

UNIVERSIDADE FEDERAL DO PARÁ
INSTITUTO DE TECNOLOGIA
PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA ELÉTRICA

CLUSTERIZAÇÃO, CLASSIFICAÇÃO E PREDIÇÃO DE “PRÉ-EFEITO ANÓDICO” DE CUBA ELETROLÍTICA DE ALUMÍNIO PRIMÁRIO

BRUNO NICOLAU MAGALHÃES DE SOUZA CONTE

DM: 24/2020

UFPA / ITEC / PPGEE
Campus Universitário do Guamá
Belém-Pará-Brasil
2020

UNIVERSIDADE FEDERAL DO PARÁ
INSTITUTO DE TECNOLOGIA
PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA ELÉTRICA

BRUNO NICOLAU MAGALHÃES DE SOUZA CONTE

CLUSTERIZAÇÃO, CLASSIFICAÇÃO E PREDIÇÃO DE “PRÉ-EFEITO ANÓDICO” DE CUBA ELETROLÍTICA DE ALUMÍNIO PRIMÁRIO

DM: 24/2020

UFPA / ITEC / PPGEE
Campus Universitário do Guamá
Belém-Pará-Brasil
2020

UNIVERSIDADE FEDERAL DO PARÁ
INSTITUTO DE TECNOLOGIA
PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA ELÉTRICA

BRUNO NICOLAU MAGALHÃES DE SOUZA CONTE

CLUSTERIZAÇÃO, CLASSIFICAÇÃO E PREDIÇÃO DE “PRÉ-EFEITO ANÓDICO” DE CUBA ELETROLÍTICA DE ALUMÍNIO PRIMÁRIO

Dissertação submetida à Banca Examinadora do Programa de Pós-Graduação em Engenharia Elétrica - PPGEE, do Instituto de Tecnologia - ITEC, da Universidade Federal do Pará - UFPA, para a obtenção do Grau Mestre em Engenharia Elétrica na área de Computação Aplicada. Linha de Pesquisa: Inteligência Artificial. Orientador: Prof. Dr. Roberto Célio Limão de Oliveira.

UFPA / ITEC / PPGEE
Campus Universitário do Guamá
Belém-Pará-Brasil

2020

Dados Internacionais de Catalogação na Publicação (CIP) de acordo com ISBD Biblioteca do ITEC/UFPA-Belém-PA

C761c Conte, Bruno Nicolau Magalhães de Souza, 1989-

Clusterização, classificação e predição de pré-efeito anódico de cuba eletrolítica de alumínio primário / Bruno Nicolau Maalhães de Souza Conte.-2020.

Orientador: Roberto Célio Limão de Oliveira

Dissertação (Mestrado) – Universidade Federal do Pará, Instituto de Tecnologia, Programa de Pós-graduação em Engenharia Elétrica, Belém, 2020.

1, Redes neurais (computação). 2. Alumínio – produção. 3. Processo de fabricação – modelos matemáticos.I. Título.

CDD 23 ed. 006.32

Elaborado por Lucicléa Silva de Oliveira – CRB-2/648

Ativa
Acesse

“CLUSTERIZAÇÃO, CLASSIFICAÇÃO E PREDIÇÃO DE “PRÉ-EFEITO ANÓDICO” DE CUBA ELETROLÍTICA DE ALUMÍNIO PRIMÁRIO”

AUTOR: **BRUNO NICOLAU MAGALHÃES DE SOUZA CONTE**

DISSERTAÇÃO DE MESTRADO SUBMETIDA À BANCA EXAMINADORA APROVADA PELO COLEGIADO DO PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA ELÉTRICA, SENDO JULGADA ADEQUADA PARA A OBTENÇÃO DO GRAU DE MESTRE EM ENGENHARIA ELÉTRICA NA ÁREA DE COMPUTAÇÃO APLICADA.

APROVADA EM: 21/08/2020

BANCA EXAMINADORA:



Prof. Dr. Roberto Célio Limão de Oliveira
(Orientador - PPGEE/UFPA)



Prof.^a Dr.^a Jasmine Priscyla Leite de Araújo
(Avaliadora Interna - PPGEE/UFPA)



Prof. Dr. Lídio Mauro Lima de Campos
(Avaliador Externo ao Programa - ICEN/UFPA)

VISTO:

Prof.^a Dr.^a Maria Emília de Lima Tostes

(Coordenadora do PPGEE/ITEC/UFPA)

Dedico este trabalho à minha mãe Maria Magalhães (in memoriam), que cuidou de mim até o último instante da sua vida.

AGRADECIMENTOS

Agradeço primeiramente a Deus, por ter me dado sabedoria e determinação para iniciar e terminar essa dissertação.

À Universidade Federal do Pará, por ter me proporcionado a oportunidade da realização deste trabalho de dissertação, além do apoio dado pelo corpo docente que sempre procurou transmitir conhecimento para o apoio a este trabalho.

Ao Professor Dr. Roberto Celio Limão De Oliveira, pela orientação eficiente e segura o que proporcionou o direcionamento desta dissertação.

Agradeço a minha mãe (in memoriam) por ter me acompanhado em toda minha trajetória, com todo seu carinho e amor incondicional, ao meu pai pelo seu apoio e compreensão, aos meus Irmãos por sempre manter a união familiar, a Karla Oliveira por sempre permanecer ao meu lado e por todo o seu amor.

Agradeço aos meus familiares e amigos, e por todos que contribuíram de maneira direta e indireta para a conclusão deste estudo.

“Pedi, e dar-se-vos-á, buscai, e achareis; batei, e abrir-se-vos-á. Porque todo o que pede, recebe; e o que busca, acha; e a quem bate, abrir-se-á. Ou qual de vós, porventura, é o homem que, se seu filho lhe pedir pão, lhe dará uma pedra? Ou, porventura, se lhe pedir um peixe, lhe dará uma serpente? Pois se vós outros, sendo maus, sabeis dar boas dádivas a vossos filhos, quanto mais vosso Pai, que está nos Céus, dará boas dádivas aos que lhes pedirem”

(Mateus, VII: 7-11)

RESUMO

O setor industrial é um dos principais responsáveis pela grave situação ambiental do planeta e também por crescentes exigências legais, com relação aos resíduos gerados. Em contrapartida, muitas empresas têm reagido pró ativamente, a partir da implantação de estratégias de gestão como: produção limpa, certificação ambiental, redução de resíduos tóxicos, reciclagem, consumo sustentável e reuso, principalmente. É oportuno ressaltar, que o processo de redução de alumínio é responsável por uma grande quantidade de emissão de gases causadores do efeito estufa e, assim, promove impactos ambientais e alterações climáticas graves. Durante o processo de redução de alumínio, a ocorrência do efeito anódico provoca um aumento extremo na tensão da cuba e, conseqüentemente, uma elevação na temperatura do banho, com temperaturas altíssimas, resultado em um distúrbio térmico, com a possibilidade de ocorrer o derretimento da camada isolante da cuba e as conseqüências finais são a perda de produção em toda a linha de cubas, sua vida útil diminuída e a produção de gases PFCs. Buscando uma estratégia apoiada na sustentabilidade, tendo em vista a problemática do agravamento do Efeito Estufa, o aumento extremo na tensão do forno e, conseqüentemente, a perda de produção em toda a linha de cubas, este trabalho propõe o uso de uma Rede Neural Artificial junto com algoritmos de Clusterização para criar automaticamente os rótulos de pré-Efeito anódico, e assim, prever o comportamento dinâmico não-linear da ocorrência do pré-efeito anódico do forno da indústria de redução de alumínio primário, com base em dados reais das cubas eletrolíticas. Com a utilização desses modelos de *Machine Learning* pode-se prever a ocorrência do pré-efeito anódico permitindo que os operadores de processos tomem medida mitigadoras de supressão do efeito anódico, evitando a perda da produção do alumínio na cuba e diminuindo a emissão de gases do efeito estufa.

PALAVRAS CHAVE: Redes Neurais Artificiais. Clusterização. Redução de Alumínio. Efeito Anódico. Previsão.

ABSTRACT

The industrial sector is one of the main responsible for the serious environmental situation on the planet and also for increasing legal requirements, in relation to the waste generated. On the other hand, many companies have reacted proactively, based on the implementation of management strategies such as: clean production, environmental certification, reduction of toxic waste, recycling, sustainable consumption and reuse, mainly. It is worth mentioning that the aluminum reduction process is responsible for a large amount of greenhouse gas emissions and, thus, promotes environmental impacts and serious climate changes. During the aluminum reduction process, the occurrence of the anodic effect causes an extreme increase in the tub tension and, consequently, an increase in the bath temperature, with very high temperatures, resulting in a thermal disturbance, with the possibility of melting the insulating layer of the vat and the final consequences are the loss of production in the entire vat line, its shortened service life and the production of PFC gases. Seeking a strategy based on sustainability, I try to take into account the problem of the worsening of the Greenhouse Effect, the extreme increase in kiln tension and, consequently, the loss of production in the entire line of vats, this work proposes the use of an Artificial Neural Network together with Clustering algorithms to automatically create anodic pre-Effect labels, and thus predict the nonlinear dynamic behavior of the primary aluminum reduction industry oven anodic pre-effect, based on actual vat data electrolytic. With the use of these Machine Learning models, it is possible to predict the occurrence of the anodic pre-effect, allowing process operators to take mitigating measures to suppress the anodic effect, avoiding the loss of aluminum production in the vat and decreasing the emission of gases from the greenhouse effect.

KEY WORDS: Artificial Neural Networks. Clustering. Aluminum Reduction. Anodic Effect. Forecast.

LISTA DE FIGURAS

Figura 1 - Produção de Alumínio por eletrólise	13
Figura 2 - Cadeia de Produção do Alumínio	14
Figura 3 - Forno Eletrolítico	20
Figura 4 - Efeito Estufa	23
Figura 5 - Características da Rede Neural	28
Figura 6 - Modelo de Neurônio Artificial	29
Figura 7 - Funções de Ativação para uma Rede Neural	29
Figura 8 - Representação das camadas de um Rede Neural	32
Figura 9 - Conjunto de dados com K-means	33
Figura 10 - Exemplo de dendrograma	35
Figura 11 - Construção da Arvore CF	36
Figura 12 - Fluxograma do modelo hibrido	42
Figura 13 - Matriz de scatter plot para os 6 atributos do conjunto de dados	44
Figura 14 - Boxplot para o atributo Voltagem Líquida	46
Figura 15 - Boxplot para o atributo Fluoreto de Alumínio	47
Figura 16 - Boxplot para o atributo Fluoreto de Cálcio	47
Figura 17 - Boxplot para o atributo Temperatura de Banho	48
Figura 18 - Boxplot para o atributo Intensidade de Corrente.....	48
Figura 19 - Boxplot para o atributo Voltagem Bruta	49
Figura 20 - Método de amostragem Holdout	50
Figura 21 - Características da Base de Dados Originais	51
Figura 22 - Características da base de dados com K-means	51
Figura 23 - Resultado da Base de Dados com Agrupamento Aglomerativo	52
Figura 24 - Base de Dados com Agrupamento Bich	53
Figura 25 - Gráfico de Avaliações Externas	54
Figura 26 - Taxa de Erro variando-se o número de neurônios para a primeira camada intermediária.....	56
Figura 27 - Taxa de Erro variando-se o número de neurônios na primeira e segunda camada intermediária	56
Figura 28 - Taxa de Erro variando-se o número de neurônios na primeira, segunda e terceira camada intermediária.....	57

Figura 29 - Taxa de Erro variando-se o número de neurônios na primeira, segunda, terceira e quarta camada intermediária.....	58
Figura 30 - Taxa de Erro variando-se o número de neurônios nas cinco camadas intermediária.....	58
Figura 31 - Ciclo de Treinamento da Rede Neural com melhor a topologia	59
Figura 32 - Topologia da Rede Neural	62
Figura 33 - Disposição no espaço tridimensional para os dados de entrada	62
Figura 34 - Categorização do conjunto de dados realizado pela Rede Neural	63
Figura 35 - Matriz de confusão da Rede Neural	64

LISTA DE TABELAS

Tabela 1 - Comparação das propriedades físicas dos metais mais utilizados	21
Tabela 2 - Principais Gases de Efeito Estufa	24
Tabela 3 - Parâmetros da RNA	41
Tabela 4 - Limite Inferior e Superior de corte para Outliers	49
Tabela 5 - Conjunto de dados	50
Tabela 6 - Resultados dos Critérios de avaliação	54
Tabela 7 - Resultados das Predições Neurais para os dados de teste	60
Tabela 8 – Conjunto de dados	61

LISTA DE ABREVIATURA E SIGLAS

ABAL - Associação Brasileira de Alumínio

Al₂O₃ - Óxido de alumínio

AMI - Adjusted Mutual Information

Birch - Algoritmo de Agrupamento

CF - Recurso de Cluster

CFCs - Clorofluorcarbonos

CH₄ - Metano

CO₂ - Gás carbônico

EA - Efeito Anódico

IA - Inteligência Artificial

IACS - International Annealed Copper Standard

MLP - Rede de Múltiplas Camadas

PFCs - Gases perfluorcarbonos

RNA - Redes Neurais Artificiais

RNMD - Redes Neurais Multicamadas Diretas

RNR - Redes Neurais Recorrentes

SEMISH - Seminário Integrado de Software e Hardware

SUMÁRIO

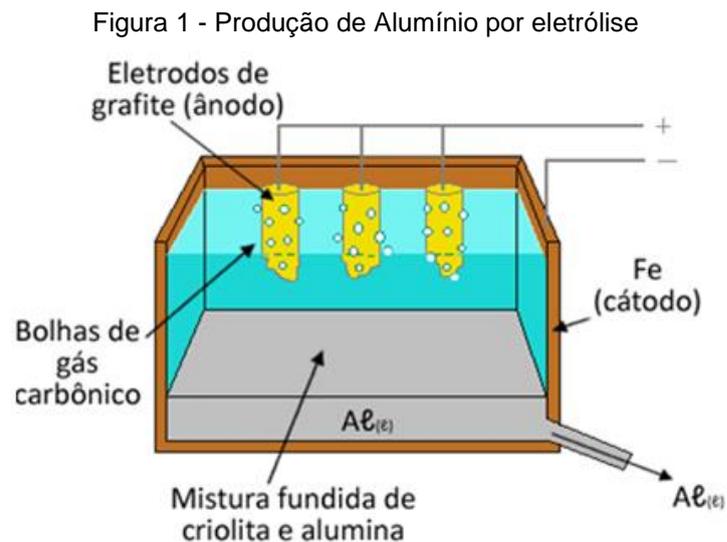
1 INTRODUÇÃO	13
1.1 Descrição Geral do Problema	13
1.2 Justificativa e Objetivo	15
1.3 Revisão Bibliográfica	16
1.4 Estrutura da Dissertação	18
2 O PROCESSO DO ALUMÍNIO E O EFEITO ANÓDICO	19
2.1 Processo de Fabricação do Alumínio	19
2.2 Efeito Anódico	23
2.3 Controle do Processo de cuba de eletrolítica	25
3 APRENDIZAGEM DE MÁQUINA.....	27
3.1 Introdução	27
3.2 Rede Neurais Artificiais	27
3.2.1 Funções de ativação	30
3.2.2 Treinamento	30
3.2.3 Arquitetura	31
3.2.4 FeedForward	32
3.3 Algoritmos de Agrupamento	32
3.3.1 Algoritmo Particional - K-means	33
3.3.2 Algoritmo Hierárquico - Agrupamento Aglomerativo	34
3.3.3 Agrupamento Bich	35
3.3.4 Avaliação para algoritmos de agrupamento	37
4 USO DE REDES NEURAIS E CLUSTERIZAÇÃO NA PREDIÇÃO DO PRE- EFEITO ANÓDICO NO PROCESSO DE REDUÇÃO DO ALUMÍNIO.....	41
4.1 Aprendizado	41
4.2 Resultados e Discussão	43
4.2.1 Base de Dados	43
4.2.2 Pré-Processamento de Dados	45
4.3 Estudo de Caso	61
5 CONCLUSÃO	65
REFERÊNCIAS	66

1 INTRODUÇÃO

1.1 Descrição Geral do Problema

O processo de produção de alumínio primário, alumínio metálico ou alumínio líquido, é chamado de Hall-Héroult e transforma a alumina oxidada em alumínio. É fundamental que este processo se mantenha estável pois qualquer mudança dos elementos do forno acaba por ocasionar o superaquecimento do forno, impactando assim na perda da aresta, alto consumo de energia elétrica e emissão de gases para atmosfera.

A Figura 1 apresenta o processo elétrico de redução da alumina (principal matéria prima do alumínio primário).



Fonte: DIEGO (2018)

Em geral, o processo Hall-Héroult fundamenta-se na obtenção do alumínio líquido por meio da reação química entre alumina e o carbono fornecido pelos ânodos das cubas eletrolíticas. Basicamente, a eletrólise trata de separar a alumina em alumínio e oxigênio, sob a passagem de corrente elétrica no banho eletrolítico, que é composto por criolita e outros elementos como sais fundentes, gases e aditivos químicos, para garantir a estabilidade/temperatura do processo em torno de 960°C (CONTE, 2015).

O forno eletrolítico é o recipiente de uso industrial, visto na Figura 1, que é uma das etapas da cadeia de produção do alumínio (Figura 2).

Figura 2 - Cadeia de Produção do Alumínio



Fonte: Conte (2015, p.14)

É propício falar, que a instabilidade do processo de redução de alumínio acaba por ocasionar o aumento da formação de gases no forno eletrolítico e posteriormente o surgimento de um fenômeno indesejado no processo produtivo do alumínio o Efeito Anódico (EA) e outros. No entanto, grande parte da sua ocorrência é ocasionada pela variação da resistência, que faz com que o banho fique mais resistivo. Assim o EA faz com que haja um aumento rápido da tensão do forno, alcançando valores até 10 vezes maiores que os valores normais (CONTE, 2015).

