



**INSTITUTO FEDERAL DE EDUCAÇÃO, CIÊNCIA E TECNOLOGIA
DO CEARÁ
IFCE CAMPUS FORTALEZA
PROGRAMA PÓS-GRADUAÇÃO EM ENGENHARIA DE
TELECOMUNICAÇÕES**

ZÉLIA MARIA CARVALHO LIMA

**MODELO DE SISTEMAS DE COMUNICAÇÕES ÓPTICAS NO
ESPAÇO LIVRE UTILIZANDO APRENDIZADO DE MÁQUINA**

**FORTALEZA
2023**

ZÉLIA MARIA CARVALHO LIMA

MODELO DE SISTEMAS DE COMUNICAÇÕES ÓPTICAS NO ESPAÇO
LIVRE UTILIZANDO APRENDIZADO DE MÁQUINA

Dissertação apresentada à Coordenação de Pós-Graduação em Engenharia de Telecomunicações do Instituto Federal do Ceará (IFCE) - Campus Fortaleza, como requisito parcial para obtenção do grau de Mestre em Engenharia de Telecomunicações. Área de concentração: Micro-Ondas e Óptica Integrada.

Orientador: Prof. Dr. Glendo de Freitas Guimarães.

Co-orientador: Dr. Yosdan Martínez Camejo

FORTALEZA
2023

Dados Internacionais de Catalogação na Publicação
Instituto Federal do Ceará - IFCE
Sistema de Bibliotecas - SIBI
Ficha catalográfica elaborada pelo SIBI/IFCE, com os dados fornecidos pelo(a) autor(a)

LIMA, ZELIA MARIA CARVALHO.

MODELO DE SISTEMAS DE COMUNICAÇÕES ÓPTICAS NO ESPAÇO LIVRE UTILIZANDO
APRENDIZADO DE MÁQUINA / ZELIA MARIA CARVALHO LIMA. - 2023.

100 f. : il. color.

Dissertação (Mestrado) - Instituto Federal do Ceará, Mestrado em Engenharia de Telecomunicações,
Campus Fortaleza, 2023.

Orientação: Prof. Dr. Glendo de Freitas Guimarães.

Coorientação: Prof. Dr. Yosdan Martínez Camejo.

1. Free Space Optics (FSO). 2. Machine Learning. 3. Optisystem. I. Título.

CDD 621.382



INSTITUTO FEDERAL DE EDUCAÇÃO, CIÊNCIA E TECNOLOGIA DO CEARÁ
Av. 13 de maio, 2081 , - Bairro Benfica - CEP 60040-531 - Fortaleza - CE - www.ifce.edu.br

AVALIAÇÃO

Processo: 23256.001230/2023-94

Interessado: Zélia Maria Carvalho Lima

ZÉLIA MARIA CARVALHO LIMA

MODELO DE SISTEMAS DE COMUNICAÇÕES ÓPTICAS NO ESPAÇO LIVRE UTILIZANDO APRENDIZADO DE MÁQUINA

Dissertação apresentada ao curso de Pós-Graduação em Engenharia de Telecomunicações do Instituto Federal de Educação, Ciência e Tecnologia do Ceará (IFCE) - *Campus Fortaleza*, como requisito parcial para obtenção do Título de Mestre em Engenharia de Telecomunicações.

Área de concentração: Sistemas de Telecomunicações

Aprovado em: 17 / 02 / 2023

BANCA EXAMINADORA

Prof. Dr. **Glendo de Freitas Guimarães** - orientador
Instituto Federal de Educação, Ciência e Tecnologia do Ceará - IFCE

Prof. Dr. **Yosdan Martínez Camejo** - coorientador
Instituto Federal de Educação, Ciência e Tecnologia do Ceará - IFCE

Prof. Dr. **Danilo Sousa Rocha**
Instituto Federal de Educação, Ciência e Tecnologia do Ceará - IFCE

Prof. Dr. **José Rubens Rodrigues de Sousa**

Universidade de Fortaleza - UNIFOR



Documento assinado eletronicamente por **Danilo Sousa Rocha, Professor do Ensino Básico, Técnico e Tecnológico**, em 08/03/2023, às 08:46, com fundamento no art. 6º, § 1º, do [Decreto nº 8.539, de 8 de outubro de 2015](#).



Documento assinado eletronicamente por **Glendo de Freitas Guimaraes, Coordenador(a) do PPGET**, em 08/03/2023, às 08:50, com fundamento no art. 6º, § 1º, do [Decreto nº 8.539, de 8 de outubro de 2015](#).



Documento assinado eletronicamente por **JOSÉ RUBENS RODRIGUES DE SOUSA, Usuário Externo**, em 08/03/2023, às 14:59, com fundamento no art. 6º, § 1º, do [Decreto nº 8.539, de 8 de outubro de 2015](#).



Documento assinado eletronicamente por **Yosdan Martínez Camejo, Usuário Externo**, em 08/03/2023, às 19:40, com fundamento no art. 6º, § 1º, do [Decreto nº 8.539, de 8 de outubro de 2015](#).



A autenticidade do documento pode ser conferida no site https://sei.ifce.edu.br/sei/controlador_externo.php?acao=documento_conferir&id_orgao_acesso_externo=0 informando o código verificador **4553807** e o código CRC **4AF11A39**.

Este trabalho é dedicado a Deus e aos meus pais, por terem me aconselhado com sabedoria ao longo dos anos. Em especial pela presença de minha amada mãe, em quase toda a minha vida, me dando todo o suporte necessário.

AGRADECIMENTOS

Agradeço a Deus em primeiro lugar, que me ajudou a superar os obstáculos, me dando forças para concluir este trabalho.

A toda minha família que sempre me incetivou e sempre esteve ao meu lado nos momentos mais difíceis. Sem o apoio deles, o sucesso deste trabalho não seria possível. Ao Instituto Federal do Ceará (IFCE), que em minha trajetória proporcionou um ambiente favorável à troca de saberes e à construção do conhecimento. Ao meu orientador Prof. Glendo de Freitas Guimarães, por ter acreditado em mim desde o início, quando repassei minha vontade de cursar o mestrado, por ser uma referência tanto profissional como humana, pelas críticas construtivas que melhoraram a estrutura desse trabalho, pelas sugestões e questionamentos levantados. Agradeço também ao meu coorientador Yosdan Martínez Camejo pelo estímulo e cooperação na formatação deste trabalho de mestrado.

Aos colegas do IFCE, Rogean Costa Silva pelo apoio valioso no estudo de Inteligência Artificial (IA) e a Valeska Fernandes que me apoiou desde o início do mestrado, estudando e incentivando durante todo o curso, sempre compartilhando conhecimentos.

A todos os meus amigos de empresa pelas colaborações e compartilhamento de conhecimentos, em especial ao colega Francis Marconcini que me auxiliou diretamente no estudo de IA, transferindo as experiências adquiridas no mestrado concluído recentemente, sendo sempre prestativo e disponível.

Aos participantes da banca examinadora e a todos os professores que me receberam em suas salas, compartilharam conhecimento e contribuíram para o meu desenvolvimento profissional.

“Não vos amoldeis às estruturas deste mundo,
mas transformai-vos pela renovação da mente,a
fim de distinguir qual é a vontade de Deus: o que
é bom, o que Lhe é agradável, o que é perfeito.”

(Bíblia Sagrada, Romanos 12, 2)

RESUMO

Considerando a escassez de dados na área de comunicações ópticas no espaço livre para a análise de multiparâmetros, este estudo teve como objetivo utilizar uma metodologia de Inteligência Artificial (IA), visando esboçar este modelo de transmissão de dados. Dois cenários de simulação foram concebidos, o primeiro cenário com quatro parâmetros de entrada: distância, atenuação do canal no espaço livre óptico, ou *Free Space Optics (FSO)*, ganho dos amplificadores no transmissor e receptor, e o segundo cenário com seis entradas, onde foram acrescentados o diâmetro de abertura do receptor e divergência do feixe. Com esta base de dados coletada, realizou-se a modelagem dos dados através de algoritmos de IA. Com o intuito de prever um valor numérico específico, algoritmos regressores foram utilizados e a efetividade do modelo treinado foi medida com base na estatística, coeficiente de determinação (R^2) e *Mean Absolute Error* (MAE). O parâmetro de saída analisado foi o máximo fator Q. Com base nos resultados foi constatado que, para o primeiro cenário, os modelos *Boosted Decision Tree Regression* (BDTR) e *AdaBoost + Decision Tree Regression* (ADTR) apresentaram uma estimativa excelente para o fator Q, com R^2 em torno de 99,995%, e, para o segundo cenário, os algoritmos *Gradient Boosting Regressor* (GBR) e *Histogram Gradient Boosting Regression* (HGBR) apresentaram ótima predição, um pouco pior que o primeiro cenário, porém com R^2 em torno de 99,8%. O conjunto de dados gerados por simulação foi comparado com os valores preditos nos *Machine Learning Algorithms* (MLAs) graficamente, confirmando a eficácia da predição do valor alvo. Destaca-se no segundo cenário o estudo de um link FSO em condições climáticas adversas com a melhoria de desempenho através da regulação dos parâmetros internos do FSO, utilizando as técnicas de ML para otimizar o desempenho da rede FSO e minimizar os custos na implementação destas redes. Esta experiência não foi encontrada na pesquisa realizada na literatura.

Palavras-chave: *Free Space Optics* (FSO). Aprendizagem de Máquina. *Optisystem*

ABSTRACT

Considering the lack of data in the free-space optical communications area for multi-parameter analysis, the study aimed to use an Artificial Intelligence (AI) methodology to sketch this data transmission model. We designed two simulation scenarios: the first scenario with four input parameters: distance, attenuation of the Free Space Optics (FSO) channel, gain of the transmitter and receiver amplifiers, and the second scenario with six inputs, where the receiver aperture diameter and beam divergence were added. We performed the data modeling using AI algorithms with this database collected. In order to predict a specific numerical value, we used regression algorithms, and we measured the trained model effectiveness based on the coefficient of determination (R^2) and Mean Absolute Error (MAE). We analyzed the maximum Q-factor as the output parameter. Based on the results, we could observe that, for the first scenario, the *Boosted Decision Tree Regression (BDTR)* and *AdaBoost + Decision Tree Regression (ADTR)* models presented an excellent estimate for the Q factor, with R^2 around 99.995%, and, for the second scenario, the algorithms *Gradient Boosting Regressor (GBR)* and *Histogram Gradient Boosting Regression (HGBR)* showed excellent prediction, worse than the first scenario, but with R^2 around 99.8%. The simulation-generated dataset was compared with the predicted values in the Machine Learning Algorithms (MLAs) graphically, confirming the effectiveness of the target value prediction. Noteworthy in the second scenario is the study of an FSO link in adverse weather conditions with performance improvement by tuning the internal parameters of the FSO, using ML techniques to optimize the performance of the FSO network and minimize the costs of the implementation of these networks. We did not find this experiment in the research conducted in the literature.

Keywords: Free-Space Optics (FSO). Machine Learning. Optisystem.

LISTA DE ILUSTRAÇÕES

Figura 1 -Camadas na atmosfera da Terra.....	25
Figura 2 - Free Space Optical (FSO) – Sistema de Comunicação.....	26
Figura 3 - Serviços suportados pela rede 5G.....	27
Figura 4 - Distribuição de intensidade de um feixe gaussiano (seção transversal)	28
Figura 5 - Hermite - Modos de feixe gaussiano (aproximado).....	28
Figura 6 - Parâmetros do feixe divergente: ω_0 é a largura da cintura do feixe ωx , profundidade do foco; θ é a dispersão angular (ângulo de divergência após a cintura).....	29
Figura 7 - Distribuição gaussiana de seção transversal circular e potência irregular distri- buição não gaussiana.	30
Figura 8 - Comprimentos de onda do Espectro Solar.....	33
Figura 9 - O espalhamento de Rayleigh é bidirecional (A), enquanto o espalhamento de Mie (B) e Geométrico (C) é principalmente unidirecional	34
Figura 10 - Efeito da miragem devido a cintilação	36
Figura 11 - Elementos que compõem um canal de comunicação FSO	38
Figura 12 - Imagem esquemática dos processos de emissão.....	39
Figura 13 - Codificação RZ e NRZ	40
Figura 14 - Princípio de funcionamento do Modulador Mach-Zhender	41
Figura 15 - Estrutura do Amplificador SOA	43
Figura 16 - Diferentes elementos podem ser usados em DFAs para amplificar bandas ópticas correspondentes	45
Figura 17 - Níveis de excitação para o érbio	46
Figura 18 - Bloco semiconductor usado como fotodetector	48
Figura 19 - Temas do estudo de IA	53
Figura 20 - Tipos de modelos de regressão utilizados no aprendizado de máquina	55
Figura 21 - Diagrama de uma árvore de decisão	58
Figura 22 - Demonstração de algoritmo baseado em histograma	60
Figura 23 - Sistema de comunicação FSO simulado no OptiSystem.....	62
Figura 24 - Representação esquemática das entradas e saídas na simulação do canal FSO	63
Figura 25a- Mapa de calor para o cenário 1 Pearson	67
Figura 25b - Mapa de calor para o cenário 1 Spearman.....	68

Figura 25c - Mapa de calor para o cenário 2 Pearson	69
Figura 25d - Mapa de calor para o cenário 2 Spearman.....	69
Figura 26 - Mapa de calor comparando Pearson vs Spearman para atenuação	71
Figura 27 - Atenuação na rede FSO	72
Figura 28 - Parâmetro interno FSO: Diâmetro de abertura receptor	73
Figura 29 - Parâmetro interno FSO: Divergência do feixe.....	74
Figura 30 - Gráfico base de treinamento vs coeficiente de determinação para os algoritmos de IA: BDTR, DFR, GBR, HGBR e AdaBoost + DRT – Cenário 1	78
Figura 31 - Gráfico de barras do coeficiente R ² para todos algoritmos por percentual de treinamento – Cenário 1	78
Figura 32 - Gráfico base de treinamento vs coeficiente de determinação R ² para os algoritmos de IA: BDTR, DFR, GBR, HGBR e AdaBoost + DRT – Cenário 2.....	79
Figura 33 - Comparação entre os desempenhos de R ² para o algoritmo GBR – Cenários 1 e 2	80
Figura 34 - Comparação entre os desempenhos de R ² para o algoritmo HGBR – Cenários 1 e 2	81
Figura 35 - Comparação entre os desempenhos de R ² para o algoritmo BDTR – Cenários 1 e 2	81
Figura 36 - Comparação entre os desempenhos de R ² para o algoritmo DFR – Cenários 1 e 2	82
Figura 37 - Comparação entre os desempenhos de R ² para o algoritmo AdaBoost + DTR – Cenários 1 e 2	82
Figura 38 - Gráfico de barras do coeficiente R ² para todos algoritmos por percentual de treinamento – Cenário 2	83
Figura 39a - Comparação real x predito – BDTR para 1 km	85
Figura 39b - Comparação real x predito – BDTR para 5 km	85
Figura 39c - Comparação real x predito – BDTR para 10 km	86
Figura 40a - Comparação real x predito – AdaBoost + DTR para 1 km.....	87
Figura 40b - Comparação real x predito – AdaBoost + DTR para 5 km.....	87
Figura 40c - Comparação real x predito – AdaBoost + DTR para 10 km.....	88
Figura 41a - Comparação entre real e predito - HGBR (km 1 – DF=1/DAR=25).....	89
Figura 41b - Comparação entre real e predito - HGBR (km 1 – DF=3/DAR=15).....	90
Figura 42a - Comparação entre real e predito - HGBR (km 10 – DF=1/DAR=25).....	91
Figura 42b - Comparação entre eal e predito - HGBR (km 10 – DF=3/DAR=15).....	91

Figura 43a - Comparação entre real e predito - GBR (km 1 – DF=1/DAR=25).....	93
Figura 43b - Comparação entre real e predito - GBR (km 1 – DF=3/DAR=15).....	93
Figura 44a - Comparação entre real e predito - GBR (km 10 – DF=1/DAR=25).....	94
Figura 44b - Comparação entre real e predito - GBR (km 10 – DF=3/DAR=15).....	94
Figura 45 - Comparação: MAE para 80% de treinamento	95

LISTA DE TABELAS

Tabela 1 - Caracterização do nevoeiro com base em dados empíricos.	32
Tabela 2 - Características LED x LASER	39
Tabela 3 - Características de fotodiodos PIN comuns.....	50
Tabela 4 - Interpretando os valores do coeficiente de Pearson	61
Tabela 5 - Valores de entradas utilizados na simulação do canal FSO	63
Tabela 6 - Coeficientes de atenuação de acordo com as condições climáticas.	64
Tabela 7 - Valores das variáveis de entradas fixados.....	72
Tabela 8 - Valores para a distância 2 km.....	72
Tabela 9 - Valores para a distância 3 km.....	73
Tabela 10 - Valores das variáveis de entradas fixados.....	74
Tabela 11 - Valores fixados.....	75
Tabela 12 - Condições de coleta dos gráficos para as novas entradas	88
Tabela 13 - Percentual de acerto depende da base de dados	92

LISTA DE ABREVIATURAS E SIGLAS

ADTR	do inglês, adaboost + decision tree regression
ASE	do inglês, amplified spontaneous emission
BDTR	do inglês, boosted decision tree regression
BER	do inglês, bit error rate
BWIS	do inglês, beam wander induced scintillation
CNN	do inglês, convolutional neural network
CW	do inglês, continuous wave
DFR	do inglês, decision forest regression
DL	do inglês, deep learning
DNN	do inglês, deep neural network
DTR	do inglês, decision tree regression
DWDM	do inglês, dense wavelength division multiplexing
EDFA	do inglês, erbium doped fiber amplifier
FSO	do inglês, free space optics
FWHM	do inglês, full-width half-maximum
GSL	do inglês, geometrical spreading loss
HGBR	do inglês, histogram gradient boosting regression
IA	do inglês,
IDE	do inglês, integrated development environment
Laser	do inglês, light amplification by stimulated emission of radiation
LED	do inglês, light emitting diode
LISP	do inglês, list processing
MAE	do inglês, mean absolute error
MLA	do inglês, machine learning algorithms
MZ	do inglês, mach-zehnder
NRZ	do inglês, no return to zero
OOK	do inglês, on-off keying
QOT	do inglês, quality of transmission
RF	do português, radiofrequência
RZ	do inglês, return to zero
SNR	do inglês, signal to noise ratio
SBR	do inglês, solar background radiation

SVM	do inglês, support vector machine
UAV	do inglês, unmanned aerial vehicle
WDM	do inglês, wavelengt division multiplexing

LISTA DE SÍMBOLOS

Γ	Letra grega Gama
Λ	Lambda
ζ	Letra grega minúscula zeta
$\in$	Pertence
n	Índice de refração do meio
c	Velocidade da luz no vácuo
α	Coefficiente de atenuação
σ^2	Variância de fotocorrente
Γ	Fator de confinamento
η	Eficiência quântica

SUMÁRIO

1	INTRODUÇÃO	21
1.1	OBJETIVOS	22
1.1.1	<i>Objetivo Geral</i>	<i>22</i>
1.1.2	<i>Objetivos específicos</i>	<i>22</i>
1.2	ORGANIZAÇÃO DA DISSERTAÇÃO.....	23
2	FUNDAMENTAÇÃO TEÓRICA	25
2.1	PROPAGAÇÃO DE LUZ EM MEIOS NÃO GUIADOS	25
2.2	FREE SPACE OPTICAL (FSO)	26
2.3	CARACTERÍSTICAS DO FEIXE DO LASER.....	27
2.3.1	<i>Comprimento de Onda</i>	<i>27</i>
2.3.2	<i>Perfil e Modos de Feixe.....</i>	<i>27</i>
2.3.3	<i>Divergência do Feixe.....</i>	<i>28</i>
2.4	EFEITOS ATMOSFÉRICOS NOS SINAIS ÓPTICOS	31
2.4.1	<i>Eletricidade Atmosférica.....</i>	<i>31</i>
2.4.2	<i>Névoa.....</i>	<i>31</i>
2.4.3	<i>Poluição do Ar.....</i>	<i>32</i>
2.4.4	<i>Chuva.....</i>	<i>32</i>
2.4.5	<i>Neve.....</i>	<i>33</i>
2.4.6	<i>Interferência Solar</i>	<i>33</i>
2.4.7	<i>Espalhamento</i>	<i>34</i>
2.4.8	<i>Cintilação.....</i>	<i>35</i>
2.4.9	<i>Vagueamento do Feixe</i>	<i>37</i>
2.5	ELEMENTOS BÁSICOS FSO	37
2.6	TRANSMISSOR ÓPTICO.....	38
2.6.1	<i>Emissão Espontânea e Estimulada.....</i>	<i>38</i>
2.6.2	<i>Laser CW.....</i>	<i>39</i>
2.6.3	<i>Gerador de pulsos.....</i>	<i>40</i>
2.6.4	<i>Modulador Mach-Zehnder</i>	<i>40</i>
2.7	AMPLIFICADOR ÓPTICO.....	41
2.7.1	<i>Amplificadores ópticos semicondutores.....</i>	<i>43</i>
2.7.2	<i>Amplificadores de Fibra Óptica.....</i>	<i>44</i>
2.7.2.1 -	<i>Amplificadores de Fibra Dopada por Érbio</i>	<i>45</i>

2.8	RECEPTOR ÓPTICO	47
2.8.1	<i>Fotossensibilidade/Responsividade e Eficiência Quântica</i>	48
2.8.2	<i>Largura de banda , tempo de subida e corrente no escuro</i>	49
2.8.3	<i>Fotodetector PIN</i>	50
2.8.4	<i>Relação Sinal Ruído</i>	50
2.8.5	<i>Taxa de erro de bit - BER</i>	51
2.9	ESTUDO DE IA	52
2.9.1	<i>História da inteligência artificial</i>	52
2.9.2	<i>Aprendizado de Máquina</i>	53
2.9.3	<i>Aprendizado Supervisionado</i>	54
2.9.4	<i>Aprendizado Não Supervisionado</i>	55
2.10	ALGORITMOS DE INTELIGÊNCIA ARTIFICIAL	56
2.10.1	<i>Boosting</i>	56
2.10.2	<i>Adaboost + Decision Tree Regression (Adaboost + DTR)</i>	56
2.10.3	<i>Boosted Decision Tree Regression (BDTR)</i>	57
2.10.4	<i>Decision Forest Regression (DFR)</i>	58
2.10.5	<i>Gradient Boosting Regression (GBR)</i>	59
2.10.6	<i>Histogram Gradient Boosting Regression (HGBR)</i>	59
2.11	ESTUDO DA CORRELAÇÃO ENTRE ATRIBUTOS	60
2.11.1	<i>Coefficiente de Pearson</i>	60
2.11.2	<i>Coefficiente de Spearman</i>	61
3	METODOLOGIA	62
3.1	SIMULAÇÃO DE UM CANAL (FSO)	62
3.2	INTELIGÊNCIA ARTIFICIAL	64
4	RESULTADOS E DISCUSSÃO	66
4.1	ANÁLISE DE CORRELAÇÃO	66
4.1.1	<i>Mapa de calor para Pearson e Spearman nos cenários 1 e 2</i>	66
4.1.2	<i>Análise de Pearson e Spearman para atenuação</i>	70
4.2	ANÁLISE DOS DADOS SINTÉTICOS DO OPTISYSTEM	71
4.2.1	<i>Análise de atenuação no canal FSO</i>	71
4.2.2	<i>Análise dos parâmetros internos do FSO: diâmetro de abertura do receptor e divergência do feixe</i>	73
4.3	TREINAMENTO DA BASE DE DADOS ATRAVÉS DA IA	75
4.3.1	<i>Treinamento da IA – Cenário 1</i>	77

4.3.2	Treinamento da IA – Cenário 2	79
4.4	COMPARAÇÃO ENTRE REAL X PREDITO	84
4.4.1	Cenário 1	84
4.4.1.1 -	<i>Algoritmo Boosted Decision Tree Regression (BDTR)</i>	84
4.4.1.2 -	<i>Algoritmo AdaBoost + Decision Tree Regression (AdaBoost + DTR)</i>	86
4.4.2	Cenário 2	88
4.4.2.1 -	<i>Histogram Gradient Boosting Regression (HGBR)</i>	89
4.4.2.2 -	<i>Gradient Boosting Regression (GBR)</i>	92
4.5	COMPARAÇÃO ENTRE OS MEAN ABSOLUTE ERROR (MAE).....	95
5	CONCLUSÕES	97
6	REFERÊNCIAS	98

1 INTRODUÇÃO

A tecnologia de comunicação FSO, é uma tecnologia emergente de banda larga ponto a ponto de alta velocidade, categorizada em comunicações ópticas sem fio e usa a luz como meio de transmissão entre os transceptores. Os sistemas FSO podem ser implantados em uma distância de link de poucos quilômetros. Nos últimos anos, a tecnologia FSO recebeu maior atenção dos pesquisadores e aceitação em redes de campus empresariais. Como vantagens a tecnologia FSO fornece alta largura de banda, alta confidencialidade, alta velocidade, baixo investimento inicial, destituído de requisitos de licença de espectro, baixo consumo de energia, imune a interferências de radiofrequência, facilidade de instalação em comparação com a fibra óptica. Por outro lado a tecnologia FSO em diferentes condições climáticas, como neblina, névoa e chuva, que causam partículas suspensas no ar, interagem com os fótons do comprimento de onda óptico, ocasionando uma dispersão do sinal. Estes fenômenos causam atenuação e reduzem a disponibilidade geral dos sistemas FSO. Além disso, o canal de comunicação FSO é afetado pela cintilação devido ao efeito da turbulência térmica. Técnicas de diversidade são consideradas para melhorar o desempenho da capacidade dos canais FSO sob forte turbulência.

Para o efetivo estudo das redes ópticas, faz-se necessário a simulação numérica, tanto para redes ópticas com fibra quanto para o espaço livre. Tem-se algumas limitações para a implementação de tais cenários, pois estas simulações podem ser demoradas, o custo temporal e computacional é complexo, e mesmo assim não contemplariam todos os cenários possíveis. Desta forma, uma das soluções para resolver este impasse seria utilizar algoritmos que possam fornecer a informação de entrada/saída sem necessidade de testar todos os cenários possíveis. O uso de algoritmos de inteligência artificial tem crescido em várias áreas de estudo, inclusive na área de comunicações ópticas. Desta forma este trabalho visa elaborar um modelo de rede óptica com FSO utilizando algoritmos de IA.

Em relação a análise do desempenho de uma rede FSO através de algoritmos de aprendizado de máquina, são encontradas diferentes abordagens na literatura. Para prover o conjunto de dados necessário para a aprendizagem dos algoritmos encontramos simulações reais em laboratório (Rahman et al., 2018), simulações pelo *Optisystem* (ALGEDIR, et al., 2020), em seguida o desempenho do sistema foi analisado pela previsão do fator Q com algoritmos de regressão ou classificação, para os dois trabalhos da literatura citados acima.

Outra abordagem encontrada foi o estudo de um link FSO em seis condições climáticas diferentes, mantendo a potência da fonte do laser no mínimo valor possível para superar o obstáculo da turbulência atmosférica em área plana. Os diâmetros das aberturas de transmissão e recepção do link FSO são variados para obter fator Q superior e diagrama de olho aceitável para um comprimento de onda específico de 1550nm usando potência de 0,1mW para uma distância fixa do link de 500m. Foi verificado que, sob diferentes condições climáticas e com potência mínima, o fator de qualidade tornava-se muito baixo, ou seja, com baixíssima potência há dificuldade para lidar com atenuações muito altas. Neste momento do estudo foi alterado o diâmetro de abertura do transmissor e do receptor e então calculado o fator Q para cada caso. Foi observado que, à medida que o diâmetro de abertura do transmissor aumentava, o fator de qualidade reduzia, e, à medida que o diâmetro de abertura do receptor aumentava, o fator de qualidade aumentava. Gráficos que comprovaram esta afirmação foram apresentados. Desta forma, com mais pesquisa e desenvolvimento, pode-se transmitir facilmente sinais de baixa potência, considerando as condições meteorológicas para melhorar o uso da óptica do espaço livre (FSO). Para a simulação da análise de desempenho foi utilizado também o *Optisystem 7.0*. Neste artigo não foi utilizada aprendizagem de máquina. (NANDI, et al., 2019).

1.1 OBJETIVOS

Levando em conta o que foi explanado até o momento, apresentam-se os objetivos pretendidos neste trabalho.

1.1.1 Objetivo Geral

O objetivo desta dissertação de mestrado é melhorar o desempenho de uma rede FSO utilizando algoritmos de inteligência artificial, visando simplificar a análise dos parâmetros resultando na melhoria da rede.

1.1.2 Objetivos específicos

- Avaliar a possibilidade do uso de técnicas de inteligência artificial para modelagem de uma rede de comunicações ópticas no espaço livre, utilizando FSO, com foco em simplificar a análise de multi-parâmetros e como consequência melhorar o desempenho

dessa rede.

- Testar uma rede óptica no espaço livre, utilizando o software Optisystem e avaliar os parâmetros de entrada: distância, atenuação do canal óptico FSO, ganho dos amplificadores no transmissor e receptor, para o primeiro cenário testado. Para o segundo cenário acrescenta-se o diâmetro de abertura do receptor e divergência do feixe e correlacionando ambos os cenários à mesma saída, o fator Q.
- Avaliar os melhores algoritmos de inteligência artificial para correlacionar parâmetros de entrada com parâmetros de saída para os dois cenários citados anteriormente.
- Comparar os valores simulados no *Optisystem* com os resultados da IA, visando testar a eficiência dos algoritmos.

1.2 ORGANIZAÇÃO DA DISSERTAÇÃO

Esta dissertação está estruturada em 5 capítulos na seguinte sequência: Introdução, Fundamentação Teórica, Metodologia, Resultados e Conclusão.

No Capítulo 2 foi discutido todo embasamento teórico que compreende a rede óptica no espaço livre, detalhando a propagação de luz em meios não guiados, as características do feixe de luz, os diversos efeitos atmosféricos para os sinais ópticos e as áreas de atuação do FSO. Também será apresentada a base conceitual dos transmissores e receptores ópticos, detalhando seus componentes, dando ênfase aos amplificadores utilizados para FSO. Foram detalhados também os algoritmos de IA, e o estudo da correlação entre as variáveis de entrada.

Em seguida, no Capítulo 3, foi montada uma rede FSO com o software OptiSystem. Nesta rede foram realizados dois cenários de simulações: o primeiro com quatro entradas de dados e o segundo com seis entradas de dados, onde foram acrescentados parâmetros internos do FSO. Como saída de dados foi levantados o Máximo Q Factor. Na sequência foi modelada esta rede através de algoritmos linguagem de máquina.

