

INSTITUTO FEDERAL DE EDUCAÇÃO, CIÊNCIA E TECNOLOGIA DE MINAS  
GERAIS – *CAMPUS* FORMIGA  
BACHARELADO EM ENGENHARIA ELÉTRICA

Caio César Branco Nunes

**ESTIMATIVA DO ESTADO DE CARGA DE BATERIAS EM VEÍCULOS  
ELÉTRICOS UTILIZANDO INTELIGÊNCIA COMPUTACIONAL**

FORMIGA – MG

2022

CAIO CÉSAR BRANCO NUNES

**ESTIMATIVA DO ESTADO DE CARGA DE BATERIAS EM VEÍCULOS  
ELÉTRICOS UTILIZANDO INTELIGÊNCIA COMPUTACIONAL**

Trabalho de Conclusão de Curso apresentado ao Curso de Bacharelado em Engenharia Elétrica do Instituto Federal de Educação, Ciência e Tecnologia de Minas Gerais – *Campus Formiga* como requisito para obtenção do título de bacharel em Engenharia Elétrica.

Orientador: Prof. Me. Marco Antônio Silva Pereira

Coorientador: Me. Carlos Antônio Rufino Júnior

FORMIGA – MG

2022

Nunes, Caio César Branco  
N972e            Estimativa do estado de carga de baterias em veículos elétricos utilizando  
inteligência computacional / Caio César Branco Nunes -- Formiga : IFMG, 2022.  
70p. : il.

Orientador: Prof. MSc. Marco Antônio Silva Pereira  
Coorientador: Prof. MSc. Carlos Antônio Rufino Júnior  
Trabalho de Conclusão de Curso – Instituto Federal de Educação,  
Ciência e Tecnologia de Minas Gerais – *Campus* Formiga.

1. Veículos elétricos. 2. Inteligência computacional. 3. Bateria de ions-lítio.  
I. Pereira, Marco Antônio Silva. II. Rufino Júnior, Carlos Antônio. III. Título.

CDD 621.3

CAIO CÉSAR BRANCO NUNES

**ESTIMATIVA DO ESTADO DE CARGA DE BATERIAS EM VEÍCULOS  
ELÉTRICOS UTILIZANDO INTELIGÊNCIA COMPUTACIONAL**

Trabalho de Conclusão de Curso apresentado ao Curso de Bacharelado em Engenharia Elétrica do Instituto Federal de Educação, Ciência e Tecnologia de Minas Gerais – *Campus Formiga* como requisito para obtenção do título de bacharel em Engenharia Elétrica.

Orientador: Prof. Me. Marco Antônio Silva Pereira

Coorientador: Me. Carlos Antônio Rufino Júnior

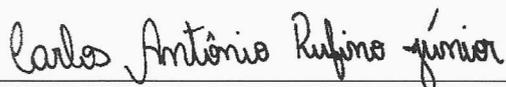
Avaliado em 20 de janeiro de 2022.

Nota: 84,0

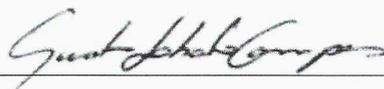
**BANCA EXAMINADORA**



Prof. Me. Marco Antônio Silva Pereira (Orientador)



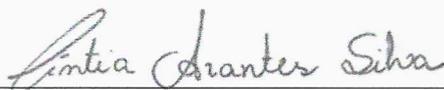
Me. Carlos Antônio Rufino Júnior (Coorientador)



Prof. Dr. Gustavo Lobato Campos



Prof. Dr. Renan Souza Moura



Prof<sup>ª</sup>. Ma. Cíntia Arantes Silva

## AGRADECIMENTOS

Aos meus pais, Alessandra e Moacir (*in memoriam*), que em todos os momentos me apoiaram e me motivaram traçar o caminho deste sonho em me graduar em engenharia. As minhas irmãs Maria Clara, Vitória e Mirela por estarem sempre ao meu lado e poder ser um exemplo a elas, que de onde viemos, o estudo é sempre o melhor caminho.

A toda minha família, como segundas mães, tia Adriana e tia Andrea, que me criaram e deram amor e carinho, ao meu tio Allysson que foi um espelho para mim durante esta jornada. Ao meu tio Carlos que admiro muito seu esforço e trabalho. As minhas primas que considero irmãs, Ana Luiza, Julia e Maria Fernanda. Agradeço também meu tio Nem e que sempre esteve ao meu lado.

Ao meu orientador Me. Marco Antônio, por todo apoio e ensinamentos para concretizar este trabalho.

Ao meu mentor e um dos melhores amigos, Rufino, que considero um irmão mais velho, que além de ter me orientado neste trabalho, foi meu companheiro de república.

Ao meu braço direito Lucas, que foi a grande amizade que formei durante a graduação, agradeço demais pelo amigo é.

Ao Ravel, que é um dos meus melhores amigos e companheiro durante o curso. Junto dos meus amigos da Lambda que foram os melhores companheiros de república, que fazem uso constante de entorpecentes, Rafael, Danilo, Leandro, Chanry e o mala do Marco Tulio. Aos meus amigos Hudson e Igor. Aos meus amigos do “projêto”.

Aos meus professores por passarem todos os ensinamentos do curso e para a vida.

A todos que colaboraram em minha jornada.

*"A fé na vitória tem que ser inabalável"*

*O Rappa*

## RESUMO

Este trabalho visa estimar o estado de carga em baterias de íons-lítio presentes em veículos elétricos, utilizando inteligência computacional. Apesar de existirem várias formas de aferição da quantidade de energia presente nestas baterias, muitas são imprecisas e não apresentam os resultados em tempo real. Diante disso, com um amplo banco de dados, *BATTERY AND HEATING DATA IN REAL DRIVING CYCLES*, dos autores Matthias Steinstraeter, Johannes Buberger, Dimitar Trifonov, que foram coletados por sensores instalados no carro BMW I3 (60hA) em viagens de teste, foi possível implementar modelos para previsão do estado de carga associado ao uso da bateria do veículo em estudo. Estes, foram desenvolvidos por técnicas de inteligência computacional no *software Jupyter*, tais como, Regressão Linear Múltipla, Árvore de Decisão e Redes Neurais Artificiais. As avaliações por métricas estatísticas mostram que os resultados obtidos são satisfatórios, apontando alto índice de acertos nas previsões realizadas. Um dos melhores modelos obteve Erro Médio Absoluto igual a 0,8106 e R-quadrado de 0,9924. Este foi implementado considerando que, a partir de uma análise exploratória dos dados, as grandezas que possuem maior relação com o estado de carga de uma bateria em veículos elétricos são: tensão da bateria (*battery voltage*), corrente da bateria (*battery current*), temperatura ambiente (*ambient temperature*), elevação (*elevation*) e velocidade (*velocity*).

**Palavras Chave:** Veículos elétricos, Inteligência Computacional, Bateria de íons-lítio.

## ABSTRACT

This work aims to estimate the state of charge in lithium-ion batteries presented in electric vehicles, using computational intelligence. Despite the amount of energy presented in these batteries, many are inaccurate and do not present real-time results. From that, with a large database, BE HEATING DATA IN DRIVING CY CY, from the authors, Matthias Buberger, Diitar Trifonov, who were able to BMW I3 (60hA), by real sensors, which they were able to test, it was possible (60hA) in travel of test. implementation of models to predict the state of charge associated with the use of the vehicle's battery under study. These were developed by computational intelligence techniques in Jupyter software, such as Multiple Linear Regression, Decision and Artificial Neural Networks. Estimates by measurements show that the results obtained from hits, a high rate in those presented. One of the best models had Mean Absolute Error equal to 0.8106 and R-squared of 0,9924. This was implemented considering that, from an analysis of the data, the quantities that have the greatest relationship with the state of charge of a battery in electric vehicles are: battery voltage, battery current, speed environment (room temperature), speed (elevation) and temperature (speed).limp

**Key words:** Electric Vehicles, Computational Intelligence, Li-Ion Battery.

## LISTA DE ILUSTRAÇÕES

Figura 1: Estrutura das baterias de íons de lítio. ....	19
Figura 2: Neurônio Biológico.....	25
Figura 3: A transmissão de impulsos nervosos entre neurônios.....	26
Figura 4: Passagem do impulso. ....	26
Figura 5: Representação de um neurônio artificial.....	27
Figura 6: Função Limiar. ....	29
Figura 7: Função Linear. ....	29
Figura 8: Função Rampa. ....	30
Figura 9: Função Sigmóide.....	30
Figura 10: Função tangente hiperbólica. ....	31
Figura 11: Exemplo de rede camada única.....	32
Figura 12: Exemplo de rede multicamadas. ....	32
Figura 13: Exemplo de rede multicamadas. ....	33
Figura 14: Diagrama de Blocos do Treinamento Supervisionado.....	34
Figura 15: Diagrama de Blocos do Treinamento não-Supervisionado. ....	34
Figura 16: Representação geométrica dos parâmetros $\beta_i$ e $\beta_0$ .....	36
Figura 17: Hiperplano referente a multivariáveis independentes .....	37
Figura 18: Exemplo de Estrutura Árvore de decisão.....	38
Figura 19: Fluxograma referente a metodologia utilizada.....	41
Figura 20: Bibliotecas auxiliares para desenvolvimento da rotina computacional. ....	42
Figura 21: Veículo BMW I3 (60hA). ....	43
Figura 22: Gráficos da distribuição de dados das variáveis com outliers. ....	45
Figura 23: Gráficos da distribuição de dados constantes. ....	46
Figura 24: Correlação linear entre SoC e as demais variáveis. ....	47
Figura 25: Correlação não linear entre SoC e as demais variáveis. ....	48
Figura 26: Curva de dados reais e valores previstos (7 variáveis). ....	54
Figura 27: Curva de dados reais e valores previstos (6 camadas de prof.).....	56
Figura 28: Curva de dados reais e valores previstos (5 neurônios nas camadas ocultas).....	58
Figura 29: Gráfico MAE. ....	59
Figura 30: Gráfico MSE. ....	59
Figura 31: Gráfico RMSE. ....	60

Figura 32: Gráfico $R^2$ .....	60
Figura 33: Histograma das métricas de validação .....	61
Figura 34: Curva de dados reais e valores previstos pelo Modelo RL-I. ....	67
Figura 35: Curva de dados reais e valores previstos pelo Modelo RL-II.....	67
Figura 36: Curva de dados reais e valores previstos pelo Modelo AD-I.....	68
Figura 37: Curva de dados reais e valores previstos pelo Modelo AD-III.....	68
Figura 38: Curva de dados reais e valores previstos pelo Modelo RNA-II.....	69
Figura 39: Curva de dados reais e valores previstos pelo Modelo RNA-III. ....	69
Figura 40: Curva de dados reais e valores previstos pelo Modelo RNA-IV. ....	70

## LISTA DE TABELAS

Tabela 1: Métodos para estimação do estado de carga (SoC) .....	20
Tabela 2: Modelos Regressão Linear Múltipla .....	52
Tabela 3: Métricas de avaliação para Modelos de Regressão Linear Múltipla.....	53
Tabela 4: Resultados do modelo de Árvore de Decisão.....	55
Tabela 5: Resultados métricas de avaliação Modelo Árvore de Decisão.....	55
Tabela 6: Resultados do modelo de Redes Neurais Artificias .....	57
Tabela 7: Resultados métricas de avaliação Modelo Redes Neurais Artificias .....	57

## LISTA DE SIGLAS E ABREVIATURAS

IA	<i>Inteligência Artificial</i>
IC	<i>Inteligência Computacional</i>
MAE	<i>Mean Absolute Error</i>
MSE	<i>Mean Square Erro</i>
$R^2$	<i>R-quadrado</i>
RMSE	<i>Mean Squared Error</i>
RNA	<i>Redes Neurais Artificiais</i>
SOC	<i>Stage of Charge</i>
SOH	<i>Stage of Health</i>

## SUMÁRIO

<b>1. INTRODUÇÃO</b> .....	<b>15</b>
<b>1.1. Hipóteses e Motivação</b> .....	<b>16</b>
<b>1.2. Objetivos</b> .....	<b>17</b>
1.2.1. <i>Objetivo geral</i> .....	<i>17</i>
1.2.2. <i>Objetivos específicos</i> .....	<i>17</i>
<b>1.3. Organização do trabalho</b> .....	<b>18</b>
<b>2. REFERENCIAL TEÓRICO</b> .....	<b>19</b>
<b>2.1. Estimativa do estado de carga de baterias em veículos elétricos</b> .....	<b>19</b>
<b>2.2. Histórico da Inteligência Artificial</b> .....	<b>21</b>
2.2.1. <i>Anos 40</i> .....	<i>22</i>
2.2.2. <i>Anos 50</i> .....	<i>22</i>
2.2.3. <i>Anos 60</i> .....	<i>23</i>
2.2.4. <i>Anos 70 a 80</i> .....	<i>23</i>
2.2.5. <i>A partir dos anos 90</i> .....	<i>23</i>
<b>2.3. Redes Neurais Artificiais</b> .....	<b>24</b>
2.3.1. <i>Neurônio Biológico</i> .....	<i>24</i>
2.3.2. <i>Neurônio Artificial</i> .....	<i>27</i>
2.3.3. <i>Funções de Ativação</i> .....	<i>28</i>
2.3.3.1. <i>Limiar</i> .....	<i>28</i>
2.3.3.2. <i>Linear</i> .....	<i>29</i>
2.3.3.3. <i>Rampa</i> .....	<i>30</i>
2.3.3.4. <i>Sigmóide</i> .....	<i>30</i>
2.3.3.5. <i>Tangente Hiperbólica</i> .....	<i>31</i>
2.3.4. <i>Arquitetura de RNA's</i> .....	<i>31</i>
2.3.4.1. <i>Redes feedforward de Camada Única</i> .....	<i>31</i>
2.3.4.2. <i>Redes feedforward de Múltiplas Camadas</i> .....	<i>32</i>
2.3.4.3. <i>Redes recorrentes</i> .....	<i>33</i>
2.3.5. <i>Treinamento</i> .....	<i>33</i>
<b>2.4. Regressão Linear</b> .....	<b>34</b>
2.4.1. <i>Regressão Linear Simples</i> .....	<i>35</i>
2.4.2. <i>Regressão Linear Múltipla</i> .....	<i>36</i>
<b>2.5. Árvore de Decisão</b> .....	<b>37</b>
<b>2.6. Métricas de Avaliação</b> .....	<b>39</b>
<b>3. METODOLOGIA</b> .....	<b>41</b>
<b>3.1. Introdução</b> .....	<b>41</b>
<b>3.2. Base de Dados</b> .....	<b>43</b>
3.2.1. <i>Análise exploratória do banco dados</i> .....	<i>45</i>

3.3.	Pseudocódigos dos modelos implementados .....	49
3.3.1.	<i>Regressão Linear Múltipla</i> .....	49
3.3.2.	<i>Árvore de Decisão</i> .....	50
3.3.3.	<i>Redes Neurais Artificiais</i> .....	51
4.	RESULTADOS E DISCUSSÕES .....	52
4.1.	Introdução .....	52
4.2.	Regressão Linear .....	52
4.3.	Árvore de Decisão .....	54
4.4.	Redes Neurais Artificiais.....	56
4.5.	Considerações finais .....	58
5.	CONCLUSÃO.....	62
	REFERÊNCIAS BIBLIOGRÁFICAS .....	64
	ANEXO I – Curvas de comparação entre os dados reais e previstos pelos modelos complementares. ....	67
	ANEXO I (a) – Modelos de Regressão Linear Múltipla. ....	67
	ANEXO I (b) – Modelos de Árvore de Decisão. ....	68
	ANEXO I (c) – Modelos de Redes Neurais Artificiais.....	69

## 1. INTRODUÇÃO

A frota de veículos utilizada no mundo atualmente é, de forma predominante, constituída por motores a combustão. No entanto, pode-se notar avanços significativos na produção de veículos movidos a eletricidade, os quais apresentam potenciais transformações na indústria automobilística e, sobretudo, nos conceitos de mobilidade urbana.

Um exemplo desta tendência, pode ser verificado pela proposta apresentada por uma comissão europeia, de acordo com AQUINO (2021), que visa a retirada de carros a combustão em um prazo de 15 anos. Tal proposta prevê a proibição da venda de veículos que emitem gás carbônico a partir de 2035, entrando em vigor com o Acordo de Paris (BRASIL,2017), sob o qual se comprometeram a tomar métodos para limitar o aumento da temperatura média global.

A utilização dos veículos elétricos apresenta inúmeras vantagens, destacando-se: a baixa emissão de poluentes e de gases de efeito estufa em seu funcionamento, reduzida emissão de ruídos e elevado rendimento energético. Apesar das vantagens apresentadas, as tecnologias empregadas em veículos elétricos ainda possuem custos elevados e as atuais baterias utilizadas nestes modelos são pouco eficientes, promovendo autonomia limitada. Neste contexto, são demandados estudos sobre as características de operação de tais baterias.

Estas baterias (Níquel e Chumbo), apresentam eficiência relativamente baixa, em torno de 15 a 18%, entretanto, as baterias de íons de lítio em veículos elétricos obtiveram uma boa performance, chegando a atingir uma eficiência entre 60 a 70% (SHAREEF; ISLAM; MOHAMED, 2016). Desta maneira, é de extrema importância iniciarem investimentos e estudos nesta área de pesquisa para o desenvolvimento e aperfeiçoamento das mesmas, convergindo também com a implementação de políticas públicas de comercialização.

As baterias de íons de lítio de veículos elétricos têm como um dos principais parâmetros o estado de carga, traduzido do inglês *state of charge* (SOC). Este representa o nível de carga da bateria, ou seja, quando o estado de carga da bateria é 100% quer dizer que a bateria está totalmente carregada e quando o estado de carga é 0% a bateria está totalmente descarregada. Esse parâmetro específico tem impactos na operação das baterias e estimar seu estado pode evitar acidentes, que podem ser desde as pequenas chamas até explosões e incêndios de grandes proporções. Além disso, uma estimativa adequada do estado de carga de uma bateria é fundamental no gerenciamento da autonomia do veículo.

Portanto estimar o estado de carga das baterias, que pode ser entendido como a quantidade de energia remanescente, é crucial para que se viabilize de maneira eficiente e segura a utilização das mesmas. Nos dias de hoje, realizar um método de estimação preciso é um desafio, sendo possível utilizar diversas metodologias baseadas, principalmente, em aprendizado de máquina. (RAHIMI-EICHI et al., 2013).

Diante desse cenário, o objetivo do trabalho é prever o estado de carga para uma bateria de íon de lítio considerando diferentes variáveis medidas em um veículo elétrico. Para alcançar esse resultado, propõe-se o desenvolvimento de modelos baseados em técnicas de inteligência computacional. Este tipo de metodologia é idealmente desenvolvido para resolução de problemas em diversas situações da vida real, a partir de aprendizado e modelamento de relações entre entradas e saídas de dados que podem ser não-lineares e complexos. Sendo assim, é possível realizar generalizações e inferências, revelar relacionamentos, padrões, previsões ocultas e modelar dados altamente voláteis. Neste contexto, utilizando dados de variáveis medidas no veículo BMW I3 (60hA), espera-se que os modelos desenvolvidos sejam capazes de estimar com precisão o estado de carga da bateria desse veículo considerando dados reais.

### **1.1. Hipóteses e Motivação**

Com o acordo de Paris em 2015 (BRASIL, 2017), que resguarda a diminuição de gases que contribuem para o efeito estufa, surge-se diversas possibilidades na área de estudos relacionados a veículos elétricos, estimulando, desta maneira, o desenvolvimento de pesquisas e tecnologia nesta área.

Nos automóveis a combustão, a gasolina presente no tanque, pode ser aferida com sensores, como por exemplo, o de nível. Entretanto, em veículos elétricos os métodos para medir a quantidade de energia restante das baterias, o estado de carga (SoC), não são tão simples. Isto se justifica, pelo fato de a bateria estar submetida a carga e descarga constantemente durante o uso do veículo. Uma vez que o estado de carga não pode ser medido por algum sensor específico com precisão, torna-se necessário estudos para desenvolvimento de modelos que possam estimá-lo para extrair informações adequadas sobre a autonomia do veículo elétrico, ou até mesmo da saúde da bateria.

Diante das considerações realizadas, são expostas as hipóteses deste trabalho:

- Seria possível estimar o estado de carga da bateria em veículos elétricos com auxílio de inteligência computacional?

Caso seja possível, surge o seguinte questionamento:

- A precisão obtida nesta estimativa, apresentaria uma confiabilidade satisfatória?

