Butterworth, onde $g_0 = 1$, $\omega_c = 1$ e N varia de 1 a 5. Portanto,

Figura 5. Circuito elétrico equivalentes: a) Rede pi b) Rede T

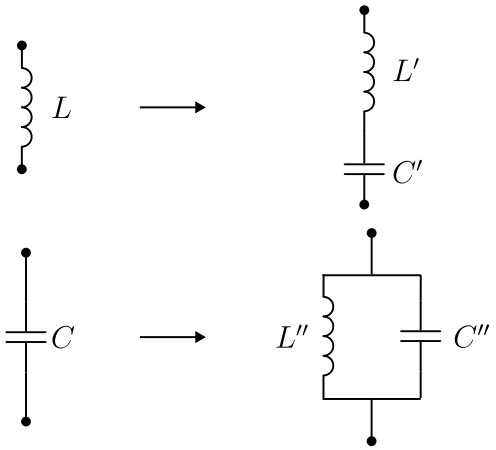


Fonte: Elaborado pelo autor

para implementar um filtro passa baixa em uma frequência de corte e impedância de fonte específicos, é necessário realizar um escalonamento, conforme descreve a seção 8.4 do (POZAR, 2012).

O desenvolvimento de um filtro passa-faixa se da ao realizar a conversão do filtro passa-baixa, onde uma indutância L por unidade de comprimento apresenta um comportamento similar a uma associação de uma indutância L' e capacitância C' em série; uma capacitância C por unidade de comprimento apresenta um comportamento similar a uma associação de indutância L'' e capacitância C'' em paralelo, conforme ilustra a Figura 6 (PENCHEL, 2024).

Figura 6. Conversão de elementos concentrados para o design de um filtro passa-faixa



Fonte: Elaborado pelo autor

Na Figura 6 são ilustradas as transformações necessárias para implementar um filtro passa-faixa com parâmetros concentrados, onde, os valores de L', C', L'' e C'' são

dados pela Equação (2.10):

$$L' = \frac{R_0 g_i}{\Delta \omega_c}, \tag{2.10a}$$

$$C' = \frac{\Delta}{\omega_c R_0 g_i}, \tag{2.10b}$$

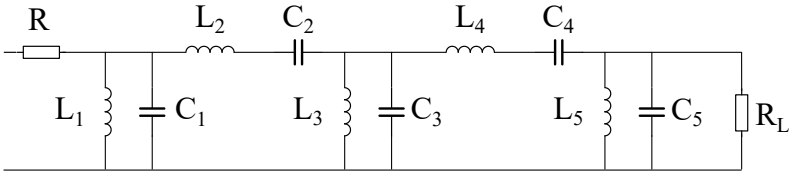
$$L'' = \frac{R_0 \Delta}{\omega_c g_i}, \tag{2.10c}$$

$$C'' = \frac{g_i}{\Delta R_0 \omega_c}. \tag{2.10d}$$

Na Equação (2.10), g_i são os parâmetros normalizados encontrados na Tabela 1. Δ é a fração de largura de banda do filtro, dada por $\Delta = \frac{\omega_2 - \omega_1}{\omega_c}$, onde ω_2 representa a frequência angular de corte superior, ω_1 a frequência angular de corte inferior, e ω_c a frequência angular central.

À luz disso, o circuito equivalente do projeto foi implementado e é representado na Figura 7 onde os valores de capacitância e indutância foram calculados a partir das equações (2.10), da Tabela 1 e uma resistência de carga (R_L) e de fonte (R_0) de 50 Ω. O resultado desses valores são mostrados na Tabela 2. Além disso, os parâmetros $|S_{11}|$ e $|S_{21}|$ são ilustrados na Figura 8.

Figura 7. Design de filtro *Butterworth* passa-faixa de 5ª ordem implementado pelo método da perda de inserção



Fonte: Elaborado pelo autor

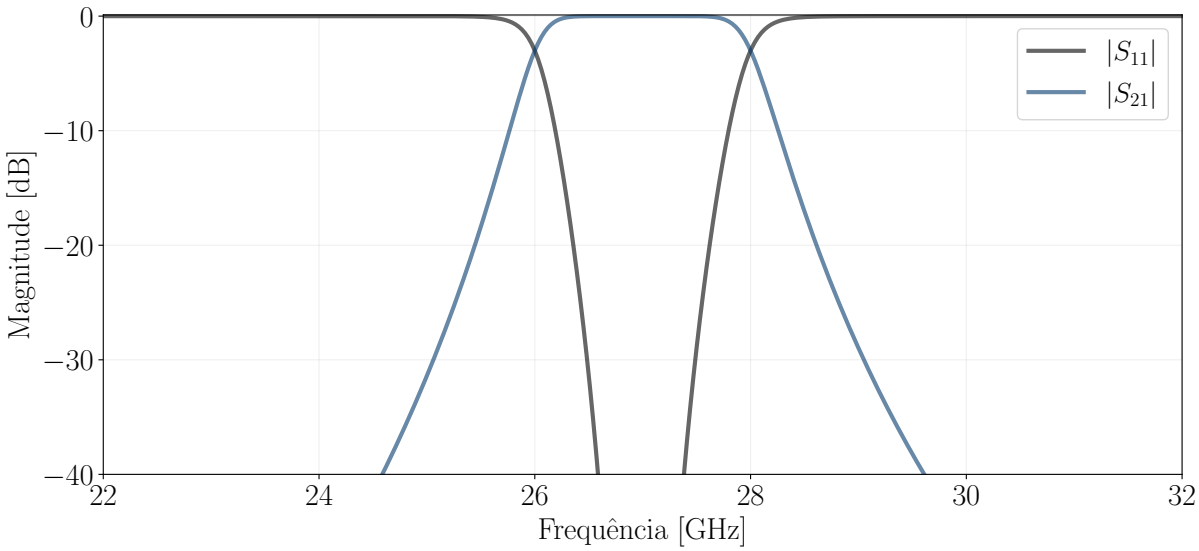
Tabela 2. Valores de indutância e capacitância para o design de um filtro passa-faixa de quinta ordem para operar entre 26 GHz e 28 GHz.

Elemento	**Capacitor**	**Indutor**
1	0,98 pF	35,37 pH
2	5,41 fF	6,43 nH
3	3,18 pF	10,93 pH
4	5,41 fF	6,43 nH
5	0,98 pF	35,37 pH

Fonte: Autor.

Na Figura 8 são mostrados a perda de inserção e o coeficiente de reflexão para o circuito elétrico equivalente. Devido ao cálculo dos parâmetros S terem sido calculados de

Figura 8. Parâmetros $|S_{11}|$ e $|S_{21}|$ para o filtro projetado na Figura 7.



Fonte: Elaborado pelo autor

forma ideal, ou seja, sem considerar as dimensões, o filtro apresenta uma resposta mais próxima do ideal, tendo uma perda de inserção igual a 0 dB na faixa de passagem e um $|S_{11}| < -40$ dB na frequência central.

Após obter o filtro com parâmetros concentrados, o próximo passo é implementar um circuito equivalente com parâmetros distribuídos. Pode-se alcançar esse objetivo utilizando as transformações de Richard, para converter os elementos concentrados em segmentos de linhas de transmissão, e as identidades de Kuroda para separar fisicamente os elementos dos filtros utilizando as linhas de transmissão (POZAR, 2012). No entanto, devido ao uso de uma topologia específica de linhas de transmissão, as linhas acopladas, a abordagem utilizada deverá levar em conta os efeitos de acoplamento entre as linhas, conforme discutido na Seção 2.3.

2.3 Filtro interdigital

Os elementos planares distribuídos, como as linhas de transmissão, têm uma influência significativa no mercado de telecomunicações para frequências elevadas (300 MHz a 300 GHz) devido à sua capacidade de miniaturizar circuitos e integrar mais dispositivos em sistemas compactos e eficientes (BERNARDO; OLIVEIRA; PENCHEL, 2023). Em frequências mais altas, a preferência pelo uso desses elementos é evidente, pois trabalhar com elementos distribuídos é mais vantajoso do que com parâmetros concentrados (HONG, 2011). Entre esses, as linhas de transmissão destacam-se na modelagem precisa de filtros, graças as suas propriedades indutivas (geradas pela corrente que passa pelas linhas) e capacitivas (geradas pela separação entre condutores) inerentes, que permitem a

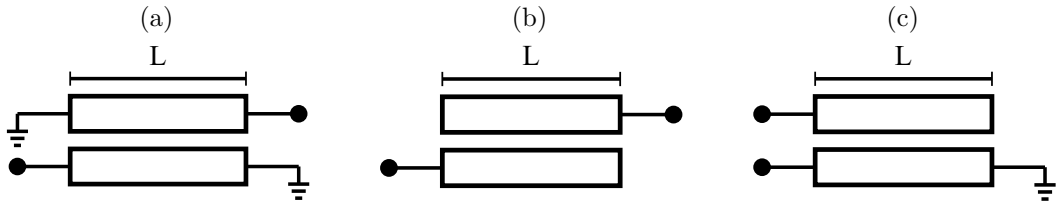
aplicação do método de perda de inserção para ajustar a reatância necessária (POZAR, 2012).