No Capítulo 4 são apresentados os resultados em forma de gráficos e tabelas, relativo ao fator de qualidade de serviço, obtido das implementações no Optisystem da rede óptica utilizando FSO. Gráficos dos algoritmos de inteligência artificial que foram utilizados são apresentados, com a análise das estatísticas R2 e do MAE (Mean Absolute Error), para a medição da eficiência. Foram

visualizadas as diferenças entre o valor real e o valor predito para os algoritmos que tiveram melhor resultado.

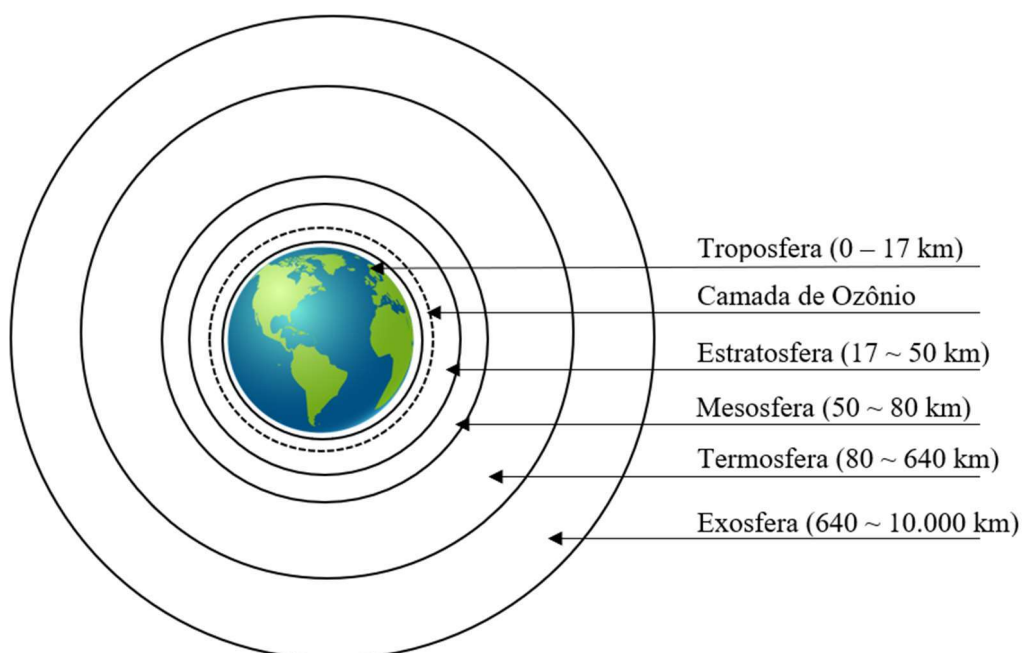
O Capítulo 5 descreve a conclusão geral, ressaltando as principais contribuições deixadas durante este trabalho de mestrado, sugestões para trabalhos futuros são apontadas.

2 FUNDAMENTAÇÃO TEÓRICA

2.1 PROPAGAÇÃO DE LUZ EM MEIOS NÃO GUIADOS

A comunicação óptica de espaço livre (FSO) é encontrada em uma variedade de aplicações em telecomunicações. Em aplicações terrestres, a tecnologia FSO emprega um infravermelho modulado irradiando através da atmosfera, mais particularmente através da troposfera (Figura 1). Nas aplicações no espaço, feixes de laser são usados para estabelecer links intersatélites (ISL). Para qualquer aplicação que utilize FSO, na troposfera ou no espaço, existem um conjunto de diferentes desafios. No espaço a distância é muito longa e os receptores de satélite podem ou não suportar o sol, enquanto na troposfera o meio não é estável. Um terceiro tipo de aplicação compreende um nó FSO estacionário se comunicando com um nó em movimento, aplicação na rede 5G para os UAVs (*Unmanned Aerial Vehicle*), como exemplo. Como tal, a engenharia FSO requer interdisciplinaridade e experiência, para que a rede FSO final forneça links de alta taxa de dados com confiabilidade e com o desempenho esperado em todas ou, pelo menos, na maioria das condições ambientais, tudo isso de forma lucrativa. As duas entidades principais da rede FSO são: o feixe de laser e o meio através do qual ele viaja.

Figura 1 -Camadas na atmosfera da Terra



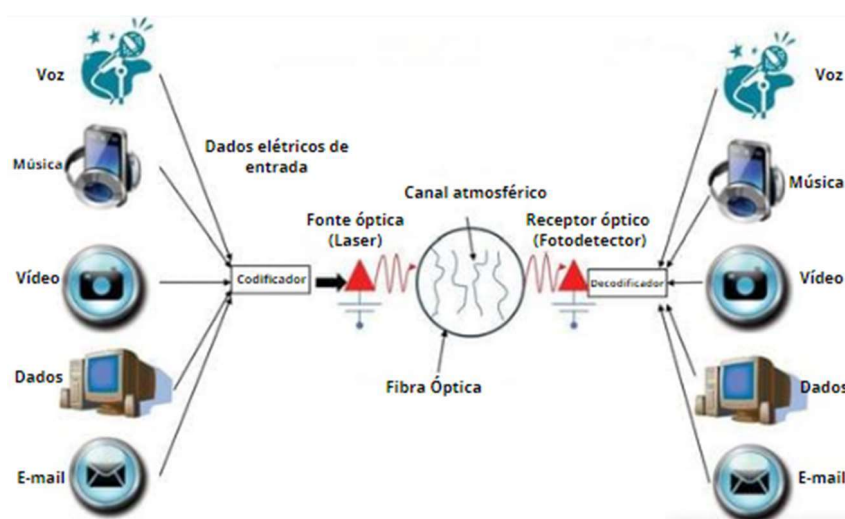
Fonte: autora

2.2 FREE SPACE OPTICAL (FSO)

O FSO chegou para solucionar o requisito de alta largura de banda para internet moderna. A comunicação FSO é a alternativa mais prática para resolver o problema de conectividade de banda larga de última milha, como também como um complemento aos links convencionais de microondas e radiofrequência (RF). Com a chegada do 5G e devido ao aumento contínuo da taxa de dados fornecida pelos links FSO, este será uma das ferramentas mais potentes para lidar com gargalos de conectividade que foram criados em redes de alta velocidade na última década, devido ao grande sucesso e aceitação continuada da Internet. Para a rede 5G a taxa de dados com a internet ultrapassará os limites da infraestrutura existente com aplicativos de alta largura de banda, como serviço de vídeo, piloto automático, manufatura industrial, cidades inteligentes, telemedicina e outros dispositivos portáteis habilitados para rede 5G. Eliminar esses gargalos é crucial para o futuro crescimento e sucesso da sociedade contemporânea da Internet. A elevada largura de banda de acesso de comunicações ópticas e redes de borda serão necessárias para atender a essas novas demandas.

A Figura 2 mostra um diagrama de comunicação FSO com conectividade sem fio para diferentes usuários. Por outro lado, os links para o uso da comunicação usando FSO de longo alcance entre plataformas móveis e aéreas (exemplo UAV, veículos aéreos não tripulados), permanecem como uma área de pesquisa e desenvolvimento nos dias atuais.

Figura 2 - Free Space Optical (FSO) – Sistema de Comunicação



Fonte: MAJUMDAR, 2015

Na Figura 3 observa-se os diferentes serviços suportados pela rede 5G, onde é percebido o quanto o FSO se adequa a essas necessidades. Para enfatizar o quanto a comunicação por canais de FSO é impactante para sistema de comunicação analógico ou digital, para o comprimento de onda de $1\mu\text{m}$ equivale a $3 \times 10^{14}\text{Hz} = 300.000\text{GHz}$, ou seja, um canal de 16 GHz equivale a somente $3.3 \times 10^{-6}\mu\text{m}$.

Figura 3 - Serviços suportados pela rede 5G



Fonte: Modificada da Huawei (2019)

2.3 CARACTERÍSTICAS DO FEIXE DO LASER

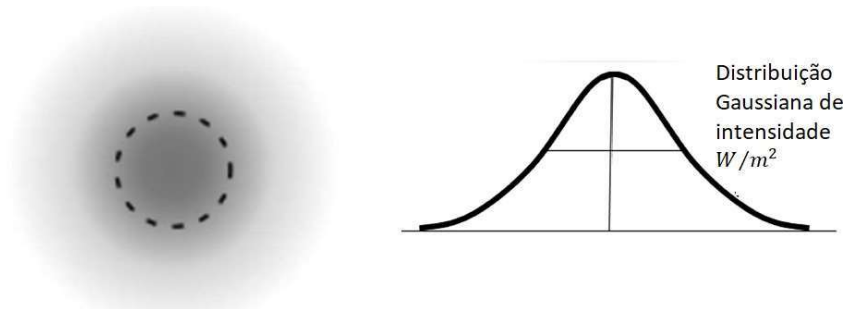
2.3.1 Comprimento de Onda

O feixe de laser usado em links FSO tem um comprimento de onda de 800 nm, 1310 nm e 1550 nm. O mais utilizado dos três é o de 1550 nm, pois suporta altos níveis de potência, taxa de dados em Gb/s, comprimentos de link mais longos e também suporta a tecnologia WDM - multiplexação por divisão de comprimento de onda.

2.3.2 Perfil e Modos de Feixe

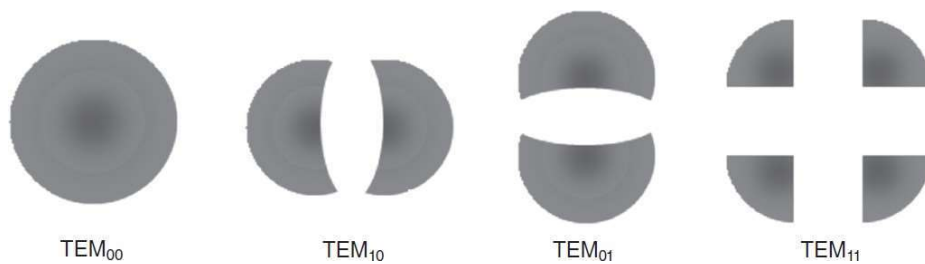
É chamada de irradiância, ou intensidade óptica, a magnitude usada para descrever a energia por unidade de superfície de um feixe luminoso (SALEH, 1991). À medida que o feixe de laser sai do dispositivo, a distribuição de intensidade ao longo do espaço livre não é uniforme e pode ter uma distribuição Gaussiana, neste caso o feixe também é denominado “Gaussiano”. Normalmente os lasers que emitem feixe com uma distribuição Gaussiana pura estão operando no modo transversal eletromagnético fundamental (TEM_{00}), Figura 4.

Figura 4 - Distribuição de intensidade de um feixe gaussiano (seção transversal)



Fonte: KARTALOPOULOS, 2011

Figura 5 - Hermite - Modos de feixe gaussiano (aproximado)



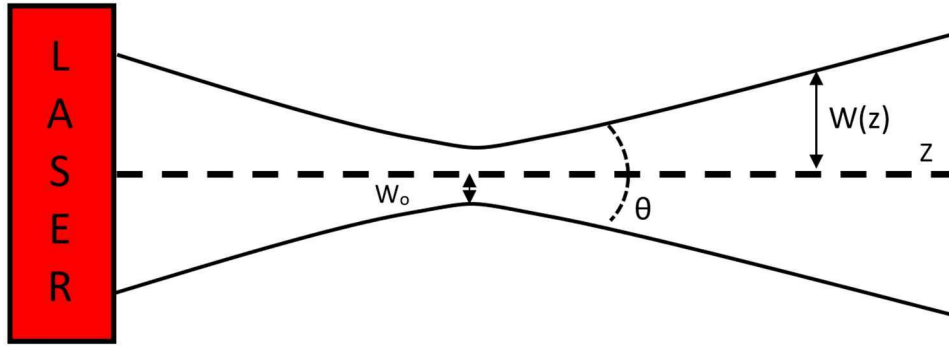
Fonte: KARTALOPOULOS, 2011

A análise do perfil do feixe óptico é complexa e as equações Hermite-Gaussian são usadas para descrever os modos de feixe, que são descritas como “ TEM_{mn} ”, onde m e n são índices polinomiais nas direções x e y . Com distribuição gaussiana pura, obtendo-se $m = n = 0$ e assim TEM_{00} . No entanto alguns lasers não são tão uniformes e operam em modos diferentes, de acordo com a Figura 5.

2.3.3 Divergência do Feixe

Por ocasião da não uniformidade da seção transversal do feixe óptico, ele não é puramente paralelo na direção de propagação, mesmo que o feixe saia paralelo, ele não permanece assim por causa da difração espacial, que faz com que o feixe no primeiro momento se estreite, em um ponto conhecido como “cintura” w_0 , na sequência divergem em um ângulo θ , conforme Figura 6.

Figura 6 - Parâmetros do feixe divergente: ω_0 é a largura da cintura do feixe ω_x , profundidade do foco; θ é a dispersão angular (ângulo de divergência após a cintura)



Fonte: Autora

Um feixe de laser se propagando na direção z e com distribuição gaussiana, matematicamente descrito pela Equação 2.1:

$$E_r = E \frac{\omega_0}{\omega_z} \exp\left(-\frac{r^2}{\omega_z^2}\right) \exp(-jkz - jk(r^2/2R_z) + j\tan^{-1}(z/z_0)), \quad (2.1)$$

onde: E é a amplitude do campo elétrico, ω_0 é a cintura mínima do feixe, ω_x é a cintura do feixe na distância z , r é o raio da cintura em z , k é o número da onda aproximado de $k \approx 2\pi/\lambda R$ é o raio da onda de curvatura na distância z e z_0 é a distância Rayleigh onde o feixe se expandiu para $\sqrt{2\omega_0}$.

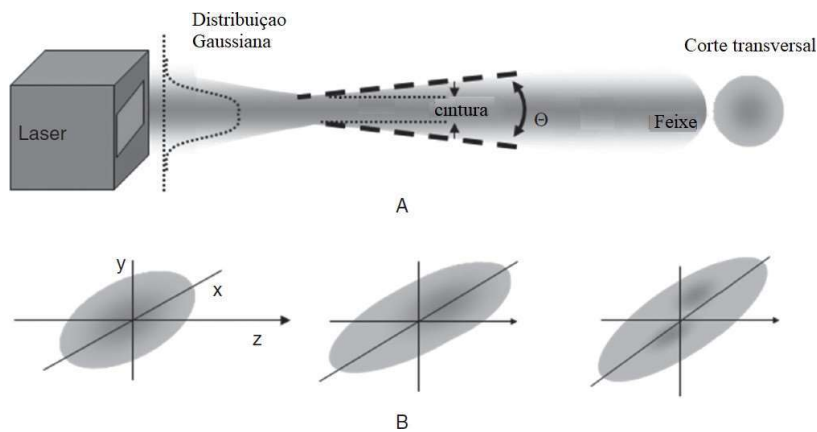
A divergência do feixe expande o diâmetro da seção transversal do feixe ao longo da distância, conhecido como espalhamento geométrico, que reduz rapidamente a densidade de potência óptica do feixe, conhecida como perda de propagação geométrica (KARTALOPOULOS, 2011). O feixe se inicia com uma seção transversal de diâmetro milimétrico na abertura do dispositivo laser, porém como resultado da geometria espalhar o feixe de laser, após alguns quilômetros o feixe terá alguns metros de diâmetro. Existe um dispositivo óptico conhecido como colimador, que ajuda a alcançar uma divergência geométrica desprezível, com um raio quase constante sobre o eixo de propagação.

A perda de propagação geométrica (GSL - *Geometrical Spreading Loss*) para feixes de laser típicos, com área de superfície de abertura de transmissão SA_T , na distância R onde o receptor está com área de superfície de abertura do receptor SA_R , com ângulo de divergência constante θ , e assumindo potência constante através do feixe, é calculado de acordo com a Equação 2.2:

$$GSL = \frac{\text{Área de superfície da abertura de recepção}}{\text{Área da superfície do feixe na distância } R} = \frac{SA_R}{SA_T + (\pi/4)(\theta R)^2}. \quad (2.2)$$

A perda de propagação acima mencionada é para um feixe ideal, porém os feixes práticos não têm uma distribuição de energia constante ao longo da propagação, mas eles podem ter uma distribuição gaussiana. Como citado acima, os feixes de laser não são puramente cônicos, mas podem ter um perfil longitudinal com uma cintura de raio ω_0 , após o que eles divergem em um ângulo quase constante. Em feixes práticos, a seção transversal e a distribuição de intensidade podem não ser circular nem gaussiana, seção transversal não circular reduz a eficiência de acoplamento do feixe na fibra e no receptor FSO devido a irregularidades na distribuição de energia e divergência irregular, como ilustrado na Figura 7.

Figura 7 - Distribuição gaussiana de seção transversal circular e potência irregular distribuição não gaussiana.



Fonte: KARTALOPOULOS, 2011

2.4 EFEITOS ATMOSFÉRICOS NOS SINAIS ÓPTICOS

A troposfera é a camada na qual muitos fenômenos climáticos ocorrem, que interagem e afetam a qualidade de propagação do sinal óptico. A tecnologia FSO em aplicações terrestres depende diretamente da propagação do laser irradiando através da troposfera, portanto esses fenômenos climáticos afetam o sinal óptico de propagação do FSO. A troposfera é um meio altamente dinâmico e instável, consiste de gases, vapores, poeira transportada pelo ar, aerossóis naturais e artificiais, poluentes e outras partículas, que com a luz do sol, variação de temperatura e pressão se movem continuamente e mudam sua consistência e características, portanto prever este modelo dinâmico da troposfera é altamente complexo. Todas as moléculas e partículas da troposfera interagem com a luz e, além de interagir quimicamente, eles causam absorção, dispersão, atenuação de neblina, atenuação de chuva, atenuação de neve, cintilação, descargas elétricas, marés atmosféricas e outros efeitos que afetam o sinal óptico. Nesta seção será observado os principais efeitos atmosféricos que afetam o canal de comunicação do FSO.

2.4.1 Eletricidade Atmosférica

A eletricidade atmosférica, entre as nuvens e entre estas e a superfície da Terra, acumulam um diferencial de tensão típico que varia de 20.000 Volts a 100 Mega-Volts, que pode causar várias descargas atmosféricas, cada uma de até 35.000 Amperes (KARTALOPOULOS, 2011). As descargas atmosféricas, geram raios que emitem flashes de ondas eletromagnéticas, com duração entre 20 a 130 ms, e caso interfira no canal de propagação do FSO a 1 Gb/s com flash de 100 ms, é equivalente a 10^8 bits (ou 12,5 Mbytes), o que corresponde a uma perda substancial de dados.

2.4.2 Névoa

O nevoeiro se caracteriza por uma nuvem em contato com o solo, distingue-se da neblina apenas por sua densidade de gotas, e é expressa em grau de visibilidade em quilômetros ou em metros. O nevoeiro tem uma densidade de gotículas maior do que a neblina e reduz a visibilidade para menos de 1 Km, enquanto a neblina reduz a visibilidade para não menos de 2 Km. Como consequência, a névoa atenua a luz que passa por ela mais do que a neblina. Onde coeficiente de atenuação (atenuação em dB/Km) não é o mesmo para todas as ondas eletromagnéticas, mas é uma função do comprimento de onda, portanto as comunicações em frequências mais baixas são menos atenuadas pelo nevoeiro, sendo mais vantajosas. No entanto, essas frequências são as frequências de rádio que não suportam a alta largura de banda e longo

comprimento de link que as frequências ópticas suportam, e estas são muito mais atenuadas por nevoeiro.

Devido aos muitos parâmetros envolvidos na caracterização do nevoeiro, modelos empíricos foram desenvolvidos para caracterizar a atenuação em termos de visibilidade com base em dados obtidos ao longo de um período para o local de interesse (Tabela 1).

Tabela 1 - Caracterização do nevoeiro com base em dados empíricos.

Tipo de Névoa	Visibilidade (m)	Atenuação (dB/km)
Densa	40-70	250-143
Espessa	70-250	143-40
Moderada	250-500	40-20
Leve	500-1000	20-9.3

Fonte: autora

2.4.3 Poluição do Ar

A poluição do ar é formada por fumaça e névoa, gerada por fábricas que emitem grandes quantidades de fumaça e dióxido de enxofre como resultado da queima de carvão. Atualmente como resultado da lei do ar limpo, esse tipo de poluente foi bastante reduzido nos países avançados. No entanto, as emissões de veículos e industriais interagem com a umidade e outras moléculas na atmosfera que quando a luz solar passa forma poluentes químicos secundários, conhecidos como poluentes fotoquímicos. A poluição fotoquímica absorve e/ou espalha a luz do laser usado na comunicação FSO.

2.4.4 Chuva

A chuva consiste em gotas de água na faixa de 100 μm a 10 mm. Assim, a chuva afeta comunicações na faixa de GHz; por exemplo, a atenuação da chuva em 4 GHz é menor que a atenuação em 12 GHz, uma frequência normalmente usada em links de satélite. A chuva também afeta as frequências ópticas, bem menos que a neve. A atenuação da chuva dos sinais de comunicação depende da intensidade da chuva. Desta forma, um sinal pode ser influenciado pela chuva, mesmo que não haja chuva no local do transmissor ou no receptor, mas no percurso entre os dois.

Granizo são gotas de chuva congeladas que, como a chuva, também afetam a integridade da comunicação também dos sinais ópticos. No entanto, o granizo tem um tamanho de 5 a 50 mm. Como a neve, os modelos empíricos de atenuação da chuva foram desenvolvidos

para prever atenuação do sinal óptico da faixa de visibilidade, $a_{rai} = 2.9/V$.

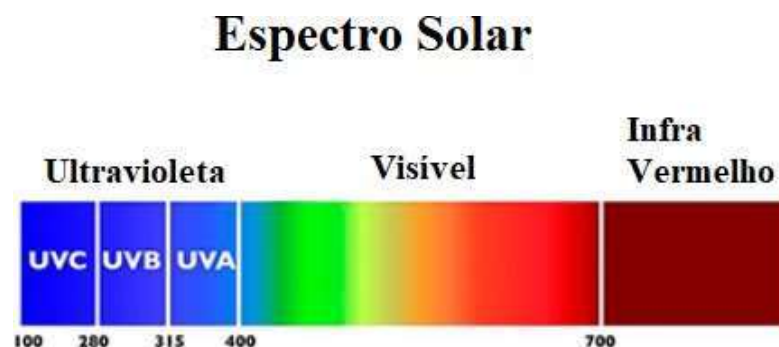
2.4.5 Neve

A neve, como o nevoeiro e a chuva, tem um efeito prejudicial na qualidade dos sinais ópticos transmitidos pelo FSO. A neve é uma forma de precipitação de água na forma de microcristais de gelo hexagonais devido à estrutura molecular da água. À medida que esses microcristais são criados, eles formam flocos e desce nas seguintes formas: flocos de neve, cristais de neve e geada. A atenuação da neve exercida nos sinais ópticos, com base na faixa de visibilidade, é aproximado pelo modelo empírico em, $a_{snow} = 58/V$.

2.4.6 Interferência Solar

No espectro solar (Figura 8) tem-se ondas eletromagnéticas de alta intensidade que são capazes de interferir na transmissão eletromagnética, incluindo as frequências ópticas. No entanto, a interferência solar ocorre se o canal de comunicação opera na radiação do espectro solar. Nos sistemas de comunicação FSO, os comprimentos de onda mais populares são cerca de 800 nm e cerca de 1550 nm, sofrendo interferência solar em potencial, embora a interferência seja mais provável em 800 nm do que em 1550 nm. No entanto, para que a interferência solar seja um problema, a luz do sol deve cair sobre o fotodetector. Além disso, quando 1550 nm é usado, um filtro óptico de passagem longa no receptor pode rejeitar comprimentos de onda abaixo de 1300 nm para reduzir significativamente a interferência da *solar background radiation* ou radiação solar de fundo (SBR). No FSO, a radiação solar de fundo (SBR) é bastante reduzida com um design de caixa adequado e com filtros ópticos.

Figura 8 - Comprimentos de onda do Espectro Solar



Fonte: autora

2.4.7 Espalhamento

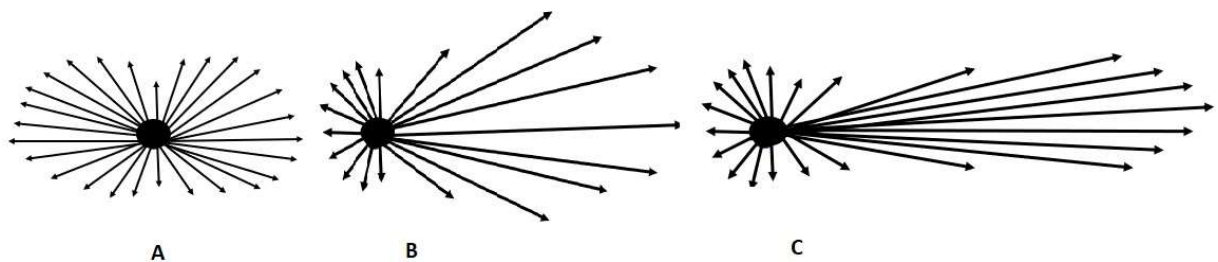
Quando a luz passa pela atmosfera, os fótons interagem com moléculas e outras partículas e são espalhados em todas as direções possíveis como bolas de bilhar, ou seja, os fótons são desviados pelas moléculas sem alterar seu comprimento de onda ou energia. Portanto, o fenômeno do espalhamento não gera redução na intensidade óptica do feixe, o que ocorre é que parte do sinal óptico transmitido não irá chegar ao receptor, deste modo, uma nova atenuação é imposta ao enlace. O mecanismo de dispersão real depende do tamanho e tipo de moléculas do ar e do comprimento de onda da luz.

Um sistema FSO é afetado por três tipos de espalhamentos diferentes gerando atenuação atmosférica:

- Espalhamento Rayleigh
- Espalhamento Mie
- Espalhamento Geométrico

Na Figura 9, são mostrados três tipos de diagrama de espalhamento onde o feixe de luz incide sobre partículas espalhadoras. Observando-se que o tipo de espalhamento depende da relação entre o comprimento de onda do feixe luminoso e do raio da partícula espalhadora.

Figura 9 - O espalhamento de Rayleigh é bidirecional (A), enquanto o espalhamento de Mie (B) e Geométrico (C) é principalmente unidirecional



Fonte: autora

Para realizar a comparação entre comprimento de onda e tamanho da partícula espalhadora e, então, definir que tipo de espalhamento ocorre, o parâmetro adimensional x_0 é apresentado (WILLEBRAND, 2001) na Equação 2.3:

$$x_0 = \frac{2\pi r}{\lambda} \quad , \quad (2.3)$$

onde r é o raio da partícula espalhadora e λ é o comprimento de onda do feixe de luz incidente.

Para $x_0 \ll 1$, tem-se o espalhamento Rayleigh, que ocorre quando as partículas na atmosfera são muito menores do que o comprimento de onda da luz que passa por eles, representado no item A da Figura 9. Como consequência do espalhamento Rayleigh, comprimentos de onda mais curtos (azul) espalham mais do que comprimentos de onda mais longos (vermelho). Como resultado, durante o dia o céu é azul em todas as direções porque o que é visto é a luz espalhada pelas moléculas e partículas da atmosfera. Ao observar o sol, no entanto, o que tem-se é luz não espalhada (com azul sendo removido por dispersão) portanto, o sol parece vermelho - amarelado. No pôr-do-sol, porque o sol está no horizonte ou além dele, os raios percorrem o caminho mais longo na atmosfera, a luz azul já foi espalhada e apenas os raios avermelhados são os que chegam aos nossos olhos.

Para $x_0 \approx 1$, tem-se o espalhamento Mie, que ocorre quando as partículas na atmosfera têm aproximadamente o mesmo tamanho que o comprimento de onda da luz que passa por ela., representado no item B da Figura 9. Como exemplos de partículas responsáveis pelo espalhamento do tipo Mie onde é encontrado poeira, pólen, fumaça e vapor de água.

Para $x_0 \gg 1$, ocorrendo o espalhamento geométrico que ocorre quando partículas na atmosfera são muito maiores que o comprimento de onda. Em aplicações FSO onde a luz do laser é de $1,5 \mu\text{m}$, o tamanho da partícula é da ordem de mm. Tais partículas podem ser solo, areia e poeira que se espalham pelo ar durante tempestades e ventos fortes em específicas partes do mundo (desertos, áreas agrícolas, atividade vulcânica, etc).

2.4.8 Cintilação

A atmosfera é uma mistura de gases, moléculas e partículas que ganham ou perdem energia ou calor continuamente. Algumas partículas ou moléculas são aquecidas mais do que outras, enquanto algumas são resfriadas mais do que outras. Desta forma ocorrendo expansão e contração destas partículas no ar, movimentos ascendentes e descendentes, assim como deslocamentos laterais. O resultado final é uma térmica turbulência em células no ar caracterizadas por mudanças não homogêneas e dinâmicos índice de refração, densidade e consistência do ar. Desta forma quando um feixe de luz se propaga através da turbulência atmosférica, a maioria de suas propriedades são afetadas. Sua polarização, refração, absorção, espalhamento e atenuação flutua aleatoriamente em uma frequência entre 0,01 Hz e 200 Hz. Nas

mesmas condições, a intensidade e a frequência das flutuações aumentam com a frequência da onda, como pode ser visto na Figura 10.

Figura 10 - Efeito da miragem devido a cintilação



Fonte: CORREA, 2016

No FSO, quando o feixe de laser atravessa a turbulência atmosférica, sua polarização e a coerência flutua devido a flutuações anisotrópicas (que assume posições diferentes em resposta à ação de estímulos externos) da massa de ar em todo percurso, e sua constante de atenuação flutua devido à perda de potência não consistente ao longo da massa de ar ao longo do caminho. Como resultado, quando o sinal chega ao receptor, sua intensidade flutua devido a flutuações aleatórias de irradiância temporal e espacial do feixe e o sinal foca e desfoca no fotodetector aleatoriamente. Essa flutuação do sinal devido à turbulência térmica é conhecida como cintilação.

As cintilações são medidas pelo índice de cintilação, índice σ , expresso na Equação 2.4:

$$\sigma_{index}^2 = \frac{\langle I^2 \rangle - \langle I \rangle^2}{\langle I \rangle^2}, \quad (2.4)$$

sendo I a irradiância da onda óptica e os colchetes se referem ao valor médio ao longo do tempo.

