

Bruno Gilmar Honnef

Redes neurais para previsão do coeficiente de condutância global de refrigeradores domésticos

Trabalho de conclusão do curso submetido(a) ao curso de Ciência da Computação da Universidade Federal de Santa Catarina para a obtenção do Grau de Bacharel em Ciência da Computação.

Universidade Federal de Santa Catarina - UFSC
Departamento de Informática e estatística - INE

Orientador: Elder Rizzon Santos

Florianópolis

2020

Ficha de identificação da obra elaborada pelo autor,
através do Programa de Geração Automática da Biblioteca Universitária da UFSC.

SobrenomeAutor, NomeAutor

Título do trabalho : Subtítulo do trabalho / NomeAutor

SobrenomeAutor ; orientador, NomeOrientador

SobrenomeOrientador ; coorientador, NomeCoorientador

SobrenomeCoorientador. - Florianópolis, SC, 2014.

100 p.

Tese (doutorado) - Universidade Federal de Santa
Catarina, Centro Tecnológico. Programa de Pós-Graduação em
Ciência da Computação.

Inclui referências

1. Ciência da Computação. 2. Exemplo de ficha
catalográfica. I. SobrenomeOrientador, NomeOrientador. II.
SobrenomeCoorientador, NomeCoorientador. III. Universidade
Federal de Santa Catarina. Programa de Pós-Graduação em
Ciência da Computação. IV. Título.

Bruno Gilmar Honnef

Redes neurais para previsão do coeficiente de condutância global de refrigeradores domésticos

Este trabalho de conclusão de curso foi julgado adequado para obtenção do Título de "Bacharel em ciência da computação" e aprovado em sua forma final pelo Curso de Ciência da Computação.

Florianópolis, 03 de Janeiro de 2020.

Prof. Dr. **Alexandre Gonçalves Silva**
Coordenador do Curso

Banca examinadora:

Prof. Dr. **Elder Rizzon Santos**
Orientador

Prof. Dr. **Alvaro Junio Pereira Franco**
Avaliador

Prof. Dr. **Rafael de Santiago**
Avaliador

Florianópolis
2020

Aos meus avós e pais.

Agradecimentos

Agradeço em primeiro lugar meus pais, Rosana e Gilmar que sempre me apoiaram e motivaram.

Aos meus avós, Eni e Leonardo, por sempre estarem comigo nessa jornada.

Aos amigos que fiz durante a graduação pela amizade, bons momentos e risadas.

Ao meu amigo Lucas Lazzaris pelo auxílio na parte técnica.

Ao meu orientador, Elder, pela paciência e por me ajudar na elaboração deste trabalho.

Aos membros da banca examinadora, pela disposição em avaliar este trabalho.

Aos professores que conheci durante essa jornada pelo conhecimento transmitido.

Ao Stack overflow (em especial), Wikipédia e outros meios de disseminação de conhecimento, só por existirem.

E a Universidade Federal de Santa Catarina pela estrutura fornecida.

Resumo

Neste trabalho, são apresentados modelos de redes neurais com o objetivo de prever o coeficiente de condutância global (UA) de refrigeradores domésticos, no qual é utilizado dados experimentais para análise de desempenho e pré-treinamento da rede com a ampliação sintética do conjunto de dados utilizando técnicas de amostragem e equações que regem o modelo físico. O desenvolvimento deste trabalho consiste em gerar os dados sintéticos, geração e ajuste dos hiper-parâmetros das redes neurais e, por fim, as redes são treinadas com a técnica padrão de ajuste ou pré-treinadas com *Stacked auto-encoders*. As técnicas apresentadas neste trabalho se tornam úteis quando é necessário lidar com conjuntos de dados pequenos ou com pouca diversidade. É visto que as redes tiveram um bom desempenho, mostrando ser viável o uso de redes neurais para previsão de propriedades físicas.

Palavras-chave: Rede neural, conjunto pequeno de dados, coeficiente de condutância global, UA , dados sintéticos, dados experimentais, técnicas de amostragem, SAE, stacked auto-encoders, pré-treinamento.

Abstract

This work shows models of neural network that will try to predict the overall heat coefficient (UA) of domestic refrigerators, where experimental data is used to check the performance and pre-training the network with the synthetic extension of the dataset using sampling techniques and equations that govern the physical model. The methodology of this work is composed by generate the synthetic data, the network's generation and hyper-parameter adjustments, and the networks are trained using the default fit function or pre-trained using stacked auto-encoders. The techniques used in this works are useful when there is a small dataset, or the dataset is not diverse. It is shown that the networks had a good performance, therefore, showing that it is viable the use of neural networks to predict physical properties.

Keywords: Neural network, small data set, overall heat transfer coefficient, UA , synthetic data, experimental data, sampling techniques, SAE, stacked auto-encoders, pre-training.

Lista de ilustrações

Figura 1 – Esquema de condução de calor através das paredes	26
Figura 2 – Teste de fluxo de calor reverso	27
Figura 3 – Modelo matemático simples de um neurônio artificial	29
Figura 4 – Exemplo de rede neural do tipo <i>feed-forward</i>	30
Figura 5 – Funções de ativação	30
Figura 6 – Esquemático do processo de pré-treinamento da rede.	32
Figura 7 – Diagrama da metodologia adotada.	39
Figura 8 – Configurações de isolantes de cada parede do gabinete do refrigerador.	42
Figura 9 – Gráfico mostrando as saídas UA_R , UA_C e UA_G , gerado utilizando proporções.	43
Figura 10 – Gráfico mostrando as saídas UA_R , UA_C e UA_G	43
Figura 11 – Gráfico mostrando as saídas UA_R , UA_C e UA_G experimentais e pseudo-experimentais.	44
Figura 12 – Ilustração dos dados de entrada da rede	46
Figura 13 – Resultados da previsão da rede dual.	49
Figura 14 – Resultados da previsão dos dados pseudo-experimentais da rede dual.	50
Figura 15 – Resultados da previsão da rede dual com uma configuração de isolantes.	51
Figura 16 – Resultados da previsão da rede única para o refrigerador.	51
Figura 17 – Resultados da previsão da rede única para o congelador.	52
Figura 18 – Gráfico mostrando as geometrias pelo UA_R	63
Figura 19 – Gráfico mostrando as geometrias pelo UA_C	64
Figura 20 – Gráfico mostrando os parâmetros geométricos.	65
Figura 21 – Gráfico mostrando as geometrias pelo UA_R , gerado utilizando proporções.	66
Figura 22 – Gráfico mostrando as geometrias pelo UA_C , gerado utilizando proporções.	66
Figura 23 – Gráfico mostrando os parâmetros geométricos, gerado utilizando proporções.	67

Lista de tabelas

Tabela 1 – Resistências térmicas das camadas de revestimento	28
Tabela 2 – Síntese das técnicas usadas nos trabalhos correlatos.	38
Tabela 3 – Resistências térmicas dos materiais isolantes utilizados	41
Tabela 4 – Valores para a busca de hiperparâmetros	46
Tabela 5 – hiperparâmetros ajustados da rede	47
Tabela 6 – Parâmetros de cada treinamento.	47
Tabela 7 – Acurácia das redes	48
Tabela 8 – Acurácia das redes com apenas uma configuração de refrigerador	50
Tabela 9 – Dicionário de nomenclatura.	69
Tabela 10 – Amostra dos dados experimentais usados.	70
Tabela 11 – Métricas sobre os dados uniformes.	70
Tabela 12 – Métricas sobre os dados proporcionais.	70

Lista de abreviaturas e siglas

CFD	<i>Computational fluid dynamics</i> , traduz-se para Dinâmica computacional de fluidos.
DNN	<i>Deep neural network</i> , traduz-se para Rede neural profunda.
SCS	<i>Solidification crack susceptibility</i> , traduz-se para Susceptibilidade à fissuração por solidificação.
SAE	<i>Stacked autoencoder</i> .
HTC	<i>Heat transfer coefficient</i> , traduz-se para Coeficiente de transferência de calor.
IHCP	<i>Inverse heat transfer problems</i> , traduz-se para Problemas de transferência de calor inverso.
SMOGN	<i>Synthetic Minority Over-Sampling Technique for Regression with Gaussian Noise</i> .
SMOTE	<i>Synthetic Minority Over-sampling Technique</i> .
SMOTER	<i>Synthetic Minority Over-sampling Technique for regression</i> .
MSE	<i>Mean squared error</i> , traduz-se para Erro médio quadrático.
MAE	<i>Mean absolute error</i> , traduz-se para Erro médio absoluto.
PU	<i>Poliuretano</i> .

Lista de símbolos

A	Área, m^2 .
h	Altura, mm .
k	Condutividade térmica, W/mK .
L	Largura, mm .
l	Espessura, mm .
p	Profundidade, mm .
q	Taxa de transferência de calor, W .
T	Temperatura, K .
U	Coefficiente global de transferência de calor, W/m^2K .
UA_C	Coefficiente de condutância do congelador, W/K .
UA_G	Coefficiente de condutância global, W/K .
UA_R	Coefficiente de condutância do refrigerador, W/K .
V	Volume, <i>litros</i> .
W	Peso atribuído a uma entrada.

Sumário

1	INTRODUÇÃO	23
1.1	Objetivos	24
1.1.1	Objetivo Geral	24
1.1.2	Objetivos Específicos	24
1.2	Organização do texto	24
2	FUNDAMENTAÇÃO TEÓRICA	25
2.1	Coeficiente de condutância global	25
2.1.1	Teste de fluxo de calor reverso	26
2.1.2	Condutância térmica das paredes	27
2.2	Redes neurais	28
2.2.1	Stacked autoencoders	30
2.2.2	Hiperparâmetros	31
2.2.2.1	Hiperparâmetros de otimização	31
2.2.2.1.1	Taxa de aprendizagem	31
2.2.2.1.2	Tamanho do conjunto	31
2.2.2.1.3	Número de épocas	33
2.2.2.2	Hiperparâmetros do modelo	33
2.2.2.2.1	Número de unidades ocultas	33
2.2.2.2.2	Primeira camada oculta	33
2.2.2.2.3	Número de camadas	33
2.2.2.2.4	Função de ativação	34
2.2.2.3	Ajuste de hiperparâmetros	34
2.2.2.3.1	Otimização Bayesiana	34
2.2.3	Treinamento	34
2.2.3.1	Supervisionada	35
2.2.3.2	Não supervisionada	35
2.2.3.3	Por reforço	35
2.3	Técnicas de amostragem	35
2.3.1	SMOGLN	35
2.3.2	Regressão não balanceados	36
2.3.3	SMOTER	36
3	TRABALHOS CORRELATOS	37
3.1	Redes neurais	37

4	DESENVOLVIMENTO	39
4.1	Metodologia	39
4.2	Dados	40
4.2.1	Geração de dados sintéticos	40
4.2.1.1	Qualidade dos dados	42
4.2.2	Dados experimentais	43
4.3	Modelo	44
4.3.1	Entradas e saídas	45
4.3.2	Ajuste dos hiperparâmetros	45
4.3.3	Treinamento	47
4.4	Resultados	47
5	CONCLUSÃO	53
5.1	Trabalhos futuros	53
	REFERÊNCIAS	55
	Glossário	59
	APÊNDICES	61
	APÊNDICE A – GRÁFICOS DA ANÁLISE DE DADOS	63
	APÊNDICE B – AMOSTRA E MÉTRICAS DOS DADOS UTILIZADOS NESTE TRABALHO	69
	APÊNDICE C – ARTIGO DESTE TRABALHO EM FORMATO SBC	71

1 Introdução

Refrigeradores são eletrodomésticos essenciais no cotidiano do ser humano. De acordo com [Barthel e Götz \(2012\)](#) existem cerca de 1.4 bilhões de refrigeradores domésticos em operação no mundo inteiro, os quais são responsáveis por cerca de 15% do consumo de energia no setor residencial. Na tabela do [Inmetro \(2020\)](#), um refrigerador doméstico frost-free típico consome cerca de 45 kWh/mês, apesar do consumo unitário ser relativamente baixo, no Brasil o consumo de energia de refrigeradores domésticos representa 11% de toda a energia produzida [Eletrobrás \(2004\)](#).

A crescente demanda por energia tem provocado o estabelecimento de metas cada vez mais ambiciosas de redução de consumo de energia e regras mais rigorosas de classificação de eletrodomésticos. Apesar de serem extensivamente estudados, o funcionamento de refrigeradores ainda está longe do ideal, sendo caracterizados por um excessivo consumo de energia e até por temperaturas acima do recomendado para a conservação de alimentos [James, Evans e James \(2008\)](#).

A maneira tradicional de avaliar o desempenho de um refrigerador consiste na realização de cálculos simplificados com base nas curvas características dos componentes, análises numéricas através de softwares comerciais de *CFD* (do inglês, *Computational Fluid Dynamics*) e de testes experimentais normalizados. Apesar de serem ferramentas importantes para o dimensionamento de componentes, as curvas características e o CFD não fornecem informações acerca do comportamento sistêmico do refrigerador, que então são obtidas através de testes em câmaras climatizadas com temperatura e umidade do ar controladas. Tais testes são custosos, além da necessidade de um protótipo, por exemplo, um teste de abaixamento de temperatura consome mais de 24 horas se considerar apenas o tempo necessário para estabilização do experimento e para efetuar as medições [Hermes \(2006\)](#).

O presente trabalho busca diminuir o custo com estes experimentos, mais especificamente com o teste de fluxo de calor reverso, o qual tem como objetivo avaliar a qualidade do isolamento térmico das paredes através da medição das condutâncias térmicas globais dos compartimentos, comumente denominadas *UA*. São analisados modelos de redes neurais tendo como entrada a geometria do refrigerador e o isolante de cada uma de suas paredes, os modelos de redes propostos se ajustam até com poucos dados. Tais dados podem ser experimentais ou gerados sinteticamente através de equações que regem tal modelo físico ou utilizando técnicas de amostragem do conjunto experimental. O treinamento da rede é feita de duas formas, a primeira utilizando de uma SAE (do inglês, *Stacked autoencoder*) para pré-treinar a rede e a segunda utilizando a técnica padrão de ajuste da rede.

1.1 Objetivos

Nas seções abaixo estão descritos os objetivos deste trabalho que estão subdivididos em específicos e gerais.

1.1.1 Objetivo Geral

O presente trabalho propõe modelos de redes neurais e técnicas de geração de dados para treinar e analisar o desempenho da rede a qual tem o propósito de prever o coeficiente de condutância global dos gabinetes de refrigeradores domésticos.

1.1.2 Objetivos Específicos

- a) Examinar e categorizar a distribuição de dados;
- b) Padronizar a entrada de dados;
- c) Propor modelos neurais para prever o coeficiente;
- d) Propor técnicas de treinamento para os modelos;
- e) Analisar o desempenho do modelo;

1.2 Organização do texto

O texto deste trabalho está dividido em quatro partes, iniciando com o [Capítulo 2](#) a qual fornece uma base de conhecimento para o melhor entendimento das partes seguintes, o [Capítulo 3](#) é focado na revisão bibliográfica e trabalhos correlatos que inspiraram a solução. No [Capítulo 4](#) é mostrada a metodologia juntamente com o desenvolvimento, ajustes na rede e os resultados obtidos com ela. No [Capítulo 5](#) é discutido os resultados, suas conclusões e o que pode ser feito para aprimorar.

2 Fundamentação teórica

Este capítulo tem como propósito explicar os fundamentos e bases teóricas necessárias para a compreensão do trabalho desenvolvido. Na seção 2.1, é explicado o significado do coeficiente de condutância global, como é obtido experimentalmente e como é estimado para cada parede. Na seção 2.2, é apresentado o conceito de redes neurais, nas subseções são elaboradas a metodologia para o treinamento das redes e suas funções.

2.1 Coeficiente de condutância global

Um dos parâmetros de grande influência no desempenho de um refrigerador é a resistência térmica do gabinete, o qual depende, principalmente, da condutividade do isolamento térmico do gabinete, e pode ser expressa na forma de um coeficiente global de transferência de calor U com unidade $[W/m^2K]$ (INCROPERA et al., 2007; COLOMBO, 2014).

O coeficiente U é uma medida da capacidade global de transferir calor em uma série de barreiras condutivas e convectivas. Para o caso de um trocador de calor, podemos calcular U usando uma expressão análoga à lei do resfriamento de Newton,

$$q = -UA\Delta T \quad (2.1)$$

onde o vetor q é a taxa de transferência de calor $[W]$ e o sinal indica a direção do calor (do corpo quente para o frio), A é a área de troca de calor $[m^2]$, e ΔT é a diferença de temperaturas entre os lados da parede.

A Figura 1 apresenta uma ilustração esquemática da condução de calor através das paredes, onde a linha vermelha representa o perfil de temperatura entre a parede externa até a interna. Em que T_a , T_i são, respectivamente, a temperatura ambiente e interna, T_{se} , T_{si} são, respectivamente, a temperatura da superfície externa e interna, A_{se} , A_{si} são a área da superfície externa e interna, l representa a espessura total da parede, l_a , l_b , l_c são as espessuras de cada parede interna podendo ter n paredes.

Quando o coeficiente global de transferência de calor está associado a uma área A de troca de calor então é designado o termo condutância global de calor UA , que é a multiplicação do termo U pela área de troca A resultando na unidade $[W/K]$. Para obter este coeficiente para um dado gabinete refrigerado é necessário um procedimento experimental denominado teste de fluxo de calor reverso Incropera et al. (2007), Colombo (2014)

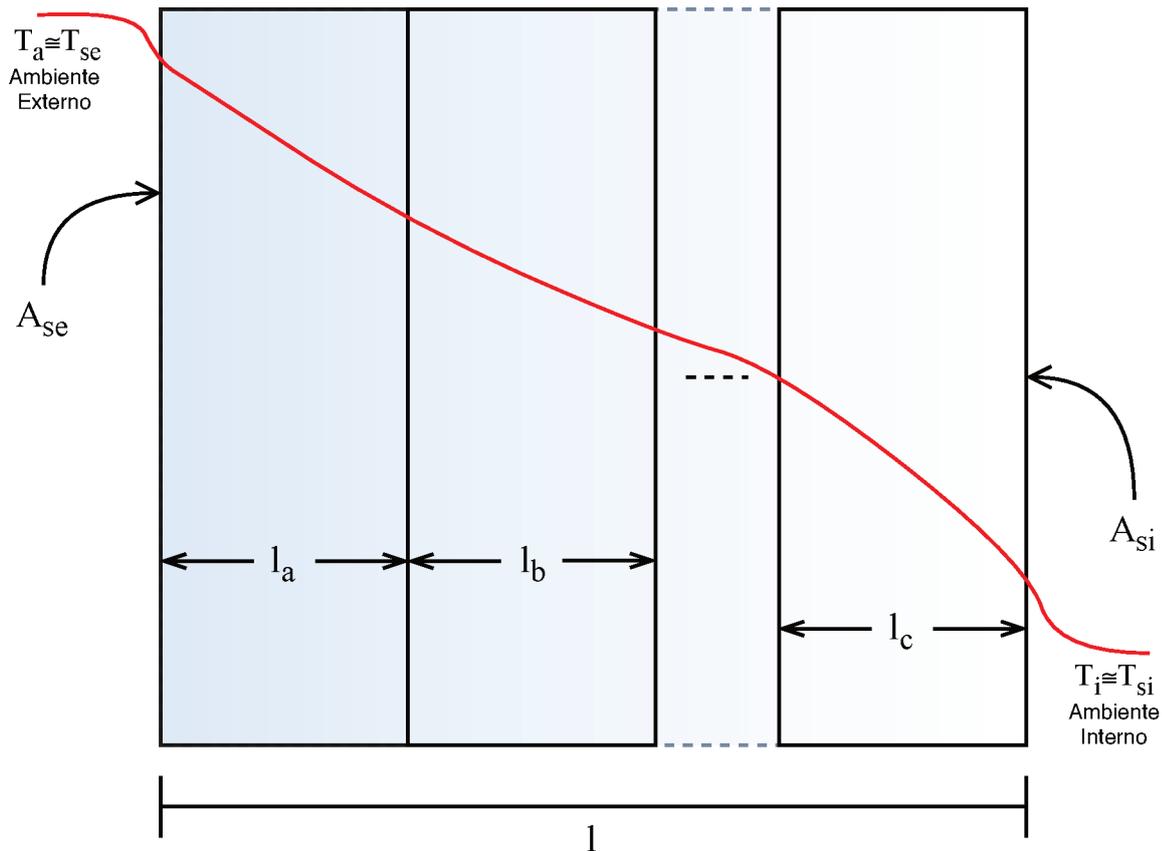


Figura 1 – Esquema de condução de calor através das paredes. Adaptado de [Hermes \(2006\)](#), [Incropera et al. \(2007\)](#).

2.1.1 Teste de fluxo de calor reverso

O principal objetivo deste teste é a avaliação da qualidade do isolamento térmico das paredes através da medição das condutâncias térmicas globais, normalmente chamadas de UA . A metodologia comumente adotada são as propostas por [Junior e Sousa \(2008\)](#) e [Sim e Ha \(2011\)](#).

O nível de isolamento térmico de um sistema de refrigeração é normalmente determinado através de experimentos estabelecidos pela norma NTB 00119. Esse teste consiste em instalar resistências elétricas, com potência inferior a 50 W, controladas por uma fonte variável de tensão. O objetivo é aquecer o interior do gabinete criando um fluxo de calor do interior para o exterior, originando o nome do experimento. A câmara de testes é mantida a temperatura e umidade constantes e o sistema de refrigeração é desligado e as resistências elétricas dissipam calor até que se atinja uma diferença de temperatura de 25°C entre os ambientes em **Regime permanente**. As resistências elétricas são distribuídas no interior de gabinete de maneira que a estratificação térmica seja minimizada. O monitoramento do teste limita-se à potência dissipada pelas resistências elétricas e as temperaturas internas e externas como mostrado na [Figura 2 Seidel \(2001\)](#).

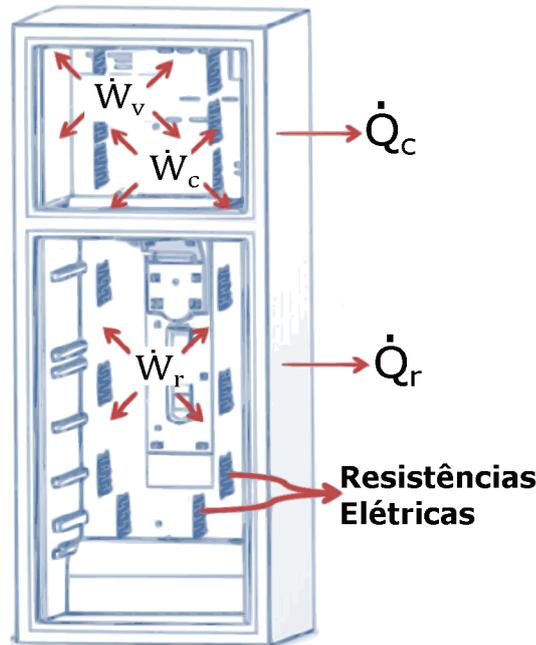


Figura 2 – Teste de fluxo de calor reverso. Adaptado de Thiessen (2015).