Uma classe especial de linhas de transmissão são as linhas acopladas, que consistem em pares de condutores posicionados próximos entre si. Essas linhas acopladas não apenas permitem a redução do tamanho físico dos filtros, mas também proporcionam vantagens adicionais em termos de desempenho e eficiência. Especificamente, a proximidade gera um aumento na capacitância mútua, pois as cargas elétricas de sinais opostos em cada condutor tendem a se acumular devido à separação estreita entre os condutores (HONG, 2011).

Além disso, os efeitos indutivos são amplificados, pois as correntes que fluem em condutores próximos interagem fortemente, gerando campos magnéticos que se acoplam entre si, permitindo um controle mais preciso sobre suas características de transmissão (HONG, 2011).

É possível implementar um sistema de duas portas a partir de uma linha acoplada ao deixar duas das quatro portas das linhas acopladas em curto-circuito ou como um circuito aberto. Como consequência, segundo Pozar (2012), existem 10 combinações possíveis que oferecem respostas passa-baixa, passa-faixa, passa tudo e rejeita tudo. Na Figura 9 são ilustrados os configurações possíveis para uma resposta passa-faixa.

Figura 9. Configurações de linhas acopladas



Fonte: Elaborado pelo autor

Na Figura 9 são mostradas configurações onde (a) os terminais de cada linha são curto-circuitados, (b) os terminais são deixados em circuito aberto e (c) uma linha é curto circuitada e outra é deixada em circuito aberto. De acordo com LibreTexts (2024), a partir dessas configurações é possível desenvolver filtros interdigitais, *combline*, *hairpin* e de linha acopladas em paralelo. Neste trabalho foi realizado um estudo com a configuração apresentada na Figura 9(a), um filtro interdigital.

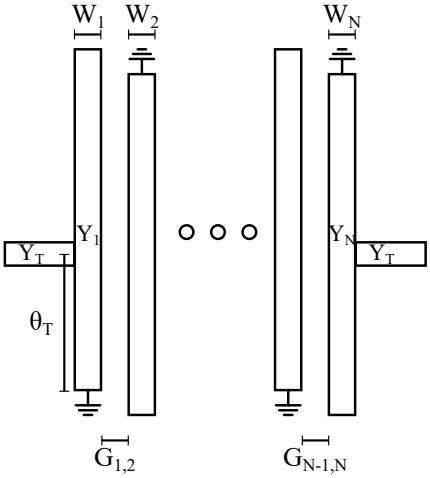
Sendo uma topologia especial de linhas acopladas, os filtros interdigitais são um conjunto de ressonadores de linhas de transmissão. De acordo com Matthaei (1962), as estruturas de linha interdigitais têm um grande destaque em estruturas de ondas lentas. Entretanto, à medida que estudos avançaram, percebeu-se que essa estrutura apresenta bons resultados quando aplicada em filtros passa-faixa. Além disso, eles trazem como

benefícios a possibilidade de serem extremamente mais compactos em área total ocupada em um substrato quando comparado a topologia de linhas acopladas, por exemplo.

A Figura 10 mostra a estrutura generalizada de de um filtro interdigital, que consiste em – de acordo com (HONG, 2011) e (MATTHAEI, 1962) – um *array* de ressonadores de linhas de transmissão com comprimento elétrico de 90º na frequência central. Onde cada ressonador possui um circuito aberto e um curto-circuito em suas extremidades.

A configuração da Figura 10 apresenta *microstrips* de entrada e saída com admitância Y_T, em que devem ter admitâncias iguais à Y_0. Além disso, θ_t é o comprimento elétrico medido em radianos que indica a distância das *microstrips* de entrada/saída em relação aos lados curto-circuitados dos ressonadores inicial e final do filtro. Por fim, há a largura dos ressonadores W_n e o espaçamento entre os ressonadores ($G_{n-1,n}$).

Figura 10. Estrutura generalizada de um filtro interdigital.



Fonte: Baseado em (HONG, 2011). Elaborado pelo autor

À luz disso, conforme Hong (2011), as equações para o cálculo das impedâncias das estruturas é dado pelas Equações (2.11) até (2.16). Nas equações (2.11) e (2.12) são apresentadas as impedâncias par e ímpar ((2.12)) entre o primeiro e o segundo ressonador. Nas equações (2.13) e (2.14) são as impedâncias par e ímpar entre os elementos i e $i+1$. Por fim, o último conjunto de impedâncias par e ímpar são apresentados nas equações (2.15) e (2.16).

$$Z_{0e1,2} = \frac{1}{Y_1 - Y_{1,2}} \tag{2.11}$$

$$Z_{0o1,2} = \frac{1}{Y_1 + Y_{1,2}} \tag{2.12}$$

$$Z_{0ei,i+1} = \frac{1}{2Y_1 - 1/Z_{0ei-1,i} - Y_{i,i+1} - Y_{i-1,i}} \text{ para } i = 2 \text{ até } n-2 \tag{2.13}$$

$$Z_{0oi,i} = \frac{1}{2Y_{i,i+1} + 1/Z_{0ei,i+1}} \text{ para } i = 2 \text{ até } n-2 \tag{2.14}$$

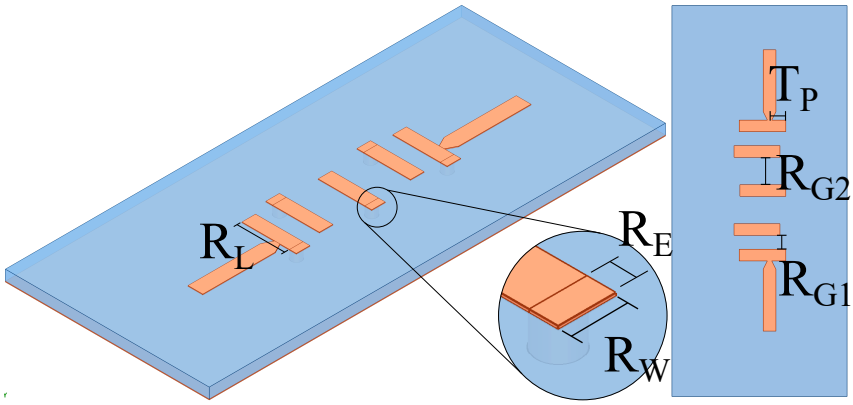
$$Z_{0en-1,n} = \frac{1}{Y_1 - Y_{n-1,n}} \tag{2.15}$$

$$Z_{0on-1,n} = \frac{1}{Y_1 + Y_{n-1,n}} \tag{2.16}$$

Com isso, a primeira parte do projeto é obter uma base de dados capaz de fornecer informações suficientes para que um modelo de multicamadas de *perceptron* (MLP) prediza as dimensões do filtro. Para isso, foi implementado um filtro de quinta ordem com o auxílio das equações (2.11) até (2.16), do software *Advanced Design System* e do *Nuhertz Designs System*.

O filtro implementado possui topologia interdigital de quinta ordem e opera entre 26 GHz até 28 GHz, para um substrato ROGERS RT/duroid 6006 com uma espessura de 0,254 mm e com um condutor de cobre com espessura de 0,18 mm. As dimensões iniciais do filtro implementado pelo software *Nuhertz*estão na Tabela 3, assim como a sua estrutura é ilustrada na Figura 11.

Figura 11. Estrutura do design do filtro interdigital



Fonte: Elaborado pelo autor

Tabela 3. Dimensões do filtro interdigital implementado pelo *Nuhertz*.

Elemento	**Sigla**	**Dimensão [μm]**
Comprimento do ressonador	R_L	1124
Largura do ressonador	R_W	409
Extensão de comprimento do ressonador	R_E	12,65
Espaçamento 1 entre os ressonadores	R_{G1}	539,2
Espaçamento 2 entre os ressonadores	R_{G2}	680
Ponto de acoplamento	T_P	197,4

Fonte: Elaborado pelo autor.