Em links de comunicação FSO, o conhecimento do índice de cintilação é importante para determinar o desempenho do sistema. O índice de cintilação aumenta quanto maior for a distância

entre o transmissor e receptor do sinal óptico, quando o sinal óptico torna-se menos coerente e o efeito de foco no receptor de potência óptica se deteriora. Para melhorar o efeito de cintilação, o diâmetro da lente coletora no receptor é aumentado para além da largura de correlação de irradiância, ρ_c , do sinal óptico recebido. Aberturas maiores que ρ_c realizam média de abertura e, como tal, a correlação largura, ρ_c é um parâmetro de projeto que ajuda a reduzir a quantidade de efeito de cintilação no fotodetector do receptor.

2.4.9 Vagueamento do Feixe

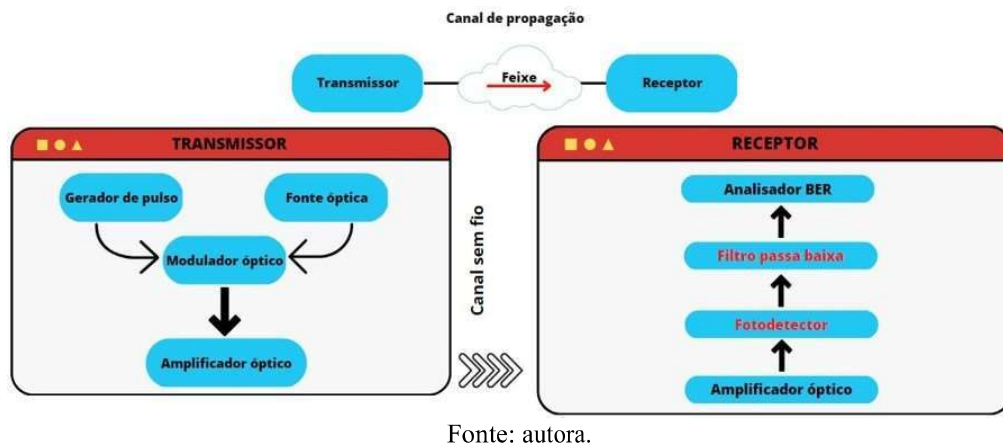
Vento é o movimento de massas de ar que ocorre na troposfera e na estratosfera. A própria luz não é afetada apenas pelo vento, mas talvez o que causa o vento (variações meteorológicas). Em aplicações FSO, o vento tem um efeito degradante no alinhamento do feixe porque os transceptores FSO são posicionados no topo de edifícios ou em postes. A forte turbulência do vento pode balançar um prédio alto e assim deslocar o feixe de seu receptor alvo. Como consequência, o efeito de ventos turbulentos no alinhamento do feixe pode ter efeitos negativos, pois o feixe de laser parece balançar no receptor, que é conhecido como desvio do feixe. Ocorrendo além dos ventos variações de temperatura no caminho do feixe (que causam cintilação), então o desvio do feixe também pode afetar o índice de cintilação, isso é conhecido como beam wander induced scintillation (BWIS) ou cintilação induzida pelo vagueamento do feixe (BWIS). Claramente, o efeito do último depende das propriedades do feixe; BWIS é insignificante se o feixe for colimado ou divergente, e é agravado se o feixe estiver focado.

O efeito do desvio do feixe depende se o feixe é colimado, é divergente ou está focado, no diâmetro do feixe, e também se os transceptores tem um sistema de rastreamento automático para manter o alinhamento mesmo se a posição de um transceptor muda alguns metros em relação ao outro.

2.5 ELEMENTOS BÁSICOS FSO

Um enlace FSO é formado por três elementos principais: transmissor, receptor e canal de propagação. A maneira como esses elementos são dispostos é apresentado na Figura 11.

Figura 11 - Elementos que compõem um canal de comunicação FSO



O objetivo de um sistema de comunicação é transferir informações de um ponto A, de origem, para um ponto B, de destino, com segurança (criptografia, elétrico versus óptico), de forma rápida, econômica e, o mais importante, com o mínimo de deterioração possível (ou seja, com a melhor qualidade possível).

2.6 TRANSMISSOR ÓPTICO

A função do transmissor óptico é converter um sinal elétrico de entrada em um sinal óptico de saída que será inserido no canal de propagação. O componente principal do transmissor óptico é a fonte óptica. Neste trabalho será utilizado o laser semiconductor, pois apresenta tamanho compacto, alta eficiência, boa confiabilidade, apropriada faixa de comprimento de onda, pequena área de emissão e ainda pode ser modulado diretamente em altas frequências.

2.6.1 Emissão Espontânea e Estimulada

Os processos físicos de emissão espontânea e estimulada ocorrem nos LEDs e lasers. Normalmente os materiais absorvem luz ao invés de emitir luz. Na Figura 12 tem-se dois níveis de energia E_1 e E_2 , nível básico e excitado respectivamente. Se um fóton de luz incidente de energia $h\nu$ incide nesse meio, o fóton será absorvido pelo átomo que passa ao estado excitado. Quando o átomo está no estado excitado, retorna ao estado básico de duas formas possíveis: pela emissão espontânea ou pela emissão estimulada. No caso da emissão espontânea, fótons são emitidos fora de fase, idênticos apenas no comprimento de onda, porém na emissão estimulada os fótons são gerados totalmente pareados, com a mesma fase, frequência e polarização, ou seja idênticos ao seu fóton criador.

Figura 12 - Imagem esquemática dos processos de emissão



Fonte: autora

Desta forma, foram identificadas as seguintes características para LED e LASER as quais estão resumidas na Tabela 2.

Tabela 2 - Características LED x LASER

Características	LED	LASER
Taxa de dados	Baixa	Alta
Tipo de fibra	Multimodo	Multimodo / Monomodo
Distância	Curta	Longa
Tempo de vida	Longo	Curto
Sensibilidade a temperatura	Baixa	Considerável
Custo	Baixo	Elevado
Tipo de interação	Luz-matéria	Luz-matéria
Gera luz do tipo	Incoerente	Coerente
Tipo de emissão	espontânea	estimulada
Largura de linha	Larga	Estreita
Possui espelhos nas extremidades	Não	Sim

Fonte: autora

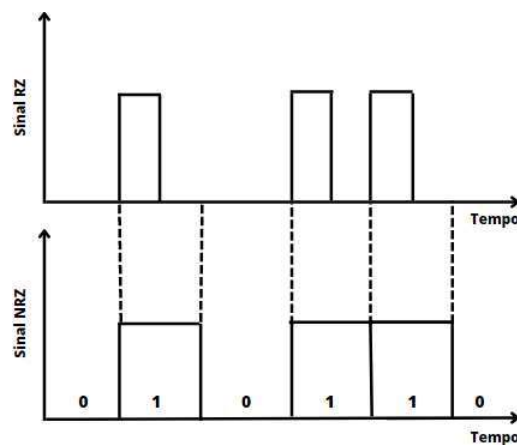
2.6.2 Laser CW

Os lasers de onda contínua, ou continuous wave, têm uma saída nominalmente constante em um intervalo definido, significa que os principais parâmetros do feixe permanecem constantes durante toda a duração do feixe. O termo onda contínua refere-se ao feixe coerente de luz monocromática emitida por meio de ganho, que também determina o comprimento de onda do laser. O comprimento de onda do laser CW é fixo e são usados em comunicações com potência de saída típica na faixa de 10 a 30 mW, largura de linha de 10 Mhz, taxa de supressão de modo único (SMSR) menor que 50 dB e estabilidade de comprimento de onda em torno de 10 pm (± 1.25 Ghz).

2.6.3 Gerador de pulsos

O grande desafio em comunicações ópticas é transmitir uma informação de entrada, que são os dados elétricos, que podem inclusive estar na forma analógica, em um sinal óptico com a mesma informação. A informação será gerada por uma sequência de bits dispostos de forma pseudoaleatória de bits 0 e 1, que normalmente são convertidos em uma sequência de bits digitais no formato RZ - *Return to Zero* ou NRZ - *No Return to Zero* (Figura 13). No nosso trabalho foi utilizado o Gerador de Pulsos NRZ.

Figura 13 - Codificação RZ e NRZ



Fonte: autora

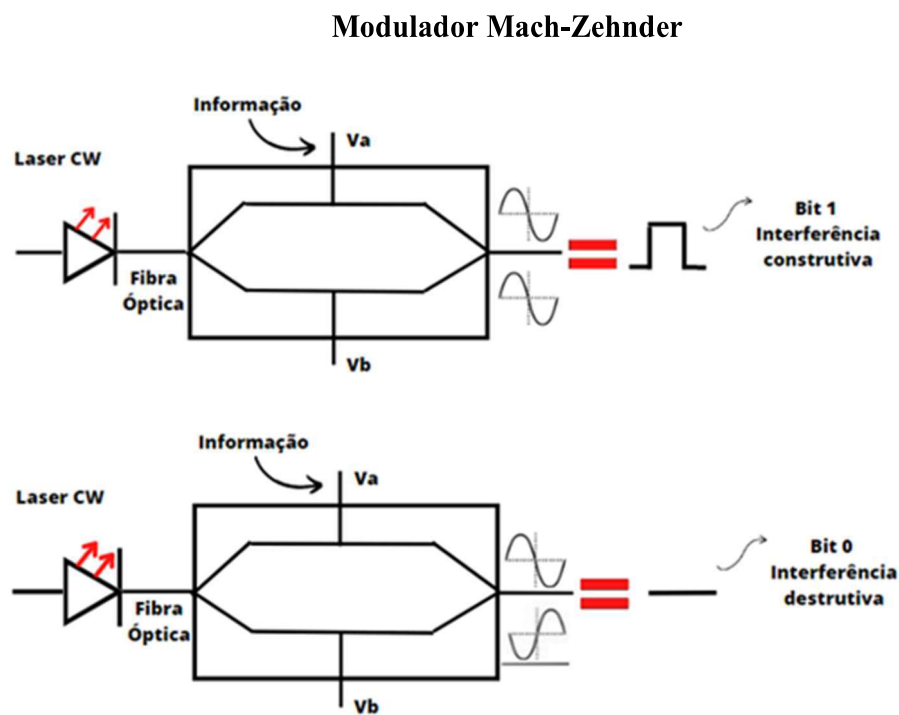
Na codificação RZ, cada pulso óptico é mais curto do que a largura do bit slot, e sua amplitude volta a zero antes que acabe a duração do bit. Na codificação NRZ, o pulso óptico permanece em 0 ou 1 em toda a duração do bit, e sua amplitude não retorna a zero entre dois ou mais bits 1 seguidos. Em consequência, a largura do pulso é variada dependendo da sequência de bits; no formato RZ, a largura dos pulsos é constante. Na codificação RZ tem-se pulsos mais estreitos no tempo e consequentemente mais largos na frequência, já para a codificação NRZ os pulsos são mais largos no tempo e mais estreitos na frequência, conforme a Figura 13.

2.6.4 Modulador Mach-Zehnder

Foi utilizada a modulação externa neste trabalho, haja visto que na modulação direta a taxas de 5 Gb/s ou mais, o chip de frequência se torna tão grande que raramente ela é utilizada. A modulação externa é amplamente utilizada em sistemas de telecomunicações com altas taxas de transmissão e para longas distâncias.

Neste projeto inicialmente, será utilizado o laser CW, de onde sai o sinal contínuo no tempo e estreito na frequência, após o laser CW tem-se o modulador MZ com dois braços, chamados de guias de onda de canal de LiNbO₃, onde o sinal CW é separado através destes dois caminhos, na sequência os sinais são recombinados na saída do MZ. Caso não haja nenhuma interferência neste processo, o sinal de saída será igual ao sinal de entrada, ou seja o mesmo sinal CW da entrada. No MZ é aplicada uma tensão a cada um dos braços, as tensões V_a e V_b . O sinal elétrico de entrada é combinado às tensões V_a e V_b , ou seja muda o índice de refração (η), consequentemente muda a velocidade do sinal. O MZ modula a fase do feixe de luz que atravessa os guias de onda de canal. O que ocorre de fato na saída do modulador MZ, é o produto da ação do sinal elétrico de entrada combinado com as tensões V_a e V_b , de tal sorte que as tensões V_a e V_b são aplicadas para a saída ser “1” ou “0”. Ocorrendo a interferência positiva quando a saída for “1”, e a interferência negativa quando a saída for “0”. Desta forma a sequência de bits elétricos aplicada na entrada do modulador produz uma réplica óptica de sequência de bits na saída do modulador, podendo ser chamado de efeito eletro-óptico do modulador MZ. Esse formato de modulação é conhecido como OOK - *On-Off Keying* (Figura 14).

Figura 14 - Princípio de funcionamento do Modulador Mach-Zehnder



Fonte: autora

2.7 AMPLIFICADOR ÓPTICO

Em enlaces FSO de curto alcance (abaixo de 2 Km), não necessitam de amplificação óptica

desde que o link tenha sido projetado para atender a sensibilidade do fotodetector para a potência emitida pelo laser e considerando a atenuação média nas condições normais do espaço livre, e outras perdas como conectores e etc. No entanto em links mais longos a amplificação do sinal óptico será necessária até atingir seu destino final.

Existem dois sistemas de amplificação, o sistema de amplificação O – E – O e o sistema de amplificação O – O. O sistema de amplificação O – E – O é um processo em várias etapas, primeiro converte o sinal óptico para eletrônico, o sinal eletrônico é reprogramado, remodelado e regenerado (operação conhecida como 3R) e na sequência é convertido para óptico. O sistema O-E-O apesar de mais simples tecnicamente em comparação com o amplificador óptico para óptico é mais caro.

Os amplificadores ópticos diretos (OA), amplificam um sinal óptico fraco, mas eles não executam todas as funções 3R dos amplificadores O-E-O. São exemplos de amplificadores ópticos diretos os amplificadores ópticos semicondutores (SOA), os amplificadores ópticos de fibra (OFA), os amplificadores de Raman estimulados e os amplificadores Brillouin estimulados, destes apenas O SOA e o OFA são aplicáveis ao FSO.

As principais características dos amplificadores ópticos são:

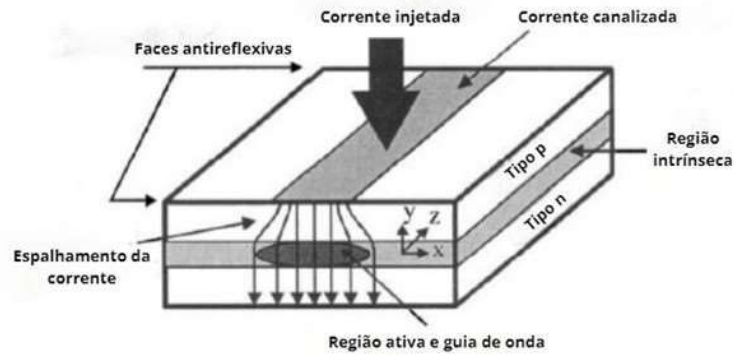
- **Ganho:** é a relação entre a potência de saída e a potência de entrada, medida em dB.
- **Ganho de eficiência:** é o ganho em função da potência de entrada (dB/mW).
- **Largura de banda de ganho:** define a faixa da largura de banda da função de amplificação.
- **Ganho de saturação:** é a potência máxima de saída do amplificador, a partir de então ele não pode aumentar o ganho, mesmo com o aumento da potência de entrada.
- **Ruído:** é uma característica inerente dos amplificadores. Em amplificadores eletrônicos, o ruído é devido à recombinação espontânea (aleatória) de pares elétron-buraco que produz um sinal indesejado adicionado ao sinal de informação a ser amplificado. Nos amplificadores ópticos, é devido à emissão espontânea de luz de íons excitados.
- **Sensibilidade de polarização:** é a dependência de ganho de amplificadores ópticos na polarização do sinal.
- **Potência de saída saturada:** é a potência de saída máxima que pode ser alcançada após a amplificação, e a potência amplificada não pode exceder essa potência, também é definida como o nível de potência de saída para o qual o ganho do amplificador caiu 3 dB.

2.7.1 Amplificadores ópticos semicondutores

Amplificadores ópticos semicondutores (SOA) se baseiam no princípio do laser convencional. Uma região de guia de onda ativa é impressada entre uma região tipo p e uma região tipo n .

Uma corrente é aplicado para excitar íons da região ativa e criar pares elétron-buraco, de modo que esses elétrons saltam para um nível de energia mais alta. Então, como luz de um comprimento de onda específico é acoplado ao guia de ondas ativo, a estimulação ocorre e faz com que os pares elétron-buraco se recombinem e gerem mais fótons, com a mesma fase, frequência, polarização e direção do fóton que chega e portanto, a amplificação óptica é alcançada. Para melhorar a eficiência de acoplamento do sinal óptico na região ativa, as faces finais do SOA são revestidas com um material antireflexivo (Figura 15).

Figura 15 - Estrutura do Amplificador SOA



Fonte: Facco Rodrigo, 2019.

O ganho do amplificador SOA é dado pela Equação 2.5:

$$G = -12 \left(\frac{\lambda - \lambda_p}{\Delta\lambda} \right)^2 + G_p, \quad (2.5)$$

onde: G_p é o ganho de pico no comprimento de onda correspondente a λ_p , $\Delta\lambda$ é o FWHM (*full – width half – maximum*) na largura de banda e o fator 12 é resultado da definição de $\Delta\lambda$.

A potência de saída de saturação de 3 db é descrita pela Equação 2.6:

$$P_s = q_s(\lambda - \lambda_p) + P_{s-p}, \quad (2.6)$$

onde, q_s é o coeficiente linear e $P_{(s-p)}$ é a saída saturada (3dB) no ganho de pico (λ_p).

Vantagens do SOA:

- Alto ganho (25 – 30 dB)
- Compacto: SOAs são feitos com InGaAsP e, portanto, são semicondutores pequenos e compactos
- facilmente integrável com outros semicondutores e componentes ópticos.
- Alta potência de saída: Potência de saturação de saída na faixa de 5 a + 13 dBm
- Ampla faixa de comprimento de onda
- Baixo custo
- Alto volume de produção

Desvantagens do SOA:

- Alta perda por acoplamento
- Depende da polarização
- Controle de temperatura
- Alta figura de ruído: mais alta que EDFA

2.7.2 Amplificadores de Fibra Óptica

Uma importante classe de amplificadores ópticos concentrados utiliza terras raras como meio de ganho por dopagem do núcleo da fibra óptica durante o processo de fabricação (AGRAWAL, 2014). A finalidade é absorver a energia óptica em uma faixa espectral e emitir energia óptica em outra, dentro da faixa de comprimento de onda útil em comunicações de fibra (800 – 900 ou 1300 – 1620 nm). Sabe-se que cada elemento tem suas próprias características de absorção e emissão. Um OFA amplifica apenas os sinais dentro de uma faixa específica que é determinada pelo elemento dopante, e não tem exatamente o mesmo ganho para todos os comprimentos de onda na faixa espectral de amplificação.

Para uma amplificação sustentada, a taxa de excitação deve ser menor ou igual à taxa de estimulação mais a taxa de emissão espontânea, como descrito na Equação 2.7.

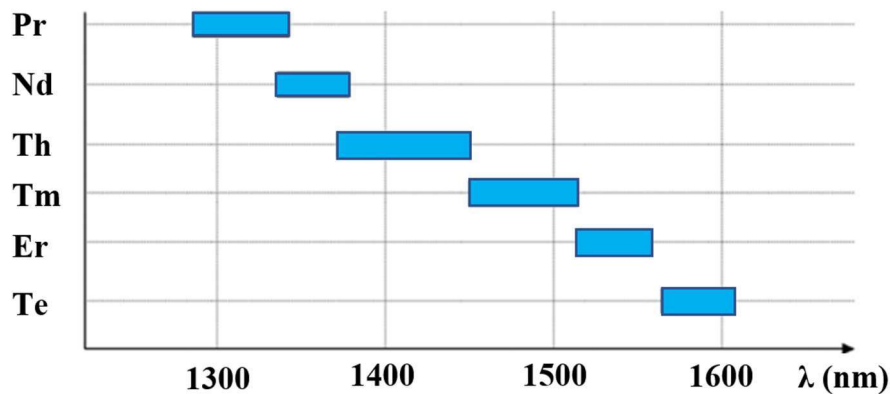
$$\frac{dN_e}{dt} \leq \frac{dN_{st}}{dt} + \frac{dN_{sp}}{dt}, \quad (2.7)$$

onde, N_e é o número de elétrons excitados para um nível de energia mais alta, N_{st} é o número de elétrons retornando a um nível de energias mais baixas por estimulação e N_{sp} é o número de elétrons retornando a um nível de energias mais baixas espontaneamente.

Portanto, o processo de excitação, estimulação e amplificação se reduz a um problema descrito por um conjunto de equações de taxas diferenciais.

O elemento químico mais utilizado para operar nos amplificadores a fibra dopada é o érbio (Er^{3+}), pois emitem na faixa de $1.5 \mu\text{m}$ de comprimento de onda. Além do érbio, outros elementos de terras raras foram utilizados nos amplificadores de fibra dopada (DFA): Nd^{3+} , Ho, Te, Th, Tm, Yb e Pr, cada um operando em várias bandas espectrais, conforme Figura 16.

Figura 16 - Diferentes elementos podem ser usados em DFAs para amplificar bandas ópticas correspondentes



Fonte: autora

2.7.2.1 - Amplificadores de Fibra Dopada por Érbio

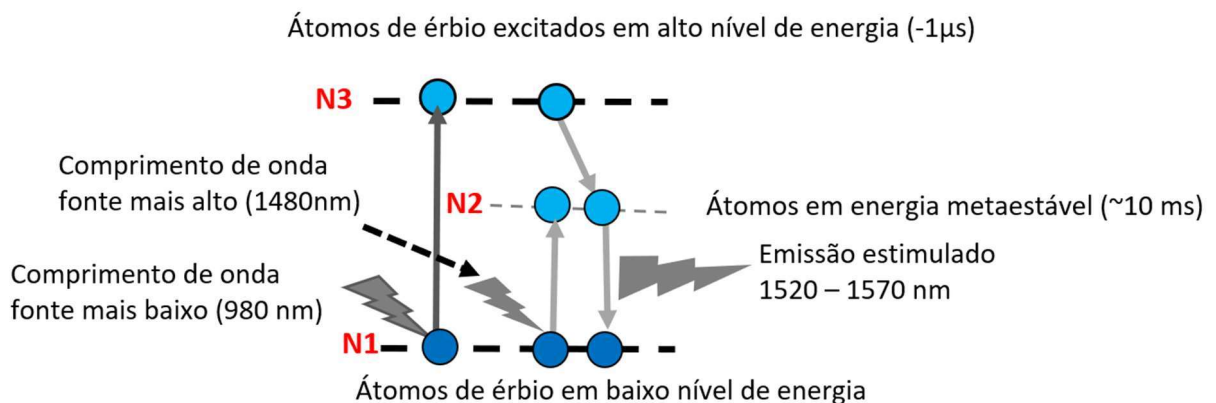
O amplificador de fibra dopada com érbio (EDFA) revolucionou a comunicação por fibra óptica após 1995, pois sua emissão espectral está entre 1530 a 1565 nm, que corresponde à banda C em DWDM. Como o érbio também é excitado por várias frequências ópticas conhecidas como bombeamento, a um comprimento de onda de 514, 532, 667, 800, 980 e 1480nm, onde os últimos dois são os bombeamentos mais populares em EDFAs usados em comunicações de fibra óptica.

O ganho do amplificador EDFA é dado pela Equação 2.8:

$$G = 10 \log_{10} \left(\frac{P_s}{P_e} \right) dB. \quad (2.8)$$

Sabe-se que quanto menor o comprimento de onda mais energético é o fóton, portanto fótons de bombeamento retiram os elétrons do estado fundamental de energia N1 para N3, porém o tempo de vida do N3 é muito curto ($1\mu s$), então os elétrons caem rapidamente para o N2, que possui um tempo de vida maior (10ms). Deste modo os elétrons vão se acumulando no nível 2 e quando o número de população dos elétrons no N2 for maior que no N1 a inversão de população acontece, neste ponto começa a amplificação do sinal. Neste momento quando um fóton de sinal chega ele vai interagir com os elétrons de N2, que decaem de N2 para N1, gerando um fóton gêmeo, por emissão estimulada com a mesma fase, frequência, polarização e direção do fóton que chega alcançando assim amplificação do sinal. Esse processo de amplificação ocorre em toda extensão da fibra dopada (Figura 17).

Figura 17 - Níveis de excitação para o érbio



Fonte: autora

EDFAs com ganho superior a 50 dB, ou seja o ganho 100.000 vezes maior que o sinal de entrada, já foram aplicados em comunicações por fibra óptica, bem como nas estações de retransmissão FSO que operam na banda C. No entanto, como os sinais FSO não carregam muitos canais ópticos, cada um em comprimento de onda diferente, mas poucos canais ópticos, o ganho EDFA não é compartilhado por muitos canais como nas comunicações por fibra óptica e, portanto, não há problema em esgotar o EDFA. Portanto em estações retransmissoras FSO pode ter um baixo ganho e características de baixo ruído EDFA, aumentando a qualidade do sinal com redução de custos.

Vantagens EDFAs:

- Têm uma alta eficiência de transferência de energia do bombeamento para o sinal (> 50%).
- Amplifica direta e simultaneamente uma ampla região de comprimento de onda acima de 80 nm (na região de 1550 nm).
- São transparentes ao formato de modulação óptica.
- Possuem uma figura de ruído baixo
- São independentes de polarização (reduzindo assim a perda de acoplamento para transmissão via fibra)
- Disponível em Banda C e L
- Não necessita de eletrônica robusta

Desvantagens EDFAs:

- Não são dispositivos pequenos e não podem ser integrado com outros dispositivos semicondutores.
- Exibem emissão de luz espontânea amplificada (ASE), esta saída é chamada de ruído espontâneo.
- Diafonia.
- Ganho de saturação
- Ganho depende do comprimento da fibra dopada
- Ganho depende do comprimento de onda
- Espectro limitado

2.8 RECEPTOR ÓPTICO

O papel de um receptor óptico é converter o sinal óptico de volta à forma elétrica e recuperar os dados transmitidos pelo sistema de onda luminosa. O receptor óptico pode conter um amplificador, um fotodetector intrínseco (PIN) e um filtro passa-baixa Bessel (LPF), o LFP usado com uma frequência de corte de $(0,75 \times \text{symbol rate})$. O seu principal componente é o fotodetector, que converte luz em eletricidade por meio do efeito fotoelétrico (AGRAWAL, 2010).

Dependendo do tipo de material do fotodiodo algumas características diferem como: tempo de resposta, sensibilidade, ou a capacidade de gerar buracos - pares de elétrons por

fóton, taxa de ruído, custo, confiabilidade e dimensões. Esses requisitos são mais bem atendidos por fotodetectores feitos de materiais semicondutores. Os fotodiodos semicondutores são parecidos com os diodos, porém é permitido que sua junção seja exposta à luz externa. Em fotodiodos semicondutores, quando um fóton interage com a junção, um elétron é excitado da banda de valência para a banda de condução, deixando um buraco positivo na banda de valência.

O mecanismo fundamental responsável pelo processo de foto-detecção é a absorção óptica (fenômeno caracterizado pela perda de intensidade de uma onda eletromagnética que passa através de um meio material, onde parte de sua energia pode ser absorvida pelos átomos e convertida em energia térmica).

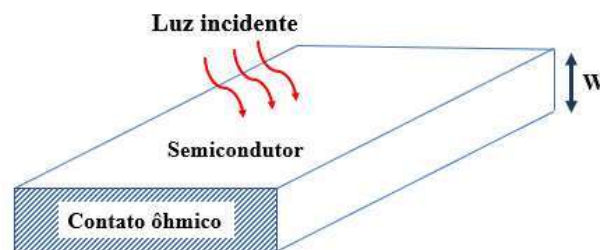
2.8.1 Fotossensibilidade/Responsividade e Eficiência Quântica

Dado um bloco semicondutor ilustrado da Figura 18, onde a energia $h\nu$ dos fótons incidentes exceder a banda proibida de energia, um par elétron-lacuna será gerado a cada vez em que um fóton for absorvido pelo semicondutor. Com uma tensão aplicada entre os polos de contatos ôhmicos, tem-se um campo elétrico, onde elétrons e lacunas são varridos ao longo do semicondutor, resultando no fluxo de uma corrente elétrica, Equação 2.9:

$$I_p = R_d \cdot P_{in}, \quad (2.9)$$

onde, I_p é a fotocorrente, R_d a responsividade do fotodetector e P_i a potência óptica incidente.

Figura 18 - Bloco semicondutor usado como fotodetector



Fonte: autora

A fotossensibilidade é a proporção da potência óptica (em Watts) incidente no dispositivo para a corrente resultante (em Amperes), também conhecida como responsividade (medida em $A/(W)$).

A grandeza fundamental η denominada *eficiência quântica* é definida como o número de pares elétrons-buracos gerados (ou seja, corrente), dividido pelo número de fótons, Equação 2.10:

$$\eta = \frac{\text{Taxa de geração de elétrons}}{\text{Taxa de incidência de fótons}} = \frac{I_p/q}{P_{in}/h\nu} = \frac{h\nu}{q} R_d, \quad (2.10)$$

sabendo que $E_g = h\nu = 1.24$ e $\lambda = c/\nu$, ambos expressos em μm , é disposto abaixo na Equação 2.11:

$$R_d = \frac{\eta q}{h\nu} \approx \frac{\eta \lambda}{1.24}, \quad (2.11)$$

a responsividade de um fotodetector é diretamente proporcional ao comprimento de onda, porém esta proporcionalidade tem um limite. Quando a energia do fóton for menor que E_g ($h\nu < E_g$), ou seja quando entrar na banda proibida, a eficiência quântica η cai a zero.

A relação de dependência de η em relação a λ , pode ser demonstrado por meio do coeficiente de absorção α . Considerando que as faces do bloco de semiconductor na Figura 18 têm revestimento antirreflexo, a potência transmitida P_{tr} , através do bloco de largura W é dada na Equação 2.12:

$$P_{tr} = P_{in} \exp(-\alpha W). \quad (2.12)$$

Por outro lado, a potência absorvida, pode ser escrita na Equação 2.13:

$$P_{abs} = P_{in} - P_{tr} = P_{in}[1 - \exp(-\alpha W)], \quad (2.13)$$

portanto, a eficiência quântica η é dada pela Equação 2.14:

$$\eta = \frac{P_{abs}}{P_{in}} = 1 - \exp(-\alpha W). \quad (2.14)$$

Desta forma η se torna zero quando $\alpha = 0$, e no caso em que $\alpha w \gg 1$, η tende a 1.

