Dissertação de Mestrado

Aprendizagem de Máquina para correção do espectro de emissão de nanopartículas usadas para nanotermometria em aplicações biológicas

por Edvonaldo Horácio dos Santos

orientada por
Prof. Dr. Bruno Almeida Pimentel
Prof. Dr. Rafael de Amorim Silva

Universidade Federal de Alagoas
Instituto de computação
Maceió, Alagoas
Julho 25, 2024

UNIVERSIDADE FEDERAL DE ALAGOAS
Instituto de computação

APRENDIZAGEM DE MÁQUINA PARA CORREÇÃO DO ESPECTRO DE EMISSÃO DE NANOPARTÍCULAS USADAS PARA NANOTERMOMETRIA EM APLICAÇÕES BIOLÓGICAS

Dissertação de Mestrado submetida pelo Instituto de computação da Universidade Federal de Alagoas como requisito parcial para a obtenção do diploma em Mestre em Informática

Edvonaldo Horácio dos Santos

Orientador: Prof. Dr. Bruno Almeida Pimentel
Coorientador: Prof. Dr. Rafael de Amorim Silva

Banca Examinadora:

Avaliador 1	Prof. Dr. Evandro de Barros Costa (UFAL)
Avaliador 2	Prof. Dr. Wagner Ferreira da Silva(UFAL)
Avaliador 3	Prof. Dr. Erving Clayton Ximendes (UAM)

Maceió, Alagoas
Julho 25, 2024

Catalogação na fonte
Universidade Federal de Alagoas
Biblioteca Central
Divisão de Tratamento Técnico
Bibliotecária: Taciana Sousa dos Santos – CRB-4 – 2062

S237a Santos, Edvonaldo Horácio dos.
Aprendizagem de máquina para correção do espectro de emissão de nanopartículas usadas para nanotermometria em aplicações biológicas / Edvonaldo Horácio dos Santos. – 2024.
113 f. : il. color.

Orientador: Bruno Almeida Pimentel.
Coorientador: Rafael de Amorim Silva.
Dissertação (Mestrado em Informática) – Universidade Federal de Alagoas. Instituto de Computação. Maceió, 2023.

Bibliografia: f. 78-84.
Apêndices: f. 85-113.

1. Nanotermometria. 2. Nanopartículas. 3. Aprendizagem de máquina. I. Título.

CDU: 004

Dedicatória

À Thaline (*in memoriam*), amada e exemplar esposa, amiga e companheira.

Agradecimentos

Agradeço a todas as pessoas que compõem o corpo do Instuto de Computação da UFAL (IC-UFAL), técnicos e docentes, especialmente aos meus orientadores, os professores Dr. Bruno Almeida Pimentel e Dr. Rafael de Amorim, que foram fundamentais para a realização deste trabalho. Sem suas orientações, direcionamentos e paciência seria impossível concluí-lo. Além de ótimos profissionais, são exemplos de humanidade. Menciono também a técnica administrativa Ana Luisa Ferreira Gomes. Semelhante aos meus orientadores, é exemplo de profissionalismo e humanidade.

Agradeço ao Insituto de Física da UFAL (IF-UFAL) e à Universidade Untonoma de Madrid (UAM) nas pessoas do prof. Dr. Wagner Ferreira e prof. Dr. Erving C. Ximenes, respectivamente por fornecerem informações que impulsionaram esta pesquisa.

Agradeço à minha esposa, Thaline, por seu amor durante toda a nossa trajetória: desde o namoro, passando pelo noivado até o casamento, que, apesar de breve, foi incrível. Ela me viu começar este trabalho, mas, infelizmente, sua partida precoce a impediu de ver presencialmente o final dessa dissertação. À ela, com sua **força** e **coragem**, dedico esta conclusão.

Aos *friends* do IC: Alfredo, Andressa, Leony, Thiago Cordeiro, os quais tive o privilégio de conhecer no instituto durante a graduação e que mantive contato ao longo do mestrado; aos *friends* da Organização do Núcleo da ABC² em Maceió; aos *friends* da minha jornada de 2023 e de 2024.

Agradeço também à minha família que agora está comigo fisicamente, especialmente minha mãe, Maria José Paulino, e um dos meus irmãos, Edvan Horácio. Não obstante todos os demais serem importantes, a menção dos anteriores vem representar bem todo o apoio que recebi desde a tenra infância. Sempre que puderam, não exitam esforços para me ajudar.

Aos grupos cristãos: Núcleo da ABC^2 em Maceió e CRU Campus em Alagoas.

Finalmente, agradeço ao Deus trino, único criador e doador de todas coisas.

Edvonaldo Horácio dos Santos

Se um dia pude ver mais longe, é porque subi no ombro de gigantes. Isaac Newton (adaptado).

Resumo

No campo da nanotermometria o uso das emissões de nanopartículas (NPs) têm se mostrado uma importante ferramenta em diversas áreas, em particular, para o diagnóstico precoce de câncer e de inflamações localizadas. Isso porque as regiões com essas características apresentam temperaturas distintas de um tecido normal. Embora o uso das emissões de NPs para atuarem como nanotermômetros seja uma área bastante promissora, há diversos obstáculos a serem superados, como as distorções espectrais causadas pelos tecidos. Este trabalho usou Aprendizagem de Máquina (AM) para correção espectral em dados de tecidos usando o coeficiente de absorção μ_a (massa cinzenta do cérebro, mama pré-menopausa, fígado, pele, pele variando a concentração de água e água), e o coeficiente de espalhamnento (μ_s) da massa cinzenta do cérebro, mama pré-menopausa, fígado e pele. Os algoritmos que tiveram melhores resultados na correção espectral foram a Árvore de Decisão, Floresta Aleatória e o k Vizinhos Mais Próximos. Eles obtiveram, em média, valores de *MAE*, *MSE*, *MAPE*, R^2 e σ dentre os melhores possíveis. A Árvore de Decisão foi o que mais se destacou em razão de seu baixo custo de tempo, bem como na sustentação da hipótese, usando o teste de Wilcoxon, de que sua capacidade preditiva para este cenário alcançou resultados exatos e precisos. Cada uma das correções espectrais também foi usada para determinação da temperatura dos tecidos. Os modelos que mais se destacaram foram a Árvore de Decisão, a Floresta Aleatória e o k Vizinhos Mais Próximos. O gráfico de violino de temperatura de cada um desses preditores mostrou exatidão e precisão, determinando corretamente a temperatura com uma casa decimal.

Palavras-chave*: *Biológicos*; *Aprendizagem de Máquina*; *Espectroscopia*; *Nanopartículas*; *Nanotermometria*; *Regressão

Abstract

In the field of nanothermometry, the use of nanoparticle (NP) emissions has proven to be an important tool in various areas, particularly for the early diagnosis of cancer and localized inflammations. This is because regions with these characteristics exhibit temperatures different from normal tissue. Although the use of NP emissions as nanothermometers is a very promising area, there are several obstacles to overcome, such as spectral distortions caused by tissues. This work used Machine Learning (ML) for spectral correction in tissue data using the absorption coefficient μ_a (brain gray matter, premenopausal breast, liver, skin, skin with varying water concentration, and water) and the scattering coefficient (μ_s) of brain gray matter, premenopausal breast, liver, and skin. The algorithms that achieved the best results in spectral correction were Decision Tree, Random Forest, and k-Nearest Neighbors. On average, they obtained *MAE*, *MSE*, *MAPE*, R^2, and σ values among the best possible. The Decision Tree stood out the most due to its low time cost and for supporting the hypothesis, using the Wilcoxon test, that its predictive capacity for this scenario achieved accurate and precise results. Each of the spectral corrections was also used to determine the temperature of the tissues. The models that stood out the most were Decision Tree, Random Forest, and k-Nearest Neighbors. The temperature violin plot of each of these predictors showed accuracy and precision, correctly determining the temperature to one decimal place.

Keywords*: *Biological*; *Machine Learning*; *Nanoparticles*; *Nanothermometry*; *Regression*; *Spectroscopy

Lista de Figuras

1.1 Ilustração de experimento *in vivo* com NPs para obtenção do espectro 21

2.1 Esquema de NPs tipo núcleo-casca sendo excitadas em 790 nm com diversas emissões 26
2.2 Espectroscopia usando nanopartículas 26
2.3 Obtenção das temperaturas usando espectro de emissão de NPs 27
2.4 Etapas do processo de Descoberta de Conhecimento em Base de Dados 28
2.5 Ilustração do *k-fold cross validation* 30
2.6 Exemplos de funções de hipóteses: (a) consistente com os dados; (b) não consistente com os dados 33
2.7 Floresta Aleatória genérica 36
2.8 Ilustração do *AdaBoost* 38

5.1 *Overview* da aplicada desde a obtenção dos dados até a obtenção avaliação da temperatura 46
5.2 Interface gráfica do *Virtual Tissue Simulator* 49

6.1 Gráfico de violino do σ obtido a partir da correção espectral usando os seis modelos de AM na **massa cinzenta do cérebro** 54
6.2 Variação da temperatura obtida pela correção espectral usando os algoritmos de AM na **massa cinzenta do cérebro** 55
6.3 Gráfico de violino do σ obtido a partir da correção espectral usando os seis modelos de AM na **mama pré-menopausa** 56
6.4 Gráfico de violino da temperatura obtido a partir da correção espectral usando os seis modelos de AM na **mama pré-menopausa** 57
6.5 Gráfico de violino do σ obtido a partir da correção espectral usando os seis modelos de AM na **fígado** 59
6.6 Gráfico de violino da temperatura obtida pela correção espectral usando os algoritmos de AM no **fígado** 59
6.7 Gráfico de violino do σ obtido a partir da correção espectral usando os seis modelos de AM na **pele** 61

6.8 Gráfico de violino da temperatura obtida pela correção espectral usando os algoritmos de AM na **pele** . . . 61
6.9 Gráfico de violino do σ obtido a partir da correção espectral usando os seis modelos de AM na **pele** variando a concentração de água . . . 63
6.10 Variação da temperatura usando os espectros corrigidos pelos algoritmos de AM na **pele** variando a concentração de água . . . 64
6.11 Gráfico de violino do σ obtido a partir da correção espectral usando os seis modelos de AM na **água** . . . 65
6.12 Gráfico de violino da temperatura obtida pela correção espectral usando os algoritmos de AM na **água** . . . 66
6.13 Gráfico de violino do σ obtido a partir da correção espectral usando os seis modelos de AM na **massa cinzenta do cérebro** . . . 68
6.14 Gráfico de violino da temperatura obtida pela correção espectral usando os algoritmos de AM na **massa cinzenta do cérebro** . . . 68
6.15 Gráfico de violino do σ obtido a partir da correção espectral usando os seis modelos de AM na **mama pré-menopausa** . . . 70
6.16 Gráfico de violino da temperatura obtida pela correção espectral usando os algoritmos de AM na **mama pré-menopausa** . . . 70
6.17 Gráfico de violino do σ obtido a partir da correção espectral usando os seis modelos de AM no **fígado** . . . 72
6.18 Gráfico de violino da temperatura obtida pela correção espectral usando os algoritmos de AM no **fígado** . . . 73
6.19 Gráfico de violino da temperatura obtida pela correção espectral usando os algoritmos de AM na **pele** . . . 75
6.20 Gráfico de violino da temperatura obtida pela correção espectral usando os algoritmos de AM na **pele** . . . 76

A.1 Correções espectrais na massa cinzenta do cérebro usando (a) *MLR* e (b) *DT*. . . 86
A.2 Correções espectrais na massa cinzenta do cérebro usando (a) *RF*, (b) *AdaBoost*, (c) *kNN* e (d) *MLP*. . . . 87
A.3 Correções espectrais na mama pré-menopausa usando (a) *MLR*, (b) *DT*, (c) *RF* e (d) *AdaBoost*. . . . 88
A.4 Correções espectrais na mama pré-menopausa usando (a) *kNN* e (b) *MLP*. . . . 89
A.5 Correções espectrais no fígado usando (a) *MLR* e (b) *DT*. . . . 89
A.6 Correções espectrais no fígado usando (a) *RF*, (b) *AdaBoost*, (c) *kNN* e (d) *MLP*. 90
A.7 Correções espectrais na pele usando (a) *MLR*, (b) *DT*, (c) *RF* e (d) *AdaBoost*. . 91
A.8 Correções espectrais na pele usando (a) *kNN* e (b) *MLP*. . . . 92
A.9 Correções espectrais na pele variando a concentração de água usando (a) *MLR* e (b) *DT*. . . . 92

A.10 Correções espectrais na pele variando a concentração de água usando (a) *RF*, (b) *AdaBoost*, (c) *kNN* e (d) *MLP*. 93
A.11 Correções espectrais na água usando (a) *MLR*, (b) *DT*, (c) *RF* e (d) *AdaBoost*. . 94
A.12 Correções espectrais na água usando (a) *kNN* e (b) *MLP*. 95
A.13 Correções espectrais na massa cinzenta do cérebro usando (a) *MLR* e (b) *DT*. . 95
A.14 Correções espectrais na massa cinzenta do cérebro usando (a) *RF*, *AdaBoost*, (c) *kNN* e (d) *MLP*. 96
A.15 Correções espectrais na mama pré-menopausa usando (a) *MLR*, (b) *DT*, (c) *RF*, (d) *AdaBoost* 97
A.16 Correções espectrais na mama pré-menopausa usando (a) *kNN* e (b) *MLP*. . . . 98
A.17 Correções espectrais no fígado usando (a) *MLR* e (b) *DT*. 98
A.18 Correções espectrais no fígado usando (a) *RF*, (b) *AdaBoost*, (c) *kNN* e (d) *MLP* 99
A.19 Correções espectrais no pele usando (a) *MLR*, (b) *DT*, (c) *RF* e (d) *AdaBoost* . . 100
A.20 Correções espectrais na pele usando (a) *kNN* e (b) *MLP*. 101

B.1 Fluxograma com extração de artigos da base 108
B.2 Quantidade de artigos em cada base 109
B.3 Produção científica anual 110

Lista de Tabelas

2.1 Vantagens e Desvantagens da Regressão Linear Múltipla 34
2.2 Vantagens de uma Árvore de Decisão 36
2.3 Desvantagens de uma Árvore de Decisão 36
2.4 Vantagens da Floresta Aleatória 37
2.5 Desvantagens da Floresta Aleatória 37
2.6 Vantagens e Desvantagens do *AdaBoost* 38
2.7 Vantagens e Desvantagens do *kNN* 39
2.8 Vantagens e Desvantagens das RNAs 40

5.1 Parâmetros dos tecidos para obtenção do μ_a (coeficiente de absorção) e qtde. de curvas 47
5.2 Parâmetros dos tecidos para obtenção do μ_s (coeficiente de espalhamento) e da respectiva quantidade de curvas 48
5.3 Treinamento usando as curvas da Massa Cinzenta do Cérebro, Mama Pré-Menopausa, Fígado e Pele com μ_a e μ_s 49
5.4 Treinamento usando as curvas da Pele com variação na concentração de água e usando o μ_a 50
5.5 Treinamento usando as curvas da Água e usando o μ_a 50
5.6 Treinamento usando as curvas da Água e usando o μ_s 50

6.1 Principais dados de avaliação de correção espectral na Massa Cinzenta do Cérebro 53
6.2 Principais dados de avaliação de correção espectral na Mama Pré-Menopausa . 56
6.3 Principais dados de avaliação de correção espectral no Fígado 58
6.4 Principais dados de avaliação de correção espectral na Pele 60
6.5 Principais dados de avaliação de correção espectral na Pele variando a concentração de água 62
6.6 Principais dados de avaliação de correção espectral na Água 65
6.7 Principais dados de avaliação de correção espectral na Massa Cinzenta do Cérebro 67
6.8 Principais dados de avaliação de correção espectral na Mama Pré-Menopausa . 69
6.9 Principais dados de avaliação de correção espectral no Fígado 71
6.10 Principais dados de avaliação de correção espectral na Pele 74

B.1 Critérios PICOC 104
B.2 Palavras-chave e respectivos sinônimos 105
B.3 Strings de Busca nas Bases de Dados 105

C.1 Valores Críticos para o Teste de Wilcoxon do Posto com Sinais 114

Lista de Acrônimos

IA	Inteligência Artificial
AM	Aprendizagem de Máquina
CART	*Classification and Regression Trees*
DT	*Decision Tree*
KDD	*Knowledge Discovery in Database*
kNN	*k Nearest Neighbors*
RF	*Random Forest*
MAE	*Mean Absolute Error*
MAPE	*Mean Absolute Percentage Error*
MLP	*Artificial Neural Network Multi-layer Perceptron*
MSE	*Mean Square Error*
NLs	Nanotermômetros Luminescentes
NPs	Nanopartículas
RNA	Rede Neural Artificial
RSL	Revisão Sistemática de Literatura

Sumário

1 Introdução **17**

1.1 Motivação e Justificativa 17

1.2 Contextualização da Proposta 20

1.3 Estrutura do Trabalho 23

2 Fundamentação Teórica **24**

2.1 Nanotermometria 24

2.2 Processo de *KDD* e Aprendizagem de Máquina 27

2.2.1 Processo de Descoberta de Conhecimento em Base de Dados 27

2.2.2 Aprendizagem de Máquina 32

2.2.3 Métodos de Regressão 34

3 Relato do Problema **41**

4 Proposta **43**

5 Metodologia de Experimentos **45**

5.1 Procedimento 45

5.1.1 Seleção, Pré-Processamento e Transformação 45

5.1.2 Mineração 48

5.1.3 Avaliação e Interpretação 49

6 Resultados e Discussão **52**

6.1 Resultados e Discussão nos tecidos com distorções usando μ_a e profundidade . 52

6.1.1 Massa Cinzenta do Cérebro 52

6.1.2 Mama Pré-Menopausa 55

6.1.3 Fígado 57

6.1.4 Pele 59

6.1.5 Pele variando a concentração de água 62

6.1.6 Água 64

6.2 Resultados e Discussão nos tecidos com distorções usando μ_s e profundidade . 66

6.2.1 Massa Cinzenta do Cérebro 66

6.2.2 Mama Pré-Menopausa 68
6.2.3 Fígado 71
6.2.4 Pele 72

7 Conclusão e Trabalhos Futuros **77**

Bibliografia **79**

A Correções espectrais **86**
A.1 Correções espectrais obtidas a partir dos espectros distorcidos usando μ_a e profundidade 86
A.2 Correções espectrais obtidas a partir dos espectros distorcidos usando μ_s e profundidade 95

B Revisão Sistemática de Literatura **102**
B.0.1 Metodologia 102
B.0.2 Análise Bibliométrica 104
B.0.3 Resultados 108
B.0.4 Conclusões da RSL 108

C Teste estatístico de Wilcoxon **111**
C.0.1 Estatística 111
C.0.2 Teste de Hipótese 112
C.0.3 *Overview* do Teste de Wilcoxon 112
C.0.4 Teste de Wilcoxon do Posto Sinalizado 113

Capítulo 1

Introdução

Este capítulo fornece uma visão geral acerca do trabalho. Na primeira seção serão apresentadas a Motivação e a Justificativa, apresentando razões que levaram ao desenvolvimento desta pesquisa. Em seguida, na seção sobre Contextualização, será abordado o cenário em que se encaixa a proposta. Por fim, na última seção será tratada sobre a estrutura desta dissertação.

1.1 Motivação e Justificativa

A Nanotecnologia é vista como uma das ciências promissoras do século XXI (ROCO, 2011); (KRUKEMEYER et al., 2015). Além desta característica, seu surgimento tratou-a como um campo de estudo multidisciplinar. (ROCO, 2011) mostra os investimentos na área ao longo dos anos e como estes evoluíram exponencialmente desde os anos de 1990. De acordo com (ROCO, 1999) ela diz respeito ao controle e a reestruturação de materiais para níveis atômicos e moleculares na escala de 1 nm a 100 nm. Este tipo de tecnologia permite desenvolver novos materiais, dispositivos e sistemas com funções e propriedades fundamentais ainda não vistas, sendo considerada a última fronteira econômica para a criação de ferramentas úteis em diversos campos, como a Medicina Molecular (ROCO, 2011) e a Nanomedicina (KUSSUL and BAIDYK, 2004).

A Nanomedicina surge, então, como uma consequência natural do aparecimento da Nanotecnologia como um campo multidisciplinar. Tal é a influência desta última no advento da primeira que ambas apareceram na década de 90 (ROCO, 1999); (KRUKEMEYER et al., 2015). (SWEENEY, 2015) exibem um overview do investimento nesta área na década de 2010, bem como trazem o conceito da *Europe Science Foundation*, que a define como "a ciência e a tecnologia de **diagnóstico**, **tratamento**, **prevenção** de doenças, lesões traumáticas, de alívio da dor e de preservação e melhoria da saúde humana, utilizando ferramentas moleculares e conhecimento molecular do corpo humano".

Dentro da Nanotecnologia existe a Nanotermometria. Dentre as aplicações biológicas, pode-se destacar a mensuração da temperatura em células e/ou tecidos biológicos a partir

do emprego de nanopartículas (NPs) como nanotermômetros. Isso é possível porque os objetos biológicos podem ser vistos como sistemas que operam dentro de faixas de temperaturas bem definidas com uma variação de poucos graus, com sensibilidiade da ordem de 10^{-1} (ZHOU et al., 2016).

Esse equilíbrio é controlado por inúmeros sistemas fisiológicos que são regulados a fim de mantere o equilíbrio do corpo e manter a vida. A perda deste equilíbrio, seja com temperaturas mais altas (hipertermia) ou temperaturas mais baixas (hipotermia), tem grandes chances de causarem desordens fisiológicas, danos nos tecidos bem como alterar as atividades metabólicas. No caso do corpo humano a temperatura de equilíbrio gira em torno dos 37,0 ºC (WALKER et al., 1990). Além disso, sabe-se que dores físicas e até o sofrimento emocional podem causar flutuações na temperatura, gerando grandes efeitos sobre o corpo, as quais podem ser sintomas de doenças sérias como infecções bacterianas ou virais, diabetes, deficiência de vitaminas, doenças autoimunues e câncer (ZHOU et al., 2016); (SANTOS et al., 2018); (XIMENDES et al., 2017).

Assim, o emprego de nanotermômetros luminescentes tem sido um objeto de estudos por pesquisadores biomédicos em estudos pré-clínicos (ZOHAR et al., 1998); (SUZUKI et al., 2007). Dentre os objetivos destaca-se o diagnóstico de doenças através da medição de temperatura para a realização de um tratamento eficaz. Em relação às principais vantagens no uso deste tipo de nanotecnologia destacam-se duas: os nanotermômetros são minimamente invasivos, minimizando distúrbios no processo metabólico e são facilmente expelidos pelo organismo. Logo, aplicações na biomedicina com o intuito de diagnóstico precoce de tumores e/ou outras inflamações e o monitoramento de temperatura aumentam a eficácia e a qualidade do tratamento, trazendo comodidade aos pacientes. Mencionando particularmente sobre o tratamento de câncer que começa com o diagnóstico, os tratamentos existentes atualmente oferecem bons resultados para os vários tipos de neoplasias existentes, mas ainda causam desgaste aos pacientes (SMETS et al., 1993); (SCHMIDT et al., 2020). Portanto, qualquer método que melhore a qualidade do tratamento é um avanço significativo.

Mesmo sendo uma tecnologia promissora e com benefícios notáveis, Yingli Shen e colaboradores nos mostram que no contexto de experimentos *in vivo* o espectro dos nanotermômetros luminescentes tem sua leitura comprometida, afetando a obtenção da temperatura correta no tecido de interesse (SHEN et al., 2020). Os autores também conduziram uma forma de realizar a correção da leitura obtida pelos nanotermômetros, mas que apesar de promissora, requer o conhecimento prévio das curvas do tecido do indivíduo em que se deseja realizar a aferição. Além disso, autores como Alexey N. Bashkatov e colaboradores, mostram como é complexa a modelagem das propriedades ópticas dos tecidos biológicos (BASHKATOV et al., 2011); (BASHKATOV et al., 2016).

Conforme se pode notar, tanto a técnica proposta por (SHEN et al., 2020) como as modelagens matemáticas exibidas por (BASHKATOV et al., 2011) e (BASHKATOV et al., 2016) mostram como é complexa e demorada a atividade de realizar uma predição espectral

correta, uma vez que depende de pelo menos dois passos: o primeiro, que diz respeito à criação de modelos para os tecidos biológicos; e o segundo, que é realizar experimentos que comprovem que os modelos correspondem à realidade. Este processo torna-se demorado, uma vez que cada organismo tem particularidades, como diferentes espessuras e tipos de tecidos. Então, o uso de Aprendizagem de Máquina (AM) visando automatizar este tipo de procedimento é uma ferramenta promissora, uma vez que seu uso em diversos tipos de problemas de regressão tem alcançado bons resultados (BISSOTO et al., 2018); (KIM et al., 2006); (GRÖMPING, 2009); (ROMERO and VENTURA, 2007)).

Nesse contexto, uma vez que deseja-se avaliar o comportamento de modelos de aprendizagem de máquina para correção de intensidade em nanotermometria, o processo de Descoberta de Conhecimento em Bases de Dados (*Knowledge Discovery in Database - KDD*), também conhecido por Mineração de Dados (*Data Mining*), proposto por (FAYYAD et al., 1996) é uma metodologia útil, porque permite seguir um fluxo bem estabelecido na literatura. Descrevendo de forma breve, tivemos o seguinte *pipeline*:

1) Seleção dos dados: as *features* correspondem aos espectros distorcidos, representando as NPs imersas em tecidos biológicos. As *labels* correspondem ao espectro correcto, ou seja, representam as NPs fora de qualquer tipo de tecido biológico.

2) Pré-processamento: verificação da qualidade das *features* e *labels*.

3) Transformação: verificação da necessidade de transformação uma vez que as *features* e *labels* são valores numéricos.

4) Mineração: aplicação dos seis modelos de aprendizagem de máquina usando o *repeated* k-*fold cross validation*.

5) Avaliação: utilização de métricas de regressão, além da diferença absoluta integral (σ) e reconstrução das curvas corrigidas usando gráficos. Outro meio de avaliação é a aplicação do teste de hipótese de Wilcoxon. Após as curvas reconstruídas, extrair a temperatura de acordo com a metodologia usada em nanotermometria.

Cada uma dessas etapas torna o trabalho passível de ser replicado, bem como permite investigações futuras. O procedimento de mineração de dados também permite avaliar de forma adequada os resultados obtidos, além de possibilitar ajustes, caso a etapa de avaliação não apresente resultados satisfatórios.

No passo 4 do *KDD* surge o uso dos algoritmos de AM, na qual uma das principais vantagens é em relação a automação do processo de correção espectral, uma vez que ele é um procedimento feito com significativa intervenção humana, tornando-o demorado. O aprimoramento da capacidade preditiva das distorções em diversos tipos de tecidos é outro proveito com respeito ao uso de AM: as distorções espectrais precisam de modelos físicos complexos

(BASHKATOV et al., 2011); (BASHKATOV et al., 2016), os quais nem sempre são conhecidos. Então, uma investigação na descoberta de padrões para melhoria das predições nesse contexto é um avanço significativo.

Outro ponto relevante em termos de vantagem é em relação ao uso de técnicas em nível de software, tornando possível o processo ficar menos custoso, uma vez que no uso dos algoritmos de AM na etapa de mineração de dados torna automática a predição. Além disso, cada uma das etapas no processo de *KDD* permite ajustes computacionais, que podem ser feitos na seleção, pré-processamento, transformação dos dados e mineração. Nesta última destaca-se a possibilidade dos ajustes dos hiperparâmetros dos modelos aliado a técnicas de treinamento, que no nosso caso foi o *repeated k-fold crosso validation*. Todos estes benefícios propiciados pela mineração de dados cria um *pipeline* para extração de padrões, além de avaliar como se comportam as técnicas de AM.

1.2 Contextualização da Proposta

A temperatura é um parâmetro físico importante e seu conhecimento é essencial em uma grande variedade de áreas. Particularmente, no caso dos organismos, ela está associada à energia térmica média das moléculas de um corpo, a qual, no contexto humano, é um valor em torno de 37,0 °C. A manutenção desse parâmetro corporal normal (normotermia) é um pré-requisito básico para o bom funcionamento do organismo humano. (FRANK et al., 1997), por exemplo, comenta como a devida manutenção da normometria ajuda a reduzir a incidência de eventos cardíacos mórbidos.

Uma temperatura corporal anormal, seja ela muito alta (hipertermia) ou muito baixa (hipotermia), pode causar danos aos tecidos, alterar as atividades metabólicas, entre outras complicações. É bem conhecido que muitas doenças ou disfunções de saúde, como inflamações e tumores, estão associadas a pequenas anomalias de temperatura (SANTOS et al., 2018); (XIMENDES et al., 2017). Nesse caso, o monitoramento local da temperatura pode ser usado como um procedimento de detecção precoce (STARK and WAY, 1974); (LIANG et al., 2013).

Assim, os pesquisadores Biomédicos e Físicos (inseridos no ambiente de biomedicina) têm como uma das suas funções o uso de nanopartículas luminescentes para obtenção da temperatura de um determinado tecido, com o intuito de obter a temperatura e auxiliar na tarefa da descoberta de doenças de forma precisa. Para realizar este procedimento um dos meios é a realização de experimentos *in vivo* e *in vitro* utlizando NPs dopadas de metais terras raras, que neste contexto são chamados de nanotermômetros luminescentes.

Para realizar o experimento e obter a temperatura no tecido desejado são necessárias algumas etapas:

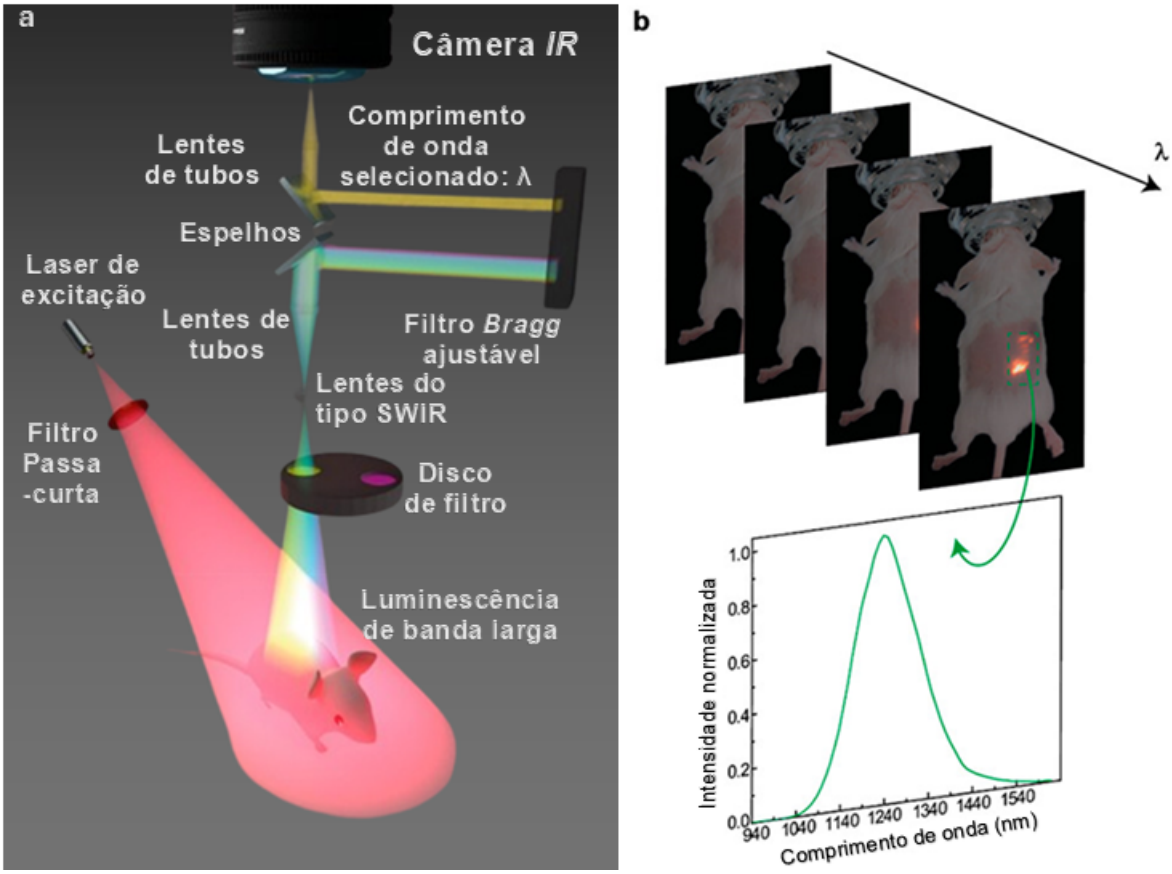
1. Escolher um ser vivo, normalmente um camundongo.

2. Preparar o material para injeção, esta que contém uma solução onde estão presentes as NPs.

3. Caracterização das propriedades luminescentes das NPs.

Após a realização destas três etapas o laser é usado para realizar a excitação das NPs a fim de se obter o espectro, que consiste de um gráfico em 2D cujo eixo das abcissas é o comprimento de onda e o eixo das ordenadas é a intensidade normalizada. Uma ilustração geral desse tipo de procedimento está na Figura (1.1).

Figura 1.1: Ilustração de experimento *in vivo* com NPs para obtenção do espectro



Fonte: SHEN, Yingli et. al. modificado (2020).

A partir da intensidade obtida, são avaliados os picos de modo a obter a temperatura. Um modo disso ser realizado é o seguinte: dadas três intensidades (picos) I_1 e I_2, ou I_2 e I_3 em comprimentos de onda bem definidos é calculada a razão entre elas. Usando outro gráfico, cujo eixo das abcissas é a temperatura e o eixo das ordenadas é a razão $\frac{I_1}{I_2}$ ou $\frac{I_2}{I_3}$, a temperatura daquele tecido onde foi extraído o espectro é obtida. Isso é possível porque este valor numérico (razão) tem uma relação bem definida com a temperatura.

No entanto, (SHEN et al., 2020) mostra que as intensidades das NPs em condições *in vivo* e fora dos tecidos possuem diferenças consideráveis o que compromete a obtenção da temperatura. Os autores mencionam ainda como alguns parâmetros que podem interferir na medida correta, tais como: intensidade luminescente, taxa da intensidade luminescente, comprimento de onda do espectro e posição do comprimento de onda de uma emissão. Além disso, o tecido no qual os NPs estão inseridos afeta a leitura do mesmo, pois cada um possui propriedades que distorcem de forma distinta a intensidade. Tais propriedades dependem de diferentes fatores, como o tipo e a espessura do tecido. Uma modelagem razoável para esse contexto que tem se mostrado promissora é o uso da Lei de Beer-Lambert modificada.

No entanto, mesmo trazendo resultados promissores, nem sempre é possível obter experimentalmente os parâmetros para uso deste modelo matemático, o que torna a tarefa do pesquisador inviável. Nesse sentido, o emprego de técnicas de AM para melhorar a capacidade preditiva e automatizar o processo é um caminho que traz uma economia de tempo, fator importante no contexto do diagnóstico precoce de doenças.

1.3 Estrutura do Trabalho

Este trabalho é composto por mais 6 capítulos. No capítulo 2 serão explicados os principais conceitos relativos à nanotermometria bem como uma descrição do que é realizado atualmente. Além disso, serão abordado sobre o *KDD*, bem como descritos de forma objetiva os seis modelos de aprendizagem utilizados, bem como suas principais vantagens e desvantagens. O capítulo 3 discorrerá de forma mais detalhada sobre o problema, levando em conta os principais trabalhos relacionados. No capítulo seguinte, isto é, no quarto, será apresentada a proposta da pesquisa.

O capítulo 5 contém a metodologia utilizada nos experimentos desde à aquisição dos dados até a avaliação dos espectros e obtenção da temperatura. O capítulo seguinte, isto é, o sexto, diz respeito aos resultados e as respectivas discussões. No sétimo serão abordadas as principais conclusões e contribuições da pesquisa, bem como limitações e trabalhos futuros. A fim de facilitar a leitura do capítulo 2, foi adicionada uma seção de Apêndice, que corresponde as figuras de correção espectral. Além deste, há também um apêndice referente ao procedimento para realização da RSL, além de outro com conceitos básicos de Estatística e, de forma mais detalhada, um sobre o Teste de Wilcoxon do posto sinalizado, que foi utilizado neste trabalho.

Capítulo 2

Fundamentação Teórica

Neste capítulo será explicitada a teoria que dá suporte a este trabalho. Ele aborda a descrição necessária sobre a área de Nanotermometria, descreve os modelos de AM usados. A divisão foi feita em subcapítulos da seguinte forma:

1) **Subcapítulo 2.1**: Descrição da área da física chamada de Nanotermometria.

2) **Subcapítulo 2.2**: Descrição sobre o processo de *KDD* e a etapa de mineração terá como foco descrever os modelos AM e as técnicas empregadas neste trabalho. Ela é subdividida em 3 subseções. Na primeira é detalhado sobre o processo de Decoberta de Conhecimento em Bases de Dados (*KDD*, do inglês *Knowledge Discovery in Database*). Na segunda são detalhados sobre os tipos de Aprendizagem clássicos. Na última serão especificados os métodos empregados para a correção dos espectros, incluindo as formulações matemáticas dos modelos.

2.1 Nanotermometria

Para nos situarmos onde a Nanotermometria se encaixaria precisamos explicitar de forma concisa sobre Nanomedicina. A Nanomedicina é das subáreas da Nanotecnologia sendo um campo de estudos muitas vezes caracterizado por sua multidisciplinaridade. A *Europe Science Foundation* conceitua esta área como "a ciência e a tecnologia de **diagnóstico**, **tratamento**, **prevenção** de doenças, lesões traumáticas, de alívio da dor e de preservação e melhoria da saúde humana, utilizando ferramentas moleculares e conhecimento molecular do corpo humano". De acordo com (KRUKEMEYER et al., 2015) as possibilidades da Nanomedicina incluem três pilares:

1. Nanomaterais e nanoinstrumentos;

2. Conhecimento de medicina molecular no campo da genética;

3. Nanotecnologias que pode ser usadas para um diagnóstico rápido e preciso, bem como para terapias.

Neste último pilar, isto é, o (3), destaca-se a Nanotermometria, que no contexto de aplicações biológicas diz respeito à mensuração de temperatura em células e/ou tecidos biológicos usando NPs como nanotermômetros luminescentes. Isso é possível porque os objetos biológicos podem ser vistos como sistemas que operam dentro de faixas de temperaturas bem definidas com uma variação de poucos graus, com sensibilidade da ordem de 10^{-1} (ZHOU et al., 2016). Uma das maneiras de realizar este tipo de mensuração é através da Espectroscopia, na qual os físicos pesquisadores estudam sobre a absorção e emissão da luz em diversos materiais, como os tecidos biológicos.