Em regime permanente, todo calor gerado pelas resistências elétricas flui através das paredes do refrigerador e, utilizando a Equação 2.1, chega-se à seguinte equação para determinação do parâmetro UA :

$$UA = \frac{q_r}{(T_e - T_i)} \quad (2.2)$$

na qual UA é o produto entre o coeficiente global de transferência de calor e a área média total do gabinete, q_r é a potência dissipada pelas resistências elétricas, T_e é a média das temperaturas externas e T_i é a média das temperaturas internas.

Vale ressaltar que este teste não distingue os fluxos de calor através das diversas paredes, da porta e da gaxeta do refrigerador além de não permitir a avaliação da distribuição de fluxo de calor sobre as superfícies.

2.1.2 Condutância térmica das paredes

De acordo com Hermes (2006), transferência de calor através das paredes de um refrigerador pode ser tratada como unidimensional, uma vez que o gradiente de temperaturas na direção normal à parede é da ordem de 100 vezes maior que nas outras direções. No ponto de vista de transferência de calor, a parede é formada por cinco resistências térmicas em série: convecção com ambiente externo, condução na chapa de aço, condução no isolamento, condução na chapa plástica e convecção com o ambiente interno. A Tabela 1 mostra as grandezas dessas resistências e podemos notar que a resistência do

isolamento de poliuretano (PU) é pelo menos uma ordem de grandeza maior que as outras, de modo que somente o PU precisa ser levado em conta no cálculo da carga térmica.

Camada	Espessura [mm]	Condutividade [W/mK]	Condutância [W/m ² K]	Resistência [m ² K/W]
Plástico	1.5	0.16	106.67	~0.01
PU	50	0.021	0.42	~1
Aço	0.5	50	10 ⁵	10 ⁻⁵
Ar interno	-	-	20	~0.1
Ar externo	-	-	10	~0.1

Tabela 1 – Resistências térmicas das camadas de revestimento. Adaptado de [Colombo \(2014\)](#)

Logo, a condutância térmica para cada parede pode ser calculada utilizando a seguinte equação:

$$UA_i = \frac{k_{iso}A_i}{l_i} \quad (2.3)$$

Onde k_{iso} é a condutividade do isolamento da parede com menor grandeza, l é a espessura dessa parede, o índice i representa cada uma das paredes do respectivo compartimento (congelador e refrigerador). O valor do UA dos compartimentos é o somatório de cada uma das paredes do compartimento e os coeficientes do congelador UA_C e do refrigerador UA_R estariam determinados. Para determinar o coeficiente de condutância global, UA_G ou simplesmente UA , de todo o gabinete basta somar os coeficientes de cada compartimento.

$$UA_G = UA_C + UA_R \quad (2.4)$$

Porém esta equação não leva em conta outras fontes de carga térmica, como a transferência de calor através das gaxetas. De modo geral, a transferência de calor através da gaxeta é calculada subtraindo as parcelas associadas à condução de calor da carga térmica total, que é obtida através do experimento descrito na [subseção 2.1.1](#).

2.2 Redes neurais

Redes neurais são sistemas computacionais que foram inspirados pelo cérebro humano. Esse sistema é capaz de aprender a realizar uma tarefa através de exemplos ou até mesmo sozinho, tais métodos são detalhados na [subseção 2.2.3](#). Este modelo de sistema é utilizado para encontrar relações entre dados não lineares, muitas vezes complexos e cheio de ruídos, e tentar reproduzir ou categorizar novas entradas nunca vistas antes, o

uso mais conhecido é para o reconhecimento de imagens, por exemplo, identificar se em uma imagem há um gato.

Redes neurais são compostas de neurônios interconectados, veja a [Figura 4](#), que tentam imitar o que acontece no cérebro humano. Cada neurônio, como mostrado na [Figura 3](#), possui conexões de entrada com um peso $W_{i,j}$ associado, conexão de entrada i do neurônio j , o qual determina a força da conexão. Junto com as entradas tem uma especial chamada *bias* que é usado para ajustar a saída e somado junto na função de entrada. Conexões de saída, que podem ser as entradas de outros neurônios ou o resultado final, uma função de entrada ([Equação 2.5](#)) que é o somatório das entradas a_i multiplicada por seu respectivo peso $W_{i,j}$. O neurônio também possui uma função de ativação g e uma saída a_j que é basicamente o resultado da [Equação 2.5](#) aplicada na função g , ficando $a_j = g(in_j)$ [Russell e Norvig \(2016\)](#).

$$in_j = \sum_{i=0}^n W_{i,j} a_i \quad (2.5)$$

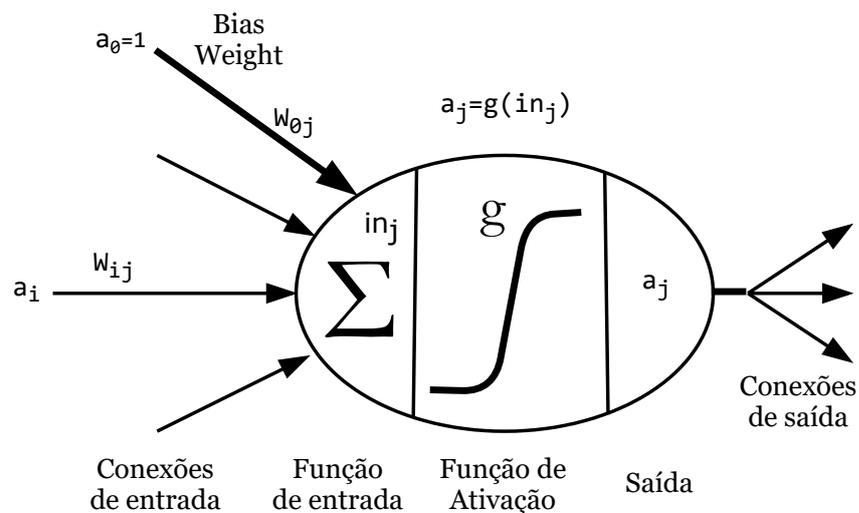


Figura 3 – Modelo matemático simples de um neurônio artificial. Adaptado de [Russell e Norvig \(2016\)](#).

A função de ativação g normalmente ou é uma função com saídas limitadas (0 ou 1), como é o caso da função passo binário da [Figura 5](#), ou é uma função logística como é o caso da sigmoide da [Figura 5](#). Os neurônios da rede podem estar conectados de duas maneiras, *feed-forward* o qual as conexões são unidirecionais, ou seja, é um grafo acíclico. Ou é uma rede recorrente, a qual alimenta a sua entrada com a sua saída [Russell e Norvig \(2016\)](#).

Entretanto, redes neurais ainda estão longe de funcionar como o cérebro humano, o que acontece é que é aplicado um algoritmo e um treinamento para alcançar um objetivo, por exemplo, uma rede criada para jogar xadrez nunca conseguirá guiar um

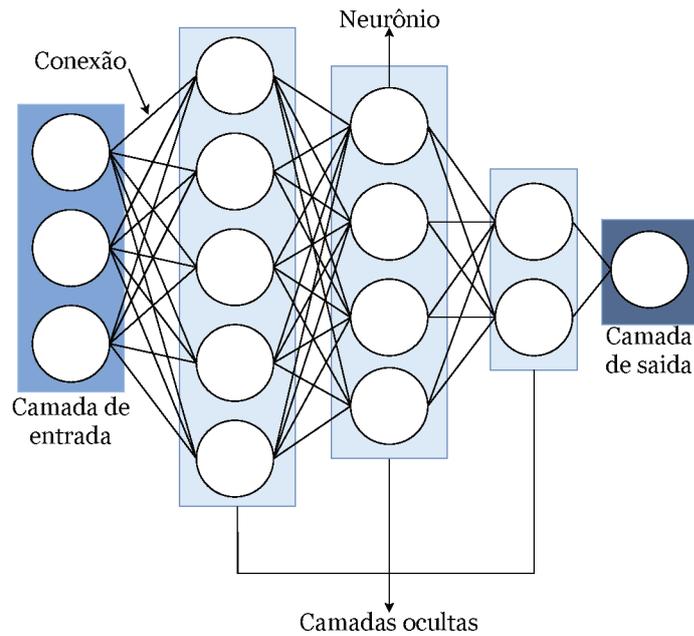


Figura 4 – Exemplo de rede neural do tipo *feed-forward*.

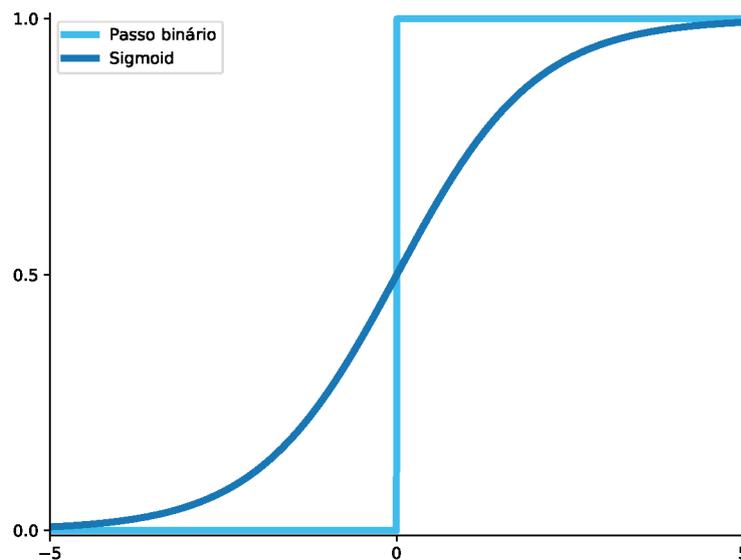


Figura 5 – Funções de ativação.

carro. Há também o elevado custo computacional para treinar uma rede, dependendo da complexidade da rede pode levar de segundos até dias de treinamento para obter resultados satisfatórios [Gerven \(2017\)](#).

2.2.1 Stacked autoencoders

Um *autoencoder* é um tipo de rede neural que visa reproduzir as entradas em suas saídas, de forma não supervisionada, ou seja, o número de neurônios na entrada é o mesmo que na saída. Esse tipo de rede é bastante utilizado para redução de dimensões, pois a rede irá criar uma cópia da entrada “compactada”, ou seja, irá ignorar traços não importantes

dos dados, como é o caso de ruídos. Géron (2017)

O pré-treinamento de uma rede neural com *SAE* ocorre que para cada camada oculta é criado um *autoencoder*, o número de neurônios de cada *autoencoder* é o mesmo que o número de neurônios da camada correspondente da rede, é apresentado um esquemático na [Figura 6](#). Tomamos como exemplo uma rede com 4 camadas ocultas 13-(7-5-4-3)-1¹, então são criados 5 *autoencoders* com as seguintes estruturas: 21-(7)-21, 7-(5)-7, 5-(4)-5, 4-(3)-4 e 3-(1)-3. O primeiro *autoencoder* pega a entrada da rede neural como sua entrada e saída e então a saída da sua camada oculta se torna a entrada do próximo *autoencoder*, em outras palavras, a saída da camada oculta do *autoencoder* é a entrada do próximo *autoencoder*.

2.2.2 Hiperparâmetros

Hiperparâmetros são propriedades que regem todo o processo de aprendizagem da rede o qual inclui variáveis que determinam a estrutura da rede (como o [número de unidades ocultas](#)) e variáveis que interferem na qualidade e velocidade do aprendizado (como a [taxa de aprendizagem](#)), tais variáveis são definidas antes de iniciar o treinamento e algumas, como é o caso das estruturais, não podem ser alteradas posteriormente, e podem ser divididas em duas categorias, hiperparâmetros de otimização e do modelo [Géron \(2017\)](#), [Leonel \(2019\)](#).

2.2.2.1 Hiperparâmetros de otimização

Estes parâmetros estão relacionados ao processo de otimização e de treinamento da rede.

2.2.2.1.1 Taxa de aprendizagem

De acordo com [Murphy \(2012\)](#) a taxa de aprendizagem, ou em inglês *learning rate*, é um hiperparâmetro que define o tamanho do passo em cada iteração enquanto está se movendo em direção ao mínimo de uma função de perda, em outras palavras, ele é quem controla o quanto ajustamos os pesos da rede de acordo com o gradiente de perda.

Se esse hiperparâmetro é muito menor que os valores de ótimo, então levará muito mais épocas para alcançar o valor ideal, já se ele for muito maior que o valor de ótimo, então ele irá ultrapassar o valor ideal e o algoritmo talvez não convirja.

2.2.2.1.2 Tamanho do conjunto

O tamanho do conjunto, ou em inglês *batch size*, influencia nos recursos requeridos do processo de treinamento, velocidade e iterações necessárias. Um número grande de amostras

¹ Os valores entre parênteses representam o número de neurônios nas camadas ocultas.

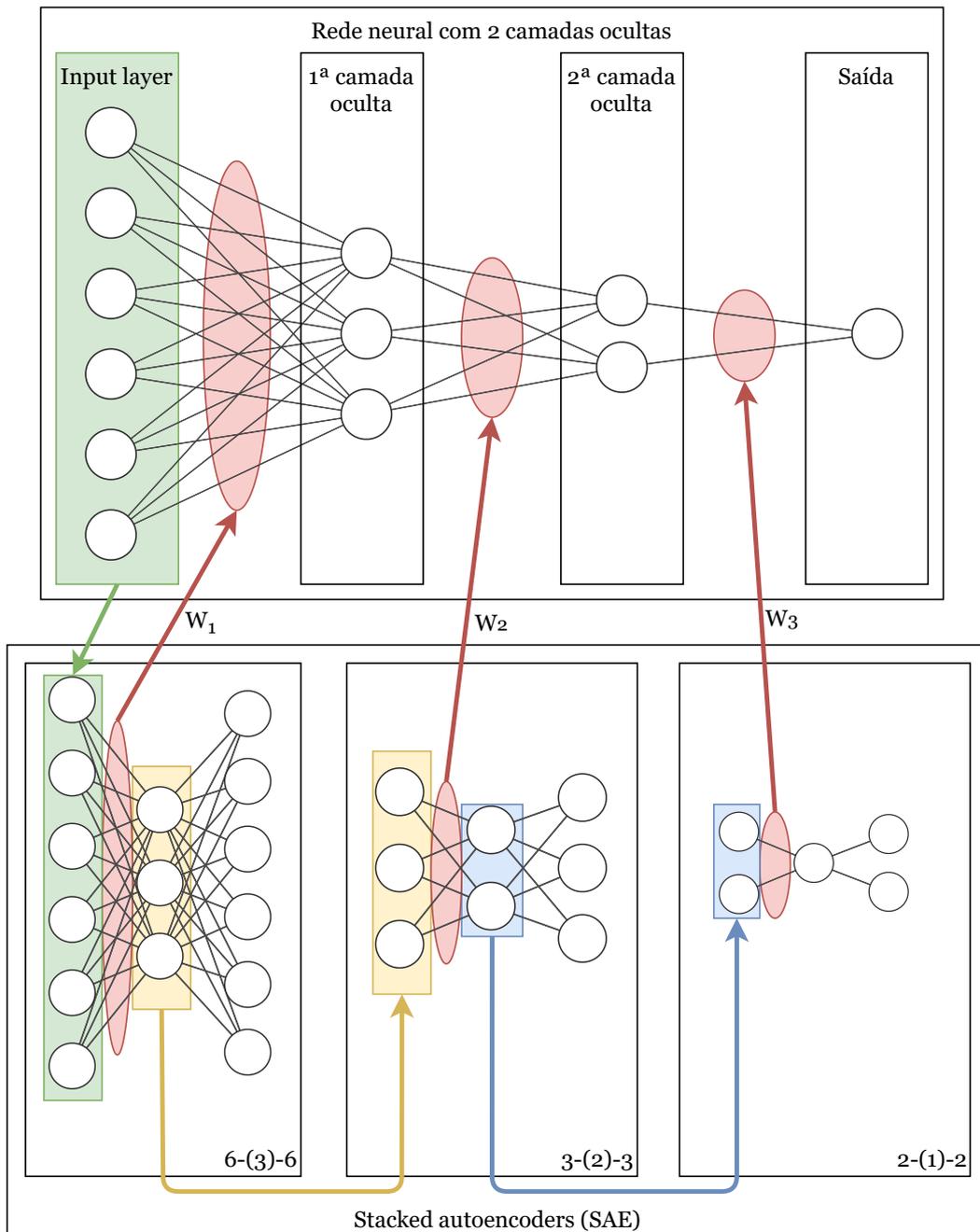


Figura 6 – Esquemático do processo de inicialização camada por camada da rede neural utilizando autoencoders.

dentro do conjunto permite um maior poder computacional que utiliza multiplicação de matrizes no processo de treinamento, porém requer muito mais memória no processo. Um número menor dentro do conjunto induz a mais ruído nos cálculos, porém ajuda a prevenir que o treinamento pare em um mínimo local [Sharma \(2017b\)](#), [Leonel \(2019\)](#).

2.2.2.1.3 Número de épocas

Uma época é caracterizada quando um conjunto de dados é passado inteiramente para frente e para trás na rede uma vez. Já que uma época requer muita memória para computar de uma vez, então é dividido em diversos conjuntos menores, chamados de *batch*. Para evitar o problema de *overfitting* é utilizada a técnica de *early stopping*, o qual, através de uma métrica (normalmente o erro), interrompe o processo de treinamento antes de chegar no total de épocas definidas pelo usuário [Sharma \(2017b\)](#).

2.2.2.2 Hiperparâmetros do modelo

Estes hiperparâmetros estão envolvidos com a estrutura do modelo.

2.2.2.2.1 Número de unidades ocultas

Também chamado de neurônio, ou em inglês *hidden unit*, esse hiperparâmetro define o número de neurônios em cada camada da rede. Um número grande de neurônios nas camadas pode resultar em dois problemas, o primeiro sendo o *overfit*, uma vez que a rede pode “memorizar” o conjunto de dados, e o outro problema sendo que quanto maior o número de neurônios maior é o número de dados necessários para treiná-la, para que então chegue a resultados satisfatórios e com menor erro nas previsões.

2.2.2.2.2 Primeira camada oculta

Esse hiperparâmetro define o número de neurônios da primeira camada após a camada de entrada, essa camada é, talvez, uma das mais importantes da rede, pois baseado em observações empíricas de [Amudha \(2019\)](#), definir um número maior de neurônios do que o da camada de entrada faz com que a rede tenha resultados melhores em certas tarefas.

2.2.2.2.3 Número de camadas

Esse hiperparâmetro define o número de camadas que a rede terá, e para cada uma deve ser definido o número de neurônios. Quanto mais camadas a rede possui, maior é o número de dados necessários para treiná-la e mais lento é o processo de aprendizagem. Se utilizar uma rede de muitas camadas para um conjunto pequeno de dados, é possível que ocorra o problema de *overfitting*, o qual a rede “memorizará” os dados de treinamento, porém se a rede possui um número maior de camadas então ela conseguirá distinguir mais características do conjunto de treino [Leonel \(2019\)](#), [Amudha \(2019\)](#).

2.2.2.2.4 Função de ativação

Esse hiperparâmetro define a função de ativação de cada camada da rede, essa função converte um sinal de entrada do neurônio para um sinal de saída, que pode ser alimentado na próxima camada. O propósito desta função é de introduzir não-linearidade na saída do neurônio [Sharma \(2017a\)](#).

2.2.2.3 Ajuste de hiperparâmetros

O ajuste de hiperparâmetros, ou otimização dos hiperparâmetros, tem como objetivo escolher o melhor conjunto de parâmetros para o algoritmo de aprendizado para então minimizar o valor da função de perda no conjunto de dados [Claesen e Moor \(2015\)](#).

2.2.2.3.1 Otimização Bayesiana

A otimização *Bayesiana* constrói um modelo probabilístico da função mapeando valores de hiperparâmetros a serem buscados para o objetivo avaliado no conjunto de validação. O algoritmo, iterativamente, avalia as configurações de parâmetros mais promissoras baseados nas buscas anteriores. [Snoek, Larochelle e Adams \(2012\)](#).

2.2.3 Treinamento

De acordo com [Russell e Norvig \(2016\)](#) aprendizagem é se um agente inteligente consegue melhorar seu desempenho em tarefas futuras depois de fazer observações sobre o mundo, ou seja, aprendizagem é melhorar as ações através do estudo de suas próprias experiências.

Para redes neurais, aprendizagem ou treinamento é a adaptação dos pesos dentro da rede para melhor acomodar as entradas dadas e aprimorar a acurácia das previsões sobre entradas nunca vistas antes. Normalmente o treinamento termina quando a taxa de erro não diminui ou a redução é pequena, isso é obtido através de uma função de custo ou função de perda que é avaliada periodicamente durante o treinamento. ([ZELL, 1994](#))

Para o treinamento de redes neurais é definido a taxa de aprendizagem, *learning rate*, que é o tamanho dos passos que o modelo dá para ajustar seus pesos, uma taxa alta diminui o tempo de treinamento, porém terá uma acurácia menor ao final do treinamento, enquanto que uma taxa baixa aumenta o tempo de treinamento porém a acurácia final é maior. Para tal é usado otimizadores como o SGD (do inglês, *Stochastic gradient descent*) que tentam acelerar a minimização do erro, enquanto outros aumentam a confiabilidade.

2.2.3.1 Supervisionada

Utiliza um conjunto de entradas e saídas conhecidos, a tarefa é produzir a saída esperada para cada entrada. Nesse caso a função de custo está relacionada a eliminação de deduções incorretas [Ojha, Abraham e Snášel \(2017\)](#). Uma função de custo bastante utilizada é a erro quadrático médio (*mean-squared error*) que tenta minimizar o erro quadrático médio entre a saída da rede e a saída conhecida.

Esse método de aprendizagem é frequentemente usado para reconhecimento de padrões (ou classificação) e regressões (aproximação de funções), é aplicado a dados sequenciais, por exemplo, reconhecer letras ou números escritos a mão, reconhecer voz e palavras faladas, gestos, etc. Este é o método que será utilizado para treinar a rede com o ajuste padrão mostrado na [subseção 4.3.3](#).