## **1.2. Objetivos**

Nesta seção são apresentados os objetivos deste trabalho de conclusão de curso. Para melhor compreensão, foram divididos em duas subseções, Objetivo Geral e Objetivos Específicos, cujas descrições são apresentadas a seguir.

### ***1.2.1. Objetivo geral***

Este trabalho tem como objetivo a estimação de carga de baterias de íons de lítio no veículo elétrico BMW I3 (60hA) através da análise de dados reais, os quais foram coletados em estudo desenvolvido por Matthias (2020). Esta análise terá o auxílio de técnicas de inteligência computacional, especificadamente com a implementação dos métodos de Regressão Linear, Árvore de decisão e Redes Neurais.

### ***1.2.2. Objetivos específicos***

Para alcançar o objetivo geral, consideram-se os seguintes objetivos específicos:

- Realizar um levantamento bibliográfico sobre inteligência computacional e aprendizado de máquina;
- Estudar o comportamento de baterias de íons de lítio;
- Analisar os dados provenientes da medição de grandezas relacionadas à operação de um veículo elétrico, conforme Matthias (2020);
- Implementar bibliotecas e funções para tratamento dos dados em estudo;

- Desenvolver e testar modelos de estimação do estado de carga a partir de técnicas de inteligência computacional: Regressão linear, Árvore de Decisão e Redes Neurais Artificiais;
- Validar modelos, implementando funções de métricas estatísticas.

### **1.3. Organização do trabalho**

Este trabalho de conclusão de curso, foi dividido em 5 capítulos. Em primeiro plano, encontra-se no Capítulo 1, a introdução, objetivo e motivação. Posteriormente é apresentado, no Capítulo 2, o referencial teórico abordando os principais conceitos que foram pilares para realizar as implementações do trabalho. No terceiro capítulo, são abordados os procedimentos e métodos utilizados para o desenvolvimento do algoritmo de previsão do estado de carga (SoC) de baterias de um veículo elétrico. Já no Capítulo 4, são expostos os resultados obtidos dos testes realizados com o algoritmo de estimação. Por fim, no Capítulo 5, são apresentadas as conclusões retiradas deste trabalho.

## 2. REFERENCIAL TEÓRICO

### 2.1. Estimativa do estado de carga de baterias em veículos elétricos

A crescente busca por uma fonte energética para a substituição do combustível fóssil, alavancou as pesquisas tecnológicas e científicas para as baterias. Em primeiro instante, houve um aumento no mercado para aplicações das baterias “recarregáveis” em notebooks, celulares, desfibriladores, câmeras digitais, etc. Atualmente, as pesquisas convergem para os estudos de viabilização do uso de baterias em automóveis híbridos e elétricos.

Com o resultado destas pesquisas, obteve-se maior interesse nas baterias a base de íons de lítio, que possuem maior densidade de energia (Wh/kg, ou watt hora por quilograma) e maior densidade de potência (W/kg, ou watts por quilograma). Sendo assim, podem oferecer a mesma quantidade de energia do que as baterias de chumbo-ácido, porém com um tamanho menor e um peso mais leve.

As baterias de íons de lítio não possuem lítio metálico como eletrodo, sendo utilizados íons de lítio presentes no eletrólito na forma de sais de lítio ( $\text{Li}^+$ ) dissolvidos. Para a circulação de corrente na bateria, como mostra a Figura 1, o dispositivo possui polos denominados como ânodo (eletrodo feito de cobre) e cátodo (eletrodo feito de alumínio), que são conectados eletroliticamente por um meio não aquoso ou sólido, sendo o circuito externo curto circuitado. O lítio é oxidado no eletrodo negativo (material reduzido) e os íons da decomposição eletrolítica percorrem circuito para o eletrodo positivo (material oxidado) (SANTOS, 2018).

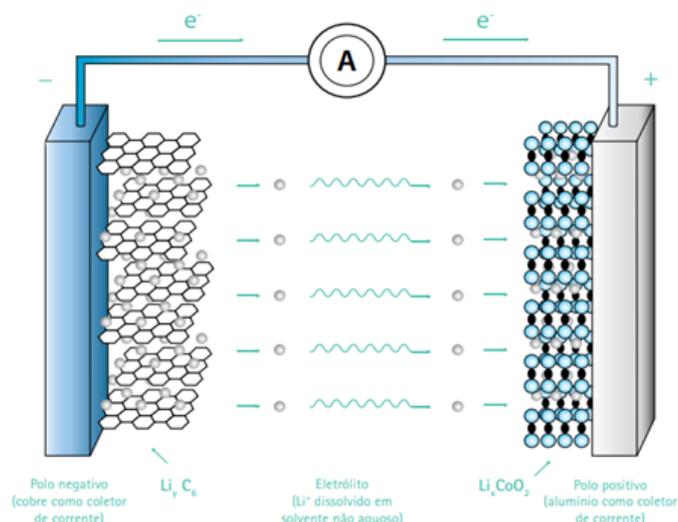


Figura 1: Estrutura das baterias de íons de lítio.

Fonte: (SANTOS, 2018).

As vantagens principais das baterias de íons de lítio, segundo (Pinho, 2014), são:

- Não possuem necessidade de manutenção;
- Possuem ciclo de vida extenso;
- Operam em vasta faixa de temperatura;
- Versatilidade em capacidade de carga.

Em contrapartida, apresentam desvantagens como, custo relativamente alto, necessita de um dispositivo de proteção, degrada em temperaturas elevadas (FILHO, 2017).

Um dos parâmetros a serem observados em uma bateria é seu estado de carga (SoC), o qual pode ser definido como um medidor de quantidade de energia presente na bateria. Em comparação aos automóveis a base de combustíveis fósseis, o SoC tem a mesma função que um medidor de gasolina de um tanque, porém afere-se a quantidade atual de carga presente em uma célula até o seu estado descarregado (HANNAN, 2017). Por ser uma medida adimensional, o estado de carga apresenta-se com valores entre 0 (descarregado) e 1 (carregado), ou também, pode ser apresentado em porcentagem.

Este parâmetro é de extrema importância para o monitoramento da bateria, sendo possível definir os ciclos de carga e quantidade de energia restante na bateria, o que permite, desta forma, medir a autonomia do veículo. Diferentemente dos automóveis de combustão interna, não é possível aferir diretamente com um sensor, fato que resulta na necessidade de implementação de outras técnicas. Existem várias metodologias para estimar o estado de carga, como apresentado na Tabela 1, as quais apresentam vantagens e desvantagens. Neste trabalho, serão aplicados métodos computacionais, como RNA's (Redes Neurais Artificiais), Árvore de decisão e Regressão Linear Múltipla.

Tabela 1: Métodos para estimação do estado de carga (SoC)

Método	Área de Aplicação	Vantagens	Desvantagens
<b>Teste de descarga</b>	Usada para determinação da capacidade no início da vida	Fácil e precisa; independente do SoH.	Offline, modifica o estado da bateria, perda de energia.
<b>Integral da corrente (Coulomb Counting)</b>	Todos os tipos de baterias, diversas aplicações	Preciso se um número suficiente de recalibrações forem feitas e com boas medições de corrente.	Sensível a reações parasitas, necessita de pontos de recalibração regulares.
<b>OCV (Open Circuit Voltage)</b>	Chumbo, Lítio, Zn/Br	Online, barato, previsão do OCV.	Necessita de longos períodos de descanso. (Corrente = 0)

Tabela 1 (cont.): Métodos para estimação do estado de carga (SoC)

Método	Área de Aplicação	Vantagens	Desvantagens
EMF	Chumbo, Lítio	Online, barato, previsão do EMF.	Necessita de longos períodos de descanso. (Corrente = 0)
Modelo Linear	Chumbo, aplicações fotovoltaicas	Online, fácil.	Necessita de dados de referência para ajustar os dados.
Espectroscopia de impedância	Todos os sistemas	Proporciona informações sobre o SoH e qualidade de energia.	Sensível a temperatura, custo elevado.
Resistência interna	Chumbo, Ni/Cd	Proporciona informações sobre o SoH, possibilidade de medições online.	Preciso, mas apenas para curtos intervalos de tempo.
Redes neurais artificiais	Todos os sistemas de baterias	Online	Precisa de dados de uma bateria similar para treinamento da rede.
Lógica fuzzy	Todos os sistemas de baterias	Online	Custo computacional elevado em aplicações no mundo real.
Filtros de Kalman	Todos os sistemas de baterias, Sistemas fotovoltaicos, aplicações dinâmicas.	Online, dinâmico	Dificuldade em implementar o algoritmo que considere todas as características do sistema real.

Fonte: (SANTOS, 2019).

## 2.2. Histórico da Inteligência Artificial

Os conceitos fundamentais sobre Inteligência Artificial (IA) foram pesquisados a partir da década de 1940 com o objetivo de sequenciar, analisar e reproduzir o comportamento do cérebro humano. No entanto, somente no ano de 1955, surgiu-se o termo “Inteligência Artificial”, dito por John McCarthy, que era um jovem professor de matemática do *Dartmouth College*. Esse termo não estava relacionado à construção de uma Inteligência Artificial propriamente dita, mas procurava meios alternativos para a utilização do computador que ainda era um protótipo.

Em um primeiro momento, o termo “*Inteligência Artificial*” causou uma certa confusão na sociedade devido à falta de definição de métodos e objetivos dessa técnica computacional. A palavra “*Inteligência*” provoca dúvidas e questionamentos sobre a capacidade de processamento dessa tecnologia e o seu uso era visto como exagerado para algumas aplicações. Diante desse cenário, diversos pesquisadores tentaram definir o que é inteligência artificial de acordo com suas perspectivas. McCarthy e Hayes definiram a Inteligência Artificial como “*uma máquina é inteligente se ela é capaz de solucionar uma classe de problemas que requerem*

*inteligência para serem solucionados por seres humanos*”. Uma outra definição foi proposta por Barr e Feigenbaum, que definiram a IA como sendo “*a parte da ciência da computação que compreende o projeto de sistemas computacionais que exibam características associadas, quando presentes no comportamento humano, inteligência*”(BARR, 1981). Uma linha temporal pode ser traçada com o objetivo de apresentar os principais avanços ao longo do tempo dessa tecnologia.

### **2.2.1. Anos 40**

Durante a Segunda Guerra Mundial houve um grande desenvolvimento tecnológico com o objetivo de desenvolver novos equipamentos para os campos de batalha a fim de vencer a guerra. Para isso, uma grande quantidade de recursos foi destinada a pesquisas científicas para desenvolvimento de diversas tecnologias, dentre elas, a IA. Os estudos na área ainda não eram abrangentes e o foco se destinava apenas em tomadas de decisão de acordo com algumas regras impostas.

### **2.2.2. Anos 50**

Na década de cinquenta iniciou-se a prova de teoremas, baseada na linha de pesquisa psicológica, utilizando a lógica de estratégia para finalidades matemáticas. A modelagem também teve início nesta época, através de regras de produção, que eram baseadas em lógicas de predicados.

A implementação da programação incluindo comandos de lógicas de predicados foi responsável por um avanço considerável para a programação de sistemas com raciocínios. Além disso, houve aperfeiçoamentos de aplicações matemáticas e simuladores.

Na linha de pesquisa biológica, essa década teve um grande passo de sucesso, pois foi implementado o primeiro simulador de Redes Neurais Artificiais (RNA) e do primeiro neurocomputador. Com a teoria de aprendizado proposta por Donald Hebb (1949) e o modelo matemático de McCulloch e Pitts (1943) foi possível a criação do modelo de rede neural artificial chamado Perceptron.

### **2.2.3. Anos 60**

O desenvolvimento na linha biológica foi aperfeiçoado com aprimoramento do modelo *Perceptron* e o surgimento de um segundo modelo, o *Adaline*, que compartilhava das mesmas ideias de rede, porém, eram diferentes na lógica de aprendizado.

Os pesquisadores da linha psicológica obtiveram verdadeira descoberta da IA, visto que nesta década, acreditavam na possibilidade de realizar tarefas humanas, como o pensamento e a compreensão da linguagem através do computador.

### **2.2.4. Anos 70 a 80**

Durante o período de 70 a 80, houve questionamentos levantados sobre o aprendizado das máquinas, mas as RNA's voltaram a ser consolidadas através do físico John Hopfield que comprovou ser possível a simulação de um sistema físico em um modelo matemático. Em 1986, uma equipe de especialistas validou sua pesquisa. Logo depois, apresentaram um algoritmo chamado *Backpropagation* que era a evolução do modelo *Perceptron*, superando as limitações do modelo anterior.

### **2.2.5. A partir dos anos 90**

A partir dos anos 90, as RNA's começaram a ser utilizadas em várias aplicações com a evolução dos modelos implementados. Nesse período, surgiram muitas propostas de aperfeiçoamentos aumentando, assim, o interesse pela área. A partir dessa época, consolidaram a inclusão das RNA's nos estudos sobre Inteligência Artificial.

Por fim, iniciaram-se, também neste período construções de sistemas híbridos, que eram sistemas em que mesclavam as concepções das duas linhas de pesquisa, resultando em grandes sistemas que prometem abranger uma forma mais completa da representação do comportamento humano, reforçando a ideia principal da IA. É importante ressaltar que existem ramificações da IA sendo que uma delas, a denominada Inteligência Computacional (IC), estuda modelos com base nas áreas complexas da natureza, como a teoria da evolução e as próprias redes de neurônios, para encontrar soluções imprecisas ou inexatas para problemas complexos que extrapolam a capacidade de processamento de informações do ser humano. Atualmente, a área de IC está em expansão motivada principalmente pelo surgimento da indústria 4.0.

## 2.3. Redes Neurais Artificiais

Considerada como uma técnica da IC, a RNA pode ser definida como uma classe de modelos matemáticos para a resolução de problemas de classificações e definições de padrões, sendo parte de estudo de fenômenos complexos.

Até 1980, a maior parte das aplicações que envolviam processamento de dados e informações eram através da abordagem da programação de computadores, que se dá pela implementação de um algoritmo para solucionar tal problema, o qual é codificado numa linguagem de programação.

Por outro lado, as RNA's são sistemas de processamento paralelo e distribuído, e tem como forma de computação distinta da tradicional, sendo definida como não algorítmica, e sim caracterizada por sistema semelhante a estrutura do cérebro humano. É por esse motivo que as RNA's não se baseiam em regras ou programas e são uma alternativa para os métodos de programação convencionais.

A capacidade de aprender e resumir informações por meio de exemplos e generalizar a informação aprendida é a principal atração da solução de problemas por meio de RNA's. Esta generalização está relacionada à capacidade da rede de aprender por meio de conjunto de exemplos e, em seguida, dar uma resposta coerente a dados não fornecidos. Isso prova que a capacidade das RNA's é mais do que um simples mapeamento entre as relações de entrada e saída. Outras funções importantes são a capacidade de auto-organização e processamento de tempo, tornando as RNA's uma ferramenta de computação extremamente poderosa e atraente para soluções de problemas complexos.

A busca por soluções em uma RNA utiliza sistemas paralelos distribuídos e compostos por unidade de processamentos simples (nodos), que calculam determinadas funções matemáticas, normalmente não-lineares. Estas unidades são compostas por uma ou mais camadas interligadas por um grande número de conexões, geralmente unidirecionais. A maioria dos modelos estão relacionados a pesos, os quais armazenam os conhecimentos representados no modelo e servem para ponderar a entrada expostas por cada neurônio (BRAGA, 2007).

### 2.3.1. *Neurônio Biológico*

Para uma melhor compreensão do funcionamento das RNA's, se faz necessário definir o neurônio biológico, visto que as RNA's se baseiam na estrutura do cérebro, especificamente,

os neurônios. Segundo Oliveira (2005), Moreto (2005), um neurônio biológico, conforme mostrado na Figura 2, é o pilar principal para o desenvolvimento de uma RNA.

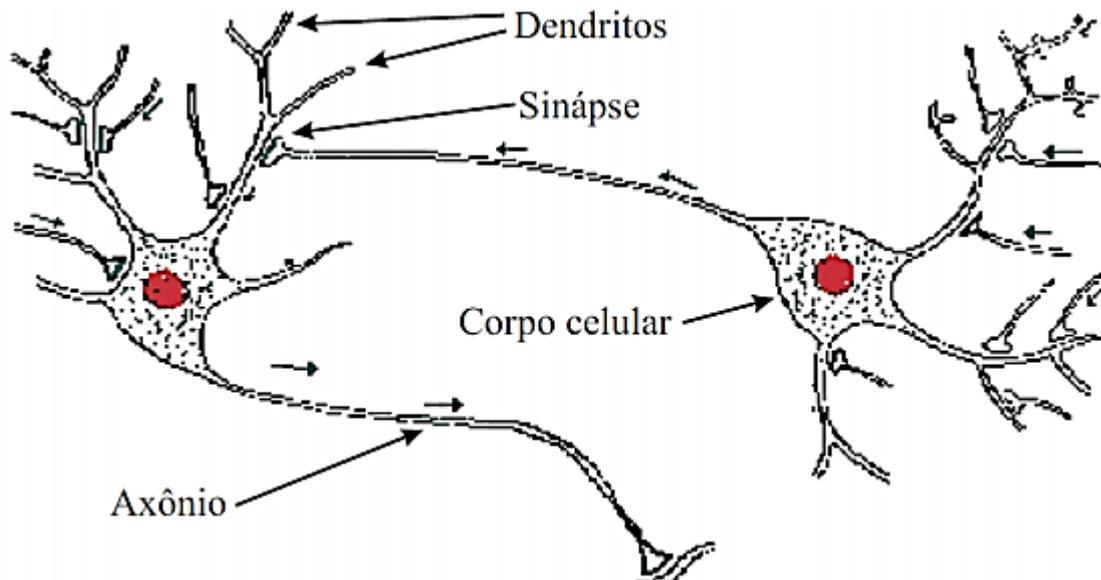


Figura 2: Neurônio Biológico.  
Fonte: (JUNGES, 2018).

Nota-se que o neurônio biológico é dividido em 3 partes: o corpo da célula, os dendritos e o axônio, sendo que cada uma possui funções específicas e complementares entre si. A função do dendrito se resume em receber informações ou impulsos nervosos provenientes de outros neurônios e, conduzi-los até o corpo celular. Estas informações são processadas, gerando assim uma resposta, em forma de impulsos e são transmitidos a outros neurônios (nodos), por meio do axônio e atingindo os dendritos em sequência (JESS, 2004).

Os impulsos de resposta são assimilados como uma perturbação elétrica circulante e sua velocidade está diretamente relacionada ao diâmetro da fibra nervosa, sendo mais veloz o impulso, em um diâmetro maior. Os impulsos recebidos por um neurônio são armazenados e associados como uma função somatória, onde seu resultado é passado para frente pelo axônio desse neurônio.

Com base nos três conceitos do corpo de um neurônio, é possível compreender a sinapse, que ocorre na região de encontro de um axônio e um dendrito, no qual o impulso é transmitido de um neurônio para o outro (OLIVEIRA, 2005). A Figura 3 mostra como ocorre a transmissão de impulsos nervosos entre neurônios.

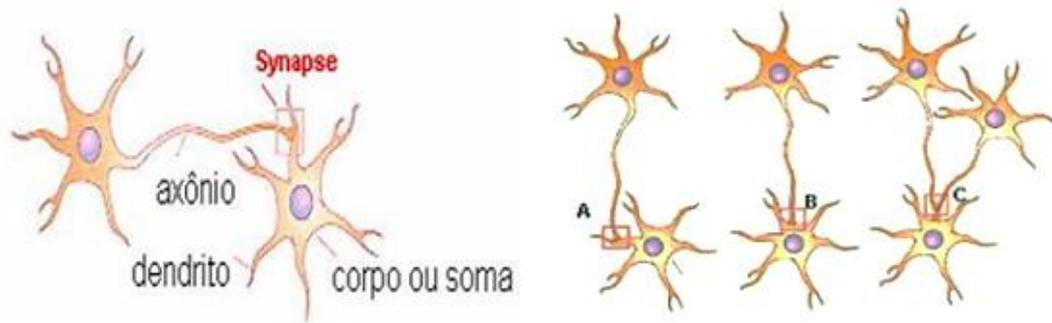


Figura 3: A transmissão de impulsos nervosos entre neurônios.  
Fonte: (GRAUPE,1997).

Os axônios de um neurônio não possuem contato direto com os dendritos de um segundo neurônio, existindo entre eles um pequeno espaço. No momento em que o impulso alcança a extremidade do axônio, são liberadas substâncias químicas que percorrem esse espaçamento. Em consequência, essas substâncias acumulam-se em torno do dendrito do segundo neurônio de forma que o impulso é transmitido de um neurônio para o outro. A Figura 4, ilustra a passagem do impulso entre neurônios.

