

**UNIVERSIDADE FEDERAL DE SÃO CARLOS
CENTRO DE CIÊNCIAS EXATAS E DE TECNOLOGIA
PROGRAMA DE PÓS-GRADUAÇÃO EM CIÊNCIA E
ENGENHARIA DE MATERIAIS**

PREVISÃO DE VISCOSIDADE E TEMPERATURA *LIQUIDUS* DE VIDROS
ÓXIDOS VIA REDES NEURAS ARTIFICIAIS

Graziela Pentean Bessa

São Carlos – SP

2020

**UNIVERSIDADE FEDERAL DE SÃO CARLOS
CENTRO DE CIÊNCIAS EXATAS E DE TECNOLOGIA
PROGRAMA DE PÓS-GRADUAÇÃO EM CIÊNCIA E
ENGENHARIA DE MATERIAIS**

PREVISÃO DE VISCOSIDADE E TEMPERATURA *LIQUIDUS* DE VIDROS
ÓXIDOS VIA REDES NEURAS ARTIFICIAIS

Graziela Pentean Bessa

Dissertação apresentada ao
Programa de Pós-Graduação em Ciência e
Engenharia de Materiais como requisito
parcial à obtenção do título de MESTRA EM
CIÊNCIA E ENGENHARIA DE MATERIAIS

Orientador: Dr. Edgar Dutra Zanotto

Co-orientador: Dr. Daniel Roberto Cassar

Agência Financiadora: CNPq (Processo: 134518/2017-9)

São Carlos-SP

2020

DEDICATÓRIA

Dedico este trabalho a todos que amo e a todos os professores que tive desde a minha infância.

VITAE

Bacharela em Engenharia de Materiais pela Universidade Federal de São Carlos (2017).



UNIVERSIDADE FEDERAL DE SÃO CARLOS

Centro de Ciências Exatas e de Tecnologia
Programa de Pós-Graduação em Ciência e Engenharia de Materiais

Folha de Aprovação

Defesa de Dissertação de Mestrado da candidata Graziela Pentean Bessa, realizada em 30/04/2020.

Comissão Julgadora:

Prof. Dr. Edgar Dutra Zanotto (UFSCar)

Prof. Dr. Marcello Rubens Barsi Andreetta (UFSCar)

Profa. Dra. Viviane Oliveira Soares (UEM)

O presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Código de Financiamento 001.

O Relatório de Defesa assinado pelos membros da Comissão Julgadora encontra-se arquivado junto ao Programa de Pós-Graduação em Ciência e Engenharia de Materiais.

AGRADECIMENTOS

Agradeço ao professor Edgar por todo o apoio fornecido nos últimos anos; ao Daniel por toda a ajuda técnica e pessoal durante o mestrado e aos colegas do LaMaV por todos os bons momentos no laboratório.

Não posso deixar de agradecer ao meu marido, Maurício, por estar comigo em todos os momentos e sempre acreditar em mim. Agradeço também a minha família e aos meus amigos Larissa, Renato, Clarissa e Bruno.

Agradeço ao Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) pelo financiamento fornecido durante o projeto. Agradeço também a Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP) pelo apoio ao CEPID (2013/07793-6).

O presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – Brasil (CAPES) – Código de Financiamento 001.

RESUMO

Diante do número gigantesco de composições vítreas possíveis ($\sim 10^{52}$), o número de composições vítreas já estudadas é muito pequeno ($\sim 10^5$). Uma possibilidade ao invés de fundir e vitrificar cada composição e medir as propriedades é utilizar simulações que pré-selecionem as composições de maior interesse para sintetizar em laboratório. Os algoritmos de aprendizado de máquina podem ser utilizados para identificar padrões nos dados disponíveis na literatura (composições versus propriedades) e, dessa maneira, colaborar com a descoberta de novos materiais vítreos. Dentre os vários algoritmos existentes, a rede neural artificial tem se destacado por possuir alta capacidade de mapeamento de sistemas não lineares e foi utilizado nesta dissertação. O algoritmo utilizado nesse trabalho otimiza os hiperparâmetros da RNA a fim de encontrar a melhor topologia para cada propriedade e foi escrito na linguagem Python. A viscosidade e a temperatura *liquidus* foram escolhidas como exemplos devido à sua fundamental importância no processo de fabricação de vidros óxidos, os mais importantes do ponto de vista comercial. Como resultado, a rede neural artificial foi capaz de prever a temperatura *liquidus* com $R^2 = 0,997$, as temperaturas T2 e T4 com $R^2 = 0,999$ e T3 com $R^2 = 0,998$. Os desvios relativos médios para as redes de viscosidade são, no máximo, 2,5%, enquanto que para T_{liq} é de 3,7%. Esses erros são considerados pequenos, da mesma ordem que os erros existentes nos dados experimentais. Mas, ao analisar o comportamento de tais propriedades, é possível observar algumas falhas na previsão, como em regiões com temperaturas eutéticas para a rede de T_{liq} . Mesmo com tais erros, os modelos resultantes foram capazes de prever razoavelmente essas propriedades, isto é, apesar dos erros, os valores previstos são bem melhores que “chutes”. Para trabalhos futuros, é possível utilizar outras propriedades como dados de entrada, bem como testar outros algoritmos.

Palavras chave: Redes neurais artificiais; vidros; temperatura liquidus; viscosidade; previsão

ABSTRACT

PREDICTION OF VISCOSITY AND LIQUIDUS TEMPERATURE OF OXIDE GLASSES USING ARTIFICIAL NEURAL NETWORKS

The number of glass compositions that have already been studied ($\sim 10^5$) is many times smaller than the potential glass compositions ($\sim 10^{52}$). The traditional method “cook and look” all of the potential compositions to identify each one has an interesting property is very time-consuming and also expensive. Another possibility is to use machine learning algorithms. This type of algorithm is capable of identifying patterns in large datasets, so they can be used to accelerate the study of glass compositions. There are several types of machine learning algorithms and artificial neural networks (also known as ANN) is one of them and have been used successfully for different types of problems and also used in this thesis. The algorithms used was written to optimize hyper-parameters to find the best topology for each neural network. Viscosity and liquidus temperature were chosen to be studied because they are important parameters for oxide glass production. The ANN was capable to predict liquidus temperature with $R^2 = 0,997$ and viscosity temperature T2 and T4 with $R^2 = 0,999$ and T3 with $R^2 = 0,998$. The mean relative errors for the viscosity ANN are equal to 2,5 at maximum, while for liquidus temperature mean relative error is equal to 3,7%. These errors are considered small and at the same order of magnitude of the scattering of dataset. It is possible to see some problems on the predictions like eutetic regions in phase diagrams for liquidus temperature. Even though the errors, the neural networks were capable of predicting both properties. For future works it is possible to select other properties of atoms to form the dataset besides the composition and the property of interest.

Keywords: Artificial neural networks; glasses; liquidus temperature; viscosity. prediction.

SUMÁRIO

FOLHA DE APROVAÇÃO.....	i
AGRADECIMENTOS	iii
RESUMO.....	v
ABSTRACT	vii
SUMÁRIO.....	ix
LISTA DE TABELAS	xi
LISTA DE FIGURAS	xiii
SÍMBOLOS E ABREVIATURAS.....	xvii
1 INTRODUÇÃO	1
2 OBJETIVOS	5
3 REVISÃO BIBLIOGRÁFICA.....	7
3.1 Materiais vítreos e suas propriedades	7
3.1.1 Definição de materiais vítreos.....	7
3.1.2 Habilidade de vitrificação de líquidos.....	11
3.1.3 Temperatura <i>liquidus</i>	17
3.1.4 Viscosidade	18
3.1.5 Fabricação de vidros.....	21
3.2 Aprendizado de máquina	21
3.2.1 Redes neurais artificiais.....	23
3.2.2 Sobreajuste e subajuste	26
3.2.3 Aplicações	28
4 MATERIAS E MÉTODOS.....	33
4.1 Coleta e tratamento de dados	33
4.2 Design e treinamento da rede neural	35
4.3 Análise dos dados obtidos	36
5 RESULTADOS E DISCUSSÃO.....	37

5.1 Temperatura <i>liquidus</i>	37
5.2 Viscosidade.....	43
6 CONCLUSÕES	59
7 TRABALHOS FUTUROS	61
8 REFERÊNCIAS BIBLIOGRÁFICA	63
APÊNDICE A: ALGORITMO UTILIZADO ESQUEMATIZADO.....	69

LISTA DE TABELAS

Tabela 1: Classificação da habilidade de vitrificação de acordo com o parâmetro JZCA.	17
--	----

LISTA DE FIGURAS

Figura 1: Esquemas bidimensionais de (a) sílica cristalina e (b) sílica vítrea, onde o círculo maior representa átomos de oxigênio e o menor, átomos de silício. Reproduzido de [18]	9
Figura 2: Representação esquemática de um ensaio de DSC de um material vítreo na faixa de temperatura onde ocorre a transição vítrea. Reproduzido de [19] com autorização do autor.	10
Figura 3: Entalpia em função da temperatura para um vidro, LSR, líquido e sólido cristalino. Adaptado de [19] com autorização do autor.....	12
Figura 4: Diagrama TTT esquemático adaptado de [21], onde T_n e t_n são, respectivamente, temperatura e tempo do nariz da curva, e T_m é a temperatura de fusão.....	13
Figura 5: Viscosidade em função da temperatura para vidro soda-cal, indicando temperaturas importantes na produção. Adaptado de [36].	18
Figura 6: Representação de um neurônio artificial. Reproduzido de [11].	23
Figura 7: Arquitetura de uma rede perceptron de multicamadas. Adaptado de [11].	24
Figura 8: Erro de predição para dados de treinamento e validação ao longo do treino.	25
Figura 9: Exemplo de sobreajuste dos dados (representados por pontos). A linha cheia é o modelo que sofre sobreajuste, enquanto a linha tracejada é o modelo ideal.....	27
Figura 10: Fluxograma das etapas deste projeto	33
Figura 11: Frequência de (a) número de diferentes elementos químicos por composição e (b) valores de T_{liq} . O eixo y está representado em escala logarítmica.....	38
Figura 12: T_{liq} prevista pela rede neural versus valores reportados na literatura. A linha preta é a reta identidade que indica o que seria a previsão perfeita. O gráfico no canto superior mostra a distribuição de erros relativos (ϕ).	38
Figura 13: Distribuição dos valores do resíduo da previsão da temperatura liquidus (em Kelvin) em função do valor reportado da liquidus. As caixas compreendem os valores dos quartis Q1 e Q3 e a barra de erro (em preto) compreende 67% dos dados (um desvio padrão). A mediana é ilustrada por uma reta laranja horizontal e seu intervalo de confiança (95%) é indicado pelos	

entalhos nas caixas. Apenas os dados reservados pela rotina de holdout foram utilizados para construir este gráfico.	39
Figura 14: Média e desvio padrão do resíduo da previsão da temperatura liquidus (em Kelvin) em função do elemento químico presente na composição. O número em parêntesis (ao lado de cada elemento) indica o número de composições que contém o elemento químico e se encontra em ordem crescente. Apenas os dados reservados pela rotina de holdout foram utilizados para construir este gráfico.	40
Figura 15: Diagrama de fases versus previsão da rede neural artificial do sistema binário $\text{SiO}_2\text{-Na}_2\text{O}$	41
Figura 16: Diagrama de fases versus previsão da rede neural artificial do sistema binário $\text{B}_2\text{O}_3\text{-Na}_2\text{O}$	42
Figura 17: Frequência de (a) número de diferentes elementos químicos por composição e (b) valores de T2. O eixo y está representado em escala logarítmica.	44
Figura 18: T2 prevista pela rede neural versus valores reportados na literatura. A linha preta é a reta identidade que indica o que seria a previsão perfeita. O gráfico no canto superior mostra a distribuição de erros relativos (ϕ).	45
Figura 19: Distribuição dos valores do resíduo da previsão da T2 (em Kelvin) em função do valor reportado. As caixas compreendem os valores dos quartis Q1 e Q3 e a barra de erro (em preto) compreende 67% dos dados (um desvio padrão). A mediana é ilustrada por uma reta laranja horizontal e seu intervalo de confiança (95%) é indicado pelos entalhos nas caixas. Apenas os dados reservados pela rotina de holdout foram utilizados para construir este gráfico.	46
Figura 20: Média e desvio padrão do resíduo da previsão da T2 (em Kelvin) em função do elemento químico presente na composição. O número em parêntesis (ao lado de cada elemento) indica o número de composições que contém o elemento químico, e se encontra em ordem crescente. Apenas os dados reservados pela rotina de holdout foram utilizados para construir este gráfico.	47
Figura 21: Frequência de (a) número de diferentes elementos químicos por composição e (b) valores de temperatura T3. O eixo y está representado em escala logarítmica.	48

Figura 22: T3 prevista pela rede neural versus valores reportados na literatura. A linha preta é a reta identidade que indica o que seria a previsão perfeita. O gráfico no canto superior mostra a distribuição de erros relativos (ϕ).	49
Figura 23: Distribuição dos valores do resíduo da previsão da T3 (em Kelvin) em função do valor reportado. As caixas compreendem os valores dos quartis Q1 e Q3 e a barra de erro (em preto) compreende 67% dos dados (um desvio padrão). A mediana é ilustrada por uma reta laranja horizontal e seu intervalo de confiança (95%) é indicado pelos entalhes nas caixas. Apenas os dados reservados pela rotina de holdout foram utilizados para construir este gráfico.	50
Figura 24: Média e desvio padrão do resíduo da previsão da T3 (em Kelvin) em função do elemento químico presente na composição. O número em parêntesis (ao lado de cada elemento) indica o número de composições que contém o elemento químico e se encontra em ordem crescente. Apenas os dados reservados pela rotina de holdout foram utilizados para construir este gráfico.	50
Figura 25: Frequência de (a) número de diferentes elementos químicos por composição e (b) valores de temperatura T4. O eixo y está representado em escala logarítmica.	51
Figura 26: T4 prevista pela rede neural versus valores reportados na literatura. A linha preta é a reta identidade que indica o que seria a previsão perfeita. O gráfico no canto superior mostra a distribuição de erros relativos (ϕ).	52
Figura 27: Distribuição dos valores do resíduo da previsão da T4 (em Kelvin) em função do valor reportado. As caixas compreendem os valores dos quartis Q1 e Q3 e a barra de erro (em preto) compreende 67% dos dados (um desvio padrão). A mediana é ilustrada por uma reta laranja horizontal e seu intervalo de confiança (95%) é indicado pelos entalhes nas caixas. Apenas os dados reservados pela rotina de holdout foram utilizados para construir este gráfico.	53
Figura 28: Média e desvio padrão do resíduo da previsão da T4 (em Kelvin) em função do elemento químico presente na composição. O número em parêntesis (ao lado de cada elemento) indica o número de composições que contém o elemento químico e se encontra em ordem crescente. Apenas os dados reservados pela rotina de holdout foram utilizados para construir este gráfico.	53
Figura 29: T2, T3 e T4 do binário SiO ₂ -Na ₂ O em função da porcentagem molar de Na ₂ O.....	55
Figura 30: T2, T3 e T4 do binário B ₂ O ₃ -Na ₂ O em função da porcentagem molar de Na ₂ O.....	55

Figura 31: T2, T3 e T4 do binário $P_2O_5-Na_2O$ em função da porcentagem molar de Na_2O 56

Figura 32: Curvas de T2, T3, T4 e liquidus do diagrama de fases da referência [61] para o sistema binário SiO_2-Na_2O 57

Figura 33: Curvas de T2, T3, T4 e liquidus do diagrama de fases da referência [62] para o sistema binário $B_2O_3-Na_2O$ 58

SÍMBOLOS E ABREVIATURAS

A	Parâmetro das equações de viscosidade VFTH e AG relacionado a energia necessária para escoamento viscoso
AD	Árvore de decisão
AG	Adam-Gibbs
DSC	Calorimetria Exploratória Diferencial (<i>Differential Scanning Calorimetry</i>)
DTA	Análise Térmica Diferencial (<i>Differential Thermal Analysis</i>)
E	Experiência
f(x)	Função na variável x
GFA	Habilidade de vitrificação (<i>Glass Forming Ability</i>)
G _L	Energia livre do líquido
G _s	Energia livre do sólido
GS	Estabilidade vítrea (<i>Glass Stability</i>)
GPR	Regressão do processo Gaussiano
g(.)	Função de ativação
JZCA	Parâmetro Jiusti-Zanotto-Cassar-Andreeta para prever GFA
K _H	Parâmetro de Hruby, 1972
LSR	Líquido super-resfriado
MVS	Máquinas de vetores de suporte
MYEGA	Equação de viscosidade proposta por Mauro-Yue-Ellison-Gupta-Allan
P	Performance
Q1	Primeiro quartil
Q3	Terceiro quartil
RBFANN	Redes neurais com função de ativação de base radial (<i>Radial basis function neural network</i>)
R _c	Taxa de resfriamento crítica
RD	Desvio médio relativo
RNA	Rede neural artificial
R ²	Coeficiente de determinação
S _c	Entropia configuracional

T	Temperatura
T_a	Tarefa
T_g	Temperatura de transição vítrea
T_{gr}	Temperatura de transição vítrea reduzida
$T_{g,12}$	Temperatura onde a viscosidade é 10^{12} Pa.s
T_{liq}	Temperatura <i>liquidus</i>
T_m	Temperatura de fusão
T_{max}	Temperatura cuja taxa de crescimento de cristal é máxima
t_n	Tempo mínimo do nariz do diagrama TTT
T_n	Temperatura do nariz do diagrama TTT
TTT	Tempo, Temperatura e Transformação
T_x	Temperatura de início de cristalização
T_0	Parâmetro da equação VFTH onde a viscosidade diverge para o infinito
T_2	Temperatura cuja viscosidade é igual a 10^2 Pa.s
T_3	Temperatura cuja viscosidade é igual a 10^3 Pa.s ou ponto de trabalho (<i>Working point</i>)
T_4	Temperatura cuja viscosidade é igual a 10^4 Pa.s ou ponto de escoamento (<i>Flow point</i>)
u	Potencial de ativação
U	Taxa de crescimento de cristal
VFTH	Equação de viscosidade proposta por Vogel-Fulcher-Tammann-Hesse
w_n	Peso de sinapse do sinal de entrada
x_n	Sinal de entrada de um neurônio artificial
y	Valor de saída de um neurônio artificial
ΔT_x	Região de líquido super-resfriado
ΔH_f	Entalpia de fusão
ϵ	Erro irreduzível
η	Viscosidade
η_∞	Parâmetro pré-exponencial da equação VFTH
Θ	Limiar de ativação

ϕ	Erro relativo
Σ	Função agregadora

1 INTRODUÇÃO

Vidros são materiais fora do equilíbrio termodinâmico, não cristalinos, que possuem apenas ordem atômica de curto alcance e apresentam o fenômeno de transição vítrea [1]. Ao contrário de materiais policristalinos, onde as propriedades dependem da microestrutura, as propriedades de materiais vítreos monofásicos fabricados pela rota de fusão e resfriamento, em geral, são pouco influenciadas pelas variáveis de processamento durante sua fabricação quando comparadas com materiais cristalinos. Elas dependem essencialmente da composição química, fato que torna os vidros excelentes candidatos para estudos via técnicas de inteligência artificial [2,3,4,5,6,7,8,9].

Existe um número muito grande de possibilidades para composições vítreas. Isto se deve por dois motivos principais:

1. A ausência de estrutura cristalina permite um maior número de arranjos atômicos para materiais vítreos do que materiais cristalinos, pois os vidros não têm a necessidade de seguir a estequiometria química como os cristais.
2. Vidros podem ser feitos por praticamente todos os elementos da tabela periódica. Existem aproximadamente 80 elementos químicos usuais (não-radioativos) que podem ser utilizados para a fabricação de vidros.

De acordo com Zanotto e Coutinho [10], de todas as possíveis combinações de elementos para formar vidros, apenas uma pequena fração foi estudada durante toda a história. Eles observaram que, de um total de mais de 10^{52} composições possíveis (utilizando certos critérios), apenas 200.000 composições de vidros inorgânicos foram fabricadas e reportadas na literatura. Mesmo considerando que este número de vidros já fabricados aumentou uma ordem de grandeza desde a publicação do artigo (o que provavelmente não é o caso), isto ainda significa que existe uma enorme quantidade de composições vítreas a serem exploradas. Dentre essas possibilidades, podem existir composições com combinações de propriedades de extremo interesse para aplicações tecnológicas, que ainda não foram descobertas. O desafio neste

contexto é a metodologia para escolher quais destas composições desconhecidas tem o maior potencial para aplicações tecnológicas, minimizando a necessidade de experimentos. Afinal, não seria possível por questão de tempo e disponibilidade de matéria-prima testar todas ou a maior parte das composições ainda inexploradas.

