



**UNIVERSIDADE ESTADUAL DA PARAÍBA**  
**CAMPUS I - CAMPINA GRANDE**  
**PRÓ-REITORIA DE PÓS-GRADUAÇÃO E PESQUISA**  
**PROGRAMA DE PÓS-GRADUAÇÃO EM CIÊNCIA E TECNOLOGIA AMBIENTAL**  
**MESTRADO EM CIÊNCIA E TECNOLOGIA AMBIENTAL**

**MARCOS RENATO BASILIO AMORIM**

**ESTUDO COMPARATIVO DA REGRESSÃO LINEAR MÚLTIPLA E REDES  
NEURAS ARTIFICIAIS NA MODELAGEM DO PROCESSO DE ADSORÇÃO**

**CAMPINA GRANDE – PB**  
**MARÇO DE 2022**

MARCOS RENATO BASILIO AMORIM

**ESTUDO COMPARATIVO DA REGRESSÃO LINEAR MÚLTIPLA E REDES  
NEURAS ARTIFICIAIS NA MODELAGEM DO PROCESSO DE ADSORÇÃO**

Dissertação apresentada ao Programa de Pós-Graduação em Ciência e Tecnologia Ambiental (PPGCTA), da Universidade Estadual da Paraíba, como requisito parcial para obtenção do título de Mestre em Ciência e Tecnologia Ambiental.

**Área de concentração: Tecnologias de  
Tratamento de Água e Resíduos**

**Orientador:** Prof. Dr. Fernando Fernandes Vieira

**CAMPINA GRANDE - PB  
MARÇO DE 2022**

É expressamente proibido a comercialização deste documento, tanto na forma impressa como eletrônica. Sua reprodução total ou parcial é permitida exclusivamente para fins acadêmicos e científicos, desde que na reprodução figure a identificação do autor, título, instituição e ano do trabalho.

A524e Amorim, Marcos Renato Basílio.

Estudo comparativo da regressão linear múltipla e redes neurais artificiais na modelagem do processo de adsorção [manuscrito] / Marcos Renato Basílio Amorim. - 2022.  
136 p. : il. colorido.

Digitado.

Dissertação (Mestrado em Ciência e Tecnologia Ambiental) - Universidade Estadual da Paraíba, Centro de Ciências e Tecnologia, 2022.

"Orientação : Prof. Dr. Fernando Fernandes Vieira , Departamento de Engenharia Sanitária e Ambiental - CCT."

1. Corantes têxteis. 2. Resíduos industriais. 3. Adsorção.  
4. Regressão linear múltipla. 5. Redes neurais artificiais. I.  
Título

21. ed. CDD 628.4

MARCOS RENATO BASILIO AMORIM

**ESTUDO COMPARATIVO DA REGRESSÃO LINEAR MÚLTIPLA E REDES  
NEURAS ARTIFICIAIS NA MODELAGEM DO PROCESSO DE ADSORÇÃO**

Dissertação apresentada ao Programa de Pós-Graduação em Ciência e Tecnologia Ambiental (PPGCTA), da Universidade Estadual da Paraíba, como requisito parcial para obtenção do título de Mestre em Ciência e Tecnologia Ambiental.

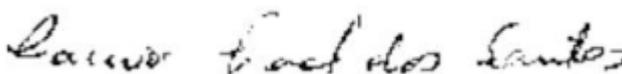
Área de concentração: Tecnologias de Tratamento de Água e Resíduos.

Aprovada em: 25/02/2022.

**BANCA EXAMINADORA**



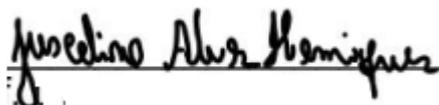
Prof. Dr. Fernando Fernandes Vieira (Orientador)  
Universidade Estadual da Paraíba (UEPB)



Prof. Dr. Laércio Leal dos Santos  
Universidade Estadual da Paraíba (UEPB)



Prof. Dr. William de Paiva  
Universidade Estadual da Paraíba (UEPB)



Prof. Dr. Juscelino Alves Henriques  
Instituto Federal de Pernambuco (IFPE)

Dedico primeiramente a Deus, pela força, esperança e fé em sempre alcançar novos objetivos e vencer os desafios que surgem. Aos meus pais, Eliete Basilio Amorim e José Carlos Amorim, por todo apoio, afeto e dedicação que me forneceram durante essa jornada. Ao meu irmão e minhas irmãs e amigos e amigas que caminharam comigo durante esses dois anos de trabalho. Ao meu orientador, prof. Dr. Fernando Fernandes Vieira, pelo aprendizado e crescimento profissional proporcionado neste tempo de convivência, do qual tenho a certeza que sem ele não poderia ser possível a conclusão deste trabalho.

## **AGRADECIMENTOS**

Ao Dr. Wilton Silva Lopes, coordenador Programa de Pós Graduação em Ciência e Tecnologia Ambiental da UEPB, por seu empenho.

Ao professor Dr. Fernando Fernandes Vieira pelas leituras sugeridas ao longo dessa orientação e pela dedicação.

A minha mãe Eliete Basilio Amorim, e ao meu pai José Carlos Amorim, ao meu irmão e minhas irmãs e aos meus sobrinhos, pela compreensão por minha ausência nas reuniões familiares.

Aos professores do Programa de Pós Graduação em Ciência e Tecnologia Ambiental da UEPB, em especial, Láercio Leal dos Santos, William de Paiva, que contribuíram ao longo de dois anos, por meio das disciplinas e debates, para o desenvolvimento desta pesquisa.

À secretária, Nadilma Alves Cunha, pela presteza e atendimento quando nos foi necessário.

Aos colegas de classe pelos momentos de amizade e apoio.

## RESUMO

O setor têxtil é um ramo de atividade que vem se destacando ao longo dos anos no Brasil. Todavia, discute-se muito sobre o consumo desses produtos, mas pouco se fala sobre os resíduos gerados ao longo de toda a produção. Resíduos estes que muitas vezes são descartados de forma inadequada e causam sérios prejuízos ao meio ambiente, em especial aos corpos hídricos. Dentre esses resíduos é possível citar os efluentes descartados pelas indústrias têxteis. O processo de adsorção pode ser utilizado como uma alternativa para o tratamento de efluente do setor. A proposta desse trabalho é a aplicabilidade da modelagem matemática por meio de duas metodologias: a Análise de Regressão que fez uso da Regressão Linear Múltipla e as Redes Neurais Artificiais em dados experimentais de adsorção de corantes têxteis, como forma de identificar um modelo que se ajuste aos dados e verifique a eficiência da adsorção. O treinamento do modelo foi realizado com o objetivo de ajustar o melhor modelo. Sete variáveis foram utilizadas na aplicação da RLM e da RNA, sendo cinco variáveis de controle: massa, concentração, pH, rotação e tempo; e duas de respostas: unidade de massa adsorvente e eficiência de remoção. A melhor resposta da RLM para a variável unidade de massa de adsorvente foi observada no tratamento ácido para treino com 40% e teste com 30% de validação. Para o treino o  $R_a^2$  apresentou uma média de aproximadamente 84% de variação para explicação do modelo, no teste foi obtido um valor médio de aproximadamente 84% para o  $R_a^2$ . A variável eficiência de remoção apresentou para a melhor resposta no tratamento alcalino com 30% de validação onde foi obtido um  $R_a^2$  de aproximadamente 58% para teste, enquanto para treino com 30% de validação foi obtido um valor de  $R_a^2$  de 56,26%. Para a RNA com uma camada foram obtidos os seguintes resultados: para a variável unidade de massa de adsorvente o tratamento in-naturra apresentou um  $R_a^2$  de 99,67% enquanto para a eficiência de remoção no tratamento ácido foi obtido um  $R_a^2$  de 99,85% ambos com 40% de validação dos dados. A RNA com duas camadas apresentou um  $R_a^2$  (Y1) de 99,73% para o tratamento in-natura com 30% de validação, enquanto o  $R_a^2$  (Y2) foi de 99,93% para o tratamento ácido com 40% de validação. A RNA com uma e duas camadas permite obter melhores valores para a variação do ajuste do modelo quando comparada com a RLM.

**PALAVRAS CHAVE:** Corantes Têxteis. Resíduos Industriais. Adsorção. Regressão Linear Múltipla. Redes Neurais Artificiais.

## ABSTRACT

The textile sector is a branch of activity that has been standing out over the years in Brazil. However, there is a lot of discussion about the consumption of these products, but little is said about the waste generated throughout the production. These residues are often improperly discarded and cause serious damage to the environment, especially to water bodies. Among these residues, it is possible to mention the effluents discarded by the textile industries. The adsorption process can be used as an alternative for the sector's effluent treatment. The purpose of this work is the applicability of mathematical modeling through two methodologies: Regression Analysis that made use of Multiple Linear Regression and Artificial Neural Networks in experimental data of adsorption of textile dyes, as a way of identifying a model that fits to the data and check the adsorption efficiency. Model training was performed in order to fit the best model. Seven variables were used in the application of RLM and ANN, being five control variables: mass, concentration, pH, rotation and time; and two responses: adsorbent mass unit and removal efficiency. The best response of the RLM for the variable adsorbent mass unit was observed in the acid treatment for training with 40% and test with 30% of validation. For training,  $R_a^2$  presented an average of approximately 84% of variation to explain the model, in the test an average value of approximately 84% was obtained for  $R_a^2$ . The removal efficiency variable presented the best response in the alkaline treatment with 30% of validation, where an  $R_a^2$  of approximately 58% was obtained for testing, while for training with 30% of validation, an  $R_a^2$  value of 56 was obtained 26%. For the RNA with one layer, the following results were obtained: for the variable adsorbent mass unit, the in-natura treatment presented an  $R_a^2$  of 99.67%, while for the removal efficiency in the acid treatment, an  $R_a^2$  was obtained. of 99.85% both with 40% of data validation. The RNA with two layers showed an  $R_a^2$ (Y1) of 99.73% for the in-natura treatment with 30% validation, while the  $R_a^2$  (Y2) was 99.93% for the acid treatment with 40 % validation. The ANN with one and two layers allows to obtain better values for the variation of the model fit when compared to the RLM.

**KEYWORDS:** Textile Dyes. Industrial Waste. Adsorption. Regression Linear Multiple. Artificial Neural Networks.

## LISTA DE FIGURAS

Figura 1 – Estrutura do Neurônio Biológico.....	28
Figura 2 – Modelo de um Neurônio Artificial.....	29
Figura 3(a) – Arquitetura <i>FeedForWard</i> de camadas simples.....	31
Figura 3(b) – Arquitetura <i>FeedForWard</i> de camadas múltiplas.....	31
Figura 3(c) – Arquitetura Realimentada.....	31
Figura 3.1–Funções de ativação: a) limiar b) sigmóide c) tangente hiperbólica.....	39
Figura 4 - Bagaço de cana de açúcar (a) antes da preparação(b) depois da preparação.....	45
Figura 4.1- (a) Rede FeedForward de Camadas Múltiplas.....	46
Figura 4.1- (b) Rede FeedForward de Camadas Múltiplas duas saídas.....	46
Figura 5 - Gráficos de resíduos e homocedasticidade de Qt .....	54
Figura 6 - Gráficos de resíduos e homocedasticidade de Rem.....	55
Figura 7 - Gráficos ácido 30% (a) e (b) e in-natura 40% (c) e (d) da Qt.....	60
Figura 8 - Gráficos ácido 30% (a) e (b) e (c) e (d) 40% da Qt.....	61
Figura 9 - Gráficos de Pareto do MAE-T e MAE-V in-natura Rem.....	64
Figura 10 - Gráficos de Pareto do $R^2$ e $R_a^2$ in-natura Rem.....	64
Figura 11 - Gráficos fatoriais dos níveis MAE-T e MAE-V in-natura Rem.....	65
Figura 12 - Gráficos fatoriais dos níveis $R^2$ e $R_a^2$ in-natura Rem.....	65
Figura 13 - Gráficos de Resíduo MAE-T e MAE-V in-natura Rem.....	66
Figura 14 - Gráficos de Resíduo $R^2$ e $R_a^2$ in-natura Rem.....	66
Figura 15 - Gráficos de Pareto do MAE-T e MAE-V ácido Rem.....	68
Figura 16 - Gráficos de Pareto do $R^2$ e $R_a^2$ ácido Rem.....	68
Figura 17 - Gráficos fatoriais dos níveis MAE-T e MAE-V ácido Rem.....	69
Figura 18 - Gráficos fatoriais dos níveis $R^2$ e $R_a^2$ ácido Rem.....	69
Figura 19 - Gráficos de Resíduos MAE-T e MAE-V ácido Rem.....	70
Figura 20 - Gráficos de Resíduos $R^2$ e $R_a^2$ ácido Rem.....	70
Figura 21 - Gráficos de Pareto do MAE-T e MAE-V alcalino Rem.....	72
Figura 22 - Gráficos de Pareto do $R^2$ e $R_a^2$ alcalino Rem.....	72
Figura 23 - Gráficos fatoriais dos níveis MAE-T e MAE-V alcalino Rem.....	73
Figura 24 - Gráficos fatoriais dos níveis $R^2$ e $R_a^2$ alcalino Rem.....	73
Figura 25 - Gráficos de Resíduos MAE-T e MAE-V alcalino Rem.....	74
Figura 26 - Gráficos de Resíduos $R^2$ e $R_a^2$ alcalino Rem.....	74
Figura 27 - Gráfico ajuste da RNA e resíduo níveis $R_a^2$ e MAE-V.....	75
Figura 28 - Gráfico ajuste da RNA e resíduo do nível MAE-T.....	76

Figura 29 - Gráfico ajuste da RNA e resíduo do nível $R_a^2$ e MAE-V.....	76
Figura 30 - Gráfico ajuste da RNA e resíduo do nível MAE-T.....	77
Figura 31 - Gráfico ajuste da RNA e resíduo dos níveis MAE-T e MAE-V.....	77
Figura 32 - Gráfico ajuste da RNA e de resíduo do nível $R_a^2$ .....	78
Figura 33 - Gráficos de Pareto do MAE-T e MAE-V in-natura Qt.....	80
Figura 34 - Gráficos de Pareto do $R^2$ e $R_a^2$ in-natura Qt.....	80
Figura 35 - Gráficos Fatoriais dos níveis MAE-T e MAE-V in-natura Qt.....	81
Figura 36 - Gráficos Fatoriais dos níveis $R^2$ e $R_a^2$ in-natura Qt.....	81
Figura 37 - Gráficos de Resíduo MAE-T e MAE-V in-natura Qt.....	82
Figura 38 - Gráficos de Resíduo $R^2$ e $R_a^2$ in-natura Qt.....	82
Figura 39 - Gráficos de Pareto do MAE-T e MAE-V ácido Qt.....	84
Figura 40 - Gráficos de Pareto do $R^2$ e $R_a^2$ ácido Qt.....	84
Figura 41 - Gráficos fatoriais dos níveis MAE-T e MAE-V ácido Qt.....	85
Figura 42 - Gráficos fatoriais dos níveis $R^2$ e $R_a^2$ ácido Qt.....	85
Figura 43 - Gráficos de Resíduos - MAE-T e MAE-V ácido Qt.....	86
Figura 44 - Gráficos de Resíduos - $R^2$ e $R_a^2$ ácido Qt.....	86
Figura 45 - Gráficos de Pareto do MAE-T e MAE-V alcalino Qt.....	89
Figura 46 - Gráficos de Pareto do $R^2$ e $R_a^2$ alcalino Qt.....	89
Figura 47 - Gráficos fatoriais de MAE-T e MAE-V alcalino Qt.....	90
Figura 48 - Gráficos fatoriais de $R^2$ e $R_a^2$ alcalino Qt.....	91
Figura 49 - Gráficos de Resíduos MAE-T e MAE-V alcalino Qt.....	91
Figura 50 - Gráficos de Resíduos $R^2$ e $R_a^2$ alcalino Qt.....	91
Figura 51- Gráfico ajuste da RNA e resíduo dos níveis MAE-T, MAE-V e $R_a^2$ .....	92
Figura 52 - Gráfico ajuste da RNA e resíduo dos níveis MAE-T e $R_a^2$ (Y1) .....	92
Figura 53 - Gráfico ajuste da RNA e resíduo dos níveis MAE-T e $R_a^2$ (Y2) .....	93
Figura 54 - Gráfico ajuste da RNA e resíduo do nível $R_a^2$ .....	93
Figura 55 - Gráfico ajuste da RNA e resíduo dos níveis $R_a^2$ e MAE-T.....	94
Figura 56 - Gráfico ajuste da RNA e resíduo do nível MAE-V.....	94
Figura 57 - Gráficos de Pareto do MAE-T e MAE-V.....	98
Figura 58 - Gráficos de Pareto do $R^2$ e $R_a^2$ ... ..	98
Figura 59 - Gráficos fatoriais dos níveis MAE-T e MAE-V.....	99
Figura 60 - Gráficos fatoriais dos níveis $R^2$ e $R_a^2$ .....	100
Figura 61 - Gráficos de Resíduo MAE-T e MAE-V.....	101
Figura 62 - Gráficos de Resíduo $R^2$ e $R_a^2$ .....	101

Figura 63- Gráficos de Pareto do MAE-T e MAE-V.....	105
Figura 64 - Gráficos de Pareto do $R^2$ e $R_a^2$ .....	105
Figura 65 - Gráficos fatoriais dos níveis MAE-T e MAE-V.....	106
Figura 66 - Gráficos fatoriais dos níveis $R^2$ e $R_a^2$ .....	107
Figura 67 - Gráficos de Resíduo MAE-T e MAE-V.....	108
Figura 68 - Gráficos de Resíduo $R^2$ e $R_a^2$ .....	108
Figura 69 - Gráficos de Pareto do MAE-T e MAE-V.....	112
Figura 70 - Gráficos de Pareto do $R^2$ e $R_a^2$ .....	112
Figura 71 - Gráficos fatoriais dos níveis MAE-T e MAE-V.....	113
Figura 72 - Gráficos fatoriais dos níveis $R^2$ e $R_a^2$ .....	114
Figura 73 - Gráficos de Resíduo MAE-T e MAE-V.....	115
Figura 74 - Gráficos de Resíduo $R^2$ e $R_a^2$ duas camadas alcalino.....	116
Figura 75 - Gráfico ajuste da RNA e resíduo dos níveis $R_a^2$ (Y1) e MAE-V.....	117
Figura 76 - Gráfico ajuste da RNA e resíduo do nível $R_a^2$ (Y2) .....	117
Figura 77 - Gráfico ajuste da RNA e resíduo do nível MAE-T.....	118
Figura 78 - Gráfico ajuste da RNA e resíduo dos níveis $R_a^2$ (Y1 e Y2) .....	119
Figura 79 - Gráfico ajuste da RNA e resíduo do nível MAE-T.....	120
Figura 80 - Gráfico ajuste da RNA e resíduo do nível MAE-V.....	121
Figura 81 - Gráfico ajuste da RNA e resíduo do nível $R_a^2$ (Y1) .....	121
Figura 82 - Gráfico ajuste da RNA e resíduo do nível $R_a^2$ (Y2) .....	122
Figura 83 - Gráfico ajuste da RNA e resíduo do nível MAE-T.....	123
Figura 84 - Gráfico ajuste da RNA e resíduo do nível MAE-V.....	124

## LISTA DE ABREVIações E SÍMBOLOS

$\Delta W_{ij}$  = Regra Delta ou Regra de Widrow-Hoff

**%REM** = eficiência na remoção

**AIC** = Critério de Informação de Akaike

**ap** = ajuste de peso.

**BIC** = Critério de Informação Bayesiano

$d_j$  = saída desejada

**DQO** = demanda química de oxigênio

**FA** = função de ativação

**HCl** = Ácido Clorídrico

**pH** = potencial hidrogeniônico

**lr** = taxa de aprendizagem

**MAE-T** = Erro quadrado médio para treino

**MAE-V** = Erro quadrado médio para validação

**MLP** = Perceptron Multicamadas

**NaOH** = Hidróxido de Sódio

**NN** = números de neurônios

**NE** = número de épocas

$n_1$  = primeira camada neural escondida

$n_2$  = segunda camada neural escondida

**RNA** = Redes Neurais Artificiais

**RLM** = Regressão Linear Múltipla

$Q_t$  (mg. g<sup>-1</sup>) = unidade de massa de adsorvente

$R^2$  = Erro Quadrado ou Coeficiente de Determinação

$R_a^2$  = Coeficiente de determinação ajustado

$t$  = tempo de correção do gradiente descendente

**tm** = termo *momentum*

$x_n$  = camada de entrada

$y_m$  = camada neural de saída

$y_a$  = equação polinomial para o tratamento ácido da variável  $Q_t$

$y_b$  = equação polinomial para o tratamento alcalino da variável  $Q_t$

$y_c$  = equação polinomial para o tratamento in-natura variável  $Q_t$

$y_d$  = equação polinomial para o tratamento ácido da variável Rem

$y_e$  = equação polinomial para o tratamento alcalino da variável rem

$y_f$  = equação polinomial para o tratamento in-natura da variável Rem

$W$  = peso;

$W_a$  = peso anterior

## LISTA DE TABELAS

Tabela 1 – Variáveis utilizadas no modelo.....	47
Tabela 2 - Planejamento simulado para uma camada.....	47
Tabela 3 - Planejamento simulado para duas camadas.....	48
Tabela 4 - Bagaço cana de açúcar com tratamento ácido.....	50
Tabela 5 - Bagaço cana de açúcar com tratamento alcalino.....	50
Tabela 6 - Bagaço cana de açúcar com tratamento in-natura.....	50
Tabela 7 – Valores do AIC e BIC da variável Qt.....	51
Tabela 8 – Valores do AIC e BIC da variável Rem.....	51
Tabela 9 – Teste de Durbin-Watson e Shapiro-Wilk.....	52
Tabela 10 – Treino do tratamento ácido Qt com 5 repetições.....	56
Tabela 11 – Teste do tratamento ácido Qt com 5 repetições.....	56
Tabela 12 – Treino do tratamento alcalino Qt com 5 repetições.....	56
Tabela 13 – Teste do tratamento alcalino Qt com 5 repetições.....	56
Tabela 14 – Treino do tratamento in-natura Qt com 5 repetições.....	56
Tabela 15 – Teste do tratamento in-natura Qt com 5 repetições.....	57
Tabela 16 – Treino do tratamento ácido Rem com 5 repetições.....	57
Tabela 17 – Teste do tratamento ácido Rem com 5 repetições.....	57
Tabela 18 – Treino do tratamento alcalino Rem com 5 repetições.....	57
Tabela 19 – Teste do tratamento alcalino Rem com 5 repetições.....	57
Tabela 20 – Treino do tratamento in-natura Rem com 5 repetições.....	58
Tabela 21 – Teste do tratamento in-natura Rem com 5 repetições.....	58
Tabela 22 - Planejamento do tratamento in-natura – Rem.....	62
Tabela 23 - Planejamento do tratamento ácido – Rem.....	67
Tabela 24 - Planejamento do tratamento alcalino – Rem.....	71
Tabela 25 - Planejamento do tratamento in-natura – Qt.....	79
Tabela 26 - Planejamento do tratamento ácido – Qt.....	83
Tabela 27 - Planejamento do tratamento alcalino – Qt.....	88
Tabela 28 - Planejamento do tratamento ácido em duas camadas.....	95
Tabela 29 - Planejamento do tratamento in-natura em duas camadas.....	102
Tabela 30 - Planejamento do tratamento alclalino em duas camadas.....	109

## SUMÁRIO

<b>1 INTRODUÇÃO</b> .....	15
<b>2 OBJETIVOS</b> .....	17
2.1 Objetivo Geral.....	17
2.1.1 Objetivos Específicos .....	17
<b>3 FUNDAMENTAÇÃO TEÓRICA</b> .....	18
3.1 Indústria Têxtil e a poluição ambiental .....	18
3.2 Tratamento de efluentes têxteis usando adsorção .....	19
3.3 Modelagem Matemática .....	19
3.3.1 Regressão Linear Múltipla .....	20
3.3.2 Redes Neurais Artificiais.....	23
3.3.3 Tipos de Redes Neurais Artificiais .....	30
3.3.4 Treinamento da RNA .....	32
3.3.5 Função de Transferência de Ativação .....	37
3.3.6 Análise da influência da função de ativação .....	38
3.3.7 Algoritmo de treinamento das RNA.....	39
3.3.8 Regra Delta.....	41
3.3.9 Estudo da Arte.....	42
<b>4 METODOLOGIA</b> .....	44
4.1 Localização.....	44
4.2 Dados Experimentais.....	44
4.3 Regressão Linear Múltipla .....	45
4.3.1 Aplicação da Regressão Linear Múltipla.....	45
4.4 Rede Neural Artificial .....	46
4.4.1 Determinação da Arquitetura da Rede Neural.....	46
4.4.2 Variáveis do Modelo .....	47
4.4.3 Planejamento simulado para uma camada.....	47
4.4.4 Planejamento simulado para uma camada.....	47
4.5 Utilização da Linguagem Python em Redes Neurais.....	48
<b>5 RESULTADOS E DISCUSSÕES</b> .....	50
5.1 Descrição Estatística do Banco de Dados experimentais.....	50
5.2 Modelagem usando a Regressão Linear Múltipla.....	51
5.3 Modelagem usando Redes Neurais Artificiais .....	62
5.3.1 Bagaço de cana de açúcar in-natura uma camada (Rem).....	62

5.4.1 Bagaço de cana de açúcar ácido uma camada (Rem).....	67
5.5.1 Bagaço de cana de açúcar alcalino uma camada (Rem).....	71
5.6.1 Otimização da RNA com uma camada (Rem).....	75
5.7.1 Bagaço de cana de açúcar in-natura uma camada (Qt).....	79
5.8.1 Bagaço de cana de açúcar ácido uma camada (Qt).....	83
5.9.1 Bagaço de cana de açúcar alcalino uma camada (Qt).....	88
5.10.1 Otimização da a RNA com uma camada (Qt).....	92
5.10.2 Tratamento do bagaço de cana de açúcar com duas camadas.....	95
5.10.3 Otimização para a RNA com duas camadas.....	116
<b>6 CONCLUSÕES .....</b>	<b>125</b>
<b>REFERÊNCIAS BIBLIOGRÁFICAS.....</b>	<b>126</b>

## 1 INTRODUÇÃO

Poluição ambiental pode ser definida como toda ação ou omissão do homem que, pela descarga de material ou energia atuando sobre as águas, o solo, o ar, causa um desequilíbrio nocivo, seja ele de curto, seja de longo prazo, sobre o meio ambiente. A definição do agente causador de poluição é dada como ser uma pessoa física ou jurídica de direito público ou privado, responsável direta ou indiretamente pela atividade causadora da degradação ambiental. (VALLE, 2004, p.49).

Em função dos processos industriais desenvolvidos na sociedade atual há um crescente aumento nos índices de poluição ambiental oriundos de práticas nocivas ao meio ambiente. Entre os vários setores industriais responsáveis pela poluição do meio ambiente, o têxtil merece um destaque especial, pois utiliza uma grande quantidade de água e lança milhares de efluentes contaminados com corantes (GHAZI MOKRI *et al.*, 2015; HYNES *et al.*, 2020).

No processo de produção têxtil são utilizadas substâncias químicas (corantes) para o tingimento dos tecidos, os quais, se lançados nos corpos hídricos, podem causar problemas ambientais. O descarte dos efluentes têxteis sem tratamento nos ambientes aquáticos podem levar rapidamente ao esgotamento do oxigênio dissolvido, tendo como consequência o desequilíbrio desse ecossistema. A presença de corantes nessas águas, impede a penetração da luz solar nas camadas mais profundas, alterando a atividade fotossintética do meio, resultando em deterioração da qualidade dessa água, diminuindo a solubilidade de oxigênio, e resultando em efeitos tóxicos sobre a fauna e flora aquática (LALNUNHLIMI; KRISHNASWAMY, 2016).

O processo de adsorção é um dos métodos mais populares e eficientes para o tratamento de efluentes, sendo utilizado principalmente na área industrial a fim de reduzir a toxicidade dos mesmos. É um método de transferência de massa em que são utilizados sólidos com capacidade de concentrar em sua superfície substâncias líquidas ou gasosas, possibilitando que ocorra a separação de fases de um fluido (MAXWELL, 2021 e NASCIMENTO, 2014).

Todos os fenômenos físicos deste universo são passíveis de modelagem matemática. Alguns deles são mais simples de modelar e outros, tão complexos que um modelo matemático aproximado já se torna um grande feito. A modelagem matemática, dotada de critérios adequados, proporciona a criação de um conjunto de

equações denominadas equações da dinâmica do movimento, ou seja, equações matemáticas que descrevem o comportamento dinâmico do fenômeno físico.

Os modelos matemáticos são determinados por meio do uso de regressão. A análise de regressão baseia-se na produção de uma análise estatística para averiguar a presença de uma conexão entre uma variável dependente com uma ou mais variáveis independentes, ou seja, significa conseguir uma equação matemática a qual explique a alternância da variável dependente, por meio da variação dos níveis das variáveis independentes (LOPES, CHAVARETTE, & COSSI, 2017).

Uma das maneiras de se fazer a modelagem matemática é por meio da análise de regressão que consiste na realização de uma análise estatística com o objetivo de verificar a existência de uma relação entre uma variável dependente com uma ou mais variáveis independentes. Esta relação é representada por meio do uso de um modelo matemático, que associa estas variáveis.

As RNA são baseadas na estrutura individual dos nodos, na topologia de suas conexões e no comportamento conjunto e dinâmico destes nodos naturais a estímulos externos. Estas podem ser operadas como um modelo de caixa preta, o qual não requer informações detalhadas sobre o sistema ou equipamento em questão. Esta ferramenta pode aprender a relação entre entradas e saídas com base nos dados de treinamento, sendo, portanto, amplamente utilizada para modelagem, predição e classificação, por possibilitar a identificação de relações presentes em fontes de dados limitados, incompletos, com ruídos, dinâmicos e não lineares. Assim, a habilidade da RNA de aprender e generalizar o comportamento de qualquer processo complexo e não linear a torna uma poderosa ferramenta de modelagem (AGHBASHLO *et al.*, 2012; MOHANRAJ; JAYARAJ; MURALEEDHARAN, 2012; YOUSSEFI; EMAMDJOMEH; MOUSAVI, 2009).

Este trabalho propõe a predição por meio da modelagem matemática baseada no conhecimento de Análise de Regressão e Redes Neurais Artificiais para o tratamento de dados de adsorção de corantes têxteis.

## **2 OBJETIVOS**

### **2.1 Objetivo Geral**

Estudar a aplicação das técnicas de modelagem matemática no estudo do processo de adsorção de corantes têxteis, usando como adsorvente o bagaço de cana de açúcar.

#### **2.1.1 Objetivos Específicos**

1. Analisar a técnica da Regressão Linear Múltipla para prever o ajuste de um modelo para os dados de adsorção.
2. Investigar as Redes Neurais Artificiais para predição do comportamento dos dados verificando a eficiência do processo de adsorção.
3. Comparar o efeito da modelagem matemática usando RLM e RNA.

### 3 FUNDAMENTAÇÃO TEÓRICA

Este presente trabalho apresenta uma divisão em tópicos, onde são abordados na seguinte sequência os temas: a indústria têxtil, a poluição ambiental, tratamento de efluentes têxteis usando adsorção, Modelagem Matemática, Regressão Linear Múltipla, Redes Neurais Artificiais, Tipos de Redes Neurais Artificiais, Treinamento das RNA, Função de Transferência/Ativação, Análise da influência da função de ativação, Algoritmo de treinamento das RNA, Regra Delta e Estudo da Arte.

#### 3.1 Indústria têxtil e a poluição ambiental

A indústria têxtil está inserida na problemática ambiental, uma vez que os seus efluentes, além de serem em grandes proporções, apresentam a agravante de ser colorido, podendo ser percebido em um corpo hídrico mesmo em pequenas quantidades como  $1 \text{ mg. L}^{-1}$  (BERTAZZOLLI; PELEGRINI, 2002).

O desenvolvimento industrial está diretamente relacionado com a poluição ambiental. A indústria têxtil possui um dos processos de maior geração de poluentes, contribuindo quantitativa e qualitativamente com carga poluidora descartada no meio ambiente, os quais, quando não corretamente tratados, são indutores de sérios problemas de contaminação ambiental.

Em se tratando da conservação ambiental, o setor industrial precisa alterar sua forma de produção buscando alternativas menos impactantes ao ambiente.

O Brasil possui um forte setor têxtil, representando uma parte importante da economia do país, porém existe uma ausência da responsabilidade ambiental dos descartes dos efluentes gerados por esse setor. Calcula-se que quase toda a carga orgânica de corantes que entram no meio ambiente é através dos efluentes industriais têxteis (REVANKAR; LELE, 2007).

A despeito do alto desenvolvimento desse setor, é muito pouco o investimento feito para tratar os resíduos gerados pela indústria têxtil. De acordo com Abreu *et al.*, (2008) as empresas investem apenas o básico no tratamento de seus resíduos, pois seu único objetivo é evitar multas, sem pensar na preocupação social e ambiental. Outra problemática é a falta de fiscalização e monitoramento dos resíduos descartados por essas indústrias, podendo ocorrer farsas ou alterações nos supostos tratamentos dos efluentes.

### **3.2 Tratamento de efluentes têxteis usando adsorção.**

Entre os vários segmentos de indústrias que necessitam de um tratamento posterior de água, está a indústria têxtil, que gera efluentes com uma forte coloração devido à utilização de corantes nos processos de tingimento, além de grande quantidade de sólidos suspensos, elevadas concentrações de DQO, significativa quantidade de metais pesados, compostos orgânicos clorados e surfactantes (ARAUJO e YOKOYAMA, 2006).

Os processos convencionais de tratamento de efluentes, em geral, não são capazes de degradar/remover as estruturas moleculares complexas dos corantes. Quando não corretamente tratadas, essas águas residuárias podem causar sérios problemas de contaminação ambiental e potencializar os riscos à saúde das pessoas, os quais estão associados à entrada de componentes tóxicos nas cadeias alimentares de animais e seres humanos, podendo ser carcinogênicos e/ou mutagênicos (KUMAR *et al.*, 2017). Nos últimos anos, têm-se buscado novas tecnologias para a degradação ou remoção dos corantes com o intuito de adequar o efluente tratado perante a legislação para lançamento nos corpos receptores.

Entre as tecnologias de tratamento, a adsorção provou ser um dos métodos físico-químicos mais eficazes para a remoção de poluentes. Recentemente, alguns pesquisadores têm se concentrado na utilização de adsorventes produzidos a partir de materiais de baixo custo e que sejam abundantes, a fim de substituir o carvão ativado convencional (KANKILIÇ *et al.*, 2016; AHMED, 2017).

### **3.3 Modelagem Matemática**

No entender de Bassanezi (2015), a Modelagem Matemática consiste na arte de transformar problemas da realidade em problemas matemáticos e resolvê-los interpretando suas soluções na linguagem do mundo real.

Para Bassanezi (2015) a Modelagem Matemática é uma metodologia utilizada para obtermos alguma explicação ou entendimento de determinadas situações reais.

Nesse contexto, a modelagem matemática surge como uma ferramenta auxiliar na aplicabilidade da Análise Regressão e RNA, no estudo do processo de adsorção de corantes têxteis.

### 3.3.1 Regressão Linear Múltipla

A Regressão Linear Múltipla é a generalização da Regressão Simples para um conjunto de preditores. Como no caso simples, também apresenta uma variável dependente ( $Y$ ), mas utiliza duas ou mais variáveis preditoras. O modelo de Regressão Linear Múltipla pode ser descrito como na Equação 1:

$$Y_i = \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + \dots + \beta_{p-1} X_{i,p-1} + e_i \quad i = 0,1,2,3, \dots, n \quad (1)$$

em que:

$Y_i$  é o valor da variável resposta na  $i$ -ésima observação;

$\beta_0, \beta_1, \dots, \beta_{p-1}$  são coeficientes da regressão;

$X_{i1}, X_{i2}, \dots, X_{i,p-1}$  são constantes conhecidas;

$e_i$  são independentes com  $N(0, \sigma^2)$ .

Cada uma das variáveis preditoras tem seu próprio coeficiente, de forma análoga à inclinação da regressão simples, estes são chamados coeficientes da regressão. Em particular, o coeficiente  $\beta_0$  é denominado constante da regressão. Os parâmetros estatísticos deste modelo podem ser estimados utilizando o Método de Mínimos Quadrados, desde que o número de observações seja maior que o número de variáveis regressoras mais um (CHARNET *et al.*, 2008).

Há necessidade de se observar, no uso de modelos de regressão, os seguintes pressupostos básicos: especificação, normalidade, homocedasticidade, não multicolineariedade, não-autocorrelação, independência e inexistência de pontos atípicos. O atendimento a tais pressupostos objetiva garantir que as avaliações sejam não tendenciosas, eficientes e consistentes (ABNT, 2011).

Levine (2014) define os pressupostos:

- Linearidade - relação entre variáveis linear.
- Independência - requer que os erros ( $\epsilon_j$ ) sejam independentes entre si (importância maior, principalmente quando a coleta de dados se dá ao longo de determinado período de tempo).
- Normalidade - requer que os erros sejam distribuídos de forma normal para cada valor de  $X$ . Os resíduos padronizados podem ser verificados pela proporção de

frequência relativa dos resíduos amostrais padronizados, na proporção de 68% para intervalos  $[-1; 1]$ , 90% para intervalos  $[-1,64; 1,64]$ , 95% para intervalos  $[-1,96; 1,96]$  (ABNT, 2011). Tal atributo pode ser verificado através de gráfico de barras com distribuição das frequências (situação melhor verificada para números amostrais suficientes).

- Igualdade de variância (homocedasticidade) - requer que a variância dos erros seja constante em relação a todos os valores de  $X$  (situação de aleatoriedade)

Dois critérios de seleção de modelos de regressão, amplamente utilizados na literatura, são: critério de informação de Akaike ( $AIC$ ) e o critério de informação Bayesiano ( $BIC$ ). Proposto por AKAIKE (1974), o critério  $AIC$  é definido por

$$AIC = -2 \log[L(\boldsymbol{\beta})] + 2p,$$

onde  $\boldsymbol{\beta}$  é o estimador de máxima verossimilhança de  $\boldsymbol{\beta}$ ,  $L(\boldsymbol{\beta})$  é a função de verossimilhança e  $p$  o número de parâmetros desconhecidos do modelo. Para detalhes sobre função de verossimilhança e estimação por máxima verossimilhança veja BOLFARINE e SANDOVAL (2010). Segundo esse critério de seleção, o modelo que apresentar menor valor  $AIC$  é o modelo melhor ajustado. O  $BIC$  foi proposto por CHWARTZ (1978), assim como o  $AIC$  o modelo que apresenta menor  $BIC$  é o melhor ajustado:

$$BIC = -2 \log[L(\boldsymbol{\beta})] + 2p \log n,$$

onde  $n$  é o tamanho amostral.

Outro aspecto importante no ajuste de modelos de regressão linear múltipla é a multicolinearidade. Tem-se como objetivo investigar se há multicolinearidade entre as variáveis regressoras, visto que a forte correlação entre elas pode resultar em efeitos negativos no ajuste do modelo. A multicolinearidade é um problema comum em regressão linear, indicando que existe uma relação de linearidade entre as variáveis independentes, prejudicando a estimação dos coeficientes da regressão (VILLA, 2016). Uma das formas de identificar a presença de multicolinearidade é avaliar o Fator de Inflação da Variância (VIF - Variance Inflation Factor). Esse fator

mede a associação entre as variáveis regressoras de acordo com o coeficiente de determinação do modelo de regressão, apenas com as variáveis independentes. De acordo com (BERK, 1977) o Fator de Inflação da Variância é definido de acordo com a Equação 2 como:

$$VIF_i = \frac{1}{1-R_i^2} \quad (2)$$

em que  $R_i^2$  é o coeficiente de determinação da regressão da variável explicativa  $X_i$  sobre as outras variáveis explicativas com  $i = 1, 2, \dots, k$ , sendo  $k$  quantidade de variáveis explicativas no modelo (VILLA, 2016). Geralmente, considera-se  $VIF > 10$  como um indicativo de problema de multicolinearidade.

De acordo com Sassi *et al.*, (2011), esta metodologia descreve a relação matemática existente entre a variável resposta e as variáveis explicativas e pode reduzir o número de variáveis com o mínimo de perda de informação, permitindo detectar padrões de similaridade, associação e correlação entre as variáveis do estudo.

Segundo (CHARNET, 1999), os resíduos de um modelo de regressão linear têm uma relação muito forte com a qualidade do ajuste, bem como com a confiabilidade dos testes estatísticos sobre os parâmetros do modelo. Nesse sentido, a análise de resíduos tem uma importância fundamental na verificação da qualidade dos ajustes de modelos. Basicamente, essa análise fornece evidências sobre possíveis violações nas suposições do modelo, tais como a de normalidade, homocedasticidade, e quando for o caso ainda fornece indícios de falta de ajuste do modelo proposto.

Para a normalidade é verificado se os resíduos possuem distribuição normal. Caso não seja atendida essa característica, haverá a possibilidade de maior erro padrão e de estimativas ineficientes (ROCHA; JÚNIOR, 2018). O Teste de Shapiro-Wilk (MONTGOMERY, 2013) tem sido o mais utilizado para atestar a normalidade de resíduos e é conformado por duas hipóteses (nula e alternativa):

- $H_0$  - as amostras vêm de uma população que segue uma distribuição normal.
- $H_1$  - as amostras não vêm de uma população que segue uma distribuição normal (p-value suficientemente baixo).

A homocedasticidade pressupõe que os resíduos tenham variâncias estatisticamente iguais entre os níveis experimentais, sendo que uma prova muito comum aplicada é o Teste de Bartlett (MONTGOMERY, 2013), com as seguintes hipóteses (ROCHA; JÚNIOR, 2018):

- $H_0$  - as variâncias entre os níveis são estatisticamente iguais, homocedásticas.
- $H_1$  - as variâncias entre os níveis são estatisticamente diferentes, heterocedásticas (pvalue suficientemente baixo).

A independência pressupõe que a correlação entre os resíduos seja zero, ou seja, uma observação tem efeito nulo sobre a outra. Quando os resíduos são dependentes uns dos outros, significa que há um vício na coleta dos dados brutos e as observações dependem umas das outras (ROCHA; JÚNIOR, 2018). O Teste de Durbin-Watson (MONTGOMERY, 2013) pode ser utilizado para detectar a presença de autocorrelação (dependência) nos resíduos.

- $H_0$  - os erros são independentes.
- $H_1$  - os erros são dependentes ou autocorrelacionados (*p-value* suficientemente baixo).

A regressão linear múltipla é uma técnica estatística utilizada quando se deseja analisar o comportamento de uma variável resposta em relação a outras variáveis que são responsáveis pela variabilidade das observações. Destarte, assume-se que há uma relação linear entre a variável resposta e as  $k$  variáveis explicativas (CORRAR; PAULO; DIAS FILHO, 2007).

De acordo com Hair Júnior *et al.*, (2009), a regressão múltipla é um modelo realista porque, no mundo em que vivemos, a previsão quase sempre depende de vários fatores. Mesmo quando estamos interessados no efeito de apenas uma das variáveis, é aconselhável incluir as outras capazes de afetar a variável dependente, efetuando uma análise de regressão múltipla.

### **3.3.2 Redes Neurais Artificiais**

As primeiras informações sobre estudos na área da “computação neural”

surgiram em 1943, em artigos criados pelo neuroanatomista e psiquiatra Warren McCulloch, do Instituto de Tecnologia de Massachusetts, e do matemático Walter Pitts, da Universidade de Illinois. Sendo que alguns anos depois, em 1949, Donald Olding Hebb publicava importante obra, o livro “The Organization of Behaviour” (A organização do Comportamento) que influenciou vários modelos de RNAs de destaque na atualidade (FERNANDES, 2005).

Redes Neurais Artificiais (GOODFELLOW; BENGIO; COURVILLE, 2016), (MELLO; PONTI, 2018) são um tipo de algoritmo de aprendizado de máquina inspirado no funcionamento do cérebro humano, tendo a capacidade de aprender uma transformação desejada de forma automática. Elas são genéricas o suficiente para que o usuário não precise fornecer regras específicas da tarefa, somente um conjunto de dados relevantes de entrada com suas respectivas saídas desejadas. Essa classe de algoritmos está sendo exaustivamente pesquisada e utilizada em produtos comerciais atualmente devido a sua capacidade de resolver problemas muito complexos, como segmentação semântica de imagens (LIU; DENG; YANG, 2019).

Elas aprendem e processam informação de uma forma similar às redes neurais biológicas. Uma Rede Neural Artificial é capaz de classificar e gerar previsões sobre um conjunto de dados de entrada, ao adaptar o processamento da entrada para que seja gerado o melhor resultado possível.

Ao passo que as tecnologias computacionais foram avançando, os pesquisadores foram buscando outras maneiras de aplicar e desenvolver os conhecimentos de RNA. Nesse contexto, outros tipos de redes neurais artificiais foram surgindo, assim também como sua forma de aplicação aliada a outras ferramentas (AGGARWAL, 2018).

De acordo com (HAYKIN, 2017), uma Rede Neural Artificial é um processador paralelamente distribuído e constituído de séries de unidades de processamento simples, que tem a propensão natural para armazenar conhecimento experimental e tornar disponível para uso, se assemelha ao cérebro, pois seu conhecimento é adquirido pela rede a partir de um ambiente, no caso o conjunto de dados exposto, caracterizando-se um processo de aprendizagem, e pelas forças de conexão entre neurônios, conhecidas como pesos sinápticos, utilizadas para armazenar o conhecimento adquirido.

A natureza de uma Rede Neural Artificial (RNA) é fundamentada em dois aspectos, seu algoritmo e sua arquitetura, a partir do momento em que a rede é criada,

esta é treinada a partir de exemplos, e o algoritmo de aprendizagem então desenvolve estes dados e armazena os parâmetros ajustáveis da rede, os chamados pesos.

Uma RNA é composta por uma série de neurônios organizados em diversas camadas, seu processamento costuma ser linear, a partir da combinação das entradas com os pesos e, em seguida pela passagem por uma função de ativação.

A modelagem e a simulação são utilizadas para entender e obter dados de um processo sem que haja a necessidade de realizar a todo o momento experimentos práticos. Uma das ferramentas mais poderosas da atualidade para prever o processo de adsorção são as Redes Neurais Artificiais - RNA (GHAEDI; VAFAEI, 2017; TANZIFI *et al.*, 2017).

Tanzifi *et al.*, (2017) utilizaram RNA para modelar e prever o percentual de remoção do corante methyl orange com relação a cinco parâmetros: pH, concentração inicial de corante na solução, quantidade de adsorvente, tempo de adsorção e temperatura. Comparando-se os dados obtidos em meio aos experimentos com os preditos pela RNA pode-se concluir que o modelo foi capaz de estimar o percentual de remoção do corante sob distintas condições de forma adequada.

O funcionamento das RNA baseia-se na estrutura neuronal biológica de organismos inteligentes, as quais adquirem conhecimento por meio de processos de aprendizado (FAGUNDES-KLEN *et al.*, 2011; GHAEDI; VAFAEI, 2017).

De modo genérico, a rede neural artificial pode ser entendida como um sistema computacional que busca modelar o funcionamento do cérebro humano para a realização de atividades específicas.

Uma Rede Neural Artificial (RNA) é um modelo matemático que se baseia no comportamento de um neurônio do sistema nervoso biológico, composto por unidades denominadas neurônios artificiais, que interagem entre si, armazenando conhecimento e gerando respostas para uso (CAMPOS *et al.*, 2016).

Sendo comparadas ao cérebro humano, as Redes Neurais Artificiais (RNAs) são consideradas uma alternativa atraente, pois a sua alta precisão na predição de dados faz com que passe a frente dos métodos convencionais (CHITEKA; ENWEREMADU, 2016). As RNA foram aplicadas em diferentes campos de estudos obtendo sucesso, como: engenharia, medicina, matemática, psicologia, economia, dentre outras áreas de grande importância para sociedade, fazendo com que as estruturas organizacionais sejam flexíveis com melhores eficiências e desempenhos,

podendo fornecer uma nova concepção de produtos e serviços complexos (STAUB *et al.*, 2015).

Uma Rede Neural Artificial é um processador distribuído paralelo, composto de processamento de unidades com habilidade para armazenar conhecimento experimental e disponibilizá-lo para uso (HEIDARI *et al.*, 2020). Assemelha-se ao cérebro humano em dois aspectos: (i) o conhecimento é adquirido pela rede a partir de seu ambiente por meio de um processo de aprendizado; (ii) os pontos fortes da conexão interneurônio, conhecidos como pesos sinápticos, são usados para armazenar o conhecimento necessário.

As RNA têm o objetivo de aprender sobre a dinâmica de um determinado processo. Para se obter sucesso da modelagem de um processamento via rede neural, as variáveis em estudo devem estar devidamente conhecidas e definidas, resultando assim em uma base dados de qualidade que integra todas informações necessárias (ATA, 2015).

Dentre os diferentes tipos de redes neurais artificiais existentes, com aplicações comprovadamente eficientes na engenharia, pode-se citar o perceptron multicamadas (MLP) (RIPLEY, 1996; BISHOP, 2006; HAYKIN, 2008; DANTAS NETO *et al.*, 2014, 2016, 2017; LEITE *et al.* 2018). Esse sistema consiste em uma rede neural com uma ou mais camadas ocultas paralelas, que tem como intuito aumentar a eficiência das ligações como ocorre no sistema nervoso humano.

As RNA possuem algumas características que as diferenciam dos demais tipos de ferramentas computacionais, sendo que estas estão intimamente relacionadas com a sua aplicação. Para tanto, Da Silva *et al.*, (2016) elencaram algumas dessas características que podem facilitar quando da necessidade de sua escolha, quais sejam: Adaptação por experiência, Capacidade de aprendizagem, Habilidade de generalização, Organização de dados, Tolerância a falhas, Armazenamento distribuído e Facilidade de prototipagem.

Dentre as características anteriormente citadas, destacam-se: a adaptação a partir de padrões assimilados pelas RNA, os quais estão diretamente relacionados aos valores dos pesos sinápticos; a capacidade de aprendizado mediante o treinamento prévio, de acordo com os valores das variáveis apresentadas; a capacidade de generalização, conforme seu treinamento, uma vez que a rede é capaz de estimar valores a partir de dados de entrada desconhecidos; e a possibilidade de transformação do processamento da RNA em um hardware, permitindo um número

maior de aplicação (DA SILVA *et al.*, 2016).

Um sistema baseado em RNA é representado pela formação de uma rede composta por neurônios, sendo que cada um desses é formado por nós, sinapses ou vieses e função de ativação; em cada nó há um peso específico, responsável pelo efeito que terá sobre a etapa seguinte (ARAM *et al.*, 2017). Em sua aplicação, são utilizadas diversas formas de arquitetura, sendo que sua escolha está condicionada à aplicabilidade da rede, tendo em vista seu desempenho mediante à tarefa a ser realizada e o processamento dos dados de entrada.