2.8.2 Largura de banda , tempo de subida e corrente no escuro

A largura de banda de um fotodetector é determinada pela velocidade com que o mesmo responde às variações na potência óptica incidente (AGRAWAL, 2010).

O tempo de subida T_r , é definido como o tempo necessário para que a corrente passe de 10% a 90% de seu valor final quando a potência óptica muda rapidamente. T_r depende do tempo que elétrons e lacunas levam para viajar até os contatos elétricos. T_r também depende do tempo de resposta do circuito elétrico usado para processar a fotocorrente.

O terceiro parâmetro importante para um fotodetector é a corrente no escuro, ou *dark current* (I_d), que se refere a corrente gerada na ausência de sinal óptico, quando o diodo é polarizado reversamente, é uma fonte de ruído sob condições de polarização reversa. Este parâmetro deve ser desprezível para um bom fotodetector ($I_d < 10\text{nm}$).

2.8.3 Fotodetector PIN

O receptor mais comum para comunicações ópticas de alta velocidade é o detector baseado em fotodiodo PIN. Os detectores PIN podem ser de banda extremamente larga, com resposta DC a 50+ GHz disponível comercialmente. Embora os receptores baseados em PIN sejam relativamente simples, eles são os menos sensíveis e requerem cuidado no design eletrônico *front-end*, para alcançar um bom desempenho a altas taxas. Segue na Tabela 3 as características para o fotodiodos PIN.

Tabela 3 - Características de fotodiodos PIN comuns

Parâmetro	Símbolo	Unidade	Si	Ge	InGaAs
Comprimento de onda	λ	μm	0,4-1,1	0,8-1,8	1,0-1,7
Responsividade	R_d	A/W	0,4-0,6	0,5-0,7	0,6-0,9
Eficiência quântica	η	%	75-90	50-55	60-70
Corrente no escuro	I_d	nA	01/out	50-500	1-20
Tempo de subida	T_r	ns	0,5-1	0,1-0,5	0,02-0,5
Largura de banda	Δf	GHz	0,3-0,6	0,5-3	1-10
Tensão de polarização	V_b	V	50-100	6-10	5-6

Fonte: AGRAWAL, 2010

2.8.4 Relação Sinal Ruído

A *shot noise current* ou corrente de ruído de tiro, $I_{(S-N)}$ de um fotodetector é representado na Equação 2.15:

$$I_{(S-N)} = \sqrt{2e(I_{dark} + I_{ph})B}, \quad (2.15)$$

onde, I_{dark} é a corrente que flui na ausência do sinal fotônico, I_{ph} é a fotocorrente gerada pelo sinal fotônico e B é a largura de banda do fotodetector (em MHz).

Da mesma forma, quando a corrente gerada flui através de um resistor de carga, *thermal noise current* ou corrente de ruído térmico é gerada, $I_{(Th-N)}$, cujo valor médio é expresso pela Equação 2.16:

$$I_{(Th-N)} = \sqrt{4kT B/R}, \quad (2.16)$$

onde, T é a temperatura (em graus Kelvin), B é a largura de banda do fotodetector, k é a constante de Boltzmann e R é a resistência de carga.

Caso a potência do sinal for expressa em termos de fotocorrente e resistor de carga, então a relação sinal-ruído, SNR, é expressa nas Equações 2.17 e 2.18:

$$SNR = 10 \log_{10}[(Signal\ power)/(Total\ noise\ power)], \quad (2.17)$$

$$SNR = 10 \log_{10}[(I_{Th-N}^2)/(I_{S-N}^2 + I_{Th-N}^2)]dB. \quad (2.18)$$

2.8.5 Taxa de erro de bit - BER

Pode-se definir as informações enviadas de uma fonte digital quando a mensagem j th é transmitido conforme Equação 2.19:

$$I_j = \log_2 \left(\frac{1}{P_j} \right) \text{ bits}, \quad (2.19)$$

onde P_j é a probabilidade de transmissão da j -enésima mensagem. Já a probabilidade de um evento A (bit erro), ou seja $P(A)$ pode ser definido em termos da frequência relativa de A ocorrer em n tentativas, e é dado na Equação 2.20:

$$P(A) = \lim_{n \rightarrow \infty} \left(\frac{n_A}{n} \right), \quad (2.20)$$

onde, $P(A)$ é Probabilidade de A (bit erro) ocorrer em n tentativas, n_A é número de vezes que A (bit erro) ocorre em n tentativas e n é número total de tentativas.

Todas as funções de probabilidade ($0 \leq P(A) \leq 1$) são cotadas entre limites fixos.

Portanto, o BER - *Bit Error Rate*, pode ser dado pela Equação 2.21:

$$BER = \frac{\text{Número de erros}}{\text{Número total de bits enviados}}, \quad (2.21)$$

2.9 ESTUDO DE IA

2.9.1 História da inteligência artificial

Em 1943 Walter Pitts e Warren McCulloch juntaram a proposta de funcionamento do neurônio biológico do ganhador Nobel de Medicina Camillo Golgi. Neste mesmo ano foi criado pelos pesquisadores, um sistema neural artificial, onde eram interpretados em sinais binários de ligado ou desligado. O objetivo era representar em funções matemáticas uma certa combinação destes neurônios.

Em 1950 foi construído em Harvard, o primeiro computador neural por Marvin Minsk e Dean Edmonds, com a finalidade de simular o comportamento de 40 neurônios. Neste mesmo ano foi criado o teste de IA, capaz de responder perguntas por escrito de um interrogador humano, conhecido como Teste de Turing. Resultados positivos para este teste somente começaram a aparecer após os anos 2000.

Em 1958 foi criada a linguagem LISP (*LIS*t *Pro*cessing), a principal linguagem de programação para trabalhar com IA que ficou sendo usada por cerca de trinta anos.

Em 1959 as primeiras implementações de algoritmos genéticos ocorreram, mas não teve muito sucesso com estes experimentos.

Nos anos 60 os primeiros sistemas inteligentes para tradução de texto foram feitos, porém os resultados até 1966 não eram satisfatórios.

No início dos anos 1970, no ramo de pesquisa da IA, foi proposto o uso de elementos racionais mais simples para formar um elemento mais complexo, conhecido como Métodos Fracos ou IA Fraca. Então surgiram na sequência os Sistemas Especialistas, que resolviam problema específicos de forma satisfatória, porém ao se estender os sistemas especialistas para área da medicina, foi visto a necessidade do sistema tratar também a incerteza do modelo. Daí se iniciaram os estudos de fatores de incerteza nas decisões que foram apresentados pelo sistema de IA.

Na década de 80 foi dada mais ênfase aos sistemas especialistas. No final desta década as pesquisas em redes neurais voltaram a ter relevância.

Em 1998 técnicas de convolução em redes neurais de múltiplas camadas foi aplicada

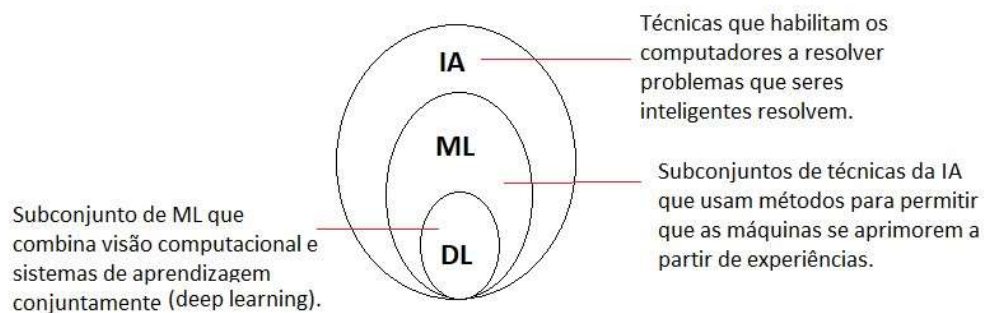
por um grupo de pesquisadores, visava extrair características de imagem. Foi possível o uso de dados não estruturados com a CNN (*Convolutional neural Network*), onde teve o início ao que ficou conhecido como *Deep Learning*.

2.9.2 Aprendizado de Máquina

Vários conceitos para aprendizado de máquina foram mostrados abaixo:

- A habilidade que uma máquina tem para melhorar seu desempenho com base em experiências acumuladas;
- Uma sub-área da inteligência artificial para obtenção de novos métodos, que podem aprender tanto pela experiência adquirida ou através da extração de conhecimento de exemplos (Figura 19).

Figura 19 - Temas do estudo de IA



Fonte: autora

- Uma área da ciência da computação que permite que o computador tenha a capacidade de aprender sem ser programado simplesmente;
- O principal objetivo do aprendizado de máquinas é desenvolver algoritmos que executem aprendizagem automática sem intervenção humana.

Existem dois principais paradigmas para entender como funciona a aprendizagem de máquina: o aprendizado supervisionado e o aprendizado não supervisionado. Em um sistema de aprendizagem de máquina, o que é fornecido para o algoritmo são as entradas que tem-se e as saídas esperadas. A ideia é deixar o algoritmo “aprender” os padrões e “desenvolver” a solução necessária para que aquelas entradas resultem naquelas saídas. O que é muito diferente do que é feito durante o desenvolvimento do software tradicional, quando são informadas às entradas como o programa vai se comportar.

“Podemos dizer que um sistema aprendeu algo quando, para realizar uma tarefa que pode ser medida, ele aprende com novas experiências e melhora o seu desempenho ao realizar

aquela tarefa.” (NOGARE, 2020), ou seja, algo só pode aumentar sua eficiência se este algo puder ser medido seu desempenho e compararmos os resultados com as novas experiências. É esta análise de desempenho que ajuda a máquina a aprender, pode ser executada de forma explícita, com alguma rotina dizendo que o resultado foi assertivo ou não, gerando novas experiências para que o sistema melhore o desempenho nos passos seguintes. Esta análise de desempenho pode ser feita também de forma implícita, quando o sistema executa uma determinada ação e é entendido que a resultante do algoritmo foi correta ou incorreta.

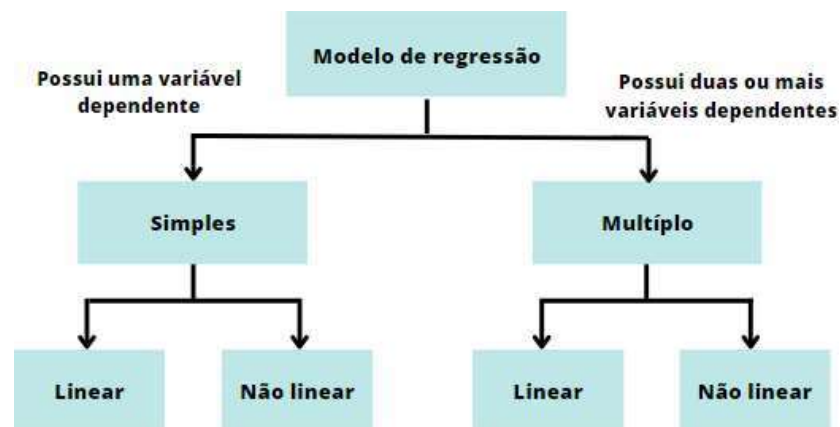
Os sistemas de aprendizagem de máquinas são aplicados para problemas complexos, quando o desenvolvimento do software tradicional não é eficaz, podem ser classificados em famílias e de acordo com a escolha do tipo de supervisão.

2.9.3 Aprendizado Supervisionado

Neste paradigma de aprendizado, a variável de saída, ou variável alvo é fornecida na base de treino, ou seja, é conhecido com exatidão quais são as entradas e saídas esperadas para um conjunto de dados. Tendo duas possíveis categorias no aprendizado supervisionado: o algoritmo de classificação e o algoritmo de regressão. Cada uma dessas categorias resolve problema diferente. No algoritmo de classificação é categorizado um conjunto de dados em classes. Seguem alguns dos algoritmos de classificação mais famosos: KNN, *Naive Bayes*, *Logistic Regression*, *Support Vector Machines* e *Decision Trees*. O algoritmo de regressão é utilizado quando se quer prever valores, ou seja trabalha para responder dados numéricos. Ambos, mesmo com alvos, ou saídas completamente diferentes, fazem parte do mesmo paradigma de aprendizado supervisionado. Alguns modelos famosos de algoritmos de regressão são: linear regression, polynomial regression, logistic regression, principal components, regression (PCR). Os modelos de regressão são diversos e suas possibilidades estão descritas abaixo.

Na Figura 20 a primeira subdivisão dos modelos de regressão se refere ao número de variáveis, para uma variável, tem-se , modelos de regressão simples, ou múltiplo para duas ou mais variáveis. Em seguida, se divide em lineares e não lineares.

Figura 20 - Tipos de modelos de regressão utilizados no aprendizado de máquina



Fonte: autora

Para criar um destes sistemas, classificador ou regressor, é necessário apresentar muitos exemplos para o algoritmo, por exemplo se possuímos as características de um veículo como: cor, quantidade de portas, tipo de carroceria, potência do motor, marca, modelo. Para trabalhar com algoritmo de classificação, que por exemplo pode responder qual é o tipo de carroceria deste veículo. O algoritmo de regressão pode responder qual a potência do motor, do veículo. Para simplificar pode-se dizer que os algoritmos de classificação tem como objetivo classificar itens ou amostras de acordo com as características observadas pelo supervisor, enquanto os algoritmos de regressão refletem a compreensão de relação da máquina, referente às variáveis para prever valores.

2.9.4 Aprendizado Não Supervisionado

Como o próprio nome sugere o aprendizado não supervisionado não possui uma variável alvo, ou saída definida, não se sabe qual será sua saída até observar os resultados. Neste paradigma tem-se vários tipos de famílias de algoritmos, como: algoritmos de agrupamento ou *Clustering*, algoritmo de redução de dimensionalidade e também regras de associação. Foram discutidas cada uma delas abaixo:

Agrupamento ou *Clustering*: consiste em apresentar grupos possíveis de dados não rotulados baseado em suas semelhanças ou diferenças. Por exemplo para melhorar a experiência dos leitores de um determinado blog em um determinado jornal, se forem coletados dados com as características dos leitores como: gênero, horário de acesso ao blog, tempo lendo o conteúdo do blog e país de origem. Pode-se criar um algoritmo de agrupamento entre as pessoas mais

assíduas e menos assíduas a este blog. Não se tem nenhuma informação sobre esses leitores, mas este algoritmo pode apresentar grupos possíveis e posso criar rótulos para esses grupos. Um dos algoritmos mais usados para esta tarefa é o *K-Means*.

Regras de associação: neste caso busca-se aprender relações que descrevem porções dos dados. Como exemplo, outra atividade para melhorar a experiência dos leitores na recomendação de conteúdos, seria utilizar os grupos separados pelo algoritmo de agrupamento, onde é possível saber quais matérias foram lidas em cada visita ao blog. Desta forma quando um leitor lê uma matéria é possível recomendar outro post, que talvez seja de interesse dele, baseado na experiência de outros visitantes. O *Apriori* é um dos algoritmos mais comuns para esta tarefa.

Redução de dimensionalidade: ao estudar um conjunto de dados, o percentual de dados utilizados para treinamento pode ser muito alto. O excesso de dados de entrada pode impactar o desempenho dos algoritmos causando, por exemplo, o *overfitting*, ou seja o modelo acertaria praticamente tudo na avaliação dos dados restantes, mas caso fossem repassados novos dados para o modelo o resultado seria bem diferente do esperado. Ao utilizar técnica de redução de dimensionalidade, será executada uma redução no número de recursos de entrada, de forma que torne-os gerenciáveis por parte do modelo, além de preservar a integridade dos dados. Existem várias técnicas que podem ser utilizadas: *Missing Values Ratio*, *Low Variance Filter*, *High Correlation Filter*, *Random Forests / Ensemble Trees*, *Principal Component Analysis (PCA)*, *Backward Feature Elimination* e *Forward Feature Construction*.

2.10 ALGORITMOS DE INTELIGÊNCIA ARTIFICIAL

2.10.1 Boosting

O *Boosting* é um método de conjunto ou *ensemble* que constroem vários modelos de aprendizagem de máquinas, utilizando o resultado de cada modelo na definição de um único resultado, com o objetivo de construir um melhor modelo de predição. Existem diversos algoritmos de *Boosting*: *AdaBoost*, *MART*, *Gradient Boosting*, *LogitBoost*, *L2Boosting*, *Gradient Boosting*.

2.10.2 Adaboost + Decision Tree Regression (Adaboost + DTR)

AdaBoost foi um dos algoritmos de Boosting, utilizados para melhorar o resultado do algoritmo *Decision Tree Regression*. É explicado em (DIETTERICH, 2000) que o algoritmo *AdaBoost* trabalha com uma ideia de manter um conjunto de pesos sobre o conjunto de dados de treinamento original S . Tais pesos sofrerão alteração posteriormente pelo próprio *AdaBoost* após cada modelo ser treinado por um algoritmo de treinamento básico (em questão, foi utilizado o *Decision Tree Regression*). À medida que o modelo erra sua previsão, o valor do peso é aumentado e, à medida que o modelo acerta, o valor do peso é decrementado. Também é feita uma descrição a respeito de como o *AdaBoost* consegue utilizar os pesos obtidos para a construção de um novo conjunto de dados de treinamento S , a ser passado ao algoritmo de treinamento básico. Para tal ação, existem dois métodos possíveis: *Boosting* por amostragem e *Boosting* por Ponderação.

No *Boosting* por amostragem primeiro, a construção de exemplos é feita a partir de S , com reposição, e com probabilidade proporcional aos seus pesos.

Já no *Boosting* por Ponderação, pode ser utilizado com algoritmos de treinamentos que conseguem aceitar diretamente o conjunto de treinamento ponderado.

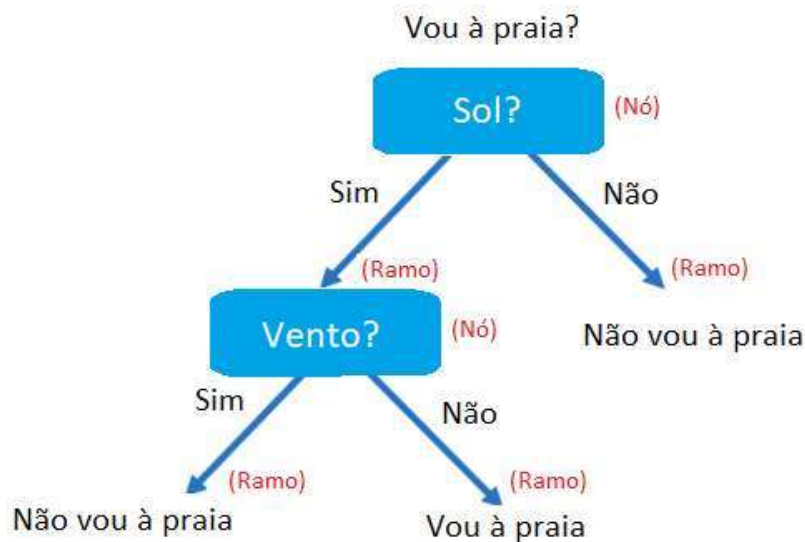
Seja com o primeiro ou com o segundo, todo o conjunto de treinamento S com pesos associados é passado para o algoritmo de treinamento básico. *AdaBoost* foi considerado um dos métodos *ensemble* mais efetivos.

2.10.3 Boosted Decision Tree Regression (BDTR)

Boosted Decision Tree Regression é um algoritmo de regressão formado a partir de duas técnicas: *Decision Tree* e *Boosting* (Jumin et al., 2021). A parte de *Boosting* é baseada em uma eficiente implementação do algoritmo *Multiple Additive Regression Trees* – MART (LIKEBUPT et al., 2022), por sua vez, trabalha produzindo sequencialmente novas árvores de decisão baseando-se na redução da existência de resíduos (ELBISY, 2021). *Decision Tree Regression* (DTR), ou árvore de decisão é um dos primeiros algoritmos aprendidos pelos iniciantes de aprendizado de máquina. Como o nome sugere, neste algoritmo vários pontos de decisão serão criados, estes pontos são os “nós” da árvore e em cada um deles o resultado da decisão será seguir por dois caminhos. Os caminhos existentes são os conhecidos “ramos”. Esta é a estrutura básica de uma árvore de decisão, nos nós temos os pontos de decisão, que irão indicar um ramo ou outro para sequência do fluxo, e estas perguntas terão como opções de resposta o sim e não, e desta forma toda a árvore estará sendo construída, partindo de um ponto comum, podendo existir várias opções de caminhos diferentes

a serem percorridos na árvore, cada um levando a um resultado. Na Figura 23 temos uma árvore de decisão extremamente simples, apenas com dois nós e poucos ramos.

Figura 21 - Diagrama de uma árvore de decisão



Fonte: autora

Neste exemplo simples pode-se concluir que não existirá uma única árvore de decisão para um mesmo problema, pois com diferentes árvores se chegará a um mesmo resultado.

2.10.4 Decision Forest Regression (DFR)

Decision Forest Regression (DFR), é um modelo de regressão que consiste em um conjunto de árvores de decisão. Cada árvore em uma floresta de decisão produz uma distribuição Gaussiana como uma previsão. Uma agregação é realizada para encontrar uma distribuição gaussiana mais próxima da distribuição combinada para todas as árvores do modelo. Cada árvore de decisão, por sua vez, possui modelo não paramétrico e realiza uma sequência de testes simples para cada instância, atravessando uma estrutura binária de dados de árvores até que um nó de ramo (decisão) seja alcançado.

As árvores de decisão têm algumas vantagens:

- São eficientes tanto na computação como no uso da memória no treino e na previsão
- Podem representar limites de decisão não lineares.
- Realizam tanto seleção quanto classificação integradas de recursos, sendo resistentes na ocorrência de características de ruídos.

2.10.5 Gradient Boosting Regression (GBR)

O algoritmo *Gradient Boosting Regression* é um algoritmo de otimização baseado na função de erro. *Gradiente boosting* é uma tecnologia de aprendizagem de máquinas usada para resolver problemas de classificação e regressão. Gerando modelo de previsão forte integrando modelos de previsão fracos, como a árvore de decisão, onde sua principal característica é que cada cálculo é feito por um modelo básico, o próximo cálculo buscará reduzir o resíduo de tais modelos e criará um novo modelo básico na direção do gradiente com resíduos reduzidos (WEI,2019).

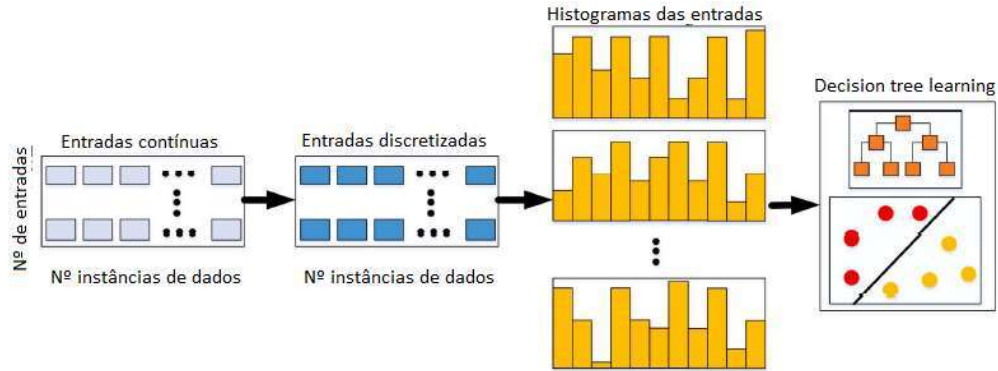
Resumidamente, *Gradient Boosting* é um algoritmo de Boosting que busca a criação de novos modelos, ajustando-os consecutivamente, a fim de obter um melhor resultado para o que se deseja observar. Para isto, busca-se que os novos aprendizes estejam com correlação maximizada em relação ao gradiente negativo da função de perda associado ao conjunto completo (NATEKIN; KNOLL, 2013).

2.10.6 Histogram Gradient Boosting Regression (HGBR)

Histogram Gradient Boosting é um algoritmo baseado em histograma para construção de uma estrutura de árvore. *Gradient Boosting* refere-se a um conjunto de modelos de aprendizado de máquina baseados em árvore de decisão como aprendizes básicos. Uma árvore subsequente depende de seus predecessores, a última árvore visa reduzir os casos mal classificados cometidos pela primeira, com esse processo a função de perda representada pela taxa de erro de todo o conjunto é gradualmente diminuída durante a fase de treinamento do modelo. O algoritmo baseado em histograma representado na Figura 24, é um método eficaz para treinamento de modelo baseado em Gradient Boosting. Com este algoritmo os intervalos de entradas contínuas usados pelas árvores de decisão são discretizadas em pequenos compartimentos (bins), esses compartimentos são subsequentemente usados para construir histogramas que representam a distribuição dos valores das entradas. Estatísticas básicas, incluindo o número de instâncias de dados e a soma dos gradientes em cada caixa pode ser calculada. Com base nestes índices, os pontos de divisão ideais usados para treinar os alunos de base podem ser determinado. O algoritmo baseado em histograma pode reduzir significativamente o custo computacional porque a fase de treinamento de uma árvore de decisão não exigem a varredura de todas as faixas de entradas para ponto de divisão avaliado. Além disso, o algoritmo baseado em histograma também pode ajudar a conseguir um melhor

desempenho porque a fase de aprendizagem é menos suscetível ao ruído.

Figura 22 - Demonstração de algoritmo baseado em histograma



Fonte: HOANG, 2023

2.11 ESTUDO DA CORRELAÇÃO ENTRE ATRIBUTOS

2.11.1 Coeficiente de Pearson

No estudo da análise de dados, algumas vezes é necessário entender a relação entre dois ou mais atributos. A análise de correlação é uma forma descritiva que mede o grau de dependência entre variáveis, ou melhor o quanto uma variável interfere em outra. Essa medida é feita através de coeficientes. O coeficiente de *Pearson*, dado na Equação 2.22, é usado para encontrar a relação entre os atributos x e y com base no valor r , nos ajudando a descobrir o quanto um atributo está associado a outros atributos (4.1).

$$r = \frac{N \cdot \sum xy - \sum x \cdot \sum y}{\sqrt{(N \cdot \sum x^2 - (\sum x)^2) \cdot (N \cdot \sum y^2 - (\sum y)^2)}}. \quad (2.22)$$

Para análise do coeficiente de Pearson foram utilizadas as bibliotecas *pandas* e *seaborn* do *Python*. A biblioteca *pandas* foi utilizada para manipulação matemática e a biblioteca *seaborn* para visualização do gráfico heatmap, ou mapa de calor referente ao mapeamento estatístico.

O coeficiente r , assume apenas valores entre -1 e 1, onde:

- Correlação positiva: quando duas variáveis possuem uma relação direta, crescem ou decrescem juntas.
- Correlação negativa: quando duas variáveis possuem uma relação indireta, quando uma variável cresce a outra decresce, ou vice-versa.

- Não possuem correlação: quando o crescimento ou decréscimo de uma variável não tem efeito sobre a outra variável.

Segue na Tabela 6 os valores possíveis para 'r' e correlação.

Tabela 4 - Interpretando os valores do coeficiente de Pearson

Intervalo de 'r'	Sinal	Correlação
0.9 a 1	Positivo ou negativo	Muito forte
0.7 a 0.89	Positivo ou negativo	Forte
0.4 a 0.69	Positivo ou negativo	Moderada
0.2 a 0.39	Positivo ou negativo	Fraca
0 a 0.19	Positivo ou negativo	Bem fraca

Fonte: autora

2.11.2 Coeficiente de Spearman

Na correlação de Spearman entre duas variáveis é idêntico ao valor de classificação de duas variáveis na correlação de Pearson, enquanto as relações lineares são avaliadas na correlação de Pearson, as relações monotônicas são avaliadas na correlação de Spearman, conforme Equação 2.23:

$$\rho = 1 - \frac{6 \sum d_i^2}{n(n^2 - 1)}, \quad (2.23)$$

onde:

ρ : é o coeficiente de Spearman.

n: é o número de amostras,

$\sum d_i^2$: é a soma do quadrado das variáveis pareadas.

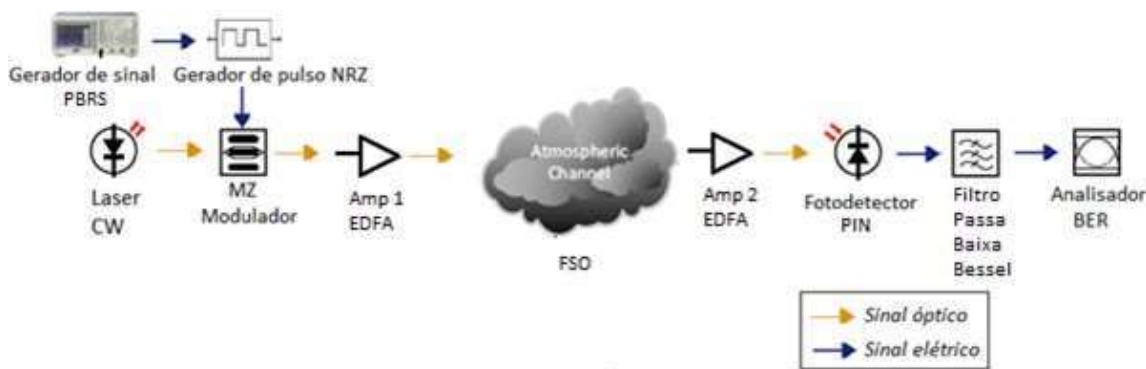
3 METODOLOGIA

Este capítulo descreve a análise de desempenho de uma rede FSO com a previsão do fator de qualidade (Factor Q). Foi utilizado o software *Optisystem*, versão 7, para simulação dos dados reais, na sequência foram utilizados algoritmos de aprendizagem de máquinas visando treinar a base de dados gerada pelo Optisystem para prever o fator de qualidade.

3.1 SIMULAÇÃO DE UM CANAL (FSO)

O software OptiSystem foi a ferramenta computacional utilizada nesse trabalho, para a simulação de uma Rede Óptica sem fio FSO (Figura 21).

Figura 23 - Sistema de comunicação FSO simulado no OptiSystem



Fonte: autora

Na representação esquemática de FSO apresentada na Figura 21, pode-se observar que foi utilizado um transmissor, um canal FSO e um receptor para a simulação do FSO.