A solução proposta neste trabalho é identificar padrões nos dados que estão disponíveis na literatura para extrapolar para composições nunca estudadas. Uma maneira de encontrar estes padrões em grandes conjuntos de dados é através de rotinas de aprendizado de máquina. Dentre os vários algoritmos disponíveis, o que tem se destacado é a rede neural artificial (RNA), que possui alta capacidade de mapeamento de sistemas não lineares, o que permite identificar padrões nos dados obtidos do sistema estudado [11]. Esse algoritmo já vem sendo utilizado para reconhecimento de objetos [12], reconhecimento facial [13], reconhecimento de discursos, entre tantos outros exemplos. Em específico, na área de ciência de materiais, as redes neurais já foram utilizadas para prever temperatura de transição vítrea [3,14], temperatura *liquidus* [15], solubilidade de vidros [16,2], módulo de Young [5,4] e densidade [7].

A viscosidade e a temperatura *liquidus* foram escolhidas para serem estudadas neste trabalho com o auxílio de redes neurais artificiais, por serem propriedades amplamente estudadas e importantes no processo de fabricação de vidros. Além disso, ambas as propriedades tem comportamento já conhecido em diversos sistemas binários, através de modelos e diagramas de fases, respectivamente. Desta maneira, ao estudá-las em sistemas binários, será possível confrontar os resultados para verificar a capacidade de previsão da rede neural artificial utilizada.

Outro fator considerado na escolha das propriedades de estudo é que ambas estão relacionadas com a habilidade de vitrificação de líquidos (do inglês, *glass forming ability* - GFA) conforme a equação de JZCA [17].

Portanto, este trabalho tem como objetivo desenvolver e treinar redes neurais artificiais para prever viscosidade e temperatura *liquidus* de composições

de vidros óxidos, uma vez que estas são as mais produzidas mundialmente. A relevância deste trabalho reside no potencial de economia de tempo e recursos no processo de desenvolvimento de novos materiais vítreos através da aplicação das redes neurais.

2 OBJETIVOS

O objetivo deste trabalho é avaliar a aplicabilidade de redes neurais artificiais para prever propriedades de vidros óxidos, mais especificamente, a viscosidade e a temperatura *liquidus*, a partir de dados encontrados na literatura. Pra isso, quatro redes neurais artificiais foram treinadas para cada propriedade: temperatura *liquidus* e viscosidade. No caso da viscosidade, foram escolhidas as temperaturas T2, T3 e T4 (temperatura onde a viscosidade do líquido é 10^2 , 10^3 e 10^4 Pa.s, respectivamente).

3 REVISÃO BIBLIOGRÁFICA

3.1 Materiais vítreos e suas propriedades

3.1.1 Definição de materiais vítreos

A definição mais recente de vidros na literatura é a de Zanotto e Mauro [1]: o vidro é um estado condensado da matéria fora do equilíbrio, não cristalino e que exhibe o fenômeno de transição vítrea. A estrutura atômica dos vidros é similar à do líquido super-resfriado (LSR) que o origina, sendo que os vidros relaxam espontaneamente para o LSR. Em uma escala infinita de tempo, o vidro tende a cristalizar.

A primeira parte da definição aborda o caráter termodinâmico do vidro. Estes são instáveis, ou seja, fora do equilíbrio termodinâmico. Acima da temperatura de fusão (T_m), o estado líquido é o estável (possui menor energia livre de Gibbs), e abaixo da T_m , o estado estável é o cristalino. Sendo assim, vidros não são o estado termodinâmico de menor energia livre de Gibbs em nenhuma faixa de temperatura. Dois importantes fenômenos podem ocorrer para reduzir a energia livre de Gibbs em um material vítreo:

- Relaxação estrutural: fenômeno onde a estrutura atômica do vidro se altera em função do tempo em direção à estrutura de equilíbrio metaestável do LSR;
- Cristalização da estrutura atômica: fenômeno onde a estrutura atômica de um vidro ou LSR se altera gerando uma região cristalina, buscando uma situação de equilíbrio termodinâmico estável.

Em temperaturas muito abaixo da temperatura de transição vítrea, no entanto, as cinéticas descritas acima são muito lentas e o tempo necessário para atingir uma situação de equilíbrio é muito grande.

A segunda parte da definição aborda o caráter estrutural do vidro, pontuando que são materiais não-cristalinos. Isto é, o arranjo dos átomos no espaço não é ordenado. Existem três possibilidades clássicas para o arranjo atômico:

1. **Arranjo Aleatório:** os átomos preenchem o espaço de maneira aleatória, sem qualquer regra (materiais amorfos);
2. **Arranjo com ordem de curto alcance:** os átomos possuem ordem apenas considerando uma curta distância, ou seja, apenas os átomos adjacentes possuem posições bem definidas (materiais vítreos);
3. **Arranjo com ordem de curto e longo alcance:** os átomos estão arranjados de forma periódica no espaço (materiais cristalinos).

Fica claro então a primeira diferença entre materiais vítreos e amorfos através do arranjo atômico. A segunda diferença é que os materiais amorfos não apresentam o fenômeno de transição vítrea, que será abordado mais adiante.

Para ilustrar a distinção entre materiais vítreos e cristalinos, a Figura 1 exemplifica graficamente a diferença entre a sílica vítrea e cristalina (SiO_2). A imagem é um esquema bidimensional das estruturas reais, visto que elas são tridimensionais. Nota-se que no caso da estrutura cristalina, Figura 1a, as posições dos átomos de silício e oxigênio ao longo do espaço são periódicas, formando anéis hexagonais que se repetem periodicamente ao longo do espaço. Vários eixos de simetria podem ser observados. Diferentemente, no caso da sílica vítrea (Figura 1b) não existe essa mesma distribuição ordenada de átomos, tendo anéis irregulares de tamanhos diferentes. Porém é possível observar que existe uma ordem de curto alcance, que são as ligações de silício e oxigênio que se repetem.

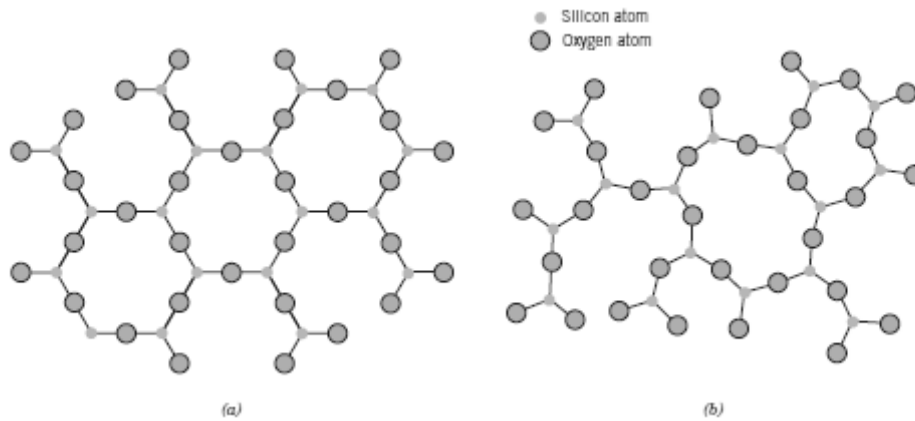


Figura 1: Esquemas bidimensionais de (a) sílica cristalina e (b) sílica vítrea, onde o círculo maior representa átomos de oxigênio e o menor, átomos de silício. Reproduzido de [18]

Por fim, o terceiro conceito importante na definição é que os vidros apresentam o fenômeno de transição vítrea. A transição vítrea é a faixa de temperatura na qual o tempo médio de relaxação da estrutura é da mesma ordem que o tempo de observação. Na prática, o material ao passar pela transição vítrea no resfriamento muda seu comportamento mecânico. Acima da faixa de temperatura da transição vítrea ele se encontra como líquido metaestável (líquido super-resfriado) e abaixo, como vidro. O vidro se comporta mecanicamente como um sólido, que é caracterizado por ter forma definida e só flui (plasticidade) se a tensão aplicada for maior que a tensão de escoamento do material; mas o vidro apresenta estrutura atômica similar à de um líquido. O estado líquido, no entanto, difere do estado vítreo por apresentar escoamento viscoso da matéria, que é capaz de se moldar pela força da gravidade de acordo com o recipiente no qual estiver.

Muito embora a transição vítrea seja representada por uma única temperatura, chamada de temperatura de transição vítrea (T_g), na realidade existe uma faixa de temperatura onde esse fenômeno ocorre quando um experimento dinâmico (isto, com taxa de variação de temperatura diferente de zero) é realizado. Este fato gera uma discrepância entre valores reportados pois existem algumas possibilidades para definir a T_g a partir da curva de um ensaio de calorimetria diferencial exploratória (DSC), como por exemplo: pelo *onset*,

pelo ponto de inflexão ou pelo *endpoint*. A Figura 2 traz um exemplo de como determinar a T_g por essas três formas, além de apresentar o comportamento típico de um vidro durante a transição vítrea.

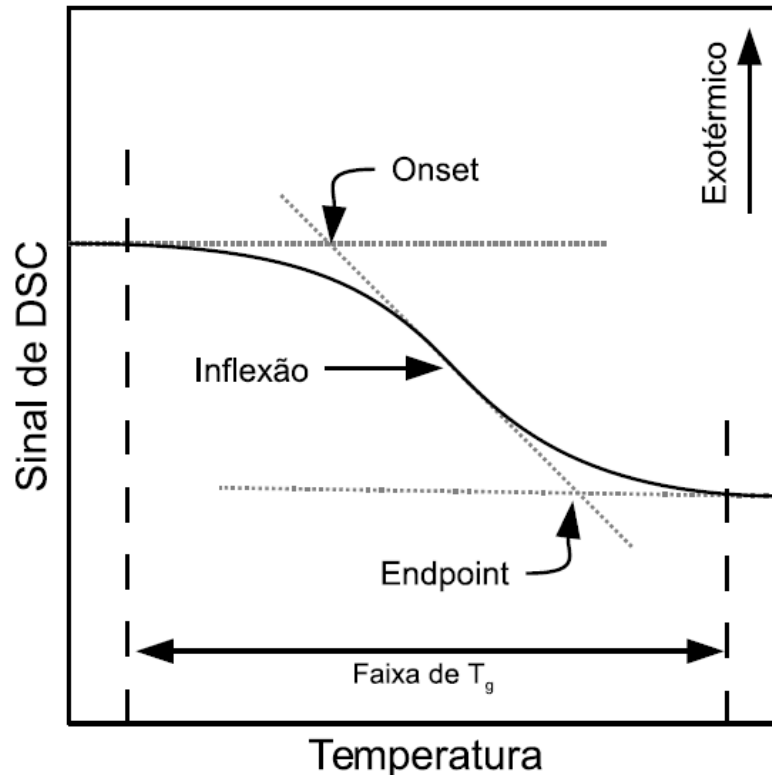


Figura 2: Representação esquemática de um ensaio de DSC de um material vítreo na faixa de temperatura onde ocorre a transição vítrea. Reproduzido de [19] com autorização do autor.

Por fim, existem diferentes composições possíveis para materiais vítreos, podendo ser óxidos, metálicos, orgânicos, inorgânicos, não-óxidos e não-metálicos. O foco deste trabalho são os vidros óxidos. Estes possuem grande importância comercial e são os vidros de maior produção em volume. Os vidros óxidos são compostos por alguns elementos estruturais, sendo eles:

1. **Formador de rede:** estes átomos participam da formação da cadeia estrutural do vidro. Os principais elementos formadores são: silício (Si), boro (B), fósforo (P), telúrio (Te) e germânio (Ge);

2. **Modificador de rede:** estes átomos são responsáveis por quebrar as cadeias estruturais, se ligando a oxigênios que passam a ser chamados de oxigênios não-ponteantes. Os elementos da família 1A e 2A, por exemplo, atuam como modificadores de rede;
3. **Oxigênio (ponteante e não ponteante):** os átomos de oxigênio também participam da cadeia, sendo eles ponteantes quando estão ligados apenas entre formadores e não ponteantes quando eles estão ligados a pelo menos um modificador e um formador.

3.1.2 Habilidade de vitrificação de líquidos

Conforme ilustra o gráfico de volume específico em função da temperatura (Figura 3), vidros só existem abaixo de T_g . É possível obter um vidro através do resfriamento de um líquido, desde que a taxa de resfriamento seja suficientemente alta que impossibilite a reordenação de seus átomos em um arranjo cristalino. Desta forma, o líquido continua existindo abaixo da T_m , onde passa a ser chamado de líquido super-resfriado. Ao prosseguir o resfriamento, o material se torna um vidro quando a energia cinética de seus átomos deixa de ser suficiente para manter o estado de equilíbrio metaestável de sua estrutura. De maneira simplificada, diz-se que o material vitrificou ao passar pela sua faixa de temperatura de transição vítrea.

A taxa de resfriamento necessária para vitrificação é inversamente proporcional à habilidade de vitrificação (GFA) da composição. A menor taxa de resfriamento capaz de vitrificar o fundido é chamada de taxa de resfriamento crítica (R_c). Bons formadores de vidros possuem R_c menor que maus formadores, sendo que estes últimos podem precisar de remoção forçada do calor para vitrificar. Diversos parâmetros influenciam na GFA, como a viscosidade do LSR, a maneira como os átomos estão arranjados na estrutura do vidro ou líquido super-resfriado, força de ligação entre os átomos e cinética de cristalização. Na teoria qualquer composição que exista no estado líquido pode ser vitrificada, sendo necessário apenas fornecer a taxa de resfriamento suficiente para impedir a cristalização do mesmo [20].

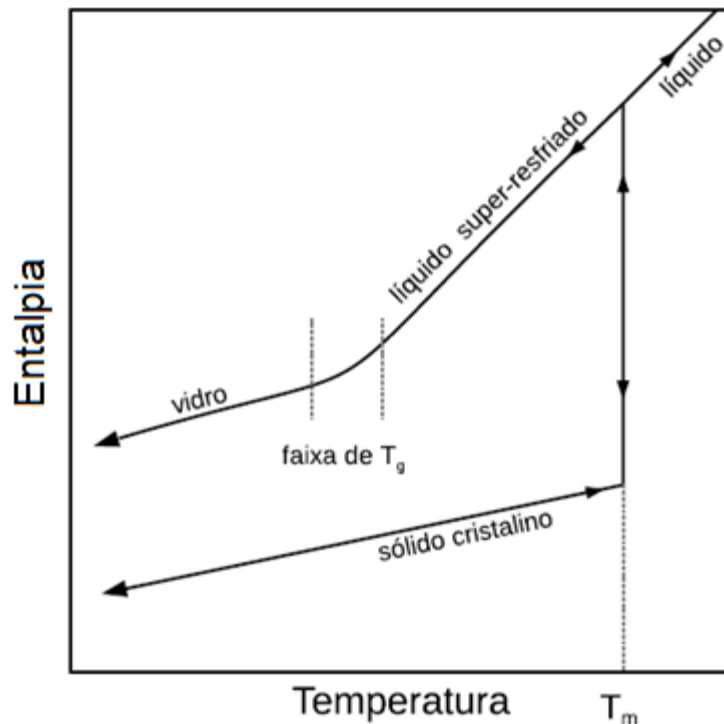


Figura 3: Entalpia em função da temperatura para um vidro, LSR, líquido e sólido cristalino. Adaptado de [19] com autorização do autor.

Mas como medir a R_c para um determinado vidro? Para tanto, pode-se construir um diagrama de tempo, temperatura e transformação, mais conhecido como diagrama TTT. Ele indica qual o tempo e temperatura necessários para conseguir obter uma fração cristalina específica. Uma fração cristalina usual para determinar a R_c é 10^{-6} .

Na Figura 4 encontra-se um diagrama TTT esquemático, onde observa-se o comportamento típico dessas curvas, que é a presença de um nariz, no qual a fração cristalina pode ser obtida no menor tempo possível (t_n) em uma determinada temperatura T_n . Através destes valores, é possível calcular a taxa de resfriamento crítica (R_c) através da Equação 1, que se resume a calcular a tangente que passa tanto por T_m quanto pelo nariz do diagrama. Taxas de resfriamento menores que a taxa crítica cruzam a linha do diagrama TTT e, então, obtêm-se um material com uma fração volumétrica de fase cristalina acima da desejada. Para taxas de resfriamento maiores que R_c o material é um

vidro, ou seja, a taxa de resfriamento foi rápida o bastante para impedir uma certa quantidade mínima de rearranjo atômico. A taxa de resfriamento crítica é a menor taxa necessária para obter o vidro, passando exatamente no nariz da curva do diagrama TTT.

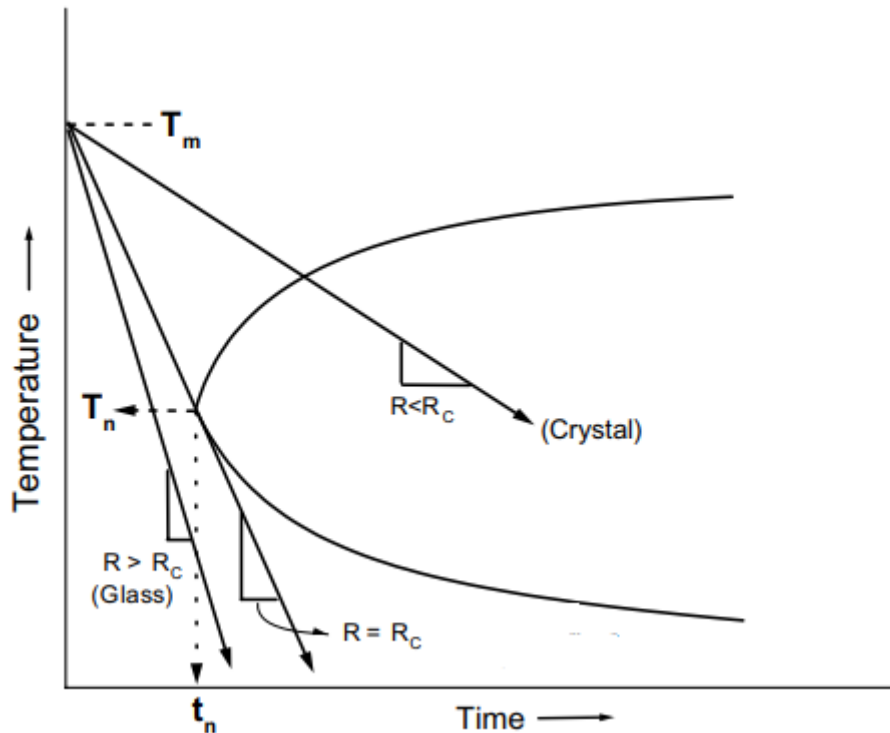


Figura 4: Diagrama TTT esquemático adaptado de [21], onde T_n e t_n são, respectivamente, temperatura e tempo do nariz da curva, e T_m é a temperatura de fusão.

$$R_c = \frac{T_m - T_n}{t_n} \quad (1)$$

Apesar de ser uma ótima ferramenta, a construção do diagrama TTT é realizada a partir de diversos experimentos de resfriamento isotérmico [22], o que demanda tempo. Além disso, a R_c aqui obtida não equivale à realidade, pois o resfriamento aplicado nos processos é contínuo e não isotérmico. Assim, a R_c obtida através do diagrama TTT é superestimada em uma ordem de grandeza [23,24]. Existe ainda os diagramas TTT de resfriamento contínuo, que

representam melhor as condições reais de resfriamento, entretanto, podem subestimar os valores de R_c por conta de levar em consideração apenas os dados de viscosidade e propriedades termodinâmicas [25].

Outra maneira possível para determinar a R_c é utilizando as análises térmicas, como a calorimetria diferencial exploratória (DSC). Neste caso o processo é realizado através de uma série de experimentos controlados através do computador. Se a taxa de resfriamento aplicada foi menor que a crítica, haverá um pico de cristalização. Ao aplicar a taxa igual ou maior que a crítica, não haverá mais o pico de cristalização, sendo possível determinar a R_c [26]. Porém, a taxa determinada por este método pode ter um erro muito grande [27]. Outra técnica é a análise térmica diferencial (DTA), que consiste em analisar se há a presença ou não de um pico de cristalização, mas diferente do DSC (que a leitura é feita no resfriamento), nesta técnica a leitura é realizada no reaquecimento de amostras resfriadas em diferentes taxas. Desta maneira o erro seria menor que o DSC [21]. No entanto, os autores [21] utilizaram cadinho de platina para vidros silicato, o que pode atuar como agente nucleante e mudar a cinética de cristalização, gerando um erro nos valores de R_c encontrados.