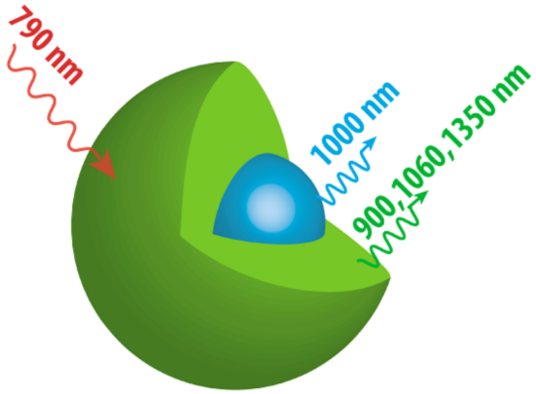
Mas o que é Espectroscopia? Ela pode ser definida como o estudo da absorção e emissão da luz e também de outras formas de radiação pela matéria e de como determinados processos dependem do comprimento de onda da radiação em estudo (HURST and CHU, 2021). Os pequisadores (ADEOLA et al., 2023) explicam ainda que ela pode ser vista como o emprego de nanomateriais, tais como metais, polímeros, materiais orgânicos termosensíveis com características luminscentes, nanocompostos dopados de elementos terras-raras, entre outros tipos com o objetivo de mensurar temperatura em um determinado ambiente. De acordo com (ZHOU et al., 2016) essa busca por nanotermômetros é amplamente impulsionada por dois campos: (i) biologia celular fundamental para compreender a termogênese celular em nível unicelular e (ii) imagens in vivo para monitorar e controlar procedimentos de ablação térmica clinicamente importantes

Dentre as aplicações das NPs usadas como nanotermômetros luminescentes, pode-se mencionar que o uso delas têm se mostrado uma tecnologia fundamental em diversas áreas, tais como Eletrônica (KEFENI et al., 2017), Engenharia (WANG et al., 2017), Biologia (SPERLING et al., 2008), Medicina (PANKHURST et al., 2003); (DE JONG and BORM, 2008); (ULBRICH et al., 2016), entre outras. Em Biomedicina merecem destaque aquelas dopadas com elementos terras-raras que são especialmente úteis para o diagnóstico precoce de doenças como o câncer e inflamações localizadas (SANCHEZ et al., 2008). Para tanto, as NPs são excitadas por um laser em um determinado comprimento de onda, sendo possível capturar suas emissões espectrais. A figura (2.1) ilustra uma NP modelo casca-núcleo sendo excitada a 790 nm e suas emissões nos seguintes comprimentos de onda 1000 nm, 900 nm, 1060 nm e 1350 nm.

Nas aplicações biológicas é possível inseri-las em um tecido biológico a fim de obter a emissão espectral naquela região. A figura (2.2) ilustra um experimento usando NPs em um tecido biológico. A partir do espectro obtido é possível determinar a temperatura naquela região. Quando usadas nesse tipo de aplicação elas são chamadas de nanotermômetros luminescentes.

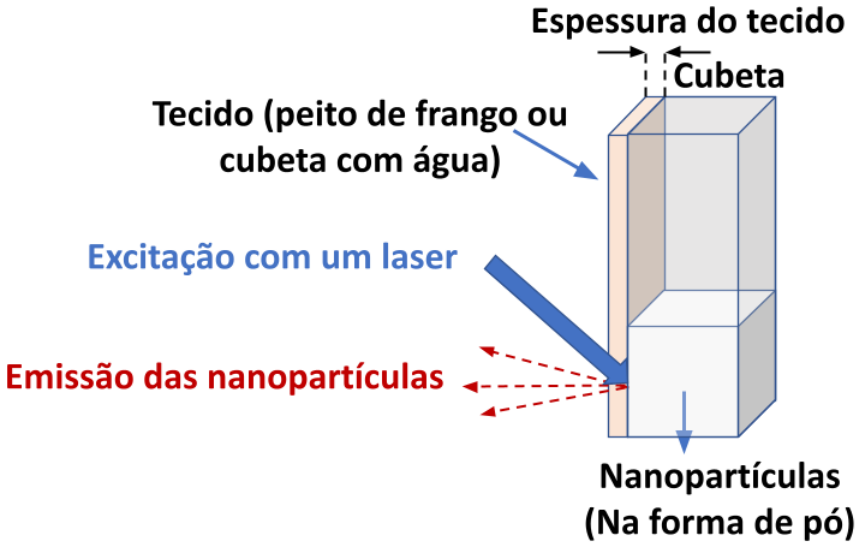
Quando os nanotermômetros são usados para explorar a relação entre as propriedades luminescentes e a temperatura, têm-se a área de estudo chamada de Nanotermometria Luminescente (JAQUE and VETRONE, 2012). O seu uso em aplicações biológicas tem sido

Figura 2.1: Esquema de NPs tipo núcleo-casca sendo excitadas em 790 nm com diversas emissões



Fonte: E. C. Ximendes et al (2015).

Figura 2.2: Espectroscopia usando nanopartículas



Fonte: E. C. Ximendes et al (2015).

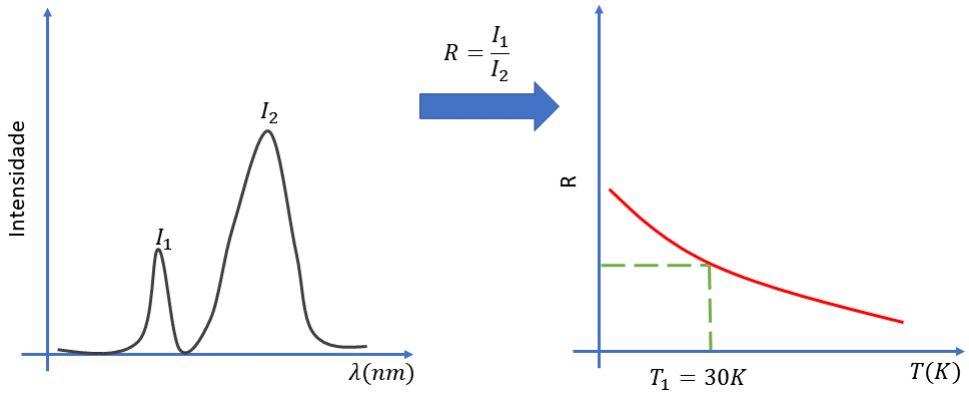
investigado nos últimos anos de forma significativa conforme se pode notar em (DEL ROSAL et al., 2017), (BEDNARKIEWICZ et al., 2020), (NEXHA et al., 2021) e (SOARES et al., 2023).

Mensurar a temperatura em contexto médico é importante, isso porque em organismos vivos ela está associada à energia térmica das moléculas, sendo um pré-requisito essencial para o bom funcionamento do sistema biológico. Se ela for muito alta (hipertermia) ou muito baixa (hipotermia), podem ocorrer diversos tipos de danos nos tecidos e com a consequente

alteração nas atividades metabólicas, algumas das quais podem levar a óbito, como é o caso de sepse.

Em relação ao diagnóstico de doenças, sabe-se que diversas inflamações e tumores (cânceres) estão associadas a pequenas variações de temperatura no nível celular (SANTOS et al., 2018); (XIMENDES et al., 2017). Isso acontece porque nestes tipos de doenças, o metabolismo celular ocorre de forma mais acelerada se comparado ao das células que estão sadias (OWENS et al., 2020). Um exemplo deste tipo de fenômeno é no câncer de mama (YAHARA et al., 2003).

Figura 2.3: Obtenção das temperaturas usando espectro de emissão de NPs



Fonte: Autor.

2.2 Processo de *KDD* e Aprendizagem de Máquina

Nesta subseção serão descritos o Processo de Descoberta de Conhecimento em Base de Dados (do inglês *Knowledge Discovery in Database*), que também pode ser visto como sinônimo de *Data Mining*, e os algoritmos de AM utilizados. Isto permitirá o leitor conhecer com maior solidez as técnicas empregadas, incluindo suas descrições matemáticas acompanhadas das respectivas vantagens e desvantagens.

2.2.1 Processo de Descoberta de Conhecimento em Base de Dados

O processo de Descoberta de Conhecimento em Base de Dados diz respeito à extração de informação relevante a partir de uma base de dados. De acordo com (FAYYAD et al., 1996) em seu famoso artigo *From Data Mining to Knowledge Discovery in Databases*, publicado em 1996 na revista AI Magazine da Associação Americana de Inteligência Artificial, o termo foi cunhado no primeiro workshop de *KDD* em 1989 por Piatetsky-Shapiro. Na época de publicação, quando os computadores pessoais e a internet estavam em expansão para o público,

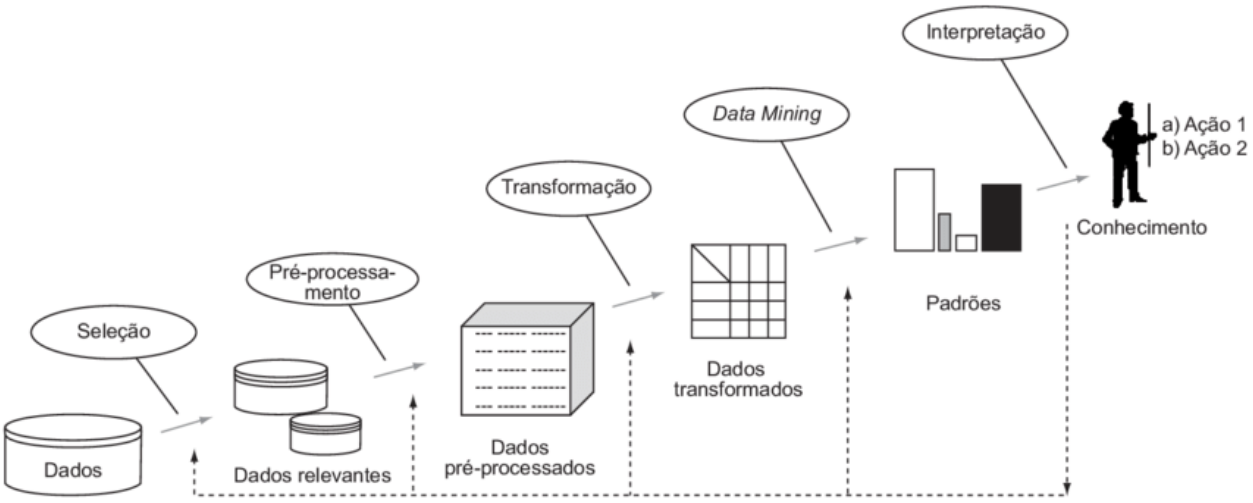
os autores mencionados anteriormente já perceberam a necessidade pela criação de teorias e ferramentas computacionais para auxiliar os pesquisadores e empresas na busca por conhecimento a partir de grandes quantidades de dados.

Em relação ao quantitativo de informação disponível, a última década mostrou um crescimento exponencial no acúmulo de dados, que agora se encontra na casa dos zetabytes (10^{21} bytes). Não é sem razão que estamos na era do *Big Data*. Quanto às técnicas computacionais, notou-se que havia um leque considerável de algoritmos que poderiam ser utilizados nos dados, tais como a Árvore de Decisão, a Máquina de Vetores de Suporte e Redes Neurais Perceptrons.

Contudo, na década de 90 o poder computacional ainda era muito limitado. Isso mudou radicalmente a partir dos anos 2010 quando surgiram processadores específicos, placas de processamento gráfico (GPUs) e até unidades próprias de processamento de frameworks voltadas às tarefas de AM e aprendizagem profunda, permitindo melhores desempenhos dos modelos, além de terem um grau de sofisticação maior em relação aos hardwares anteriores. Esse incremento no poder de processamento segue a ideia proposta por (MOORE, 2006), que já esperava um aumento significativo no desempenho dos computadores com o passar do tempo. Dentre as tecnologias de processamento, podemos citar as Unidades de Processamento de Tensores (*Tensor Process Unit* - TPUs) cuja manifestação pública se deu em 2016. Essas unidades de processamento são utilizadas no framework *TensorFlow* (ABADI et al., 2015) em tarefas de mineração de dados com modelos de AM e também os modelos conexionistas de aprendizagem profunda.

De acordo com ((FAYYAD et al., 1996)), o processo de Descoberta de Conhecimento em Base de Dados envolve 5 etapas. Elas estão ilustrados na Figura (2.4).

Figura 2.4: Etapas do processo de Descoberta de Conhecimento em Base de Dados



Fonte: Fayyad et. al (1996), adaptado

Vale destacar que ((FAYYAD et al., 1996)) menciona a relevância de definir bem o domínio da aplicação, a partir de um conhecimento prévio relevante com foco em definir o(s) objetivo(s), declarando este passo como o primeiro. Iremos, então, apresentar de forma breve e relacionando com a nossa pesquisa, cada uma das 5 etapas do processo de *KDD*.

Seleção de Dados

A etapa 1, que corresponde à seleção dos dados, são coletadas as informações a partir das quais se deseja extrair informação relevante, ou seja, conhecimento. Eles podem ser encontrados de diversas formas, desde bases de dados não estruturadas, como imagens, texto, e/ou aquelas estruturadas, geralmente organizadas em tabelas. Além disso, essa seleção de informações pode ser extraídas de uma única base de dados ou de várias bases. Cabe àqueles que estão envolvidos no processo tomar essa decisão. As variáveis escolhidas são os recursos (features) a partir dos quais se deseja encontrar padrões. No nosso trabalho as features correspondem às curvas distorcidas (espectros), na qual cada uma é um par ordenado do tipo (λ, I(λ)) e as *labels* são as intensidades não distorcidas, referentes aos espectros não distoricdos das NPs de TiO_2 Nd^{3+}.

Pré-Processamento

A etapa número 2 envolve aquela que é considerada uma das mais importantes do processo. Isso porque diz respeito a uma investigação sobre a qualidade dos dados: se os dados não possuirem boa qualidade, mesmo com os melhores algoritmos de AM, os resultados não serão satisfatórios e/ou não generalizarão de forma adequada (SINGH and SINGH, 2020); (ZHU and WU, 2004); (BATISTA and MONARD, 2003).É nesta fase do processo de descoberta de conhecimento que é realizada uma investigação tão minuciosa quanto possível, buscando explorar a qualidade dos dados. Operações como remoção de ruídos, preenchimento de dados ausentes entre outros, podem ser feitas. É bastante comum utilizar técnicas de estatísticas, tais como uso de média, mediana, moda, frequências e análise por quantis apenas para citar algumas. No presente trabalho não foi utilizada nenhuma técnica de pré-processamento, pois as features estavam com boa qualidade.

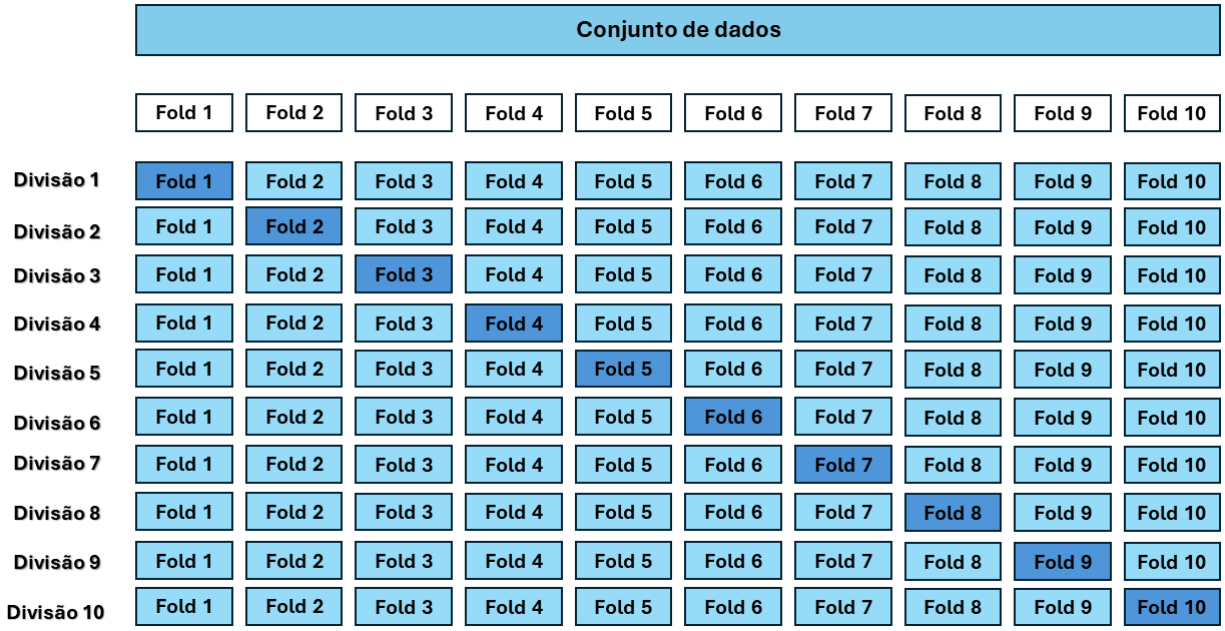
Transformação

A etapa 3, transformação, diz respeito à etapa subsequente ao pré-processamento. Ela envolve realizar o emprego de técnicas para tornar os dados possíveis de utilizar na etapa de mineração. Destacam-se métodos como *one-hot encoding* e a normalização de dados. Em nosso trabalho não foi efetuada nenhuma transformação e/ou normalização dos dados, de modo que foi empregado os valores de *features* e *labels* de forma direta.

Mineração

O passo número 4 envolve a mineração, ou seja, o *Data Mining*. Ele consiste na aplicação de um ou mais modelos de AM aos dados. Vale destacar que, antes da aplicação da mineração, os dados precisam ser divididos em treino e teste. Para isso são empregadas diversas técnicas. Neste trabalho foi aplicado o *repeated k-fold cross validation*, com um número de *folds* iguais a 10 e de repetições iguais a 30. A ideia básica dessa técnica é dividir o conjunto de dados em k partes iguais. Enquanto k-1 partes são utilizadas como treino, a outra é utilizada para teste, até completar o uso das k partições. Esse processo é repetido de forma interativa, que para o nosso caso foram 30 vezes. Ideia é observar para qual valor médio as predições de cada feature estão resultando. A Figura (2.6) ilustra o processo com 1 repetição.

Figura 2.5: Ilustração do *k-fold cross validation*



Fonte: Autor (2024).

Na etapa de mineração de dados foram utilizados os seguintes algoritmos de aprendizagem de máquina:

1. Regressão Linear Múltipla (*MLR*)
2. Árvore de Decisão (*DT*)
3. Floresta Aleatória (*RF*)
4. Adaptative Bosting (*AdaBoost*)
5. k Vizinhos Mais Próximos (*kNN*)

6. Rede Neural Multi-Layer Perceptron (*MLP*)

A partir deles foram geradas as métricas de avaliação, bem como os espectros preditos por cada modelo. O intuito é observar os valores de métricas de cada técnica de AM, bem como realizar a reconstrução de intensidade a fim de obter a temeperatura.

Avaliação e Interpretação

O último passo é a avaliação resultante da mineração de dados. Nada mais é do que o emprego de métricas de desempenho que expressam por meio números a aprendizagem dos algoritmos. Destacam-se o Erro Médio Absoluto (*MAE*), Erro Quadrático Médio (MSE), Erro Percentual Médio Absoluto (*MAPE*) e o R² Score.

Seja N o número de amostras do model, y_i o valor de *label* verdadeiro, $\hat{y}_i$ o valor de *label* predito pelo algoritmo. Então, as métricas de desempenho mencionadas são definidas da seguinte forma:

$$MAE = \frac{1}{N}\sum_{i=1}^{N}(y_i - \hat{y}_i) \tag{2.1}$$

$$MSE = \frac{1}{N}\sum_{i=1}^{N}(y_i - \hat{y}_i)^2 \tag{2.2}$$

$$MAPE = \frac{1}{N}\sum_{i=1}^{N}\frac{|y_i - \hat{y}_i|}{max(\epsilon, |y_i|)} \tag{2.3}$$

$$R^2(y, \hat{y}) = 1 - \frac{\sum_{i=1}^{N}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{N}(y_i - \bar{y})^2} \tag{2.4}$$

onde $\bar{y} = \frac{1}{n}\sum_{i=1}^{n}$ e na Equação 2.3 o é um valor que definido no código da (PEDREGOSA et al., 2011) para quando o denominador for nulo ou menor que $\bar{y}$.

Para o *MAE*, *MSE* e *MAPE* melhor é o desempenho do modelo quanto mais próximos esses valores forem de 0. Já o R² indica que o modelo não está aprendendo quando seus valores são menores que 0 e indica que o algoritmo tem aprendizado quando passa a ter valores maiores que 0. O melhor valor possível é 1 e, de modo geral, quando está entre $0 \leq$ R² ≤ 1 pode ser visto como diretamente proporcional ao grau de explicação das *features*.

Além destas métricas, Erving C. Ximendes e colaboradores definiram uma equação para avaliar a distorção espectral, definida do seguinte modo:

$$\sigma_{tecido} = \int |I^{norm}_{tecido}(\lambda) - I^{norm}_{sem-tecido}(\lambda)|\, d\lambda \tag{2.5}$$

onde $I^{norm}_{tecido}(\lambda) = \frac{I(\lambda)}{\int I(\lambda) d\lambda}$ e indica o espectro normalizado; I^{norm}_{tecido} e $I^{norm}_{sem-tecido}$ é o espectro obtido com o tecido e sem o tecido, respectivamente. Quanto mais próximo de 0 o valor de

σ_{tecido} melhor é o resultado.

2.2.2 Aprendizagem de Máquina

A AM diz respeito ao emprego de algoritmos que extraem padrões de dados. Nesse contexto, é importante mensurar quando o modelo de AM está aprendendo. Nesse sentido os pesquisadores (RUSSELL and NORVIG, 2013) assim conceituam o aprendizado de um agente:

"Um agente estará aprendendo se melhorar o seu desempenho nas tarefas futuras de aprendizagem após fazer observações sobre o mundo."

Os pesquisadores mencionados anteriormente destacam ainda que as melhorias e técnicas para construir um agente que aprende com dados depende de quatro fatores principais:

1) Que componente deve ser melhorado.

2) O conhecimento prévio que o agente já tem.

3) Que representação é usada para os dados e para o componente.

4) Que *feedback* está disponível para a aprendizagem. Os valores reais são chamados de **y** e são usados na aprendizagem, enquanto que o valor predito pelo agente normalmente é chamado de $\hat{\mathbf{y}}$, no qual o *feedback* é o erro, que pode ser medido por para a i-ésima amostra como $e_i = y_i - \hat{y}_i$.

Os autores destacam ainda que em relação ao primeiro fator (componente) são incluídos:

(i) A estruturação das condições do estado atual para realizar as ações.

(ii) Como o conhecimentos sobre o mundo em que o agente está inserido evolui e sobre os resultados de ações possíveis.

(iii) Objetivos que descrevem classes de estados, entre outros. Para o segundo e o terceiro fatores (conhecimento prévio e representação) podem-se incluir vários componentes: sentenças na forma de lógicas, redes bayesianas, problemas de otimização e outros.

Há ainda um quarto fator, que é de onde surge a classificação tradicionald dos tipos de aprendizagem de máquina:

1) **Aprendizagem Por Reforço**

2) **Aprendizagem Não-Supervisionada**

3) **Aprendizagem Supervisionada**

Detalharemos apenas a aprendizagem supervisionada, uma vez que os demais tipos não foram utilizados neste trabalho.

Aprendizagem Supervisionada

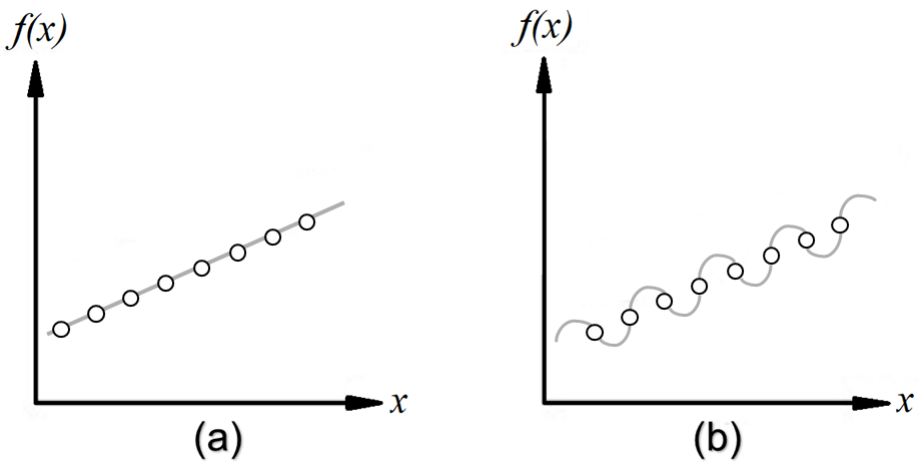
Neste tipo de aprendizagem os padrões dos dados são extraídos fixando-se rótulos de forma prévia. As variáveis de entrada a partir da qual se deseja econtrar algum padrão são chamadas de *features*, e as variáveis de saída são chamados de *labels*.

Os autores (RUSSELL and NORVIG, 2013) definem o que é aprendizagem supervisionada da seguinte forma:

Definição 2.2.1. Dado um conjunto de treinamento de *N* pares de exemplos de entrada e saída $(x_1, y_1), (x_2, y_2), ..., (x_n, y_n)$, com $x_i, y_i \in \mathbb{R}^n$, onde cada y_i foi gerado por uma função y = f(x_i), há uma função h (chamada de hipótese) que mais se aproxima do comportamento de f.

Para mensurar a qualidade da a função hipótese é necessário fornecer um conjunto de testes de exemplos que são distintos do conjunto de treinamento. Nesse aspecto, quando a saída y_i é um conjunto finito de valores (por exemplo: sim, não; ótimo, bom, regular; classe A, classe B, classe C), o problema é chamado de **classificação**. Caso a saída y_i seja um número real, como temperatura, notas de uma prova, o problema é de **regressão**. Vamos focar no processo de aprendizagem por regressão, descrevendo os seis algoritmos usados no trabalho.

Figura 2.6: Exemplos de funções de hipóteses: (a) consistente com os dados; (b) não consistente com os dados



Fonte: (RUSSELL and NORVIG, 2013).

Algoritmos clássicos neste tipo de abordagem são: *SVR*, *DT*, *RF*, *AdaBoost*, RNAs *Multilayer Perceptron* (*MLP*), Regressão Linear Múltipla (*MLR*) e regressão *Lasso*. Dentre as aplicações em tarefas de classificação podemos citar a classificação de imagens, classificação de textos; em problemas de regressão podemos exemplificar a predição de notas de alunos, valores de imóveis.

2.2.3 Métodos de Regressão

Nesta subseção serão descritos de forma sucinta a respeito dos métodos de AM utilizados neste trabalho, bem como suas respectivas vantagens e desvantagens.

Regressão Linear Múltipla

A Regressão Linear Múltipla (*MLR*) é uma técnica bem consoldada na Matemática e na Estatística (FREEDMAN, 2009). Ela foi um dos primeiros tipos de modelos de regressão a serem analisados de forma rigorosa (YAN and SU, 2009). Dentre suas características mais notáveis está o fato de predizer variáveis com alto grau de dependência linear.

Sejam W = (w_1, w_2, ..., w_n) e **x** = (x_1, x_2, ..., x_n), **b** = (b_1, b_2, ..., b_n) os vetores que correspondem aos coeficientes, variáveis depententes e valores a serem verdadeiros, respectivamente. A *MLR* consiste em aplicar o Método dos Mínimos Quadrados, que nada mais é do que uma técnica que procura encontrar os valores tais que a soma residual dos quadrados entre os dados verdadeiros **b** e o hiperplano encontrado sejam os menores possíveis. Em linguagem matemática podemos escrever este problema da seguinte forma:

$$\mathbf{Wx} = \mathbf{b} \tag{2.6}$$

Ao final, a resolução desta equação linear pode ser expressa pela seguinte equação:

$$min_{\boldsymbol{x}} |\mathbf{Wx} - \mathbf{b}|^2 \tag{2.7}$$

Algumas vantagens e desvantagens deste método estão na Tabela 2.1:

Tabela 2.1: Vantagens e Desvantagens da Regressão Linear Múltipla

Vantagens	**Desvantagens**
Alto grau de explicabilidade em problemas lineares	Assume a linearidade nos dados
Simplicidade	Problemas de multicolinearidade quando mais de duas variáveis são colineares
Desempenho computacional eficiente	Sensibilidade a outliers

Fonte: Autor (2023).

Árvore de Decisão

Árvore de Decisão é um algoritmo de aprendizado de máquina supervisionado e não paramétrico, sendo aplicada tanto para problemas de classificação quanto de regressão. Seu objetivo geral é criar um modelo que consiga predizer uma variável a partir de regras de decisão inferidas usando o conjunto de testes.

Para alcançar a decisão, este modelo de AM executa uma sequência de testes. Cada nó (seja a raíz seja um nó interno) corresponde a um teste do valor dos atributos da entrada, e

as arestas dos nós são as classificações de acordo com os possíveis valores do atributo. Assim a condição é formada pelo nó e pela aresta. Cada caminho da raiz até a folha representa uma regra, que são conjunções de condições percorridas que geram um valor, que no caso da árvore de decisão por regressão, é um número.

Dado que a entrada consiste de diversos atributos (categóricos, numéricos ou ambos), para escolher o melhor critério são usados diversos algoritmos, os quais merecem citar: ID3 (*iterative dichotomiser* 3), C4.5, C5.0 e o CART (*Classification and Regression Tree*). Como a biblioteca *scikit-learn* usa o *CART* na sua implementação, os parágrafos seguintes descreverão este algoritmo.

Baseado-se no trabalho de (BREIMAN et al., 1984) sobre arvores de classificação e regressão, o *Classification and Regression Trees* (*CART*) é um algoritmo que gera uma árvore binária de decisão construída dividindo o nó em dois outros nós (filhos) repetidamente, iniciando da raiz da árvore, que contém os critérios de aprendizado. Para problemas de regressão o erro quadrático médio (*MSE*) é usado no algoritmo a fim de mensurar o critério de impureza e realizar a divisão, que é realizada em 3 passos:

1. Encontre a feature que proporciona a melhor divisão;
2. Encontre o melhor nó para a divisão: entre as divisões do passo anterior, escolhe-se àquela que maximiza o critério de divisão (maior valor do *MSE*);
3. Realize a divisão usando o nó encontrado no passo 2 até que o critério de parada seja satisfeito: *MSE* abaixo de um *threshold* ou homogêneo.

Sendo Y = $\{y_1, y_2, ..., y_i\}$ o conjunto que designa a variável dependente, com i = {1, 2, 3, ..., n} onde n $\in$ N, isto é, o número total de variáveis dependentes, e = $\{\hat{y}_1, \hat{y}_2, ..., \hat{y}_i\}$ o conjunto de *labels*, o *MSE* é dado por:

$$MSE = \frac{1}{N}\sum_{i=1}^{N}(y_i - \hat{y}_i)^2 \tag{2.8}$$

Algumas das vantagens e desvantagens da Árvore de Decisão, de acordo com (JAMES et al., 2013), se encontram nas Tabelas (2.2) e (2.3).

Floresta Aleatória

Uma Floresta Aleatória (*Random Forest*) é formada por um número elevado de *DTs*. A aleatoriedade no algoritmo pode ser vista de duas formas possíveis:

i) Cada árvore é baseada é um subconjunto aleatório de observações;

ii) Cada divisão em cada árvore é criada a partir de um conjunto aleatório de variáveis candidatas.

Tabela 2.2: Vantagens de uma Árvore de Decisão

Vantagens
Simples entendimento e fácil interpretação.
Requer uma preparação de dados pouco exaustiva.
O custo no uso da árvore é logarítmico.
Uso de variáveis categóricas e numéricas.
Validação dos testes usando testes estatísticos.
Apresenta um bom desempenho.

Fonte: JAMES et al. (2013).

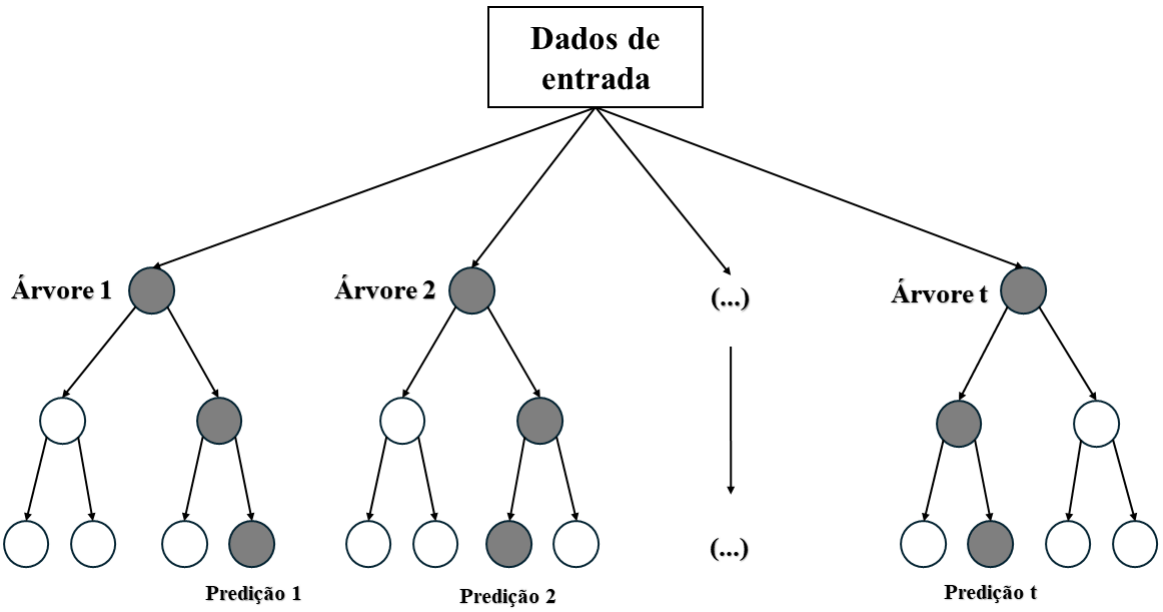
Tabela 2.3: Desvantagens de uma Árvore de Decisão

Desvantagens
O processo de aprendizado pode ser complexo e não representar bem os dados.
Podem ser instáveis.
O algoritmo é guloso e não garante a solução ótima.
O aprendizado pode criar um viés nas variáveis dominantes.

Fonte: JAMES et al. (2013).

Devido a essa aleatoriedade característica, as árvores individuais podem ser bastante instáveis, criando predições distintas em cada árvore. Nessa técnica o valor predito é dado pela predição média do conjunto formado pelas árvores individuais. Podemos visualizar uma Floresta Aleatória genérica com número **t** de árvores na Figura 2.7.

Figura 2.7: Floresta Aleatória genérica



Fonte: Autor (2023).

A Floresta Aleatória é construída usando o algoritmo *CART* da Árvore de Decisão,

que é aplicado em cada uma das ávores (BREAIMAN, 2001). Após selecionar os nós de divisão de forma aleatória, cerca média de 1/3 dos dados não são usadas em nenhuma árvore individual. Eles são "*out of bag (OOB)*". Levando isso em consideração e, de forma semelhante para a Árvore de Decisão, a acurácia de predição para a Floresta Aleatória é dada por:

$$OOB - MSE = \frac{1}{N}\sum_{i=1}^{N}(y_i - \hat{y}_{iOOB})^2 \tag{2.9}$$

onde $\hat{y}_{iOOB}$ é predição para a i-ésima observação de todas as árvores para as quais essa observação foi *OOB*, y_i é a i-ésima amostra de treino e N é a quantidade de *features*.

As principais vantagens e desvantagens mencionadas por (BUSKIRK, 2018) a respeito da floresta aleatória estão nas Tabelas (2.4) e (2.5), respectivamente.

Tabela 2.4: Vantagens da Floresta Aleatória

Vantagens
Eficiência em grandes bancos de dados.
Reduz o problema de *overfitting* e da variância nas árvores de decisão.
Funciona bem com variáveis categóricas e contínuas.
Florestas aleatórias são eficazes na predição de funções complexas com grande interações ou de funções não lineares.

Fonte: BUSKIRK (2018).

Tabela 2.5: Desvantagens da Floresta Aleatória

Desvantagens
A Floresta Aleatória pode ser computacionalmente cara.
A importância de variável pode ser enviesada caso as variáveis estejam correlacionadas.
Os dados ausentes devem ser tratados antes da execução do algoritmo.

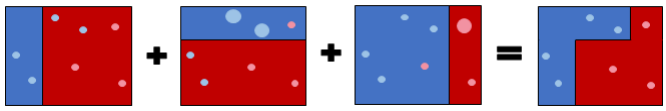
Fonte: BUSKIRK (2018).

Adaptive Boosting

O *Adaptive Boosting* (*AdaBoost*) faz parte de uma família de algoritmos chamados de *ensemble*. Ele foi criado por (FREUND and SCHAPIRE, 1995) para resolver problemas de classificação. De forma concomitante à publicação dos criadores, os pesquisadores (DRUCKER, 1997) adaptaram a ideia para o contexto de regressão.

Seu objetivo básico é combinar o resultado de diversos regressores com resultados ruins para generalizar melhor o conjunto de dados. A ideia chave dessa técnica é construir hipóteses sucessivas que tendem a ser melhoradas de acordo com a atribuição de pesos (w) para cada uma dessas hipóteses.

Figura 2.8: Ilustração do *AdaBoost*



Fonte: Julia Nulkuski (2020).

O raciocínio em problemas de regressão é análogo à resolução da classificação. As principais vantagens desse algoritmo são a sua fácil implementação, flexibilidade e maior capacidade de generalizar os dados. Por outro lado, outliers podem fazê-lo prejudicar sua capacidade de modelagem de dados causando *overfitting*. As principais vantagens de desvantagens deste modelo de AM podem ser vistas na Tabela 2.6.

Tabela 2.6: Vantagens e Desvantagens do *AdaBoost*

Vantagens	**Desvantagens**
Aumento de desempenho	Pode ocasionar *overfitting*
Redução de viés e variância	Sensibilidade a *outliers*
Capacidade de lidar com dados desbalanceados	Pode resultar em alto custo computacional

Fonte: Autor (2023).

k Vizinhos Mais Próximos

O k Vizinhos Mais Próximos (*kNN*) é um algoritmo de AM supervisionado criado em 1951 por (FIX and HODGES, 1951). A ideia essencial deste modelo de AM consiste em predizer a *label* usando uma interpolação associada aos vizinhos mais próximos do valor desejado. Dentre os hiperparâmetros do modelo que são usados na tarefa de interpolação, destacam-se os seguintes:

1) Vizinhança: refere-se à quantidade de rótulos que circundam a *label*.

2) Algoritmo: o algoritmo para construir a vizinhança. Ele pode ser uma *Ball Tree* (OMOHUNDRO, 1989), *KD Tree* (BENTLEY, 1975) ou usar o método de força bruta.

3) Métrica: corresponde ao tipo de distância utilizada para mensurar os vizinhos mais próximos. Este tipo de medida pode ser de vários tipos, sendo o padrão a distância *Minkowski*, que é uma generalização da distância Euclidiana. Uma lista com as distâncias possíveis que podem ser usadas neste algoritmo de AM pode ser encontrada na biblioteca *Scipy* (VIRTANEN et al., 2020) e *Scikit-learn* (PEDREGOSA et al., 2011).

As principais vantagens e desvantagens deste modelo são mencionadas na Tabela 2.7:

Tabela 2.7: Vantagens e Desvantagens do *kNN*

Vantagens	**Desvantagens**
Fácil de entender e implementar	Pode resultar em alto custo computacional quando a vizinhança é comparada com um alto número de dados não rotulados
Treinamento rápido do modelo	Se o algoritmo de vizinhança é lento, pode demorar no treinamento
É robusto para dados com muitos ruídos	Pode resultar em alto consumo de memória
Em geral tem bom desempenho nas tarefas de classificação e regressão	Sensível a estrutura dos dados

Fonte: Jadhav et. al. (2016).