2.2.3.2 Não supervisionada

Na aprendizagem não supervisionada a entrada é dada junto com uma função custo que é dependente da tarefa e a rede deve encontrar os padrões na entrada. Os problemas frequentes que usam esse modelo de aprendizagem são na área de estimativas, o que inclui o *clustering* que é a detecção de potenciais conjuntos úteis dentro do conjunto de entrada, por exemplo, estimar distribuições estatísticas, compressões e filtros [Russell e Norvig \(2016\)](#).

2.2.3.3 Por reforço

O objetivo da aprendizagem por reforço é ajustar os pesos da rede para realizar ações que minimizem os custos a longo prazo. Em cada ponto no tempo o agente realiza uma ação e o ambiente gera uma observação e então um custo instantâneo de acordo com certas regras [Russell e Norvig \(2016\)](#).

2.3 Técnicas de amostragem

Técnicas de amostragem são algoritmos que geram novas amostras baseadas em amostras já existentes.

2.3.1 SMOGN

Synthetic Minority Over-Sampling Technique for Regression with Gaussian Noise ou seu acrônimo *SMOGN* é um algoritmo que tenta resolver problemas de regressão não balanceados em que os casos mais importantes não são bem representados no conjunto de dados. O algoritmo combina dois outros algoritmos de sub-amostragem aleatórios que é o *SMOTER* e a introdução do *ruído Gaussiano*. A ideia chave do *SMOGN* é combinar as

duas estratégias para geração de amostras sintéticas com o objetivo simultâneo de limitar o risco que o *SMOTER* pode incluir usando da estratégia conservadora do ruído *Gaussiano Branco*, Torgo e Ribeiro (2017).

SMOBN gera novas amostras utilizando *SMOTER* apenas quando a amostra *seed* e a *k-próxima* amostra estão “próximas o suficiente”, quando as amostras estão “distantes” então o ruído *Gaussiano* é utilizado Branco, Torgo e Ribeiro (2017).

2.3.2 Regressão não balanceados

Tais problemas caem no contexto de tarefas de predição, a qual tem como objetivo aproximar uma função $Y = f(x)$ desconhecida. Para encontrar tal função é utilizado um conjunto de dados com N amostras. A predição se chama regressão quando a variável Y é contínua, e chamada de classificação quando Y é nominal, ou, como o nome sugere, uma classe Branco, Torgo e Ribeiro (2017).

Regressões não balanceadas é uma classe particular de problemas de regressão que são caracterizadas por duas propriedades: o usuário possui preferências não uniformes no conjunto de dados e, as características mais importantes no conjunto não estão bem representadas. Isso significa que em uma regressão não balanceada o usuário atribui mais importância para características menos frequentes que outras Branco, Torgo e Ribeiro (2017).

2.3.3 SMOTER

SMOTER é a versão do algoritmo *SMOTE* (do inglês, *Synthetic Minority Over-sampling Technique*) para regressões. O algoritmo *SMOTE* é um método de amostragem que trata de problemas de classificação com a distribuição de classes não balanceadas. Esse método combina sub-amostragem das classes mais frequentes com sobre-amostragem das classes menos frequentes Torgo et al. (2013).

3 Trabalhos correlatos

Este capítulo faz uma revisão da bibliografia existente sobre os pontos abordados no Capítulo 4, como é o caso do pré-treinamento com SAE, utilização de redes neurais para prever o resultado de problemas físicos e geração de dados sintéticos, listando trabalhos correlatos e pesquisas alinhadas com o desenvolvimento deste trabalho.

3.1 Redes neurais

Feng, Zhou e Dong (2019) utilizam uma DNN (do inglês, *deep neural network*) pré-treinada através de uma SAE (do inglês, *Stacked autoencoder*) para prever defeitos em materiais, por exemplo, segregações, poros de contração, rachaduras, os quais são problemas não lineares complexos de múltiplas variáveis. No artigo, eles utilizam uma rede neural para prever especificamente a susceptibilidade à fissuração por solidificação (*SCS*).

Como é um problema complexo e importante da área de materiais, então é extensamente investigado por métodos experimentais e alguns métodos numéricos, porém ainda não resolvidos. Investigações experimentais sobre esse problema são caros e demorados, portanto, apenas algumas ligas metálicas são estudadas e os dados são esparsos e espalhados, o que indica que coletar grandes conjuntos de dados sobre *SCS* se mostra um grande desafio.

Os autores argumentam que por conta da natureza dos problemas na área de materiais terem número de entradas menores que os encontrados nas áreas de *data mining* e reconhecimento de imagens, por exemplo para processar uma imagem de 100x100 pixels é necessário uma entrada de pelo menos 10^4 variáveis enquanto que problemas na área de materiais dificilmente possuem mais de 100 variáveis de entrada, esse é o principal motivo que se faz necessário grandes conjuntos de dados para treinamento e validação. Com isso em mente os autores demonstram a viabilidade de utilizar uma *DNN* mesmo com um número limitado de dados, no caso desse artigo eles trabalharam com um conjunto de 575 dados coletados da literatura.

Szénási et al. (2018) utilizam uma rede neural para estimar o coeficiente de transferência de calor (*HTC*), a tarefa de prever o *HTC* pertence aos problemas de transferência de calor inverso (*IHCP*) os quais não podem ser resolvidos por métodos numéricos diretos. As maneiras comuns de resolver o *IHCP* são baseadas em algoritmos de busca com heurísticas, como algoritmos genéticos, ou Otimização por enxame de partículas e outros métodos usando teoria de enxame. O problema destes métodos é que são computacionalmente caros, tornando o processo demorado.

A solução proposta pelos autores é utilizar uma rede neural com uma função aproximadora universal, como a rede possui uma quantidade relativamente grande de entradas (120) e

saídas (101) então é necessário um número grande de amostras. Logo, eles optaram por criar as amostras utilizando milhares de simulações com *HTCs* teóricos.

Trabalho	Feng, Zhou e Dong (2019)	Szénási et al. (2018)
Problema	Prever defeitos em materiais	Prever HTC
Solução	DNN pré-treinada com SAE	Função aproximadora universal baseada em rede neural
Tipo do conjunto de dados	Experimental	Sintético
Tamanho do conjunto de dados	Pequeno, em torno de 575 amostras	Grande, pode gerar quantas amostras necessárias

Tabela 2 – Síntese das técnicas usadas nos trabalhos correlatos.

4 Desenvolvimento

Neste capítulo é descrita a solução do problema proposto. Na [seção 4.1](#), é apresentada a solução proposta, em seguida na [seção 4.2](#) é discutido como é gerado os dados sinteticamente junto com sua análise. Na [seção 4.3](#), é apresentado como os modelos finais foram gerados e treinados. Por fim, na [seção 4.4](#), são discutidos os resultados obtidos com cada rede e seus respectivos treinamentos.

4.1 Metodologia

O desenvolvimento deste trabalho consiste em examinar e categorizar a distribuição de dados, normalizar a entrada de dados, propor modelos de redes neurais para prever os dados e, por fim, analisar o desempenho dos modelos gerados.

Foi utilizada uma rede neural para prevermos o coeficiente de condutância global (UA) de refrigeradores domésticos usando dados sintéticos para treiná-la, a utilização de tais dados é porque não há muitos dados experimentais disponíveis para treinamento, enquanto que, mesmo perdendo acurácia, os dados sintéticos podem ser gerados em grandes quantidades.

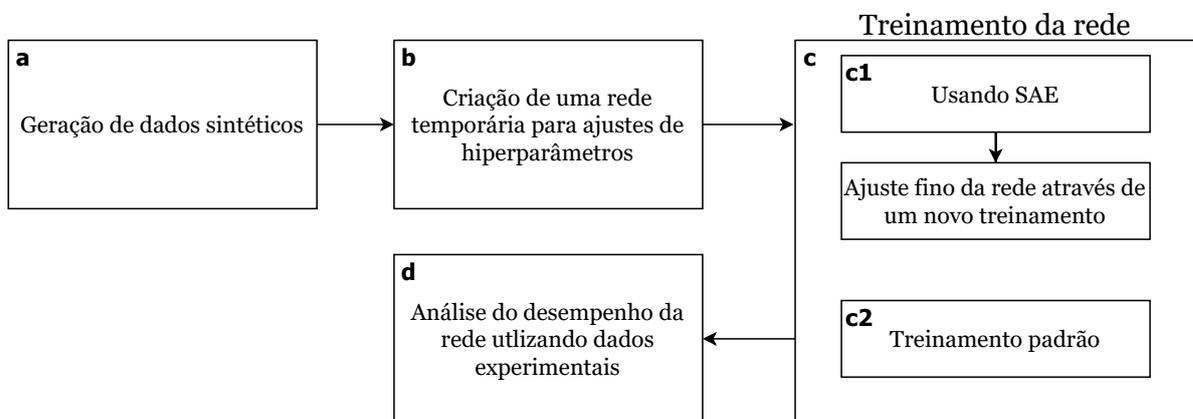


Figura 7 – Diagrama da metodologia adotada.

A metodologia adotada para o desenvolvimento deste trabalho pode ser dividida em quatro etapas, como ilustra o diagrama da [Figura 7](#):

- a) Primeiro são gerados os dados sintéticos e pseudo-experimentais que, como será discutido na [subseção 4.2.2](#), os dados experimentais obtidos não são suficientes para um treinamento, então tais dados foram submetidos à uma técnica de amostragem e então é gerado em torno de 200 amostras sintéticas para simular como se fossem os experimentais.

- b) É criada uma rede preliminar em que será submetida às técnicas de ajuste de hiperparâmetros utilizando os dados sintéticos gerados anteriormente. Esta etapa indica o número de neurônios, camadas ocultas, funções de ativação, taxa de aprendizagem, etc.
- c) O treinamento da rede é dividido em dois:
1. É criado um *Stacked autoencoder (SAE)* para inicializar os pesos e *biases* da rede neural para que fiquem próximos da solução global e, depois ela é submetida a um treinamento padrão utilizando os dados pseudo-experimentais. A [subseção 2.2.1](#) discute como é feito o pré-treinamento.
 2. A rede é treinada apenas com dados sintéticos de maneira padrão (sem pré-treinamento), o método padrão utilizado é o treinamento supervisionado, apresentado na [subseção 2.2.3.1](#).
- d) Os dados experimentais são utilizados para verificar a acurácia da rede e, por fim, são extraídas as conclusões do projeto.

4.2 Dados

Talvez uma das partes mais importantes no desenvolvimento de uma rede neural é a correta análise da qualidade, da variabilidade e o apropriado pré-processamento dos dados envolvidos no treinamento e validação da rede para iniciar os hiperparâmetros mais adequados (ainda não otimizados) do problema. Para o problema desenvolvido neste trabalho são utilizados dois grupos de dados, os dados sintéticos, que são gerados a partir de teses, dissertações e produtos encontrados no mercado, e os dados experimentais, que são obtidos a partir dos experimentos documentados em teses e dissertações.

4.2.1 Geração de dados sintéticos

São duas formas de geração de dados. A primeira (referido no texto como *uniforme*) é utilizando distribuições uniformes para a altura, largura e profundidade. Tais valores são retirados de web sites de compras que indicam os parâmetros geométricos mínimos e máximos. As espessuras também são escolhidas através de uma distribuição uniforme, porém os intervalos de mínimo e máximo são retirados de trabalhos acadêmicos. Cada parede possui um intervalo; as paredes da esquerda e direita são iguais; a traseira normalmente é mais espessa e a da porta geralmente é a menos espessa. As paredes do congelador são mais espessas do que as do refrigerador. A capacidade de um recipiente com paredes de espessura l pode ser calculada utilizando a seguinte equação,

$$V = (h - 2l) * (L - 2l) * (p - 2l) \quad (4.1)$$

que para o caso de um compartimento i (congelador ou refrigerador) podemos escrever como,

$$V_i = \frac{(h - l_{si} + l_{Ii}) * (L - 2l_{ei}) * (p - l_{pi} + l_{ti})}{10^6} \quad (4.2)$$

onde os subíndices s, I, e, p, t representam, respectivamente, parede superior, inferior, esquerda (a direita tem a mesma espessura), porta e traseira. Divide-se por 10^6 para obter o volume V_i em litros.

A capacidade de um gabinete é dividida em dois, um para o refrigerador e um para o congelador, na [Figura 12](#) é ilustrado um gabinete *top-down*, há uma linha dividindo os compartimentos que é chamada de travessa, entra no cálculo de ambos como espessura superior (ou inferior no caso do congelador). Cada círculo representa um parâmetro l da [Equação 4.2](#), e os outros parâmetros tendo a mesma denominação, com exceção da altura, a qual é gerada uma altura para cada compartimento e depois ambas são somadas junto com as espessuras superior, inferior e da travessa para obter a altura total do refrigerador.

Após gerar os dados base, que são a altura, largura, profundidade e a espessura de cada parede, é então escolhida uma configuração de isolantes, como é ilustrado na [Figura 8](#), na qual cada cor indica um isolante diferente, neste caso o máximo de isolantes diferentes é 2. Então é escolhido, aleatoriamente, um isolante da [Tabela 3](#), a qual enumera os isolantes mais comuns em refrigeradores domésticos. Por fim, utilizando a [Equação 2.3](#), é calculado o UA_C e o UA_R que servirão como *rótulo* (do inglês, *label*) para o treinamento da rede.

Isolante	Condutividade [W/mK]
PU	0.028
VIP com núcleo de fibra de vidro	0.0015

Tabela 3 – Resistências térmicas dos materiais isolantes utilizados. Adaptado de [Thiessen \(2015\)](#).

A segunda forma (referido no texto como *proporcional*) de geração de dados é utilizando proporção o qual divide a altura total pela altura de cada compartimento, a altura total também é dividida pela largura e profundidade, já as espessuras e os isolantes são geradas da mesma forma descrita anteriormente e após a geração das espessuras, então é calculado os volumes de cada compartimento e seus respectivos UAs .

São gerados cerca de 10 mil dados “uniformes” e 20 mil dados “proporcionais” para realizar o treinamento das redes e o pré-treinamento dos SAE , é possível ver a distribuição e exemplos dos dados “uniformes” nas [Figuras 10, 18, 19 e 20](#) e exemplos dos dados proporcionais gerados nas [Figuras 9, 21, 22 e 23](#).

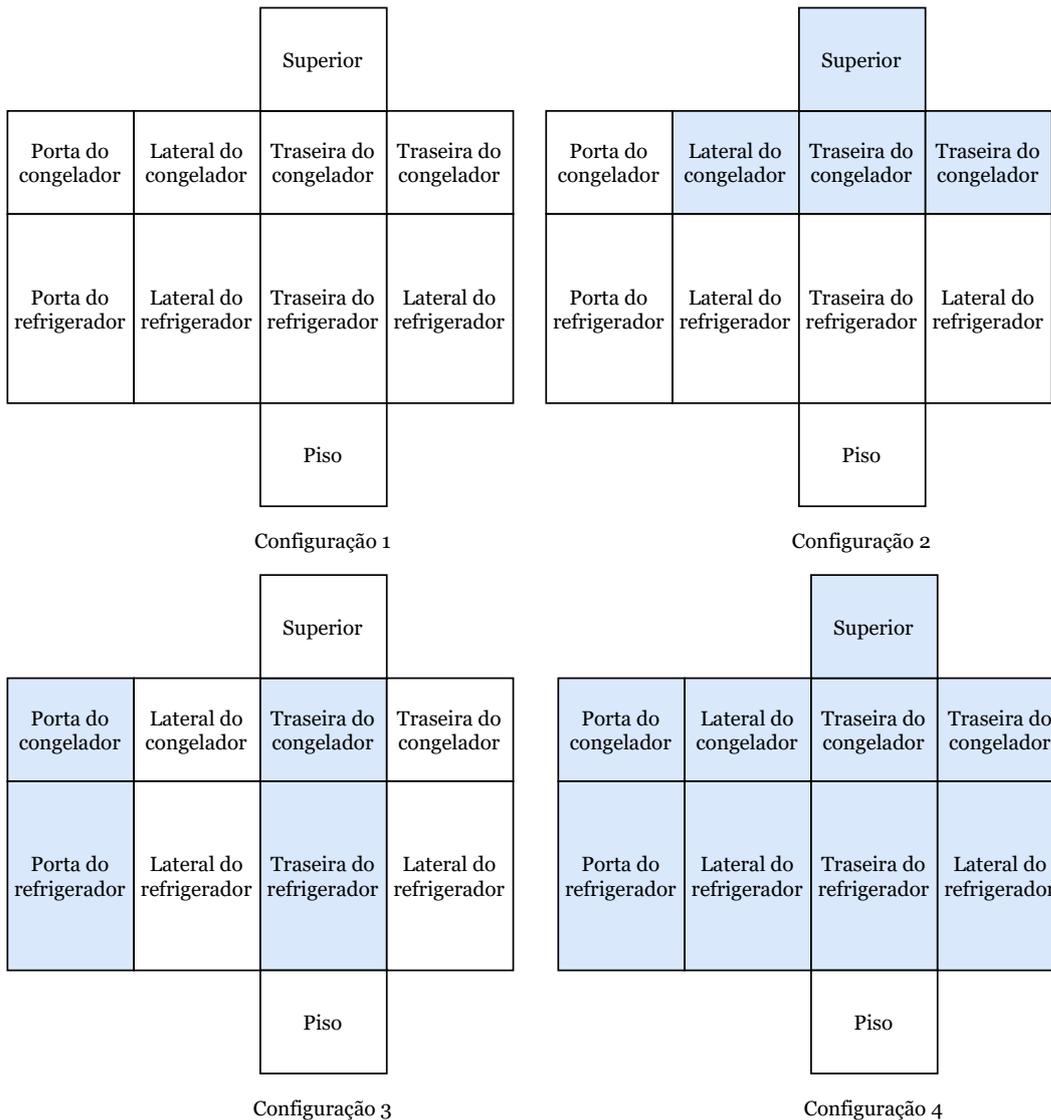


Figura 8 – Configurações de isolantes de cada parede do gabinete do refrigerador. Adaptado de Thiessen (2015).

4.2.1.1 Qualidade dos dados

Os gráficos exibidos nas Figuras 21, 22, 23 (que se encontram no Apêndice A) e 9 mostram os dados gerados através de proporção, nota-se claramente uma divisão linear entre os dados, o motivo é porque foram utilizados dois refrigeradores de proporções diferentes, o número de divisões aumentaria com diferentes proporções. Já nos gráficos das Figuras 18, 19, 20 (que se encontram no Apêndice A) e 10 mostram dados mais esparsos por conta que não possuem restrição quanto à sua geração. As Tabelas 11 e 12 no Apêndice C, respectivamente, uniforme e proporcional, apresentam as métricas dos dados sintéticos gerados.

Em uma análise preliminar visual dos dados e resultados, mais aprofundados na seção 4.4, é possível presumir que o treinamento com os dados de tendência linear (como

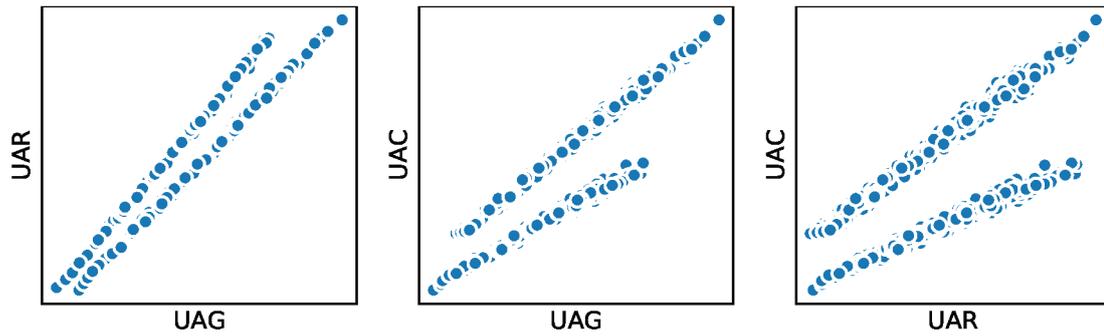


Figura 9 – Gráfico mostrando as saídas UA_R , UA_C e UA_G , gerado utilizando proporções.

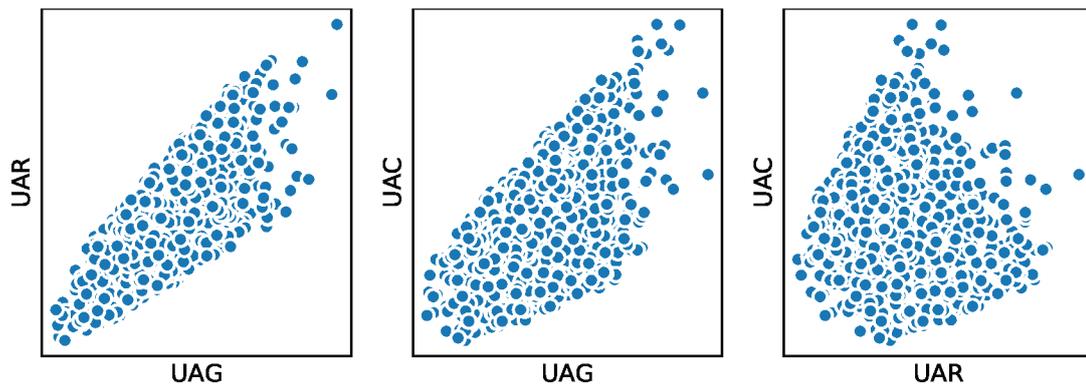


Figura 10 – Gráfico mostrando as saídas UA_R , UA_C e UA_G , caso uniforme.

mostrado na Figura 9) possuem melhores predições quando as entradas¹ estão dentro ou próximas desta tendência, enquanto que quando treinada com os dados mais esparsos (mostrado na Figura 10) a rede mostra mais generalização, isso significa que a maioria das entradas têm uma predição relativamente boa porém para entradas que estão dentro das tendências lineares o conjunto proporcional de treino é melhor.

4.2.2 Dados experimentais

Os dados experimentais foram retirados de dissertações, artigos e teses. Infelizmente não há uma base de dados de fácil acesso que contenha tais informações e, por conta da dificuldade de obtenção, não há muitos dados disponíveis para este trabalho.

Porém os poucos dados obtidos são submetidos a uma técnica de amostragem chamada de *Synthetic Minority Over-Sampling Technique for Regression with Gaussian Noise*, ou simplesmente *SMOGRN*, apresentado na subseção 2.3.3, o qual gera, baseado nos dados fornecidos, mais dados, chamaremos no texto de *dados pseudo-experimentais*. Apesar de poder criar quantas amostras desejar, é necessário levar em conta que não haverá muita variabilidade neles. Por tanto foram utilizados para verificar a acurácia da rede.