O transmissor inclui um gerador de sinais (PBRs), um gerador de pulso NRZ, uma fonte de laser CW, um modulador *Mach Zehnder* e um pré amplificador EDFA. Os dados mostrados pelo gerador PBRs são os dados elétricos que se deseja transmitir no projeto a uma taxa de 10 Gbps. Estes dados são equivalentes a uma sequência de bits na forma pseudoaleatória de 0 e 1, que são convertidos em outra sequência de bits digitais no formato NRZ - No Return to Zero. O laser CW atua como fonte portadora do sinal óptico, com comprimento de onda de 1550nm e potência de 20dBm. No modulador MZ, uma sequência de bits elétricos é aplicada na entrada do mesmo, e combinada com o sinal óptico da fonte de laser CW, produzindo uma réplica óptica de sequência de bits na saída do modulador, conhecido como efeito eletro-óptico do modulador MZ. Essa luz modulada é amplificada, usando o pré-amplificador EDFA, e transmitida pelo canal FSO através do espaço livre e será enviada diretamente ao receptor óptico.

De fato, o canal óptico no espaço livre se resume a dois telescópios com o canal FSO entre eles. O diâmetro de abertura do receptor FSO será um dos valores alterados nesse projeto, assim como a divergência do feixe.

Os sinais ópticos do canal FSO são amplificados novamente na recepção através de outro EDFA, amplificador de fibra dopada com érbio. Na sequência, este sinal será recebido pelo fotodetector PIN que converte o sinal óptico de volta à forma elétrica e recupera os dados transmitidos pelo sistema de onda luminosa, depois tem-se um filtro passa-baixa Bessel (LPF), o LFP usado com uma frequência de corte de $(0,75 \times \text{symbol rate})$ serve para filtrar os sinais de ruído. Da biblioteca de visualizadores elétricos do Optisystem foi utilizado um analisador de BER na saída do sistema FSO para medição da qualidade do serviço através do fator Q.

Neste trabalho foram empregadas como entrada de dados seis (6) variáveis, e, como saída, duas variáveis, conforme descrito na Figura 22.

Figura 24 - Representação esquemática das entradas e saídas na simulação do canal FSO



Fonte: autora

Os dados de entrada na simulação variaram conforme Tabela 4.

Tabela 5 - Valores de entradas utilizados na simulação do canal FSO

Entradas	Valores
Distância (km)	1 a 10
Atenuação (dB/km)	0.11 ; 4 ; 9 ; 13 ; 16 ; 22
Amplificadores (dB)	6 a 10 ; 16 a 20 ; 26 a 30
Diâmetro de abertura receptor (cm)	15 ; 20 ; 25
Divergência do feixe (mrad)	1 ; 2 ; 3

Fonte: autora

A atenuação do espaço livre variou com o nível da visibilidade de acordo com as

condições climáticas diversas descritas na Tabela 5.

Tabela 6 - Coeficientes de atenuação de acordo com as condições climáticas.

Atenuação (dB/km)	Visibilidade FSO
0.11 dBm	Tempo Claro
4 dBm	Neblina
9 dBm	Neblina Fina
13 dBm	Neblina Leve
16 dBm	Neblina Forte
22 dBm	Neblina Espessa

Fonte: autora

Considerando o que foi visto anteriormente, foram realizadas duas simulações implementadas no software Optisystem variando os parâmetros de entrada:

- 9.000 simulações com as entradas: distância, atenuação, ganho amp1 e ganho amp2;
- 16.000 simulações com as entradas: distância, atenuação, ganho amp1, ganho amp2, diâmetro da abertura do receptor e divergência do feixe. As duas últimas entradas foram inseridas para estudo da relação entre o diâmetro da abertura do receptor e a divergência do feixe.

Posteriormente, os dados obtidos das simulações foram utilizados para o treinamento da base de dados em questão, usando vários modelos de aprendizagem de máquinas.

3.2 INTELIGÊNCIA ARTIFICIAL

Técnicas de aprendizado de máquina foram utilizadas para prever o parâmetro fator Q da rede FSO. Da plataforma Azure ML Classic foram usados os algoritmos regressores *Boosted Decision Tree Regression* (BDTR) e *Decision Forest Regression* (DFR), já os algoritmos *Histogram Gradient Boosting Regression* (HGBR), *Gradient Boosting Regression* (GBR) e *AdaBoost + Decision Tree Regression* (AdaBoost + DTR) foram usados da plataforma Anaconda Navigator, mais especificamente no IDE Jupyter notebook, com a linguagem Python versão 3.9. A plataforma Google Colaboratory foi também utilizada, pois utiliza a internet reduzindo o tempo de processamento.

Antes da aplicação do algoritmo para treinamento do modelo que realizaria a predição,

dividiu-se a base de dados em dois blocos: o primeiro para treinamento do modelo e a segunda para avaliação do mesmo. De maneira geral, costuma-se utilizar uma relação de 70% da base para treinamento e 30% da base para avaliação ou, ainda, 80% da base para treinamento e os 20% restantes para teste. Neste estudo foi separado, 80% da base para treinamento e os 20% restantes para teste. A parte separada para avaliação será utilizada para verificação por meio de parâmetros estatísticos. Estes parâmetros estatísticos representam o quanto o modelo, após a fase de aprendizado, está se adequando ao problema que lhe é apresentado. No estudo foram utilizadas a estatística R^2 ou coeficiente de determinação e MAE (Mean Absolute Error).

Também foi necessária a escolha de quais colunas da tabela seriam utilizadas como base para a previsão do valor de outra coluna. Neste ponto foram selecionadas as colunas base (as chamadas features) referentes a: distância (km), atenuação, ganho AMP1 e ganho AMP2. Já no tocante à coluna que é necessário prever o valor (chamada de alvo), foi escolhida a relacionada ao fator Q de início. Posteriormente foram incluídas mais duas colunas: diâmetro de abertura do receptor e divergência do feixe, que serão utilizadas como features adicionais. A coluna alvo que deseja-se prever o valor permanece a mesma: fator Q. Adiante será comparada a eficiência para os dois casos em questão.

4 RESULTADOS E DISCUSSÃO

No presente capítulo são apresentados os resultados em forma de gráficos e tabelas, relativo ao fator de qualidade de serviço: *MaxQFactor* (Máximo Q Fator) obtido das implementações no Optisystem da rede óptica utilizando FSO. Analisa-se a correlação entre as variáveis de entrada e o alvo (valor que se deseja prever) utilizadas na IA, assim como a implementação de gráficos comparando os diversos tipos de algoritmos de Inteligência Artificial utilizados. Foi observada a efetividade dos algoritmos de inteligência artificial através do coeficiente de determinação R^2 e do MAE (*Mean Absolute Error*). Foram verificadas também as diferenças entre o valor real e o valor predito para os algoritmos que tiveram melhores resultados nas duas simulações implementadas no Optisystem.

4.1 ANÁLISE DE CORRELAÇÃO

4.1.1 Mapa de calor para Pearson e Spearman nos cenários 1 e 2

Foi apresentado o mapa de calor para os coeficientes de Pearson e Spearman, verificando a correlação entre as variáveis de entrada, ou seja tudo aquilo que se tomará como base para encontrar o valor do alvo e o alvo, aquilo que se busca prever o valor. Considerando o primeiro cenário, onde foram usados os atributos abaixo:

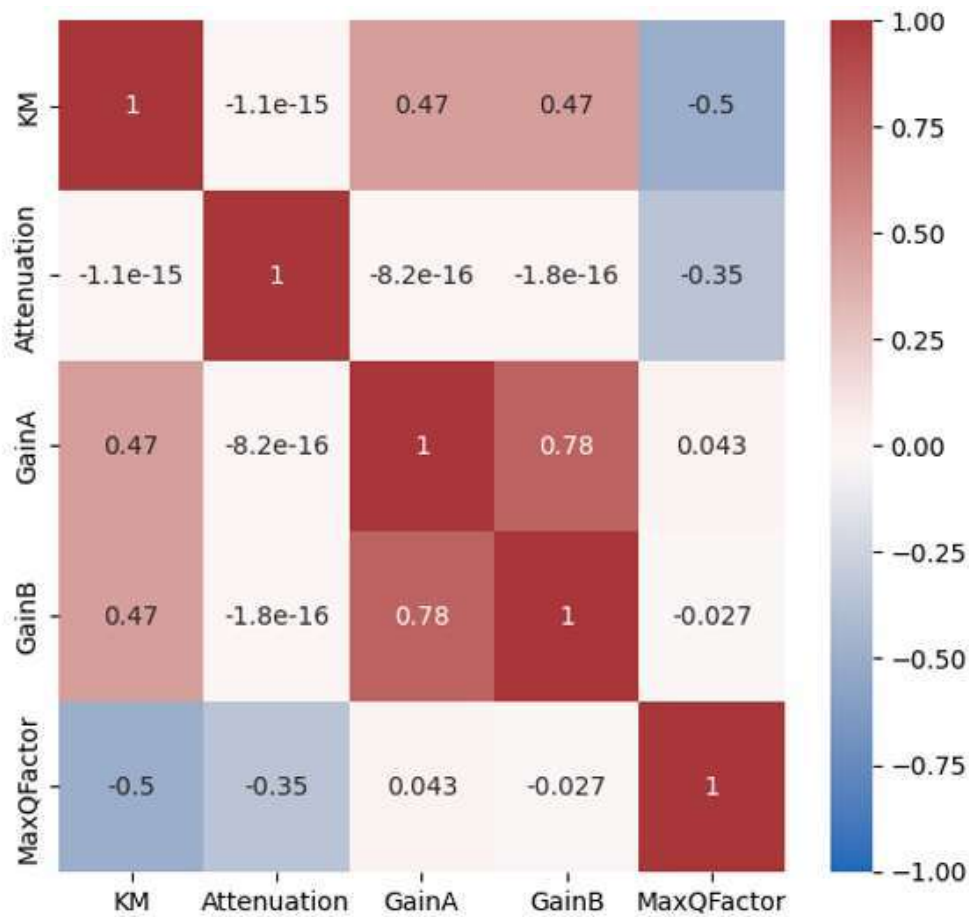
- Entradas: Distância (km), atenuação (dB/km), ganho 1 (dB), ganho 2 (dB)
- Saída: Max Q fator

De acordo com as Figuras 25a e 25b, foi demonstrado que os coeficientes de Pearson e Spearman possuem resultados bem parecidos. Checando a correlação entre o Max Q fator e as quatro entradas tem-se:

- possui com a distância (km): correlação moderada negativa (r) e moderado relacionamento monotonicamente decrescente (ρ)
- possui com a atenuação(dB/km): correlação fraca negativa (r) e forte relacionamento monotonicamente decrescente (ρ) . Note que aqui ocorreu uma diferença entre os dois coeficientes.
- possui com os ganhos 1 e 2: correlação bem fraca positiva e negatva, respectivamente (r) e bem fraco relacionamento monotonicamente crescente e decrescente, respectivamente (ρ) .

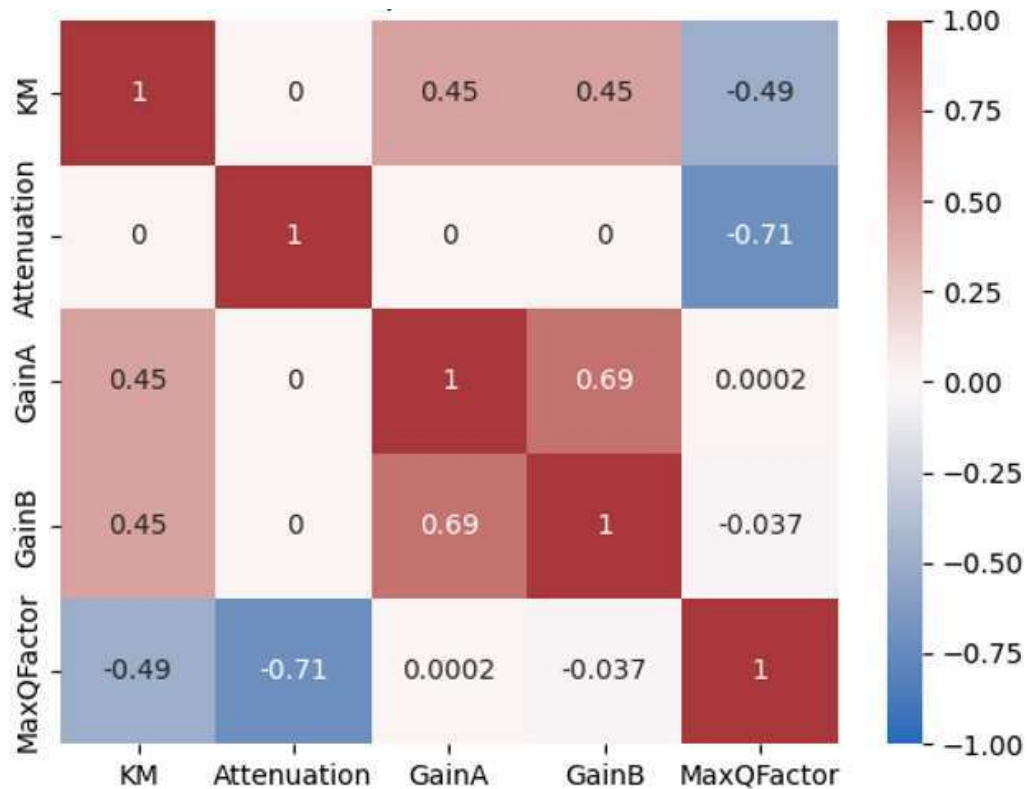
Pode ser destacado que o coeficiente de Spearman se mostrou mais robusto ao inferir a não correlação entre as variáveis, observe que foi apresentado ‘ $\rho = 0$ ’ para diversos pontos do mapa de calor, enquanto que Pearson apresenta valores próximos a zero, nessas condições.

Figura 25a- Mapa de calor para o cenário 1 Pearson



Fonte: autora

Figura 25b - Mapa de calor para o cenário 1 Spearman



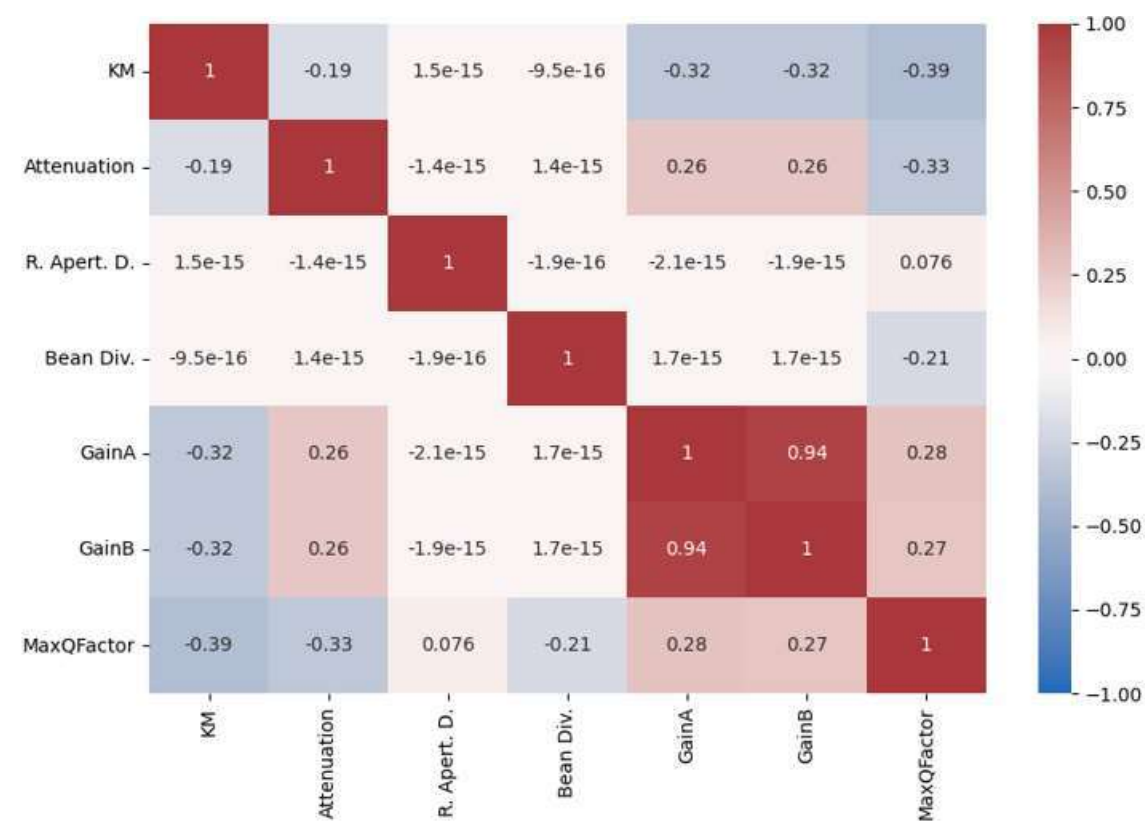
Fonte: autora

Analizando agora, o mapa de calor gerado para quando foi trabalhado com os dados do cenário 2, ou seja acrescentando as entradas: diâmetro de abertura do receptor e divergência do feixe. Foi concluído pelas Figuras 25c e 25d que a saída Max Q Fator possui:

- com a distância e atenuação: correlação fraca negativa (r) e fraco relacionamento monotonicamente decrescente (ρ)
- com os ganhos: correlação fraca positiva (r) e fraco relacionamento monotonicamente crescente (ρ)
- com a divergência do feixe: correlação fraca negativa (r) e fraco relacionamento monotonicamente decrescente (ρ)
- com o diâmetro de abertura do receptor: correlação muito fraca positiva e muito fraco relacionamento monotonicamente crescente (ρ)

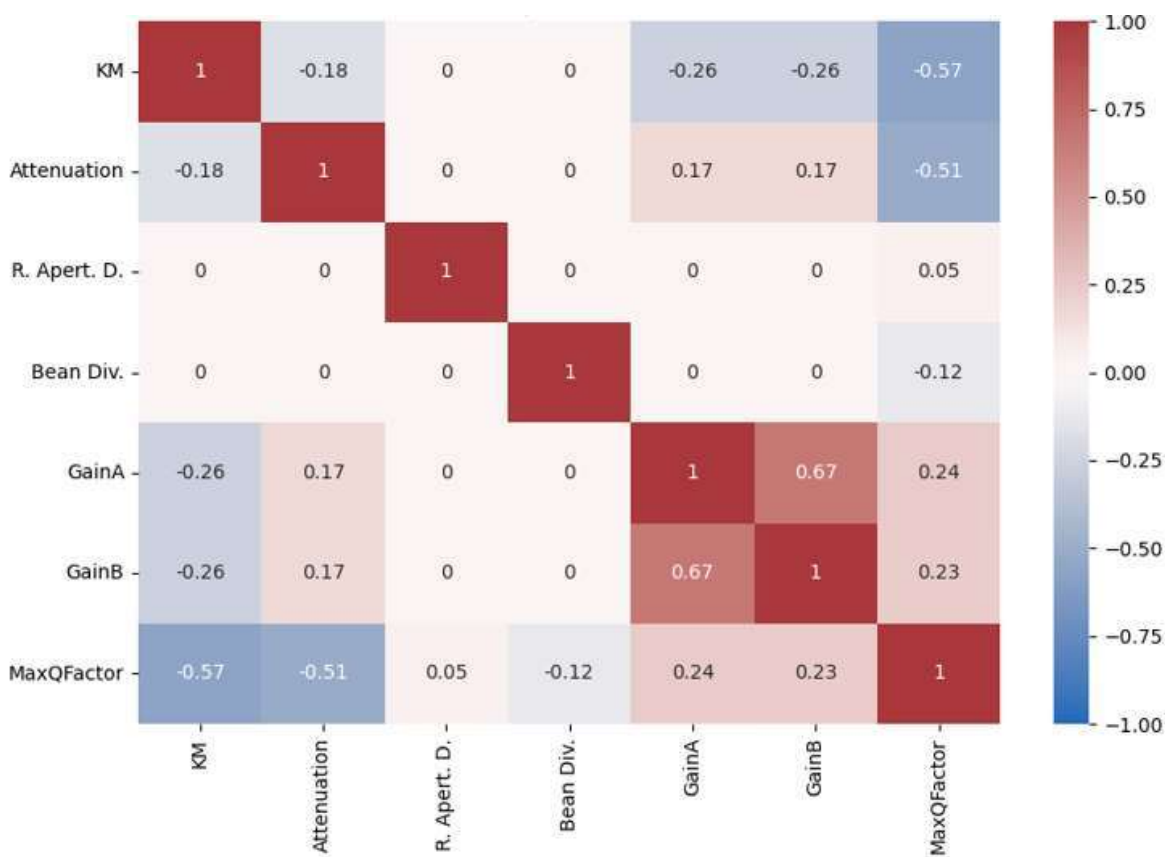
Novamente, pode ser observado que a matriz de correlação de Spearman foi mais robusta em relação a de Pearson na hora de inferir a não correlação entre as variáveis.

Figura 25c - Mapa de calor para o cenário 2 Pearson



Fonte: autora

Figura 25d - Mapa de calor para o cenário 2 Spearman

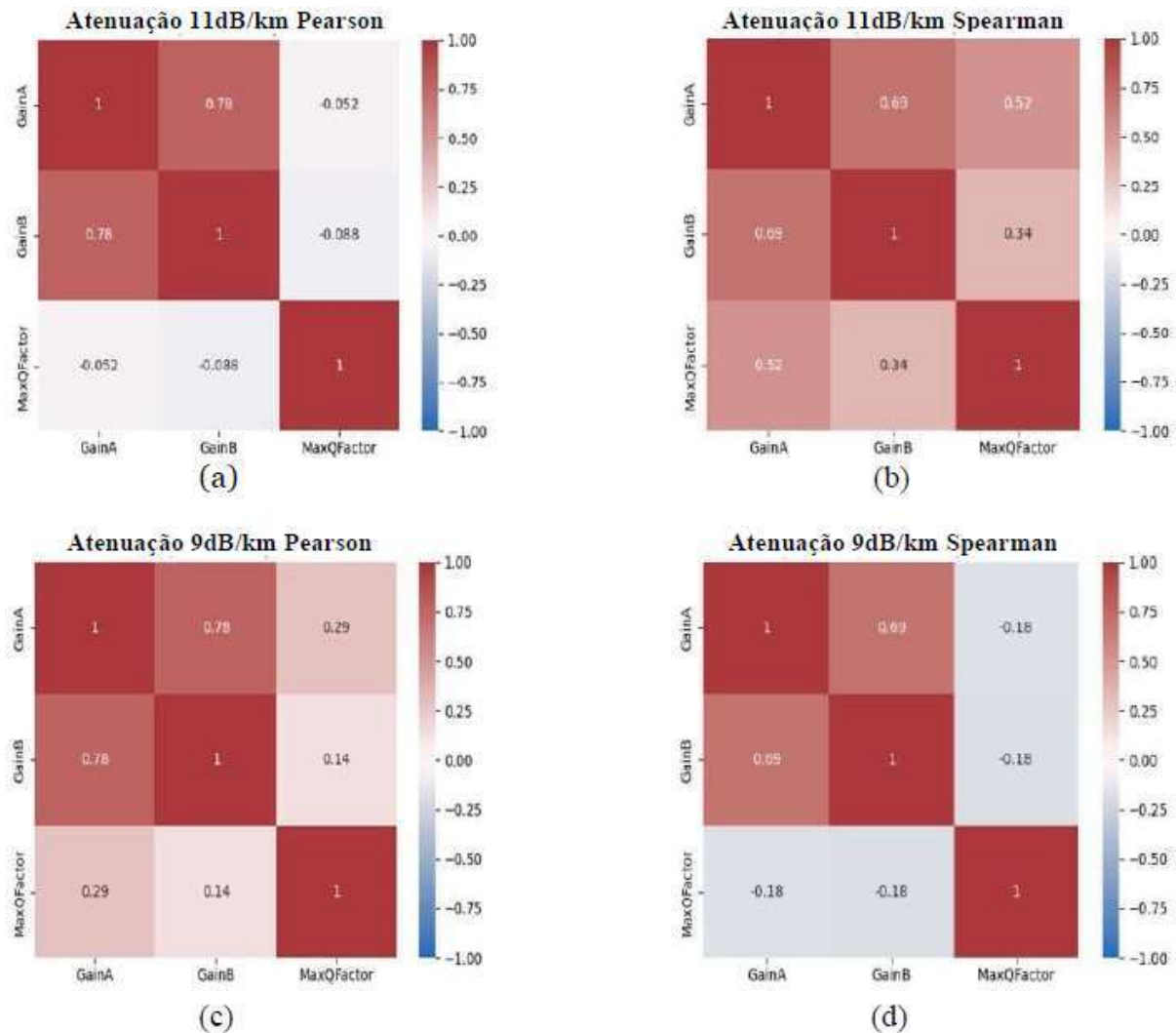


Fonte: autora

4.1.2 Análise de Pearson e Spearman para atenuação

Comparando o coeficiente de Pearson e Spearman para a atenuação 0.11 e 9.0 dB/km, e monitorando as variáveis de entrada: ganho A e ganho B e atributo de saída max Q fator, apresentado nas Figuras 26a, 26b para atenuação 0.11 dB/km e 26c e 26d para atenuação de 9dB/km. Como se pode observar entre as entradas ganho A e ganho B os valores de r e ρ para atenuação de 0.11 e 9 dB/km estão no intervalo de 0.7 a 0.9, portanto possuem uma forte correlação (r) ou forte relacionamento monotonicamente crescente (ρ). Por outro lado, observando os valores obtidos para a correlação das variáveis ganho A e ganho B em relação ao *MaxQFactor*, foi verificado que a correlação está diretamente ligada à atenuação do canal, ou seja para a atenuação de 0.11 dB/km tem-se os valores de 0.29 e 0.14 (Pearson) e 0.52 e 0.34 (Spearman), mas quando a atenuação aumenta para 9dB/km esses valores caem bastante para os dois coeficientes. Provando que a correlação está diretamente ligada a atenuação do canal, ou seja quanto maior a atenuação do canal FSO menor o coeficiente de Pearson, como também menor o coeficiente de Spearman. Importante observar este comportamento semelhante para os dois coeficientes.

Figura 26 - Mapa de calor comparando Pearson vs Spearman para atenuação



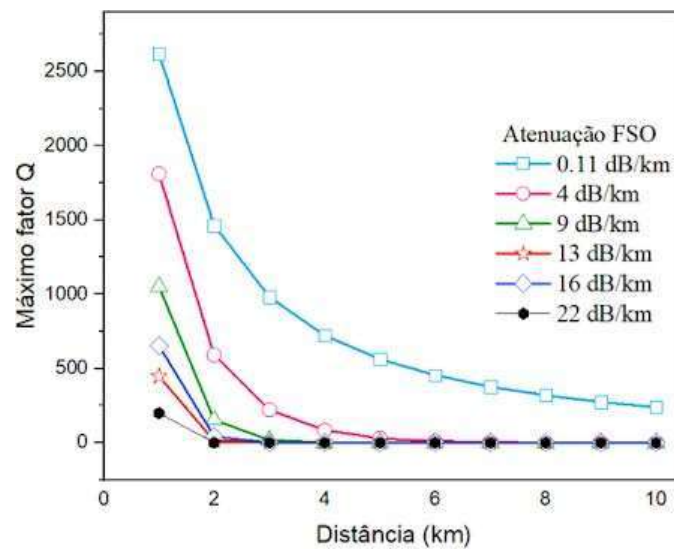
Fonte: autora

4.2 ANÁLISE DOS DADOS SINTÉTICOS DO OPTISYSTEM

4.2.1 Análise de atenuação no canal FSO

As curvas de atenuação simuladas no Optisystem são apresentadas na Figura 27 para condições diversas de neblina (Distância vs. Máximo Q Fator). Verificou-se que em condições climáticas adversas o canal FSO foi significativamente degradado.

Figura 27 - Atenuação na rede FSO



Fonte: autora

Nesta análise, os valores dos componentes foram fixados seguindo a Tabela 7.

Tabela 7 - Valores das variáveis de entradas fixados

Componentes	Valores Fixados
PBRS	10 Gbps
Laser CW	1550nm / 20 dBm
Amp1	6 dB
Amp2	10 dB
Diâmetro de abertura receptor (cm)	25
Divergência do feixe (mrad)	1

Fonte: autora

Investigando com mais afinco os valores para a distância de 2 km com atenuação de 22 dB/km, Tabela 8, percebe-se que a qualidade do serviço está zerada, não sendo possível o funcionamento do canal FSO nessas condições.

Tabela 8 - Valores para a distância 2 km

KM	Atenuação (dB/km)	Max. Q Factor
2	0.11	1.463,209516267330
2	4	589,712959664391
2	9	152,787243783064
2	13	41,258463539366
2	16	12,843519182866
2	22	0,000000000000

Fonte: autora

Para a distância de 3 km, com as atenuações setadas em 13, 16 e 22 dB/km, a qualidade do serviço também foi perdida nessas condições, não sendo possível também o funcionamento do canal FSO (Tabela 9):

Tabela 9 - Valores para a distância 3 km

Distância (km)	Attenuation (dB/km)	Max. Q Factor
3	0.11	980,981994852176
3	4	220,690357386678
3	9	17,545514296881
3	13	0,000000000000
3	16	0,000000000000
3	22	0,000000000000

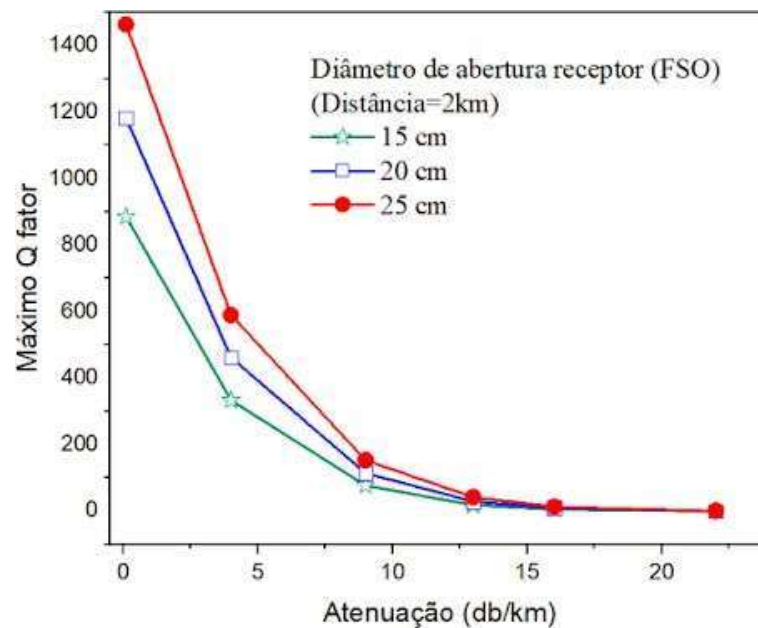
Fonte: autora

Com isso, demonstrou-se que a comunicação no espaço livre (FSO) é uma solução interessante para entrega da comunicação de última milha, garantindo uma alta taxa de dados. No entanto, o desempenho sofre bastante com as alterações climáticas na interface aérea. O termo última milha se refere às extremidades de uma rede de telecomunicação, onde acontece o consumo do serviço contratado pelo cliente final.