Rede Neural Perceptron

As Redes Neurais Artificiais (RNAs) são um grupo de algoritmos de AM chamados de conexionistas e bioinspirados, pois buscam imitar o comportamento do que se entendia ser o funcionamento do cérebro humano (RUSSELL and NORVIG, 2013). Atualmente existe uma variedade de arquiteturas de RNAs, na qual a primeira arquitetura consistiu na utilização dos neurônios artificiais do tipo Perceptron. Este modelo de neurônio foi expresso matematicamente por (MCCULLOCH and PITTS, 1943), conforme se pode observar na Equação 2.10:

$$a_j = g(\sum_{i=0}^{n} w_{i,j} a_i), \tag{2.10}$$

onde a_i é a ativação de saída da unidade i e $w_{i,j}$ é o peso sobre a ligação i com essa unidade e a_j é a função de ativação da unidade.

As RNAs Percepetron (*MLP*) são compostas por vários neurônios que são conectados por ligações direcionadas. A ligação de um neurônio i para um neurônio j serve para propagar a ativação a_i de i para j. Cada uma dessas ligações possui um peso associado, $w_{i,j}$ que procura representar a intensidade e o sinal de conexão. Há também uma parcela de entrada chamado de *bias*, a_0 = 1 com peso associado $w_{0,j}$. Equacionando, pode-se representar esse conceito da seguinte forma:

$$in_j = \sum_{0}^{n} w_{i,j} a_i, \tag{2.11}$$

onde in_j é a entrada do j-ésimo neurônio.

Em seguida, aplica-se uma **função de ativação** g, chamada de a_j, a essa soma para obter a saída, que pode ser expressa como:

$$a_j = g(in_j) = g(\sum_{i=0}^{n} w_{i,j} a_i) \tag{2.12}$$

A função de ativação pode ser de vários tipos. Pode-mos citar: sigmoide, retificadora linear (ReLU) e tangente hiperbólica. Sem o uso de uma função de ativação, este modelo de AM conexionista não seria capaz de prever dados com padrões não lineares. A ideia chave da aprendizagem em uma RNA Perceptron consiste na atualização dos pesos. Os algoritmos utilizados são o **Gradiente Descendente Estocástico**, **Adam** e o *Broyden–Fletcher–Goldfarb–Shanno* de Memória Limitada (*BFGS*). Além disso, outros hiperparâmetros podem ser definidos, tais como a taxa de aprendizado, o número de neurônios por camadas e a quantidade de camadas. Algumas das vantanges de desvantages de uma RNA Perceptron podem ser vistas na Tabela 2.8.

Tabela 2.8: Vantagens e Desvantagens das RNAs

Vantagens	**Desvantagens**
Capacidade de predizer modelos não lineares	Sensível a escala das *features*
Capactidade de aprendizagem em tempo real	Ajuste de hiperparâmetros pode ser complexo
Aplicável em problemas de Regressão e Classificação	RNAs com muitas camadas pode ter funções de perda com muitos mínimos locais

Fonte: Autor (2023).

Capítulo 3

Relato do Problema

A nanotermometria luminescente é um campo que emergiu na última década. Por meio do uso de NPs, é possível mensurar a temperatura de um tecido biológico de forma minimamente invasiva (ZOHAR et al., 1998); (SUZUKI et al., 2007). Contudo os pesquisadores (SHEN et al., 2020) mostraram como o uso de nanotermômetros luminescentes em tecidos biológicos têm sua leitura comprometida. Isso acontece porque cada tecido tem características que causam distorções espectrais e a modelagem matemática dessas distorções é complexa (BASHKATOV et al., 2011); (BASHKATOV et al., 2016).

Uma forma de modelar as distorções espectrais em tecidos é usando a lei de Beer-Labert modificada(SHEN et al., 2020), que é definida como:

$$I_{det}(\lambda) = I_o(\lambda)e^{-g(\lambda,L)-L[\sum_j \mu j(\lambda)*DPF(\lambda)]} \tag{3.1}$$

onde:

1. $I_o(\lambda)$: a intensidade luminosa no comprimento de onda λ do NLs sem a imersão no tecido.

2. **g**: é uma função que depende da geometria do tecido e que mede a quantidade de perda luminosa.

3. μ_i: é o coeficiente de extinção do j-ésimo cromóforo encontrado no tecido.

4. **DPF**: é o fator de comprimento diferencial (*differential pathlength factor*).

Contudo, esse tipo de modelagem também possui limitações, tal como a questão na variação dos parâmetros dos tecidos. Citando e explicando a alguns destes parâmetros, podemos mencionar:

1 A espessura: busca-se variar a espessura do tecido (em mm) simulando as NPs imersas em diversas profundidades.

2 Escolha entre o coeficiente de distorção ou espalhamento e não ambos, limitando assim a superposição entre eles.

3 Variação da concentração de água de água, gordura, entre outros são valores cujas combinações podem ser muitas, mesmo levando em conta valores fisicamente possíveis.

Os parâmetros mencionados acima causam distorções na medição do espectro e, conforme se poderá ver nos resultados de reconstrução de intensidades usando os modelos de AM, mesmo uma variação mínima nas intensidades resulta em valores de temperatura discrepantes do real. Isso é um prejuízo significativo do ponto de vista do uso dos espectros por pesquisadores, uma vez que busca-se usar essa métrica para auxiliar no diagnóstico de doenças. A precisão e exatidão, então, são dois fatores importantes nesse contexto. Obter a temperatura com a exatidão de uma casa decimal é fundamental em aplicações biológicas, uma vez que a variação deste dado físico nesses ambientes é relevante.

Buscando contornar o problema de distorção espectral causado por diversos fatores, (SHEN et al., 2020) fez o uso de curvas de calibração. Esse método, contudo, é demorado. Além disso, na prática é inviável obter as curvas não distorcidas uma vez que há diversos tipos de tecidos biológicos que podem ser utilizados com o intuito de mapear a temperatura. Além disso, cada um destes tecidos possue parâmetros com valores específicos, alguns dos quais mencionamos anteriormente ao descrever a Equação (5.2).

Os autores (DONG and ZINK, 2014) fizeram um estudo dos nanoambientes para obtenção de temperatura a partir da relação de intensidade e temperatura. Consiste em um método para mensurar a temperatura no interior de NPs que estão aquecidas usando-se um campo magnético. Apesar de promissor para aplicações físicas e biomédicas, não foi encontrado relato de seu uso em tecidos biológicos, bem como não é uma abordagem que utiliza algoritmos de aprendizagem de máquina.

Em outro trabalho, os pesquisadores (WU et al., 2023) utilizaram aprendizagem profunda para obter imagens volumétricas na segunda janela biológica. As imagens obtidas por meio do uso de rede neural tem mostrado riqueza nos detalhes. Apesar de um trabalho promissor, demonstrando a capacidade preditiva de uma RNA em biomedicina, não utiliza NPs em tecidos biológicos.

Outro grupo de pesquisa, os cientistas (LABOUTA et al., 2019) propuseram o uso de mineração de dados usando algoritmos de AM para detectar a citotoxidade de NPs em cerca de 3 mil amostras. O uso de árvore de decisão demonstrou que este classificador possui boas características preditivas, além de ser um modelo caixa-branca.

Conforme se pode notar nos relatos anteriores existem poucas abordagens efetivas em relação ao uso do processo de *KDD*, e mais particularmente ao emprego de técninas de AM em nanotermometria. É nesta lacuna que este trabalho pretende avançar e trazer contribuições, mostrando que os métodos de AM podem ser eficazes. Outros trabalhos relacionados bem como a forma como estas pesquisas foram obtidas está no apêndice deste trabalho (B).

Capítulo 4

Proposta

Neste trabalho propomos o uso de seis algoritmos de AM para realizar a correção do espectro de NPs em tecidos biológicos: regressão linear múltipla, árvore de decisão, floresta aleatória, *adaptative boost*, k-vizinho mais próximo, rede neural artificial *multi-layer perceptron*. O processo de descoberta de conhecimento em base de dados (*KDD*) (FAYYAD et al., 1996) foi adaptado para realizar o treinamento e a validação. Seguindo o *KDD*, tivemos o seguinte *pipeline*:

1) Seleção dos dados: os espectros não distorcidos e distorcidos forma coletados.

2) Pré-processamento: foi realizada uma verificação na qualidade dos dados.

3) Transformação: foi analisada a necessidade de algum tipo de transformação.

4) Mineração: foi usado o *reapeated k-fold cross validation* com k=10 e n=30 para treinamento e obtenção das predições de intensidade.

5) Avaliação: foram usadas as métricas de AM (*MAE*, *MSE*, *MAPE*, R^2) e da física (σ) aliadas ao teste de hipótese de Wilcoxon. As hipóteses serão:

   (a) Hipótese nula: H_0 $y_i = \hat{y}_i$ (o valor médio da intensidade de cada modelo em um comprimento de onda é igual ao valor real).

   (b) Hipótese alternativa: H_1 $y_i \neq \hat{y}_i$ (o valor médio da intensidade de cada modelo em um comprimento de onda é diferente do valor real).

   Além disso, foram plotados para cada um dos seis modelos em cada tecido biológico os espectros corrigidos, bem como gráficos de violino do σ e da temperatura obtida a partir das correções espectrais. Vale relembrar que a temperatura é mensurada a partir das razões dos picos de intensidade I_1 (895 nm) e I_2 (1060 nm) ou I_2 (1060 nm) e I_3 (1360 nm).

Além disso, vale destacar outras propostas específicas desta pesquisa:

- Estabelecer o processo de *KDD* como referência na construção do *pipeline* para o problema com o intuito de realizar investigações futuras: tratando-se do uso de técnicas de AM para nanotermometria, construir uma metodologia é fundamental para a validação e trabalhos futuros.

- Descobrir quais algoritmos são efetivos para a correção dos espectros: uma das etapas na mineração de dados é encontrar as melhores técnicas para cada contexto de aplicação.

- Comparar a capacidade preditiva dos modelos em relação a obtenção da temperatura a partir das curvas de calibração corrigidas pelas técnicas de AM com o que é utilizado na física: estabelecer o quanto elas são diferentes é fundamental, uma vez que mostrará quais grupo de algoritmos são piores e melhores na correção dos espectros. Além disso, deve-se levar em conta fatores importantes relacionados aos modelos: tempo de treinamento, métricas de avaliação de AM e da física.

Capítulo 5

Metodologia de Experimentos

Este capítulo descreverá a metodologia abordada, bem como de que forma os dados serão adiquiridos, treinados e avaliados. Seu entendimento depende da compreensão referente à **Fundamentação Teórica**.

5.1 Procedimento

5.1.1 Seleção, Pré-Processamento e Transformação

A Figura 5.1 ilustra o procedimento geral para aquisição dos dados. Os espectros na fase 1-a foram obtidos por meio dos colaboradores. Eles correspondem às NPs em temperatura ambiente, que corresponde à 301.2 K. Os dados consistem de 1121 pontos de um gráfico 2D de pares (λ, I(λ), no qual o eixo das abcissas é o comprimento de onda medido em nanômetros (nm) e o eixo das ordenadas é a intensidade. O comprimento de onda são valores do intervalo fechado de 850.0 nm até 1410.0 nm, com passos de 0.5 nm, ou seja, 850.0, 850.5, 851.0, ..., 1410.

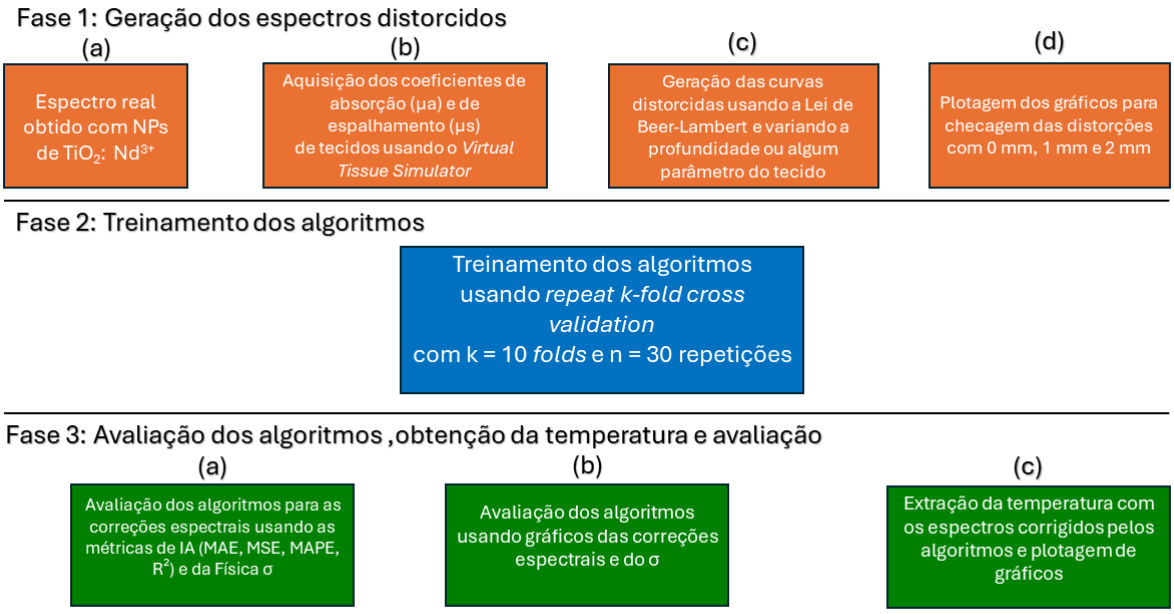
Na etapa 1-b foram obtidos os seguintes dados para os comprimentos de onda presentes no intervalo fechado descrito anteriormente.

1) coeficientes de absorção (μ_a) dos seguintes tecidos: pele (*skin*), massa cinzenta do cérebro (*brain gray matter*), mama pré-menopausa (*breast pre menopausa*), e fígado (*liver*) com valores padrão. O da pele, além dos dados com valores padrão, foram extraídos diversos gráficos, no qual cada um foi variada a concentração de H_2O.

2) coeficientes de espalhamento (μ_s') dos tecidos: pele (*skin*), massa cinzenta do cérebro (*brain gray matter*), mama pré-menopausa (*breast pre menopause*), e fígado (*liver*) com valores padrão.

3) coeficiente de absorção da água foi obtido através do trabalho de (JACQUES and MCAULIFFE, 1991) que pode ser encontrado aqui. Em seguida foi realizada uma interpolação de grau 3 para obtenção dos 1121 pares de pontos (I(μa, λ).

Figura 5.1: *Overview* da aplicada desde a obtenção dos dados até a obtenção avaliação da temperatura



Fonte: Autor (2024).

Os casos 1 e 2 foram aquisições do *Virtual Tissue Simulator* (HAYAKAWA et al., 2022). A Figura (5.2) exibe a interface gráfica do software.

Na fase 1-c, para distorcer as curvas dos tecidos usando ambos os coeficientes, foi utilizada a lei de Beer-Lambert (BOUGUER, 1729); (LAMBERT, 1760), que é dada por:

$$I_{det}(\lambda) = I_o(\lambda)e^{-g(\lambda,L)-L[\sum_j \mu j(\lambda) x DPF(\lambda)]} \tag{5.1}$$

onde:

1. $I_o(\lambda)$: a intensidade luminosa no comprimento de onda λ das NPs sem a presença do tecido.

2. **g**: é uma função que depende da geometria do tecido e que mede a quantidade de perda luminosa.

3. μ_i: é o coeficiente (de absorção ou espalhamento) do i-ésimo cromóforo encontrado no tecido.

4. **DPF**: é o fator de comprimento diferencial (*differential pathlength factor*).

Para o contexto dos experimentos a equação foi reescrita como:

$$I_{det}(\lambda) = I_o(\lambda)e^{-\alpha(\lambda)*L_i} \tag{5.2}$$

onde:

1. $I_o(\lambda)$: a intensidade luminosa no comprimento de onda λ das NPs sem a presença do tecido.
2. α: coeficiente de absorção (μ_a) ou de espalhamento (μ_s) no comprimento de onda λ
3. L_i: espessura do tecido.

Os parâmetros de cada tecido para o coeficiente de absorção μ_a e as respectivas quantidades de curvas geradas estão na Tabela (5.1):

Tabela 5.1: Parâmetros dos tecidos para obtenção do μ_a (coeficiente de absorção) e qtde. de curvas

Tecido	**Hemoglobina (μM)**	**Oxiemoglobina (μM)**	**Água** (vol. frac.)	**Gordura (vol. frac.)**	**Melanina (vol. frac.)**	**Número de curvas**
Massa cinzenta do cérebro	24.0	56.0	0.8	0.12	Não possui	100
Mama pré-menopausa	6.9	19.6	0.345	0.41	Não possui	100
Pele	28.4	22.4	0.7	0.0	0.0051	100
Pele com H2O modificado	28.4	22.4	Variável	0.0	0.0051	60
Fígado	66.0	124.0	0.87	0.02	Não possui	100

Fonte: Autor (2024).

Para os tecidos em que foram obtidos 100 curvas a espessura do mesmo variou de 0 mm (sem tecido) até 2 mm com incremento de aproximadamente 0.02 mm. Já para a pele, em que foam obtidas 60 curvas, foi alterada a concentração de H_2O de 0.4 a 0.71 com incrementos de 0.005 (Tabela 5.1).

As curvas obteidas com os tecidos relativos ao coeficiente de espalhamento mus' e a respectiva quantidade de curvas estão na Tabela (5.2).

Em seguida, na fase 1-d, foram plotados gráficos para checar se a variação espectral estava aceitável considerando um contexto real em termos de aplicação práticas, bem como visualizar se a alteração nos espectros era adequada para realizar o treinamento. Vale mencionar

Tabela 5.2: Parâmetros dos tecidos para obtenção do μ_s (coeficiente de espalhamento) e da respectiva quantidade de curvas

Tecido	**Hemoglobina (μM)**	**Oxiemoglobina (μM)**	**Água** (vol. frac.)	**Gordura (vol. frac.)**	**Melanina (vol. frac.)**	**Número de curvas**
Massa cinzenta do cérebro	24.0	56.0	0.8	0.12	Não possui	100
Mama pré-menopausa	6.9	19.6	0.345	0.41	Não possui	100
Pele	28.4	22.4	0.7	0.0	0.0051	100
Fígado	66.0	124.0	0.87	0.02	Não possui	100

Fonte: Autor (2024).

também que a variação da espessura de 0 mm a 2 mm simula as NPs presentes em diferentes profundidades nos tecidos.

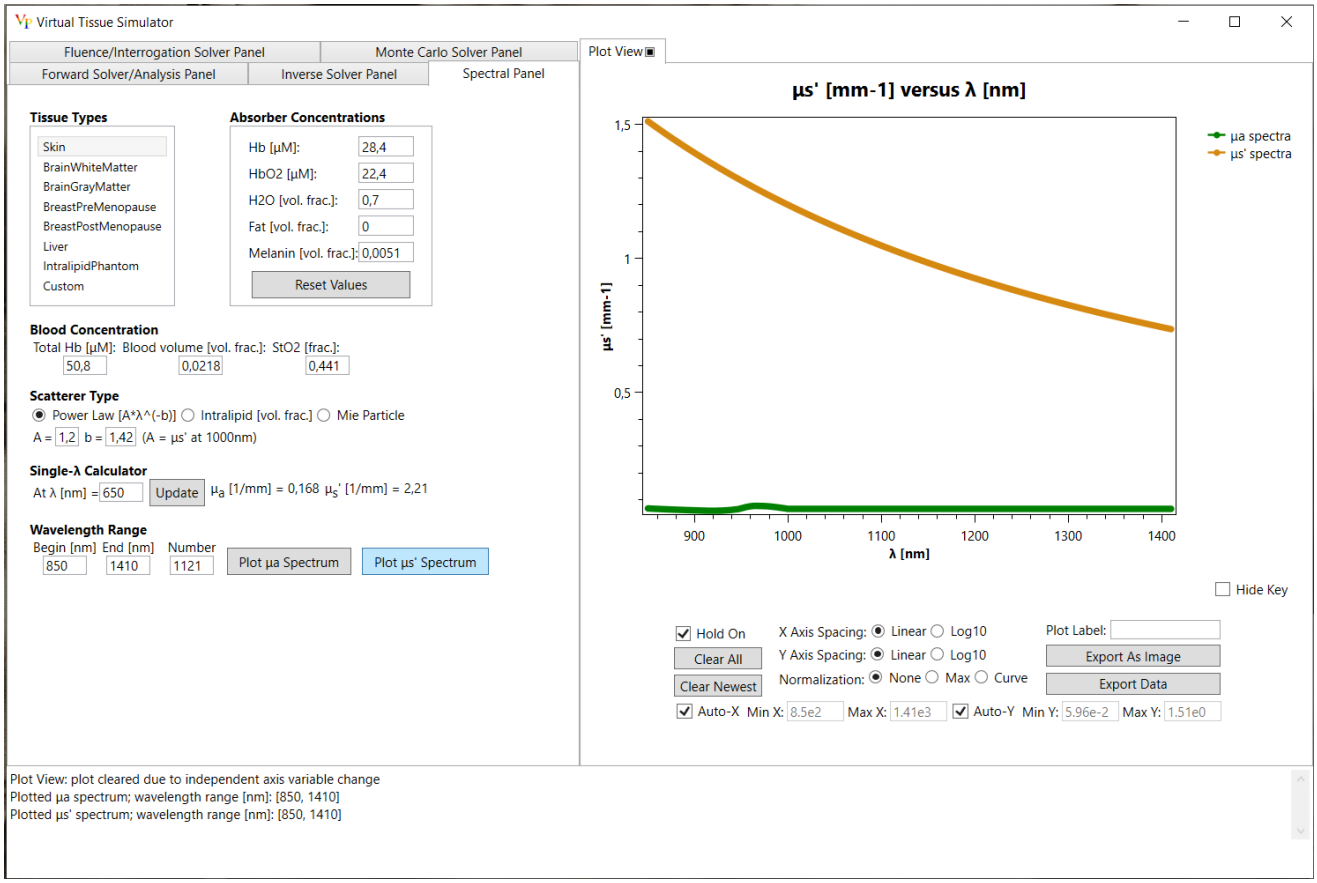
5.1.2 Mineração

Na fase 2 foi realizado o treinamento dos modelos usando o *repeated k-fold cross validation* com n = 30 repetições e k = 10 *folds*. Tais valores correspondem àqueles já consolidados na literatura em outros tipos de problemas. Em relação aos hiperparâmetros de cada um dos 6 modelos, via de regra eles foram mantidos em seus valores padrão. À excessão se deu para o algoritmo de Floresta Alatória, cujo número de árvores foi ajustado para ser a raiz quadrada da quantidade de amostras, sendo um dos possíveis valores sugestivos mencionados por (GRÖMPING, 2009).

Os algoritmos são escritos em python a partir da biblioteca *scikit-learn* (PEDREGOSA et al., 2011). O código está disponibilizado no seguinte repositório: Experimentos para Correção Espectral.

O hardware utilizado consiste de uma máquina com 16 GB de RAM, processador i5-10500H com 6 cores, 12 threads e uma placa de vídeo GTX 1650 com 4 GB de memória dedicada. Os treinamentos com os seguintes tecidos distorcidos usando o coeficiente de absorção μ_a foram realizados em Windows, à exceção do algoritmo de Floresta Aleatória: pele, mama pré-menopausa, fígado, pele, pele variando H2O e água. Os treinamentos dos tecidos realizados com dados gerados a partir dos coeficientes de espalhamento μ_s foram realizado em Ubuntu 22.04 *LTS*. Esse fato vale a pena ser mencionado, pois resultou em tempo de treinamento consideravelmente menor neste último sistema operacional, sendo cerca de três vezes mais rápido. Um resumo geral desses detalhes estão nas Tabelas (5.3, 5.4, 5.6).

Figura 5.2: Interface gráfica do *Virtual Tissue Simulator*



Fonte: Autor (2024).

Tabela 5.3: Treinamento usando as curvas da Massa Cinzenta do Cérebro, Mama Pré-Menopausa, Fígado e Pele com μ_a e μ_s

Algoritmo	**SO de treinamento**
Regressão Linear Múltipla (*MLR*)	Windows
Árvore de Decisão (*DT*)	Windows
Floresta Aleatória (*RF*)	Ubuntu 22.02
AdaBoost	Windows
k Vizinhos Mais Próximos (*kNN*)	Windows
Rede Neural Artificial Multi-Layer Perceptron (*MLP*)	Windows

Fonte: Autor (2024).

5.1.3 Avaliação e Interpretação

Na fase 3-a, para cada *feature* (λ, I) com a respectiva predição, foram extraídas as métricas de desempenho de IA: *MAE*, *MSE*, *MAPE*, R^2; e também a métrica de avaliação σ. Nos tecidos com 100 curvas temos 100 predições possíveis para cada *feature* que, em virtude das

Tabela 5.4: Treinamento usando as curvas da Pele com variação na concentração de água e usando o μ_a

Algoritmo	**SO de treinamento**
Regressão Linear Múltipla (*MLR*)	Windows
Árvore de Decisão (*DT*)	Windows
Floresta Aleatória (*RF*)	Ubuntu 22.02
AdaBoost (*ADA*)	Windows
k Vizinhos Mais Próximos (*kNN*)	Windows
Rede Neural Artificial Multi-Layer Perceptron (*MLP*)	Windows

Fonte: Autor (2024).

Tabela 5.5: Treinamento usando as curvas da Água e usando o μ_a

Algoritmo	**SO de treinamento**
Regressão Linear Múltipla (*MLR*)	Ubuntu 22.02
Árvore de Decisão (*DT*)	Ubuntu 22.02
Floresta Aleatória (*RF*)	Ubuntu 22.02
AdaBoost	Ubuntu 22.02
k Vizinhos Mais Próximos (*kNN*)	Windows
Rede Neural Artificial Multi-Layer Perceptron (*ANN MLP*)	Ubuntu 22.02

Fonte: Autor (2024).

Tabela 5.6: Treinamento usando as curvas da Água e usando o μ_s

Algoritmo	**SO de treinamento**
Regressão Linear Múltipla (*MLR*)	Windows
Árvore de Decisão (*DT*)	Windows
Floresta Aleatória (*RF*)	Ubuntu 22.02
AdaBoost (*ADA*)	Windows
k Vizinhos Mais Próximos (*kNN*)	Windows
Rede Neural Artificial Multi-Layer Perceptron (*ANN MLP*)	Windows

Fonte: Autor (2024).

30 repetições, acumulam 3000 resultados em cada comprimento de onda λ. De modo análogo, no tecido com 60 curvas temos 60 predições possíveis para cada *featura* que acumulam 1800 resultados possíveis pelo mesmo motivo anterior (n = 30 repetições).

A fase 3-b foi feita usando 2D, em que o eixo da abcissa é o comprimento de onda e o eixo da ordenada é a média das diferenças de intensidades no comprimento de onda λ_i, chamado de ΔI. Para cada tecido, foram plotados os resultados dos algoritmos. Para os valores de σ foi também extraído o gráfico de violino por tecido, exibindo os resultados de cada um dos

6 algoritmos utilizados.

A última etapa, 3-c, foi extraída a temperatura com os espectros corrigidos. Os dados para extração da curva de temperatura foram fornecidos pelos colaboradores. Os pontos são pares ordenados do tipo (temperatura (em Kelvin), razão (adimensional)), onde a razão é o quociente entre os picos I_2 e I_1 nos comprimentos de onda 895.0 nm e 1067 nm, respectivamente, ou dos picos I_2 e I_3, este último que corresponde ao valor de 1340 nm. Foram obtidos ps valores correspondendo a temperatura média de cada algoritmo em cada um dos tecidos, bem como o gráfico de violino. Para efeitos de comparação os resultados foram colocados numa tabela do seguinte modo: uma coluna correspondendo à distorção com máxima profundidade ou, no caso da pele, com valor de H_2O em seu valor mínimo.

Noutra coluna foi colocado o valor de temperatura com o espectro corrigido pelo algoritmo, considerando seu valor médio. Para cada um destes casos foi calculada a variação percentual média considerando a temperatura das NPs sem tecido, cujo valor de referência foi de 301.2 K ($\approx$ 28.2 ºC) usando a razão das intensidades I_1 e I_2, que foi aplicada nos tecidos usando o coeficiente de absorção (μ_a); e de 304.1 K ($\approx$ 29.1 ºC) obtida pela razão das intensidades I_2 e I_3, que foi usada nos tecidos com o coeficiente de espalhamento (μ_s).

Capítulo 6

Resultados e Discussão

Neste capítulo serão discutidos os resultados de correção espectral para cada um dos modelos de AM, seguido pelos valores obtidos do σ e das temperaturas. Para essas duas últimas medidas, serão usados dois tipos de gráficos:

1) O primeiro será um *plot* 2D usando os pares (λ, I). Nele, serão mostrados o espectro não distorcido, o espectro corrigido pelo algoritmo usando a média aritmética para cada comprimento de onda e uma margem indicando as variações nas predições de cada modelo. Todos estes gráficos estão no Apêndice A (A).

2) O segundo é um gráfico do tipo violino do σ e das temperaturas obtidas a partir dos espectros corrigidos por cada modelo de AM.

Além disso, o leitor também deverá ficar atento para as tabelas: os melhores valores em cada coluna estão destacados em negrito e na cor azul. No caso em que houver valores iguais na mesma coluna, ambos serão destacados. Ainda em relação a temperatura, as amostras de NPs terão dois valores:

1) Para os tecidos com distorções usando μ_a e profundidade a temperatura obtida a partir do espectro não distorcido é de 301.2 K.

2) Para os tecidos com distorções usando μ_a e profundidade a temperatura obtida a partir do espectro não distorcido é de 304.1 K.

6.1 Resultados e Discussão nos tecidos com distorções usando μ_a e profundidade

6.1.1 Massa Cinzenta do Cérebro

Os algoritmos treinados com a Massa Cinzenta do Cérebro resultaram em métricas com valores próximos do melhor caso possível, considerando o contexto de problemas de regressão. A *MLR* realizou o treinamento mais rápido, com cerca de 28 segundos e valores médios

de *MAE*, *MSE*, *MAPE* e R² de 0,0458, 0,0157, 0,2164 e 0,9991, respectivamente. O σ resultou em um valor de 0,0069. Em contraste, com relação ao tempo de treinamento, o algoritmo que mais demorou foi o *RF*, usando um tempo de quase 52 minutos. Para este algoritmo os valores médios de *MAE*, *MSE*, *MAPE* e R² foram de 0,0003, 0,0, 0,0004 e 1,0, respectivamente, enquanto que o valor médio do σ foi de 0,0002. Resultados similares nas métricas foram obtidos nos algoritmos de *DT* e *kNN*, com tempo de treinamento consideravelmente menor, sendo de 1 minuto e 10 segundos para a Árvore de Decisão e de 38 segundos para o *kNN*. Os piores resultados nas métricas de desempenho foram do *AdaBoost*, com valores médios de *MAE*, *MSE*, *MAPE*, R² e σ de 0,2688, 0,0923, 5,8684, 0,9948, 0,1829, respectivamente; e da *MLP*, com os seguintes valores médios de *MAE*, *MSE*, *MAPE*, R² e σ: 0,2688, 0,0923, 5,8684, 0,9948, 0,1829. A *MLP* foi cerca de duas vezes mais rápida que o *AdaBoost* no tempo de treinamento, sendo de aproximadamente 10 minutos para o primeiro e de 22 minutos e 50 segundos para o segundo. Os valores de métricas com os respectivos valores de cada desvio padrão estão resumidos na Tabela (6.1).

A *MLR* se destacou negativamente nas métricas de desempenho, pois se trata de um modelo linear tentando resolver um problema não linear. Já o desempenho insatisfatório do *AdaBoost* pode ser explicado pelo baixo número de estimadores, que no nosso caso foi de 50, além do uso de uma função de perda linear. Para efeitos comparativos, o *RF* usou 105 estimadores (árvores de decisão), enquanto o critério de seleção foi o *MSE*. Em relação à *MLP*, seu baixo desempenho pode ser explicado pelo número elevado de camadas ocultas, no qual um dos problemas comuns é o otimizador ficar em um mínimo local. Além disso, o uso de uma função de ativação linear, a ReLU, mantém a linearidade neste modelo conexionista, o que pode ser explicado pelo fato de que aplicar uma transformação linear, ou seja, usar a ReLU como função de ativação, preserva a linearidade na busca de uma solução.

Tabela 6.1: Principais dados de avaliação de correção espectral na Massa Cinzenta do Cérebro

Algoritmo	**Tempo**	***MAE* médio e *std***	***MSE* médio e *std***	***MAPE* médio e *std***	**R² médio e *std***	**σ médio e *std***
MLR	**28s**	0,0458 +/- 0,0010	0,0157 +/- 0,0007	0,2164 +/- 0,0163	0,9991 +/- 0,0	0,0233 +/- 0,0034
DT	1min10s	0,0002 +/- 0,0001	**0,0 +/- 0,0**	0,0002 +/- 0,0	**1,0 +/- 0,0**	0,0002 +/- 0,0002
RF	51min58s	0,0003 +/- 0,0	**0,0 +/- 0,0**	0,0004 +/- 0,0	**1,0 +/- 0,0**	0,0002 +/- 0,0002
AdaBoost	22min49s	0,2688 +/- 0,0225	0,0923 +/- 0,0146	5,8684 +/- 0,6327	0,9948 +/- 0,0009	0,1829 +/- 0,0146
kNN	38s	**0,0 +/- 0,0**	**0,0 +/- 0,0**	**0,0 +/- 0,0**	**1,0 +/- 0,0**	**0,0 +/- 0,0**
MLP	9min49s	0,0809 +/- 0,0382	0,0185 +/- 0,0112	0,6502 +/- 0,5293	0,9990 +/- 0,0007	0,0441 +/- 0,0186

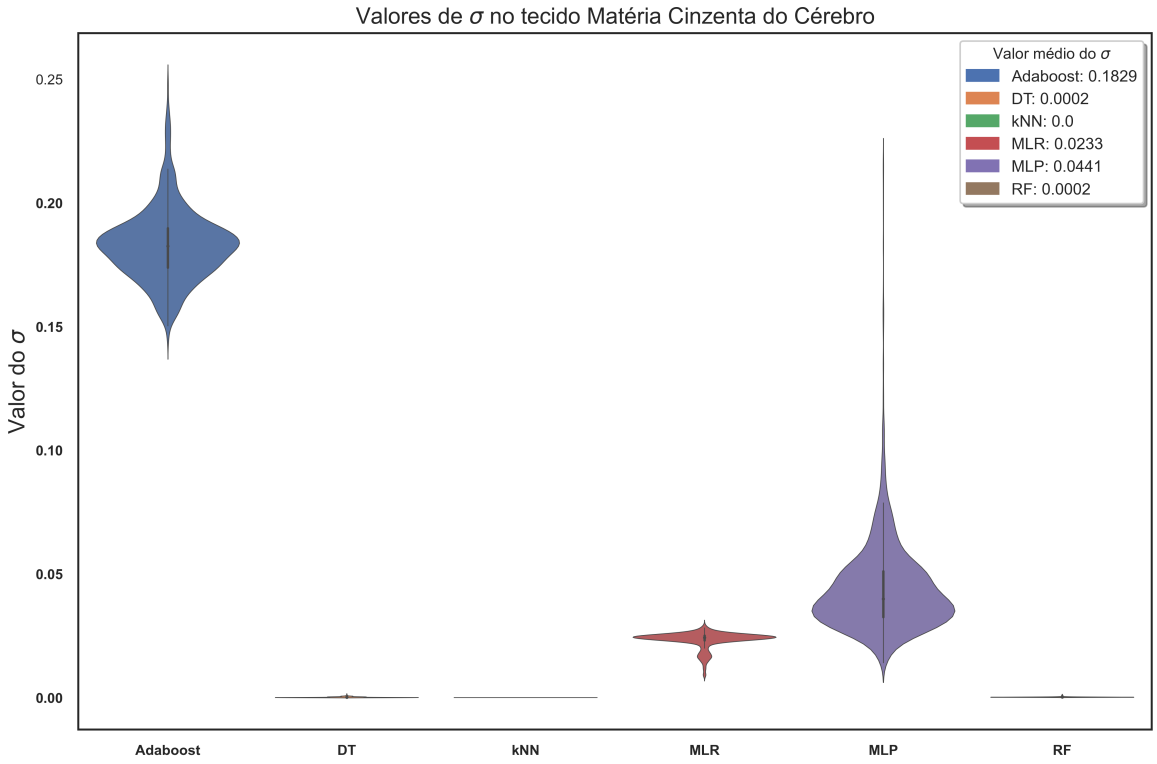
Fonte: Autor (2024).

No teste de hipótese de Wilcoxon, considerando o valor médio de cada *label*, obteve

um p-valor de 0,0 para a Regressão Linear Múltipla, 0,0 para o *AdaBoost*, *MLP* de 0,0 e de 0,001 para o *kNN*. Isso significa que para estes algoritmos o valor médio predito de intensidade I para o comprimento de onda λ não é o real. Os demais algoritmos, *DT* e *RF* resultaram no p-valor de 0,1821 e 0,8690, respectivamente, ou seja, o valor médio predito por eles é o valor real.

A reconstrução de intensidade das curvas para este tecido nos seis algoritmos pode ser encontrada nas Figuras (A.1) e (A.2).

Figura 6.1: Gráfico de violino do σ obtido a partir da correção espectral usando os seis modelos de AM na **massa cinzenta do cérebro**



Fonte: Autor (2024).

A variação mais acentuada do σ foi no *AdaBoost*, seguido pela *MLP*. Na *MLR* foi visualmente significativa, porém menor que nos outros dois algoritmos anteriores. Os melhores resultados foram dos *kNN*, seguidos pela *DT* e *RF*, cujo valor do σ foi igual. Maiores detalhes podem ser encontrados na Figura (6.1).

O gráfico de violino da temperatura (Figura 6.2) mostraram uma variação considerável no *MLR*, *MLP* e *AdaBoost*. Já a temperaura obtida usando as curvas corrigidas pela *DT*, pelo *RF* e através do *kNN* foi altamente precisa e exata.

Figura 6.2: Variação da temperatura obtida pela correção espectral usando os algoritmos de AM na **massa cinzenta do cérebro**

Fonte: Autor (2024).

6.1.2 Mama Pré-Menopausa

Os regressores aplicados na Mama Pré-Menopausa resultaram em métricas bem próximas dos melhores valores possíveis, considerando outros problemas de regressão. A *MLR* foi menos custoso no tempo de treinamento, gastando apenas 43 segundos e com valores médios de *MAE*, *MSE*, *MAPE*, R² e σ de 0,0194, 0,0228, 0,0963, 0,9998, 0,0099, respectivamente. Por outro lado tempo de treinamento do *RF* foi de 2 horas e 30 minutos com valores médios de *MAE*, *MSE*, *MAPE*, R² e σ de 0,0002, 0,0001, 0,0001, 1,0 e 0,0002, respectivamente. Resultados similares nos valores médios das métricas foram obtidos nos algoritmos de *DT* e *kNN*, com tempo de treinamento menor, sendo de 2 minutos e 33 segundos para a Árvore de Decisão e de apenas 55 segundos para o *kNN*. Os piores resultados nas métricas de desempenho foram do *AdaBoost*, com valores médios de*MAE*, *MSE*, *MAPE*, R² e σ de 0,2530, 0,0820, 5,0835, 0,9953, 0,1608, respectivamente; e da *MLP*, com os seguntes valores de *MAE*, *MSE*, *MAPE*, R² e σ: 0,0525, 0,0055, 0,5818, 1,0, 0,0331. O *AdaBoost* foi cerca de duas vezes mais lento que a *MLP* no tempo de treinamento. Os valores com os respectivos desvios padrões estão resumidos na Tabela (6.2).

