

UNIVERSIDADE FEDERAL DE SÃO CARLOS
CENTRO DE CIÊNCIAS EXATAS E DE TECNOLOGIA
DEPARTAMENTO DE ENGENHARIA DE MATERIAIS

PREVISÃO DA HABILIDADE DE FORMAR VIDROS (GFA)
VIA ALGORITMOS DE INTELIGÊNCIA ARTIFICIAL

ANDRÉ FIGUEIRA LOPES

SÃO CARLOS -SP
2024

**PREVISÃO DA HABILIDADE DE FORMAR VIDROS (GFA) VIA
ALGORITMOS DE INTELIGÊNCIA ARTIFICIAL**

Trabalho de conclusão de curso apresentado ao Departamento de Engenharia de Materiais da Universidade Federal de São Carlos, como requisito para obtenção do título de bacharel em Engenharia de Materiais.

Orientador: Edgar Dutra Zanotto

SÃO CARLOS -SP

2024



ATA DE DEFESA DE TRABALHO DE CONCLUSÃO DE CURSO (TCC)

NOME: Andre Figueira Lopes

RA: 744365

TÍTULO: Previsão da habilidade de formar vidros (GFA) via algoritmos de inteligência artificial

ORIENTADOR(A): Prof. Dr. Edgar Dutra Zanotto

CO-ORIENTADOR(A):

DATA/HORÁRIO: 30/01/2024, 10h30

BANCA – NOTAS:

	Monografia	Defesa
Prof. Dr. Edgar Dutra Zanotto	10	10
Profa. Dra. Ana Cândida Martins Rodrigues	10	10
Média	10	10

BANCA – ASSINATURAS:

Prof. Dr. Edgar Dutra Zanotto

Profa. Dra. Ana Cândida Martins Rodrigues

AGRADECIMENTOS

Gostaria, primeiramente, de agradecer aos meus pais (Milena e Luís) por terem sido fundamentais na minha formação, me ensinando a ser quem sou. À minha mãe, agradeço por ter sido presença e carinho. Ao meu pai, agradeço por me apoiar nas minhas escolhas, estar ao meu lado e por me ensinar a ver a vida com um toque de humor. Ao meu irmão (Henrique), por ser meu companheiro desde pequeno: nas brincadeiras, nas férias, nos jogos, nas viagens e em tantos outros momentos. Por ser meu irmãozinho sempre, com quem posso contar para tudo. Aos meus avós (Fani e Pedro), que dentre muitas outras coisas, me ensinaram o amor ao conhecimento e à leitura.

À minha namorada e companheira, Amanda, por todas as conversas, risadas e por seu apoio. Por seu carinho comigo e pelas pequenas grandes coisas que vivemos juntos, e por me fazer reviver a arte na sua companhia. Aos amigos, agradeço pelo apoio nos momentos em que precisei, mas sobretudo pelo bons momentos de convívio, trazendo leveza e alegria.

Gostaria de agradecer ao prof. Dr. Edgar D. Zanotto pelos ensinamentos, orientações e pela dedicação à esse projeto. Agradeço também ao Rafael Baságua pelas suas contribuições relevantes para o desenvolvimento deste trabalho. Por fim, agradeço ao departamento de engenharia de materiais e à UFSCar por proporcionar um ambiente propício ao aprendizado e à pesquisa. Agradeço aos que lutam para que a universidade continue a ser pública e para que se torne cada vez mais acessível.

RESUMO

Vidros são frequentemente formados a partir do super-resfriamento de líquidos. A habilidade de formar vidros, conhecida como *glass-forming ability* (GFA) é uma propriedade fundamental no desenvolvimento de novos vidros, no entanto sua obtenção por métodos experimentais é custosa. Portanto, há grande interesse científico no desenvolvimento de métodos indiretos para obtenção de GFA. Estudos anteriores apontaram dois parâmetros fundamentais para a previsão da GFA, sendo eles a temperatura *liquidus* e a viscosidade em $T_{liquidus}$. No presente trabalho, a viscosidade em $T_{liquidus}$ foi obtida para 4950 composições de vidros óxidos por meio do ajuste da equação Vogel–Fulcher–Tammann–Hesse (VFTH, $\log_{10} \eta(T) = A + B / (T - T_0)$, onde η em Pa.s e T em K).

Esse trabalho propõe a integração de modelos físicos e de *Machine Learning* (ML) para prever a GFA de vidros óxidos por duas vias. Pela via indireta, é gerado um modelo ML para prever a $T_{liquidus}$ e outro modelo para prever a viscosidade em $T_{liquidus}$ calculada pelo modelo VFTH. Em seguida, os dois parâmetros são utilizados para calcular GFA por meio do modelo físico "Jezica", onde $GFA = \eta(T_l) / T_l^2$. A influência de cada componente na previsão da viscosidade foi avaliada de acordo com os papéis dos óxidos formadores de rede, intermediários e modificadores, lançando uma nova perspectiva na discussão de teorias estruturais e em seu papel nas propriedades. Por sua vez, pela via direta, a $T_{liquidus}$ e a viscosidade em $T_{liquidus}$ são utilizadas para calcular GFA e o modelo de ML permite prever diretamente a GFA. Essas abordagens buscam viabilizar a previsão de GFA baseada exclusivamente na composição de formadores de vidros óxidos, bem como estender a compreensão dessa propriedade, acelerando o desenvolvimento de novos vidros.