O evento efeito anódico é associado ao surgimento de um revestimento de gás em que a corrente elétrica não consegue circular com facilidade, e com essa “capa isolante” o banho eletrolítico não pode mais “tocar” a superfície anódica resultando em diversas consequências, entre elas, perda de produção em toda a linha do forno e uma enorme contribuição para o efeito estufa. (CONTE Thiago, 2017; CONTE Bruno; 2017 e OLIVEIRA, Roberto, 2017)

Em linhas gerais, o processo industrial de obtenção de alumínio com até 99,9 % de pureza é um processo complexo, pois envolve a participação de vários elementos como sais fundentes, gases e aditivos químicos para transformação da alumina no metal primário e para manter a estabilidade. Assim, o uso das técnicas de aprendizagem de máquina poderão ser utilizadas para prever o comportamento

do pré-efeito anódico, mantendo a estabilidade no forno eletrolítico e, minimizando a problemática do agravamento da concentração de gases do efeito estufa.

Portanto, o presente trabalho propõe a utilização de técnicas de algoritmos de agrupamento do tipo K-means Clustering, Aglomerativo, Bich e o uso de Rede Neurais Artificiais para diagnosticar o comportamento do pré-efeito anódico em fornos eletrolítico para a produção de alumínio primário.

1.2 Justificativa e Objetivo

A complexidade do processo de redução de alumínio primário, associado a ações na minimização das perdas de produção e o impacto dos recursos tecnológicos no meio ambiente, tornam a tarefa dos engenheiros e projetistas responsáveis pela identificação, controle e supervisão cada vez mais difícil, uma vez que as ferramentas matemáticas tradicionais são deficientes em modelar algumas características típicas destes sistemas, tais como não-linearidades. Torna-se clara a necessidade de ferramentas computacionais mais sofisticadas capazes de criar uma hipótese, ou função para resolver o problema que se deseja tratar.

O principal objetivo desta dissertação é disponibilizar às fábricas de alumínio primário um modelo para o gerenciamento das informações de controle para monitorar as variações do pré-efeito anódico, a fim de melhorar as alterações do processo do forno eletrolítico e garantir a qualidade do alumínio. A proposta é dividida em partes fundamentais, descritas mais especificamente abaixo:

- a) Definir e elaborar algoritmos de aprendizagem de máquina como Redes Neurais Artificiais, agrupamento K-means, agrupamento Aglomerativo e Birch;
- b) Avaliar o potencial de diagnosticar padrões de pré-efeito anódico das Redes Neurais Artificiais em dados do forno eletrolítico de produção de alumínio primário;

- c) Validar a capacidade de clusterização com K-means, agrupamento Aglomerativo e Birch no auxílio da predição de pré-efeito anódico por meio dos critérios de avaliações externas.

1.3 Revisão Bibliográfica

Referente as aplicações de Redes Neurais Artificiais (RNAs) em sistemas industriais de redução de alumínio primário, podem ser citados trabalhos no que concerne a modelagem do Pré-efeito Anódico. Nesse sentido Conte T., Conte B. e Oliveira (2017) desenvolveram um trabalho que avalia dois tipos de Redes Neurais Artificiais na tarefa de predição do Efeito Anódico que ocorre nos fornos eletrolítico de alumínio primário. A proposta é utilizar Redes Neurais Multicamadas Diretas (RNMD) e Redes Neurais Recorrentes (RNR) para identificar a ocorrência do Efeito Anódico do forno eletrolítico.

Conte T., Conte B. e Oliveira (2018) compararam as RNAs com as técnicas de mineração de dados Árvore de decisão e Random Florest para avaliar a sua capacidade de diagnosticar padrões na tarefa de identificar o Pré-Efeito Anódico dos fornos eletrolítico de alumínio primário.

Silva (2009), a partir dos sinais de tensão e corrente apresenta a investigação de modelos paramétricos Autoregressive Model Exogenous Inputs (ARX) e Output Error (OE) que melhor representa o funcionamento do forno eletrolítico para possibilitar a predição do efeito anódico no processo produtivo do alumínio.

Chen L., Chen Y., Gui W. (2019) contribuem para a pesquisa e prática de previsão de Efeito Anódico (EA) de duas maneiras específicas. Primeiro, é proposto um modelo de previsão de EA, baseado no algoritmo CTFM para previsão de EA, que pode melhorar o tempo de execução e a precisão da previsão de EA. Segundo, a eficácia do modelo de previsão de EA é verificada usando um número suficientemente grande de amostras extraídas dos dados de produção da planta eletrolítica de alumínio Jinglian na Mongólia Interior.

Zhou et al. (2015), a fim de melhorar a precisão do efeito anódico em células de eletrólise de alumínio, propôs o método de previsão por meio da Rede Neural de Regressão Generalizada (GRNN).

Entretanto, o diferencial desse trabalho atual, em relação aos artigos mencionados acima é realizar primeiro uma tarefa de aprendizado descritivo, ou não supervisionado, para identificar informações relevantes nos dados sem a presença de um elemento externo que guiaria o aprendizado, e posteriormente realizar um aprendizado preditivo, ou supervisionado, treinando e criando um estimador a partir dos rótulos da tarefa de agrupamento anterior, e por fim avaliar o modelo de rede preditiva com as etiquetas do banco de dados originais.

1.4 Estrutura da Dissertação

Este trabalho está dividido em 5 capítulos, estruturado da seguinte forma: O atual capítulo descreve a introdução do trabalho envolvendo descrição geral do problema, justificativa, objetivo e a revisão bibliográfica utilizadas como modelos de propagação existentes na literatura. Nos capítulos seguintes, serão descritos sobre a arquitetura proposta, bem como a metodologia de avaliação utilizada e dos resultados obtidos.

No capítulo 2 é descrito as etapas para obtenção do alumínio, a estrutura física e funções da cuba eletrolítica, apresentando algumas das causas do efeito anódico.

Por outro lado, o capítulo 3 apresentará as técnicas de aprendizagem de máquina tais como: redes neurais artificiais e algoritmos de agrupamento.

Enquanto no capítulo 4 é descrita a metodologia utilizada no trabalho e os resultados obtidos pela modelagem a partir do método proposto para detecção do pré-efeito anódico.

O capítulo 5 descreve as conclusões do trabalho e algumas possibilidades de trabalhos futuros.

2 O PROCESSO DO ALUMÍNIO E O EFEITO ANÓDICO

2.1 Processo de Fabricação do Alumínio

Segundo a Associação Brasileira de Alumínio (ABAL, 2019) o processo de obtenção do alumínio é feita a partir da extração do mineral bauxita da natureza, pois o alumínio não é encontrado na crosta terrestre, e a mesma deve apresentar no mínimo 30% de óxido de alumínio (Al_2O_3) para que a produção seja economicamente viável.

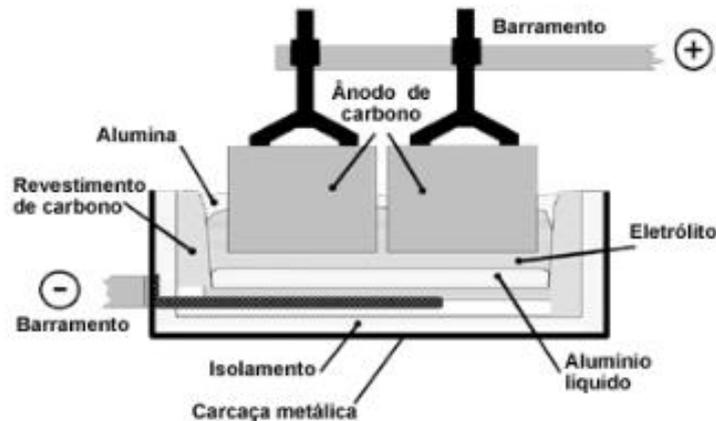
A bauxita pode ser encontrada em três principais grupos climáticos como Mediterrâneo, Tropical e Subtropical. E, destacamos as reservas brasileiras de bauxitas, que estão entre as maiores do mundo, além da ótima qualidade do minério. A obtenção do alumínio se dá em três partes: Mineração, Refinaria e Redução.

Na primeira parte, basicamente a mineração é caracterizada pela remoção planejada da vegetação e do solo orgânico, retirada das camadas superficiais, lavagem do minério e outros.

Na etapa da refinaria, a alumina oxidada (ou alumina) passa por um processo de purificação e, é distribuída para diversas aplicações, como por exemplo, tratamento de água, polimento, materiais refratários e insumo para obtenção do alumínio primário.

Na última etapa, em altas temperaturas e com a passagem de alta corrente elétrica, ocorre o processo de redução conhecida como Hall-Hérolt em cubas eletrolíticas (ver Figura 3).

Figura 3 - Forno Eletrolítico



Fonte: ABAL (2019)

A cuba eletrolítica é uma construção de aço preparada para ocorrer o processo de obtenção do alumínio. A transformação da alumina em alumínio, basicamente consiste em dissolver a alumina em um banho de criolita fundida com a participação de outros elementos como sais fundentes, gases e aditivos químicos, sobre a passagem de uma alta corrente elétrica, e promovendo a combinação do oxigênio com o ânodo de carbono, desprendendo-se na forma de dióxido de carbono, e em alumínio líquido que se precipita no fundo da cuba eletrolítica. O metal líquido (alumínio primário) é retirado das cubas e encaminhado ao lingotamento, no qual são produzidos os lingotes, isto é, placas de metal primário.

Conforme mostrado na Figura 3 tem-se um exemplo de uma cuba eletrolítica, onde identifica-se:

- a) **Anodo de carbono:** Estrutura formada por derivado do petróleo e piche sólido conectados à hastes;
- b) **Alumina:** Insumo para obtenção do alumínio primário;
- c) **Alumínio Líquido:** É o metal obtido pela quebra da molécula de alumina, através da eletrólise;
- d) **Isolamento:** São elementos diretamente ligados à proteção térmica.
- e) **Revestimento de carbono:** Composição semelhante ao anodo, estando situado na parte inferior da cuba;
- f) **Eletrólito:** Dissociação das moléculas de alumina (Al_2O_3) em alumínio e gás carbônico (CO_2).

Nas fábricas de alumínio primário há centenas de cubas eletrolíticas produzindo o alumínio, sendo normalmente distribuídas em salas chamadas de 'Redução' onde as cubas estão ligadas eletricamente em série.

As características químicas e físicas do alumínio são descritas de forma resumida a seguir e o tornam um dos metais mais utilizados no mundo (ABAL, 2019):

- a) **Peso específico:** É um metal leve, por volta de 35% do peso do aço e 30% do peso do cobre;
- b) **Resistência a corrosão:** Resistente a reações químicas devido uma camada de autoproteção de óxido;
- c) **Reciclagem:** 100% reciclável sem perda de suas características físico-químicas;
- d) **Condutividade elétrica:** O Alumínio primário possui condutividade de 62% da IACS (International Annealed Copper Standard).

A Tabela 1 a seguir apresenta uma comparação das propriedades físicas típicas dos três metais mais usados pela sociedade.

Tabela 1 - Comparação das propriedades físicas dos metais mais utilizados

Características	Alumínio	Aço	Cobre
Densidade(g/cm ³)	2,70	7,86	8,96
Temperatura de fusão(°C)	660	1500	1083
Módulo de elasticidade térmica (L/°C)	23.1 0-6	11,7.10-6	16,5.1 0-6
Condutibilidade térmica 250C (Cal/cm/°C)	0,53	0,12	0,94
Condutibilidade elétrica (%IACS)	61	14,5	100

Fonte: Elaborada pelo autor do trabalho.

É notório que a combinação das propriedades do alumínio faz dele um metal de ampla aplicação em vários setores da sociedade, no entanto, o alumínio não aparece na sua forma elementar na natureza, logo, é necessário combinar com outros elementos formando o composto.

Desse modo, os principais insumos para produção do alumínio primário durante o processo de Redução como alumina, criolita, fluoreto de alumínio, coque de petróleo e outros dissolvidos no banho eletrolítico contido nas cubas eletrolítica apresentam reações como a formação de gases na fabricação do alumínio primário.

Essas reações ou efeitos são indesejados, isto é, gases de efeito estufa que tem uma vida atmosférica muito longa, gases perfluorcarbonos ou PFCs, por também provocar em perda de eficiência no processo de fabricação e aumento de consumo de energia.

Os gases PFCs liberados pela indústria de alumínio acontece eventualmente durante o processo de redução de alumínio primário, em eventos chamados efeito anódico. Os efeitos anódicos são provocados quando há uma porcentagem baixa de alumina (Al_2O_3) dissolvida no banho eletrolítico, ocasionando elevada resistência elétrica e aumento de voltagem (acima da operação normal).

No processo de redução de alumínio, encontram-se controles e tecnologias de redução de alumínio que colaboram de forma operacional para a diminuição de emissões de gases, tais como (ROCHA, 2010):

- a) Controle do nível de banho eletrolítico;
- b) Tecnologia e procedimento para controle dos efeitos anódicos;
- c) Tecnologia e procedimentos para predição e supressão dos efeitos anódicos;
- d) Tecnologia e procedimentos de adição de alumina nas cubas eletrolíticas;
- e) Formulação dos insumos componentes do anodo;
- f) Tecnologia e procedimentos para colocação e remoção de pinos no anodo;
- g) Eficiência do sistema de exaustão das cubas eletrolíticas.

Segundo ANNA NEVES (2014), essas ações resultam, aproximadamente, uma redução 78% das emissões de PFCs no panorama das indústrias brasileiras, com investimento na modernização tecnológica dos aspectos acima.

Portanto, o uso de RNAs e técnicas de agrupamento podem ajudar a reduzir as emissões, contribuindo para o processo de produção de alumínio primário e,

assim, ajudam positivamente para antecipar e suprimir a ocorrência dos efeitos anódicos.

2.2 Efeito Anódico

O chamado efeito anódico, é a formação de gases, durante o processo de transformação da alumina em alumínio, como o gás carbônico (CO_2), o metano (CH_4) e os gases perfluorcarbonos, conhecidos como PFCs, que são mais potentes que o gás carbônico (CO_2) para formação do efeito estufa (ALCOA, 2020).

O processo de fabricação do alumínio é responsável por uma enorme quantidade de emissão de gases causadores do efeito estufa. Esse fenômeno natural é de extrema importância para a sobrevivência, pois assegura as temperaturas médias globais, garantindo o equilíbrio térmico e condições para o desenvolvimento dos seres vivos. Basicamente, o sol desprende calor à terra. Parte desse calor é absorvido pela superfície terrestre e pelos mares, e o restante é retornado para o espaço. Entretanto, parte da emissão solar irradiada pela superfície os gases retêm a energia solar na atmosfera impedindo que esse calor seja devolvido totalmente ao espaço (Brasil Escola, 2020). (Ver Figura 4).

Figura 4 - Efeito Estufa



Fonte: Brasil Escola (2020)

A maioria dos gases do efeito estufa são gerados naturalmente, porém os gases mais prejudiciais são os originados pelo homem (Ver Tabela 2).

Tabela 2 - Principais Gases de Efeito Estufa

Gases do Efeito Estufa	Principais Atividades para emissão
Dióxido de Carbono	Queimas de combustíveis fósseis
Gás Metano	Aterros sanitários e lixões
Óxido Nitroso	Atividades agrícolas
Hidrofluorcarbonetos	Sistema de aquecimento e refrigeração
Hexafluoreto	Indústria eletrônica
Perfluorcarbono	Produção de alumínio
Clorofluorcarbonos (CFCs)	composto baseado em carbono que contem cloro e flúor

Fonte: Elaborada pelo autor do trabalho.

Ao longo dos anos, há um crescimento considerável da concentração de gases do efeito estufa e tem provocado mudanças irreversíveis no comportamento climático do planeta a saber (ABAL, 2019):

- a) Elevação na frequência de secas;
- b) Derretimento das calotas polares e aumento do nível do mar;
- c) Extinção de espécies e danos a diversos ecossistemas;
- d) Diminuição a reserva de água doce do mundo;
- e) Agravamento da segurança alimentar, prejudicando as coletas na agricultura e a pesca.

Por isso, líderes de todo planeta estão empenhados em ações para reduzir progressivamente a emissão desses gases de efeito estufa, que já é uma realidade em alguns países, com uma das possibilidades do uso de fontes alternativas de energia, renováveis e limpas, trocando o uso de combustíveis fósseis, projetos que envolvem Inteligência Artificial (IA) para busca de descobertas de matérias mais duráveis e menos poluentes, estratégias de monitoramento e controle, questões como redução de desperdício e, além disso, atividades cotidianas como exemplo:

- a) Reduzir a utilização de transportes em pequenos trajetos;
- b) Optar pelo uso de bicicletas ou de transportes coletivo;
- c) Incentivar a coleta seletiva.

Nesse sentido, desde 2012 o IBAMA prevê que os processos de licenciamento ambiental de atividades capazes de emitir a chamada emissão de gases “fugitivos” do setor industrial deverão conter propostas pelos empreendedores de medidas para mitigarem suas emissões, em atenção aos compromissos assumidos pelo Brasil, que serão avaliados pelo órgão ambiental (IBAMA, 2012). Também como medida preventiva, em 2000, a ANP já havia lançado o “Programa de ajustes para redução da queima de gás natural na Bacia de Campos” (ANP, 2019). No caso da indústria de produção alumínio primário, as emissões fugitivas advêm principalmente na etapa de redução e podem ser atribuídas a: vazamentos em fornos e escapamento de gases.

2.3 Controle do Processo de cuba eletrolítica

A operação das cubas é realizada em relação ao banho eletrólito, no início do funcionamento, o insumo principal do banho é a criolita, mas durante as operações das cubas eletrolíticas, os insumos iniciais do banho sofrem alterações que necessitam constantemente ser ajustados para manter as condições adequadas ao processo eletrolítico. Este ajuste é feito pela adição programada de fluoreto de alumínio (AlF_3). O controle da cuba eletrolítica esta ligado diretamente com a temperatura de fusão do banho, sua eficiência de corrente e a resistividade do material. Caso a temperatura do banho comece a aumentar, é adicionado o Fluoreto de Cálcio (CaF_2) que atua como fundente, abaixando a temperatura de fusão do banho e assim mantendo as propriedades físico-químicas do banho.

A diminuição da eficiência de corrente está diretamente relacionada com o aumento da solubilidade do alumínio denominado de “ocorrência da reação de volta”, onde parte do alumínio formado é reoxidado segundo a equação:



A perda de alumínio causada pela reação inversa pode ser minimizada pela diminuição da temperatura de operação, utilizando um eletrólito com excesso de Fluoreto de alumínio (AlF_3).