Martins *et al.*, (2016) afirmam que uma das principais vantagens das redes neurais artificiais sobre outros métodos, como o modelo clássico de regressão, é a sua facilidade em adicionar variáveis qualitativas de forma eficiente, além de detectar relações não lineares entre a variável de saída e as variáveis de entrada durante todo o seu processo de aprendizagem.

Os neurônios são organizados em uma estrutura de redes formada por camadas contendo um número definido de neurônios. Na modelagem, é utilizado o processo de aprendizagem por treinamento (MULERO *et al.*, 2017). A RNA consiste em extrair informações de um banco de dados de variáveis contínuas e categóricas para identificar padrões entre elementos de entrada e um elemento de interesse (COSENZA *et al.*, 2017).

A eficiência, adaptabilidade e principalmente a capacidade de resolver problemas torna as RNA uma ferramenta extremamente importante em diversas análises (TKÁČ & VERNER, 2016).

As RNA são comumente utilizadas na resolução de questões complexas, onde os modos das variáveis não são rigorosamente conhecidos. Uma de suas principais características é a capacidade de aprender através de exemplos e de generalizar a informação aprendida, gerando um modelo não linear e, tornando sua aplicação na análise espacial bastante eficiente (SPÖRL *et al.*, 2011 apud FLECK, *et al.*, 2016).

Há ainda outros elementos que podem ser configurados ou predeterminados na arquitetura de uma rede, como: a quantidade de camadas de neurônios e a quantidade de neurônios de cada camada; a função de transferência e treinamento e o método de aprendizagem e medição de desempenho. Definidas estas configurações, uma RNA está preparada para ser treinada com base em uma amostra, composta por entrada e saída (HAVKIN, 2001).

Como no cérebro humano as Redes Neurais Artificiais apresentam uma

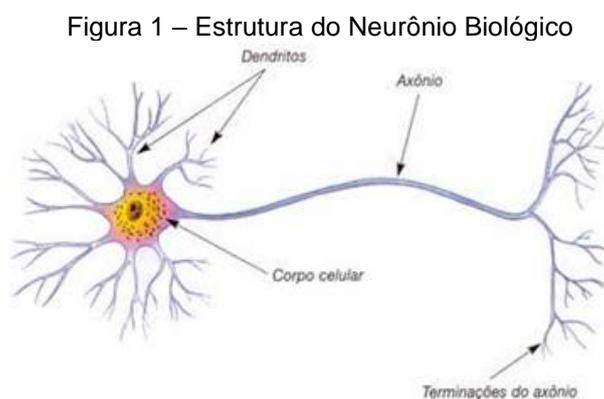
organização em lâminas, chamadas também de camadas, sendo que estas camadas apresentam contatos entre si e estão subdivididas em camadas de entrada, camadas intermediárias ou escondidas e camadas de saída. As redes de múltiplas camadas possuem uma ou mais destas camadas intermediárias.

A conexão destas camadas se faz pelas sinapses neuronais sendo que a plasticidade sináptica, que é a capacidade das sinapses sofrerem modificações e com isso modularem o processamento das informações constitui a base do processo cognitivo tanto das redes neuronais biológicas como das RNA.

Um sistema de associar pesos com as informações armazenadas pelo neurônio e a adaptação desses pesos aos estímulos do ambiente é como se dá esse processo que permite a RNA trabalhar com as habilidades semelhantes à percepção, raciocínio e memória das redes biológicas (SILVA, 2019).

A capacidade de aprender das RNA na realidade trata-se da habilidade de se adaptar alterando seu desempenho ao longo do tempo. O tipo de aprendizado é determinado pela técnica empregada no processo de ajustes dos pesos sinápticos. Um conjunto bem determinado destas regras constitui um algoritmo de aprendizagem ou treinamento, sendo que a maneira como o ambiente influencia a rede ela terá um aprendizado supervisionado, não supervisionado ou por reforço (SILVA, 2019).

Para entender uma Rede Neural Artificial (RNA), é necessário fazer um paralelo entre um neurônio biológico e artificial. O neurônio biológico é formado pela soma que é o corpo do neurônio, pelo axônio que conduz o impulso nervoso de um neurônio para outro, por dendritos que são ramificações que fazem a comunicação entre um neurônio e as sinapses, que por sua vez propagam o sinal neural ativando a comunicação com o próximo neurônio como é possível observar na Figura 1.

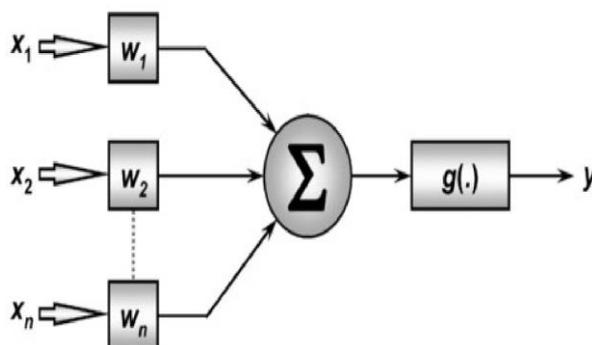


Fonte: Italo José, 2022.

O neurônio é a unidade fundamental da rede neural. Este elemento de processamento interage com outros por meio de conexões, que funcionam como caminhos de condução de sinais de único sentido.

O neurônio artificial é uma estrutura lógico-matemática que simula de forma simplificada o comportamento e as funções de um neurônio biológico (GUYON,1991) como pode ser observado na Figura 2.

Figura 2 - Modelo de um Neurônio Artificial



Fonte: Palmiere, 2016.

O neurônio artificial é o componente de construção das Redes Neurais Artificiais projetados para simular a função de um neurônio biológico. Os sinais que chegam, chamados entradas, multiplicadas pelos pesos de conexão são primeiro somados e depois passam por uma função de transferência para produzir a saída para esse neurônio. A função de ativação é a soma ponderada das entradas do neurônio e a função de transferência mais comumente usada é a função sigmóide. No neurônio artificial os dendritos são representados pelos valores de entrada, os pesos são as sinapses, a função de soma faz o papel do corpo do neurônio e a função de transferência representa o axônio.

As Redes Neurais Artificiais (RNA), de acordo com Zahedi e Rounaghi (2015), são modelos computacionais, inspirados no sistema nervoso central, utilizados para estimar ou aproximar funções que podem depender de um grande número de entradas, geralmente desconhecidas, sendo comumente apresentadas como sistemas de neurônios interconectados, que podem calcular valores de entradas, de uma maneira que os neurônios de entrada alimentam os valores para cada um dos neurônios na camada oculta, e a camada oculta fornece-os para a camada de saída.

As Redes Neurais são compostas por uma determinada quantidade de entradas e unidades de processamento, as quais são ligadas através de pesos sinápticos. As entradas são propagadas através da topologia da RNA, sendo transformadas pelos pesos sinápticos e pela função de ativação (AF) dos neurônios (MACHADO E FONSECA JÚNIOR, 2013).

Em analogia as Redes Neurais Biológicas, uma Rede Neural Artificial é um agregado de unidades computacionais conhecidas por neurônios artificiais. Cada neurônio possui terminais de entrada similares aos dendritos dos neurônios biológicos, por onde as informações são inseridas, posteriormente computadas, resultando um valor de saída que será propagado para outras unidades até a última camada provendo a resposta da rede (KO *et al.*, 2010; KOCAMAZ *et al.*, 2016; GOMES *et al.*, 2016; SILVA *et al.*, 2016, WANG *et al.*, 2016).

A partir da estrutura e funcionamento do neurônio biológico, pesquisadores tentaram simular este sistema em computador. O modelo mais bem aceito foi proposto por McCulloch e Pitts, conhecido como Perceptron, o qual implementa de maneira simplificada os componentes e o funcionamento de um neurônio (BRAGA *et al.*, 2000).

A modelagem e a simulação são utilizadas para entender e obter dados de um processo sem que haja a necessidade de realizar a todo o momento experimentos práticos. A modelagem matemática e a simulação de processos reduzem a necessidade de experimentação dispendiosa e demorada em escala laboratorial (BORBA *et al.*, 2011; NEVES *et al.*, 2017).

Os modelos empíricos geralmente são obtidos a partir de simples correlações matemáticas dos dados experimentais (SINGH & KULSHRESTHA, 1987) e os seus parâmetros, normalmente, não possuem significado físico.

Uma das ferramentas mais poderosas da atualidade para prever o processo de adsorção são as Redes Neurais Artificiais - RNA (GHAEDI; VAFAEI, 2017; TANZIFI *et al.*, 2017).

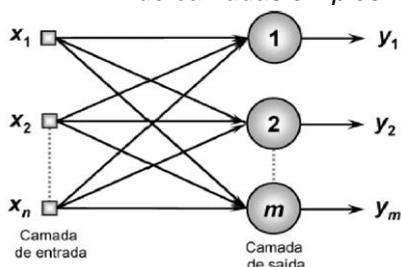
### **3.3.3 Tipos de Redes Neurais Artificiais**

A arquitetura de uma RNA é a forma como os neurônios estão dispostos na rede. Essa disposição dos neurônios também acontece no cérebro humano, fazendo como que os dados possam ser processados adequadamente. Existem diversas maneiras de organizar os neurônios de uma rede, e por essa razão há uma necessidade de se classificar as redes. A classificação de redes neurais é feita

conforme seu número de camadas, podendo ser de camada única ou de múltiplas camadas.

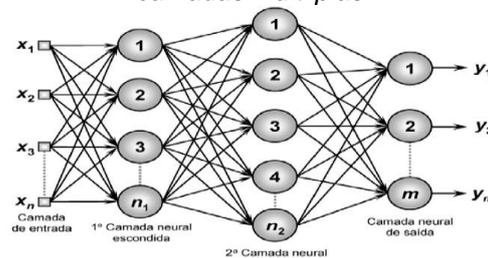
Dentro das redes múltiplas há outra classificação conforme sua alimentação, se as redes são cíclicas (*feedback*) ou acíclicas (*feedforward*) (PEREIRA, 2016) como pode-se observar nas Figuras 3(a) Arquitetura FeedForward de camadas simples, 3(b) Arquitetura Feed Forward de camadas múltiplas e 3(c) Arquitetura Realimentada.

Figura 3(a) - Arquitetura FeedForward de camadas simples



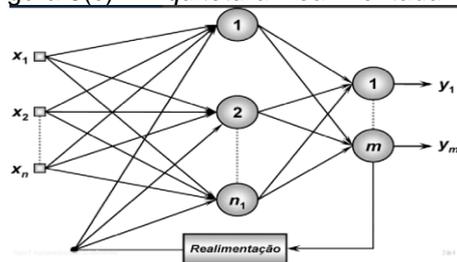
Fonte: Silva *et al.*, (2010).

Figura 3(b) - Arquitetura FeedForward de camadas múltiplas



Fonte: Silva *et al.*, (2010).

Figura 3(c) - Arquitetura Realimentada



Fonte: Silva *et al.*, (2010).

As Redes Neurais Recorrentes são uma técnica de aprendizado de máquina que apresentam neurônios recorrentes. Como estas incluem *loops*, elas podem armazenar informações ao processar novas entradas.

As Redes Recorrentes se distinguem das redes alimentadas por terem pelo menos um laço de realimentação, com a presença ou não de neurônios ocultos. A presença de laços de realimentação tem um impacto profundo na capacidade de aprendizagem da rede e no seu desempenho.

As Redes Multilayer Perceptron podem ser consideradas as Redes Neurais Artificiais mais populares, conhecida também como redes feedforward, elas respondem problemas não lineares.

As RNA recorrentes têm laços de realimentação local e/ou global em sua estrutura. Redes *feedforward* (i.e. sem laços de realimentação) podem ser adaptadas para processar séries temporais por meio de uma entrada com linha de atrasos com derivações (*tapped delay line*) e treinada pelo algoritmo backpropagation, elas podem também ser facilmente convertidas em arquiteturas recorrentes simples por realimentações de ativações dos neurônios das camadas escondidas ou de saída, dando origem as redes de Elman e de Jordan, respectivamente.

As Arquiteturas de uma rede neural são de diversas formas, conforme a necessidade da aplicação, estando diretamente ligadas com o algoritmo de aprendizagem usado para treinar a rede. Basicamente, os itens que compõem a estrutura de uma rede neural são, portanto, sujeito a modificações, sendo os seguintes: Camadas intermediárias, Quantidade de neurônios, Função de transferência e Algoritmo de aprendizado.

Uma das maiores dificuldades encontradas no uso das redes neurais é a escolha da melhor arquitetura, uma vez que esse processo é experimental e demanda um grande tempo de execução. Na prática, o processo deve ser aplicado com o intuito de testar vários métodos de aprendizado e as diferentes configurações que uma rede possa ter para a resolução de um problema em específico (MIRANDA *et al.*, 2009). Em geral, identificam-se três classes de arquiteturas de rede fundamentalmente diferentes: redes alimentadas adiante com camada única, redes alimentadas diretamente com múltiplas camadas e redes recorrentes.

### **3.3.4 Treinamento das RNA**

Uma rede neural artificial (ou simplesmente uma rede neural ou rede) é um modelo computacional biologicamente inspirado, constituído por elementos de processamento simples (neurônios artificiais) que aplicam uma determinada função matemática aos dados (função de ativação) gerando uma única resposta, são dispostos em camadas e ligados entre si, sendo estas conexões, geralmente, associadas a coeficientes denominados de pesos. O ajuste destes pesos é realizado por um processo chamado treinamento ou aprendizado, sendo responsável pela extração das características dos dados e armazenamento de conhecimento das redes.

As Redes Neurais Artificiais podem ser treinadas para aprenderem sobre determinado padrão, sendo capazes de generalizar, ou seja, estimar dados

incompletos ou desconhecidos, o que difere das técnicas tradicionais. É uma solução matemática capaz de lidar com problemas não-lineares e complexos (KO *et al.*, 2010; KOCAMAZ *et al.*, 2016; GOMES *et al.*, 2016; SILVA *et al.*, 2016, WANG *et al.*, 2016).

O treinamento começa no momento em que os dados chegam na entrada, a rede passa a alterar seus parâmetros, de polarização e sinapses, e atua de uma nova forma. O treinamento pode encerrar em três hipóteses: pelo número de épocas, que são o número de iterações do treinamento, por um valor de erro mínimo aceitável ou no caso do Toolbox de Redes Neurais do MATLAB, quando há um incremento no erro da entropia cruzada durante a validação (MATHWORKS, 2017).

A propriedade mais importante de uma rede neural é a sua capacidade de aprendizado a partir de um ambiente. Uma rede neural aprende acerca do seu ambiente através de um processo iterativo de ajustes aplicados a seus parâmetros livres (pesos sinápticos). Um conjunto pré-estabelecido de regras bem definidas para a solução de um problema de aprendizagem é denominado algoritmo de treinamento. Não existe um algoritmo de aprendizagem único para o projeto de redes neurais. Em vez disso, tem-se um “conjunto de ferramentas” representado por uma variedade de algoritmos de treinamento, cada qual oferecendo vantagens específicas.

O treinamento de uma RNA ocorre por meio de uma amostra de dados ocorridos no passado ( $X_1...X_n$ ) ou de padrões relacionados a determinado objeto ou situação. Os dados inseridos no sistema alimentam os terminais de entrada do neurônio artificial. Cada dado é multiplicado por um peso específico ( $W_1...W_n$ ), o que irá resultar em uma soma ponderada. Esse cálculo subtraído do limiar de ativação do neurônio resulta no potencial de ativação a ser computado pela função apropriada que por sua vez irá replicar o resultado para os demais elementos da rede (KO *et al.*, 2010; KOCAMAZ *et al.*, 2016; GOMES *et al.*, 2016; SILVA *et al.*, 2016, WANG *et al.*, 2016).

Para que uma RNA possa ser utilizada, faz-se necessário primeiramente que a mesma seja treinada. O processo de treinamento ou aprendizado nada mais é do que apresentar à rede uma sequência de conjuntos de dados relativos ao processo ao qual a rede será aplicada, sendo, então, extraída uma relação matemática entre os valores de entrada e os de saída do conjunto de treinamento e armazenada na forma de pesos sinápticos ao longo da estrutura interna da rede neural (SILVA, 2013), durante esta fase de treinamento, ocorre o seu aprendizado.

A aprendizagem é uma das características notáveis das RNAs. No aprendizado supervisionado, o treinamento possui um conhecimento sobre os dados, uma espécie

de instrutor que confere o quanto a rede está próxima de uma solução aceitável, este conhecimento, está representado sob forma de um conjunto de amostras de entrada-saída. O processo de treinamento modifica os pesos das RNAs com a finalidade de melhorar um determinado critério de desempenho, de tal forma que, para o conjunto de entrada informado, a rede seja capaz de calcular uma saída o mais próximo possível da saída desejada.

A aprendizagem de uma Rede Neural Artificial é um processo onde os parâmetros livres são adaptados através de um processo de estimulação pelo ambiente em que a rede está inserida. Com isso, o tipo de aprendizagem é determinado com base na maneira pela qual a modificação dos parâmetros ocorre. Em síntese, tem-se a seguinte sequência de eventos: a) a rede neural é estimulada por um ambiente; b) a rede neural sofre modificações nos seus parâmetros livres como resultado desta estimulação; c) a rede neural responde de uma maneira nova ao ambiente, devido às modificações ocorridas na sua estrutura interna (Haykin, 2001).

A capacidade de aprender das RNAs na realidade trata-se da habilidade de se adaptar alterando seu desempenho ao longo do tempo. O tipo de aprendizado é determinado pela técnica empregada no processo de ajustes dos pesos sinápticos. Um conjunto bem determinado destas regras constitui um algoritmo de aprendizagem ou treinamento, sendo que a maneira como o ambiente influencia a rede ela terá um aprendizado supervisionado, não supervisionado ou por reforço (Silva, 2019).

De acordo com Silva (2013) e Haykin (2001), há várias técnicas de aprendizagem de uma rede neural, sendo que as mais empregadas são a supervisionada (ou com um professor) e a não-supervisionada (ou sem professor).

Quando se emprega a aprendizagem supervisionada, são apresentados dados de entrada à rede, que calcula o valor de saída relativo a cada valor de entrada e compara aquele valor com a saída verdadeira que a rede deveria apresentar, sendo a diferença entre estes, conhecida como erro, utilizada para corrigir os pesos sinápticos, a fim de reduzir cada vez mais o erro obtido até um valor aceitável.

O processo de aprendizagem pode ser feito a partir de diversas técnicas, sendo a mais comum a técnica de aprendizagem backpropagation. Trata-se de um método de treinamento do tipo supervisionado, que pode ser completado através do processo de propagação e atualização de peso.

Neupane e Shakya (2017) descrevem a sequência de operação da rede por meio do algoritmo backpropagation: primeiramente apresenta-se um padrão à camada

de entrada da rede, a qual será transmitida a cada camada da rede, até que a resposta chegue à camada de saída. Após essa etapa, compara-se a saída obtida à saída desejada e então, se calcula o erro. O erro percorre o caminho inverso, ou seja, da camada de saída para a camada de entrada, promovendo modificações nos pesos das conexões. Esse método é repetido até que se obtenham erros aceitáveis na saída da rede.

No caso de não haver a figura do professor, como é o caso da aprendizagem não-supervisionada, o conjunto de treinamento não apresenta os valores da saída a serem atingidos, apenas os dados de entrada, o que força a rede a se auto-organizar de acordo com as particularidades apresentadas por cada dado de entrada e a ajustar dos pesos de forma a refletir esta organização. A RNA deve descobrir por si padrões, características, regularidades, correlações ou categorias nos dados de entrada. As unidades e conexões devem, portanto, apresentar algum grau de auto-organização.

De acordo com (BUDAMA, 2017), o computador foi projetado para fazer duas coisas muito bem: executar cálculos extremamente rápido e seguir ordens explícitas. Porém, no mundo real existem problemas mais complexos que não são possíveis de solucionar realizando um passo-a-passo bem definido.

É para resolver esses problemas que se utilizam as técnicas de aprendizado de máquina. O aprendizado de máquina nada mais é do que o uso de métodos computacionais para encontrar padrões em um determinado conjunto de dados e então conseguir tomar decisões ou realizar previsões no encontro de novos dados. Ou seja, a máquina é treinada usando uma quantidade de dados, métodos e técnicas que dão a ela a habilidade de aprender como executar uma tarefa, sem ser explicitamente programada para isso.

Existem três maneiras diferentes de realizar o treinamento de uma máquina, são eles:

- **Aprendizado supervisionado:** é utilizado quando se quer realizar a previsão de uma saída dada uma entrada e têm-se exemplos de pares de entrada/saída. A máquina então procura aprender qual é a regra da relação entrada/saída (GUIDO, 2016). Esta é a metodologia utilizada neste trabalho.

- **Aprendizado não supervisionado:** não são fornecidas as respostas, logo a tarefa da inteligência não é bem definida e seu desempenho não consegue ser claramente avaliada. A máquina deve tentar entender as diferentes estruturas e padrões e classificar ou predizer por conta própria (PATEL, 2019).

- **Aprendizado por reforço:** é aprender o que fazer como mapear situações em determinadas ações, para maximizar o sinal de recompensa. O computador não recebe a instrução de qual ação escolher, mas, em vez disso, deve descobrir quais ações geram mais recompensa, tentando executá-las (SUTTON, 2018).

Uma vez completo o treinamento, a rede neural deverá ser capaz de fornecer as saídas desejadas não somente para entradas conhecidas, mas também a resposta mais plausível para qualquer entrada relacionada aos padrões treinados.

Atualmente, muitas formas de RNAs são propostas e usadas com variações, principalmente no número de neurônios por camada, tipo de função de ativação dos neurônios, número de camadas ocultas e tipo de conexão entre os neurônios. Outro fator que interfere significativamente no tempo de treinamento e na eficiência de uma rede são os algoritmos utilizados, bem como os parâmetros e coeficientes utilizados em seu treinamento.

As RNA oferecem maneiras diferentes de resolver muitos problemas não-lineares que são difíceis de serem resolvidos por meio de técnicas convencionais. Muitos algoritmos são aplicados para a minimização e melhoria da aprendizagem da RNA. Esses algoritmos podem ser classificados como de minimização local e minimização global. Algoritmos de minimização local, tal qual o método gradiente descendente, são rápidos, mas convergem para mínimos locais. Por outro lado, os algoritmos de minimização global usam estratégias heurísticas para escapar dos mínimos locais (SHANG E BENJAMIN, 1996).

Em geral, sabe-se que nenhum dos algoritmos de Treinamento de Redes Neurais Artificiais é completo. Alguns algoritmos apresentam boas características, tais como alta velocidade de convergência. Entretanto, o erro quadrático médio na saída pode ainda ser relativamente alto. Outros algoritmos atingem pequeno erro quadrático médio na saída, porém apresentam em geral baixa velocidade de convergência.

A definição do algoritmo de treinamento a ser utilizado, bem como de seus parâmetros influenciam principalmente na fuga de mínimos locais, no desempenho da tarefa desejada e no tempo de treinamento. Por exemplo, taxas de aprendizado muito baixas acarretam um aprendizado muito lento, enquanto que taxas de aprendizado muito altas geram oscilações no processo de treinamento, dificultando e até mesmo impedindo a convergência do algoritmo de treinamento error backpropagation, pois a busca no espaço de soluções “pula” o erro mínimo.

### 3.3.5 Função de Transferência ou Ativação

As funções de ativação são uma característica extremamente importante das redes neurais artificiais. Elas basicamente decidem se um neurônio deve ser ativado ou não, ou seja, se a informação que o neurônio está recebendo é relevante ou deve ser ignorada (NWANKPA, 2018).

A função de ativação representa o efeito que a entrada interna e o estado atual de ativação exercem na definição do próximo estado de ativação da unidade.

A função de ativação tem como objetivo determinar um limite mínimo e máximo para a amplitude da saída do neurônio, para tal, é determinado um intervalo de ativação, descrito como intervalo unitário fechado, variando de 0 a 1 ou de -1 a 1, e posteriormente a função de ativação produz uma resposta de saída. Outra característica da função de ativação é como ela é ativada, podendo ser uma função limiar (degrau), linear por partes ou sigmóide.

As funções de ativação não lineares são diferenciáveis para que os parâmetros da rede possam ser ajustados usando backpropagation (KHAN *et al.*, 2018). Em outras palavras, para aprender relacionamentos complexos, precisa-se usar neurônios que empregam algum tipo de não linearidade (BUDUMA; LOCASCIO, 2017).

As três funções de ativação mais utilizadas são a função sigmoide, tangente hiperbólica, e ReLU (Rectified Linear Unit) (SEWAK *et al.*, 2018), alguns estudos têm mostrado a importância das funções de ativação para o aprendizado da rede neural.

Hornik (1991 & 1993) utilizou funções de ativação não-polinomiais. Singh (2003) e Chandra (2003) propuseram uma classe de funções sigmóides. Skoundrianos (2004) e Tzafestas (2004) propuseram uma nova função de ativação sigmoide com bons resultados para modelagem de sistemas dinâmicos de tempo discreto. Ma (2005) e Khorasani (2005) usaram a função polinomial Hermite com resultados satisfatórios. Wen (2005) e Ma (2005) propuseram a função Max-Piecewise-Linear (MPWL) para aproximação de funções. Gomes (2008) e Ludermir (2008) usaram as funções complemento log-log e probit para mostrar que quando os dados seguem uma distribuição binomial com características complemento log-log e probit o uso da função sigmóide logística na modelagem de redes neurais é inadequado. Gomes (2008) e Ludermir (2008) usaram as funções complemento loglog e probit em redes neurais para aproximação e regressão.

Estes e outros artigos citados mostram que a escolha da função de ativação é

considerada, por muitos especialistas, tão importante quanto a arquitetura e o algoritmo de aprendizagem da rede neural.

Segundo Haykin (2001) existem diversos tipos de funções de ativação, sendo as mais utilizadas:

Função limiar: normalmente restringe a saída da RNA em valores binários [0,1]. A saída do neurônio assume valor 0 quando seu resultado for negativo, e 1 caso contrário, como é possível observar na Equação 3:

$$f(u) = \begin{cases} 0, & \text{se } u \leq 0 \\ 1, & \text{se } u \geq 0 \end{cases} \quad (3)$$

Função sigmoideal: trata-se da função mais comum. É definida como uma função crescente com balanceamento adequado entre o comportamento linear e não linear e assume um intervalo de variação entre 0 e 1, é possível verificar na Equação 4.

$$f(u) = \frac{1}{1 + e^{(-au)}} \quad (4)$$

Função tangente hiperbólica: pelo fato da função logística apresentar valores de ativação apenas no intervalo (0, 1), em muitos casos ela é substituída pela função tangente hiperbólica, que preserva a forma sigmoideal da função logística, mas assume valores positivos e negativos, como pode-se observar na Equação 5.

$$f(u) = \tanh(u) \quad (5)$$

### 3.3.6 Análise da influência da Função de Ativação

Um dos componentes mais importantes do neurônio artificial é a sua Função de Ativação ou Transferência. Ela tem por objetivo limitar a amplitude válida do sinal de saída do neurônio a um valor finito. Normalmente, esta amplitude normalizada se encontra em um intervalo fechado unitário [0, 1] ou, em alguns casos, [-1, 1].

Cada camada tem uma função específica. A camada de saída recebe os

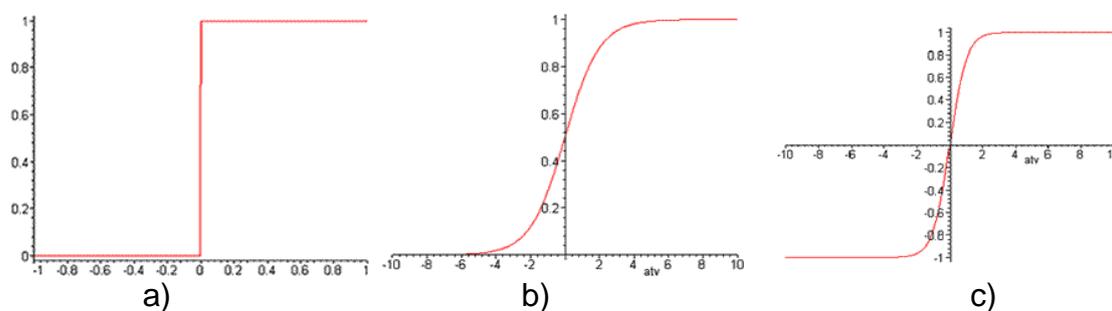
estímulos da camada intermediária (escondida) e constrói o padrão de resposta. As camadas intermediárias funcionam como extratoras de características. Seus pesos são uma codificação de características apresentadas nos padrões de entrada e permitem que a rede crie sua própria representação, mais rica e complexa, do problema.

A função de ativação Sigmóide apresenta um comportamento semelhante ao da função degrau, com a vantagem de apresentar uma curva suave e continuamente diferenciável, o que facilita o processo de treinamento dando direcionamento para a adequação dos pesos sinápticos com base no erro do sistema. Como apresenta valor de saída entre 0 e 1 é bastante empregada em camadas que precisam ter como saída apenas números positivos (REIS, 2016).

Uma função semelhante à Sigmóide é a tangente hiperbólica. A tangente hiperbólica é uma função escalonada da Sigmóide (Data Science Academy, 2019), mantendo as mesmas características, sendo muito aplicada, tendo o diferencial de possuir como saída valores entre -1 e 1.

Seguem algumas das funções não lineares mais conhecidas:

Figura 3.1 – Funções de ativação: a) limiar b) sigmóide c) tangente hiperbólica



Fonte: LNCC (2010).

### 3.3.7 Algoritmo de Treinamento das RNA

Os algoritmos de treinamento podem ser vistos como algoritmos de otimização que promovem a minimização da função erro, definida em relação aos valores esperados na saída, obedecendo aos limites de tolerância pré-estabelecidos. Dentre as diferentes técnicas de otimização aplicáveis, podem ser citados: a regra delta, o algoritmo de Boltzmann, o algoritmo *back-propagation*, o *simulated annealing*, algoritmos baseados em cadeias de Markov. Maiores detalhes podem ser encontrados em Kartalopoulos (1996).

Durante o treinamento o objetivo principal da rede é adequar os valores de peso da rede a fim de convergir à série de cálculos não lineares às saídas desejadas. Desta forma, o sentido do cálculo de ajuste dos pesos é contrário ao caminho da informação propagada pelas amostras de entrada na rede, pois o início dos cálculos de ajuste acontece a partir da camada de saída em direção à camada de entrada. Por isso este ajuste é conhecido como *back propagation*, ou retropropagação de erros.

O algoritmo de treinamento *back-propagation* é conceitualmente, uma generalização do algoritmo de aprendizado de Widrow e Hoff - "*Least Mean Square*" (LMS), também conhecido como Regra Delta. Ver Kosco (1992), Kovács (1996) e Kartalopoulos (1996). Este algoritmo enquadra-se na categoria de aprendizagem supervisionada.

O Algoritmo Backpropagation é o algoritmo de aprendizado supervisionado mais conhecido para treinamento de redes multicamadas. Esse algoritmo é utilizado para ajustar os pesos das camadas intermediárias com o objetivo de minimizar a função de custo da saída da rede. O treinamento ocorre em duas fases, que percorrem sentidos opostos: *forward* (para frente) e *backward* (para trás). A fase *forward* utiliza um padrão de entrada para definir a saída da rede. Os dados seguem o sentido *forward* da entrada para a saída. A fase *backward* usa a saída desejada e a saída da rede para ajustar os pesos das conexões. Os erros seguem o sentido *backward* da saída para a entrada (BRAGA *et al.*, 2000).

O Algoritmo de *Backpropagation* é uma das ferramentas mais utilizadas para o treinamento de Redes Neurais Artificiais. Entretanto, em algumas aplicações práticas, ele pode ser muito lento. Para permitir uma utilização mais ampla, muitas técnicas têm sido discutidas para acelerar o seu desempenho.

O problema de Treinamento de Redes Neurais Artificiais pode ser visto como um problema de otimização, onde é desejado minimizar o erro quadrático médio entre a saída desejada e aquela produzida pela Rede Neural Artificial. Existem vários tipos de algoritmos para treinamento das Redes Neurais, alguns com forte supervisão e outros com necessidade reduzida de informações provida pelo meio ambiente.

Existem diversos tipos de RNA e cada uma possui características próprias para melhor se adaptar a determinados problemas. Entre os diversos tipos de RNA destacam-se as Redes Perceptron de Múltiplas Camadas - Multilayer Perceptron (MLP) e as Redes de Funções de Base Radiais - Radial Basis Function (RBF). As RNA's podem ser usadas em diversos casos de previsão e predição, os quais se

estendem desde a caracterização de materiais até balanços hídricos de sistemas agrícolas (TAVARES *et al.*, 2020; SANTOS & SILVA, 2020).

### 3.3.8 Regra Delta

A Regra Delta ou regra de *Widrow-Hoff* é um processo de aprendizado por correção de erro também conhecido como, “*least mean square*”, tem sido aplicado para redes com uma única camada com funções de ativação puramente lineares.

O objetivo da Regra Delta é determinar um conjunto de pesos ( $W$ ) que minimize o erro quadrático sobre determinado número de amostras de treinamento. A obtenção de um ponto mínimo é alcançada por meio de um processo de iteração local, utilizando um conjunto de amostras. A Regra Delta é computacionalmente complexa e quando o tamanho das amostras é grande, o tempo de aprendizado pode se tornar longo (BENNEDIKTISSON, J. A. *et al.*, 1991).

A regra delta possui uma saída desejada  $d_j$ , assim o peso será proporcional à saída, (Equação 6).

$$\Delta W_{ij} = (d_j - a_j)a_i \quad (6)$$

A Regra Delta depende da função de ativação dos neurônios e minimiza o erro entre a saída desejada e o valor de ativação do neurônio. Para isso segue-se alguns passos:

- Iniciar todas as conexões com pesos aleatórios;
- Repetir até que o erro seja satisfatoriamente pequeno;
- Para cada par de treinamentos faça:
- Calcular a resposta obtida;
- Se o erro não for satisfatoriamente pequeno então:
- Atualizar os pesos: peso novo = peso anterior + taxa de aprendizado.

De acordo com Baughman (1995) uma limitação do algoritmo *Back propagation* é a demanda de tempo para o treinamento da rede, dependendo do tamanho da rede neural, o treinamento pode durar horas ou até dias dependendo da velocidade do processamento de dados do computador. Para sanar esta limitação pesquisadores investigaram novos procedimentos de treinamento. Desta forma, um

destes procedimentos é utilizar o algoritmo *vanilla Back propagation*, baseado no aprendizado de gradiente descendente (Equação 7).

$$[W + 1] = [W_a] + [lr] \times [e] \times [t] \quad (7)$$

Onde:

$W$  = peso;

$W_a$  = peso anterior;

$lr$  = taxa de aprendizagem;

$e$  = entrada de variáveis;

$t$  = tempo de correção do gradiente descendente.

Outro método de treinamento usando o aprendizado de gradiente descendente é a regra delta, um método iterativo que minimiza o erro de média quadrática. Este algoritmo é semelhante ao algoritmo de *vanilla Back propagation*, mas possui várias diferenças entre elas pode - se citar a adição de um *momentum* (Equação 8). O parâmetro momento é um peso extra adicionado aos fatores de peso quando esses são ajustados. Este parâmetro é adicionado ao treinamento com o objetivo de acelerar o treinamento da rede uma vez que quando se muda os fatores do peso melhora-se o rendimento do treinamento da rede (BAUGHMAM *et al.*, 1995).

$$[W + 1] = [W_a] + [lr] \times [e] \times [t] + [tm] \times [ap] \quad (8)$$

Onde:

$tm$  = termo *momentum*;

$ap$  = ajuste de peso.

Nas resoluções das equações não lineares da RNA o uso do método do gradiente descendente com o *Back propagation vanilla*, pode ocorrer desse algoritmo não alcançar a soma do erro médio quadrático mínimo. Assim, se faz necessário a adição de um termo *momentum*.

### 3.3.9 Estudo de Estado da Arte

Na área da adsorção diversos estudos foram promissores na degradação de corantes utilizando RNA, para as análises, Assefi *et al.*, (2013) e Dil *et al.*, (2015),

fizeram um comparativo entre várias tipologias de redes, os resultados analisados estatisticamente com metodologias de superfícies de resposta obtiveram excelente grau de aceitação.

Diversos trabalhos têm surgido estudando o processo de adsorção para a remoção de corantes fazendo uso de diferentes biomassas, como o realizado por Soares *et al.* (2019) que fez uso do carvão ativado do bagaço de cana-de-açúcar na adsorção do corante Erionyl Yellow A-R. Outro foi o realizado por Arriel *et al.*, (2019) que fez uso do bagaço de malte bruto e lavado na remoção do corante reativo azul BF-5G, obtendo resultados promissores. O resíduo oriundo do gengibre também foi estudado para a remoção de corantes obtendo bons resultados para a adsorção do corante alaranjado de metila (DA SILVA *et al.*, 2020).

Para lidar com essas limitações, têm sido desenvolvidos modelos usando pontos de dados experimentais precisos para compreender e prever um processo. Nesse sentido, modelos inteligentes como redes neurais artificiais (RNAs) têm sido empregados com eficiência para prever processos de adsorção (ELEMEN *et al.*, 2012; GHAEDI *et al.*, 2013; NIA *et al.*, 2014; GHAEDI *et al.*, 2017).

Os dados experimentais disponíveis no processo de adsorção fornecem informações valiosas sobre o entendimento básico do comportamento e desenvolvimento desses processos, onde as avaliações experimentais nem sempre são possíveis. Técnicas para obter dados no processo de adsorção podem ser difíceis, caras e demoradas (GHAEDI *et al.*, 2015).

Atualmente existem vários processos de descontaminação e degradação de corantes têxteis em efluentes industriais (PUCHANA-ROSETO *et al.*, 2018; SARKER *et al.*, 2019; SATHASIVAM *et al.*, 2019; TOMASSONI *et al.*, 2019; DALARI *et al.*, 2020), dentre eles a adsorção (QAMAR *et al.*, 2020; KAUSAR *et al.*, 2018; AFROZE; ZEN, 2018).

Dutta *et al.*, (2010) propuseram a aplicação de redes neurais para prever a eficiência e capacidade de adsorção de corantes tóxicos presentes em águas residuárias. A rede foi treinada com as respostas obtidas variando o pH da fase líquida, concentração de TiO<sub>2</sub> e concentração de corante. Após diversas tentativas sobre as funções de transferência e os algoritmos de treinamento, percebeu-se que a combinação de retropropagação resiliente com a função linear positiva mostrou melhores resultados, com  $R^2$  maior que 0,98.

## 4 METODOLOGIA

### 4.1 Localização

O trabalho foi desenvolvido no Laboratório de Pesquisa em Ciências Ambientais (LAPECA), que pertence ao Departamento de Engenharia Sanitária e Ambiental (DESA) localizado no Centro de Ciências e Tecnologia (CCT) da Universidade Estadual da Paraíba (UEPB).

### 4.2 Dados Experimentais

Em todo o processo de modelagem, a familiaridade com os dados disponíveis é de máxima importância.

Os dados utilizados neste trabalho são provenientes da pesquisa desenvolvida por Xavier (2020), em que ela, usou materiais lignocelulósicos (bagaço de cana de açúcar) para adsorver corantes têxteis, presentes em águas residuárias.

O bagaço da cana-de-açúcar utilizado para estudo foi adquirido no comércio da cidade de Campina Grande – PB. O material foi coletado logo após a etapa de moagem da cana, de forma que não ocorresse a degradação microbológica das amostras do bagaço. Foi cortado em pedaços uniformes com o intuito de se ter um material homogêneo durante a secagem. Em seguida, foi lavado por três vezes em água destilada para a retirada de possíveis impurezas e reservado para secagem.

Com o objetivo de avaliar as possíveis mudanças no processo de adsorção realizaram-se tratamentos químicos no material adsorvente de forma que fosse possível comparar os materiais *in natura* e após os tratamentos químicos.

Xavier (2020) na primeira etapa realizou os experimentos e analisou o planejamento experimental, onde o material foi seco, triturado e submetido ao tratamento em solução de ácido clorídrico sob constante agitação mecânica, durante um período de 24 horas à temperatura ambiente. Após o período pré-estabelecido, o bagaço de cana-de-açúcar foi tratado com ácido e lavado com água destilada até a total remoção dos íons de cloreto, esta remoção foi monitorada por meio de nitrato de prata até que a solução não se apresentasse turva na presença deste, como pode ser observado na figura 4.

No trabalho, foi usado o adsorvente na forma *in-natura* e após a ativação química e alcalina, usando soluções HCl e NaOH 0,1 mol. L<sup>-1</sup>.

Figura 4 – Bagaço de cana de açúcar (a) antes e (b) depois da preparação.



Fonte: Xavier (2020).

Para o estudo foi utilizado um conjunto de dados composto por sete variáveis, sendo cinco variáveis independentes e duas respostas. As variáveis independentes correspondem a: massa, concentração, pH, Rotação e Tempo; as respostas: unidade de massa adsorvente e eficiência de remoção, as variáveis independentes foram submetidas a experimentação por Xavier (2020), onde foi utilizado o material de forma in-natura, um tratamento ácido e um alcalino. Para o material na forma in-natura foram empregados um conjunto de dados com 138 valores para cada uma das variáveis independentes, no tratamento ácido foram utilizados 92 valores, e no alcalino 46.

### 4.3 Regressão Linear Múltipla

#### 4.3.1 Aplicação da Regressão Linear Múltipla

A RLM foi realizada levando em consideração o conjunto de dados. Este conta com a presença de cinco variáveis de entrada e duas variáveis de saída. Por definição, com a utilização de apenas uma variável de entrada, é possível a implementação da RLS, como o conjunto de dados consta de cinco variáveis de entrada é preciso fazer uso da RLM. O conjunto de dados é formado por cinco variáveis independentes, massa, concentração, pH, Rotação e tempo, e duas variáveis dependentes, Qt e Rem como pode ser observado na Equação 9.

O modelo

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_k x_k + \varepsilon \quad \text{Equação (9)}$$

$x_1$  : Massa

$x_2$ : Concentração

$x_3$ : pH

$x_4$ : Rotação

$x_5$ : Tempo

$y_1$ : Qt

$y_2$ : Rem

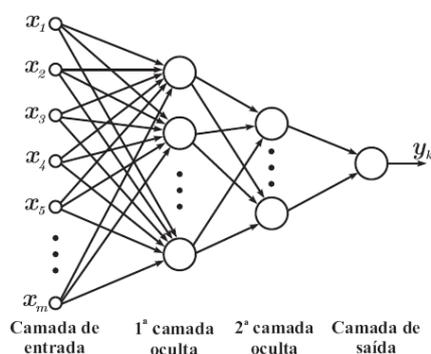
é chamado de modelo de regressão linear múltipla com  $k$  variáveis regressoras. Os parâmetros  $\beta_i$  ( $i = 0$  a  $k$ ) são chamados de coeficientes de regressão parciais. As variáveis independentes são: massa, concentração, pH, Rotação e tempo e as dependentes Qt e Rem.

#### 4.4 Rede Neural Artificial

##### 4.4.1 Determinação da Arquitetura da Rede Neural

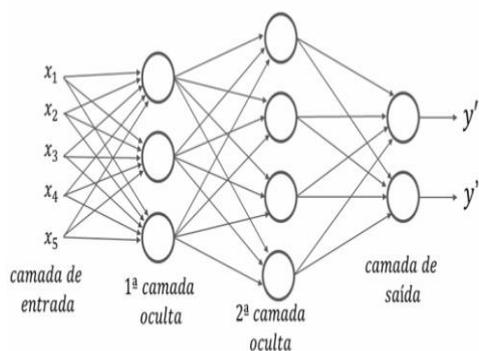
Na criação da rede, considera-se uma arquitetura do tipo *Feedforward* múltipla camada que pode ser observado nas figuras 4.1(a) e 4.2(b).

Figura 4.1(a) - Rede *Feedforward* de Camadas Múltiplas.



Fonte: J3sus Neto, 2022.

Figura 4.1(b) - Rede *Feedforward* de Camadas Múltiplas duas saídas



Fonte: IGTIBlog, 2022.

#### 4.4.2 Variáveis do Modelo

Na Tabela 1, podem ser observadas as variáveis utilizadas para a realização da pesquisa.

**Tabela 1 - Variáveis Utilizadas no Modelo**

Variáveis de Entrada	Variáveis de Saída
Massa de Adsorvente	Unidade de Massa de Adsorvente
Concentração do Solutivo	Eficiência de Remoção
pH da Solução	
Tempo de Contato	
Agitação/Rotação	

Fonte: o autor, 2022.

#### 4.4.3 Planejamento Simulado para uma camada

Na tabela 2, podem ser observados os fatores e os níveis utilizados no planejamento simulado. Em cada nível, foram realizados 1 bloco de 12 experimentos com 5 repetições para cada tratamento empregado. Por exemplo, para o nível 1, o número de neurônios permaneceu fixo igual a 20, o número de épocas variou a cada 4 experimentos de 2000, 4000 e 8000, a função de ativação alternou entre Sigmóide e Relu e a validação dos dados alternou em 30 e 40%.

**Tabela 2 - Planejamento simulado para uma camada**

Fator	Nível 1	Nível 2	Nível 3
Neurônios	20	40	80
Épocas	2000	4000	8000
Função	Sigmoid	Relu	Sigmoid
Divisão (%)	30	40	30

Fonte: o autor, 2022.

#### 4.4.4 Planejamento Simulado para duas camadas

Na tabela 3, podem ser observados os fatores e os níveis utilizados no planejamento simulado. Em cada nível, foram realizados 1 bloco com 36 experimentos com 5 repetições para cada tratamento empregado. Por exemplo, para o nível 1, o número de neurônios permaneceu fixo igual a 30 para a primeira camada, o número de neurônios da segunda camada variou a cada 12 experimentos de 30, 60 e 90, o número de épocas variou a cada 4 experimentos de 2000, 4000 e 8000, a função de ativação alternou entre Sigmóide e Relu e a validação dos dados alternou em 30 e 40%.

**Tabela 3 - Planejamento simulado para duas camadas**

Fator	Nível 1	Nível 2	Nível 3
Neurônios	30-30	30-60	30-90
Épocas	2000	4000	8000
Função	Sigmóide	Relu	Sigmóide
Divisão (%)	30	40	30

Fonte: o autor, 2002.

#### 4.5 Utilização da Linguagem Python em Redes Neurais

O Python também é multiplataforma, e caso apareça alguma que ainda ele não suporte, os desenvolvedores podem modificar o código para que seja possível compilar para aquela linguagem, pois o Python também é uma linguagem livre. Assim, todos esses fatores fazem com que o Python seja uma das preferidas linguagens para o trabalho com inteligência artificial, data mining e machine learning.

Os fatores que tornaram decisiva a escolha de tal linguagem para o projeto foram: A facilidade de aprendizagem e a variedade de bibliotecas voltadas para análise de dados.

Considerada como sendo uma linguagem de alto nível na solução de cálculos matemáticos complexos por envolver grande número de interação. Será utilizado na pesquisa para o treinamento das redes neurais artificiais, ajuste dos pesos e os cálculos do erro das funções de ativação.

Para a construção do modelo foram consideradas as seguintes bibliotecas do Python:

Numpy - Essa é uma biblioteca que faz o processamento de matrizes e vetores (objetos do tipo "array"). Ela fornece um conjunto de frameworks para manipular e gerenciar essas informações de forma rápida e eficiente. Permite ainda a integração não somente com diversas bibliotecas Python, mas também com outras linguagens de programação como C e C++.

Pandas – é uma das bibliotecas Python mais importantes para a ciência de dados. Ela se destaca porque, além de ser completa, é fácil de utilizar.

Scikit Learn - Essa é uma biblioteca simples que trata de um assunto cada vez mais relevante: machine learning. Ela também é escrita em outras linguagens, como C e C++, porém a maior parte da biblioteca padrão é programada na linguagem Python.

Matplotlib - é uma biblioteca Python que usa o Python Script para escrever gráficos e plotagens bidimensionais.

Scipy - é uma biblioteca Python de código aberto usada para computação científica e técnica. É uma biblioteca Python gratuita. É muito adequado para aprendizado de máquina. No entanto, o cálculo não é a única tarefa que torna o scipy especial. Também é muito popular para manipulação de imagens.

Seaborn – atua em cima do matplotlib e ajuda a melhorar o visual dos gráficos, dando uma aparência mais bem acabada.

Statsmodels – Biblioteca com uma grande variedade de funções estatísticas.

TensorFlow – é uma das bibliotecas Python mais famosas. É fácil de aprender, gratuita, open source e traz diversas ferramentas para programação de machine learning. Ela se apresenta como uma solução para diversos problemas dessa área.

Keras - é outra biblioteca de código aberto escrita em Python que traz funcionalidades muito interessantes. Ela é muito utilizada na inteligência artificial, e é ótima para o monitoramento de redes detalhadas. Seu objetivo é ser fácil de usar, extensível e modular. Ela traz diversas funções que constroem partes fundamentais das redes neurais, como camadas, otimizadores, objetivos, entre outras.

## 5 RESULTADOS E DISCUSSÕES

### 5.1 Descrição estatística dos bancos de dados experimentais

Para o conjunto de dados experimentais (in-natura, ácido e alcalino) foram obtidas as estatísticas descritivas dos parâmetros analisados que podem ser observadas nas tabelas 4, 5 e 6.

**Tabela 4 – Bagaço de cana de açúcar com tratamento ácido**

Variável	N	Média	DesvPad	Variância	CoefVar	Soma de quadrados
Massa (g)	92	0,250	0,088	0,008	35,58	6,470
concentração	92	150,000	29,650	879,12	19,77	2150000,000
pH	92	5,000	1,186	1,407	23,72	2428,000
Rotação	92	50,000	5,930	35,165	11,86	233200,000
Tempo (s)	92	45,000	8,895	79,121	19,77	193500,000
Qt (mg. g <sup>-1</sup> )	92	15,727	7,422	55,089	47,19	27769,213
Rem (%)	92	90,543	7,390	54,613	8,16	759195,774

Fonte: o autor, 2022.

**Tabela 5 – Bagaço de cana de açúcar tratamento alcalino**

Variável	N	Média	DesvPad	Variância	CoefVar	Soma de quadrados
Massa (g)	46	0,250	0,089	0,008	35,780	3,235
concentração	46	150,000	29,810	888,890	19,880	1075000,000
pH	46	5,000	1,1930	1,422	23,850	1214,000
Rotação	46	50,000	5,963	35,556	11,930	116600,000
Tempo (s)	46	45,000	8,940	80,000	19,880	96750,000
Qt (mg. g <sup>-1</sup> )	46	11,390	5,2420	27,481	46,030	7203,821
Rem (%)	46	67,870	15,720	247,200	23,170	222984,770

Fonte: o autor, 2022.