Na Figura 11 é apresentada a estrutura do filtro interdigital, onde é possível observar os principais elementos da estrutura: as dimensões do ressonador são determinadas pelos

valores de R_L para o comprimento e R_W para a largura. Esta estrutura apresenta simetria em relação ao ressonador central, resultando em apenas dois tamanhos distintos de espaçamento entre ressonadores, nomeadamente R_{G1} e R_{G2}. Para alcançar uma resposta aprimorada, um comprimento extra, denominado R_E, é introduzido para os ressonadores. Além disso, a distância elétrica entre o ponto de derivação é T_P. e, por fim, vale salientar que para realizar o curto-circuito, existem as vias através do substrato (TSV), sendo essencial para um filtro interdigital.

Esse filtro servirá como base para o treinamento do modelo MLP, que será abordado no próximo capítulo, focando em como as redes neurais artificiais podem ser aplicadas na modelagem precisa das dimensões do filtro a partir das características operacionais desejadas. Para que isso ocorra, foram geradas 1004 amostras. Essas amostras vieram da variação de $\pm$ 20 % dos parâmetros de espaçamento R_{G1} e R_{G2}.

3 Modelagem de filtro utilizando redes neurais artificiais

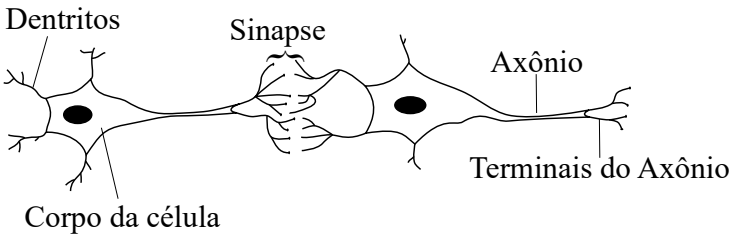
Segundo Nielsen (2015), ANNs são estruturas inspiradas no funcionamento do cérebro humano. Elas são capazes de identificar/aprender padrões complexos a partir de dados, permitindo serem utilizadas como ferramentas para predição e classificação de fenômenos na natureza. Entre os diversos tipos de ANNs, o MLP destaca-se pela sua capacidade de realizar aproximações não lineares, o que é crucial para modelar sistemas como filtros de micro-ondas.

Portanto, a fim de melhor contextualizar o projeto, neste capítulo serão apresentados os fundamentos das redes neurais artificiais, com foco no perceptron multicamadas, além da modelagem do projeto.

3.1 Redes neurais artificiais

As redes neurais artificiais, inspiradas pelos mecanismos de aprendizagem natural, possuem os *perceptrons* que imitam o comportamento biológico dos neurônios (VADAPALLI, 2022). Isto é, em um neurônio, principal componente de uma rede neural, é composto de três elementos essenciais: o corpo da célula, os axônios e os dendritos. Os dendritos são responsáveis por receber os sinais advindos de outros neurônios, enquanto os axônios são responsáveis por transmitirem os sinais para outros neurônios. O ponto de conexão entre um terminal do axônio e um dendrito é definido como sinapse, conforme ilustra a Figura 12

Figura 12. Estrutura de um neurônio



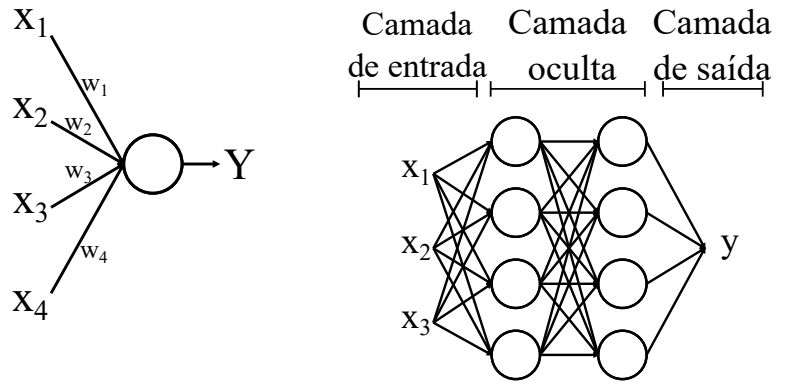
Fonte: Baseado em Vadapalli (2022). Elaborado pelo autor

Dessa maneira, conforme Vadapalli (2022), uma representação artificial de um neurônio pode ser introduzida com o desenvolvimento dos perceptrons. Conforme ilustra a Figura 13a. Esses perceptrons recebem várias entradas ou características (x_1, x_2, ..., x_n) e geram uma única saída binária. Assim como em um organismo, onde cada sinapse possui intensidades de estímulos diferentes, cada entrada em um perceptron é associada a um peso

Figura 13. Definição de perceptron

(a) Representação de um perceptron

(b) Rede de perceptron com múltiplas camadas



Fonte: Elaborado pelo autor

(w_1, w_2, ..., w_n). Esses pesos determinam o grau de influência de cada entrada na geração da saída. Além disso, conforme Aggarwal (2018), a predição feita por um perceptron é binária, assumindo valores de 0 ou 1. Esta predição é baseada na soma ponderada das entradas ($\sum_j w_j x_j$), que é então passada por uma função de ativação para determinar a saída final. Além dos pesos, o perceptron pode incluir um termo de *bias* (b), que é adicionado à soma ponderada das entradas. O termo de *bias* desloca a função de ativação, permitindo que o perceptron modele relações que não passam pela origem, o que pode ser crucial para melhorar a precisão da predição. Nesse caso, a soma ponderada passa a ter o *bias* como fator auxiliar ($\sum_j w_j x_j + b$).

Embora um único perceptron seja capaz de realizar classificações simples, sua limitação está na complexidade dos padrões que pode aprender. Assim como um organismo não possui apenas um neurônio, mas sim uma vasta rede de neurônios interconectados, as redes neurais precisam de múltiplas camadas de perceptrons para processar informações complexas de maneira eficiente. Dessa forma, uma rede neural é composta de diversos perceptrons dispostos em camadas, onde a saída de um neurônio alimenta a entrada de outro, formando uma estrutura em cascata influenciada pelos pesos, conforme ilustrado na Figura 13b.

Nessa arquitetura, as decisões tomadas pela camada inicial são simplistas. No entanto, à medida que camadas adicionais são introduzidas, as decisões se tornam progressivamente mais complexas. As camadas situadas entre a entrada e a saída são comumente chamadas de “camadas ocultas”, pois os cálculos que ocorrem nessas camadas permanecem ocultos dos usuários. Isso permite que a rede neural modele relações complexas e acomode múltiplas saídas, expandindo a capacidade de aprendizado muito além do que um único perceptron poderia alcançar (NIELSEN, 2015).

Portanto, o processo de aprendizagem de uma rede neural artificial gira em torno da

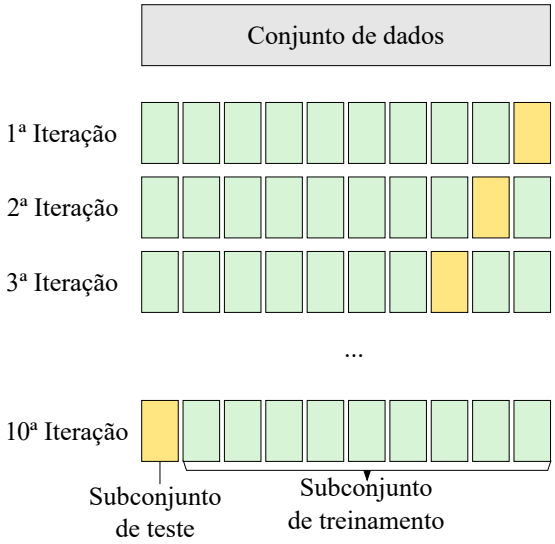
busca de um melhor peso para as conexões entre os neurônios. Esse processo é realizado pelo conjunto de dados de treinamento que engloba dados de saída e de entrada correspondentes. Nos organismos, escolhas incorretas geram respostas indesejadas, de forma equivalente ocorre nas redes neurais artificiais, onde, nos dados de treinamento, são avaliados as discrepâncias entre as saída prevista e a saída real e a partir disto, os pesos são ajustados. Esse mecanismo de avaliar a diferença é denominado de função de perda e o processo de comparar a saída prevista e a saída real e ajustar os pesos de forma a minimizar essa diferença é denominada de *backpropagation* (AGGARWAL, 2018).

Nesse contexto, segundo Terven et al. (2023) uma função de perda é o erro quadrático médio (MSE) que é uma métrica comumente utilizada para obter a média do quadrado da diferença entre os valores preditos e os reais, entre outras palavras, $\frac{1}{n}\sum(y_i - f(x_i))^2$, onde n é o número total de iterações, y_i é a saída desejada e $f(x_i)$ é o valor predito. Essa medida é tipicamente utilizada no conjunto de dados de teste, que é utilizado para avaliar a precisão do modelo gerado pelo conjunto de dados de treinamento.