¹ Que nunca foram vistas pela rede antes.

A Figura 11 mostra os dados experimentais originais, em azul, e os gerados através da técnica, em laranja. É possível notar que os dados gerados são próximos dos originais, o desvio padrão é baixo (em torno de 0.18 para a média de 1.73 do UA_G), porém por maior que seja o número de dados gerados, eles sempre estarão próximos dos já existentes, logo não é possível utilizá-los para treinamento de uma rede por conta da pouca variabilidade. A Tabela 10 do Apêndice C apresenta uma parte dos dados experimentais colhidos.

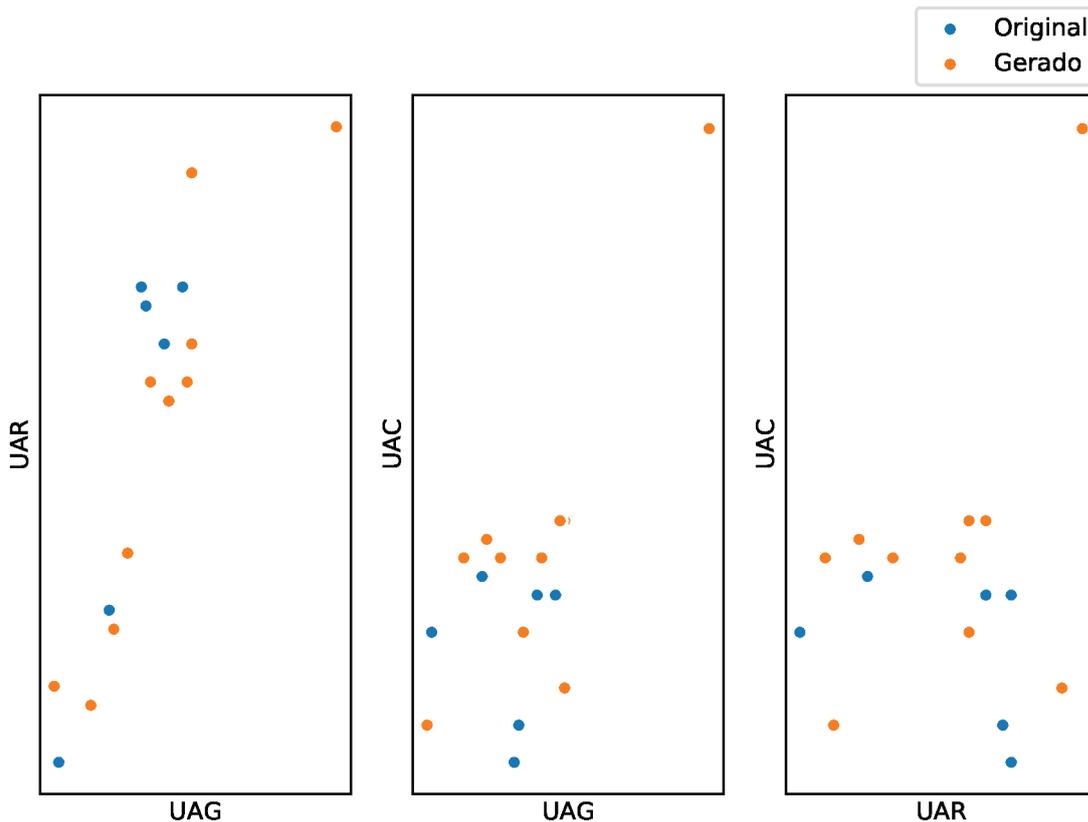


Figura 11 – Gráfico mostrando as saídas UA_R , UA_C e UA_G . Experimentais e pseudo-experimentais.

4.3 Modelo

Nesta seção, é exposto como é criado o modelo, o ajuste dos hiperparâmetros e seu treinamento. O modelo foi construído utilizando a biblioteca *tensorflow*, Abadi et al. (2015), da Google. Foi selecionada devido o número de implementações existentes e por ser uma biblioteca que está sempre em aprimoramento.

Há dois modelos de rede, a primeira chamada *rede dual* porque ela tem ambos isolantes, refrigerador e congelador, como entradas e os UA como saída. A outra rede, chamada de *rede única*, possui apenas as entradas e saídas respectivas de cada compartimento. No quesito de treinamento, há duas técnicas adotadas, a primeira sendo a implementação de *SAEs* para pré-treinar e depois ajustar com dados pseudo-experimentais e a segunda

é a técnica de ajuste (do inglês, *fit*) padrão de uma rede neural utilizando apenas dados sintéticos, na [subseção 4.3.3](#) é detalhado o processo de treinamento de cada rede.

4.3.1 Entradas e saídas

Os dados de entrada para a rede são, como ilustra a [Figura 12](#), a geometria do refrigerador que inclui altura (h) [mm], largura (L) [mm], profundidade (p) [mm], capacidade do refrigerador e congelador [litros] e os isolantes [W/mK] de cada parede representados pelos círculos na imagem (são 10), totalizando 15 entradas. Não é utilizada a espessura das paredes como entrada por três motivos:

- Simplificação da rede com menos entradas, conseqüentemente é necessário menos dados para o treinamento;
- As espessuras podem ser inferidas a partir da capacidade dos compartimentos, largura e profundidade;
- Os fabricantes de refrigeradores não fornecem dados de espessura nos manuais disponíveis no seu web site, ficando a busca limitada a dados encontrados em teses, artigos e dissertações.

A rede gera duas saídas que são os coeficientes UA_C e UA_R que então são somados para obter o UA_G ([Equação 2.4](#)).

4.3.2 Ajuste dos hiperparâmetros

A flexibilidade das redes neurais vem com um dos seus maiores inconvenientes que é a grande quantidade de hiperparâmetros para ajustar. Tais parâmetros tem grande importância no sucesso da arquitetura da rede, uma vez que impactam diretamente o treinamento e seu desempenho. Para uma rede é possível alterar o número de camadas, número de neurônios por camada, o tipo de função de ativação de cada camada, a lógica de inicialização dos pesos, e muito mais [Géron \(2017\)](#).

O ajuste foi realizado utilizando a técnica de otimização *bayesiana* (do inglês, *Bayesian Optimization*) da biblioteca *Keras Tuner*, [Chollet et al. \(2015\)](#), com o erro médio quadrado como função de perda e as métricas observadas são tanto o erro médio quadrático (*MSE*) quanto o erro médio absoluto (*MAE*).

Na [Tabela 4](#), são apresentados os valores que são buscados pelo algoritmo, tais valores são definidos manualmente pelo usuário, os valores são, se necessário, reajustados para uma nova busca, uma vez que esta etapa é bastante empírica e computacionalmente custosa, por exemplo, o usuário pode ver que a rede não obteve um bom desempenho com cinco camadas, então esse valor é reajustado e submetido ao algoritmo novamente.

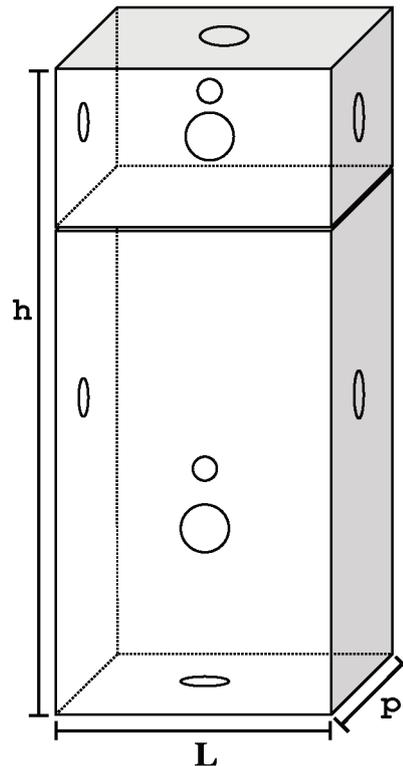


Figura 12 – Ilustração dos dados de entrada da rede.

hiperparâmetro	Valores de busca
Otimizadores	Adadelta, Adagrad, Nadam, SGD, Ftrl, RMSprop, Adam, Adamax
Funções de ativação	Tanh, sigmoid, exponential, relu, elu, hard sigmoid, linear, selu, softmax, softplus, softsign
Camadas ocultas	de 2 a 10
Neurônios nas camadas ocultas	de 3 a 24

Tabela 4 – Valores para a busca de hiperparâmetros.

A Tabela 5 apresenta os hiperparâmetros utilizados para cada rede, os hiperparâmetros que foram submetidos à busca são os otimizadores, todas as funções de ativação e o *layout* da camada oculta. Os valores na linha “Camada oculta” representam os *layouts* das camadas ocultas, por exemplo, [6,5,4,3] indica que são 4 camadas ocultas (tamanho da lista) e a primeira tem seis neurônios, a segunda cinco, e assim por diante, sendo cada camada do tipo densa, ou seja, todos os neurônios da primeira camada são ligados com os da segunda e todos da segunda camada ligados com da terceira e assim por diante.

hiperparâmetro	Rede dual	Rede Dual 1 config	Rede congelador	Rede refrigerador
Épocas	1000	1000	1000	1000
Taxa de aprendizado	0.05	0.001	0.001	0.01
Otimizador	SGD	Adagrad	Adamax	RMSprop
Função de ativação da camada de entrada	Relu	Tanh	Softsign	Softplus
Função de ativação das camadas ocultas	Relu	Tanh	Relu	Selu
Função de ativação da camada de saída	Exponential	Tanh	Selu	Elu
Neurônios na camada de entrada	15	15	9	9
Camada oculta	[6, 21, 24, 15]	[6, 5, 4, 3]	[12,12,12,3,3,3]	[12,18,3,3,3,3,3,3]
Neurônios na camada de saída	2	2	1	1

Tabela 5 – hiperparâmetros ajustados da rede.

4.3.3 Treinamento

As redes foram submetidas a dois tipos de treinamento, o primeiro pré-treinando utilizando *SAE* (subseção 2.2.1) com um conjunto sintético de dados depois fazendo um ajuste fino utilizando o conjunto pseudo-experimental, e o segundo utilizando a função de ajuste padrão da rede (treinamento supervisionado, como apresentado na subseção 2.2.3.1) utilizando apenas dados sintéticos. Os parâmetros utilizados para cada treinamento podem ser vistos na Tabela 6. Na coluna *SAE* há dois valores: o primeiro sendo para o treinamento do *SAE* e o segundo é para o ajuste fino. A Tabela 7 da seção 4.4 é mostrada a acurácia da rede juntamente com discussões acerca dos resultados obtidos.

Parâmetro	SAE	Padrão
Épocas	1000/100	1000
Número de Amostras conjunto uniforme	10 000/200	10 000
Número de Amostras conjunto proporcional	10 000/200	10 000

Tabela 6 – Parâmetros de cada treinamento.

4.4 Resultados

Na Tabela 7 é mostrada a rede, com a técnica usada e o conjunto de dados o qual foi submetido ao treino. A acurácia é determinada pela função de erro médio quadrático (*MSE*). Quanto menor esse valor maior é a precisão da rede, para ser considerado um

Rede	Técnica de treinamento	Conjunto de dados	Acurácia Conjunto de testes	Acurácia conjunto Pseudo-Experimentais
Rede dual	SAE	Uniforme	0.4165	0.623
		Proporcional	0.5509	0.627
	Padrão	Uniforme	0.836	0.405
		Proporcional	1.830	0.786
Rede única Congelador	SAE	Uniforme	0.028	0.602
		Proporcional	0.048	0.316
	Padrão	Uniforme	0.0067	0.373
		Proporcional	0.0022	0.279
Rede única Refrigerador	SAE	Uniforme	0.014	0.497
		Proporcional	0.606	0.109
	Padrão	Uniforme	0.006	0.540
		Proporcional	0.156	0.915

Tabela 7 – Acurácia das redes.

desempenho bom é necessário que o erro, sobre ambos conjuntos, seja menor que 0.2. Esse valor de erro é dependente do conjunto de dados, por exemplo, se o conjunto é bastante esparso então a tendência é do erro ser maior do que em um conjunto normal. No caso desta rede em específico é definido, através de observações empíricas sobre os resultados das diferentes redes e comparando com os dados experimentais, que, para a predição ser relativamente boa, a rede deva atingir um erro menor que 0.2.

Nota-se que a rede dual teve o pior desempenho entre todas, enquanto que as redes de saída única possuem os melhores resultados, com exceção da rede única que prevê para o refrigerador treinada com o conjunto proporcional, talvez com hiper-parâmetros diferentes a acurácia seria maior. De maneira geral, as que obtiveram os melhores resultados são aquelas pré-treinadas usando o *SAE*.

Esperava-se que as redes pré-treinadas teriam uma acurácia maior sobre o conjunto experimental por conta do seu segundo treinamento usando os dados pseudo-experimentais, mesmo usando poucas épocas no treinamento. Também era esperado que o conjunto proporcional, de maneira geral, tivesse uma acurácia maior do que o uniforme, pelo menos sobre o conjunto pseudo-experimental.

As Figuras 13, 14, 15, 16 e 17 apresentam um gráfico dos *valores esperados* x *valores previstos* com os resultados do treinamento padrão, em azul, e usando *SAE*, em laranja, a linha vermelha traça os valores ideais, ou seja, quanto mais próximo da linha melhor é a previsão que a rede fez, em outras palavras, menor é o erro. É esperado que o gráfico apresente os valores de forma linear, seguindo a tendência da linha vermelha (mesmo que estejam um pouco fora do ideal). Quando os valores estão muito dispersos, sem tendências, isso quer dizer que a rede não conseguiu se adequar, ou se ajustar, com os dados.

Na [Figura 13](#), é apresentado os resultados da previsão, sendo que na primeira coluna é exibido os valores previstos para UA_R e na segunda os valores para UA_C . Note que, mesmo que o erro esteja grande, os valores previstos para o UA_C utilizando SAE com o conjunto uniforme são melhores que o padrão. Já na [Figura 14](#) é mostrado os valores previstos sobre o conjunto pseudo-experimental, os valores obtidos não são satisfatórios, pois estão totalmente dispersos.

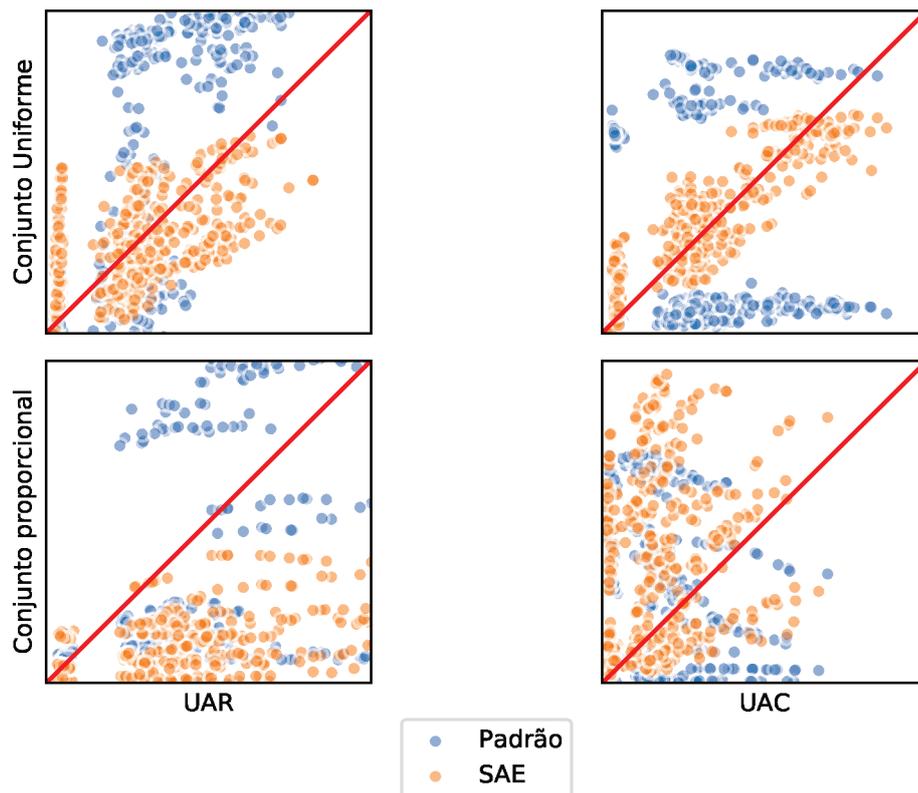


Figura 13 – Resultados da previsão da rede dual.

Infelizmente, não foi possível chegar a um resultado satisfatório com a *rede dual* utilizando as quatro configurações de refrigeradores simultaneamente, porém se utilizar apenas uma configuração por vez é possível obter resultados melhores, como é mostrada na [Tabela 8](#) e as previsões feitas na [Figura 15](#), note como os dados estão seguindo uma tendência linear e estão poucos dispersos. O treinamento foi realizado utilizando apenas a configuração 1 da [Figura 8](#) e com os parâmetros mostrados na [subseção 4.3.3](#).

Nas [Figuras 16](#) e [17](#) são apresentados os resultados das previsões sobre os conjuntos de dados do refrigerador e congelador, respectivamente. As previsões da primeira coluna são das redes treinadas no conjunto uniforme enquanto que na segunda coluna é treinado utilizando o conjunto proporcional. Note que o conjunto uniforme possui mais generalização, logo a previsão sobre esse conjunto se torna boa, como é mostrada no primeiro gráfico de cada rede, já quando a mesma rede é treinada sobre um conjunto proporcional ela perde um pouco da generalização, porém a previsão sobre o conjunto pseudo-experimental se

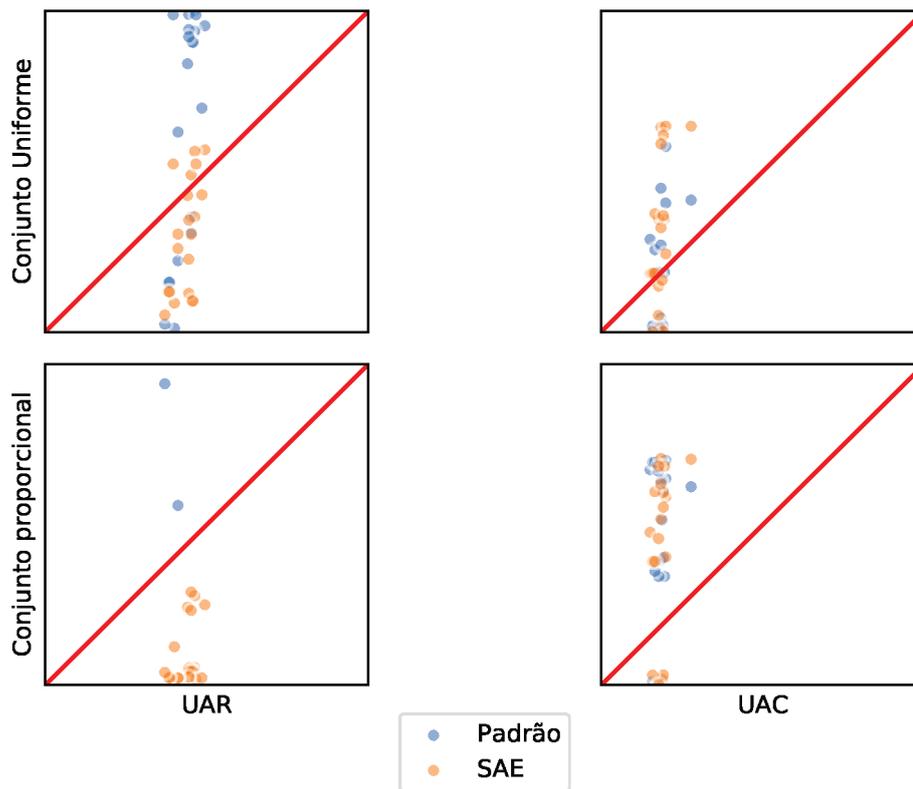


Figura 14 – Resultados da previsão dos dados pseudo-experimentais da rede dual.

Rede	Técnica Treinamento	Conjunto de dados	Acurácia Conjunto de testes	Acurácia Conjunto pseudo-experimental
Rede dual	Padrão	Uniforme	0.014	0.168
		Proporcional	0.018	0.157

Tabela 8 – Acurácia das redes com apenas uma configuração de refrigerador.

torna melhor. Talvez fosse possível obter uma melhor generalização se esta rede passasse pelo processo de ajuste de hiper-parâmetros utilizando os dados proporcionais para tal.

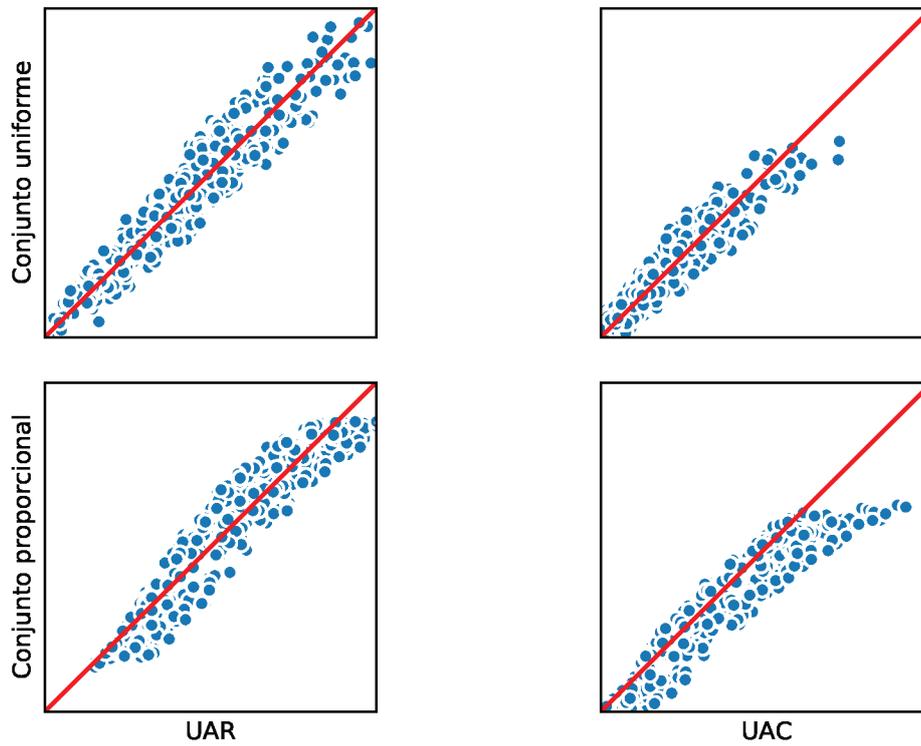


Figura 15 – Resultados da previsão da rede dual com uma configuração de isolantes.