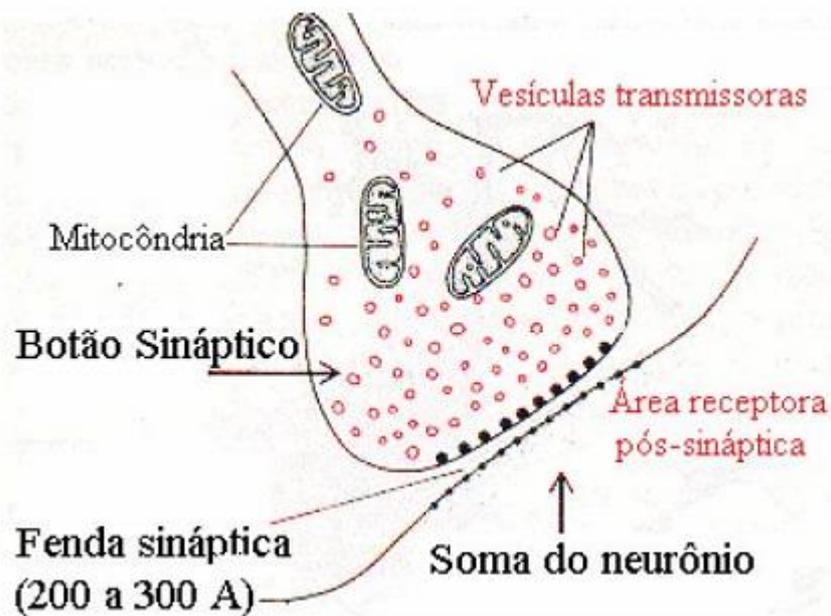


Figura 4: Passagem do impulso.  
Fonte: (GRAUPE,1997).

### 2.3.2. Neurônio Artificial

As RNA's buscam simular o funcionamento do cérebro humano utilizando computadores com a finalidade de tornar possível o aprendizado e a tomada de decisão de forma inteligente. As estruturas de interconexão dos neurônicos de uma RNA também são baseadas na estrutura do cérebro.

De maneira semelhante ao modelo biológico, o artificial também é dividido em 3 partes, como mostra a Figura 5, que são definidas segundo Oliveira (2005) e Moreto (2005):

- Sinapses: é o nome dado as entradas de sinais do neurônio, que possuem seus respectivos pesos, os quais são multiplicados aos parâmetros injetados em cada neurônio. Estes pesos possuem o objetivo de determinar a influência do sinal para o neurônio, possuindo valor positivo (sinapse excitatória) ou negativo (sinapse inibitória);
- Somador: tem como função realizar a soma ponderada (pelos pesos sinápticos) das entradas e, adicionalmente, ajusta-se sua saída por um parâmetro denominado bias ( $b_k$ ). Desta forma, é possível aumentar ou diminuir a saída do somador, para melhor adaptação do modelo aos dados fornecidos.
- Função de ativação: é um modelo matemático que limita o sinal que sai do somador para valores normalizados, como por exemplo entre  $[-1, 1]$  ou  $[0, 1]$ .

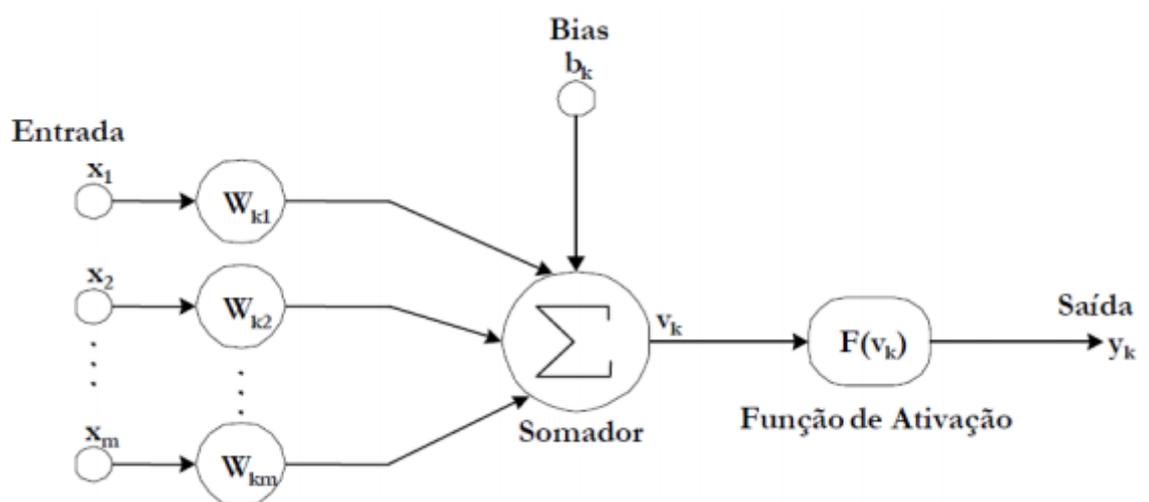


Figura 5: Representação de um neurônio artificial.

Fonte: (OLIVEIRA, 2005).

Matematicamente, Moreto (2005) modelou a saída (do somador) de um neurônio artificial conforme mostrado na Equação (1). A Equação (2), modela a saída do neurônio após a aplicação da função de ativação.

$$v_k = \sum_{j=1}^m w_{kj} \cdot x_j + b_k \quad (1)$$

$$y_k = \varphi(v_k) \quad (2)$$

Onde:

$j$ = índice de soma;

$m$ = número de entrada do neurônio;

$w_{kj}$ = pesos das sinapses;

$x_j$ = sinal de entrada;

$b_k$ = bias;

$v_k$ = saída do somador;

$\varphi$ = função de ativação;

$y_k$ = saída do neurônio.

É importante destacar que os neurônios biológicos são complexos e possuem diversos mecanismos ainda não explicados matematicamente, por isso, o modelo de neurônio artificial ainda é muito primitivo quando comparado com o modelo biológico (OLIVEIRA, 2005).

### 2.3.3. Funções de Ativação

A função de ativação é definida por uma função matemática que limita a saída do somador para uma classe de valores normalizada. Existem diversos tipos de funções de ativação, tais como limiar, linear, função rampa, sigmóide e tangente hiperbólica.

#### 2.3.3.1. Limiar

De acordo Moreto (2005), com a implementação do modelo de neurônio artificial desenvolvido por McCulloch & Pitts, obteve-se a função limiar. Sendo uma função binária que se restringe em apenas duas saídas que podem assumir dois valores, sendo eles: zero ou um. A

Figura 6 mostra o gráfico de uma função limiar a qual é descrita matematicamente pela Equação (3).

$$F(v_k) = \{0, v_k < 0; 1, v_k \geq 0\} \quad (3)$$

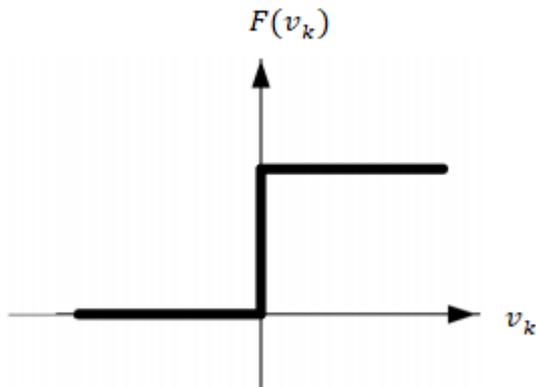


Figura 6: Função Limiar.  
Fonte: (JESS, 2004).

#### 2.3.3.2. Linear

A função linear, diferentemente da função limiar, não limita a saída do neurônio. A Figura 7 mostra o gráfico da função linear e a função matemática correspondente dessa função é descrita por meio da Equação (4).

$$F(v_k) = v_k \quad (4)$$

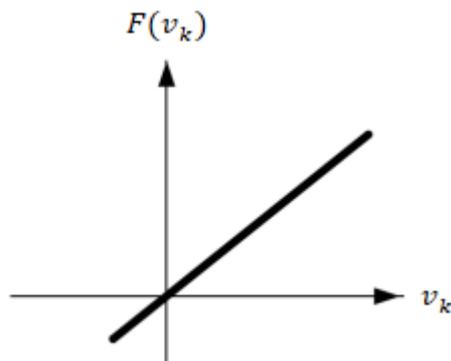


Figura 7: Função Linear.  
Fonte: (JESS, 2004).

### 2.3.3.3. Rampa

A função rampa, possui intervalos que limitam a saída do neurônio entre um valor inferior e outro superior. No entanto, os valores entre o intervalo  $[-s, +s]$ , recebem na saída o valor de entrada. A Figura 8 mostra o gráfico da função rampa e Equação 5 descreve matematicamente essa função.

$$F(v_k) = \begin{cases} 0, & v_k \leq -s; \\ \sum_{j=1}^m w_{kj} \cdot x_j + b_k, & -s < v_k < +s; \\ 1, & v_k \geq +s; \end{cases} \quad (5)$$

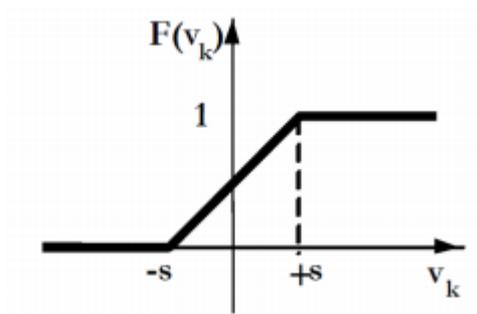


Figura 8: Função Rampa.

Fonte: (JESS, 2004).

### 2.3.3.4. Sigmóide

A função Sigmóide é descrita como uma função estritamente crescente, que demonstra um balanceamento aceitável entre o comportamento linear e não linear (OLIVEIRA, 2005). A Figura 9 mostra o gráfico da função sigmóide e a Equação (6) mostra a formulação matemática dessa função.

$$F(v_k) = \frac{1}{1 + e^{-\frac{v_k}{T}}} \quad (6)$$

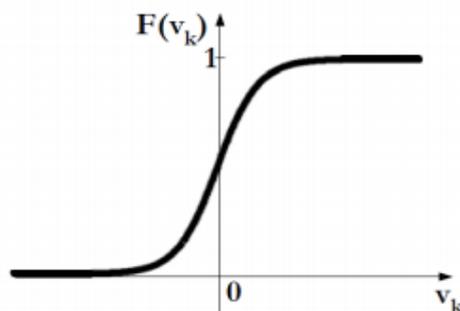


Figura 9: Função Sigmóide.

Fonte: (JESS, 2004).

### 2.3.3.5. Tangente Hiperbólica

Esta função, ilustrada na Figura 10, é semelhante a função sigmóide, porém, enquanto a função sigmóide possui saídas limitadas entre  $[-1, 1]$ , a tangente hiperbólica se limita em saídas entre  $[0, 1]$ . A Equação 7 descreve matematicamente essa função.

$$F(v_k) = \frac{e^t - e^{-t}}{e^t + e^{-t}} \quad (7)$$

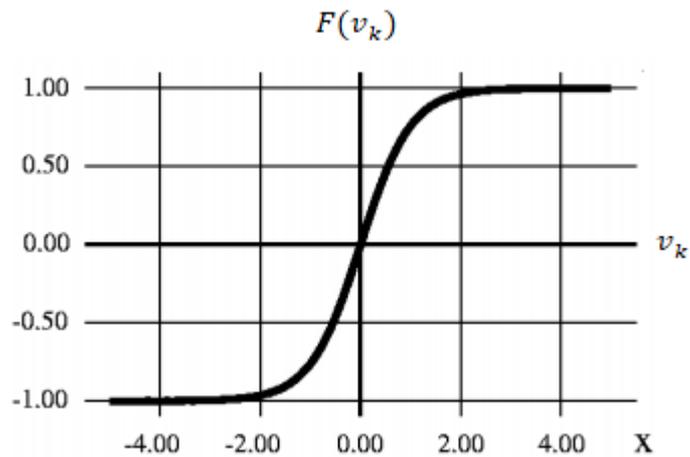


Figura 10: Função tangente hiperbólica.  
Fonte: (GALLO, 2016).

### 2.3.4. Arquitetura de RNA's

Uma RNA não é implementada utilizando apenas um neurônio artificial, mas por um conjunto de neurônios interligados, formando uma arquitetura específica. A arquitetura de uma RNA é definida com base no número de camadas e o tipo de conexão entre elas e pode assumir três tipos de classificação: redes *feedforward* de camada única, redes *feedforward* de múltiplas camadas e redes recorrentes.

#### 2.3.4.1. Redes *feedforward* de Camada Única

Esta é a forma mais simples de uma rede neural, os neurônios são expostos a uma única camada e a saída de cada neurônio possui apenas as saídas da rede. Nesta configuração, a rede possui uma camada de nós que alimenta a camada de saídas com os sinais de entrada. A Figura 11 mostra um exemplo de rede de camada única.

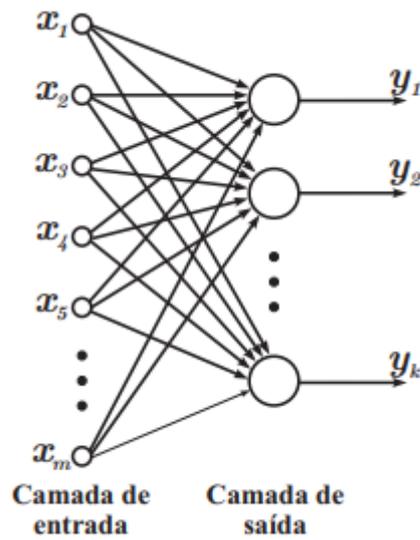


Figura 11: Exemplo de rede camada única.  
Fonte: (MORETO, 2005).

#### 2.3.4.2. Redes *feedforward* de Múltiplas Camadas

Esta configuração de rede é distinta da rede anterior, devido ao acréscimo de mais camadas na sua configuração, sendo elas chamadas de camadas de neurônios escondidos ou ocultos. São definidos como camada de entrada, os neurônios que recebem sinais de excitação por meio externo. As camadas escondidas, são as que possuem neurônios intermediários, e por fim, as de saídas, possuem neurônios que estão à saída da rede. A Figura 12 mostra um exemplo de rede multicamadas.

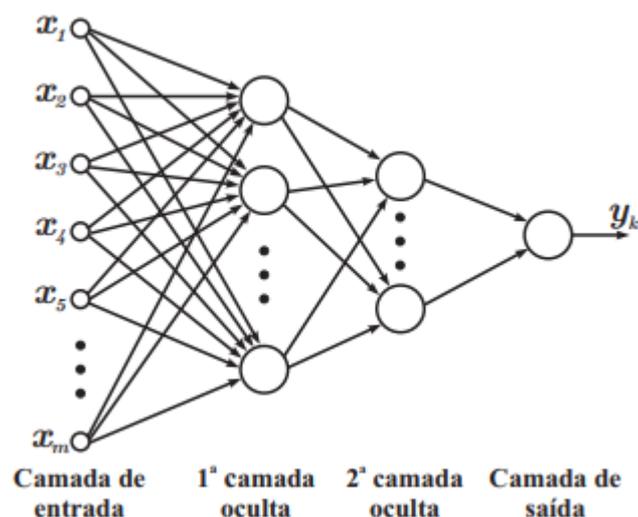


Figura 12: Exemplo de rede multicamadas.  
Fonte: (MORETO, 2005).

### 2.3.4.3. Redes recorrentes

As redes recorrentes, são baseadas em realimentações entre camadas distintas, desta forma, pode-se observar na Figura 13, o envio de sinal de um neurônio de uma camada posterior para uma camada anterior.

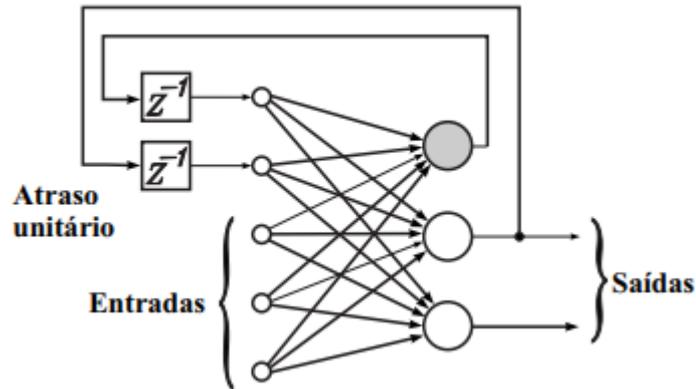


Figura 13: Exemplo de rede multicamadas.  
Fonte: (MORETO, 2005).

### 2.3.5. Treinamento

Após ter compreendido o formato da arquitetura das RNA's, torna-se fundamental compreender o treinamento do modelo, cujo objetivo consiste em atribuir pesos às sinapses. Nesta atribuição, os valores são alterados com o intuito de encontrar uma solução genérica para o problema apresentado, tendo o menor erro possível na saída da rede.

Os tipos de treinamentos, em geral, podem ser classificados em duas classes: a aprendizagem supervisionada e a aprendizagem não-supervisionada, que serão explicadas em seguida.

O modelo de treinamento supervisionado, apresenta um fator externo, como um supervisor que aponta a saída desejada da rede para a entrada inserida, conforme mostrado na Figura 14. Desta forma, este fator externo reajusta os parâmetros da RNA, através do erro, que é a diferença do valor calculado e o apresentado pelo supervisor.

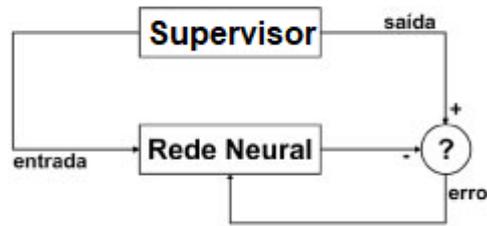


Figura 14: Diagrama de Blocos do Treinamento Supervisionado.  
Fonte: (FIORIN et al., 2011)

O modelo de treinamento não supervisionado, não apresenta uma saída da rede, diante disso, a mesma necessita extrair dados importantes dos impulsos para classificá-los em grupos com características semelhantes que são pré-determinadas. Entretanto, em um possível caso que o dado não é relacionado em nenhum grupo, a rede cria uma nova classe para o mesmo. Este comportamento pode ser visualizado pela Figura 15.

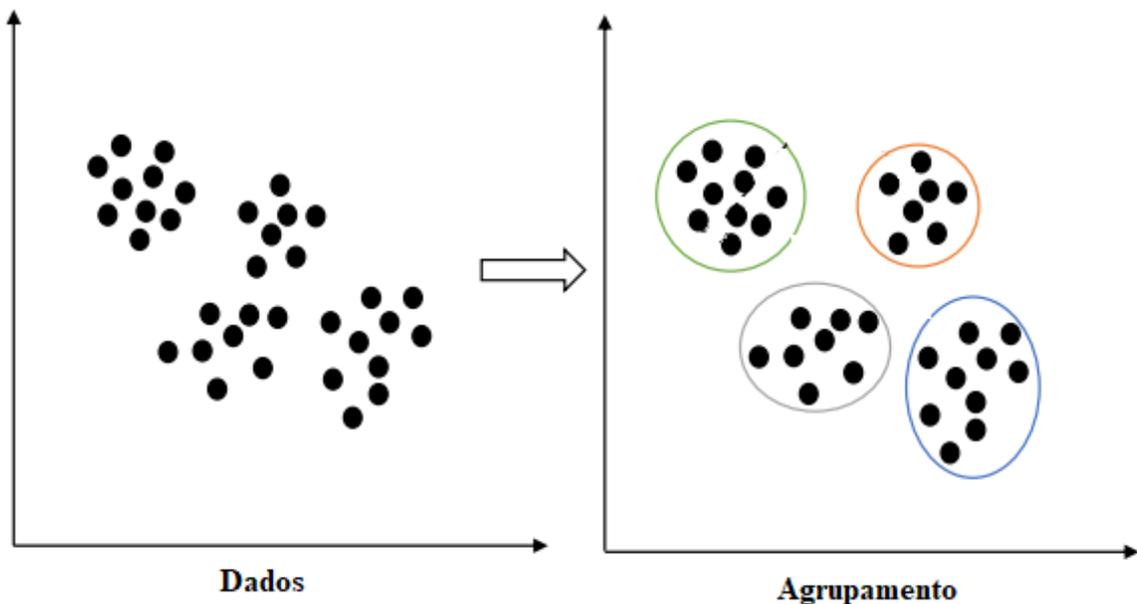


Figura 15: Diagrama de Blocos do Treinamento não-Supervisionado.  
Fonte: (EZUGWU et al., 2020).