A *MLR* apresentou um desempenho inferior nas métricas, devido ao fato de ser um modelo linear aplicado a um problema não linear. O desempenho insatisfatório do *AdaBoost* pode ser atribuído ao baixo número de estimadores, que no nosso caso foi de 50, além do uso de uma função de perda linear. Para comparação, o *RF* utilizou 105 estimadores (árvores de decisão), com o critério de seleção sendo o *MSE*. No caso da *MLP*, seu desempenho insatisfatório pode ser explicado pelo grande número de camadas ocultas, onde um dos problemas recorrentes é o otimizador ficar preso em um mínimo local. Além disso, a função de ativação linear, ReLU, contribui para a manutenção da linearidade nesse modelo conexionista, o que pode ser entendido pelo fato de que ao aplicar uma transformação linear, como a ReLU, a linearidade da solução é preservada.

No teste de hipótese de Wilcoxon, considerando o valor médio de cada *label*, obteve um p-valor de 0,0 para a Regressão Linear Múltipla, 0,0 para o *AdaBoost*, *MLP* de 0,0 e de 0,001 para o *kNN*. Isso significa que para estes algoritmos o valor médio predito de intensidade I para o comprimento de onda λ não é o real. Os demais algoritmos, *DT* e *RF* resultaram no p-valor de 0,9123 e 0,8691, respectivamente, ou seja, o valor médio predito por eles é o valor real.

A reconstrução de intensidade das curvas para este tecido nos seis algoritmos pode

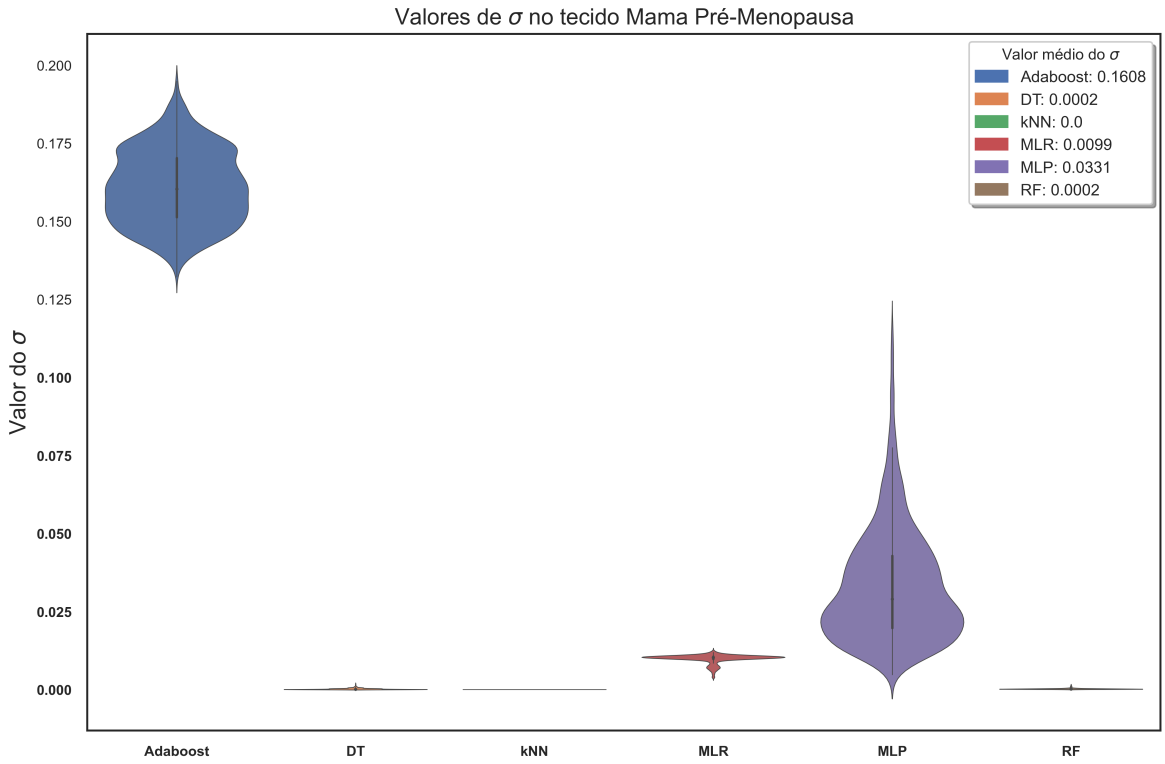
Tabela 6.2: Principais dados de avaliação de correção espectral na Mama Pré-Menopausa

Algoritmo	**Tempo**	***MAE* médio e *std***	***MSE* médio e *std***	***MAPE* médio e *std***	**R² médio e *std***	**σ médio e *std***
MLR	**43s**	0,0194 +/- 0,0004	0,0028 +/- 0,0001	0,0963 +/- 0,0073	0,9998 +/- 0,0	0,0099 +/- 0,0014
DT	2min33s	0,0002 +/- 0,0001	0,0001 +/- 0,0	0,0001 +/- 0,0	**1,0 +/- 0,0**	0,0002 +/- 0,0002
RF	2h30min43s	0,0003 +/- 0,0	**0,0 +/- 0,0**	0,0002 +/- 0,0	**1,0 +/- 0,0**	0,0002 +/- 0,0002
AdaBoost	22min49s	0,253 +/- 0,0168	0,0820 +/- 0,0091	5,0835 +/- 0,5539	0,9953 +/- 0,0005	0,1608 +/- 0,0119
kNN	55s	**0,0 +/- 0,0**	**0,0 +/- 0,0**	**0,0 +/- 0,0**	**1,0 +/- 0,0**	**0,0 +/- 0,0**
MLP	13min33,0s	0,0525 +/- 0,0293	0,0055 +/- 0,0042	0,5818 +/- 0,4438	0,9997 +/- 0,0002	0,0331 +/- 0,0181

Fonte: Autor (2024).

ser encontrada nas Figuras (A.3) e (A.4).

Figura 6.3: Gráfico de violino do σ obtido a partir da correção espectral usando os seis modelos de AM na **mama pré-menopausa**



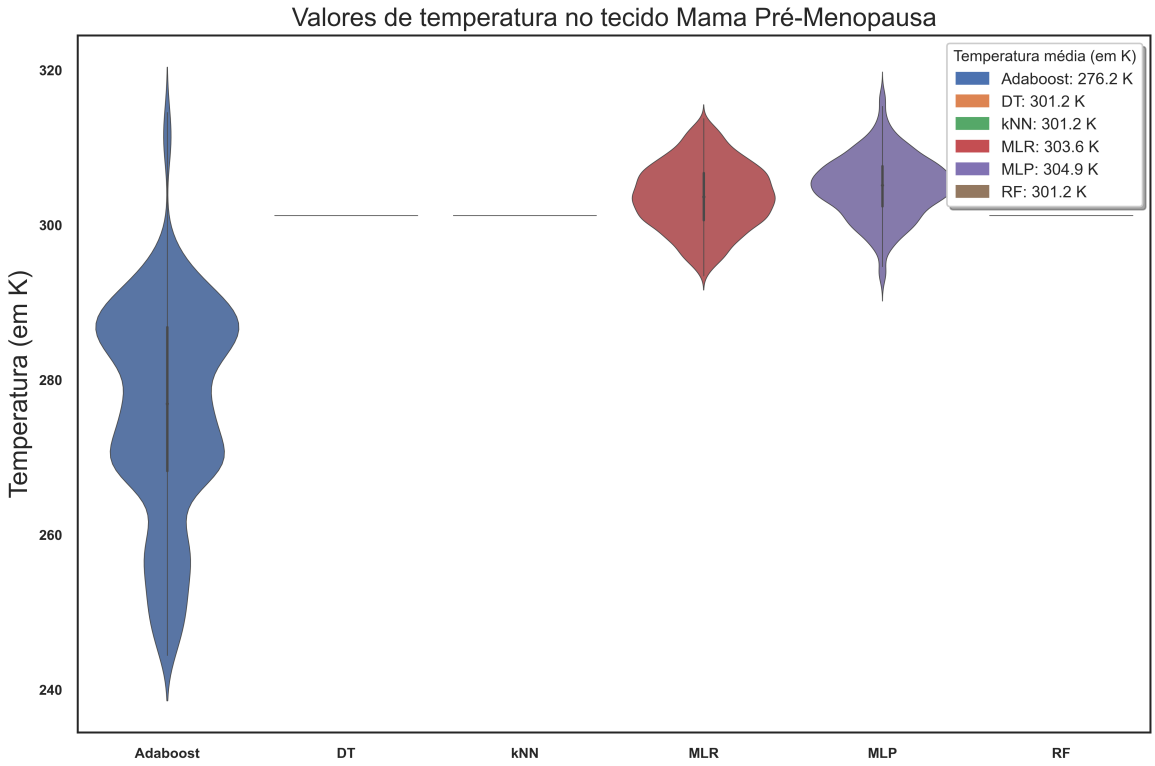
Fonte: Autor (2024).

A variação mais acentuada do σ foi no *AdaBoost*, seguido pela *MLP*. Na *MLR* foi visualmente significativa, porém menor que nos outros dois algoritmos anteriores. Os melhores

resultados foram dos *kNN*, seguidos pela *DT* e *RF*, cujo valor do σ foi igual. Maiores detalhes podem ser encontrados na Figura (6.3).

O gráfico de violino da temperatura (Figura 6.4) mostrou uma variação considerável na *MLP* e no *AdaBoost*. Já a temperaura obtida usando as curvas corrigidas pelo *AdaBoost*, apesar de não ter sido preciso, teve valores exatos. Os melhores resultados dessa grandeza física foi obtida pela *DT*, *RF* e *kNN*, que além exatos, foram precisos.

Figura 6.4: Gráfico de violino da temperatura obtido a partir da correção espectral usando os seis modelos de AM na **mama pré-menopausa**



Fonte: Autor (2024).

6.1.3 Fígado

No Fígado, o algoritmo que mais se destacou em tempo de treinamento foi a*MLR* gastando apenas 40 segundos. Suas os valores médios das métricas de desempenho foram de 0,0631, 0,0325, 0,1354, 0,9982, 0,0318, que correspondem aos valores de *MAE*, *MSE*, *MAPE*, R^2 e σ, respectivamente. Já o tempo de treinamento do *RF* foi de quase 2 horas e 52 minutos com valores médios de *MAE*, *MSE*, *MAPE*, R^2 e σ de 0,0003, 0,0, 0,0005, 1,0, 0,0002, respectivamente. Resultados semelhantes nas métricas foram obtidos nos algoritmos de *DT* e *kNN*, com tempo de treinamento menor, sendo de 2 minutos e 40 segundos para a Árvore de Decisão e de 1 minuto para o *kNN*. Os piores resultados nas métricas de desempenho foram do *AdaBoost*,

com valores médios de *MAE*, *MSE*, *MAPE*, R² e σ de 0,3249, 0,1389, 6,9437, 0,9921, 0,2217, respectivamente; e da *MLP*, com os seguintes valores médios de *MAE*, *MSE*, *MAPE*, R² e σ: 0,1554, 0,0525, 1,4516, 0,9970, 0,0876. O *AdaBoost* resultou em um tempo de treinamento de cerca de 28 minutos, enquanto a *MLP* obteve um custo no tempo de quase 15 minutos. Os valores com os respectivos desvios padrões estão resumidos na Tabela (6.3).

O desempenho da *MLR* foi insatisfatório nas métricas, pois trata-se de um modelo linear sendo utilizado em um problema não linear. O desempenho abaixo do esperado do *AdaBoost* pode ser explicado pelo número reduzido de estimadores, que no nosso caso foi de 50, além da adoção de uma função de perda linear. Para efeito de comparação, o *RF* empregou 105 estimadores (árvores de decisão) e utilizou o *MSE* como critério de seleção. No que se refere à *MLP*, seu desempenho insatisfatório pode ser atribuído ao elevado número de camadas ocultas, o que frequentemente leva o otimizador a se prender em um mínimo local. Além disso, a função de ativação ReLU, por ser linear, mantém a linearidade no modelo conexionista, o que ocorre porque a aplicação de uma transformação linear, como a ReLU, preserva a linearidade na busca pela solução.

Tabela 6.3: Principais dados de avaliação de correção espectral no Fígado

Algoritmo	**Tempo**	***MAE* médio e *std***	***MSE* médio e *std***	***MAPE* médio e *std***	**R² médio e *std***	**σ médio e *std***
MLR	**40s**	0,0631 +/- 0,0015	0,0325 +/- 0,0015	0,1354 +/- 0,0094	0,9982 +/- 0,0001	0,0318 +/- 0,0055
DT	2min40s	0,0002 +/- 0,0001	**0,0 +/- 0,0**	0,0002 +/- 0,0	**1,0 +/- 0,0**	0,0002 +/- 0,0
RF	2h51min53s	0,0003 +/- 0,0	**0,0 +/- 0,0**	0,0005 +/- 0,0	**1,0 +/- 0,0**	0,0002 +/- 0,0001
AdaBoost	28min6s	0,3249 +/- 0,0365	0,1389 +/- 0,0280	6,9437 +/- 0,9395	0,9921 +/- 0,0016	0,2217 +/- 0,0025
kNN	1min	**0,0 +/- 0,0**	**0,0 +/- 0,0**	**0,0 +/- 0,0**	**1,0 +/- 0,0**	**0,0 +/- 0,0**
MLP	14min56s	0,1554 +/- 0,0914	0,0525 +/- 0,0559	1,4516 +/- 1.2658	0,9970 +/- 0,0032	0,0876 +/- 0,0433

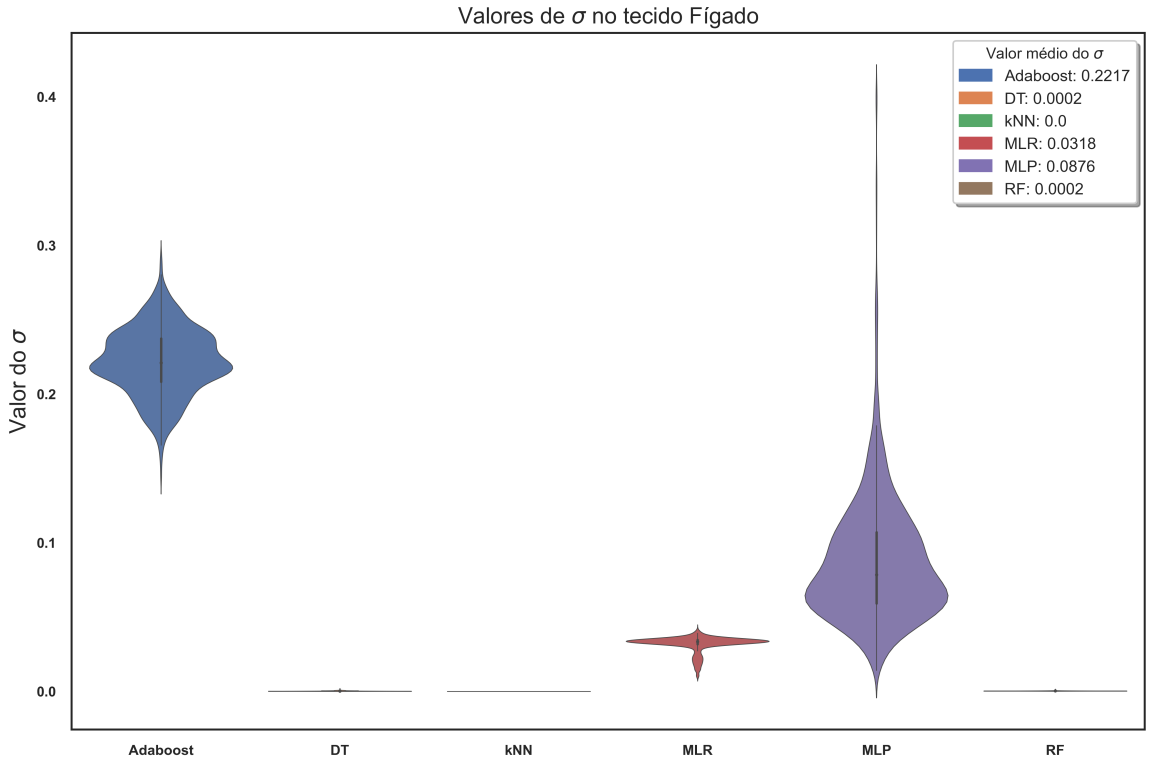
Fonte: Autor (2024).

No teste de hipótese, considerando o valor médio de cada *label*, a Regressão Linear Múltipla obteve um p-valor de 0,0, o *AdaBoost* de 0,0235, a *MLP* de 0,0 e o *kNN* de 0,001. Isso significa que para estes algoritmos o valor médio predito de intensidade I para o comprimento de onda λ não é o real. Por outro lado, a *DT* e o *RF* resultaram em um p-valor de 0,6866 e 0,3156, respectivamente, ou seja, o valor médio predito por eles é o valor real.

A reconstrução de intensidade das curvas para este tecido nos seis algoritmos pode ser encontrada nas Figuras (A.5) e (A.6).

A variação mais acentuada do σ foi no *MLP*, seguido pelo *AdaBoost* e *MLR*, que foi consideravelmente menor que os dois primeiros. Os melhores resultados foram dos *kNN*, seguidos pela *DT* e *RF*, cujo valor do σ foi igual. Maiores detalhes podem ser encontrados na

Figura 6.5: Gráfico de violino do σ obtido a partir da correção espectral usando os seis modelos de AM na **fígado**



Fonte: Autor (2024).

Figura (6.5).

O gráfico de violino da temperatura (Figura 6.6) mostrou uma alta variação na temperatura na *MLP*, *AdaBoost* e *MLR*. Os melhores resultados da temperatura usando as curvas corrigidas foram dos algoritmos: *DT*, *RF* e *kNN*, que foram exatos e precisos.

Figura 6.6: Gráfico de violino da temperatura obtida pela correção espectral usando os algoritmos de AM no **fígado**

Fonte: Autor (2024).

6.1.4 Pele

Para a Pele, o *MLR* foi o algoritmo mais rápido no tempo de treinamento, gastando apenas 40 segundos e com valores médios de *MAE*, *MSE*, *MAPE*, R^2 e σ de 0,0622, 0,1358, 0,0431, 0,9982, 0,0312, respectivamente. Em contrapartida o tempo de treinamento do *RF*

foi de aproximadamente 2 horas e 53 minutos com valores médios de *MAE*, *MSE*, *MAPE*, R² e σ de 0,0002, 0,0001, 0,0002, 1,0, 0,0002, respectivamente. Resultados semelhantes nas métricas foram obtidos nos algoritmos de *DT* e *kNN*, com tempo de treinamento menor, sendo de 2 minutos e 40 segundos para a Árvore de Decisão e de 55 segundos para o *kNN*. Os piores resultados nas métricas de desempenho foram do *AdaBoost*, com valores médios de *MAE*, *MSE*, *MAPE*, R² e σ de 0,3165, 0,1325, 6,7287, 0,9925, 0,2141, respectivamente; e da *MLP*, com os seguintes valores médios de *MAE*, *MSE*, *MAPE*, R² e σ: 0,1462, 0,0471, 1,3736, 0,9974, 0,0824. O *AdaBoost* resultou em um tempo de treinamento de aproximadamente 23 minutos, enquanto a *MLP* gastou cerca 13 minutos. Os valores com os respectivos desvios padrões estão resumidos na Tabela (6.4).

O desempenho inferior da *MLR* nas métricas pode ser atribuído ao fato de ser um modelo linear aplicado a um problema não linear. O desempenho aquém do esperado do *AdaBoost*, se comparado à *DT*, ao *RF* e ao *kNN*, pode ser justificado pelo número limitado de estimadores, que no nosso caso foi de 50, além do uso de uma função de perda linear. Em contrapartida, o *RF* utilizou 105 estimadores e adotou o *MSE* como critério de seleção. Quanto à *MLP*, seu desempenho insatisfatório pode ser explicado pelo grande número de camadas ocultas, o que frequentemente resulta no otimizador se prendendo em um mínimo local. Além disso, a função de ativação ReLU, por ser linear, contribui para a manutenção da linearidade no modelo conexionista, o que ocorre porque a aplicação de uma transformação linear, como a ReLU, preserva a linearidade durante a busca pela solução.

Tabela 6.4: Principais dados de avaliação de correção espectral na Pele

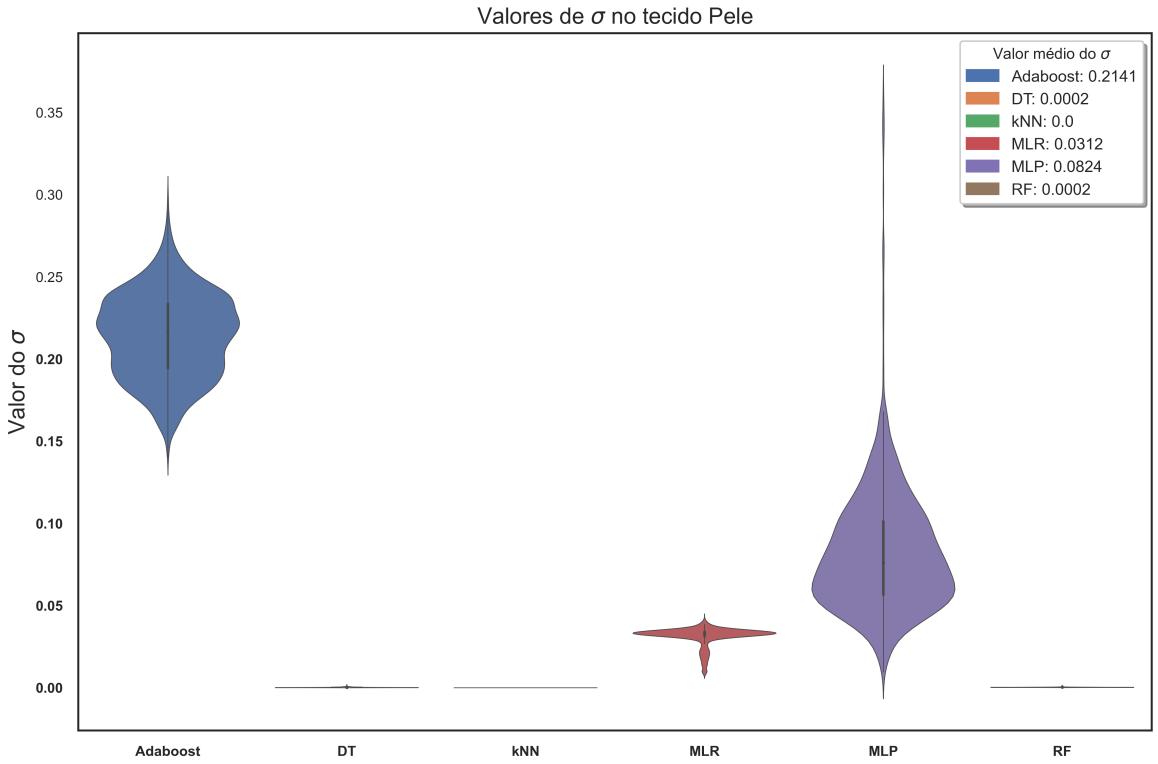
Algoritmo	**Tempo**	***MAE* médio e *std***	***MSE* médio e *std***	***MAPE* médio e *std***	**R² médio e *std***	**σ médio e *std***
MLR	**40s**	0,0622 +/- 0,0015	0,1316 +/- 0,0015	0,0431 +/- 0,0026	0,9982 +/- 0,0001	0,0312 +/- 0,0058
DT	2min40s	0,0002 +/- 0,0001	**0,0 +/- 0,0**	0,0002 +/- 0,0	**1,0 +/- 0,0**	0,0002 +/- 0,0002
RF	2h53mi24s	0,0003 +/- 0,0	**0,0 +/- 0,0**	0,0002 +/- 0,0	**1,0 +/- 0,0**	0,00002 +/- 0,0002
AdaBoost	26min13s	0,3165 +/- 0,0402	0,1325 +/- 0,0298	6,7287 +/- 1,0246	0,9925 +/- 0,0017	0,2141 +/- 0,0258
kNN	55s	**0,0 +/- 0,0**	**0,0 +/- 0,0**	**0,0 +/- 0,0**	**1,0 +/- 0,0**	**0,0 +/- 0,0**
MLP	13min17s	0,1462 +/- 0,0797	0,0471 +/- 0,0026	1,3736 +/- 1,1222	0,9973 +/- 0,0026	0,0824 +/- 0,0385

Fonte: Autor (2024).

Em relação ao de hipótese, considerando o valor médio de cada *label*, a Regressão Linear Múltipla obteve um p-valor de 0,0008, o *AdaBoost* de 0,0235, a *MLP* de 0,0 e o *kNN* de 0,001. Isso significa que nestes algoritmos o valor médio predito de intensidade I para o comprimento de onda λ não é o real. Por outro lado, a *DT* e o *RF* resultaram em um p-valor de 0,6432 e 0,9060, respectivamente, ou seja, o valor médio predito por eles é o valor real.

A reconstrução de intensidade das curvas para este tecido nos seis algoritmos pode ser encontrada nas Figuras (A.7) e (A.8).

Figura 6.7: Gráfico de violino do σ obtido a partir da correção espectral usando os seis modelos de AM na **pele**



Fonte: Autor (2024).

A variação mais acentuada do σ foi no *MLP*, seguido pelo *AdaBoost* e *MLR*, que foi consideravelmente menor que os dois primeiros. Os melhores resultados foram dos *kNN*, seguidos pela *DT* e *RF*, cujo valor do σ foi igual. Maiores detalhes podem ser encontrados na Figura (6.7).

O gráfico de violino da temperatura (Figura 6.8) mostrou uma variação considerável na temperatura usando as curvas corrigidas pela *MLP* e pelo *AdaBoost*. Já a *MLR* resultou em valores exatos, mas imprecisos. Os melhores resultados da temperatura usando as curvas corrigidas foram dos algoritmos: *DT*, *RF* e *kNN*, que foram exatos e precisos.

Figura 6.8: Gráfico de violino da temperatura obtida pela correção espectral usando os algoritmos de AM na **pele**

Fonte: Autor (2024).

6.1.5 Pele variando a concentração de água

No tecido Pele variando a concentração de água, o algoritmo mais rápido no tempo de treinamento a *MLR* foi, gastando 18 segundos e com valores médios de *MAE*, *MSE*, *MAPE*, R^2 e σ de 0,0055, 0,0003, 0,0051, 1,0, 0,0029, respectivamente. Por outro lado o tempo de treinamento do *RF* foi de aproximadamente 19 minutos e 32 segundos com valores médios de *MAE*, *MSE*, *MAPE*, R^2 e σ de 0,0001, 0,0, 0,0, 1,0, 0,0001, respectivamente. Resultados semelhantes nas métricas foram obtidos nos algoritmos de *DT* e *kNN*, com tempo de treinamento menor, sendo de 37 segundos para a Árvore de Decisão e de 22 segundos para o *kNN*. Os piores resultados nas métricas de desempenho foram do *AdaBoost*, com valores médios de *MAE*, *MSE*, *MAPE*, R^2 e σ de 0,2241, 0,0827, 4,555, 0,9953, 0,1467, respectivamente; e da *MLP*, com os seguintes valores médios de *MAE*, *MSE*, *MAPE*, R^2 e σ: 0,0411, 0,0027, 0,5110, 1,0, 0,0269. O *AdaBoost* resultou em um tempo de treinamento de aproximadamente 3 minutos, enquanto a *MLP* demorou cerca 6 minutos. Os valores com os respectivos desvios padrões estão resumidos na Tabela (6.5).

O desempenho insatisfatório da *MLR* nas métricas pode ser explicado pelo fato de ser um modelo linear aplicado a um problema de natureza não linear. O rendimento abaixo do esperado do *AdaBoost*, em comparação com a *DT*, *RF* e *kNN*, pode ser atribuído ao número reduzido de estimadores, que em nosso caso foi de 50, além da utilização de uma função de perda linear. Por outro lado, a *RF* utilizou 105 estimadores e adotou o *MSE* como critério de seleção. Em relação à *MLP*, seu desempenho inferior pode ser justificado pelo elevado número de camadas ocultas, o que frequentemente faz com que o otimizador se prenda em um mínimo local. Além disso, a função de ativação ReLU, por ser linear, tende a manter a linearidade no modelo conexionista, o que se dá porque a aplicação de uma transformação linear, como a ReLU, preserva essa característica durante a busca pela solução.

Tabela 6.5: Principais dados de avaliação de correção espectral na Pele variando a concentração de água

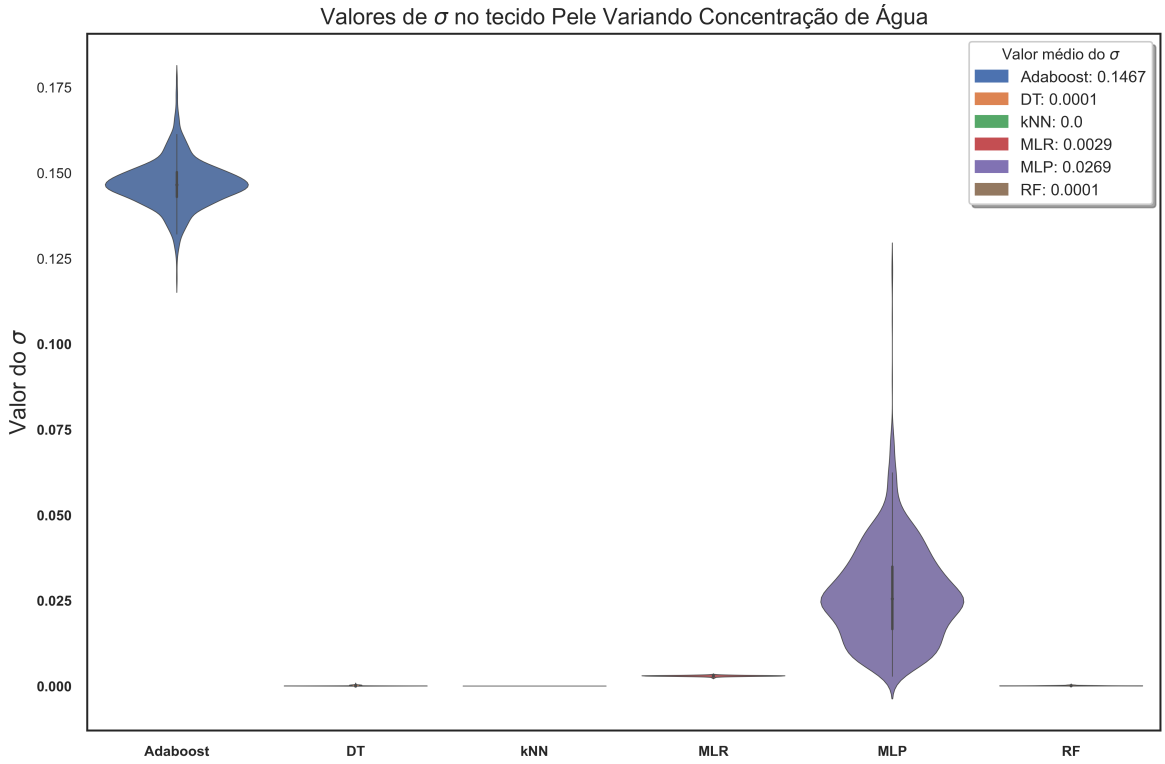
Algoritmo	**Tempo**	***MAE* médio e *std***	***MSE* médio e *std***	***MAPE* médio e *std***	**R^2 médio e *std***	**σ médio e *std***
MLR	**18s**	0,0055 +/- 0,0002	0,0003 +/- 0,0	0,0051 +/- 0,0004	**1,0 +/- 0,0**	0,0029 +/- 0,0002
DT	37s	0,0001 +/- 0,0001	**0,0 +/- 0,0**	0,0001 +/- 0,0	**1,0 +/- 0,0**	0,0001 +/- 0,0001
RF	19min32s	0,0001 +/- 0,0	**0,0 +/- 0,0**	0,0001 +/- 0,0	**1,0 +/- 0,0**	0,0001 +/- 0,0001
AdaBoost	3min16s	0,2441 +/- 0,0135	0,0827 +/- 0,0090	4,5550 +/- 0,4254	0,9953 +/- 0,0005	0,1467 +/- 0,0069
kNN	22s	**0,0 +/- 0,0**	**0,0 +/- 0,0**	**0,0 +/- 0,0**	**1,0 +/- 0,0**	**0,0 +/- 0,0**
MLP	6min18s	0,0411 +/- 0,0269	0,0027 +/- 0,0032	0,5110 +/- 0,3679	0,9998 +/- 0,0002	0,0269 +/- 0,0140

Fonte: Autor (2024).

Para o teste de hipótese considerando o valor médio de cada *label*, a Regressão Linear Múltipla obteve um p-valor de 0,0, o *AdaBoost* de 0,0, a *MLP* de 0,0 e o *kNN* de 0,043. Logo nestes algoritmos o valor médio predito de intensidade I para o comprimento de onda λ não é o real. Por outro lado, a *DT* e o *RF* resultaram em um p-valor de 0,8652 e 0,5810, respectivamente, ou seja, o valor médio predito por eles é o valor real.

A reconstrução de intensidade das curvas para este tecido nos seis algoritmos pode ser encontrada nas Figuras (A.9) e (**??**).

Figura 6.9: Gráfico de violino do σ obtido a partir da correção espectral usando os seis modelos de AM na **pele** variando a concentração de água



Fonte: Autor (2024).

A variação mais acentuada do σ foi no *MLP*, seguido pelo *AdaBoost*. O algoritmo de *MLR* obteve acentuada precisão no sigma. Os melhores resultados foram dos *kNN*, seguidos pela *DT* e *RF*, cujo valor do σ foi igual. Maiores detalhes podem ser encontrados na Figura (6.9).

O gráfico de violino da temperatura (Figura 6.10) mostrou uma variação considerável na temperatura usando as curvas corrigidas pela *MLP*, *AdaBoost* e *MLR*. Já a *MLR* resultou em valores exatos, mas imprecisos. Os melhores resultados da temperatura usando as curvas corrigidas foram dos algoritmos: *DT*, *RF* e *kNN*, sendo exatos e precisos.

Figura 6.10: Variação da temperatura usando os espectros corrigidos pelos algoritmos de AM na **pele** variando a concentração de água

Fonte: Autor (2024).

6.1.6 Água

Na Água o algoritmo foi mais rápido no tempo de treinamento foi a *MLR*, com 27 segundos e valores médios de *MAE*, *MSE*, *MAPE*, R² e σ de 0,0316, 0,0025, 0,3015, 1,0, 0,0167, respectivamente. Por outro lado o tempo de treinamento do *RF* foi de aproximadamente 52 minutos e 55 segundos com valores médios de *MAE*, *MSE*, *MAPE*, R² e σ de 0,0002, 0,0, 0,0004, 1,0, 0,0002, respectivamente. Resultados parecidos nas métricas foram obtidos nos algoritmos de *DT* e *kNN*, com tempo de treinamento menor, sendo de 1 minuto e 12 segundos para a Árvore de Decisão e de 38 segundos para o *kNN*. Os piores resultados nas métricas de desempenho foram do *AdaBoost*, com valores médios de *MAE*, *MSE*, *MAPE*, R² e σ de 0,2507, 0,0834, 4,8292, 0,9953, 0,1550, respectivamente; e da *MLP* com os seguintes valores médios de *MAE*, *MSE*, *MAPE*, R² e σ: 0,0734, 0,0140, 0,9042, 0,9992, 0,0535. O *AdaBoost* resultou em um tempo de treinamento de aproximadamente 8 minutos, enquanto a *MLP* demorou cerca 9 minutos. Os valores com seus respectivos desvios padrões estão resumidos na Tabela (6.6).

O desempenho inferior da *MLR* nas métricas, se comparado com os modelos de AM com melhores resultados, pode ser atribuído ao fato de que é um modelo linear sendo aplicado a um problema que não é linear. O desempenho insatisfatório do *AdaBoost*, quando comparado à *DT*, *RF* e *kNN*, pode ser explicado pelo baixo número de estimadores, que em nosso caso foi de 50, além do uso de uma função de perda linear. Em contrapartida, a *RF* utilizou 105 estimadores e adotou o *MSE* como critério de seleção. No caso da *MLP*, seu desempenho abaixo do esperado pode ser explicado pelo elevado número de camadas ocultas, o que frequentemente faz com que o otimizador fique preso em um mínimo local. Além disso, a função de ativação ReLU, por ser linear, tende a manter a linearidade no modelo conexionista, o que ocorre porque a aplicação de uma transformação linear, como a ReLU, preserva essa linearidade durante a busca por uma solução.

Para o teste de hipótese considerando o valor médio de cada *label*, a Regressão Linear Múltipla obteve um p-valor de 0,0001, o *AdaBoost* de 0,0, a *MLP* de 0,0, o *kNN* de 0,001, a *DT* de 0,035 e o *RF* de 0,014. Portanto o valor médio predito de intensidade I para o comprimento de onda λ não é o real. A reconstrução de intensidade das curvas para este tecido nos seis algoritmos pode ser encontrada nas Figuras (A.11) e (A.12).