Por conta da dificuldade de medir experimentalmente a R_c , diversos autores [17,28,29,30] buscam entender quais os parâmetros que afetam a GFA em busca de uma maneira mais fácil de quantificar esta propriedade, tanto para vidros óxidos quanto para vidros metálicos. Alguns autores [31,24,32,33] começaram a estabelecer uma relação entre estabilidade vítrea (GS) dos vidros com a GFA, mas estas são propriedades diferentes que podem não estar correlacionadas em alguns casos. A estabilidade vítrea é a resistência que o vidro apresenta a cristalizar durante seu aquecimento. Normalmente se assume que quanto menor a GFA, menor a GS, o que não é verdade para vidros de aluminato de cálcio, por exemplo, que apesar de uma baixa GFA apresentam uma GS razoável [26]. De acordo com Nascimento *et al.* [24], parâmetros de estabilidade vítrea e a taxa de resfriamento crítico estão correlacionadas quando o vidro em questão apresenta nucleação predominantemente heterogênea na escala de tempo laboratorial.

Na literatura existem diversos critérios para quantificar a GS de vidros [32]. Normalmente, se baseia na diferença de temperatura de início de cristalização (T_x) e temperatura de transição vítrea (T_g), sendo que alguns autores defendem que este valor deve ser normalizado por T_g , T_x , T_m , entre outros. Todas essas temperaturas podem ser obtidas por meio de medidas de DSC do vidro, entretanto, para T_x pode haver uma discrepância entre valores reportados (conforme explicado na Figura 3 pela maneira de determinar a temperatura).

Um critério que é bem conhecido e estudado é o proposto por Hrubý [32], que é calculado pela Equação 2. Composições com K_H próximo a 0,1 são maus formadores, próximos de 0,5 são bons formadores e acima de 1 possuem alto grau de polimerização. De acordo com Cabral *et al.* [31] este parâmetro pode ser usado para estimar a GFA se o vidro em questão apresentar nucleação de cristais volumétrica (homogênea).

$$K_H = \frac{T_x - T_g}{T_{liq} - T_x} \quad (2)$$

Justi *et al.* [17] desenvolveram e testaram dois novos parâmetros para prever a GFA. O primeiro deles, chamado de parâmetro mãe, é dado pela Equação 3, que foi derivado a partir da Equação 1. Entretanto, a dificuldade em utilizar esta equação é devido ao fato que dados de crescimento de cristais (U) são encontrados apenas para alguns vidros, impossibilitando um estudo mais global.

$$R_C = \frac{1}{GFA} \propto T_L U(T_{max}) \quad (3)$$

Através desta equação, para o modelo de crescimento por discordâncias no cristal, a GFA pode ser aproximada para a Equação (4), batizada de JZCA, muito mais interessante para estudo pois na literatura há muito mais dados para

temperatura *liquidus* e viscosidade do que dados para taxa de crescimento de cristais.

$$GFA \propto \frac{\eta(T_{liq})}{T_{liq}^2} \quad (4)$$

Como premissa para ambos os parâmetros, assumiu-se que a nucleação heterogênea é quem de fato controla a habilidade de vitrificação, embora existam na natureza algumas composições com alta taxa de nucleação homogênea. Esta afirmação se prova razoável por demonstração no artigo devido a comparações feitas com os vidros de anortita e fresnoita.

Ambas as equações foram testadas utilizando dados de binários e ternários de vidros estequiométricos e não estequiométricos provenientes da literatura, com diferentes habilidades de formação de vidros, desde muito boas até extremamente ruins. No fim, ambos os parâmetros apresentaram boa correlação com a taxa de resfriamento críticas. A equação JZCA também está de acordo com o conceito de que composições que apresentam maior viscosidade na temperatura *liquidus* e uma T_{liq} menor vitrificam mais fácil. Os máximos e mínimos valores estimados para $\log(JZCA)$ estão entre +1 e -11, e, conforme a Tabela 1, quanto maior o valor do parâmetro, melhor a habilidade de vitrificação do líquido.

Tabela 1: Classificação da habilidade de vitrificação de acordo com o parâmetro JZCA.

Log(JZCA)	Classificação
> -1	Ótimo formador de vidro
-1 a -4	Bom formador de vidro (vidros comerciais)
-4 a -7	Relutantes e maus formadores
< -7	Maus formadores

3.1.3 Temperatura *liquidus*

A temperatura *liquidus* é a maior temperatura onde uma ou mais fases cristalinas coexistem em equilíbrio termodinâmico, com pelo menos uma fase líquida. É uma propriedade muito importante no processo de fabricação de vidros pois determina qual a temperatura mínima do processo de fusão que garante que nenhum cristal esteja presente no fundido. Se o líquido precursor do vidro não estiver completamente fundido e homogêneo, pode gerar defeitos visuais na peça final [34].

É uma propriedade difícil de medir para composições de vidros comerciais, pois a cristalização observada é muito pequena. O método mais comum para medir tal propriedade é cristalizar o vidro a uma temperatura próxima da *liquidus* esperada e observar via microscopia a cristalização da amostra. Outra técnica é utilizar um forno gradiente, onde uma mesma amostra é submetida a diferentes temperaturas, o que acelera o processo mencionado anteriormente. Também é possível utilizar a técnica de DSC para medir, porém é necessário garantir a cristalização durante a corrida, então, por vezes, é necessário um tratamento prévio, conforme Ferreira *et al.* [35].

Existem ainda programas de computador como FactSage e Thermo-Calc que são baseados no método CALPHAD (*Calculation of phase diagrams* – do inglês: cálculo de diagramas de fases), capazes de determinar diagramas de

fases com base em um banco de dados. Assim, ele efetua cálculos de energia livre de Gibbs para as fases presentes no sistema em estudo (seja ele binário ou multicomponente) através de modelos matemáticos [36,37].

3.1.4 Viscosidade

Viscosidade é uma medida da resistência do líquido à deformação plástica por cisalhamento. Sua dependência varia tanto com a temperatura (chegando a 12 ordens de magnitude do vidro fundido até a peça final) quanto com a composição, sendo que mesmo impurezas podem alterar drasticamente a viscosidade (como por exemplos impurezas em vidro de sílica puro). Por ser uma propriedade muito importante em todas as etapas do processo de fabricação de vidros, desde a fusão até o recozimento da peça, há um grande interesse em conseguir prever esta propriedade [38]. A Figura 5 ilustra diferentes etapas do processo de fabricação de vidros em função da viscosidade e temperatura do material.

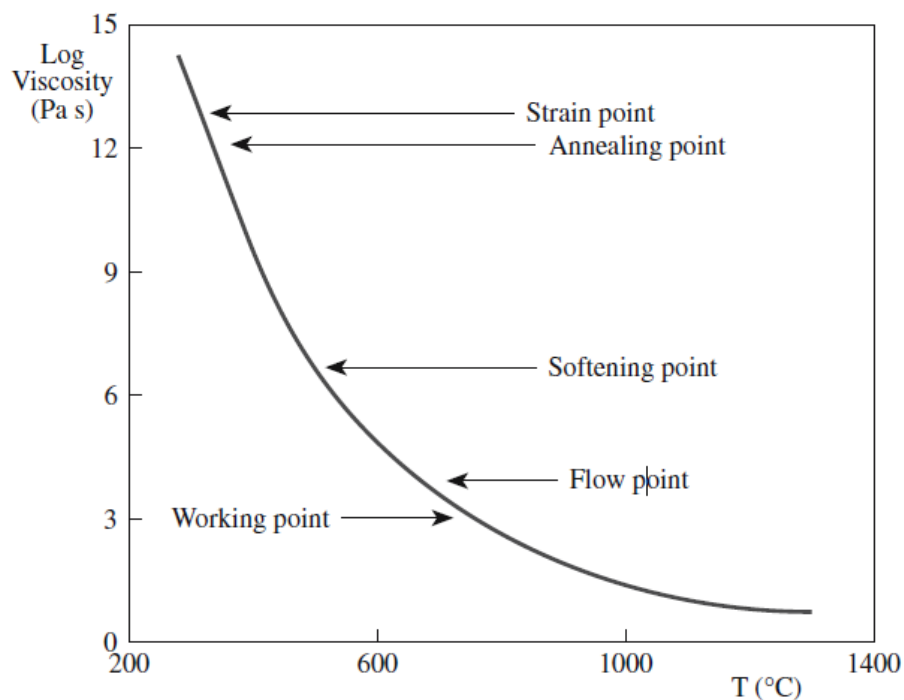


Figura 5: Viscosidade em função da temperatura para vidro soda-cal, indicando temperaturas importantes na produção. Adaptado de [36].

A temperatura de deformação (do inglês, *strain point*) é a temperatura onde abaixo dela o vidro tem comportamento frágil e ocorre quando a viscosidade está na ordem de 10^{14} Pa.s. Em 10^{12} Pa.s ocorre a temperatura de recozimento (do inglês, *annealing point*) onde é possível aliviar as tensões no material decorrentes do processo de conformação. A temperatura de amolecimento (do inglês, *softening point*) é o início da chamada região de conformação ou fusão do vidro, onde ocorre a maioria dos processos industriais de conformação de vidros e a viscosidade é da ordem de 10^7 Pa.s. Já o fim da região de conformação é dado pela temperatura de trabalho (do inglês, *working point*), cuja viscosidade é da ordem de 10^3 Pa.s (que, neste trabalho, será tratada também como T3). É na temperatura de trabalho que a maioria dos processos industriais ocorre, pois é mais fácil de conformar o material. Entre essas duas temperaturas, ocorre a *flow point* na faixa de 10^4 Pa.s, na qual o material começa a escoar de maneira irrestrita. Esta temperatura é conhecida como T4. Para viscosidades abaixo de 10^2 , conhecida como T2, o vidro já se comporta como um líquido. Em resumo, é entre T4 e T2 que temos uma mudança significativa do material, partindo de um líquido na T2 para um material que consegue manter o formato em T4, muito embora ele não se comporte como um sólido pois ainda é possível moldá-lo.

Existem diversas equações que descrevem a variação da viscosidade em função da temperatura. Um modelo bastante utilizado é o VFTH, que foi obtido empiricamente e independentemente por Vogel, Fulcher, Tammann e Hesse entre os anos de 1921 e 1926 [39,40,41]. A Equação 6 é a forma geral dessa expressão, onde $\eta(T)$ é a viscosidade, T_0 a temperatura para qual o modelo diverge e o parâmetro A, relacionado a energia necessária para escoamento viscoso.

$$\eta(T) = \eta_{\infty} 10^{\frac{A}{T-T_0}} \quad (6)$$

Mauro *et al.* [42] reescreveram a equação de maneira que todos os parâmetros tivessem sentido físico. A Equação 7 é a equação VFTH reescrita,

onde $T_{g,12}$ é a temperatura onde a viscosidade é 10^{12} Pa.s e m é o índice de fragilidade do vidro. Apesar do nome ser “fragilidade”, esse valor não está relacionado as propriedades mecânicas do vidro, mas sim ao comportamento da viscosidade na região de $T_{g,12}$. Fluidos fortes apresentam comportamento linear em gráfico do logaritmo da viscosidade pelo inverso da temperatura, enquanto que fluidos frágeis desviam da linearidade. A Equação 8 é utilizada para calcular o índice de fragilidade.

$$\eta(T) = \eta_{\infty} 10^{\frac{(12 - \log(\eta_{\infty}))^2}{m \left(\frac{T}{T_{g,12}} - 1 \right) + (12 - \log(\eta_{\infty}))}} \quad (7)$$

$$m = \frac{d \log(\eta(T))}{d \left(\frac{T_{g,12}}{T} \right)}_{T=T_{g,12}} \quad (8)$$

A equação de Adam-Gibbs (AG), de 1965 [43], também é um modelo muito importante por conectar a viscosidade, uma propriedade dinâmica, com a entropia configuracional, uma propriedade termodinâmica. A entropia configuracional é a diferença entre a entropia do líquido e a entropia do vidro em uma dada temperatura. É possível calcular a entropia configuracional com dados de calorimetria do material.

$$\eta(T) = \eta_{\infty} e^{\frac{A}{T S_c(T)}} \quad (9)$$

Um outro modelo, de 2006, foi desenvolvido através de configurações físicas por Mauro-Yue-Ellison-Gupta-Allan [44] a partir da equação de Adam-Gibbs. Ficou conhecido como modelo MYEGA (Equação 10) e sua vantagem sobre os demais modelos é uma melhor descrição da viscosidade em função da temperatura, especialmente para temperaturas menores que a temperatura de transição vítrea.

$$\log \eta(T) = \log \eta_{\infty} + \frac{B}{T} e^{\frac{C}{T}} \quad (10)$$

Para todos os modelos, η_{∞} possui valor entre -2 e -5 (quando a viscosidade é representada em Pa.s). As outras constantes precisam ser encontradas para cada material, a partir de medidas experimentais, para encontrar a equação completa. Com os parâmetros de viscosidade é possível extrapolar o comportamento viscoso para todas as temperaturas (incluindo temperaturas inacessíveis experimentalmente).

3.1.5 Fabricação de vidros

A técnica mais aplicada comercialmente para a fabricação de vidros é a fusão seguida de resfriamento. Entretanto, existem outras técnicas de fabricação de vidros, como deposição química de vapor e sol-gel [45,46], e estas rotas de fabricação podem alterar o valor de certas propriedades dos materiais vítreos quando comparados a fusão e resfriamento. A base de dados SciGlass® que foi utilizada neste trabalho (que será apresentada na seção de materiais e métodos) pode conter composições sintetizadas por essas rotas não usuais. Porém, elas são minoria frente ao número de dados que utilizaram a rota usual de fusão e resfriamento.

3.2 Aprendizado de máquina

Aprendizado de máquina (do inglês, *machine learning*) é um campo da ciência da computação que estuda e constrói algoritmos capazes de aprender padrões e fazer previsões a partir de um conjunto de dados. Uma definição mais formal é dada por Mitchell [47], que diz que um programa de computador consegue aprender pela experiência E, executando uma tarefa T_a com uma performance P, se a performance P na tarefa T_a melhora com a experiência E.

Ou seja, o algoritmo consegue melhorar a sua performance em uma determinada tarefa conforme ele executa essa tarefa e ganha experiência.

Existem diferentes maneiras pelas quais um algoritmo pode aprender, de acordo com os dados que são fornecidos inicialmente. Os principais tipos de aprendizado são [11]:

1. **Aprendizado supervisionado:** os dados de entrada são fornecidos junto com os dados de saída desejados, assim, a máquina é capaz de aprender comparando a diferença entre os valores gerados por ela com o de saída, visando sempre diminuir o erro de predição;
2. **Aprendizado não supervisionado:** não é fornecido qualquer dado de saída, apenas de entrada, assim, a máquina é livre para encontrar uma estrutura de agrupamento com base apenas nas características dos dados de entrada fornecidos;
3. **Aprendizado por reforço:** o programa deve desempenhar um determinado objetivo, sendo que os acertos são reforçados por meio de recompensas e os erros recebem punições.

Alguns autores ainda definem também o aprendizado semi supervisionado, que fornece, ao mesmo tempo, dados de entrada com e sem seus respectivos dados de saída, sendo utilizado em casos onde os dados disponíveis são incompletos [48].

Algoritmos do tipo aprendizado de máquina são importantes nos casos de classificação de dados (onde as classes são previamente conhecidas), *clustering* (diferente da classificação, as classes não são previamente conhecidas), problemas de regressão e estimativa de probabilidade. Além disso, podem ser combinadas com outros tipos de algoritmos, por exemplo, algoritmos genéticos, em busca de melhores resultados a depender do problema em questão [49].

Entre os mais comuns tipos de algoritmos, temos máquinas de vetores de suporte (MVS), árvores de decisão (AD) e redes neurais artificiais (RNA). Cada uma deve ser aplicada de acordo com o resultado que se deseja obter. MVSs possuem aprendizado do tipo supervisionado e são capazes de prever em qual

das duas classes um novo dado de entrada (também chamado de exemplo) pertence. ADs, também um algoritmo do tipo supervisionado, utilizam a árvore de decisão como modelo para prever os dados de saída. As RNA serão abordadas na próxima sessão deste texto.

3.2.1 Redes neurais artificiais

As redes neurais artificiais são inspiradas no sistema nervoso de seres vivos, capazes de adquirir conhecimento, possuindo aprendizado do tipo supervisionado ou não supervisionado, a depender da maneira como é construído o algoritmo. É composto por unidades de processamento chamados neurônios artificiais que tentam simular o funcionamento de um neurônio. O funcionamento de um neurônio artificial ilustrado pela Figura 6 começa pelo recebimento dos sinais de entrada (x_1, x_2, \dots, x_n), aos quais serão atribuídos pesos de sinapse de acordo com sua relevância (w_1, w_2, \dots, w_n). Então, todas as entradas, com seus respectivos pesos, são somadas por uma função agregadora (Σ) e o valor obtido é comparado com o limiar de ativação (Θ , também conhecido como viés do neurônio). A diferença entre Σ e Θ gera o potencial de ativação (u), que se positivo emite um sinal excitante, se negativo um sinal inibidor. Então aplica-se uma função de ativação $g(\cdot)$ que limita o valor de saída (y).

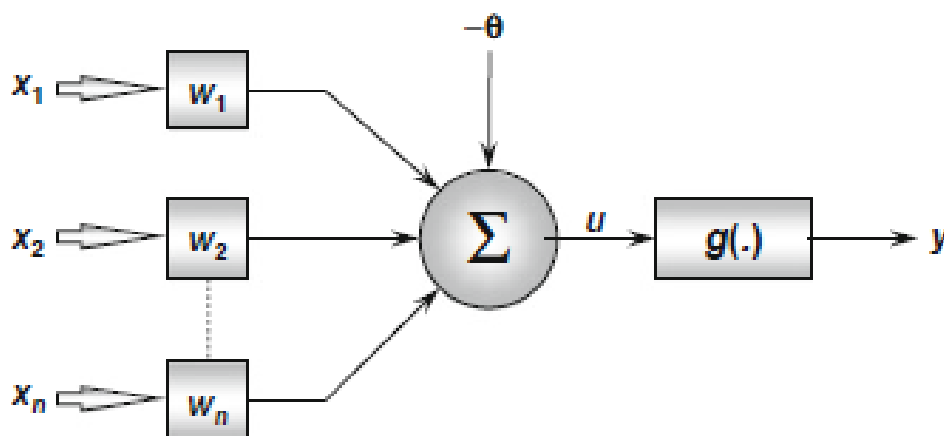


Figura 6: Representação de um neurônio artificial. Reproduzido de [11].

A arquitetura ou topologia da rede neural diz respeito a como os neurônios estão organizados no espaço. Basicamente, os neurônios se organizam em camadas de um dos três tipos: camada de entrada, camada intermediária (ou oculta) e a camada de saída. A camada de entrada é a responsável por receber os exemplos e normalizar os valores dentro do limite que é previsto pela função de ativação. A camada intermediária e a de saída são compostas por neurônios e são responsáveis, respectivamente, por aprender com os valores de entrada e gerar os valores de saída. O número de camadas varia de acordo com o tipo de algoritmo de rede neural, por exemplo, redes perceptron possuem apenas uma camada de entrada e uma camada intermediária, que também atua como camada de saída, enquanto que a perceptron de multicamadas apresenta pelo menos uma camada intermediária, além da camada de saída. Neste trabalho, será utilizada a rede perceptron de multicamadas com duas camadas intermediárias, que está ilustrada pela Figura 7.