## 2.4. Regressão Linear

Implementado por Francis Galton, em uma pesquisa para medir a altura das pessoas, os fundamentos da regressão linear foram apresentados, inicialmente, como um fenômeno

definido como lei de regressão universal de Galton e, posteriormente, por Pearson (SOARES, 2011). Trata-se de uma técnica estatística, pois os modelos matemáticos são obtidos pela análise da relação entre uma ou mais variáveis, em que a dependência de aspectos de interesse possibilita estimar determinados valores apenas com o conhecimento das variáveis que os afetam.

Atualmente, utilizado na programação, a Regressão Linear é um algoritmo de aprendizado de máquinas supervisionado, normalmente aplicado no estudo de um banco de dados específico. O modelo possibilita a análise entre a correlação existente entre uma variável alvo, que pode ser definida como variável dependente (a ser estimada), com outras variáveis ditas como explicativas, que também são conhecidas como variáveis de entrada ou variáveis independentes (ROCHA, 2013).

#### **2.4.1. Regressão Linear Simples**

A Regressão Linear Simples, é o modelo primitivo da regressão linear, que estuda a relação entre a variável dependente e possuindo apenas uma variável independente como parâmetro, no caso, ( $X_1$ ).

Matematicamente o modelo é representado através da seguinte equação:

$$Y = \beta_0 + \beta_1 X_1 \quad (8)$$

Em que:

$Y$  = variável Alvo (dependente);

$X_1$  = variável independente;

$\beta_1$  = Coeficiente de regressão, representa a variação em no eixo das ordenadas em função dos valores da abscissa;

$\beta_0$  = Constante de regressão, representa o valor de intercepção da reta no eixo das ordenadas.

Considerando os parâmetros descritos no modelo matemático, este pode ser interpretado, graficamente, pela Figura 16, a qual ilustra uma reta de regressão genérica. Supondo que existe efetivamente uma relação linear entre  $X_1$  e  $Y$ , os parâmetros  $\beta_0$  e  $\beta_1$  são ajustados computacionalmente para minimizar os erros em comparação com valores conhecidos de  $Y$ .

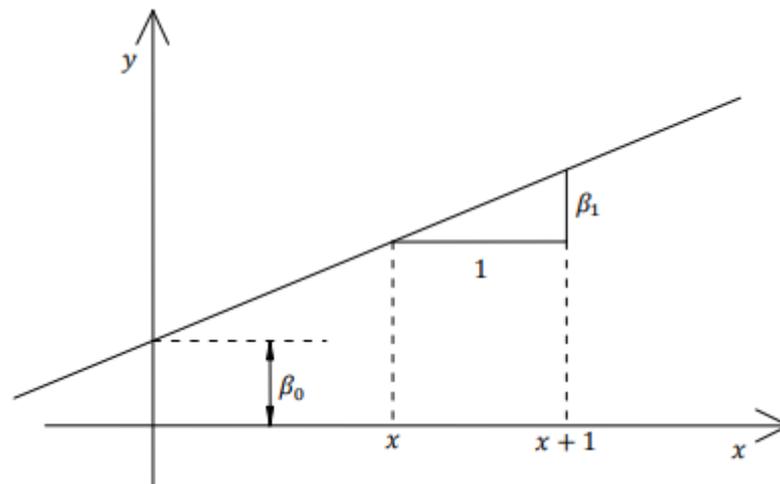


Figura 16: Representação geométrica dos parâmetros  $\beta_i$  e  $\beta_0$   
 Fonte: (RODRIGUES, 2012).

#### 2.4.2. Regressão Linear Múltipla

A técnica de Regressão Linear Múltipla é utilizada para estudo da relação entre uma única variável dependente com multivariáveis independentes (BAPTISTELLA, 2005). Semelhante ao modelo de Regressão Linear Simples, o modelo com mais variáveis é representado pela seguinte equação:

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_i X_i \quad (9)$$

Em que:

$Y$  = variável Alvo (dependente);

$X_i$  = variáveis independentes;

$\beta_i$  = Coeficientes de regressão, relacionado cada uma das variáveis de entrada;

$\beta_0$  = Constante de regressão, representa o valor de intercepção da reta no eixo das ordenadas.

Devido as múltiplas variáveis explicativas de entrada, diferente da regressão linear simples, que se obtém uma reta, a regressão linear múltipla obtém-se um hiperplano, exemplificado pelo gráfico mostrado na Figura 17.

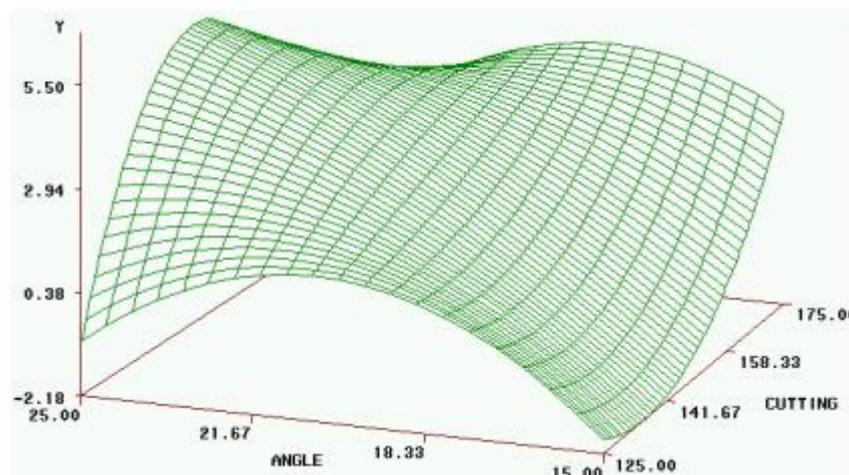


Figura 17: Hiperplano referente a multivariáveis independentes  
Fonte: (RODRIGUES, 2012).

## 2.5. Árvore de Decisão

É uma técnica gráfica em formato simbólico de uma árvore, na qual a tomada de decisões é baseada na análise de possíveis consequências (riscos, custos, prejuízos, etc.) de determinadas ações, sendo possível estabelecer planos para alcançar um objetivo proposto.

Especificamente, por ser uma técnica de mineração de dados, é bastante comum em soluções de problemas de classificação, regressão e previsão de variáveis. Utilizando a metodologia de indução de regras, as árvores apresentam os resultados hierarquicamente. Um exemplo disso, são os nós, que constituem os atributos de maior importância para a decisão. Desta forma, tendo como consideração estes “nós”, de acordo com sua importância, são tomadas as decisões para obter tal resultado (FREITAS, 2015).

A estrutura da árvore de decisão é definida como:

1. Cada nó, é um atributo dos dados utilizados;
2. O atributo-classe são as folhas da árvore;
3. Um ramo ligado de um nó-filho a um nó pai, tem seu valor atribuído ao nó-pai;
4. Um atributo que está presente em um nó, não pode ser apresentado em seus nós descendentes.

A Figura 18, ilustra a estrutura de uma árvore de decisão. Esta estrutura apresenta uma hierarquia para classificar as informações de um conjunto de dados sendo que, uma saída é

definida como variável alvo, no caso o SoC, e, em sequência, é realizada uma busca no conjunto de dados de entrada das variáveis que estão mais relacionadas com a saída, determinando-a como o primeiro ramo. Em seguida são definidos os nós posteriores com os demais fatores que podem ter influência sobre a saída até que se alcance a folha, que é o último nível, sendo o resultado que se busca prever.

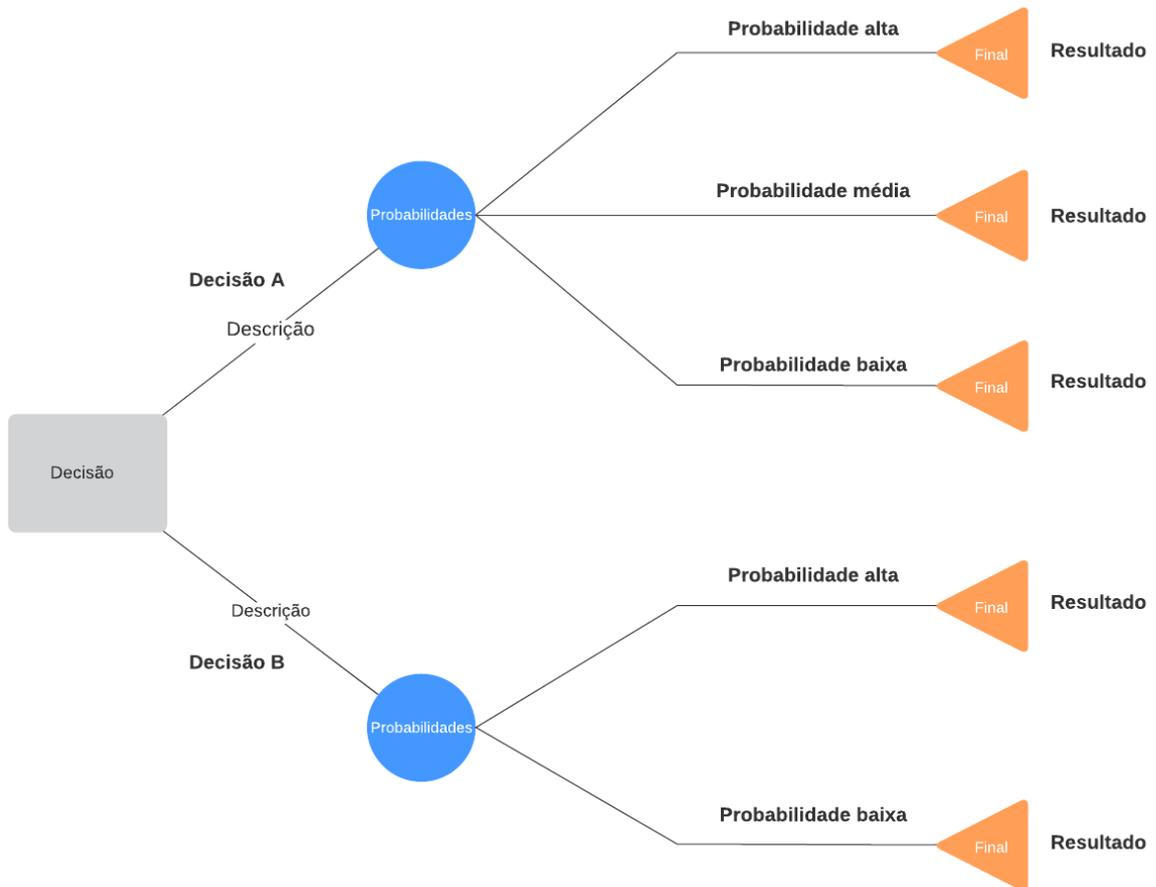


Figura 18: Exemplo de Estrutura Árvore de decisão.  
Fonte: (HOMEM MAQUINA, 2020).

No exemplo mostrado na Figura 18, a decisão a ser tomada, ou seja, o caminho a ser percorrido na árvore, é baseada nas probabilidades descritas pelas variáveis de entrada. Por exemplo, caso uma variável de entrada, cuja medida se encontra em determinada faixa de valores, tenha probabilidade maior de gerar um resultado que consta nos ramos da Decisão A, este será o caminho a ser percorrido, descartando-se o caminho da Decisão B. Em sequência, outra variável de entrada, menos impactante que a primeira, é posta em análise, de forma que o próximo caminho a ser tomado é definido pela probabilidade associada ao valor desta variável. No exemplo, o resultado é estabelecido por apenas duas tomadas de decisão, porém é possível estabelecer outros parâmetros/variáveis, aumentando o tamanho da árvore de decisão.

Implementar um modelo de árvore de decisão, especificadamente neste trabalho, tornou-se um desafio, visto que, a variável que se busca prever (SoC), requer o uso de árvores grandes e complexas, tendo a possibilidade de obter *overfitting*, que é um fenômeno que ocorre quando a rede memoriza os dados de entrada e não consegue se adaptar a pequenas variações dos padrões de entrada. Sendo assim, os resultados para dados distintos aos de treinamento, a performance pode não ser satisfatória.

Desta maneira, para prevenção do problema citado, foram utilizados alguns parâmetros, sendo eles:

- Números máximos de nós;
- Profundidade máxima da árvore;

Utilizando estes dois parâmetros, é possível limitar o tamanho da árvore que será implementada, obtendo um modelo com menor complexidade, reduzindo a probabilidade de obter a memorização dos dados expostos ao modelo.

## 2.6. Métricas de Avaliação

Para avaliar o desempenho dos modelos de previsão desenvolvidos, é necessário empregar as denominadas métricas de avaliação. Dentre as métricas existentes, destaca-se a raiz quadrada do erro-médio, traduzido do inglês *Root Mean Squared Error* (RMSE), que representa a raiz quadrática da diferenças entre os valores obtidos pelos modelos de estimação de SoC implementados em relação a curva real dos dados. O cálculo desta métrica de avaliação é realizado conforme descrito pela Equação (8).

$$RMSE = \frac{1}{N} \sum_{t=1}^N \sqrt{\mathbf{d}_t - \mathbf{y}_t} \quad (8)$$

onde, N é a quantidade de amostras,  $\mathbf{d}_t$  é valor real do SoC e  $\mathbf{y}_t$  é o valor encontrado pelo modelo.

Um segundo método de avaliação utilizado é o Erro Quadrático Médio, *Mean Square Erro* (MSE), descrito pela Equação 9, que possui uma forma de validação diferente da anterior, ao punir os erros maiores encontrados, pela elevação de seus valores ao quadrado.

$$MSE = \frac{1}{N} \sum_{t=1}^N (\mathbf{d}_t - \mathbf{y}_t)^2 \quad (9)$$

Em terceiro plano, é utilizado o Erro Absoluto Médio, *Mean Absolute Error* (MAE), representado pela Equação 10. Esta métrica, em termos gerais, descreve a distância média entre os valores reais e os encontrados pelo modelo.

$$MAE = \frac{1}{N} \sum_{t=1}^N |d_t - y_t| \quad (10)$$

Desta maneira, as métricas citadas anteriormente, MAE, RMSE e MSE representam a magnitude do erro, assim, quanto menor os valores calculados, melhor é o desempenho do modelo implementado.

Um quarto índice utilizado, foi o R-quadrado, que expressa a variância dos dados resultados do modelo desenvolvido, ou seja, é calculado a porcentagem da variância presente, podendo ser um valor entre 0 a 1. Desta forma, um valor de R-quadrado mais próximo da unidade, expressa um modelo mais assertivo.

$$R^2 = 1 - \sum_{t=1}^N \frac{(y_i - \hat{y}_i)^2}{(y_i - \bar{y})^2} \quad (11)$$

### 3. METODOLOGIA

#### 3.1. Introdução

Neste capítulo será abordada a metodologia empregada para o alcance dos objetivos do presente trabalho, assim como, os dados utilizados, modelos e técnicas. Em resumo, a metodologia proposta consiste na análise de um banco de dados que contém informações sobre diversas grandezas físicas que podem influenciar o estado de carga de baterias utilizadas em veículos elétricos. Esta análise é realizada com o auxílio de técnicas de inteligência computacional, conforme fluxograma apresentado na Figura 19.

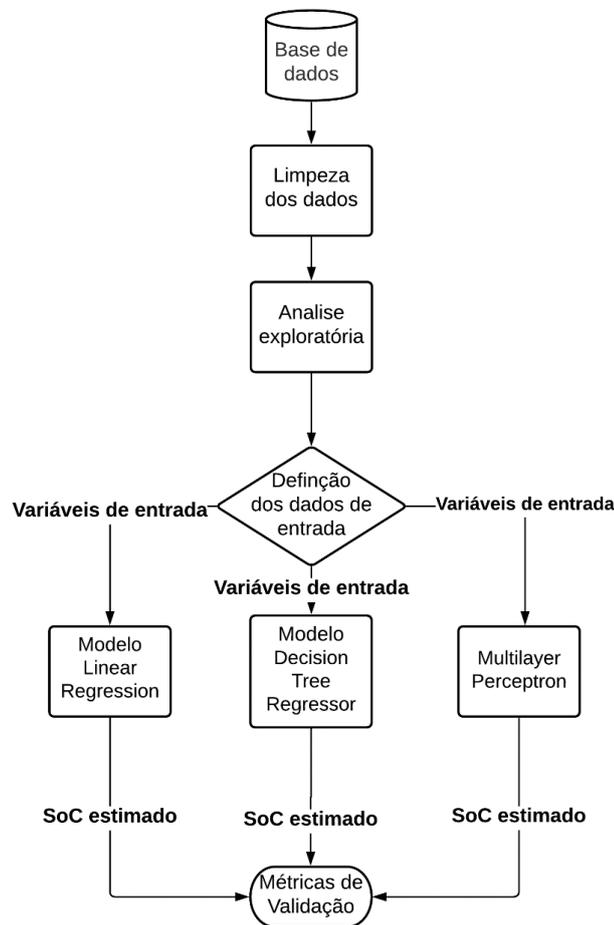


Figura 19: Fluxograma referente a metodologia utilizada.  
Fonte: Autor, 2021.

O fluxograma apresentado, descreve a sequência de instruções definida para a rotina computacional desenvolvida neste trabalho, as quais podem ser divididas nas etapas expostas a seguir.

1. Limpeza dos dados;
2. Análise exploratória;
3. Desenvolvimento do modelo *Linear Regression*;
4. Desenvolvimento do modelo *Decision Tree Regressor*;
5. Desenvolvimento do modelo *Multilayer Perceptron*;
6. Comparação e análise estatística dos modelos.

A primeira etapa consiste na limpeza de dados contidos no banco em análise. Este tratamento é necessário para possibilitar a manipulação adequada das informações. Basicamente, as informações contidas no banco de dados foram cuidadosamente avaliadas e, as grandezas que apresentavam medições sem valores válidos, ou que manifestavam escalas discrepantes entre valores mínimos e máximos, foram excluídas das análises. Desta forma, a próxima etapa, descrita como análise exploratória, se torna mais eficiente para identificar correlações, padrões e hipóteses. Nas etapas finais, os modelos de Regressão Linear, Árvore de decisão e RNA foram desenvolvidos e submetidos às métricas de avaliação, de modo a analisar a performance dos resultados. O treinamento dos modelos foi realizado considerando-se 70% dos dados, sendo que os restantes (30%) foram utilizados para os testes de validação.

O *software* utilizado para implementação da rotina computacional, foi o *Jupyter*, o qual utiliza a linguagem de programação *Python*. É importante destacar que foi adotada a utilização de algumas bibliotecas, as quais possuem funções específicas que facilitam a manipulação de dados e estruturação dos modelos previstos. Dentre estas, destaca-se o uso das bibliotecas indicadas na Figura 20.

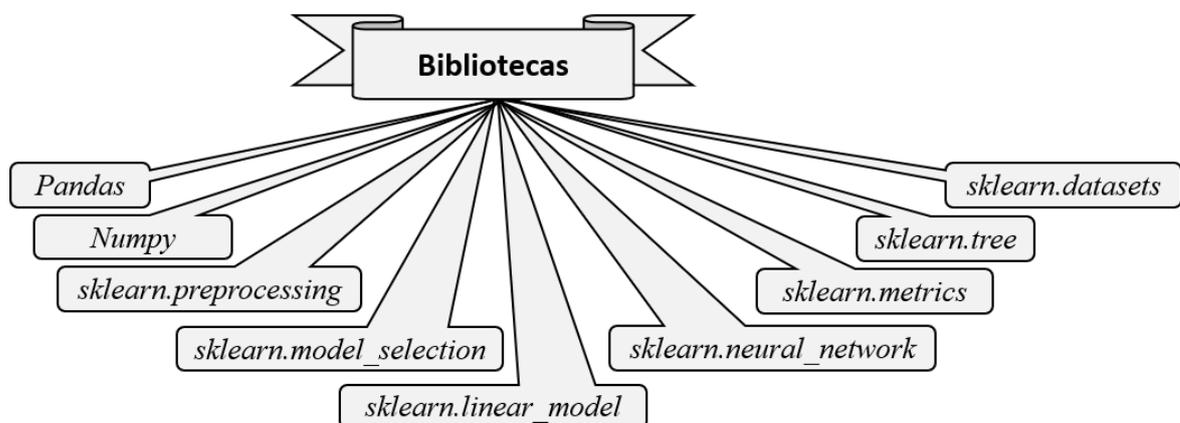


Figura 20: Bibliotecas auxiliares para desenvolvimento da rotina computacional.

Fonte: Autor, 2021.