Embora essa abordagem ofereça algum conhecimento sobre o desempenho do modelo, ela pode ser enganosa se os dados de treinamento tiverem uma distribuição que não represente bem o problema geral (TERVEN et al., 2023). Como consequência, o modelo pode ficar enviesado, comprometendo sua capacidade de generalizar para novos dados. Uma forma de contornar isso é usar técnicas de validação cruzada, que re-amostram os dados de forma a avaliar o modelo de maneira mais robusta. A validação cruzada *k-fold* é uma dessas alternativas, onde o conjunto de dados é dividido em k subconjuntos. Em cada iteração, um dos subconjuntos é reservado para teste, enquanto o restantes é utilizado para treinamento. Esse procedimento é repetido k vezes e o MSE médio ao longo das iterações oferece uma avaliação mais abrangente (AGGARWAL, 2018). A Figura 14 ilustra um exemplo de validação cruzada, onde um conjunto de dados é dividido em 10 segmentos. Desses 10 segmentos, em cada iteração, 9 são usados para treinamento e 1 é usado para teste, onde o MSE é calculado. Esse procedimento é repetido 10 vezes, e a média do MSE obtida nas 10 iterações é usada como a avaliação final do modelo.

Segundo Aggarwal (2018), elemento essencial para o aprendizado do modelo, a função de ativação auxilia a transformar as entradas em saída(s). Esse procedimento inicia com a soma ponderada das entradas e então são passadas por uma função de ativação. De acordo com Sharma, Sharma e Athaiya (2020), sem a aplicação de uma função de ativação, uma rede neural apenas realiza uma combinação linear das entradas ponderadas pelos pesos. Isso torna o modelo equivalente a uma regressão linear, que é fácil de resolver, mas incapaz de capturar e aprender relações complexas dos dados. Portanto, sem funções de ativação, uma rede neural se comporta essencialmente como um modelo de regressão linear, o que limita significativamente seu potencial (SHARMA; SHARMA; ATHAIYA, 2020). Além disso, um aspecto crucial das funções de ativação é a sua derivada, que influencia

Figura 14. Validação cruzada para 10 segmentos



Fonte: Elaborado pelo autor

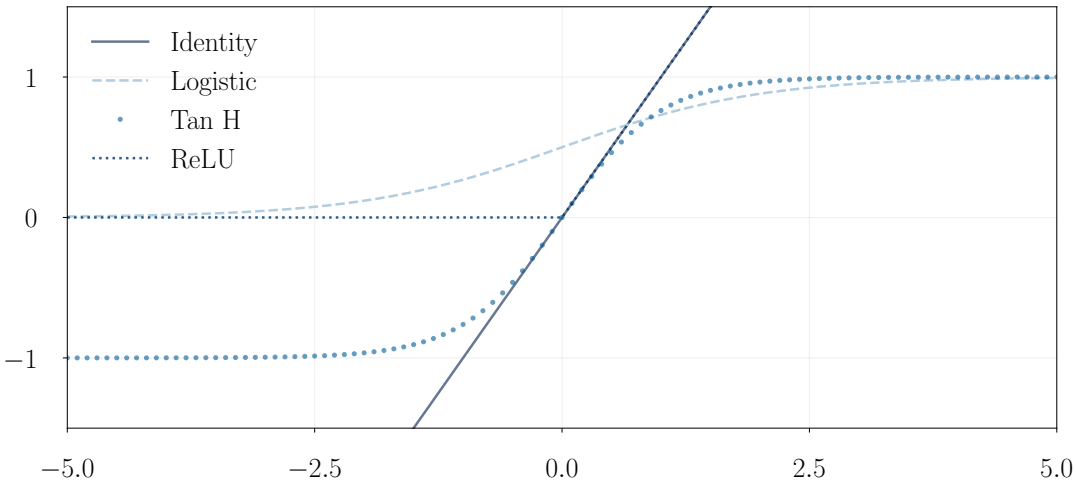
diretamente a atualização dos pesos e dos *bias* durante o treinamento. A derivada da função de ativação é essencial para o algoritmo de *backpropagation*, pois permite calcular os gradientes necessários para ajustar os parâmetros da rede. O cálculo dos gradientes, facilitado pela derivada das funções de ativação, permite ajustar os pesos e *biases*[1] da rede para minimizar a função de perda durante o treinamento (DING; QIAN; ZHOU, 2018).

A Figura 15 apresenta uma variedade de funções de ativação, onde a escolha da função adequada depende de critérios específicos, como por exemplo, o tipo de problema e a arquitetura da rede. A função *identity* (linear) é geralmente usada em problemas de regressão na camada de saída, onde a previsão precisa ser um valor contínuo, mas raramente utilizadas em problemas com camadas ocultas devido a simplicidade. Para previsões de classes binária, a função *sigmoid* é recomendada. A ReLU é amplamente adotada, pois melhora a eficiência do treinamento ao ativar seletivamente os neurônios de maneira simples e eficaz. (AGGARWAL, 2018), (DING; QIAN; ZHOU, 2018).

Como a função *identity* ($f(x) = x$) é linear, seu gradiente é constante, o que não permite a captura de relações complexas nos dados. Isso limita seu uso em redes neurais, já que técnicas de aprendizado baseadas em gradiente, como o *backpropagation*, requerem não linearidade para modelar relações complexas. Por outro lado, a função de ativação *sigmoid* ($f(x) = \frac{1}{1+e^{-x}}$), que mapeia entradas na faixa de 0 a 1, introduz desafios de treinamento, pois não é centrada em zero, o que pode introduzir um viés nas ativações, dificultando o treinamento eficiente das redes profundas. Em contraste, a função de ativação *tanh* ($f(x) = \frac{e^{x}-e^{-x}}{e^{x}+e^{-x}}$) produz uma faixa de saída de -1 a 1,pois é uma versão escalada da *sigmoid*, centrada em zero, o que facilita a aprendizagem de redes neurais porque os valores de

[1] *Biases*: plural de *bias*. *Bias* é um parâmetro que permite a rede neural ajustar melhor o modelo para um mapeamento adequado do processo.

Figura 15. Funções de ativação



Fonte: Elaborado pelo autor

saída não são todos positivos. Isso pode melhorar a convergência durante o treinamento em comparação com a *sigmoid*, pois distribui melhor os gradiente (SHARMA; SHARMA; ATHAIYA, 2020).

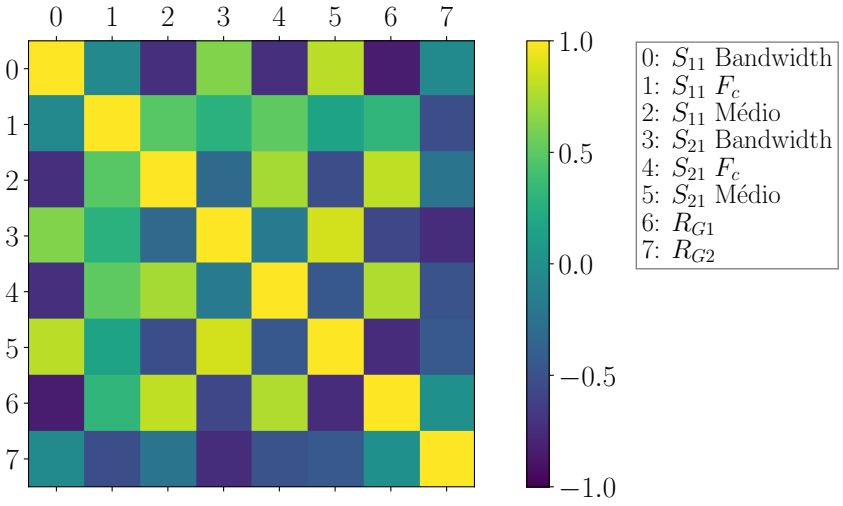
A característica distintiva da função de ativação ReLU ($f(x) = \max(0, x)$) é sua capacidade de ativar seletivamente os neurônios. Quando a entrada é negativa ou zero, a saída da ReLU é zero, efetivamente tornando esses neurônios inativos, ou seja, não contribuem para o cálculo das saídas da camada seguinte. A função derivada da ReLU é simples e eficiente para o algoritmo de retropropagação, pois é 1 para entradas positivas e 0 para entradas negativas. Essa simplicidade reduz o custo computacional durante o treinamento da rede neural.