4.2.2 Análise dos parâmetros internos do FSO: diâmetro de abertura do receptor e divergência do feixe

Serão apresentadas nesta seção as curvas do diâmetro de abertura do receptor simuladas no *Optisystem*, Figura 28, variando de 15 a 25 cm (Atenuação vs. Máximo Q fator). Foi verificado que quanto maior o parâmetro de entrada ‘Diâmetro de abertura do receptor’ melhor a qualidade do sinal, ou seja maior o parâmetro de saída ‘Máximo Q fator’.

Figura 28 - Parâmetro interno FSO: Diâmetro de abertura receptor



Fonte: autora

Tabela 10 - Valores das variáveis de entradas fixados

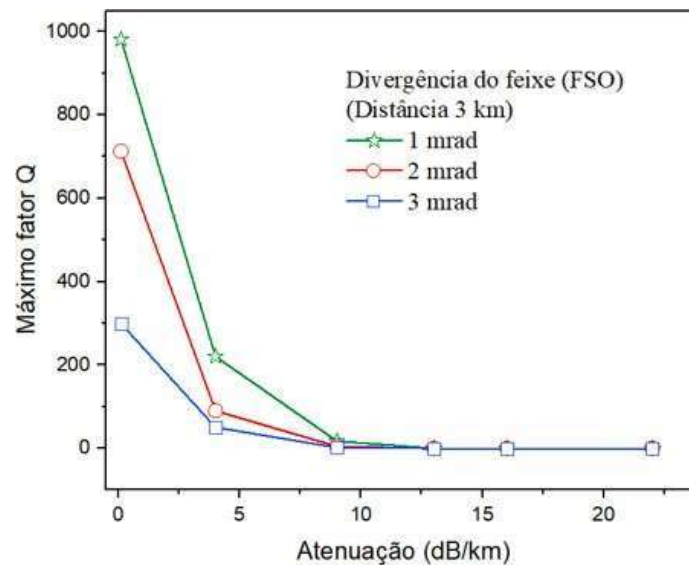
Entradas	Valores Fixados
PBRs	10 Gbps
Laser CW	1550nm / 20 dBm
Amp1	6 dB
Amp2	10 dB
Distância (km)	2
Divergência do feixe (mrad)	1

Fonte: autora

Nesta análise, os valores para os componentes foram fixados segundo a Tabela 10 .

Na sequência, foram apresentadas, na Figura 29, as curvas de divergência do feixe simuladas no *Optisystem*, variando de 1 a 3 mrad (Atenuação x Máximo Q Fator). Nota-se que quanto menor o parâmetro de entrada ‘Divergência do feixe’ melhor o desempenho do link, ou seja maior o parâmetro de saída ‘Máximo Q fator’.

Figura 29 - Parâmetro interno FSO: Divergência do feixe



Fonte: autora

Tabela 11 - Valores fixados

Entradas	Valores Fixados
PBRS	10 Gbps
CW	1550nm / 20 dBm
Amp1	6 dB
Amp2	10 dB
Distância (km)	3
Diâmetro de abertura receptor (cm)	25

Fonte: autora

Nesta análise, os valores dos componentes foram fixados segundo a Tabela 11.

O comportamento dos gráficos mostrados nas Figuras 27, 28 e 29, pode ser explicado pela Equação 4.1 da potência no receptor:

$$P_R = P_T \exp(-\alpha L) \frac{A_{RX}}{(\theta L)^2}, \quad (4.1)$$

onde:

P_R = potência recebida

P_T = potência transmitida

A_{RX} = área de abertura do receptor

θ = ângulo de divergência

α = atenuação atmosférica

L = distância entre o transmissor e o receptor

Foi observado na Equação 4.1, que a potência recebida é diretamente proporcional à potência transmitida e ao ângulo de abertura do receptor, ou diâmetro de abertura do receptor, porém é inversamente proporcional ao ângulo de divergência, ou divergência do feixe, e à distância entre o transmissor e receptor. A parte exponencial da expressão é referente à atenuação atmosférica, que tem um forte impacto na qualidade do canal FSO. Com o estudo criterioso desses dois parâmetros internos, pode-se melhorar o desempenho da rede óptica utilizando FSO. Inclusive dispõe-se de artigo que foca neste tema (ALI et al., 2020).

4.3 TREINAMENTO DA BASE DE DADOS ATRAVÉS DA IA

Os dados obtidos, por intermédio da modelagem do FSO no *Optisystem*, foram utilizados em cinco diferentes algoritmos de IA:

- *Boosted Decision Tree Regression* (BDTR)
- *Decision Forest Regression* (DFR)
- *Gradient Boosting Regressor* (GBR)
- *Histogram Gradient Boosting Regression* (HGBR)
- *AdaBoost + Decision Tree Regression* (AdaBoost + DTR)

A efetividade do modelo treinado foi medida com base na estatística R^2 , coeficiente de determinação que é definido na Equação 4.2:

$$R^2 = \frac{SQ_{Tot} - SQ_{Rest}}{SQ_{Tot}} = \frac{SQ_{Reg}}{SQ_{Tot}} = 1 - \frac{SQ_{Rest}}{SQ_{Tot}}, \quad (4.2)$$

onde:

$$\begin{aligned} SQ_{Tot} &= \sum_{i=1}^n (y_i - \bar{y})^2 \text{ é a soma de quadrados total,} \\ SQ_{Rest} &= \sum_{i=1}^n (y_i - \hat{y})^2 \text{ a soma de quadrados do erro,} \\ SQ_{Reg} &= \sum_{i=1}^n (\hat{y}_i - \bar{y})^2 \text{ a soma de quadrados da regressão.} \end{aligned}$$

O coeficiente R^2 mede a porcentagem da variação total dos dados, em relação a sua média, explicado pelo modelo de regressão. Deve-se salientar que o coeficiente R^2 possui valores no intervalo $[0, 1]$ e, quando mais perto de 1 são os valores, a porcentagem da variação dos dados é menor em relação a sua média. Este coeficiente de determinação deve ser acompanhado de outras ferramentas para a avaliação do ajuste, pois o mesmo não consegue identificar se todas as suposições do modelo são compatíveis com os dados sob investigação.

O MAE (*Mean Absolute Error*), ou erro absoluto médio, em estatística é uma medida de erros entre observações paralelas que expressam o mesmo evento. Exemplos de Y *versus* X incluem comparações de previsto *versus* medido, tempo seguinte *versus* tempo inicial e uma técnica de medição *versus* uma técnica alternativa de medição. O MAE é calculado como a soma dos erros absolutos dividido pelo tamanho da amostra, conforme Equação 4.3:

$$MAE = \frac{\sum_{i=1}^n |y_i - x_i|}{n} = \frac{\sum_{i=1}^n |e_i|}{n}, \quad (4.3)$$

MAE é uma média aritmética dos erros absolutos, onde:

$$|e_i| = |y_i - x_i|,$$

y_i é o valor predito,

x_i é o valor real.

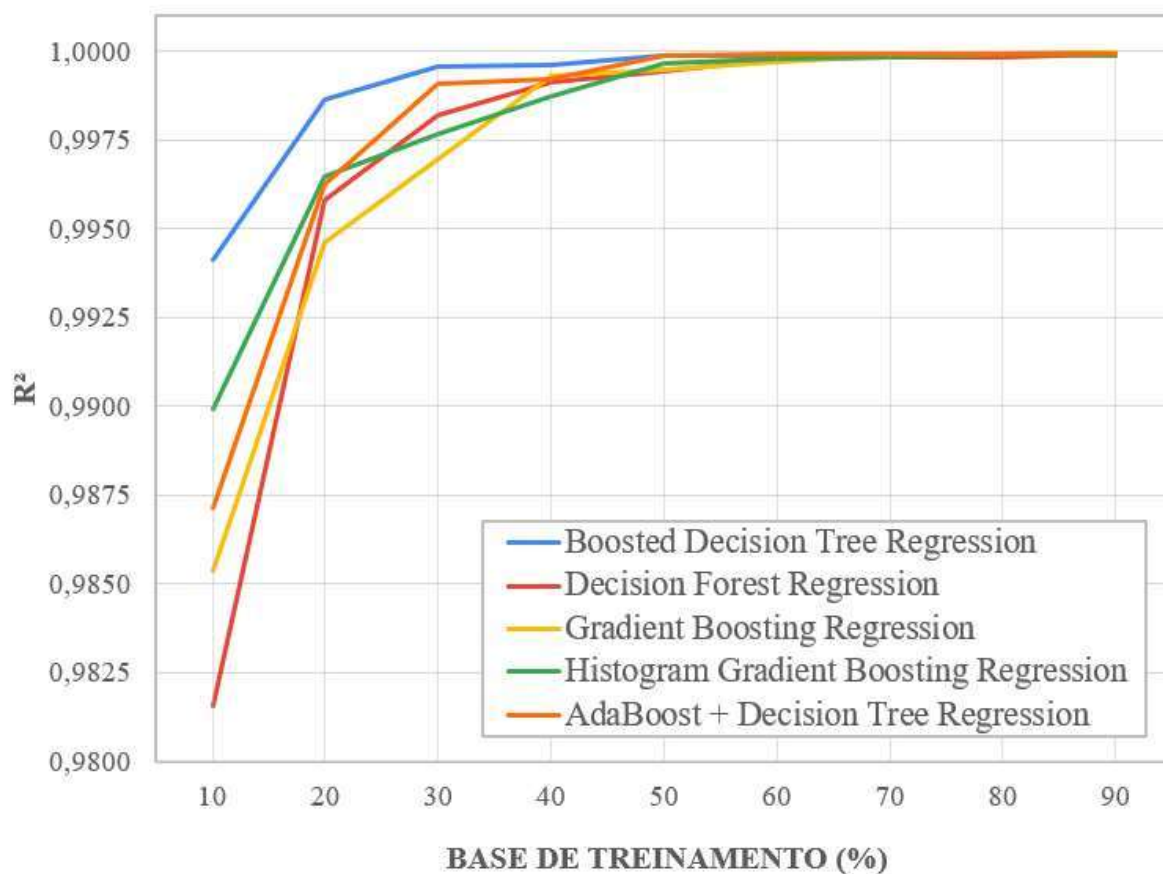
4.3.1 Treinamento da IA – Cenário 1

Foram levadas em consideração as seguintes premissas para o cenário 1:

- Foi utilizada 80% da base para treinamento e 20% restante para avaliação
- Foram usadas as entradas: distância (km), atenuação (dB/km), ganho 1 (dB), ganho 2 (dB)
- Saída: Max Q fator

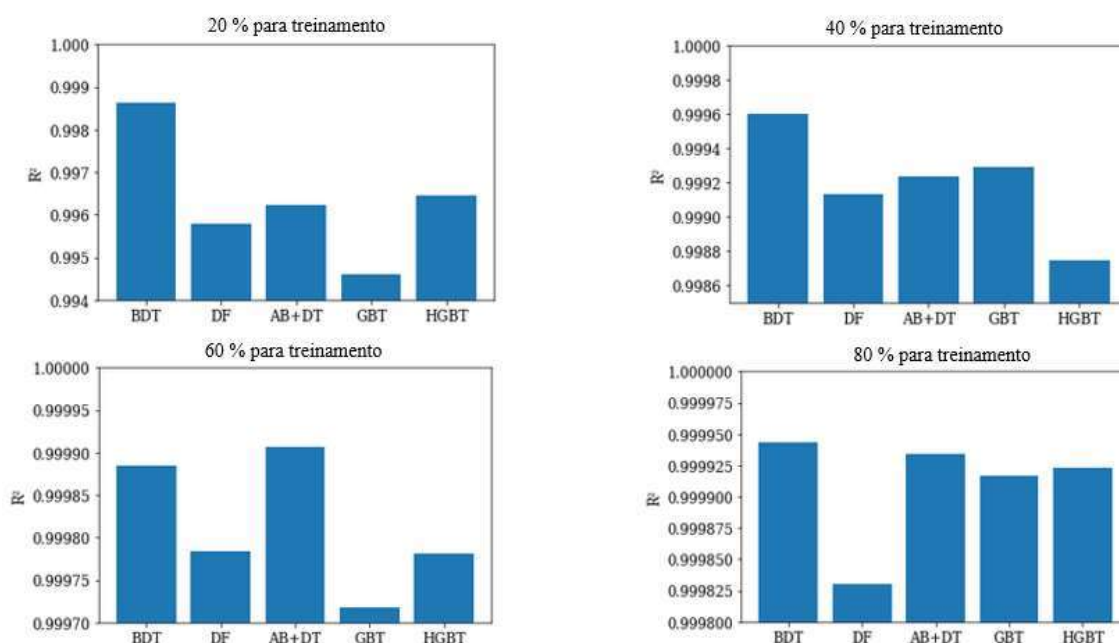
No gráfico apresentado na Figura 30, com o comparativo entre os cinco algoritmos de IA utilizados para o cenário 1, em termos da porcentagem do banco de dados que foi utilizada para treinamento do modelo e o coeficiente de determinação R^2 . Neste cenário os cinco algoritmos tiveram excelente coeficiente de determinação R^2 , com melhor resultado para os algoritmos BDTR e AdaBoost + DTR. Foi demonstrado na Figura 31 um gráfico de barras comparativo entre os R^2 de cada um dos algoritmos utilizados para treinamento, relacionando as porcentagens 20, 40, 60 e 80%, onde foi visualizado que os algoritmos BDTR e AdaBoost + DTR tiveram os melhores desempenhos. Essa conclusão será validada na sessão 4.4, quando serão comparados os valores REAL e PREDITO, caso necessário o estudo do MAE será inserido.

Figura 30- Gráfico base de treinamento vs coeficiente de determinação para os algoritmos de IA: BDTR, DFR, GBR, HGBR e AdaBoost + DRT – Cenário 1



Fonte: autora

Figura 31 - Gráfico de barras do coeficiente R2 para todos algoritmos por percentual de treinamento – Cenário 1



Fonte: autora

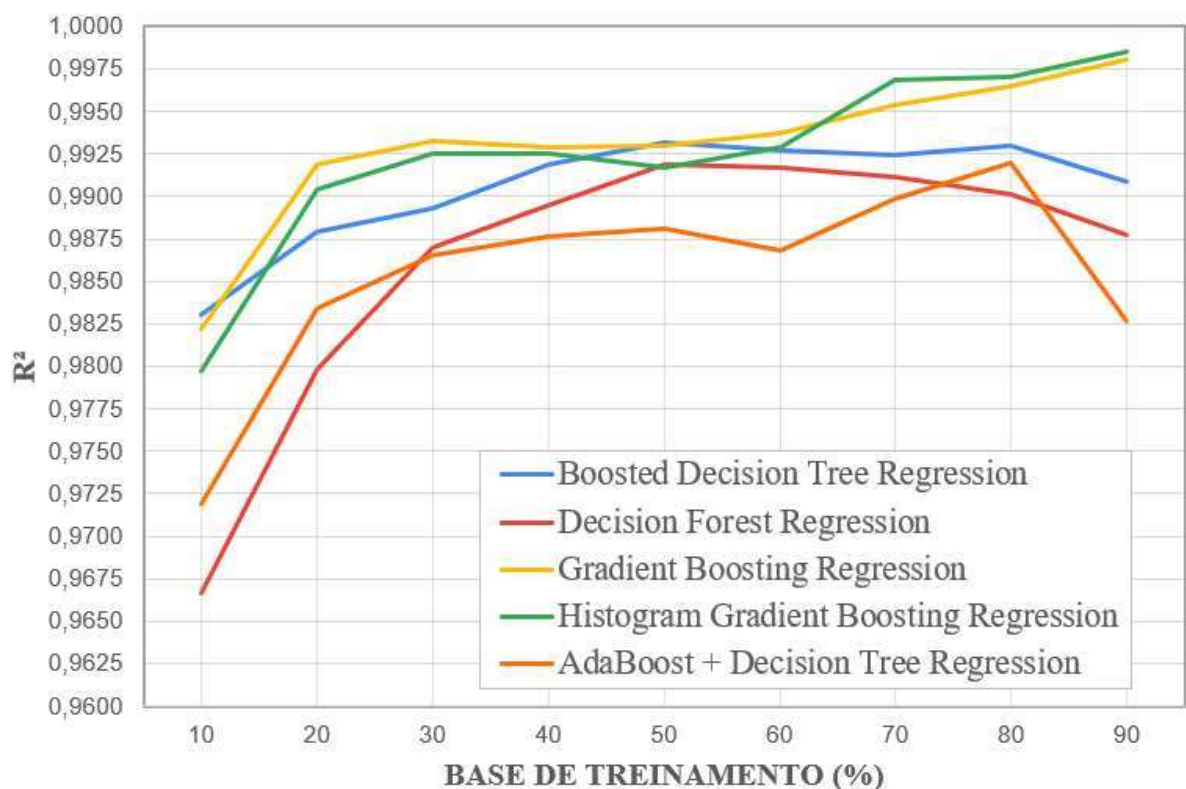
4.3.2 Treinamento da IA – Cenário 2

Foram levadas em consideração as seguintes premissas para o cenário 2:

- Foi utilizada 80% da base para treinamento e 20% restante para avaliação.
- Foram usadas as entradas: distância (km), atenuação (dB/km), ganho 1 (dB), ganho 2 (dB), diâmetro de abertura do receptor (cm) e divergência do feixe (mrad).
- Saída: Max Q fator

No gráfico apresentado na Figura 32, tem-se o comparativo entre os cinco algoritmos de IA utilizados para o cenário 2, em termos da porcentagem do banco de dados que foi utilizada para treinamento do modelo e o coeficiente de determinação R^2 . Quando são incluídas estas duas colunas (diâmetro de abertura do receptor e divergência do feixe), são percebidas algumas dificuldades no aprendizado, de uma maneira geral houve uma piora no coeficiente de determinação se comparado com o que estava sendo obtido sem a inclusão dessas duas novas entradas.

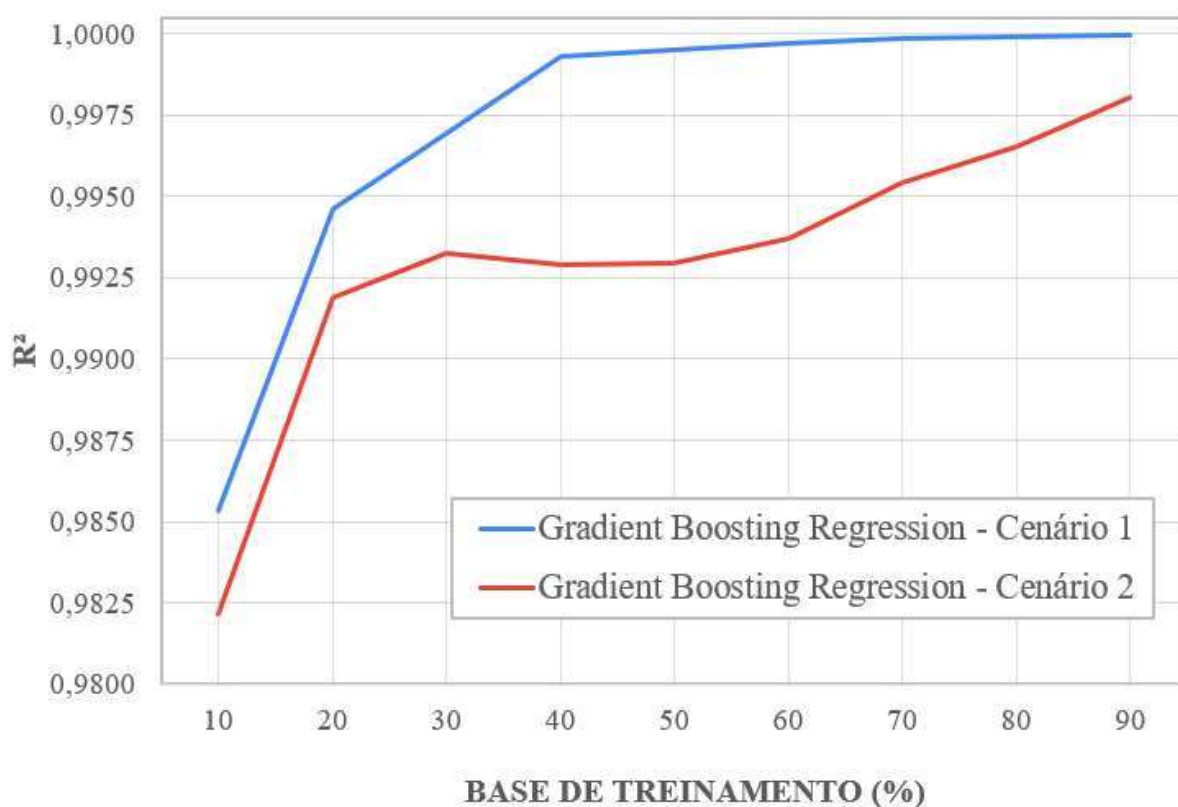
Figura 32 - Gráfico base de treinamento vs coeficiente de determinação R^2 para os algoritmos de IA: BDTR, DFR, GBR, HGBR e AdaBoost + DRT – Cenário 2



Fonte: autora

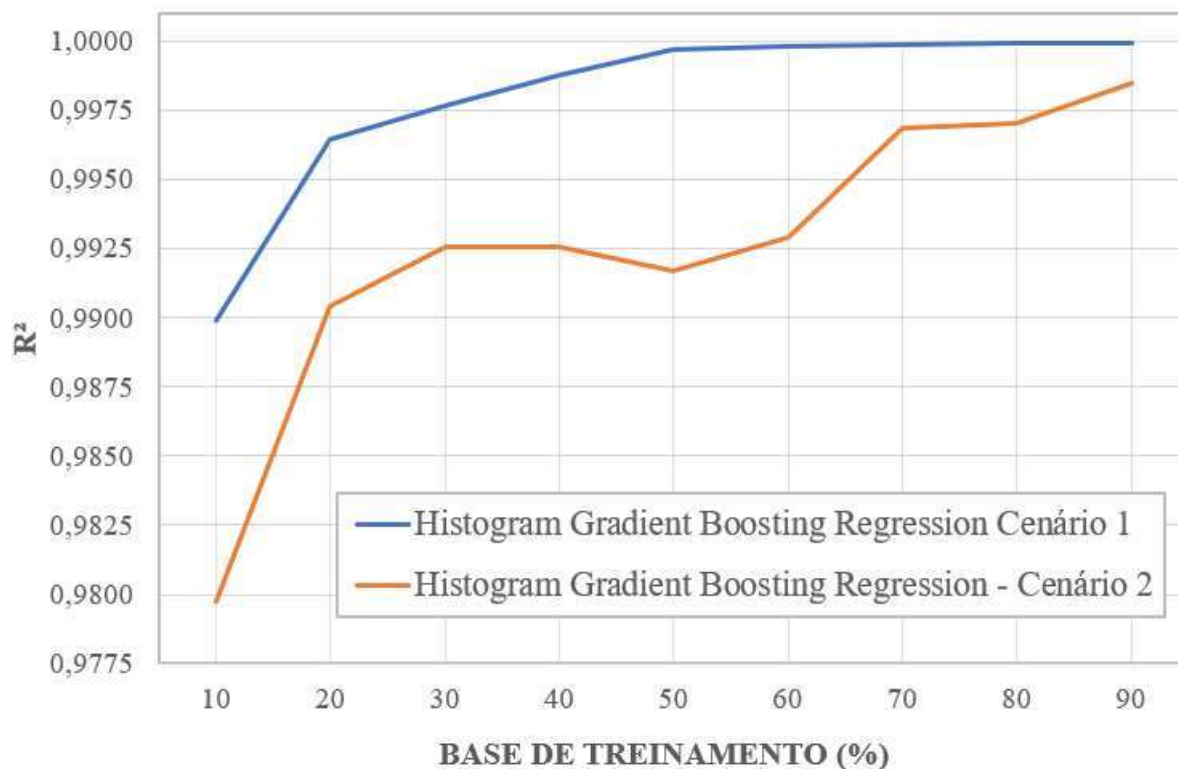
Comparando as figuras 30 e 32, há discrepâncias quanto a relação entre o percentual da base de dados utilizada e o quão bom o modelo se mostra após seu aprendizado. A partir de certa porcentagem, percebe-se a presença de uma maior saturação para o(s) modelo(s) apresentado(s). Desta forma, surge a necessidade de comparação de maneira direta dos resultados obtidos para um mesmo algoritmo com e sem estas duas novas entradas (Figuras 33, 34, 35, 36, 37). Foram observadas diferenças quanto à porcentagem da base de dados necessária para um bom aprendizado do modelo.

Figura 33 - Comparação entre os desempenhos de R^2 para o algoritmo GBR – Cenários 1 e 2



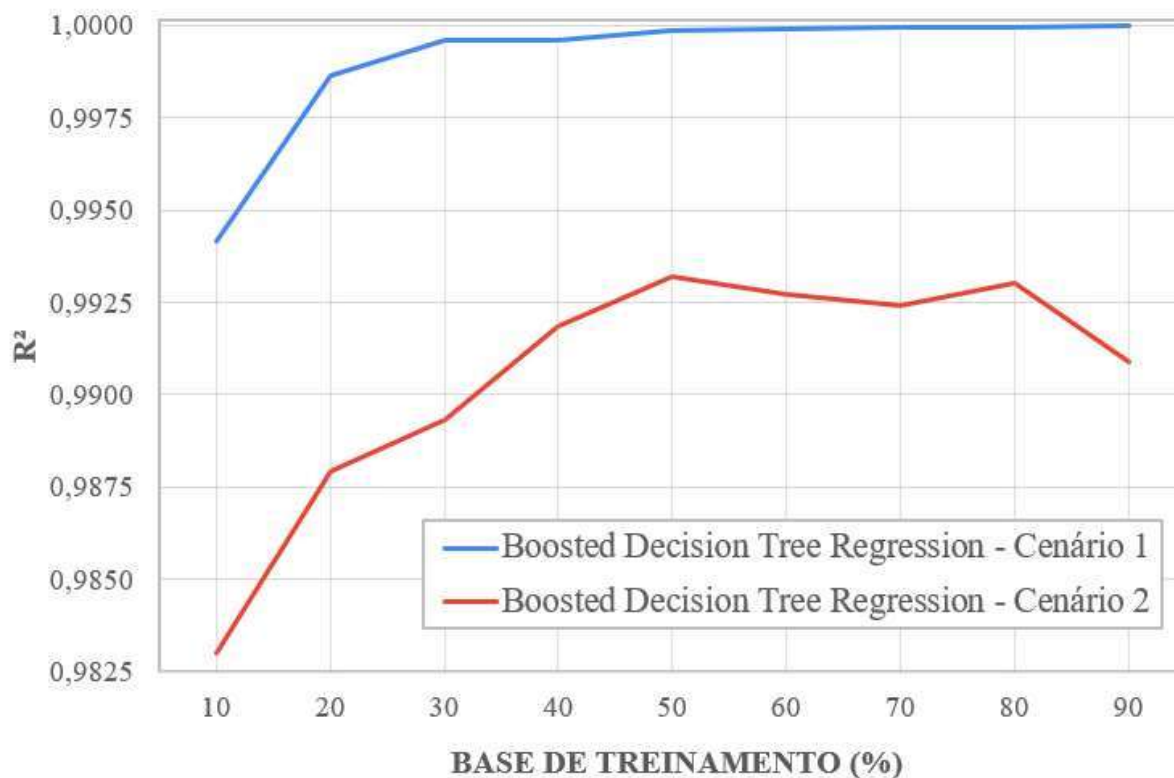
Fonte: autora

Figura 34 - Comparação entre os desempenhos de R2 para o algoritmo HGBR – Cenários 1 e 2