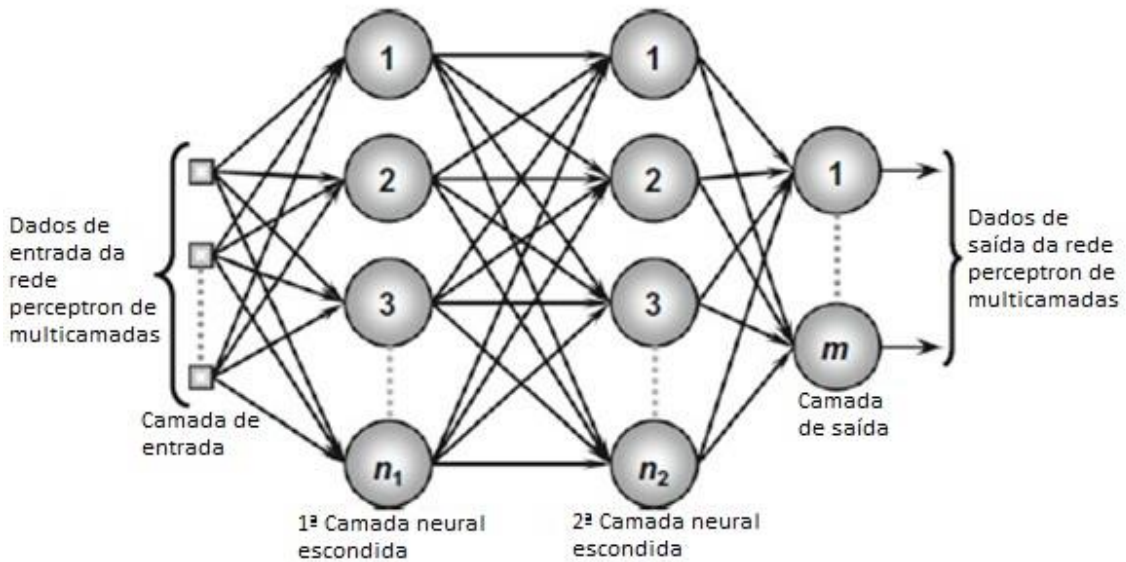


Figura 7: Arquitetura de uma rede perceptron de multicamadas. Adaptado de [11].

O primeiro passo antes de qualquer tentativa de obter um modelo preditivo através de aprendizado de máquina é chamado de *holdout*. Trata-se do processo

de separar parte dos dados que serão reservados e não participarão do processo de treino. Neste trabalho optamos por separar 20% dos dados seguindo este procedimento. Estes dados são de enorme importância, uma vez que em nenhum momento os algoritmos terão acesso a eles. Esta é uma forma de testar como que o modelo treinado se comporta em uma situação de previsão de dados jamais vistos. Espera-se que um bom modelo seja capaz de generalizar os padrões observados e prever com razoável precisão estes dados reservados.

O segundo passo consiste em particionar novamente os dados de treino de maneira para que seja possível medir a capacidade da rede de prever os dados durante os cálculos. Desta maneira, foi utilizada uma partição de dados, onde os dados de entrada da rede são subdivididos em 10 grupos de tamanho similar e de maneira aleatória. Oito grupos são utilizados nas sessões de treino, um na sessão de validação e um na sessão de teste. Os valores do grupo de treino são utilizados pela rede para ajustar os pesos e limiares de ativação a cada sessão de treino. Ao ver todos os dados de treino, a rede calcula, utilizando os mesmos parâmetros, o valor de saída para os dados de validação. Ao comparar os erros do grupo de treino e validação, é possível ver se a rede está realmente melhorando seu modelo, quando os dois erros diminuem, ou se está apenas sobreajustando os dados do grupo de teste, quando o erro do grupo de validação aumenta enquanto de treino diminui. A Figura 8 representa graficamente os dois erros.

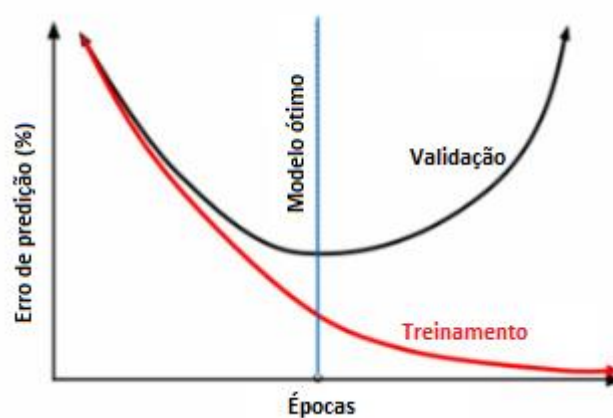


Figura 8: Erro de predição para dados de treinamento e validação ao longo do treino.

Certos parâmetros de uma rede neural devem ser definidos antes do processo de treino. Estes parâmetros são conhecidos como hiperparâmetros e eles não se alteram durante a fase de treino. Estes podem ser ajustados manualmente ou com algoritmos de otimização de hiperparâmetros. A vantagem deste último é a automatização deste processo. Os principais hiperparâmetros de redes neurais são:

1. Número de neurônios nas camadas ocultas;
2. Porcentagem de *dropout* (abandono) de cada camada: a porcentagem de *dropout* corresponde ao número de neurônios que serão desligados durante uma sessão de treinamento em cada época, ou seja, que recebem e armazenam informação, mas não enviam o sinal adiante.
3. Tamanho do lote utilizado em cada sessão de treinamento
4. Paciência: número de épocas nas quais o erro do grupo de validação não diminui

3.2.2 Sobreajuste e subajuste

Um dos problemas mais recorrentes citados na literatura é sobreajuste (*overfitting*) e subajuste (*underfitting*) dos dados. Enquanto o subajuste é caracterizado pela incapacidade da rede de identificar o padrão e extrapolar, o sobreajuste é o extremo oposto, ou seja, a rede se ajusta tanto aos dados que modela até mesmo ruídos e erros presentes neles. Desta maneira, não há uma boa capacidade de generalização dos dados de entrada em ambos os casos. A linha contínua na Figura 9 ilustra um caso de sobreajuste para uma função $f(x)$.

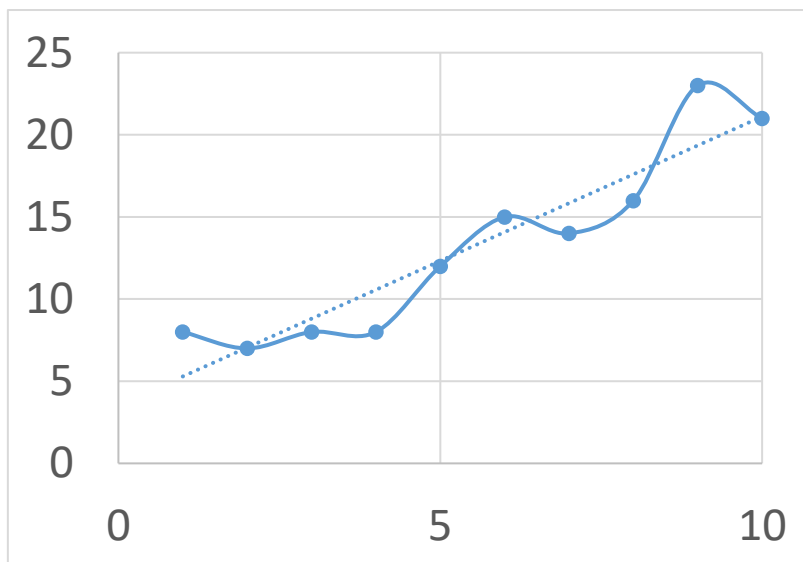


Figura 9: Exemplo de sobreajuste dos dados (representados por pontos). A linha cheia é o modelo que sofre sobreajuste, enquanto a linha tracejada é o modelo ideal.

O valor do erro para uma situação de sobreajuste tende a zero, enquanto que, para subajuste, é mais alto que no caso de um ajuste correto. Isto porque, um dos casos mais comuns de subajuste é utilizar uma função inadequada para representar os dados, por exemplo, uma função linear para representar dados de uma função logarítmica.

O caso mais comum nas redes neurais é o sobreajuste. Isto porque, sem mecanismos corretos de interrupção do treinamento da rede, ela vai continuar executando cálculos até que o erro tenda a zero. Para evitar, é necessário acompanhar os erros de previsão tanto dos dados do grupo de treino quanto de validação. Conforme a rede ajusta seus parâmetros, o erro de ambos os grupos tende a diminuir por melhorar a previsão da rede, até o momento em que o erro do grupo de validação não diminui mais, ou passa a aumentar, enquanto o erro do grupo de treino continuar a cair (Figura 9). Neste momento, fica claro que a rede está tendo sobreajuste nos dados de treino [48].

A paciência é um importante hiperparâmetro para prevenir isto, encerrando o treino. Uma paciência com valor alto aumenta tempo computacional e pode promover o sobreajuste, enquanto que uma paciência

com valor baixo pode encerrar a rotina de treino antes de chegar na melhor combinação de pesos de sinapse e limiar de ativação. Além disso, a porcentagem de *dropout* é um outro parâmetro para evitar com que a rede apresente sobreajuste. Ao desligar alguns neurônios, de maneira aleatória, obriga a rede a dividir melhor a tarefa de identificar padrões entre os neurônios [50] e a torna sua previsão mais robusta.

3.2.3 Aplicações

As redes neurais artificiais e outros algoritmos de aprendizado de máquina tem cada vez mais ganhado destaque em pesquisas de diferentes áreas do conhecimento e já tem sua eficiência comprovada para diversas aplicações como identificar objetos [12] e reconhecimento facial [13]. Mesmo na área de Ciência de Materiais já existem alguns trabalhos que obtiveram sucesso em prever propriedades e abrem possibilidade para novas pesquisas ampliando mais ainda a aplicação delas na área. Como por exemplo o trabalho de Chen *et al.* [14], no qual utilizaram RNA para prever a temperatura de transição vítrea para polímeros. Como o foco deste trabalho é para composições de óxidos formadoras de vidros e nas propriedades de temperatura *liquidus* e viscosidade, segue uma revisão mais detalhada sobre redes neurais artificiais focados em vidros óxidos e suas propriedades de interesse.

Dreyfus *et al.* [15] utilizaram redes neurais artificias para prever temperatura *liquidus* de composições de óxidos formadores de vidros. Em seu trabalho ele utilizou uma rede do tipo perceptron multicamadas, variando o número de neurônios nas camadas intermediárias de 0 a 8, para quatro diferentes grupos de dados. Cada grupo foi definido como dados de um determinado sistema: $\text{Al}_2\text{O}_3\text{-K}_2\text{O-SiO}_2$ (384 dados), $\text{CaO-K}_2\text{O-SiO}_2$ (155 dados), $\text{Na}_2\text{O-CaO-Al}_2\text{O}_3\text{-SiO}_2$ (893 dados) e $\text{Na}_2\text{O-CaO-Al}_2\text{O}_3\text{-MgO-SiO}_2$ (309 dados) e todos os dados foram levantados a partir do SciGlass®. A rede foi capaz de estimar bem a temperatura *liquidus* para esses sistemas, com erro de até 160 K. Outra conclusão deste trabalho foi que quanto maior o número de neurônios nas camadas intermediárias, melhor foi a previsão da rede.

Brauer *et al.* [16] em 2007 modelaram a solubilidade de vidros do sistema P_2O_5 -CaO-MgO- Na_2O - TiO_2 a partir de dados de 31 vidros feitos pelos próprios autores com uma rede neural artificial de uma camada intermediária. A rede foi capaz de prever os dados, com um excelente coeficiente de correlação de 0,9996.

Cai *et al.* [51,52,53] trabalharam com redes neurais artificiais com funções de ativação de base radial (radial basis function neural network – RBFANN) para prever a temperatura de transição vítrea reduzida (T_{gr}) [51] (representada pela Equação 11), região de líquido super-resfriado (ΔT_x) [52] (representada pela Equação 12) e taxa de resfriamento crítico (R_c) [53] para ligas metálicas formadoras de vidro. As três propriedades escolhidas pelo autor estão relacionadas com a habilidade de vitrificação. As principais diferenças entre a rede utilizada por esses autores e a rede perceptron multicamadas são que a primeira apresenta uma única camada intermediária e utiliza como função de ativação a função Gaussiana [54].

$$T_{gr} = \frac{T_g}{T_{liq}} \quad (11)$$

$$\Delta T_x = T_x - T_g \quad (12)$$

Para prever T_{gr} , eles utilizaram dados da literatura (porém não informaram a quantidade de dados) como dados de entrada. Já no caso do ΔT_x , utilizaram onze diferentes propriedades como dados de entrada, que são fortemente correlacionadas com ΔT_x , como entropia de mistura, raio atômico médio, eletronegatividade média, temperatura de fusão, entre outros, além do valor de saída desejado, o ΔT_x . Por fim, para a R_c , de maneira análoga a ΔT_x , foram escolhidos 13 parâmetro (sendo alguns iguais ao de ΔT_x). Não divulgaram mais detalhes a respeito da coleta e do conjunto de dados. A rede neural artificial foi capaz de prever os dados com boa acurácia para as três propriedades, sendo ainda capaz de identificar quais parâmetros têm maior relevância.

Mauro *et al.* [38] em 2016 utilizaram rede neural artificial para prever a temperatura *liquidus* de silicatos. Os dados utilizados foram 851 composições com até oito componentes, todas formadoras de vidro. Então, eles encontraram o melhor design da rede (número de camadas intermediárias, número de neurônios e função de ativação) e, em seguida, treinaram uma rede com tal design. O design selecionado para a segunda etapa foi o que apresentou o menor desvio padrão dos dados de validação. Para treinar a rede neural, os dados foram divididos em dois grupos, sendo 2/3 para treinamento e 1/3 para validação. A rede foi capaz de prever a temperatura *liquidus* com um erro de 10,5 K, um excelente resultado considerando que esta propriedade não varia linearmente com a composição.

Krishnan *et al.* [2] investigaram quatro diferentes algoritmos de aprendizado de máquina, entre elas a rede neural artificial, para prever a dissolução de vidros do sistema $\text{SiO}_2\text{-Al}_2\text{O}_3\text{-Na}_2\text{O}$. Foram utilizados 299 dados a partir da literatura e os dados de entrada para RNA foram a composição do vidro, o valor inicial do pH da solução e o pH da solução após um dado tempo. Como resultado, eles observaram que todas as máquinas de aprendizado utilizadas foram capazes de prever com certa acurácia, porém a RNA exibiu um potencial para extrapolar maior que as demais e foi a melhor entre todas.

Ren *et al.* [55] utilizaram um algoritmo de aprendizado de máquina para acelerar o processo de desenvolvimento de novos vidros metálicos, ou seja, medir a habilidade de vitrificação destes. O algoritmo foi treinado com 5313 composições diferentes que contém 51 elementos. Para realizar o treino parte dos dados foram excluídos para validação. Foram utilizados diferentes algoritmos de aprendizado de máquina e o estudo considerou parâmetros de duas teorias de formação de vidros metálicos, onde uma se baseia na entalpia e entropia do líquido, enquanto a segunda trata sobre formação de cluster de átomos em pequena escala no líquido. O uso de mais parâmetros aumentou a acurácia da rede.

Cassar *et al.* [3] utilizaram redes neurais artificiais para prever a temperatura de transição vítrea de vidros óxidos. Foram utilizados 55150 exemplos de entrada para a rede neural, retirados do SciGlass®. No total,

havia 45 elementos químicos. A RNA previu 95% dos dados com erro menor que 9%, com erro abaixo de 6% para 90% dos dados. Este é um excelente resultado, pois 6% é equivalente a dispersão dos dados de entrada.

Bishnoi *et al.* [4] utilizaram redes neurais e regressão do processo Gaussiano (GPR) para prever módulo de Young de vidros óxidos. GPR é um algoritmo de aprendizado supervisionado que parte da premissa que os dados de entrada apresentam uma distribuição Gaussiana com média e desvio padrão bem definidos. Este método é melhor para grupos pequenos de dados de entrada, como foi o caso que os autores utilizaram. No fim, GPR obteve um melhor desempenho, apresentando resultados de R^2 acima da rede neural, justamente por conta do número de dados de entrada utilizado, visto que a rede neural apresenta bom desempenho para grupos maiores. Outro trabalho que também previu módulo de Young através de diferentes algoritmos (entre eles, RNA) é de Yang *et al.* [5] que utilizou dados obtidos através de dinâmica molecular para vidros óxidos. A RNA apresentou desempenho superior que os demais algoritmos.

Outro estudo que também envolve propriedades elásticas de vidros óxidos e densidade é de Deng [7]. A rede neural se mostra um algoritmo com bom potencial, porém possui desempenho um pouco inferior que floresta aleatória e k vizinhos mais próximos. Todos os três possuem R^2 acima de 0,95.

4 MATERIAS E MÉTODOS

O estudo foi dividido nas etapas descritas no fluxograma (Figura 10): coleta de dados, tratamentos dos dados, design das redes neurais, treinamento das redes neurais e análise dos resultados obtidos. Cada etapa será detalhada a seguir.

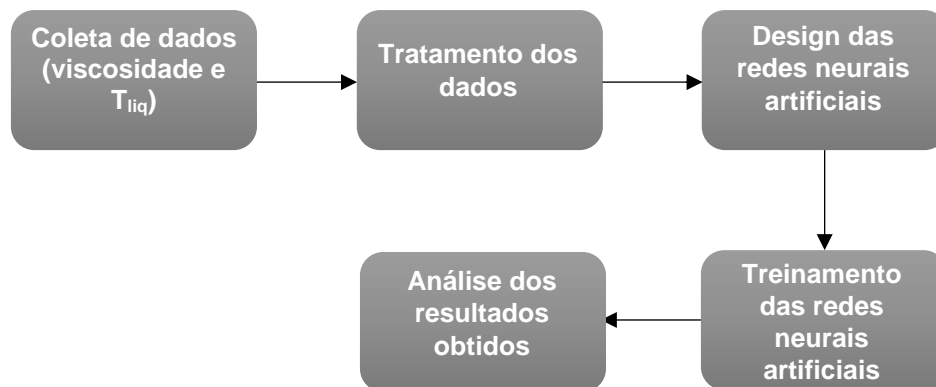


Figura 10: Fluxograma das etapas deste projeto

4.1 Coleta e tratamento de dados

Os dados de temperatura *liquidus* e viscosidade utilizados para treino e teste das redes neurais artificiais foram coletados a partir do banco de dados do SciGlass®, recentemente disponibilizado gratuitamente com uma licença permissiva para uso científico e comercial [56]. Para a temperatura *liquidus*, também foram considerados dados de diagramas de fases da literatura, utilizando o programa Phase Equilibria Diagrams Database Version 4.0 [57] para busca e seleção dos diagramas de fase, tendo preferência por diagramas mais recentes. Nos casos onde existem dois diagramas de fases divergentes, foram considerados ambos

A coleta no SciGlass foi realizada com as seguintes restrições:

- Apenas vidros com fração atômica de oxigênio maior que 0,3 (a fim de limitar a busca apenas para vidros óxidos, eliminando assim vidros calcogenetos e fluoretos);
- Apenas vidros que não contenham os seguintes elementos: S, H, C, Pt, Au, F, Cl, N, Br e I (a fim de evitar elementos que possam ocupar a posição do oxigênio no vidro);

Após a coleta, checamos a frequência de cada elemento químico nos exemplos coletados. Todos os vidros contendo elementos químicos presentes em menos que 1% dos exemplos foram eliminados. Esta regra é para garantir a representatividade de cada elemento químico com um número mínimo de exemplos, para melhorar o treino e a expectativa de generalização do aprendizado da rede.

Todos os vidros disponíveis com valores de T₂, T₃ e T₄ (temperaturas onde a viscosidade é, respectivamente, 10², 10³ e 10⁴ Pa.s) e de temperatura *liquidus* foram coletados seguindo as restrições acima. Para reduzir o erro de previsão da T_{liq}, os dados de binários e óxidos puros do SciGlass foram excluídos e apenas os dados provenientes de diagramas de fases foram considerados. Para composições ternárias ou mais óxidos, os dados são todos do SciGlass.

Todas as composições (em fração atômica) foram arredondadas para três casas decimais. O descritor químico obtido durante a coleta foi a fração atômica de cada elemento químico que compõe os vidros e não porcentagem de óxidos, o que seria mais comum. O motivo desta estratégia foi a redução da dimensionalidade do problema, pois existem muito mais óxidos possíveis do que elementos químicos, devido a diferentes valências dos cátions. Com a redução da dimensionalidade do problema, espera-se uma melhor capacidade de generalização da rede neural artificial. Por fim, dados duplicados foram removidos e substituídos pelo valor da mediana da propriedade.

4.2 Design e treinamento da rede neural

Todo o processo computacional deste trabalho foi realizado em ambiente de programação Python (versão 3.7) utilizando uma máquina com processador Ryzen 7 2700X de 16Gb de RAM. Para escrever os programas, foram utilizados módulos, que são arquivos de código Python que podem ser importados com as suas funções já escritas e prontas para uso. Os módulos escolhidos foram: numpy para álgebra linear, pandas para gerenciamento dos dados, tensorflow para criar e treinar a RNA, hyperopt para otimizar os hiperparâmetros, matplotlib para elaboração de gráficos ao final da rotina.