3.2 Modelagem utilizando MLP

Dados os conceitos vistos na seção 3.1 e das amostras geradas na seção 2.3, buscou-se implementar um modelo de regressão utilizando uma MLP que fosse capaz de predizer os parâmetros de dimensão R_{G1} e R_{G2} do filtro a partir do filtro. Portanto, o primeiro passo envolveu a análise dos dados gerados para estabelecer as entradas do sistema. Elas consistem na largura de banda de S_{11}, definida como o intervalo de frequência onde $|S_{11}|$ é inferior a -10 dB; a média de $|S_{11}|$ dentro dessa largura de banda; F_c de S_{11}, que representa a frequência central dessa largura de banda; a largura de banda de S_{21}, definida como o intervalo de frequência onde $|S_{21}|$ está 3 dB abaixo do seu valor de pico; a média de $|S_{21}|$ dentro dessa largura de banda; e F_c de S_{21}, que indica a frequência central dessa largura de banda. As saídas do modelo são R_{G1} e R_{G2}. A correlação entre os parâmetros é ilustrado na Figura 16.

Na Figura 16, as correlações entre as variáveis do sistema são representadas por cores.

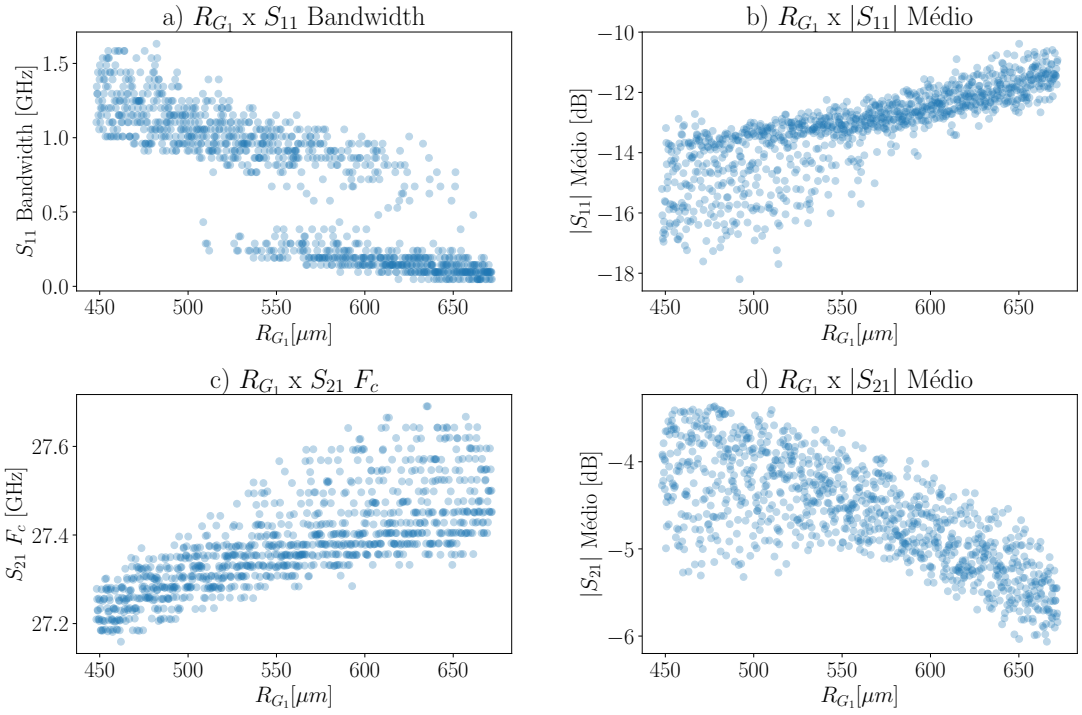
Figura 16. Matriz de correlação entre as variáveis do sistema



Fonte: Elaborado pelo autor

As cores mais próximas do amarelo indicam uma forte correlação positiva, significando que, à medida que uma variável aumenta, a outra também tende a aumentar. Também é possível observar a correlação entre as variáveis através das Figuras 17 e 18.

Figura 17. Correlação entre R_{G1} e os parâmetros de espalhamento S



Fonte: Elaborado pelo autor

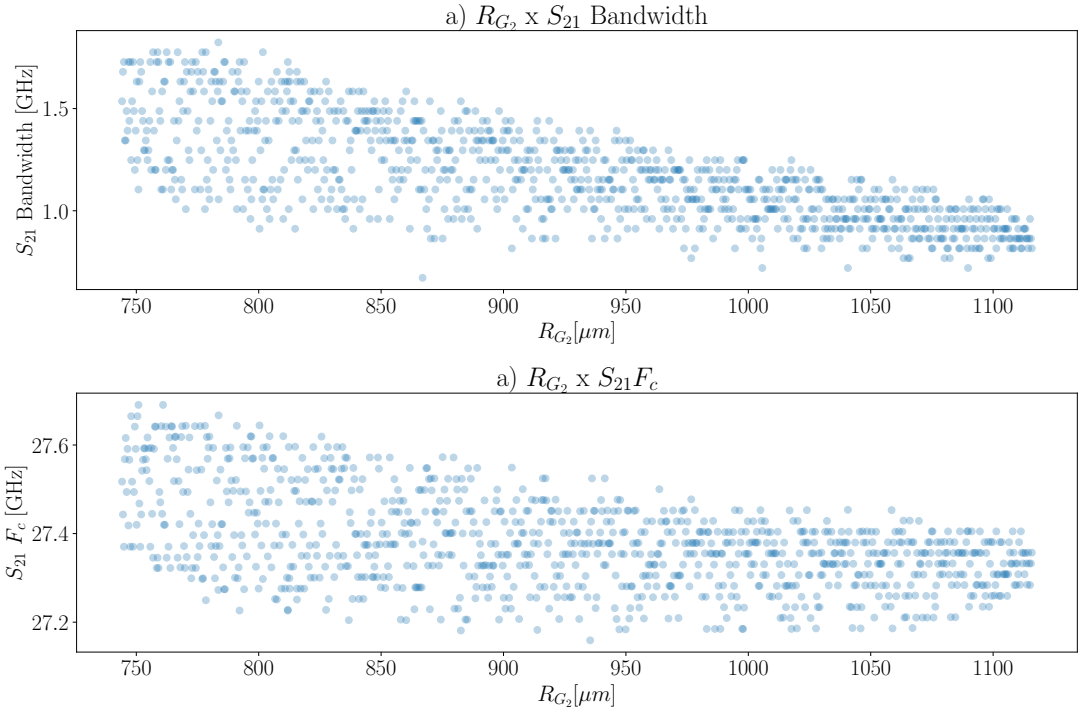
Um exemplo claro dessa correlação é observado na diagonal principal da matriz, que mostra a correlação de uma variável consigo mesma, sempre igual a 1. Outros exemplos de forte correlação positiva estão nas Figuras 17.b) e 17.c), que ilustram a relação entre

R_{G1} e o valor médio de $|S_{11}|$, e entre R_{G1} e a frequência central de S_{21}, respectivamente.

Em contrapartida, cores próximas ao azul escuro indicam uma forte correlação negativa, onde as variáveis são inversamente proporcionais. Isso é claramente visível na relação entre a largura de banda de S_{11} e R_{G1} (Figura 17.a), e entre o valor médio de $|S_{21}|$ e R_{G1} (Figura 17.d). A largura de banda de S_{21} e R_{G2} também apresentam uma forte correlação negativa, como mostrado na Figura 18.a).

Por fim, algumas variáveis exibem correlações mais fracas. Um exemplo é a relação entre a frequência central de S_{21} e R_{G2}, ilustrada na Figura 18.b).

Figura 18. Correlação entre R_{G2} e os parâmetros de espalhamento S



Fonte: Elaborado pelo autor

Para a construção do modelo, foi utilizada a biblioteca *scikit-learn*, uma ferramenta em Python que fornece diversos recursos para análise preditiva de dados (SCIKIT-LEARN, 2024). No desenvolvimento deste modelo, utilizou-se os seguintes pacotes do *scikit-learn*: *MLPRegressor*, *train_test_split*, *StandardScaler*, *Pipeline*, *cross_val_score*, *KFold* e *GridSearchCV*. O *MLPRegressor* permite criar um modelo de regressão baseado em um MLP. Para separar os dados em conjuntos de treino e teste, fez-se uso do *train_test_split*, que, por padrão, divide os dados em 75% para treinamento e 25% para teste, embora essa proporção possa ser ajustada conforme necessário.

O *StandardScaler* é uma ferramenta de pré-processamento essencial que normaliza os dados. Ele transforma cada característica para ter média zero e desvio padrão um, criando uma distribuição normal padrão para cada entrada. Esta normalização é crucial no treinamento de redes neurais MLP, pois dados com escalas muito diferentes podem

fazer com que o processo de otimização por gradiente (*backpropagation*) oscile e demore para convergir. A padronização ajuda o aprendizado ao evitar que características com maiores variâncias dominem o processo de aprendizado (SCIKIT-LEARN, 2024).