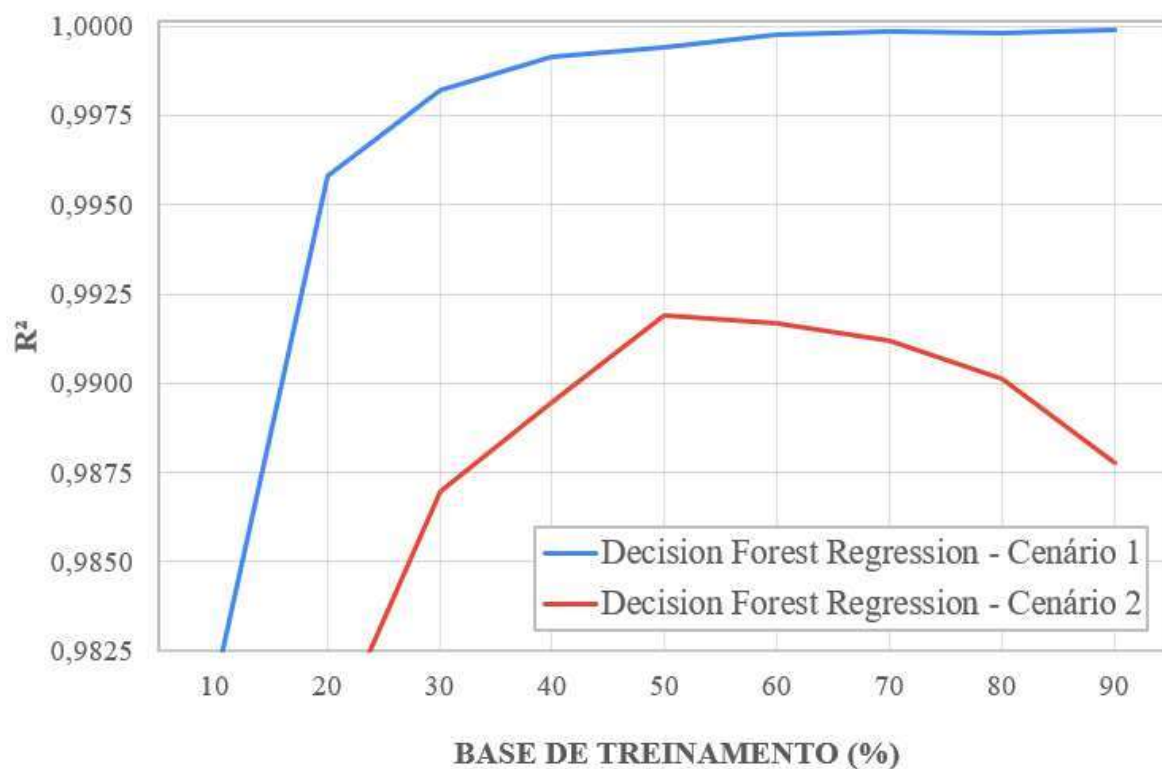


Fonte: autora

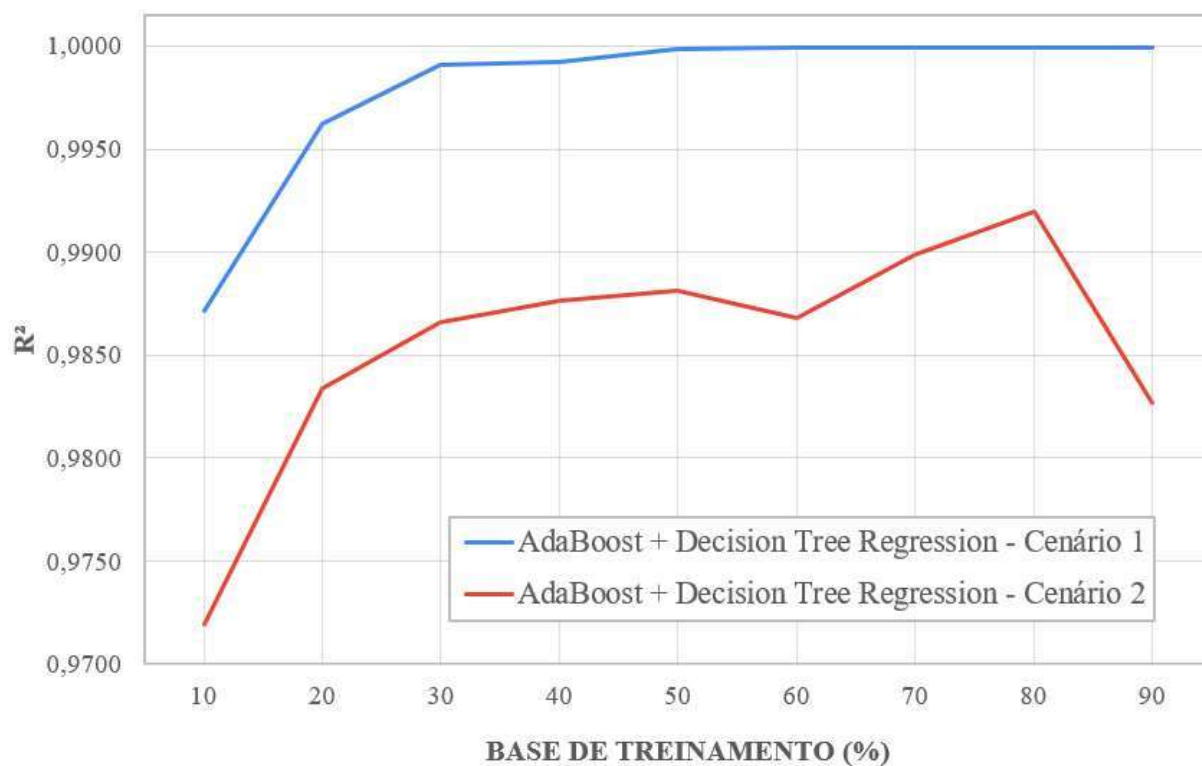
Figura 35 - Comparação entre os desempenhos de R2 para o algoritmo BDTR – Cenários 1 e 2



Fonte: autora

Figura 36 - Comparação entre os desempenhos de R^2 para o algoritmo DFR – Cenários 1 e 2

Fonte: autora

Figura 37 - Comparação entre os desempenhos de R^2 para o algoritmo AdaBoost + DTR – Cenários 1 e 2

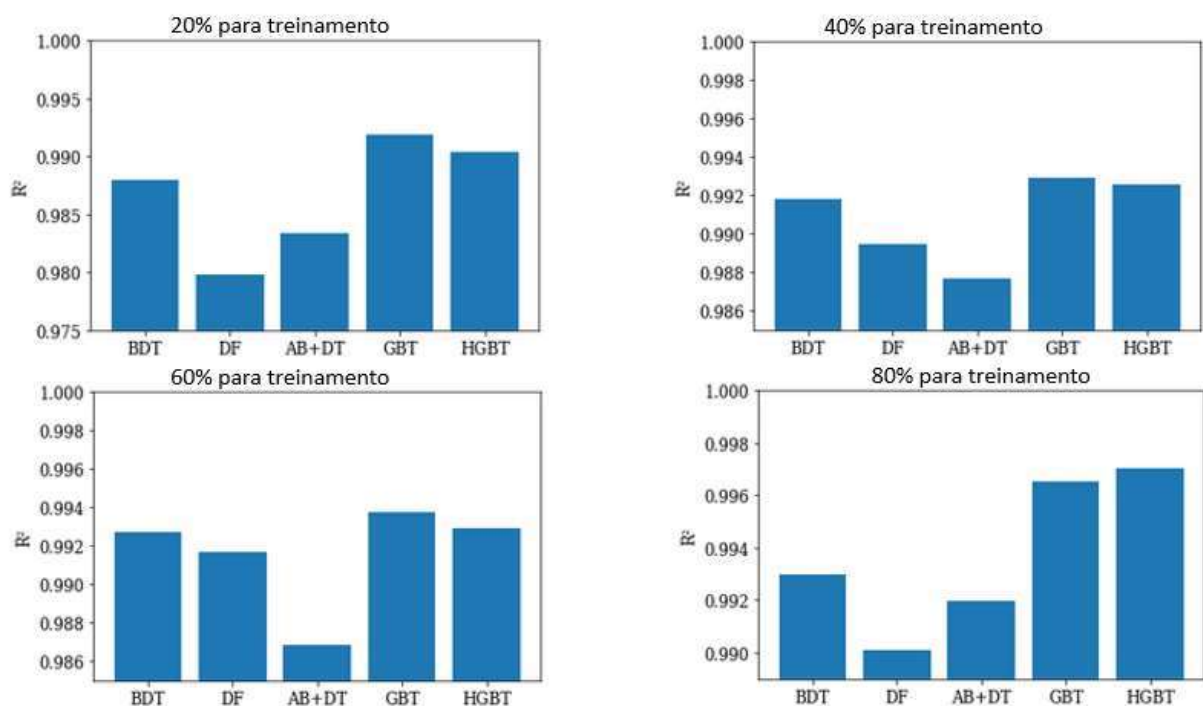
Fonte: autora

O comportamento dos gráficos da Figura 32, nos leva a concluir que quando adicionamos as entradas: diâmetro de abertura do receptor e divergência do feixe a nossos modelos de aprendizado estes tendem a apresentar dificuldades. Observa-se que em todos os casos existe a presença de uma saturação a qual se apresenta para diferentes porcentagens da base usada para o treinamento dependendo do algoritmo.

Os valores obtidos para R^2 apresentam uma diferença substancial quando a base de dados tem ou não a inclusão das duas novas variáveis. Por exemplo, nos métodos BDTR, DFR e AdaBoost + DTR quando a base de treinamento está 50%, o modelo reduz drasticamente o coeficiente de determinação, para os métodos GBR e HGBR esta redução ocorre nos 30% e 40%, respectivamente, mesmo assim o coeficiente de determinação alcançou um valor favorável para implementação da IA.

Foi mostrado na Figura 38 o gráfico de barras comparativo entre os R^2 de cada um dos algoritmos utilizados para treinamento, relacionando as porcentagens 20, 40, 60 e 80%, onde foi visualizado que os algoritmos HGBR e GBR tiveram o melhor desempenho. Essa conclusão será validada na sessão 4.4, quando serão comparados os valores REAL e PREDITO, caso necessário o estudo do MAE será inserido.

Figura 38 - Gráfico de barras do coeficiente R^2 para todos algoritmos por percentual de treinamento – Cenário 2



Fonte: autora

4.4 COMPARAÇÃO ENTRE REAL X PREDITO

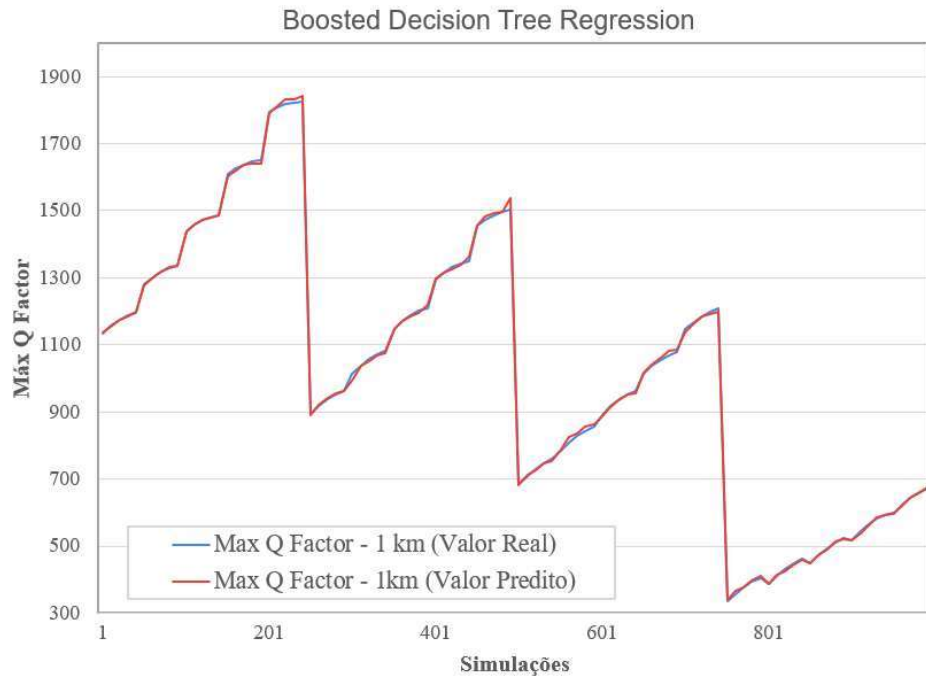
4.4.1 Cenário 1

Nesta seção, foram visualizadas as diferenças gráficas, entre o valor real e o valor predito, ou seja valor que o modelo treinado pelos algoritmos demonstraram o mais alto coeficiente de determinação R^2 . Foi trabalhado aqui com o modelo que foi treinado sobre a metodologia *lock-on* que é quando você utiliza 80% dos dados que se tem para treinamento do modelo e os 20% restantes para avaliação do mesmo. Os gráficos apresentam uma relação entre o máximo Q fator e número da linha da tabela de simulação do Optisystem, para as entradas: distância, atenuação e ganho dos dois amplificadores. Serão abordados os algoritmos BDTR e AdaBoost + DTR que tiveram o melhor desempenho.

4.4.1.1 - Algoritmo Boosted Decision Tree Regression (BDTR)

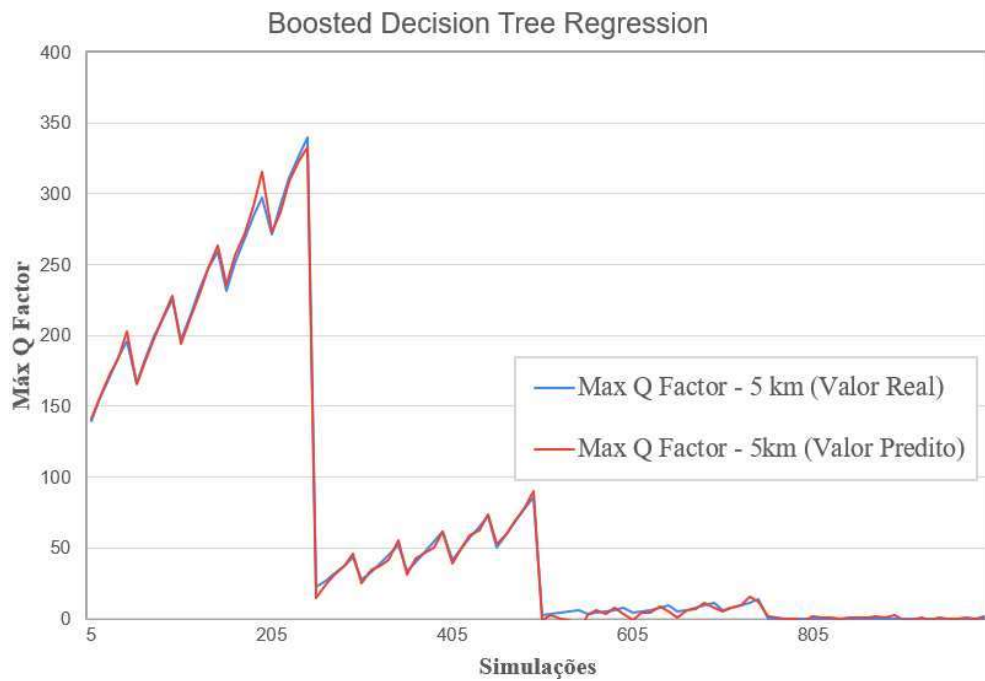
Para o algoritmo BDTR foram visualizados os valores predito, em vermelho e o valor real, em azul para a saída ‘maximo Q fator’. No eixo x do gráfico, é apresentada a linha da base de dados coletada da simulação do optisystem. Na Figura 39 estão representadas as distâncias 1, 5 e 10 km do enlace FSO, respectivamente nas Figuras 39a, 39b e 39c. Os valores dispostos nos gráficos se referem ao *zoom* na região de maior distúrbio entre estes valores. Através desta análise visual, foi constatado que o algoritmo BDTR apresentou um resultado satisfatório com este método de aprendizado de máquina.

Figura 39a - Comparação real x predito – BDTR para 1 km



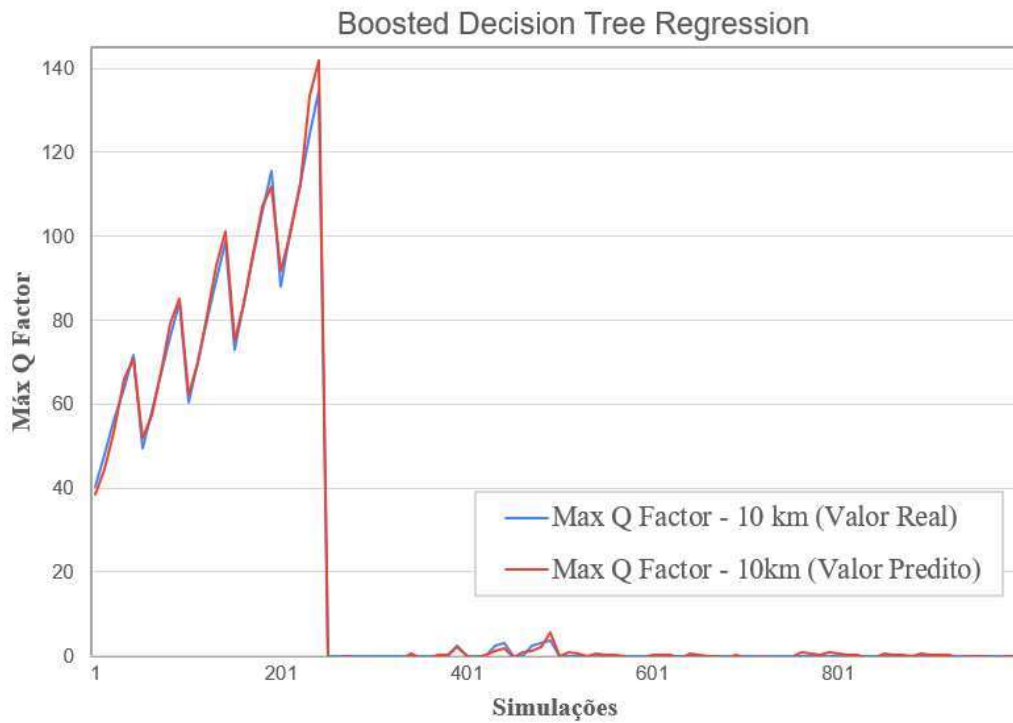
Fonte: autora

Figura 39b - Comparação real x predito – BDTR para 5 km



Fonte: autora

Figura 39c- Comparação real x predito – BDTR para 10 km

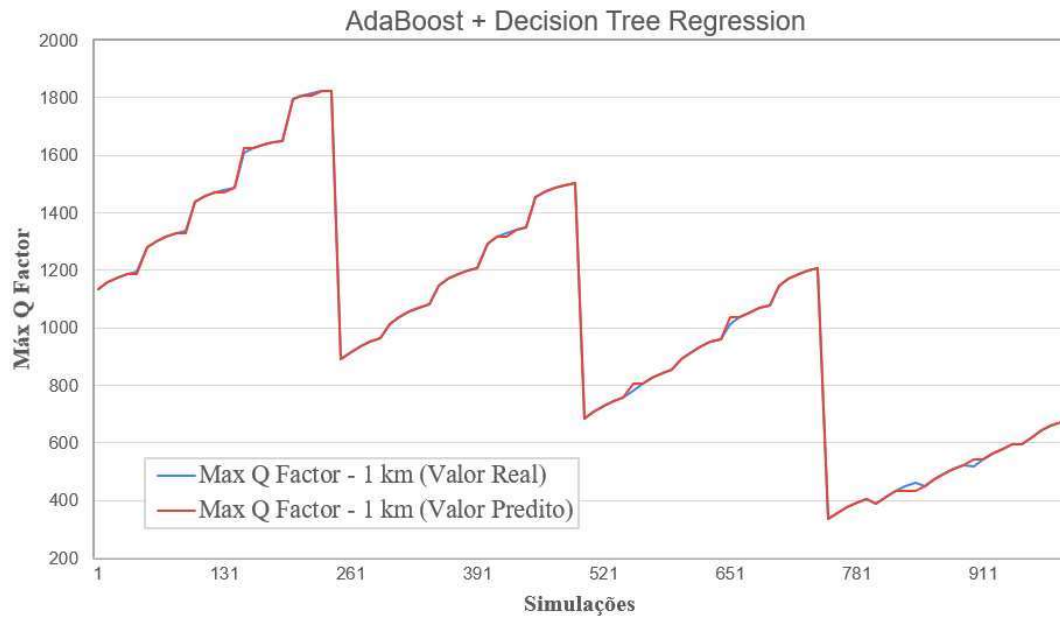


Fonte: autora

4.4.1.2 - Algoritmo AdaBoost + Decision Tree Regression (AdaBoost + DTR)

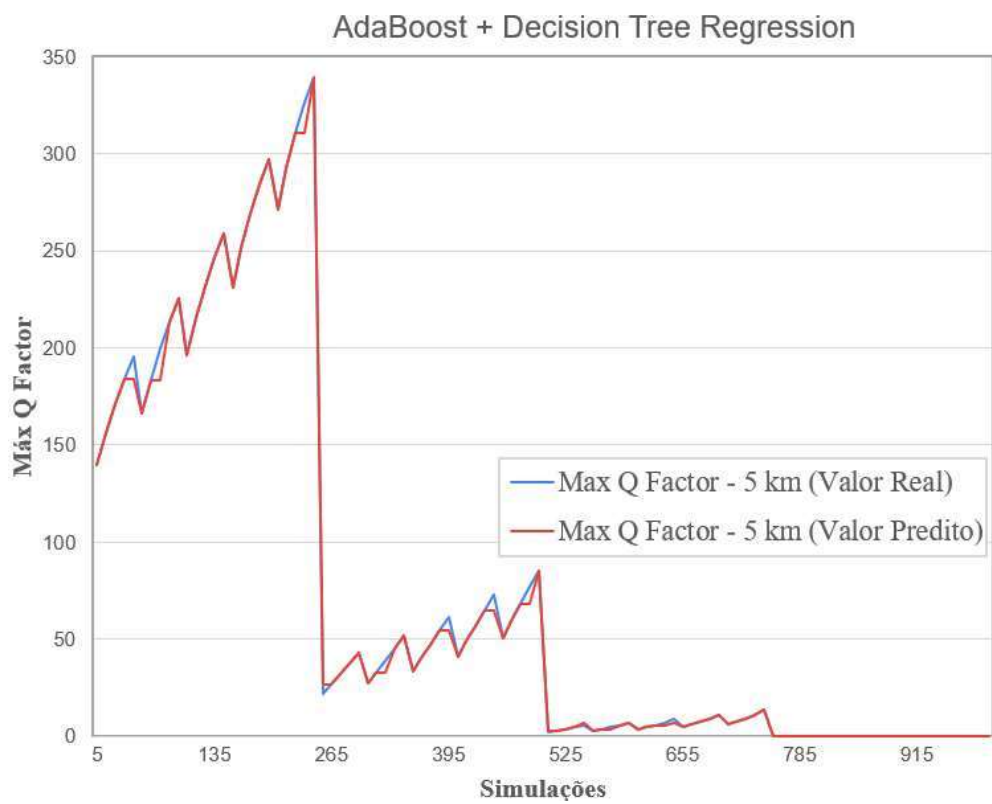
Para o algoritmo AdaBoost + DTR, também foram visualizados, no gráfico o valor predito, em vermelho e o valor real, em azul. Na Figura 40 estão representadas as distâncias 1, 5 e 10 km do enlace FSO, respectivamente nas Figuras 40a, 40b e 40c. Os valores dispostos nos gráficos se referem ao *zoom* na região de maior distúrbio entre estes valores. Através desta análise visual, foi verificado que o algoritmo AdaBoost + DTR também apresentou um resultado satisfatório com este método de aprendizado de máquina, porém em comparação ao Boosted Decision Tree Regression teve uma qualidade levemente reduzida. Portanto, esta análise ratifica o comportamento do gráfico de barras do coeficiente R^2 na Figura 31, ou seja o algoritmo BTDR apresentou melhor desempenho do que o algoritmo AdaBoost + DTR.

Figura 40a - Comparação real x predito – AdaBoost + DTR para 1 km



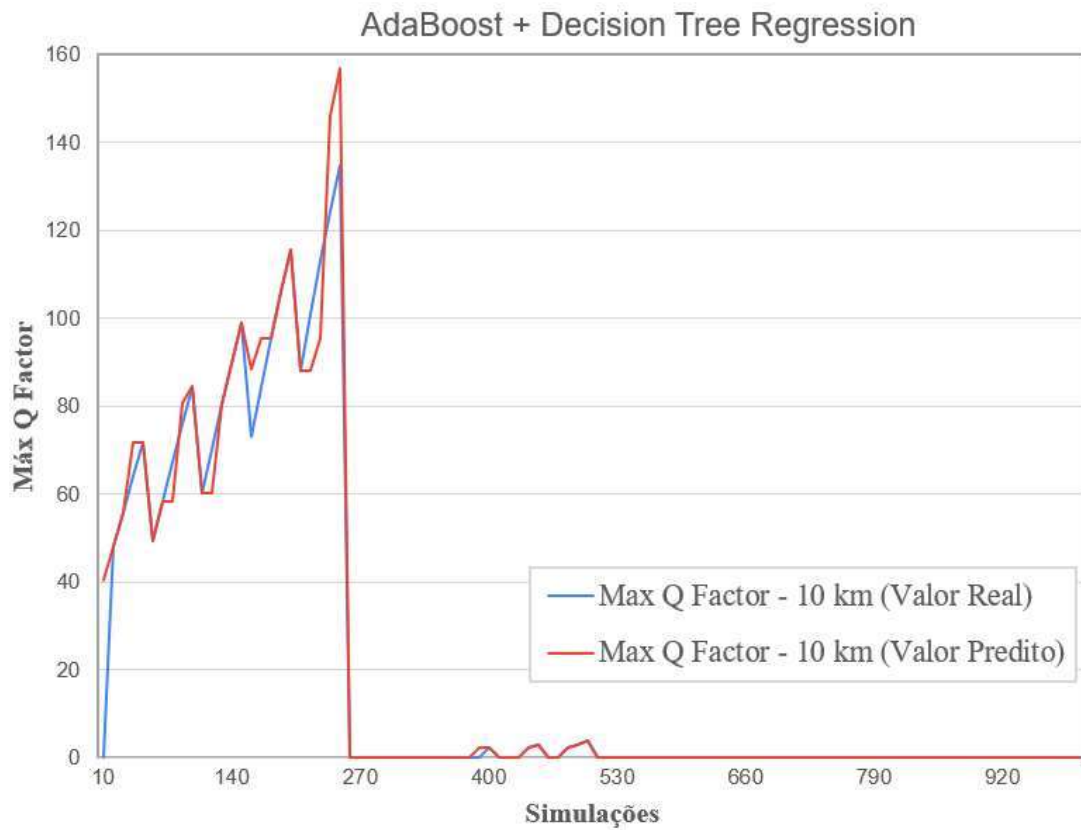
Fonte: autora

Figura 40b - Comparação real x predito – AdaBoost + DTR para 5 km



Fonte: autora

Figura 40c - Comparação real x predito – AdaBoost + DTR para 10 km



Fonte: autora

4.4.2 Cenário 2

Nesta seção, foi realizada a mesma análise do item 4.4.1, porém foram acrescentadas as entradas: diâmetro de abertura do receptor e divergência do feixe, para o treinamento do modelo, totalizando seis entradas. Foram coletadas as duas entradas nas condições descritas na Tabela 12, para os dois melhores algoritmos: **Histogram Gradient Boosting Regression (HGBR)** e **Gradient Boosting Regression (GBR)**, conforme Figura 38 do gráfico de barras do coeficiente R^2 .

Tabela 12 - Condições de coleta dos gráficos para as novas entradas

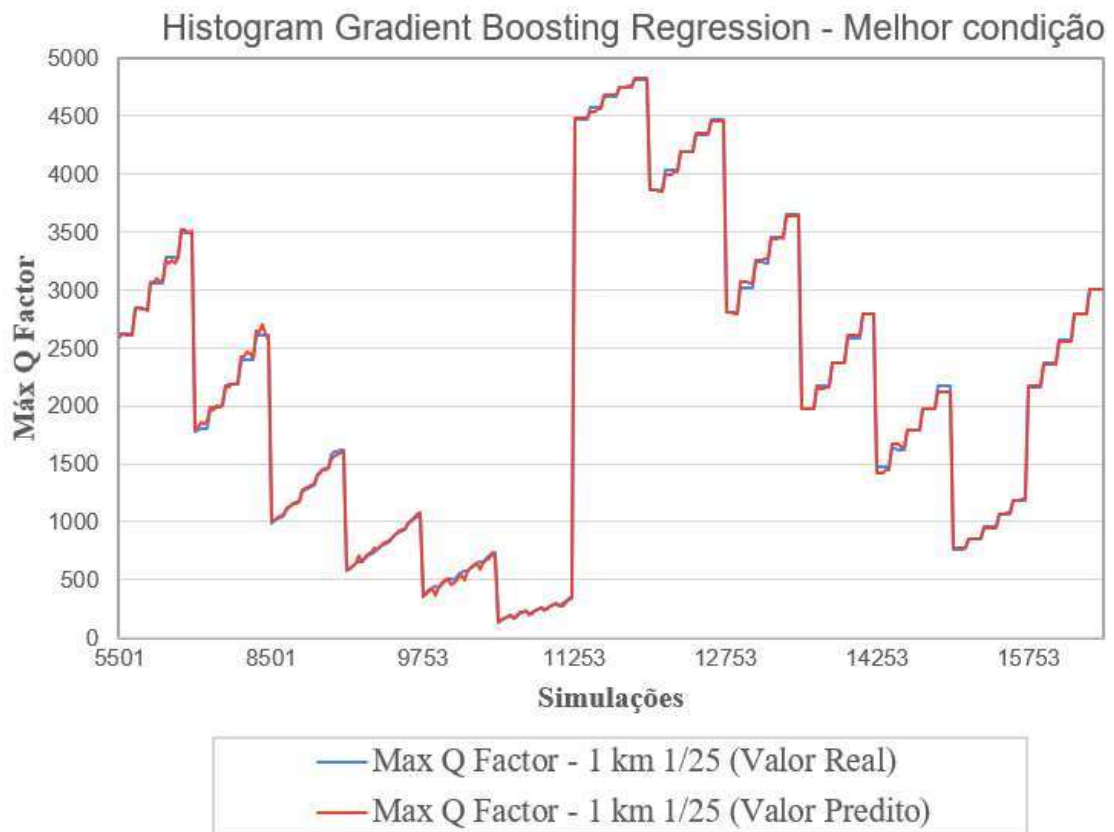
Distância (km)	Divergência do feixe (DF)	Diâmetro de abertura receptor (DAR)	Classificação
1	1	25	Melhor condição
1	3	15	Pior condição
10	1	25	Melhor condição
10	3	15	Pior condição

Fonte: autora

4.4.2.1 - Histogram Gradient Boosting Regression (HGBR)

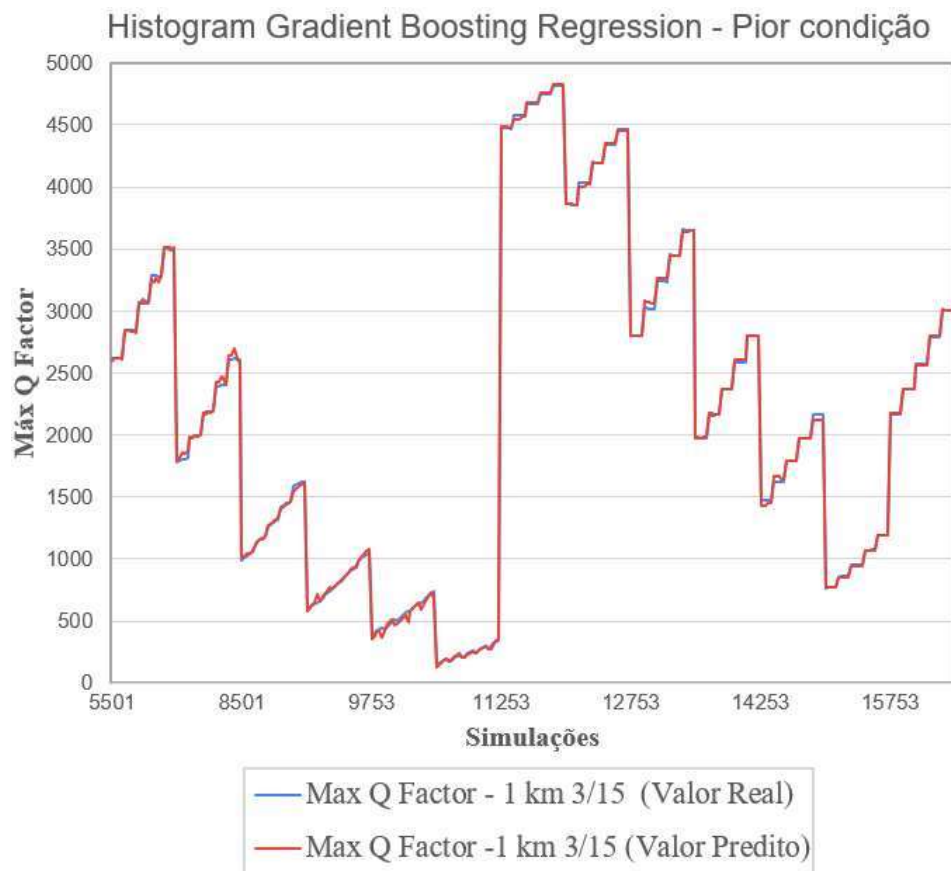
O Algoritmo HGBR foi um dos que apresentaram melhor desempenho, considerando o coeficiente de determinação R^2 , por este motivo serão analisados os valores predito, em vermelho e os valores reais, em azul, para a saída ‘máximo Q fator’. Nas Figuras 41a e 41b, para a distância de 1 km do enlace FSO, foram apresentadas a melhor e a pior condição para as entradas: divergência do feixe e diâmetro de abertura do receptor, correspondente às linhas 1 e 2 da Tabela 12.