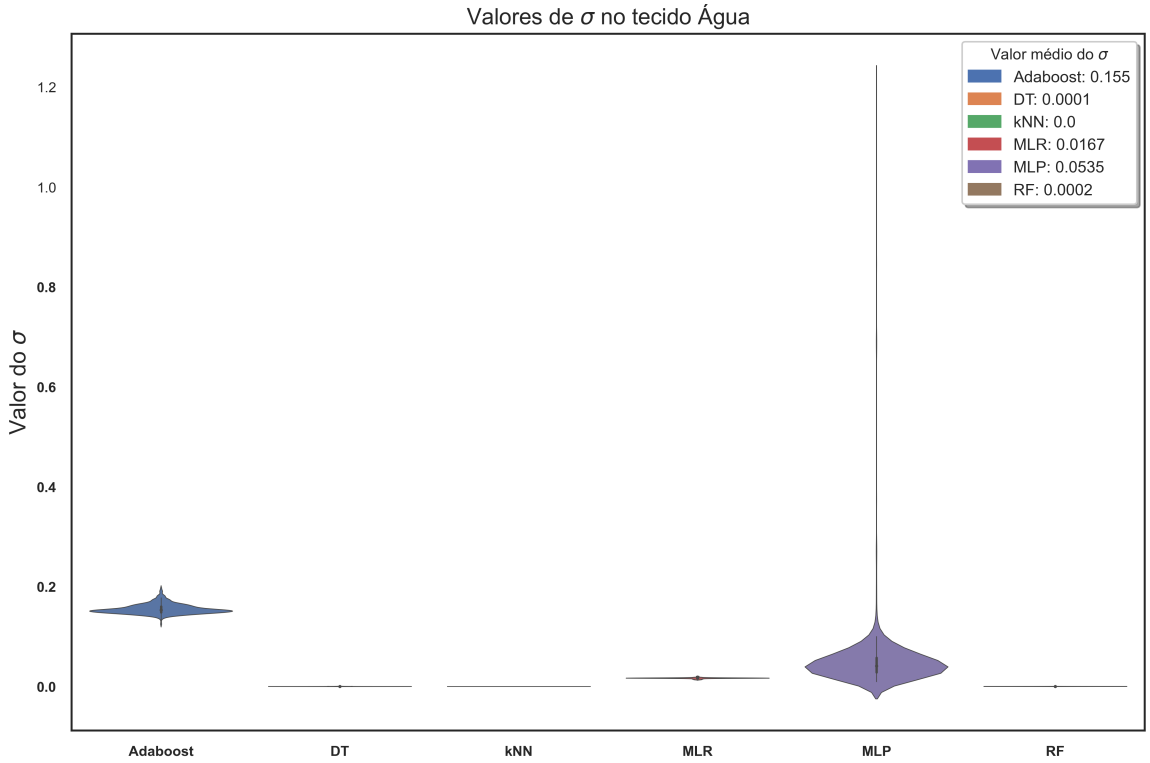
A variação mais acentuada do σ foi no *MLP*, seguido pelo *AdaBoost*. O algoritmo de *MLR* obteve acentuada precisão no sigma. Os melhores resultados foram dos *kNN*, seguidos pela *DT* e *RF*, cujos valores do σ foram praticamente iguais. Maiores detalhes podem ser

Tabela 6.6: Principais dados de avaliação de correção espectral na Água

Algoritmo	**Tempo**	***MAE* médio e *std***	***MSE* médio e *std***	***MAPE* médio e *std***	**R^2 médio e *std***	**σ médio e *std***
MLR	**27s**	0,0316 +/- 0,0004	0,0025 +/- 0,0001	0,3015 +/- 0,0178	0,9999 +/- 0,0	0,0167 +/- 0,0010
DT	1min12s	0,0002 +/- 0,0001	**0,0 +/- 0,0**	0,0002 +/- 0,0	**1,0 +/- 0,0**	0,0001 +/- 0,0002
RF	52min55s	0,0002 +/- 0,0	**0,0 +/- 0,0**	0,0004 +/- 0,0	**1,0 +/- 0,0**	0,0002 +/- 0,0001
AdaBoost	7min50s	0,2507 +/- 0,0159	0,0834 +/- 0,0090	4,8292 +/- 0,4930	0,9953 +/- 0,0005	0,1550 +/- 0,0098
kNN	38s	**0,0 +/- 0,0**	**0,0 +/- 0,0**	**0,0 +/- 0,0**	**1,0 +/- 0,0**	**0,0 +/- 0,0**
MLP	9min12s	0,0734 +/- 0,0806	0,0140 +/- 0,0629	0,9042 +/- 1,1032	0,9992 +/- 0,0035	0,0535 +/- 0,0807

Fonte: Autor (2024).

Figura 6.11: Gráfico de violino do σ obtido a partir da correção espectral usando os seis modelos de AM na **água**



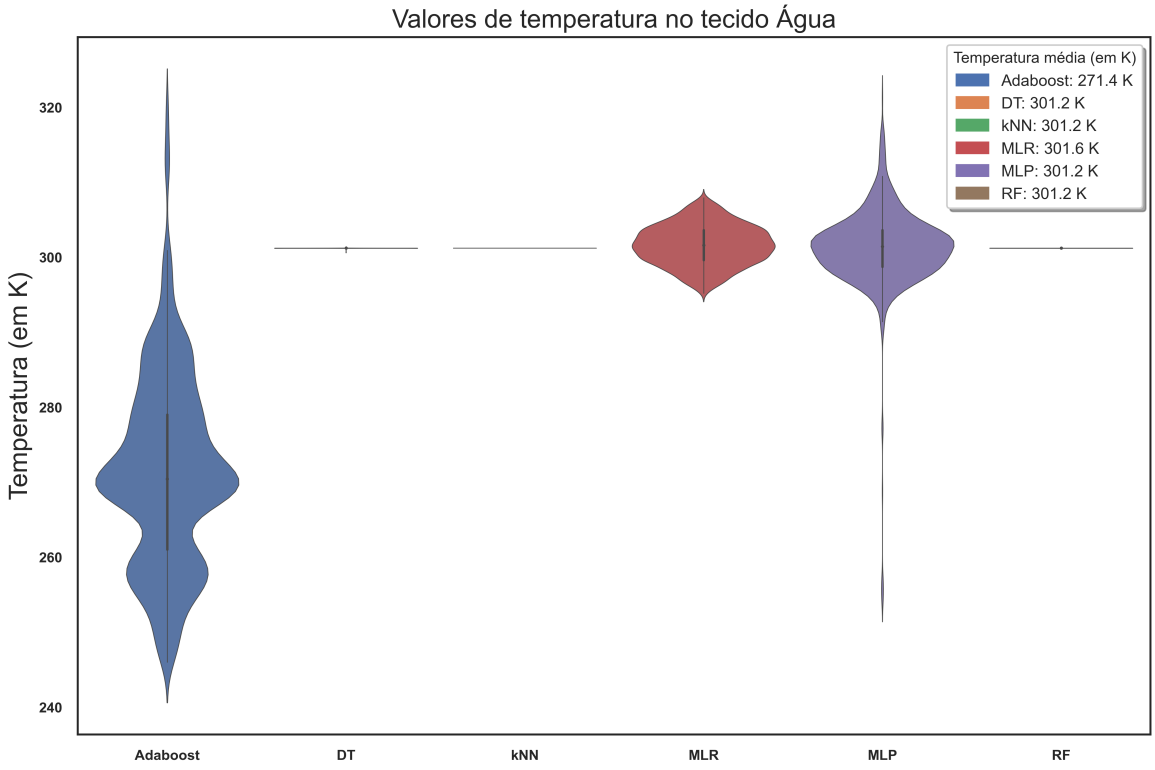
Fonte: Autor (2024).

encontrados na Figura (6.11).

O gráfico de violino da temperatura (Figura 6.12) mostrou uma variação considerável na temperatura usando as curvas corrigidas pela *MLP* e *MLR*, eqnauto o *AdaBoost* resultou

numa maior precisão, apesar de inexato. Os melhores resultados da temperatura usando as curvas corrigidas foram dos algoritmos: *DT*, *RF* e *kNN*, sendo exatos e precisos.

Figura 6.12: Gráfico de violino da temperatura obtida pela correção espectral usando os algoritmos de AM na **água**



Fonte: Autor (2024).

6.2 Resultados e Discussão nos tecidos com distorções usando μ_s e profundidade

6.2.1 Massa Cinzenta do Cérebro

Os algoritmos treinados com Massa Cinzenta do Cérebro resultou em métricas próximas dos melhores valores possíveis considerando o contexto de problemas de regressão. A *MLR* realizou o treinamento mais rápido, com cerca de 30 segundos e com valores médios de *MAE*, *MSE*, *MAPE* e R² de 0,5494, 1,6887, 4,5803 e 0,2818, respectivamente. O σ médio resultou em um valor de 0,1333. Em contraste com relação ao tempo de treinamento, o algoritmo que mais demorou foi o *RF*, com um tempo de 58 minutos. Os valores médios de *MAE*, *MSE*, *MAPE* e R² foram de 0,0002, 0,0, 0,0006 e 1,0, respectivamente, enquanto que o valor médio do σ foi de 0,0001. Resultados próximos nas métricas foram obtidos nos algoritmos de *DT* e

kNN, com tempo de treinamento consideravelmente menor, de 1 minuto e 14 segundos para a Árvore de Decisão e de apenas 53 segundos para o *kNN*. Os piores resultados nas métricas de desempenho foram do *AdaBoost*, com valores médios de *MAE*, *MSE*, *MAPE*, R^2 e σ de 0,5707, 0,48, 8,7674, 0,9728, 0,3170, respectivamente; e da *MLP*, com os seguintes valores médios de *MAE*, *MSE*, *MAPE*, R^2 e σ: 0,2688, 0,0923, 5,8684, 0,9948, 0,3659. A *MLP* foi cerca de duas vezes mais rápida que o *AdaBoost* no tempo de treinamento, sendo de cerca de 10 minutos e 22 minutos e 50 segundos, respectivamentes. Os valores médios com os respectivos desvios padrões estão resumidos na Tabela (6.7).

O desempenho insatisfatório do *AdaBoost*, em comparação com *DT*, *RF* e *kNN*, pode ser atribuído ao pequeno número de estimadores, que no nosso caso foi de 50, além da adoção de uma função de perda linear. Por outro lado, a *RF* utilizou 105 estimadores e adotou o *MSE* como critério de seleção. Quanto à *MLP*, seu desempenho inferior pode ser explicado pelo grande número de camadas ocultas, o que muitas vezes resulta no otimizador se prendendo em um mínimo local. Além disso, a função de ativação ReLU, sendo linear, tende a manter a linearidade no modelo conexionista, o que ocorre porque aplicar uma transformação linear, como a ReLU, preserva essa característica durante a busca por uma solução.

Tabela 6.7: Principais dados de avaliação de correção espectral na Massa Cinzenta do Cérebro

Algoritmo	**Tempo**	***MAE* médio e *std***	***MSE* médio e *std***	***MAPE* médio e *std***	**R^2 médio e *std***	**σ médio e *std***
MLR	**30s**	0,5494 +/- 0,0105	1,6867 +/- 0,0742	4,5803 +/- 0,3402	0,9043 +/- 0,0029	0,2818 +/- 0,0363
DT	1min14s	**0,0001 +/- 0,0**	**0,0 +/- 0,0**	0,0002 +/- 0,0001	**1,0 +/- 0,0**	**0,0001 +/- 0,0001**
RF	58min	0,0002 +/- 0,0	**0,0 +/- 0,0**	0,0006 +/- 0,0001	**1,0 +/- 0,0**	**0,0001 +/- 0,0001**
AdaBoost	18min30s	0,5707 +/- 0,0655	0,48 +/- 0,0582	8.7674 +/- 1,5548	0,9728 +/- 0,0034	0,3170 +/- 0,0321
kNN	53s	0,0004 +/- 0,0001	**0,0 +/- 0,0**	**0,0 +/- 0,0**	**1,0 +/- 0,0**	0,0003 +/- 0,0002
MLP	16min48s	0,6177 +/- 0,1794	1,1987 +/- 0,2120	3,9043 +/- 2,8931	0,9320 +/- 0,0121	0,3659 +/- 0,1626

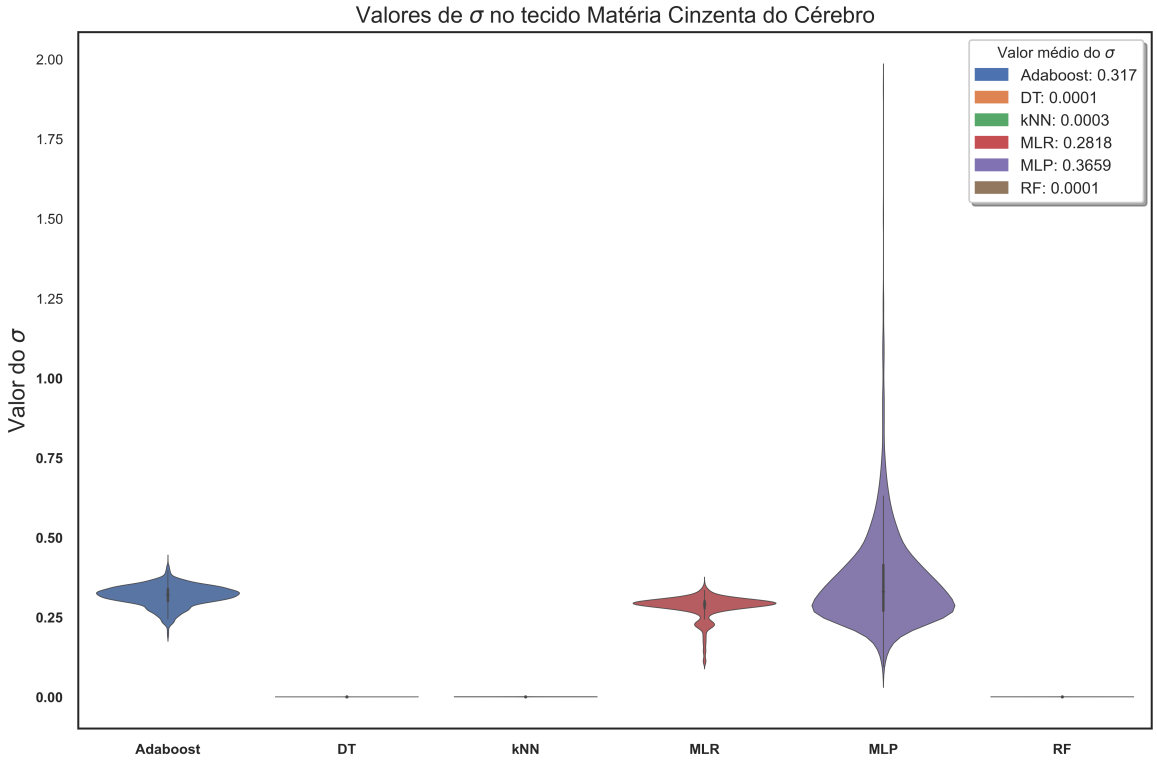
Fonte: Autor (2024).

No teste de hipótese, considerando o valor médio de cada *label*, obteve um p-valor de 0,0 para a Regressão Linear Múltipla, 0,0 para o *AdaBoost*, *MLP* de 0,0 e de 0,0001 para o *kNN* e 0,0001 no *RF*. Isso significa que para estes algoritmos o valor médio predito de intensidade I para o comprimento de onda λ não é o real. Já a *DT* resultou no p-valor de 0,2393, ou seja, o valor médio predito nesse algortimo é o valor real.

A reconstrução de intensidade das curvas para este tecido nos seis algoritmos pode ser vista nas Figuras (A.13) e (A.14)

A variação mais acentuada do σ foi na *MLP*, seguido pelo *AdaBoost* e *MLR*. Os

Figura 6.13: Gráfico de violino do σ obtido a partir da correção espectral usando os seis modelos de AM na **massa cinzenta do cérebro**



Fonte: Autor (2024).

melhores resultados foram dos *kNN*, ,*DT* e *RF*, cujos valores do σ foram próximos uns dos outros. Maiores detalhes podem ser encontrados na Figura (6.13).

O gráfico de violino da temperatura (Figura 6.2) mostrou uma alta variação dessa grandeza no *MLR*, *MLP* e *AdaBoost*. Já a temperaura obtida usando as curvas corrigidas pela *DT*, pelo *RF* e através do *kNN* foi altamente precisa e exata.

Figura 6.14: Gráfico de violino da temperatura obtida pela correção espectral usando os algoritmos de AM na **massa cinzenta do cérebro**

Fonte: Autor (2024).

6.2.2 Mama Pré-Menopausa

Os regressores aplicados na Mama Pré-Menopausa resultaram em métricas bem próximas dos melhores valores possíveis, considerando outros problemas de regressão. A *MLR* foi

menos custoso no tempo de treinamento, gastando apenas 27 segundos e com valores médios de*MAE*, *MSE*, *MAPE*, R² e σ de 0,6650, 2,3926, 5,8287, 0,8643, 0,3426, respectivamente. Por outro lado tempo de treinamento do *RF* foi de 58 minutos e 31 segundos com valores médios de *MAE*, *MSE*, *MAPE*, R² e σ de 0,0002, 0,0, 0,0006, 1,0 e 0,0001, respectivamente. Resultados similares nas métricas foram obtidos nos algoritmos de *DT* e *kNN*, com tempo de treinamento menor, sendo de 1 minuto e 14 segundos para a Árvore de Decisão e de 46 segundos para o *kNN*. Os piores resultados nas métricas de desempenho foram do *AdaBoost*, com valores médios de *MAE*, *MSE*, *MAPE*, R² e σ de 0,5297, 0,4566, 7,6711, 0,9741, 0,2920, respectivamente; e da *MLP*, com os seguintes valores médios de *MAE*, *MSE*, *MAPE*, R² e σ: 0,7415, 1,6728, 4,8283, 0,9051, 0,4386. O *AdaBoost* demorou mais que a *MLP* no tempo de treinamento. Os valores com os respectivos desvios padrões estão resumidos na Tabela (6.8).

O menor desempenho do *AdaBoost* em relação ao *DT*, *RF* e *kNN*, pode ser explicado pelo número reduzido de estimadores, que foi de 50 no nosso caso, além da escolha de uma função de perda linear. Em contraste, a *RF* utilizou 105 estimadores e usou o *MSE* como critério de seleção. No caso da *MLP*, seu desempenho insatisfatório pode ser atribuído ao número elevado de camadas ocultas, o que frequentemente leva o otimizador a ficar preso em mínimos locais. Além disso, a função de ativação ReLU, por ser linear, ajuda a manter a linearidade no modelo conexionista, pois a aplicação de uma transformação linear como a ReLU preserva essa linearidade durante a busca por uma solução.

Tabela 6.8: Principais dados de avaliação de correção espectral na Mama Pré-Menopausa

Algoritmo	**Tempo**	***MAE* médio e *std***	***MSE* médio e *std***	***MAPE* médio e *std***	**R² médio e *std***	**σ médio e *std***
MLR	**27s**	0,6650 +/- 0,0124	2,3926 +/- 0,1048	5,8287 +/- 0,4247	0,8643 +/- 0,0041	0,3426 +/- 0,0418
DT	1min14s	**0,0001 +/- 0,0**	**0,0 +/- 0,0**	0,0003 +/- 0,0002	**1,0 +/- 0,0**	**0,0001 +/- 0,0001**
RF	58min31s	0,0002 +/- 0,0	**0,0 +/- 0,0**	0,0006 +/- 0,0001	**1,0 +/- 0,0**	**0,0001 +/- 0,0001**
AdaBoost	18min7s	0,5297 +/- 0,0822	0,4566 +/- 0,0664	7.6711 +/- 2,0317	0,9741 +/- 0,0038	0,2920 +/- 0,0428
kNN	46s	0,0007 +/- 0,0001	0,0001 +/- 0,0	**0,0 +/- 0,0**	**1,0 +/- 0,0**	0,0004 +/- 0,0003
MLP	16min4s	0,7415 +/- 0,2081	1,6727 +/- 0,2932	4,8283 +/- 3,3669	0,9051 +/- 0,0166	0,4386 +/- 0,5901

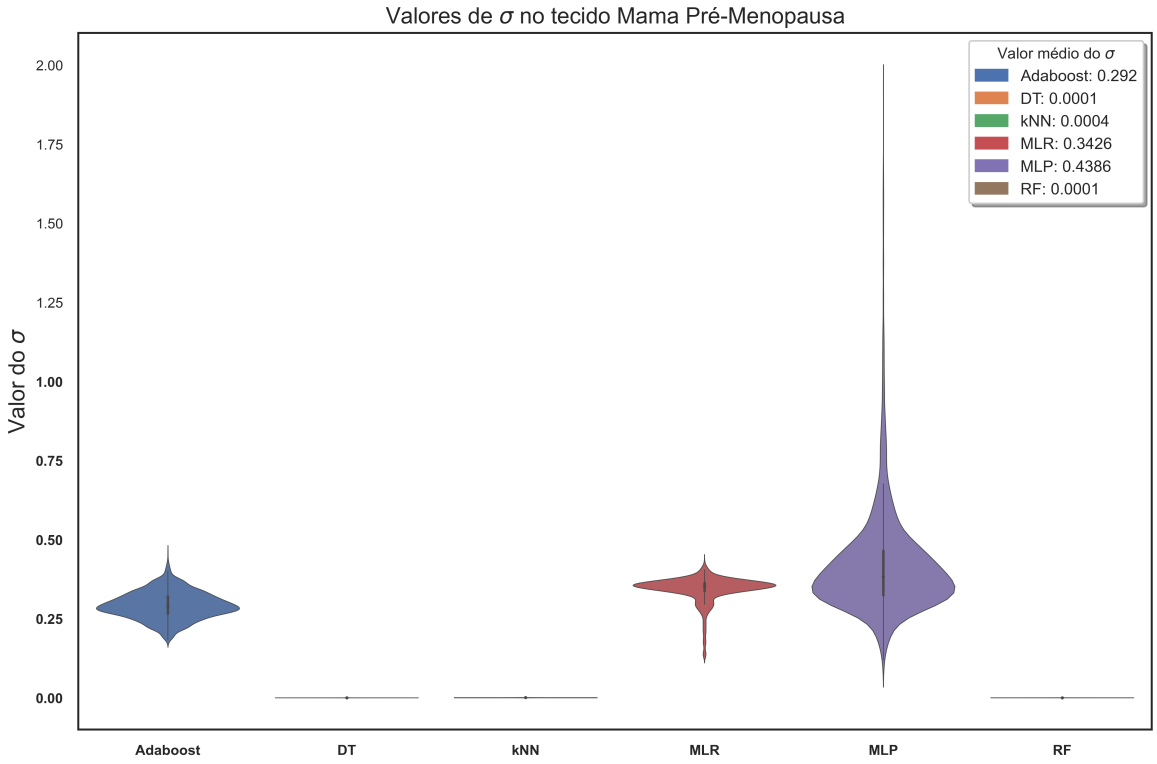
Fonte: Autor (2024).

No teste de hipótese, considerando o valor médio de cada *label*, obteve um p-valor de 0,0 para a Regressão Linear Múltipla, 0,0 para o *AdaBoost*, de 0,001 para o *kNN* e de 0,0001 no *RF*. Isso significa que para estes algoritmos o valor médio predito de intensidade I para o comprimento de onda λ não segue a mesma distribuição. Os demais algoritmos, *MLP*, *DT* e *RF* resultaram no p-valor de de 0,6908, 0,7872, respectivamente, ou seja, o valor médio predito por

eles não segue a mesma distribuição.

A reconstrução de intensidade das curvas para este tecido nos seis algoritmos pode ser encontrada nas Figuras (A.15) e (A.16)

Figura 6.15: Gráfico de violino do σ obtido a partir da correção espectral usando os seis modelos de AM na **mama pré-menopausa**



Fonte: Autor (2024).

A variação mais acentuada do σ foi na *MLP*, seguido pelo *AdaBoost*. É possível notar uma baixa variância nos resultados, tendo boa exatidão e precisão. Contudo os melhores resultados foram alcançados pela *MLR*, *kNN*, *DT* e *RF*, cujo valor do σ foi próximo de 0. Maiores detalhes podem ser encontrados na Figura (6.15).

O gráfico de violino da temperatura (Figura 6.16) mostrou alta variabilidade nas temperaturas obtidas pelos espectros corrigidos usando *MLR*, *AdaBoost* e *MLP*. Por outro lado, a temperatura obtida usando os valores preditos pela *DT*, *RF* e *kNN*, foram exatos e precisos.

Figura 6.16: Gráfico de violino da temperatura obtida pela correção espectral usando os algoritmos de AM na **mama pré-menopausa**

Fonte: Autor (2024).

6.2.3 Fígado

No Fígado, o algoritmo que mais se destacou em tempo de treinamento foi a*MLR* gastando apenas 40 segundos. Suas métricas de desempenho foram de 0,8373, 3,5677, 8,0134, 0,7976, 0,4348, que correspondem aos valores médios de *MAE*, *MSE*, *MAPE*, R² e σ, respectivamente. Já o tempo de treinamento do *RF* foi de aproximadamente 1 hora com valores médios de *MAE*, *MSE*, *MAPE*, R² e σ de 0,0001, 0,0, 0,0003, 1,0, 0,0002, respectivamente. Resultados semelhantes nas métricas foram obtidos nos algoritmos de *DT* e *kNN*, com tempo de treinamento menor, sendo de 1 minuto e 15 segundos para a Árvore de Decisão e de 53 segundos para o *kNN*. Os outros algoritmos que não resultaram em um desempenho satisfatório nas métricas de desempenho foram do *AdaBoost*, com valores médios de *MAE*, *MSE*, *MAPE*, R² e σ de 0,5132, 0,4063, 7,4934, 0,9769, 0,2870, respectivamente; e da *MLP*, com os seguntes valores médios de *MAE*, *MSE*, *MAPE*, R² e σ: 0,9151, 2,5121, 5,9349, 0,8574, 0,5590. O *AdaBoost* resultou em um tempo de treinamento de 22 minutos, enquanto a *MLP* obteve um custo no tempo de quase 13 minutos. Os valores estão resumidos na Tabela (6.9).

O desempenho inferior do *AdaBoost* comparado ao *DT*, *RF* e *kNN* pode ser atribuído ao número reduzido de estimadores, que foi de 50 no nosso experimento, e à escolha de uma função de perda linear. Em contraste, o *RF* fez uso de 105 estimadores e adotou o *MSE* como critério de seleção. No que diz respeito à *MLP*, o desempenho aquém do esperado pode ser causado pelo grande número de camadas ocultas, o que frequentemente resulta em o otimizador se prender a mínimos locais. Além disso, a função de ativação ReLU, sendo linear, contribui para a manutenção da linearidade no modelo conexionista, já que a aplicação de uma transformação linear como a ReLU preserva essa característica durante a busca por uma solução.

Tabela 6.9: Principais dados de avaliação de correção espectral no Fígado

Algoritmo	**Tempo**	***MAE* médio e *std***	***MSE* médio e *std***	***MAPE* médio e *std***	**R² médio e *std***	**σ médio e *std***
MLR	**26s**	0,8373 +/- 0,0148	3,5677 +/- 0,1533	8,0184 +/- 0,5742	0,7976 +/- 0,0059	0,4348 +/- 0,0489
DT	1min15s	**0,0001 +/- 0,0001**	**0,0 +/- 0,0**	0,0003 +/- 0,0002	**1,0 +/- 0,0**	**0,0001 +/- 0,0001**
RF	1h0min4s	0,0002 +/- 0,0	**0,0 +/- 0,0**	0,0007 +/- 0,0001	**1,0 +/- 0,0**	0,0002 +/- 0,0001
AdaBoost	22min	0,5132 +/- 0,0538	0,4063 +/- 0,0498	7,4934 1,2719	0,9769 +/- 0,0028	0,2870 +/- 0,0291
kNN	53s	0,0009 +/- 0,0	0,0001 +/- 0,0	**0,0001 +/- 0,0**	**1,0 +/- 0,0**	0,0006 +/- 0,0003
MLP	13min10s	0,9151 +/- 0,2372	2,5121 +/- 0,3696	5,9349 +/- 3,8687	0,8574 +/- 0,0208	0,5590 +/- 0,2682

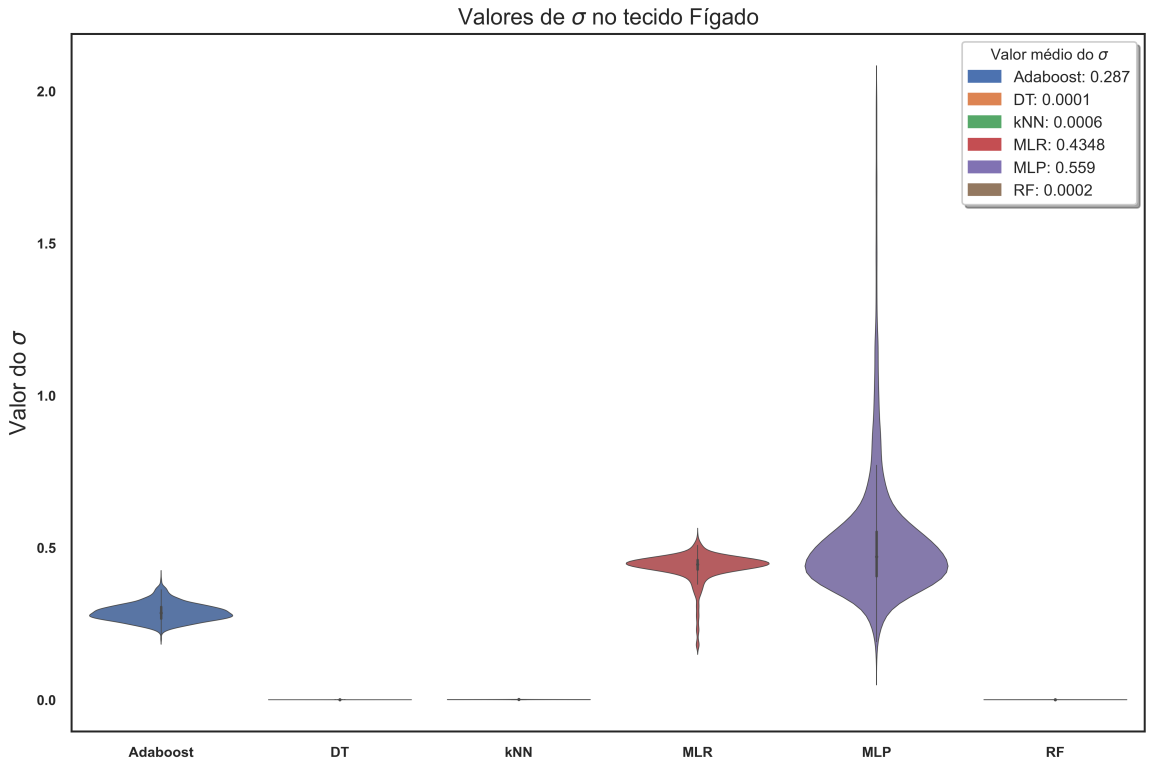
Fonte: Autor (2024).

No teste de hipótese, considerando o valor médio de cada *label*, a Regressão Linear Múltipla obteve um p-valor de 0,0, o *AdaBoost* de 0,0235, a *MLP* de 0,0, de 0,013 para o *kNN*

e de 0,018 no *RF*. Isso significa que para estes algoritmos o valor médio predito de intensidade I para o comprimento de onda λ não é o real. Por outro lado a *DT* resultou em um p-valor de 0,1328, respectivamente, ou seja, o valor médio predito por eles é o valor real.

A reconstrução de intensidade das curvas para este tecido nos seis algoritmos pode ser encontrada nas Figuras (A.17) e (A.18)

Figura 6.17: Gráfico de violino do σ obtido a partir da correção espectral usando os seis modelos de AM no **fígado**



Fonte: Autor (2024).

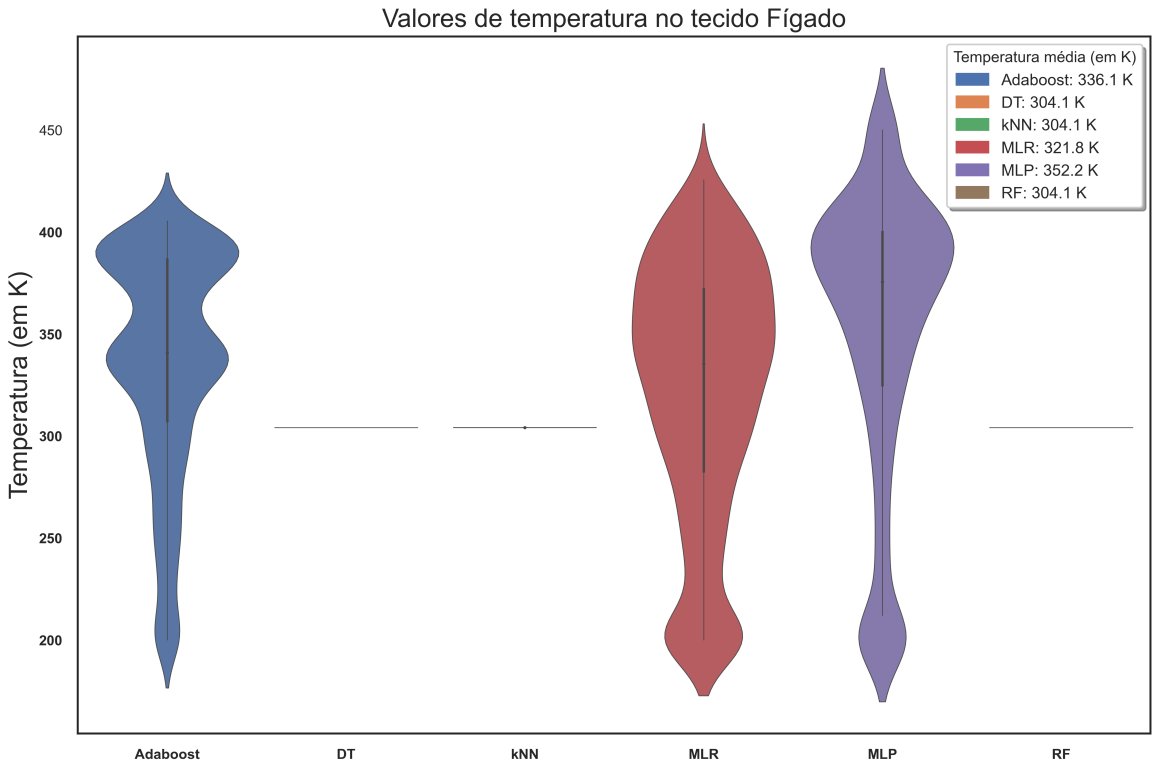
A variação mais acentuada do σ foi no *MLP* e na *MLR*. Já os σs obtidos através das curvas preditas pelos demais algoritmos tiveram valores próximos de 0. Maiores detalhes podem ser encontrados na Figura (6.17).

O gráfico de violino da temperatura (Figura 6.18) mostrou uma alta variação na temperatura obtida a partir dos espectros corrigidos usando *MLR*, *MLP* e *AdaBoost*. Os melhores resultados foram aqueles obtidos usando os seguintes algoritmos: *DT*, *RF* e *kNN*.

6.2.4 Pele

Para a Pele, o *MLR* foi o algoritmo mais rápido no tempo de treinamento, gastando apenas 25 segundos e com valores médios de *MAE*, *MSE*, *MAPE*, R^2 e σ de 1,0912, 5,4118, 12,9144, 0,6930, 0,5730, respectivamente. Em contrapartida o tempo de treinamento do *RF* foi

Figura 6.18: Gráfico de violino da temperatura obtida pela correção espectral usando os algoritmos de AM no **fígado**



Fonte: Autor (2024).

de aproximadamente 1 hora com valores médios de *MAE*, *MSE*, *MAPE*, R^2 e σ de 0,0002, 0,0, 0,0006, 1,0, 0,0001, respectivamente. Resultados semelhantes nas métricas foram obtidos nos algoritmos de *DT* e *kNN*, com tempo de treinamento menor, sendo de 1 minutos e 16 segundos para a Árvore de Decisão e de 1 minuto para o *kNN*. Resultados menos satisfatórios nas métricas de desempenho são do *AdaBoost*, com valores médios de *MAE*, *MSE*, *MAPE*, R^2 e σ de 0,7507, 0,7463, 13,5014, 0,9576, 0,4239, respectivamente; e da *MLP*, com os seguintes valores médios de *MAE*, *MSE*, *MAPE*, R^2 e σ: 1,1078, 3,8311, 6,3750, 0,7836, 0,9457. O *AdaBoost* resultou em um tempo de treinamento de aproximadamente 21 minutos, enquanto a *MLP* gastou cerca 14 minutos. Os valores estão resumidos na Tabela (6.10).

O desempenho abaixo do esperado do *AdaBoost* em comparação com *DT*, *RF* e *kNN* pode ser explicado pelo número limitado de estimadores, que foi de 50 em nossa análise, e pela escolha de uma função de perda linear. Em contraste, o *RF* utilizou 105 estimadores e empregou o *MSE* como critério de seleção. Quanto à *MLP*, o desempenho insatisfatório pode ser atribuído ao elevado número de camadas ocultas, que frequentemente faz com que o otimizador fique preso em mínimos locais. Além disso, a função de ativação ReLU, por ser linear, ajuda a manter a linearidade no modelo conexionista, pois aplicar uma transformação linear como a

ReLU preserva essa característica durante a busca pela solução.

Tabela 6.10: Principais dados de avaliação de correção espectral na Pele

Algoritmo	**Tempo**	***MAE* médio e *std***	***MSE* médio e *std***	***MAPE* médio e *std***	**R^2 médio e *std***	**σ médio e *std***
MLR	**25s**	1,0912 +/- 0,0175	5,4118 +/- 0,220	12,6144 +/- 0,9160	0,6930 +/- 0,0083	0,5730 +/- 0,0608
DT	1min16s	**0,0001 +/- 0,0001**	0,0001 +/- 0,0001	0,0003 +/- 0,0003	**1,0 +/- 0,0**	0,0001 +/- 0,0002
RF	1h0min28s	0,0002 +/- 0,0	**0,0 +/- 0,0**	0,0006 +/- 0,0001	**1,0 +/- 0,0**	**0,0001 +/- 0,0001**
AdaBoost	21min46s	0,7507 +/- 0,1118	0,7463 +/- 0,1142	13,5014 +/- 3.2267	0,9576 +/- 0,0066	0,4239 +/- 0,0514
kNN	1min4s	0,0012 +/- 0,0001	0,0001 +/- 0,0	0,0001 +/- 0,0	**1,0 +/- 0,0**	0,0007 +/- 0,0003
MLP	14min50s	1,1078 +/- 0,2305	3,8311 +/- 0,4268	6.3750 +/- 3,8150	0,7826 +/- 0,0241	0,9457 +/- 0,3286

Fonte: Autor (2024).

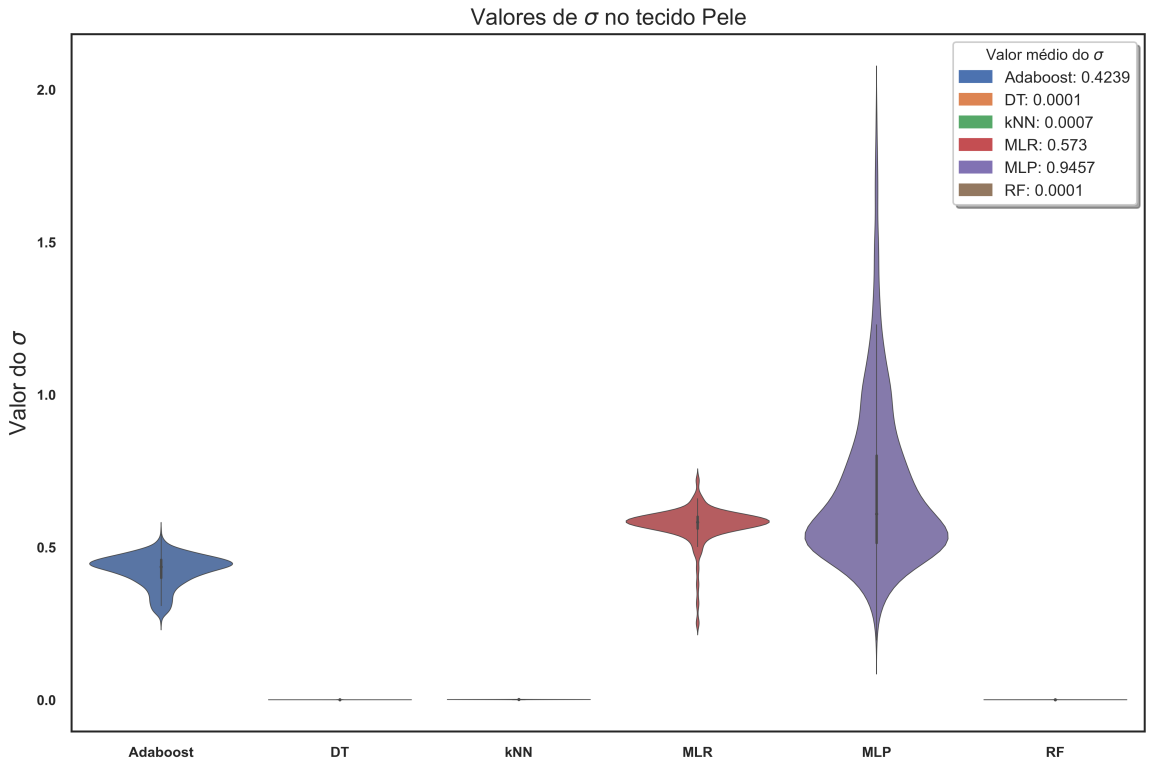
No teste de hipótese considerando o valor médio de cada *label*, a Regressão Linear Múltipla obteve um p-valor de 0,0, o *AdaBoost* de 0,0, a *MLP* de 0,0, o *kNN* de 0,0 e o *RF* de 0,0169. Logo nestes algoritmos o valor médio predito de intensidade I para o comprimento de onda λ não é o real. Por outro lado a *DT* resultou em um p-valor de 0,8125, ou seja, o valor médio predito por ele é o valor real.