Durante o processo de otimização de hiperparâmetros, 200 topologias de rede foram testadas para a previsão de T2, T3 e T4 e 75 topologias foram testadas para a previsão da temperatura *liquidus* (o número menor de topologias para temperatura *liquidus* é um reflexo do custo computacional mais elevado por conta da maior quantidade de dados). Utilizou-se de segmentação *holdout* com 80% dos dados selecionados aleatoriamente para compor o grupo de treino e 20% restantes selecionados para o teste final do modelo.

Durante o treino em si é feita uma partição local dos dados em treino, validação e teste. Importante deixar claro que 10% das composições do grupo de treino são reservadas para compor o grupo de validação e 10% são reservadas para compor um grupo de teste local (observe que esta partição de teste local não é a partição de dados reservados pela rotina de holdout, descrita no parágrafo anterior). Dados no grupo de validação não são utilizados para alterar os parâmetros da rede durante o treino, no entanto servem como “termômetro” para observar e evitar um problema bastante comum em redes neurais, o sobreajuste (descrito na Seção 3.2.2 na página 25). Em outras palavras, o grupo de validação serve para encerrar o processo de treino. Após o processo de treino ter sido encerrado, o erro da predição do grupo de teste local é computado (novamente, não confundir este grupo de teste local com os 20% de exemplos que foram reservados na rotina de holdout).

A melhor topologia (isto é, com o menor erro de previsão do grupo de teste local) obtida na seção anterior foi selecionada e utilizada para o treino da rede neural final, onde 100% dos dados foram utilizados para treino.

O parâmetro R^2 foi calculado através da Equação 13 [58,59]. Este parâmetro reflete a fração da variância da previsão da RNA que é explicada pelo modelo, sendo que este varia de 0 a 1, onde 1 é o ajuste perfeito. Esta equação se trata do R^2 para quando a hipótese testada é que existe uma correlação linear entre x e y seguindo a equação $y = ax + \varepsilon$ (onde ε é o erro irreduzível, o ruído de sua medida, enquanto que o a é um parâmetro de ajuste).

$$R^2 = 1 - \frac{\sum_i (y_i - \hat{y}_i)^2}{\sum_i y_i^2} \quad (13)$$

4.3 Análise dos dados obtidos

Com base nos gráficos gerados após o treinamento da rede neural artificial, com os erros de previsão para composições com diferentes números de elementos químicos e com o parâmetro R^2 do ajuste feito, foi avaliado a capacidade de previsão da rede neural artificial para cada propriedade. Gráficos com as propriedades foram estudados a fim de avaliar a qualidade da previsão da rede neural, bem como os erros por elemento químico e temperatura.

5 RESULTADOS E DISCUSSÃO

5.1 Temperatura *liquidus*

Foram utilizados 30.127 dados de composição química versus temperatura *liquidus* para o treinamento da rede neural artificial. Os dados são compostos por 37 elementos químicos diferentes, entre eles, estão o oxigênio, os principais elementos formadores (Si, Ge, B, P, Te e As, Bi, Sb), o alumínio, zircônio, titânio, ítrio (intermediários), oxigênio, elementos modificadores de rede da família 1 e 2 (Li, Na, K, Rb, Cs, Mg, Ca, Sr, Ba) e outros elementos químicos (Ce, Cr, Fe, Gd, La, Mn, Mo, Nb, Nd, Pb, Sn, Ta, V, W e Zn).

As Figuras 11a e 11b mostram uma análise estatística dos dados de entrada para a rede neural, com a frequência do número de diferentes elementos químicos por composição e a frequência de valores de T_{liq} , respectivamente. Dois elementos químicos representam um óxido puro, como SiO_2 por exemplo, e três elementos químicos representam um binário. Observa-se que são mais comuns composições ternárias, entretanto, há dados consideráveis entre composições binárias até composições com 13 diferentes óxidos. A partir de 13 óxidos, a frequência diminui e não há composições com 17 ou mais diferentes óxidos. Já a respeito da distribuição dos valores de T_{liq} , temos uma distribuição cuja moda é 1400 K e os valores variam entre 500 e 3000 K.

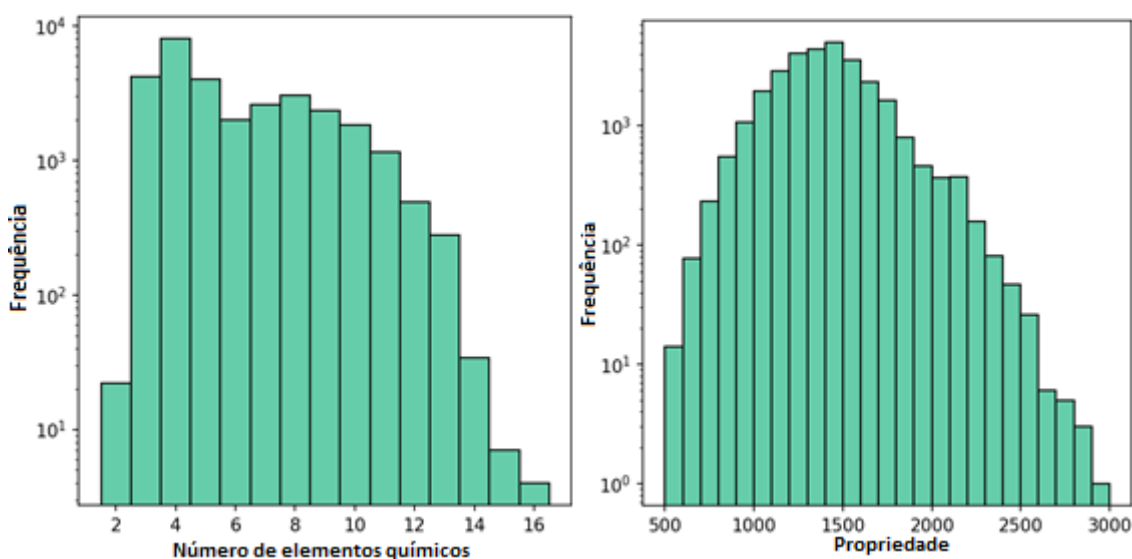


Figura 11: Frequência de (a) número de diferentes elementos químicos por composição e (b) valores de T_{liq} . O eixo y está representado em escala logarítmica.

O desempenho da rede neural está resumido na Figura 12. O R^2 da previsão do modelo foi de 0,997, o que é muito bom, pois o valor está próximo a 1, que significa que o modelo foi capaz de capturar grande parte da variabilidade dos dados. Porém, ao analisar o RD (desvio relativo médio), é possível ver que há um erro médio de 3,7% para os dados que a rede calcula. E, por ser médio, tal erro pode ser maior para casos específicos, como determinados elementos químicos ou faixas de valores da T_{liq} . Então, é preciso avaliar mais a fundo as origens do erro para tentar compreender melhor qual a capacidade de previsão da RNA para este caso.

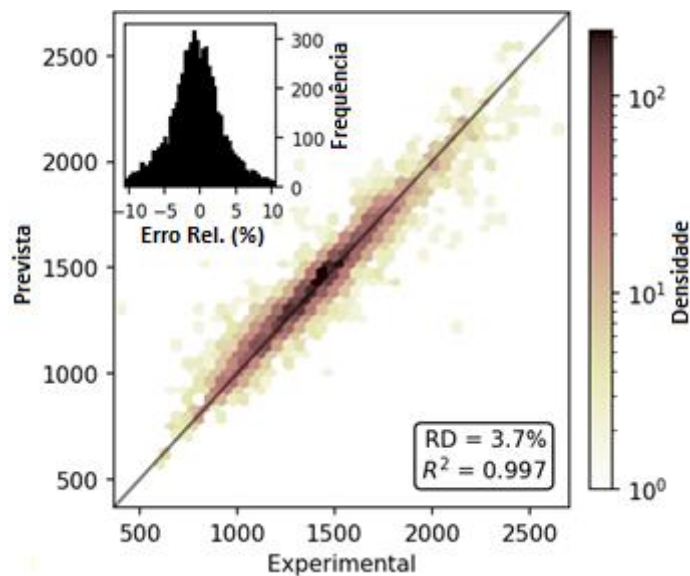


Figura 12: T_{liq} prevista pela rede neural versus valores reportados na literatura. A linha preta é a reta identidade que indica o que seria a previsão perfeita. O gráfico no canto superior mostra a distribuição de erros relativos (ϕ).

A Figura 13 mostra o resíduo da predição para a T_{liq} , em função do valor da temperatura *liquidus*. Nota-se que os resíduos são maiores para valores extremos de temperatura e menores para temperaturas intermediárias. Existe uma tendência crescente do valor mediano dos resíduos em função da temperatura *liquidus*. Valores negativos dos resíduos significam valores onde a previsão de T_{liq} é maior que o valor real reportado; vice-versa. A tendência crescente do resíduo pode indicar um certo vício da rede neural ao tentar atenuar temperaturas extremas, como eutéticos e pontos de máxima temperatura. Isto será verificado adiante ao plotar os dados de um determinado sistema comparados com a respectiva *liquidus* proveniente de um diagrama de fases da literatura.

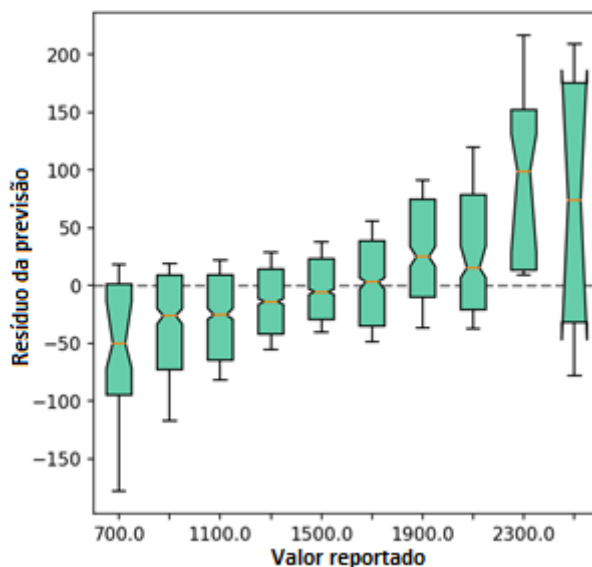


Figura 13: Distribuição dos valores do resíduo da previsão da temperatura *liquidus* (em Kelvin) em função do valor reportado da *liquidus*. As caixas compreendem os valores dos quartis Q1 e Q3 e a barra de erro (em preto) compreende 67% dos dados (um desvio padrão). A mediana é ilustrada por uma reta laranja horizontal e seu intervalo de confiança (95%) é indicado pelos entalhos nas caixas. Apenas os dados reservados pela rotina de *holdout* foram utilizados para construir este gráfico.

A Figura 14 mostra a média do resíduo e seu desvio padrão em função do elemento químico presente no vidro. Este gráfico indica que o resíduo de determinados elementos químicos (Vanádio, Telúrio e Fósforo) tem um desvio padrão maior que o usual. Entretanto, observamos que o desvio padrão do resíduo não está relacionado com o número de composições que contém o elemento químico (os elementos químicos da Figura 16 estão ordenados por número de exemplos no *dataset* reservado pela rotina de *holdout*). O resultado está ligado ao fato de que a rede não conseguiu generalizar bem para estes elementos. Ainda, é uma possibilidade que existam dados com grandes erros (outliers) para esses 3 elementos, sejam eles de qualquer origem, desde medidas até leitura do indivíduo que os pegou de um artigo e os inseriu na base de dados SciGlass. Para refinar a previsão do modelo resultante, é possível excluir estes elementos com maior erro dos dados de entrada, ou melhor, realizar estudos estatísticos para verificar a presença e eliminar apenas os outliers. O resíduo tão alto no caso do Vanádio é um bom indicativo que há outliers em composições que contém este elemento químico.

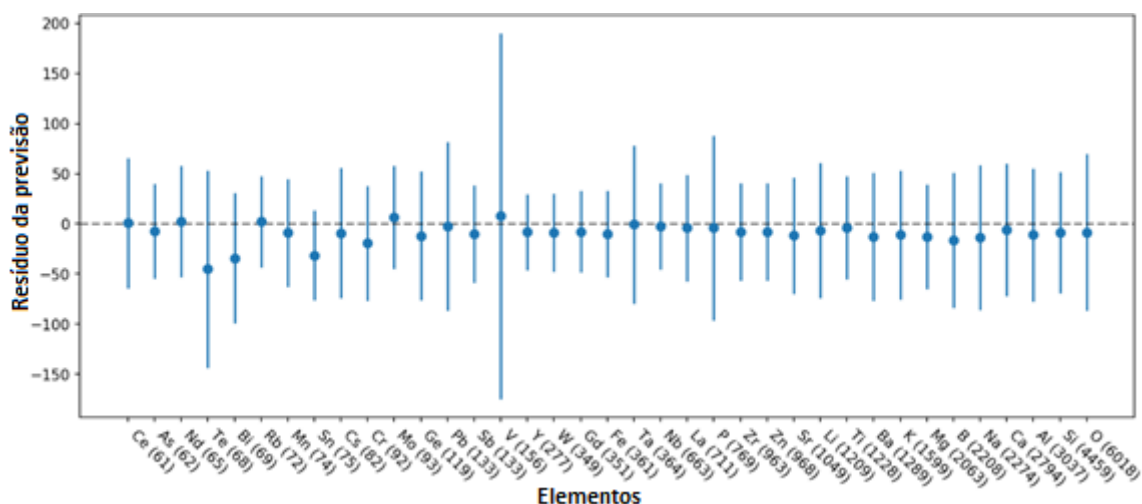


Figura 14: Média e desvio padrão do resíduo da previsão da temperatura *liquidus* (em Kelvin) em função do elemento químico presente na composição. O número em parênteses (ao lado de cada elemento) indica o número de composições que contém o elemento químico e se encontra em ordem crescente. Apenas os

dados reservados pela rotina de *holdout* foram utilizados para construir este gráfico.

Para ilustrar melhor o que está acontecendo com a previsão da RNA, as Figuras 15 e 16 representam a previsão em conjunto com dados de um diagrama de fases, que foi utilizado o treinamento da rede, para os binários $\text{SiO}_2\text{-Na}_2\text{O}$ e $\text{B}_2\text{O}_3\text{-Na}_2\text{O}$. O erro da previsão depende de muitos fatores, e é diferente para cada elemento, conforme discutido anteriormente. Porém, é interessante notar o comportamento da rede neural em pontos eutéticos, onde a temperatura diminui bruscamente. A rede tende a amortecer essa mudança brusca. Desta maneira, para valores mais alto e baixos da temperatura, o erro da previsão é maior e isso reforça a ideia de que o erro apresentado na Figura 14 é maior para temperaturas extremas.

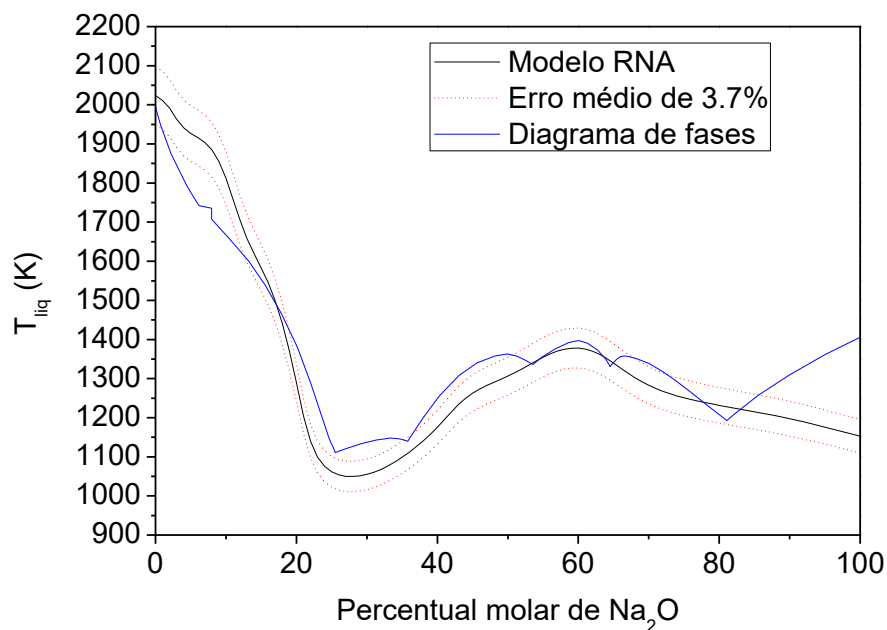


Figura 15: Diagrama de fases versus previsão da rede neural artificial do sistema binário $\text{SiO}_2\text{-Na}_2\text{O}$.

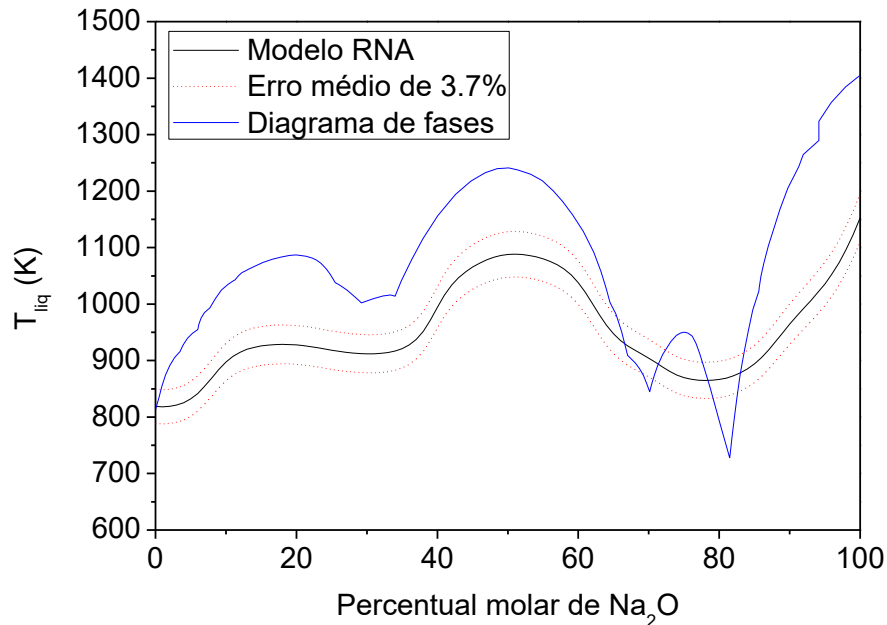


Figura 16: Diagrama de fases versus previsão da rede neural artificial do sistema binário $B_2O_3-Na_2O$.

Em sistemas binários, a rede neural consegue fazer uma previsão da temperatura *liquidus* sem conseguir capturar todas as suas nuances, como eutéticos e pontos de máxima temperatura. É um início, porém, necessita de refinamento para conseguir prever de maneira mais precisa os valores de T_{liq} . Obviamente, ao tentar prever uma composição com uma grande quantidade de elementos químicos diferentes, como 16, ela consegue prever melhor que um palpite de um profissional na área, pois há muitas interações que nenhuma pessoa poderia correlacionar. Mas, esta medida serviria apenas como base para experimentos no laboratório, por conta de seu erro associado.

Para tentar melhorar o modelo, a alternativa mais simples é realizar o estudo com os dados de entrada. Neste trabalho não foi realizado um estudo estatístico para avaliar a presença de outliers nos dados de T_{liq} . Apenas algumas regras foram criadas e testadas a fim de tentar garantir dados mais representativos (conforme foi detalhado na Seção 4.1).

Outra questão é a respeito de utilizar apenas a composição química como dado de entrada para prever uma propriedade termodinâmica. De uma maneira geral, ela foi sim eficiente, mas talvez por isso ainda apresente um certo erro,

principalmente na questão dos eutéticos. Isto porque outros parâmetros como força de ligação são importantes neste caso. Para tentar melhorar os resultados aqui obtidos, poderia utilizar além da composição química outras propriedades dos elementos químicos como números de elétrons na camada de valência, eletronegatividade e eletropositividade, raio atômico e iônico. Certamente, o problema seria muito mais complexo com mais variáveis, mas talvez a recompensa melhorando o modelo seja válida.