**Tabela 6 – Bagaço de cana de açúcar tratamento in-natura**

Variável	N	Média	DesvPad	Variância	CoefVar	Soma de quadrados
Massa (g)	138	0,250	0,088	0,007	35,510	9,705
concentração	138	150,000	29,600	875,910	19,730	3225000,000
pH	138	5,000	1,184	1,401	23,680	3642,000
Rotação	138	50,000	5,919	35,036	11,840	349800,000
Tempo (s)	138	45,000	8,879	78,832	19,730	290250,000
Qt (mg. g <sup>-1</sup> )	138	14,091	6,513	42,421	46,220	33213,009
Rem (%)	138	82,178	10,005	100,093	12,170	945646,664

Fonte: o autor, 2022.

## 5.2 Modelagem Usando a Regressão Linear Múltipla

É possível verificar nas tabelas 7 e 8, os valores encontrados do *AIC* e o *BIC*, estes são critérios utilizados para a escolha do melhor modelo da regressão linear múltipla, para cada tratamento aplicado das variáveis *Qt* e *Rem*. Foram realizadas 5 repetições para cada experimento, ao realizar as repetições, o programa seleciona e fornece como resposta o menor valor para o critério, logo, tanto no *AIC* quanto no *BIC* nas tabelas 7 e 8 podem ser verificados o menor valor obtido para o modelo ajustado em cada tratamento aplicado.

AKPA e UNUABONAH (2011) concluíram que o uso de critérios de informação, tais como o critério de Akaike ou Baysiano, são melhores funções para ranquear modelos de isotermas de adsorção que apresentam diferente número de parâmetros. Estes critérios penalizam modelos com grande número de parâmetros e grande número de dados experimentais. Esta penalização faz sentido, uma vez que a explicação mais simples possível é procurada.

As métricas *AIC* e *BIC* são dois índices usados para uma comparação entre modelos na qual possuem muitos parâmetros e constantes em análise comparativa com outros modelos com menos parâmetros e constantes, a proposta dessas métricas são de penalizar outros modelos que possuem muitos parâmetros (STHDA, 2018).

**Tabela 7 – Valores do AIC e BIC da variável Qt**

Tratamento	AIC	BIC
Ácido	469,8	485,0
Alcalino	237,1	248,1
In-natura	686,2	703,8

Fonte: o autor, 2022.

**Tabela 8 – Valores do AIC e BIC da variável Rem**

Tratamento	AIC	BIC
Ácido	559,0	574,2
Alcalino	354,3	365,3
In-natura	937,7	955,3

Fonte: o autor, 2022.

De acordo com Garson (2011), o pesquisador pode utilizar o teste de Durbin-Watson, *d*, para detectar a presença de autocorrelação em seus dados. A estatística *d* varia entre 0 e 4 de tal modo quanto mais perto de 0 maior é a autocorrelação positiva e quanto mais perto de 4 maior é autocorrelação negativa. Valores entre 1,5 e 2,5 sugerem independência das observações. Assim, caso não exista autocorrelação o valor esperado de *d* é 2; valores significativamente inferiores a 2

indicam autocorrelação positiva; e, valores significativamente superiores a 2 indicam autocorrelação negativa (SAS Institute, 1993).

Na tabela 9, pode ser observado o valor do teste de Durbin-Watson, a autocorrelação a independência dos dados e o teste de Shapiro-Wilk. Pode ser verificado que de acordo com o teste de Shapiro-Wilk, para todos os tratamentos aplicados os dados não são similares a uma distribuição normal, já pelo teste de Durbin-Watson pode ser verificado a existência de autocorrelação positiva nas seguintes situações, para dois tratamentos ácidos, tanto para a variável Qt quanto a Rem, e para o tratamento in-natura da variável Rem.

**Tabela 9 - Teste de Durbin-Watson e Shapiro-Wilk**

Tratamento	Teste Durbin-Watson	Teste Shapiro-Wilk	Auto Correlação	Observações
Ácido – Qt	1,171	0,747	Positiva	Dependentes
Alcalino - Qt	2,118	0,747	Pouca/nenhuma	Independentes
In-natura Qt	1,660	0,747	Pouca/nenhuma	independentes
Ácido – Rem	1,245	0,747	Positiva	Dependentes
Alcalino - Rem	1,706	0,747	Pouca/Nenhuma	Independentes
In-natura - Rem	1,034	0,747	Positiva	Dependentes

Fonte: o autor, 2022.

Segundo Bisquera, Sarierra e Martinez (2004) duas ou mais distribuições de valores apresentam homocedasticidade quando as respectivas variações são equivalentes em linhas estatísticas, ou seja, diversas distribuições de valores possuem homocedasticidade ou homogeneidade de variâncias quando as diferenças notadas entre as variâncias não são estatisticamente significativas. Contrariamente, quando tal fato não acontece, afirma-se que ocorre heterocedasticidade.

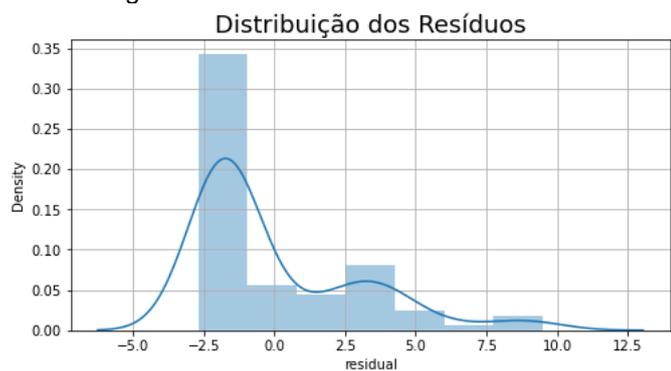
Outra suposição apresentada por Hair Jr. *et al* (2009) em análise de regressão múltipla é a variância constante do termo de erro. De acordo com Hair Jr. *et al.*, (2009) a presença de variâncias desiguais (heterocedasticidade) é uma das violações mais comuns de suposição.

O modelo de regressão linear múltipla deve ser avaliado quanto a algumas suposições, em relação ao resíduo, para verificação da validade da equação. O resíduo é a principal medida de avaliação da variável estatística em relação à qualidade do ajuste, e é analisado quanto à sua homocedasticidade, independência e normalidade conforme HAIR *et al.*, (2010).

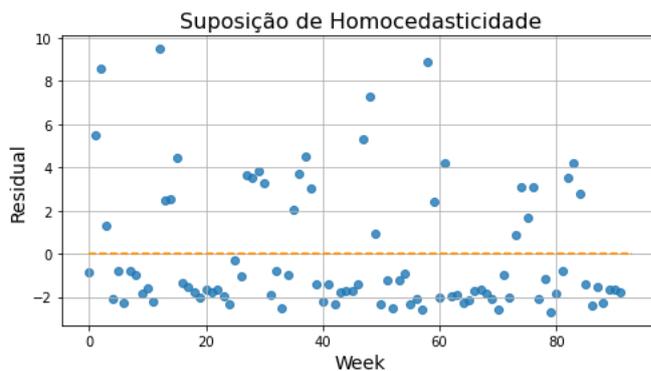
- **Homocedasticidade:** A homocedasticidade garante que a variância do erro seja constante para todas as variáveis. Para o teste de homocedasticidade, é aplicado o teste de Breusch-Pagan e deseja-se não rejeitar a hipótese  $H_0$ , ou seja, o resultado para o *p-valor* do teste deve ser maior que o nível de significância considerado.
- **Independência:** O teste de independência, busca assumir que os resíduos do modelo de regressão linear múltipla estimado, não pode ser modelado por nenhuma outra técnica estatística preditiva. Graficamente, a aleatoriedade dos resíduos, é demonstrada por uma distribuição aleatória no gráfico de resíduos ou por meio do teste de hipótese de Durbin-Watson. Deseja-se que o teste gere um resultado para o *p-valor* maior que o nível de significância estabelecido, ou seja, não se pode rejeitar  $H_0$  para se garantir a independência dos resíduos.
- **Normalidade:** A normalidade dos resíduos, assim como os outros testes de significância para os resíduos, deve ser testada para garantir a confiabilidade do ajuste do modelo de regressão. Um dos testes mais utilizados para verificação da normalidade de um conjunto de dados é o teste de Shapiro-Wilk, que deve resultar em um *p-valor* maior que o nível de significância prefixado, ou seja, deseja-se não rejeitar a hipótese  $H_0$  para se verificar a normalidade.

Na figura 5 (a) e (b) tratamento ácido, (c) e (d) tratamento alcalino e (e) e (f) in-natura podem ser observados os gráficos de resíduos e homocedasticidade para a variável Qt. Na Figura 6 (a) e (b) tratamento ácido, (c) e (d) tratamento alcalino e (e) e (f) in-natura podem ser observados os gráficos de resíduos e homocedasticidade para a variável Rem referente aos tratamentos utilizados.

Figura 5 – Gráficos de Resíduos e Homocedasticidade da variável Qt



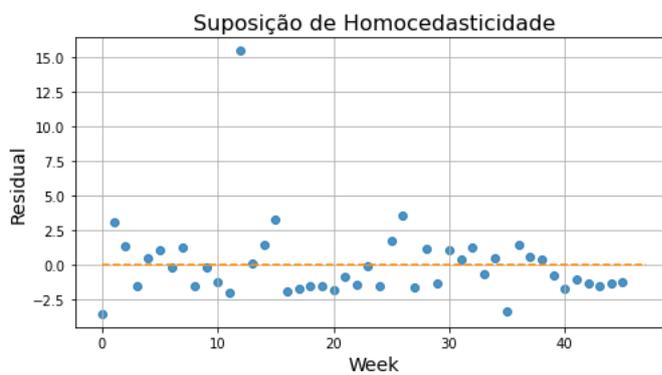
(a)



(b)



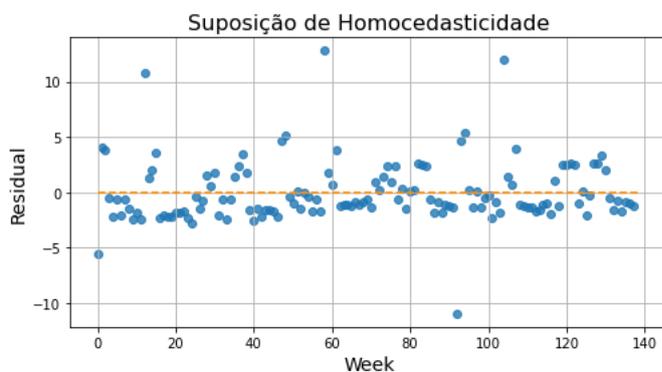
(c)



(d)



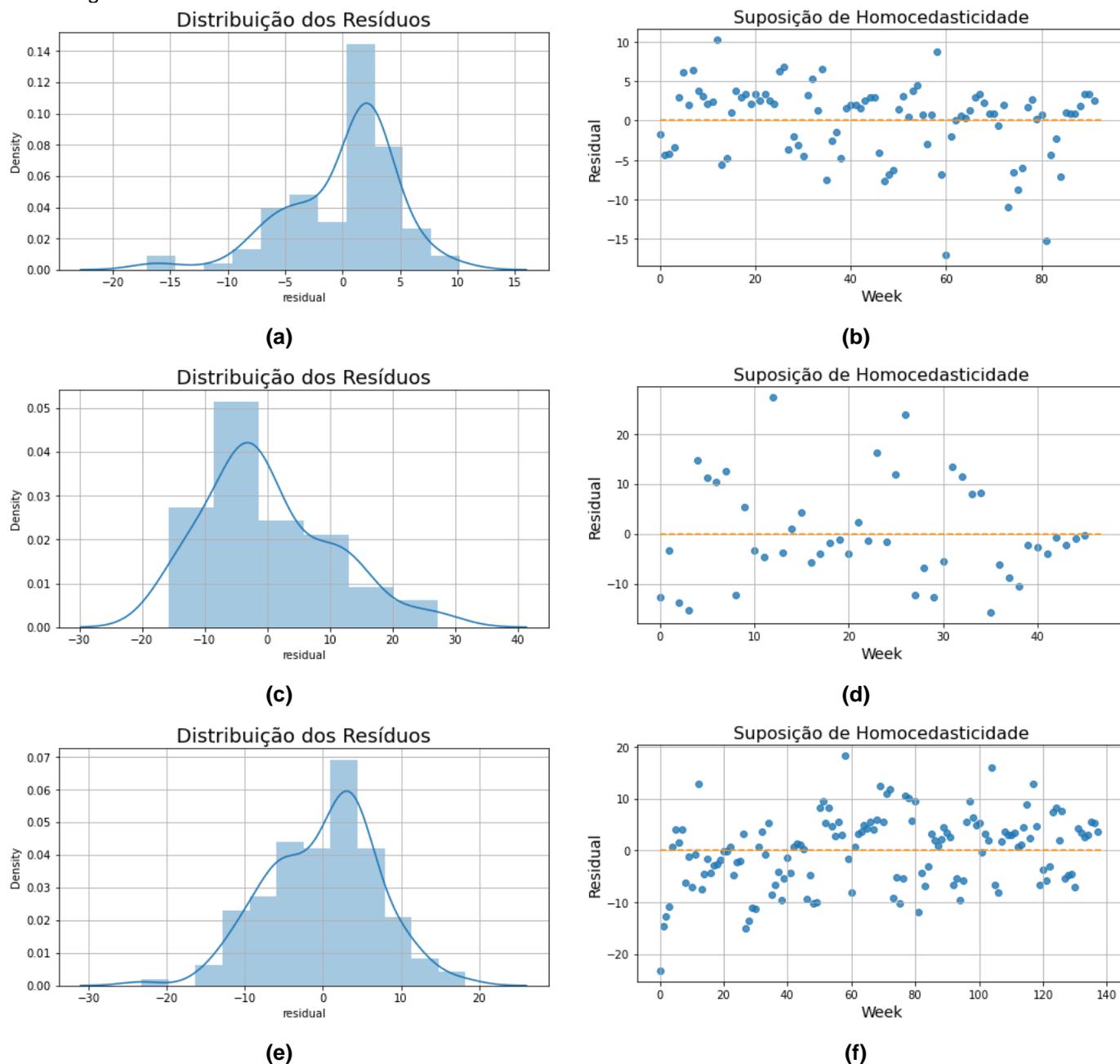
(e)



(f)

Fonte: o autor, 2022.

Figura 6 – Gráficos de Resíduos e Homocedasticidade da variável Rem



Fonte: o autor, 2022.

Nas tabelas 10 a 21 podem ser observados os valores médios do  $R^2$  e  $R_a^2$  para treino e teste dos tratamentos aplicados com 30 e 40% de validação, gerados por meio de 5 repetições de cada experimento, onde foi obtido o ajuste do modelo por meio da Análise de Regressão, através da técnica de RLM, para as variáveis Qt e Rem. Pode ser verificado, as melhores respostas para a variável Qt, no treino e teste do tratamento ácido, no treino com 40% foi obtido um valor de 84,04%, já para o teste com 30% foi observado um valor de 84,16%, no caso da variável Rem, as melhores respostas obtidas foram para o tratamento alcalino, ambas com 30% para treino e

teste, no treino foi obtido um valor de 56,26% e no teste foi verificado um valor de 57,90%.

**Tabela 10 – Treino do tratamento ácido Qt com 5 repetições**

Qt	$R^2$ -30%	$R_a^2$ -30%	$R^2$ -40%	$R_a^2$ -40%
1	0,8495	0,8366	0,8322	0,8151
2	0,8514	0,8386	0,8514	0,8362
3	0,8524	0,8397	0,8565	0,8419
4	0,8483	0,8352	0,8856	0,8740
5	0,8435	0,8300	0,8502	0,8349
Média	0,8490	0,8360	0,8551	0,8404

Fonte: o autor, 2022.

**Tabela 11 – Teste do tratamento ácido Qt com 5 repetições**

Qt	$R^2$ -30%	$R_a^2$ -30%	$R^2$ -40%	$R_a^2$ -40%
1	0,8501	0,8160	0,8637	0,8418
2	0,8878	0,8624	0,8620	0,8397
3	0,8811	0,8541	0,8629	0,8408
4	0,8674	0,8373	0,8487	0,8244
5	0,8684	0,8385	0,8615	0,8391
Média	0,8790	0,8416	0,8597	0,8371

Fonte: o autor, 2022.

**Tabela 12 – Treino do tratamento alcalino Qt com 5 repetições**

Qt	$R^2$ -30%	$R_a^2$ -30%	$R^2$ -40%	$R_a^2$ -40%
1	0,9117	0,8947	0,9107	0,8894
2	0,6782	0,6163	0,8649	0,8327
3	0,9051	0,8869	0,6807	0,6047
4	0,7246	0,6716	0,9058	0,8833
5	0,8682	0,8429	0,9122	0,8913
Média	0,8175	0,7824	0,8548	0,8202

Fonte: o autor, 2022.

**Tabela 13 – Teste do tratamento alcalino Qt com 5 repetições**

Qt	$R^2$ -30%	$R_a^2$ -30%	$R^2$ -40%	$R_a^2$ -40%
1	0,7783	0,6398	0,5652	0,3980
2	0,8932	0,8265	0,7851	0,7025
3	0,6722	0,4673	0,8745	0,8262
4	0,9513	0,9209	0,6825	0,5604
5	0,7357	0,5706	0,7217	0,6147
Média	0,8061	0,6850	0,7258	0,6203

Fonte: o autor, 2022.

**Tabela 14 – Treino do tratamento in-natura Qt com 5 repetições**

Qt	$R^2$ -30%	$R_a^2$ -30%	$R^2$ -40%	$R_a^2$ -40%
1	0,8035	0,7925	0,7456	0,7289
2	0,8081	0,7974	0,7910	0,7773
3	0,7857	0,7738	0,8169	0,8048
4	0,8017	0,7907	0,7460	0,7293
5	0,8144	0,8041	0,8226	0,8109
Média	0,8026	0,7917	0,7844	0,7702

Fonte: o autor, 2022.

**Tabela 15 – Teste do tratamento in-natura Qt com 5 repetições**

Rem	$R^2$ -30%	$R_a^2$ -30%	$R^2$ -40%	$R_a^2$ -40%
1	0,9097	0,8971	0,9162	0,9078
2	0,8577	0,8379	0,8697	0,8567
3	0,8883	0,8728	0,8253	0,8078
4	0,8793	0,8626	0,9085	0,8994
5	0,8812	0,8647	0,8490	0,8339
Média	0.8036	0,7917	0,7844	0,7702

Fonte: o autor, 2022.

**Tabela 16 - Treino do tratamento ácido Rem com 5 repetições**

Rem	$R^2$ -30%	$R_a^2$ -30%	$R^2$ -40%	$R_a^2$ --40%
1	0,7090	0,6428	0,6441	0,5868
2	0,5754	0,4790	0,6756	0,6233
3	0,6668	0,5910	0,4905	0,4084
4	0,6006	0,5098	0,5447	0,4713
5	0,6469	0,5667	0,6260	0,5657
Média	0,6397	0,5578	0,5961	0,5311

Fonte: o autor, 2022.

**Tabela 17 – Teste do tratamento ácido Rem com 5 repetições**

Rem	$R^2$ -30%	$R_a^2$ -30%	$R^2$ -40%	$R_a^2$ --40%
1	0,5603	0,5224	0,5722	0,5285
2	0,6242	0,5918	0,5425	0,4958
3	0,5585	0,5204	0,6477	0,6117
4	0,6288	0,5968	0,6185	0,5796
5	0,5705	0,5335	0,5429	0,4962
Média	0,5884	0,5529	0,5847	0,5423

Fonte: o autor, 2022.

**Tabela 18 – Treino do tratamento alcalino rem com 5 repetições**

Rem	$R^2$ -30%	$R_a^2$ -30%	$R^2$ -40%	$R_a^2$ --40%
1	0,6152	0,5412	0,6015	0,5067
2	0,6477	0,5799	0,7555	0,6973
3	0,6058	0,5300	0,7082	0,6388
4	0,6239	0,5516	0,4877	0,3658
5	0,6734	0,6106	0,5305	0,4187
Média	0,6332	0,5626	0,6166	0,5254

Fonte: o autor, 2022.

**Tabela 19 - Teste do tratamento alcalino Rem com 5 repetições**

Rem	$R^2$ -30%	$R_a^2$ -30%	$R^2$ -40%	$R_a^2$ --40%
1	0,6701	0,4639	0,6479	0,5126
2	0,8616	0,7751	0,6923	0,5740
3	0,6891	0,4949	0,5622	0,3939
4	0,8396	0,7394	0,8046	0,7295
5	0,6442	0,4219	0,7648	0,6743
Média	0,7409	0,5790	0,6943	0,5768

Fonte: o autor, 2022.

**Tabela 20 – Treino do tratamento in-natura Rem com 5 repetições**

Rem	$R^2$ -30%	$R_a^2$ -30%	$R^2$ -40%	$R_a^2$ -40%
1	0,5539	0,5291	0,4685	0,4336
2	0,5305	0,5045	0,4719	0,4372
3	0,5472	0,5220	0,4409	0,4042
4	0,5390	0,5134	0,5841	0,5568
5	0,5531	0,5283	0,5400	0,5098
Média	0,5447	0,5194	0,5010	0,4683

Fonte: o autor, 2022.

**Tabela 21 – Teste do tratamento in-natura Rem com 5 repetições**

Rem	$R^2$ -30%	$R_a^2$ -30%	$R^2$ -40%	$R_a^2$ -40%
1	0,4782	0,4058	0,6561	0,6217
2	0,5055	0,4369	0,6743	0,6417
3	0,5303	0,4650	0,6608	0,6269
4	0,5054	0,4368	0,4468	0,3915
5	0,5108	0,4429	0,5188	0,4706
Média	0,5060	0,4374	0,5913	0,5504

Fonte: o autor, 2022.

A análise de regressão foi aplicada e apresentou o resultado dos coeficientes de determinação  $R^2$  do modelo. Este coeficiente é uma medida do modelo linear generalizado em relação aos valores observados. É utilizado para regressões lineares e seu valor varia entre 0 e 1, que pode ser associado a valores percentuais (MANN, 2007).

Para a variável Qt com 40% de validação para treino do tratamento ácido a RLM apresentou a melhor média para o  $R_a^2$ , com 84,04% para 30% de validação foi obtido o valor de 84,16% o teste. O tratamento alcalino apresentou 82,02% de valor para o  $R_a^2$  no treino com 40% de validação, já no teste foi obtido 68,50% com 30% de validação. No tratamento in-natura com 30 e 40% de validação dos dados foi obtido o mesmo resultado para o coeficiente de determinação ( $R_a^2$ ) para treino e teste, para o treino foi obtido o valor de 77,02% e o teste um valor de 79,17%.

Já para a variável Rem, no tratamento alcalino foram obtidos os melhores valores para o  $R_a^2$  com 30% de validação dos dados, no treino o valor obtido foi de 56,26% e no teste 57,90%. O tratamento ácido também apresentou as melhores respostas com 30% de validação dos dados, mas os valores obtidos foram menores do que no tratamento alcalino, no treino foi obtido o valor de  $R_a^2$  igual a 55,78% e no teste 57,90%. O tratamento in-natura apresentou a melhor resposta para o treino com 30% de validação e um valor de  $R_a^2$  igual a 51,94, já no teste com 40% de validação foi obtido o valor de  $R_a^2$  igual a 55,04%

Em se tratando de dados de adsorção, um dos problemas encontrados foi a tentativa de ajuste do modelo por meio da análise de regressão, de início foi pensado na retirada do parâmetro  $\beta_0$  onde foi percebido que o valor para o  $R^2$  aumentava consideravelmente, embora nas bibliografias pesquisadas não foi encontrado nada que justificasse a retirada deste parâmetro, logo, foi resolvido utilizá-lo para a obtenção das duas variáveis respostas, considerar os valores obtidos na análise de regressão e comparar com os valores da RNA.

Em estudos de adsorção, são comprovados que alguns erros na análise de modelagem de dados aparecem constantemente. Em vista disso, vários autores abordam a importância da interpretação correta dos dados experimentais (BOLSTER; HORNBERGER, 2008; EL-KHAIARY; MALASH, 2011; SUBRAMANYAM; DAS, 2014; TRAN *et al.*, 2017).

Apesar de facilmente ser interpretado o  $R^2$ , a aplicação deste nem sempre é apropriada para avaliar o ajuste e a comparação de modelos (HO; CHIU; WANG, 2005; HADI *et al.*, 2012; ZHANG *et al.*, 2013).

Nas figuras 7 e 8, podem ser verificados os gráficos para os dados globais das variáveis Qt e Rem. A regressão polinomial permitiu obter uma equação que foi mais adequada para a representatividade do comportamento dos dados, permitindo assim o melhor ajuste para o modelo.

Para cada tratamento utilizado, podem ser observados abaixo as equações 10, 11, 12, 13, 14 e 15 representativa de sua função para o modelo ajustado.

$$y_a = 18,512 + 0,155*x_1 + 0,079*x_2 - 0,009*x_3 + 0,0004*x_4 - 6,57*x_5 \quad (10)$$

$$y_b = 9,018 - 5,157*x_1 + 1,490*x_2 - 0,167*x_3 - 0,008*x_4 - 0,0001*x_5 \quad (11)$$

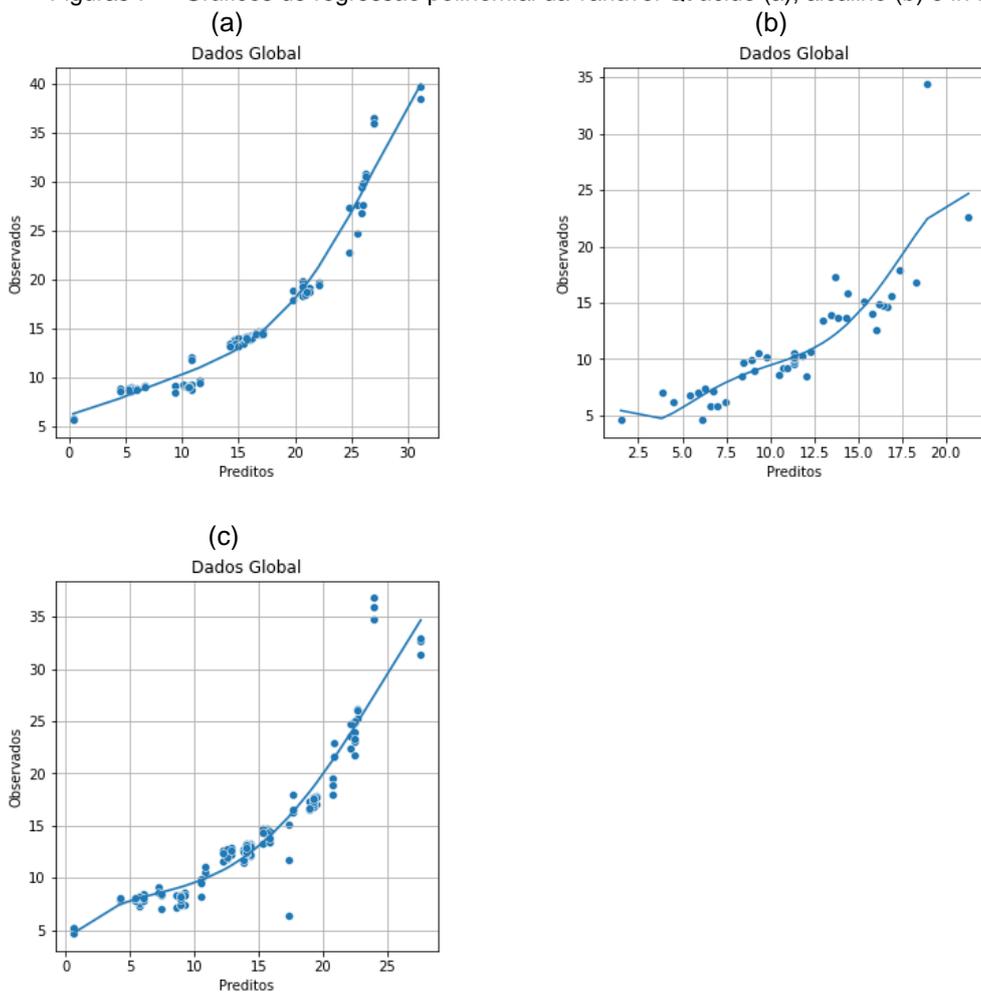
$$y_c = 15,673 - 2,060*x_1 + 0,683*x_2 - 0,069*x_3 + 0,003*x_4 - 4,498*x_5 \quad (12)$$

$$y_d = 81,041 - 17390,524*x_1 + 382,695*x_2 - 4,194*x_3 + 0,022*x_4 - 4,998*x_5 \quad (13)$$

$$y_e = 36,507 + 24,993*x_1 - 1,069*x_2 + 0,021*x_3 - 0,0001*x_4 + 6,659*x_5 \quad (14)$$

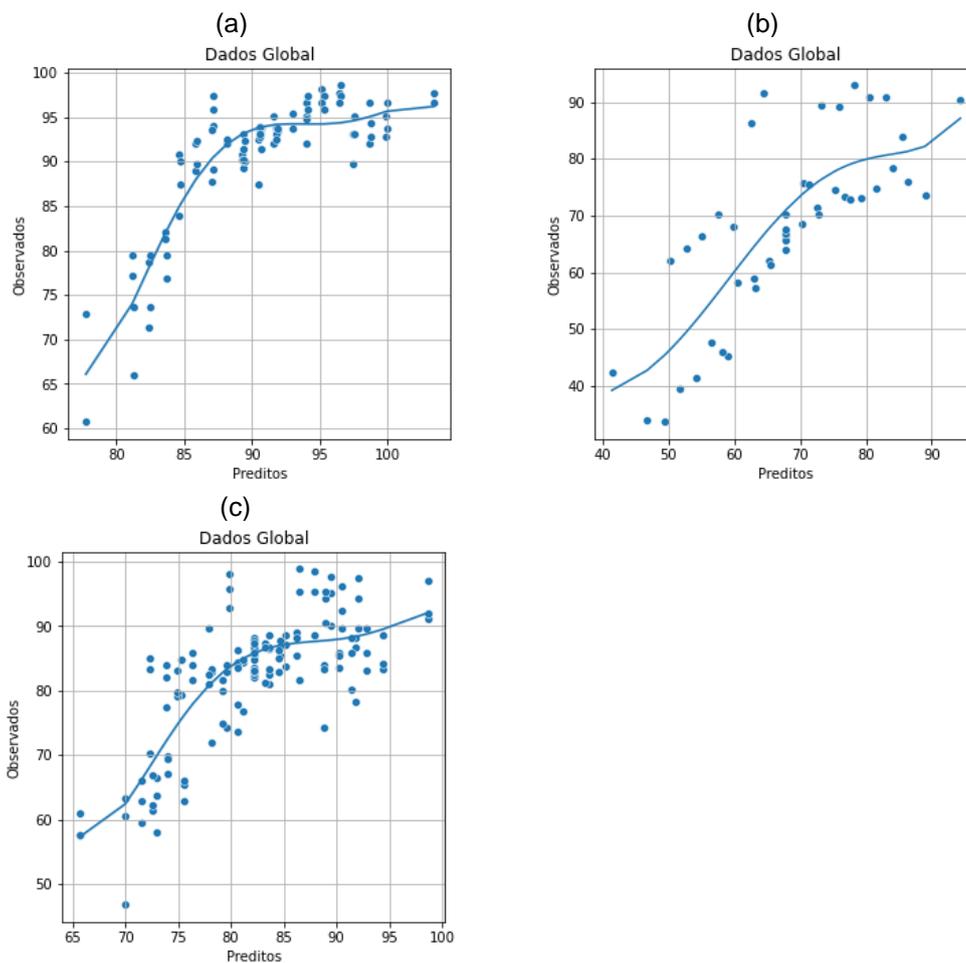
$$y_f = 73,138 - 3066,517*x_1 + 73,266*x_2 - 0,867*x_3 + 0,0005*x_4 - 1,191*x_5 \quad (15)$$

Figuras 7 – Gráficos de regressão polinomial da variável Qt ácido (a), alcalino (b) e in-natura (c)



Fonte: o autor, 2022.

Figura 8 – Gráficos de regressão polinomial da variável Rem ácido (a), alcalino (b) e in-natura (c)



Fonte: o autor, 2022.

### 5.3 Modelagem usando Redes Neurais Artificiais

#### 5.3.1 Bagaço de cana de açúcar in-natura uma camada (Rem)

Na tabela 22, são apresentados os fatores, as respostas MAE-T, MAE-V,  $R^2$  e  $R_a^2$  e o tempo de CPU utilizados no planejamento experimental da RNA.

**Tabela 22 - Planejamento do tratamento in-natura Rem**

Exp	NN	NE	FA	DIV (%)	MAE(T)	MSE(T)	$R^2$	$R_a^2$	Tempo(s)
1	30	2000	Sigmoid	30	7,2499	91,0685	0,9879	0,9878	517,72
2	30	2000	Sigmoid	40	7,3030	93,3643	0,9881	0,9881	521,24
3	30	2000	Relu	30	7,5121	88,4222	0,9874	0,9873	510,53
4	30	2000	Relu	40	8,0289	99,3416	0,9862	0,9861	514,93
5	30	4000	Sigmoid	30	5,4242	44,5288	0,9941	0,9940	1023,18
6	30	4000	Sigmoid	40	5,3310	44,0997	0,9944	0,9944	1011,39
7	30	4000	Relu	30	5,9710	58,1923	0,9919	0,9919	1013,56
8	30	4000	Relu	40	6,3962	67,5241	0,9914	0,9913	1025,04
9	30	8000	Sigmoid	30	3,5993	19,1640	0,9975	0,9975	2020,39
10	30	8000	Sigmoid	40	3,7058	19,4863	0,9972	0,9972	2016,66
11	30	8000	Relu	30	5,6615	50,5514	0,9932	0,9931	1999,26
12	30	8000	Relu	40	6,0014	57,1682	0,9929	0,9929	2011,84
13	60	2000	Sigmoid	30	7,2499	91,0685	0,9879	0,9878	471,98
14	60	2000	Sigmoid	40	7,3030	93,3643	0,9881	0,9881	477,15
15	60	2000	Relu	30	7,5121	88,4222	0,9874	0,9873	471,13
16	60	2000	Relu	40	8,0289	99,3416	0,9862	0,9861	473,90
17	60	4000	Sigmoid	30	5,1917	41,7965	0,9944	0,9944	933,37
18	60	4000	Sigmoid	40	5,3310	44,0997	0,9944	0,9944	940,50
19	60	4000	Relu	30	5,9710	58,1923	0,9919	0,9919	931,64
20	60	4000	Relu	40	6,3962	67,5241	0,9914	0,9913	936,41
21	60	8000	Sigmoid	30	3,5993	19,1640	0,9975	0,9975	1858,54
22	60	8000	Sigmoid	40	3,7058	19,4863	0,9972	0,9972	1865,23
23	60	8000	Relu	30	5,6615	50,5514	0,9932	0,9931	1852,34
24	60	8000	Relu	40	6,0014	57,1682	0,9929	0,9929	1867,43
25	120	2000	Sigmoid	30	7,3775	99,7908	0,9844	0,9842	265,44
26	120	2000	Sigmoid	40	6,2136	70,1089	0,9910	0,9909	262,57
27	120	2000	Relu	30	6,2673	62,9494	0,9909	0,9908	260,01
28	120	2000	Relu	40	7,0240	78,4636	0,9894	0,9894	262,25
29	120	4000	Sigmoid	30	4,5856	32,2519	0,9957	0,9956	516,64
30	120	4000	Sigmoid	40	4,8781	34,3093	0,9955	0,9954	521,99
31	120	4000	Relu	30	5,8575	54,8061	0,9925	0,9925	514,41
32	120	4000	Relu	40	5,8198	55,3786	0,9928	0,9928	518,25
33	120	8000	Sigmoid	30	3,4165	16,7608	0,9978	0,9978	1034,79
34	120	8000	Sigmoid	40	3,4378	16,6815	0,9976	0,9976	1037,08
35	120	8000	Relu	30	5,1598	41,4637	0,9943	0,9943	1026,67
36	120	8000	Relu	40	5,3693	44,5176	0,9942	0,9941	1035,12

Fonte: o autor, 2022.

Nas figuras 9 e 10, são apresentados os gráficos de Pareto dos seguintes níveis: MAE-T, MAE-V,  $R^2$  e  $R_a^2$ . Nos níveis, MAE-T e MAE-V, pode ser observado que os fatores épocas, a função de ativação e o número de neurônios apresentaram efeitos estatisticamente significativo para o ajuste do modelo, já o fator divisão para o

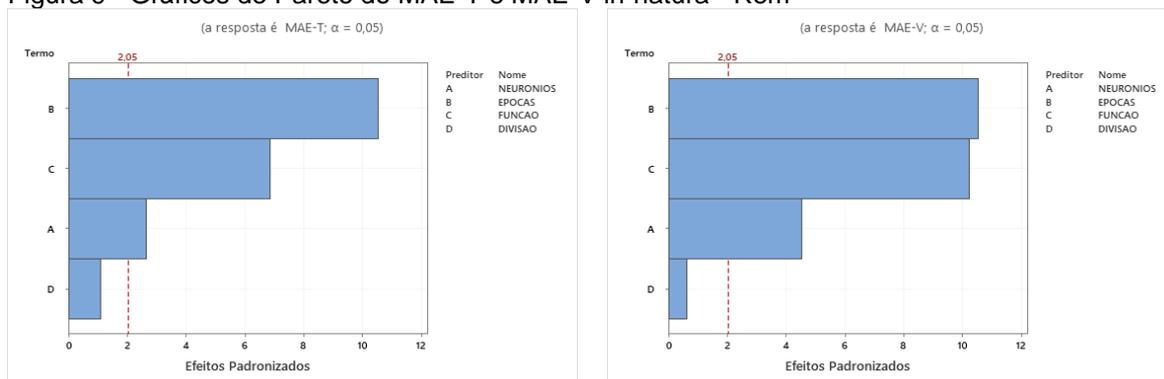
conjunto de dados não é importante. Em relação aos  $R^2$  e  $R_a^2$  as épocas e as funções de ativação apresentaram efeitos estatisticamente significativo.

O fator com maior impacto na performance da rede é o número de neurônios na camada oculta (SHAODUAN, 2005). Tal parâmetro deve ser proporcional à complexidade do problema (SHAODUAN, 2005), estando relacionada com a quantidade de dados usada para ajuste, porém, na figura 10 pode ser verificado que tanto para  $R^2$  e  $R_a^2$  o número de neurônios e a divisão dos dados não apresentaram efeito estatisticamente significativo para o ajuste do modelo.

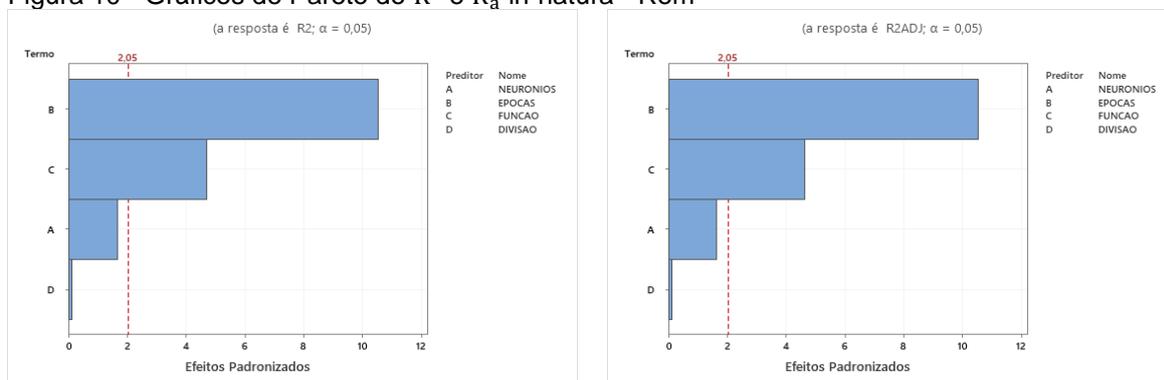
Em um bom ajuste espera-se que os valores dos parâmetros  $R^2$  das etapas de treinamento, validação, teste sejam similares e o mais próximos possível de 1. Um baixo número de neurônios leva a uma performance ruim com valores de  $R^2$  longe do valor ideal. Já um número excessivo de neurônios faz com que os pesos da rede fiquem bem ajustados aos dados de treinamento, mas não sejam capazes de prever os dados de validação e teste, levando a  $R^2$  diferentes em ordem de grandeza, com a fase de treinamento apresentando valores próximos a 1 e as demais fases apresentando valores baixos. Tal fenômeno é conhecido como overfitting (BEALE, 2010).

Uma das etapas importantes no desenvolvimento de uma rede neural diz respeito à divisão do conjunto de dados disponível (BASHEER; HAJMMER, 2000). Geralmente, os desenvolvedores de RNA dividem de forma aleatória e uniforme os dados de entrada e saída em um processo denominado de estimacão do conjunto amostral (HAYKIN, 2001). Os dados de treinamento são utilizados para estimar os parâmetros desconhecidos da rede, denominados de pesos e bias. O processo de validação consiste em uma subdivisão de dados retirados aleatoriamente do conjunto de treinamento. Esse processo de validação é utilizado como critério de parada do processo de treinamento da rede, e conseqüentemente contribui para que não ocorra um sobreajuste (overfitting). Por fim, o conjunto de teste consiste de uma amostra não pertencente ao conjunto de treinamento e validação, e sua utilização está direcionada para a verificação em relação ao desempenho da eficiência do modelo treinado conforme as circunstâncias reais de utilização da rede (ZIGGAH, *et al.* 2019). Devido a essas três categorias serem interligadas, o processo de divisão dos dados tem influência direta no desempenho da rede (ZIGGAH, *et al.* 2019).

Figura 9 - Gráficos de Pareto do MAE-T e MAE-V in-natura - Rem



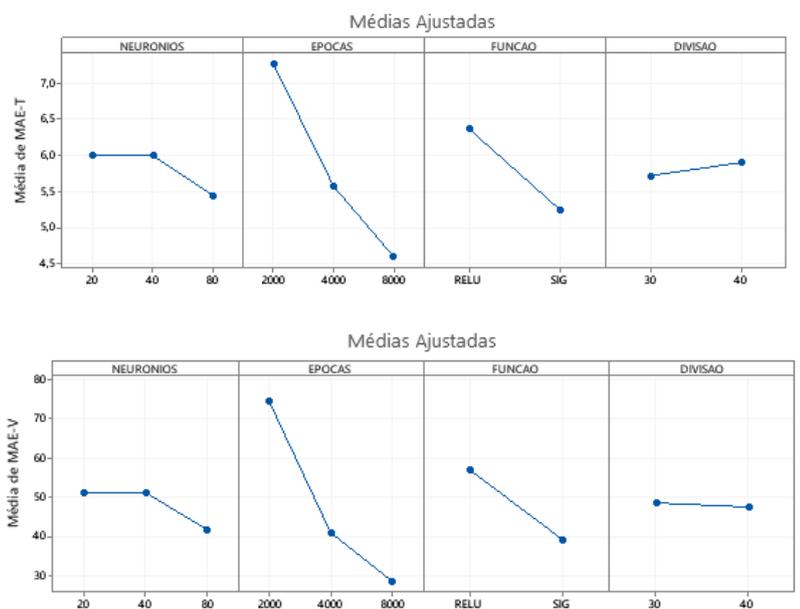
Fonte: o autor, 2022.

Figura 10 - Gráficos de Pareto do  $R^2$  e  $R_a^2$  in-natura - Rem

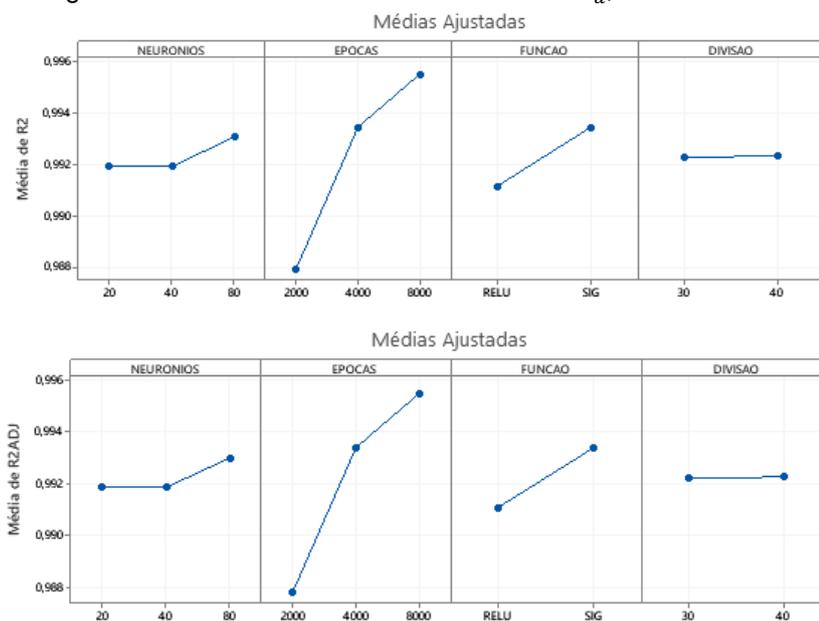
Fonte: o autor, 2022.

Ao analisar os gráficos fatoriais nas figuras 11 e 12, é possível identificar quais fatores tem maior influência na aplicação da RNA, e ainda pode ser observado que tanto para o treinamento quanto para a validação os gráficos apresentam semelhanças entre os fatores épocas e divisão. As diferenças nos fatores número de neurônios e função de ativação se deve a influência que o fator vai ocasionar na RNA. Os modelos de efeitos principais aditivos e interação multiplicativa são a melhor abordagem de análise, no entanto, na presença de dados ausentes é inadmissível sua aplicabilidade (GAUCH e ZOBEL, 1990; YAN, 2013).

Figura 11 - Gráficos fatoriais dos níveis MAE-T e MAE-V in-natura - Rem



Fonte: o autor, 2022.

Figura 12 - Gráficos fatoriais dos níveis  $R^2$  e  $R_a^2$ , in-natura - Rem

Fonte: o autor, 2022.

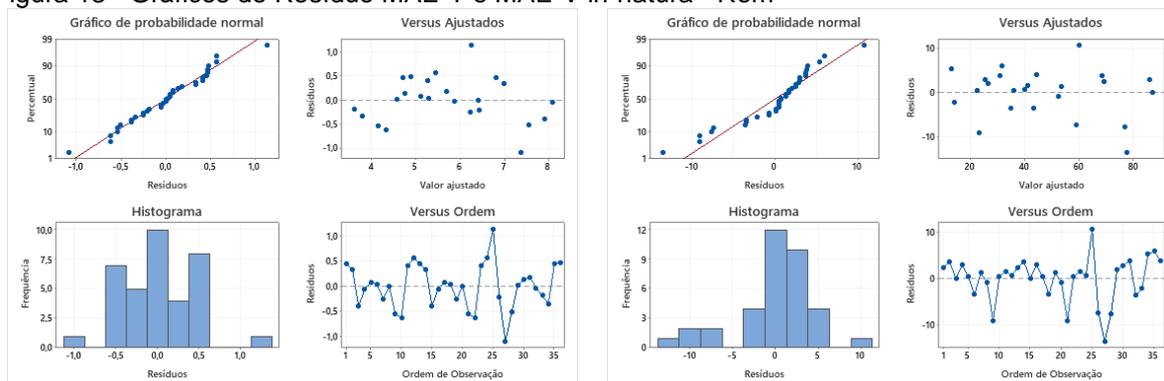
Nas figuras 13 e 14, são apresentados os gráficos de resíduos para os níveis MAE-T e MAE-V, assim também como para o  $R^2$  e o  $R_a^2$ , também é apresentado o gráfico de probabilidade normal dos resíduos, apresentando uma distribuição normal.

No gráfico de probabilidade normal (figura 13), quanto mais próximos os pontos estiverem da linha reta, mais os dados se aproximam de uma distribuição normal para representar os dados. Logo, o pressuposto de normalidade dos resíduos foi atendido.

Já na distribuição de resíduos em função dos valores ajustados (figura 12) observa-se visualmente que a pressuposta homocedasticidade (variâncias iguais) também é atendida, visto que não há padrão e os resíduos estão dispersos aleatoriamente em torno do zero (MONTGOMERY, 2009).

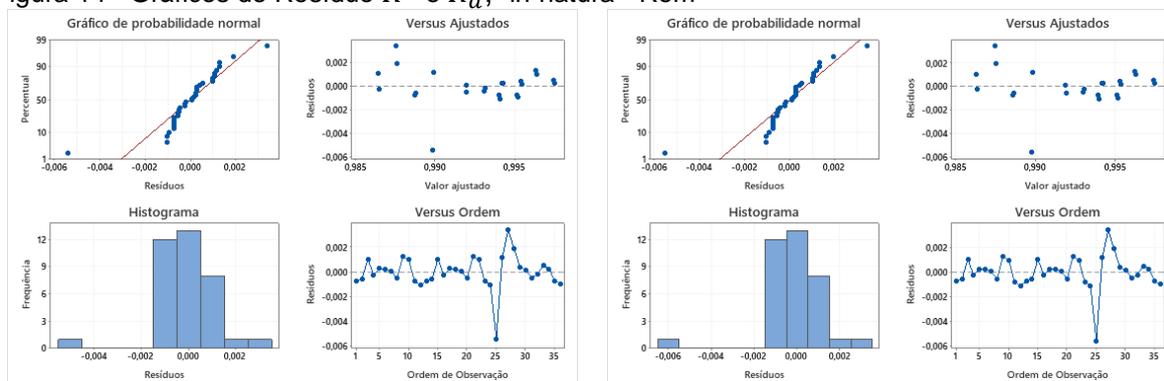
No gráfico de Resíduos versus ordem é verificada a dependência entre os resíduos.

Figura 13 - Gráficos de Resíduo MAE-T e MAE-V in-natura - Rem



Fonte: o autor, 2022.

Figura 14 - Gráficos de Resíduo  $R^2$  e  $R_a^2$ , in-natura - Rem



Fonte: o autor, 2022.

#### 5.4.1 Bagaço de cana de açúcar ácido uma camada (Rem)

Na tabela 23, são apresentados os fatores, os níveis, as respostas MAE-T, MAE-V,  $R^2$  e  $R_a^2$  e o tempo de CPU utilizados no planejamento experimental da RNA.