Dado o exposto, a concentração de alumina dissolvida no banho é um dos parâmetros mais importantes para o controle da redução. Se esta quantidade é muito baixa, pode ocorrer o efeito anódico. Caso seja muito alta, a eficiência da célula diminui e a alumina não dissolvida é depositada no catodo, atuando como uma camada isolante elétrica, assim, aumentando o consumo de energia e causando problemas térmicos na cuba eletrolítica.

Desta forma, é muito importante que as variáveis: Voltagem Líquida, Fluoreto de Alumínio (AlF_3), Fluoreto de Cálcio (CaF_2), Temperatura de Banho, Intensidade de Corrente e Voltagem Bruta estejam presentes na modelagem dos algoritmos de agrupamento e Rede Neural Artificial.

3 APRENDIZAGEM DE MÁQUINA

3.1 Introdução

Este capítulo apresenta os principais conceitos de redes neurais artificiais e algoritmos de agrupamento. Inicialmente será abordado os conceitos de rede neural artificial com detalhes das suas principais características como: arquitetura e topologia.

Em seguida, serão descritos os preceitos sobre algoritmos de Clusterização, abordando alguns métodos de agrupamento de dados.

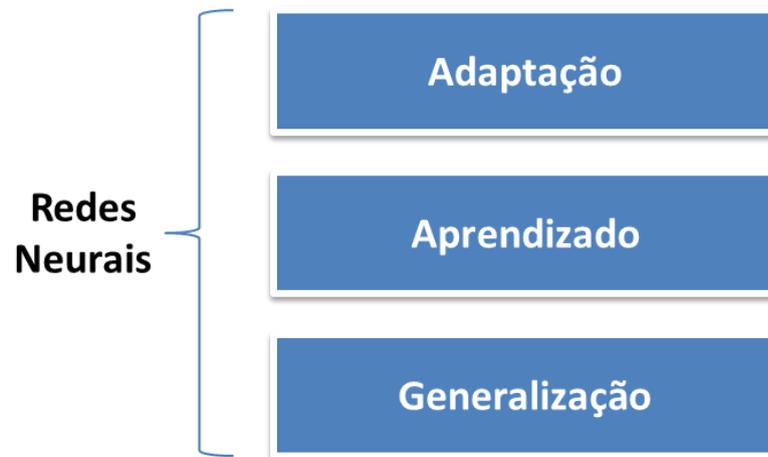
3.2 Rede Neurais Artificiais

Rede neurais artificiais são modelos computacionais inspirados no sistema nervoso de seres vivos, possuem a capacidade de aquisição, manutenção do conhecimento e podem ser definidos como um conjunto de unidades de processamento, caracterizadas por neurônios artificiais, que são interligados por um grande número de interconexões.

As Redes Neurais Artificiais (RNA's) visam, em sua maioria, solucionar problemas complexos, modelando sistemas através de conexões que possam simular o sistema nervoso humano, abrangendo a capacidade que o mesmo possui de aprender e agir perante as mais adversas situações apresentadas, adquirindo conhecimento por meio da experiência e da observação (HAYKIN, 2009).

A Figura 5 apresenta uma das principais características das redes neurais artificiais.

Figura 5 - Características da Rede Neural

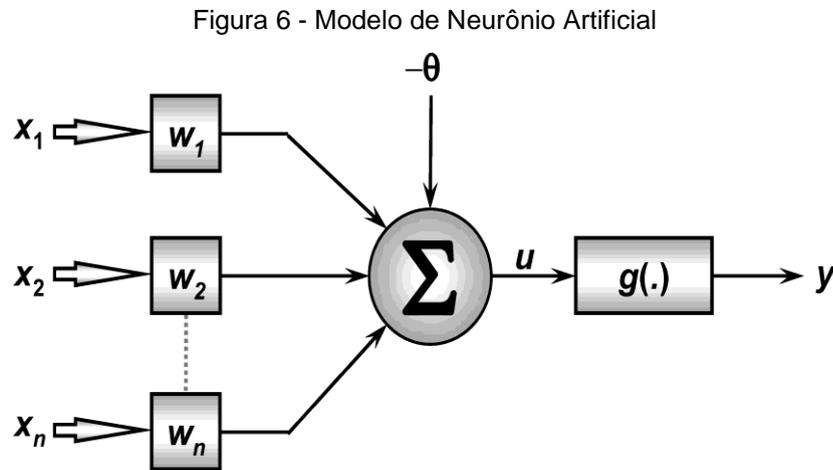


Fonte: Elaborado pelo autor do trabalho

- a) A adaptabilidade é a mudança das configurações internas da rede, como por exemplo os seus pesos sinápticos que são ajustados a partir da apresentação sucessiva de exemplos (dados, amostras, medidas) relacionados ao comportamento do problema;
- b) O aprendizado por intermédio da aplicação de um método de treinamento, a rede, esta consegue extrair o relacionamento existente entre as diversas variáveis que compõem o problema;
- c) Generalização: após o processo de treinamento da rede, esta consegue generalizar o conhecimento adquirido, estimando soluções que eram até então desconhecidas.

O neurônio é a unidade de processamento de uma rede neural. Cada unidade de processamento desempenha um papel muito simples. Cada ponto de entrada do neurônio recebe um valor. Os valores recebidos são ponderados e enviados para uma função matemática de soma, depois enviado para uma função de ativação. A saída da função é a resposta do neurônio que pode ser utilizada como entrada de

outros neurônios. A Figura 6 a seguir demonstra um modelo completo de um neurônio (SILVA, 2010).



Fonte: Silva (2010)

Os sinais de entrada $\{x_1, x_2, \dots, x_n\}$ são dados enviados ao neurônio externamente. Através dos pesos sinápticos $\{w_1, w_2, \dots, w_n\}$, esses sinais são ponderados. A seguir, o somador Σ (combinador linear) agrega todos os sinais de entrada com seus respectivos pesos sinápticos para gerar o potencial de ativação u .

O limiar de ativação \emptyset (também conhecido como bias) é um sinal externo que tem a função de aumentar ou diminuir a saída do combinador linear. O potencial de ativação u passa pela função de ativação g , responsável por restringir a amplitude do sinal de saída y do neurônio dentro de um intervalo pré-estabelecido. O resultado obtido pelo neurônio pode ser também apresentado pela Equação (3.1) e (3.2):

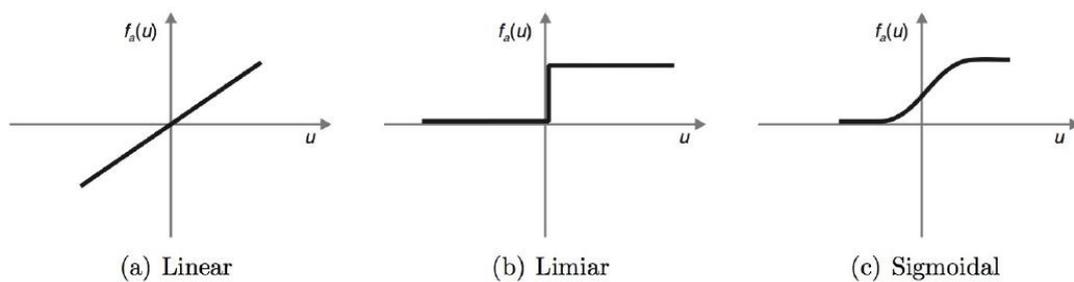
$$u = \sum_{i=1}^n w_i * x_i \tag{3.1}$$

$$y = g(u + \emptyset) \tag{3.2}$$

3.2.1 Funções de ativação

As funções de ativação são chamadas logo após a etapa realizada pelo somador Σ (combinador linear). A saída da função é a etapa final de um neurônio, que pode fazer parte do resultado final da rede, ou servir como entrada em outro neurônio. Várias funções de ativação têm sido propostas na literatura. A Figura 7 mostra o formato de três dessas funções, as funções linear, limiar e sigmoide.

Figura 7 - Funções de Ativação para uma Rede Neural



Fonte: Carvalho (2011)

Com já dito, a saída do neurônio é definida por meio da aplicação de uma função de ativação á entrada total. Na Figura 7 acima o uso da função linear (Figura 7(a)) implica retornar como saída o valor de u . Na função limiar (Figura 7(b)) o valor do limiar define quando o resultado da função limiar será igual a 1 ou 0. Na função sigmoide (Figura 7(b)) diferentes inclinações podem ser usadas. A função sigmoide representa uma aproximação contínua e diferencial da função limiar.

3.2.2 Treinamento

Para Braga (2000), uma das propriedades mais importantes de uma RNA é a capacidade de aprender por intermédio de exemplos e fazer inferências sobre o que aprendeu, melhorando gradativamente o seu desempenho. Esse processo se refere ao conhecimento obtido pela rede a partir de seu ambiente, sendo armazenado nas forças de conexão entre os neurônios artificiais – conhecidas como pesos sinápticos. O processo de treinamento de uma rede neural é denominado de algoritmo de aprendizagem e pode-se dividir em 2 grupos:

- a) **Treinamento supervisionado** – Este modelo de treinamento possui a presença de um “supervisor externo”, que conhece os rótulos de saída desejada para cada conjuntos de valores informados a rede. Com isso o supervisor externo pode avaliar a capacidade da hipótese. A partir de uma Tabela de dados de entrada e saída, a rede é capaz de aprender e estabelecer hipóteses sobre o sistema. Os pesos sinápticos são continuamente ajustados, buscando reduzir os desvios entre as saídas produzidas pela rede das saídas alvo desejadas;
- b) **Treinamento não supervisionado** - Diferente do método anterior, o treinamento não-supervisionado treina a rede fornecendo apenas dados de entrada. Os algoritmos de AM utilizados nessas tarefas não fazem uso do atributo de saída. Uma tarefa de agrupamentos de dados, por exemplo, tem por meta encontrar grupos de objetos semelhantes no conjunto de dados.

O aprendizado supervisionado é o mais comum no treinamento das RNA's. É chamado de aprendizado supervisionado porque a entrada e saída desejadas para a rede são fornecidas por um supervisor chamado de professor, externo. O objetivo é ajustar os parâmetros da rede, de forma a encontrar uma ligação entre os pares de entrada e saída fornecidos.

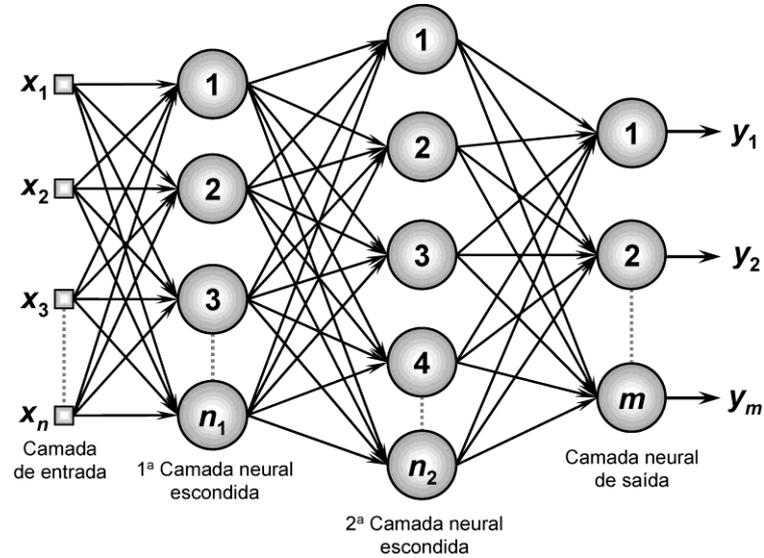
3.2.3 Arquitetura

Na arquitetura de uma RNA, os neurônios podem estar dispostos em uma ou mais camadas. O tipo de arquitetura escolhida está diretamente ligado ao algoritmo de treinamento adotado para a rede. Basicamente as RNAs são estruturadas em camadas: camada de entrada, camada(s) escondidas(s) e camada de saída.

A Figura 8 a seguir mostra o grafo da arquitetura de um perceptron de múltiplas camadas com duas camadas ocultas e uma camada de saída. A camada de entrada é responsável por receber os sinais do meio externo. Na camada

intermediária são extraídas as características relacionadas á aplicação da rede. A camada de saída gera o sinal de saída da RNA para o mundo externo

Figura 8 - Representação das camadas de um Rede Neural



Fonte: Silva (2010)

Podemos dizer que uma RNA é caracterizada por dois aspectos básicos: arquitetura e aprendizado. Enquanto a arquitetura estar relacionada ao tipo e número de unidades de processamento e à forma como os neurônios estão conectados, o aprendizado diz respeito às regras utilizadas para o ajuste dos pesos da rede e que informação é utilizada pelas regras (Silva, 2010).

3.2.4 FeedForward

As redes neurais FeedForward são os tipos mais comuns em aplicações reais. A rede consiste de um conjunto de neurônios que constituem a camada de entrada e uma camada de saída.

3.3 Algoritmos de Agrupamento

A Clusterização de Dados ou tarefa de aprendizado descritivo é uma técnica de mineração de dados multivariados que a partir das informações de entrada de um determinado contexto tem por objetivo agrupar automaticamente, por aprendizado não supervisionado, os exemplos da base de dados em determinado número de

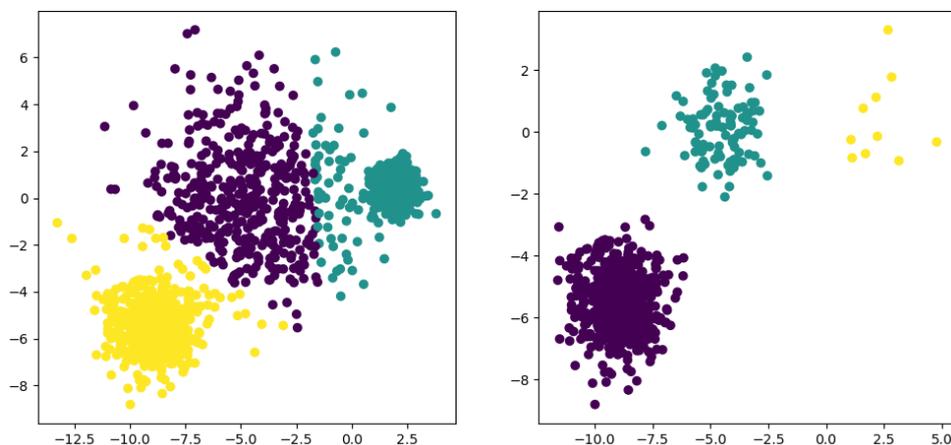
grupos, geralmente disjuntos denominados clusters ou agrupamentos. O aprendizado descritivo se referem a identificação de informação relevantes nos dados sem a presença de um elemento externo para guiar o aprendizado.

Os modelos não supervisionados utilizam uma abordagem diferente do método supervisionado. Isso ocorre porque como não há uma variável específica a ser explicada (ou seja, não há um Target), então não há sentido em treinar o conjunto de dados, pois também não será possível avaliar a assertividade do modelo, assim, pode-se dizer que o aprendizado de máquina não supervisionado procura encontrar padrões, perfis, itens semelhantes.

3.3.1 Algoritmo Particional - K-means

O K-means é um algoritmo particional que divide a base de dados em k-grupos, onde o número k é dado pelo usuário. Inicialmente, o algoritmo escolhe um determinado numero de objetos como sendo os centros dos k clusters. Cada objeto do conjunto de dados são alocados entre os k clusters de acordo com a medida de similaridade adotada, de modo que cada objeto fique no cluster que forneça a menor distancia entre o objeto e o centro do mesmo. Então, o algoritmo utiliza uma estratégia iterativa de controle para determinar que objetos devem mudar de cluster, para que a função objetivo usada seja otimizada. A Figura 9 ilustra o resultado da aplicação do algoritmo K-means ao conjunto de dados.

Figura 9 - Conjunto de dados com K-means



Fonte: Machine Learning in Python (2018)

O algoritmo K-means divide o conjunto de n amostras x para dentro k aglomerados disjuntos c , cada um descrito pela média μ_j das amostras no cluster. O algoritmo visa escolher centróides que minimizem a inércia ou o critério da soma dos quadrados dentro do cluster, que podem ser calculadas pela Equação 3.3:

$$\sum_{i=0}^n \min_{\mu_j \in C} (\|x_i - \mu_j\|^2)$$

(3.3)

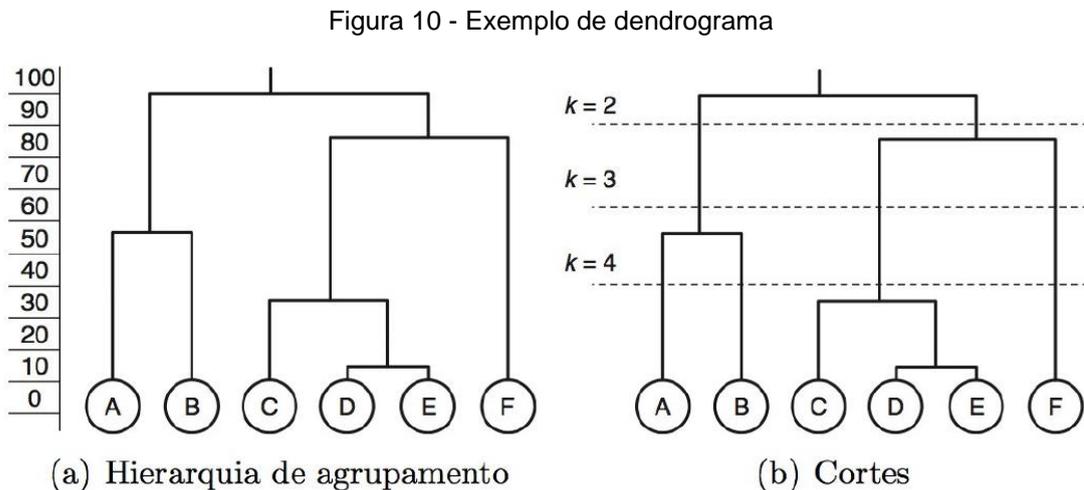
Em termos básicos, o algoritmo possui três etapas. O primeiro passo escolhe aleatoriamente quais amostras do conjunto de dados formarão os centróides iniciais. Após a inicialização, o K-means consiste em loop entre as duas outras etapas. No segundo passo o algoritmo atribui cada amostra ao centroide mais próximo. Na terceira etapa, cria novos centróides tomando o valor médio de todas as amostras atribuídas a cada centroide anterior. A diferença entre o antigo e o novo centroide é calculada e o algoritmo repete essas duas últimas etapas até que esse valor seja menor que um limite definido, ou seja, ele se repete até que os centróides não se movam significativamente (ARTHUR; VASSILVITSKII, 2007).