O *Pipeline* facilita a automatização do fluxo de trabalho, combinando várias etapas de pré-processamento e modelagem em uma sequência coerente. Isso agiliza o processo de construção e validação de modelos, permitindo um manuseio mais eficiente (SCIKIT-LEARN, 2024). Para avaliação de modelos, foi utilizada a técnica de validação cruzada através dos pacotes *cross_val_score* e *KFold*. Por fim, o *GridSearchCV* permite a otimização dos hiperparâmetros (são as configurações ajustáveis) do modelo através de uma busca exaustiva em um espaço de parâmetros predefinidos. Essa técnica é fundamental para encontrar a combinação ideal de hiperparâmetros que maximiza a performance do modelo (SCIKIT-LEARN, 2024).

Dessa maneira, para encontrar o melhor modelo, foi necessário realizar ajustes nos hiperparâmetros. No caso, o *MLPRegressor* oferece a possibilidade de alterar diversas configurações, como por exemplo a função de ativação, sendo *ReLU*, *tanh*, *sigmoid*, *identity*.

A taxa de aprendizado inicial é outro hiperparâmetro, que é responsável por determinar o quanto os pesos são atualizados em cada iteração do treinamento. Se a taxa for muito alta, o modelo pode não convergir, oscilando em torno do mínimo. Se for muito baixa, o treinamento pode ser extremamente lento, exigindo muitas iterações para alcançar uma solução adequada.

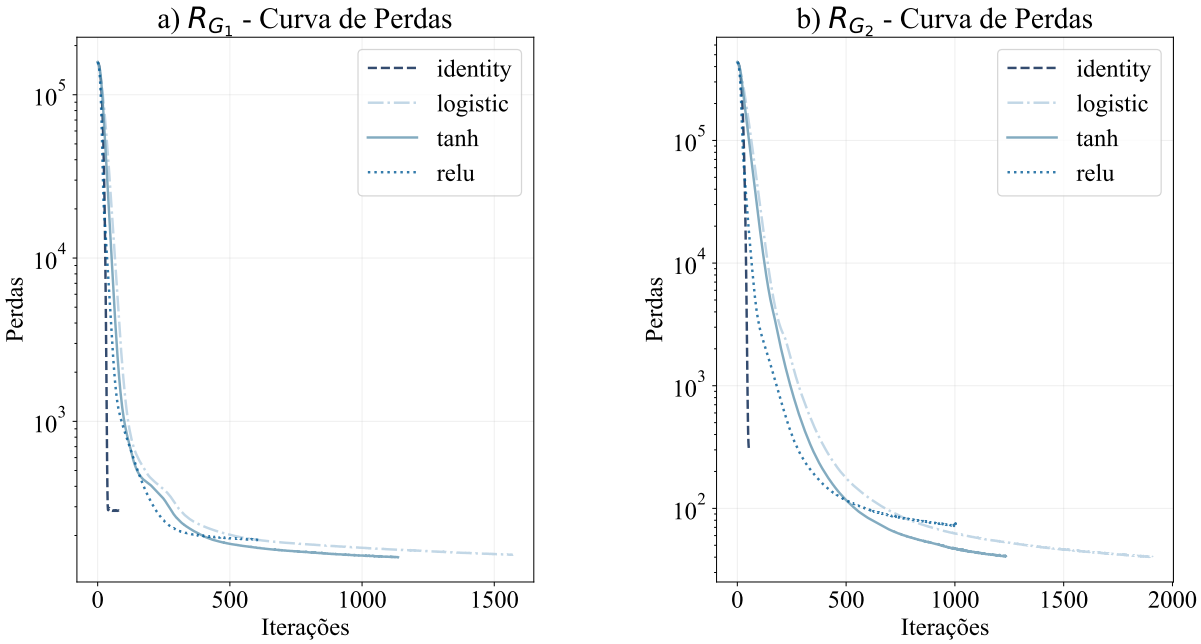
O número de neurônios por camada escondida também é um hiperparâmetro, que especifica quantas camadas escondidas possui o modelo e quantos neurônios estão presentes nela(s).

O hiperparâmetro *alpha* controla a intensidade da regularização L2, que penaliza pesos altos para evitar o *overfitting* (aprendizado enviesado). Um *alpha* alto significa maior penalização, o que ajuda a evitar que o modelo se ajuste demais aos dados de treinamento. No entanto, se *alpha* for muito alto, pode resultar em *underfitting*, em outras palavras, o modelo não consegue aprender adequadamente os padrões dos dados.

A primeira análise foi da função de ativação, onde na Figura 19 são mostradas as performances do modelo, para cada uma das saídas (R_{G1} e R_{G2}) diante da variação das funções de ativação.

É possível constatar através da Figura 19.a) e 19.b) que a função tangente hiperbólica obteve uma menor perda, em conjunto com a função *logistic* (sigmoid). Isso se deve ao fato de ambas as funções não serem lineares e poderem lidar melhor com problemas mais complexos. Vale a ressalva que a função de ativação *tanh* também obteve um melhor resultado com um total de 1198 iterações para o modelo de R_{G1} e 1235 iterações para o modelo de R_{G2}. Além disso, ao analisar o resultado da validação cruzada mostrado

Figura 19. Análise da função de perdas do modelo variando as funções de ativação



Fonte: Elaborado pelo autor

na Tabela 4, é possível constatar que tanto a função *logistic*, quanto a *tanh* apresentam resultados mais próximos de 1, portanto, através do *GridSearch* foi feita uma análise nos hiperparâmetros supracitados a fim de encontrar o melhor modelo.

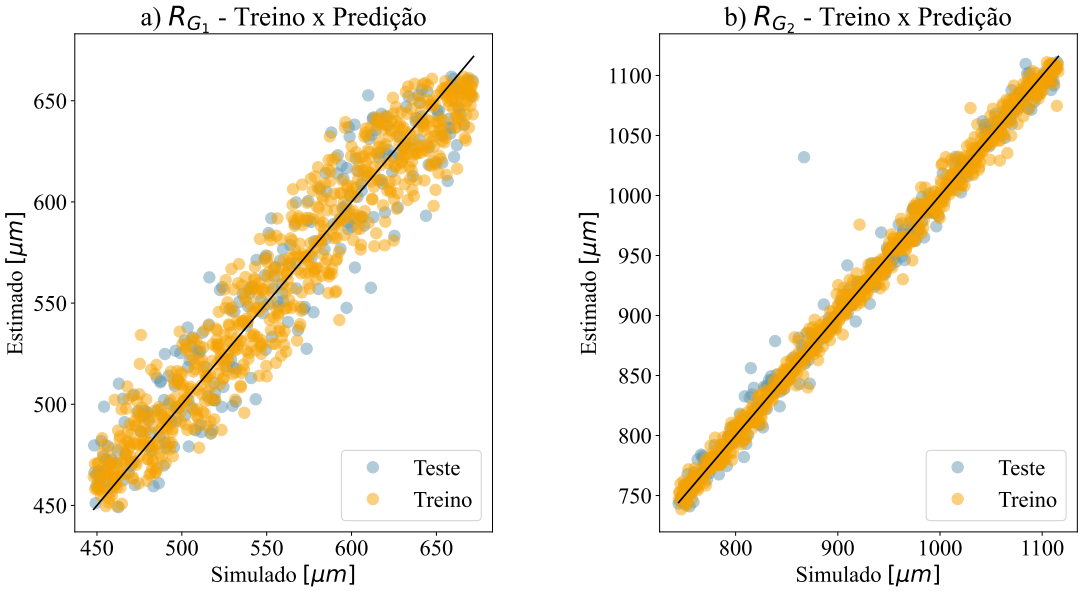
Tabela 4. Resultados da validação cruzada para diferentes funções de ativação.

Função de ativação	R_{G1} **- Validação cruzada**	R_{G2} **- Validação cruzada**
identity	0,85	0,94
logistic	0,89	0,98
tanh	0,89	0,98
ReLU	0,88	0,96

Fonte: Elaborado pelo autor.

Inicialmente fez-se uma busca mais ampla com variações com o número de neurônios na camada oculta variando de 50, 100 e 200. Obteve-se que o melhor resultado foi para 200 neurônios. Então, após essa busca, fez-se outra análise para uma busca mais refinada a fim de reduzir um possível aprendizado enviesado e obter uma melhor capacidade de predição. A busca pelos melhores hiperpâmetros para o modelo abrangeram: número de neurônios entre 190 até 200, com variação de 1 em 1; a taxa de aprendizado inicial de 0,001 até 0,01, variando de 0,001 em 0,001; *alpha* variando de 0,001 até 0,1 variando de 0,01. Essa busca gerou mil possibilidades para cada saída, onde o resultado da melhor é ilustrado na Figura 20, cujo obteve um *score* de 0,90 para R_{G1} e 0,98 para R_{G2}, onde a perda final é de 153,0 para R_{G1} e 38,69 R_{G2}.