**Tabela 23 - Planejamento Experimental do tratamento ácido Rem**

Exp	NN	NE	FA	DIV (%)	MAE(T)	MSE(T)	$R^2$	$R_a^2$	Tempo(s)
1	30	2000	Sigmoid	30	29,8817	997,1577	0,9934	0,9933	493,62
2	30	2000	Sigmoid	40	26,5515	758,4931	0,9934	0,9933	504,57
3	30	2000	Relu	30	7,8879	98,0364	0,9884	0,9883	485,12
4	30	2000	Relu	40	7,7308	90,2023	0,9884	0,9883	489,07
5	30	4000	Sigmoid	30	5,5426	54,6926	0,9934	0,9933	953,23
6	30	4000	Sigmoid	40	5,1028	43,3110	0,9934	0,9933	966,31
7	30	4000	Relu	30	6,2872	62,6440	0,9926	0,9925	947,85
8	30	4000	Relu	40	6,0935	56,5788	0,9924	0,9924	975,62
9	30	8000	Sigmoid	30	3,7048	28,0526	0,9963	0,9963	1912,06
10	30	8000	Sigmoid	40	3,2832	20,6270	0,9964	0,9963	1921,46
11	30	8000	Relu	30	3,9479	25,2395	0,9970	0,9969	1889,09
12	30	8000	Relu	40	3,7538	20,3529	0,9969	0,9968	1933,17
13	60	2000	Sigmoid	30	7,5333	75,3245	0,9934	0,9933	246,40
14	60	2000	Sigmoid	40	6,2090	49,5644	0,9934	0,9933	256,61
15	60	2000	Relu	30	7,6466	91,6818	0,9891	0,9890	243,65
16	60	2000	Relu	40	7,4383	83,3789	0,9892	0,9891	255,35
17	60	4000	Sigmoid	30	4,8600	48,7993	0,9938	0,9937	483,48
18	60	4000	Sigmoid	40	4,7565	42,0256	0,9934	0,9934	506,52
19	60	4000	Relu	30	5,4797	47,2175	0,9944	0,9944	480,79
20	60	4000	Relu	40	4,8198	35,7338	0,9950	0,9950	505,51
21	60	8000	Sigmoid	30	2,9020	15,3396	0,9978	0,9978	960,88
22	60	8000	Sigmoid	40	2,7163	11,3294	0,9976	0,9976	1008,82
23	60	8000	Relu	30	3,6569	21,6095	0,9972	0,9972	957,12
24	60	8000	Relu	40	3,4856	16,8141	0,9972	0,9972	1003,13
25	120	2000	Sigmoid	30	5,0190	51,8563	0,9934	0,9934	247,45
26	120	2000	Sigmoid	40	4,7352	42,1173	0,9934	0,9934	259,15
27	120	2000	Relu	30	6,8525	74,5435	0,9912	0,9911	246,44
28	120	2000	Relu	40	6,8336	70,5871	0,9907	0,9906	258,45
29	120	4000	Sigmoid	30	4,1543	36,4656	0,9953	0,9952	490,94
30	120	4000	Sigmoid	40	3,7760	28,6136	0,9953	0,9953	512,47
31	120	4000	Relu	30	4,3547	29,5225	0,9965	0,9965	485,76
32	120	4000	Relu	40	3,9745	24,6025	0,9964	0,9964	504,84
33	120	8000	Sigmoid	30	2,2186	9,2095	0,9986	0,9986	1218,91
34	120	8000	Sigmoid	40	1,7750	5,1437	0,9985	0,9985	1024,82
35	120	8000	Relu	30	3,4451	19,5372	0,9975	0,9974	982,76
36	120	8000	Relu	40	3,0816	14,1456	0,9973	0,9973	982,76

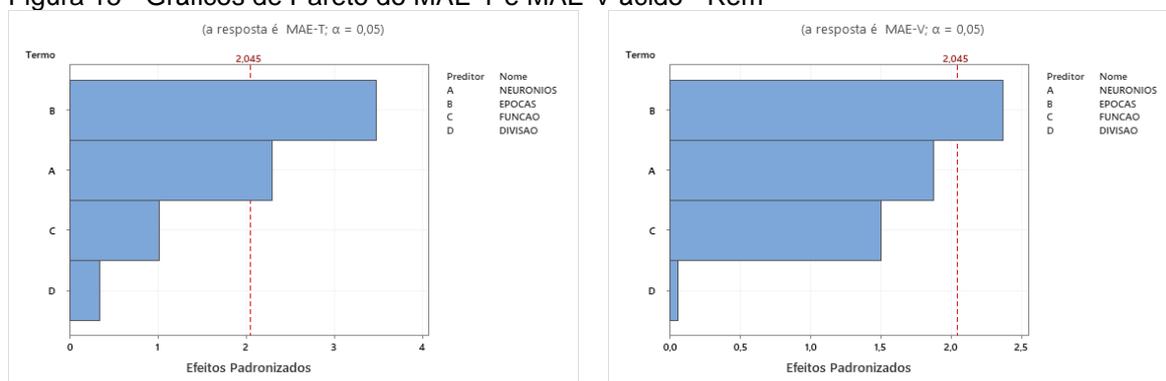
Fonte: o autor, 2022.

As figuras 15 e 16, apresentam os gráficos de Pareto para os seguintes níveis: MAE-T, MAE-V,  $R^2$  e  $R_a^2$ . Para o nível MAE-T é observado que os fatores épocas e o número de neurônio apresentaram efeitos estatisticamente significativo para o ajuste do modelo, já os fatores função de ativação e divisão para o conjunto de dados não apresentaram importância. Já para o nível MAE-V apenas o número de épocas apresentou efeito estatisticamente significativo.

Em relação aos  $R^2$  e  $R_a^2$  as épocas, os números de neurônios e as funções de ativação apresentaram efeitos estatisticamente significativo, enquanto a divisão de dados não foi importante.

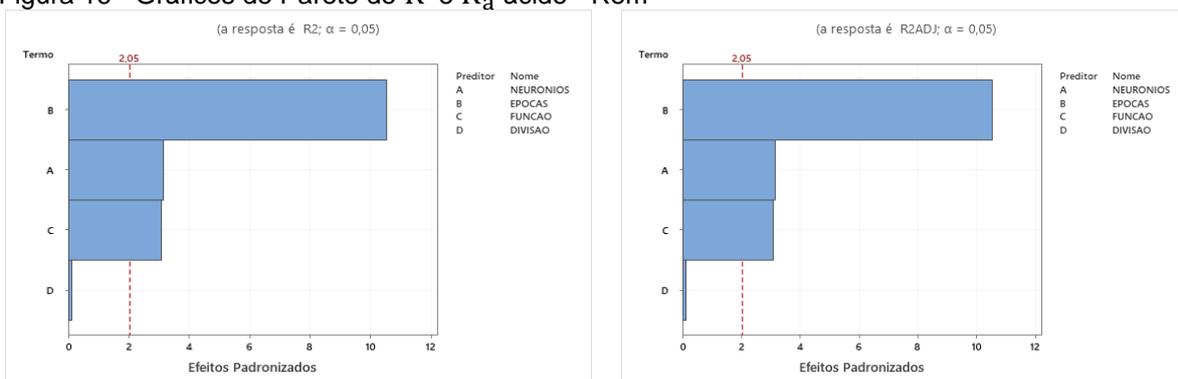
As saídas dos nós de uma camada são as entradas dos nós nas camadas seguintes, de forma que as entradas nos neurônios são combinadas linearmente e ponderadas por pesos, e o resultado é modificado por uma função não linear, chamada de função de ativação, para gerar a saída (HYNDMAN; ATHANASOPOULOS, 2018),

Figura 15 - Gráficos de Pareto do MAE-T e MAE-V ácido - Rem



Fonte: o autor, 2022.

Figura 16 - Gráficos de Pareto do  $R^2$  e  $R_a^2$  ácido - Rem



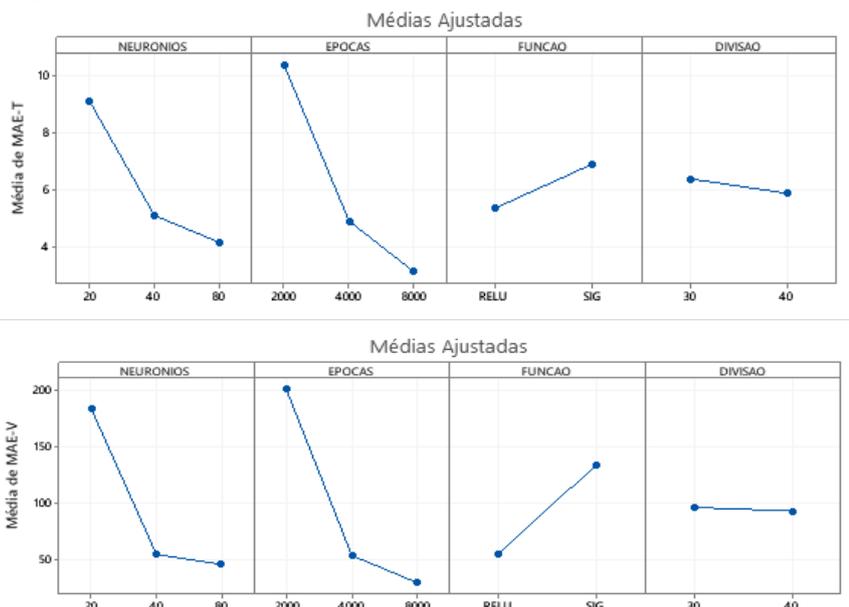
Fonte: o autor, 2022.

Ao analisar os gráficos fatoriais nas figuras 17 e 18, pode ser verificado os fatores que tem maior influência na aplicação da RNA, e ainda pode ser observado que tanto o aumento do número de épocas quanto o de neurônios permitem que seja obtido os melhores resultados para os coeficientes de determinação.

A função de ativação influencia de forma significativa a complexidade da aprendizagem e o desempenho de uma rede neural artificial, uma vez que é o

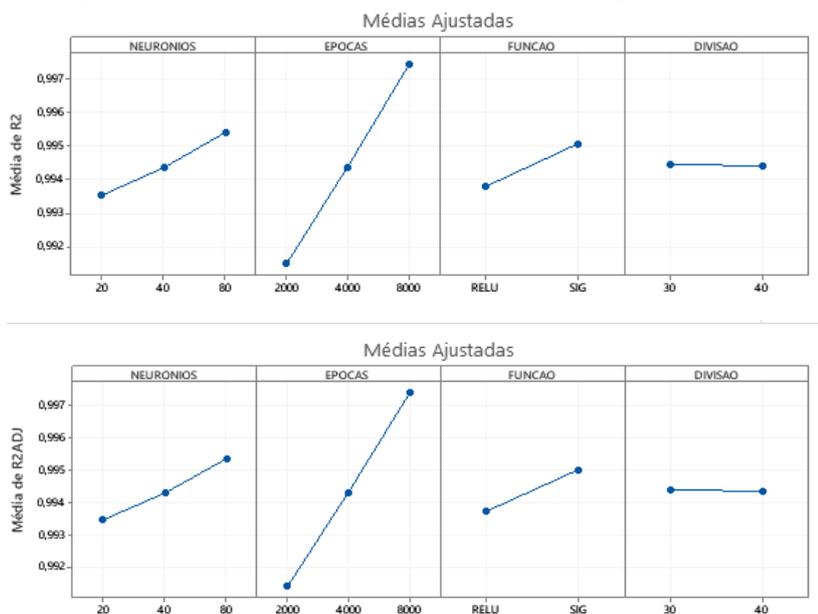
componente responsável por introduzir não-linearidade na RNA (LECUN *et al.*, 2015; MOHAMMED *et al.*, 2019).

Figura 17 - Gráficos fatoriais dos níveis MAE-T e MAE-V ácido - Rem



Fonte: o autor, 2022.

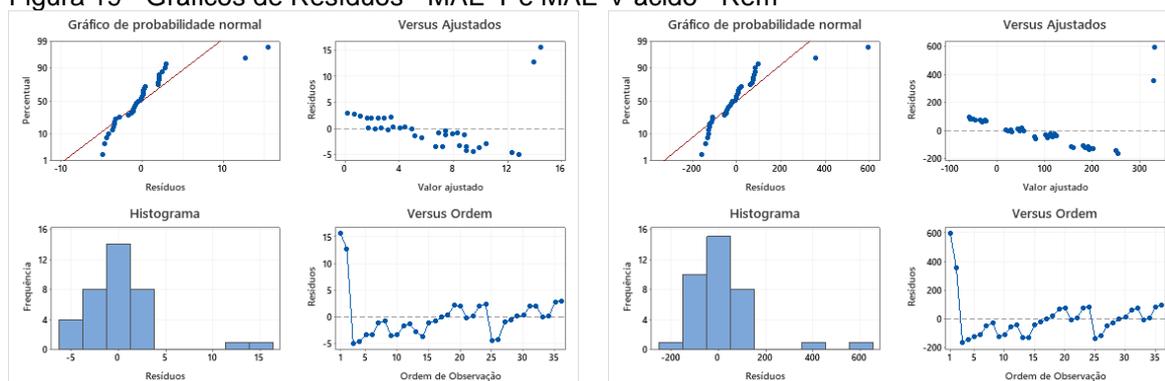
Figura 18 - Gráficos fatoriais dos níveis  $R^2$  e  $R_a^2$  ácido - Rem



Fonte: o autor, 2022.

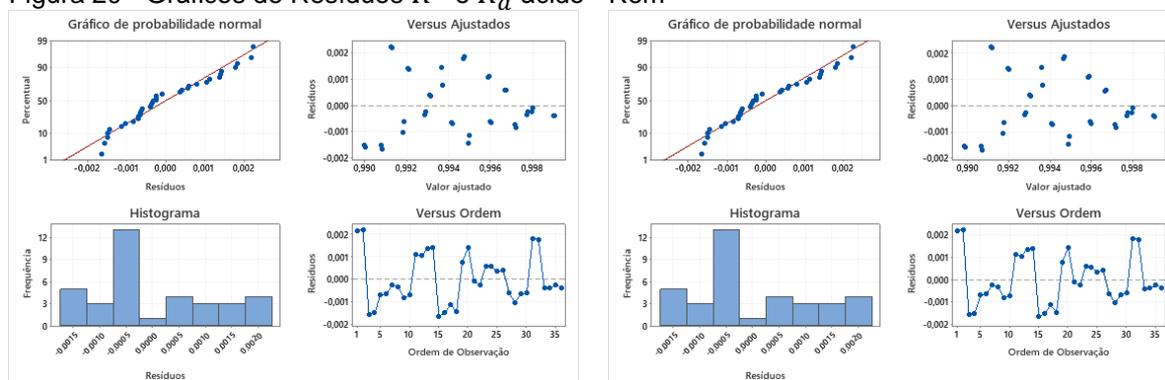
Nas figuras 19 e 20, são apresentados os gráficos de Resíduos para os níveis MAE-T, MAE-V,  $R^2$  e o  $R_a^2$ .

Figura 19 - Gráficos de Resíduos - MAE-T e MAE-V ácido - Rem



Fonte: o autor, 2022.

Figura 20 - Gráficos de Resíduos  $R^2$  e  $R_a^2$  ácido - Rem



Fonte: o autor, 2022.

No gráfico de probabilidade normal (figura 20) para o  $R^2$  e o  $R_a^2$  quanto mais próximos os pontos estiverem da linha reta, mais os dados se aproximam de uma distribuição normal. Logo, o pressuposto de normalidade dos resíduos foi atendido. Já na distribuição de resíduos em função dos valores ajustados (figura 20) observa-se visualmente que a pressuposta homocedasticidade (variâncias iguais) também é atendida, visto que não há padrão e os resíduos estão dispersos aleatoriamente em torno do zero (MONTGOMERY, 2009).

### 5.5.1 Bagaço de cana de açúcar alcalino uma camada (Rem).

Na tabela 24, são apresentados os fatores, os níveis, as respostas MAE-T, MAE-V,  $R^2$  e  $R_a^2$  e o tempo de CPU utilizados no planejamento experimental da RNA.

**Tabela 24 - Planejamento do tratamento alcalino Rem**

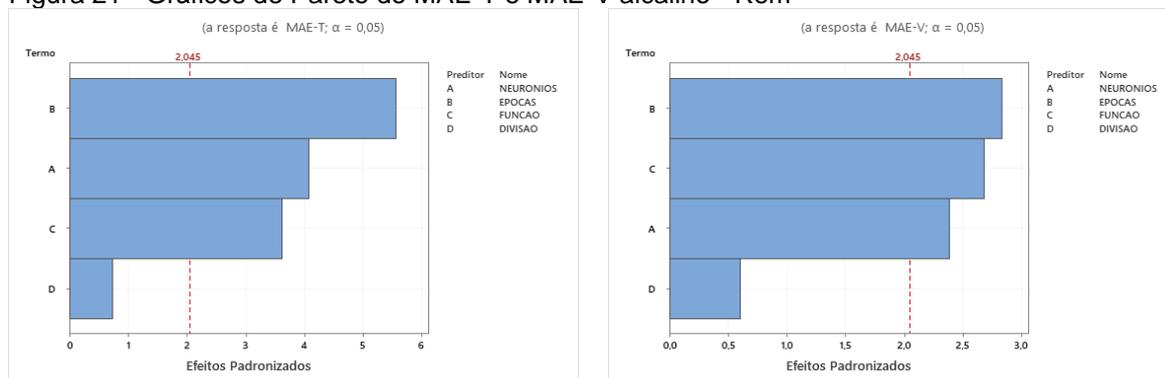
Exp	NN	NE	FA	DIV (%)	MAE(T)	MSE(T)	$R^2$	$R_a^2$	Tempo(s)
1	30	2000	Sigmoid	30	34,9271	1507,795	0,9501	0,9490	495,24
2	30	2000	Sigmoid	40	34,6205	1434,828	0,9501	0,9490	487,31
3	30	2000	Relu	30	12,2198	250,1861	0,9552	0,9542	470,85
4	30	2000	Relu	40	11,0920	206,8485	0,9553	0,9543	471,38
5	30	4000	Sigmoid	30	20,1871	526,7319	0,9501	0,9490	953,45
6	30	4000	Sigmoid	40	15,1802	337,8749	0,9505	0,9494	938,13
7	30	4000	Relu	30	10,8325	182,4701	0,9664	0,9656	933,01
8	30	4000	Relu	40	9,9944	149,8449	0,9661	0,9653	935,36
9	30	8000	Sigmoid	30	10,1767	169,9926	0,9672	0,9665	1864,59
10	30	8000	Sigmoid	40	8,9481	126,9807	0,9664	0,9657	1894,78
11	30	8000	Relu	30	6,2923	61,6283	0,9868	0,9865	1906,81
12	30	8000	Relu	40	7,1232	77,7775	0,9831	0,9827	1883,12
13	60	2000	Sigmoid	30	18,5576	454,4661	0,9501	0,9490	239,13
14	60	2000	Sigmoid	40	16,2606	372,7625	0,9503	0,9492	239,30
15	60	2000	Relu	30	11,6737	217,4866	0,9604	0,9595	238,75
16	60	2000	Relu	40	10,5530	175,8123	0,9603	0,9594	237,50
17	60	4000	Sigmoid	30	11,4393	202,5419	0,9626	0,9617	468,19
18	60	4000	Sigmoid	40	10,2799	174,6752	0,9592	0,9583	466,07
19	60	4000	Relu	30	9,5116	141,5531	0,9736	0,9730	463,53
20	60	4000	Relu	40	9,0690	126,4617	0,9709	0,9702	464,46
21	60	8000	Sigmoid	30	6,9008	85,1966	0,9795	0,9790	932,76
22	60	8000	Sigmoid	40	5,7747	66,1871	0,9794	0,9789	927,38
23	60	8000	Relu	30	4,8548	38,7715	0,9896	0,9894	929,22
24	60	8000	Relu	40	4,2932	30,0660	0,9903	0,9901	928,89
25	120	2000	Sigmoid	30	12,4998	244,9157	0,9553	0,9543	301,78
26	120	2000	Sigmoid	40	10,7587	192,1905	0,9571	0,9561	301,81
27	120	2000	Relu	30	10,9765	185,3746	0,9657	0,9650	267,17
28	120	2000	Relu	40	10,1441	156,3770	0,9641	0,9633	232,59
29	120	4000	Sigmoid	30	8,5780	128,1081	0,9741	0,9736	455,77
30	120	4000	Sigmoid	40	7,6434	101,0012	0,9720	0,9713	454,90
31	120	4000	Relu	30	6,5569	64,5593	0,9872	0,9869	453,89
32	120	4000	Relu	40	7,22588	83,2583	0,9807	0,9803	457,41
33	120	8000	Sigmoid	30	4,04161	32,1168	0,9878	0,9875	899,99
34	120	8000	Sigmoid	40	3,19397	20,9690	0,9904	0,9902	1066,96
35	120	8000	Relu	30	3,72291	24,1183	0,9906	0,9904	925,39
36	120	8000	Relu	40	3,56826	22,5691	0,9903	0,9901	913,77

Fonte: o autor, 2022.

Nas figuras 21 e 22, são apresentados os gráficos de Pareto para os seguintes níveis: MAE-T, MAE-V,  $R^2$  e  $R_a^2$ .

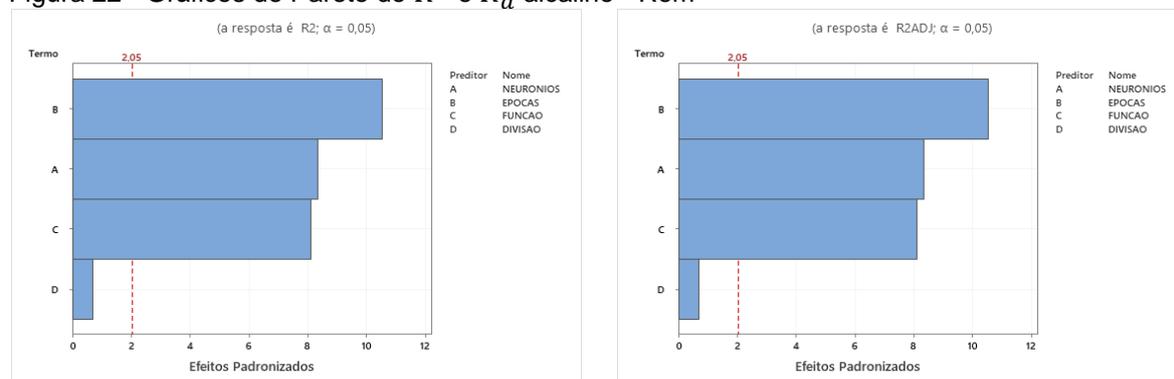
Nos níveis, MAE-T, MAE-V,  $R^2$  e  $R_a^2$  é observado que os fatores épocas, a função de ativação e o número de neurônio apresentaram efeitos estatisticamente significativo para o ajuste do modelo, já o fator divisão para o conjunto de dados não é importante.

Figura 21 - Gráficos de Pareto do MAE-T e MAE-V alcalino - Rem



Fonte: o autor, 2022.

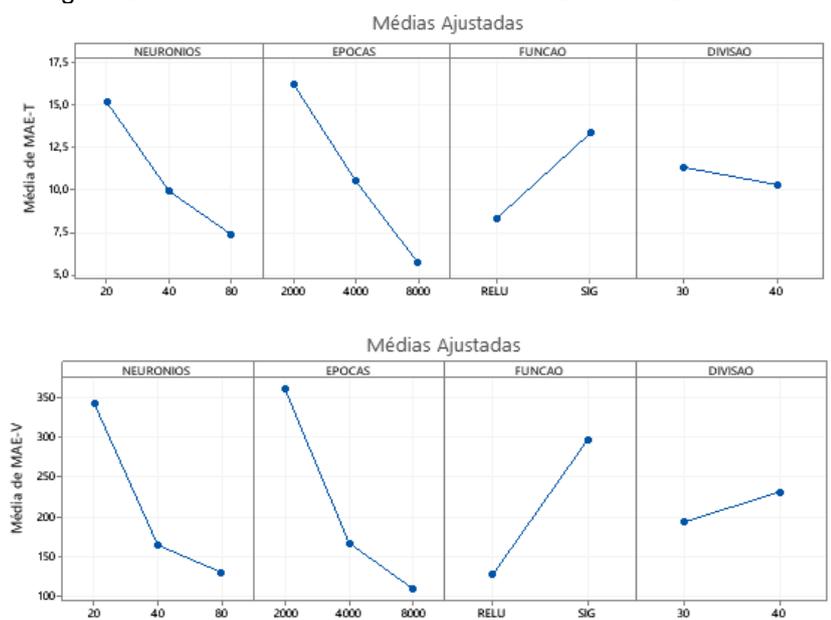
Figura 22 - Gráficos de Pareto do  $R^2$  e  $R_a^2$  alcalino - Rem



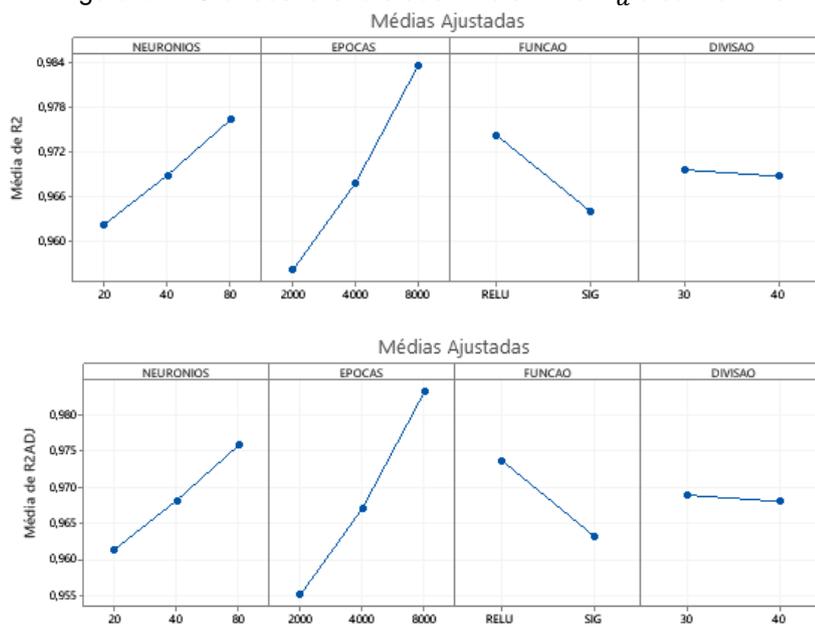
Fonte: o autor, 2022.

Os gráficos fatoriais nas figuras 23 e 24, exibem quais fatores tem maior efeito estatístico significativo na aplicação da RNA. Pode ser observado que tanto para o número de neurônios quanto o número de épocas, à medida que estes números foram aumentados foi obtido o maior valor para o coeficiente de determinação.

Figura 23 - Gráficos fatoriais dos níveis MAE-T e MAE-V alcalino - Rem



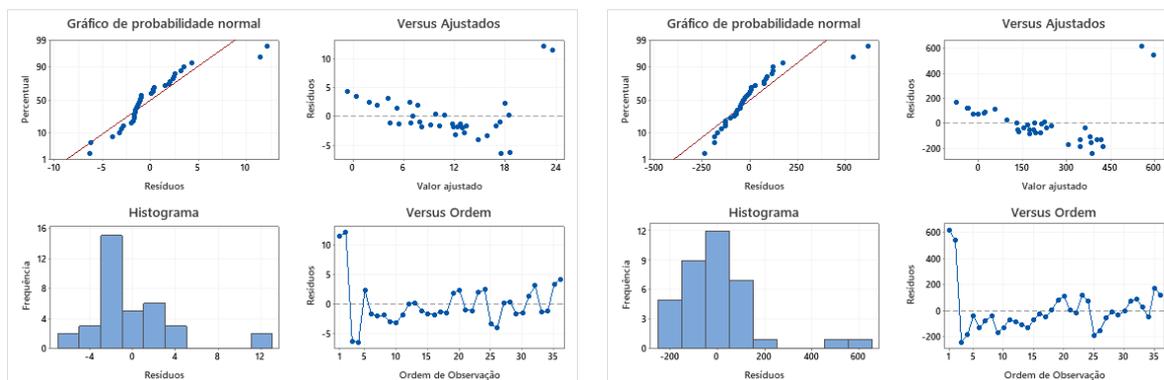
Fonte: o autor, 2022.

Figura 24 - Gráficos fatoriais dos níveis  $R^2$  e  $R_a^2$  alcalino - Rem

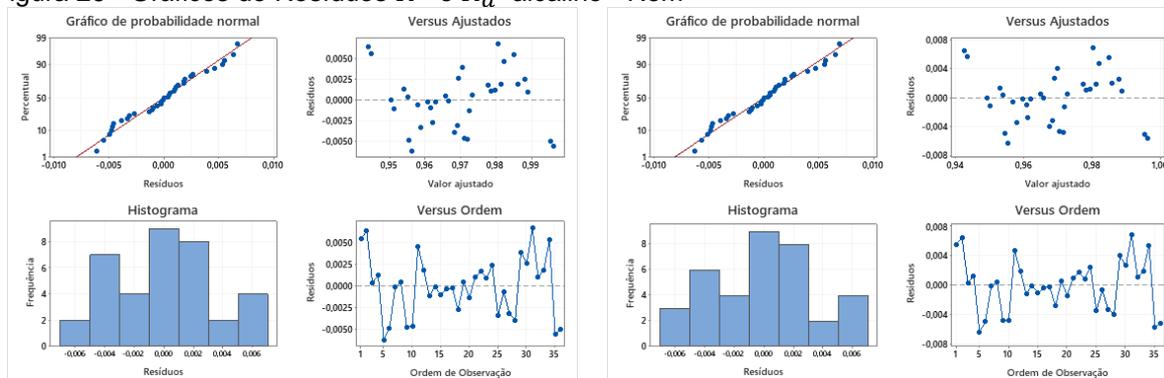
Fonte: o autor, 2022.

Nas figuras 25 e 26, são apresentados os gráficos de Resíduos para os níveis MAE-T e MAE-V, assim também como para o  $R^2$  e o  $R_a^2$ .

Figura 25 - Gráficos de Resíduos MAE-T e MAE-V alcalino - Rem



Fonte: o autor, 2022.

Figura 26 - Gráficos de Resíduos  $R^2$  e  $R_a^2$  alcalino - Rem

Fonte: o autor, 2022.

No gráfico de probabilidade normal (figuras 25 e 26), são observados que os pontos estão próximos a linha reta, indicando que os dados se aproximam de uma distribuição normal. Logo, pela análise dos resíduos é possível identificar que os dados se aproximam de uma distribuição normal. Para a distribuição de resíduos em função dos valores ajustados (figuras 25 e 26) é observado que a pressuposição de homocedasticidade (variâncias iguais) também é atendida, visto que não há padrão e os resíduos estão dispersos aleatoriamente em torno do zero (MONTGOMERY, 2009).

Para o MAE-T e o MAE-V, o gráfico de probabilidade normal pressupõe que os valores dos resíduos se aproximam de uma distribuição normal, embora, a distribuição apresente valores de outliers. Para o gráfico de resíduos versus ajustados pode ser verificado a pressuposição que os resíduos são aleatoriamente distribuídos e com variância constante, e ainda pode ser verificado a presença de outliers. No histograma, é possível verificar que os dados não seguem distribuição normal. No gráfico de resíduos versus ordem verifica-se que há independência entre os resíduos.

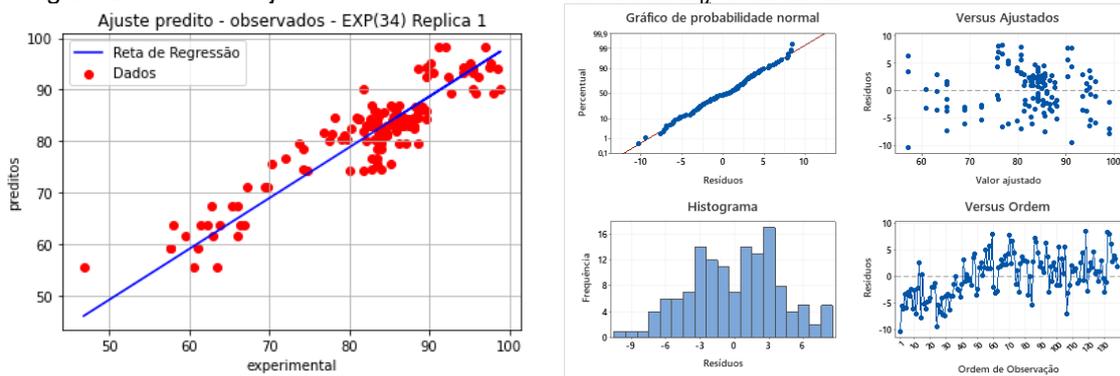
No caso do  $R^2$  e  $R_a^2$  os requisitos de normalidade dos resíduos e homocedasticidade das amostras foram atendidos como observado na figura 26.

### 5.6.1 Otimização da RNA com uma camada (Rem)

A otimização é uma parte inerente do treinamento de redes neurais e outras máquinas de aprendizado, dado que este treinamento consiste da determinação dos parâmetros internos da máquina (pesos sinápticos), através da minimização de uma função custo (BRUNTON *et al.*, 2020)

Na figura 27, pode ser observado que o experimento 34 aponta o valor otimizado de  $R_a^2$  e MAE-V para a RNA com uma camada, 80 neurônios, 8.000 épocas, função de ativação: Sigmoid e 40% para validação dos dados com um tempo de CPU igual a 411,16 segundos para o tratamento inatura da variável Rem.

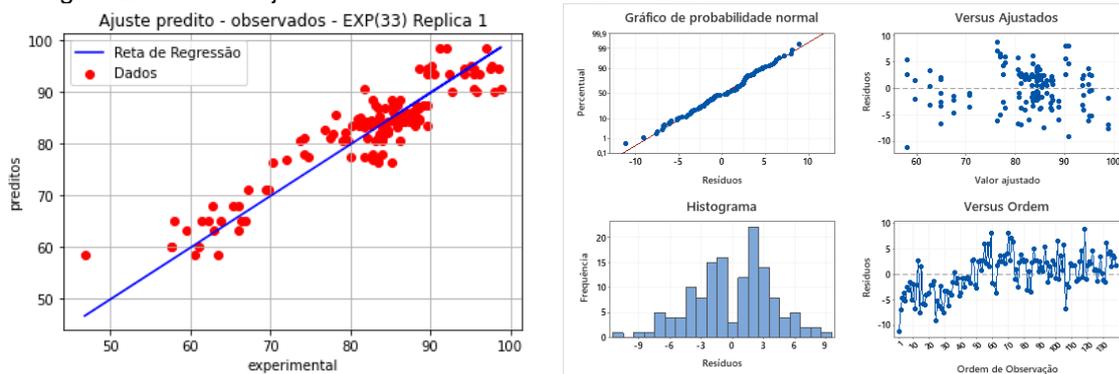
Figura 27 - Gráfico ajuste da RNA e resíduo dos níveis  $R_a^2$  e MAE-V.



Fonte: o autor, 2022.

Na figura 28, pode ser observado que no experimento 33 temos o gráfico otimizado do MAE-T para a RNA com uma camada, 80 neurônios, 8.000 épocas, função de ativação: Sigmoid e 30% para validação dos dados com um tempo de CPU igual a 419, 20 segundos para o tratamento inatura.

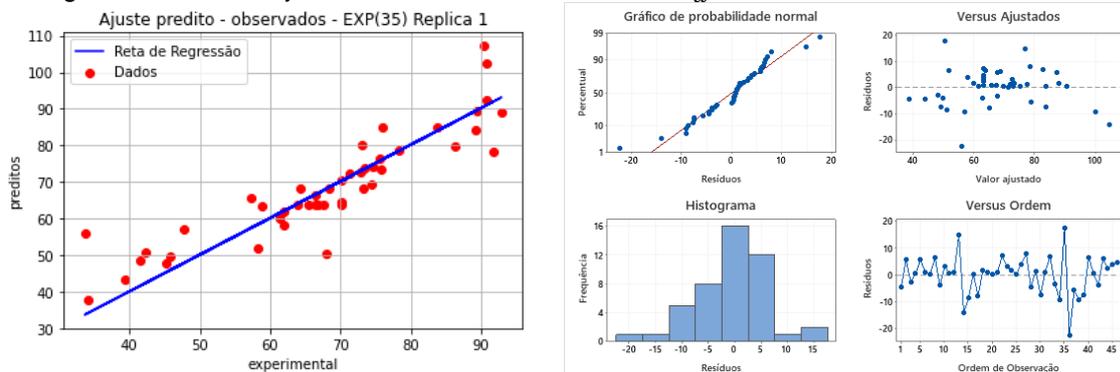
Figura 28 - Gráfico ajuste da RNA e resíduo do nível MAE-T.



Fonte: o autor, 2022.

Na figura 29, pode ser verificado o experimento 35, onde é observada a resposta otimizada do  $R_a^2$  e do MAE-V para a RNA com uma camada, 80 neurônios, 8.000 épocas, função de ativação: Relu e 30% para validação dos dados com um tempo de CPU igual a 392,46 segundos para o tratamento alcalino.

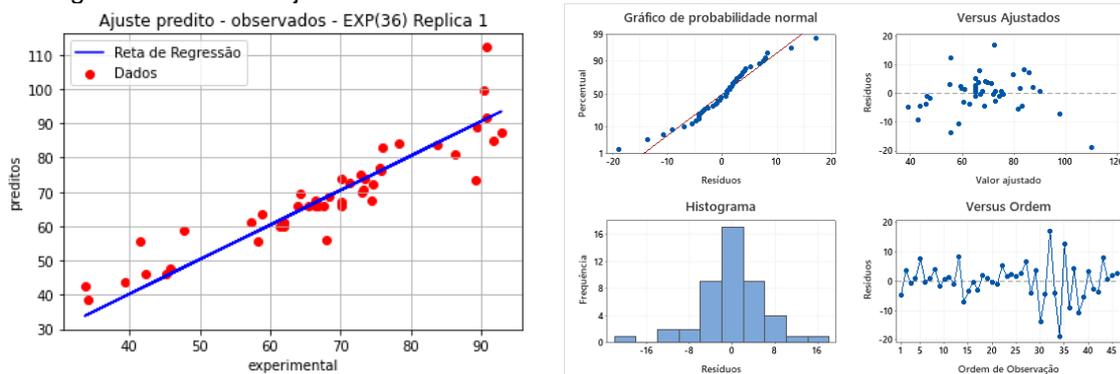
Figura 29 - Gráfico ajuste da RNA e resíduo do nível  $R_a^2$  e MAE-V.



Fonte: o autor, 2022.

A figura 30, apresenta o experimento 36 que aponta a otimização do MAE-T para a RNA com uma camada, 80 neurônios, 8.000 épocas, função de ativação: Relu e 40% para validação dos dados com um tempo de CPU igual a 386,33 segundos.

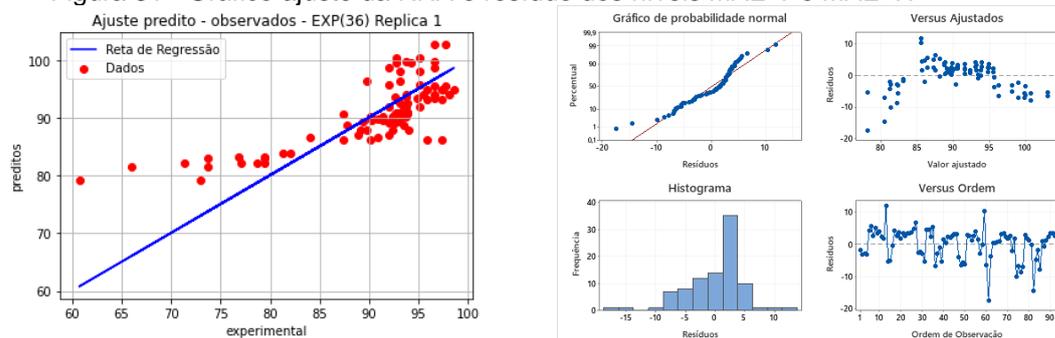
Figura 30 - Gráfico ajuste da RNA e resíduo do nível MAE-T.



Fonte: o autor, 2022.

Na figura 31 pode ser observado que no experimento 34 é verificado a otimização do MAE-T e do MAE-V para a RNA com uma camada, 80 neurônios, 8.000 épocas, função de ativação: Sigmoid e 40% para validação dos dados com um tempo de CPU igual a 421,69 segundos para o tratamento ácido.

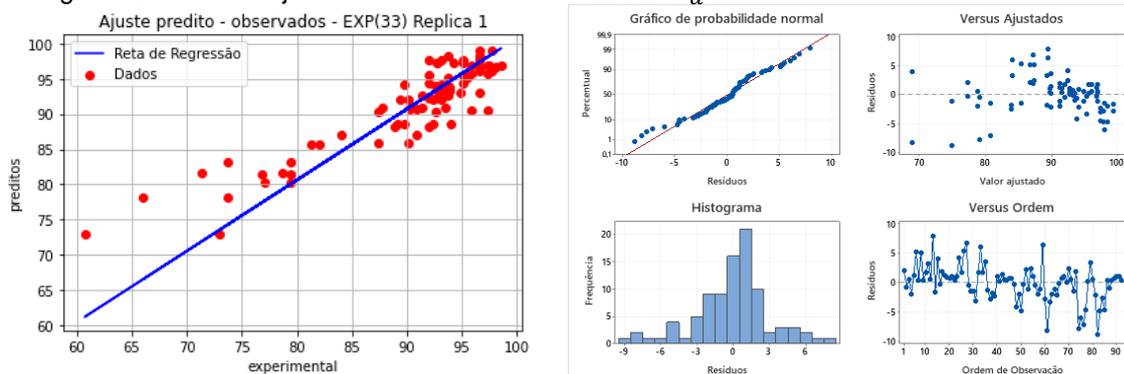
Figura 31 - Gráfico ajuste da RNA e resíduo dos níveis MAE-T e MAE-V.



Fonte: o autor, 2022.

Na figura 32, pode ser verificado que o experimento 33 aponta o valor otimizado do  $R_a^2$  para a RNA com uma camada, 80 neurônios, 8.000 épocas, função de ativação: Sigmoid e 30% para validação dos dados com um tempo de CPU igual a 449,48 segundos.

Figura 32 - Gráfico ajuste da RNA e resíduo do nível  $R_a^2$ .



Fonte: o autor, 2022.

Para a otimização da variável Rem, pode ser observado que o gráfico de probabilidade normal indica que os três tratamentos apresentaram aproximação a uma distribuição normal, logo, o pressuposto de normalidade dos resíduos foi atendido. Ainda nos três tratamentos, pode ser verificado em diversos pontos consecutivos acima e abaixo da linha central, indicando um comportamento sistêmico e comprovando a dependência dos resíduos.

### 5.7.1 Bagaço de cana de açúcar in-natura uma camada (Qt)

Na tabela 25, são apresentados os fatores, os níveis e as respostas MAE-T, MAE-V,  $R^2$  e  $R_a^2$  e o tempo de CPU utilizados no planejamento experimental da RNA.

**Tabela 25 - Planejamento Experimental do tratamento in-natura Qt**

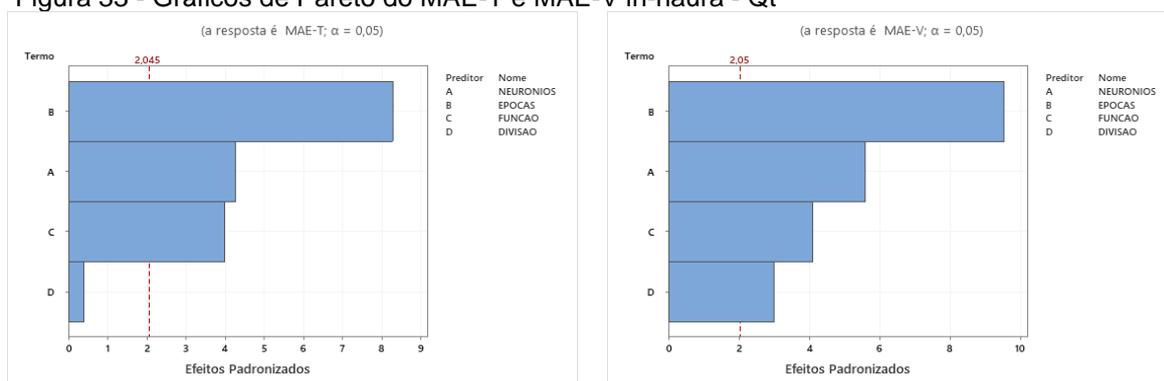
Exp	NN	NE	FA	DIV (%)	MAE(T)	MSE(T)	$R^2$	$R_a^2$	Tempo(s)
1	30	2000	Sigmoid	30	0,9572	0,9562	0,9577	0,9574	257,50
2	30	2000	Sigmoid	40	0,8983	0,8960	0,9530	0,9527	258,62
3	30	2000	Relu	30	0,9284	0,9269	0,9552	0,9549	257,15
4	30	2000	Relu	40	0,9324	0,9309	0,9567	0,9564	257,80
5	30	4000	Sigmoid	30	0,9461	0,9449	0,9873	0,9873	507,47
6	30	4000	Sigmoid	40	0,9420	0,9407	0,9879	0,9878	516,93
7	30	4000	Relu	30	0,9580	0,9570	0,9843	0,9842	507,79
8	30	4000	Relu	40	0,9611	0,9602	0,9720	0,9718	509,94
9	30	8000	Sigmoid	30	0,9885	0,9883	0,9970	0,9969	1013,96
10	30	8000	Sigmoid	40	0,9703	0,9697	0,9761	0,9759	1016,93
11	30	8000	Relu	30	0,9765	0,9760	0,9740	0,9738	1011,47
12	30	8000	Relu	40	0,0100	0,0102	0,9697	0,9695	1016,19
13	60	2000	Sigmoid	30	0,9303	0,9288	0,9761	0,9759	255,95
14	60	2000	Sigmoid	40	0,9319	0,9304	0,9740	0,9738	257,66
15	60	2000	Relu	30	0,9385	0,9371	0,9697	0,9695	256,07
16	60	2000	Relu	40	0,9505	0,9494	0,9697	0,9695	257,69
17	60	4000	Sigmoid	30	0,9737	0,9732	0,9922	0,9922	510,13
18	60	4000	Sigmoid	40	0,9771	0,9766	0,9901	0,9900	511,36
19	60	4000	Relu	30	0,9666	0,9659	0,9761	0,9759	507,99
20	60	4000	Relu	40	0,9563	0,9553	0,9803	0,9801	510,54
21	60	8000	Sigmoid	30	0,9790	0,9785	0,9974	0,9974	1013,96
22	60	8000	Sigmoid	40	0,9808	0,9804	0,9966	0,9966	1020,11
23	60	8000	Relu	30	0,9779	0,9774	0,9857	0,9856	1013,81
24	60	8000	Relu	40	0,9640	0,9632	0,9948	0,9948	1019,46
25	120	2000	Sigmoid	30	0,9591	0,9582	0,9803	0,9802	257,88
26	120	2000	Sigmoid	40	0,9617	0,9608	0,9796	0,9794	259,22
27	120	2000	Relu	30	0,9473	0,9461	0,9720	0,9718	256,84
28	120	2000	Relu	40	0,9524	0,9514	0,9726	0,9724	260,05
29	120	4000	Sigmoid	30	0,9804	0,9800	0,9924	0,9923	511,08
30	120	4000	Sigmoid	40	0,9733	0,9727	0,9909	0,9909	513,28
31	120	4000	Relu	30	0,9745	0,9740	0,9907	0,9906	509,12
32	120	4000	Relu	40	0,9641	0,9633	0,9802	0,9801	512,81
33	120	8000	Sigmoid	30	0,9850	0,9847	0,9975	0,9975	1015,03
34	120	8000	Sigmoid	40	0,9728	0,9722	0,9968	0,9967	1019,26
35	120	8000	Relu	30	0,9621	0,9613	0,9950	0,9950	1013,23
36	120	8000	Relu	40	0,9631	0,9623	0,9957	0,9957	1019,26

Fonte: o autor, 2022.

Nas figuras 31 e 32, são apresentados os gráficos de Pareto para os seguintes níveis: MAE-T, MAE-V,  $R^2$  e  $R_a^2$ . Para o MAE-T, pode ser observado que os fatores épocas, a função de ativação e o número de neurônio apresentaram efeitos estatisticamente significativo, já o fator divisão para o conjunto de dados não apresentou efeito estatisticamente significativo para ajuste do modelo. Já para o MAE-

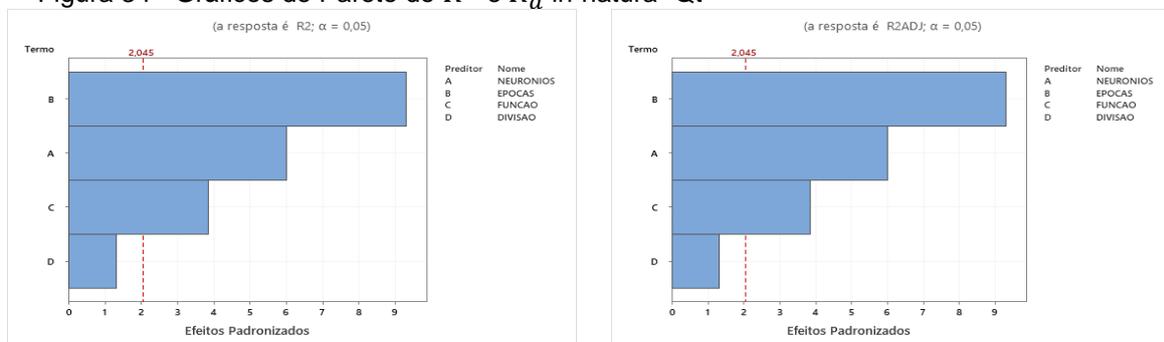
V, os fatores épocas, a função de ativação, o número de neurônio e a divisão dos dados apresentaram efeitos estatisticamente significativo. Em relação aos  $R^2$  e  $R_a^2$  as épocas, o número de neurônios e as funções de ativação apresentaram efeitos estatisticamente significativo. A divisão dos dados, não apresentou efeito estatístico significativo no ajuste do modelo. Ainda pode ser identificado, que o aumento do número de neurônios e de épocas resulta nos maiores coeficientes de determinação.

Figura 33 - Gráficos de Pareto do MAE-T e MAE-V in-natura - Qt



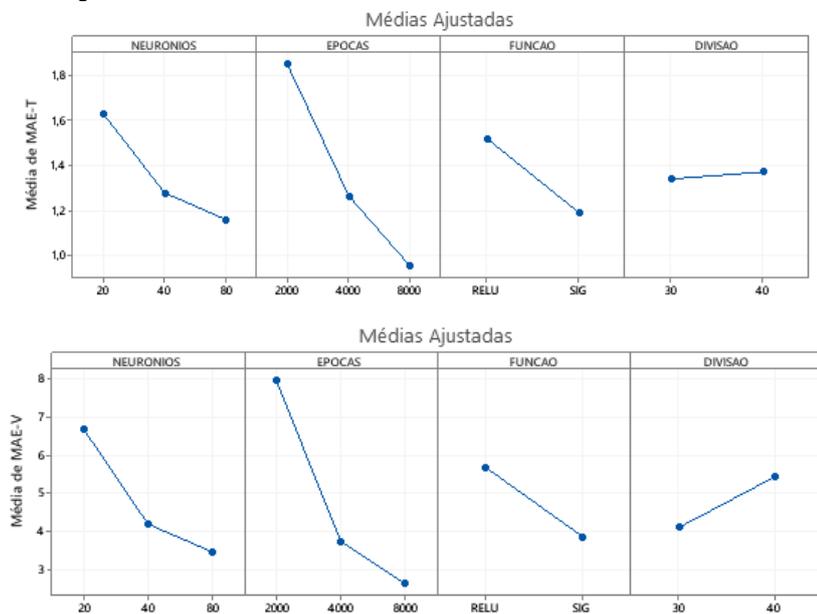
Fonte: o autor, 2022.