5.2 Viscosidade

É importante notar que diferente da temperatura *liquidus*, os valores de T2, T3 e T4 não foram necessariamente reportados nos artigos, mas certas vezes obtidos por interpolação dos dados feita pelos desenvolvedores do software SciGlass. Esta interpolação traz uma incerteza adicional nos dados. No entanto, não acreditamos que esta incerteza seja muito maior que a própria incerteza das medidas de viscosidade, uma vez que as interpolações foram realizadas apenas dentro do domínio de temperaturas onde existem dados experimentais.

A primeira rede neural de viscosidade foi feita para prever a T2, temperatura na qual a viscosidade do líquido é 10^2 Pa.s. Foram utilizados 13.177 dados de temperatura e composição, o que é aproximadamente um terço do utilizado para a T_{liq} , porém, ainda é um número adequado para treinar uma rede neural artificial. Os dados são compostos por 24 elementos químicos, ou seja, um banco de dados menor e com menos variedade de elementos químicos. Os elementos são desde formadores de rede (P, B, Si, As, Sb), o alumínio, zircônio, titânio e ítrio (elementos intermediários), oxigênio e elementos modificadores de rede como Ba, Ca, Ce, Fe, K, La, Li, Mg, Mn, Na, Pb, Sn, Sr e Zn. Chama-se atenção para a ausência do germânio e telúrio, importantes elementos formadores de vidros.

A Figura 17 apresenta informações acerca dos dados de entrada para a rede neural. A temperatura varia entre 500 até 2600 K para composições com

dois até 16 elementos químicos. Novamente, há uma carência de dados para composições com 17 ou mais elementos químicos.

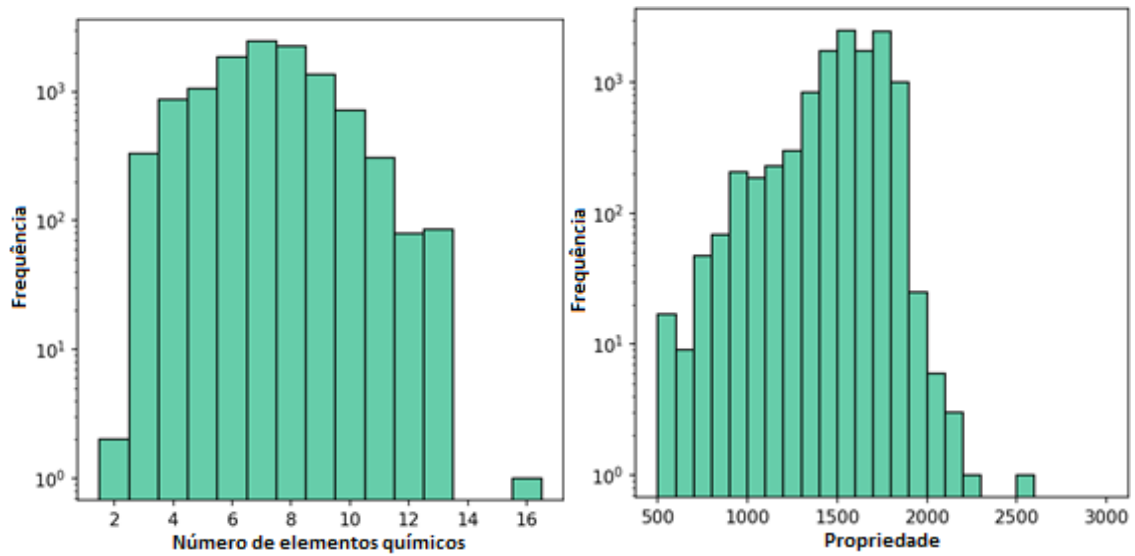


Figura 17: Frequência de (a) número de diferentes elementos químicos por composição e (b) valores de T2. O eixo y está representado em escala logarítmica.

O desempenho da rede neural está resumido na Figura 18. Percebe-se que aqui a rede teve uma melhor capacidade de previsão, obtendo um valor de desvio relativo médio melhor que no caso da T_{liq} , 1,5% contra 3,7%. Ainda, o R^2 é igual a 0,999 o que representa uma forte correlação linear entre dados previstos e os dados experimentais. Então, a capacidade de previsão da rede neural não está diretamente ligada ao número de dados de entrada fornecidos, se não, a rede de T_{liq} seria melhor. Um fato é que o menor número de elementos químicos presentes nos dados de entrada dessa rede neural reduziu o número de variáveis no modelo, o que pode ter minimizado o erro. Além disso, é possível que a qualidade dos dados de viscosidade seja melhor que da temperatura *liquidus*. Outra explicação para melhor precisão da rede de T2 (e das demais

redes de viscosidade) é que a propriedade em si não depende do arranjo cristalino do material, diferentemente de T_{liq} . Esta é uma questão importante, pois a rede neural recebe apenas a fração atômica dos elementos químicos como informação de entrada, e é conhecido que apenas com esta informação é bastante difícil prever características cristalográficas do respectivo material cristalino.

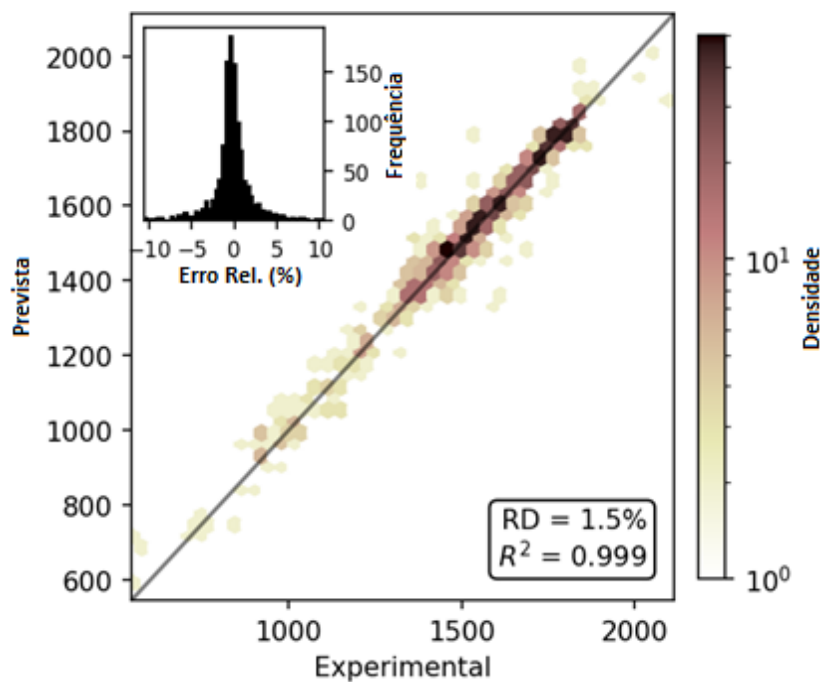


Figura 18: T_2 prevista pela rede neural versus valores reportados na literatura. A linha preta é a reta identidade que indica o que seria a previsão perfeita. O gráfico no canto superior mostra a distribuição de erros relativos (ϕ).

A Figura 19 mostra o resíduo da previsão de T_2 , onde observa-se a tendência de um erro maior para temperaturas abaixo de 1200 K, com exceção para 600 K. Este erro maior se deve a menor quantidade de exemplos para o treino da rede, que pode ser observada na Figura 19b (chamamos a atenção do leitor para o fato que o eixo y desta figura está em escala logarítmica).

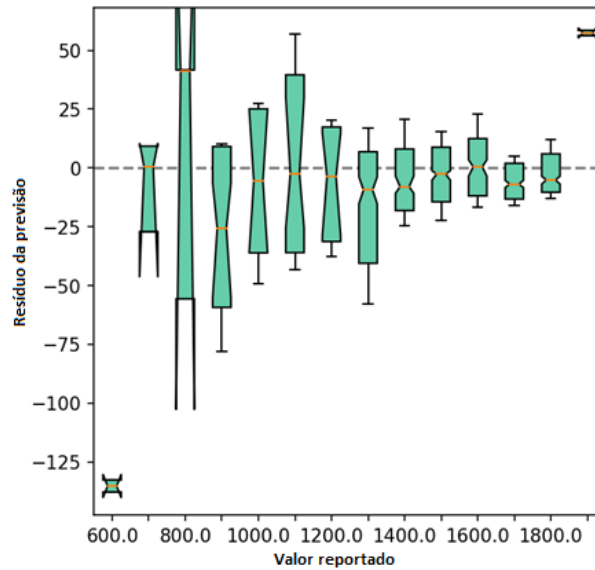


Figura 19: Distribuição dos valores do resíduo da previsão da T2 (em Kelvin) em função do valor reportado. As caixas compreendem os valores dos quartis Q1 e Q3 e a barra de erro (em preto) compreende 67% dos dados (um desvio padrão). A mediana é ilustrada por uma reta laranja horizontal e seu intervalo de confiança (95%) é indicado pelos entalhes nas caixas. Apenas os dados reservados pela rotina de *holdout* foram utilizados para construir este gráfico.

A Figura 20 mostra o resíduo médio e seu respectivo desvio padrão para vidros contendo os diferentes elementos químicos presentes no *dataset* reservado durante o *holdout*. Quanto aos elementos químicos, igual na rede anterior, não se observa uma relação com o número de exemplos que contém o elemento. O erro é maior para o elemento ferro (Fe), mas é próximo aos erros dos demais elementos, de maneira geral.

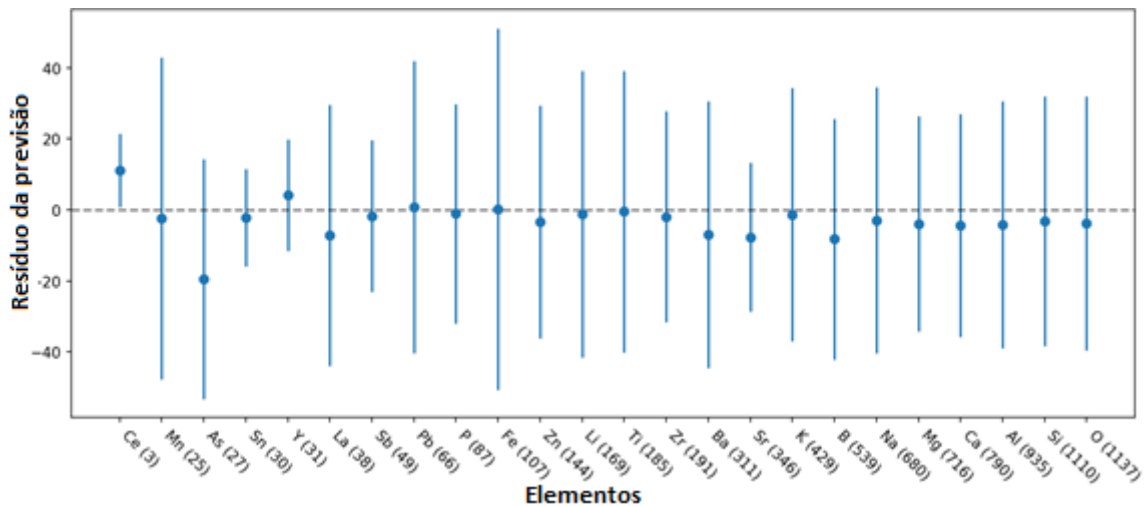


Figura 20: Média e desvio padrão do resíduo da previsão da T2 (em Kelvin) em função do elemento químico presente na composição. O número em parêntesis (ao lado de cada elemento) indica o número de composições que contém o elemento químico, e se encontra em ordem crescente. Apenas os dados reservados pela rotina de *holdout* foram utilizados para construir este gráfico.

A segunda rede de viscosidade foi feita para encontrar em qual temperatura para uma dada composição a viscosidade tem o valor de 10^3 Pa.s, sendo essa temperatura chamada T3, também conhecida como temperatura de trabalho (*working point*, em inglês). Foram utilizados 11.817 dados de temperatura e composição, um terço do utilizado para a T_{liq} , porém, ainda adequado para treinar uma rede neural. Os dados são compostos também por 24 elementos químicos, os mesmos elementos para a T2.

A Figura 21 apresenta informações acerca dos dados de entrada para a rede neural. A temperatura varia entre 500 até 2600 K nos dados de entrada para composições com dois até 16 elementos químicos. Novamente, há uma carência de dados para composições com 17 ou mais elementos químicos.

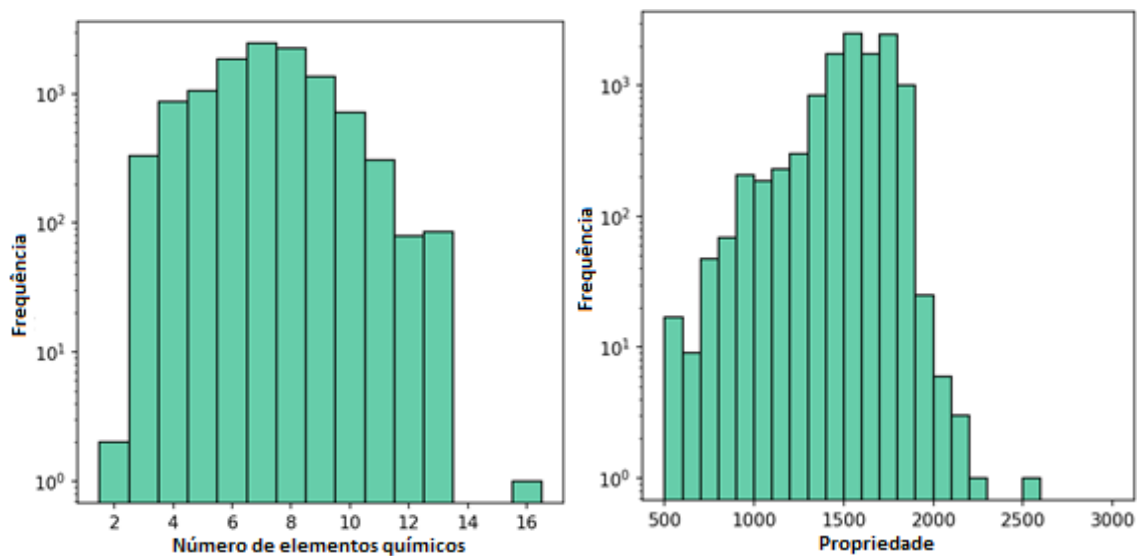


Figura 21: Frequência de (a) número de diferentes elementos químicos por composição e (b) valores de temperatura T3. O eixo y está representado em escala logarítmica.

O desempenho da rede neural está resumido na Figura 22. Tal como a rede neural para a T2, a rede neural para T3 também teve um excelente desempenho, tendo um desvio relativo médio de 1,7%, e com R^2 de 0,999, o que era esperado por possuir mesma ordem de números de exemplos e elementos químicos praticamente iguais.

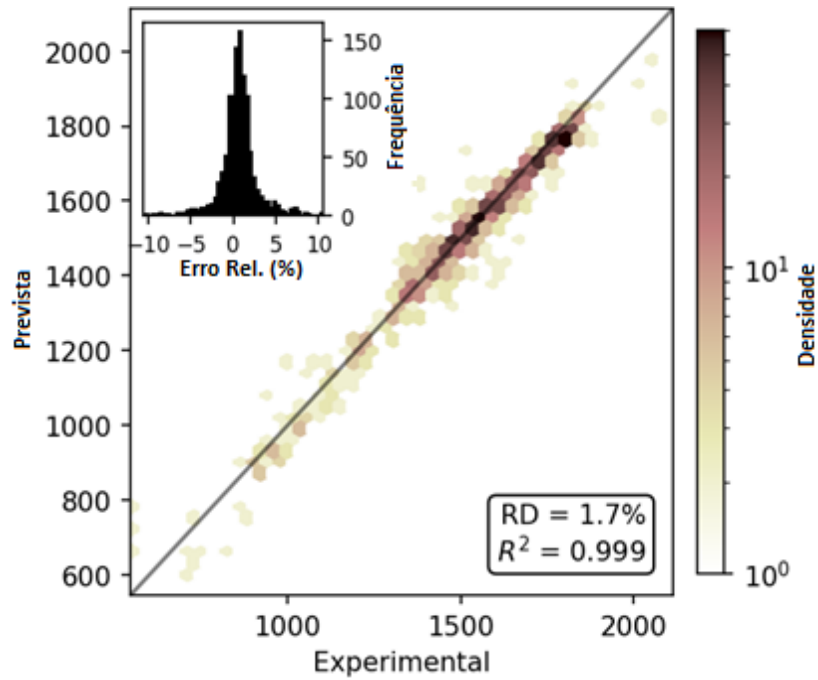


Figura 22: T3 prevista pela rede neural versus valores reportados na literatura. A linha preta é a reta identidade que indica o que seria a previsão perfeita. O gráfico no canto superior mostra a distribuição de erros relativos (ϕ).

A Figura 23 mostra o resíduo da previsão de T3, que também apresenta um erro maior para temperaturas abaixo de 1000 K. Quanto aos elementos químicos (Figura 24), igual na rede anterior, não se observa uma relação com o número de exemplos que contém o elemento. O erro é maior para o ferro e manganês; para os demais elementos é menor e não varia muito.

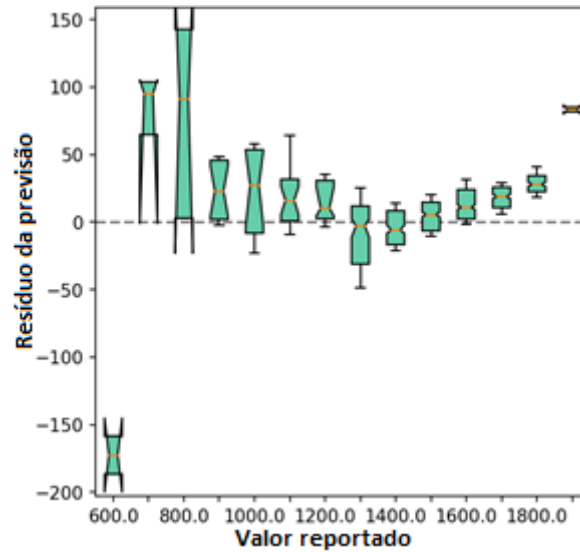


Figura 23: Distribuição dos valores do resíduo da previsão da T3 (em Kelvin) em função do valor reportado. As caixas compreendem os valores dos quartis Q1 e Q3 e a barra de erro (em preto) compreende 67% dos dados (um desvio padrão). A mediana é ilustrada por uma reta laranja horizontal e seu intervalo de confiança (95%) é indicado pelos entalhes nas caixas. Apenas os dados reservados pela rotina de holdout foram utilizados para construir este gráfico.

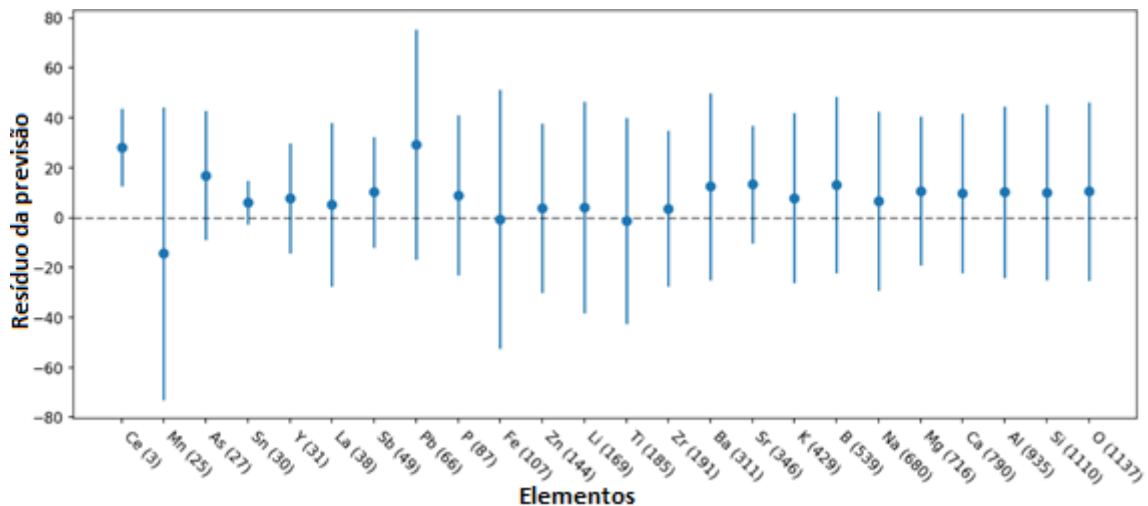


Figura 24: Média e desvio padrão do resíduo da previsão da T3 (em Kelvin) em função do elemento químico presente na composição. O número em parêntesis (ao lado de cada elemento) indica o número de composições que contém o elemento químico e se encontra em ordem crescente. Apenas os dados reservados pela rotina de *holdout* foram utilizados para construir este gráfico.