A reconstrução de intensidade das curvas para este tecido nos seis algoritmos pode ser encontrada nas Figuras (A.19) e (A.20)

A variação mais acentuada do σ foi no *MLP*, seguido pelo *AdaBoost* e *MLR*. Os melhores resultados foram dos *kNN*, seguidos pela *DT* e *RF*, cujo valor do σ foi igual. Maiores detalhes podem ser encontrados na Figura (6.19).

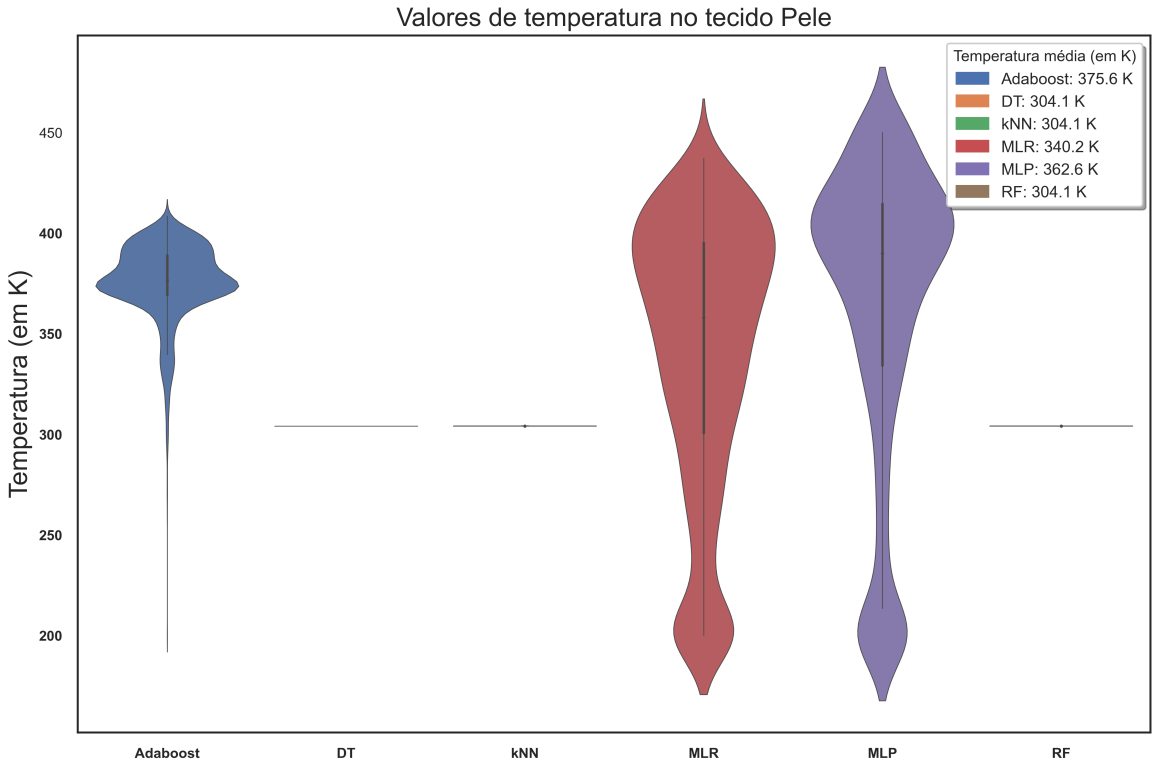
O gráfico de violino da temperatura mostrou uma variação considerável na temperatura usando as curvas corrigidas pela *MLR* e *MLP*, enquanto o *AdaBoost* obteve uma variância menor, sendo mais preciso que os anteriores. Os melhores resultados da temperatura usando as curvas corrigidas foram dos algoritmos: *DT*, *RF* e *kNN*, sendo exatos e precisos. Maiores detalhes podem ser encontrados na Figura (6.20).

Figura 6.19: Gráfico de violino da temperatura obtida pela correção espectral usando os algoritmos de AM na **pele**



Fonte: Autor (2024).

Figura 6.20: Gráfico de violino da temperatura obtida pela correção espectral usando os algoritmos de AM na **pele**



Fonte: Autor (2024).

Capítulo 7

Conclusão e Trabalhos Futuros

A nanotecnolgia é uma das tecnologias mais promissoras do século XXI. Neste contexto, o uso de NPs como nanotermômetros luminescentes tem grande potencial a ser explorado pela medicina, seja para mapeamento de tecidos através de imagens, seja para o diagnóstico e o tratamento de doenças. Contudo, neste último tipo de aplicação ainda há um cenário com diversos problemas a serem superados. No que se refere ao dianóstico e tratamento de doenças, pode-se mencionar as distorções espectrais causadas pelos tecidos quando as NPs estão contida neles, gerando distorções espectrais, bem como em uma medida incorreta da temperatura na região de interesse.

Neste trabalho foram usados 6 algoritmos de AM: *MLR*, *DT*, *RF*, *AdaBoost*, *kNN* e *MLP*. Quando as distorções espectrais foram geradas pelo coeficiente de absorção (μ_a) e profundidade, foram utilizados os seguintes tecidos: **massa cinzenta do cérebro**, **mama pré-menopausa**, **fígado**, **pele** e a **água**. Fixando a profundidade das NPs e variando o coeficiente de absorção foram geradas distorções na **pele**, na qual o parâmetro modificado foi a concentração de água. De modo análogo, quando as distorções espectrais foram geradas pelo coeficiente de absorção (μ_s) e pela profundidade, foram feitas correções usando os modelos de AM nos seguintes tecidos biológicos: **massa cinzenta do cérebro**, **mama pré-menopausa**, **fígado** e **pele**.

Os resultados mostraram métricas de desempenho próximas dos melhores valores possíveis. Em todos os tipos de decidos, sejam aqueles com as distorções causadas pelo coeficiente de absorção ou pelo coeficiente de espalhamento, obtiveram um R^2 médio maior ou igual a 0,78, havendo aqueles que atingiram o valor máximo possível: 1,0. Resultados similares em termos de desempenho foram obtidos com o *MAE*, *MSE* e *MAPE*, com valores nulos ou próximos de zero com arredondamento de 4 casas decimais e, com rara exceção, valor nulo na casa das unidades, ou seja, menores que 0,9. O σ teve resultados menores que 0,5, a maioria dos quais foram próximos de 0. Além disso os valores obtidos foram compatíveis com as demais métricas utilizadas (*MAE*, *MSE* e *MAPE*).

Dentre os algoritmos que obtiveram pior desempenho na correção espectral usando o coeficiente de absorção e a profundidade estão a *MLR*, *MLP* e *AdaBoost*. O teste de hipótese

também sustentou que tais modelos de AM nesse contexto não predizem o valor real do espectro. No cenário de bom desempenho, mas cujo teste de hipótese em alguns casos não sustentou que seja um bom preditor para este problema está o *kNN*. Por outro lado, a *DT* e o *RF*, além de obterem valores máximos de métricas e um baixo desvio padrão, sustentaram a hipótese que são bons preditores na correção espectral. Os resultados foram análogos usando o coeficiente de espalhamento (μ_s).

Outrossim, vale destacar que, apesar de ser um bom preditor, o *RF* é computacionalmente caro, deixando como melhor algoritmo em termos de desempenho computacional e resultados o algoritmo Árvore de Decisão. Vale também destacar que para os algoritmos com bom desempenho o *kNN* também pode ser explorado, ainda que o teste de hipótese não sustentou a hipótese nula.

Como o objetivo das correções espectrais é a obtenção de temperatura, os piores algoritmos foram a *MLR*, a *MLP* e o *AdaBoost*, cujos gráficos de violino do σ e da temperatura mostraram valores com uma alta variância e/ou baixa exatidão com imprecisão. Por outro lado, os outros três algoritmos, *DT*, *RF* e *kNN* tiveram uma alta extatidão e precisão, acertando a temperatura com 1 casa decimal. Nesse caso, os resultados foram melhores que o uso de curvas de calibração proposto por (XIMENDES et al., 2017), cuja diferença possui discrepâncias variando de 6 ºC a 11 ºC, enquanto que no caso das correções usando AM, os 3 algoritmos obtiveram valores corretos de temperatura com 1 casa decimal.

Os resultados aliados ao teste de hipótese Wilcoxon comprovaram que as técnicas de AM tem uma capacidade preditiva com precisão. Contudo, dentre as principais limitações, podemos destacar a utilização de dados gerados por uma lei bem definida, quando pode ser factível usar dados que não possuam nenhuma lei física conhecida, ou mesmo outras formas de combinação de distorções. Além disso, o número de tecidos utilizados, considerando um contexto de aplicações reais, é limitado.

Como trabalhos futuros, vale a pena mencionar o uso dos modelos com melhores resultados num cenário real, ou seja, em laboratório. Também podem ser verificado se os algoritmos com piores resultados na correção espectral podem ser melhorados a partir da busca de hiperparâmetros, como *Grid Search* e *Random Search*, especialmente para os modelos de regressão *MLR*, *AdaBoost* e *MLP*.

Por fim, este trabalho mostrou como o uso de algoritmos de AM para correção espectral e obtenção de temperatura em tecidos biológicos podem melhorar a capacidade preditiva, sendo tão efetiva quanto a proposta de (XIMENDES et al., 2017), além de automatizar o processo. Um avanço neste sentido poderá proporcionar o diagnóstico precoce de determinadas doenças e um tratamento mais eficaz, como no caso de terapias no contexto de tumores a base de hipertermia.

Bibliografia

ABADI, M., AGARWAL, A., BARHAM, P., BREVDO, E., CHEN, Z., CITRO, C., CORRADO, G. S., DAVIS, A., DEAN, J., DEVIN, M., GHEMAWAT, S., GOODFELLOW, I., HARP, A., IRVING, G., ISARD, M., JIA, Y., JOZEFOWICZ, R., KAISER, L., KUDLUR, M., LEVENBERG, J., MANÉ, D., MONGA, R., MOORE, S., MURRAY, D., OLAH, C., SCHUSTER, M., SHLENS, J., STEINER, B., SUTSKEVER, I., TALWAR, K., TUCKER, P., VANHOUCKE, V., VASUDEVAN, V., VIÉGAS, F., VINYALS, O., WARDEN, P., WATTENBERG, M., WICKE, M., YU, Y., and ZHENG, X. (2015). TensorFlow: Large-scale machine learning on heterogeneous systems. Software available from tensorflow.org.

ADEOLA, A. O., CLERMONT-PAQUETTE, A., PIEKNY, A., and NACCACHE, R. (2023). Advances in the design and use of carbon dots for analytical and biomedical applications. *Nanotechnology*, 35(1):012001.

ARIA, M. and CUCCURULLO, C. (2017). bibliometrix: An r-tool for comprehensive science mapping analysis. *Journal of informetrics*, 11(4):959–975.

BASHKATOV, A. N., GENINA, E. A., and TUCHIN, V. V. (2011). Optical properties of skin, subcutaneous, and muscle tissues: a review. *Journal of Innovative Optical Health Sciences*, 4(01):9–38.

BASHKATOV, A. N., GENINA, E. A., and TUCHIN, V. V. (2016). Tissue optical properties. In *Handbook of biomedical optics*, pages 87–122. CRC Press.

BATISTA, G. E. and MONARD, M. C. (2003). An analysis of four missing data treatment methods for supervised learning. *Applied artificial intelligence*, 17(5-6):519–533.

BEDNARKIEWICZ, A., MARCINIAK, L., CARLOS, L. D., and JAQUE, D. (2020). Standardizing luminescence nanothermometry for biomedical applications. *Nanoscale*, 12(27):14405–14421.

BENTLEY, J. L. (1975). Multidimensional binary search trees used for associative searching. *Communications of the ACM*, 18(9):509–517.

BISSOTO, A., PEREZ, F., RIBEIRO, V., FORNACIALI, M., AVILA, S., and VALLE, E. (2018). Deep-learning ensembles for skin-lesion segmentation, analysis, classification: Recod titans at isic challenge 2018. *arXiv preprint arXiv:1808.08480*.

BOUGUER, P. (1729). *Essai d'optique sur la gradation de la lumière*. Claude Jombert.

BREAIMAN, L. (2001). Random forests. *Machine learning*, 45:5–32.

BREIMAN, L., FRIEDMAN, J., OLSHEN, R., and STONE, C. (1984). Cart. *Classification and regression trees*.

BUSKIRK, T. D. (2018). Surveying the forests and sampling the trees: an overview of classification and regression trees and random forests with applications in survey research. *Survey Practice*, 11(1):1–13.

DE JONG, W. H. and BORM, P. J. (2008). Drug delivery and nanoparticles: applications and hazards. *International journal of nanomedicine*, 3(2):133–149.

DEL ROSAL, B., XIMENDES, E., ROCHA, U., and JAQUE, D. (2017). In vivo luminescence nanothermometry: from materials to applications. *Advanced Optical Materials*, 5(1):1600508.

DONG, J. and ZINK, J. I. (2014). Taking the temperature of the interiors of magnetically heated nanoparticles. *ACS nano*, 8(5):5199–5207.

DRUCKER, H. (1997). Improving regressors using boosting techniques. In *Icml*, volume 97, pages 107–115. Citeseer.

FAYYAD, U., PIATETSKY-SHAPIRO, G., and SMYTH, P. (1996). From data mining to knowledge discovery in databases. *AI magazine*, 17(3):37–37.

FIX, E. and HODGES, J. L. (1951). Discriminatory analysis, nonparametric discrimination.

FRANK, S. M., FLEISHER, L. A., BRESLOW, M. J., HIGGINS, M. S., OLSON, K. F., KELLY, S., and BEATTIE, C. (1997). Perioperative maintenance of normothermia reduces the incidence of morbid cardiac events: a randomized clinical trial. *Jama*, 277(14):1127–1134.

FREEDMAN, D. A. (2009). *Statistical models: theory and practice*. cambridge university press.

FREUND, Y. and SCHAPIRE, R. E. (1995). A decision-theoretic generalization of on-line learning and an application to boosting.

GRÖMPING, U. (2009). Variable importance assessment in regression: linear regression versus random forest. *The American Statistician*, 63(4):308–319.

HAYAKAWA, C. K., MALENFANT, L., RANASINGHESAGARA, J., CUCCIA, D. J., SPANIER, J., and VENUGOPALAN, V. (2022). Mccl: an open-source software application for monte carlo simulations of radiative transport. *Journal of biomedical optics*, 27(8):083005–083005.

HUANG, X., LIN, J., and DEMNER-FUSHMAN, D. (2006). Evaluation of pico as a knowledge representation for clinical questions. In *AMIA annual symposium proceedings*, volume 2006, page 359. American Medical Informatics Association.

HURST, GEORGE SAMUEL, G. J. D. S. J. O. and CHU, S. (2021). Spectroscopy. urllhttps://www.britannica.com/science/spectroscopy.

JACQUES, S. L. and MCAULIFFE, D. J. (1991). The melanosome: threshold temperature for explosive vaporization and internal absorption coefficient during pulsed laser irradiation. *Photochemistry and photobiology*, 53(6):769–775.

JAMES, G., WITTEN, D., HASTIE, T., and TIBSHIRANI, R. (2013). *An introduction to statistical learning*, volume 112. Springer.

JAQUE, D. and VETRONE, F. (2012). Luminescence nanothermometry. *Nanoscale*, 4(15):4301–4326.

KEELE, S. and OTHERS (2007). Guidelines for performing systematic literature reviews in software engineering.

KEFENI, K. K., MSAGATI, T. A., and MAMBA, B. B. (2017). Ferrite nanoparticles: synthesis, characterisation and applications in electronic device. *Materials Science and Engineering: B*, 215:37–55.

KIM, S.-B., HAN, K.-S., RIM, H.-C., and MYAENG, S. H. (2006). Some effective techniques for naive bayes text classification. *IEEE transactions on knowledge and data engineering*, 18(11):1457–1466.

KRUKEMEYER, M., KRENN, V., HUEBNER, F., WAGNER, W., and RESCH, R. (2015). History and possible uses of nanomedicine based on nanoparticles and nanotechnological progress. *J. Nanomed. Nanotechnol*, 6(6):336.

KUSSUL, E. and BAIDYK, T. (2004). Improved method of handwritten digit recognition tested on mnist database. *Image and Vision Computing*, 22(12):971–981.

LABOUTA, H. I., ASGARIAN, N., RINKER, K., and CRAMB, D. T. (2019). Meta-analysis of nanoparticle cytotoxicity via data-mining the literature. *ACS nano*, 13(2):1583–1594.

LAMBERT, J. H. (1760). *Photometria sive de mensura et gradibus luminis, colorum et umbrae*. sumptibus vidvae E. Klett, typis CP Detleffsen.

LARSEN, T., SCHMID, S., VILLANUEVA, L. G., and BOISEN, A. (2013). Photothermal analysis of individual nanoparticulate samples using micromechanical resonators. *Acs Nano*, 7(7):6188–6193.

LIANG, R., TIAN, R., SHI, W., LIU, Z., YAN, D., WEI, M., EVANS, D. G., and DUAN, X. (2013). A temperature sensor based on cdte quantum dots–layered double hydroxide ultrathin films via layer-by-layer assembly. *Chemical Communications*, 49(10):969–971.

MCCULLOCH, W. S. and PITTS, W. (1943). A logical calculus of the ideas immanent in nervous activity. *The bulletin of mathematical biophysics*, 5:115–133.

MONTGOMERY, D. C., RUNGER, G. C., and CALADO, V. (2000). *Estatística Aplicada E Probabilidade Para Engenheiros* . Grupo Gen-LTC.

MOORE, G. E. (2006). Cramming more components onto integrated circuits, reprinted from electronics, volume 38, number 8, april 19, 1965, pp. 114 ff. *IEEE solid-state circuits society newsletter*, 11(3):33–35.

NEXHA, A., CARVAJAL, J. J., PUJOL, M. C., D´AZ, F., and AGUILÓ, M. (2021). Lanthanide doped luminescence nanothermometers in the biological windows: strategies and applications. *Nanoscale*, 13(17):7913–7987.

OMOHUNDRO, S. M. (1989). *Five balltree construction algorithms*. International Computer Science Institute Berkeley.

OWENS, A., GODBOLE, M., DABYDEEN, D., MEDEIROS, L., PHATAK, P., and KANDLIKAR, S. (2020). A comparative analysis of the tumor pathology and the metabolic heat generation of growing malignant tumors. In *International Conference on Nanochannels, Microchannels, and Minichannels*, volume 83693, page V001T12A007. American Society of Mechanical Engineers.

PANKHURST, Q. A., CONNOLLY, J., JONES, S. K., and DOBSON, J. (2003). Applications of magnetic nanoparticles in biomedicine. *Journal of physics D: Applied physics*, 36(13):R167.

PEDREGOSA, F., VAROQUAUX, G., GRAMFORT, A., MICHEL, V., THIRION, B., GRISEL, O., BLONDEL, M., PRETTENHOFER, P., WEISS, R., DUBOURG, V., VANDERPLAS, J., PASSOS, A., COURNAPEAU, D., BRUCHER, M., PERROT, M., and DUCHESNAY, E. (2011). Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research*, 12:2825–2830.

RICHARDSON, W. S., WILSON, M. C., NISHIKAWA, J., and HAYWARD, R. S. (1995). The well-built clinical question: a key to evidence-based decisions. *ACP journal club*, 123(3):A12–A13.

ROCO, M. (1999). Nanoparticles and nanotechnology research. *Journal of Nanoparticle Research*, 1(1):1.

ROCO, M. C. (2011). The long view of nanotechnology development: the national nanotechnology initiative at 10 years.

ROMERO, C. and VENTURA, S. (2007). Educational data mining: A survey from 1995 to 2005. *Expert systems with applications*, 33(1):135–146.

RUSSELL, S. and NORVIG, P. (2013). Inteligência artificial, tradução de regina célia simille. *Rio de Janeiro: Campus Elsevier*.

SANCHEZ, B. M., LESCH, M., BRAMMER, D., BOVE, S. E., THIEL, M., and KILGORE, K. S. (2008). Use of a portable thermal imaging unit as a rapid, quantitative method of evaluating inflammation and experimental arthritis. *Journal of pharmacological and toxicological methods*, 57(3):169–175.

SANTOS, H. D., XIMENDES, E. C., IGLESIAS-DE LA CRUZ, M. D. C., CHAVES-COIRA, I., DEL ROSAL, B., JACINTO, C., MONGE, L., RUBIA-RODR´GUEZ, I., ORTEGA, D., MATEOS, S., et al. (2018). In vivo early tumor detection and diagnosis by infrared luminescence transient nanothermometry. *Advanced Functional Materials*, 28(43):1803924.

SCHARDT, C., ADAMS, M. B., OWENS, T., KEITZ, S., and FONTELO, P. (2007). Utilization of the pico framework to improve searching pubmed for clinical questions. *BMC medical informatics and decision making*, 7:1–6.

SCHMIDT, M. E., HERMANN, S., ARNDT, V., and STEINDORF, K. (2020). Prevalence and severity of long-term physical, emotional, and cognitive fatigue across 15 different cancer entities. *Cancer medicine*, 9(21):8053–8061.

SHEN, Y., LIFANTE, J., FERNANDEZ, N., JAQUE, D., and XIMENDES, E. (2020). In vivo spectral distortions of infrared luminescent nanothermometers compromise their reliability. *ACS nano*, 14(4):4122–4133.

SIDNEY, S. (1957). Nonparametric statistics for the behavioral sciences. *The Journal of Nervous and Mental Disease*, 125(3):497.

SINGH, D. and SINGH, B. (2020). Investigating the impact of data normalization on classification performance. *Applied Soft Computing*, 97:105524.

SMETS, E., GARSSEN, B., SCHUSTER-UITTERHOEVE, A., and DE HAES, J. (1993). Fatigue in cancer patients. *British journal of cancer*, 68(2):220–224.

SOARES, A. C., SALES, T. O., XIMENDES, E. C., JAQUE, D., and JACINTO, C. (2023). Lanthanide doped nanoparticles for reliable and precise luminescence nanothermometry in the third biological window. *Nanoscale Advances*.

SPERLING, R. A., GIL, P. R., ZHANG, F., ZANELLA, M., and PARAK, W. J. (2008). Biological applications of gold nanoparticles. *Chemical Society Reviews*, 37(9):1896–1908.

STARK, A. M. and WAY, S. (1974). The use of thermovision in the detection of early breast cancer. *Cancer*, 33(6):1664–1670.

SUZUKI, M., TSEEB, V., OYAMA, K., and ISHIWATA, S. (2007). Microscopic detection of thermogenesis in a single hela cell. *Biophysical journal*, 92(6):L46–L48.

SWEENEY, A. E. (2015). Nanomedicine concepts in the general medical curriculum: initiating a discussion. *International journal of nanomedicine*, pages 7319–7331.

ULBRICH, K., HOLA, K., SUBR, V., BAKANDRITSOS, A., TUCEK, J., and ZBORIL, R. (2016). Targeted drug delivery with polymers and magnetic nanoparticles: covalent and noncovalent approaches, release control, and clinical studies. *Chemical reviews*, 116(9):5338–5431.

VIRTANEN, P., GOMMERS, R., OLIPHANT, T. E., HABERLAND, M., REDDY, T., COURNAPEAU, D., BUROVSKI, E., PETERSON, P., WECKESSER, W., BRIGHT, J., VAN DER WALT, S. J., BRETT, M., WILSON, J., MILLMAN, K. J., MAYOROV, N., NELSON, A. R. J., JONES, E., KERN, R., LARSON, E., CAREY, C. J., POLAT, İ., FENG, Y., MOORE, E. W., VANDERPLAS, J., LAXALDE, D., PERKTOLD, J., CIMRMAN, R., HENRIKSEN, I., QUINTERO, E. A., HARRIS, C. R., ARCHIBALD, A. M., RIBEIRO, A. H., PEDREGOSA, F., VAN MULBREGT, P., and scipy 1.0 contributors (2020). SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python. *Nature Methods*, 17:261–272.

WALKER, H. K., HALL, W. D., and HURST, J. W. (1990). Clinical methods: the history, physical, and laboratory examinations.

WANG, H., LHUILLIER, E., YU, Q., ZIMMERS, A., DUBERTRET, B., ULYSSE, C., and AUBIN, H. (2017). Transport in a single self-doped nanocrystal. *ACS nano*, 11(2):1222–1229.

WILCOXON, F. (1947). Probability tables for individual comparisons by ranking methods. *Biometrics*, 3(3):119–122.

WOHLIN, C., RUNESON, P., HÖST, M., OHLSSON, M. C., REGNELL, B., and WESSLÉN, A. (2012). *Experimentation in software engineering*. Springer Science & Business Media.

WU, S., YANG, Z., MA, C., ZHANG, X., MI, C., ZHOU, J., GUO, Z., and JIN, D. (2023). Deep learning enhanced nir-ii volumetric imaging of whole mice vasculature. *Opto-Electronic Advances*.

XIMENDES, E. C., ROCHA, U., DEL ROSAL, B., VAQUERO, A., SANZ-RODR´GUEZ, F., MONGE, L., REN, F., VETRONE, F., MA, D., GARC´A-SOLÉ, J., et al. (2017). In vivo ischemia detection by luminescent nanothermometers. *Advanced Healthcare Materials*, 6(4):1601195.

YAHARA, T., KOGA, T., YOSHIDA, S., NAKAGAWA, S., DEGUCHI, H., and SHIROUZU, K. (2003). Relationship between microvessel density and thermographic hot areas in breast cancer. *Surgery today*, 33:243–248.

YAN, X. and SU, X. (2009). *Linear regression analysis: theory and computing*. world scientific.

ZHOU, H., SHARMA, M., BEREZIN, O., ZUCKERMAN, D., and BEREZIN, M. Y. (2016). Nanothermometry: from microscopy to thermal treatments. *ChemPhysChem*, 17(1):27–36.

ZHU, X. and WU, X. (2004). Class noise vs. attribute noise: A quantitative study. *Artificial intelligence review*, 22:177–210.

ZIEGLER, M., DATHE, A., POLLOK, K., LANGENHORST, F., HUBNER, U., WANG, D., and SCHAAF, P. (2020). Metastable atomic layer deposition: 3d self-assembly toward ultra-dark materials. *ACS nano*, 14(11):15023–15031.

ZOHAR, O., IKEDA, M., SHINAGAWA, H., INOUE, H., NAKAMURA, H., ELBAUM, D., ALKON, D. L., and YOSHIOKA, T. (1998). Thermal imaging of receptor-activated heat production in single cells. *Biophysical journal*, 74(1):82–89.

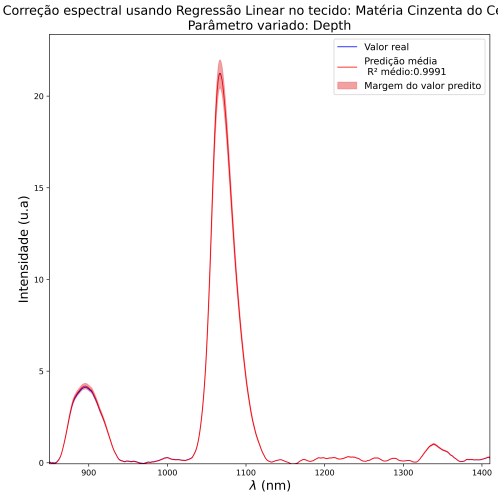
Apêndice A

Correções espectrais

A.1 Correções espectrais obtidas a partir dos espectros distorcidos usando μ_a e profundidade

Figura A.1: Correções espectrais na massa cinzenta do cérebro usando (a) *MLR* e (b) *DT*.

(a)



(b)

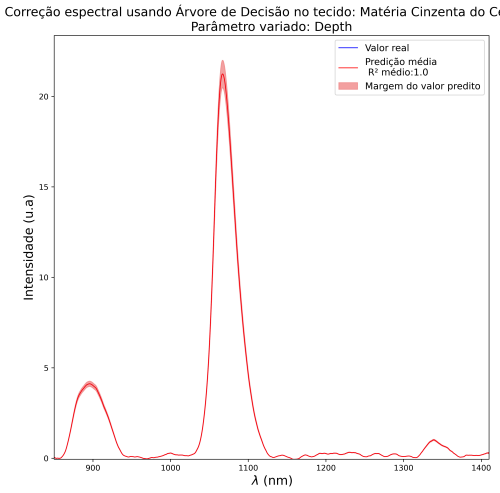
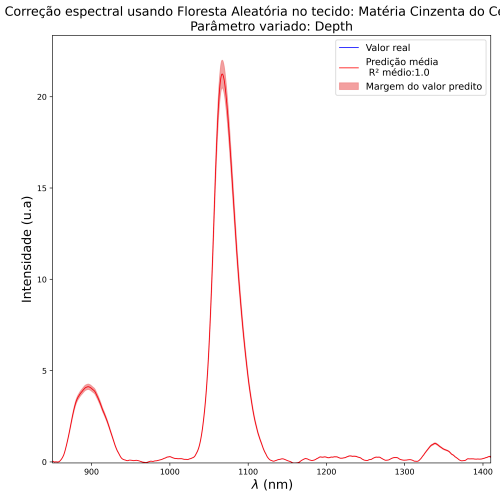
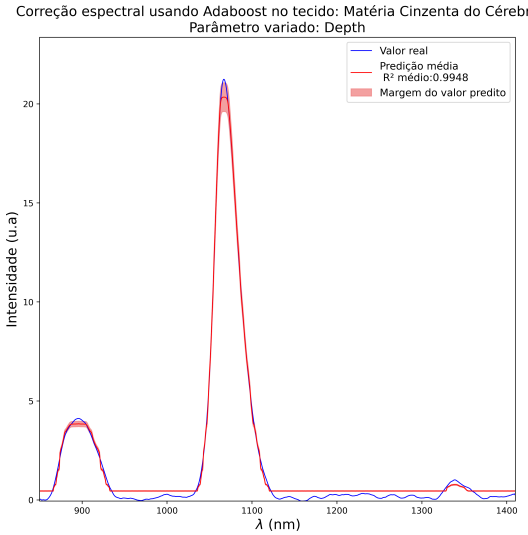


Figura A.2: Correções espectrais na massa cinzenta do cérebro usando (a) *RF*, (b) *AdaBoost*, (c) *kNN* e (d) *MLP*.

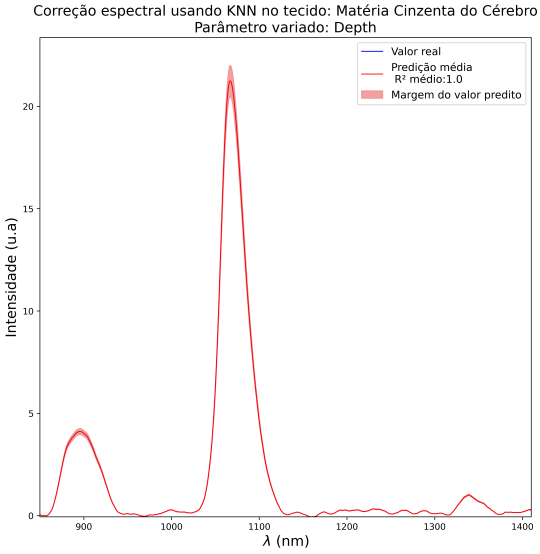
(a)



(b)



(c)



(d)

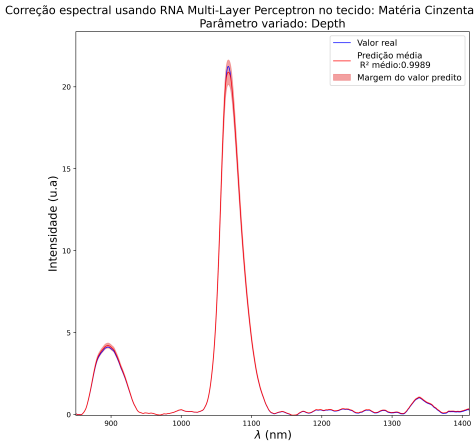
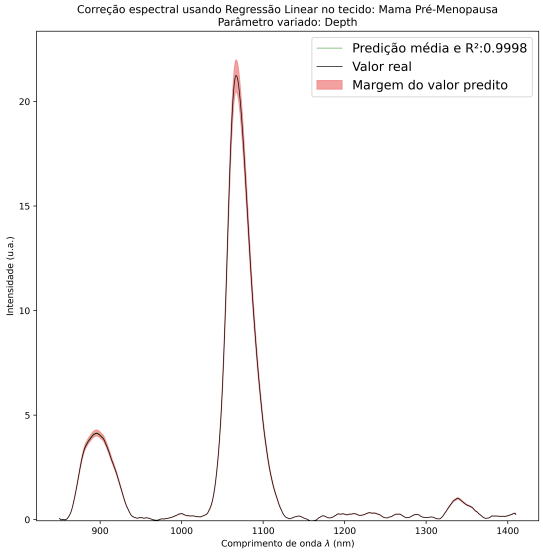
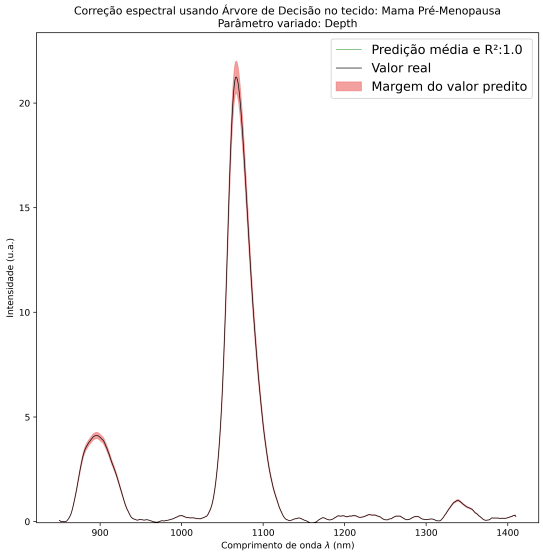


Figura A.3: Correções espectrais na mama pré-menopausa usando (a) *MLR*, (b) *DT*, (c) *RF* e (d) *AdaBoost*.

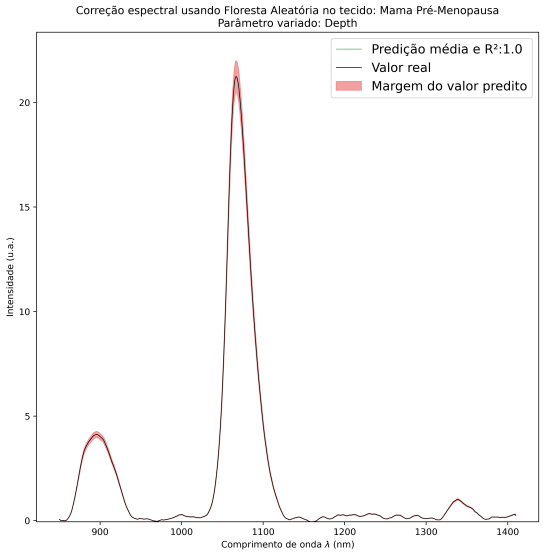
(a)



(b)



(c)



(d)

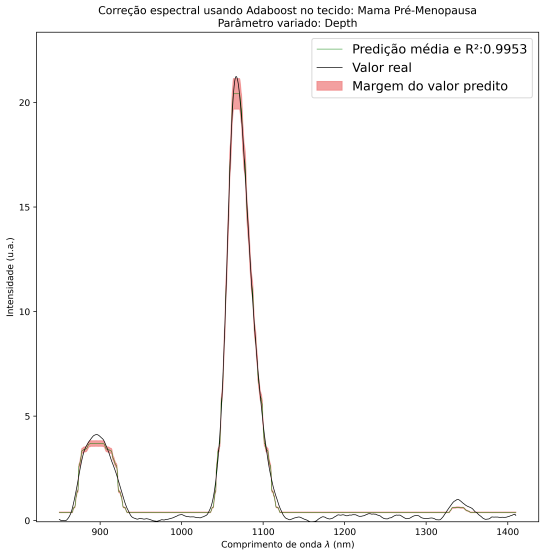
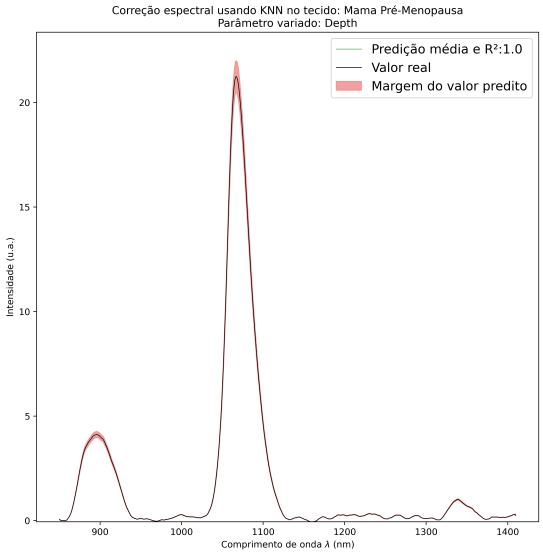


Figura A.4: Correções espectrais na mama pré-menopausa usando (a) *kNN* e (b) *MLP*.