Figura 41a - Comparação entre real e predito - HGBR (km 1 – DF=1/DAR=25)



Fonte: autora

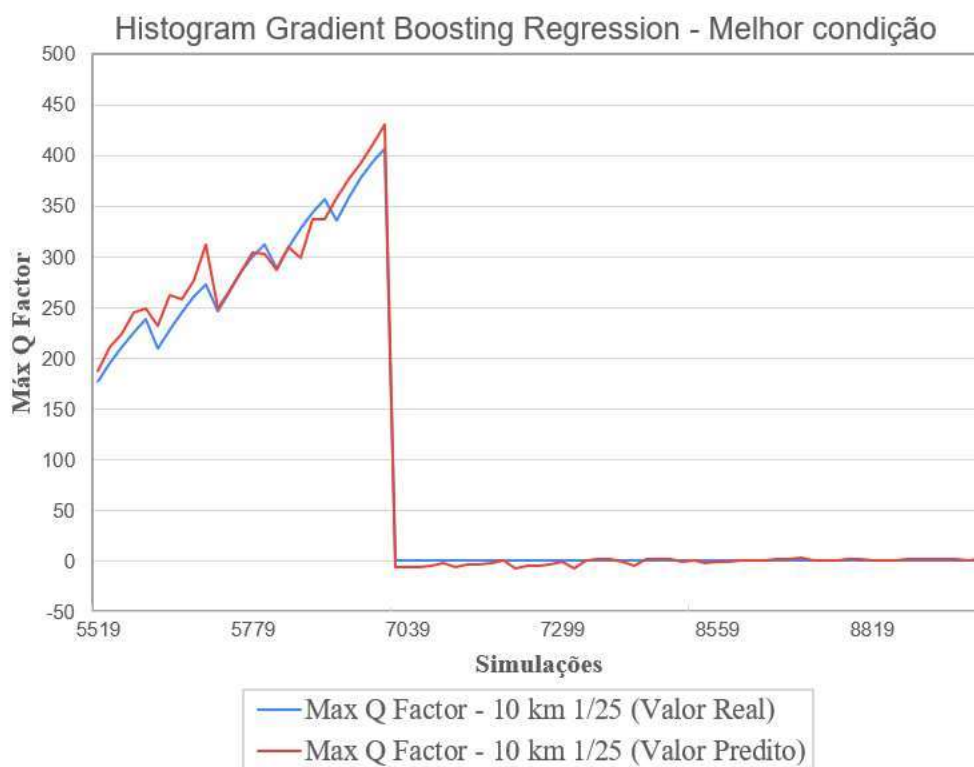
Figura 41b - Comparação entre real e predito - HGBR (km 1 – DF=3/DAR=15)



Fonte: autora

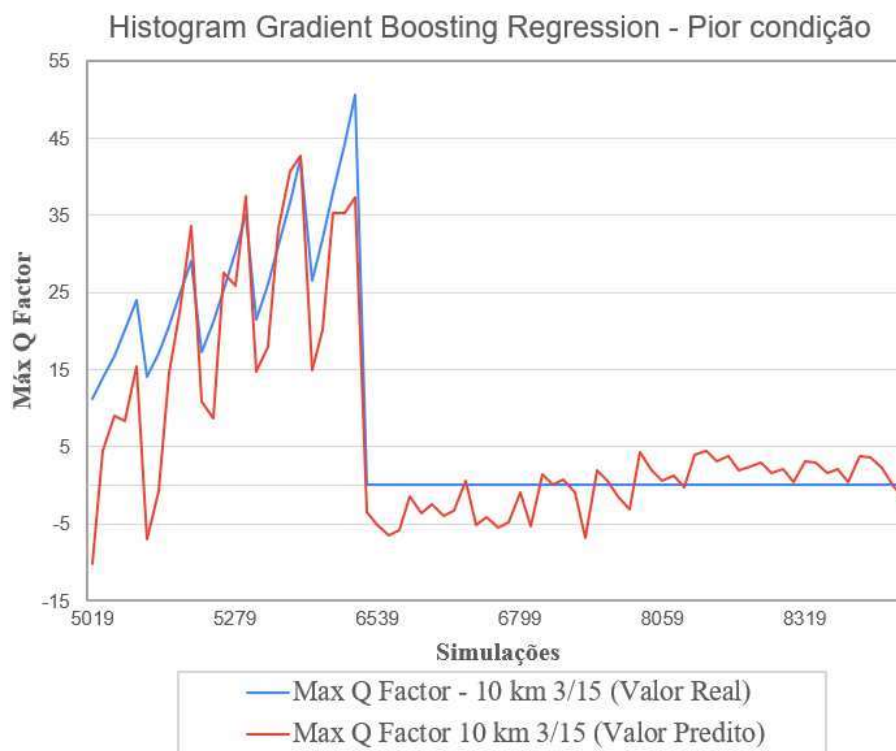
Nas Figuras 42a e 42b, para a distância de 10 km do enlace FSO, foram apresentadas também a melhor e a pior condição para as entradas: divergência do feixe e diâmetro de abertura do receptor, agora correspondente às linhas 3 e 4 da Tabela 12.

Figura 42a - Comparação entre real e predito - HGBR (km 10 – DF=1/DAR=25)



Fonte: autora

Figura 42b - Comparação entre eal e predito - HGBR (km 10 – DF=3/DAR=15)



Fonte: autora

Foi verificado no algoritmo HGBR que para a melhor condição, ou seja divergência do feixe=1 e diâmetro de abertura receptor=25, a comparação entre o real e o predito ficaram próximas, piorando o coeficiente de determinação com o aumento da distância. Em contrapartida quando foi analisada a pior condição, ou seja, divergência do feixe=3 e diâmetro de abertura receptor=15, foi verificado claramente o aumento do erro, com o aumento da distância.

Outro ponto importante que foi observado, é que quanto maior a base de dados para o modelo aprender, maior será o acerto. Para os maiores valores de atenuação a base de dados será menor, pois a maioria das saídas, serão zeradas, como pode ser observado na Tabela 13.

Tabela 13 - Percentual de acerto depende da base de dados

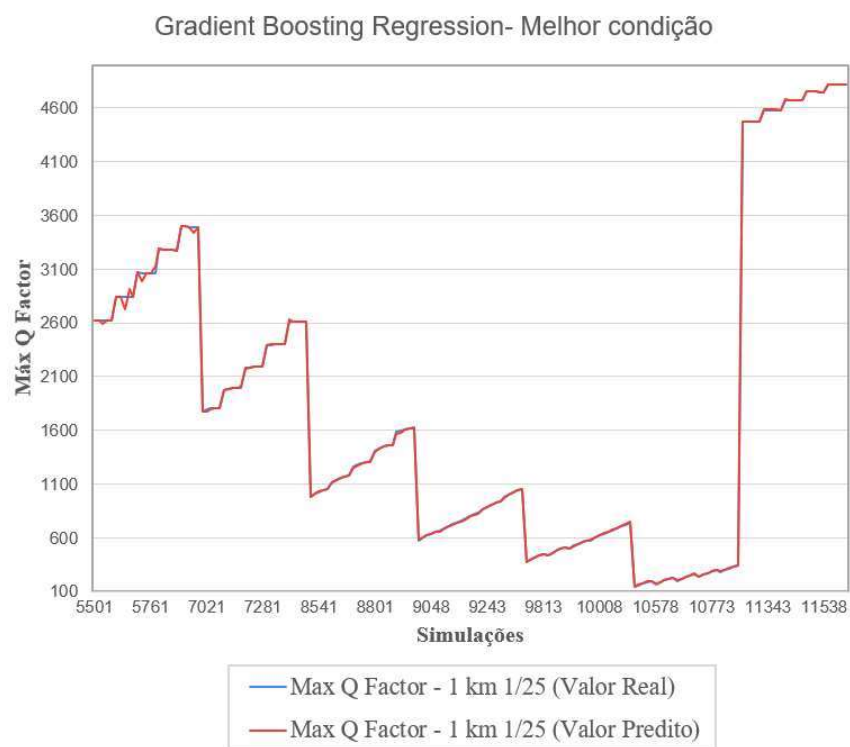
Distância	Atenuação	Diâmetro de abertura do receptor	Divergência do feixe	Ganho Amp1	Ganho Amp2	MaxQFator	MaxQFatorPred	Variação	% Erro	% Acerto	Base Válida (Linhas)
1	0,11	25	1	20	16	4817,865111	4826,979093	9,113982747	0,189171	99,81083	1300
1	22	25	1	6	6	142,3659196	125,9347133	16,43120632	11,54153	88,45847	
1	0,11	15	3	20	16	2625,429248	2624,481753	0,947494842	0,036089	99,96391	
1	22	15	3	6	6	9,007978135	13,5976713	4,589693164	49,04858	49,04858	
5	0,11	25	1	20	16	2543,534521	2546,753693	3,219171716	0,126563	99,87344	431
5	9	25	1	16	16	4,180900192	5,653903469	1,473003278	35,23173	64,76827	
5	0,11	15	3	20	20	605,595017	623,5957065	18,00068945	2,972397	97,0276	
5	4	15	3	8	10	2,421016129	1,704197717	0,716818412	29,60816	70,39184	
10	0,11	25	1	6	6	177,2835385	187,4629139	10,17937539	5,741862	94,25814	100
10	0,11	25	1	10	10	407,2278712	430,6155158	23,38764457	5,743135	94,25687	
10	0,11	15	3	10	10	50,67541623	37,42565213	13,2497641	26,14634	73,85366	
10	0,11	15	3	6	6	11,28255596	-10,13972908	10,13972908	89,87085	10,12915	

Fonte: autora

4.4.2.2 - Gradient Boosting Regression (GBR)

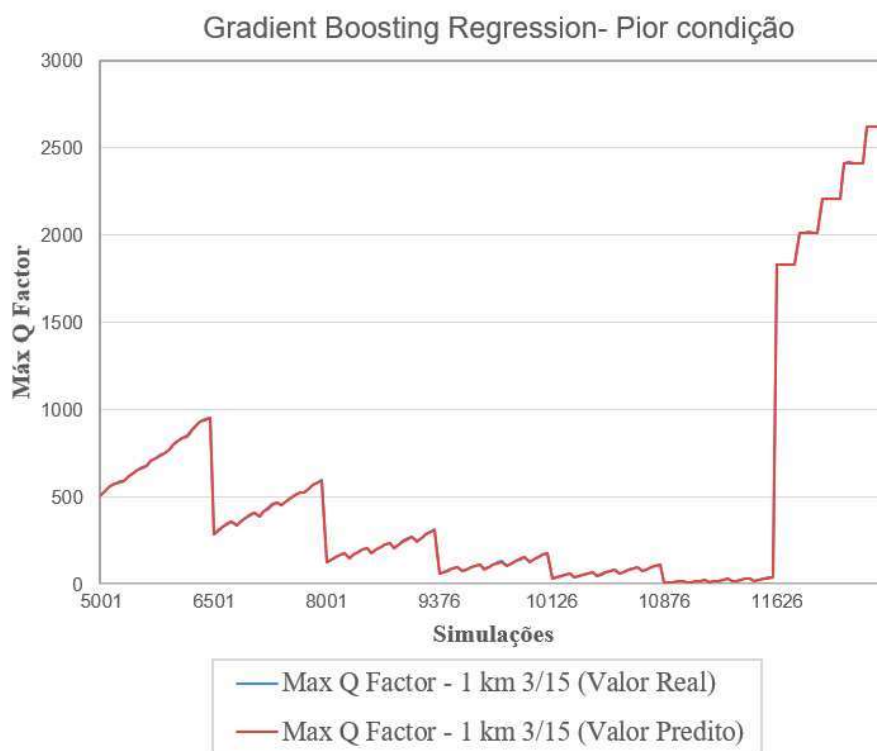
O Algoritmo GBR também foi o outro modelo que apresentou bom desempenho, considerando o coeficiente de determinação R^2 , por este motivo foram visualizados os valores predito, em laranja e os valores reais, em azul, para a saída ‘maximo Q fator’. Nas figuras 43a e 43b, para a distância de 1 km do enlace FSO, foram apresentadas a melhor e a pior condição para as entradas: divergência do feixe e diâmetro de abertura do receptor, correspondente às linhas 1 e 2 da Tabela 12.

Figura 43a - Comparação entre real e predito - GBR (km 1 – DF=1/DAR=25)



Fonte: autora

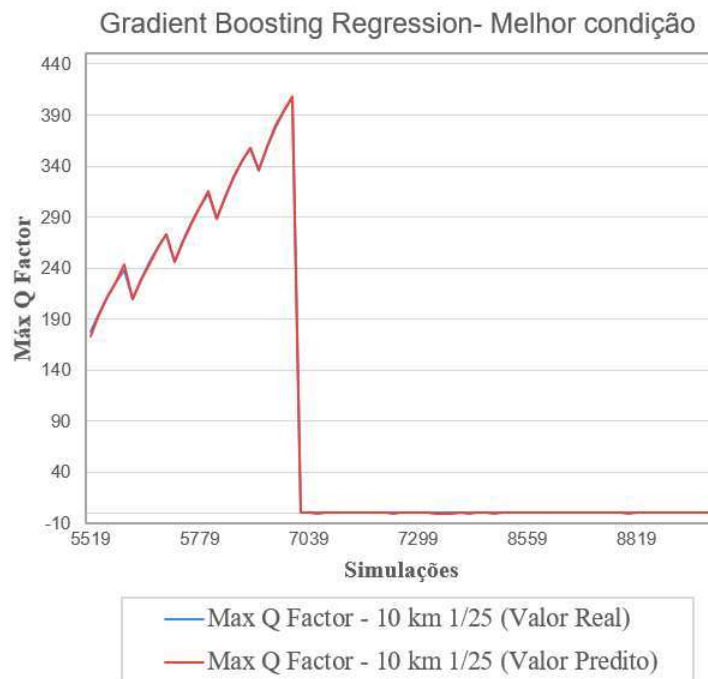
Figura 43b - Comparação entre real e predito - GBR (km 1 – DF=3/DAR=15)



Fonte: autora

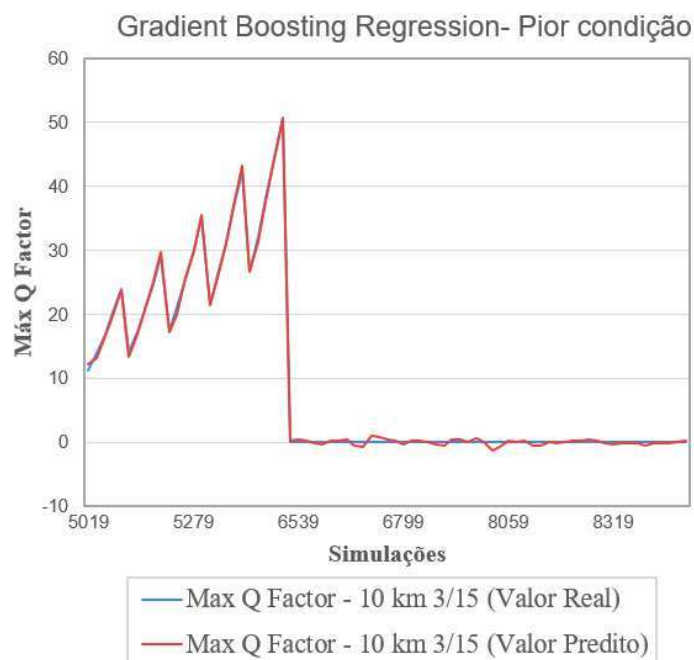
Nas figuras 44a e 44b, para a distância de 10 km do enlace FSO, foram apresentadas também a melhor e a pior condição para as entradas: divergência do feixe e diâmetro de abertura do receptor, correspondente às linhas 3 e 4 da Tabela 12.

Figura 44a - Comparação entre real e predito - GBR (km 10 – DF=1/DAR=25)



Fonte: autora

Figura 44b - Comparação entre real e predito - GBR (km 10 – DF=3/DAR=15)



Fonte: autora

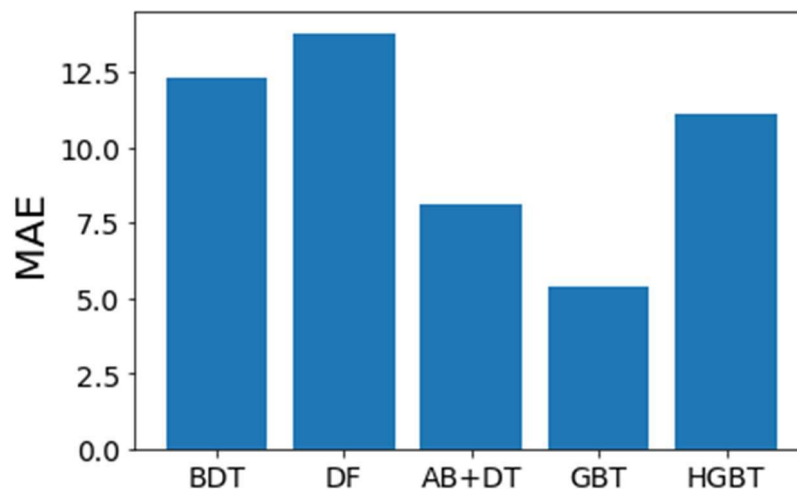
Foi verificado no algoritmo GBR, da mesma forma que no HGBR, para a melhor condição, ou seja divergência do feixe=1 e diâmetro de abertura receptor=25, a comparação entre o real e o predito ficou bem próxima, piorando apenas com o aumento da distância. Em contrapartida quando foi analisada a pior condição, ou seja, divergência do feixe=3 e diâmetro de abertura receptor=15, foi verificado claramente o aumento do erro, que foi amplificado ainda mais com o aumento da distância. Estranhamente de acordo com os gráficos o erro do GBR foi menor que no HGBR, pois de acordo com o estudo do coeficiente de determinação o HGBR apresentou melhor desempenho. No próximo tópico foi analisado o MAE - *Mean Absolute Error*, ou erro absoluto médio para análise deste cenário.

4.5 COMPARAÇÃO ENTRE OS MEAN ABSOLUTE ERROR (MAE)

A fim de justificar o motivo de, apesar de um modelo possuir R^2 melhor que outro, os resultados de predição saírem piores, foi apresentado nesta seção através da comparação entre os valores de MAE que foram obtidos para os modelos treinados com 80% do banco de dados, quando possuímos as colunas diâmetro de abertura receptor e divergência do feixe.

Nota: Quanto menor melhor nesse caso.

Figura 45 - Comparação: MAE para 80% de treinamento



Fonte: autora

Foi verificado na Figura 45 que o MAE chegou a ser duas vezes maior no HGBT do que no GBR, por esse motivo apesar do algoritmo HGBT apresentar melhor valor de R^2 , este algoritmo apresentou erro maior nos gráficos da seção 4.4.3. Portanto para o segundo cenário,

ou seja com seis entradas o algoritmo que apresentou melhor desempenho foi o GBR - *Gradient Boosting Regression*.

5 CONCLUSÕES

Neste trabalho de mestrado foi possível utilizar algoritmos de inteligência artificial para análise de multiparâmetros de uma rede de comunicação óptica sem fio.

Foi possível realizar uma excelente correlação utilizando os parâmetros de entrada; distância, atenuação do canal óptico FSO, ganho dos amplificadores no transmissor e receptor e o parâmetro de saída fator Q com os modelos Boosted Decision Tree Regression (BDTR) e AdaBoost + Decision Tree Regression (ADTR) que apresentaram uma estimativa excelente para o fator Q, com R^2 em torno de 99,995%. Atinge-se também um ótimo resultado para o segundo cenário com dois parâmetros de entrada adicionados: diâmetro de abertura do receptor e divergência do feixe, onde os algoritmos Gradient Boosting Regressor (GBR) e Histogram Gradient Boosting Regression (HGBR) apresentaram predição um pouco pior que o primeiro cenário, porém com o coeficiente de determinação em torno de 99,8%.

Quando analisadas as correlações entre as variáveis de entrada ($GainA$, $GainB$) e saída ($MaxQFactor$) é observado que estas estão parcialmente descorrelacionadas e tem uma forte dependência da atenuação do canal. Por outro lado a correlação de Spearman se mostrou mais robusta que a de Pearson ao inferir a não correlação entre as variáveis.

Ao final desse trabalho concluímos que os modelos de IA podem auxiliar na predição dos valores de saída e assim contribuir na elaboração de uma arquitetura de rede FSO eficaz. Como sugestão para trabalhos futuros, a elaboração de uma rede FSO de múltiplos canais, com sistemas de modulação que visem minimizar os efeitos de atenuação do canal, assim como adicionar parâmetros de entrada como: potência de entrada, diâmetro de abertura do transmissor, perdas diversas no FSO como perdas geométricas, perdas no transmissor, perdas no receptor, pois são parâmetros importantes para uma qualidade superior no sinal de saída.

6 REFERÊNCIAS

- AGRAWAL, G. P. *Fiber-optic Communication Systems*. [S.l.]: New Jersey: [s.n.], p. 29 – 92. ISBN 978-0-470-50511-3., 2010.
- ALGEDIR Amal A., ELGANIMI Taissir Y.. **Machine Learning Models for Predicting the Quality Factor of FSO Systems with Multiple Transceivers**. 2020 2nd Global Power, Energy and Communication Conference (IEEE GPECOM2020), October 20-23, 2020, Online Conference. 2020.
- AMIRABADI, M. A.; KAHAEI, M. H.; NEZAMALHOSSEINI, S. A. **Novel suboptimal approaches for hyperparameter tuning of deep neural network** [under the shelf of optical communication]. Iran University of Science and Technology (IUST), Tehran, Iran, 27 de abril de 2020, publicado em artigo na Physical Communication - Elsevier B.V. 2020.
- AZIZ Siti Hajar Ab, NOR Norhanis Aida M., ZABIDI Suriza Ahmad. **Performance Analysis of Free Space Optics Link Under the Effect of Rain Attenuation**. 2021 8th International Conference on Computer and Communication Engineering (ICCCE). Artigo acessado em 12 de outubro de 2022.
- BENSHUANG Yu, Yong Yao, Yufeng Zhao, Chao Liu, Xuelian Yu. **Simulation research of medium-short distance free-space optical communication with optical amplification based on polarization shift keying modulation**. Yu. Department of Electronic and Information Engineering, Shenzhen Graduate School, Harbin Institute of Technology, Shenzhen, China. 18 de dezembro de 2013, publicado em artigo na Optik - Elsevier B.V.
- BÜHLMANN, P.; YU, B. Boosting. WIREs Computational Statistics, v. 2, n. 1, p. 69–74, 31 dez. 2009. (<https://doi-org.ez138.periodicos.capes.gov.br/10.1002/wics.55>)
- CHANDHINI S.A. *Analysing the Concrete Compressive Strength using Pearson and Spearman* International Conference on Electronics, Communication and Aerospace Technology. India, 2017.
- CHAUDHARY Sushank, BANSAL Preety, LUMB Manisha. **Effect of Beam Divergence on WDM-FSO Transmission System**. International Journal of Computer Applications (0975 – 8887). Volume 93 – No 1, maio 2014. China. 2014
- CORREA, J. G. P. S. *Estudo dos efeitos da Turbulência em um enlace FSO*. Dissertação (Mestrado) — Instituto Militar de Engenharia, 2016.
- DIETTERICH, T. G. *An Experimental Comparison of Three Methods for Constructing Ensembles of Decision Trees: Bagging, Boosting, and Randomization*. *Machine Learning*40, 139-157. [S.l.: s.n.], 2000. (<https://doi.org/10.1023/A:1007607513941>)
- DUBEY Arjun, SINGH Harmeet. **Performance Analysis of MIMO-FSO communication link with Hybrid Modulation**. IEEE 7th International Conference on Smart Structures and Systems ICSSS 2020. 2020.
- ELBISY, M. S.; ELBISY, A. M. S. *Prediction of significant wave height by artificial neural networks and multiple additive regression trees*. Ocean Engineering, v. 230, p. 109077, 2021. Disponível em: <<https://www-sciencedirect.ez138.periodicos.capes.gov.br/science/>

article/pii/S0029801821005126?via=ihub>.

HOSSAIN Farhana, AFROZE Zeenat. **Eliminating the Effect of Fog Attenuation on FSO Link by Multiple TX/RX System with Travelling Wave Semiconductor Optical Amplifier**. Proceedings of 2013 2nd International Conference on Advances in Electrical Engineering (ICAEE 2013). Dhaka, Bangladesh. 2013

HUAWEI TECHNOLOGIES Co. **HCIA – 5G**. Curso de Certificação em tecnologia 5G. Realizado em março de 2021.

JUMIN, E.; BASARUDDIN, F. B.; YUSOFF, Y. B. M. *Solar radiation prediction using boosted decision tree regression model: A case study in Malaysia*. . Environ Sci Pollut Res 28, 26571-26583, 2021. Disponível em: <<https://doi-org.ez138.periodicos.capes.gov.br/10.1007/s11356-021-12435-6>>.

KARTALOPOULOS, S. V. *Free Space Optical Networks for ultra -Board Band Services*. [S.l.]: A Jonh Wiley and Son, in., 2011.

LIKEBUPT, e. a. *Boosted Decision Tree Regression: Component Reference - Azure Machine Learning*. [s.n.]. Disponível em: <<https://learn.microsoft.com/en-us/azure/machine-learning/component-reference/boosted-decision-tree-regression>>.

LIMA, A. d. S. e. a. **CLASSIFICADOR ENSEMBLE: UMA ABORDAGEM NÃO PARAMÉTRICA APLICADO À DETECÇÃO DE DIABETES**. In: *Revista do Seminário Internacional de Estatística com R*, v. 4, n. 2. [s.n.], 2019. Disponível em: <<https://periodicos.uff.br/anaisdoser/article/view/29318/17028>>.

LINKTREE DIDÁTICA TECH. **Inteligência Artificial & Data Science**. Disponível em: <https://didatica.tech/como-funciona-o-algoritmo-arvore-de-decisao/>. Acesso em: 29 de janeiro de 2023.

MAJUMDAR, A. K. *Advance Free Space Optics (FSO)*. [S.l.]: Springer Series in Optical Sciences, 2015.

Mean absolute error. In WIKIPEDIA – a enciclopédia livre. Florida: Wikipedia Foundation, 2021. Disponível em: https://en.wikipedia.org/wiki/Mean_absolute_error Acesso em: 29 de janeiro de 2023.

MICROSOFT – LEARN. **Componente de regressão florestal de decisão (2022)**. Disponível em: <https://learn.microsoft.com/pt-pt/azure/machine-learning/component-reference/decision-forest-regression> Acesso em: 29 de janeiro de 2023.

NANDI Moinak, ROY eshita, CHAKRABORTY subhalaxmi, BASU tanusree. **Study of variation of q-factor using minimum power under different weather conditions**. Artigo acessado em 12 de outubro de 2022

NATEKIN, A.; KNOLL, A. *Gradient boosting machines, a tutorial*. *Frontiers inNeurorobotics*, v. 7. [S.l.: s.n.], 2013. (<https://doi.org/10.3389/fnbot.2013.00021>)

NETO, A. **Apredizagem de Máquinas– Lab. Inteligêncial Artificial (LIA)**. CE In: Programa de Pos Graduação em Ciencia da Computação PPGCC/IFCE, 2022 Fortaleza.

- NOGARE, D. [Inteligência Artificial & Machine Learning, 2020, na internet] Disponível em: <<https://diegonogare.net/2020/04/como-funciona-a-aprendizagem-de-maquina/>> Acesso em: 28 de janeiro de 2023.
- RAHMAN, K. F.; AFRIN, F.; TAIP, F. S. bin. **Optical Network Performance Analysis through Prediction of Q Factor on Simulated and Validated Dataset.** *IEEE*, 2018. Artigo acessado em 12 de outubro de 2022
- SABZEVARI, M.; MARTÍNEZ-MUÑOZ, G.; SUÁREZ, A. *Vote-boosting ensembles. Pattern Recognition*, v. 83, p. 119-133. [S.l.: s.n.], 2018. (<https://doi-org.ez138.periodicos.capes.gov.br/10.1016/j.patcog.2018.05.022>)
- SADZALI Nur Syahidah, Norhanis NOR Aida M., AZIZ Siti Hajar Binti Ab. **Performance Analysis of FSO Systems using Different Modulation Techniques under the Influence of Atmospheric Turbulence.** 2021 8th International Conference on Computer and Communication Engineering (ICCCE). 2021. 2021.
- WEI, L. e. a. **An Improved Gradient Boosting Regression Tree Estimation Model for Soil Heavy Metal (Arsenic) Pollution Monitoring Using Hyperspectral Remote Sensing.** *Applied Sciences*, v. 9, n. 9, p. 1943, 2019.
- XUE Donglin, HAN Pengchao, LIU Yajing, SHA Zijie, LIU Yejun, GUO Lei. **Low-Cost Free-Space-Optical Communication System with Federated Learning-based Channel Prediction.** Asia Communications and Photonics Conference 2020, China 2020.
- YUVARAJ P. *Analyzing User Knowledge by Pearson and Spearman Method.* International Conference on Trends in Electronics and Informatics India, 2017.