3.3.2 Algoritmo Hierárquico - Agrupamento Aglomerativo.

Algoritmos de clusterização baseados no método hierárquico organizam um conjunto de dados em uma estrutura hierárquica de acordo com a proximidade entre os dados. Os resultados de um algoritmo Hierárquico são normalmente mostrados como uma árvore binária ou dendrograma, que é uma árvore que iterativamente divide a base de dados em subconjuntos menores. A raiz do dendrograma representa o conjunto de dados inteiro e os nós folhas representam os indivíduos. O resultado da clusterização pode ser obtido cortando-se o dendrograma em diferentes níveis de acordo com o número de cluster k desejado. Esta forma de representação fornece descrições informativas e visualização para as estruturas de grupos em potencial.

A Figura 10 ilustra o funcionamento do algoritmo de agrupamento hierárquico do tipo aglomerativo baseado em métricas de integração que são os mais utilizados.

A Figura 10(a) ilustra um dendograma com a hierarquia de agrupamento para o conjunto de objetos. Cortes em cada nível do dendograma representam diferentes partições dos dados, com diferentes números de clusters. A Figura 10(b) ilustra cortes no dendograma representando as partições.



Fonte: Carvalho (2011)

O Aglomerativo Hierárquico executa um cluster hierárquico usando uma abordagem bottom/up: cada observação começa em seu próprio cluster e os clusters são mesclados sucessivamente. O critério de ligação determina a métrica usada para a estratégia de mesclagem (CARVALHO, 2011).

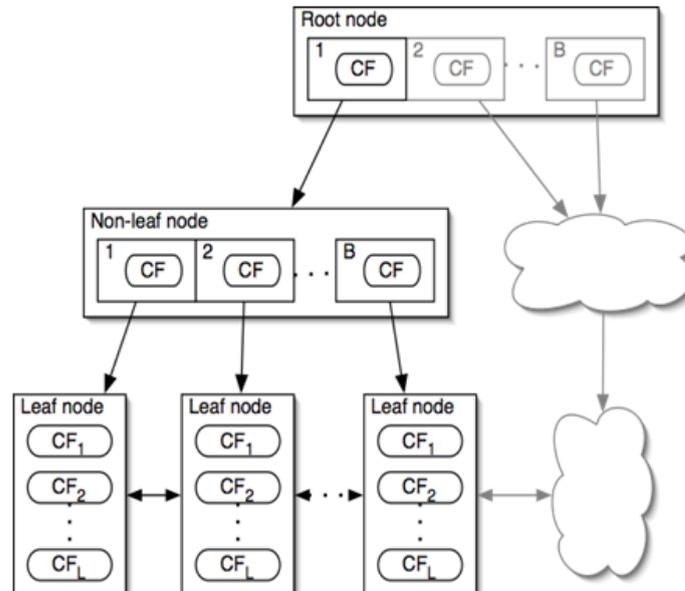
3.3.3 Agrupamento Bich.

Os métodos de agrupamento de dados existentes não tratam adequadamente o problema de processar grandes conjuntos de dados com uma quantidade limitada de recursos. Consequentemente, à medida que o tamanho do conjunto de dados aumenta, eles são reduzidos em termos de tempo de execução e qualidade dos resultados.

O BIRCH tenta minimizar os requisitos de memória de grandes conjuntos de dados resumindo as informações contidas nas regiões densas como entradas do Recurso de Cluster (CF) (MAKLIN, 2019).

A Figura 11 ilustra a construção da árvore CF que é uma representação muito compacta do conjunto de dados porque cada entrada em um nó folha não é um único ponto de dados, mas um sub-cluster.

Figura 11 - Construção da Arvore CF



Fonte: MAKLIN (2019)

Este algoritmo primeiro gera um resumo mais compacto que retém o máximo possível de informações dos dados e, em seguida, agrupando o resumo de dados em vez do conjunto de dados original.

A seguir uma explicação mais detalhada sobre as etapas de processamento do algoritmo Bich.

Na primeira fase, o Bich cria uma estrutura de dados em árvore CF com altura balanceada. Na segunda etapa, o algoritmo percorre todas as entradas da folha na árvore CF inicial para reconstruir uma árvore CF menor. Na etapa três, um algoritmo de agrupamento existente é usado para agrupar todas as entradas de folha. Aqui, um algoritmo de cluster hierárquico aglomerado é aplicado diretamente aos sub-cluster representados por seus vetores CF. Ele também fornece a flexibilidade de permitir que o usuário especifique o número desejado de clusters ou o limite de diâmetro desejado para clusters. Assim é obtido um conjunto de clusters que captura o padrão de distribuição principal nos dados.

Diante dos modelos apresentados, esta dissertação propõe a utilização dos algoritmos de agrupamento K-means, Aglomerativo e Bich junto com a rede neural de múltiplas camadas com algoritmo de treinamento “Adam”. A tarefa de clusterização será responsável por identificar as categorias de “efeito anódico” e “não efeito anódico”, sem a presença de um elemento externo para guiar o aprendizado. A partir dos rótulos gerados pelo algoritmo de agrupamento, uma rede Neural Artificial é treinada por um algoritmo supervisionado.

3.3.4 Avaliação para algoritmos de agrupamento

O objetivo dos algoritmos de agrupamento é encontrar uma estrutura de clusters nos dados em que os objetos pertencentes a cada clusters compartilham alguma característica de similaridade relevante para problema em estudo. Porém não se sabe se a solução é representativa para o conjunto de dados analisados.

A análise e a comparação de algoritmos de agrupamento são tarefas complexas, pois dependem muito do domínio da aplicação e das técnicas de agrupamento empregadas (CARVALHO, 2011).

Sendo assim, existem diversas técnicas na análise do desempenho de algoritmos de agrupamento que possibilitam entender o grau de escalabilidade/eficiência para o conjunto de dados analisado, a robustez relativa a ruídos e outliers e se para cada execução diferente do algoritmo os dados são alocados aos mesmos clusters.

O uso de critérios de avaliação nos permite conhecer o contexto nos quais um algoritmo funciona melhor do que o outro. A validação de um agrupamento e com base em índices estatísticos, que julgam, de maneira quantitativa e objetiva, o mérito das estruturas encontradas.

Utilizaremos nesse trabalho três tipos de critérios de avaliação externa sendo: Adjusted Mutual Information (AMI) (VINH; EPPS; BAILEY, 2010), Homogeneity Metric, Completeness Metric (ROSENBERG; HIRSCHBERG, 2007) e Índice Rand (HUBERT; ARABIE, 1985).

A Adjusted Mutual information é um ajuste da pontuação da Mutual Information (MI) para levar em conta o acaso. Essa métrica é independente dos

valores absolutos dos rótulos. Para dois agrupamentos U e V, o AMI é dado pela equação 3.4:

$$AMI = \frac{MI - E[MI]}{\text{mean}(H(U), H(V)) - E[MI]} \quad (3.4)$$

Onde sua entropia é dada pela equação 3.5:

$$H(U) = - \sum_{i=1}^{|U|} P(i) \log(P(i)) \quad (3.5)$$

Onde $P(i)$ é a probabilidade de um objeto escolhido aleatoriamente de U cair na U_i , da mesma forma para V, conforme a equação 3.6:

$$H(V) = - \sum_{j=1}^{|V|} P'(j) \log(P'(j)) \quad (3.6)$$

A Mutual Information (MI) entre U e V é calculada por:

$$MI(U, V) = \sum_{i=1}^{|U|} \sum_{j=1}^{|V|} P(i, j) \log \left(\frac{P(i, j)}{P(i)P'(j)} \right) \quad (3.7)$$

Onde $P_{(i,j)}$ é a probabilidade de um objeto escolhido aleatoriamente cair em ambas às classes U_i e V_j .

E finalmente usando o valor esperado, a informação mútua ajustada pode ser calculada usando a equação 3.8:

$$AMI = \frac{MI - E[MI]}{\text{mean}(H(U), H(V)) - E[MI]} \quad (3.8)$$

A Homogeneity Metric é utilizada para analisar a homogeneidade do agrupamento, a homogeneidade é considerada elevada se todos os clusters contiverem apenas pontos de dados que são membros de uma única classe. Enquanto que A Completeness Metric analisa se todos os membros de uma determinada classe são atribuídos ao mesmo cluster. O cálculo da Homogeneity Metric e Completeness Metric podem ser calculados pelas equações 3.9 e 3.10, respectivamente:

$$h = 1 - \frac{H(C|K)}{H(C)} \quad (3.9)$$

$$c = 1 - \frac{H(K|C)}{H(K)} \quad (3.10)$$

Onde $H(C|K)$ é a entropia condicional das classes dada às atribuições dos clusters, representada pela equação 3.11:

$$H(C|K) = - \sum_{c=1}^{|C|} \sum_{k=1}^{|K|} \frac{n_{c,k}}{n} \cdot \log \left(\frac{n_{c,k}}{n_k} \right) \quad (3.11)$$

E $H(C)$ é a entropia das classes e é dada por:

$$H(C) = - \sum_{c=1}^{|C|} \frac{n_c}{n} \cdot \log \left(\frac{n_c}{n} \right) \quad (3.12)$$

Onde o n é o número total de amostras, n_c e n_k o número de amostras pertencentes respectivamente à classe C e do cluster k , e finalmente $n_{c,k}$ o número de amostras da classe C atribuído ao cluster k .

O Índice Rand calcula a similaridade entre dois agrupamentos considerando todos os pares de amostras, que pode ser dada pela equação 3.13:

$$RI = \frac{a + b}{C_2^{n_{samples}}} \quad (3.13)$$

Onde a representa o número de pares de elementos que estão no mesmo conjunto em C e no mesmo conjunto em K e b o número de pares de elementos que estão em conjuntos diferentes em C e em conjuntos diferentes em K , e finalmente C_2^n que é o número total de pares possíveis no conjunto de dados.

4 USO DE REDES NEURAIS E CLUSTERIZAÇÃO NA PREDIÇÃO DO PRE-EFEITO ANÓDICO NO PROCESSO DE REDUÇÃO DO ALUMÍNIO.

4.1 Aprendizado

Partindo do estudo de Redes Neurais Artificiais (RNAs), as mesmas foram utilizadas com o fim de modelar a predição do pré-efeito anódico em Python, com a biblioteca específica chamada Keras (KERAS DOCUMENTATION, 2020). Uma das características importantes do uso das RNAs é treinar a rede neural com os dados de entrada e sua classe categóricas previamente disponíveis, assim, a rede ajusta seus pesos a cada interação, baseado nos rótulos de saída de cada registro.

Para os parâmetros da RNA (ver Tabela 3), o desafio aqui proposto é obter uma eficiência na predição supondo que essas classes de saídas não existam em um primeiro momento. Ou seja, realizamos primeiro um aprendizado não supervisionado para rotular nossos dados de entrada por meio das técnicas de agrupamento 'K-means', 'Agrupamento Aglomerativo' e 'Bich' e posteriormente treinamos a nossa rede neural com os rótulos definidos anteriormente, e por fim, avaliamos a saída da rede neural com as classes originais.

Tabela 3 - Parâmetros da RNA

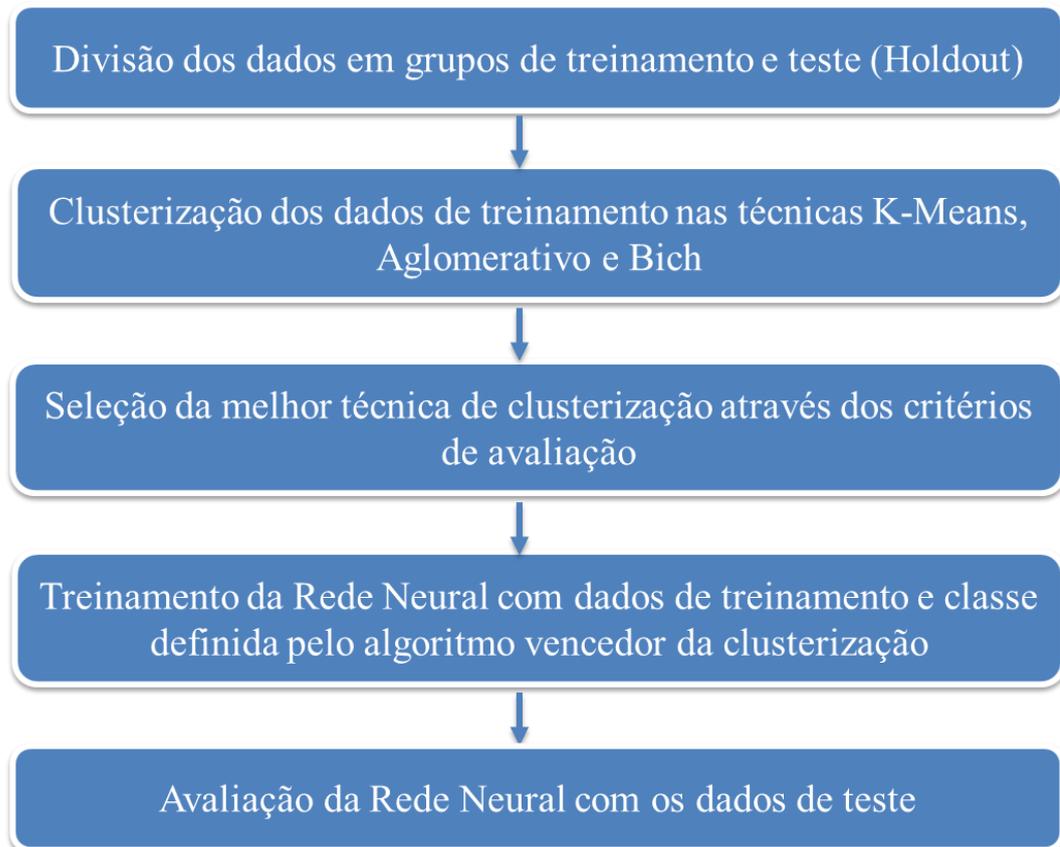
Arquitetura da Rede Neural	
Tipo de Rede	Rede de Múltiplas Camadas (MLP) com algoritmo de treinamento Adam.
Estrutura da Rede	A camada de entrada de dados é composta por 6 neurônios. Os dados são então processados e enviados para as camadas intermediárias e 1 neurônio na camada de saída.
Treinamento	Foi “treinada” com os dados de treinamento e seus respectivos rótulos, definidos pela melhor técnica de agrupamento sendo duas classes (0) indicador de “não pré-efeito anódico” e o numeral (1)

para “pré-efeito anódico”.

Fonte: Elaborado pelo autor do trabalho

Assim, a Figura 12 mostra o cenário adotado para o aprendizado:

Figura 12 - Fluxograma do Modelo Híbrido



Fonte: Elaborado pelo autor do trabalho

Na primeira etapa, é escolhido o melhor resultado das técnicas de Clusterização K-means, Agrupamento Aglomerativo e Birch, utilizadas neste trabalho e, que tem como objetivo a identificação de propriedades intrínsecas aos dados de entradas.

Na segunda etapa é utilizado uma Rede Neural Artificial com treinamento supervisionado, onde cada neurônio produzirá um valor de saída que é então comparado ao valor de saída especificado pelo cluster, estimado na etapa anterior.

Na última etapa, após o treinamento da rede, realizamos os testes de validação dos padrões treinados, e a partir disso, são calculados os erros da rede

neural artificial com as amostras da classe alvo original (dados originais). Os testes de treinamento efetuados correspondem a variável Voltagem Líquida, Fluoreto de Alumínio (no Banho), Fluoreto de Cálcio (% no Banho), Temperatura de Banho, Intensidade de Corrente, Voltagem Bruta e Pré-Efeito Anódico.

4.2 Resultados e Discussão

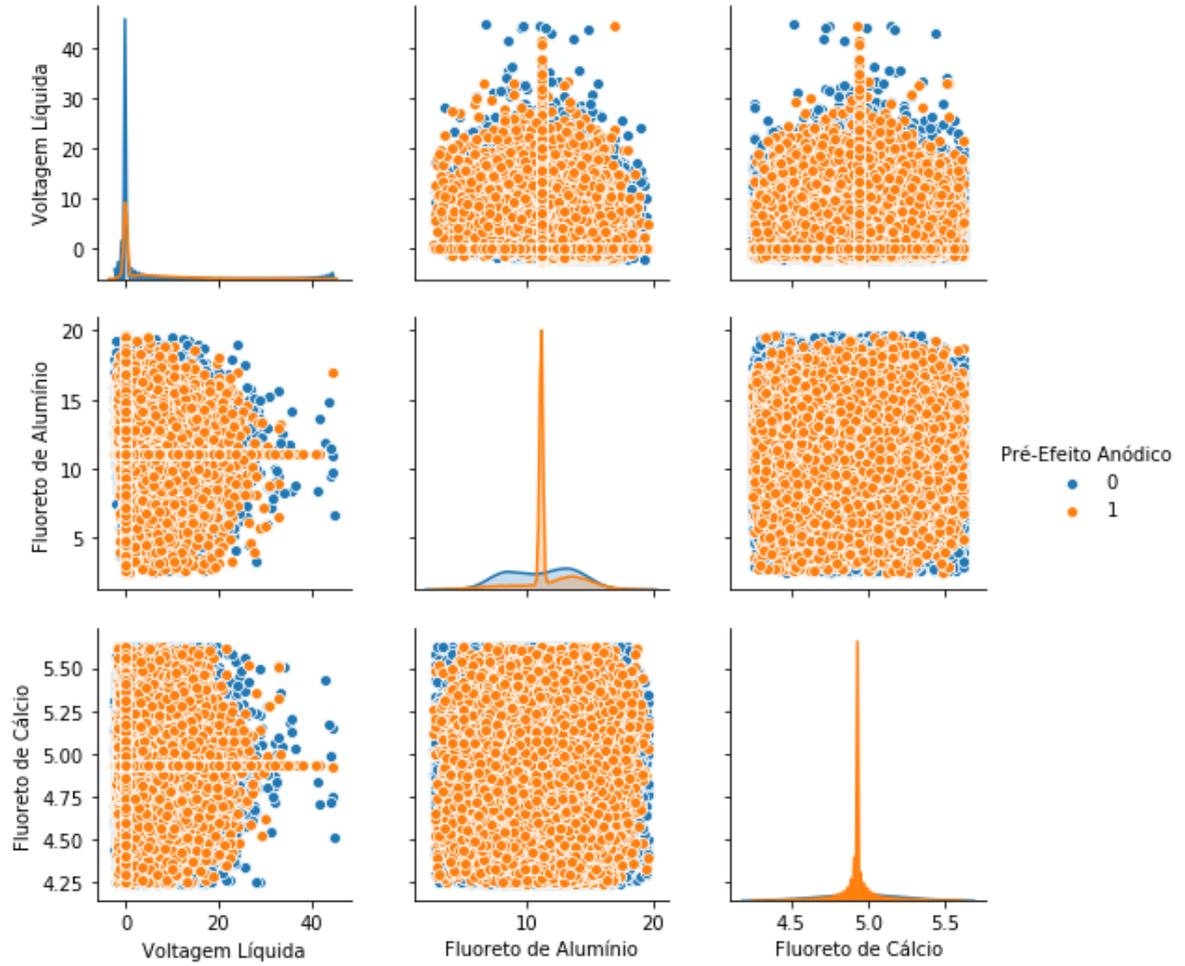
4.2.1 Base de Dados

A base de dados obtida é composta por dados multivariados, pois corresponde por 6 atributos de entradas: Voltagem Líquida, Fluoreto de Alumínio (no Banho), Fluoreto de Cálcio (% no Banho), Temperatura de Banho, Intensidade de Corrente e Voltagem Bruta. Cada objeto, constituído pelos atributos anteriores, são usados para descrever o comportamento da classe Pré-Efeito Anódico.