Figura 34 - Gráficos de Pareto do  $R^2$  e  $R_a^2$  in-natura -Qt

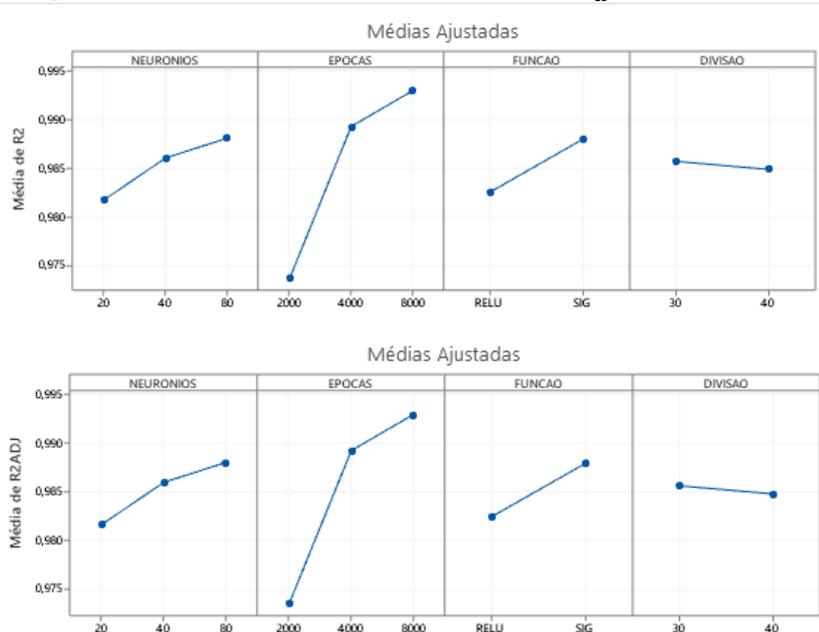


Fonte: o autor, 2022.

Figura 35 - Gráficos Fatoriais dos níveis MAE-T e MAE-V in-natura - Qt



Fonte: o autor, 2022.

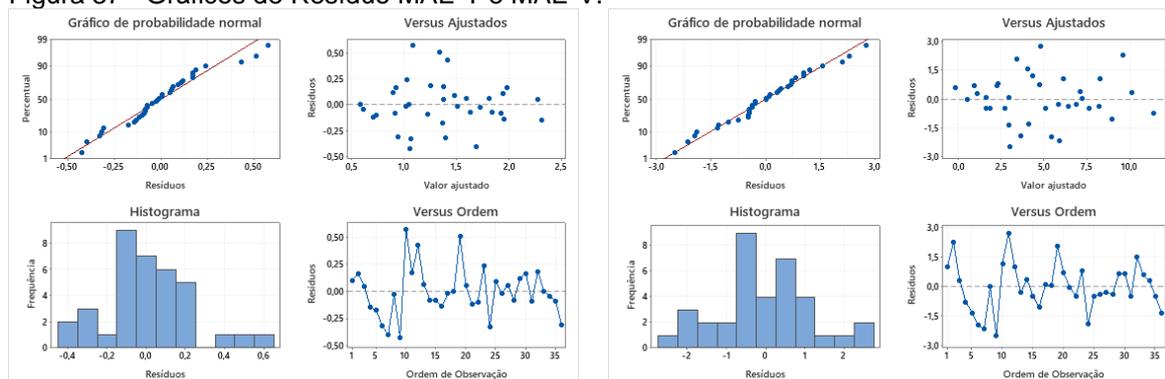
Figura 36 - Gráficos Fatoriais dos níveis  $R^2$  e  $R_a^2$  in-natura -Qt

Fonte: o autor, 2022.

Nas figuras 37 e 38, são apresentados os gráficos de resíduos para os níveis MAE-T, MAE-V,  $R^2$  e  $R_a^2$ , também pode ser observado que no gráfico de probabilidade normal nas duas figuras, seguem distribuição normal (MONTGOMERY, 2009). Logo, o pressuposto de normalidade dos resíduos foi atendido. Já na distribuição de resíduos em função dos valores ajustados, é verificado visualmente que a pressuposta

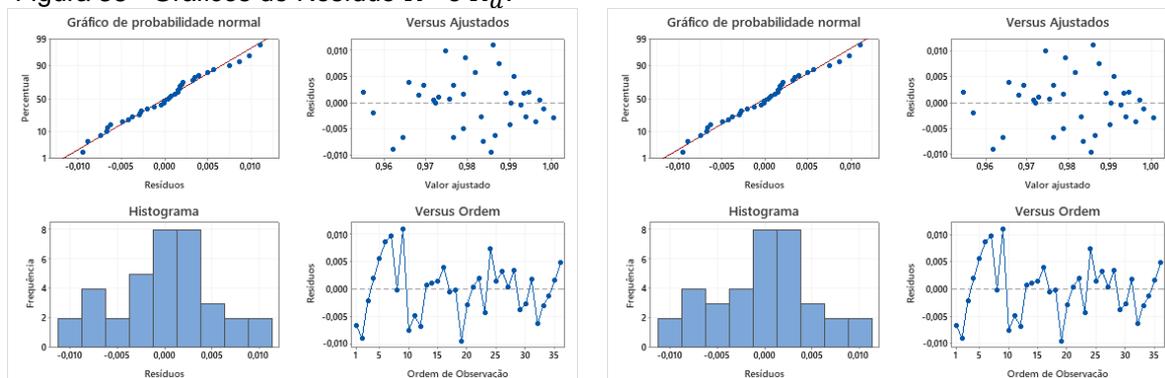
homocedasticidade (variâncias iguais) também é atendida, visto que não há padrão e os resíduos estão dispersos aleatoriamente em torno do zero. O que se espera de um modelo em relação ao grau de ajuste é valor de  $R^2$  seja o mais próximo possível de 1, que o valor de os desvios seja o mais próximo possível de zero e que os resíduos apresentem distribuição normal em torno de zero (CAPP & NIENOV, 2020).

Figura 37 - Gráficos de Resíduo MAE-T e MAE-V.



Fonte: o autor, 2022.

Figura 38 - Gráficos de Resíduo  $R^2$  e  $R_a^2$ .



Fonte: o autor, 2022.

### 5.8.1 Bagaço de cana de açúcar ácido uma camada (Qt)

Na tabela 26, são apresentados os fatores, as respostas MAE-T, MAE-V,  $R^2$  e  $R_a^2$  e o tempo de CPU utilizados no planejamento experimental da RNA.

**Tabela 26 - Planejamento do tratamento ácido Qt**

Exp	NN	NE	FA	DIV (%)	MAE(T)	MSE(T)	$R^2$	$R_a^2$	Tempo(s)
1	30	2000	Sigmoid	30	3,8807	36,1429	0,9004	0,8993	238,00
2	30	2000	Sigmoid	40	4,0669	39,4666	0,8885	0,8872	248,62
3	30	2000	Relu	30	3,1827	20,6517	0,9416	0,9409	237,77
4	30	2000	Relu	40	3,2137	20,4083	0,9394	0,9387	246,67
5	30	4000	Sigmoid	30	1,8007	6,28175	0,9817	0,9815	471,51
6	30	4000	Sigmoid	40	2,1091	9,23819	0,9730	0,9727	488,73
7	30	4000	Relu	30	2,0792	8,5296	0,9761	0,9758	470,13
8	30	4000	Relu	40	2,4177	10,1548	0,9690	0,9686	488,18
9	30	8000	Sigmoid	30	0,5937	0,6254	0,9975	0,9974	936,33
10	30	8000	Sigmoid	40	0,5644	0,5734	0,9970	0,9969	971,36
11	30	8000	Relu	30	1,3662	4,2429	0,9875	0,9873	936,44
12	30	8000	Relu	40	1,8568	6,1980	0,9805	0,9803	972,94
13	60	2000	Sigmoid	30	2,6445	17,3619	0,9524	0,9519	972,94
14	60	2000	Sigmoid	40	2,9861	19,0451	0,9468	0,9462	247,50
15	60	2000	Relu	30	2,6367	14,3346	0,9597	0,9592	237,95
16	60	2000	Relu	40	2,5074	13,5652	0,9603	0,9599	248,56
17	60	4000	Sigmoid	30	1,5766	4,0854	0,9874	0,9873	469,32
18	60	4000	Sigmoid	40	1,5561	3,9261	0,9870	0,9868	490,84
19	60	4000	Relu	30	1,9463	6,1959	0,9822	0,9820	470,29
20	60	4000	Relu	40	1,6938	5,6857	0,9814	0,9812	489,67
21	60	8000	Sigmoid	30	0,5120	0,5054	0,9982	0,9982	932,51
22	60	8000	Sigmoid	40	0,4665	0,3930	0,9977	0,9977	973,78
23	60	8000	Relu	30	1,7233	5,7122	0,9828	0,9826	938,70
24	60	8000	Relu	40	1,2743	3,6870	0,9881	0,9880	973,58
25	120	2000	Sigmoid	30	2,2674	9,8886	0,9723	0,9720	239,16
26	120	2000	Sigmoid	40	2,2597	9,9726	0,9710	0,9706	248,15
27	120	2000	Relu	30	2,4213	10,5246	0,9692	0,9688	238,71
28	120	2000	Relu	40	2,5173	11,1400	0,9671	0,9667	248,32
29	120	4000	Sigmoid	30	1,4053	3,0863	0,9899	0,9898	472,40
30	120	4000	Sigmoid	40	1,5085	3,3353	0,9885	0,9884	492,43
31	120	4000	Relu	30	1,7009	4,8553	0,9861	0,9859	471,50
32	120	4000	Relu	40	1,5973	4,6122	0,9846	0,9845	490,03
33	120	8000	Sigmoid	30	0,4557	0,4223	0,9982	0,9981	939,04
34	120	8000	Sigmoid	40	0,4754	0,3753	0,9978	0,9978	982,19
35	120	8000	Relu	30	1,3931	4,0248	0,9875	0,9874	940,14
36	120	8000	Relu	40	0,4856	0,4852	0,9974	0,9974	977,46

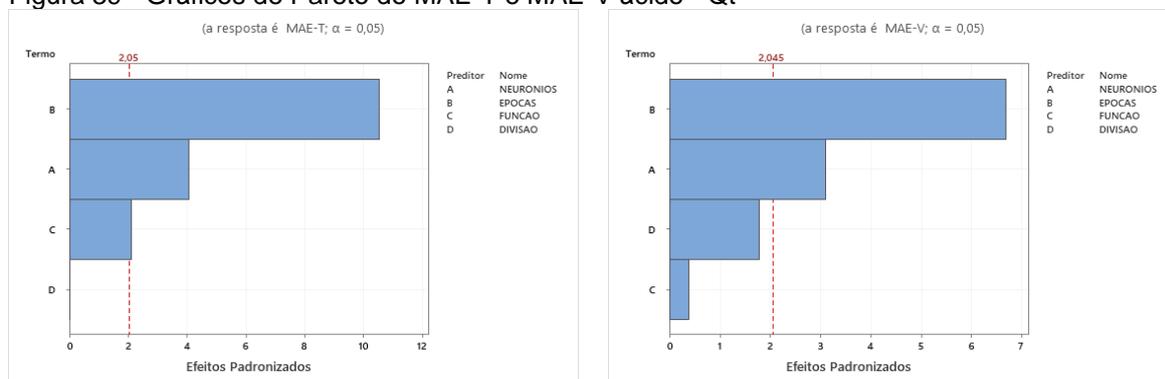
Fonte: o autor, 2022.

Nas figuras 39 e 40, são apresentados os gráficos de Pareto para os seguintes níveis: MAE-T, MAE-V,  $R^2$  e  $R_a^2$ . Para os níveis MAE-T e MAE-V pode ser observado que os fatores épocas e o número de neurônio apresentaram efeitos estatisticamente significativo para o ajuste do modelo, já os fatores função de ativação e divisão para o conjunto de dados não apresentaram efeito estatístico satisfatório.

Nas MLPs, função de ativação, números de épocas, otimizador, taxa de aprendizagem, função de custo, método de inicialização, camadas ocultas, quantidades de neurônios em cada camada e batch size são exemplos de hiperparâmetros. Eles exercem um impacto significativo no desempenho dos modelos (GOODFELLOW; BENGIO; COURVILLE, 2016).

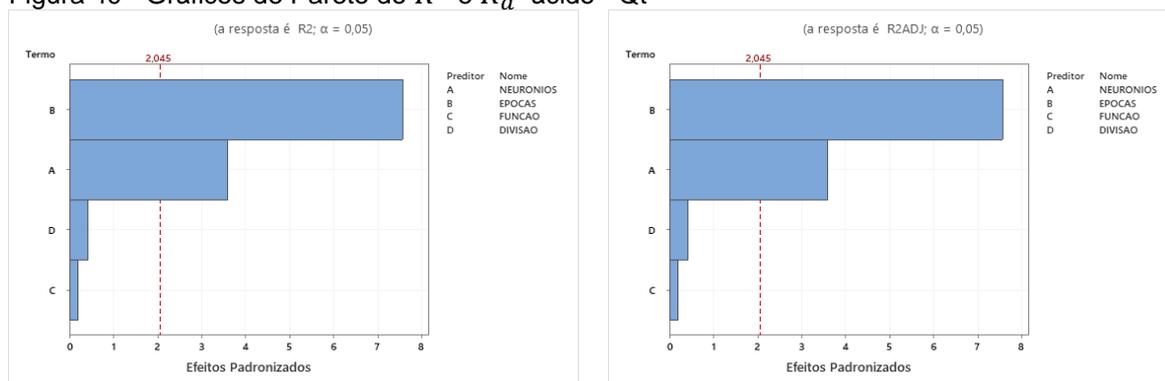
Em relação aos  $R^2$  e  $R_a^2$ , as épocas e os números de neurônios apresentaram efeitos estatisticamente significativo, enquanto, que tanto a função de ativação quanto a divisão dos dados não apresentaram efeito estatisticamente significativo.

Figura 39 - Gráficos de Pareto do MAE-T e MAE-V ácido - Qt



Fonte: o autor, 2022.

Figura 40 - Gráficos de Pareto do  $R^2$  e  $R_a^2$  ácido - Qt

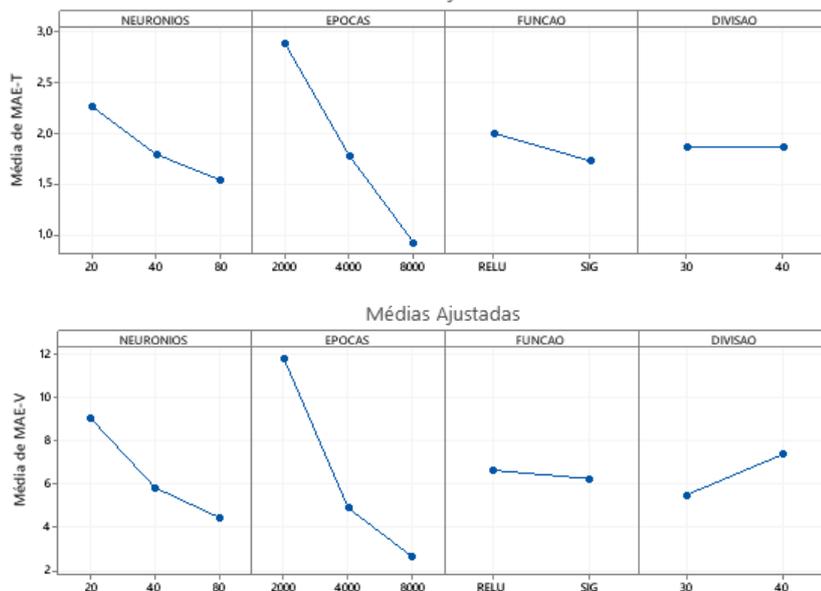


Fonte: o autor, 2022.

Ao analisar os gráficos fatoriais nas figuras 41 e 42, pode ser observado que o tanto o aumento do número de neurônios quanto o número de épocas ocasionam um menor erro, e também que esse aumento apresenta o maior resultado para o coeficiente de determinação.

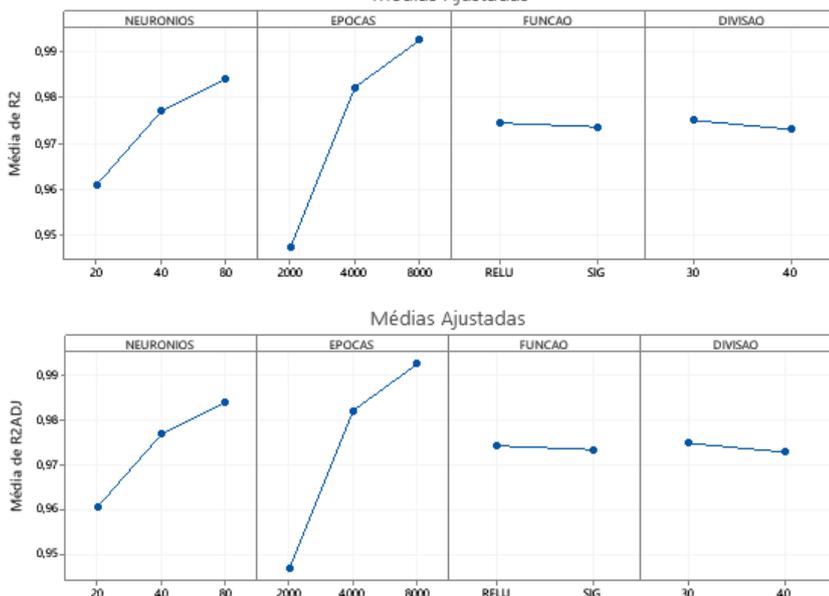
No estudo da linearidade são recomendados testes de premissas para uso apropriado do MQM, como normalidade, homocedasticidade, independência, significância da regressão e desvio da linearidade (SOUZA e JUNQUEIRA, 2005).

Figura 41 - Gráficos fatoriais dos níveis MAE-T e MAE-V ácido - Qt  
Médias Ajustadas



Fonte: o autor, 2022.

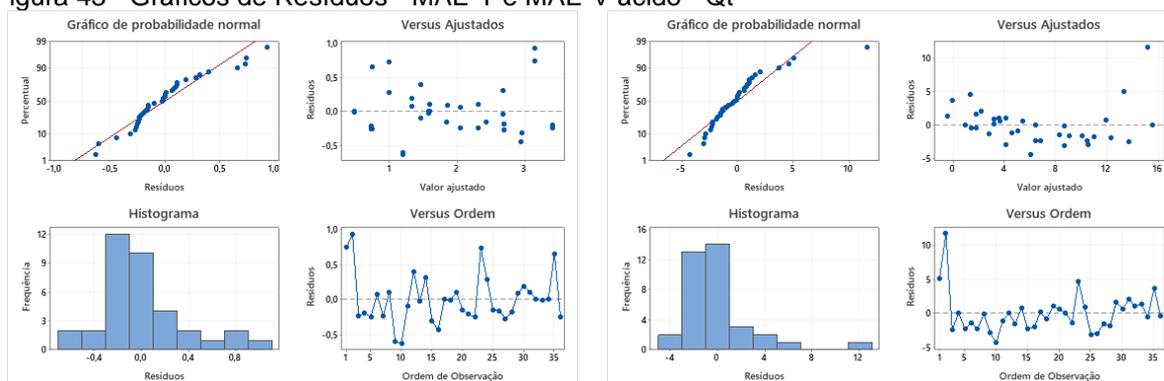
Figura 42 - Gráficos fatoriais dos níveis  $R^2$  e  $R_a^2$  ácido - Qt  
Médias Ajustadas



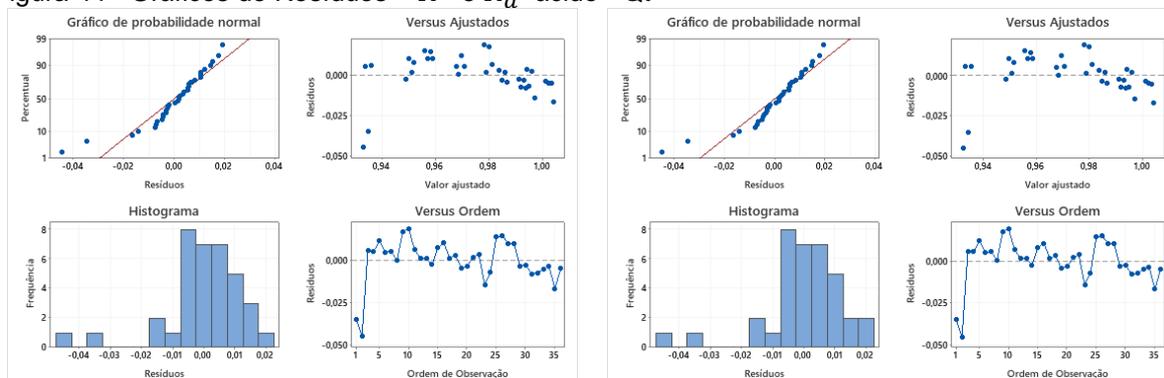
Fonte: o autor, 2022.

Nas figuras 43 e 44, são apresentados os gráficos de Resíduos para os níveis MAE-T e MAE-V, assim também como para o  $R^2$  e  $R_a^2$ .

Figura 43 - Gráficos de Resíduos - MAE-T e MAE-V ácido - Qt



Fonte: o autor, 2022.

Figura 44 - Gráficos de Resíduos –  $R^2$  e  $R_a^2$  ácido - Qt

Fonte: o autor, 2022.

Para o MAE-T e o MAE-V, o gráfico de probabilidade normal dos resíduos indica que os mesmos estão distribuídos normalmente embora a distribuição apresente valores de outliers. Para o gráfico de resíduos versus ajustados pode ser observado que não há a presença de aleatoriedade. No histograma para o parâmetro MAE-T, pode ser notado a presença de distribuição normal, enquanto para o MAE-V há a presença de outliers uma vez que há barras distantes das outras. No gráfico de Resíduos versus Ordem é verificada a independência dos dados.

Para o  $R^2$  e o  $R_a^2$  o gráfico de probabilidade normal dos resíduos indica que os resíduos são distribuídos normalmente com presença de outliers. Para o gráfico de resíduos versus ajustados é observado aleatoriedade na distribuição dos resíduos com variância constante e percebemos também a presença de outliers. No histograma, é possível verificar que os dados são assimétricos à direita. No gráfico de Resíduos versus Ordem pode ser observado a independência entre os resíduos.

Pinheiro e Bates (2006) propuseram o uso do gráfico de probabilidade normal dos resíduos condicionais para avaliar a suposição de normalidade e o gráfico dos resíduos condicionais versus os valores ajustados para avaliar a suposição de homocedasticidade.

A presença de outliers pode resultar na rejeição da suposição de normalidade multivariada e gerar uma distribuição normal multivariada contaminada (Barbosa *et al.*, 2018), o que afeta todas as inferências posteriores baseadas no pressuposto de normalidade (SCHWAGER & MARGOLIN, 1982).

Se houver pontos extremos ou outliers, um histograma permite identificar a frequência e a ordem de grandeza desses pontos (MONTGOMERY e RUNGER, 2018).

A utilização de histogramas é interessante a fim de representar graficamente a distribuição das frequências absolutas ou relativas de um determinado conjunto de valores quantitativos (FREUND, 2006). Tais gráficos, comumente, são considerados como uma das principais ferramentas de qualidade em análises estatísticas, uma vez que geram informações visuais úteis sobre o comportamento de atividades e processos (CORRÊA, 2012).

### 5.9.1 Bagaço de cana de açúcar alcalino uma camada (Qt)

Na tabela 27, são apresentados os fatores, os níveis e as respostas MAE-T, MAE-V,  $R^2$  e  $R_a^2$  e o tempo de CPU utilizados no planejamento experimental da RNA.

**Tabela 27 - Planejamento do tratamento alcalino Qt**

Exp	NN	NE	FA	DIV (%)	MAE(T)	MSE(T)	$R^2$	$R_a^2$	Tempo(s)
1	30	2000	Sigmoid	30	2,9838	23,7926	0,8805	0,8778	229,53
2	30	2000	Sigmoid	40	2,7411	23,1520	0,8901	0,8877	230,18
3	30	2000	Relu	30	2,5425	17,0294	0,9149	0,9130	228,49
4	30	2000	Relu	40	2,7271	20,1471	0,9086	0,9065	229,73
5	30	4000	Sigmoid	30	1,7207	10,8213	0,9434	0,9421	452,32
6	30	4000	Sigmoid	40	1,7207	10,8213	0,9434	0,9421	455,56
7	30	4000	Relu	30	1,7181	8,7006	0,9542	0,9532	453,71
8	30	4000	Relu	40	1,9286	10,6436	0,9509	0,9498	454,35
9	30	8000	Sigmoid	30	0,7368	2,4011	0,9749	0,9743	902,96
10	30	8000	Sigmoid	40	0,6218	1,9452	0,9661	0,9654	902,17
11	30	8000	Relu	30	1,3757	4,1492	0,9698	0,9691	900,25
12	30	8000	Relu	40	1,2152	3,6143	0,9539	0,9529	902,39
13	60	2000	Sigmoid	30	2,3679	16,4231	0,9153	0,9134	231,34
14	60	2000	Sigmoid	40	2,2415	17,6861	0,9150	0,9131	229,11
15	60	2000	Relu	30	2,3433	15,2416	0,9244	0,9228	228,44
16	60	2000	Relu	40	2,1601	13,0430	0,9415	0,9402	229,54
17	60	4000	Sigmoid	30	1,0392	4,8723	0,9661	0,9654	453,62
18	60	4000	Sigmoid	40	1,2449	6,4595	0,9622	0,9613	452,03
19	60	4000	Relu	30	1,4885	6,3185	0,9641	0,9633	451,01
20	60	4000	Relu	40	1,9850	11,2196	0,9475	0,9463	452,59
21	60	8000	Sigmoid	30	0,1410	0,0570	0,9838	0,9834	899,65
22	60	8000	Sigmoid	40	0,1298	0,0475	0,9746	0,9740	901,24
23	60	8000	Relu	30	1,3556	4,1158	0,9620	0,9612	900,51
24	60	8000	Relu	40	1,3556	4,1158	0,9620	0,9612	901,28
25	120	2000	Sigmoid	30	1,6359	9,6989	0,9459	0,9447	229,68
26	120	2000	Sigmoid	40	1,5527	10,190	0,9481	0,9469	230,68
27	120	2000	Relu	30	1,9608	10,928	0,9446	0,9434	229,62
28	120	2000	Relu	40	1,9633	9,3930	0,9540	0,9530	230,38
29	120	4000	Sigmoid	30	0,5253	1,1189	0,9789	0,9784	454,95
30	120	4000	Sigmoid	40	0,3829	0,7061	0,9796	0,9791	455,63
31	120	4000	Relu	30	1,4023	4,4773	0,9660	0,9653	453,81
32	120	4000	Relu	40	1,3181	4,1000	0,9659	0,9651	454,72
33	120	8000	Sigmoid	30	0,0435	0,0144	0,9841	0,9838	904,49
34	120	8000	Sigmoid	40	0,0808	0,0244	0,9782	0,9777	906,54
35	120	8000	Relu	30	1,1080	2,9632	0,9656	0,9649	911,15
36	120	8000	Relu	40	1,3571	4,7689	0,9562	0,9552	904,69

Fonte: o autor, 2022.

Nas figuras 45 e 46, são apresentados os gráficos de Pareto para os seguintes níveis: MAE-T, MAE-V, MSE-V,  $R^2$  e  $R_a^2$ . Para os níveis, MAE-T e MAE-V pode ser observado que os fatores épocas, a função de ativação e o número de neurônio apresentaram efeitos estatisticamente significativo, já o fator divisão para o conjunto de dados não apresentou efeito satisfatório. Já para os níveis,  $R^2$  e  $R_a^2$  pode ser notado que apenas as épocas e o número de neurônios apresentaram efeito estatisticamente

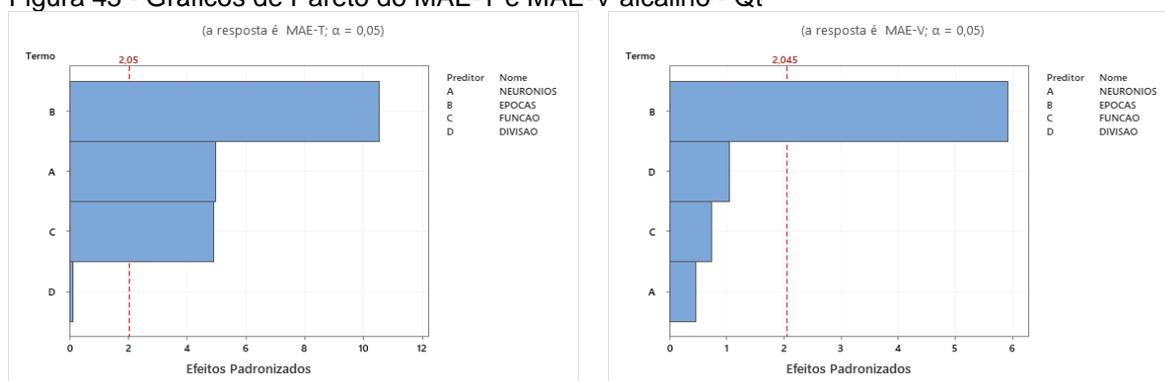
significativo, já os fatores função de ativação e divisão dos dados não apresentaram efeito estatístico para o ajuste.

A função de ativação, números de épocas e quantidades de neurônios em cada camada são exemplos de hiperparâmetros. Eles exercem um impacto significativo no desempenho dos modelos (GOODFELLOW; BENGIO; COURVILLE, 2016).

As funções de ativação são utilizadas na saída de um neurônio tendo como principal objetivo a transformação de uma entrada em um valor de saída (chamada de ativação) dentro de um intervalo determinado de forma que ajude na solução do problema (FRIEDMAN, 2017).

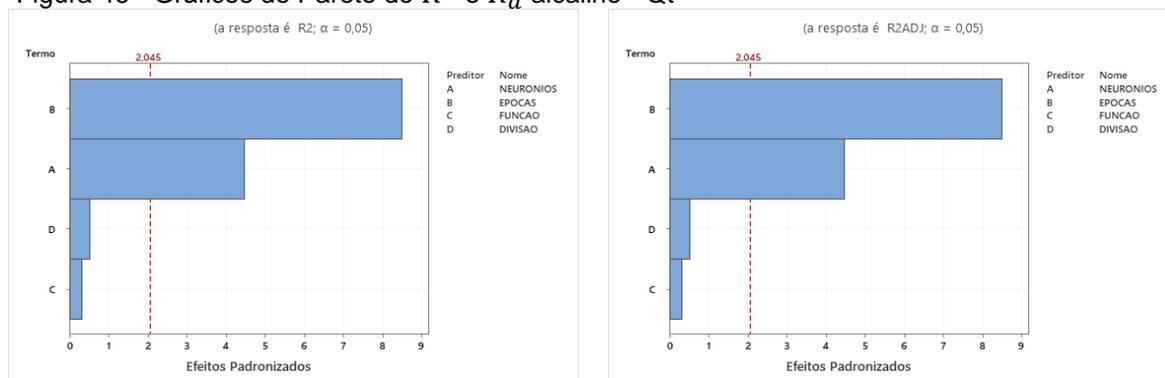
ReLU é a função de ativação mais amplamente utilizada, não linear e permite que apenas alguns neurônios são ativados, tornando a rede esparsa e eficiente (AVINASH, 2017). E a função linear é uma função de ativação que realiza uma transformação linear da entrada, é ideal para tarefas simples ou utilização na camada de saída para problemas de regressão (AVINASH, 2017).

Figura 45 - Gráficos de Pareto do MAE-T e MAE-V alcalino - Qt



Fonte: o autor, 2022.

Figura 46 - Gráficos de Pareto do  $R^2$  e  $R_a^2$  alcalino - Qt

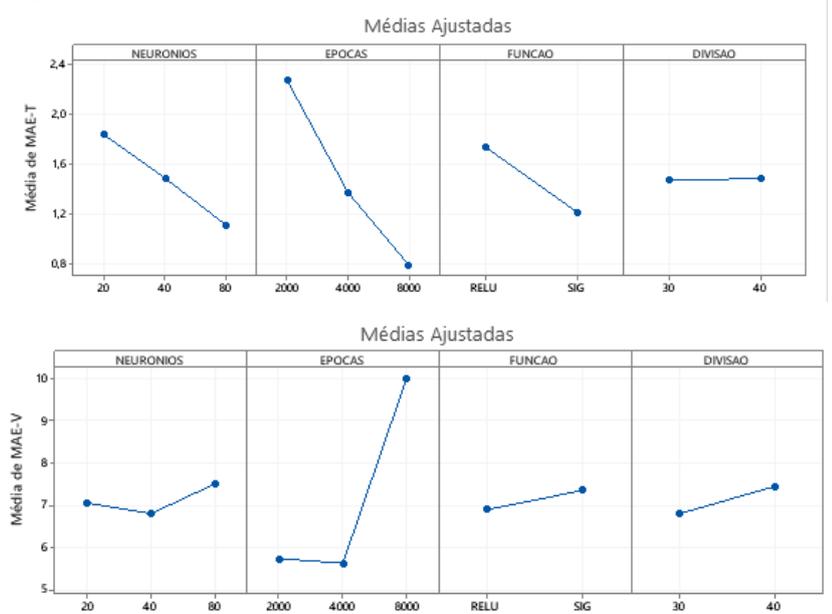


Fonte: o autor, 2022.

Ao analisar os gráficos fatoriais nas figuras 47 e 48, pode ser verificado que a medida que o número de neurônios aumenta, assim também como o aumento do número de épocas são obtidos os melhores resultados para os coeficientes de determinação. Esse mesmo aumento no número de épocas e neurônios permite que seja obtido o menor erro.

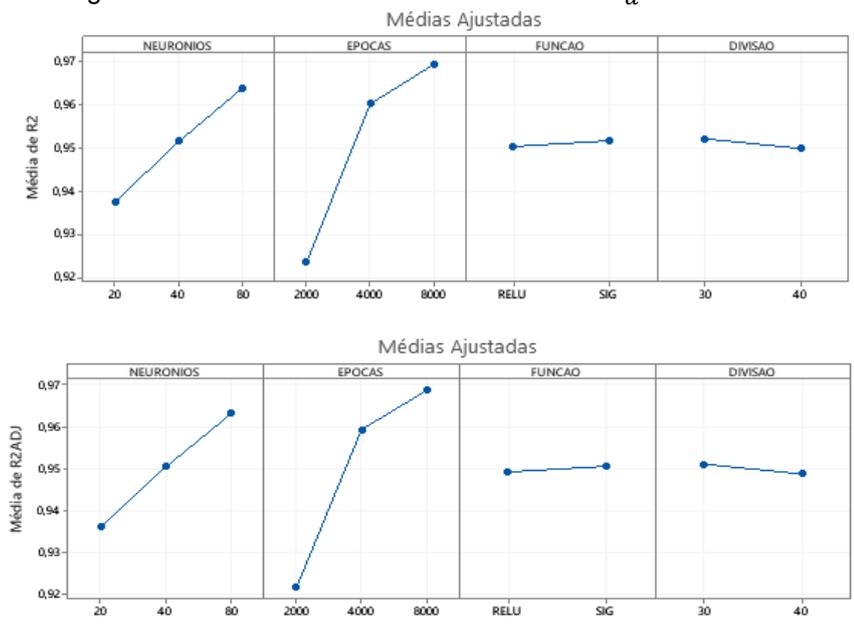
Em deep learning, principalmente em redes convolucionais, neurônios são ativados com ReLU ou alguma de suas variantes. Entre outras vantagens, ReLU é mais simples de calcular, por conseguinte impondo muito menos esforço computacional, e permite maior rapidez na convergência de valores ótimos de peso (LIU, 2017). Além disso, essa função adiciona não linearidade aos dados (SUPERDATASCIENCE TEAM, 2018).

Figura 47 - Gráficos fatoriais dos níveis MAE-T e MAE-V alcalino - Qt



Fonte: o autor, 2022.

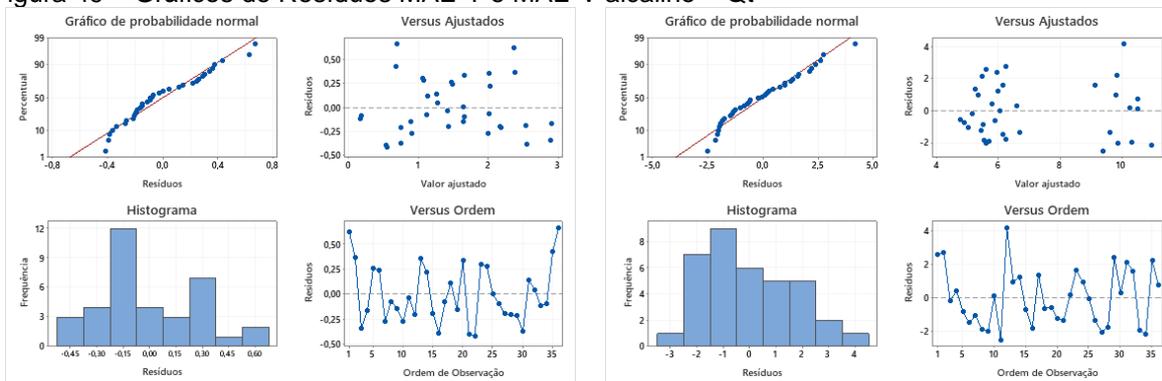
Figura 48 - Gráficos fatoriais dos níveis  $R^2$  e  $R_a^2$  alcalino - Qt



Fonte: o autor, 2022.

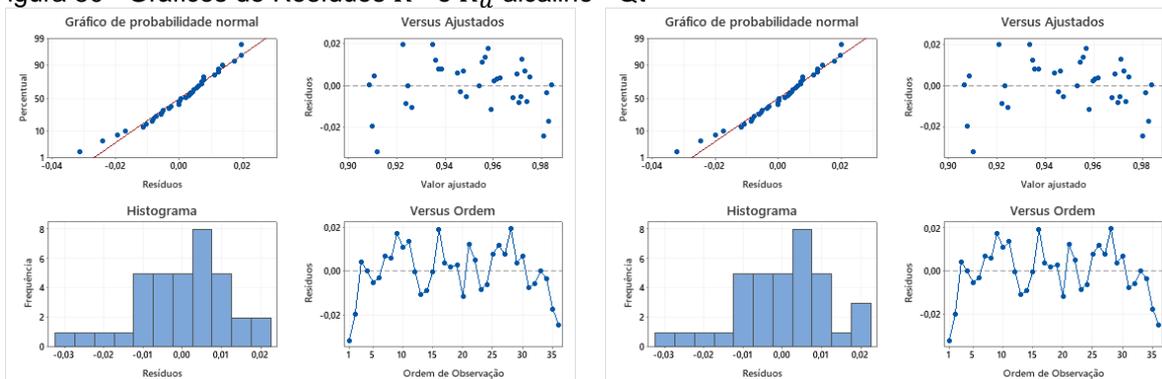
Nas figuras 49 e 50, são apresentados os gráficos de Resíduos para os parâmetros MAE-T e MAE-V, assim também como para o  $R^2$  e o  $R_a^2$ .

Figura 49 – Gráficos de Resíduos MAE-T e MAE-V alcalino – Qt



Fonte: o autor, 2022.

Figura 50 - Gráficos de Resíduos  $R^2$  e  $R_a^2$  alcalino - Qt

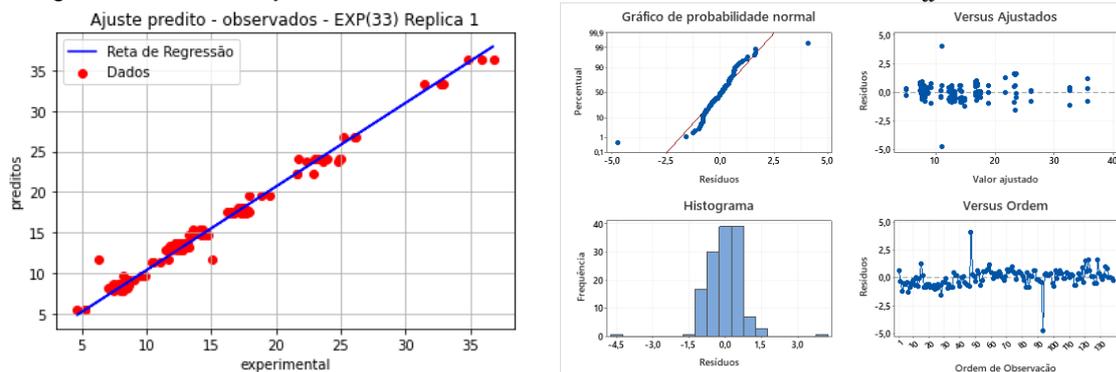


Fonte: o autor, 2022.

### 5.10.1 Otimização da RNA com uma camada (Qt)

Na figura 51, pode ser observado o experimento 34 que indica o valor otimizado do MAE-T, MAE-V e  $R_a^2$  para a RNA com uma camada, 80 neurônios, 8.000 épocas, função de ativação: Sigmoid e 30% para validação dos dados com um tempo de CPU igual a 455,53 segundos para o tratamento in-natura da variável Qt.

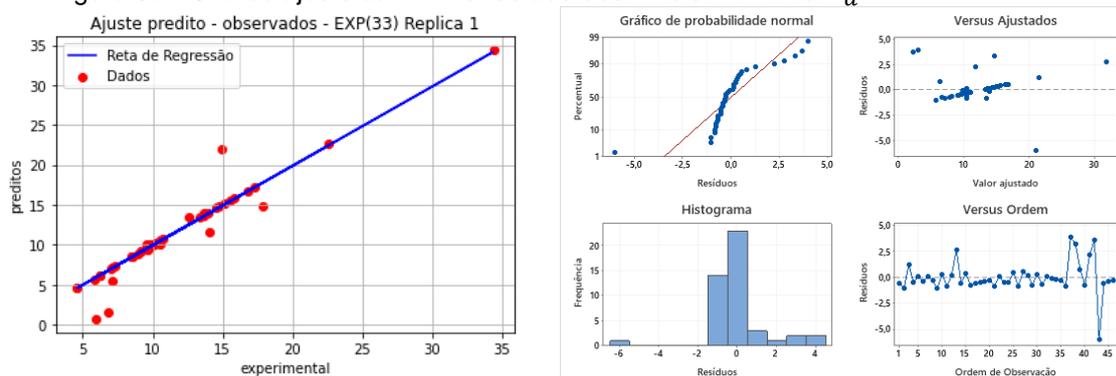
Figura 51 - Gráfico ajuste da RNA e resíduo dos níveis MAE-T, MAE-V e  $R_a^2$



Fonte: o autor, 2022.

Na figura 52, pode ser observado para o experimento 33 que aponta a otimização para o valor do MAE-T e  $R_a^2$  para a RNA com uma camada, 80 neurônios, 8.000 épocas, função de ativação: Sigmoid e 30% para validação dos dados com um tempo de CPU igual a 432,64 segundos para o tratamento alcalino da variável Qt.

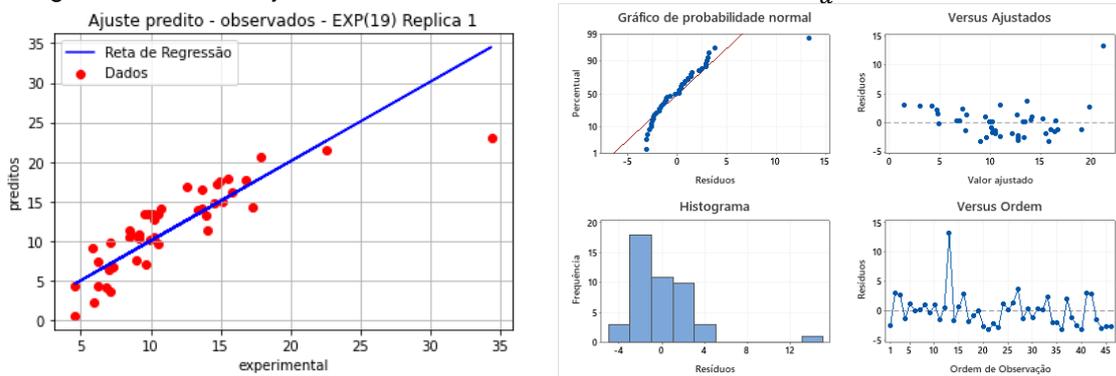
Figura 52 - Gráfico ajuste da RNA e resíduo dos níveis MAE-T e  $R_a^2$ .



Fonte: o autor, 2022.

Na figura 53, pode ser observado o experimento 19 onde é verificado a otimização do valor do MAE-V para a RNA com uma camada, 40 neurônios, 4.000 épocas, função de ativação: Relu e 30% para validação dos dados com um tempo de CPU igual a 213,16 segundos para o tratamento alcalino da variável Qt.

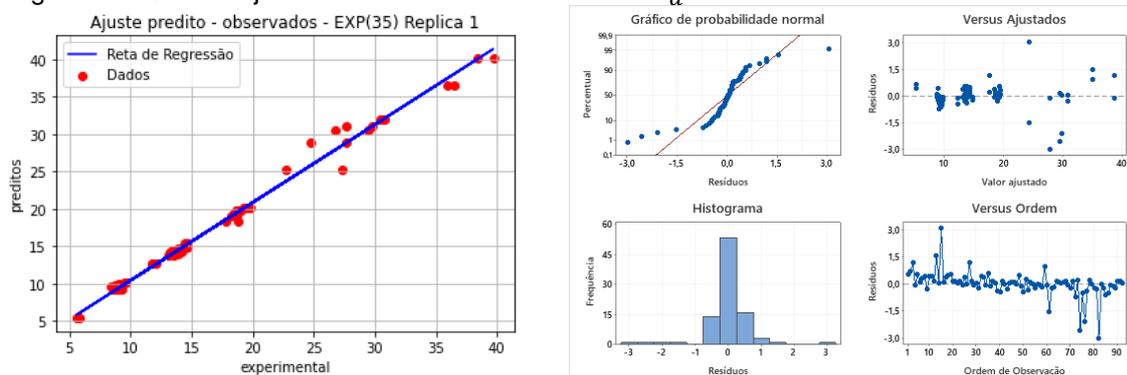
Figura 53 - Gráfico ajuste da RNA e resíduo dos níveis MAE-T e  $R_a^2$



Fonte: o autor, 2022.

Na figura 54, pode ser identificado o experimento 35 que detecta a otimização do  $R_a^2$  com uma camada, 80 neurônios, 8.000 épocas, função de ativação: Relu e 30% para validação dos dados com um tempo de CPU igual a 462,55 segundos para o tratamento ácido da variável Qt.

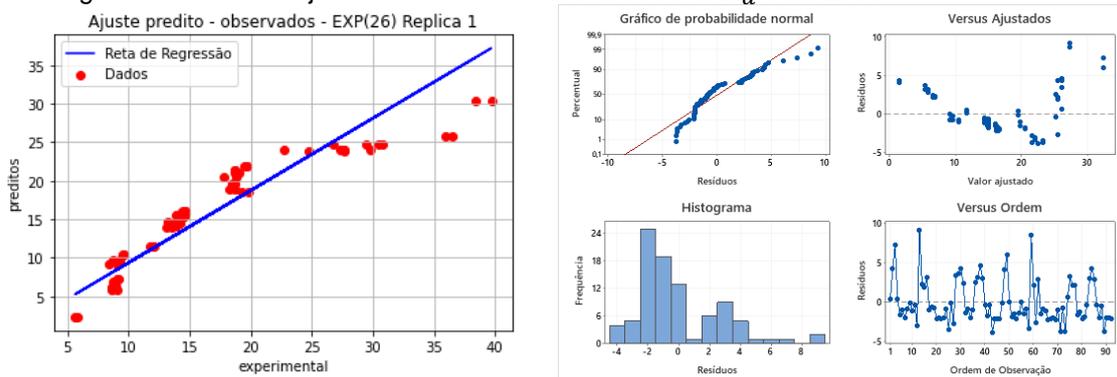
Figura 54 - Gráfico ajuste da RNA e resíduo do nível  $R_a^2$ .



Fonte: o autor, 2022.

A figura 55, apresenta o experimento 26 onde é verificado a otimização do MAE-T com uma camada, 80 neurônios, 2.000 épocas, função de ativação: Sigmoid e 40% para validação dos dados com um tempo de CPU igual a 122,04 segundos para o tratamento ácido da variável Qt.

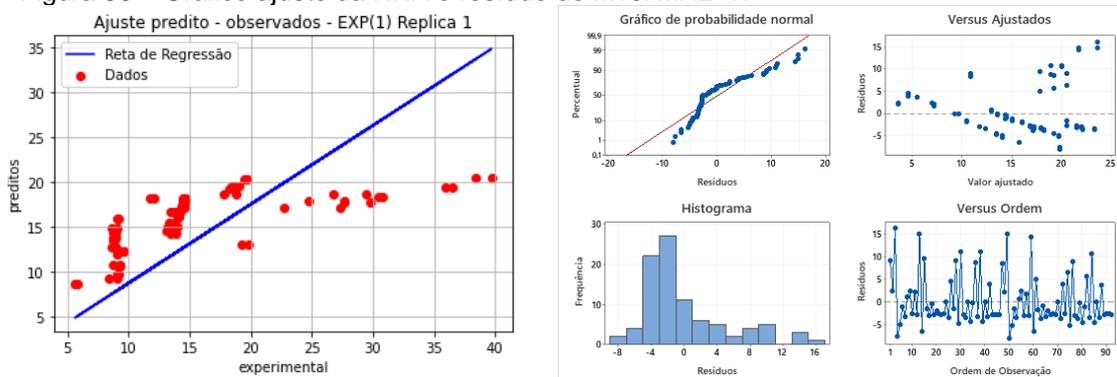
Figura 55 – Gráfico ajuste da RNA e resíduo dos níveis  $R_a^2$  e MAE-T.



Fonte: o autor, 2022.

A figura 56, exibe o experimento 1 que apresenta a otimização do MAE-V com uma camada, 20 neurônios, 2.000 épocas, função de ativação: Relu e 30% para validação dos dados com um tempo de CPU igual a 108.36 segundos para o tratamento ácido da variável Qt.

Figura 56 – Gráfico ajuste da RNA e resíduo do nível MAE-V.



Fonte: o autor, 2022

### 5.10.2 Tratamento do bagaço de cana de açúcar com duas camadas

Na tabela 28, são apresentados os fatores, os níveis e as respostas MAE-T, MAE-V,  $R^2$  e  $R_a^2$  e o tempo de CPU utilizados no planejamento experimental da RNA.