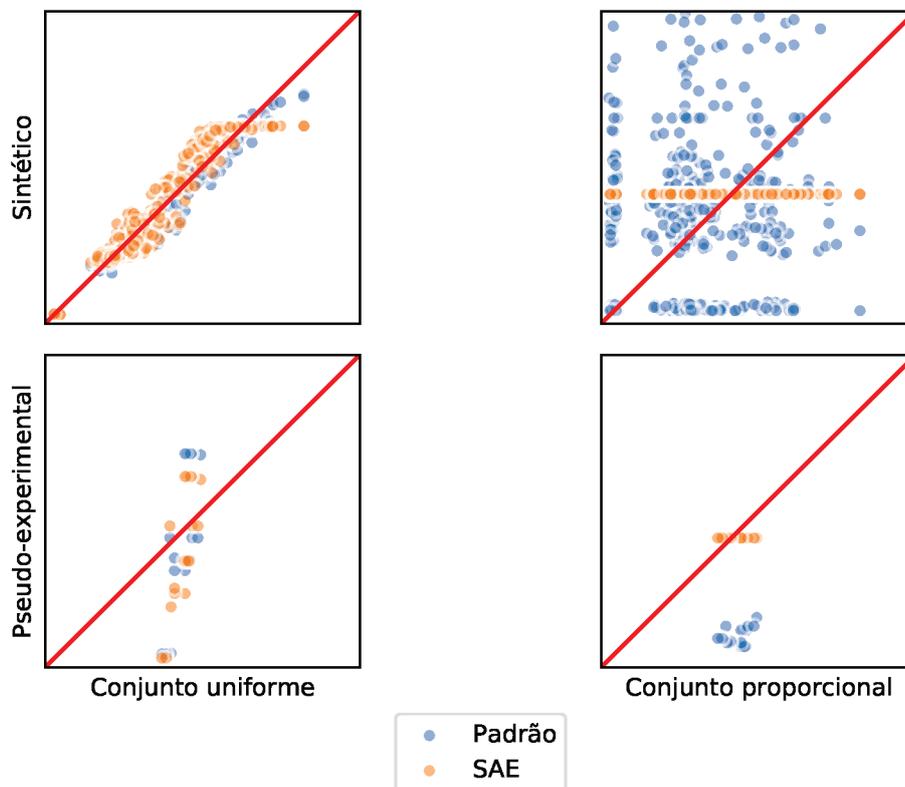


Figura 16 – Resultados da previsão da rede única para o refrigerador.

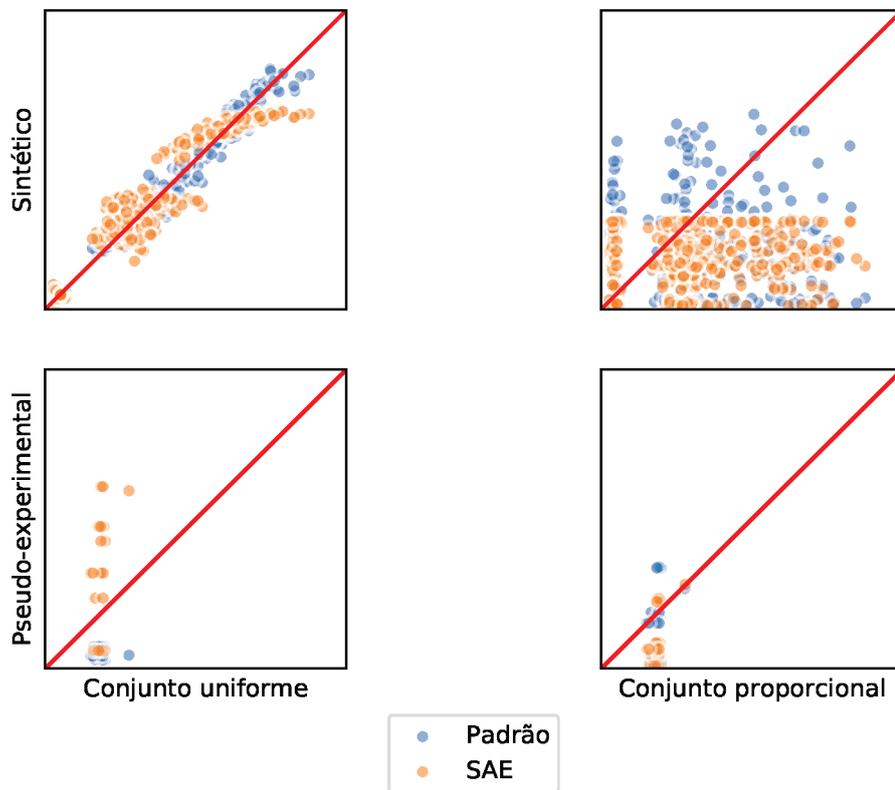


Figura 17 – Resultados da previsão da rede única para o congelador.

5 Conclusão

As técnicas apresentadas neste trabalho se tornam úteis quando é necessário lidar com conjuntos de dados pequenos ou com pouca diversidade. Foram expostas ideias, técnicas e algoritmos para que se amplie tal conjunto e que também seja possível ajustar uma rede neural ao problema.

Ambas técnicas de treinamento, *SAE* e a padrão se mostraram bastante eficientes porém, a *SAE* apresentou um bom desempenho quando se trata do conjunto experimental. Quando é necessário maior generalização, como é o caso do conjunto uniforme, ela tem o desempenho degradado e é preferível a técnica padrão. Talvez a maior vantagem da *SAE*, que foi mais explorada por [Feng, Zhou e Dong \(2019\)](#), é o treinamento com um conjunto pequeno de dados.

O treinamento sobre o conjunto proporcional se mostrou bastante eficaz quando se trata de prever os dados que estão dentro ou próximos desta proporção, como é o caso dos experimentais, cujas proporções foram utilizadas para criar o conjunto proporcional. Já o treinamento sobre o conjunto uniforme se mostrou o mais genérico. Porém, é aquele que tem o menor desempenho médio considerando a acurácia sobre o conjunto experimental.

O conjunto de configuração única, que não foi muito explorado nesse trabalho, se mostrou bastante útil quando foi necessário diminuir a dimensão e variabilidade das entradas do modelo. Por conta disso a rede dual teve um bom desempenho quando treinada sobre este conjunto. A desvantagem é que o número de modelos cresceria junto com o número de configurações, talvez tornando inviável esta solução.

Por fim, as redes tiveram um desempenho relativamente bom, mostrando ser viável o uso de redes neurais para previsão de propriedades físicas.

5.1 Trabalhos futuros

Nesta seção, são listadas algumas sugestões de trabalhos futuros que queiram utilizar este trabalho, seja como ponto de partida para outro trabalho ou continuação do mesmo.

- Obter mais dados experimentais para melhores resultados com a metodologia proposta neste trabalho ou até mesmo testar esta mesma metodologia para prever um outro problema envolvendo previsão de propriedades físicas.
- Outro ponto interessante seria fazer uso de outra técnica de ajuste dos hiperparâmetros, especialmente para as redes que não obtiveram resultados satisfatórios, como é

o caso da rede *dual*.

- Elaborar uma comparação com outros métodos de aproximação, por exemplo, rede neural *Bayesiana*, rede *Bayesiana*, árvore de regressão, entre outras técnicas, pois é possível que estas técnicas tenham um melhor desempenho para prever este tipo de problema.

Referências

- ABADI, M. et al. *TensorFlow: Large-Scale Machine Learning on Heterogeneous Systems*. 2015. Software available from tensorflow.org. Disponível em: <<https://www.tensorflow.org/>>. Citado na página 44.
- AMUDHA, V. *Why First Hidden Layer Is Very Important In Building a Neural Network Model And Relation Between Vanishing Gradient And First Hidden Layer*. 2019. Disponível em: <<https://bit.ly/2I2cPKm>>. Citado na página 33.
- BARTHEL, C.; GÖTZ, T. The overall worldwide saving potential from domestic refrigerators and freezers. *With results detailed for*, v. 11, 2012. Citado na página 23.
- BRANCO, P. O.; TORGO, L.; RIBEIRO, R. P. Smogn: a pre-processing approach for imbalanced regression. 2017. Citado na página 36.
- CHOLLET, F. et al. *Keras*. 2015. <<https://keras.io>>. Citado na página 45.
- CLAESEN, M.; MOOR, B. D. Hyperparameter search in machine learning. *arXiv preprint arXiv:1502.02127*, 2015. Citado na página 34.
- COLOMBO, E. G. *Investigação numérica e experimental do desempenho de condensadores do tipo skin em refrigeradores domésticos*. Dissertação (Mestrado) — Universidade Federal de Santa Catarina, Centro Tecnológico, Programa de Pós-Graduação em Engenharia Mecânica, 2014. Citado 2 vezes nas páginas 25 e 28.
- ELETROBRÁS. *PROCEL - Programa Nacional de Conservação de Energia Elétrica*. [S.l.], 2004. Disponível em: <<http://www.eletronbras.gov.br/procel/>>. Citado na página 23.
- FENG, S.; ZHOU, H.; DONG, H. Using deep neural network with small dataset to predict material defects. *Materials & Design*, Elsevier, v. 162, p. 300–310, 2019. Citado 3 vezes nas páginas 37, 38 e 53.
- GÉRON, A. *Hands-On Machine Learning with Scikit-Learn and TensorFlow: Concepts, Tools, and Techniques to Build Intelligent Systems*. O'Reilly Media, 2017. ISBN 9781491962244. Disponível em: <<https://books.google.com.br/books?id=bRpYDgAAQBAJ>>. Citado 2 vezes nas páginas 31 e 45.
- GERVEN, M. van. Editorial: Artificial neural networks as models of neural information processing. *front. comput. neurosci.* 11: 114. doi: 10.3389/fncom.2017.00114 editorial: Artificial neural networks as models of neural information processing marcel van gerven 1* and sander bohte 2 1 department of artificial intelligence. *Donders Institute for Brain, Cognition and Behaviour, Radboud University Nijmegen, Nijmegen, Netherlands, 2 Department of Machine Learning, Centrum Wiskunde and Informatica, Amsterdam, Netherlands*, 2017. Citado na página 30.
- HERMES, C. J. L. *Uma metodologia para a simulação transiente de refrigeradores domésticos*. Tese (Doutorado) — Universidade Federal de Santa Catarina, Centro Tecnológico, Programa de Pós-Graduação em Engenharia Mecânica, 2006. Citado 3 vezes nas páginas 23, 26 e 27.

- INCROPERA, F. P. et al. *Fundamentals of heat and mass transfer*. [S.l.]: Wiley, 2007. Citado 3 vezes nas páginas 25, 26 e 59.
- INMETRO. *Refrigeradores, frigobares, combinados, combinados frost-free*. [S.l.], 2020. Disponível em: <<http://www.inmetro.gov.br/consumidor/pbe/refrigeradores.pdf>>. Citado na página 23.
- JAMES, S.; EVANS, J.; JAMES, C. A review of the performance of domestic refrigerators. *Journal of Food Engineering*, Elsevier, v. 87, n. 1, p. 2–10, 2008. Citado na página 23.
- JUNIOR, A. A. G.; SOUSA, A. R. de. *Fundamentos de metrologia científica e industrial*. [S.l.]: Manole, 2008. Citado na página 26.
- LEONEL, J. *Hyperparameters in Machine/Deep Learning*. 2019. Disponível em: <<https://medium.com/@jorgesleonel/hyperparameters-in-machine-deep-learning-ca69ad10b981>>. Citado 3 vezes nas páginas 31, 32 e 33.
- MURPHY, K. *Machine Learning: A Probabilistic Perspective*. MIT Press, 2012. (Adaptive Computation and Machine Learning series). ISBN 9780262018029. Disponível em: <<https://books.google.com.br/books?id=NZP6AQAAQBAJ>>. Citado na página 31.
- OJHA, V. K.; ABRAHAM, A.; SNÁŠEL, V. Metaheuristic design of feedforward neural networks: A review of two decades of research. *Engineering Applications of Artificial Intelligence*, Elsevier, v. 60, p. 97–116, 2017. Citado na página 35.
- RUSSELL, S. J.; NORVIG, P. *Artificial intelligence: a modern approach*. [S.l.]: Malaysia; Pearson Education Limited,, 2016. Citado 3 vezes nas páginas 29, 34 e 35.
- SEIDEL, R. *Análise da performance e da distribuição de fluxo de calor num refrigerador doméstico*. Dissertação (Mestrado) — Universidade Federal de Santa Catarina, Centro Tecnológico, Programa de Pós-Graduação em Engenharia Mecânica, 2001. Citado na página 26.
- SHARMA, A. *Understanding Activation Functions in Neural Networks*. 2017. Disponível em: <<https://medium.com/the-theory-of-everything/understanding-activation-functions-in-neural-networks-9491262884e0>>. Citado na página 34.
- SHARMA, S. *Epoch vs Batch Size vs Iterations*. 2017. Disponível em: <<https://towardsdatascience.com/epoch-vs-iterations-vs-batch-size-4dfb9c7ce9c9>>. Citado 2 vezes nas páginas 32 e 33.
- SIM, J. S.; HA, J. S. Experimental study of heat transfer characteristics for a refrigerator by using reverse heat loss method. *International Communications in Heat and Mass Transfer*, Elsevier, v. 38, n. 5, p. 572–576, 2011. Citado na página 26.
- SNOEK, J.; LAROCHELLE, H.; ADAMS, R. P. Practical bayesian optimization of machine learning algorithms. In: *Advances in neural information processing systems*. [S.l.: s.n.], 2012. p. 2951–2959. Citado na página 34.
- SZÉNÁSI, S. et al. Estimating the heat transfer coefficient using universal function approximator neural network. In: IEEE. *2018 IEEE 12th International Symposium on Applied Computational Intelligence and Informatics (SACI)*. [S.l.], 2018. p. 000401–000404. Citado 2 vezes nas páginas 37 e 38.

THIESSEN, S. *Caracterização experimental de gabinetes refrigerados isolados com painéis de vácuo*. Dissertação (Mestrado) — Universidade Federal de Santa Catarina, Centro Tecnológico, Programa de Pós-Graduação em Engenharia Mecânica, 2015. Citado 3 vezes nas páginas 27, 41 e 42.

TORGO, L. et al. Smote for regression. In: . [S.l.: s.n.], 2013. v. 8154, p. 378–389. Citado na página 36.

ZELL, A. *Simulation neuronaler netze*. [S.l.]: Addison-Wesley Bonn, 1994. v. 1. Citado na página 34.

Glossário

Nessa seção do documento são definidos alguns termos para maior entendimento do texto deste trabalho.

Regime permanente ou regime estacionário é definido quando a incógnita, em cada ponto do sistema, for independente do tempo. 26, *Incropera et al. (2007)*

Regime transiente é definido quando a incógnita varia de acordo com o tempo. *Incropera et al. (2007)*

Apêndices

APÊNDICE A – Gráficos da análise de dados

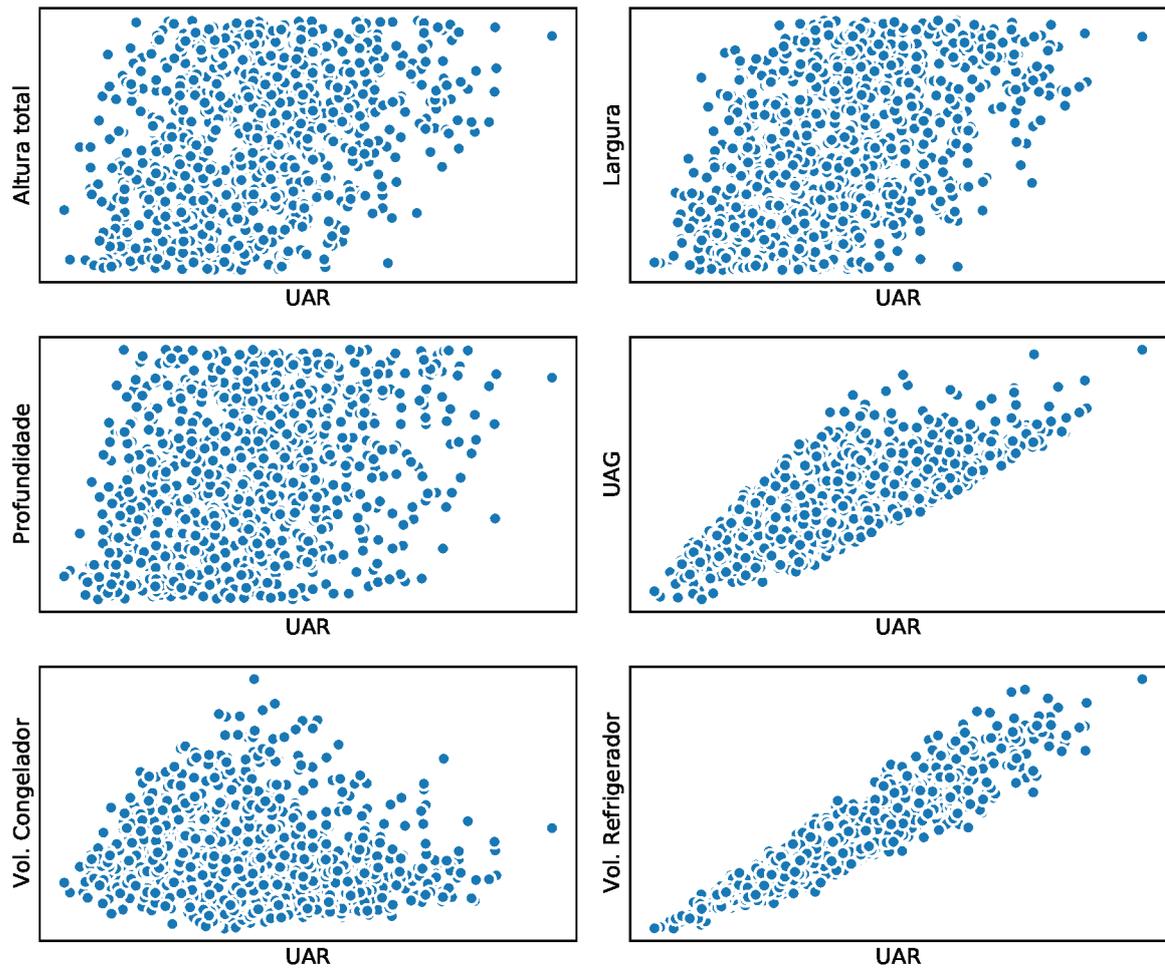


Figura 18 – Gráfico mostrando as geometrias pelo $U A_R$.

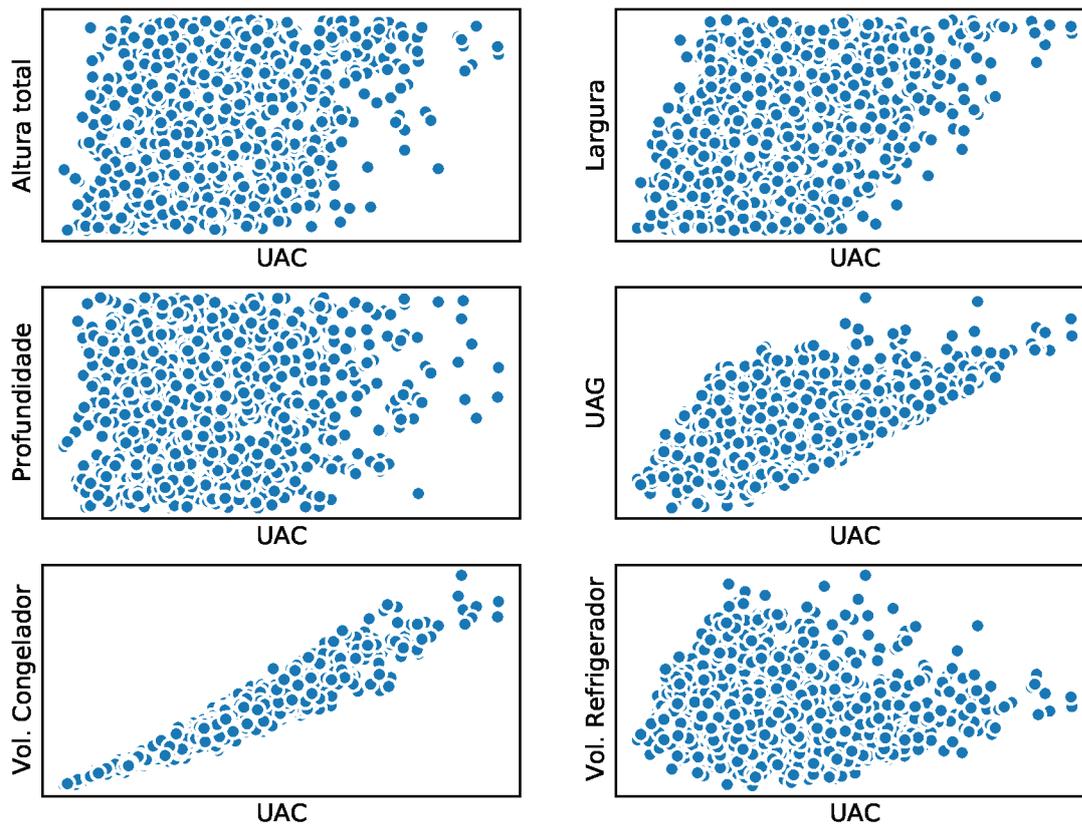


Figura 19 – Gráfico mostrando as geometrias pelo UA_C .