### 3.2. Base de Dados

Os dados utilizados neste trabalho foram disponibilizados pela IEEE DataPort™, que se trata de um repositório de dados online desenvolvido pelo IEEE (*Institute of Electrical and Electronics Engineers*), permitindo aos usuários armazenar, pesquisar, acessar e gerenciar dados diversos. Especificadamente, foi utilizado, como objeto de estudo, o conjunto de dados *BATTERY AND HEATING DATA IN REAL DRIVING CYCLES*, dos autores Matthias Steinstraeter, Johannes Buberger, Dimitar Trifonov do *Technical University of Munich, Institute of Automotive Technology*.

Para obtenção deste banco de dados, foram realizados registros de viagens reais de condução do automóvel BMW I3 (60hA) com variações de terreno, horário, condutor, etc. Em cada viagem, diversas grandezas foram medidas a cada 100ms a partir de sensores instalados no veículo citado, conforme ilustrado pela Figura 21.

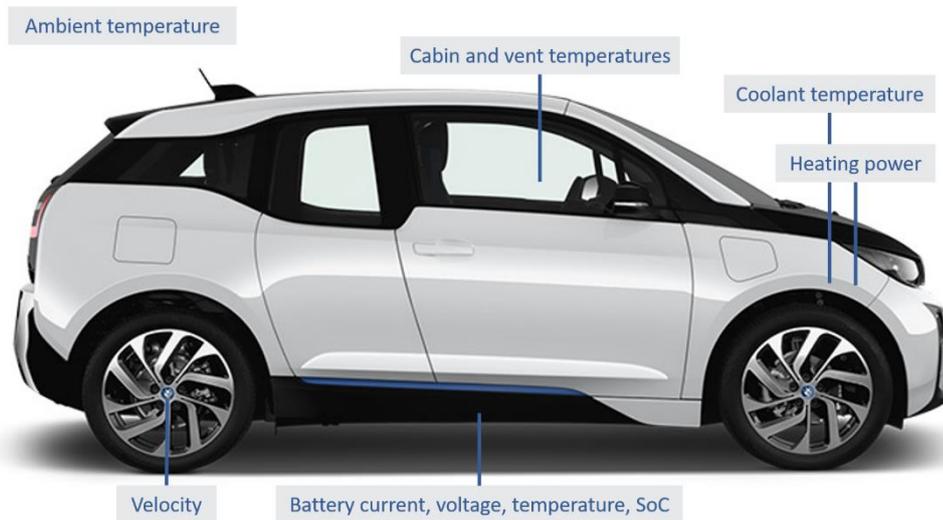


Figura 21: Veículo BMW I3 (60hA).  
Fonte: (MATTHIAS et al.,2020).

Dentre as grandezas medidas, podem se destacar: dados ambientais (temperatura, altitude, etc.); dados do veículo (velocidade, aceleração, etc.); dados da bateria (tensão, corrente, temperatura, SoC); dados do circuito de aquecimento (temperatura interna, energia de aquecimento, etc.). Especificamente, trata-se de 48 variáveis, as quais estão listadas no Quadro 1, que foram amostradas em intervalos de 1ms para viagens denominadas pelos autores das medições como *Trip B*. A *Trip B* se refere à registros de 38 viagens que, como mencionado anteriormente, variam as condições do trajeto e, como objeto de estudo, foram agrupadas informações de 4 destas viagens, totalizando 61458 registros para análise. É importante ressaltar que o banco de dados completo também apresenta outro agrupamento de viagens, denominado como *Trip A*, que não foi utilizado no presente estudo por se tratar de dados simulados.

Quadro 1: Lista de variáveis do banco de dados.

Nº	Grandeza	Nº	Grandeza
1	Tempo[s]	25	Núcleo do aquecedor de temperatura do refrigerante [°C]
2	Velocidade [km/h]	26	Temperatura Solicitada do Refrigerante [°C]
3	Elevação [m]	27	Entrada de temperatura do refrigerante [°C]
4	Aceleração [%]	28	Fluxo de volume do refrigerante +500[l/h]
5	Torque do Motor [Nm]	29	Temperatura do trocador de calor [°C]
6	Aceleração Longitudinal [m/s <sup>2</sup> ]	30	Sensor de temperatura da cabine [°C]
7	Sinal de frenagem regenerativa	31	Entrada do aquecedor de refrigerante de temperatura [°C]
8	Tensão da Bateria [V]	32	Saída do aquecedor do refrigerante de temperatura [°C]
9	Corrente da Bateria [A]	33	Saída do trocador de calor de temperatura[°C]
10	Temperatura da Bateria [°C]	34	Temperatura Descongelamento lateral esquerdo[°C]
11	Max. Temperatura da Bateria [°C]	35	Temperatura Descongelamento lateral direita[°C]
12	SoC [%]	36	Temperatura de descongelamento central [°C]
13	Display SoC [%]	37	Temperatura descongelamento central esquerda [°C]
14	Min. SoC [%]	38	Temperatura de descongelamento central direita [°C]
15	Max. SoC [%]	39	Temperatura <i>Footweel Driver</i> [°C]
16	Potência de Aquecimento CAN [kW]	40	Temperatura <i>Footweel Co-Driver</i> [°C]
17	Potência de Aquecimento LIN [W]	41	Temperatura <i>Feetvent Co-Driver</i> [°C]
18	Potência de aquecimento solicitada [W]	42	Temperatura <i>Feetvent Driver</i> [°C]
19	Potência do Ar Condicionado [kW]	43	Temperatura <i>Head Co-Driver</i> [°C]
20	Sinal Aquecedor	44	Temperatura <i>Head Driver</i> [°C]
21	Tensão do aquecedor [V]	45	Temperatura Ventilação direita [°C]
22	Corrente do aquecedor[A]	46	Temperatura Ventilação central direita [°C]
23	Temperatura ambiente [°C]	47	Temperatura Ventilação central esquerda[°C]
24	Sensor de temperatura ambiente [°C]	48	Temperatura Ventilação esquerda [°C]

Fonte: Adaptado de (MATTHIAS et al.,2020).

### 3.2.1. Análise exploratória do banco dados

Após realizada a limpeza dos dados, iniciou-se a análise exploratória, onde é feito um estudo aprofundado do conjunto de dados com o intuito de colher as suas principais características. Assim, permite-se compreender padrões, anomalias, gerar hipóteses e suposições. Considerando-se o banco de dados em estudo, o primeiro passo desta análise consistiu em gerar gráficos para verificar as variações de cada grandeza. Estes gráficos permitem investigar a distribuição dos dados para identificação daqueles que se diferenciam drasticamente dos outros, os denominados *outliers*. Estes pontos considerados fora da curva devem ser eliminados das análises para não influenciar negativamente nos resultados da implementação. Na Figura 22, é possível observar a distribuição de dados das grandezas que possuem *outliers*. Tomando como exemplo a grandeza denominada como *Heater Voltage [V]*, percebe-se que aproximadamente 50000 amostras do banco de dados (eixo vertical) se apresentam com valores na faixa entre 0 e 400 Volts (eixo horizontal). No entanto, uma pequena parcela das amostras apresenta valores mais elevados, chegando até a aproximadamente 1000 Volts, o que caracteriza uma discrepância em relação à maioria dos dados amostrados para esta grandeza.

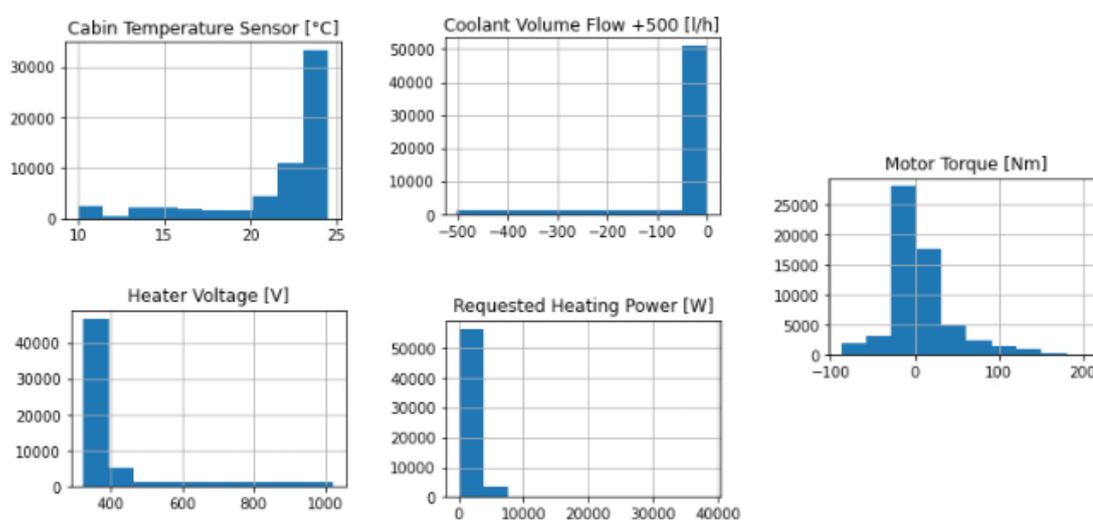


Figura 22: Gráficos da distribuição de dados das variáveis com *outliers*.

Fonte: Autor, 2021.

Em um segundo instante, observou-se que algumas grandezas são constantes e não influenciam, portanto, na estimação do estado de carga. Na Figura 23 está ilustrada a

distribuição de dados para as grandezas constantes, as quais foram removidas. Seguindo o mesmo raciocínio de análise dos *outliers*, percebe-se que todas as amostras (61458) das grandezas ilustradas a seguir, se apresentam com mesmo valor numérico.

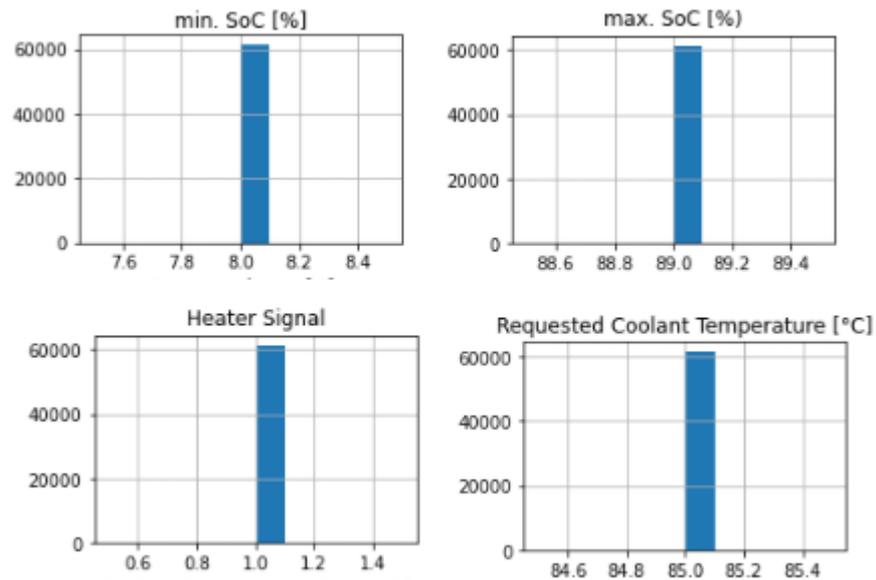


Figura 23: Gráficos da distribuição de dados constantes.

Fonte: Autor, 2021.

Em terceiro plano, foi implementado uma função para relacionar a variável a ser estimada (SoC) com as demais variáveis do banco de dados. Neste contexto, duas análises foram realizadas: correlação linear e não-linear entre as variáveis e o SoC (estado de carga). As figuras 24 e 25 apresentam estas correlações, em que as variáveis com maiores valores positivos (cores claras), possuem maior correlação positiva com o estado de carga da bateria. Em contrapartida, as variáveis mais escuras, possuem valores negativos, pois, são as variáveis que possuem correlação inversa. A partir destas análises, foi possível definir as possíveis variáveis de entradas para aplicar os modelos de Regressão Linear, Árvore de Decisão e Redes Neurais Artificiais.

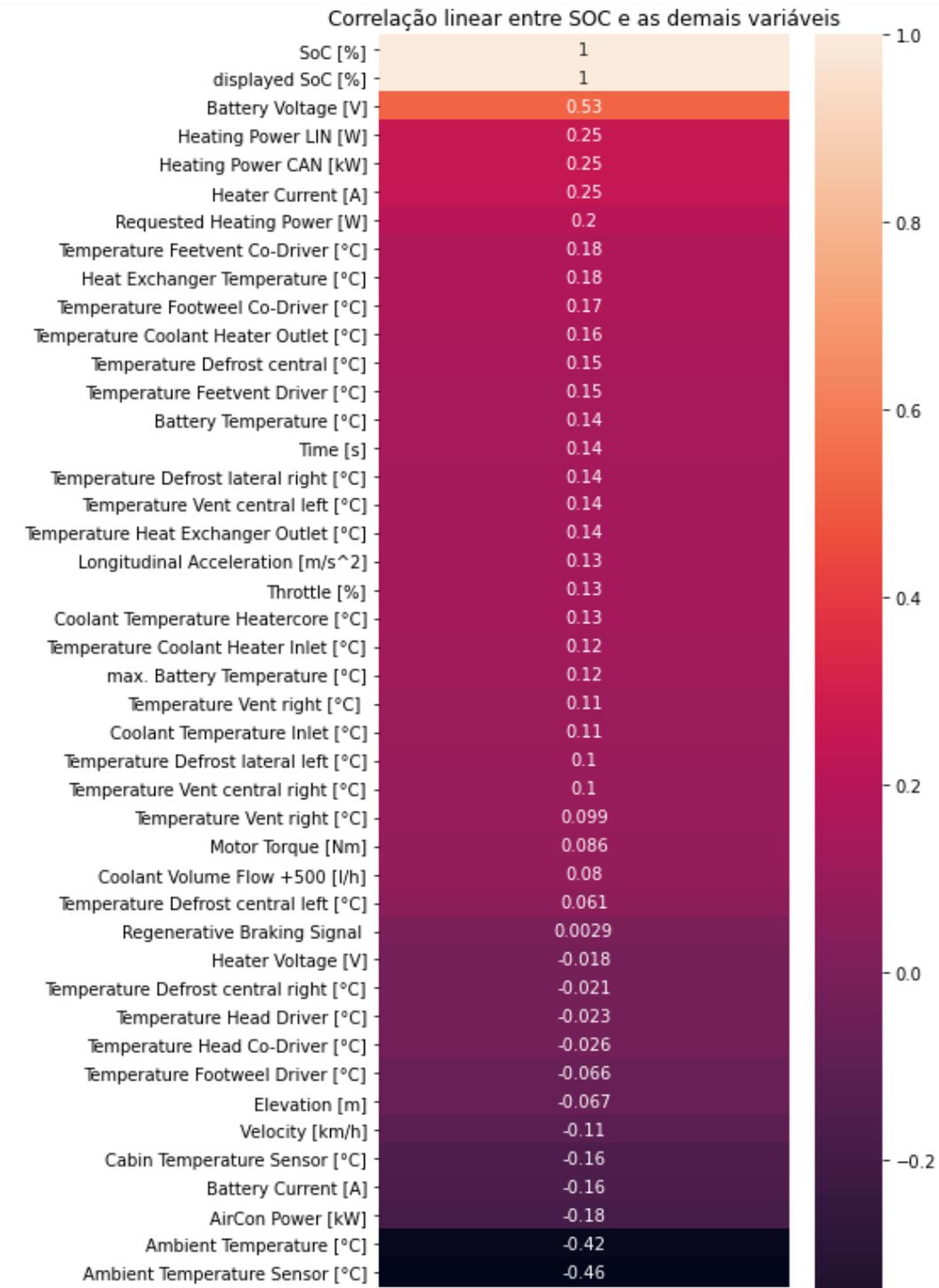


Figura 24: Correlação linear entre SoC e as demais variáveis.

Fonte: Autor, 2021.

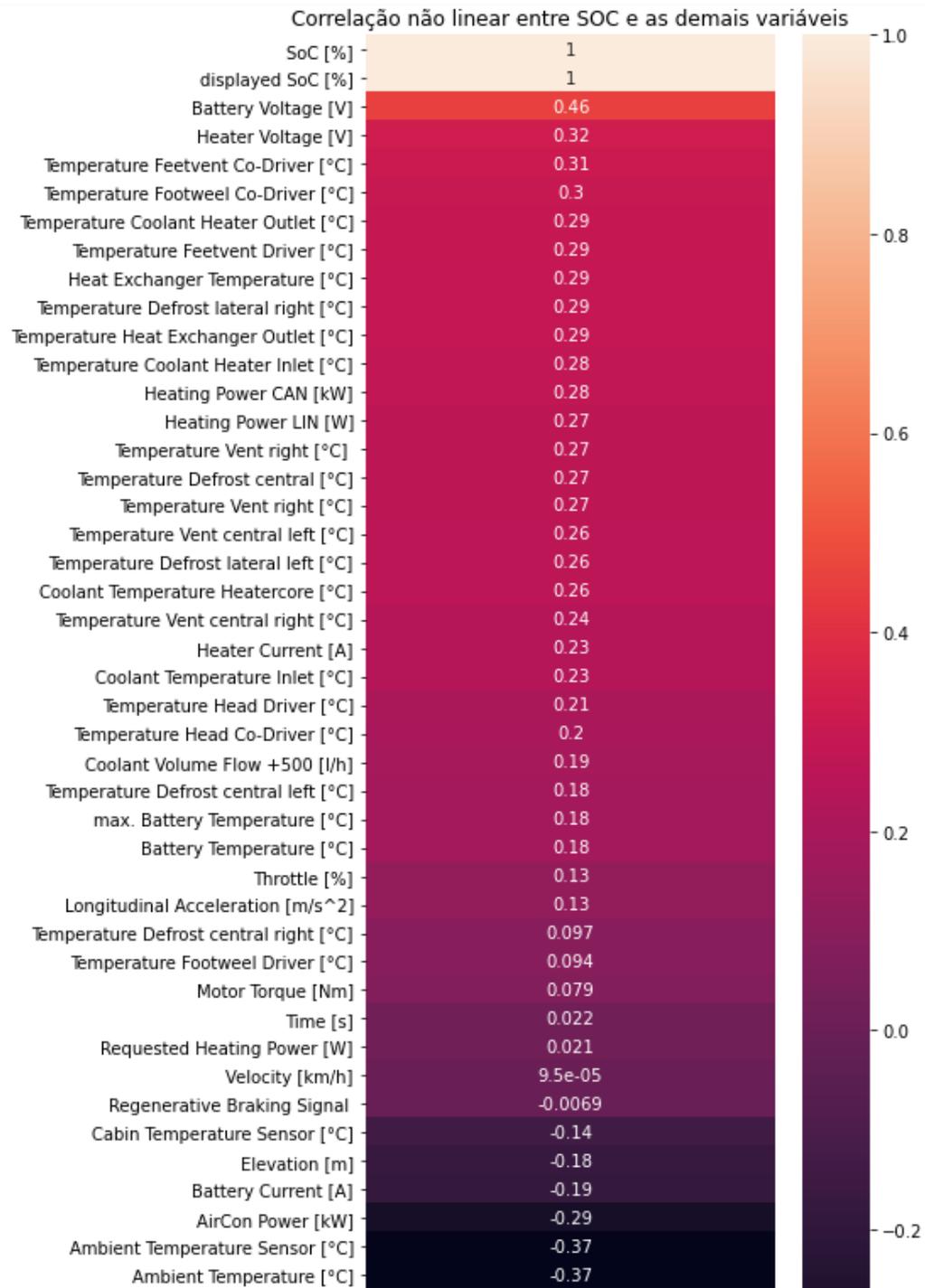


Figura 25: Correlação não linear entre SoC e as demais variáveis.

Fonte: Autor, 2021.

### 3.3. Pseudocódigos dos modelos implementados

Conforme mencionado no início deste capítulo, as rotinas computacionais desenvolvidas para implementação dos modelos de previsão do estado de carga, foram realizadas com o auxílio de bibliotecas específicas no *software Jupyter*, o qual utiliza a linguagem de programação *Python*. Para que seja possível compreender a sequência de ações que cada modelo requer, são apresentados pseudocódigos referentes aos modelos de Regressão Linear Múltipla, Árvore de Decisão e Redes Neurais Artificiais, respectivamente nas seções 3.3.1, 3.3.2 e 3.3.3.