**Tabela 28 - Planejamento Experimental do tratamento ácido em duas camadas**

(continua)

Exp	NN	NN	NE	FA	DIV (%)	MAE(T)	MSE(T)	$R^2$ (Y1)	$R_a^2$ (Y1)	$R^2$ (Y2)	$R_a^2$ (Y2)	Tempo(s)
1	30	30	2000	Sigmoid	30	5,4548	57,1977	0,8194	0,8174	0,9934	0,9933	481,20
2	30	30	2000	Sigmoid	40	5,2219	50,9269	0,8238	0,8219	0,9934	0,9934	500,05
3	30	30	2000	Relu	30	3,2816	19,6293	0,9631	0,9627	0,9970	0,9969	477,41
4	30	30	2000	Relu	40	3,3042	18,7973	0,9562	0,9557	0,9965	0,9965	497,32
5	30	30	4000	Sigmoid	30	1,4384	4,4520	0,9910	0,9909	0,9991	0,9990	948,74
6	30	30	4000	Sigmoid	40	1,2354	3,6307	0,9885	0,9883	0,9990	0,9990	989,87
7	30	30	4000	Relu	30	1,6910	6,2216	0,9904	0,9903	0,9987	0,9987	946,56
8	30	30	4000	Relu	40	1,0070	2,4737	0,9929	0,9928	0,9989	0,9989	983,95
9	30	30	8000	Sigmoid	30	0,6572	1,2869	0,9987	0,9987	0,9994	0,9994	1887,93
10	30	30	8000	Sigmoid	40	0,3742	0,3440	0,9985	0,9985	0,9993	0,9993	1968,41
11	30	30	8000	Relu	30	1,4057	5,1666	0,9906	0,9905	0,9989	0,9989	1883,78
12	30	30	8000	Relu	40	0,8149	1,4495	0,9953	0,9953	0,9992	0,9992	1968,05
13	30	60	2000	Sigmoid	30	4,9422	48,5986	0,8410	0,8392	0,9945	0,9945	478,82
14	30	60	2000	Sigmoid	40	4,4683	39,8495	0,8674	0,8659	0,9946	0,9945	501,53
15	30	60	2000	Relu	30	2,8021	15,221	0,9747	0,9744	0,9975	0,9975	479,47
16	30	60	2000	Relu	40	2,5603	12,7024	0,9754	0,9752	0,9972	0,9972	500,19
17	30	60	4000	Sigmoid	30	0,8447	1,6681	0,9981	0,9980	0,9994	0,9994	952,32
18	30	60	4000	Sigmoid	40	0,6305	0,8025	0,9978	0,9978	0,9993	0,9993	994,17
19	30	60	4000	Relu	30	1,0594	2,5868	0,9967	0,9967	0,9993	0,9993	950,12
20	30	60	4000	Relu	40	1,2945	4,0314	0,9913	0,9912	0,9987	0,9987	992,35
21	30	60	8000	Sigmoid	30	0,6398	1,2880	0,9987	0,9987	0,9994	0,9994	1912,24
22	30	60	8000	Sigmoid	40	0,3569	0,3238	0,9985	0,9985	0,9994	0,9993	1991,96
23	30	60	8000	Relu	30	0,7864	1,6811	0,9987	0,9986	0,9994	0,9994	1903,11
24	30	60	8000	Relu	40	0,6417	1,0105	0,9967	0,9966	0,9993	0,9993	1976,83
25	30	90	2000	Sigmoid	30	5,2014	52,7517	0,8248	0,8228	0,9942	0,9942	482,11
26	30	90	2000	Sigmoid	40	5,0090	46,5549	0,8390	0,8373	0,9940	0,9940	503,21
27	30	90	2000	Relu	30	2,4556	12,1102	0,9825	0,9823	0,9979	0,9978	479,56
28	30	90	2000	Relu	40	1,8100	7,4047	0,9890	0,9888	0,9979	0,9979	500,56
29	30	90	4000	Sigmoid	30	0,7880	1,5681	0,9984	0,9984	0,9994	0,9994	952,33
30	30	90	4000	Sigmoid	40	0,4612	0,4384	0,9983	0,9983	0,9994	0,9994	991,86
31	30	90	4000	Relu	30	0,8834	1,8167	0,9981	0,9980	0,9994	0,9994	951,66
32	30	90	4000	Relu	40	0,7354	1,1734	0,9977	0,9977	0,9993	0,9993	991,82
33	30	90	8000	Sigmoid	30	0,6165	1,2254	0,9987	0,9987	0,9994	0,9994	1894,81
34	30	90	8000	Sigmoid	40	0,3530	0,3111	0,9985	0,9985	0,9994	0,9993	1985,91
35	30	90	8000	Relu	30	0,7782	1,5350	0,9984	0,9984	0,9994	0,9994	1891,09
36	30	90	8000	Relu	40	0,6173	0,9762	0,9963	0,9963	0,9993	0,9993	1973,12

(...).

Tabela 28 – Planejamento do tratamento ácido em duas camadas

(continuação)

Exp	NN	NN	NE	FA	DIV (%)	MAE(T)	MSE(T)	$R^2(Y1)$	$R_a^2(Y1)$	$R^2(Y2)$	$R_a^2(Y2)$	Tempo(s)
37	60	30	2000	Sigmoid	30	5,4833	57,2186	0,8194	0,8174	0,9934	0,9933	483,56
38	60	30	2000	Sigmoid	40	5,2566	51,7645	0,8194	0,8174	0,9934	0,9933	510,21
39	60	30	2000	Relu	30	3,4287	20,3705	0,9583	0,9578	0,9970	0,9969	481,49
40	60	30	2000	Relu	40	2,7864	14,5855	0,9724	0,9721	0,9969	0,9969	502,64
41	60	30	4000	Sigmoid	30	2,2415	16,0954	0,9499	0,9493	0,9979	0,9979	956,82
42	60	30	4000	Sigmoid	40	1,1208	3,8302	0,9866	0,9865	0,9991	0,9991	1000,09
43	60	30	4000	Relu	30	1,3360	3,7327	0,9907	0,9906	0,9993	0,9993	953,15
44	60	30	4000	Relu	40	0,9852	2,3911	0,9932	0,9932	0,9991	0,9991	994,23
45	60	30	8000	Sigmoid	30	0,6512	1,3223	0,9987	0,9987	0,9994	0,9994	1902,73
46	60	30	8000	Sigmoid	40	0,3576	0,3288	0,9985	0,9985	0,9994	0,9993	1984,83
47	60	30	8000	Relu	30	0,8028	1,6101	0,9983	0,9983	0,9994	0,9994	1899,37
48	60	30	8000	Relu	40	0,8191	1,6515	0,9959	0,9959	0,9992	0,9992	1982,99
49	60	60	2000	Sigmoid	30	5,4532	57,2029	0,8194	0,8174	0,9934	0,9933	483,37
50	60	60	2000	Sigmoid	40	5,2466	51,7222	0,8197	0,8177	0,9934	0,9933	506,03
51	60	60	2000	Relu	30	1,9351	7,7959	0,9917	0,9916	0,9983	0,9983	483,40
52	60	60	2000	Relu	40	2,0178	8,3990	0,9824	0,9822	0,9979	0,9978	503,30
53	60	60	4000	Sigmoid	30	0,7983	1,6290	0,9984	0,9984	0,9994	0,9994	957,25
54	60	60	4000	Sigmoid	40	0,5796	0,6892	0,9981	0,9981	0,9993	0,9993	1005,32
55	60	60	4000	Relu	30	0,9773	2,0985	0,9982	0,9982	0,9994	0,9994	958,30
56	60	60	4000	Relu	40	0,8609	1,9399	0,9936	0,9935	0,9991	0,9991	997,91
57	60	60	8000	Sigmoid	30	0,6248	1,2386	0,9987	0,9987	0,9994	0,9994	1901,18
58	60	60	8000	Sigmoid	40	0,4022	0,4682	0,9984	0,9984	0,9994	0,9993	1988,99
59	60	60	8000	Relu	30	0,7264	1,4648	0,9987	0,9987	0,9994	0,9994	1897,83
60	60	60	8000	Relu	40	0,5134	0,6205	0,9984	0,9983	0,9993	0,9993	1981,12
61	60	90	2000	Sigmoid	30	5,3032	54,395	0,8301	0,8282	0,9936	0,9936	484,57
62	60	90	2000	Sigmoid	40	4,4580	42,137	0,8539	0,8523	0,9944	0,9944	506,59
63	60	90	2000	Relu	30	2,2416	9,6965	0,9863	0,9861	0,9982	0,9982	483,09
64	60	90	2000	Relu	40	1,9630	7,5053	0,9889	0,9888	0,9978	0,9978	503,28
65	60	90	4000	Sigmoid	30	0,7025	1,3912	0,9986	0,9986	0,9994	0,9994	956,40
66	60	90	4000	Sigmoid	40	0,4049	0,3710	0,9984	0,9984	0,9994	0,9994	999,60
67	60	90	4000	Relu	30	0,8938	1,9443	0,9984	0,9984	0,9994	0,9994	955,80
68	60	90	4000	Relu	40	0,7342	1,1682	0,9967	0,9966	0,9992	0,9992	997,81
69	60	90	8000	Sigmoid	30	0,6617	1,2817	0,9987	0,9987	0,9994	0,9994	1907,35
70	60	90	8000	Sigmoid	40	0,3419	0,3161	0,9985	0,9985	0,9993	0,9993	1991,03
71	60	90	8000	Relu	30	0,7231	1,4612	0,9987	0,9987	0,9994	0,9994	1899,78
72	60	90	8000	Relu	40	0,3950	0,3732	0,9984	0,9984	0,9993	0,9993	1987,71

(....)

Tabela 28 - Planejamento do tratamento ácido em duas camadas

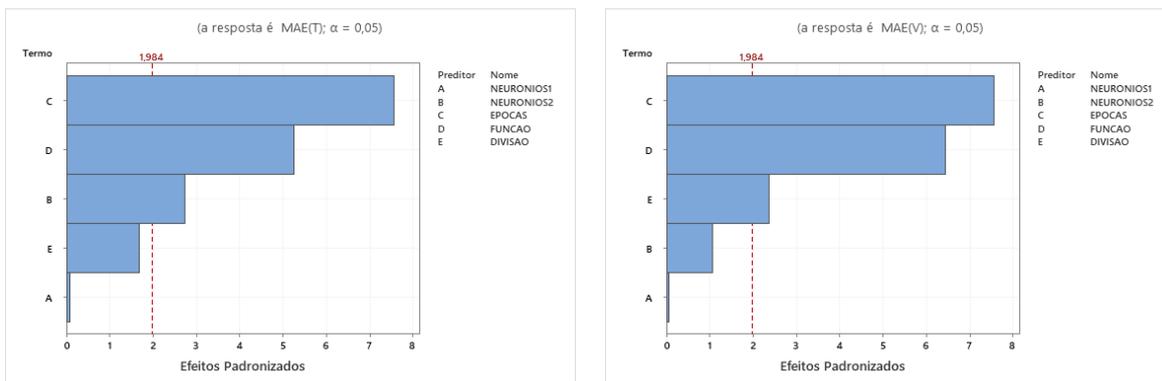
(conclusão)

Exp	NN	NN	NE	FA	DIV (%)	MAE(T)	MSE(T)	$R^2(Y1)$	$R_a^2(Y1)$	$R^2(Y2)$	$R_a^2(Y2)$	Tempo(s)
73	90	30	2000	Sigmoid	30	5,4719	57,2053	0,8194	0,8174	0,9934	0,9933	484,77
74	90	30	2000	Sigmoid	40	5,6184	54,2444	0,8194	0,8174	0,9934	0,9933	507,80
75	90	30	2000	Relu	30	2,8309	13,8290	0,9754	0,9752	0,9976	0,9976	484,23
76	90	30	2000	Relu	40	2,2818	10,186	0,9797	0,9795	0,9976	0,9975	503,98
77	90	30	4000	Sigmoid	30	1,5767	6,4669	0,9792	0,9790	0,9990	0,9990	959,07
78	90	30	4000	Sigmoid	40	2,2802	17,816	0,9356	0,9349	0,9975	0,9974	1002,90
79	90	30	4000	Relu	30	1,6934	6,4617	0,9913	0,9912	0,9986	0,9986	956,31
80	90	30	4000	Relu	40	1,1527	2,8685	0,9917	0,9916	0,9989	0,9989	999,05
81	90	30	8000	Sigmoid	30	0,6593	1,2831	0,9987	0,9987	0,9994	0,9994	1905,94
82	90	30	8000	Sigmoid	40	0,3826	0,3468	0,9985	0,9985	0,9994	0,9994	1994,51
83	90	30	8000	Relu	30	1,6290	5,9553	0,9907	0,9906	0,9987	0,9987	1903,85
84	90	30	8000	Relu	40	0,7290	1,1814	0,9959	0,9958	0,9993	0,9993	1988,35
85	90	60	2000	Sigmoid	30	4,7117	46,4234	0,8556	0,8540	0,9946	0,9945	487,98
86	90	60	2000	Sigmoid	40	5,2443	51,7620	0,8194	0,8174	0,9934	0,9933	508,76
87	90	60	2000	Relu	30	2,0691	9,1152	0,9889	0,9888	0,9982	0,9982	485,58
88	90	60	2000	Relu	40	1,8399	7,0956	0,9898	0,9897	0,9979	0,9978	507,62
89	90	60	4000	Sigmoid	30	0,7413	1,4801	0,9984	0,9984	0,9994	0,9994	960,93
90	90	60	4000	Sigmoid	40	0,4948	0,5159	0,9982	0,9982	0,9994	0,9993	1008,14
91	90	60	4000	Relu	30	1,2178	3,4466	0,9967	0,9967	0,9993	0,9993	959,63
92	90	60	4000	Relu	40	0,8200	1,2703	0,9966	0,9966	0,9992	0,9992	1000,40
93	90	60	8000	Sigmoid	30	0,6464	1,2859	0,9988	0,9987	0,9994	0,9994	1904,31
94	90	60	8000	Sigmoid	40	0,3655	0,3289	0,9985	0,9985	0,9993	0,9993	2000,60
95	90	60	8000	Relu	30	0,8269	1,6528	0,9984	0,9984	0,9994	0,9994	1902,30
96	90	60	8000	Relu	40	0,6013	0,7226	0,9981	0,9981	0,9993	0,9993	1988,09
97	90	90	2000	Sigmoid	30	4,9786	48,9509	0,8493	0,8476	0,9942	0,9941	485,57
98	90	90	2000	Sigmoid	40	5,2555	51,7123	0,8195	0,8176	0,9934	0,9933	511,47
99	90	90	2000	Relu	30	1,9482	8,0601	0,9888	0,9887	0,9984	0,9984	484,90
100	90	90	2000	Relu	40	1,6603	5,6800	0,9884	0,9882	0,9982	0,9982	508,11
101	90	90	4000	Sigmoid	30	0,6708	1,3175	0,9986	0,9986	0,9994	0,9994	959,08
102	90	90	4000	Sigmoid	40	0,4137	0,3883	0,9984	0,9984	0,9994	0,9994	1005,8
103	90	90	4000	Relu	30	0,9907	2,3951	0,9975	0,9975	0,9993	0,9993	959,55
104	90	90	4000	Relu	40	0,9609	2,1244	0,9958	0,9957	0,9992	0,9992	1002,16
105	90	90	8000	Sigmoid	30	0,6179	1,2285	0,9987	0,9987	0,9994	0,9994	1906,29
106	90	90	8000	Sigmoid	40	0,3360	0,3094	0,9985	0,9985	0,9993	0,9993	1997,99
107	90	90	8000	Relu	30	0,7488	1,4999	0,9987	0,9987	0,9994	0,9994	1908,46
108	90	90	8000	Relu	40	0,5502	0,7969	0,9984	0,9984	0,9994	0,9993	1995,23

Fonte: o autor, 2022.

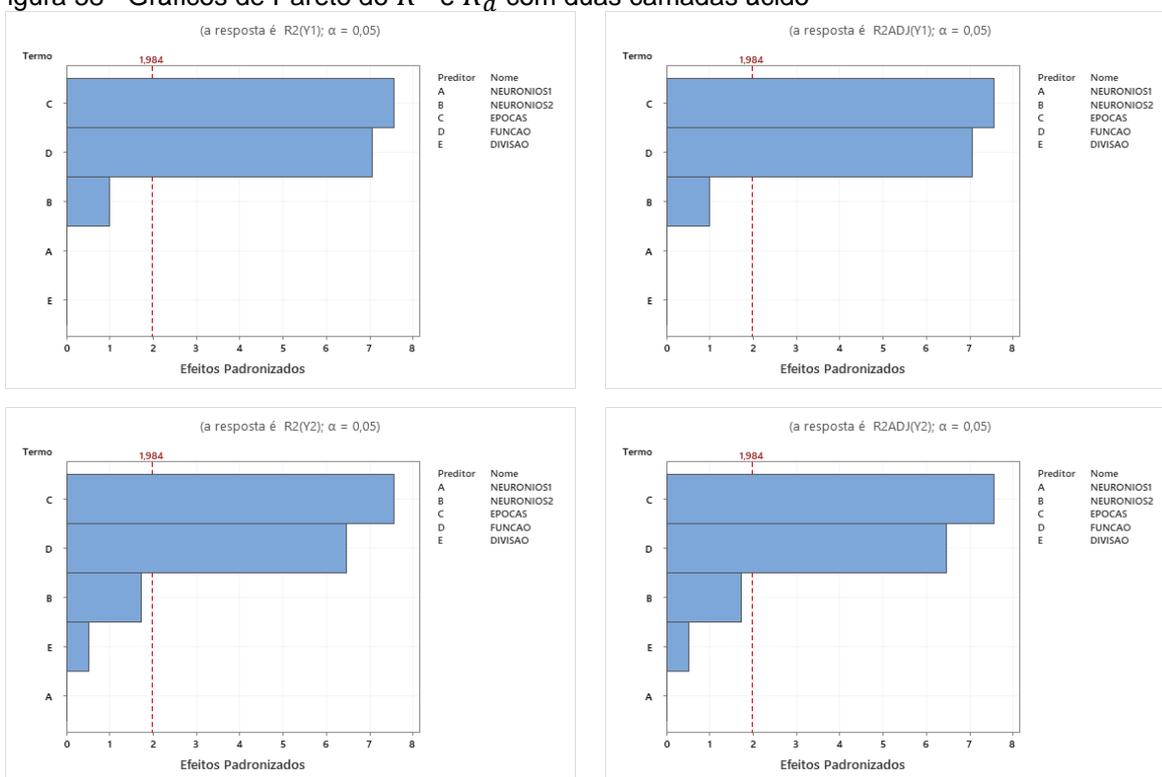
Nas figuras 57. e 58, são apresentados os gráficos de Pareto dos seguintes níveis: MAE-T, MAE-V,  $R^2$  e  $R_a^2$ . Nos níveis, MAE-T e MAE-V, pode ser observado que os fatores épocas, a função de ativação e o número de neurônios apresentaram efeitos estatisticamente significativo para o ajuste do modelo, já o fator divisão para o conjunto de dados não apresentou efeito estatístico significativo. Em relação aos  $R^2$  e  $R_a^2$  as épocas e as funções de ativação apresentaram efeitos estatisticamente significativo. O número de neurônios e a divisão não apresentam efeito satisfatório para o ajuste do modelo.

Figura 57 - Gráficos de Pareto do MAE-T e MAE-V com duas camadas ácido



Fonte: o autor, 2022.

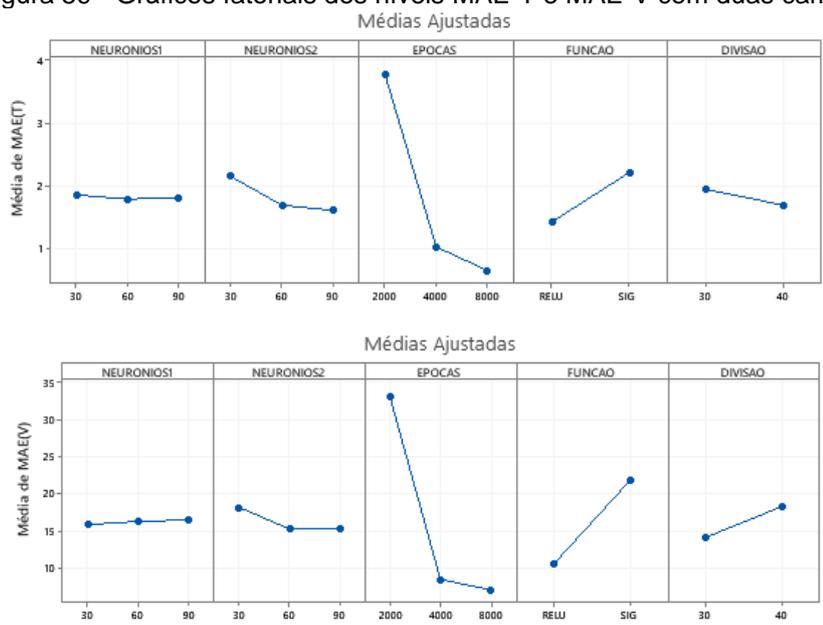
Figura 58 - Gráficos de Pareto do  $R^2$  e  $R^2_\alpha$  com duas camadas ácido



Fonte: o autor, 2022.

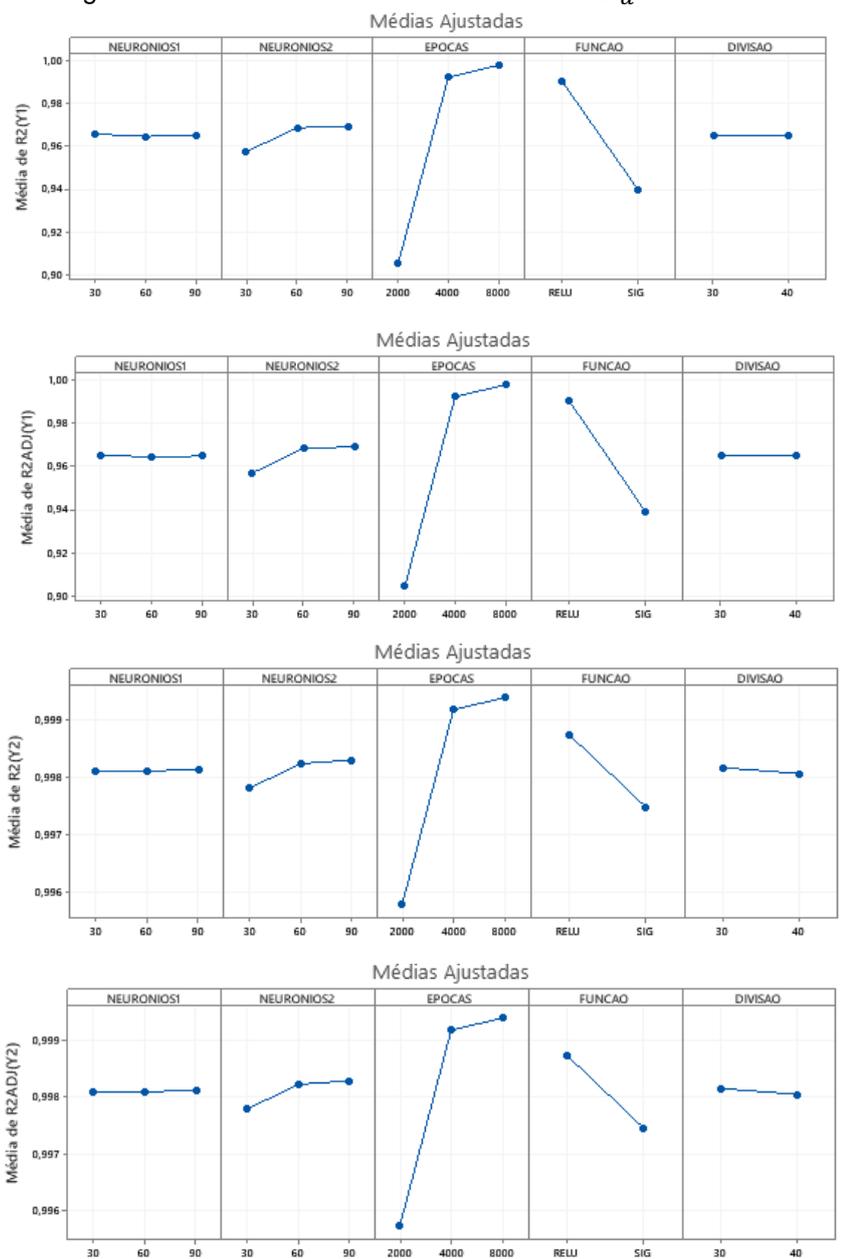
Ao analisar os gráficos fatoriais nas figuras 59 e 60, pode ser observado que o aumento do número de épocas apresentou o menor valor para o erro obtido. No caso do coeficiente de determinação, o aumento do número de épocas permitiu que fosse obtido a melhor resposta.

Figura 59 - Gráficos fatoriais dos níveis MAE-T e MAE-V com duas camadas ácido



Fonte: o autor, 2022.

Figura 60 - Gráficos fatoriais dos níveis  $R^2$  e  $R_a^2$  com duas camadas ácido



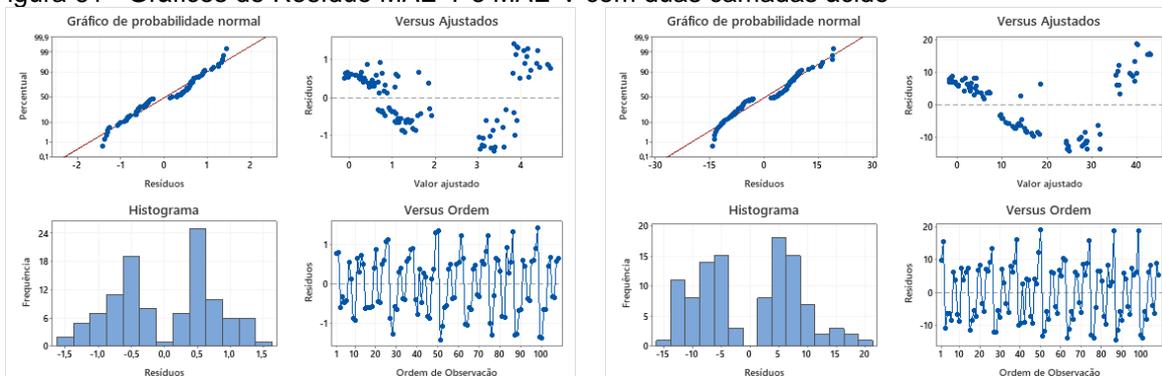
Fonte: o autor, 2022.

Nas figuras 61 e 62, são apresentados os gráficos de Resíduos para os níveis MAE-T e MAE-V,  $R^2$  e o  $R_a^2$ , também é apresentado o gráfico de probabilidade normal dos resíduos, apresentando uma distribuição normal, visto que a maioria dos pontos estão próximos a linha de referência, o que também valida a análise dos dados para o método de monitoramento proposto.

Para o gráfico de resíduos versus ajustados, é possível verificar a pressuposição que os resíduos são aleatoriamente distribuídos e com variância

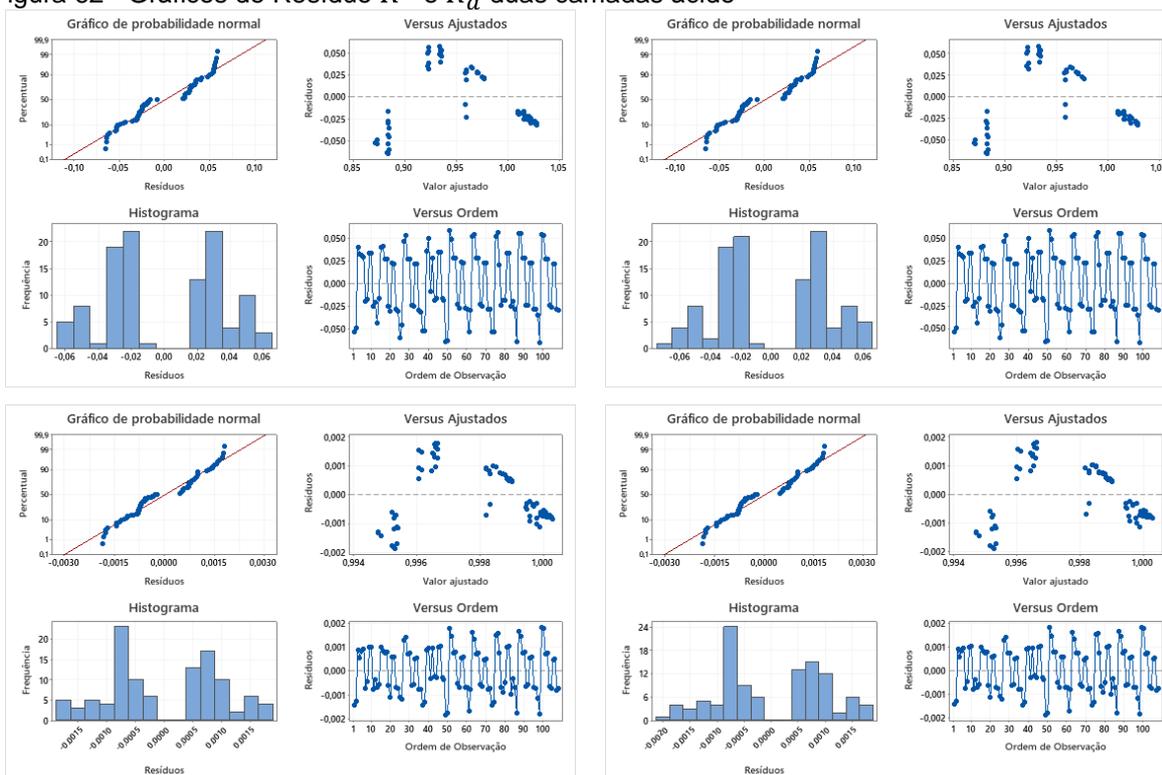
constante e também é percebido a presença de outliers. No gráfico de Resíduos versus ordem verifica-se que há independência entre os resíduos.

Figura 61 - Gráficos de Resíduo MAE-T e MAE-V com duas camadas ácido



Fonte: o autor, 2022.

Figura 62 - Gráficos de Resíduo  $R^2$  e  $R_a^2$  duas camadas ácido



Fonte: o autor, 2022.

Na tabela 29, são apresentados os fatores, níveis e as respostas MAE-T, MAE-V,  $R^2$  e  $R_a^2$  e o tempo de CPU utilizados no planejamento experimental da RNA para o tratamento in-natura.

Tabela 29 - Planejamento do tratamento in-inatura em duas camadas

(continua)

Exp	NN	NN	NE	FA	DIV (%)	MAE(T)	MSE(T)	$R^2(Y1)$	$R_a^2(Y1)$	$R^2(Y2)$	$R_a^2(Y2)$	Tempo(s)
1	30	30	2000	Sigmoid	30	3,2634	22,3970	0,9507	0,9504	0,9959	0,9959	520,95
2	30	30	2000	Sigmoid	40	4,1064	30,6867	0,8887	0,8879	0,9957	0,9957	523,68
3	30	30	2000	Relu	30	3,4022	25,7514	0,9660	0,9657	0,9940	0,9939	518,78
4	30	30	2000	Relu	40	3,0944	21,9601	0,9721	0,9719	0,9952	0,9951	522,21
5	30	30	4000	Sigmoid	30	1,8944	7,23439	0,9969	0,9969	0,9982	0,9982	1032,11
6	30	30	4000	Sigmoid	40	1,9549	7,46639	0,9955	0,9954	0,9980	0,9980	1038,58
7	30	30	4000	Relu	30	2,1974	9,81280	0,9922	0,9921	0,9977	0,9977	1025,74
8	30	30	4000	Relu	40	2,4693	12,2991	0,9860	0,9859	0,9971	0,9971	1032,77
9	30	30	8000	Sigmoid	30	1,8580	6,99128	0,9974	0,9974	0,9983	0,9982	2048,37
10	30	30	8000	Sigmoid	40	1,7900	6,68525	0,9968	0,9968	0,9981	0,9981	2064,39
11	30	30	8000	Relu	30	1,9617	7,6768	0,9940	0,9939	0,9982	0,9982	2048,48
12	30	30	8000	Relu	40	1,8524	7,3650	0,9963	0,9962	0,9980	0,9980	2067,69
13	30	60	2000	Sigmoid	30	2,2457	10,1454	0,9911	0,9910	0,9977	0,9976	514,69
14	30	60	2000	Sigmoid	40	2,4115	11,3819	0,9893	0,9893	0,9972	0,9972	522,50
15	30	60	2000	Relu	30	2,9921	19,0340	0,9768	0,9766	0,9955	0,9954	519,18
16	30	60	2000	Relu	40	2,8796	18,3206	0,9776	0,9774	0,9959	0,9959	521,38
17	30	60	4000	Sigmoid	30	1,8668	7,2486	0,9972	0,9972	0,9982	0,9982	1031,80
18	30	60	4000	Sigmoid	40	1,8871	7,1517	0,9965	0,9964	0,9981	0,9981	1037,36
19	30	60	4000	Relu	30	1,9615	7,6512	0,9961	0,9961	0,9982	0,9981	1030,30
20	30	60	4000	Relu	40	2,0635	8,5745	0,9935	0,9934	0,9980	0,9979	1035,45
21	30	60	8000	Sigmoid	30	1,8149	6,8363	0,9975	0,9975	0,9983	0,9982	2055,58
22	30	60	8000	Sigmoid	40	1,7699	6,6144	0,9967	0,9967	0,9981	0,9981	2068,36
23	30	60	8000	Relu	30	2,0552	9,3574	0,9922	0,9922	0,9979	0,9978	2054,83
24	30	60	8000	Relu	40	1,8242	6,9597	0,9964	0,9964	0,9980	0,9980	2064,37
25	30	90	2000	Sigmoid	30	2,1047	9,1317	0,9944	0,9943	0,9979	0,9978	521,10
26	30	90	2000	Sigmoid	40	2,1848	9,5085	0,9927	0,9927	0,9977	0,9977	524,36
27	30	90	2000	Relu	30	2,3599	12,6160	0,9876	0,9875	0,9969	0,9969	520,46
28	30	90	2000	Relu	40	2,2448	10,9035	0,9928	0,9927	0,9973	0,9972	521,49
29	30	90	4000	Sigmoid	30	1,8621	6,9631	0,9972	0,9971	0,9982	0,9982	1034,59
30	30	90	4000	Sigmoid	40	1,8124	6,9279	0,9964	0,9964	0,9981	0,9981	1039,50
31	30	90	4000	Relu	30	1,9068	7,3550	0,9963	0,9963	0,9982	0,9982	1032,40
32	30	90	4000	Relu	40	1,9297	8,0679	0,9953	0,9953	0,9980	0,9980	1036,03
33	30	90	8000	Sigmoid	30	1,8701	7,1540	0,9974	0,9974	0,9982	0,9982	2059,95
34	30	90	8000	Sigmoid	40	1,7939	6,7132	0,9969	0,9968	0,9981	0,9981	2079,45
35	30	90	8000	Relu	30	1,8404	6,9361	0,9973	0,9973	0,9982	0,9982	2055,99
36	30	90	8000	Relu	40	1,7965	6,7741	0,9968	0,9967	0,9981	0,9981	2069,60

(...)

Tabela 29 - Planejamento do tratamento inatura em duas camadas

(continuação)

Exp	NN	NN	NE	FA	DIV (%)	MAE(T)	MSE(T)	$R^2(Y1)$	$R_a^2(Y1)$	$R^2(Y2)$	$R_a^2(Y2)$	Tempo(s)
37	60	30	2000	Sigmoid	30	4,6684	41,5901	0,8876	0,8868	0,9931	0,9931	522,86
38	60	30	2000	Sigmoid	40	4,3945	38,6102	0,9026	0,9019	0,9934	0,9933	523,96
39	60	30	2000	Relu	30	2,9981	17,5789	0,9752	0,9750	0,9962	0,9961	518,44
40	60	30	2000	Relu	40	2,9997	19,068	0,9757	0,9755	0,9958	0,9958	521,83
41	60	30	4000	Sigmoid	30	1,8698	7,1777	0,9973	0,9973	0,9982	0,9982	1033,80
42	60	30	4000	Sigmoid	40	1,8756	7,1948	0,9965	0,9964	0,9981	0,9981	1039,56
43	60	30	4000	Relu	30	2,0777	8,4352	0,9907	0,9906	0,9981	0,9981	1027,49
44	60	30	4000	Relu	40	2,4617	11,9648	0,9819	0,9818	0,9973	0,9973	1034,56
45	60	30	8000	Sigmoid	30	1,8226	6,8275	0,9975	0,9974	0,9982	0,9982	2057,70
46	60	30	8000	Sigmoid	40	1,7749	6,6372	0,9969	0,9969	0,9981	0,9981	2068,29
47	60	30	8000	Relu	30	1,8721	7,3686	0,9968	0,9968	0,9982	0,9982	2052,49
48	60	30	8000	Relu	40	1,9459	7,6014	0,9949	0,9948	0,9980	0,9980	2063,68
49	60	60	2000	Sigmoid	30	2,2562	10,7755	0,9922	0,9921	0,9975	0,9974	522,43
50	60	60	2000	Sigmoid	40	2,2568	10,465	0,9925	0,9924	0,9975	0,9975	525,63
51	60	60	2000	Relu	30	2,4682	13,280	0,9868	0,9867	0,9968	0,9968	521,58
52	60	60	2000	Relu	40	2,5405	13,947	0,9858	0,9857	0,9969	0,9969	523,26
53	60	60	4000	Sigmoid	30	1,8822	7,0461	0,9973	0,9973	0,9982	0,9982	1040,20
54	60	60	4000	Sigmoid	40	1,8163	6,8478	0,9966	0,9966	0,9981	0,9981	1044,34
55	60	60	4000	Relu	30	1,9277	7,6345	0,9966	0,9966	0,9982	0,9981	1033,18
56	60	60	4000	Relu	40	2,0634	8,2712	0,9900	0,9900	0,9980	0,9980	1040,48
57	60	60	8000	Sigmoid	30	1,8135	6,7640	0,9975	0,9975	0,9982	0,9982	2066,36
58	60	60	8000	Sigmoid	40	1,7867	6,6919	0,9969	0,9968	0,9981	0,9981	2076,82
59	60	60	8000	Relu	30	1,8541	7,1948	0,9971	0,9971	0,9982	0,9982	2057,78
60	60	60	8000	Relu	40	1,8296	6,9600	0,9967	0,9967	0,9981	0,9980	2067,88
61	60	90	2000	Sigmoid	30	2,0162	8,5265	0,9960	0,9960	0,9980	0,9980	522,389
62	60	90	2000	Sigmoid	40	2,0736	8,8747	0,9951	0,9951	0,9977	0,9977	526,513
63	60	90	2000	Relu	30	2,1859	10,063	0,9933	0,9932	0,9978	0,9978	522,601
64	60	90	2000	Relu	40	2,2651	10,778	0,9891	0,9890	0,9974	0,9973	524,818
65	60	90	4000	Sigmoid	30	1,8726	7,1513	0,9975	0,9974	0,9983	0,9982	1038,39
66	60	90	4000	Sigmoid	40	1,9004	7,1103	0,9968	0,9967	0,9981	0,9981	1042,75
67	60	90	4000	Relu	30	1,8751	7,3018	0,9970	0,9970	0,9982	0,9982	1036,41
68	60	90	4000	Relu	40	1,8794	7,0700	0,9958	0,9958	0,9980	0,9980	1044,01
69	60	90	8000	Sigmoid	30	1,8554	6,9549	0,9975	0,9975	0,9982	0,9982	2068,45
70	60	90	8000	Sigmoid	40	1,7881	6,7817	0,9969	0,9969	0,9981	0,9981	2081,11
71	60	90	8000	Relu	30	1,8482	7,0551	0,9974	0,9974	0,9982	0,9982	2065,11
72	60	90	8000	Relu	40	1,8098	7,0068	0,9968	0,9968	0,9981	0,9980	2077,79

(...)

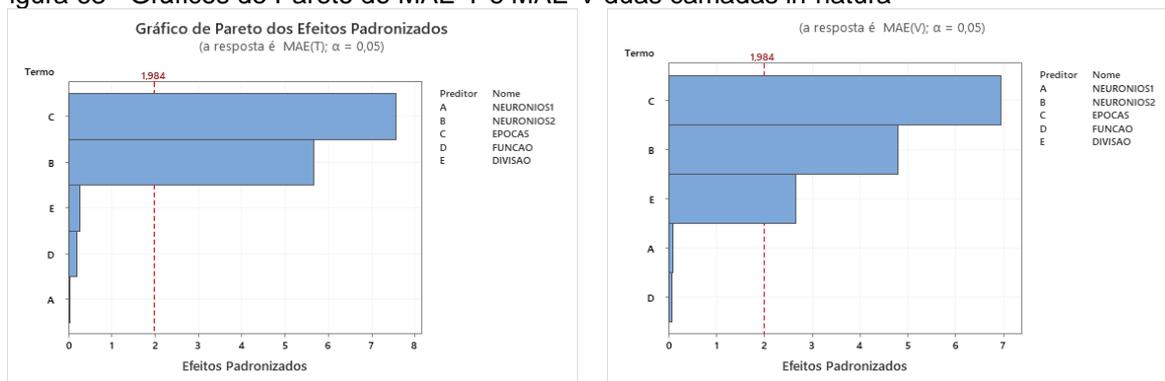
Tabela 29 - Planejamento do tratamento inatura em duas camadas

Exp	NN	NN	NE	FA	DIV (%)	MAE(T)	MSE(T)	$R^2(Y1)$	$R_a^2(Y1)$	$R^2(Y2)$	$R_a^2(Y2)$	(conclusão)
												Tempo(s)
73	90	30	2000	Sigmoid	30	5,0019	49,6373	0,8805	0,8797	0,9909	0,9908	524,73
74	90	30	2000	Sigmoid	40	5,4013	57,4829	0,8679	0,8670	0,9897	0,9897	525,90
75	90	30	2000	Relu	30	2,8522	17,5050	0,9788	0,9786	0,9959	0,9959	520,11
76	90	30	2000	Relu	40	3,1968	20,6048	0,9693	0,9691	0,9956	0,9956	524,35
77	90	30	4000	Sigmoid	30	1,8275	6,9497	0,9974	0,9974	0,9982	0,9982	1037,68
78	90	30	4000	Sigmoid	40	1,8172	6,9091	0,9966	0,9966	0,9981	0,9981	1042,22
79	90	30	4000	Relu	30	2,1684	9,3953	0,9895	0,9894	0,9979	0,9979	1032,26
80	90	30	4000	Relu	40	2,1555	9,2790	0,9907	0,9906	0,9976	0,9975	1041,38
81	90	30	8000	Sigmoid	30	1,8140	6,7556	0,9975	0,9975	0,9983	0,9982	2062,73
82	90	30	8000	Sigmoid	40	1,7906	6,6192	0,9968	0,9968	0,9981	0,9981	2077,93
83	90	30	8000	Relu	30	1,9028	7,6605	0,9972	0,9972	0,9982	0,9982	2060,07
84	90	30	8000	Relu	40	1,8909	7,3909	0,9964	0,9963	0,9980	0,9980	2073,52
85	90	60	2000	Sigmoid	30	1,9434	8,0437	0,9966	0,9966	0,9981	0,9980	525,13
86	90	60	2000	Sigmoid	40	2,4070	12,0951	0,9900	0,9899	0,9971	0,9971	528,85
87	90	60	2000	Relu	30	2,4461	12,6710	0,9886	0,9886	0,9973	0,9973	525,17
88	90	60	2000	Relu	40	2,2458	9,6213	0,9884	0,9883	0,9977	0,9977	527,10
89	90	60	4000	Sigmoid	30	1,8315	6,9207	0,9974	0,9974	0,9982	0,9982	1044,08
90	90	60	4000	Sigmoid	40	1,8128	6,8554	0,9967	0,9967	0,9981	0,9981	1046,39
91	90	60	4000	Relu	30	1,9200	7,4778	0,9969	0,9969	0,9982	0,9982	1036,75
92	90	60	4000	Relu	40	1,9010	7,1893	0,9955	0,9955	0,9980	0,9980	1043,28
93	90	60	8000	Sigmoid	30	1,8449	6,9010	0,9974	0,9974	0,9982	0,9982	2075,02
94	90	60	8000	Sigmoid	40	1,7534	6,6216	0,9969	0,9969	0,9981	0,9981	2082,89
95	90	60	8000	Relu	30	1,8445	7,2403	0,9974	0,9974	0,9982	0,9982	2065,54
96	90	60	8000	Relu	40	1,8814	7,0095	0,9963	0,9962	0,9980	0,9980	2081,05
97	90	90	2000	Sigmoid	30	1,9781	8,1801	0,9967	0,9967	0,9981	0,9981	524,38
98	90	90	2000	Sigmoid	40	1,9793	8,0168	0,9958	0,9958	0,9979	0,9979	527,94
99	90	90	2000	Relu	30	2,1472	9,4577	0,9945	0,9944	0,9980	0,9980	522,96
100	90	90	2000	Relu	40	2,0085	8,2546	0,9927	0,9927	0,9978	0,9978	527,10
101	90	90	4000	Sigmoid	30	1,8636	7,0293	0,9974	0,9974	0,9982	0,9982	1040,59
102	90	90	4000	Sigmoid	40	1,7877	6,7140	0,9968	0,9968	0,9981	0,9981	1049,64
103	90	90	4000	Relu	30	1,8857	7,2628	0,9971	0,9971	0,9982	0,9982	1040,52
104	90	90	4000	Relu	40	1,8811	7,2411	0,9963	0,9963	0,9980	0,9980	1043,46
105	90	90	8000	Sigmoid	30	1,8071	6,7630	0,9976	0,9975	0,9983	0,9982	2082,83
106	90	90	8000	Sigmoid	40	1,7524	6,5583	0,9969	0,9969	0,9981	0,9981	2091,86
107	90	90	8000	Relu	30	1,8509	7,0671	0,9974	0,9974	0,9982	0,9982	2074,41
108	90	90	8000	Relu	40	1,8154	7,1574	0,9967	0,9967	0,9980	0,9980	2085,60

Fonte: o autor, 2022.

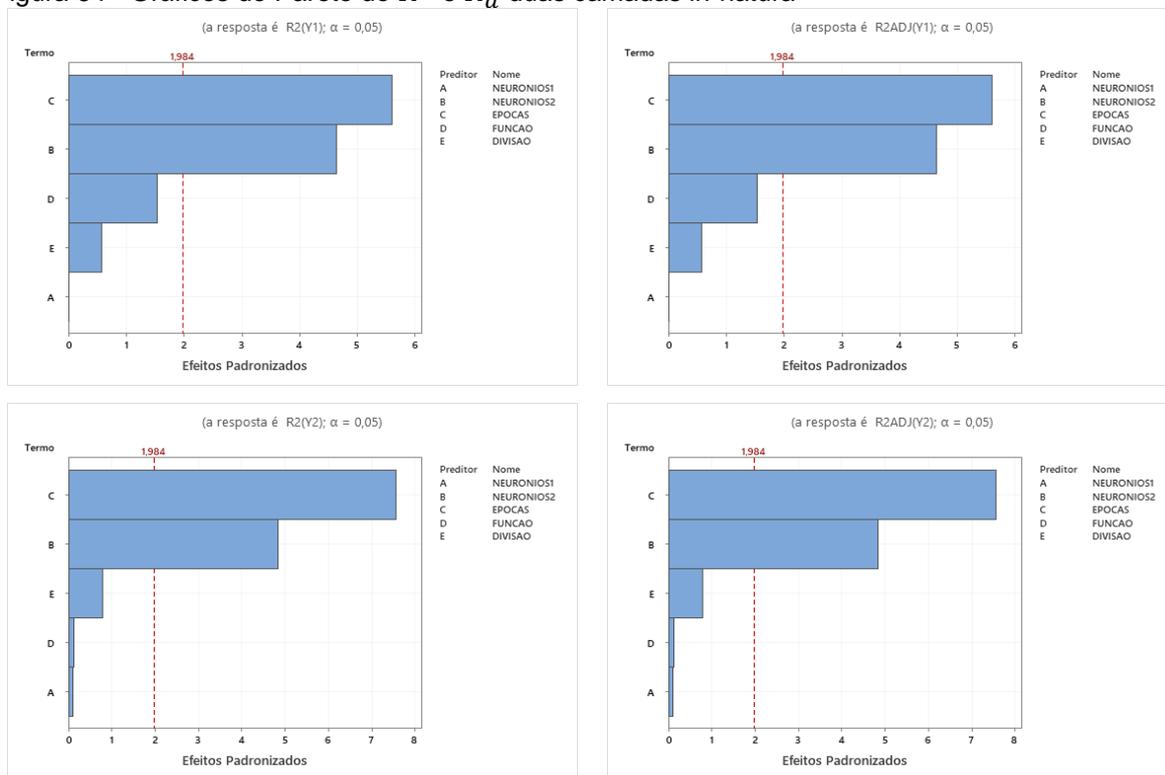
Nas figuras 63 e 64, são apresentados os gráficos de Pareto para os seguintes níveis: MAE-T, MAE-V,  $R^2$  e  $R_a^2$ . Nos Níveis, MAE-T e MAE-V, pode ser observado que os fatores épocas, a função de ativação e o número de neurônios apresentaram efeitos estatisticamente significativo para o ajuste do modelo, já o fator divisão para o conjunto de dados não apresentou efeito estisticamnete significativo. Em relação aos  $R^2$  e  $R_a^2$ , as épocas e as funções de ativação apresentaram efeitos estatisticamente significativo. O número de neurônios e a divisão não apresentam efeito satisfatório para ajuste do modelo.

Figura 63 - Gráficos de Pareto do MAE-T e MAE-V duas camadas in-natura



Fonte: o autor, 2022.

Figura 64 - Gráficos de Pareto do R<sup>2</sup> e R<sub>α</sub><sup>2</sup> duas camadas in-natura



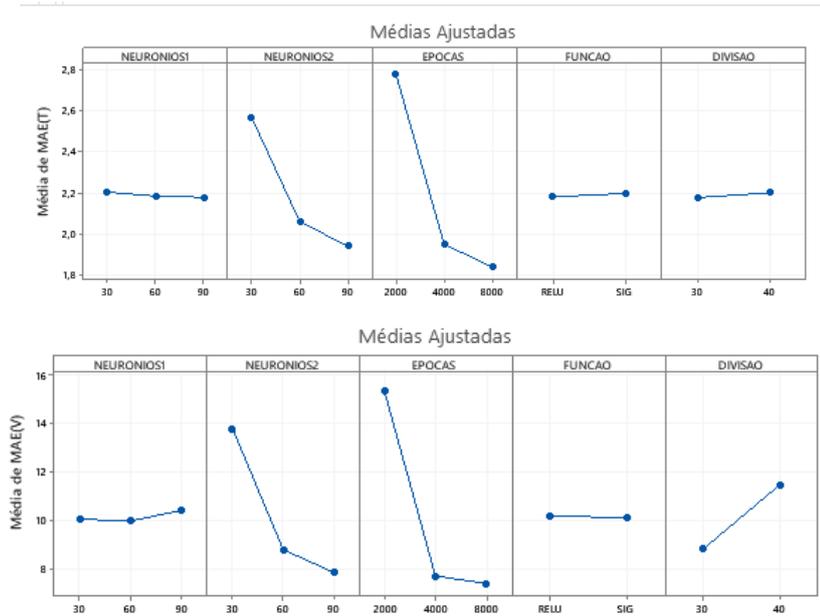
Fonte: o autor, 2022.