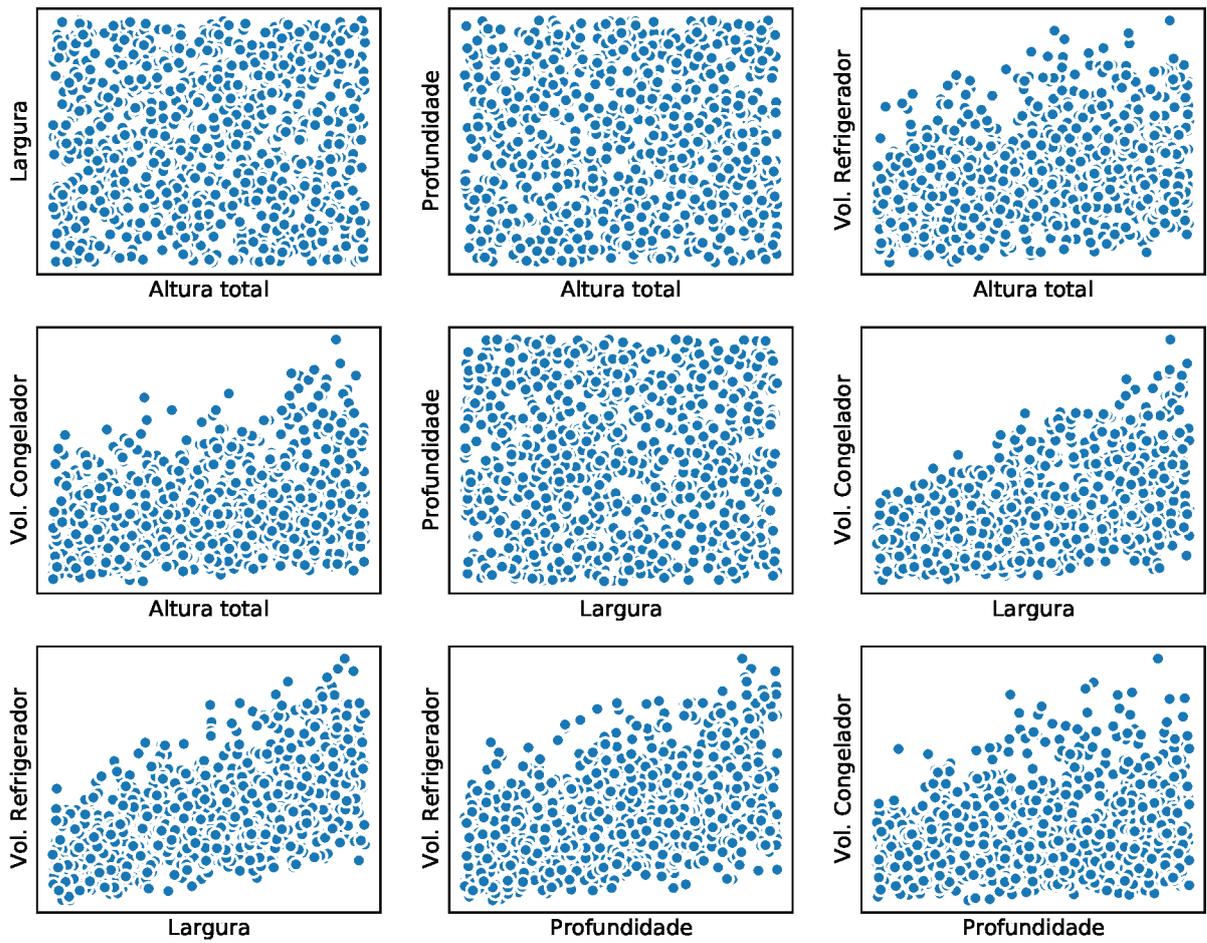


Figura 20 – Gráfico mostrando os parâmetros geométricos.

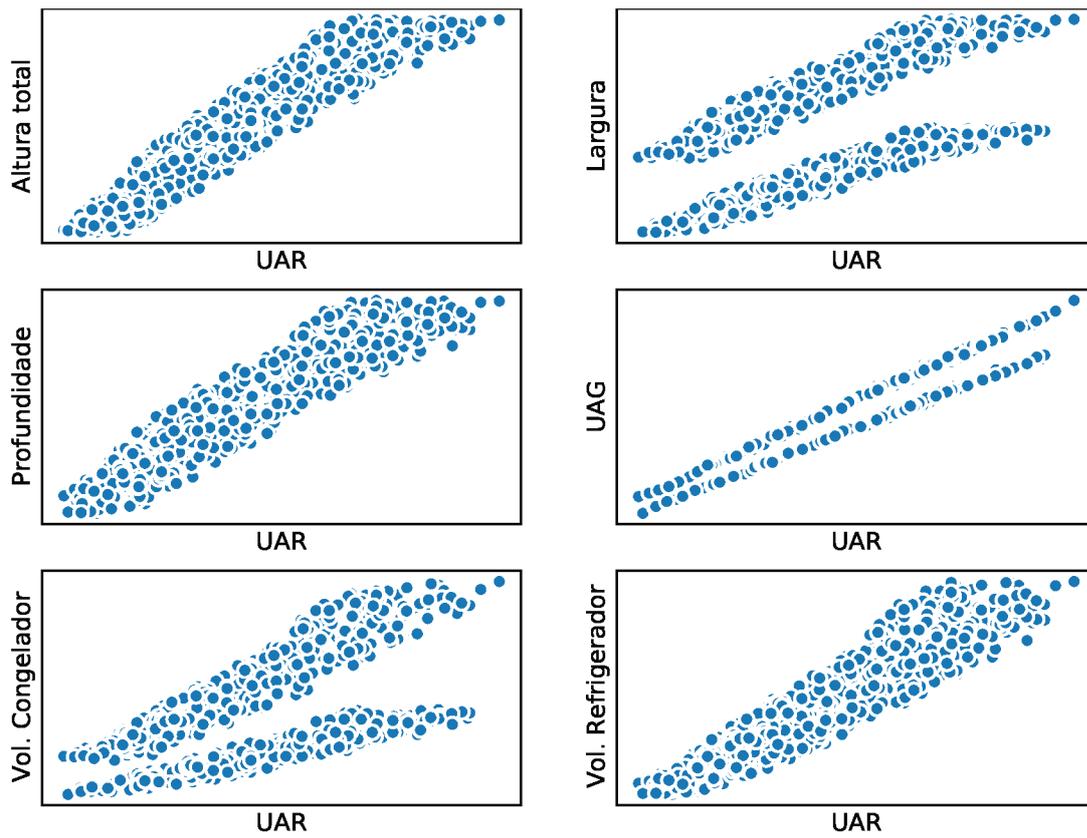


Figura 21 – Gráfico mostrando as geometrias pelo UA_R , gerado utilizando proporções.

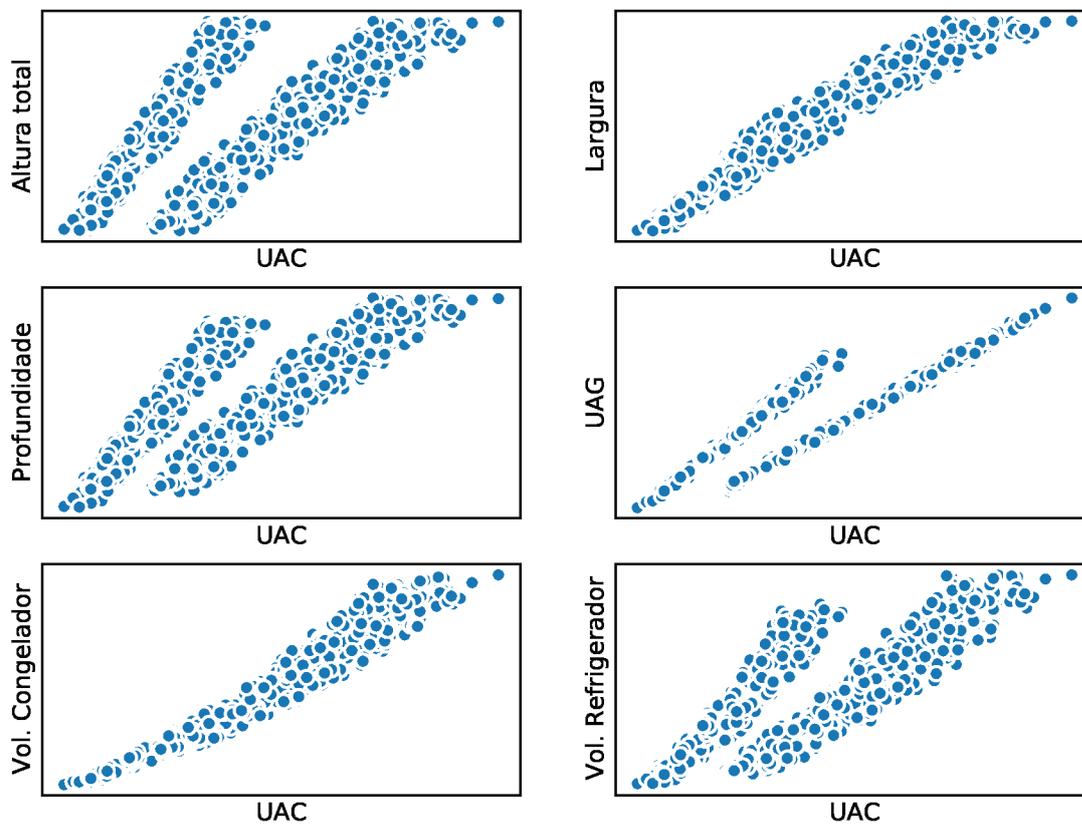


Figura 22 – Gráfico mostrando as geometrias pelo UA_C , gerado utilizando proporções.

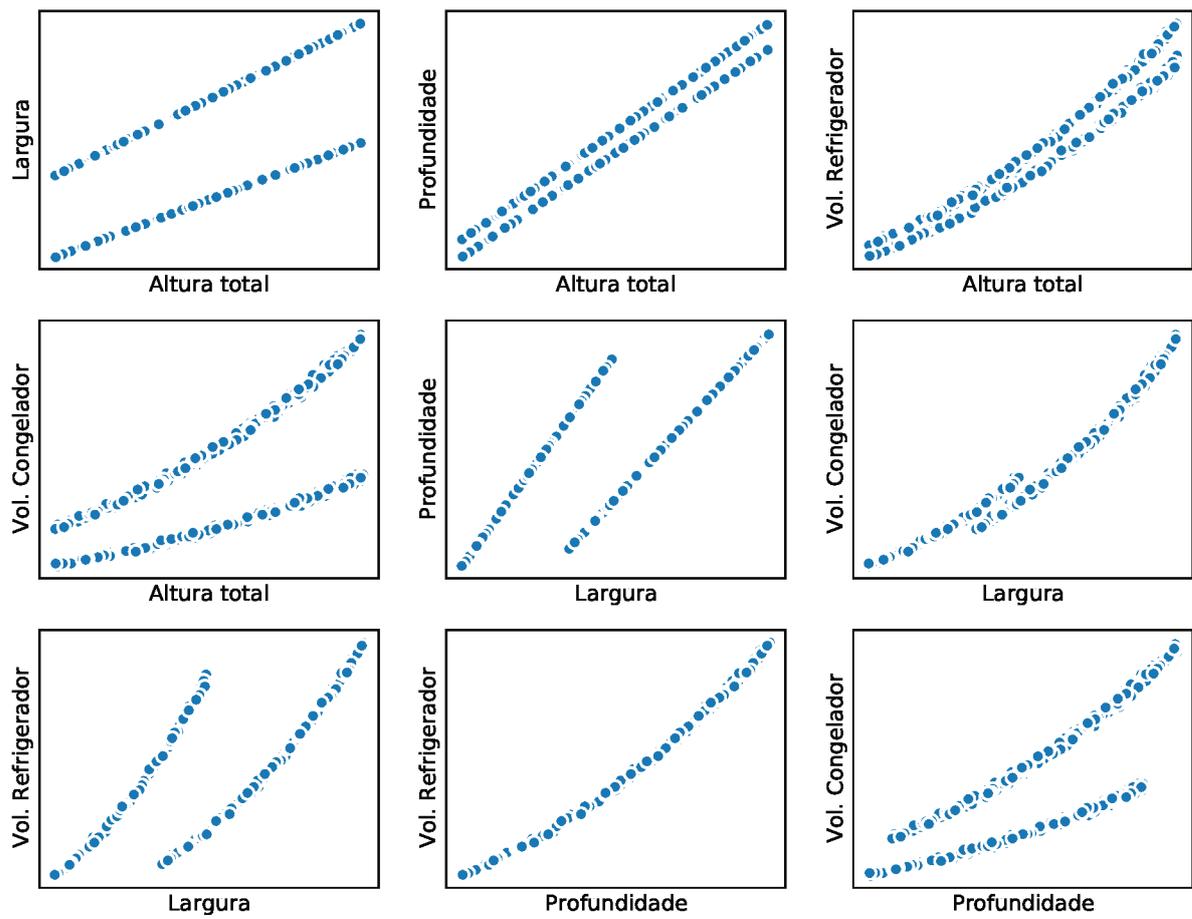


Figura 23 – Gráfico mostrando os parâmetros geométricos, gerado utilizando proporções.

APÊNDICE B – Amostra e métricas dos dados utilizados neste trabalho

ISR	Isolante superior do refrigerador
IPR	Isolante da porta do refrigerador
ITR	Isolante traseiro do refrigerador
IER	Isolante esquerdo do refrigerador
IDR	Isolante direito do refrigerador
ISC	Isolante superior do congelador
IPC	Isolante da porta do congelador
ITC	Isolante traseiro do congelador
IEC	Isolante esquerdo do congelador
IDC	Isolante direito do congelador
h	Altura
L	Largura
P	Profundidade
CIR	Capacidade interna do refrigerador
CIC	Capacidade interna do congelador
UAR	Coefficiente de condutância do refrigerador
UAC	Coefficiente de condutância do congelador
UAG	Coefficiente de condutância global

Tabela 9 – Dicionário de nomenclatura.

ISR	0.021	0.0015	0.0015	0.0015	0.0015
IPR	0.021	0.0015	0.0015	0.0015	0.0015
ITR	0.021	0.021	0.0015	0.0015	0.0015
IER	0.021	0.0015	0.0015	0.0015	0.0015
IDR	0.021	0.0015	0.0015	0.0015	0.0015
ISC	0.021	0.0015	0.0015	0.0015	0.0015
IPC	0.021	0.0015	0.021	0.021	0.0015
ITC	0.021	0.0015	0.0015	0.0015	0.0015
IEC	0.021	0.0015	0.0015	0.0015	0.0015
IDC	0.021	0.0015	0.0015	0.0015	0.0015
h	1800.0	1868.0	1868.0	1868.0	1868.0
L	670.0	710.0	710.0	710.0	710.0
P	800.0	741.0	741.0	741.0	741.0
CIR	320	329	329	329	329
CIC	100	108	108	108	108
UAR	1.334	1.08	1.03	1.04	1.0
UAC	0.752	0.51	0.52	0.43	0.48
UAG	2.086	1.59	1.55	1.47	1.48

Tabela 10 – Amostra dos dados experimentais usados. Cada coluna representa um gabinete.

Dado	Média	Desvio padrão	Mínimo	Máximo
h	1727.9	188.9	1400	2060
L	742.6	134.1	509	972
P	783.7	100.5	610	959
CIR	399.7	146.5	112	1028
CIC	241.9	113.0	32	730
UAC	1.26	0.38	0.44	2.96
UAR	1.87	0.47	0.78	3.81
UAG	3.13	0.57	1.72	5.26

Tabela 11 – Métricas sobre os dados uniformes.

Dado	Média	Desvio padrão	Mínimo	Máximo
h	1730.5	191.0	1400	2060
L	750.9	136.2	521	1022
P	783.8	87.9	622	951
CIR	495.5	185.2	201	931
CIC	164.1	94.1	37	417
UAC	0.96	0.32	0.41	1.93
UAR	2.20	0.50	1.24	3.58
UAG	3.16	0.76	1.66	5.51

Tabela 12 – Métricas sobre os dados proporcionais.

APÊNDICE C – Artigo deste trabalho em formato SBC

Início na próxima página.

Redes neurais para previsão do coeficiente de condutância global de refrigeradores domésticos

Bruno Gilmar Honnef¹

¹ Departamento de informática e estatística (INE)
Universidade Federal de Santa Catarina (UFSC)
Florianópolis, SC – Brazil

Abstract. *This work shows models of neural network that will try to predict the overall heat coefficient (UA) of domestic refrigerators, where experimental data is used to check the performance and pre-training the network with the synthetic extension of the dataset using sampling techniques and equations that govern the physical model. The methodology of this work is composed by generate the synthetic data, the network's generation and hyper-parameter adjustments, and the networks are trained using the default fit function or pre-trained using stacked auto-encoders. The techniques used in this works are useful when there is a small dataset, or the dataset is not diverse. It is shown that the networks had a good performance, therefore, showing that it is viable the use of neural networks to predict physical properties.*

Resumo. *Neste trabalho, são apresentados modelos de redes neurais com o objetivo de prever o coeficiente de condutância global (UA) de refrigeradores domésticos, no qual é utilizado dados experimentais para análise de desempenho e pré-treinamento da rede com a ampliação sintética do conjunto de dados utilizando técnicas de amostragem e equações que regem o modelo físico. O desenvolvimento deste trabalho consiste em gerar os dados sintéticos, geração e ajuste dos hiperparâmetros das redes neurais e, por fim, as redes são treinadas com a técnica padrão de ajuste ou pré-treinadas com Stacked auto-encoders. As técnicas apresentadas neste trabalho se tornam úteis quando é necessário lidar com conjuntos de dados pequenos ou com pouca diversidade. É visto que as redes tiveram um bom desempenho, mostrando ser viável o uso de redes neurais para previsão de propriedades físicas.*

1. Introdução

Refrigeradores são eletrodomésticos essenciais no cotidiano do ser humano. Existem cerca de 1.4 bilhões de refrigeradores domésticos em operação no mundo inteiro, os quais são responsáveis por cerca de 15% do consumo de energia no setor residencial [Barthel and Götz 2012]. Um refrigerador doméstico frost-free típico consome cerca de 45 kWh/mês, [Inmetro 2020], apesar do consumo unitário ser relativamente baixo, no Brasil o consumo de energia de refrigeradores domésticos representa 11% de toda a energia produzida [Eletrobrás 2004].

A maneira tradicional de avaliar o desempenho de um refrigerador consiste na realização de cálculos simplificados com base nas curvas características dos componentes, análises numéricas através de softwares comerciais de *CFD* (do inglês, *Computational Fluid Dynamics*) e de testes experimentais normalizados. Apesar de serem ferramentas

importantes para o dimensionamento de componentes, as curvas características e o CFD não fornecem informações acerca do comportamento sistêmico do refrigerador, que então são obtidas através de testes em câmaras climatizadas com temperatura e umidade do ar controladas. Tais testes são custosos, além da necessidade de um protótipo, por exemplo, um teste de abaixamento de temperatura consome mais de 24 horas se considerar apenas o tempo necessário para estabilização do experimento e para efetuar as medições [Hermes 2006].

O presente trabalho busca diminuir o custo com estes experimentos, mais especificamente com o teste de fluxo de calor reverso, o qual tem como objetivo avaliar a qualidade do isolamento térmico das paredes através da medição das condutâncias térmicas globais dos compartimentos, comumente denominadas UA . São analisados modelos de redes neurais tendo como entrada a geometria do refrigerador e o isolante de cada uma de suas paredes, os modelos de redes propostos se ajustam até com poucos dados. Tais dados podem ser experimentais ou gerados sinteticamente através de equações que regem tal modelo físico ou utilizando técnicas de amostragem do conjunto experimental. O treinamento da rede é feita de duas formas, a primeira utilizando de uma SAE (do inglês, *Stacked autoencoder*) para pré-treinar a rede e a segunda utilizando a técnica padrão de ajuste da rede.

1.1. Coeficiente de condutância global

Um dos parâmetros de grande influência no desempenho de um refrigerador é a resistência térmica do gabinete, o qual depende, principalmente, da condutividade do isolamento térmico do gabinete, e pode ser expressa na forma de um coeficiente global de transferência de calor U com unidade $[W/m^2K]$ [Incropera et al. 2007, Colombo 2014].

O coeficiente U é uma medida da capacidade global de transferir calor em uma série de barreiras condutivas e convectivas. Para o caso de um trocador de calor, podemos calcular U usando uma expressão análoga à lei do resfriamento de Newton,

$$q = -UA\Delta T \quad (1)$$

onde o vetor q é a taxa de transferência de calor $[W]$ e o sinal indica a direção do calor (do corpo quente para o frio), A é a área de troca de calor $[m^2]$, e ΔT é a diferença de temperaturas entre os lados da parede.

Quando o coeficiente global de transferência de calor está associado a uma área A de troca de calor então é designado o termo condutância global de calor UA , que é a multiplicação do termo U pela área de troca A resultando na unidade $[W/K]$. Para obter este coeficiente para um dado gabinete refrigerado é necessário um procedimento experimental denominado teste de fluxo de calor reverso [Incropera et al. 2007, Colombo 2014].

Logo, a condutância térmica para cada parede pode ser calculada utilizando a seguinte equação:

$$UA_i = \frac{k_{iso}A_i}{l_i} \quad (2)$$

Onde k_{iso} é a condutividade do isolamento da parede com menor grandeza, l é a espessura dessa parede, o índice i representa cada uma das paredes do respectivo compartimento (congelador e refrigerador). O valor do UA dos compartimentos é o somatório de cada uma das paredes do compartimento e os coeficientes do congelador UA_C e do refrigerador UA_R estariam determinados. Para determinar o coeficiente de condutância global, UA_G ou simplesmente UA , de todo o gabinete basta somar os coeficientes de cada compartimento.

1.2. Teste do fluxo de calor reverso

O principal objetivo deste teste é a avaliação da qualidade do isolamento térmico das paredes através da medição das condutâncias térmicas globais, normalmente chamadas de UA . [Junior and de Sousa 2008, Sim and Ha 2011].

Esse teste consiste em instalar resistências elétricas, com potência inferior a 50 W, controladas por uma fonte variável de tensão. O objetivo é aquecer o interior do gabinete criando um fluxo de calor do interior para o exterior, originando o nome do experimento. A câmara de testes é mantida a temperatura e umidade constantes e o sistema de refrigeração é desligado e as resistências elétricas dissipam calor até que se atinja uma diferença de temperatura de 25°C entre os ambientes em regime permanente. As resistências elétricas são distribuídas no interior de gabinete de maneira que a estratificação térmica seja minimizada. O monitoramento do teste limita-se à potência dissipada pelas resistências elétricas e as temperaturas internas e externas como mostrado na Figura 1. [Seidel 2001].

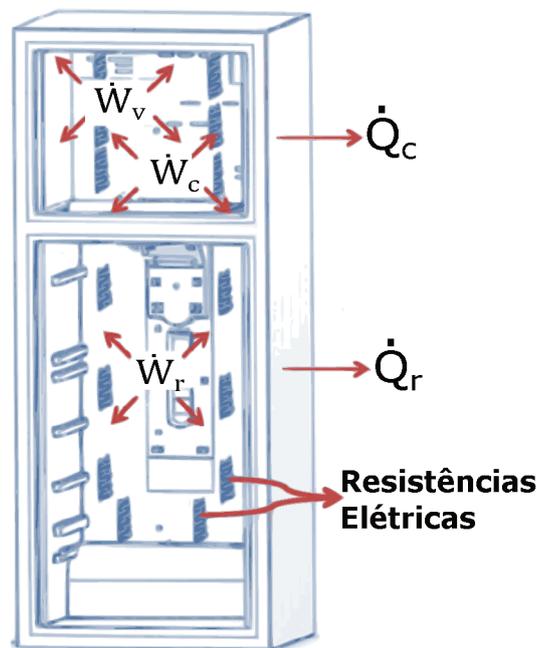


Figura 1. Teste de fluxo de calor reverso. Adaptado de [Thiessen 2015].