#### 3.3.1. Regressão Linear Múltipla

Para estimar o estado de carga (SOC) através do algoritmo de Regressão Linear, foi necessário realizar os seguintes passos:

**I:** Importar as bibliotecas necessárias:

- *pandas as pd*
- *numpy as np*
- *matplotlib.pyplot as plt*
- *sklearn.metrics as metrics*
- *sklearn.preprocessing import StandardScaler*
- *sklearn.model\_selection import train\_test\_split*
- *sklearn.metrics import r2\_score*
- *sklearn.linear\_model import LinearRegression*

**II:** Importar as variáveis independentes e dependente do banco de dados;

**III:** Separar o conjunto de treinamento e teste em variáveis específicas:

*x\_train, x\_teste, y\_train, y\_test.*

**IV:** Implementar a função Regressão Linear Múltipla com base nestas variáveis:

- *reglin = LinearRegression()*
- *reglin.fit(x\_train, y\_train);*
- *y\_pred = reglin.predict(x\_test);*

**V:** Imprimir os coeficientes lineares;

**VI:** Implementar as funções das métricas de validação, tendo como entrada as seguintes variáveis:

*y\_test, y\_pred;*

**VII:** Plotar os gráficos dos resultados, tendo como variáveis de saída:

*y\_test, y\_pred;*

### 3.3.2. *Árvore de Decisão*

O algoritmo de *Árvore de Decisão* foi implementado através dos seguintes passos:

**I:** Importar as bibliotecas necessárias:

- *pandas as pd*
- *numpy as np*
- *matplotlib.pyplot as plt*
- *seaborn as sns*
- *sklearn.metrics as metrics*
- *sklearn.preprocessing import StandardScaler*
- *sklearn.model\_selection import train\_test\_split*
- *sklearn.metrics import r2\_score*
- *sklearn.linear\_model import SGDRegressor*
- *sklearn.model\_selection import cross\_validate*

**II:** Importar as variáveis independentes e dependente do banco de dados;

**III:** Separar o conjunto de treinamento, teste e definir profundidade da árvore (*noise*):

```
n_train= 150
n_test= 1000
noise=0,1

x_train, x_teste, y_train, y_test.
```

**IV:** Implementar a função *Árvore de Decisão* e configurar seus parâmetros:

- *bdt = BaggingRegressor(DecisionTreeRegressor()),fit(x\_train, y\_train)*
- *bdt\_predict = bdt.predict(x\_test)*

**V:** Implementar as funções das métricas de validação, tendo como entrada as seguintes variáveis:

```
y_test, y_pred;
```

**VI:** Plotar os gráficos dos resultados, tendo como variáveis:

```
y_test, y_pred;
```

### 3.3.3. Redes Neurais Artificiais

Para realizar a predição do estado de carga utilizando Redes Neurais artificiais, foram feitas as seguintes etapas:

**I:** Importar as bibliotecas necessárias:

- *pandas as pd*
- *numpy as np*
- *matplotlib.pyplot as plt*
- *seaborn as sns*
- *sklearn.metrics as metrics*
- *sklearn.preprocessing import StandardScaler*
- *sklearn.model\_selection import train\_test\_split*
- *sklearn.metrics import r2\_score*
- *from sklearn.neural\_network import MLPRegressor*

**II:** Importar as variáveis independentes e depende do banco de dados;

**III:** Separar o conjunto de treinamento e teste em variáveis específicas:

*x\_train, x\_teste, y\_train, y\_test.*

**IV:** Implementar a função Redes Neurais e configurar seus parâmetros:

- *rna = MLPRegressor(random\_state=1,hidden\_layer\_sizes=(50,50),max\_iter=1000)*
- *rna.fit(X\_train, y\_train)*
- *y\_pred = rna.predict(X\_test)*

**V:** Implementar as funções das métricas de validação, tendo como entrada as seguintes variáveis:

*y\_test, y\_pred;*

**VI:** Plotar os gráficos dos resultados, tendo como variáveis:

*y\_test, y\_pred;*

## 4. RESULTADOS E DISCUSSÕES

### 4.1. Introdução

Com base na metodologia proposta anteriormente, no atual capítulo são apresentados os resultados obtidos com as implementações dos modelos para a estimação de carga, SoC. Os testes foram realizados através dos dados obtidos do veículo elétrico BMW I3 (60hA) e que foram devidamente avaliados por uma análise exploratória. Em um primeiro instante, são apresentados os resultados obtidos pelo desenvolvimento do modelo de Regressão Linear. Em seguida, é apresentado o modelo de Árvore de Decisão e seus respectivos testes de variações no modelo. Em terceiro plano, são expostos os resultados obtidos através do algoritmo desenvolvido de Redes Neurais Artificiais e suas variações.

### 4.2. Regressão Linear

A seguir, são apresentados os resultados obtidos através do modelo de Regressão Linear Múltipla, que utiliza como entrada, duas ou mais variáveis preditoras. Desta forma, foram realizados testes alterando-se a quantidade de variáveis de entradas, para que fosse encontrado um melhor modelo de previsão. As tabelas 2 e 3 contém os resultados das métricas de validação do modelo.

Tabela 2: Modelos Regressão Linear Múltipla

Modelo	Modelo RL-I	Modelo RL-II	Modelo RL-III
<b>Dados de Entrada</b>	Battery Voltage [V] Heating Power CAN [kW] Time [s] Heating Power LIN [W] Heater Voltage [V] Temperature Feet vent Co-Driver [°C] Battery Temperature [°C] Ambient Temperature [°C] Elevation [m] Temperature Defrost lateral right [°C] Temperature Footweel Co-Driver [°C]	Battery Voltage [V] AirCon Power [kW] Time [s] Battery Temperature [°C] Ambient Temperature [°C] Battery Current [A] Heater Current [A] Heating Power CAN [kW] Time [s] Heating Power LIN [W] Heater Voltage [V] Temperature Feetvent Co-Driver [°C] Elevation [m] Temperature Defrost lateral right [°C] Temperature Footweel Co-Driver [°C]	Battery Voltage [V] Heating Power CAN [kW] Temperature Feetvent Co-Driver [°C] Battery Current [A] Ambient Temperature [°C] Elevation [m] Velocity [km/h]
<b>N° de Variáveis</b>	11	15	7

Fonte: Autor, 2021.

Tabela 3: Métricas de avaliação para Modelos de Regressão Linear Múltipla

Modelo	MAE	MSE	RMSE	R <sup>2</sup>
Modelo RL-I	4,3581	29,2492	5,4082	0,8441
Modelo RL-II	1,9547	7,2164	2,6863	0,9615
Modelo RL-III	<b>2,0323</b>	<b>6,6645</b>	<b>2,5815</b>	<b>0,9644</b>

Fonte: Autor, 2021.

A partir nas métricas de validação apresentadas na Tabela 3, pode-se verificar a performance dos modelos desenvolvidos. Desta maneira, observa-se que, em um primeiro teste, foram utilizadas as variáveis com base na correlação linear feita na Figura 24. Assim, foi testado um Modelo RL-I com os dados de entradas, que possuíam variáveis com maior correlação linear positiva. Entretanto, os valores das métricas não foram satisfatórios, visto que, foi calculado um R<sup>2</sup> de 0,8441 e uma MAE de 4,3581. Diante disto, a performance deste modelo, foi considerada não satisfatória.

Em um segundo teste, foi acrescentado ao modelo, variáveis com maior correlação linear positiva e maior correção linear negativa, que totalizaram 15 variáveis de entrada, como visto na Tabela 2. Este modelo, obteve resultados satisfatórios de acordo com as métricas calculadas. Foi observado uma melhora significativa no R<sup>2</sup> de 0,9615 e uma diminuição do MAE para 1,9547, que pode ser justificado pelo fato das variáveis com correlação negativa também acrescentarem informações relevantes para a predição do modelo.

O terceiro teste, foi realizado uma análise exploratória das variáveis verificando a multicolinearidade entre elas, e analisando os coeficientes lineares de teste anteriores. Portanto, foram eliminadas algumas variáveis e definidas as mais impactantes para predição do modelo. Feito isso, foi possível desenvolver um modelo com uma acurácia e performance satisfatória. Ao reduzir o número de variáveis de entrada e resultar num menor processamento computacional, é obtido um modelo mais simplificado com êxito em suas predições. Na Figura 26 é apresentado o gráfico das curvas dos dados reais e o valor estimado pela melhor variação do modelo de Regressão Linear testado (RL-III). As curvas referentes aos demais modelos podem ser verificadas no Anexo I.

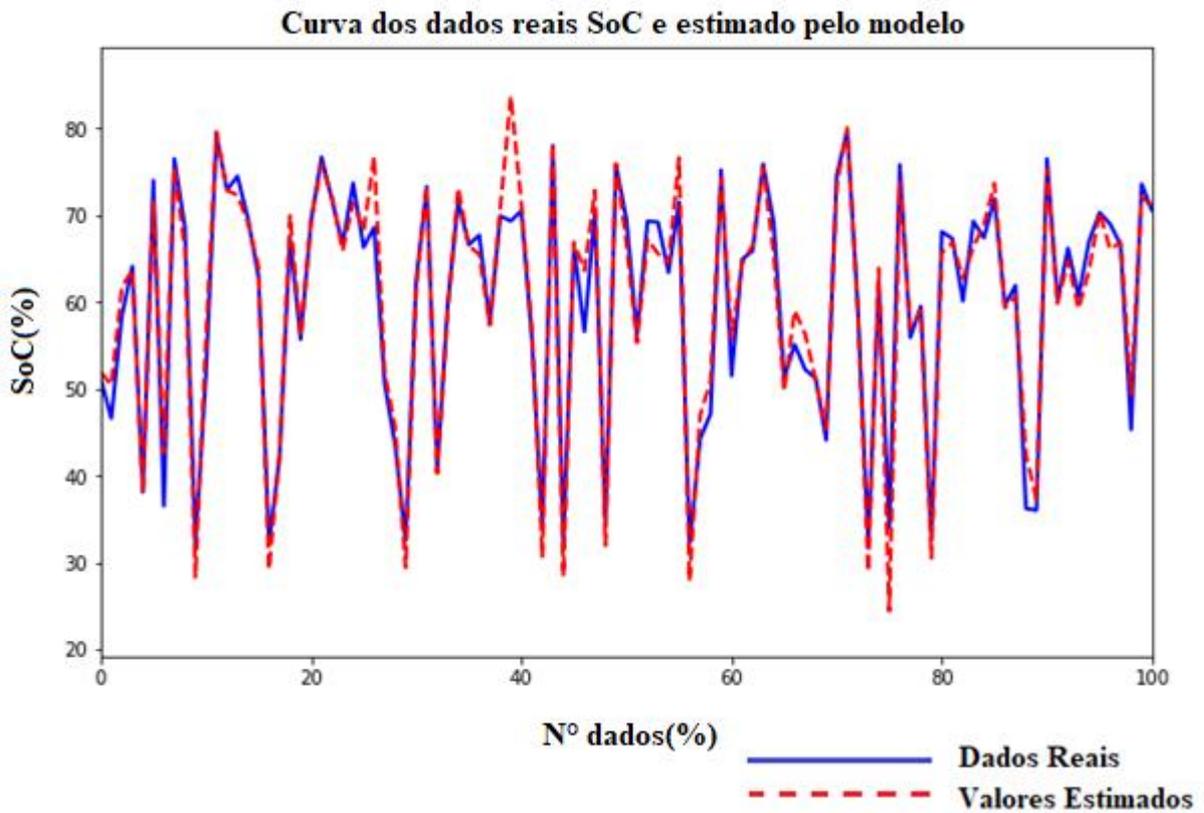


Figura 26: Curva de dados reais e valores previstos (7 variáveis).  
 Fonte: Autor, 2021.

### 4.3. Árvore de Decisão

Em um segundo momento, foi implementado o algoritmo de Árvore de Decisão, que é um modelo mais complexo e que exige um maior processamento computacional. Entretanto, este modelo apresentou resultados mais eficientes. Na Tabela 4, são mostradas as variáveis de entrada e a profundidade de cada modelo de árvore testado. As métricas de validação destes modelos podem ser avaliadas pela Tabela 5.

Por apresentar uma complexidade maior, nestes modelos foram feitos testes variando-se apenas o parâmetro definido como profundidade da árvore. Como dados de entrada, foram utilizadas as sete variáveis mais impactantes que obtiveram melhores performance no modelo de Regressão Linear.

Tabela 4: Resultados do modelo de Árvore de Decisão

Modelo	Modelo AD-I	Modelo AD-II	Modelo AD-III
<b>Dados de Entrada</b>	Battery Voltage [V] Heating Power CAN [kW] Temperature Feetvent Co-Driver [°C] Battery Current [A] Ambient Temperature [°C] Elevation [m] Velocity [km/h]	Battery Voltage [V] Heating Power CAN [kW] Temperature Feetvent Co-Driver [°C] Battery Current [A] Ambient Temperature [°C] Elevation [m] Velocity [km/h]	Battery Voltage [V] Heating Power CAN [kW] Temperature Feetvent Co-Driver [°C] Battery Current [A] Ambient Temperature [°C] Elevation [m] Velocity [km/h]
<b>Profundidade da Árvore</b>	3	6	10

Fonte: Autor, 2021.

Tabela 5: Resultados métricas de avaliação Modelo Árvore de Decisão

Modelo	MAE	MSE	RMSE	R <sup>2</sup>
<b>Modelo AD-I</b>	4,3584	35,3780	5,9479	0,8114
<b>Modelo AD-II</b>	<b>1,4112</b>	<b>5,5173</b>	<b>2,3488</b>	<b>0,9705</b>
<b>Modelo AD-III</b>	0.1573	0.1044	0.3231	0.9994

Fonte: Autor, 2021.

Diante os resultados apresentados na Tabela 5, é visto no primeiro teste, em que, foi limitado a profundidade da árvore em 3 camadas, obteve-se imprecisões significativas, já que as métricas calculadas resultaram em um R<sup>2</sup> de 0,8114 e MAE de 4,35. Portanto, conclui-se que esta configuração, por apresentar uma árvore rasa, não seria a melhor escolha para a solução deste problema, na qual compõe de várias possibilidades de dados de saída.

Com o intuito de melhorar o modelo, foi aumentado para 6 o número máximo de profundidade da árvore. Desta forma, como visto na terceira linha da Tabela 5, foram encontrados resultados mais satisfatórios, apresentando uma precisão de R<sup>2</sup> igual a 0,9705 e MAE de 1,1412.

O terceiro teste foi com uma quantidade maior de camadas de profundidade da árvore, porém, observou-se que as métricas possuem resultados tão satisfatórios que podem ter levado o modelo ao estado de “*overfitting*”, que é um problema muito comum da técnica de árvore de decisão quando não se limita os parâmetros na tentativa de buscar melhoras no modelo.

Portanto, o teste com 6 camadas de profundidade é a melhor variação do modelo de Árvore de Decisão, tendo em vista que apresentou métricas excelentes para estimação do estado de carga, e ao variar sua quantidade de variáveis, o modelo apresenta-se alguma imprecisão, indicando a baixa tendência de “*overfitting*”. Na Figura 27, é ilustrado o gráfico das curvas dos dados reais e o valor estimado pela melhor variação do modelo de Árvore de Decisão testado. As curvas referentes aos demais modelos podem ser verificadas no Anexo I.

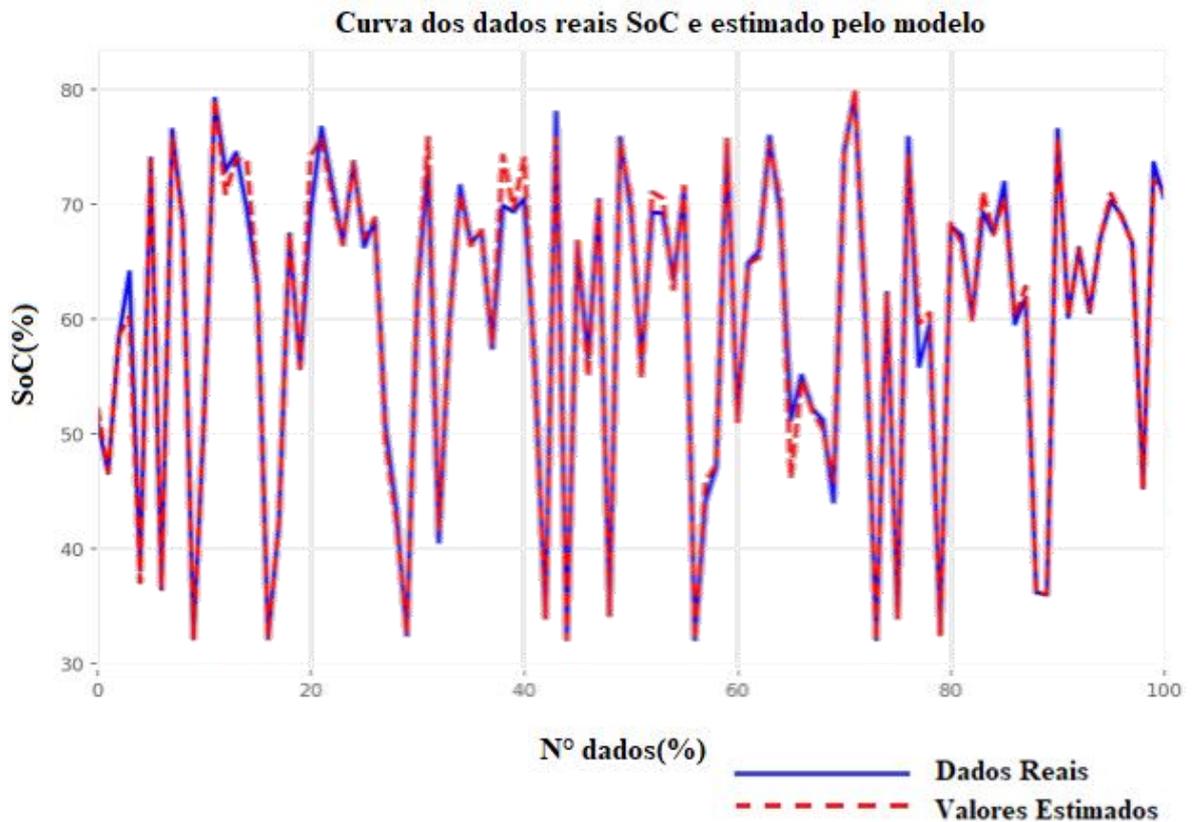


Figura 27: Curva de dados reais e valores previstos (6 camadas de prof.).  
Fonte: Autor, 2021.

#### 4.4. Redes Neurais Artificiais

Foi implementado um modelo de *Perceptron* multicamadas com testes na mesma linha de raciocínio do modelo de Árvore de Decisão, mantendo-se as sete variáveis mais impactantes como dados de entrada e variando-se apenas a arquitetura da RNA. Nas Tabelas 6 e 7, estão apresentados os resultados obtidos.

Tabela 6: Resultados do modelo de Redes Neurais Artificiais

Modelo	Modelo RNA-I	Modelo RNA -II	Modelo RNA -III	Modelo RNA -III
<b>Dados de Entrada</b>	Battery Voltage [V] Heating Power CAN [kW] Temperature Feetvent Co-Driver [°C] Battery Current [A] Ambient Temperature [°C] Elevation [m] Velocity [km/h]	Battery Voltage [V] Heating Power CAN [kW] Temperature Feetvent Co-Driver [°C] Battery Current [A] Ambient Temperature [°C] Elevation [m] Velocity [km/h]	Battery Voltage [V] Heating Power CAN [kW] Temperature Feetvent Co-Driver [°C] Battery Current [A] Ambient Temperature [°C] Elevation [m] Velocity [km/h]	Battery Voltage [V] Heating Power CAN [kW] Temperature Feetvent Co-Driver [°C] Battery Current [A] Ambient Temperature [°C] Elevation [m] Velocity [km/h]
<b>N° de neurônios 1° camada oculta</b>	5	5	10	50
<b>N° de neurônios 2° camada oculta</b>	5	20	20	50

Fonte: Autor, 2021.

Tabela 7: Resultados métricas de avaliação Modelo Redes Neurais Artificiais

Modelo	MAE	MSE	RMSE	R <sup>2</sup>
<b>Modelo RNA-I</b>	<b>1,8198</b>	<b>5,2971</b>	<b>2,3015</b>	<b>0,9717</b>
<b>Modelo RNA-II</b>	1,4181	3,2728	1,8091	0,9825
<b>Modelo RNA-III</b>	1,1544	2,2725	1,5075	0,9878
<b>Modelo RNA-IV</b>	0,8106	1,4156	1,1897	0,9924

Fonte: Autor, 2021.

Como pode-se observar nas Tabelas 6 e 7, os testes com o modelo de RNA também são considerados assertivos e satisfatórios, visto que as métricas do R<sup>2</sup> foram todas próximas de 1 e os valores de MAE, foram relativamente baixos. Esta performance é justificada pelo fato do modelo de RNA, como citado anteriormente, ser robusto e apresentar soluções precisas para problemas complexos.