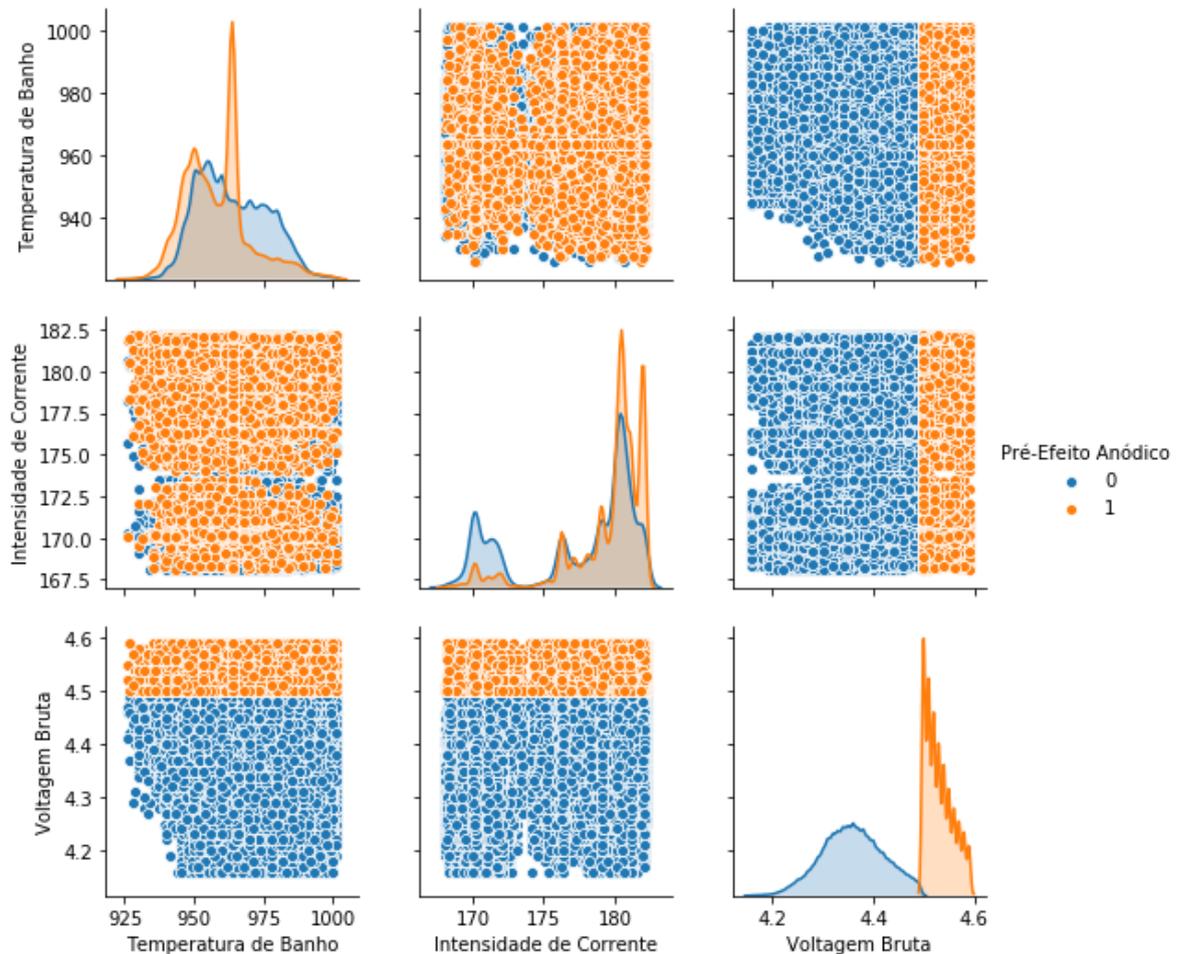
A fim de compreender a relação entre os atributos, a cada objeto, é associado um ponto em um plano bidimensional. Os valores dos atributos definem a coordenadas desse ponto.

As Figuras 13(a) e 13(b) ilustram a correlação para as diferentes combinações dos atributos do conjunto de dados.

Figura 13 - Matriz de scatter plot para 6 atributos do conjunto de dados



(a) Voltagem Líquida, Fluoreto de Alumínio e Fluoreto de Cálcio



(b) Temperatura Banho, Intensidade de Corrente e Voltagem Bruta

Fonte: Elaborada pelo autor do trabalho (2020)

Na Figura 13(a) e 13(b), as classes dos objetos 0 (Sem Pré-Efeito Anódico) em azul e 1 (Com Pré-Efeito anódico) em laranja são disponibilizados para investigar o grau com que 2 atributos estão separando as classes.

4.2.2 Pré-processamento de Dados

Inicialmente temos 209978 mil dados divididos em duas categorias, 0 (Sem Pré-Efeito Anódico) e 1 (Com Pré-Efeito anódico). Uma das formas simples de visualizar a distribuição dos dados é utilizar a técnica de visualização por boxplot.

Assim para a remoção de dados outliers foi escolhido o método Tukey. O método de Tukey ou mais conhecido como boxplot define limites inferior (Eq. 4.2) e

superior (Eq. 4.3) a partir do interquartil (*IQR*) (Eq. 4.1) e o primeiro e terceiro quartis.

$$IQR = Q3 - Q1 \quad (4.1)$$

$$L_{inf} = Q3 - (1.5 * IQR) \quad (4.2)$$

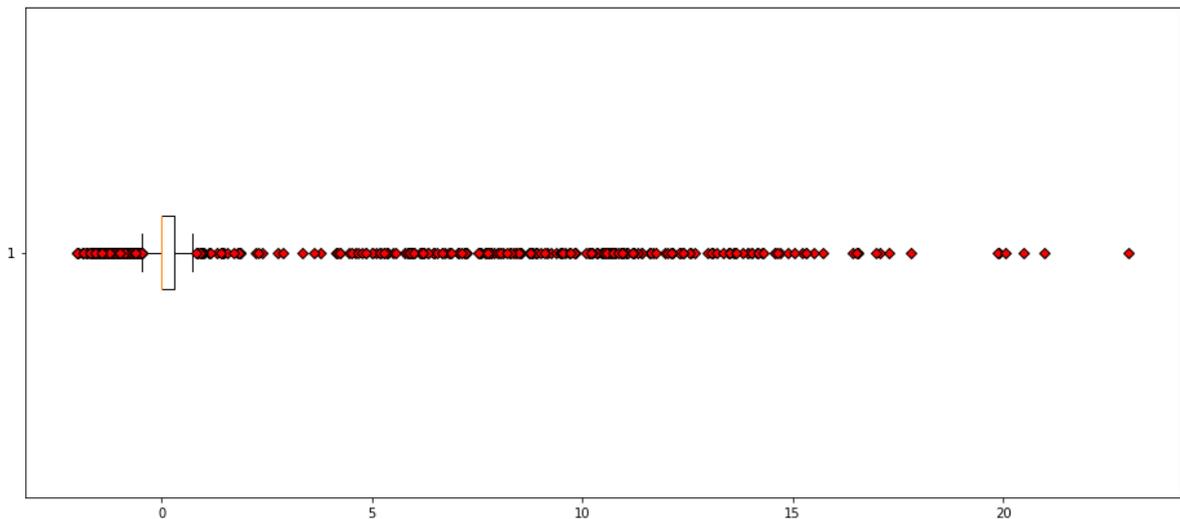
$$L_{sup} = Q3 + (1.5 * IQR) \quad (4.3)$$

Onde *Q1* é o primeiro quartil, *Q3* o terceiro quartil, *L_{sup}* limite superior e o *L_{inf}* limite inferior.

Os dados que estiverem fora desses limites serão considerados como outliers (ROUSSEEUW; HUBERT, 2017).

A Figura 14 ilustra o gráfico boxplot para o atributo Voltagem Líquida.

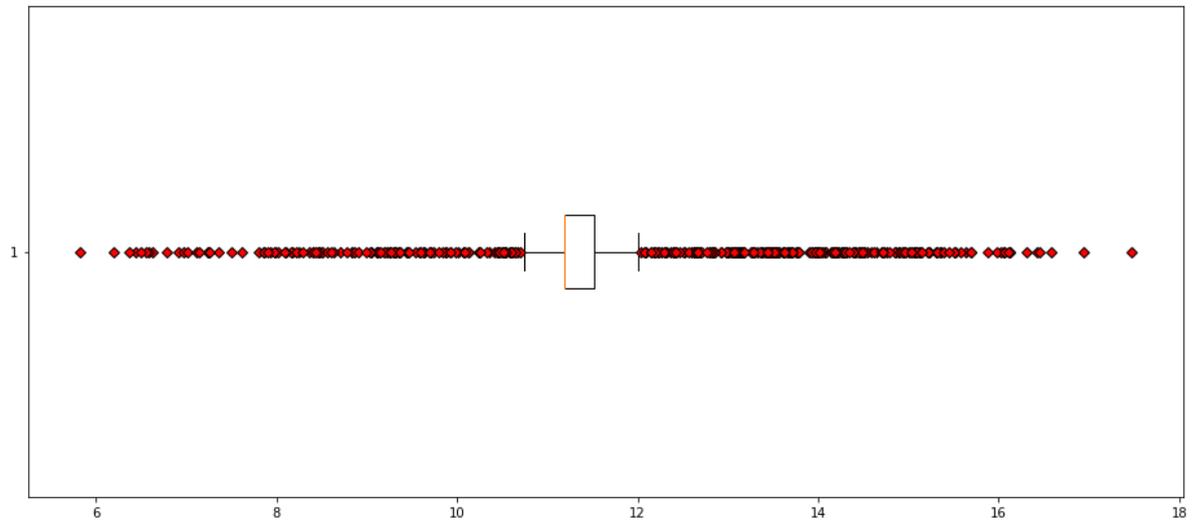
Figura 14 – Boxplot para o atributo Voltagem Líquida



Fonte: Elaborado pelo autor do trabalho

A Figura 15 ilustra o gráfico boxplot para o atributo Fluoreto de Alumínio.

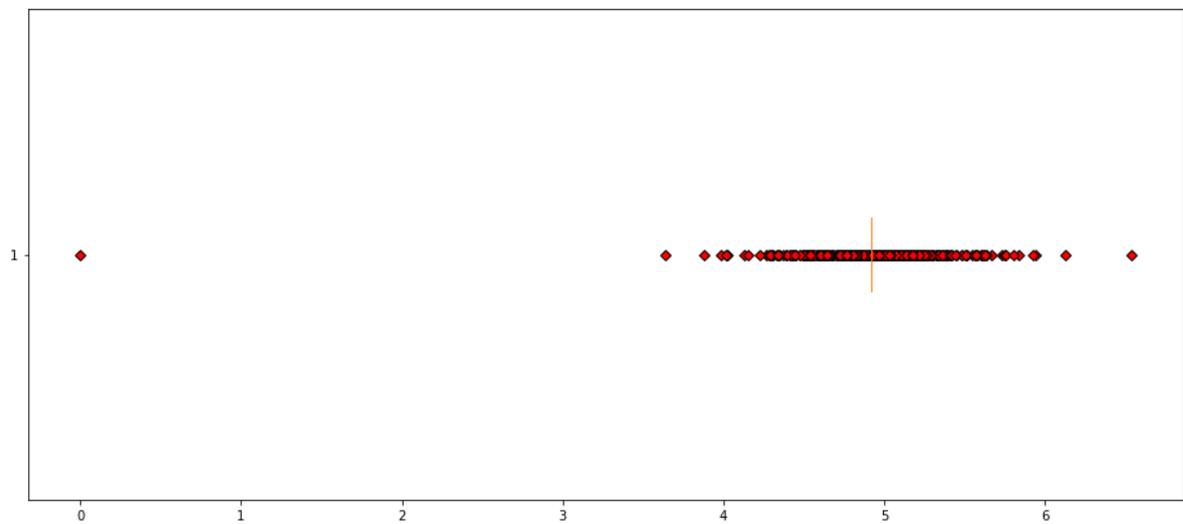
Figura 15 – Boxplot para o atributo Fluoreto de Alumínio



Fonte: Elaborado pelo autor do trabalho

A Figura 16 ilustra o gráfico boxplot para o atributo Fluoreto de Cálcio.

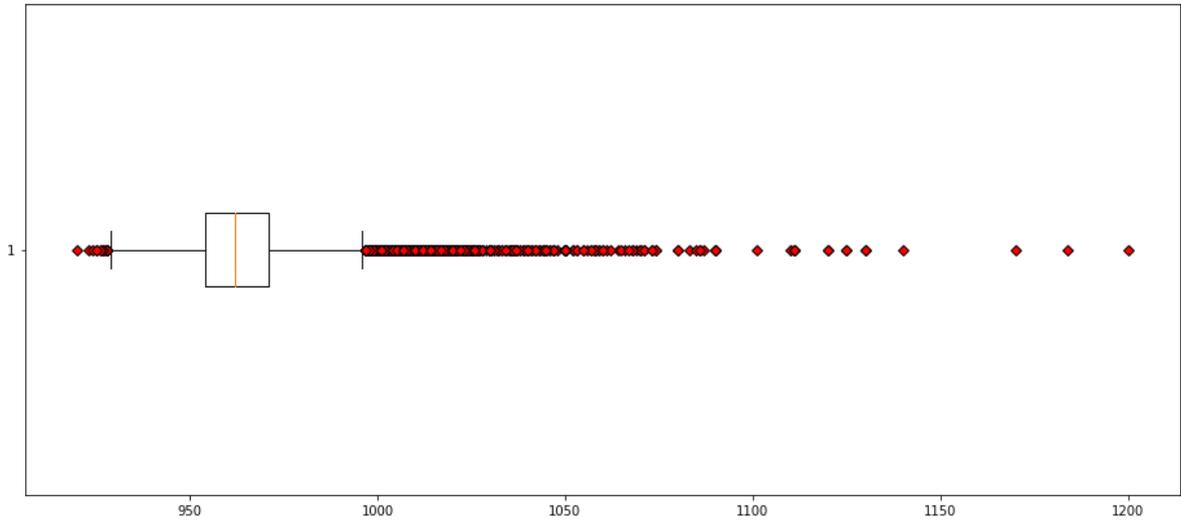
Figura 16 – Boxplot para o atributo Fluoreto de Cálcio



Fonte: Elaborado pelo autor do trabalho

A Figura 17 ilustra o gráfico boxplot para o atributo Temperatura de Banho.

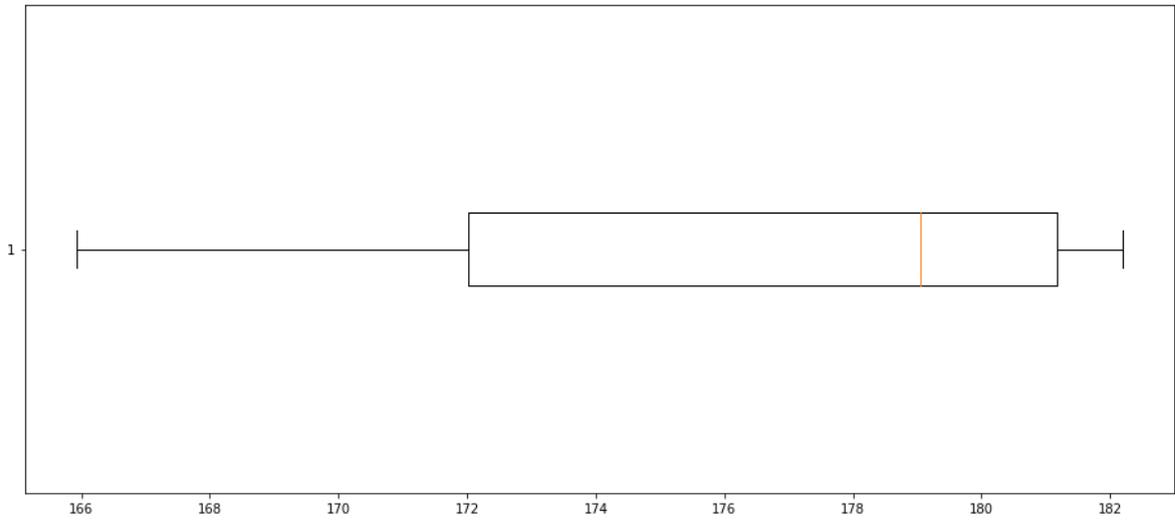
Figura 17 – Boxplot para o atributo Temperatura de Banho



Fonte: Elaborado pelo autor do trabalho

A Figura 18 ilustra o gráfico boxplot para o atributo Intensidade de Corrente.