(a)



(b)

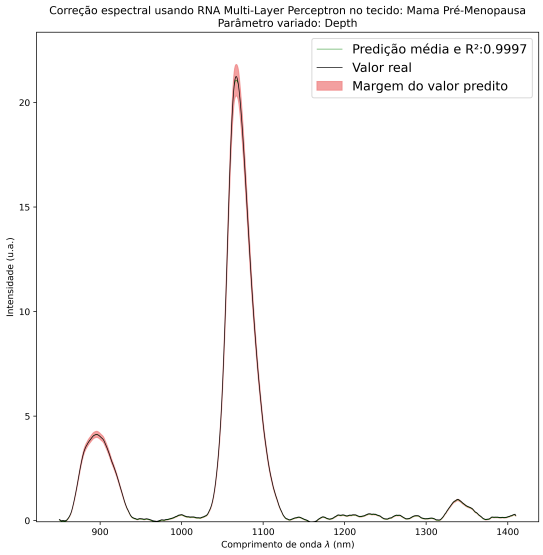
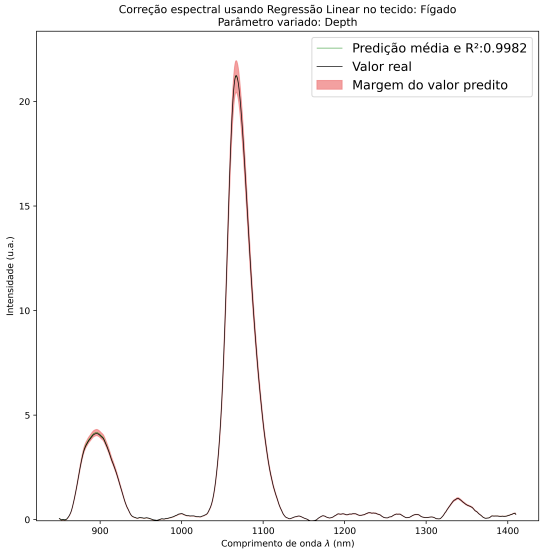


Figura A.5: Correções espectrais no fígado usando (a) *MLR* e (b) *DT*.

(a)



(b)

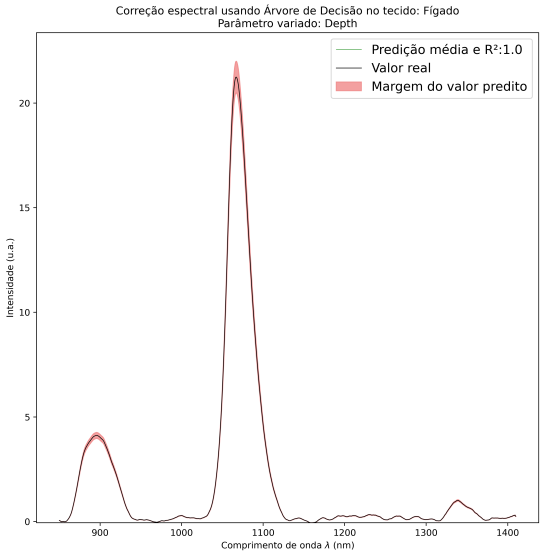
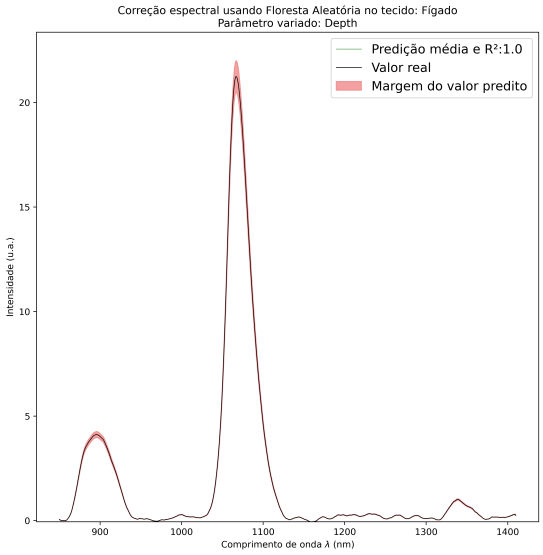
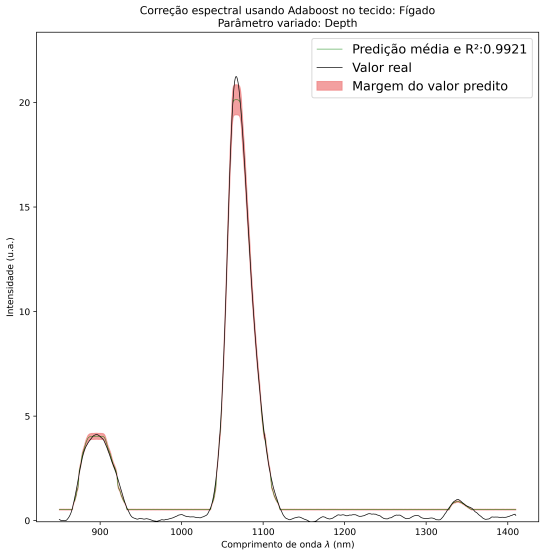


Figura A.6: Correções espectrais no fígado usando (a) *RF*, (b) *AdaBoost*, (c) *kNN* e (d) *MLP*.

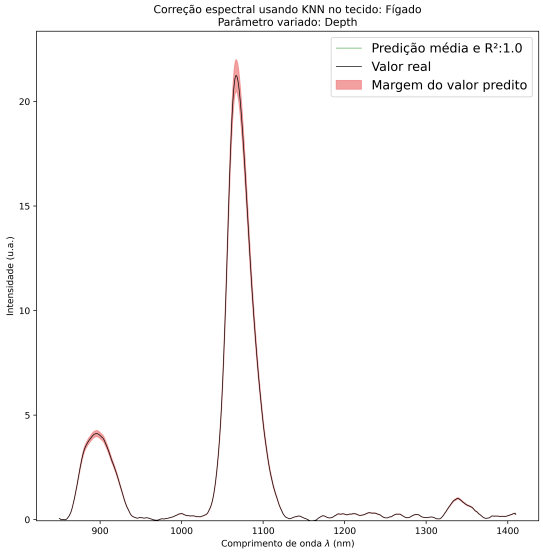
(a)



(b)



(c)



(d)

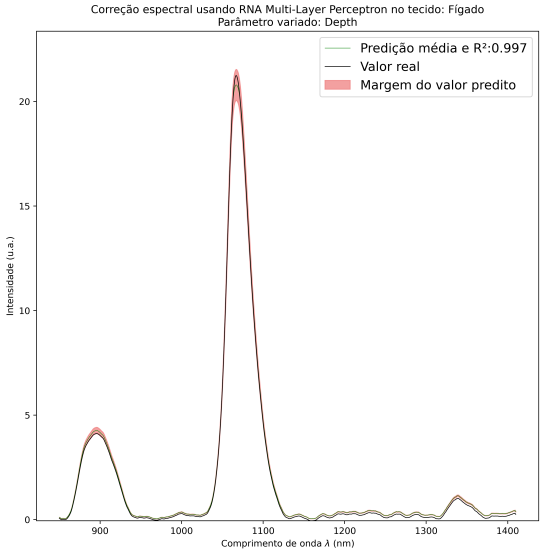
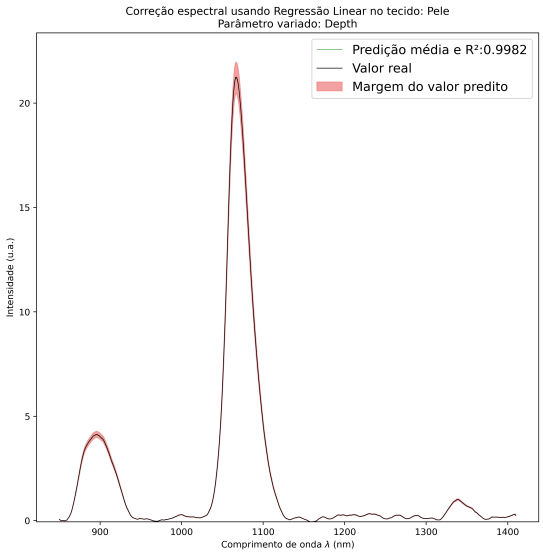
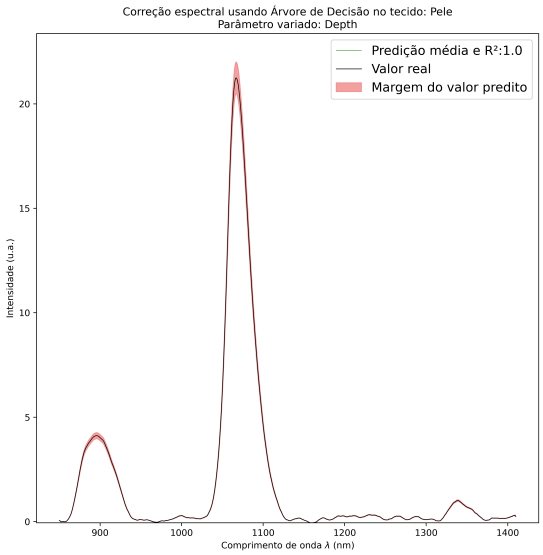


Figura A.7: Correções espectrais na pele usando (a) *MLR*, (b) *DT*, (c) *RF* e (d) *AdaBoost*.

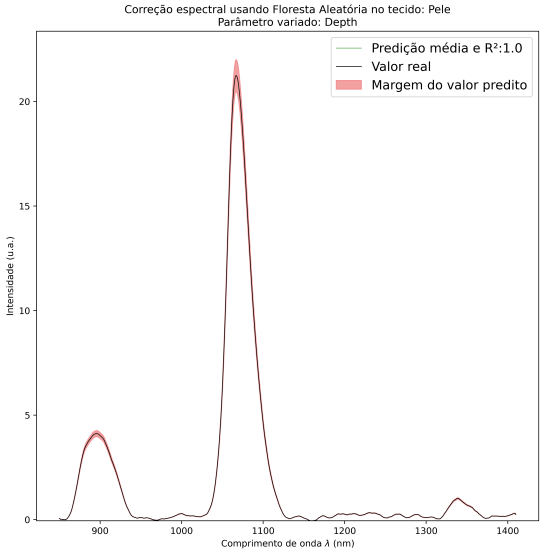
(a)



(b)



(c)



(d)

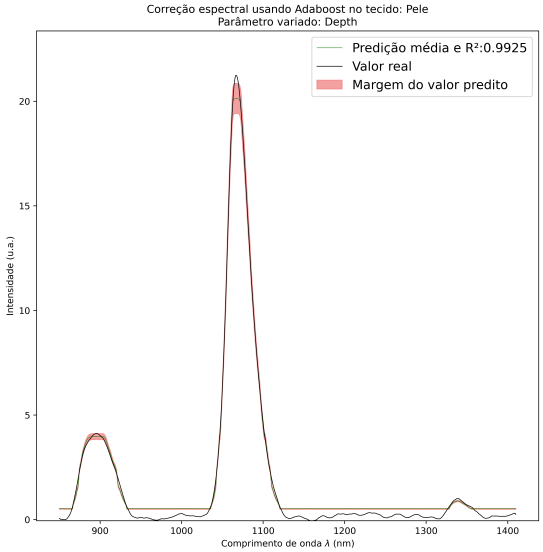


Figura A.8: Correções espectrais na pele usando (a) *kNN* e (b) *MLP*.

(a) (b)

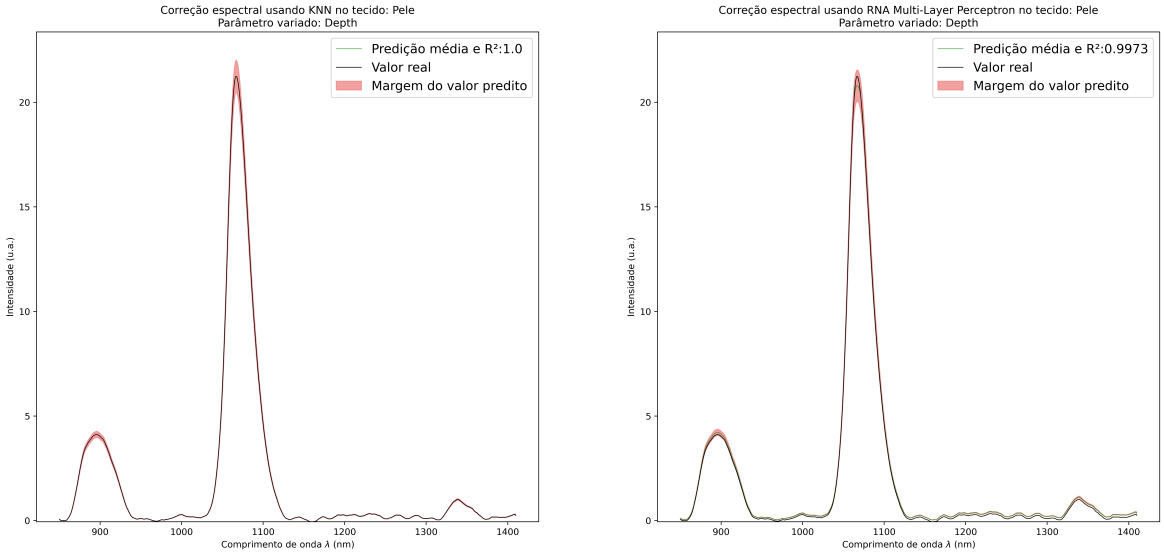


Figura A.9: Correções espectrais na pele variando a concentração de água usando (a) *MLR* e (b) *DT*.

(a) (b)

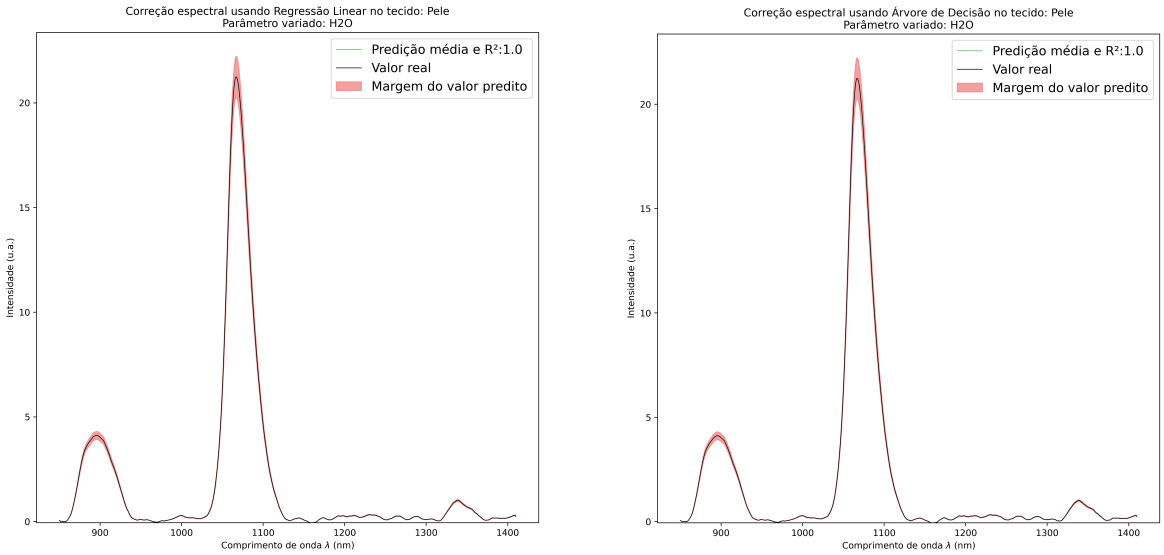
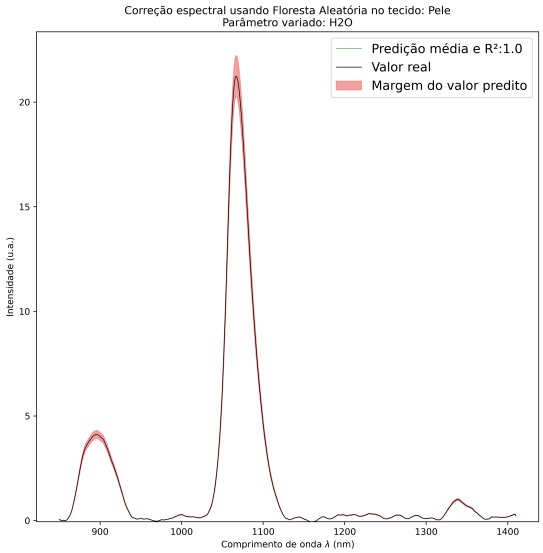
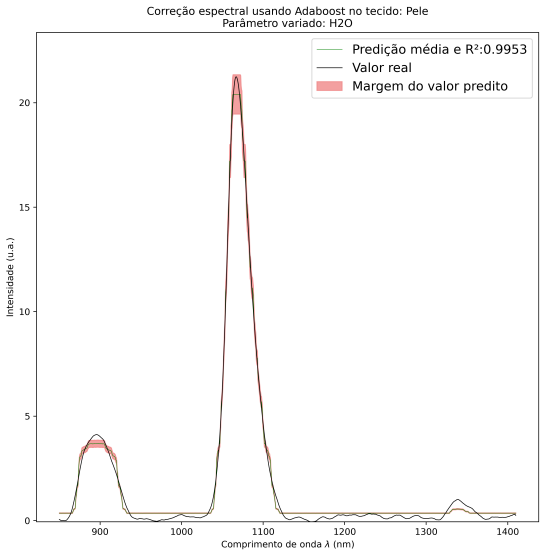


Figura A.10: Correções espectrais na pele variando a concentração de água usando (a) *RF*, (b) *AdaBoost*, (c) *kNN* e (d) *MLP*.

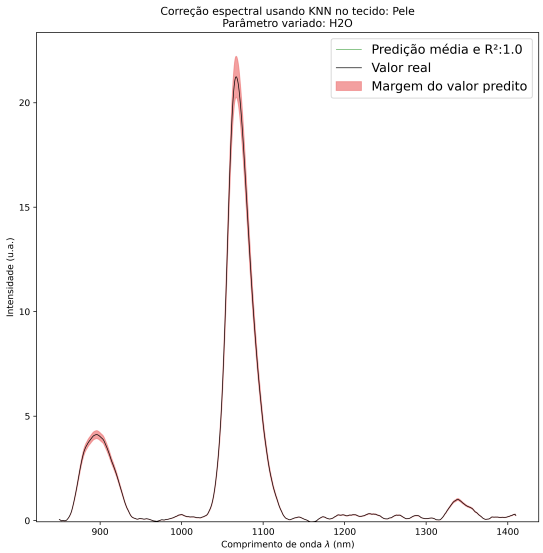
(a)



(b)



(c)



(d)

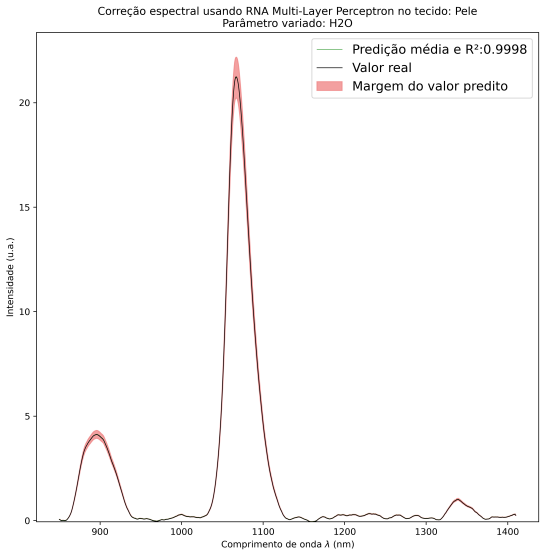
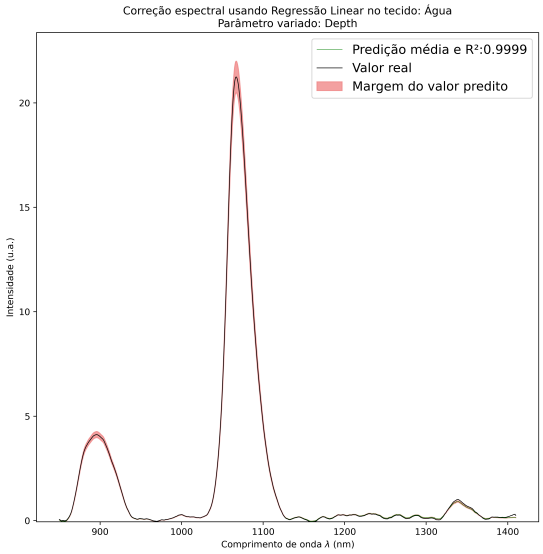
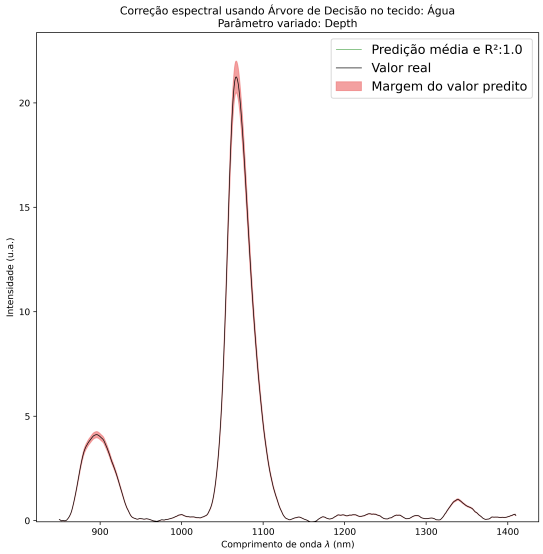


Figura A.11: Correções espectrais na água usando (a) *MLR*, (b) *DT*, (c) *RF* e (d) *AdaBoost*.

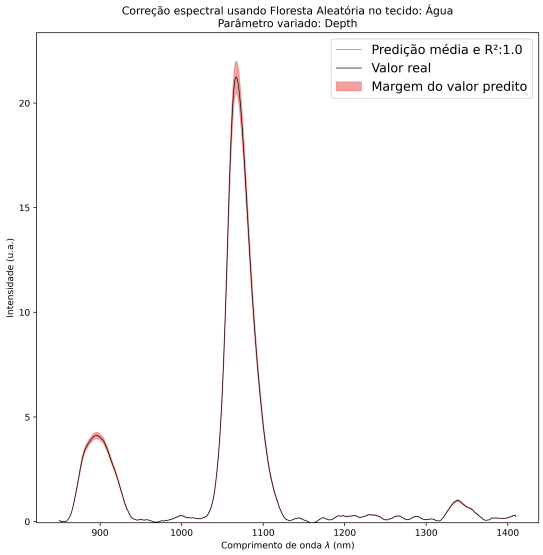
(a)



(b)



(c)



(d)

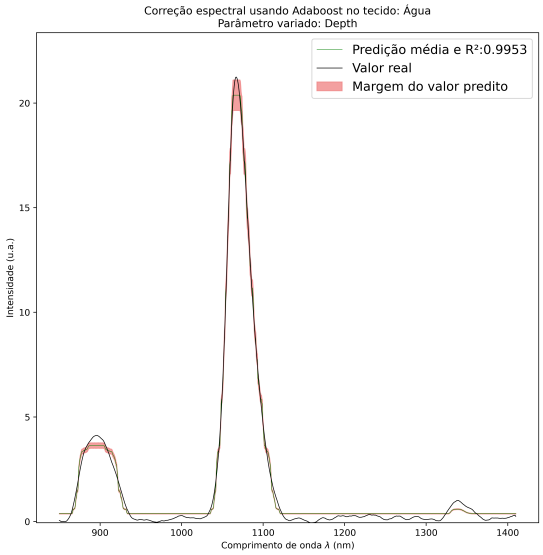
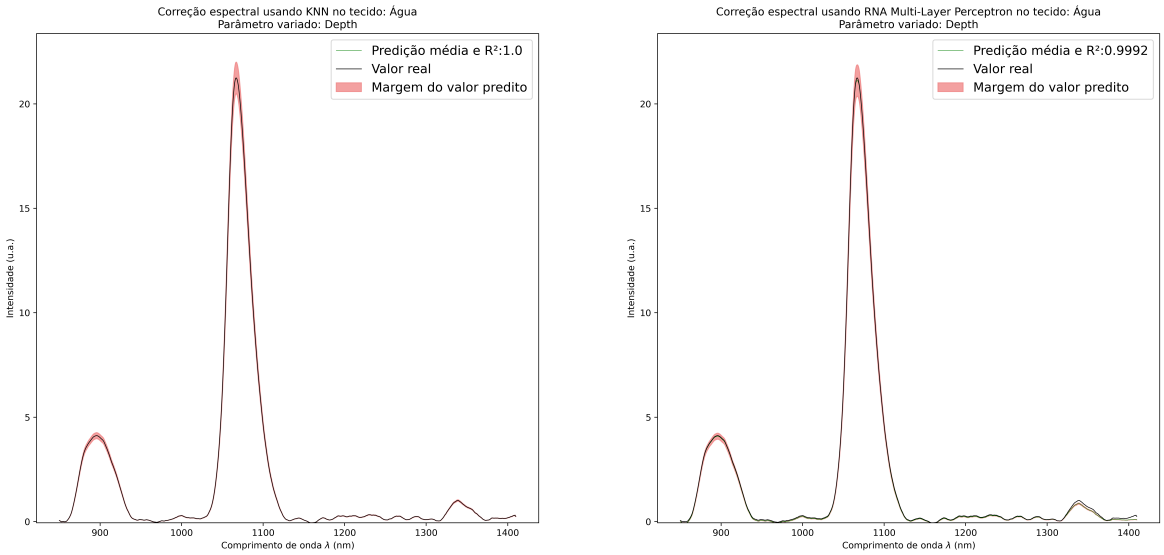


Figura A.12: Correções espectrais na água usando (a) *kNN* e (b) *MLP*.

(a) (b)



A.2 Correções espectrais obtidas a partir dos espectros distorcidos usando μ_s e profundidade

Figura A.13: Correções espectrais na massa cinzenta do cérebro usando (a) *MLR* e (b) *DT*.

(a) (b)

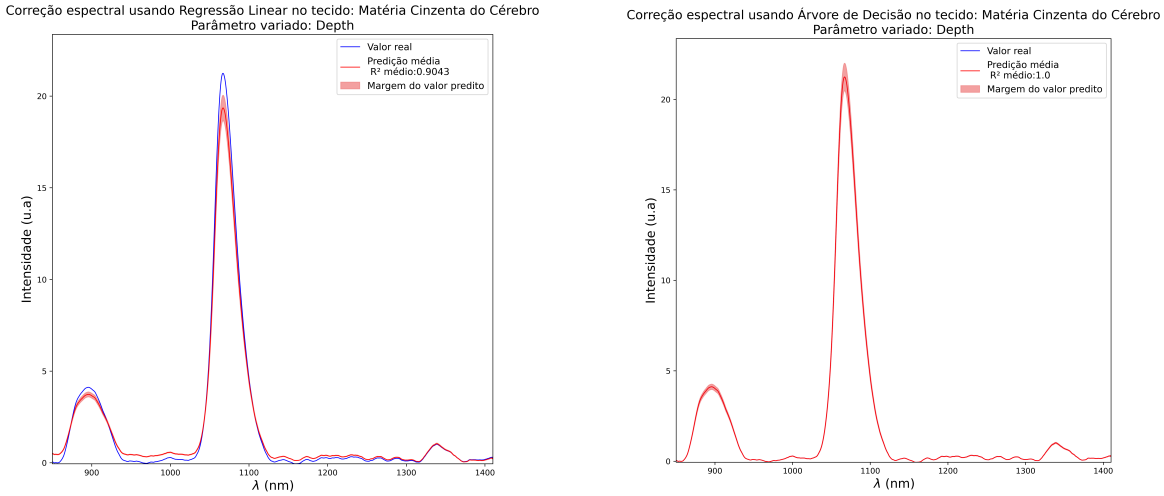
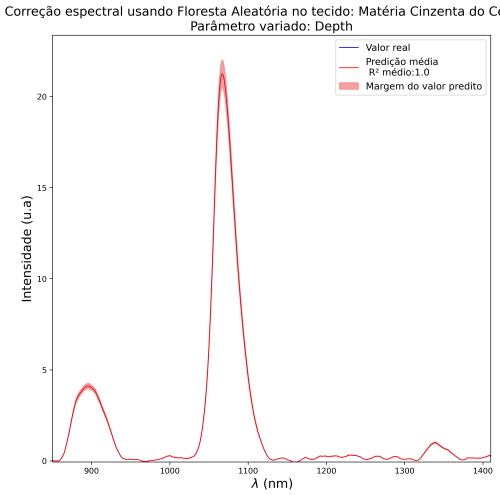
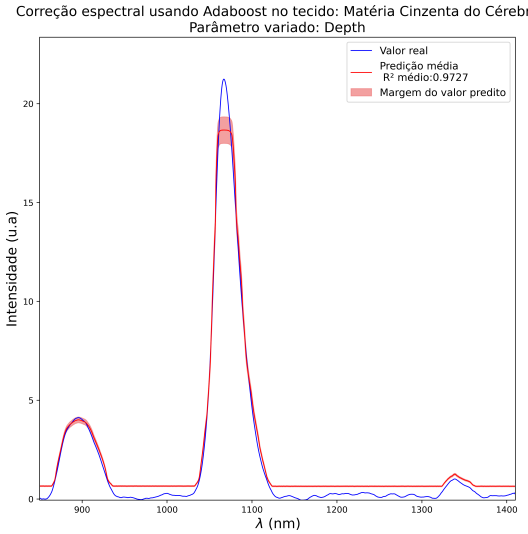


Figura A.14: Correções espectrais na massa cinzenta do cérebro usando (a) *RF*, *AdaBoost*, (c) *kNN* e (d) *MLP*.

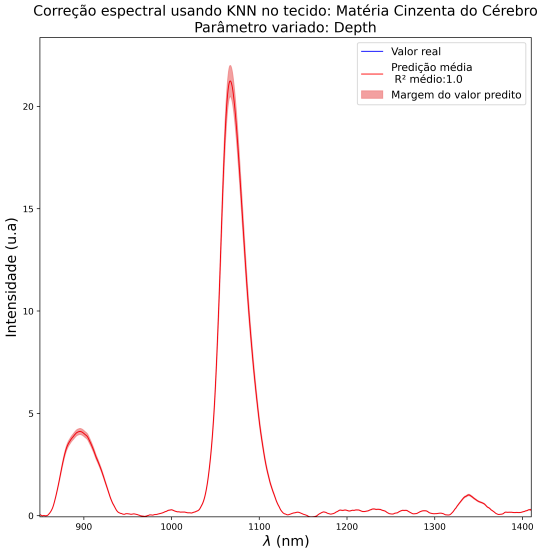
(a)



(b)



(c)



(d)

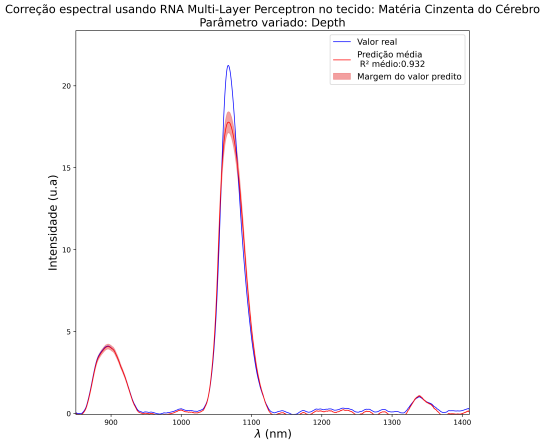
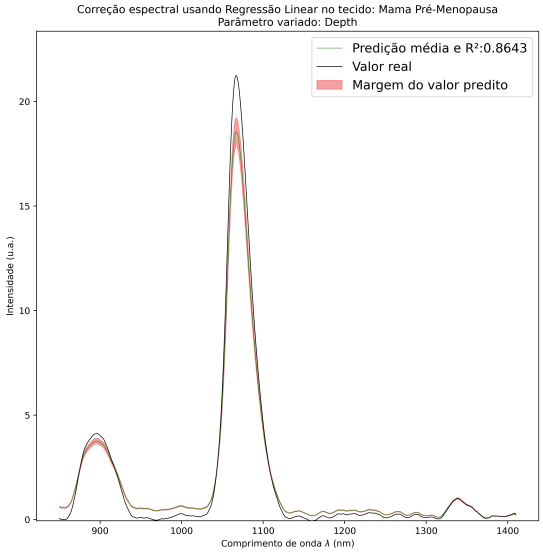
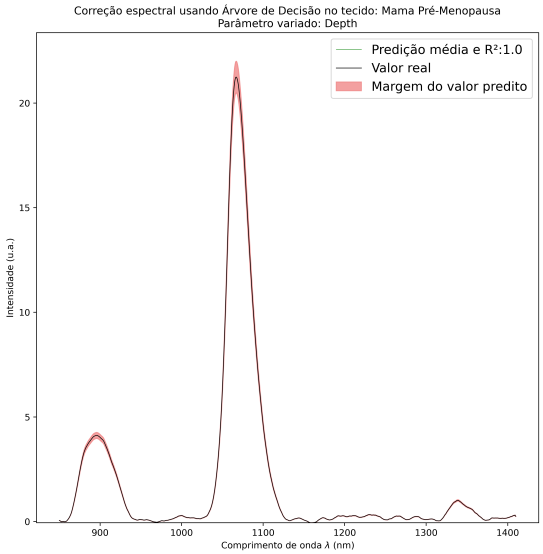


Figura A.15: Correções espectrais na mama pré-menopausa usando (a) *MLR*, (b) *DT*, (c) *RF*, (d) *AdaBoost*

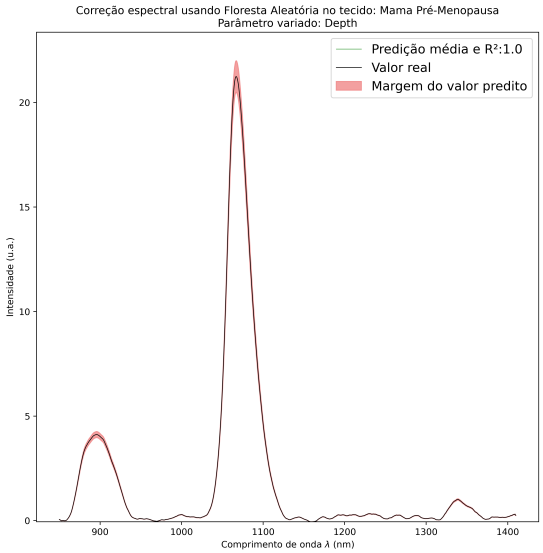
(a)



(b)



(c)



(d)

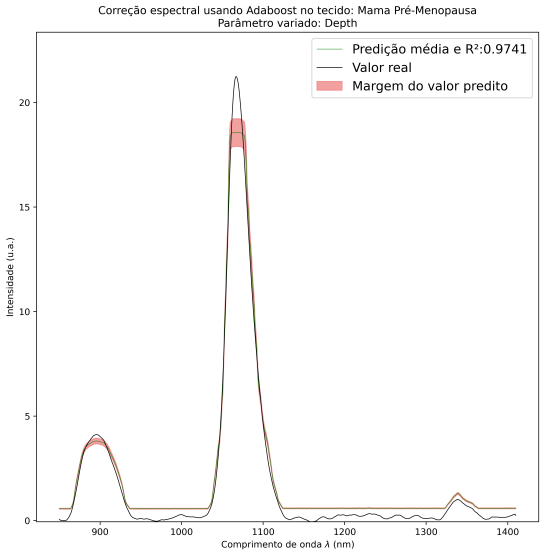


Figura A.16: Correções espectrais na mama pré-menopausa usando (a) *kNN* e (b) *MLP*.

(a) (b)

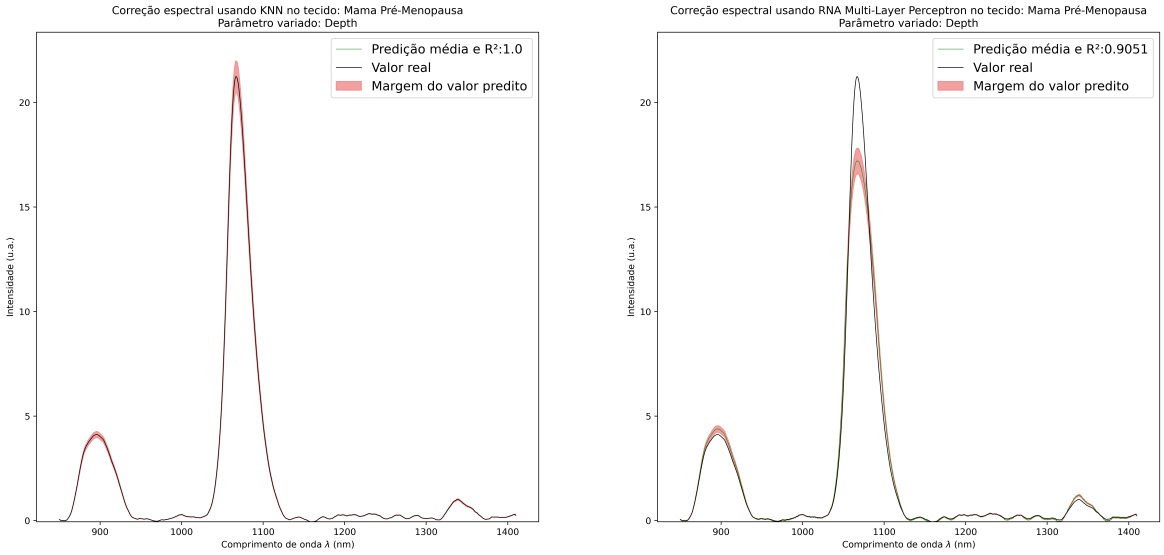


Figura A.17: Correções espectrais no fígado usando (a) *MLR* e (b) *DT*.