Entretanto, apresentam a necessidade de grande processamento computacional. Com base nisto, a melhor variação considerada para o modelo RNA de estimação do estado de carga, é o de cinco neurônios na primeira e segunda camada. Pois o mesmo, contém uma precisão satisfatória e é executado de maneira mais ágil, por conter uma menor quantidade de neurônios, o que resulta em uma menor quantidade de cálculo e conseqüentemente a redução do esforço computacional. Na Figura 28, é ilustrado o gráfico das curvas dos dados reais e o valor estimado

pela melhor variação do modelo de Redes Neurais Artificiais. As curvas referentes aos demais modelos podem ser verificadas no Anexo I.

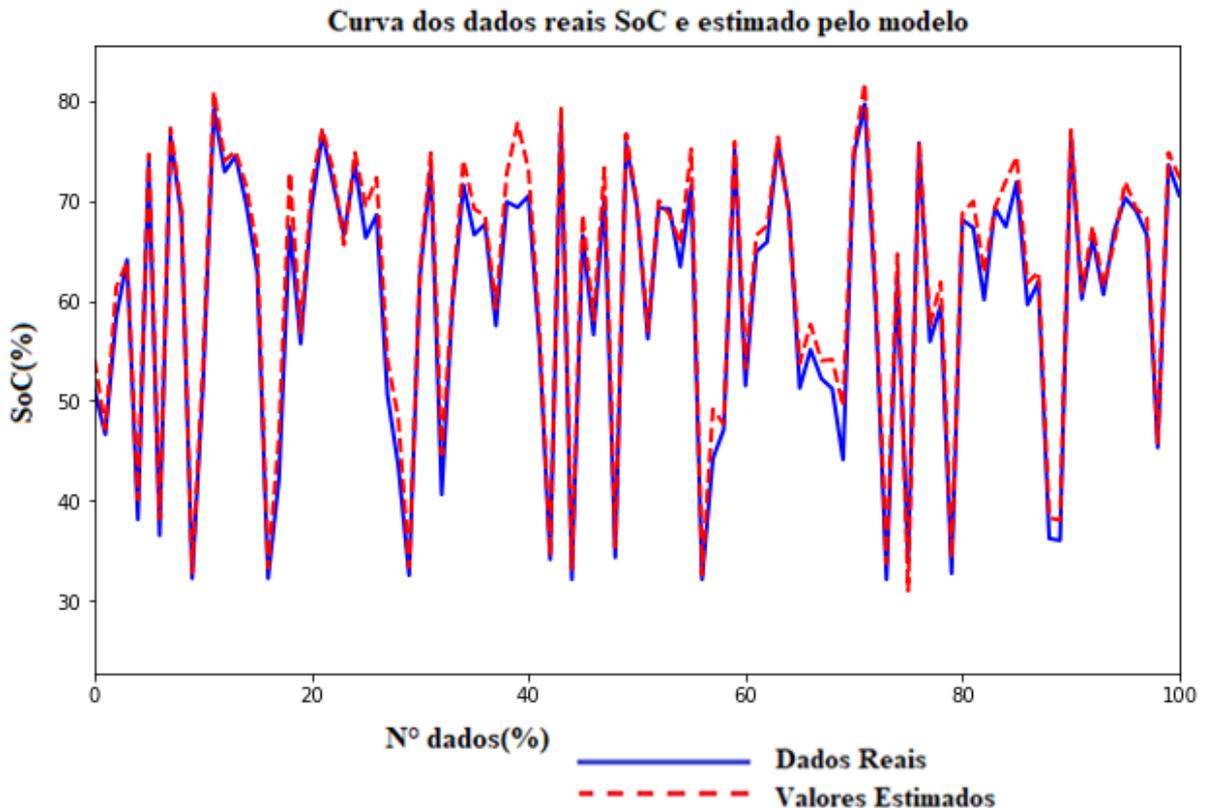


Figura 28: Curva de dados reais e valores previstos (5 neurônios nas camadas ocultas).  
Fonte: Autor, 2021.

#### 4.5. Considerações finais

Pode-se admitir que, a utilização de inteligência computacional e análise de dados proposta neste trabalho, apresentou resultados satisfatórios na estimativa do estado de carga da bateria em veículos elétricos. Isto porque, de acordo com suas métricas, além dos modelos fornecerem resultados aceitáveis, foi possível implantá-los utilizando poucas variáveis e obter modelos simplificados e versáteis. Nas figuras a seguir, estão plotadas as métricas obtidas dos modelos desenvolvidos de Regressão Linear, Árvore de Decisão e RNA.

A métricas MAE, MSE e RMSE, as quais estão relacionadas aos erros de previsão de cada modelo em relação aos dados reais, estão expostas graficamente nas Figuras 29, 30 e 31. Desta forma, é possível comparar cada modelo e sua performance através da análise do quão distante estão dos dados reais. Observa-se que os erros tendem a diminuir ao desenvolvimento dos modelos. Isso se dá pelo fato de testes e ajustes dos parâmetros em busca de melhores performance.



Figura 29: Gráfico MAE.  
Fonte: Autor, 2021.

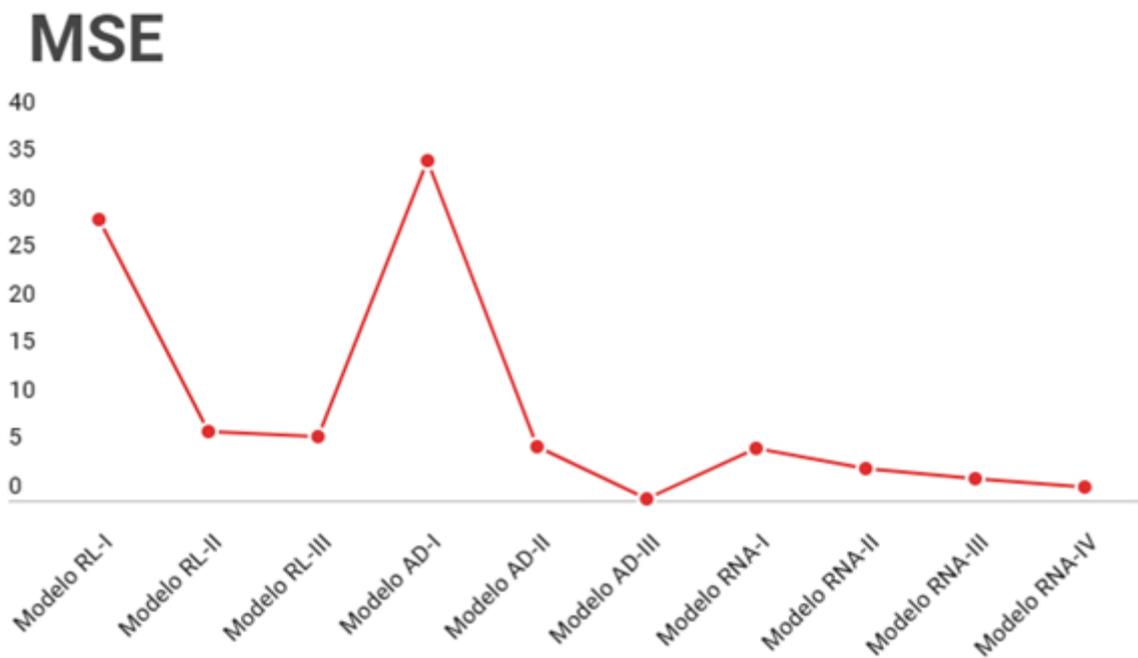


Figura 30: Gráfico MSE.  
Fonte: Autor, 2021.

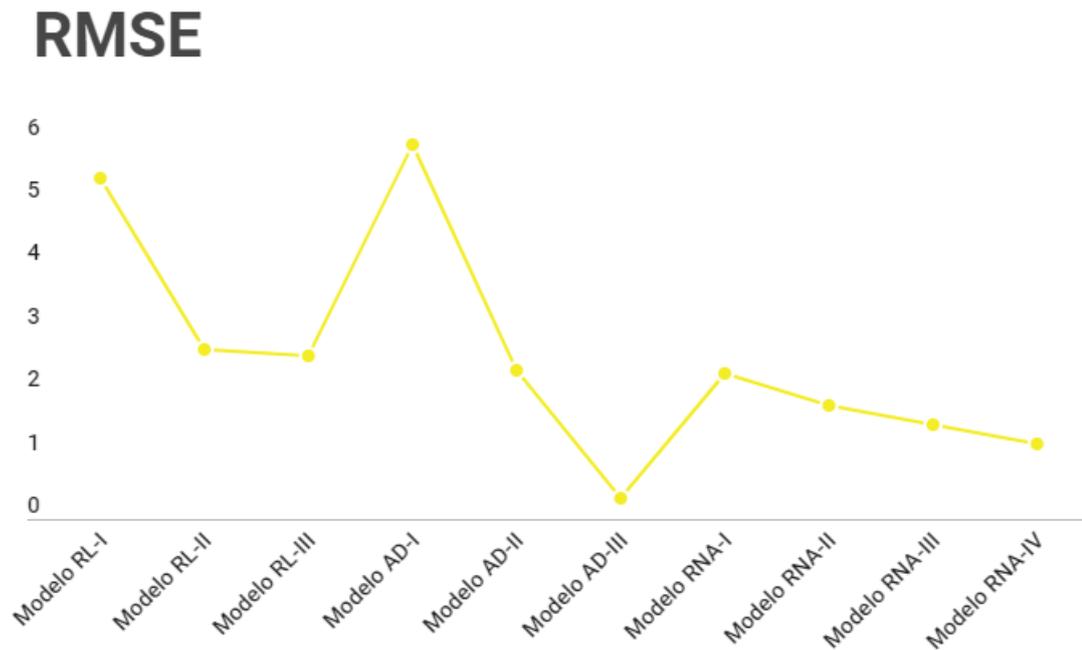


Figura 31: Gráfico RMSE.  
Fonte: Autor, 2021.

A métrica R-quadrado é apresentada na Figura 32, descrevendo a variância de cada modelo implementado. Juntamente com as demais métricas citadas, é possível analisar o potencial de assertividade, sendo que quanto mais próximo do valor unitário, mais precisos são os resultados obtidos.

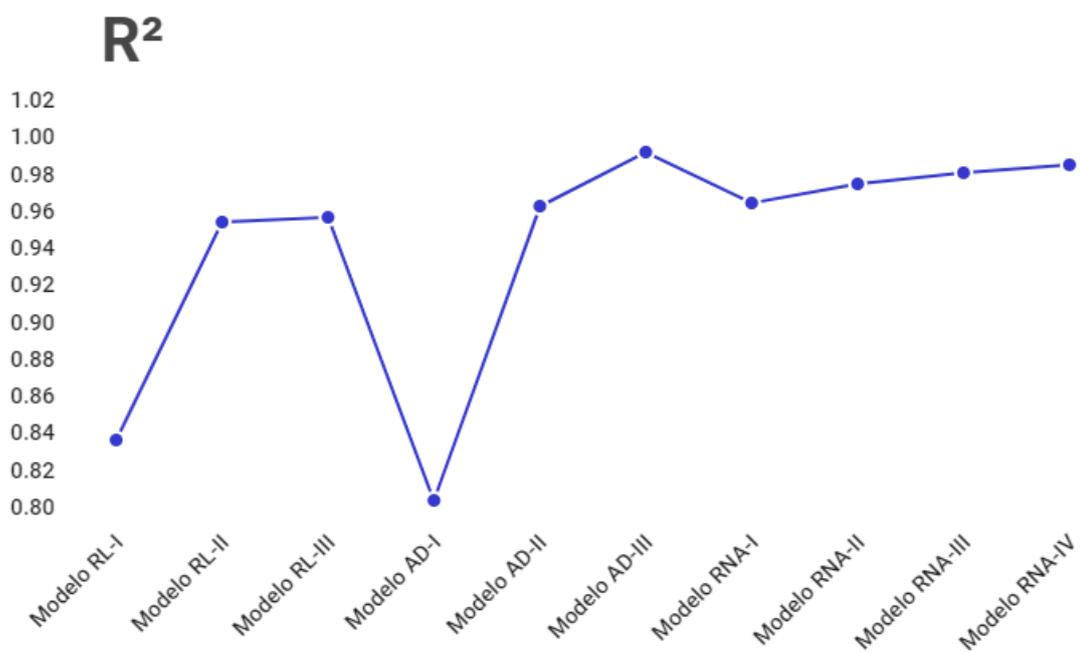


Figura 32: Gráfico R<sup>2</sup>.  
Fonte: Autor, 2021.

Por fim, o histograma da Figura 33 apresenta o agrupamento de todas as métricas de avaliação, utilizadas neste trabalho, as quais estão associadas a cada modelo desenvolvido.

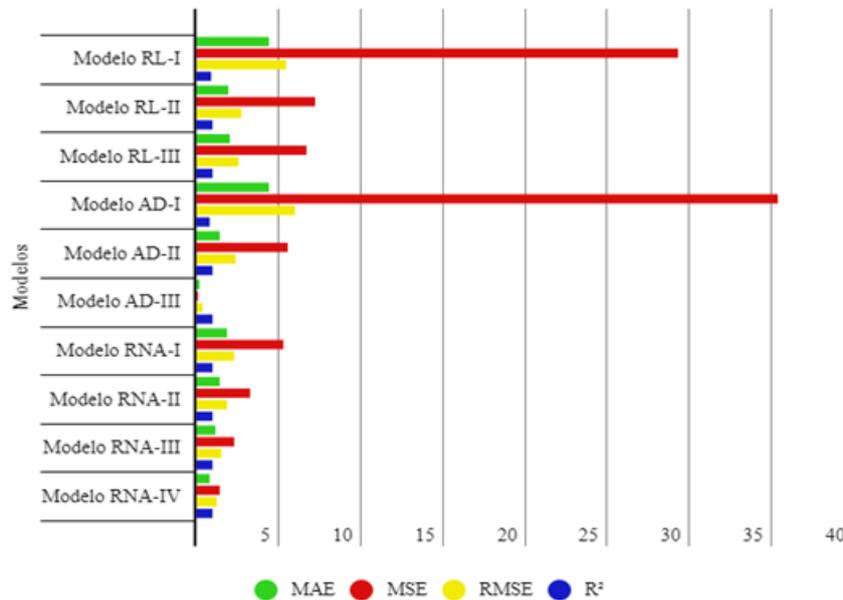


Figura 33: Histograma das métricas de validação  
Fonte: Autor, 2021

Diante dos resultados apresentados, é visto que a maioria dos modelos apresentou resultados satisfatórios. O modelo RL-I e o AD-I não se adaptaram ao problema de previsão proposto. O primeiro por ser um teste inicial, em que as variáveis utilizadas não contribuíram para definição do estado de carga de uma bateria e suas definições ainda estavam em análise. O segundo por apresentar ramos de decisão insuficientes para se adequar à variedade de soluções possíveis do problema apresentado.

## 5. CONCLUSÃO

O presente trabalho apresentou a proposta de implementação de modelos, baseados em inteligência computacional, para estimar o estado de carga (SoC) em baterias de íons de lítio em veículos elétricos. De modo a atender o objetivo proposto, a metodologia utilizou, como objeto de estudo, dados reais disponibilizados pela IEEE DataPort™, em que grandezas, associadas ao funcionamento do automóvel BMW I3 (60hA) em condições diversas, foram medidas por sensores específicos.

A implementação dos modelos propostos foi precedida por uma análise exploratória, visto que, o tratamento prévio dos dados auxilia a busca eficiente para o problema. Foi possível descartar variáveis constantes, dados faltantes e a remoção de *outliers*, que prejudicam a performance de qualquer algoritmo de previsão. Outro fator agregado pelo estudo dos dados, foi compreender a multicolinearidade entre variáveis e destacar as mais impactantes, sendo elas a tensão da bateria (*battery voltage*), corrente da bateria (*battery current*), temperatura ambiente (*ambient temperature*), elevação (*elevation*) e velocidade (*velocity*).

O modelo de Regressão Linear apresentou bons resultados, com um  $R^2$  de 0,9615 e um MAE de 1,9547 no modelo com melhor performance. Estes valores, que se referem às métricas de avaliação, são considerados satisfatórios, levando-se em consideração que o algoritmo é normalmente utilizado para solução de problemas mais simples do que as demais metodologias implementadas.

Já o modelo de Árvore de Decisão obteve resultados melhores, sendo que o cálculo das métricas de avaliação de um dos modelos obteve um  $R^2$  igual a 0,9705 e MAE de 1,1412. Entretanto, durante a realização dos testes, o algoritmo apresentou sinais de *overfitting*, sendo necessário avaliar o limite para o parâmetro de profundidade da árvore que, no caso, foi estabelecido com valor igual a 6.

A implementação de um algoritmo utilizando redes neurais também se mostrou eficaz, apresentando um  $R^2$  de 0,9717 e um MAE de 1,8198 no modelo mais simplificado. Por ser um algoritmo robusto e complexo, mesmo com poucos neurônios, a rede foi capaz de prever a variável SoC com uma boa performance. Os demais testes apresentaram melhoria nas métricas de avaliação, porém não foram considerados satisfatórios devido ao elevado esforço computacional associado ao aumento da quantidade de neurônios em teste.

Portanto, torna-se evidente o fato de que o desenvolvimento de técnicas de inteligência computacional, associado à análise de dados, é de extrema utilidade na engenharia. Isto se justifica pela aplicabilidade no reconhecimento de padrões, previsões ocultas e modelagem de dados altamente voláteis, a qual pode contribuir para encontrar soluções inovadoras para diversos tipos de problema. No caso específico do presente trabalho, tais técnicas se mostraram eficientes na previsão do estado de carga de baterias de veículos elétricos, podendo caracterizar, por exemplo, a definição de parâmetros como autonomia de um veículo elétrico, assim como alertar sobre o estado de saúde da bateria.

Adicionalmente, para validar os modelos propostos, seria necessário avaliar suas performances em relação a outros bancos de dados. Desta maneira, recomenda-se para desenvolvimento de trabalhos futuros, implementar métodos mesclados para estimar o estado de carga, desenvolvimento de software de gerenciamento de baterias em veículos elétricos e aplicação dos métodos em outros tipos de baterias.

## REFERÊNCIAS BIBLIOGRÁFICAS

Aprenda como montar uma árvore de decisão. **Homem Maquina**. 2020. Disponível em: < <https://www.homemmaquina.com.br/como-montar-uma-arvore-de-decisao/> >

AQUINO, Vagner. **Carros elétricos serão obrigatórios na Europa em 2035**. Disponível em: < <https://jornaldocarro.estadao.com.br/carros/carros-eletricos-serao-obrigatorios-na-europa-em-2035/> >

BAPTISTELLA, Marisa. **O uso de redes neurais e regressão linear múltipla na engenharia de avaliações: determinação dos valores venais de imóveis urbanos**. Curitiba, 2005. Disponível em: < <https://acervodigital.ufpr.br/bitstream/handle/1884/3153/DissertMarisa%5B1%5D.pdf?sequ> >

BARR, A.; FEIGENBAUM, E. . *The Handbook of artificial intelligence, volume 1*. Repr. ed. Stanford: Stanford, Calif.: HeurisTech Press ; Los Altos, Calif. : William Kaufmann, 1981. v. 1.

BRAGA, Antonio de Pádua; LUDERMIR, Teresa Bernarda; CARVALHO, André Carlos Ponce de Leon Ferreira. **Redes neurais artificiais: teoria e aplicações**. 2007.

BRASIL. Decreto-lei nº 9.073, de 5 de junho de 2017. **Promulga o Acordo de Paris sob a Convenção-Quadro das Nações Unidas sobre Mudança do Clima, celebrado em Paris, em 12 de dezembro de 2015, e firmado em Nova Iorque, em 22 de abril de 2016**. Disponível em: <[http://www.planalto.gov.br/ccivil\\_03/\\_ato2015-2018/2017/decreto/d9073.html](http://www.planalto.gov.br/ccivil_03/_ato2015-2018/2017/decreto/d9073.html)>.