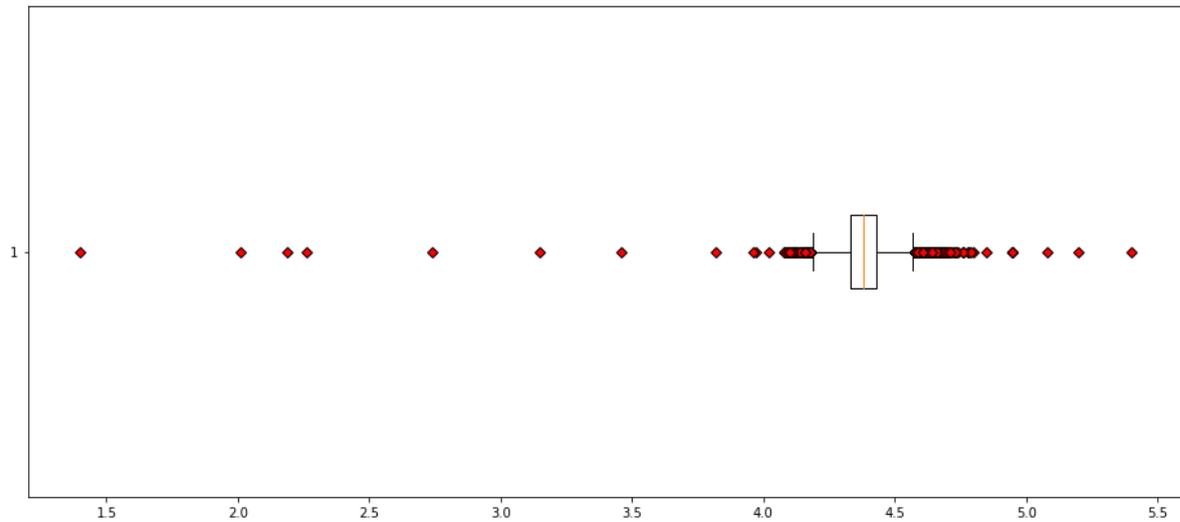
Figura 18 – Boxplot para o atributo Intensidade de Corrente



Fonte: Elaborado pelo autor do trabalho

A Figura 19 ilustra o gráfico boxplot para o atributo Voltagem Bruta.

Figura 19 – Boxplot para o atributo Voltagem Bruta



Fonte: Elaborado pelo autor do trabalho

Nos gráficos boxplots apresentados anteriormente, os lados esquerdo (inferior) e direito (superior) do retângulo representam o 1º quartil e o 3º quartil, respectivamente. A linha vertical no interior do retângulo é o 2º quartil, ou mediana. Os valores do limite superior (linha direita) e abaixo do limite inferior (linha esquerda) são considerados dados outliers (símbolos em losango).

A Tabela 4 apresenta os limites inferiores e superiores para realizar a remoção dos dados Outliers.

Tabela 4 – Limite Inferior e Superior de corte para Outliers

	Voltagem Líquida	Fluoreto de Alumínio	Fluoreto de Cálcio	Temperatura de Banho	Intensidade de Corrente	Voltagem Bruta
Limite Superior	--	17.9025	5.563	996.5	194.8	4.57
Limite Inferior	--	4.5465	4.283	928.5	158.48	4.17

Fonte: Elaborado pelo autor do trabalho

Logo, foi necessária a remoção desses outliers, uma vez que, apresentam um grande afastamento dos demais dados. A Tabela 5 mostra em detalhes as características dos dados antes e depois da remoção dos outliers.

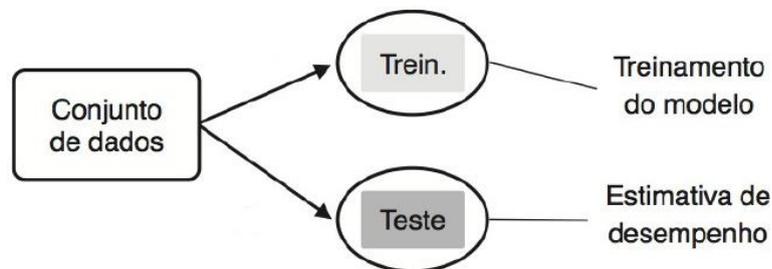
Tabela 5 - Conjunto de dados

Dados normais	Dados com Outliers Removidos	Classe
150 611	129 391	0 (Sem Pré-Efeito Anódico)
59 367	48 428	1 (Com Pré-Efeito Anódico)
209 978	177 819	Todos

Fonte: Elaborado pelo autor do trabalho

Após a remoção dos dados Outliers foi realizado a técnica de amostragem Holdout, onde os dados foram subdivididos em 75% para treinamento e 25% para teste. Os 75% dos dados foram utilizados para a fase de agrupamento e na indução da rede neural e, os exemplos de testes para a avaliação da rede neural, pois apresentam novos objetos ao modelo. A Figura 20 ilustra a divisão dos dados.

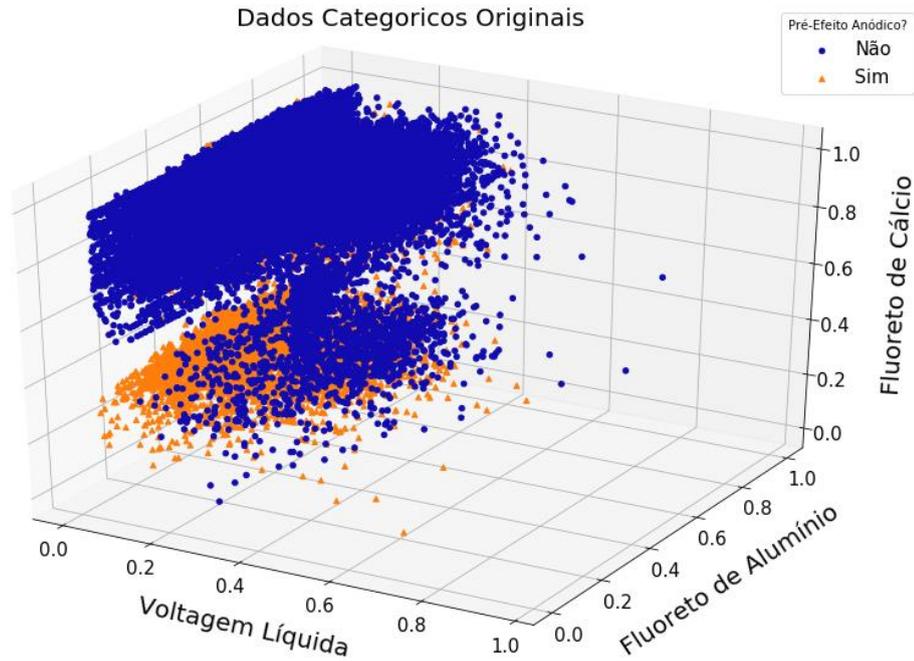
Figura 20 – Método de amostragem Holdout



Fonte: Carvalho (2011)

O mapa representado na Figura 21 é a representação gráfica simplificada do conhecimento dos dados categóricos de treinamento para três das seis variáveis utilizadas. O mapa pode ser visto como pontos de uma superfície onde representam as variáveis de entrada.

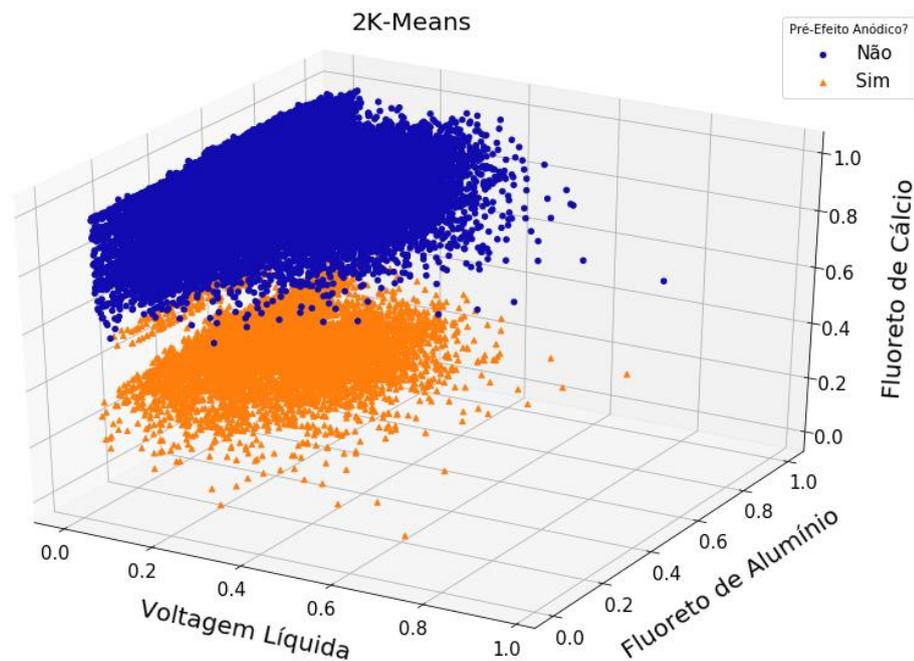
Figura 21 - Características da Base de Dados Originais



Fonte: Elaborado pelo autor do trabalho

A partir da base de dados, demonstrada anteriormente, foi utilizada a técnica K-means parametrizada em: 2 clusters e medida de distância euclidiana. A Figura 22 apresenta os resultados obtidos após a fase de agrupamento para 75% dos dados.

Figura 22 - Características da base de dados com K-means

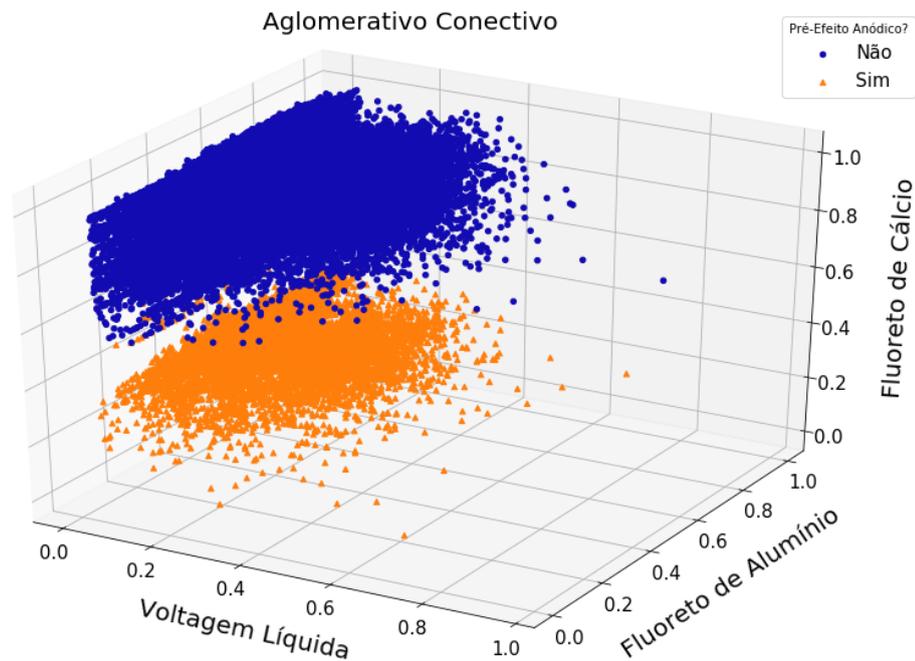


Fonte: Elaborado pelo autor do trabalho

A fim de verificar a desempenho do K-means, foi efetuado a clusterização utilizando o conjunto de treinamento descrito anteriormente com a técnica de Agrupamento Aglomerativo. O algoritmo foi então configurado para encontrar 2 cluster, medida de distancia euclidiana com critério de vinculação “Ward” (para minimiza a variação dos clusters que estão sendo mesclados) e adicionado restrição de conectividade. A Restrição de conectividade é imposta por meio de uma matriz de conectividade: uma matriz esparsa escassa que possui elementos apenas na interseção de uma linha e uma coluna com índices do conjunto de dados que devem ser conectados (SCIKIT-LEARN, 2018). Essas restrições são úteis para impor uma determinada estrutura local e tornar o algoritmo mais rápido comparado aos sem restrições, especialmente para o nosso número de amostras alto.

Na Figura 23 pode-se observar o resultado do experimento para o agrupamento Aglomerativo.

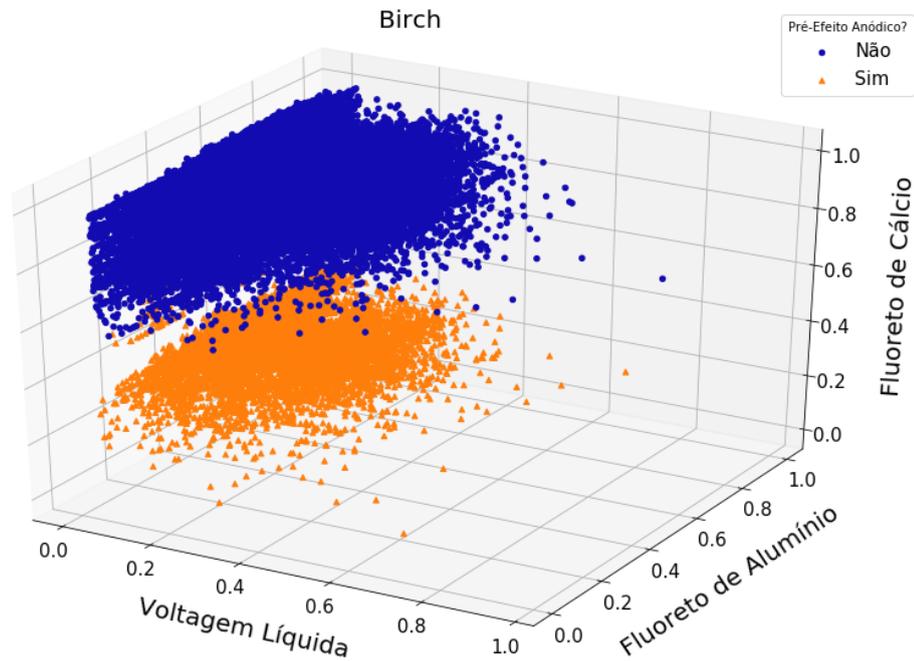
Figura 23 - Resultado da Base de Dados com Agrupamento Aglomerativo



Fonte: Elaborado pelo autor do trabalho

Quanto ao Agrupamento Birch (ver Figura 24), possui o fator de ramificação de 50 sub-cluster em cada nó e, configuração para 2 cluster.

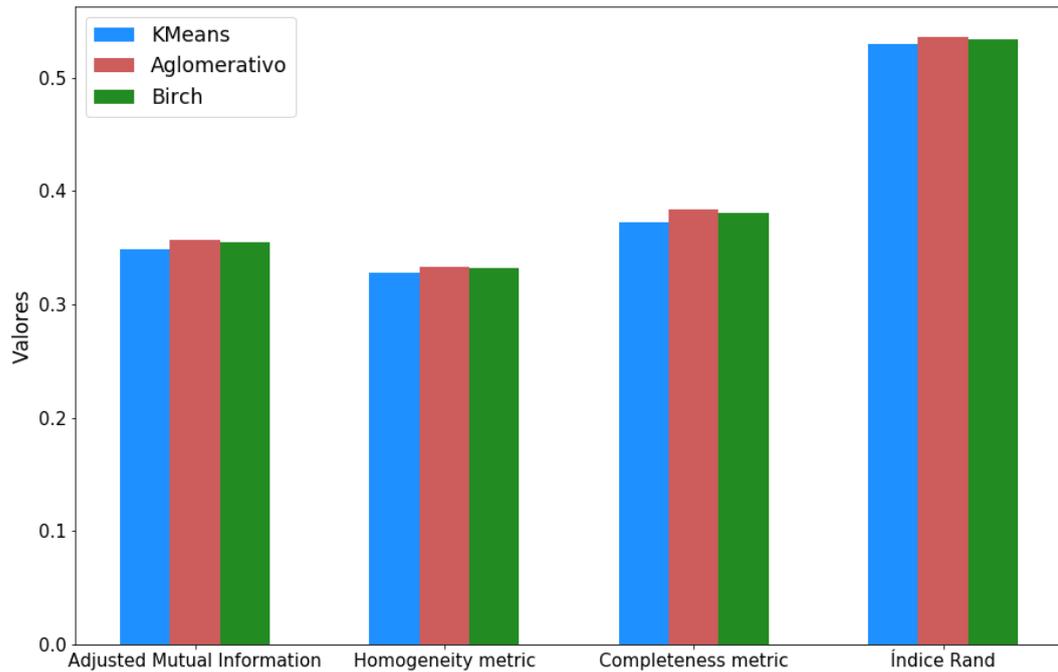
Figura 24 - Base de Dados com Agrupamento Bich



Fonte: Elaborado pelo autor do trabalho

Dentre os resultados, a técnica de clusterização que apresentou o resultado mais promissor foi o Agrupamento Aglomerativo com melhores resultados ao ser submetido nos critérios de avaliação Adjusted Mutual Information (AMI) (VINH; EPPS; BAILEY, 2010), Homogeneity Metric, Completeness Metric (ROSENBERG; HIRSCHBERG, 2007) e Índice Rand (HUBERT; ARABIE, 1985), conforme a Figura 25.

Figura 25 - Gráfico de Avaliações Externas



Fonte: Elaborado pelo autor do trabalho

A Tabela 6 ilustra os valores referentes aos gráficos anteriores para a avaliação dos agrupamentos

Tabela 6 – Resultados dos Critérios de avaliação

	Adjusted Mutual Information	Homogeneity	Completeness	Índice Rand
Aglomerativo	0.3566	0.3333	0.3834	0.5360
Bich	0.3488	0.3316	0.3812	0.5342
K-means	0.3547	0.3277	0.3729	0.5296

Fonte: Elaborado pelo autor do trabalho

Após a etapa de clusterização, seguimos para a construção do modelo da Rede Neural. O modelo preditivo será treinado com os dados de treinamento e com os rótulos do algoritmo de agrupamento Aglomerativo.