Em regime permanente, todo calor gerado pelas resistências elétricas flui através das paredes do refrigerador e, utilizando a Equação 1, chega-se à seguinte equação para determinação do parâmetro UA :

$$UA = \frac{q_r}{(T_e - T_i)} \quad (3)$$

na qual UA é o produto entre o coeficiente global de transferência de calor e a área média total do gabinete, q_r é a potência dissipada pelas resistências elétricas, T_e é a média das temperaturas externas e T_i é a média das temperaturas internas.

1.3. Redes neurais

Redes neurais são compostas de neurônios interconectados que tentam imitar o que acontece no cérebro humano. Cada neurônio, como mostrado na Figura 2, possui conexões de entrada com um peso $W_{i,j}$ associado, conexão de entrada i do neurônio j , o qual determina a força da conexão. Junto com as entradas tem uma especial chamada *bias* que é usado para ajustar a saída e somado junto na função de entrada. Conexões de saída, que podem ser as entradas de outros neurônios ou o resultado final, uma função de entrada (4) que é o somatório das entradas a_i multiplicada por seu respectivo peso $W_{i,j}$. O neurônio também possui uma função de ativação g e uma saída a_j que é basicamente o resultado da 4 aplicada na função g , ficando $a_j = g(in_j)$ [Russell and Norvig 2016].

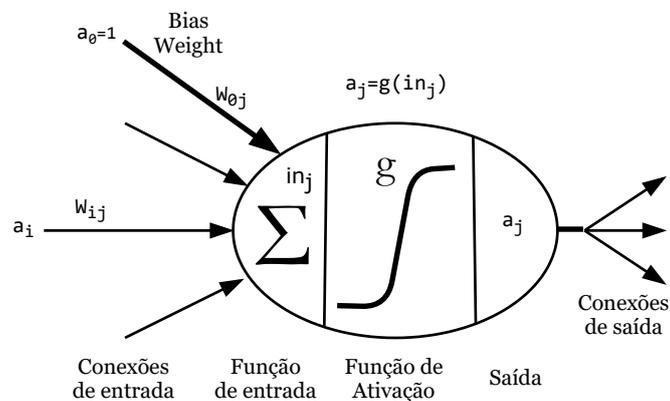


Figura 2. Modelo matemático simples de um neurônio artificial. Adaptado de [Russell and Norvig 2016].

$$in_j = \sum_{i=0}^n W_{i,j} a_i \quad (4)$$

1.3.1. Stacked autoencoders

Um *autoencoder* é um tipo de rede neural que visa reproduzir as entradas em suas saídas, de forma não supervisionada, ou seja, o número de neurônios na entrada é o mesmo que na saída. Esse tipo de rede é bastante utilizado para redução de dimensões, pois a rede irá criar uma cópia da entrada “compactada”, ou seja, irá ignorar traços não importantes dos dados, como é o caso de ruídos. [Géron 2017].

O pré-treinamento de uma rede neural com *SAE* ocorre que para cada camada oculta é criado um *autoencoder*, o número de neurônios de cada *autoencoder* é o mesmo que o número de neurônios da camada correspondente da rede, é apresentado um esquemático na Figura 3. Tomamos como exemplo uma rede com 4 camadas ocultas 13-(7-5-4-3)-1¹, então são criados 5 *autoencoders* com as seguintes estruturas: 21-(7)-21,

¹Os valores entre parênteses representam o número de neurônios nas camadas ocultas.

7-(5)-7, 5-(4)-5, 4-(3)-4 e 3-(1)-3. O primeiro *autoencoder* pega a entrada da rede neural como sua entrada e saída e então a saída da sua camada oculta se torna a entrada do próximo *autoencoder*, em outras palavras, a saída da camada oculta do *autoencoder* é a entrada do próximo *autoencoder*.

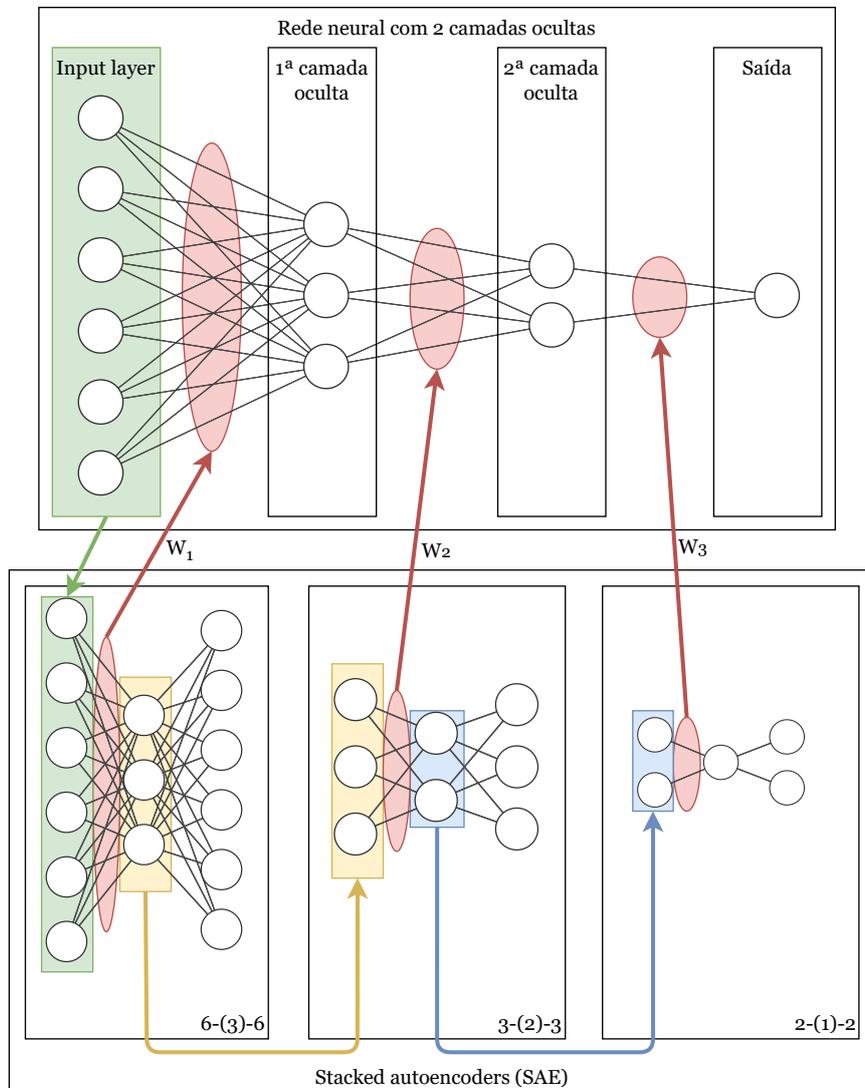


Figura 3. Esquemático do processo de inicialização camada por camada da rede neural utilizando autoencoders.

1.4. Ajuste de hiperparâmetros

Hiperparâmetros são propriedades que regem todo o processo de aprendizagem da rede o qual inclui variáveis que determinam a estrutura da rede (como é o caso do número de camadas ocultas) e variáveis que interferem na qualidade e velocidade do aprendizado (como a taxa de aprendizagem), tais variáveis são definidas antes de iniciar o treinamento e algumas, como é o caso das estruturais, não podem ser alteradas posteriormente. [Géron 2017, Leonel 2019].

O ajuste de hiperparâmetros, ou otimização dos hiperparâmetros, tem como objetivo escolher o melhor conjunto de parâmetros para o algoritmo de aprendizado para

então minimizar o valor da função de perda no conjunto de dados. A otimização *Bayesiana* constrói um modelo probabilístico da função mapeando valores de hiperparâmetros a serem buscados para o objetivo avaliado no conjunto de validação. O algoritmo, iterativamente, avalia as configurações de parâmetros mais promissoras baseadas nas buscas anteriores. [Snoek et al. 2012, Claesen and De Moor 2015].

O ajuste foi realizado utilizando a técnica de otimização *bayesiana* (do inglês, *Bayesian Optimization*) da biblioteca *Keras Tuner*, [Chollet et al. 2015], com o erro médio quadrado como função de perda e as métricas observadas são tanto o erro médio quadrático (*MSE*) quanto o erro médio absoluto (*MAE*).

Na Tabela 1, são apresentados os valores que são buscados pelo algoritmo, tais valores são definidos manualmente pelo usuário, os valores são, se necessário, reajustados para uma nova busca, uma vez que esta etapa é bastante empírica e computacionalmente custosa, por exemplo, o usuário pode ver que a rede não obteve um bom desempenho com cinco camadas, então esse valor é reajustado e submetido ao algoritmo novamente.

hiperparâmetro	Valores de busca
Otimizadores	Adadelta, Adagrad, Nadam, SGD, Ftrl, RMSprop, Adam, Adamax
Funções de ativação	Tanh, sigmoid, exponential, relu, elu, hard sigmoid, linear, selu, softmax, softplus, softsign
Camadas ocultas	de 2 a 10
Neurônios nas camadas ocultas	de 3 a 24

Tabela 1. Valores para a busca de hiperparâmetros.

A Tabela 2 apresenta os hiperparâmetros utilizados para cada rede, os hiperparâmetros que foram submetidos à busca são os otimizadores, todas as funções de ativação e o *layout* da camada oculta. Os valores na linha “Camada oculta” representam os *layouts* das camadas ocultas, por exemplo, [6,5,4,3] indica que são 4 camadas ocultas (tamanho da lista) e a primeira tem seis neurônios, a segunda cinco, e assim por diante, sendo cada camada do tipo densa, ou seja, todos os neurônios da primeira camada são ligados com os da segunda e todos da segunda camada ligados com da terceira e assim por diante.

1.5. Técnicas de amostragem

Técnicas de amostragem são algoritmos que geram novas amostras baseadas em amostras já existentes. *Synthetic Minority Over-Sampling Technique for Regression with Gaussian Noise* ou seu acrônimo *SMOGRN* é um algoritmo que tenta resolver problemas de regressão não balanceados em que os casos mais importantes não são bem representados no conjunto de dados. O algoritmo combina dois outros algoritmos de sub-amostragem aleatórios que é o *SMOTER* e a introdução do *ruído Gaussiano*. A ideia chave do *SMOGRN* é combinar as duas estratégias para geração de amostras sintéticas com o objetivo simultâneo de limitar o risco que o *SMOTER* pode incluir usando da estratégia conservadora do *ruído Gaussiano* [Branco et al. 2017].

hiperparâmetro	Rede dual	Rede Dual 1 config	Rede congelador	Rede refrigerador
Épocas	1000	1000	1000	1000
Taxa de aprendizado	0.05	0.001	0.001	0.01
Otimizador	SGD	Adagrad	Adamax	RMSprop
Função de ativação da camada de entrada	Relu	Tanh	Softsign	Softplus
Função de ativação das camadas ocultas	Relu	Tanh	Relu	Selu
Função de ativação da camada de saída	Exponential	Tanh	Selu	Elu
Neurônios na camada de entrada	15	15	9	9
Camada oculta	[6, 21, 24, 15]	[6, 5, 4, 3]	[12,12,12,3,3,3]	[12,18,3,3,3,3,3,3,3]
Neurônios na camada de saída	2	2	1	1

Tabela 2. hiperparâmetros ajustados da rede.

Para ampliar o número de dados experimentais deste trabalho, mais aprofundado na Seção 2.2.3, foi utilizado o algoritmo *SMOTER*, que é a versão do algoritmo *SMOTE* para regressões. O algoritmo *SMOTE* é um método de amostragem que trata de problemas de classificação com a distribuição de classes não balanceadas. Esse método combina sub-amostragem das classes mais frequentes com sobre-amostragem das classes menos frequentes [Torgo et al. 2013].

2. Desenvolvimento

O desenvolvimento deste trabalho consiste em examinar e categorizar a distribuição de dados, normalizar a entrada de dados, propor modelos de redes neurais para prever os dados e, por fim, analisar o desempenho dos modelos gerados.

Foi utilizada uma rede neural para prevermos o coeficiente de condutância global (*UA*) de refrigeradores domésticos usando dados sintéticos para treiná-la, a utilização de tais dados é porque não há muitos dados experimentais disponíveis para treinamento, enquanto que, mesmo perdendo acurácia, os dados sintéticos podem ser gerados em grandes quantidades.

2.1. Metodologia

A metodologia adotada para o desenvolvimento deste trabalho pode ser dividida em quatro etapas, como ilustra o diagrama da Figura 4:

- a) Primeiro são gerados os dados sintéticos e pseudo-experimentais que, como será discutido na Seção 2.2.3, os dados experimentais obtidos não são suficientes para um treinamento, então tais dados foram submetidos à uma técnica de amostragem e então é gerado em torno de 200 amostras sintéticas para simular como se fossem os experimentais.

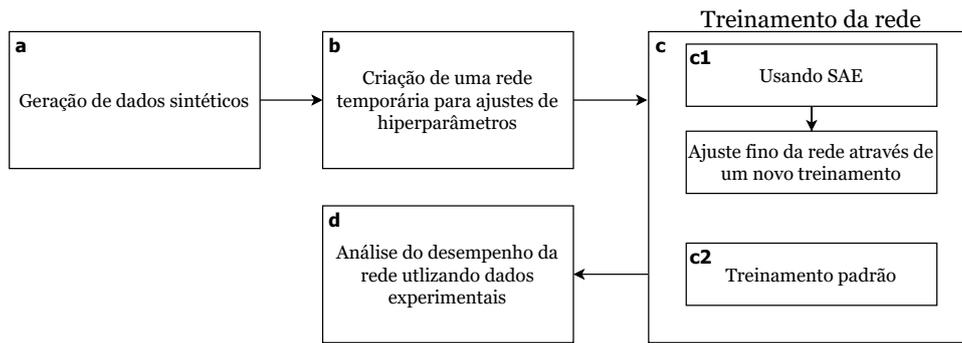


Figura 4. Diagrama da metodologia adotada.

- b) É criada uma rede preliminar em que será submetida às técnicas de ajuste de hiperparâmetros utilizando os dados sintéticos gerados anteriormente. Esta etapa indica o número de neurônios, camadas ocultas, funções de ativação, taxa de aprendizagem, etc.
- c) O treinamento da rede é dividido em dois:
1. É criado um *Stacked autoencoder (SAE)* para inicializar os pesos e *biases* da rede neural para que fiquem próximos da solução global e, depois ela é submetida à um treinamento padrão utilizando os dados pseudo-experimentais. A Seção 1.3.1 discute como é feito o pré-treinamento.
 2. A rede é treinada apenas com dados sintéticos de maneira padrão (sem pré-treinamento).
- d) Os dados experimentais são utilizados para verificar a acurácia da rede e, por fim, são extraídas as conclusões do projeto.

2.2. Dados

Para o problema desenvolvido neste trabalho são utilizados dois grupos de dados, os dados sintéticos, que são gerados a partir de teses, dissertações e produtos encontrados no mercado, e os dados experimentais, que são obtidos a partir dos experimentos documentados em teses e dissertações.

2.2.1. Geração de dados sintéticos

São duas formas de geração de dados. A primeira (referido no texto como *uniforme*) é utilizando distribuições uniformes para a altura, largura e profundidade. Tais valores são retirados de web sites de compras que indicam os parâmetros geométricos mínimos e máximos. As espessuras também são escolhidas através de uma distribuição uniforme, porém os intervalos de mínimo e máximo são retirados de trabalhos acadêmicos. Cada parede possui um intervalo; as paredes da esquerda e direita são iguais; a traseira normalmente é mais espessa e a da porta geralmente é a menos espessa. As paredes do congelador são mais espessas do que as do refrigerador. A capacidade de um recipiente com paredes de espessura l pode ser calculada utilizando a seguinte equação,

$$V = (h - 2l) * (L - 2l) * (p - 2l) \quad (5)$$

que para o caso de um compartimento i (congelador ou refrigerador) podemos escrever como,

$$V_i = \frac{(h - l_{si} + l_{Ii}) * (L - 2l_{ei}) * (p - l_{pi} + l_{ti})}{10^6} \quad (6)$$

onde os subíndices s, I, e, p, t representam, respectivamente, parede superior, inferior, esquerda (a direita tem a mesma espessura), porta e traseira. Divide-se por 10^6 para obter o volume V_i em litros.

A capacidade de um gabinete é dividida em dois, um para o refrigerador e um para o congelador, na Figura 9 é ilustrado um gabinete *top-down*, há uma linha dividindo os compartimentos que é chamada de travessa, entra no cálculo de ambos como espessura superior (ou inferior no caso do congelador). Cada círculo representa um parâmetro l da Equação 6, e os outros parâmetros tendo a mesma denominação, com exceção da altura, a qual é gerada uma altura para cada compartimento e depois ambas são somadas junto com as espessuras superior, inferior e da travessa para obter a altura total do refrigerador.

Após gerar os dados base, que são a altura, largura, profundidade e a espessura de cada parede, é então escolhida uma configuração de isolantes, como é ilustrado na Figura 5, na qual cada cor indica um isolante diferente, neste caso o máximo de isolantes diferentes é 2. Então é escolhido, aleatoriamente, um isolante da Tabela 3, a qual enumera os isolantes mais comuns em refrigeradores domésticos. Por fim, utilizando a Equação 2, é calculado o UA_C e o UA_R que servirão como *rótulo* (do inglês, *label*) para o treinamento da rede.

Isolante	Condutividade [W/mK]
PU	0.028
VIP com núcleo de fibra de vidro	0.0015

Tabela 3. Resistências térmicas dos materiais isolantes utilizados. Adaptado de [Thiessen 2015].

A segunda forma (referido no texto como *proporcional*) de geração de dados é utilizando proporção o qual divide a altura total pela altura de cada compartimento, a altura total também é dividida pela largura e profundidade, já as espessuras e os isolantes são geradas da mesma forma descrita anteriormente e após a geração das espessuras, então é calculado os volumes de cada compartimento e seus respectivos UAs .

2.2.2. Qualidade dos dados

Os gráficos exibidos na Figura 6 mostram os dados gerados através de proporção, nota-se claramente uma divisão linear entre os dados, o motivo é porque foram utilizados dois refrigeradores de proporções diferentes, o número de divisões aumentaria com diferentes proporções. Já nos gráficos da Figura 7 mostram dados mais esparsos por conta que não possuem restrição quanto à sua geração.

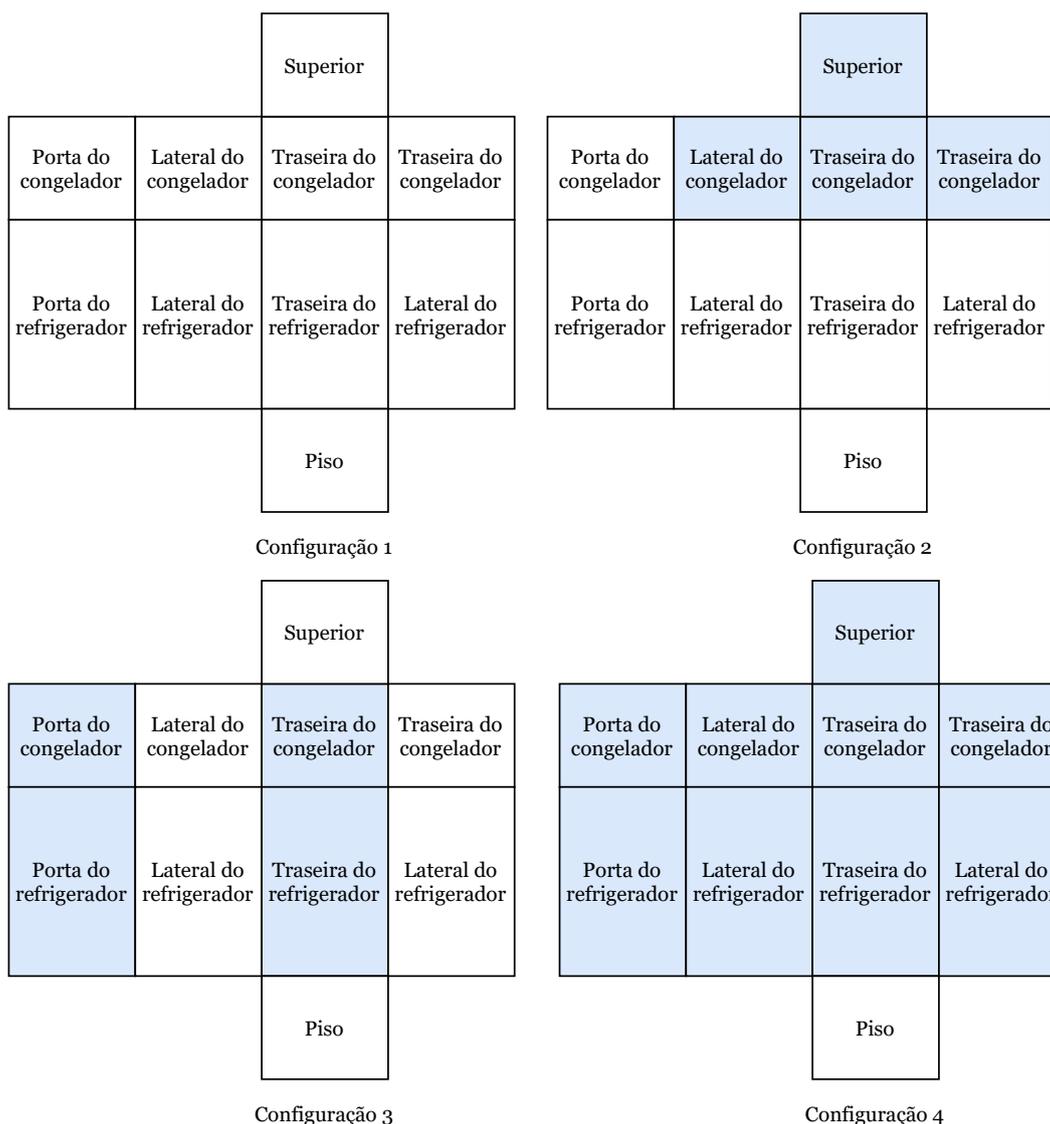


Figura 5. Configurações de isolantes de cada parede do gabinete do refrigerador. Adaptado de [Thiessen 2015].