(a) (b)

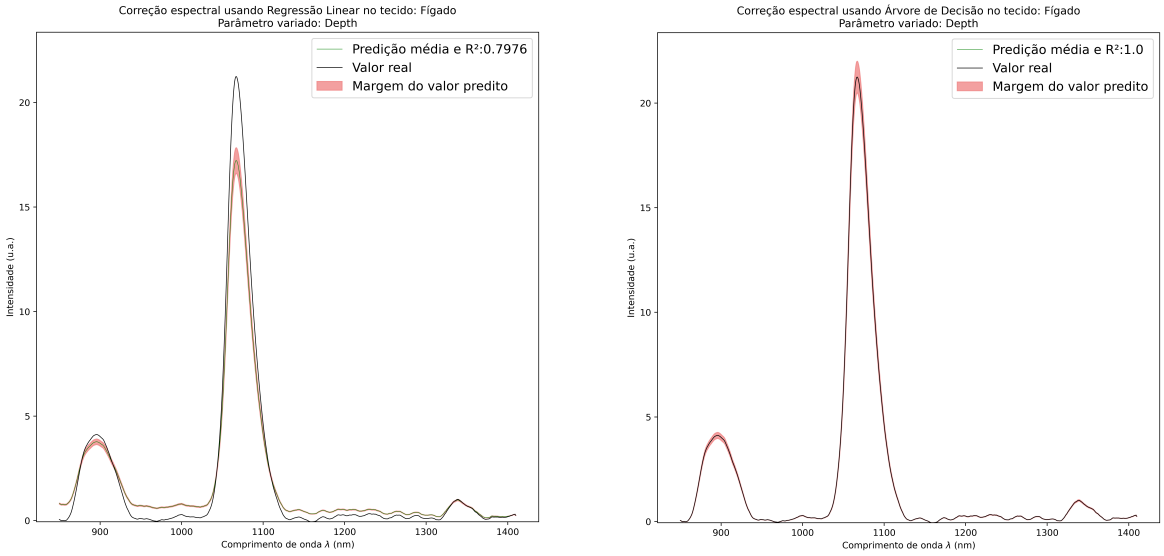
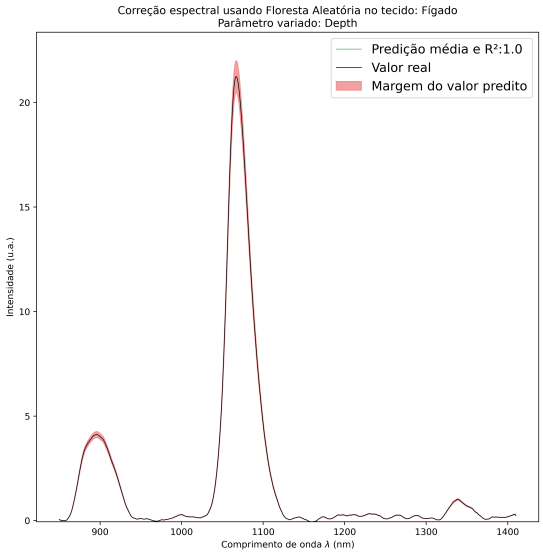
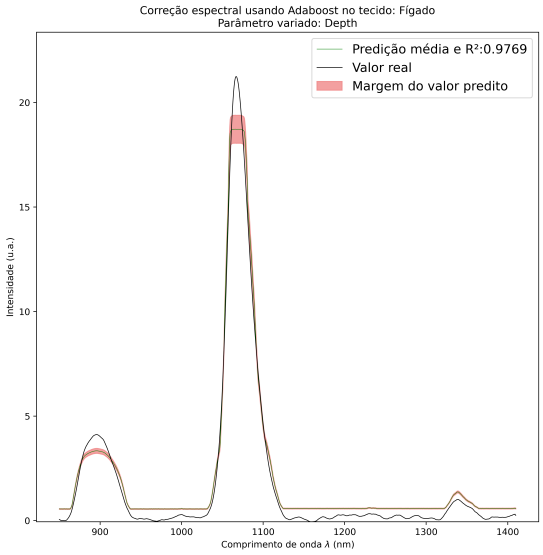


Figura A.18: Correções espectrais no fígado usando (a) *RF*, (b) *AdaBoost*, (c) *kNN* e (d) *MLP*

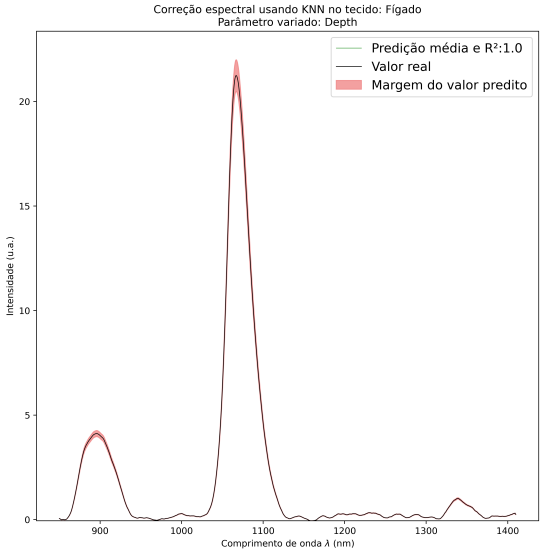
(a)



(b)



(c)



(d)

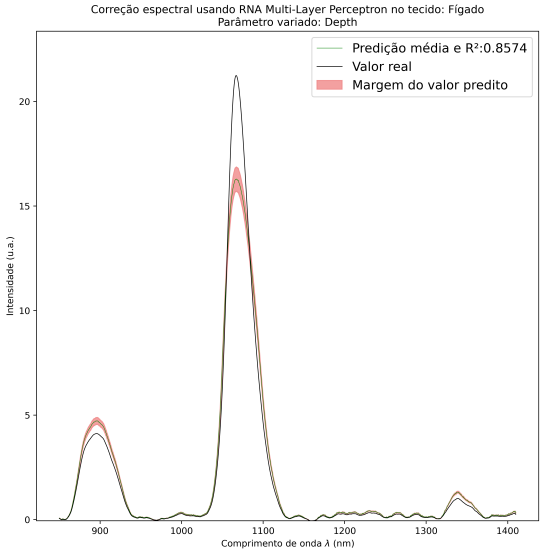
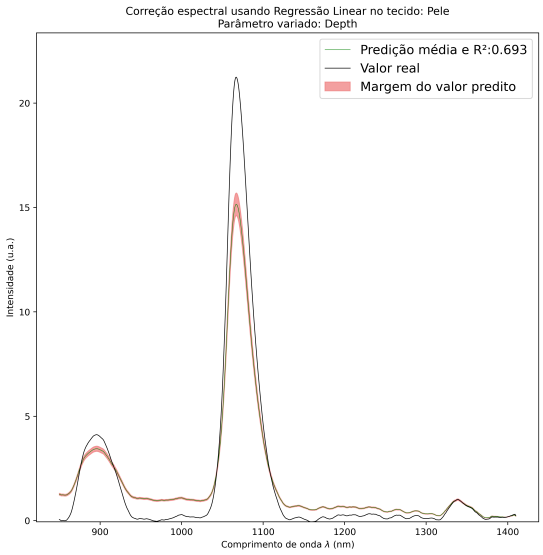
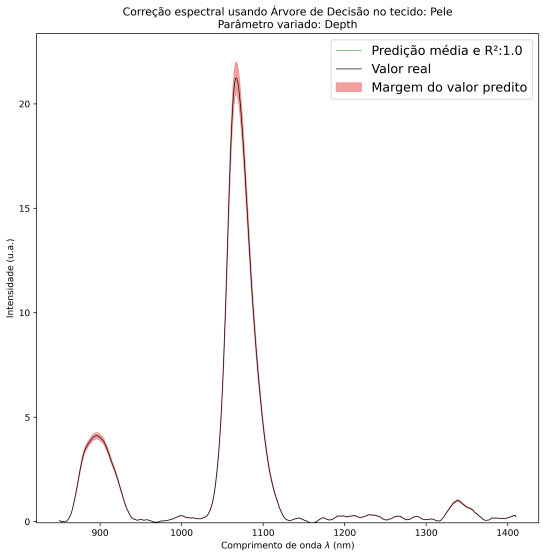


Figura A.19: Correções espectrais no pele usando (a) *MLR*, (b) *DT*, (c) *RF* e (d) *AdaBoost*

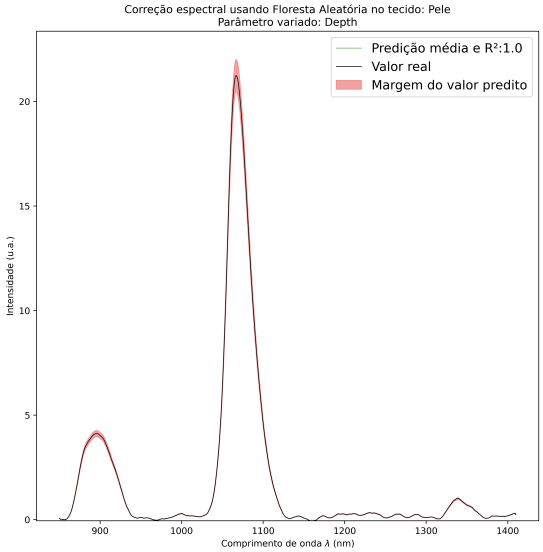
(a)



(b)



(c)



(d)

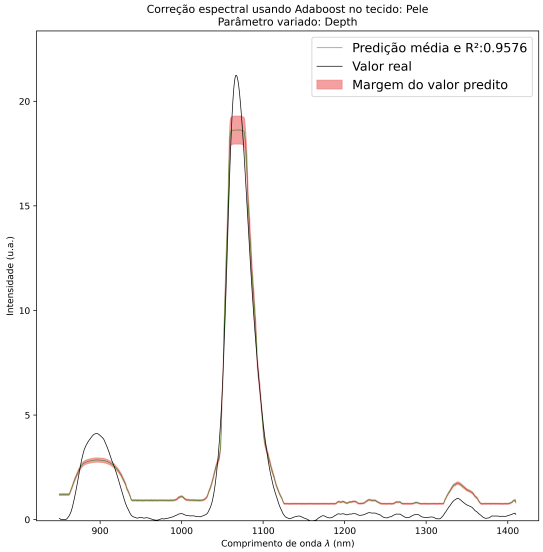
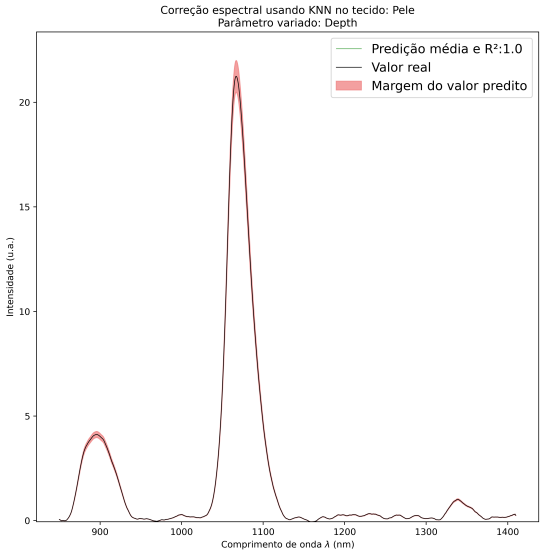
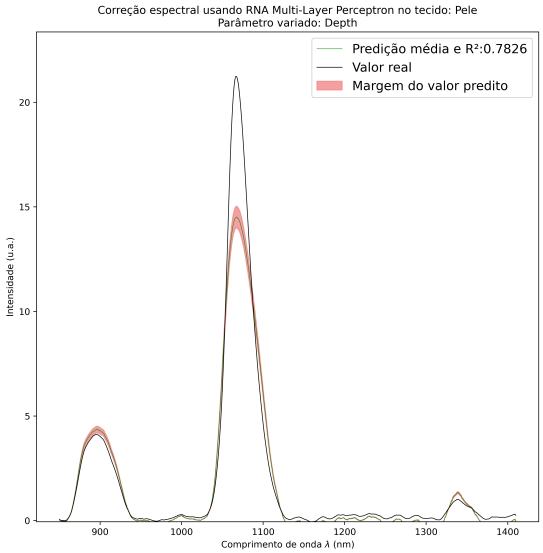


Figura A.20: Correções espectrais na pele usando (a) *kNN* e (b) *MLP*.

(a)



(b)



Apêndice B

Revisão Sistemática de Literatura

B.0.1 Metodologia

Para realizar uma RSL adequada é preciso primeiro elaborar perguntas de pesquisa que sejam adequadas ao problema que se deseja resolver. O intuito é realizar um mapeamento preliminar para saber os principais tópicos que são cobertos pela pesquisa. Neste trabalho houve uma investigação envolvendo duas áreas: Nanotermometria (Física) e Aprendizagem de Máquina (Ciência da Computação).

O segundo passo é usar o critério ou método de PICOC, que é uma adaptação do critério PICO. Este último é um critério consolidado nas pesquisas envolvendo áreas médicas (RICHARDSON et al., 1995); (SCHARDT et al., 2007); (HUANG et al., 2006). Esta técnica foi sugerida e adapatada para a Engenharia de Software por (KEELE and OTHERS, 2007). Seu intuito foi trazer uma metodologia de investigação de outra área (a Medicina), e adequá-la em um dos campos da Ciência da Computação (Engenharia de Software). Obviamente, ela pode ser acomodada outras áreas do conhecimento humano que exijam uma metodologia de revisão. Vale destacar ainda o acréscimo do Contexto (correspondente à letra C do acrônimo PICOC) que diz respeito à adoção da ferramenta para a realização da RSL, que é chamada de Parsifal (WOHLIN et al., 2012). O acrônimo PICOC tem seu significado resgatado do inglês cujo significado de cada letra é mostrado a seguir:

1. P: do inglês *Population*, diz respeito a qual população de interesse a pesquisa está se tratando.

2. I: do inglês *Intervention*, corresponde aos tipos de tecnologias, procedimentos e/ou ferramentas para a resolução do problema.

3. C: do inglês *Comparison*, técnicas para comparação das soluções buscadas na resolução do problema.

4. O: do inglês *Outcome*, o tipo de saída desejada na pesquisa enquanto propostas de resolução para o problema.

5. C: do inglês *Context*, significa o contexto em que a pesquisa está sendo trabalhada.

Após a elaboração das perguntas de pesquisa adequadas, o próximo passo é usar palavras ou termos coerentes que descrevam a investigação científica. O objetivo específico deste tipo de procedimento é obter um conjunto de palavras-chave que possa ser usado em bases científicas de dados científicas e, uma vez aplicadas outras etapas, trazer os trabalhos mais relevantes e recentes da problemática tratada, bem como uma descrição tão completa e objetiva quanto possível das técnicas e ferramentas mencionadas, trazendo à tona lacunas que possam ser preenchidas. Nesta fase de investigação da RSL, será um trabalho mais fácil encontrar *gaps*, auxiliando em uma fundamentação e justifica sólidas no trabalho que está sendo desenvolvido.

Dando sequência a metodologia de pesquisa usando o Parsifal, as strings de busca são inseridas nas bases de dados escolhidas. Para isso são feitas adaptações textuais levando em conta a organização e as exigências de cada base. Vale salientar ainda que a escoha das bases de busca deve ser feita levando em consideração a qualidade dos trabalhos e o contexto em que a pesquisa está inserida.

O próximo passo na RSL é elaborar uma seleção dos estudos usando os critérios de inclusão e exclusão adequados. O intuito é realizar uma triagem para extrair os artigos mais relevantes e que podem contribuir com a investigação do problema. Nesta etapa os trabalhos resultantes trazem bastante informação a respeito das lacunas, técnicas e outros aspectos sobre a pesquisa, uma vez que é feita após o acesso aos textos completos e uma leitura panorâmica e/ou do abstract. O Parsifal propicia essa seleção de artigos de forma fácil, mesmo sendo um processo manual. Ao final irão restar os artigos para a avaliação de qualidade, boa parte dos quais serão usados como base para a pesquisa.

A avaliação de qualidade tem o intuito de realizar um estudo mais criterioso acerca dos artigos escolhidos. É nessa fase que é realizada uma leitura aprofundada dos artigos escolhidos. O objetivo é elencar técnicas e abordagens que procurar resolver o problema. Nesta etapa atribui-se uma nota aos artigos previamente selecionados. A escala de pontuação é da preferência do pesquisador ou grupo de pesquisa envolvidos. A ideia chave é checar quais *papers* podem contribuir com a problemática escolhida.

Na etapa de extração dos dados é realizada uma extração de informação sobre os artigos que passaram pela fase de avaliação de qualidade. Um resumo das informações relevantes dos artigos que restaram é fixada e pode ser utilizada para os fins necessários. Uma visão bem mais detalhada acerca de informações relevantes dos artigos é vista de forma nítida.

A última etapa diz respeito a uma discussão da RSL. Ela deve abordar técnicas, ferramentas e outros aspectos cobertos pelos artigos selecionados. Os artigos que chegam nesta fase devem nortear toda a pesquisa.

B.0.2 Análise Bibliométrica

A nossa análise bibliográfica dentro da RSL é feita seguindo os passos mencionados anteriormente. Isso torna a RSL sólida e auditável, além de uma metodologia que é de fato sistemática.

Perguntas de Pesquisa

A partir do título da nossa pesquisa foram elaboradas as seguintes perguntas de estudo:

RQ1 - Quais métodos existem para correção de distorção espectral de nanopartículas como nanotermômetros em tecidos biológicos?

RQ2 - Quais métodos existentes para corrigir distorções espectrais de nanotermômetros em tecidos biológicos usaram aprendizado de máquina ou aprendizado profundo?

RQ3 - Quais métodos existem para correção espectral de nanopartículas como nanotermômetros em tecidos biológicos foram mais eficazes?

Critério PICOC Após a elaboração cuidadosa das perguntas de pesquisa, foram escolhidas as palavras-chave que descrevessem de forma tão adequada quanto possível as questões levantadas anteriormente. Os resultados são indicados na Tabela B.1.

Tabela B.1: Critérios PICOC

Population	nanoparticle spectrum
Intervention	machine learning
Comparison	None
Outcome	correct spectrum
Context	tissues

Fonte: Autor (2023)

Na ferramenta Parsifal é possível inserir palavras sinônimos para cada uma das palavras escolhidas no critério PICOC. O intuito é explorar de forma abrangente e direcionada trabalhos relevantes para o trabalho. Nesse sentido e considerando tanto as perguntas de pesquisa quanto as palavras do critério PICOC, foram escolhidos sinônimos exibidos na Tabela B.2.

Palavras-Chave e Sinônimos

Strings de Busca e Bases As strings de busca gerada pelo Parsifal de acordo com o critério PICOC e a Tabela B.2 foram as seguintes: ("nanoparticle spectrum"OR "nanoparticles"OR "nanothermometers spectrum"OR "nanothermometry") AND ("machine learning"OR "artificial intelligence"OR "method") AND ("correct spectrum"OR "reliability spectrum"OR "spectrum"). As strings de busca em cada uma das bases estão na Tabela B.3.

Tabela B.2: Palavras-chave e respectivos sinônimos

Palavra-chave	**Sinônimo**	**Relação**
correct spectrum	reliability spectrum; spectrum	Outcome
machine learning	artificial intelligence; method	Intervention
nanoparticle spectrum	nanoparticles; nanothermometers spectrum; nanothermometry	Population

Fonte: Autor (2023)

As bases para realizar a busca foram: American Chemical Society (ACS), Science Direct, Scopus e Web of Science. Elas foram escolhidas devido à sua credibilidade e envolvimento na área de pesquisa. Vale destacar ainda que para cada base de dados foi necessária realizar uma adaptação para realizar a busca. Isso ocorreu porque cada uma tem variações próprias para inserção das strings de busca, bem como no uso de palavras reservadas e conectores lógicos.

Tabela B.3: Strings de Busca nas Bases de Dados

Base	**String de Busca final**
ACS	[[Abstract: "nanoparticle spectrum"] OR [Abstract: "nanoparticles"] OR [Abstract: "nanothermometers spectrum”] OR [Abstract: "nanothermometry”]] AND [[Abstract: "machine learning"] OR [Abstract: "artificial intelligence"] OR [Abstract: "method”]] AND [[Abstract: "correct spectrum"] OR [Abstract: "reliability spectrum"] OR [Abstract: "spectrum”]]
Science Direct	("nanoparticle spectrum"OR "nanoparticles"OR "nanothermometers spectrum"OR "nanothermometry") AND ("machine learning"OR "artificial intelligence"OR "method") AND ("correct spectrum"OR "reliability spectrum")

Scopus	TITLE-ABS-KEY (("nanoparticle spectrum"OR "nanoparticles"OR "nanothermometers spectrum"OR "nanothermometry") AND ("machine learning"OR "artificial intelligence"OR "method") AND ("correct spectrum"OR "reliability spectrum"OR "spectrum")) AND (LIMIT-TO (SUBJAREA , "MEDI")) AND (LIMIT-TO (DOCTYPE , "ar")) AND (LIMIT-TO (LANGUAGE , "English")) AND (LIMIT-TO (EXACTKEYWORD , "Nanoparticles") OR LIMIT-TO (EXACTKEYWORD , "Sensitivity And Specificity")) AND (LIMIT-TO (SRCTYPE , "j")) AND (LIMIT-TO (EXACTSRCTITLE , "Nanomedicine Nanotechnology Biology And Medicine") OR LIMIT-TO (EXACTSRCTITLE , "Physics In Medicine And Biology") OR LIMIT-TO (EXACTSRCTITLE , "Nanomedicine") OR LIMIT-TO (EXACTSRCTITLE , "Physica Medica") OR LIMIT-TO (EXACTSRCTITLE , "Journal Of Nanomedicine And Nanotechnology") OR LIMIT-TO (EXACTSRCTITLE , "Journal Of Biomedical Nanotechnology") OR LIMIT-TO (EXACTSRCTITLE , "Nanomedicine Research Journal")) AND (LIMIT-TO (PUBSTAGE , "final")) AND (LIMIT-TO (OA , "all"))
Web of Science	AB=(("nanoparticle spectrum"OR "nanoparticles"OR "nanothermometers spectrum"OR "nanothermometer") AND ("machine learning"OR "artificial intelligence"OR "method") AND ("correct spectrum"OR "reliability spectrum"OR "spectrum")) OR TI=(("nanoparticle spectrum"OR "nanoparticles"OR "nanothermometers spectrum"OR "nanothermometer") AND ("machine learning"OR "artificial intelligence"OR "method") AND ("correct spectrum"OR "reliability spectrum"OR "spectrum"))

Fonte: Autor (2023)

Critérios de Inclusão e Exclusão Os critérios de inclusão e exclusão são utilizados para incluir e excluir trabalhos, respectivamente. Os de inclusão foram:

- Acesso e estágio de publicação: Gold, Green Published, Published
- Categorias: Nanoscience, Nanotechnology, Chemstry, Medicine and
- Estudos em inglês

- Outros: nanoparticles, sensitivity, specificity
- Tipo de documento: artigo, artigo de revisão
- Técnica de correção espectral de NPs em tecidos
- Títulos de Journals
- Uso do espectro de NPs em área médica
- Uso do espectro de NPs

Enquanto isso, os de exclusão são listados abaixo:

- NPs em materiais não biológicos
- Não usa NPs
- Uso de NPs em área médica, mas sem emprego de espectro
- Uso de NPs em área não médica

Avaliação de Qualidade Para o processo de avaliação de qualidade (após aplicação dos critérios de inclusão e exclusão), foram adotadas as seguintes questões:

1. O artigo descreve o método de forma clara?
2. O artigo traz contribuições sobre nanotermometria?
3. O artigo faz uso de algum tipo de espectro de NPs?
4. O artigo usa algum método para correção do espectro de NPs?

O valor de cada resposta foi 1.0, 0.5 e 0.0 para a resposta Sim, Parcialmente e Não, respectivamente. Vale destacar também que nesta etapa foi realizada a leitura completa dos artigos, ccom foco nas técnicas que estejam utilizando.

Extração de Dados Para realizar o processo de extração de dados foi criado um template de tópicos a serem seguidos, os quais foram:

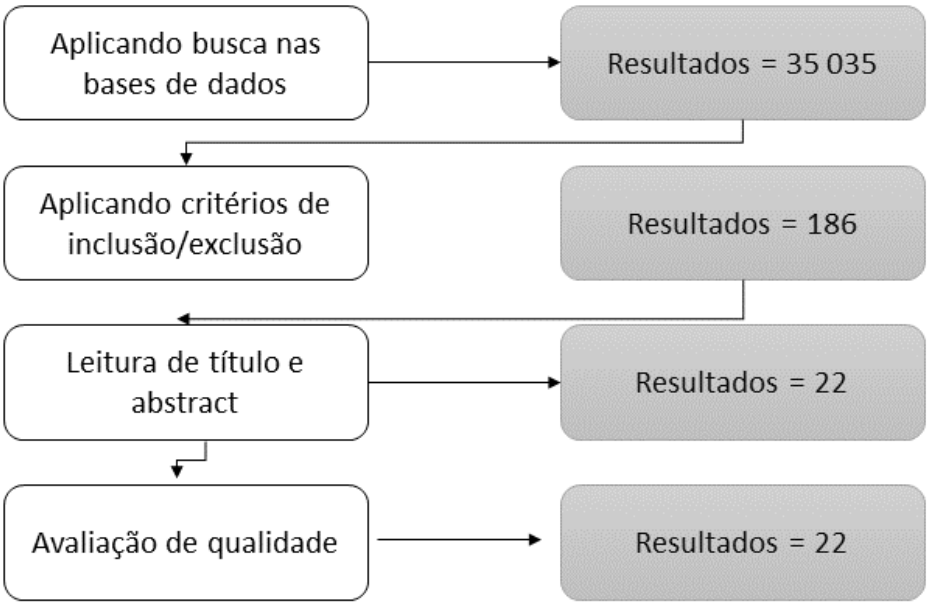
1. Título: String
2. Nome dos Autores: String
3. Ano da Publicação: Inteiro
4. Proposta: Método

Usando a ferramenta Parsifal os artigos foram submetidos, objetivando-se a extração dos dados. Isso foi muito útil para encontrar lacunas presente na literatura.

B.0.3 Resultados

A Figura B.1 mostra o *pipeline* até a obtenção dos artigos finais. Dos 35 035 que formaram a busca nas 4 bases escolhidas, ao final foram selecionados apenas 22.

Figura B.1: Fluxograma com extração de artigos da base



Fonte: Autor (2023)

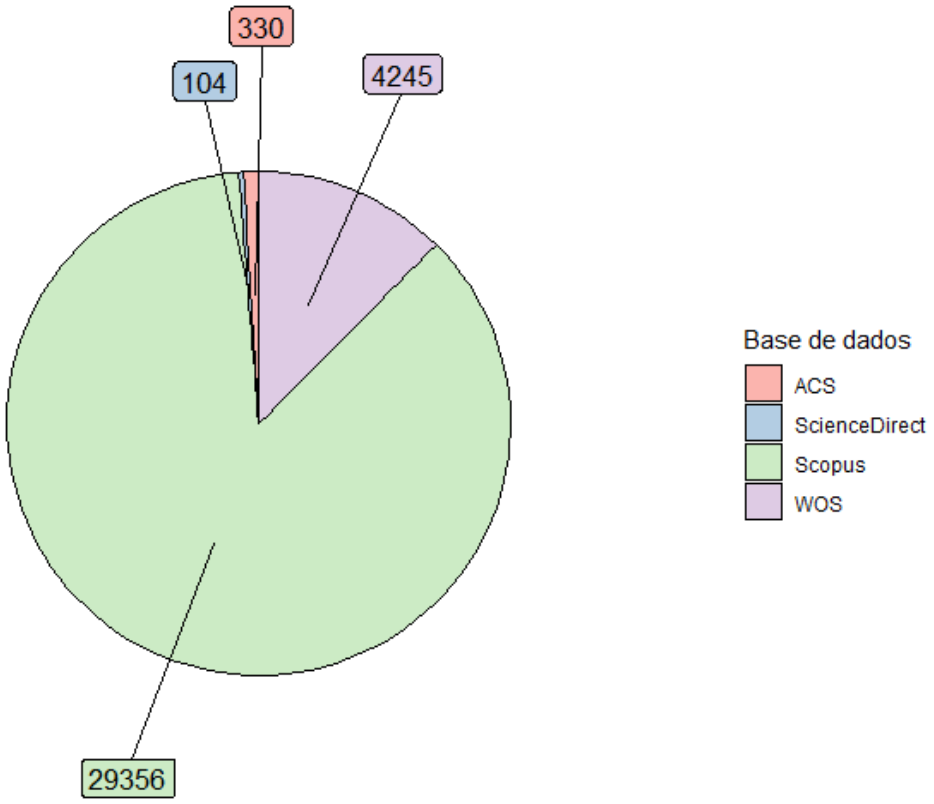
Na primeira busca realizada em cada uma das bases foram obtigos pouco mais de 35 mil artigos. A Figura B.2 exibe a quantidade presente em cada um dos conjuntos de pesquisa usados. Para uma análise bibliométrica foi utilizada a linguagem de programação R, bem como uma de suas bibliotecas muito usadas no contexto de revisões sistemáticas, a bibliometrix (ARIA and CUCCURULLO, 2017).

Pode-se notar que a produção científica anual relacionada ao emprego de NPs tem crescido de forma significativa (Figura B.3). Este gráfico foi feito após o emprego dos critérios de inclusão disponíveis nas próprias bases de dados. Após essa etapa, foi realizada uma leitura de título a abstract buscando selecionar aqueles artigos que tem mais potencial de contribuirem com a pesquisa. A Figura B.1 exibe o antes e o depois da quantidade de artigos em cada uma das bases após esse processo. Na ACS, o número foi de 330 para 14, na Science Direct foi de 104 para 49, na Scopus foi de 29 356 para 12 e na Web of Science esse número foi de 5245 para 111.

B.0.4 Conclusões da RSL

Após a avaliação de qualidade restaram apenas 22 artigos (Figura B.1). Foram escolhidos aqueles com nota maior ou igual a 2.0, cujas notas podem pertencer ao intervalo fechado [0.0;

Figura B.2: Quantidade de artigos em cada base

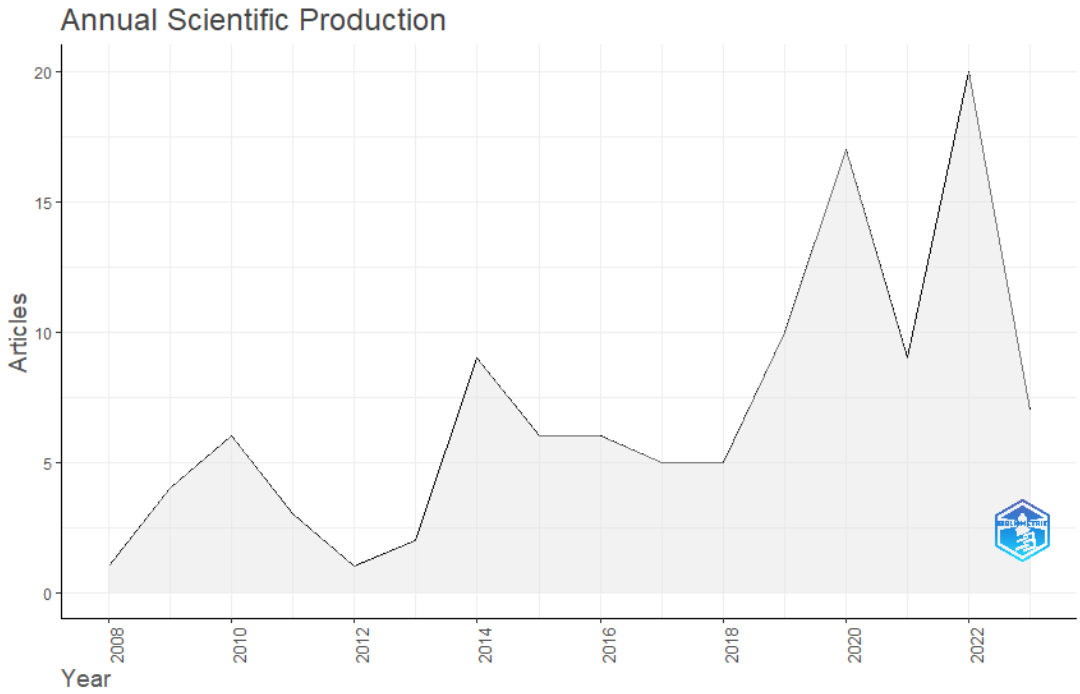


Fonte: Autor (2023)

4.0]. Os artigos tratam sobre o uso de NPs no contexto médico ou para análise espectral. Os seguintes trabalhos (WANG et al., 2017), (ZIEGLER et al., 2020) e (LARSEN et al., 2013) são aqueles cujo foco mais se aproximam dos objetivos da correção espectral de NPs em tecidos biológicos. Nas bases escolhidas não foi encontrado nenhum *paper* que trate sobre correção espectral usando AM no contexto de tecidos biológicos. Isso comprova que há uma lacuna no uso de IA neste contexto.

Assim, a RSL mostrou a existência de *gaps* sobre o uso do espectro de NPs em tecidos biológicos. Dentre eles, podemos destacar a uso das mesmas em aplicações biológicas e médicas (Figura B.3), que mesmo sendo uma área em ascenção, ainda há pesquisas a serem desenvolvidas neste campo. Outrossim, mesmo com o uso difundido da AM nos últimos anos e de técnicas cada vez mais sofisticadas de inteligência artificial, seu emprego em outras áreas como a nanotermometria é um objeto de estudo a ser explorado.

Figura B.3: Produção científica anual



Fonte: Autor (2023)

Apêndice C

Teste estatístico de Wilcoxon

Nesta subseção será abordado acerca do que é Estatística, suas divisões básicas, bem como será abordado sobre testes estatísticos. Este último terá enfoque no Teste de Wilcoxon, descrevendo sua metodologia e aplicabildiade no contexto do problema em questão. Os fundamentos desta subseção é baseada no trabalho de (MONTGOMERY et al., 2000).

C.0.1 Estatística

Pode-se dizer de forma simples que Estatística é uma área da Matemática que tem por objetivo realizar coletas, reduções, manipulações, análise, modelagem de dados em que um dos objetivos mais fundamentais é realizar inferências. Sempre que um experimento envolvendo dados está presente, esse ramo da Matemática é fundamental para uma análise objetiva, rigorosa e confiável. Ela pode ser dividida em duas:

1) **Estatística Descritiva**: utiliza um conjunto de técnicas para planejamento, organização, coleta, resumo, classificação, descrição e análise dos dados por meio de tabelas, gráficos, diagramas, histogramas entre outros recursos visuais. Além disso, o cálculo de certos parâmetros que representam esses dados também costuma estar presente, sendo fundamental para a interpretação coerente das informações. Dentro dela não há a extração de conclusões a partir de dedução ou inferência.

2) **Estatística Indutiva ou Inferêncial**: utiliza um certo conjunto de técnicas em que a partir de uma amostra estabele hipóteses e tira conclusões sobre a população de origem. A formulação dessas previsões fundamenta-se, dentre outras coisas, na teoria das probabilidades e com base na análise e interpretação de parâmetros como a média, moda, desvio padrão, variância entre outros. É nessa subárea que se encontram os Testes Estatísticos, tais como o *t-Student*, *Anova*, *Wilcoxon*.

A Estatística está presente em diversas áreas do conhecimento humano: nas humanidades podemos destacar seu aparecimento na História, Ciências Sociais, Antropologia - para análise

de dados históricos, fenômenos sociais, entre outros, usando a Estatística Descritiva e a Estatística Inferencial como testes de hipóteses; na área de exatas podemos destacar seu uso frequente nas Engenharias, na Matemática e na Aprendizagem de Máquina, que também fazem uso das técnicas de Estatística Descritiva e Indutiva.

Na Aprendizagem de Máquina no contexto de problemas de Regressão os testes são importantes, pois asseguram se os valores preditos são confiáveis em termos de precisão e exatidão, entre outros parâmetros. Por precisão queremos dizer na capacidade que o modelo tem de prever valores similares dentro de uma faixa aceitável. E por exatidão pode-se entender como o quanto o modelo de fato está acertando em relação ao parâmetro real.

C.0.2 Teste de Hipótese

Os métodos de teste de hipótese e de intervalo de confiança baseados em amostras aleatórias provienientes de uma população normal são chamados de **métodos paramétricos**. Assim, elas dependem da suposição de uma **normalidade** na distribuição dos dados. Contudo, dependendo do contexto do problema que está sendo tratado essa suposição pode não ser muito útil. É nesse ambiente que surgem os métodos chamados de **não-paramétricos** ou **livres da ditribuição**.

De modo geral eles têm um nível de significância α ou de 100 * (1 - α)%. De acordo com (MONTGOMERY et al., 2000) uma das vantagens destes tipos de técnicas está na não necessidade dos dados serem quantitiativos, ou seja, eles podem ser usados em dados categóricos (como o par sim-não) ou dados ordenados, bem como são mais rápidos e fáceis de serem feitos. Dentre os testes existentes, merecem destaque: o Teste Qui-Quadrado de Pearson, o Teste de Kruskal-Wallis, o Teste dos Sinais e o Teste de Wilcoxon. Iremos focar na descrição deste último, uma vez que ele será utilizado neste trabalho.

C.0.3 *Overview* do Teste de Wilcoxon

Este teste foi proposto pelo Químico e Estatístico Frank Wilcoxon (1892 - 1965) em seu trabalho *Probability Tables for Individual Comparisons by Ranking Methods* e publicado em 1947 (WILCOXON, 1947). Anos depois ficou uma técnica popularizada no livro de Sidney Siegel (1916 - 1961) intitulado *Nonparametric Statistics For The Behavioral Sciences* em 1956 (SIDNEY, 1957). De acordo com (MONTGOMERY et al., 2000) este teste pode ser dividido em dois:

1) **Teste de Wilcoxon do Posto Sinalizado**: ele usa somente dois sinais mais e menos das diferenças entre as observações e a mediana $\bar{\mu_0}$.

2) **Teste de Wilcoxon da Soma dos Postos**: este tipo pode ser usado para testar a hipótese H_0: $\mu_1 = \mu_2$.

C.0.4 Teste de Wilcoxon do Posto Sinalizado

Este teste é aplicado no caso em que as distribuições são **contínuas simétricas**. A partir desse pressuposto, sabe-se que a média é igual a mediana e essa técnica pode ser usada para testar H_0: $\mu = \mu_0$.

Consideremos que X_1, X_2, ..., X_n seja uma amostra aleatória cuja origem seja uma distribuição contínua e simétrica com valores de media e media iguais a μ. O teste de Wilcoxon do Posto Sinalizado pode ser organizado da seguinte maneira:

1) Calcule as diferenças X_i - μ_0, com i = 1, 2, ..., n.

2) Ordene as diferenças absolutas $|X_i$ - $\mu_0|$, com i = 1, 2, ..., n em ordem crescente e então forneça aos postos os sinais de suas diferenças correspondentes.

3) Faça W^+ ser a soma dos postos positivos, W^- ser o valor absoluto da soma dos postos negativos e W = min(W^+, W^-)

4) Se

   4.1) Teste bilateral: Caso a hipótese alternativa for H_1: $\mu \neq \mu_0$, então o valor observado da estatistica w $\leq w_a^*$ e a hipótese nula H_0: μ μ_0 será rejeitada.

   4.2) Teste unilateral: Caso a hipótese alternativa for H_1: $\mu > \mu_0$, rejeite H_{0+}: $\mu = \mu_0$, se $w^- \neq w_\alpha^*$.

   4.3) Teste unilateral: Caso a hipótese alternativa for H_1: $\mu < \mu_0$, rejeite H_{0+}: $\mu = \mu_0$, se $w^+ \neq w_\alpha^*$.

A Tabela C.1 contém os valores críticos de W, de significância de α = 0.1, α = 0.02, α = 0.01 para o teste bilateral; e α = 0.05, 0.025, 0.01, 0.05.

Caso o número de amostras seja grande, isto é, n > 20, então defini-se que W^+ (ou W^-) tem aproximadamente uma distribuição normal com média:

$$\mu_{W^+} = \frac{n(n+1)}{4} \tag{C.1}$$

e variância:

$$\sigma^2_{W^+} = \frac{n(n+1)(2n+1)}{24} \tag{C.2}$$

Deste modo, o teste de H_0: $\mu = \mu_0$, pode estar baseado da seguinte forma:

$$Z_O = \frac{\frac{W^+ - n(n+1)}{4}}{\sqrt{\frac{n(n+1)(2n+1)}{24}}} \tag{C.3}$$

Tabela C.1: Valores Críticos para o Teste de Wilcoxon do Posto com Sinais

n^{α} \ α	0.10 0.05	0.05 0.025	0.02 0.01	0.01 0.005
4				
5	0			
6	2	0		
7	3	2	0	
8	5	3	1	0
9	8	5	3	1
10	10	8	5	3
11	13	10	7	5
12	17	13	9	7
13	21	17	12	9
14	25	21	15	12
15	30	25	19	15
16	35	29	23	19
17	41	34	27	23
18	47	40	32	27
19	53	46	37	35
20	60	52	43	37
21	67	58	49	42
22	75	65	55	48
23	83	73	62	54
24	91	81	69	61
25	100	89	76	68

Fonte: Autor.