Durante o desenvolvimento da rede neural foram realizados diversos treinamentos variando-se o número de camadas intermediárias, bem como o número de neurônios em cada camada. A métrica de comparação do desempenho para a escolha da melhor topologia foi o erro médio absoluto.

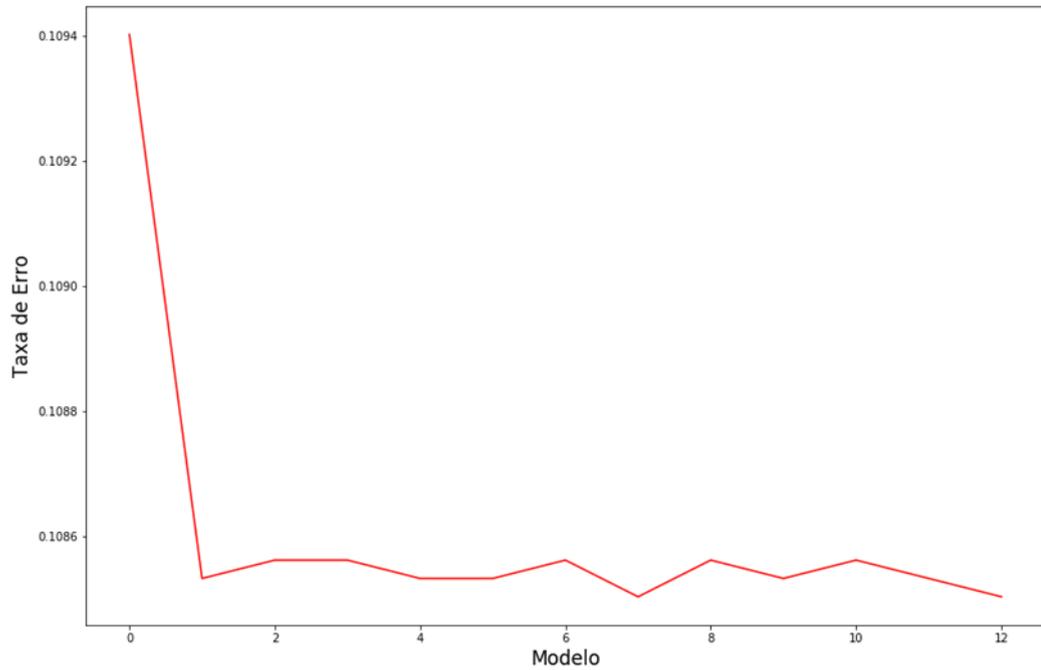
Para a topologia da Rede Neural Multicamada Direta as execuções foram realizadas considerando as seguintes topologias:

- a) Uma camada intermediária com variação da quantidade de neurônio em 3, 6, 9, 12, 15, 18, 21, 24, 27, 30, 33, 36 e 39 desta camada;
- b) Duas camadas intermediárias com variação da quantidade de neurônio na primeira camada em 3, 6, 9, 12, 15, 18, 21, 24, 27, 30 e na segunda camada em 3, 6, 9, 12, 15, 18, 21, 24, 27 e 30;
- c) Três camadas intermediárias com variação da quantidade de neurônio na primeira camada em 7, 15, 24, 32, 40 e na segunda camada em 5, 12, 19, 26, 33 e na terceira camada em 4, 12, 20, 28, 36, 28, 32;
- d) Quatro camadas intermediárias com variação da quantidade de neurônio na primeira camada em 1, 7, 15, 23, 31, 39 e na segunda camada em 5, 12, 19, 26, 33 e na terceira camada em: 1, 9, 17, 25, 33, 41 e na quarta em 4, 11, 18, 14, 21, 28 ;
- e) Quinta camada intermediária com variação da quantidade de neurônio na primeira camada em 1, 7, 15, 23, 31, 39 e na segunda camada em 5, 12, 19, 26, 33 e na terceira camada em: 1, 9, 17, 25, 33, 41 e na quarta em 4, 11, 18, 14, 21, 28 e na quinta camada em 2, 9, 16, 33, 40, 47 ;

As Figuras 26, 27, 28, 29, 30 ilustram as diversas topologias e suas taxas de erro sobre o conjunto de dados de teste.

A Figura 26 demonstra as taxas de erro variando-se os números de neurônios em somente uma camada intermediária.

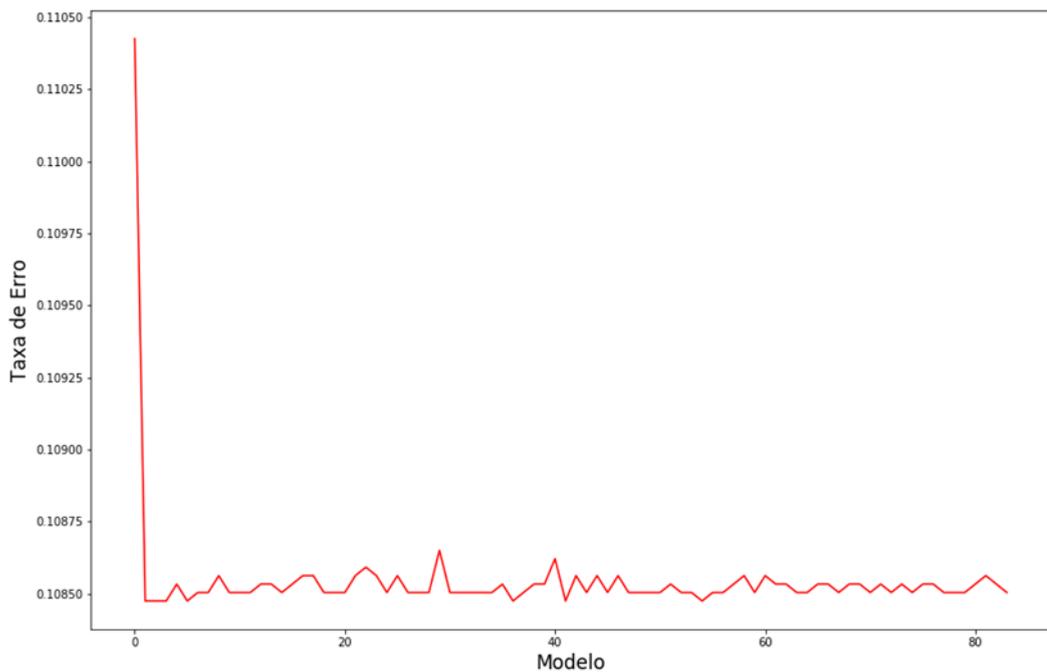
Figura 26 - Taxa de Erro variando-se o numero de neurônios para a primeira camada intermediaria



Fonte: Elaborado pelo autor do trabalho

A Figura 27 ilustra as taxas de erro das diversas redes neurais com somente 2 camadas intermediarias e variações de neurônios nessas camadas .

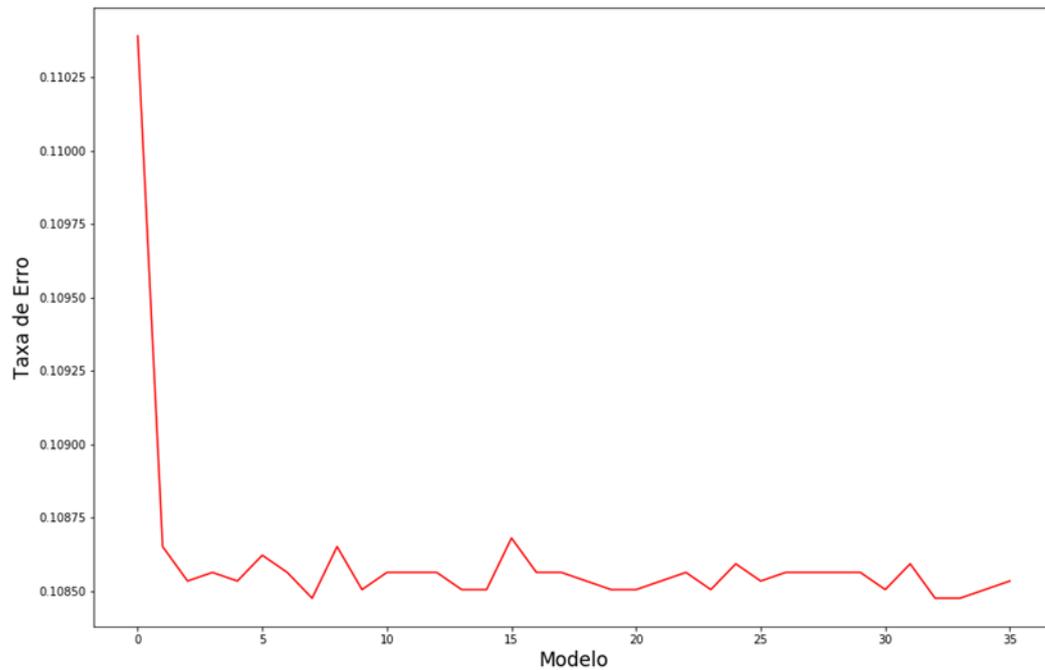
Figura 27 - Taxa de Erro variando-se o numero de neurônios na primeira e segunda camada intermediaria



Fonte: Elaborado pelo autor do trabalho

De forma semelhante, a Figura 28 demonstra as topologias para 3 camadas intermediárias.

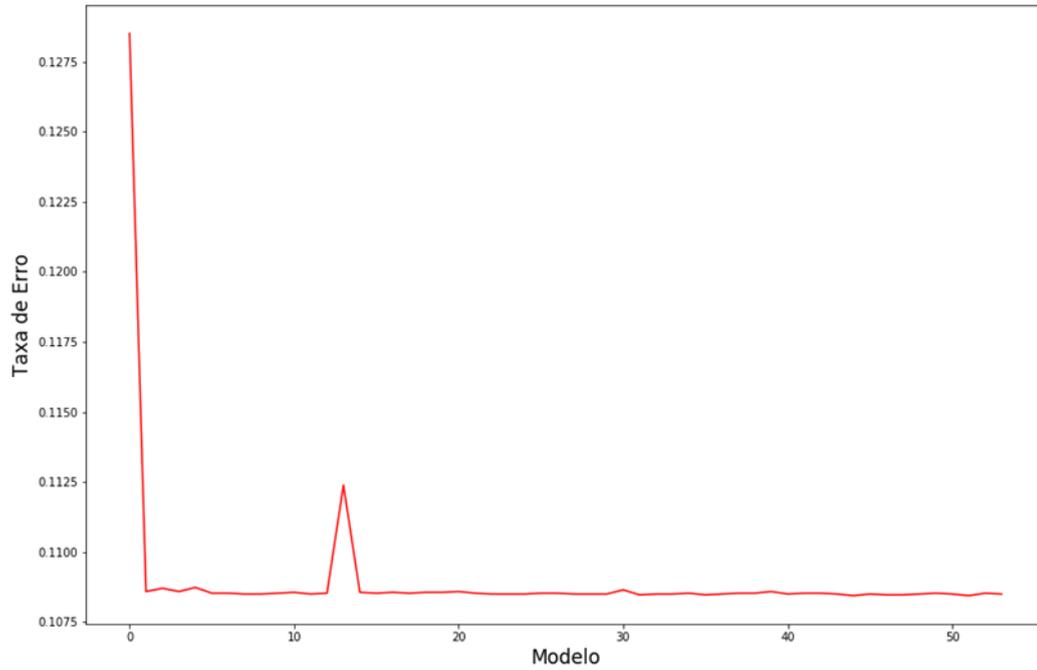
Figura 28 - Taxa de Erro variando-se o numero de neurônios na primeira, segunda e terceira camada intermediária



Fonte: Elaborado pelo autor do trabalho

A Figura 29 representa as topologias com 4 camadas intermediárias.

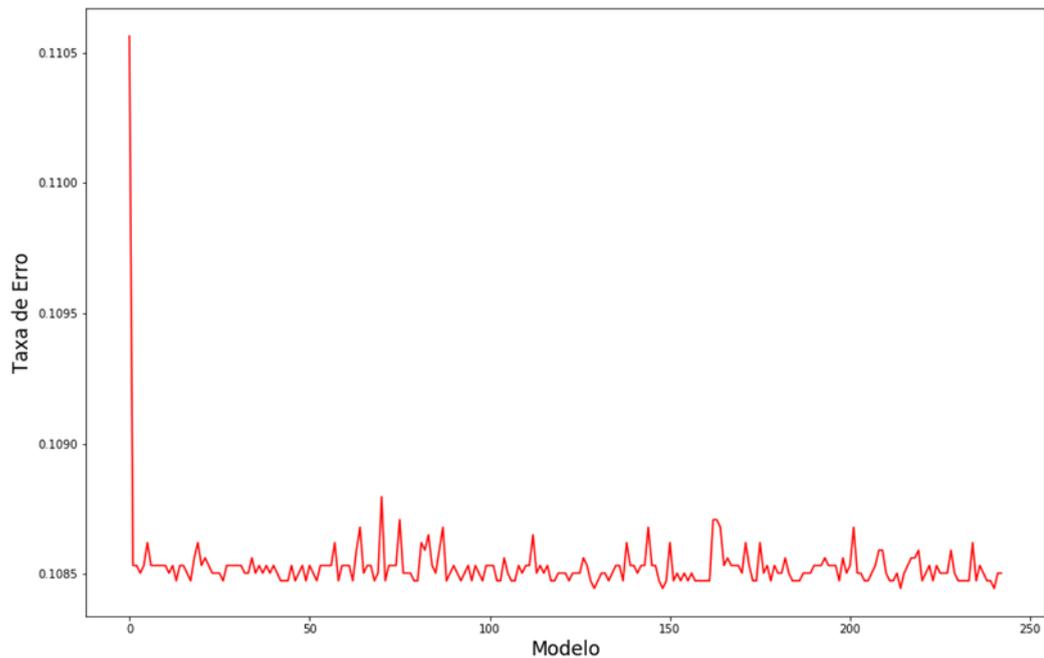
Figura 29 – Taxa de Erro variando-se o número de neurônios na primeira, segunda, terceira e quarta camada intermediária



Fonte: Elaborado pelo autor do trabalho

E finalmente, a Figura 30 demonstra as taxas de erro de uma rede neural com 5 camadas intermediárias e suas respectivas variações de neurônios.

Figura 30 – Taxa de Erro variando-se o número de neurônios nas cinco camadas intermediárias



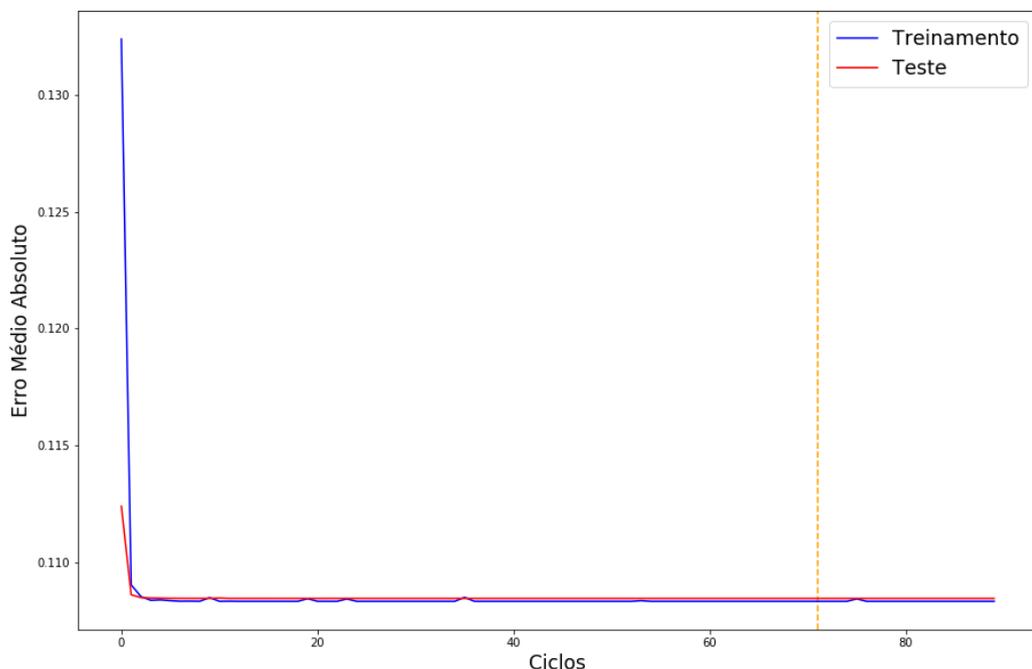
Fonte: Elaborado pelo autor do trabalho

A menor taxa de erro alcançada foi de 0.1084 e uma rede neural com 4 camadas intermediárias e 23, 12, 17, 11 neurônios em cada camada respectivamente.

Para cada modelo de rede neural, os ciclos de apresentação dos dados de treinamento são iterados até que seja atingido um critério de parada. Diferentes critérios de parada podem ser utilizados, como por exemplo, um número máximo de ciclos ou uma taxa máxima de erro. Neste trabalho, para reduzir a ocorrência de overfitting, inserimos concomitante aos dados de treino, o conjunto de teste à rede, em cada ciclo.

Durante o treinamento, para cada modelo, as taxas de erro para os dados de treinamento e teste tendem a cair. Assim, em um determinado momento a taxa de erro de validação pode começar a subir, é então, que é realizada a parada do aprendizado, pois a rede está se tornando superajustada. A Figura 30 ilustra o ciclo de treinamento da redes neural com a melhor topologia escolhida representada pela linha vertical tracejada em laranja.

Figura 31 - Ciclo de Treinamento da Rede Neural com melhor topologia



Fonte: Elaborado pelo autor do trabalho

Após a fase de treinamento da rede neural, esta será submetida a etapa de avaliação do algoritmo. A avaliação é realizada por meio da análise do desempenho do preditor, gerado por ele na rotulação do conjunto de teste.