Ao analisar os gráficos fatoriais nas figuras 65 e 66, pode ser verificado que, tanto o aumento do número de neurônios na segunda camada, quanto as épocas, indicaram os menores erros obtidos, também, pode ser verificado que esse aumento do número de épocas e do número de neurônios permite que seja obtido o melhor valor para o coeficiente de determinação.

O principal motivo para o uso das RNA, fundamenta-se na sua capacidade de reconhecimento de padrões, generalização, auto-organização e processamento temporal, o que viabiliza a solução de diversos problemas com os mais variados níveis

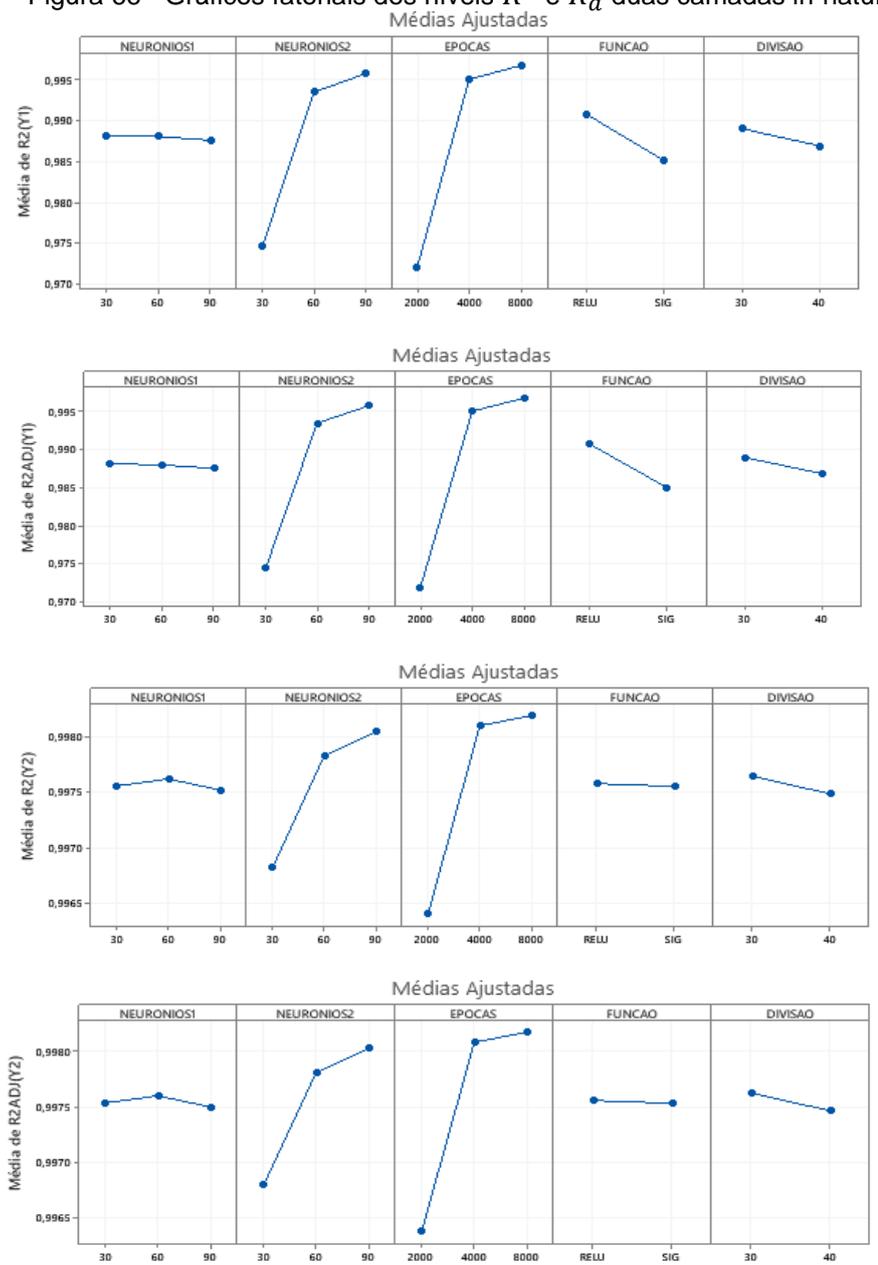
de complexidade e apresenta bons resultados em problemas que possuam elevado grau de não linearidade, contudo as RNAs não apresentam a mesma eficiência para todos os problemas. (FIORIN *et.al.* 2011).

Figura 65 - Gráficos fatoriais dos níveis MAE-T e MAE-V duas camadas in-natura



Fonte: o autor, 2022.

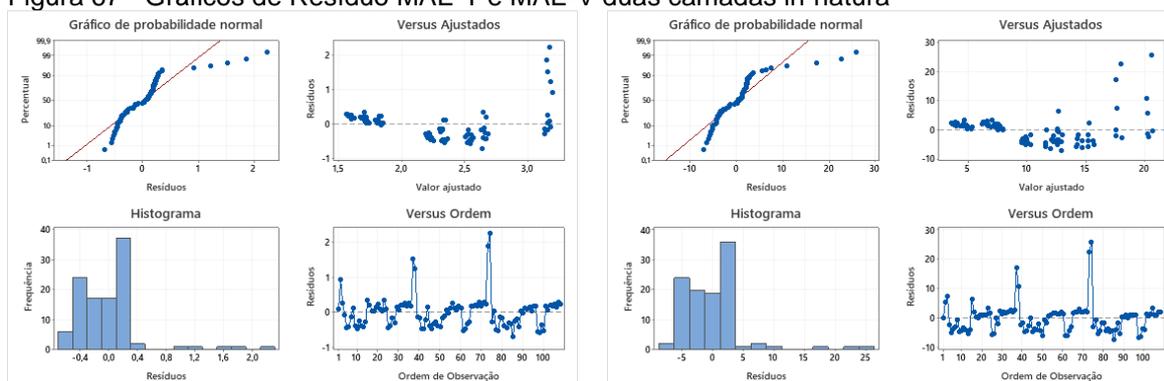
Figura 66 - Gráficos fatoriais dos níveis  $R^2$  e  $R_a^2$  duas camadas in-natura



Fonte: o autor, 2022.

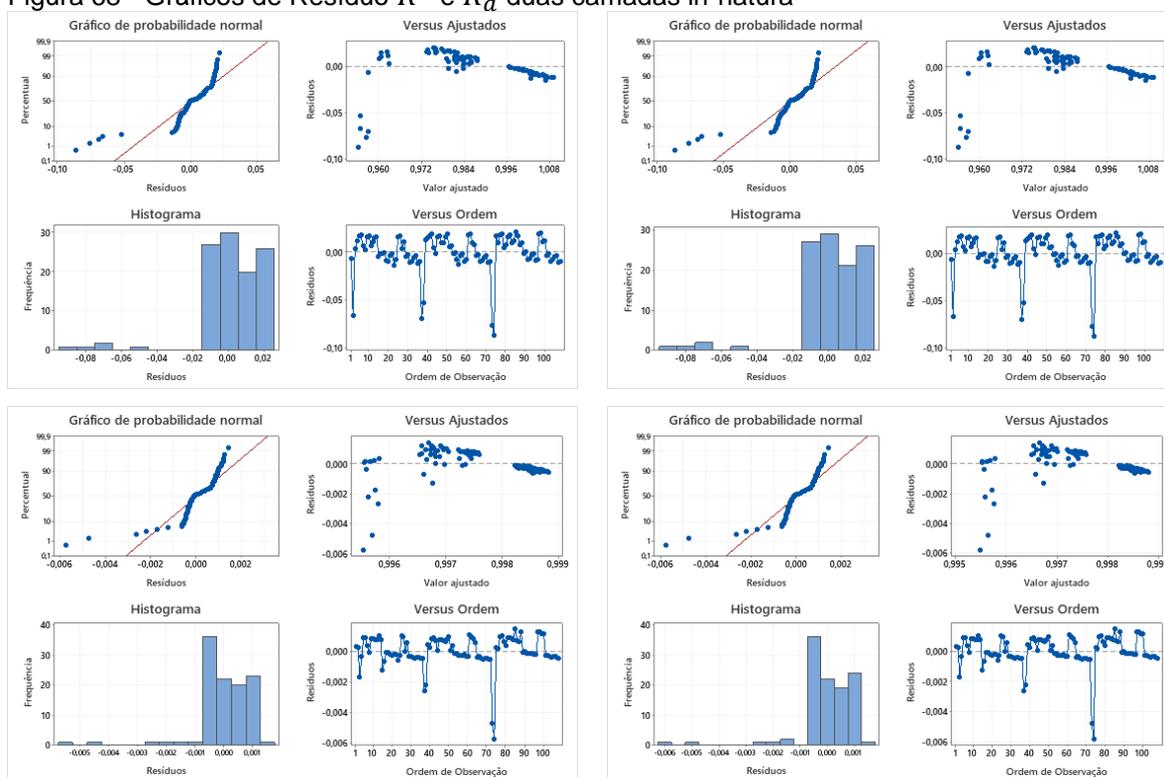
As figuras 67 e 68, apresentam os gráficos de Resíduos para os níveis MAE-T, MAE-V,  $R^2$  e  $R_a^2$ , é verificado também, o gráfico de probabilidade normal dos resíduos, mostrando que os dados não seguem uma distribuição normal. Para o gráfico de resíduos versus ajustados, é observado a pressuposição que os resíduos são aleatoriamente distribuídos e com variância constante e também é verificado a presença de outliers. No gráfico de Resíduos versus ordem verifica-se que há independência entre os resíduos.

Figura 67 - Gráficos de Resíduo MAE-T e MAE-V duas camadas in-natura



Fonte: o autor, 2022.

Figura 68 - Gráficos de Resíduo  $R^2$  e  $R_a^2$  duas camadas in-natura



Fonte: o autor, 2022.

Na tabela 30, são apresentados os fatores, os níveis utilizados, as respostas MAE-T, MAE-V,  $R^2$  e  $R_a^2$  e o tempo de CPU utilizados no planejamento experimental da RNA para o tratamento alcalino.

Tabela 30 - Planejamento do tratamento alcalino em duas camadas

(continua)

Exp	NN	NN	NE	FA	DIV (%)	MAE(T)	MSE(T)	$R^2(Y1)$	$R_a^2(Y1)$	$R^2(Y2)$	$R_a^2(Y2)$	Tempo(s)
1	30	30	2000	Sigmoid	30	9,1129	161,4186	0,8283	0,8245	0,9501	0,9490	462,17
2	30	30	2000	Sigmoid	40	8,3823	142,4987	0,8283	0,8245	0,9501	0,9490	461,74
3	30	30	2000	Relu	30	4,0936	36,9402	0,9393	0,9380	0,9874	0,9871	459,68
4	30	30	2000	Relu	40	3,9469	30,7024	0,9350	0,9335	0,9886	0,9884	460,78
5	30	30	4000	Sigmoid	30	4,9524	63,9051	0,8686	0,8657	0,9789	0,9784	913,16
6	30	30	4000	Sigmoid	40	5,9056	87,4187	0,8599	0,8567	0,9653	0,9645	914,48
7	30	30	4000	Relu	30	2,1699	11,6707	0,9631	0,9623	0,9922	0,9920	912,72
8	30	30	4000	Relu	40	1,4758	6,0585	0,9449	0,9437	0,9825	0,9821	914,53
9	30	30	8000	Sigmoid	30	0,2997	0,5231	0,9912	0,9910	0,9972	0,9972	1818,72
10	30	30	8000	Sigmoid	40	0,2830	0,4684	0,9682	0,9675	0,9879	0,9877	1830,13
11	30	30	8000	Relu	30	0,9324	2,2850	0,9486	0,9475	0,9846	0,9843	1818,60
12	30	30	8000	Relu	40	0,8187	1,9346	0,9419	0,9407	0,9809	0,9804	1820,11
13	30	60	2000	Sigmoid	30	8,6363	151,5415	0,8283	0,8245	0,9501	0,9490	460,55
14	30	60	2000	Sigmoid	40	6,9110	101,4546	0,8502	0,8469	0,9593	0,9584	462,47
15	30	60	2000	Relu	30	3,8421	30,7771	0,9471	0,9459	0,9871	0,9868	460,44
16	30	60	2000	Relu	40	3,3047	22,8848	0,9486	0,9474	0,9893	0,9891	461,43
17	30	60	4000	Sigmoid	30	1,4297	4,6549	0,9729	0,9723	0,9966	0,9966	914,09
18	30	60	4000	Sigmoid	40	0,6759	1,92606	0,9776	0,9771	0,9896	0,9894	917,72
19	30	60	4000	Relu	30	1,5961	5,4024	0,9767	0,9762	0,9936	0,9935	914,35
20	30	60	4000	Relu	40	1,2712	3,7834	0,9530	0,9520	0,9878	0,9875	916,70
21	30	60	8000	Sigmoid	30	0,1868	0,3283	0,9876	0,9873	0,9959	0,9958	1820,68
22	30	60	8000	Sigmoid	40	0,2412	0,4113	0,9758	0,9753	0,9866	0,9863	1826,41
23	30	60	8000	Relu	30	0,8669	2,4736	0,9654	0,9646	0,9930	0,9928	1822,32
24	30	60	8000	Relu	40	0,2532	0,5512	0,9464	0,9452	0,9880	0,9877	1822,52
25	30	90	2000	Sigmoid	30	6,4833	93,1435	0,8660	0,8630	0,9700	0,9693	462,26
26	30	90	2000	Sigmoid	40	4,7789	50,5884	0,8709	0,8681	0,9763	0,9758	462,79
27	30	90	2000	Relu	30	2,7360	17,1506	0,9569	0,9559	0,9890	0,9887	460,80
28	30	90	2000	Relu	40	2,2253	10,9201	0,9527	0,9517	0,9906	0,9904	461,34
29	30	90	4000	Sigmoid	30	0,8161	1,6412	0,9893	0,9890	0,9961	0,9960	916,79
30	30	90	4000	Sigmoid	40	0,2987	0,5341	0,9781	0,9776	0,9897	0,9894	918,66
31	30	90	4000	Relu	30	1,7603	7,966	0,9626	0,9618	0,9889	0,9886	914,53
32	30	90	4000	Relu	40	0,6328	1,0895	0,9562	0,9552	0,9859	0,9856	916,52
33	30	90	8000	Sigmoid	30	0,2680	0,5421	0,9894	0,9892	0,9972	0,9971	1822,69
34	30	90	8000	Sigmoid	40	0,1504	0,3720	0,9800	0,9795	0,9907	0,9905	1822,35
35	30	90	8000	Relu	30	0,2767	0,4003	0,9792	0,9787	0,9929	0,9928	1818,01
36	30	90	8000	Relu	40	0,6342	1,4146	0,9476	0,9465	0,9864	0,9861	1819,52

(...)

Tabela 30 - Planejamento do tratamento alcalino em duas camadas

(continuação)

Exp	NN	NN	NE	FA	DIV (%)	MAE(T)	MSE(T)	$R^2(Y1)$	$R_a^2(Y1)$	$R^2(Y2)$	$R_a^2(Y2)$	Tempo(s)
37	60	30	2000	Sigmoid	30	9,6292	177,6242	0,8283	0,8245	0,9501	0,9490	462,46
38	60	30	2000	Sigmoid	40	10,1060	200,3999	0,8283	0,8245	0,9501	0,9490	463,26
39	60	30	2000	Relu	30	3,0543	19,4944	0,9458	0,9446	0,9912	0,9910	459,89
40	60	30	2000	Relu	40	3,2807	23,9950	0,9352	0,9338	0,9866	0,9863	461,39
41	60	30	4000	Sigmoid	30	7,3687	123,7118	0,8516	0,8484	0,9591	0,9582	914,33
42	60	30	4000	Sigmoid	40	7,0439	113,6878	0,8401	0,8365	0,9580	0,9570	917,39
43	60	30	4000	Relu	30	1,2207	3,5728	0,9660	0,9652	0,9868	0,9865	916,13
44	60	30	4000	Relu	40	1,1895	4,3593	0,9524	0,9514	0,9873	0,9870	914,36
45	60	30	8000	Sigmoid	30	0,3134	0,4378	0,9900	0,9897	0,9961	0,9960	1820,44
46	60	30	8000	Sigmoid	40	0,2236	0,4099	0,9778	0,9773	0,9872	0,9869	1826,59
47	60	30	8000	Relu	30	0,8120	1,9222	0,9743	0,9737	0,9897	0,9895	1819,65
48	60	30	8000	Relu	40	0,7835	1,7199	0,9432	0,9419	0,9817	0,9813	1822,95
49	60	60	2000	Sigmoid	30	8,6373	151,5605	0,8283	0,8245	0,9501	0,9490	461,157
50	60	60	2000	Sigmoid	40	8,1585	138,0483	0,8283	0,8245	0,9501	0,9490	462,77
51	60	60	2000	Relu	30	2,7956	17,7272	0,9576	0,9566	0,9892	0,9890	460,88
52	60	60	2000	Relu	40	2,7618	16,3469	0,9530	0,9520	0,9873	0,9870	461,82
53	60	60	4000	Sigmoid	30	1,7949	11,848	0,9550	0,9540	0,9941	0,9940	917,24
54	60	60	4000	Sigmoid	40	0,3131	0,5070	0,9845	0,9841	0,9878	0,9875	918,19
55	60	60	4000	Relu	30	0,9758	2,2683	0,9595	0,9586	0,9921	0,9920	912,83
56	60	60	4000	Relu	40	0,8206	2,0572	0,9460	0,9448	0,9848	0,9845	916,15
57	60	60	8000	Sigmoid	30	0,3033	0,4694	0,9913	0,9911	0,9953	0,9952	1823,10
58	60	60	8000	Sigmoid	40	0,1723	0,3783	0,9850	0,9847	0,9891	0,9888	1829,14
59	60	60	8000	Relu	30	0,5832	1,2365	0,9660	0,9652	0,9915	0,9913	1828,62
60	60	60	8000	Relu	40	0,4716	0,8874	0,9279	0,9263	0,9859	0,9856	1827,36
61	60	90	2000	Sigmoid	30	7,4966	124,85	0,8425	0,8390	0,9588	0,9579	463,57
62	60	90	2000	Sigmoid	40	5,0013	62,753	0,8661	0,8632	0,9747	0,9741	464,55
63	60	90	2000	Relu	30	1,8976	8,0203	0,9667	0,9659	0,9921	0,9919	462,43
64	60	90	2000	Relu	40	1,9259	9,4866	0,9504	0,9493	0,9893	0,9891	464,10
65	60	90	4000	Sigmoid	30	0,6249	1,1543	0,9893	0,9891	0,9958	0,9957	916,06
66	60	90	4000	Sigmoid	40	0,3528	0,5708	0,9886	0,9883	0,9896	0,9893	927,52
67	60	90	4000	Relu	30	0,7979	1,5190	0,9587	0,9578	0,9897	0,9895	917,72
68	60	90	4000	Relu	40	0,6152	1,0781	0,9474	0,9462	0,9901	0,9899	916,37
69	60	90	8000	Sigmoid	30	0,1378	0,3150	0,9903	0,9900	0,9945	0,9944	1822,91
70	60	90	8000	Sigmoid	40	0,1441	0,3718	0,9867	0,9864	0,9885	0,9882	1830,32
71	60	90	8000	Relu	30	0,3143	0,4042	0,9641	0,9633	0,9911	0,9909	1821,39
72	60	90	8000	Relu	40	0,3461	0,5081	0,9212	0,9195	0,9855	0,9852	1827,86

(...)

Tabela 30 - Planejamento do tratamento alcalino em duas camadas

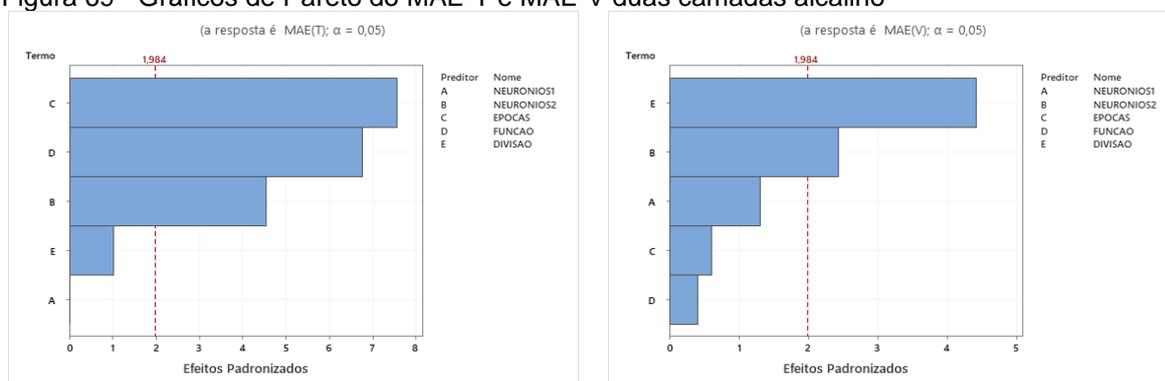
Exp	NN	NN	NE	FA	DIV (%)	MAE(T)	MSE(T)	$R^2(Y1)$	$R_a^2(Y1)$	$R^2(Y2)$	$R_a^2(Y2)$	(conclusão)
												Tempo(s)
73	90	30	2000	Sigmoid	30	11,7212	273,3380	0,8283	0,8245	0,9501	0,9490	482,67
74	90	30	2000	Sigmoid	40	9,5442	178,0524	0,8283	0,8245	0,9501	0,9490	462,64
75	90	30	2000	Relu	30	2,8994	18,6866	0,9523	0,9513	0,9910	0,9908	459,64
76	90	30	2000	Relu	40	2,9266	18,3566	0,9413	0,9400	0,9895	0,9892	462,01
77	90	30	4000	Sigmoid	30	8,6374	151,560	0,8283	0,8245	0,9501	0,9490	912,50
78	90	30	4000	Sigmoid	40	6,3630	99,8261	0,8666	0,8636	0,9612	0,9604	916,58
79	90	30	4000	Relu	30	0,8089	1,5833	0,9674	0,9666	0,9898	0,9895	911,09
80	90	30	4000	Relu	40	1,1583	3,0501	0,9181	0,9163	0,9834	0,9830	916,42
81	90	30	8000	Sigmoid	30	0,2139	0,3455	0,9924	0,9922	0,9967	0,9967	1818,18
82	90	30	8000	Sigmoid	40	0,2378	0,4296	0,9828	0,9824	0,9868	0,9865	1821,87
83	90	30	8000	Relu	30	0,9260	2,4146	0,9708	0,9701	0,9885	0,9883	1811,63
84	90	30	8000	Relu	40	0,5141	0,7813	0,9353	0,9339	0,9813	0,9809	1820,11
85	90	60	2000	Sigmoid	30	7,5218	120,5920	0,8445	0,8410	0,9610	0,9601	460,96
86	90	60	2000	Sigmoid	40	8,1595	138,0462	0,8283	0,8245	0,9501	0,9490	462,65
87	90	60	2000	Relu	30	2,6076	15,5733	0,9583	0,9574	0,9911	0,9909	460,09
88	90	60	2000	Relu	40	2,6206	14,5947	0,9294	0,9278	0,9837	0,9834	460,48
89	90	60	4000	Sigmoid	30	1,3231	4,2946	0,9833	0,9829	0,9965	0,9965	915,09
90	90	60	4000	Sigmoid	40	1,1887	4,5719	0,9627	0,9619	0,9874	0,9871	923,97
91	90	60	4000	Relu	30	0,8372	1,6550	0,9764	0,9758	0,9930	0,9929	913,76
92	90	60	4000	Relu	40	0,7701	1,3673	0,9451	0,9439	0,9845	0,9841	916,89
93	90	60	8000	Sigmoid	30	0,1469	0,3164	0,9934	0,9933	0,9965	0,9964	1823,66
94	90	60	8000	Sigmoid	40	0,1451	0,3718	0,9871	0,9868	0,9862	0,9859	1828,06
95	90	60	8000	Relu	30	0,7472	1,5385	0,9646	0,9638	0,9894	0,9891	1824,99
96	90	60	8000	Relu	40	0,1959	0,3811	0,9471	0,9460	0,9900	0,9897	1823,48
97	90	90	2000	Sigmoid	30	7,8822	129,6836	0,8303	0,8265	0,9572	0,9563	461,68
98	90	90	2000	Sigmoid	40	8,1560	138,0436	0,8283	0,8245	0,9501	0,9490	464,16
99	90	90	2000	Relu	30	1,6228	5,5504	0,9673	0,9665	0,9919	0,9917	461,60
100	90	90	2000	Relu	40	1,4787	4,7347	0,9581	0,9572	0,9901	0,9898	464,94
101	90	90	4000	Sigmoid	30	0,5612	0,9391	0,9923	0,9922	0,9961	0,9961	918,38
102	90	90	4000	Sigmoid	40	0,2402	0,4363	0,9874	0,9871	0,9870	0,9868	918,38
103	90	90	4000	Relu	30	0,6627	1,0040	0,9817	0,9813	0,9925	0,9923	915,736
104	90	90	4000	Relu	40	0,5714	0,8387	0,9487	0,9476	0,9843	0,9840	918,545
105	90	90	8000	Sigmoid	30	0,2188	0,3457	0,9918	0,9916	0,9956	0,9955	1826,66
106	90	90	8000	Sigmoid	40	0,1544	0,3728	0,9872	0,9870	0,9897	0,9894	1852,39
107	90	90	8000	Relu	30	0,4755	0,7033	0,9553	0,9543	0,9928	0,9926	1821,81
108	90	90	8000	Relu	40	0,4641	0,8616	0,9386	0,9373	0,9834	0,9830	1824,66

Fonte: o autor, 2022.

Nas figuras 69 e 70, são apresentados os gráficos de Pareto para os seguintes níveis: MAE-T, MAE-V,  $R^2$  e  $R_a^2$ . Para o nível MAE-T pode ser observado que os fatores épocas, a função de ativação e o número de neurônios da segunda camada apresentaram efeitos estatisticamente significativo para o ajuste do modelo, já a divisão para o conjunto de dados e os neurônios da primeira camada não são importantes. Já para o nível MAE-V pode ser observado que apenas a divisão para o conjunto de dados e os neurônios da segunda camada apresentaram efeitos estatisticamente significativos.

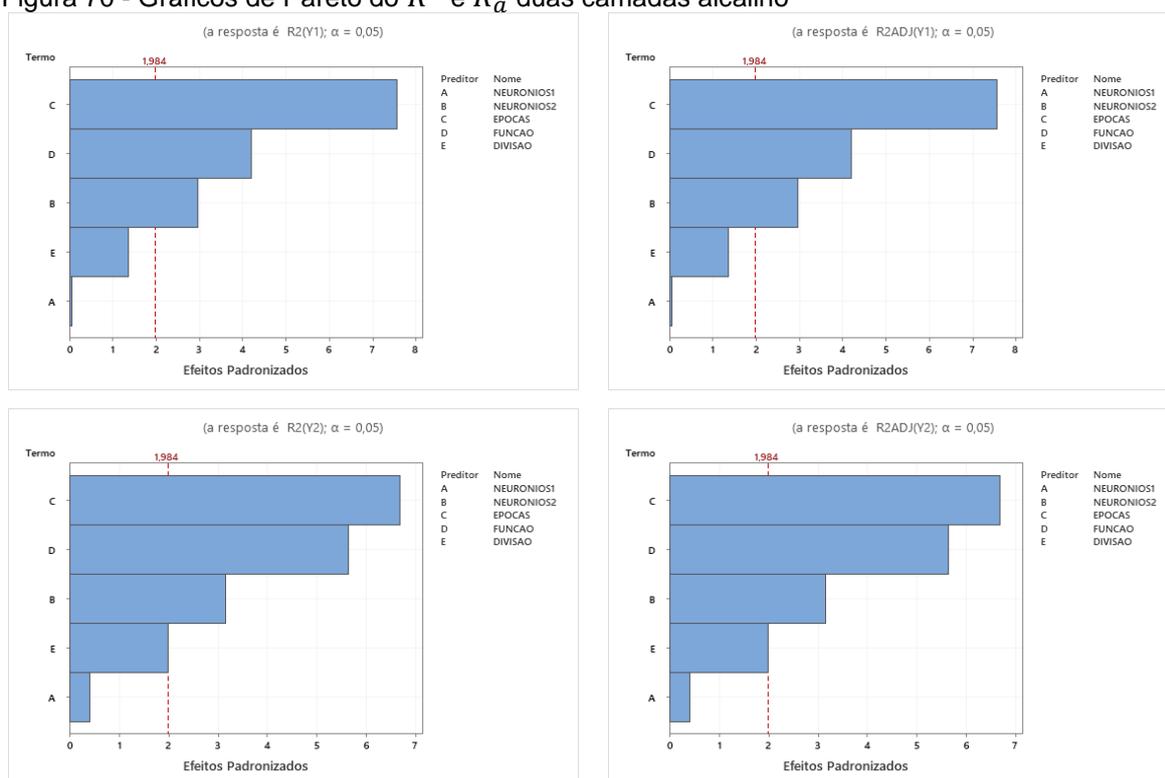
A figura 70, apresenta os fatores com efeitos estatisticamente significativos para os níveis  $R^2$  e  $R_a^2$  para as duas saídas da RNA no tratamento alcalino, os fatores são: épocas, função de ativação e número de neurônios na segunda camada.

Figura 69 - Gráficos de Pareto do MAE-T e MAE-V duas camadas alcalino



Fonte: o autor, 2022.

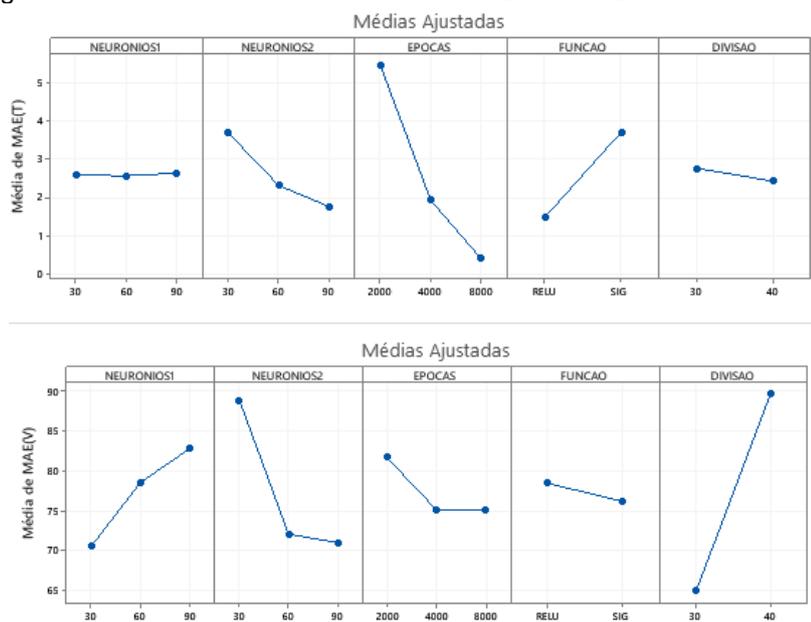
Figura 70 - Gráficos de Pareto do  $R^2$  e  $R_a^2$  duas camadas alcalino



Fonte: o autor, 2022.

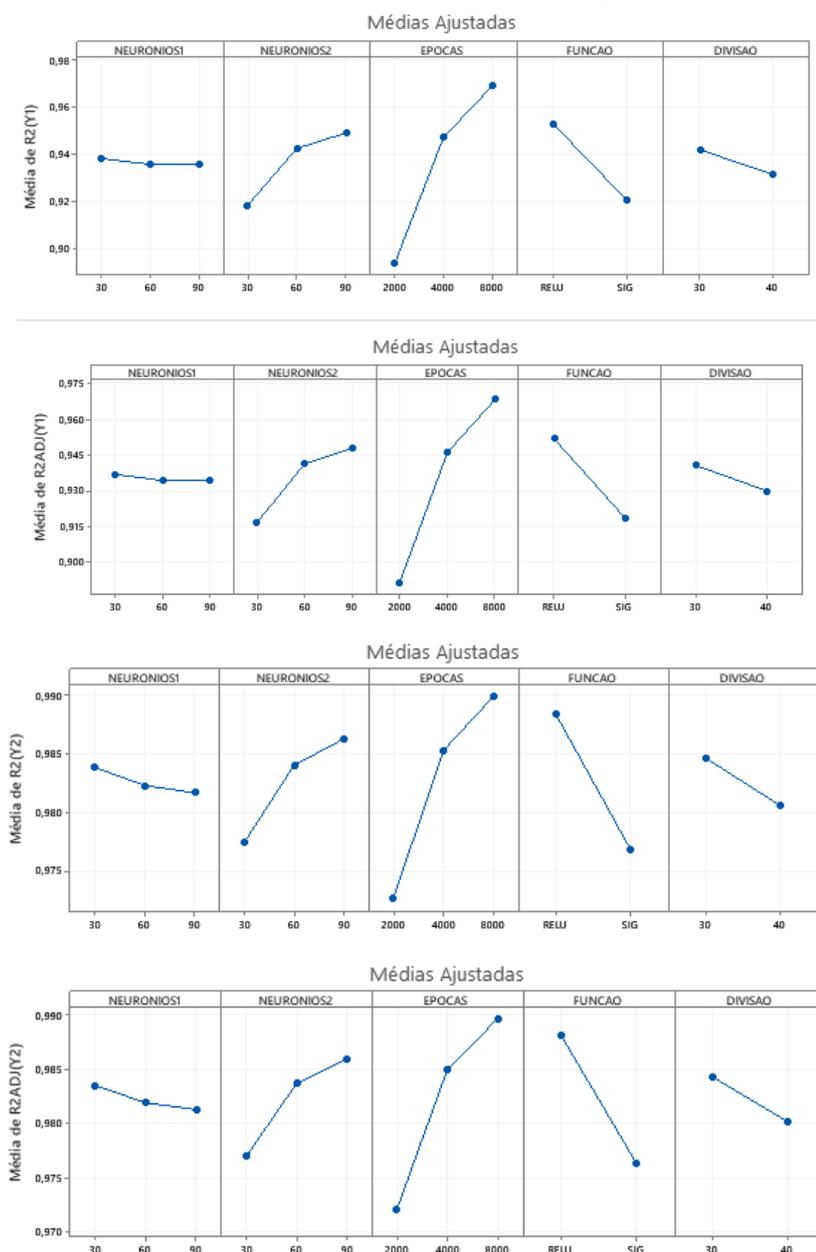
Nas figuras 71 e 72, pode ser verificado para os níveis MAE-T, MAE-V,  $R^2$  e  $R_a^2$ , que o aumento do número de épocas e também do número de neurônios apresentaram o menor erro obtido e também os melhores valores para os coeficientes de determinação.

Figura 71 – Gráficos fatoriais dos níveis MAE-T e MAE-V duas camadas alcalino



Fonte: o autor, 2022.

Figura 72 – Gráficos fatoriais dos níveis  $R^2$  e  $R_a^2$  duas camadas alcalino



Fonte: o autor, 2022.

Nas figuras 73 e 74, podem ser observados os gráficos de Resíduos para MAE-T, MAE-V,  $R^2$  e  $R_a^2$ , também é apresentado o gráfico de probabilidade normal dos resíduos, com distribuição normal, uma vez que a maioria dos pontos estão próximos a linha de referência, o que valida a análise dos dados para o método de monitoramento proposto.

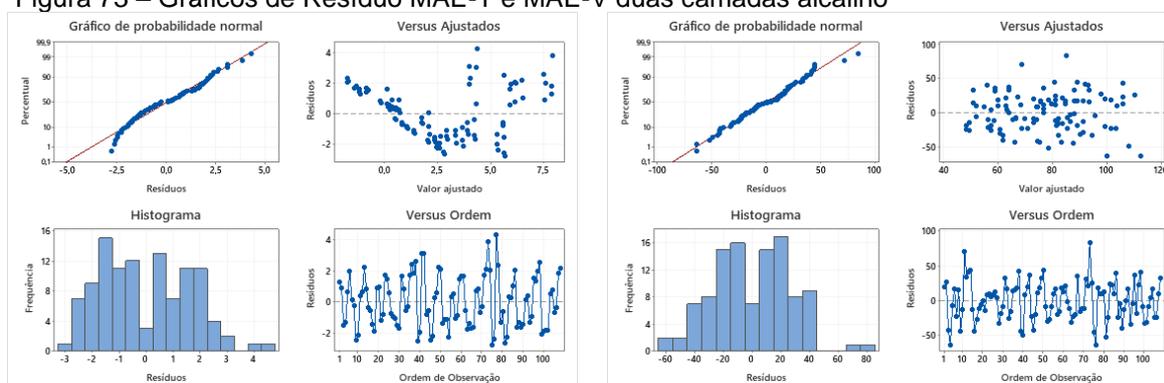
Conforme descrito por Levine (2014), os gráficos de dispersão permitem verificar as relações entre as variáveis.

De acordo com Levine (2014), o gráfico dos resíduos serve de estratégia para verificação dos pressupostos para a regressão (linearidade, independência, normalidade, homocedasticidade).

Para o gráfico de resíduos versus ajustados, pode ser verificado a pressuposição que os resíduos são aleatoriamente distribuídos com variância constante e também pode ser notada a presença de outliers.

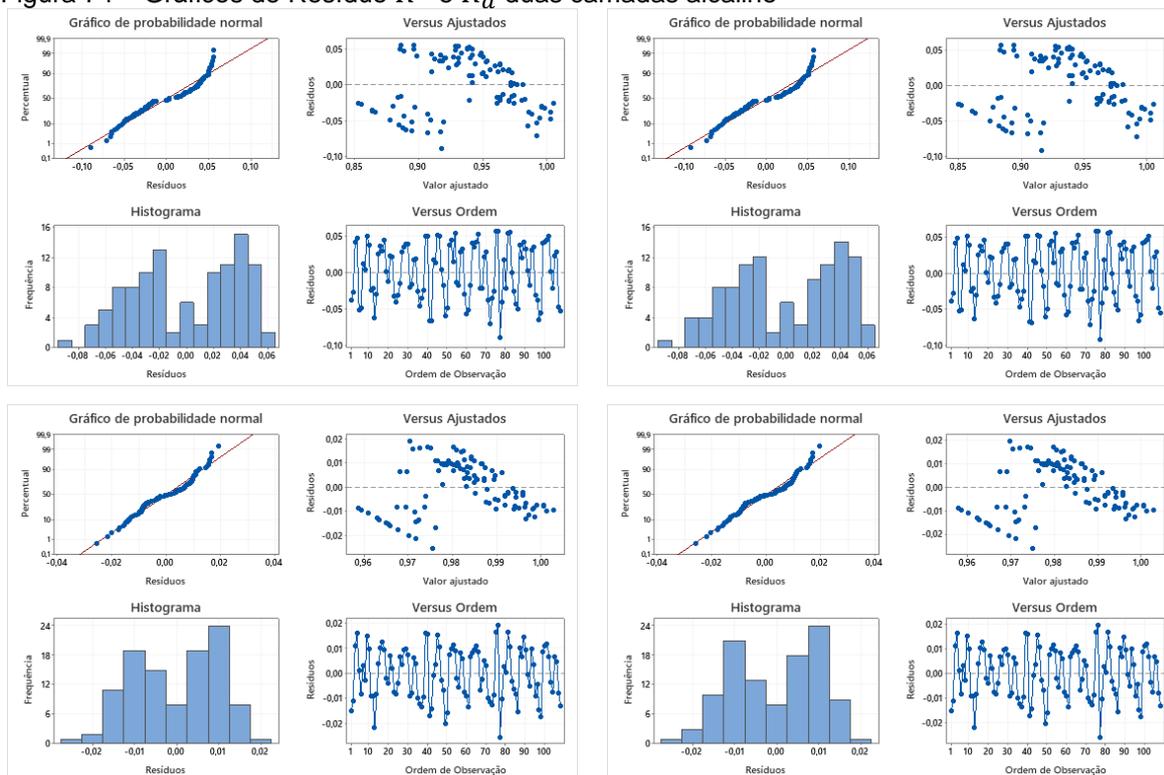
A homocedasticidade significa que a variância é constante. A análise gráfica dos resíduos pelos valores ajustados serve de verificação deste parâmetro em que os pontos devem estar dispostos de forma aleatória. (ABNT, 2011).

Figura 73 – Gráficos de Resíduo MAE-T e MAE-V duas camadas alcalino



Fonte: o autor, 2022.

Figura 74 – Gráficos de Resíduo  $R^2$  e  $R_a^2$  duas camadas alcalino

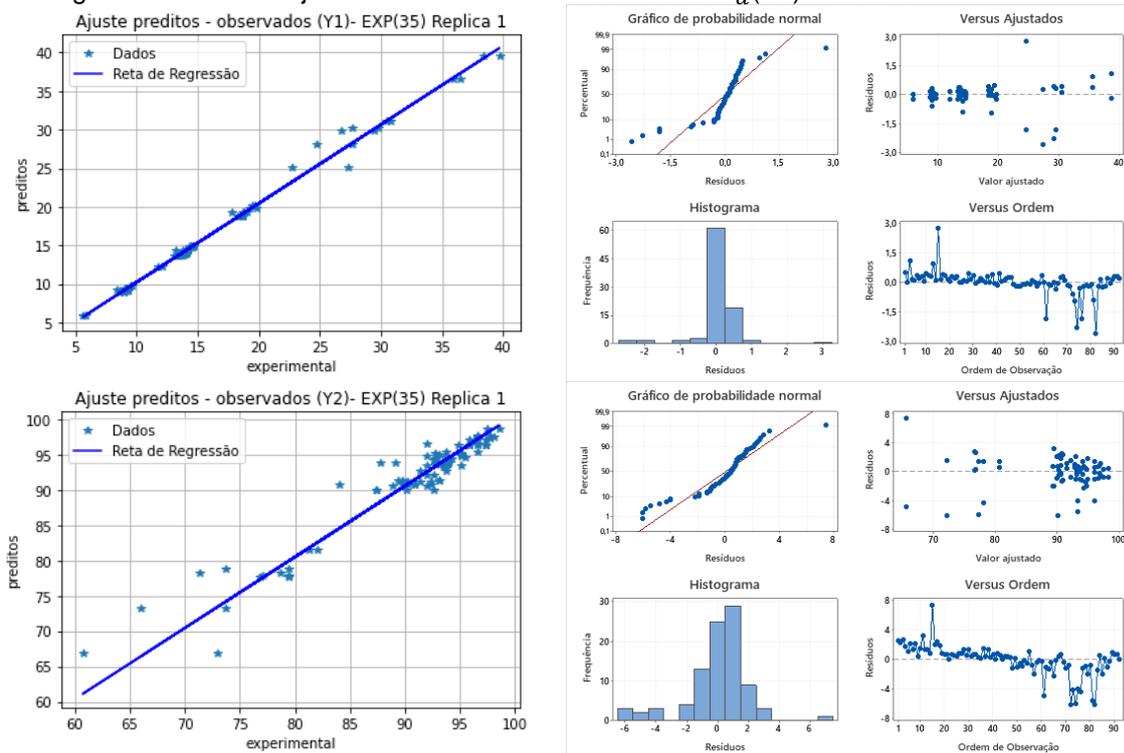


Fonte: o autor, 2022.

### 5.10.3 Otimização para a RNA com duas camadas

Na figura 75, pode ser observado que o experimento 35 apresentou a otimização para o  $R_a^2(Y1)$  e MAE-T, com 30 neurônios na primeira camada, 90 na segunda, 8.000 épocas, função de ativação: Relu e 30% para validação dos dados com um tempo de CPU igual a 828,39 segundos para o tratamento ácido das duas variáveis de saída.

Figura 75 – Gráfico ajuste da RNA e resíduo dos níveis  $R_a^2(Y1)$  e MAE-V.

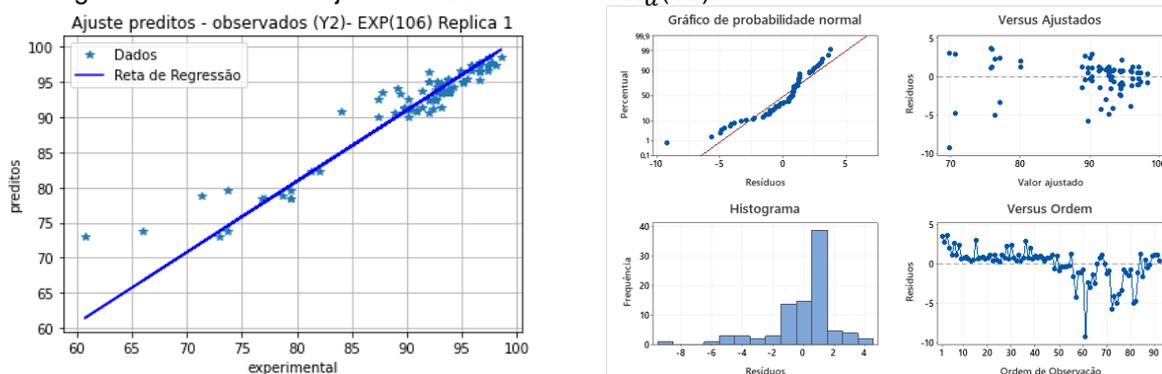


Fonte: o autor, 2022.

No gráfico de probabilidade normal (Figura 75), quanto mais próximos os pontos estiverem da linha reta, mais adequada é a distribuição normal para representar os dados (MONTGOMERY, 2009). Logo, o pressuposto de normalidade dos resíduos foi atendido.

Na figura 76, pode ser observado que o experimento 106 indica o gráfico otimizado do  $R_a^2(Y2)$  para a RNA com duas camadas, 90 neurônios na primeira, 90 na segunda, 8.000 épocas, função de ativação: Sigmóide e 40% para validação dos dados com um tempo de CPU igual a 818,90 segundos para o tratamento ácido.

Figura 76 - Gráfico de ajuste da RNA e do nível  $R_a^2(Y2)$ .

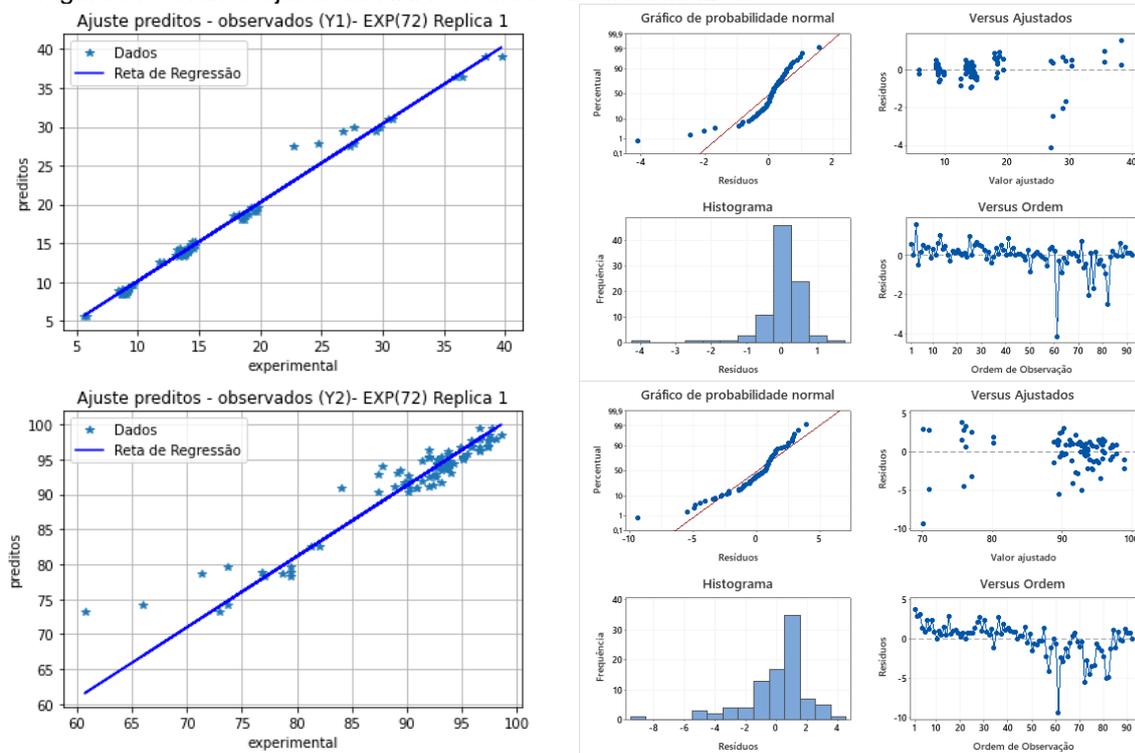


Fonte: o autor, 2022.

Pode ser observado que os pontos experimentais estão próximos a linha contínua, corroborando o afirmado por Calado (2003) que quanto mais próximo os pontos experimentais estiverem da linha contínua, mais será válida a suposição de normalidade dos resíduos. Com relação a figura 76 pode-se dizer que os resíduos seguem uma distribuição normal.

Na figura 77, pode ser observado que o experimento 72 indica o gráfico otimizado do MAE-T para a RNA com duas camadas, 60 neurônios na primeira, 90 na segunda, 8.000 épocas, função de ativação: Relu e 40% para validação dos dados com um tempo de CPU igual a 849,69 segundos para o tratamento ácido.

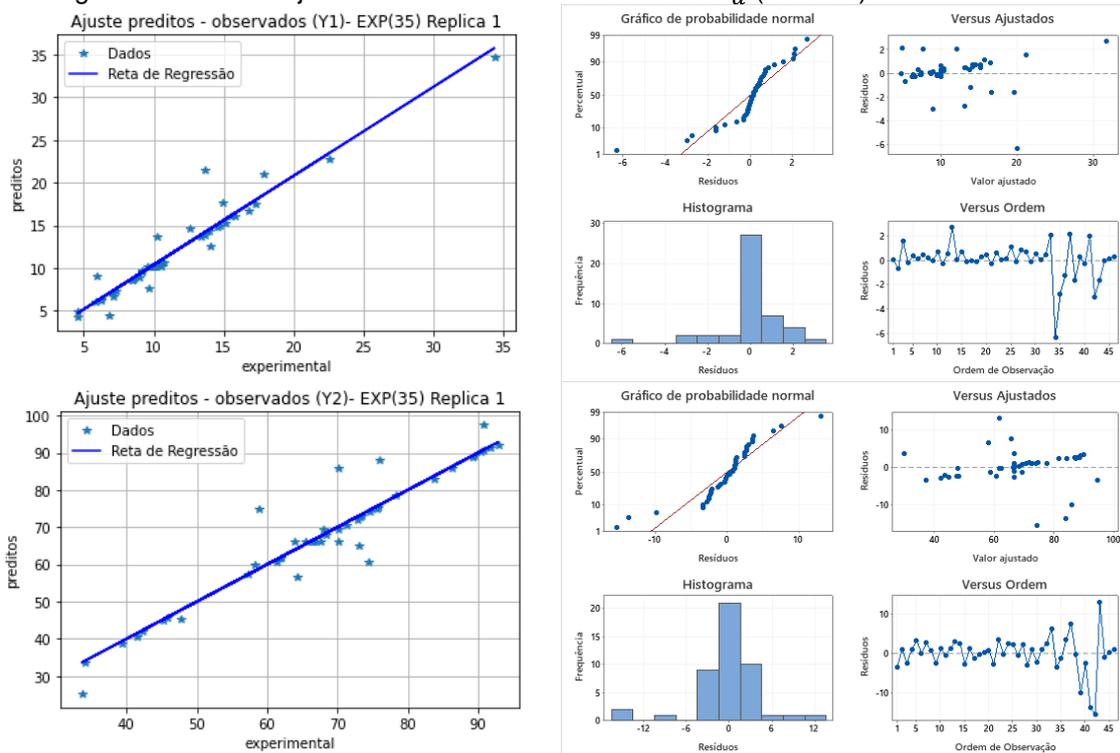
Figura 77- Gráfico ajuste da RNA e resíduo do nível MAE-T.