A terceira e última rede neural deste trabalho é a rede para prever a T4, temperatura cuja viscosidade tem valor de 10^4 Pa.s. Foram utilizados nesta rede 6.764 exemplos, quase metade dos dados das redes de T2 e T3. A respeito dos elementos químicos, temos os mesmos 24 elementos químicos da rede de T2. Assim, comparar o desempenho dessa rede com as outras é importante para ver a influência do número de exemplos. A Figura 25 apresenta informações acerca dos dados de entrada para a rede neural. Os dados são muito próximos das redes anteriores, com mesmo intervalo de temperatura e de números de elementos químicos por composição.

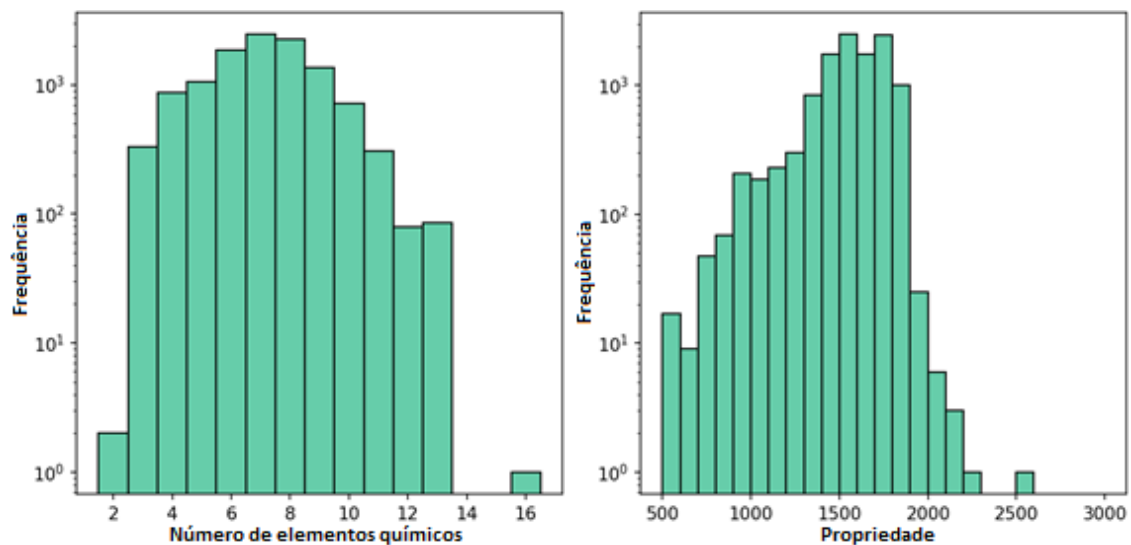


Figura 25: Frequência de (a) número de diferentes elementos químicos por composição e (b) valores de temperatura T4. O eixo y está representado em escala logarítmica.

Mesmo tendo sido treinada com metade dos dados que as demais, o modelo resultante ainda apresentou um desempenho similar, conforme a Figura 26. O desvio relativo médio é de 1,5% e o R^2 é igual a 0,999. Quanto aos erros, o erro em relação a temperatura se repete (Figura 29) abaixo de 1200 K. Isso ocorre, pois, todas as redes têm em comum o baixo número de exemplos para estas temperaturas. Já o erro com respeito ao elemento químico, aqui temos um

erro próximo para a maioria dos elementos, mas que é maior para o ferro (como as duas redes anteriores) e para manganês e chumbo.

A Figura 27 mostra o resíduo da previsão de T4, que também apresenta um erro maior para temperaturas abaixo de 1200 K. Quanto aos elementos químicos (Figura 30), o erro é maior para o elemento ferro (Fe) e chumbo (Pb), e para os demais elementos é menor e não varia muito.

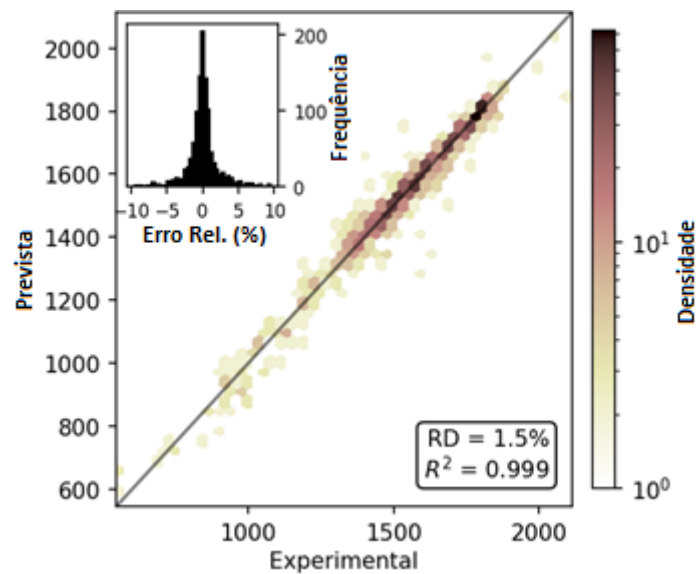


Figura 26: T4 prevista pela rede neural versus valores reportados na literatura. A linha preta é a reta identidade que indica o que seria a previsão perfeita. O gráfico no canto superior mostra a distribuição de erros relativos (ϕ).

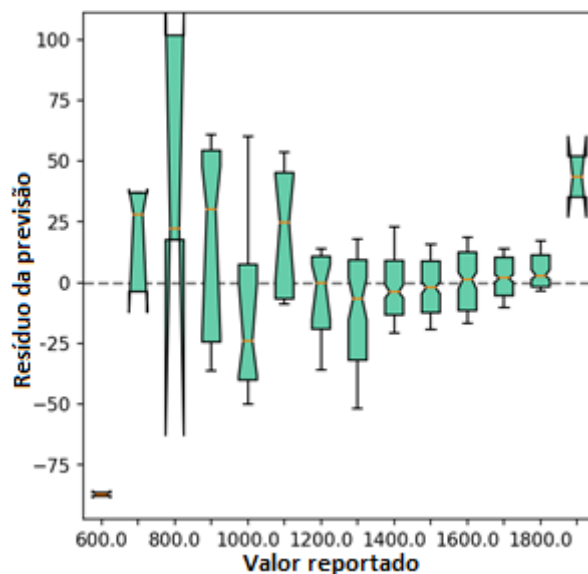


Figura 27: Distribuição dos valores do resíduo da previsão da T4 (em Kelvin) em função do valor reportado. As caixas compreendem os valores dos quartis Q1 e Q3 e a barra de erro (em preto) compreende 67% dos dados (um desvio padrão). A mediana é ilustrada por uma reta laranja horizontal e seu intervalo de confiança (95%) é indicado pelos entalhes nas caixas. Apenas os dados reservados pela rotina de *holdout* foram utilizados para construir este gráfico.

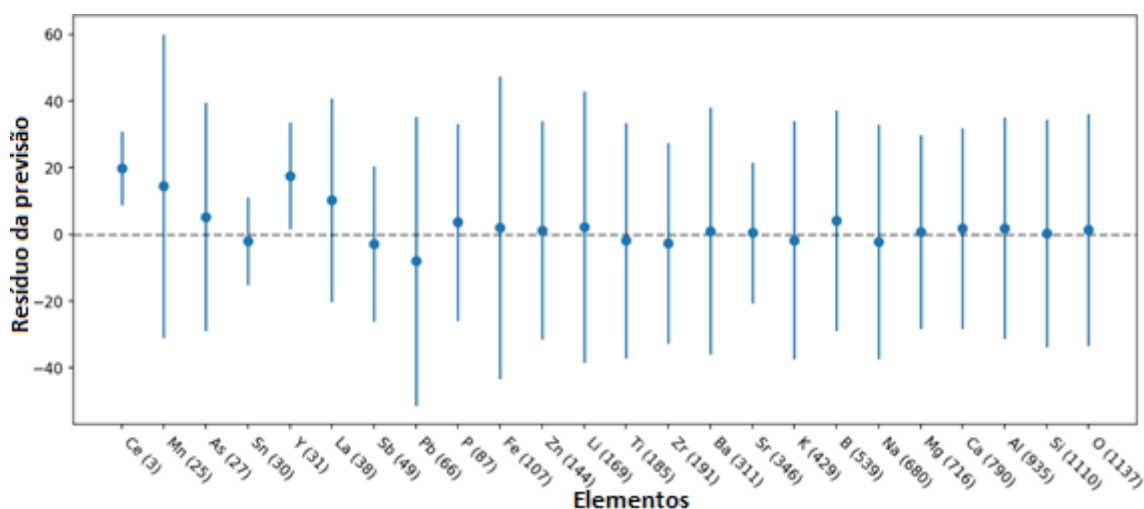


Figura 28: Média e desvio padrão do resíduo da previsão da T4 (em Kelvin) em função do elemento químico presente na composição. O número em parêntesis (ao lado de cada elemento) indica o número de composições que contém o elemento químico e se encontra em ordem crescente. Apenas os dados reservados pela rotina de *holdout* foram utilizados para construir este gráfico.

Como conclusão parcial, a rede neural artificial conseguiu prever muito melhor as temperaturas T2, T3 e T4 do que T_{liq} . O desvio relativo médio e o coeficiente R^2 foram significativamente melhores para as redes de viscosidade que para a *liquidus*.

A respeito do maior erro com o elemento ferro em todas as composições, Osugi *et al.* [60] discutem o efeito dos diferentes estados de oxidação, ou seja, a valência do ferro na viscosidade de sistemas de sílica com álcalis. O estado Fe^{2+} (FeO) se comporta como um modificador de rede, enquanto que o estado Fe^{3+} se comporta como um intermediário, tal qual a alumina. Desta maneira, a viscosidade de líquidos que contêm ferro dependem da razão Fe^{3+}/Fe^{2+} , que por sua vez depende tanto da pressão parcial de oxigênio (quanto maior, maior a razão Fe^{3+}/Fe^{2+}) como da basicidade do sistema (quanto maior a basicidade, maior a razão). Desta maneira, como existem diferentes dados de diferentes sistemas, os dados podem ter uma maior variação intrínseca devido a esses fatores. Como a rede neural só recebe a fração atômica, o erro associado a este elemento acaba aumentando pelo fato de que a rede não tem dados a respeito dessa mudança de valência e deve inferir isso através da composição total do vidro. O mesmo pode acontecer para outros elementos que apresentam mais de um estado de oxidação, como o vanádio, que possui valências +3, +4 e +5.

Para analisar o comportamento dos resultados de viscosidade, gráficos das três temperaturas (T2, T3 e T4) foram plotados em função do teor de Na_2O para os três óxidos formadores de rede: B_2O_3 , P_2O_5 e SiO_2 (Figuras 29, 30 e 31). Com a adição de óxido de sódio, fica evidente a queda de temperatura (para cada uma das 3 viscosidades) com aumento da % de óxido modificador para os três formadores. É possível observar que as linhas se cruzam em dois dos três gráficos. O comportamento esperado (correto) é que elas não se toquem, visto que T2 deve ser sempre maior que T3, que por sua vez é sempre maior que T4. No gráfico de faixa de erros, é possível notar como elas se sobrepõem, principalmente para o gráfico de SiO_2 - Na_2O .

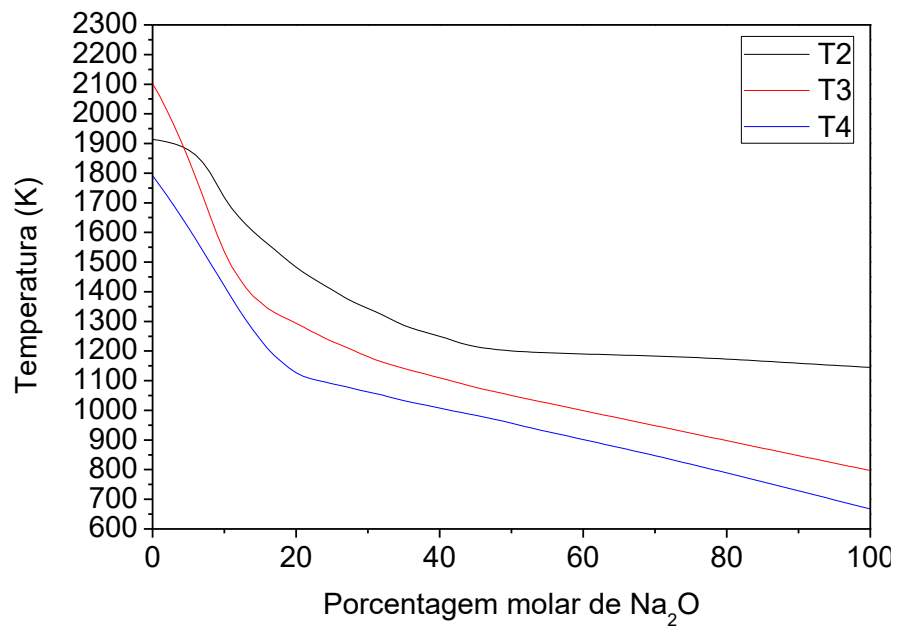


Figura 29: T2, T3 e T4 do binário SiO₂-Na₂O em função da porcentagem molar de Na₂O.

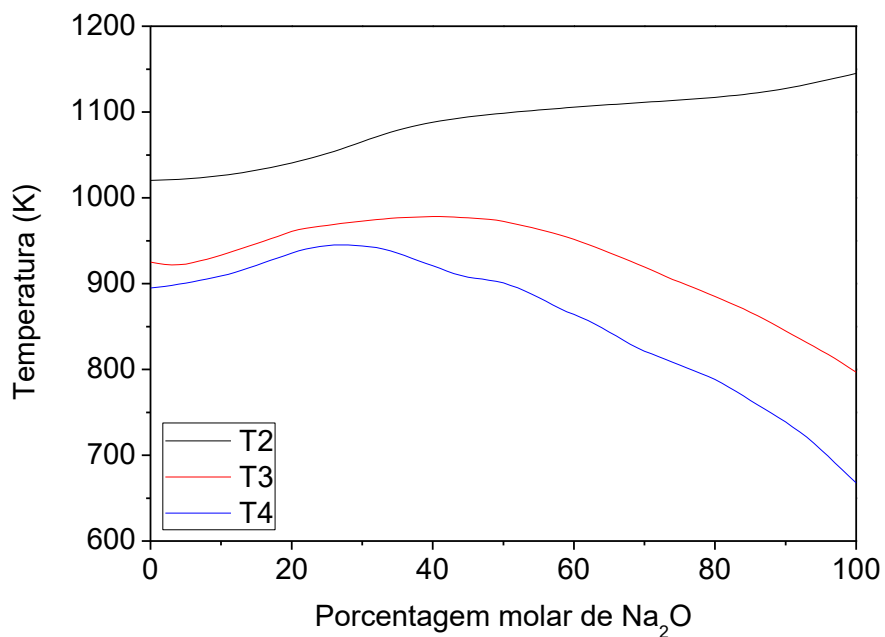


Figura 30: T2, T3 e T4 do binário B₂O₃-Na₂O em função da porcentagem molar de Na₂O.

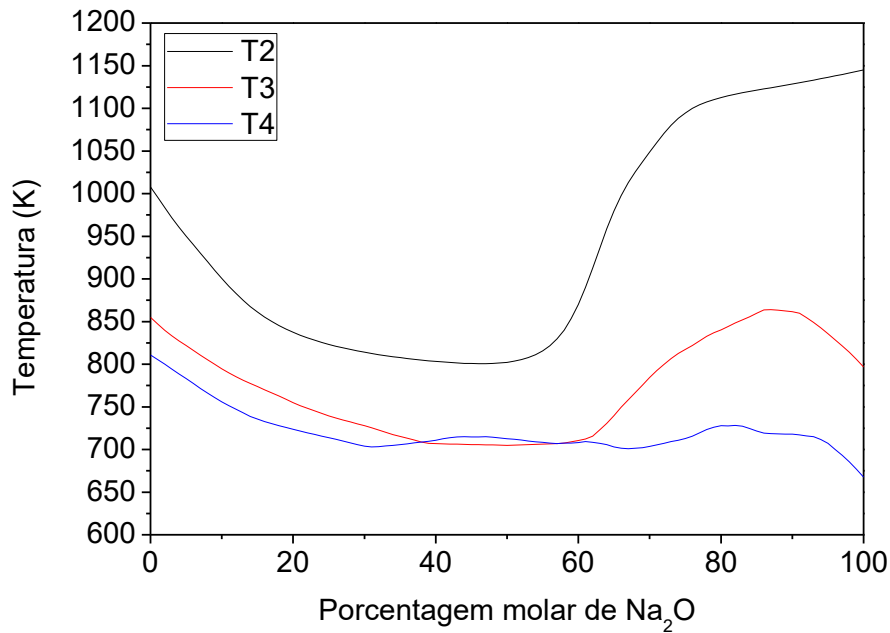


Figura 31: T2, T3 e T4 do binário P₂O₅-Na₂O em função da porcentagem molar de Na₂O.

Portanto, é possível afirmar que a rede neural foi capaz de prever razoavelmente o comportamento da viscosidade ao prever as temperaturas nas quais a viscosidade é igual a 10², 10³ e 10⁴ Pa.s, possuindo tanto R² como desvio relativo médio bons. Entretanto, na prática é possível ver que tal previsão não é perfeita, pois os gráficos de T2, T3 e T4 não seguem o comportamento esperado. Na realidade T2 é sempre maior que T3, que por sua vez é maior que T4, desta maneira, as três linhas não se cruzam. Por conta do erro na previsão, os intervalos se sobrepõem, o que possibilita que tais linhas se cruzem nos resultados obtidos. Para melhorar essa rede, assim como a de temperatura liquidus, seria importante selecionar dados de entrada “limpos”, sem a presença de *outliers*.

A fim de investigar a correlação entre a temperatura liquidus e as temperaturas T2, T3 e T4, os gráficos dos binários SiO₂-Na₂O (Figura 32) e B₂O₃-Na₂O₃ (Figura 33) foram plotados em conjunto com a T_{liq} da literatura. Poderia ser utilizado a liquidus prevista, porém, ao comparar curvas que possuem incertezas não teríamos um referencial correto. Além disso, foi considerado apenas de 0 a 50% do óxido de sódio pois sabe-se que esta é uma faixa onde

as composições vitrificam. Pode-se notar que para o sistema de sílica o comportamento da viscosidade, bem como da temperatura liquidus, é diminuir conforme é adicionado o óxido de sódio, isso porque este tem ponto de fusão menor que a sílica pura e atua como fundente. O mesmo não ocorre para o sistema com boro, pois a temperatura de fusão do boro é menor, o que leva a um aumento da temperatura liquidus. De maneira geral, o comportamento das curvas de viscosidade se aproxima do comportamento da liquidus em ambos os casos.