Figura 20. Valores estimados x valores treinados



Fonte: Elaborado pelo autor

Na Figura 20 são ilustrados os parâmetros $|S_{11}|$ e $|S_{21}|$ para um design é possível observar que o modelo foi capaz de predizer melhor os dados de R_{G2} do que R_{G1}. O resultado de um bom *score* e uma perda alta pode ser resultado de uma base de dados que não é suficiente para que o modelo possa aprender de forma mais generalista ou pode ser resultado de outros fatores além das entradas desse modelo contribuírem para R_{G1} e R_{G2}, fazendo com que o modelo não possa mapear o comportamento do modelo de forma tão eficaz, principalmente para R_{G1}. Por fim, na Figura 21 mostra o resultado de $|S_{11}|$ e $|S_{21}|$ para os parâmetros de R_{G1} e R_{G2} obtidos através do modelo de regressão MLP para os valores apresentados na Tabela 5.

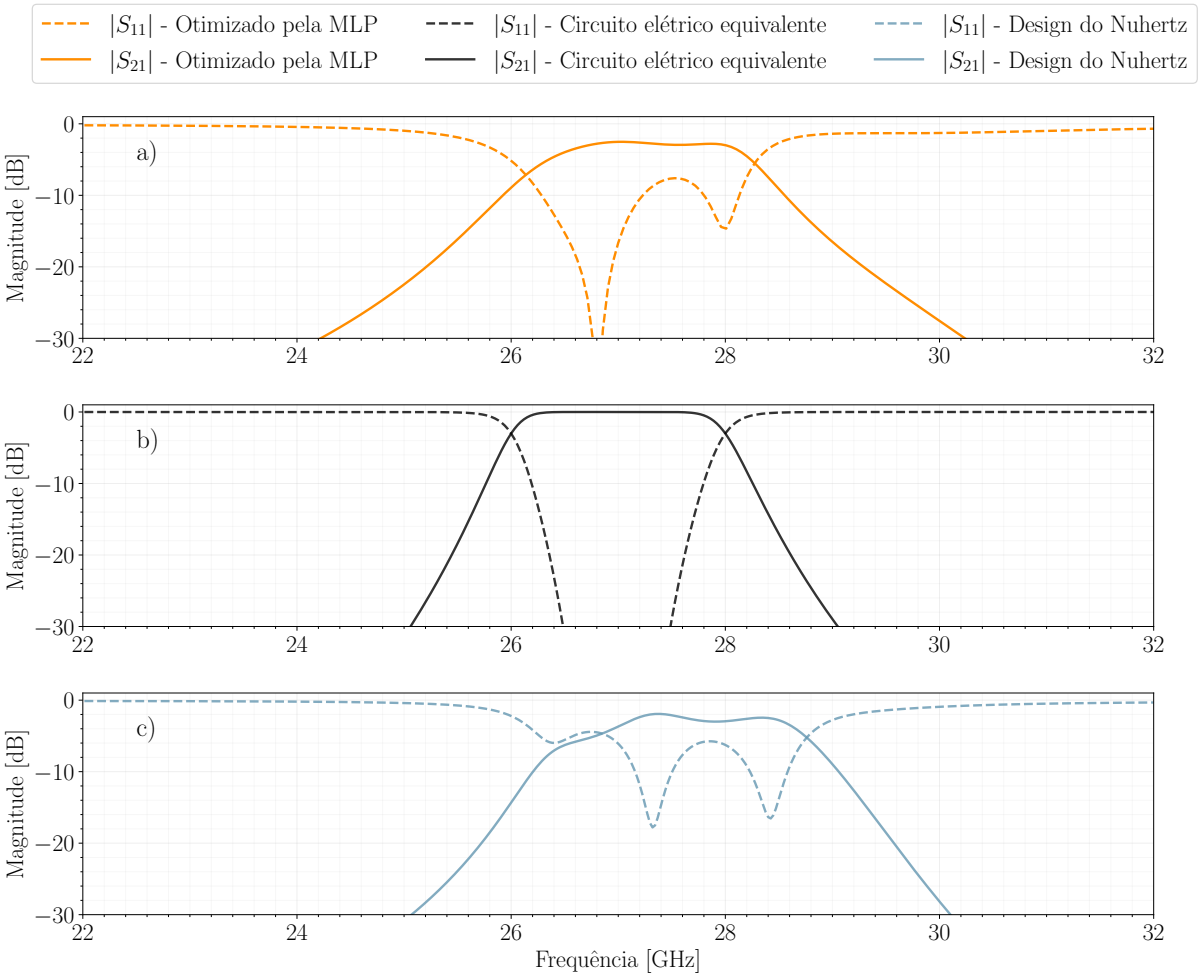
Tabela 5. Dimensões finais do filtro interdigital.

Elemento	**Sigla**	**Dimensão [μm]**
Comprimento do ressonador	R_L	1124
Largura do ressonador	R_W	409
Extensão de comprimento do ressonador	R_E	12,65
Espaçamento 1 entre os ressonadores	R_{G1}	401
Espaçamento 2 entre os ressonadores	R_{G2}	787
Ponto de acoplamento	T_P	197,4

Fonte: Elaborado pelo autor.

A partir do resultado da análise da Figura 21.a) é possível observar o modelo otimizado pela MLP é capaz de de abranger quase a mesma faixa de $|S_{11}| < -10$ dB que o modelo de circuito equivalente (Figura 21.b)). Ainda sobre essa faixa, observa-se também que há um momento em que $|S_{11}| > -10$ dB, isso ocorre pois, no momento de

Figura 21. Valores estimados x valores treinados



Fonte: Elaborado pelo autor

modelar a entrada de largura de banda de $|S_{11}|$, esse fator não foi considerado. Por fim, ao comparar com o design implementado com o software Nuhertz (Figura 21.c)) observa-se uma melhora significativa na largura de banda do filtro para o modelo final sendo de 26,30 GHz até 28,05 GHz ($\Delta = 1,75\text{GHz}$). A mínima perda de reflexão dentro desse intervalo é de 7,8 dB e a perda de inserção na frequência central é de 2,8 dB. Evidentemente, demonstra-se que o modelo MLP apresenta uma resposta mais plana na largura de banda inteira quando comparado ao *design* inicial do filtro (que mostrou uma instabilidade ao longo da magnitude na operação). Desta forma, verifica-se que o modelo MLP apresentou características de performance mais fidedignas ao modelo ideal, na comparação gráfica pela Figura 21.a e b.

4 Conclusão

O objetivo deste estudo foi atingido com êxito uma vez que buscou-se explorar a implementação mais eficaz de elementos para ondas milimétricas, com foco em um filtro interdigital. A motivação vem da necessidade de otimizar processos que atualmente podem ser demorados, especialmente ao simular estruturas complexas.

Foi realizada uma revisão bibliográfica sobre as principais topologias de filtros, cálculo teórico dos principais parâmetros de análise e design, com isso foi destacado o papel de filtro interdigital na literatura e sua contribuição e uso para este trabalho.

O estudo abordou detalhadamente o processo de desenvolvimento do projeto, desde o design inicial do filtro até a análise da base de dados e a escolha dos melhores parâmetros para um modelo de rede neural MLP. A análise revelou que a qualidade da base de dados é crucial para o desempenho do modelo, pois pode influenciar diretamente problemas como o *overfitting*.

Os resultados obtidos pelo modelo de regressão MLP foram encorajadores, uma vez que o modelo foi capaz de atingir uma acurácia de 0,90 para a predição de R_{G1} e 0,98 para a predição de R_{G2}. O modelo não apenas demonstrou uma capacidade de predição sólida, mas também superou ferramentas especializadas, como o software *Nuhertz*, em termos de precisão nos testes realizados. Enquanto o design obtido pelo *Nuhertz* apresentou o primeiro valor de $|S_{11}| < -10$ dB em 28,2GHz, e o último valor de $|S_{11}| < -10$ dB ocorreu em 28,56GHz, o modelo otimizado pela MLP alcançou esses valores em uma frequência de 26,30GHz e 28,05GHz, correspondendo melhor ao intervalo proposto de 26GHz a 28GHz.