Fonte: o autor, 2022.

Na figura 78, pode ser verificado que o experimento 35 apresenta o gráfico otimizado de  $R_a^2(Y1)$  e  $R_a^2(Y2)$  com duas camadas, 30 neurônios na primeira e 90 na segunda, 8.000 épocas, função de ativação: Relu e 30% para validação dos dados com um tempo de CPU igual a 788.89 segundos para o tratamento alcalino.

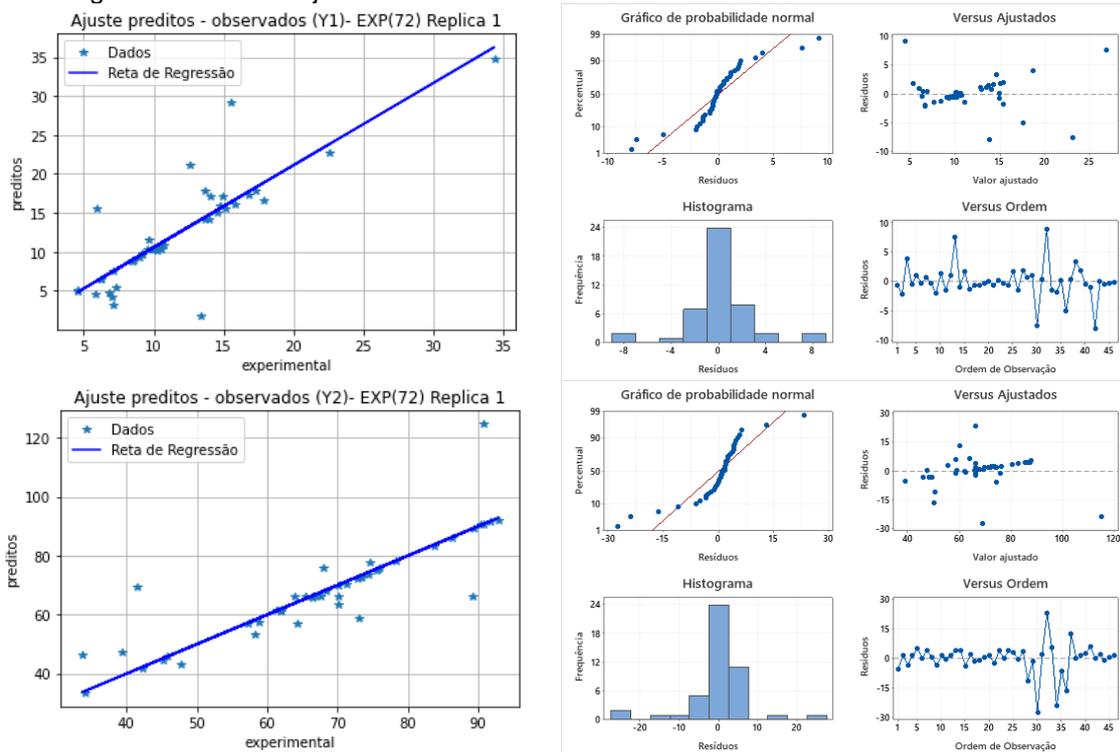
Figura 78 - Gráfico ajuste da RNA e resíduo dos níveis  $R_a^2$  (Y1 e Y2).



Fonte: o autor, 2022.

Na figura 79, pode ser observado que o experimento 72 apresenta a otimização do MAE-T com 60 neurônios na primeira camada, 90 na segunda, 8.000 épocas, função de ativação: Relu e 40% para validação dos dados com um tempo de CPU igual a 784,87 segundos para o tratamento alcalino.

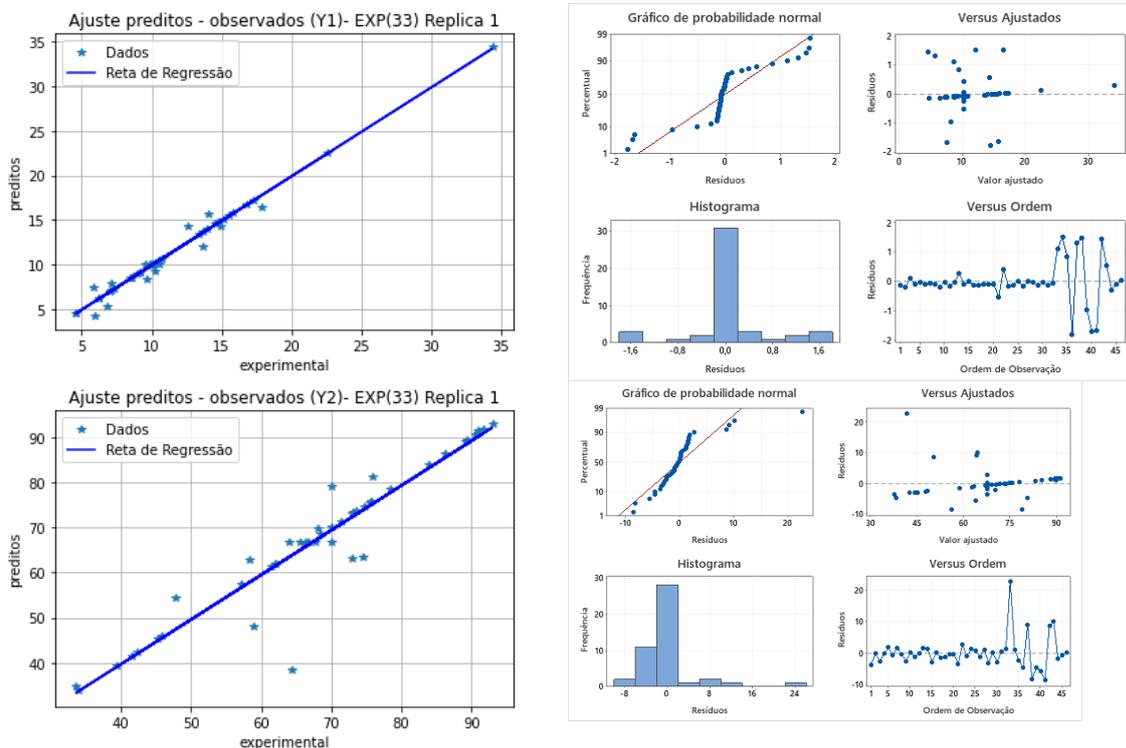
Figura 79 - Gráfico ajuste da RNA e resíduo do nível MAE-T.



Fonte: o autor, 2022.

Na figura 80, pode ser verificado que o experimento 33 discrimina o gráfico de otimização do MAE-T com 30 neurônios na primeira camada, 90 na segunda, 8.000 épocas, função de ativação: Sigmóide e 30% para validação dos dados com um tempo de CPU igual a 772,47 segundos para o tratamento alcalino.

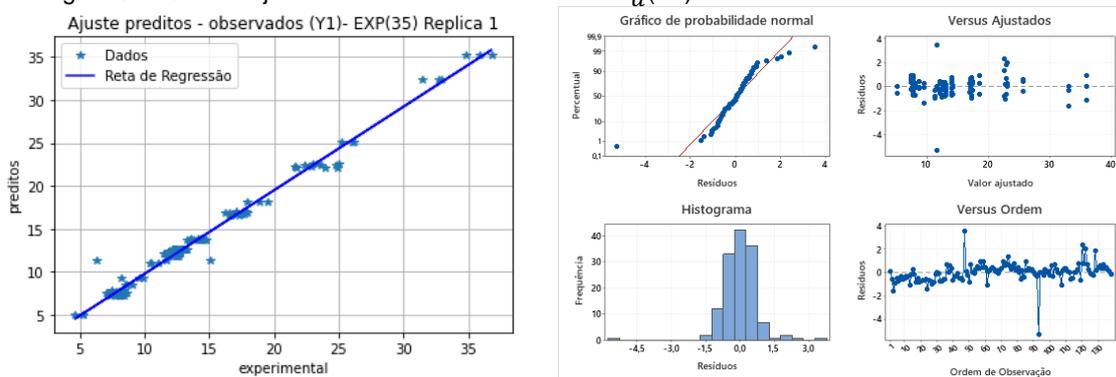
Figura 80 - Gráfico ajuste da RNA e resíduo do nível MAE-V.



Fonte: o autor, 2022.

Na figura 81, pode ser observado que o experimento 35 apresenta o gráfico otimizado do  $R_a^2(Y1)$  com 30 neurônios na primeira camada, 90 na segunda, 8.000 épocas, função de ativação: Relu e 30% para validação dos dados com um tempo de CPU igual a 861,61 segundos para o tratamento in-natura.

Figura 81 - Gráfico ajuste da RNA e resíduo do nível  $R_a^2(Y1)$ .

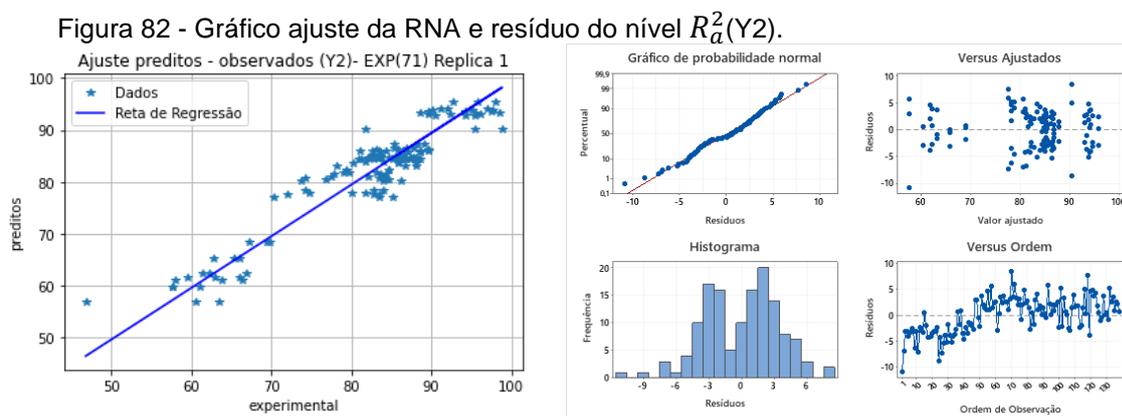


Fonte: o autor, 2022.

Na figura 82 pode ser observado para o experimento 71 aponta o gráfico otimizado do  $R_a^2(Y2)$  com 60 neurônios na primeira camada, 90 na segunda, 8.000 épocas, função de ativação: Relu e 30% para validação dos dados com um tempo de

CPU igual a 840,58 segundos para o tratamento in-natura. Para o  $R_a^2(Y2)$  pressupõe-se que há normalidade.

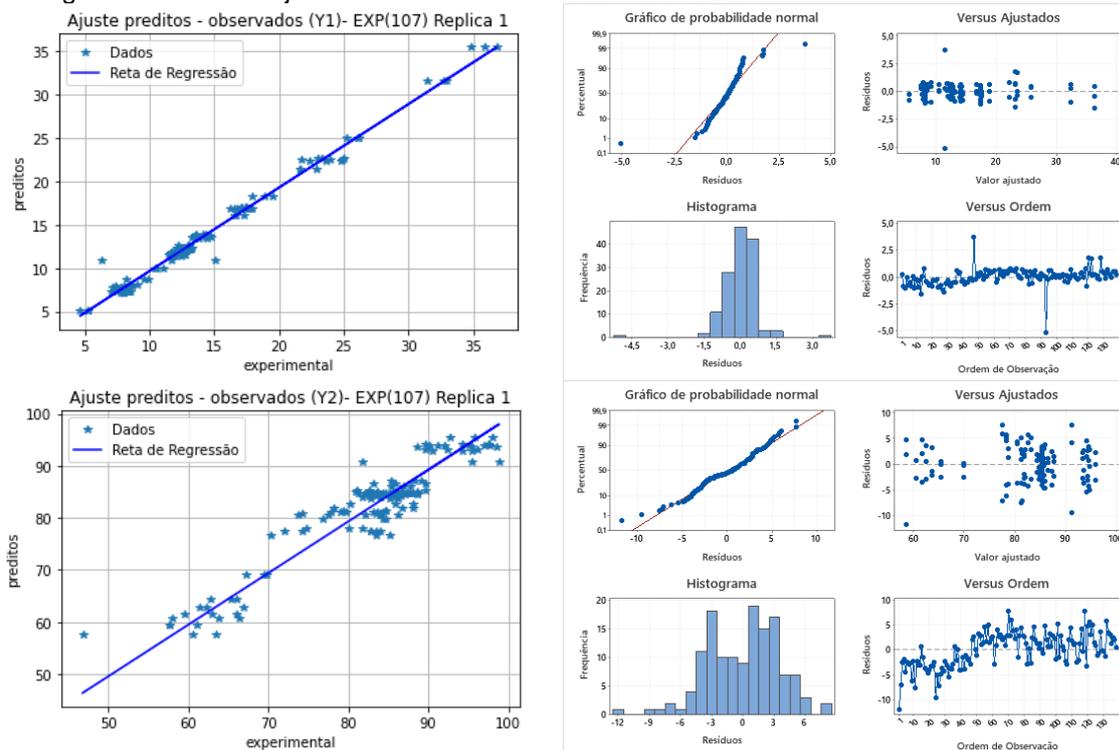
Faraway (2016) sugere a análise da homocedasticidade pelo gráfico de resíduos por valores ajustados.



Fonte: o autor, 2022.

Na figura 83, pode ser observado para o experimento 107 indica o gráfico de otimização do MAE-T com 60 neurônios na primeira camada, 90 na segunda, 8.000 épocas, função de ativação: Relu e 30% para validação dos dados com um tempo de CPU igual a 841,66 segundos para o tratamento in-natura. E a pressuposição de normalidade.

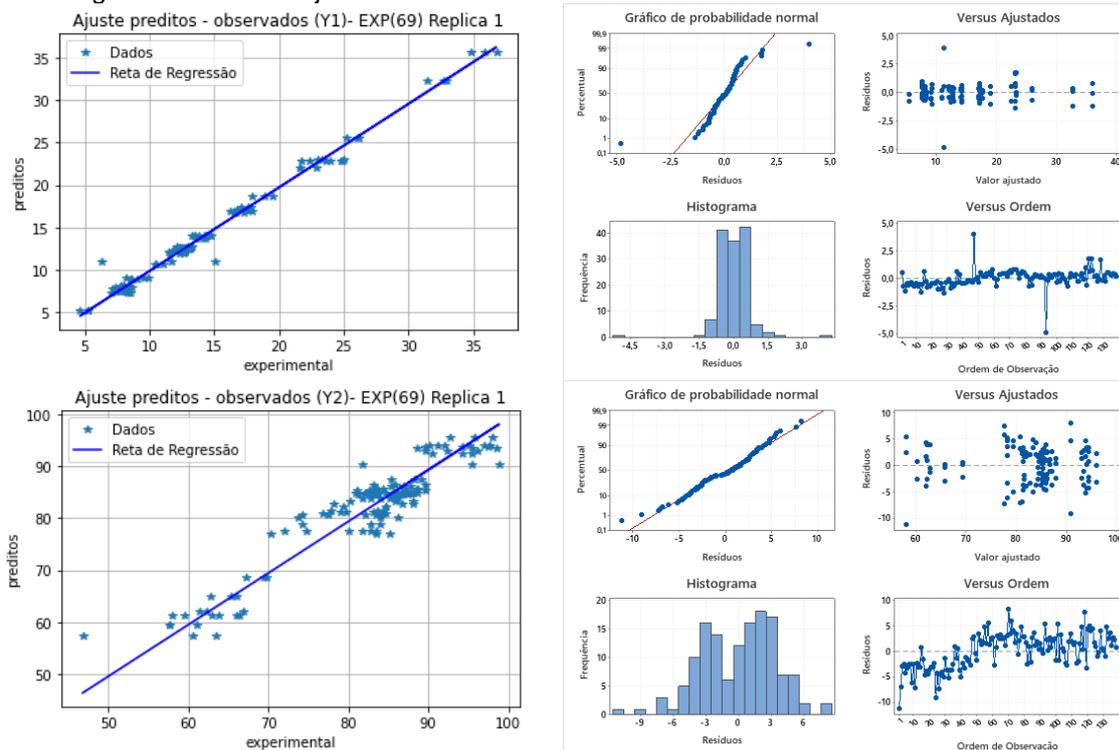
Figura 83 - Gráfico ajuste da RNA e resíduo do nível MAE-T.



Fonte: o autor, 2022.

Na figura 84, pode ser observado que o experimento 69 que apresenta o melhor valor de MAE-V, com 60 neurônios na primeira camada, 90 na segunda, 8.000 épocas, função de ativação: Sigmóide e 30% para validação dos dados com um tempo de CPU igual a 851,81 segundos para o tratamento in-natura. Nos gráficos de probabilidade normal, pode ser observado que os pontos estão bem próximos a reta, indicando que os dados seguem uma proximidade a distribuição normal, logo, o pressuposto de normalidade foi atendido, em relação a distribuição de resíduos em função dos valores ajustados observa-se visualmente que a pressuposição de homocedasticidade é atendida. Os resíduos estão dispersos aleatoriamente em torno do zero (MONTGOMERY, 2009).

Figura 84 - Gráfico ajuste da RNA e resíduo do nível MAE-V.



Fonte: o autor, 2022.

A presença destes pontos destoantes ou outliers é relacionada por Sales *et al.*, (2018) a erros do operador na obtenção dos dados em levantamentos in situ ou erros de leitura dos equipamentos utilizados nos levantamentos como podem ser observados nas figuras 83 e 84. Entretanto, para este estudo, não houve retirada de pontos considerados inconsistentes, para que a interpretação do modelo não fosse prejudicada. Faraway (2016) adiciona que modelos com efeitos fixos são particularmente mais sensíveis a pontos destoantes ou outliers, de maneira que este tipo de gráfico é uma boa ferramenta para identificá-los.

## 6 CONCLUSÕES

Nesta dissertação foram aplicadas a técnica RLM para prever o ajuste de um modelo para os dados de adsorção, como também a aplicação das RNA para predição do comportamento dos dados verificando a eficiência do processo de adsorção e também foi realizado um estudo comparativo entre a modelagem matemática usando RLM e RNA.

A aplicação da RLM se mostrou satisfatória para a unidade de massa de adsorvente obtendo um coeficiente de determinação com média 85% da variação do modelo, por outro lado para a eficiência de remoção o coeficiente obtido explica aproximadamente 50% da variação do ajuste, logo, a Análise de Regressão não é interessante para o ajuste das duas variáveis de saída.

Na utilização da RNA com uma camada foram obtidos coeficientes de determinação que explicam mais de 98% da variação do ajuste do modelo, tanto para a unidade de massa de adsorvente quanto para a eficiência de remoção, em todos os tratamentos aplicados, realizando a otimização foi possível identificar qual tratamento apresentou o melhor coeficiente de determinação para as variáveis de saída, logo a RNA com uma camada foi satisfatória para o ajuste do modelo.

Ao fazer o estudo comparativo entre a modelagem matemática por meio da RLM e da RNA pode ser observado que a RNA é mais apropriada para a predição dos dados das variáveis de saída unidade de massa de adsorvente e eficiência de remoção quanto para verificar a eficiência do processo de adsorção uma vez que estas variáveis apresentaram melhores resultados quando da utilização da RNA em relação a RLM.

## REFERÊNCIAS BIBLIOGRÁFICAS

\_\_\_\_\_. NBR 14653-2. Avaliação de bens - parte 2: imóveis urbanos. Rio de Janeiro: 2011.

ABREU, C. S. A.; SILVA, J. C. L.; OLIVEIRA B. C.; HOLANDA, F. L. Perfis estratégicos de conduta social e ambiental: estudos na indústria têxtil nordestina. *Gestão de Produção*, v. 15, n. 1, p. 159-172, 2008.

AGGARWAL, C. C. **Neural Networks and Deep Learning: A Textbook**. USA: Springer, 2018.

AGHBASHLO, M., MOBLI, H., RAFIEE, S., & MADADLOU, A. (2012). The use of artificial neural network to predict exergetic performance of spray drying process: A preliminary study. *Computers and Electronics in Agriculture*, 88, 32–43. doi:10.1016/j.compag.2012.06.007

AHMED, M., 2017. Application of raw and activated *Phragmites australis* as potential adsorbents for wastewater treatments. *Ecological Engineering*, 102, 262-269.

AKAIKE, H. A new look at the statistical model identification. **IEEE Transactions on Automatic Control**, v.19, p.716–723, 1974.

AKPA, O. M.; UNUABONAH, E. I. Small-Sample Corrected Akaike Information Criterion: An appropriate statistical tool for ranking of adsorption isotherm models. *Desalination*, v. 272, n. 1–3, p. 20–26, 2011.

ARAM, Z. *et al.* **Using chaotic artificial neural networks to model memory in the brain. Commun Nonlinear Sci Numer Simulat**, v. 44, p. 449-459, 2017.

ARAUJO, F. V. F.; YOKOYAMA, L. Remoção de cor em soluções de corantes reativos por oxidação com H<sub>2</sub>O<sub>2</sub>/UV. *Revista Química Nova*, v. 29, n. 1, p. 11-14, 2006.

ARRIEL, L. F. P. *et al.* Adsorção do corante reativo azul BF-5G em bagaço de malte bruto e lavado. **Brazilian Journal of Development**, v. 5, n. 10, p. 18496-18503, 2019.

ASSEFI, P. *et al.* **Journal of Industrial and Engineering Chemistry Artificial neural network optimization for removal of hazardous dye Eosin Y from aqueous solution using Co<sub>2</sub>O<sub>3</sub>-NP-AC: Isotherm and kinetics study.** *Journal of industrial and Engineering Chemistry*, 2013.

AVINASH, V. S. Understanding activation functions in neural networks (Machine Learning Mastery). 2017. Disponível em: <https://medium.com/the-theory-of-everything/understanding-activation-functions-in-neural-networks-9491262884e0>. Acesso em: set. 2021.

BARBOSA, J. J., PEREIRA, T. M., & OLIVEIRA, F. L. P. (2018). Uma proposta para identificação de outliers multivariados. *Ciência e Natura*, 40, 40. <http://dx.doi.org/10.5902/2179460X29535>.

BASHEER, I. A.; HAJMEER, M. Artificial neural networks: fundamentals, computing, design, and application. *Journal of microbiological methods*, v. 43, n. 1, p. 3-31, 2000.

BEALE, H Mark; Hagan T. Martin; Demuth, B. Howard. *Neural Network Toolbos user guide*. 2010, MathWorks.

BERK, K.: Tolerance and Condition in Regression Computations. *Journal of the American Statistical Association*, 72(360):863866, 1977.

BERTAZOLLI, R.; PELEGRINI, R. Descolação e degradação de poluentes orgânicos em soluções aquosas através do processo fotoeletroquímico. *Química Nova*, São Paulo, v. 25, n. 3, p. 477-482, maio, 2002.

BISHOP, C. M. **Pattern Recognition and Machine Learning**. 1. ed. New York, USA: Springer, 2006.

BISQUERA, R.; SARRIERA, J. C.; MARTINEZ, F. **Introdução à Estatística: enfoque informático com o pacote estatístico SPSS**. 1. ed. Porto Alegre: Artmed, 2004.

BOLFARINE, H.; SANDOVAL, M. C. **Introdução à Inferência Estatística**, 2ª ed., Rio de Janeiro, SBM, 2010.

BOLSTER, C. H.; HORNBERGER, G. M. On the Use of Linearized Langmuir Equations. *Soil Science Society of America Journal*, v. 72, n. 6, p. 1848, 2008

BORBA, C. E; *et al.* Application of the mass action law to describe ion exchange equilibrium in a fixed-bed column. **Chemical Engineering Journal** 172, p.312– 320, 2011.

BRAGA, A.P *et al.* **Redes Neurais Artificiais: Teoria e Aplicações**. Rio de Janeiro: LTC, 2000.

BRAGA, A. P.; LUDERMIR, T. B.; CARVALHO, A. C. P. L. F. **Redes Neurais Artificiais: teoria e aplicações**. 2º edição. Editora LTC: Rio de Janeiro, 2000.

BRUNTON, Steven L.; NOACK, Bernd R.; KOUMOUTSAKOS, Petros, Machine learning for fluid mechanics, *Annual Review of Fluid Mechanics*, v. 52, p. 477-508, 2020.

BUDAMA, N. LOCASCIO, N. **Fundamentals of Deep Learning, Designing Next-Genetarion Machine Intelligence Algorithms**. O'REILLY Media, Inc, 1st edition, 2017.

BASSANEZI, Rodney Carlos. *Modelagem Matemática: teoria e prática*. São Paulo: Contexto, 2015.

CALADO, V. & MONTGOMERY, D. C. Planejamento de Experimentos usando o Statistica. E-Papers Serviços Editoriais, 2003. El-Naas, M. H. (2011). Reject Brine Management. In M. Schorr (Ed.), *Desalination, Trends and Technologies* (pp. 237-252).

CAMPOS, B. P. F *et al.* **Predição da altura total de árvores em plantios de diferentes espécies por meio de redes neurais artificiais**. *Pesquisa Florestal Brasileira*, v. 36, n. 88, p. 375-385, 2016.

CAPP, E., & NIENOV, O. H. (2020). *Bioestatística quantitativa aplicada*. Universidade Federal do Rio Grande do Sul, UFRGS, 260p.

CHARNET, R., FREIRE, C.A.L., CHARNET, E. M. R, BONVINO, H., **Análise de Modelos de Regressão Linear com Aplicações**. 2nd ed. São Paulo, Editora Unicamp, 368 p., 2008.

CHARNET, R., L. FREIRE, C. DE CHARNET, E. E BOVINO, H.: *Análise de Modelos de Regressão Linear com Aplicações*. Editora da Unicamp, 1a ed., 1999.

CHEN, Y.-Y. *et al.* Design and implementation of cloud analytics-assisted smart power meters considering advanced artificial intelligence as edge analytics in demand-side management for smart homes. 2019.

CHITEKA, K.; ENWEREMADU, C. C. Prediction of global horizontal solar irradiance in Zimbabwe using artificial neural networks. **Journal of Cleaner Production**, v. 135, p. 701–711, 2016.

CORRAR, L. J.; PAULO, E.; DIAS FILHO, J. M. (Coords.) *Análise multivariada: para os cursos de administração, ciências contábeis e economia*. São Paulo: Atlas, 2007.

COSENZA, D. N *et al.* **Site classification for eucalypt stands using Artificial Neural Network based on environmental and management features**. *CERNE*, 23(3), 310-320, 2017.

DALARI, B. L. S. K.; *et al.* Application of heterogeneous photo-fenton process using chitosan beads for textile wastewater treatment. **Journal of Environmental Chemical Engineering**, 8, 4,1-7, 2020.

DANTAS NETO, S. A.; AMÂNCIO, L. B.; JACQUES, G. **Pile Settlement Modeling with Multilayer Perceptrons**. *EJGE*, v.19, p. 4517-4528, 2014.

DA SILVA, I. N.; SPATTI, D. H.; FLAUZINO, R. A. **Redes Neurais Artificiais: Para Engenharia e Ciências Aplicadas**. 2 ed. São Paulo: Artliber, 2016.

DA SILVA, A. C., BÔAS, R. N. V., COSTA, M. A. S., MENDES, M. F. Estudo da capacidade adsorvente do gengibre (*Zingiber officinale* Roscoe) no tratamento de

efluentes da indústria têxtil. Ciências Tecnológicas, Exatas e da Terra e seu Alto Grau de Aplicabilidade 2. 2ªed.: Atena Editora, p. 46-57, 2020.

DATA SCIENCE ACADEMY. **Deep Learning Book, 2019**. Disponível em: Acesso em: abr. 2019.

DUTTA, Suman et al. "Development of an artificial neural network model for adsorption and photocatalysis of reactive dye on TiO<sub>2</sub> surface". Expert Systems with Applications, [s. l.], v. 37, n. 12, p. 8634–8638, 2010.

EL-KHAIARY, M. I.; MALASH, G. F. Common data analysis erros in batch adsorption studies. Hydrometallurgy, v. 105, n. 3-4, p. 314- 320, 2011.

FAGUNDES-KLEN, MÁRCIA REGINA *et al.* **Modelagem do efeito do pH na biossorção de metais pela alga marinha Sargassum filipendula**. Acta Scientiarum. Technology v. 33, n. 4, p. 439–446, 2011.

FARAWAY, J J. Extending the linear model with R: generalizedlinear, mixedeffects andnonparametric regression models. Second edition. Boca Raton: CRCpress, 2016.

FERNANDES, A. Inteligência Artificial. 2. ed. Florianópolis: VisualBooks, 2005.

FIORIN, D.V., MARTINS, F.R, SCHUCH, N.J., PEREIRA, E.B. Aplicações de Redes Neurais e Previsões de disponibilidade de Recursos Energéticos Solares. Revista Brasileira de Ensino de Física, v. 33, n. 1, 1309, 2011.

FRIEDMAN, J. H. The elements of statistical learning: Data mining, inference, and prediction. Springer Open, 2017.

GARSON, David (2011). Statnotes: Topics in Multivariate Analysis, by G. David Garson. Disponível em: <http://faculty.chass.ncsu.edu/garson/PA765/statnote.htm>. Acessado em 24 de agosto de 2011.

GAUCH, H. E ZOBEL, R. W. (1990). Imputing missing yield trial data. Theoretical and Applied Genetics, 79(6):753–761.

GHAEDI, M. **Principal component analysis-adaptive neuro-fuzzy inference system modeling and genetic algorithm optimization of adsorption of methylene blue by activated carbon derived from Pistacia khinjuk**. Ecotoxicology and environmental safety, v. 96, p. 110-117, 2013.

GHAZI MOKRI, H. S.*et al.* Adsorption of C.I. Acid Red 97 dye from aqueous solution onto walnut shell: kinetics, thermodynamics parameters, isotherms. **International Journal of Environmental Science and Technology**, 12, 1401-1408, 2015.

GHAEDI, Abdol Mohammad; VAFAEI, Azam. "Applications of artificial neural networks for adsorption removal of dyes from aqueous solution: A review". Advances in Colloid and Interface Science, [s. l.], v. 245, n. April, p. 20–39, 2017.

GOODFELLOW, I.; BENGIO, Y.; COURVILLE, A. **Deep Learning**. [S.I.]: MIT Press, 2016.

GOMES, LUIZ F.A.M.; *et al.* J. D. **Time Series Forecasting with Neural Networks and Choquet Integral**. **Information Technology and Quantitative Management (ITQM)**. *Procedia Computer Science* v. 91. pp. 1119 – 1129, 2016.

GUIDO, S. MÜLLER. C. A. **Introduction to Machine Learning with Python**. O'REILLY Media, Inc, 2016.

GUYON I. Neural network and applications tutorial. **Physics Reports**, v. 17n<sup>o</sup>. 3, p. 215-259, 1991.

HADI, M. *et al.* Prediction of optimum adsorption isotherm: comparison of chi-square and Log-likelihood statistics. *Desalination and Water Treatment*, v. 49, p. 81-94, 2012.

HAIR Jr., J.F.; BLACK, W.C.; BABIN, B.J.; ANDERSON, R.E. & TATHAM, R.L. **Análise multivariada de dados**. 6.ed. Porto Alegre, Bookman, 2009.

HAYKIN, S. **Neural Networks**. Virginia, USA: Prentice Hall. 842p 1999.

HAYKIN, S. **Redes Neurais- Princípios e Práticas**. BOOKMAN, São Paulo, 2<sup>a</sup> ed. 900 p. 2001.

HAYKIN, S. **Redes Neurais: Princípios e Prática**. [S.I.]: Bookman, 2017.

HYNES, N. R. J.; KUMAR, J. S.; KAMYAB, H.; SUJANA, J. A. J.; AL-KHASHMAN, O. A.; KUSLU, Y.; ENE, A.; KUMAR, B. S. Modern Enabling Techniques and Adsorbents based Dye Removal with Sustainability Concerns in Textile Industrial Sector-A comprehensive review. **MA**, 272, 1-17, 2020.

HEIDARI, A. A. *et al.* **Ant lion optimizer: theory, literature review, and application in multi-layer perceptron neural networks**. In: *Nature-Inspired Optimizers*. [S.I.]: Springer, 2020. p. 23–46.

HYNDMAN, R. J.; ATHA NASOPOULOS, G. **Forecasting: Principles and Practice**. 2. ed. OTexts, 2018.

HO, Y.; CHIU, W.-T.; WANG, C.-C. Regression analysis for the sorption isotherms of basic dyes on sugarcane dust. *Bioresource Technology*, v. 96, p. 1285-1291, 2005.

HOFFMAN, R. VIEIRA, S. **Análise de Regressão. Uma introdução à econometria**. 2ed. HUCITEC, São Paulo, 1983.

KANKILIÇ, G.B., METIN, A.U., TUZUN, I., 2016. *Phragmites australis*: An alternative biosorbent for basic dye removal. *Ecological Engineering*, 86, 85-94.

KATHERESAN, V.; KANSEDO, J.; LAU, S. Y. Efficiency of various recent wastewater dye removal methods: **A review. Journal of Environmental Chemical Engineering**, v. 6, n. 4, p. 4676–4697, 2018.

KAUSAR, A.; *et al.* **Dyes adsorption using clay and modified clay: A review. Journal of Molecular Liquids**, 256, 395-407, 2018.

KHAN, S.; RAHMANI, H.; SHAH, S. A. A.; BENNAMOUN, M. **A guide to convolutional neural networks for computer vision. Synthesis Lectures on Computer Vision, Morgan & Claypool Publishers**, v. 8, n. 1, p. 1–207, 2018. Disponível em: Acesso em: 24 de abril de 2018.

KINGMA, P. D. BA, J. L. ADAM: Um Método para Otimização Estocástica, 2014.

KO, M.; TIWARI, A.; MEHNEN, J. **A Review of Soft Computing Applications in Supply Chain Management. Applied Soft Computing** v. 10(3), pp.661–674, 2010.

KOCAMAZ, U. E. *et al.* Control and Synchronization of chaotic supply chains using inteligente approaches. **Computers & Industrial Engineering** vol. 102, pp. 476-487, 2016.

KUMAR, M.A., VIGNESHWARAN, G., PRIYA, M.E., SEENUVASAN, M., KUMAR, V.V., ANURADHA, D., SIVANESAN, S. Concocted bacterial consortium for the detoxification and mineralization of azoic-cum-sulfonic textile mill effluent. *Journal of Water Process Engineering*, v.16, p.199-205, 2017.

LALNUNHLIMI, S.; KRISHNASWAMY, V. Decolorization of azo dyes (Direct Blue 151 and Direct Red 31) by moderately alkaliphilic bacterial consortium. *Brazilian Journal of Microbiology*, v. 47, p. 39-46, 2016.

LECUN, Y., BENGIO, Y., & HINTON, G. Deep learning. *Nature*, 521(7553), 436-444, 2015. PMID:26017442. <http://dx.doi.org/10.1038/nature14539>.

LEVINE, D. M. *et al.* *Estatística: teoria e aplicações usando o Microsoft Excel em português*. 6.ed. Rio de Janeiro: LTC, 2014.

LEITE, A. R. S.; DANTAS NETO, S. A.; POLEMIS JÚNIOR, K.; OLIVEIRA, C. M. O. Study 139 on the shear behaviour of rock joints using different models of artificial neural networks. In: **INTERNATIONAL CONGRESS ON ROCK MECHANICS AND ROCK ENGINEERING**, 2018, Foz do Iguaçu, Proceedings... Foz do Iguaçu, Paraná: ISRM, p. 2707-2714. 2018.

LEVINE, D. M. *et al.* *Estatística: teoria e aplicações usando o Microsoft Excel em português*. 6.ed. Rio de Janeiro: LTC, 2014.

LIU, D. A Practical Guide to ReLU. *Medium*, 30 nov. 2017. Disponível em: Acesso em: 05 jun. 2020.

LIU, X.; DENG, Z.; YANG, Y. **Recent progress in semantic image segmentation. Artificial Intelligence Review, Springer**, v. 52, n. 2, p. 1089–1106, 2019.

LOPES, M. L. M., CHAVARETTE, F. R., & COSSI, A. M. (2017). Avaliação do modelo de regressão linear múltipla e redes neurais artificiais na previsão do ganho de massa em animais. *Revista Brasileira de Engenharia de Biosistemas*, 11(1), 01-17.

LUIZ, T.B.P; GAIARDO, G.F.; SILVA, J.L.S. **Utilização de máquina de vetor de suporte para previsão de níveis de água subterrânea**. *Revista Águas Subterrâneas*. V. 32, n. 1, 2018.

MANN, P. S. *Introductory statistics*. John Wiley & Sons. 2007.

MARTINS, E. R. Configuração de redes neurais artificiais para estimação da altura total de árvores de eucalipto. **Revista Brasileira de Ciências Agrárias**, v. 11, n. 2, 2016.

MATHWORKS. [https://www.mathworks.com/help/pdf\\_doc/nnet/nnet\\_ug.pdf](https://www.mathworks.com/help/pdf_doc/nnet/nnet_ug.pdf). **Neural Network Toolbox User's Guide (R2017b)**, 2017. Acesso em: 28 outubro 2017.

MAXWELL. Adsorção. Disponível em: Acesso em: 12 setembro de 2021.

MELLO, R.; PONTI, M. **A. Machine Learning: A Practical Approach on the Statistical Learning Theory**. [S.l.]: Springer, 2018.

MENDES, C. R.; DILARRI, G.; PELEGRINI, R. T. **Aplicação da biomassa *Saccharomyces cerevisiae* como agente adsorvente do corante Direct Orange 2GL e os possíveis mecanismos de interações adsorbato/adsorvente**. *Matéria* (Rio de Janeiro), v. 20, n. 4, p. 898-908, 2015.

MCCULLOCH, W. S., PITTS, W. A logical calculus of the ideas immanent in nervous activity. **The bulletin of mathematical biophysics**, 5(4):115–133, Dec 1943.

MCKINNEY, W. **Python for Data Analysis**. [S.l.]: O'Reilly Media, 2013.

MIRANDA, F. A.; FREITAS, S. R. C.; FAGGION, P. L. Integração e interpolação de dados de anomalias ar livre utilizando-se a técnica de RNA e krigagem. **Boletim de Ciências Geodésicas**, v.15, n.3, p. 428-443, 2009.

MOHAMMED, M. A., NAJI, T. A., & ABDULJABBAR, H. M. The effect of the activation functions on the classification accuracy of satellite image by artificial neural network. *Energy Procedia*, 157, 164-170, 2019.  
<http://dx.doi.org/10.1016/j.egypro.2018.11.177>.

MOHANRAJ, M., JAYARAJ, S., & MURALEEDHARAN, C. (2012). Applications of artificial neural networks for refrigeration, air-conditioning and heat pump systemsA review. *Renewable and Sustainable Energy Reviews*, 16(2), 1340–1358.  
doi:10.1016/j.rser.2011.10.015.

MONTGOMERY D. C., RUNGER G. C., “Applied Statistics and Probability for Engineers”, 7a edição, Wiley Books, n. 134-138 e p. 291-299, 2018.

MONTGOMERY, D. C. *Design and analysis of experiments*. 8. ed. [S.l.]: John Wiley & Sons, 2013.

MONTGOMERY, D. C. *Introdução ao controle estatístico da qualidade*. 4. ed. Rio de Janeiro: LTC, 2009.

MULERO, Á. *et al.* **An Artificial Neural Network for the surface tension of alcohols. *Fluid Phase Equilibria***. 449, 28-40, 2017.

NASCIMENTO, Ronaldo Ferreira do. *Adsorção: aspectos teóricos e aplicações ambientais* - Fortaleza: Imprensa Universitária, 2014.

NEUPANE, N.; SHAKYA, S. Comparative analysis of backpropagation algorithm variants for network intrusion detection. In: *Computing, Communication and Automation (ICCCA), 2017 International Conference on*. IEEE, p. 726-729, 2017.

NEVES, C. V.; *et al.* **Phenomenological modeling of reactive dye adsorption onto fish scales surface in the presence of electrolyte and surfactant mixtures. *Environmental Technology***, 39, 1–17, 2017.

NIA, R. H.; GHAEDI, M.; GHAEDI, A. M. Modeling of reactive orange 12 (RO 12) adsorption onto gold nanoparticle-activated carbon using artificial neural network optimization based on an imperialist competitive algorithm. ***Journal of Molecular Liquids***, v. 195, p. 219-229, 2014.

NWANKPA, C. IJOMAH, W. GACHAGAN, A., MARSHALL, S. **Activation Functions: Comparison of trends in Practice and Research for Deep Learning**. 2018.

PATEL, A. A. **Hands-On Unsupervised Learning Using Python: How to Build Applied Machine Learning Solutions from Unlabeled Data**. O'REILLY Media, Inc, 1st edition, 2019.

PEREIRA, H. **Prótese Mioelétrica para membro superior implementada em FPGA**. Universidade Federal de Santa Catarina. Araranguá. 2016.

PINHEIRO, J.; BATES, D. *Mixed-effects models in S and S-PLUS*. [S.l.]: Springer Science & Business Media, 2006.

PUCHANA-ROSETO, M. J. *et al.*; **A Coagulation-Flocculation Process Combined With Adsorption Using Activated Carbon Obtained From Sludge For Dye Removal From Tannery Wastewater. *J. of the Chilean Chem. Society***, 63, 1, 3867–3874, 2018.

QAMAR, S. A. *et al.* **Chitosan-based hybrid materials as adsorbents for textile dyes—A review**. *Case Studies in Chemical and Environmental Engineering*, 1-21, 2020.

REIS, B. **Redes neurais - Funções de ativação**. 2016. Disponível em: Acesso em: set. 2019.

REVANKAR, M. S.; LELE, S. S. Synthetic dye decolorization by *Ganoderma* sp. WR-1. *Bioresource Technology*, v. 98, p. 775-780, 2007.

ROCHA, K. R. da; JÚNIOR, A. J. B. Anova medidas repetidas e seus pressupostos: Análise passo a passo de um experimento. *Revista Eletrônica Perspectivas da Ciência e Tecnologia*, v. 10, p. 29, 2018. ISSN 1984-5693. Disponível em: <http://revistascientificas.ifrj.edu.br:8080/revista/index.php/revistapct/article/view/955>  
Acesso em: 26 dez. 2018.

SALES, R. de S; MAGALHÃES, L. M.; NOBRE, Ga. A. R. Análise de correlação entre defeitos estruturais e funcionais em rodovias brasileiras em revestimento de CAUQ. 33 Congresso de Pesquisa e Ensino em Transporte da ANPET, Balneário do Camboriú, 2019. Disponível em: [http://www.anpet.org.br/anais/documentos/2019/Infraestrutura/Dimensionamento,%20Avalia%C3%A7%C3%A3o%20e%20Gest%C3%A3o%20de%20Pavimentos%20VI/1\\_547\\_AC.pdf](http://www.anpet.org.br/anais/documentos/2019/Infraestrutura/Dimensionamento,%20Avalia%C3%A7%C3%A3o%20e%20Gest%C3%A3o%20de%20Pavimentos%20VI/1_547_AC.pdf). Acesso em: 26 jan 2021.

SARKER, R.; CHOWDHURY, M.; DEB, A. K. **Reduction of Color Intensity from Textile Dye Wastewater Using Microorganisms: A Review**. *Int. J. Curr. Microbiol. App. Sci*, 8, 2, 3407-3415, 2019.

SAS INSTITUTE. 1993. SAS/ETS software applications guide2: econometric modeling, simulation and forecasting. Version. 6. SAS Institute Inc., Cary. 429 p.

SCHWAGER, S. J., & MARGOLIN, B. H. (1982). Detection of multivariate normal outliers. *Annals of Statistics*, 10(3), 943-954.  
<http://dx.doi.org/10.1214/aos/1176345884>.

SEWAK, M.; KARIM, M. R.; PUJARI, P. **Practical Convolutional Neural Networks: Implement advanced deep learning models using Python**. Packt Publishing Ltd, 2018.

SHANG Y.; BENJAMIN W. W. "Global Optimization for Neural Network Training", *IEEE Computer*, Vol. 29, no.3, pp. 45-54, March, 1996.

SHAODUAN, OU; ACHENIE E. K. LUKE; A hybrid neural network model for PEM fuel cells. *Journal of Power Sources* 140 (2005) 319–330

SILVA, IVAN N *et al.* **Redes Neurais Artificiais para engenharia e ciências aplicadas. Fundamentos teóricos e aspectos práticos**. 2ª edição, revisada e ampliada. São Paulo. Artliber, 2016.

SILVA, L. N. de; VON ZUBEN, F. J. *Redes Neurais Artificiais*. Disponível na internet em: Acesso em 15 de agosto de 2019.

SILVA, A. C. da, **Estudo da capacidade adsorvente do gengibre (*Zingiber officinale* Roscoe) no tratamento de efluentes da indústria têxtil**. *Ciências Tecnológicas, Exatas e da Terra e seu Alto Grau de Aplicabilidade* 2. 2ªed.: Atena Editora, p. 46-57, 2020.

SINGH, B. P. N.; KULSHRESTHA, S. P. Kinetics of water sorption by soybean and pigeonpea grains. **Journal of Food Science**, v. 52, n. 6, p. 1538-1544, 1987.

SOARES, L. A., *et al.* Otimização da adsorção do Erionyl Yellow AR em carvão ativado do bagaço da cana-de-açúcar por meio de planejamento experimental. **Brazilian Journal of Development**, v. 5, n. 12, p. 28781-28797, 2019.

SOUZA, S.V.C.; JUNQUEIRA, R. G. A procedure to assess linearity by ordinary least squares method. *Anal. Chim. Acta*, v. 552, n. 1-2, p. 25-35, 2005.

SPÖRL, C.; CASTRO, E. G.; LUCHIARI, A. Aplicação de Redes Neurais Artificiais na construção de modelos de fragilidade ambiental. **Revista do Departamento de Geografia**, v. 21, n.1, p. 113-135, 2011.

STAUB, S. *et al.* **Artificial Neural Network and Agility**. *Procedia - Social and Behavioral Sciences*, v. 195, p. 1477–1485, 2015.

STHDA, Regression Model Accuracy Metrics: R-square, AIC, BIC, Cp and more. Disponível em: [http://www.sthda.com/english/articles/38-regression-](http://www.sthda.com/english/articles/38-regression)  
SUTTON, R. S., BARTO, A. G., BACH, F. **Reinforcement Learning - An introduction**. MIT Press, second edition, 2018.

SUPERDATASCIENCE TEAM. The Ultimate Guide to Convolutional Neural Networks (CNN). SuperDataScience, 27 ago. 2018. Disponível em: Acesso em: 26 abr. 2020.

TANZIFI, Marjan *et al.* Artificial neural network optimization for methyl orange adsorption onto polyaniline nano-adsorbent: Kinetic, isotherm and thermodynamic studies. **Journal of Molecular Liquids** v. 244, p. 189–200, 2017.

TAVARES, D. S. “**Use of artificial neural networks to predict concrete compression strength**”. *Braz. J. of Develop.*, Curitiba, v.6, n.7, p. 42185-42826, jul. 2020.

TKÁČ, M.; VERNER, R. Artificial neural networks in business: Two decades of research. **Applied Soft Computing Journal**, v.38, p.788-804, 2016.

TOMASSONI, F.; *et al.* **Otimização da eletrocoagulação aplicada em efluente têxtil**. *Revista DAE*, 219, 67, 1-14, 2019.

TRAN, H. N. *et al.* Mistakes and inconsistencies regarding adsorption of contaminants from aqueous solutions: A critical review. *Water Research*, v. 120, p. 88-116, 2017.

TURAN, N. G.; MESCI, B.; OZGONENEL, O. The use of artificial neural networks (ANN) for modeling of adsorption of Cu (II) from industrial leachate by pumice. **Chemical Engineering Journal**, v. 171, n. 3, p. 1091–1097, 2011.

TURING, A. M. **Computing machinery and intelligence**. *Mind*, 59(236):433–460, 1950.

Y. SINGH and P. CHANDRA, "A class +1 sigmoidal activation functions for FFANNs," *Journal of Economic Dynamics and Control*, vol. 28, no. 1, pp. 183–187, 2003.

YAN, W. Biplot analysis of incomplete two-way data. *Crop Science*, 53(1):48–57, 2013.

YOUSSEFI, S., EMAM-DJOMEH, Z., & MOUSAVI, S. M. (2009). Comparison of Artificial Neural Network (ANN) and Response Surface Methodology (RSM) in the prediction of quality parameters of spray-dried pomegranate juice. *Drying Technology*, 27(April 2016), 910–917. doi:10.1080/07373930902988247

SASSI, C. P. et al. Modelo de regressão linear múltipla utilizando os softwares R e Statistica: uma aplicação a dados de conservação de frutas. [2011?]. Disponível em: . Acesso em: 06 jul. 2016.

THIBAUT, J., BREUSEGEM, V. V., CHÉRUUY, A. **On-line prediction of fermentation variables using neural networks.** *Biotechnology and Bioengineering*, v. 36, p. 1041-1048, 1990.

VALLE, C. E. **Qualidade Ambiental ISO 14000.** São Paulo: Senac, 2004.

VILLA, T. E. D.: Predição do Custo do Milho por Meio de Modelos de Regressão Linear Múltipla. Universidade Federal de Uberlândia, 2016.

WANG, J.; TSAPAKIS, I.; ZHONG, C. **A space–time delay neural network model for travel time prediction.** *Engineering Applications of Artificial Intelligence*. vol. 52, pp. 145–160, 2016.

XAVIER, C. S. F.; **Avaliação da remoção de corantes têxteis presentes em águas residuárias utilizando bagaço da cana-de-açúcar como adsorvente.** 2020. 77 f. Dissertação (Mestre em Ciência e Tecnologia Ambiental). Programa de Pós-Graduação em Ciência e Tecnologia Ambiental Universidade Estadual da Paraíba, 2020.

ZAHEDI, J., & ROUNAGHI, M. M. **Application of artificial neural network models and principal component analysis method in predicting stock prices on Tehran Stock Exchange.** *Physica A: Statistical Mechanics and its Applications*, 438, 178-187, 2015.

ZHANG, Z. *et al.* Comparative study on adsorption of two cationic dyes by milled sugarcane bagasse. *Industrial Crops & Products*, v. 42, p. 41-49, 2013.

ZIGGAH, Y. Y., YOUJIAN, H., TIERRA, A. R., AND LAARI, P. B. Coordinate transformation between global and local datums based on artificial neural network with k-fold cross-validation: A case study, ghana. *Earth Sciences Research Journal*, v. 3, n. 1, p. 67–77, 2019.