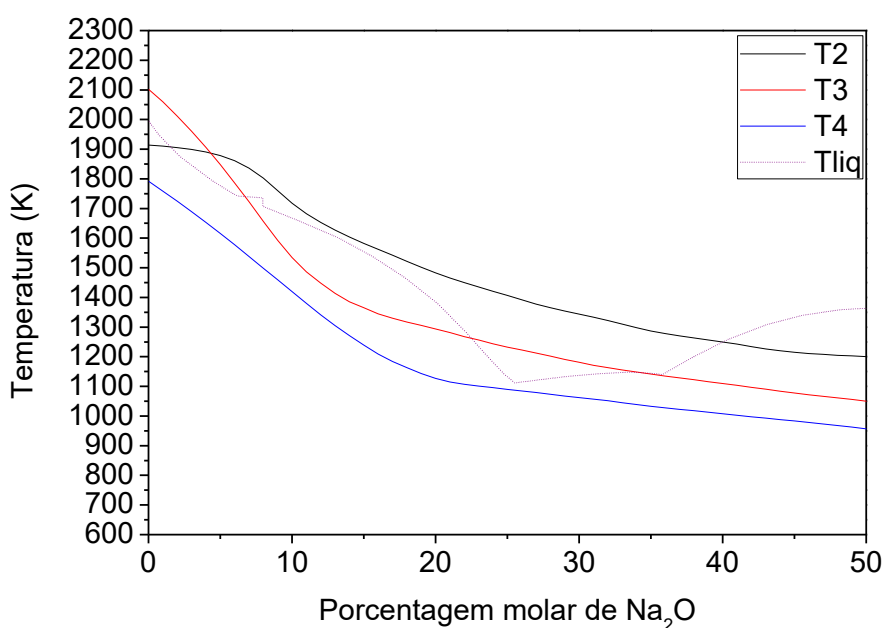


Figura 32: Curvas de T2, T3, T4 e liquidus do diagrama de fases da referência [61] para o sistema binário SiO₂-Na₂O

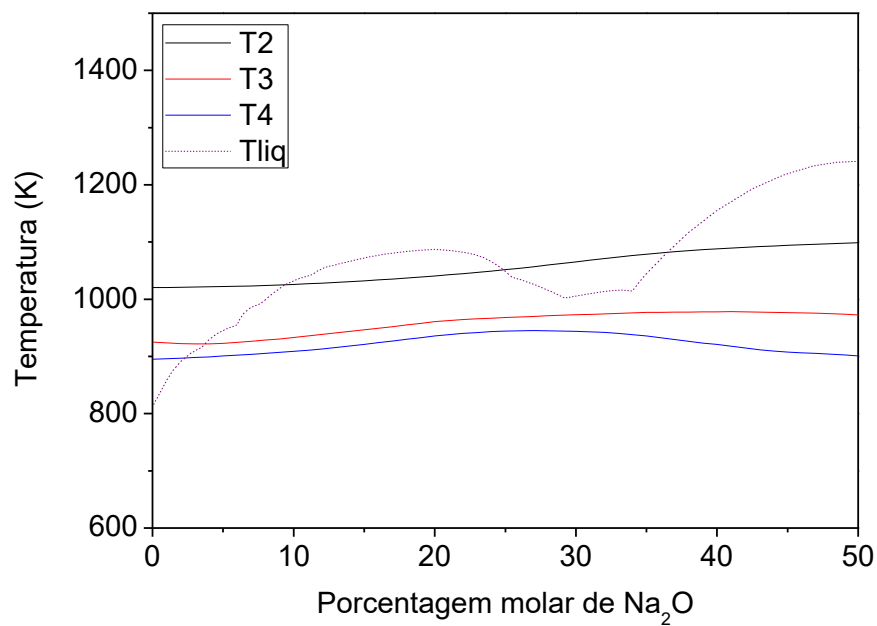


Figura 33: Curvas de T2, T3, T4 e liquidus do diagrama de fases da referência [62] para o sistema binário B₂O₃-Na₂O

6 CONCLUSÕES

Os modelos resultantes das redes neurais artificiais aqui desenvolvidas foram capazes de prever a temperatura *liquidus* com $R^2 = 0,997$, e as temperaturas T2, T3 e T4 com $R^2 = 0,999$. Os desvios relativos médios para as redes de viscosidade são, no máximo, 1,7%, enquanto que para T_{liq} é de 3,7%. Ao analisar o comportamento de tais propriedades, vemos que a rede neural tende a amortecer mudanças bruscas de temperatura, tais como em eutéticos, além de levar a erros altos para algumas composições específicas, especialmente para valores extremos.

Outra questão é que, mesmo nas redes de viscosidade, é possível notar outras falhas, visto que $T_2 > T_3 > T_4$, porém, nos gráficos resultantes nem sempre isso ocorre. Além disso, a rede apresenta erro um pouco maior em temperaturas com menor número de dados, e para composições contendo alguns elementos químicos, como o ferro.

7 TRABALHOS FUTUROS

O estudo de redes neurais artificiais para previsão de propriedades de vidros está no começo e há muito que pode ser trabalhado. Em um primeiro momento, é possível filtrar os dados de entrada utilizando métodos estatísticos para remoção de *outliers*. É necessário realizar a limpeza desses dados, porém sem excluir dados corretos do conjunto. A rede neural não consegue por si prever qualquer propriedade com um erro menor que o espalhamento inicial dos dados, ou seja, se há dados com erros, estes diminuem o poder de previsão do modelo. E mesmo quando não há um espalhamento significativo, ela ainda previu dados com um erro significativo (como com a temperatura *liquidus* de binários). Entretanto, não é possível retirar todos os dados que se encontram com erro, mesmo que sejam pequenos: é esperado que a rede seja, ao menos, um pouco robusta e consiga prever o comportamento mesmo na presença de pequenos erros.

Outro caminho é inserir novas variáveis para a rede neural. Neste trabalho, a rede recebeu apenas dados de composição e da propriedade que seria estudada. É possível inserir outros dados, como informações a respeito dos átomos (sua massa, número de elétrons, eletronegatividade, entre outras) para que a rede tenha mais informações sobre os vidros para usar na descrição das suas propriedades. Quanto mais informações relevantes ela receber, maior seria a capacidade de prever corretamente, porque estas outras variáveis de fato interferem nas propriedades. Obviamente, ao fazer isso, tornaria o problema mais complexo.

A última possibilidade é o uso de diferentes algoritmos de aprendizado de máquina. Talvez para a viscosidade a RNA seja muito boa, mas para a T_{liq} não. Seria necessário testar um algoritmo que teste simultaneamente diferentes modelos e compare seus erros para eleger o melhor método, rotina conhecida como auto-ml, similar ao que foi feito neste trabalho com os hiperparâmetros.

Esta ainda é uma área pouco explorada e que é capaz de trazer inúmeros avanços no campo de desenvolvimento de novos materiais.

8 REFERÊNCIAS BIBLIOGRÁFICA

1. ZANOTTO, E. D.; MAURO, J. C. The glassy state of matter: Its definition and ultimate fate. **Journal of Non-Crystalline Solids** , v. 471, p. 490-495, Setembro 2017.
2. KRISHNAN, N. M. A. et al. Predicting the dissolution kinetics of silicate glasses using machine learning. **Journal of Non-Crystalline Solids**, v. 487, p. 37–45, 2018.
3. CASSAR, D. R.; CARVALHO, A. P. L.; ZANOTTO, E. D. Predicting glass transition temperatures using neural networks. **Acta Materialia**, 2018.
4. BISHNOIA, S. et al. Predicting Young's modulus of oxide glasses with sparse datasets using machine learning. **Journal of Non-Crystalline Solids**, v. 524, Novembro 2019.
5. YANG, K. et al. Predicting the Young's Modulus of Silicate Glasses using High-Throughput Molecular Dynamics Simulations and Machine Learning.. **Scientific Reports** , v. 9, Junho 2019.
6. ALCOBAÇA, E. et al. Explainable Machine Learning Algorithms For Predicting Glass Transition Temperatures. **Acta Materialia**, v. 188, p. 92-100, Abril 2020.
7. DENG, B. Machine Learning on DENSITY and Elastic Property of Oxide Glasses Driven by Large Dataset. **Journal of Non-Crystalline Solids**, v. 529, Fevereiro 2020.
8. LIU, H. et al. Machine learning for glass science and engineering: A review. **Journal of Non-Crystalline Solids**, Julho 2019.
9. TANDIA, A.; ONBASLI, M. C.; MAURO, J. C. Machine Learning for Glass Modeling. In **Springer Handbook of Glass**, edited by J. David Musgraves, Juejun Hu and Laurent Calvez, 2019.
10. ZANOTTO, E. D.; COUTINHO, F. A. B. How many non-crystalline solids can be made from all the elements of the periodic table? **Journal of Non-Crystalline Solids**, v. 347, p. 285-288, Novembro 2004.
11. SILVA, I. N. et al. **Artificial neural networks - A practical course**. [S.l.]: Springer International Publishing, 2017.
12. LEE, J.-D. Object recognition using a neural network with optimal feature extraction. **Mathematical and Computer Modelling**, v. 25, p. 105-117, 1997.
13. HASHIM, M. F. et al. Multiple neural networks for Human Face Recognition. **Biomed 06, IFMBE Proceedings**, v. 15, p. 77-81, 2007.

14. CHEN, X.; SZTANDERA, L.; CARTWRIGHT, H. M. A neural network approach to prediction of glass transition temperature of polymers. **International Journal of Intelligent Systems** , v. 23, p. 22-32, 2008.
15. DREYFUS, C.; DREYFUS, G. A machine learning approach to the estimation of the liquidus temperature of glass-forming oxide blends. **Journal of Non-Crystalline Solids**, v. 318, p. 63–78, 2003.
16. BRAUER, D. S.; RUSSEL, C.; KRAFT, J. Solubility of glasses in the system P2O5–CaO–MgO–Na2O–TiO2: Experimental and modeling using artificial neural networks. **Journal of Non-Crystalline Solids** , v. 353 , p. 263–270, 2007.
17. JIUSTI, J. et al. Viscosity and liquidus-based predictor of glass-forming ability of oxide glasses. **J Am Ceram Soc.**, p. 1-12, 2019.
18. CALLISTER JR., W. D.; REFHWISCH, D. G. **Callister Materials Science Engineering: An Introduction**. Oitava edição. ed. [S.I.]: John Wiley & Sons, Inc., 2010.
19. CASSAR, D. R. **Nucleação, crescimento de cristais, relaxação e escoamento viscoso em vidros de dióxido de silício e diborato de lítio**. Tese (Doutorado em Ciência e Engenharia de Materiais) - Universidade Federal de São Carlos. São Carlos. 2014.
20. LEEDE, G.; WAAL, H. Evaluation of glass formation criteria. **Journal of Non-Crystalline Solids**, v. 104, p. 45-51, 1988.
21. RAY, C. S.; REIS, S. T.; HOLAND, W. A new DTA method for measuring critical cooling rate for glass formation. **Journal of Non-Crystalline Solids**, v. 351, p. 1350-1358, 2010.
22. RODRIGUES, F. L. **Construção de curvas tempo-temperatura-transformação para um aço supermartensítico através de difração de raios-X in situ**. UFRJ. Rio de Janeiro, p. 33. 2017.
23. ONORATO, P. I. K.; UHLMANN, D. R. Nucleating heterogeneities and glass formation. **Journal of Non-Crystalline Solids**, v. 22, p. 367- 378 , 1976.
24. NASCIMENTO, M. L. F. et al. Can glass stability parameters infer glass forming ability? **Journal of Non-Crystalline Solids**, v. 351, p. 3296-3308, 2005.
25. JIUSTI, J. **Influência de elementos modificadores e formadores de rede na habilidade de vitrificação**. Projeto de tese - Universidade Federal de São Carlos. São Carlos. 2018.
26. SHELBY, J. E. **Introduction to glass science and technology**. Segunda edição. ed. [S.I.]: The Royal Society of Chemistry, 2005.

27. KRUGER, S.; DEUBENET, J. "Stochastic nature of the liquid-to-crystal heterogeneous nucleation of supercooled lithium disilicate liquid,". **Journal of Non-Crystalline Solids**, v. 388, p. 6-9, 2014.
28. LU, Z. P.; LIU, C. T. A new glass-forming ability criterion for bulk metallic glasses. **Acta Materialia**, n. 50, p. 3501-3512, 2002.
29. ZHILIN, L.; DING, Y. H.; INOUE, A. A new criterion for predicting the glass-forming ability of bulk metallic glasses. **Journal of Alloys and Compounds**, n. 475, p. 207-219, Maio 2009.
30. CHATTOPADHYAY, C. et al. Critical evaluation of glass forming ability criteria. **Materials Science and Technology**, v. 32, n. 4, p. 380-400, 2016.
31. CABRAL, A. A.; FREDERICCI, C.; ZANOTTO, E. D. A test of the Hruby parameter to estimate glass-forming ability. **Journal of Non-Crystalline Solids**, v. 219, p. 182-186, 1997.
32. HRUBÝ, A. Evaluation of glass-forming tendency by means of DTA. **Czechoslovak Journal of Physics**, v. 22, p. 1187-1193, 1972.
33. CABRAL, A. A.; CARDOSO, A.; ZANOTTO, E. D. Glass-forming ability versus stability of silicate glasses I Experimental test. **Journal of Non-Crystalline Solids**, v. 320, p. 1-8, 2003.
34. WALLENBERGER, F. T. E. A. The Liquidus Temperature; Its Critical Role in Glass Manufacturing. **International Journal of Applied Glass Science**, n. 1 [2], p. 151-163, 2010.
35. FERREIRA, E. B.; LIMA, M. L.; ZANOTTO, E. D. DSC method for determining the liquidus temperature of glass-forming systems. **Journal of the American Ceramic Society**, 2010.
36. CARTER, C. B.; NORTON, M. G. **Ceramic materials: science and engineering**. [S.l.]: Springer, 2007.
37. GARZON, C. M.; TSCHIPTSCHIN, A. P. Modelamento termodinâmico e cinético por meio do método Calphad do processamento térmico e termoquímico de aços. **Matéria**, Rio de Janeiro, v. 11, n. 2, 2006.
38. MAURO, J. C. et al. Accelerating the Design of Functional Glasses through Modeling. **Chemistry of Materials**, v. 28, p. 4267-4277, 2016.
39. VOGEL, H. **Phys. Z.**, v. 22, p. 645, 1921.
40. FULCHER, G. S. **Journal of American Ceramic Society**, v. 8, p. 339, 1925.
41. TAMMANN, G.; HESSE, W.; ANORG, Z. **Allg. Chem.**, v. 156, p. 245, 1926.
42. MAURO, J. C. et al. Viscosity of glass-forming liquids. **Proceedings of the National Academy of Sciences of the United States of America**, v. 106, p. 19780-19784, 2009.

43. ADAM, G.; GIBBS, I. H. On the temperature dependence of cooperative relaxation properties in glass-forming liquids. **J. Chem. Phys.**, n. 43, 1965.
44. YUE, Y. et al. Viscosity of glass-forming liquids. **Proceedings of the National Academy of Sciences**, v. 106, n. 47, p. 19780–19784, 2009.
45. DIMITRIEV, Y. . I. Y. . A. I. R. History of sol-gel science and technology. **Journal of the University of Chemical Technology and Metallurgy** , v. 43, p. 181–192, 2008.
46. KAJIHARA, K. Recent advances in sol–gel synthesis of monolithic silica and silica-based glasses. **Journal of Asian Ceramic Societies**, v. 1, p. 121–133, 2013.
47. MITCHELL, T. M. Machine Learning and Data Mining. **Communications of the Acm** , 1999, v. 42, p. 31-36.
48. SCHMIDT, J. et al. Recent advances and applications of machine learning in solid-state materials science. **npj Computational materials**, 2019.
49. LIU, Y. et al. Materials Discovery and Design using Machine Learning. **Journal of Materiomics**, 2017.
50. SRIVASTAVA, N. et al. Dropout: a simple way to prevent neural networks from overfitting. **Journal of Machine Learning Research**, v. 15, p. 1929-1958, 2014.
51. CAI, A. H. et al. Artificial neural network modeling of reduced glass transition temperature of glass forming alloys. **Applied Physics Letters**, v. 92, 2008.
52. CAI, A. H. et al. Artificial neural network modeling for undercooled liquid region of glass forming alloys. **Computational Materials Science**, v. 48, p. 109-114, 2010.
53. CAI. et al. Prediction of critical cooling rate for glass forming alloys by artificial neural network. **Materials and Design**, v. 52, p. 671–676, 2013.
54. CHIDDARWAR, S. S.; BABU, N. R. Comparison of RBF and MLP neural networks to solve inverse kinematic problem for 6R serial robot by a fusion approach. **Engineering Applications of Artificial Intelligence**, v. 23, p. 1083-1092, 2010.
55. REN, F. et al. Accelerated discovery of metallic glasses through iteration of machine learning and high-throughput experiments. **Science Advances**, n. 4, 2018.
56. Disponível em: <<https://github.com/epam/SciGlass>>. Acesso em: 19 nov. 2019.
57. Disponível em: <<https://ceramics.org/publications-resources/phase-equilibrium-diagrams>>. Acesso em: 2020.

58. Disponível em: <<https://lmfit.github.io/lmfit-py/model.html>>. Acesso em: 2020.
59. CALCULATING R-squared (coefficient of determination) with centered vs. un-centered sums of squares. **StackExchange**, 2013. Disponível em: <<https://stats.stackexchange.com/questions/36064/calculating-r-squared-coefficient-of-determination-with-centered-vs-un-center/37442#37442>>. Acesso em: abril 2020.
60. SUKENAGA, S. et al. Effect of oxidation state of iron ions on the viscosity of silicate melts. **ISIJ International**, n. 53(2), p. 185-190, Janeiro 2013.
61. ALLENDORF; M. D., S. K. E. **J. Electrochem. Soc.**, v. 148, n. 2, p. B59-B67, 2001.
62. WANG, C. et al. **J. Phase Equilib.**, v. 24, n. 1, p. 12-20, 2003.

APÊNDICE A: ALGORITMO UTILIZADO ESQUEMATIZADO

início CÓDIGO_BASE

recebe variável DADOS

para cada composição com $O < 0,3$

remover a composição

para cada composição com $S > 0$ ou $H > 0$ ou $C > 0$ ou $Pt > 0$ ou $Au > 0$ ou $F > 0$ ou $Cl > 0$ ou $N > 0$ ou $Br > 0$ ou $I > 0$

remover a composição

aplicar rotina REPRESENTATIVIDADE_ELEMENTOS na variável DADOS

aplicar rotina REMOÇÃO_DE_DUPLICADOS na variável DADOS

Remover aleatoriamente 20% dos materiais contidos em DADOS e armazenar em DADOS_HOLDOUT

Aplicar rotina TESTE_DE_TOPOLOGIA

Com a melhor topologia (ERRO menor dentre todas as topologias testadas), iniciar treino da RNA com todos os dados

Retornar ERRO_FINAL da rede calculado a partir dos DADOS_HOLDOUT

fim CÓDIGO_BASE

início REPRESENTATIVIDADE_ELEMENTOS

recebe variável DADOS

se VARIABEL_CONTROLE igual a 1

VARIABEL_CONTROLE = 0

contar o número total de composições T

para cada elemento E em DADOS, contar quantas vezes a fração atômica do elemento é > 0 e criar variável E_ELEMENTO[NOME] com esse valor

checar todos os E_ELEMENTO[NOME]

se $E_ELEMENTO[NOME] < 0,01 * T$

excluir composição e VARIABEL_CONTROLE = 1

fim REPRESENTATIVIDADE_ELEMENTOS

início REMOÇÃO_DE_DUPLICADOS

recebe variável DADOS

arredonda todas as colunas referentes à composição de DADOS para terem 3 algarismos significativos

criar variável DADOS_SEM_DUPLICADOS

para cada material M em DADOS,

se M não existe em DADOS_SEM_DUPLICADOS,

criar variável DADOS_DUPLICADOS
 verificar se existem outros materiais de mesma composição
 M em DADOS e armazenar todos eles em
 DADOS_DUPLICADOS

tirar a mediana do valor da propriedade dos materiais
 armazenados em DADOS_DUPLICADOS

acrescentar material M em DADOS_SEM_DUPLICADOS
 com a mediana da propriedade

retorna para o usuário a variável DADOS_SEM_DUPLICADOS

fim REMOÇÃO_DE_DUPLICADOS

início TESTE_DE_TOPOLOGIA

Selecionar aleatoriamente 10% dos materiais e armazenar em
 GRUPO_VALIDACAO

Selecionar aleatoriamente 10% dos materiais e armazenar em
 TESTE_LOCAL

Se $P < 200$ para T1q e $P = 75$ para T2, T3 e T4.

P=0

Sortear NeuroniosCamada1 entre 20 e 300

Sortear NeuroniosCamada2 entre 20 e 300

Sortear DropoutCamada1 entre 0 e 0,3

Sortear DropoutCamada2 entre 0 e 0,3

Sortear TamanhoLote entre 64, 128, ou 256

Sortear Paciencia entre 10 e 14

Treinar a RNA com os parâmetros RNA (NeuroniosCamada1, NeuroniosCamada2, DropoutCamada1, DropoutCamada2, TamanhoLote, Paciência)

A cada época, a rede calcular ERRO_DADOS e ERRO_GRUPO_VALIDACAO e compara. Se ERRO_GRUPO_VALIDACAO aumentar com relação ao ERRO_DADOS, somar NPACIENCIA = NPACIENCIA + 1

O treinamento interrompe quando NPACIENCIA = PACIENCIA ou quando 500 épocas se passarem

Computar ERRO ao final do treino, ao calcular com TESTE_VALIDACAO

P=P+1

Classificar todos os dados de ERRO, escolhendo o menor de todos.

Fim TESTE_DE_TOPOLOGIA