Este estudo demonstrou o potencial das redes neurais MLP na otimização do design e análise de filtros interdigitais para ondas milimétricas. As descobertas indicam que, com uma base de dados adequada e uma cuidadosa seleção de hiperparâmetros, é possível alcançar resultados otimizados quando comparados aos de softwares especializados. No futuro, expandir e diversificar a base de dados, explorar combinações de modelos, integrar conhecimentos físicos ao processo de modelagem serão passos cruciais para aprimorar ainda mais a eficácia e a precisão desses modelos. Além disso, comparar este modelo com outras ferramentas de *machine learning* para melhor visualizar qual é o melhor modelo.

Referências

AGGARWAL, C. C. *Neural Networks and Deep Learning: A Textbook*. [S.l.]: Springer International Publishing, 2018. Citado 4 vezes nas páginas 13, 29, 30 e 31.

ALDAYA, I.; CARDINAL, L.; MENZINGER, J.; PENCHEL, R. A.; SANTOS, M. O.; ABBADE, M.; OLIVEIRA, J. A.; SANTOS, M.; SáNCHEZ, G. G. Integration of a* and k-means clustering for star network design in environments considering obstacles. *Anais do XL Simpósio Brasileiro de Telecomunicações e Processamento de Sinais*, 2022. Citado na página 13.

ANALYSIS, C. S. *The Insertion Loss Method of Lumped Element Filter Design*. 2024. Website. About the Author: Solving electromagnetic, Author: A. the, electromagnetic, S., Visit Website: More Content by Cadence System Analysis. Disponível em: <https://resources.system-analysis.cadence.com/blog/msa2021-the-insertion-loss-method-of-lumped-element-filter-design>. Citado na página 19.

ATEYA, A.; MUTHANNA, A.; MAKOLKINA, M.; KOUCHERYAVY, A. Study of 5g services standardization: Specifications and requirements. In: *Proceedings of the International Conference on Ubiquitous Computing and Communications*. [S.l.: s.n.], 2018. p. 1–6. Citado na página 12.

BERNARDO, H. S.; OLIVEIRA, E. M.; PENCHEL, R. A. Análise de geometrias distribuídas de dispositivos de microondas por equivalentes em circuitos elÉtricos. *Tendências e Avanços Científicos Nas Engenharias: Aeronáutica, Aeroespacial, Eletrônica e de Telecomunicações*, p. 50–68, 2023. Disponível em: <https://doi.org/10.51859/amplla.tac393.1323-5>. Citado na página 23.

BRAUN, A. *History of IoT: A Timeline of Development - IoT Tech Trends — iottechtrends.com*. 2024. <https://www.iottechtrends.com/history-of-iot/>. [Accessed 17-06-2024]. Citado na página 11.

DEJEN, A.; RIDWAN, M.; ANGUERA, J.; JAYASINGHE, J. Millimeter wave cellular communication performances and challenges: A survey. *ICST Transactions on Mobile Communications and Applications*, v. 7, n. 21, 2022. Citado na página 12.

DING, B.; QIAN, H.; ZHOU, J. Activation functions and their characteristics in deep neural networks. In: . [S.l.: s.n.], 2018. p. 1836–1841. Citado na página 31.

ERICSSON. *Uncovering the Revenue Opportunity of 5G*. 2020. Ericsson Consumer and Market Insight Report. Disponível em: <https://www.ericsson.com/en/reports-and-papers/consumerlab/reports/harnessing-the-5g-consumer-potential>. Citado 2 vezes nas páginas 11 e 12.

ERICSSON. *Mobile Data Traffic Forecast – Mobility Report*. 2023. Ericsson Mobility Report. Disponível em: <https://www.ericsson.com/en/reports-and-papers/mobility-report/dataforecasts/mobile-traffic-forecast>. Citado 2 vezes nas páginas 11 e 12.

FIORE, J. *11.4: Filter order and Poles*. Libretexts, 2021. Disponível em: <https://eng.libretexts.org/Bookshelves/Electrical_Engineering/Electronics/Operational_Amplifiers_and_Linear_Integrated_Circuits_-_Theory_and_Application_%28Fiore%29/11%3A_Active_Filters/11.04%3A_Section_4->. Citado na página 18.

HONG, J.-S. *Microstrip filters for RF/Microwave Applications*. [S.l.]: Wiley, 2011. Citado 6 vezes nas páginas 15, 16, 17, 23, 24 e 25.

INTELLIGENCE, G. T. *5G: Timeline — verdict.co.uk*. 2024. <https://www.verdict.co.uk/5g-timeline/?cf-view>. [Acessado 17-06-2024]. Citado na página 11.

LIBRETEXTS. *Electronics*. 2024. Acessado em 16 de Maio de 2024. Disponível em: <https://eng.libretexts.org/Bookshelves/Electrical_Engineering/Electronics>. Citado na página 24.

Lé, J. E. G.; OUVRIER-BUFFET, M.; GOMES, L. G.; PENCHEL, R. A.; SERRANO, A. L. C.; REHDER, G. P. Integrated antennas on mnm interposer for the 60 ghz band. *Journal of Microwaves, Optoelectronics and Electromagnetic Applications*, Sociedade Brasileira de Microondas e Optoeletrônica e Sociedade Brasileira de Eletromagnetismo, v. 21, n. 1, p. 184–193, Jan 2022. ISSN 2179-1074. Disponível em: <https://doi.org/10.1590/2179-10742022v21i1253671>. Citado na página 12.

MATTHAEI, G. Interdigital band-pass filters. *IRE Transactions on Microwave Theory and Techniques*, v. 10, n. 6, p. 479–491, 1962. Citado 2 vezes nas páginas 24 e 25.

NAQA, I. E.; MURPHY, M. J. What is machine learning? In: _____. *Machine Learning in Radiation Oncology*. [S.l.: s.n.], 2015. p. 3–11. Citado na página 13.

NIELSEN, M. A. *Neural Networks and Deep Learning*. [S.l.]: Determination Press, 2015. Citado 2 vezes nas páginas 28 e 29.

NOKIA. *5G Report: The Value of 5G Services and the Opportunity for CSPs*. 2023. Nokia Thought Leadership Report. Disponível em: <https://www.nokia.com/thought-leadership/research/5g-consumer-market-research/>. Citado na página 12.

OGATA, K. *Engenharia de Controle moderno*. [S.l.]: Pearson Prentice Hall, 2011. Citado na página 17.

PENCHEL, R. A. *Notas de aula - Dispositivos e Circuitos de RF*. 2024. Citado 4 vezes nas páginas 16, 19, 20 e 21.

POPOVSKI, P.; TRILLINGSGAARD, K. F.; SIMEONE, O.; DURISI, G. 5g wireless network slicing for embb, urllc, and mmtc: A communication-theoretic view. *IEEE Access*, v. 6, p. 55765–55779, 2018. Citado na página 12.

POZAR, D. M. *Microwave engineering*. 4. ed. [S.l.]: Wiley, 2012. Citado 7 vezes nas páginas 15, 16, 19, 20, 21, 23 e 24.

QUEIROZ, M. M.; WAMBA, S. F.; JABBOUR, C. J. C.; JABBOUR, A. B. Lopes de S.; MACHADO, M. C. Adoption of industry 4.0 technologies by organizations: a maturity levels perspective. *Annals of Operations Research*, Springer Science and Business Media LLC, out. 2022. ISSN 1572-9338. Disponível em: <http://dx.doi.org/10.1007/s10479-022-05006-6>. Citado na página 11.

SCIKIT-LEARN. *scikit-learn: machine learning in Python*. 2024. <https://scikit-learn.org/stable/>. [Accessed 14-06-2024]. Citado 2 vezes nas páginas 34 e 35.

SHARMA, S.; SHARMA, S.; ATHAIYA, A. Activation functions in neural networks. *International Journal of Engineering Applied Sciences and Technology*, v. 4, n. 12, p. 310–316, 2020. Citado 2 vezes nas páginas 30 e 32.

TERVEN, J.; CORDOVA-ESPARZA, D. M.; RAMIREZ-PEDRAZA, A.; CHAVEZ-URBIOLA, E. A. *Loss Functions and Metrics in Deep Learning*. 2023. Citado na página 30.

UNIVERSITY, M. *Butterworth and Chebyshev Filters*. MEHRAN UNIVERSITY OF ENGINEERING AND TECHNOLOGY, JAMSHORO INSTITUTE OF INFORMATION AND COMMUNICATION TECHNOLOGIES, [s.d.]. Disponível em: <https://tl.muet.edu.pk/~aftab/docs/adsp_lab7.pdf>. Citado na página 18.

VADAPALLI, P. *Biological Neural Network: Importance, Components & Comparison*. 2022. UpGrad Blog. Disponível em: <https://www.upgrad.com/blog/biological-neural-network/>. Citado na página 28.