A medida de desempenho empregada na avaliação do classificador é o cálculo de precisão, que pode ser calculado pela equação 4.4:

$$accuracy(y, \hat{y}) = \frac{1}{n} \sum_{i=0}^{n-1} 1(\hat{y}_i = y_i) \quad (4.4)$$

Onde \hat{y}_i é o valor previsto da i-ésima amostra e y_i é o valor verdadeiro correspondente ao longo das amostras n .

A Tabela 7 apresenta o resultado de desempenho da Rede Neural Artificial (treinada com os rótulos de saída, gerado pelo algoritmo de agrupamento Aglomerativo).

Tabela 7 – Resultados das Predições Neurais para os dados de teste

Modelo	Erro Médio Absoluto	Predição
Aglomerativo conectivo / Rede Neural	0.1084453818437473	0.8915546181562527

Fonte: Elaborado pelo autor do trabalho

Assim criamos uma estratégia de modelagem para descobrir as relações, padrões, regularidades e categorias nos dados que lhe foram apresentados e por fim o treinamos sobre esses padrões identificados. Permitindo que o pré-efeito anódico seja identificado e, realizado previamente ajustes nas variáveis Voltagem Líquida, Fluoreto de Alumínio, Fluoreto de Cálcio, Temperatura de Banho, Intensidade de Corrente, Voltagem Bruta de forma a mitigar as consequências do efeito anódico no forno de redução de alumínio primário.

4.3 Estudo de caso

Para uma análise mais completa, foi realizado um teste com a rede neural desenvolvida conforme mostrado no subcapítulo anterior em “Resultados e discussões”. Assim, foi obtido um conjunto de dados com 40.960 amostras. A Tabela 6 mostra em detalhes as características dos dados já com os outliers removidos.

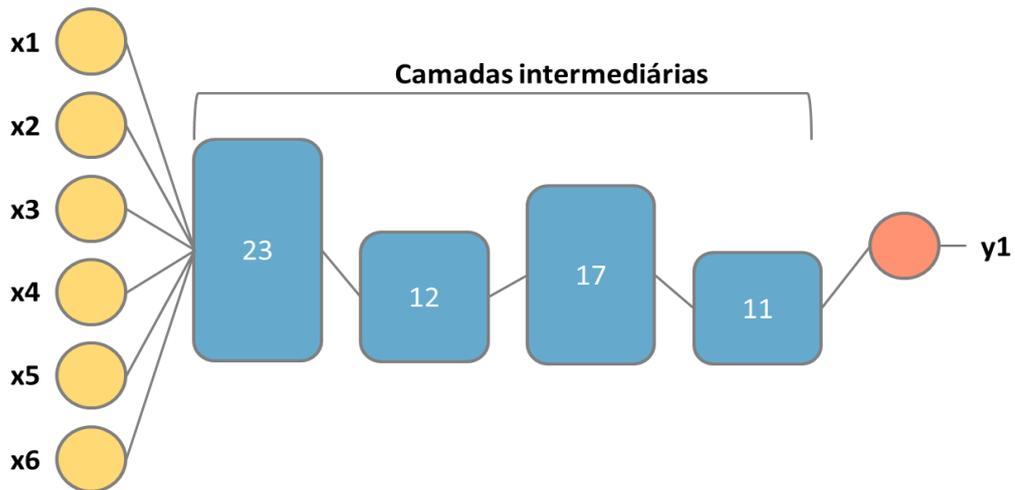
Tabela 8 - Conjunto de dados

Dados	Classe
32 510	0 (Sem Pré-Efeito Anódico)
8 454	1 (Com Pré-Efeito Anódico)
40 960	Todos

Fonte: Elaborado pelo autor do trabalho

A arquitetura de Rede Neural utilizada, para realizar a predição desses conjuntos de dados, possui 6 neurônios na camada de entrada de dados, e 23, 12, 17, 11 neurônios nas primeira, segunda, terceira e quarta camada intermediária respectivamente. A Figura 32 ilustra a arquitetura da Rede Neural utilizada.

Figura 32 - Topologia da Rede Neural

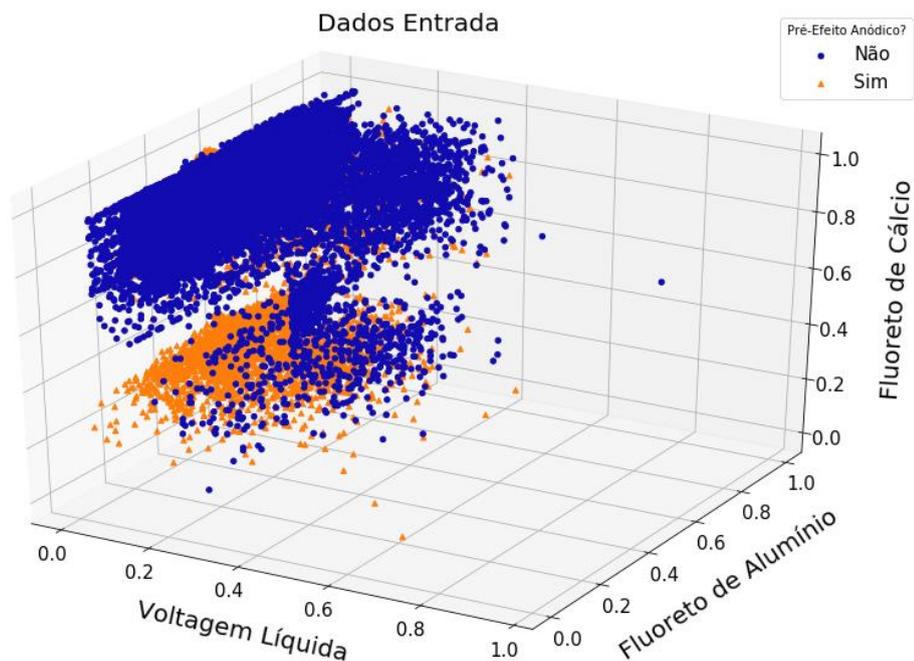


Fonte: Elaborado pelo autor do trabalho

Na Figura 32, as entradas x1, x2, x3, x4, x5, x6 representam as 6 variáveis: Voltagem Líquida, Fluoreto de Alumínio, Fluoreto de Cálcio, Temperatura de Banho, Intensidade de Corrente e Voltagem Bruta. A saída y1 indica a classe predita sendo (0) indicador de “não pré-efeito anódico” e o numeral (1) para “Pré-Efeito anódico”.

A Figura 33 ilustra a disposição dos dados de entrada no espaço tridimensional com suas respectivas classes.

Figura 33 - Disposição no espaço tridimensional para os dados de entrada

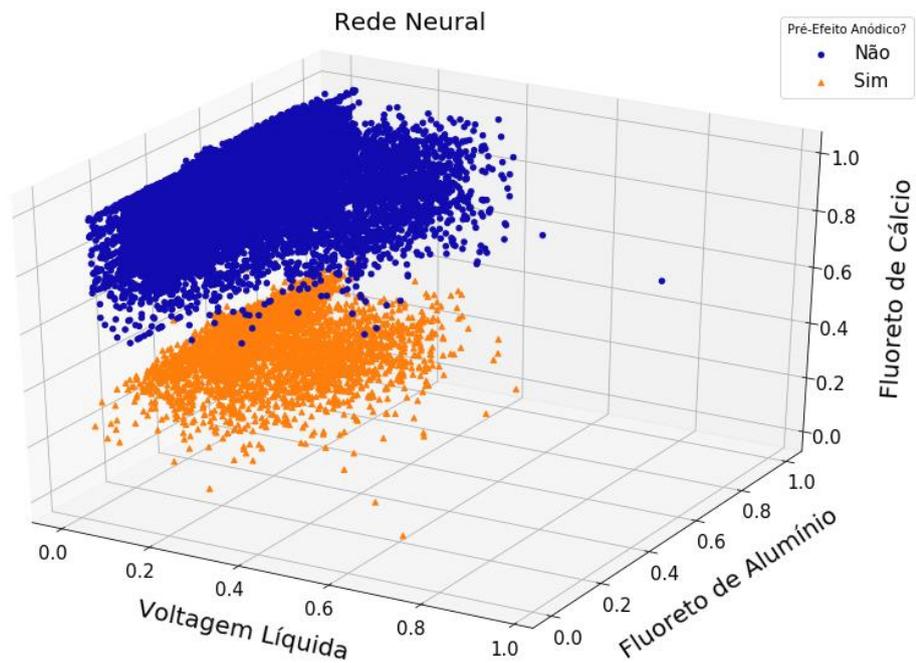


Fonte: Elaborado pelo autor do trabalho

Assim, os 40.960 registros foram submetidos à rede neural para que seja realizada a predição e obter os rótulos desse conjunto de dados.

A Figura 34 é a representação gráfica da categorização do conjunto de dados realizado pela rede neural.

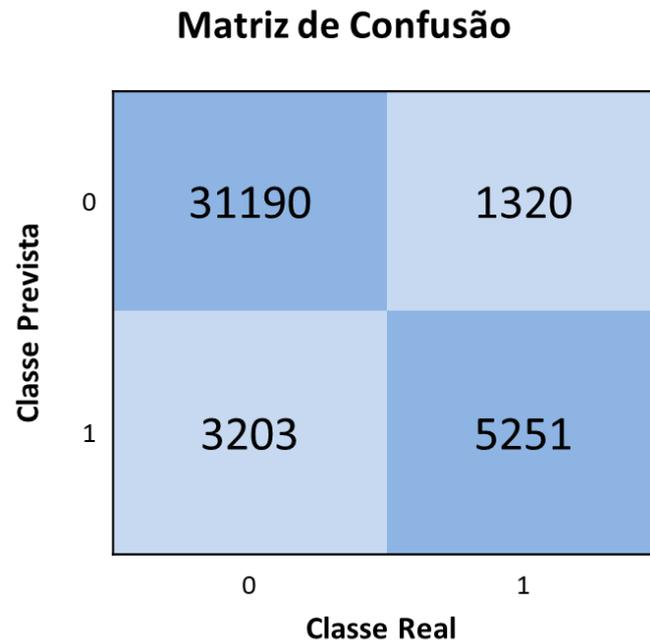
Figura 34 - Categorização do conjunto de dados realizado pela Rede Neural



Fonte: Elaborado pelo autor do trabalho

A Figura 35 apresenta o resultado da rede neural em uma matriz de confusão.

Figura 35 - Matriz de confusão da Rede Neural



Fonte: Elaborado pelo autor do trabalho

A acurácia alcançada por este modelo nestas características foi de 0.8895859779318426. Portanto, pode-se inferir que a arquitetura da rede neural, selecionada entre varias arquiteturas treinadas, manteve sua eficácia para um novo conjunto de dados submetidos a rede.

5 CONCLUSÃO

Este trabalho apresentou a proposta de um sistema de classificação de ocorrência do pré-efeito anódico do forno da indústria de redução de alumínio baseado em uma estrutura de Rede Neural Artificial junto com algoritmos de clusterização. Para o desenvolvimento do algoritmo utilizou-se a base de dados de uma empresa brasileira, internacionalmente reconhecida pelos padrões de qualidade do metal produzido, composta por dados de Voltagem Líquida, Fluoreto de Alumínio, Fluoreto de Cálcio, Temperatura de Banho, Intensidade de Corrente, Voltagem Bruta e Pré-efeito Anódico. Com os dados de entrada e sua classe categórica para modelar computacionalmente padrões da ocorrência do pré-efeito anódico foi usado primeiro um aprendizado não supervisionado para rotular nossos dados de entrada e posteriormente treinamos a nossa rede neural com os rótulos definidos na primeira etapa, e por fim, avaliamos a saída da rede neural com as classes originais.

Os bons resultados, através do uso combinado de algoritmos de clusterização e de redes neurais, acima de 89,15%, obtidos a partir do modelo predictor demonstram a sua eficiência e aplicabilidade.

Considerando o uso de clusterização, podemos considerar que os resultados obtidos demonstram a aplicabilidade desta estrutura para problemas de classificação. Outro ponto de destaque é a melhoria para o modelo final, que se torna possível quando se agrega diferentes técnicas em prol do objetivo a ser alcançado.

O trabalho publicado no qual este estudo foi fundamentado se intitula “Clusterização, Classificação e Predição de Pré-Efeito Anódico de cuba eletrolítica de alumínio primário”, publicado nos anais do Seminário Integrado de Software e Hardware (SEMISH), e pode ser acessado no endereço <https://sol.sbc.org.br/index.php/semish>.

Sugere-se para trabalhos futuros combinar (ensemble) estes e outros métodos de agrupamento, criando uma estratégia para aproveitar as melhores características de cada algoritmo na tentativa de superar as limitações de cada técnica individual. Outro trabalho futuro é usar outros modelos de RNA's para a tarefa de classificação de pré-efeito anódico.

REFERÊNCIAS

- ALCOA. **Sustentabilidade**. 2020. Disponível em: <https://www.alcoa.com/brasil/> Acesso em: 07 janeiro 2020.
- ANP. Nota Técnica Nº 041/2019/SDP. Disponível em: http://www.anp.gov.br/arquivos/cap/2019/cap22/cp22-2019_notatecnica-041.pdf. Acesso em: 08 março 2020.
- ANNA NEVES. **Redução de Emissões de GEE**. 2014. Disponível em: <https://www.pensamentoverde.com.br/sustentabilidade>. Acesso em: 07 janeiro 2020.
- ARTHUR, DAVID; VASSILVITSKII, SERGEI. **K-means++: The Advantages of Careful Seeding**, Proceedings of the eighteenth annual ACM-SIAM symposium on Discrete algorithms, Society for Industrial and Applied Mathematics (2007).
- ASSOCIAÇÃO BRASILEIRA DO ALUMÍNIO - ABAL. **Alumínio brasileiro: soluções para uma vida sustentável**. 2017. Disponível em: <http://abal.org.br>. Acesso em: 01 agosto 2019.
- BRAGA, A. P.; CARVALHO, A. C. P. L. F.; LUDEMIR, T. B. **Redes neurais artificiais: teoria e aplicações**. Rio de Janeiro: LTC, 2000.
- BRASIL ESCOLA, **Efeito Estufa**. 2020. Disponível em: <https://brasilecola.uol.com.br/geografia/efeito-estufa.htm> . Acesso em: 07 janeiro 2020.
- CARVALHO, ANDRÉ, et al. **Inteligência Artificial: uma abordagem de aprendizado de máquina**. Rio de Janeiro: LTC, 2011.
- CONTE, THIAGO. **Modelagem da Resistência Elétrica dos Fornos de Redução do Alumínio**. Dissertação (Mestrado em Computação Aplicada) – Engenharia Elétrica, Universidade Federal do Para, Belém. 2015.
- CONTE, THIAGO; CONTE, BRUNO; OLIEVIRA, ROBERTO. **Comparação de Modelos Neurais Aplicados para Predição do Efeito Anódico de Fornos de Redução do Alumínio Primário**. In: Conferência Brasileira de Dinâmica, Controle e Aplicações, 2017, São José do Rio Preto – SP.
- CONTE, THIAGO; CONTE, BRUNO; OLIVEIRA, ROBERTO. **Uso de Técnicas de Mineração de Dados para Classificação do Efeito Anódico dos Fornos de Redução de Alumínio Primário**. In: 80 Congresso Internacional do Alumínio, 2018, São Paulo.
- CHEN, L; CHEN, Y.; GUI, W. **Anode Effect Prediction Based On Collaborative Two-Dimensional Forecast Model In Aluminium Electrolysis Production**. *Journal of Industrial and Management Optimzation*, v. 15, n. 2, p. 595-612, 2019.

DIEGO. **Produção de Alumínio: Extração da Bauxita, Obtenção na Eletrolise**. Disponível em:

https://culturalivre.com/producao_de_aluminio_extracao_da_bauxita_obtencao_eletr olise/. Acesso em: 27 março 2018.

HAYKIN, S., **Neural Networks and Learning Machines**, 3 ed., Pearson, Hamilton, Canadá, 2009.

HUBERT, L.; ARABIE, P. **Comparing partitions**. *Journal of Classification*, v. 2, p. 193-218, 1985.

IBAMA. **Instrução Normativa IBAMA Nº 10 DE 07/12/2012**. 2012. Disponível em: <https://www.legisweb.com.br/legislacao/?id=277984>. Acesso em: 08 março 2020.

KERAS DOCUMENTATION. **Keras: The Python Deep Learning library**, 2020. Disponível em: <https://keras.io>. Acesso em: 07 janeiro 2020.

ROSENBERG, A.; HIRSCHBERG, J. **V-Measure: A conditional entropy-based external cluster evaluation measure**, New York, 2007.

SCIKIT-LEAM. **Machine Learning in Python**. 2018. Disponível em: <https://scikit-learn.org/stable/modules/clustering.html>. Acesso em: 08 janeiro 2020.

MAKLIN, CORY. **BIRCH Clustering Algorithm Example In Python**. 2019. Disponível em: <https://towardsdatascience.com/machine-learning-birch-clustering-algorithm-clearly-explained-fb9838cbeed9>. Acesso em: 7 janeiro. 2020.

ROCHA, INGRID. **Emissões de Gases de Efeito Estufa no Processos Industriais**. Ministério da Ciência e Tecnologia. 2010.

ROUSSEEUW, Peter J.; HUBERT, Mia. **Anomaly detection by robust statistics**. *WIREs Data Mining and Knowledge Discovery*. Lovaina, Belgium, p. 3-4, 2017.

SILVA, ANTÔNIO JOSÉ. **Modelagem Paramétrica de Cubas Eletrolíticas para Predição do Efeito Anódico**. 2009. 102 f. Curso de Engenharia de Eletricidade, Universidade Federal do Maranhão, São Luís, 2009.

SILVA, I. N.; SPATTI, D. H.; FLAUZINO, R. A., **Redes Neurais Artificiais para Engenharia e Ciências Aplicadas**, São Paulo, Brasil, Editora Artliber, 2010.

VINH, NGUYEN; EPPS, JULIEN; BAILEY, JAMES, **Information Theoretic Measures for Clusterings Comparison: Variants, Properties, Normalization and Correction for Chance** . *JMLR*, 2010.

ZHOU, KAIBO et al. **Anode Effect Prediction of Aluminum Electrolysis Using GRNN**. Chinese Automation Congress (CAC). Wuhan, China: IEEE, 2015.