EZUGWU, Absalom; SHUKLA, Amit; AGBAJE, Moyinoluwa; ADÁN, José García; OLAIDE, Oyelade; AGUSHAKA, Ovre. **Automatic clustering algorithms: a systematic review and bibliometric analysis of relevant literature**. *Neural Computing and Applications*. 2020. Disponível em: < [https://www.researchgate.net/figure/Clustering-example-with-intra-and-inter-clustering-illustrations\\_fig1\\_344590665](https://www.researchgate.net/figure/Clustering-example-with-intra-and-inter-clustering-illustrations_fig1_344590665)>

FILHO, Gustavo José Luna. **Previsão da autonomia de baterias de chumbo-ácido aplicadas a sistemas híbridos de geração de energia utilizando o método KiBaM**. Recife, 2017. Disponível em: < <https://repositorio.ufpe.br/handle/123456789/26342> >

FIORIN, Daniel V.; MARTINS, Fernando R.; SCHUCH , Nelson J.; PEREIRA, Enio B. **Aplicações de redes neurais e previsões de disponibilidade de recursos energéticos solares**. Disponível em: < <https://www.scielo.br/j/rbef/a/dxMmkDfLHjWKssfQDbtWfF/abstract/?lang=en> >

FREITAS, Liel Neves. **Aplicação De Técnicas De Árvore De Decisão Em Dados De Sistema De Potência**. Disponível em : < [https://bdm.ufpa.br:8443/jspui/bitstream/prefix/654/1/TCC\\_AplicacaoTecnicasArvore.pdf](https://bdm.ufpa.br:8443/jspui/bitstream/prefix/654/1/TCC_AplicacaoTecnicasArvore.pdf) >

GALLO, Crescenzo. *Reti Neurali Artificiali Tutorial*. 2016. Disponível em: < [https://www.researchgate.net/publication/289522531\\_Reti\\_Neurali\\_Artificiali\\_-\\_Tutorial](https://www.researchgate.net/publication/289522531_Reti_Neurali_Artificiali_-_Tutorial) >

GRAUPE, Daniel. *Principles of Artificial Neural Networks*. Singapore: World Scientific Publishing, 1997

HANNAN, M. A. et al. *A review of lithium-ion battery state of charge estimation and management system in electric vehicle applications: Challenges and recommendations. Renewable and Sustainable Energy Reviews*, 2017. Disponível em: < <https://www.sciencedirect.com/science/article/abs/pii/S1364032117306275> >

JESS, Gil Marcos. **Inteligência Artificial E Tecnologias Da Inteligência Um Repensar Segundo Os Processos De Elaboração Matemática**. Curitiba, 2004 Disponível em: < <https://acervodigital.ufpr.br/bitstream/handle/1884/32969/R%20-%20D%20-%20GIL%20MARCOS%20JESS.pdf?sequence=1&isAllowed=y> >

JUNGES, Estevan Luiz. **Detecção E Classificação De Faltas Em Linhas De Transmissão Utilizando Redes Neurais Artificiais**. Lajeado, 2018. Disponível em: < <https://www.univates.br/bdu/bitstream/10737/2235/1/2018EstevanJunges.pdf> >

MATTHIAS, Steintraeter, JOHANNES, Buberger, DIMITAR, Trifonov. *Battery and Heating Data in Real Driving Cycles*, 2020 .Disponível em: < <https://iee-dataport.org/open-access/battery-and-heating-data-real-driving-cycles> >

MCCARTHY J. and HAYES P.J. *Some philosophical problems from the standpoint of artificial intelligence*. Edinburgh University Press, Edinburgh, GB, 1969.

MORETO, Miguel. **Localização de Faltas de Alta Impedância em Sistemas de Distribuição de Energia: Uma Metodologia Baseada em Redes Neurais Artificiais**. Porto Alegre, 2005. Disponível em: < <https://www.lume.ufrgs.br/handle/10183/7602> >

OLIVEIRA, Ângelo Rocha. **Redes Neurais Artificiais Aplicadas na Detecção, Classificação e Localização de Defeitos em Linhas de Transmissão**. 2005

PINHO, João Tavares et al. **Manual de Engenharia para Sistemas Fotovoltaicos**. Rio de Janeiro: CEPEL – CRESESB, 2014. Disponível em: < <https://www.portal-energia.com/downloads/livro-manual-de-engenharia-sistemas-fotovoltaicos-2014.pdf> >

RAHIMI-EICHI, H. et al. **Battery management system: An overview of its application in the smart grid and electric vehicles**. Disponível em: < <https://ieeexplore.ieee.org/document/6532486> >

ROCHA, Germana Adrielle. **Proposta De Um Novo Modelo De Regressão Linear Para Previsão E Controle Do Indicador De Eficiência Energética De Uma Empresa Ferroviária**. Disponível em: < [https://www2.ufjf.br/engenhariadeproducao//files/2014/09/Germana-Adrielle-Rocha\\_2013.pdf](https://www2.ufjf.br/engenhariadeproducao//files/2014/09/Germana-Adrielle-Rocha_2013.pdf) >

RODRIGUES, Sandra Cristina Antunes. **Modelo de Regressão Linear e suas Aplicações**. Covilhã, 2012. Disponível em: < <https://ubibliorum.ubi.pt/bitstream/10400.6/1869/1/Tese%20Sandra%20Rodrigues.pdf> >

SANTOS, Célia Aparecida Lino. **Baterias de íons lítio para veículos elétricos**. São Paulo, 2018 Disponível em: < <http://revista.ipt.br/index.php/revistaIPT/article/view/71> >

SANTOS, Rafael de Melo Silva. **Estimação do Estado de Carga de Baterias de Lítio-íon em Diferentes Condições de Temperaturas Utilizando Filtros de Kalman**. 2019. Disponível em: < [https://repositorio.ufpb.br/jspui/handle/123456789/19499?locale=pt\\_BR](https://repositorio.ufpb.br/jspui/handle/123456789/19499?locale=pt_BR) >

SHAREEF, H.; ISLAM, M. M.; MOHAMED, A. *A review of the stage-of-the-art charging technologies, placement methodologies, and impacts of electric vehicles. Renewable and Sustainable Energy Reviews.* Disponível em: <<https://www.sciencedirect.com/science/article/abs/pii/S1364032116302568> >

SILVA, Nelson Alexandre Galione. **Influência do hidrócloro de polialilamina e nanopartículas de prata na eletro-inserção de íons lítio em matrizes hospedeiras de V2O.** Disponível em: <<https://www.teses.usp.br/teses/disponiveis/59/59138/tde-23012012-184127/pt-br.php>>

VIEIRA, Rafaela Martins. **Classificação de Mamografias Utilizando Redes Neurais Convolucionais.** Disponível em: <<https://www.formiga.ifmg.edu.br/documents/2020/Biblioteca/Rafaela-monografia-versao-final.pdf>>

**ANEXO I – Curvas de comparação entre os dados reais e previstos pelos modelos complementares.**

**ANEXO I (a) – Modelos de Regressão Linear Múltipla.**

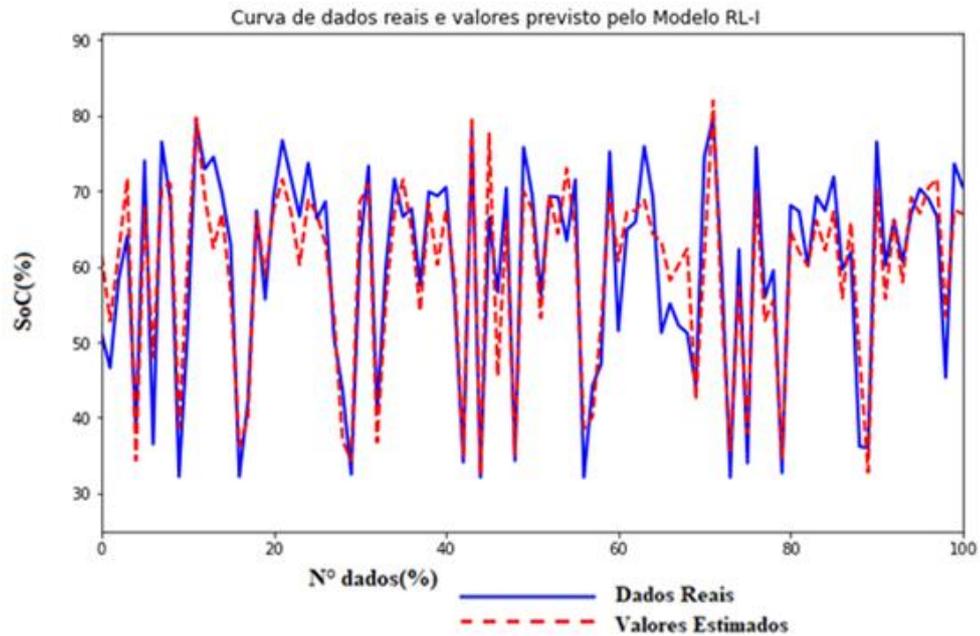


Figura 34: Curva de dados reais e valores previstos pelo Modelo RL-I.  
Fonte: Autor, 2021.

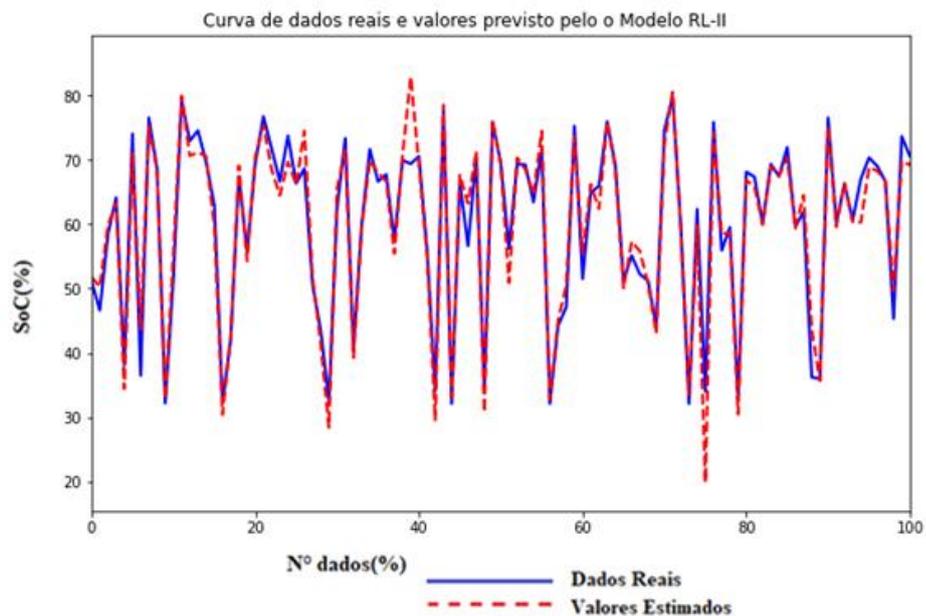


Figura 35: Curva de dados reais e valores previstos pelo Modelo RL-II.  
Fonte: Autor, 2021.

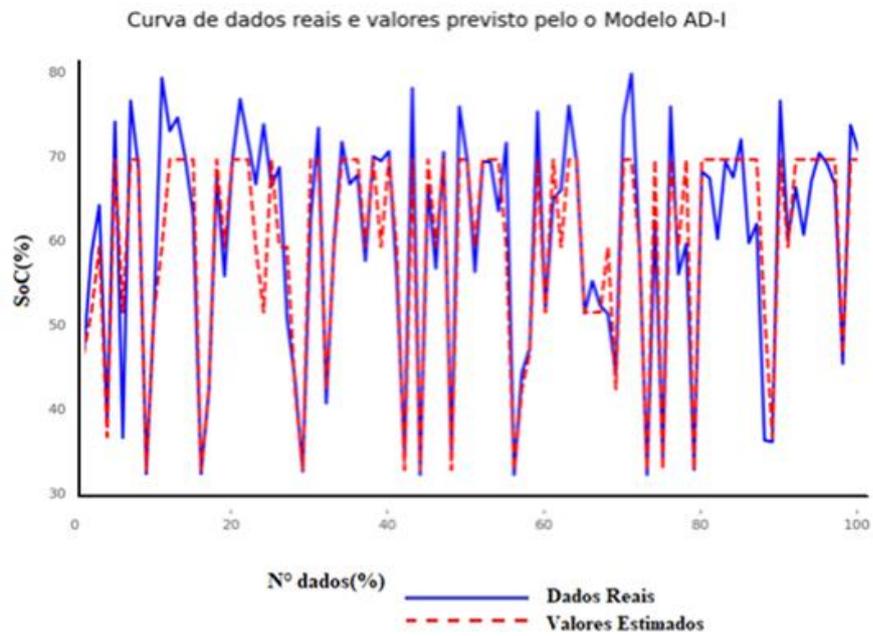
**ANEXO I (b) – Modelos de Árvore de Decisão.**

Figura 36: Curva de dados reais e valores previstos pelo Modelo AD-I.  
Fonte: Autor, 2021.

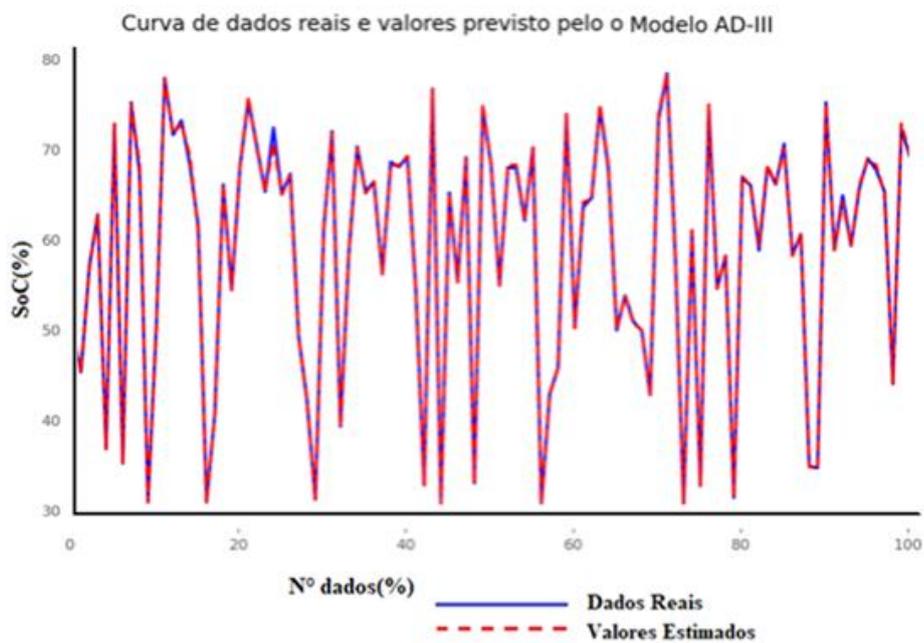


Figura 37: Curva de dados reais e valores previstos pelo Modelo AD-III.  
Fonte: Autor, 2021.

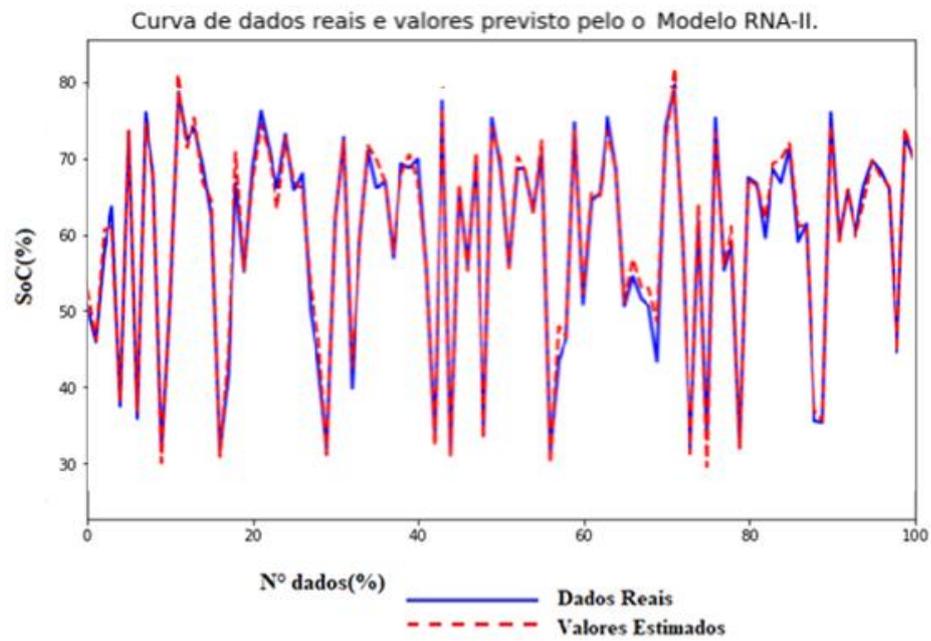
**ANEXO I (c) – Modelos de Redes Neurais Artificiais.**

Figura 38: Curva de dados reais e valores previstos pelo Modelo RNA-II.  
Fonte: Autor, 2021.

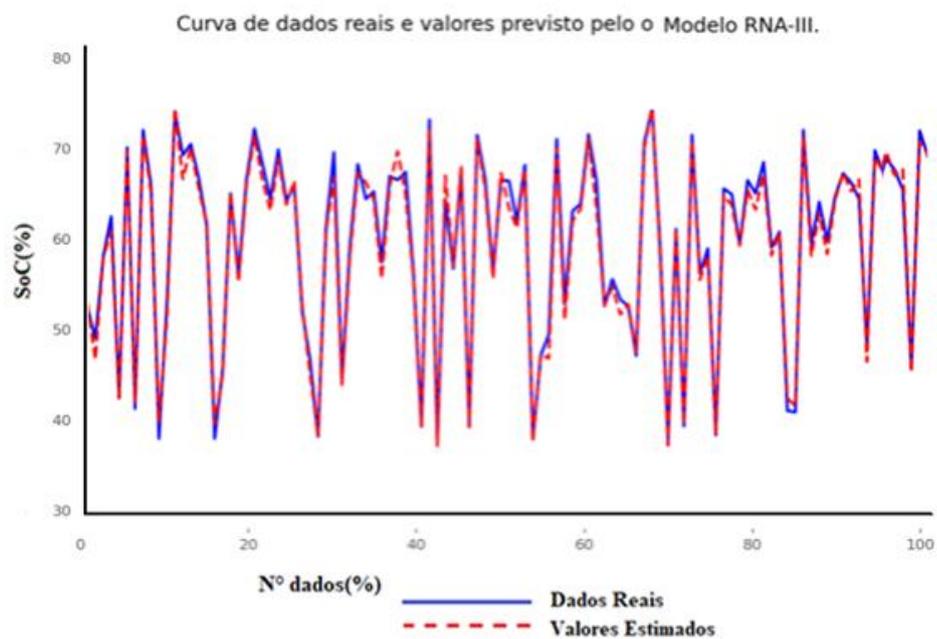


Figura 39: Curva de dados reais e valores previstos pelo Modelo RNA-III.  
Fonte: Autor, 2021.

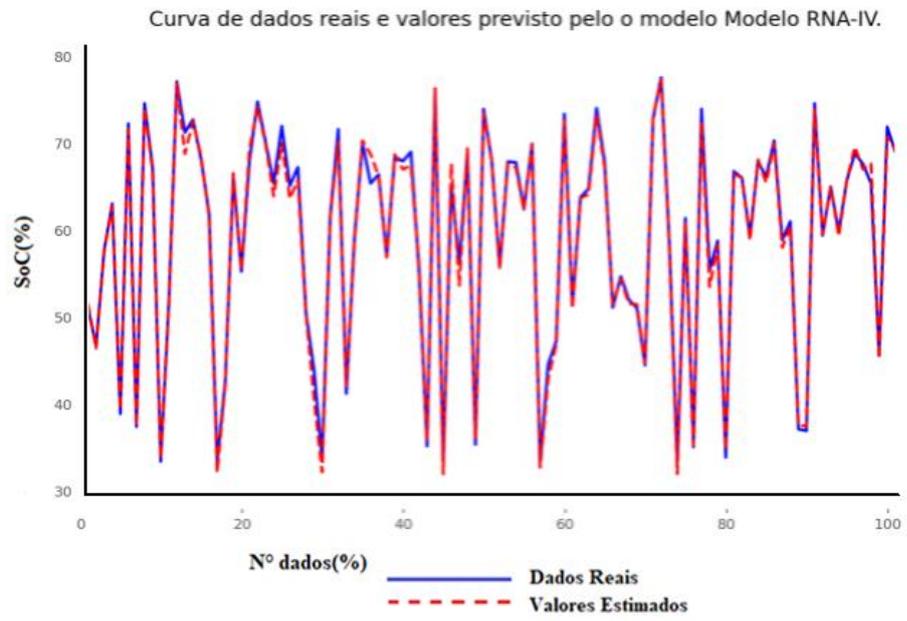


Figura 40: Curva de dados reais e valores previstos pelo Modelo RNA-IV.  
Fonte: Autor, 2021.