Em uma análise preliminar visual dos dados e resultados, mais aprofundados na Seção 2.4, é possível presumir que o treinamento com os dados de tendência linear (como mostrado na Figura 6) possuem melhores previsões quando as entradas² estão dentro ou próximas desta tendência, enquanto que quando treinada com os dados mais esparsos (mostrado na Figura 7) a rede mostra mais generalização, isso significa que a maioria das entradas têm uma previsão relativamente boa, porém, para entradas que estão dentro das tendências lineares, o conjunto proporcional de treino é melhor.

²Que nunca foram vistas pela rede antes.

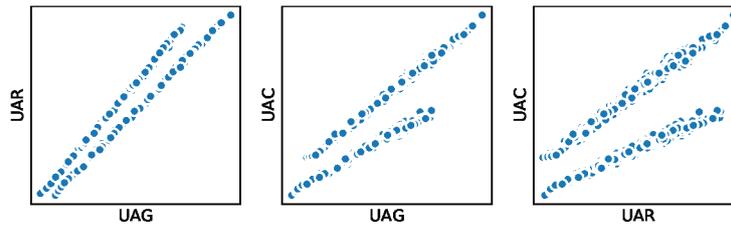


Figura 6. Gráfico mostrando as saídas UA_R , UA_C e UA_G , gerado utilizando proporções.

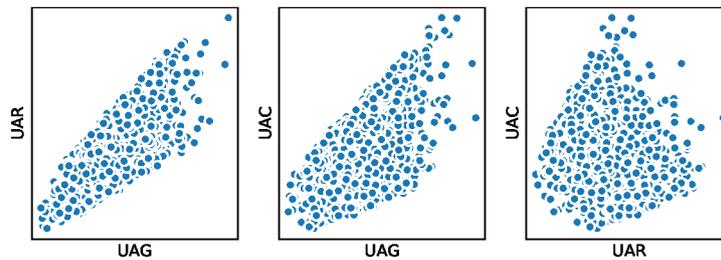


Figura 7. Gráfico mostrando as saídas UA_R , UA_C e UA_G , caso uniforme.

2.2.3. Dados experimentais

Os dados experimentais foram retirados de dissertações, artigos e teses. Infelizmente não há uma base de dados de fácil acesso que contenha tais informações e, por conta da dificuldade de obtenção, não há muitos dados disponíveis para este trabalho.

Porém os poucos dados obtidos são submetidos a uma técnica de amostragem *SMOBN*, apresentado na Seção 1.5. Apesar de poder criar quantas amostras desejar, é necessário levar em conta que não haverá muita variabilidade neles. Por tanto foram utilizados para verificar a acurácia da rede.

A Figura 8 mostra os dados experimentais originais, em azul, e os gerados através da técnica, em laranja. É possível notar que os dados gerados são próximos dos originais, o desvio padrão é baixo (em torno de 0.18 para a média de 1.73 do UA_G), porém por maior que seja o número de dados gerados, eles sempre estarão próximos dos já existentes, logo não é possível utilizá-los para treinamento de uma rede por conta da pouca variabilidade.

2.3. Modelo

Há dois modelos de rede, construídos utilizando a biblioteca *tensorflow*, [Abadi et al. 2015], da Google. A primeira rede é chamada *rede dual* porque ela tem ambos isolantes, refrigerador e congelador, como entradas e os UA como saída. A outra rede, chamada de *rede única*, possui apenas as entradas e saídas respectivas de cada compartimento. No quesito de treinamento, há duas técnicas adotadas, a primeira sendo a implementação de *SAEs* para pré-treinar e depois ajustar com dados pseudo-experimentais e a segunda é a técnica de ajuste (do inglês, *fit*) padrão de uma rede neural utilizando apenas dados sintéticos, nas Seção 2.3.2 é detalhado o processo de treinamento de cada rede.

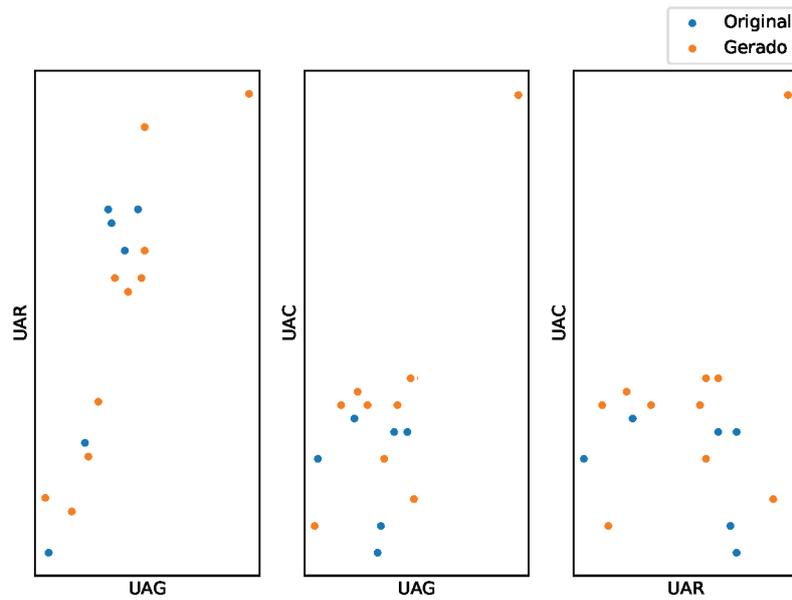


Figura 8. Gráfico mostrando as saídas UA_R , UA_C e UA_G . Experimentais e pseudo-experimentais.

2.3.1. Entradas e saídas

Os dados de entrada para a rede são, como ilustra a Figura 9, a geometria do refrigerador que inclui altura (h) [mm], largura (L) [mm], profundidade (p) [mm], capacidade do refrigerador e congelador [litros] e os isolantes [W/mK] de cada parede representados pelos círculos na imagem (são 10), totalizando 15 entradas. A rede gera duas saídas que são os coeficientes UA_C e UA_R que então são somados para obter o UA_G .

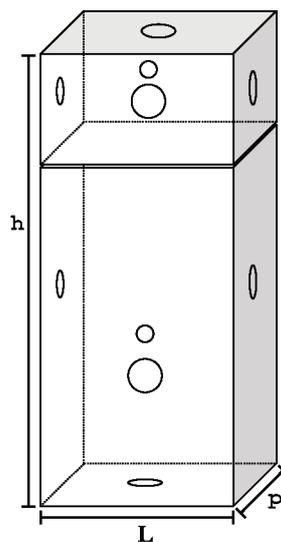


Figura 9. Ilustração dos dados de entrada da rede.

2.3.2. Treinamento

As redes foram submetidas a dois tipos de treinamento, o primeiro pré-treinando utilizando *SAE* com um conjunto sintético de dados depois fazendo um ajuste fino utilizando o conjunto pseudo-experimental, e o segundo utilizando a função de ajuste padrão da rede utilizando apenas dados sintéticos. Os parâmetros utilizados para cada treinamento podem ser vistos na Tabela 4. Na coluna *SAE* há dois valores: o primeiro sendo para o treinamento do *SAE* e o segundo é para o ajuste fino. A Tabela 5 da Seção 2.4 é mostrada a acurácia da rede juntamente com discussões acerca dos resultados obtidos.

Parâmetro	SAE	Padrão
Épocas	1000/100	1000
Número de Amostras conjunto uniforme	10 000/200	10 000
Número de Amostras conjunto proporcional	10 000/200	10 000

Tabela 4. Parâmetros de cada treinamento.

2.4. Resultados

Na Tabela 5 é mostrada a rede, com a técnica usada e o conjunto de dados o qual foi submetido ao treino. A acurácia é determinada pela função de erro médio quadrático (*MSE*). Quanto menor esse valor maior é a precisão da rede, para ser considerado um desempenho bom é necessário que o erro, sobre ambos conjuntos, seja menor que 0.2. Esse valor de erro é dependente do conjunto de dados, por exemplo, se o conjunto é bastante esparsa então a tendência é do erro ser maior do que em um conjunto normal. No caso desta rede em específico é definido, através de observações empíricas sobre os resultados das diferentes redes e comparando com os dados experimentais, que, para a predição ser relativamente boa, a rede deva atingir um erro menor que 0.2.

Nota-se que a rede dual teve o pior desempenho entre todas, enquanto que as redes de saída única possuem os melhores resultados, com exceção da rede única que prevê para o refrigerador treinada com o conjunto proporcional, talvez com hiper-parâmetros diferentes a acurácia seria maior. De maneira geral, as que obtiveram os melhores resultados são aquelas pré-treinadas usando o *SAE*.

Esperava-se que as redes pré-treinadas teriam uma acurácia maior sobre o conjunto experimental por conta do seu segundo treinamento usando os dados pseudo-experimentais, mesmo usando poucas épocas no treinamento. Também era esperado que o conjunto proporcional, de maneira geral, tivesse uma acurácia maior do que o uniforme, pelo menos sobre o conjunto pseudo-experimental.

As Figuras 10, 11, 12, 14 e 15 apresentam um gráfico dos *valores esperados x valores previstos* com os resultados do treinamento padrão, em azul, e usando *SAE*, em laranja, a linha vermelha traça os valores ideais, ou seja, quanto mais próximo da linha melhor é a previsão que a rede fez, em outras palavras, menor é o erro. É esperado que o gráfico apresente os valores de forma linear, seguindo a tendência da linha vermelha (mesmo que estejam um pouco fora do ideal). Quando os valores estão muito dispersos,

Rede	Técnica de treinamento	Conjunto de dados	Acurácia Conjunto de testes	Acurácia conjunto Pseudo-Experimentais
Rede dual	SAE	Uniforme	0.4165	0.623
		Proporcional	0.5509	0.627
	Padrão	Uniforme	0.836	0.405
		Proporcional	1.830	0.786
Rede única Congelador	SAE	Uniforme	0.028	0.602
		Proporcional	0.048	0.316
	Padrão	Uniforme	0.0067	0.373
		Proporcional	0.0022	0.279
Rede única Refrigerador	SAE	Uniforme	0.014	0.497
		Proporcional	0.606	0.109
	Padrão	Uniforme	0.006	0.540
		Proporcional	0.156	0.915

Tabela 5. Acurácia das redes.

Rede	Técnica Treinamento	Conjunto de dados	Acurácia Conjunto de testes	Acurácia Conjunto pseudo-experimental
Rede dual	Padrão	Uniforme	0.014	0.168
		Proporcional	0.018	0.157

Tabela 6. Acurácia das redes com apenas uma configuração de refrigerador.

sem tendências, isso quer dizer que a rede não conseguiu se adequar, ou se ajustar, com os dados.

Na Figura 10, é apresentado os resultados da previsão, sendo que na primeira coluna é exibido os valores previstos para UA_R e na segunda os valores para UA_C . Note que, mesmo que o erro esteja grande, os valores previstos para o UA_C utilizando *SAE* com o conjunto uniforme são melhores que o padrão. Já na Figura 11 é mostrado os valores previstos sobre o conjunto pseudo-experimental, os valores obtidos não são satisfatórios, pois estão totalmente dispersos.

Infelizmente, não foi possível chegar a um resultado satisfatório com a *rede dual* utilizando as quatro configurações de refrigeradores simultaneamente, porém se utilizar apenas uma configuração por vez é possível obter resultados melhores, como é mostrada na Tabela 6 e as previsões feitas na Figura 12, note como os dados estão seguindo uma tendência linear e estão poucos dispersos. O treinamento foi realizado utilizando apenas a configuração 1 da Figura 5 e com os parâmetros mostrados na Seção 2.3.2.

Nas Figuras 14 e 15 são apresentados os resultados das previsões sobre os conjuntos de dados do refrigerador e congelador, respectivamente. As previsões da primeira coluna são das redes treinadas no conjunto uniforme enquanto que na segunda coluna é treinado utilizando o conjunto proporcional. Note que o conjunto uniforme possui mais

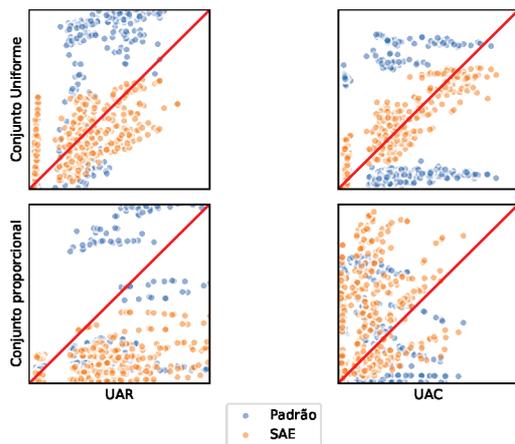


Figura 10. Resultados da previsão da rede dual.

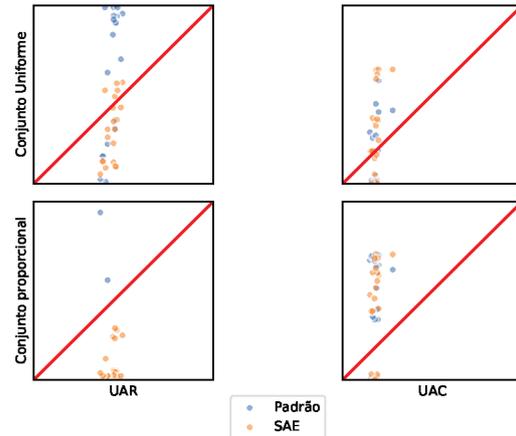


Figura 11. Resultados da previsão dos dados pseudo-experimentais da rede dual.

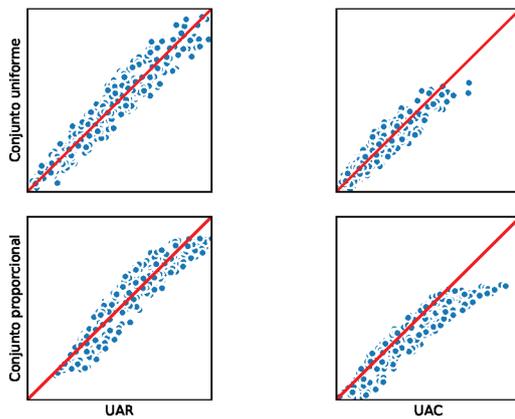


Figura 12. Resultados da previsão da rede dual com uma configuração de isolantes.

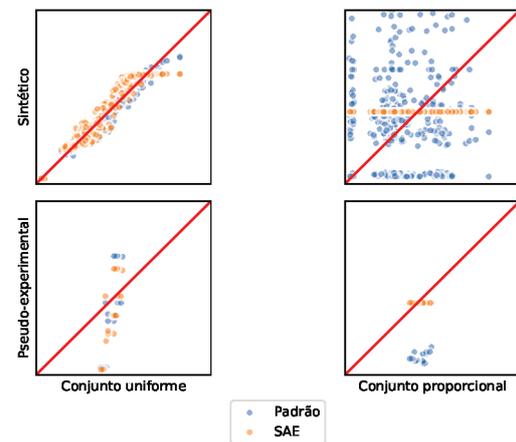


Figura 13. Resultados da previsão da rede única para o refrigerador.

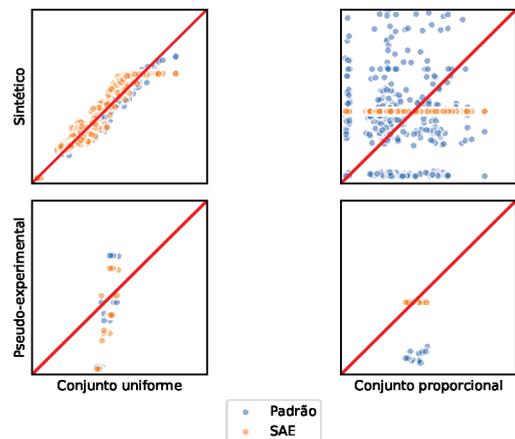


Figura 14. Resultados da previsão da rede única para o refrigerador.

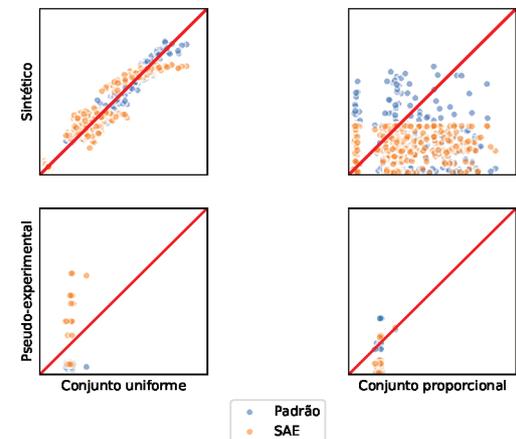


Figura 15. Resultados da previsão da rede única para o congelador.

generalização, logo a previsão sobre esse conjunto se torna boa, como é mostrada no primeiro gráfico de cada rede, já quando a mesma rede é treinada sobre um conjunto proporcional ela perde um pouco da generalização, porém a previsão sobre o conjunto pseudo-experimental se torna melhor. Talvez fosse possível obter uma melhor generalização se esta rede passasse pelo processo de ajuste de hiper-parâmetros utilizando os dados proporcionais para tal.

3. Conclusões

As técnicas apresentadas neste trabalho se tornam úteis quando é necessário lidar com conjuntos de dados pequenos ou com pouca diversidade. Foram expostas ideias, técnicas e algoritmos para que se amplie tal conjunto e que também seja possível ajustar uma rede neural ao problema.

Ambas técnicas de treinamento, *SAE* e a padrão se mostraram bastante eficientes porém, a *SAE* apresentou um bom desempenho quando se trata do conjunto experimental. Quando é necessário maior generalização, como é o caso do conjunto uniforme, ela tem o desempenho degradado e é preferível a técnica padrão. Talvez a maior vantagem da *SAE* é o treinamento com um conjunto pequeno de dados.

O treinamento sobre o conjunto proporcional se mostrou bastante eficaz quando se trata de prever os dados que estão dentro ou próximos desta proporção, como é o caso dos experimentais, cujas proporções foram utilizadas para criar o conjunto proporcional. Já o treinamento sobre o conjunto uniforme se mostrou o mais genérico. Porém, é aquele que tem o menor desempenho médio considerando a acurácia sobre o conjunto experimental.

O conjunto de configuração única, que não foi muito explorado nesse trabalho, se mostrou bastante útil quando foi necessário diminuir a dimensão e variabilidade das entradas do modelo. Por conta disso a rede dual teve um bom desempenho quando treinada sobre este conjunto. A desvantagem é que o número de modelos crescerá junto com o número de configurações, talvez tornando inviável esta solução.

Por fim, as redes tiveram um desempenho relativamente bom, mostrando ser viável o uso de redes neurais para previsão de propriedades físicas.

Referências

- Abadi, M., Agarwal, A., Barham, P., Brevdo, E., Chen, Z., Citro, C., Corrado, G. S., Davis, A., Dean, J., Devin, M., Ghemawat, S., Goodfellow, I., Harp, A., Irving, G., Isard, M., Jia, Y., Jozefowicz, R., Kaiser, L., Kudlur, M., Levenberg, J., Mané, D., Monga, R., Moore, S., Murray, D., Olah, C., Schuster, M., Shlens, J., Steiner, B., Sutskever, I., Talwar, K., Tucker, P., Vanhoucke, V., Vasudevan, V., Viégas, F., Vinyals, O., Warden, P., Wattenberg, M., Wicke, M., Yu, Y., and Zheng, X. (2015). TensorFlow: Large-scale machine learning on heterogeneous systems. Software available from tensorflow.org.
- Barthel, C. and Götz, T. (2012). The overall worldwide saving potential from domestic refrigerators and freezers. *With results detailed for*, 11.
- Branco, P. O., Torgo, L., and Ribeiro, R. P. (2017). Smogn: a pre-processing approach for imbalanced regression.
- Chollet, F. et al. (2015). Keras. <https://keras.io>.

- Claesen, M. and De Moor, B. (2015). Hyperparameter search in machine learning. *arXiv preprint arXiv:1502.02127*.
- Colombo, E. G. (2014). Investigação numérica e experimental do desempenho de condensadores do tipo skin em refrigeradores domésticos. Master's thesis, Universidade Federal de Santa Catarina, Centro Tecnológico, Programa de Pós-Graduação em Engenharia Mecânica.
- Eletrobrás (2004). Procel - programa nacional de conservação de energia elétrica. Technical report, Eletrobrás - Empresa Brasileira Energia Elétrica S.A.
- Géron, A. (2017). *Hands-On Machine Learning with Scikit-Learn and TensorFlow: Concepts, Tools, and Techniques to Build Intelligent Systems*. O'Reilly Media.
- Hermes, C. J. L. (2006). *Uma metodologia para a simulação transiente de refrigeradores domésticos*. PhD thesis, Universidade Federal de Santa Catarina, Centro Tecnológico, Programa de Pós-Graduação em Engenharia Mecânica.
- Incropera, F. P., Lavine, A. S., Bergman, T. L., and DeWitt, D. P. (2007). *Fundamentals of heat and mass transfer*. Wiley.
- Inmetro (2020). Refrigeradores, frigobares, combinados, combinados frost-free. Technical report, Instituto nacional de metrologia, qualidade tecnologia - Inmetro.
- Junior, A. A. G. and de Sousa, A. R. (2008). *Fundamentos de metrologia científica e industrial*. Manole.
- Leonel, J. (2019). Hyperparameters in machine/deep learning.
- Russell, S. J. and Norvig, P. (2016). *Artificial intelligence: a modern approach*. Malaysia; Pearson Education Limited,.
- Seidel, R. (2001). Análise da performance e da distribuição de fluxo de calor num refrigerador doméstico. Master's thesis, Universidade Federal de Santa Catarina, Centro Tecnológico, Programa de Pós-Graduação em Engenharia Mecânica.
- Sim, J. S. and Ha, J. S. (2011). Experimental study of heat transfer characteristics for a refrigerator by using reverse heat loss method. *International Communications in Heat and Mass Transfer*, 38(5):572–576.
- Snoek, J., Larochelle, H., and Adams, R. P. (2012). Practical bayesian optimization of machine learning algorithms. In *Advances in neural information processing systems*, pages 2951–2959.
- Thiessen, S. (2015). Caracterização experimental de gabinetes refrigerados isolados com painéis de vácuo. Master's thesis, Universidade Federal de Santa Catarina, Centro Tecnológico, Programa de Pós-Graduação em Engenharia Mecânica.
- Torgo, L., Ribeiro, R., Pfahringer, B., and Branco, P. (2013). Smote for regression. volume 8154, pages 378–389.