Palavras-chave: Vidros óxidos. Habilidade de formar vidros. *Machine Learning*. Modelagem híbrida. Viscosidade.

ABSTRACT

Glasses are often formed through the supercooling of liquids. The ability to form glasses, known as glass-forming ability (GFA), is a fundamental property in the development of new glasses; however, obtaining it through experimental methods is costly. Therefore, there is significant scientific interest in the development of indirect methods for GFA acquisition. Previous studies have identified two fundamental parameters for predicting GFA: the liquidus temperature and the viscosity at the liquidus temperature. In this work, viscosity at $T_{liquidus}$ was determined for 4950 oxide glass compositions by fitting the Vogel–Fulcher–Tammann–Hesse equation (VFTH, $\log_{10} \eta(T) = A + B / (T - T_0)$), where η is in Pa.s and T is in K).

This work proposes the integration of physical models and Machine Learning (ML) to predict the GFA of oxide glasses through two approaches. In the indirect pathway, an ML model is generated to predict the $T_{liquidus}$, and another model predicts the viscosity at $T_{liquidus}$ calculated by the VFTH model. Subsequently, these two parameters are used to calculate GFA through the "Jezica" physical model, where $GFA = \eta(T_l) / T_l^2$. The influence of each component in predicting viscosity was evaluated based on the roles of network-forming, intermediate, and modifier oxides, providing a new perspective in structural theories and their role in properties. On the other hand, in the direct pathway, $T_{liquidus}$ and viscosity at $T_{liquidus}$ are used to calculate GFA, and the ML model allows for the direct prediction of GFA. These approaches aim to enable the prediction of GFA solely based on the composition of oxide glass formers and to enhance the understanding of this property, thereby accelerating the development of new glasses.

Keywords: Oxide glasses. Glass-forming ability. Machine Learning. Hybrid modeling. Viscosity.

LISTA DE FIGURAS

Figura 1.1 – Esquema para prever GFA utilizando a composição de vidros por duas vias.	2
Figura 3.1 – Formação da base de dados contendo a composição e viscosidade em sua respectiva temperatura de medida.	9
Figura 3.2 – Diagrama de Venn representando a junção das bases 1 e 2 para formar a base de dados 3.	10
Figura 3.3 – Histograma contendo compostos que estão presentes em ao menos 20 composições de vidros.	11
Figura 3.4 – Histograma contendo compostos que estão presentes em menos de 20 composições de vidros.	12
Figura 3.5 – Histograma da frequência dos 25 óxidos presentes nas composições de vidros.	13
Figura 3.6 – Histograma do número de vidros com determinado número de compostos óxidos.	13
Figura 3.7 – Metodologia para criação dos modelos para prever GFA, por meio da incorporação de modelos físicos e de <i>Machine Learning</i>	16
Figura 4.1 – <i>Box plot</i> da distribuição da viscosidade em diferentes temperaturas. . .	17
Figura 4.2 – <i>Box plot</i> da distribuição de $T_{liquidus}$	18
Figura 4.3 – Histograma da quantidade de medidas de viscosidade em função da quantidade de temperaturas de medida da viscosidade.	19
Figura 4.4 – Ajuste da função VFTH para um vidro.	20
Figura 4.5 – Distribuição dos parâmetros A, B e T_0 da equação VFTH obtidos para diversas composições de vidros.	21
Figura 4.6 – Ajuste da função VFTH para vidros contendo diferentes quantidades de pontos de medida da viscosidade em função da temperatura.	22
Figura 4.7 – Ajuste de uma função linear para prever a viscosidade de vidros cuja viscosidade foi medida em pontos próximo de $T_{liquidus}$	23
Figura 4.8 – Ajuste de uma função linear para prever a viscosidade de vidros cuja viscosidade foi medida em pontos próximo de $T_{liquidus}$	24
Figura 4.9 – Gráfico dos valores previstos X valores reais de viscosidade em $T_{liquidus}$ utilizando o modelo <i>Random Forest</i>	26
Figura 4.10 – Performance do modelo <i>XGBoost</i> para prever a viscosidade em $T_{liquidus}$ em função dos hiperparâmetros selecionados.	27
Figura 4.11 – Impacto absolutos dos compostos na viscosidade em $T_{liquidus}$ avaliado pela metodologia <i>Shapley</i>	28

Figura 4.12–Impacto relativo de cada composto na viscosidade em $T_{liquidus}$ avaliado pela metodologia <i>Shapley</i>	29
Figura 4.13–Previsão de $T_{liquidus}$ utilizando modelo <i>Random Forest</i>	31
Figura 4.14–Previsão de GFA utilizando o modelo <i>Random Forest</i>	32

LISTA DE TABELAS

Tabela 3.1 – Estatística descritiva dos valores de viscosidade e temperatura <i>liquidus</i>	14
Tabela 3.2 – Classificação da habilidade de formar vidros	16
Tabela 4.1 – Classificação das composições segundo a habilidade de formar vidros .	24
Tabela 4.2 – Estatística descritiva da fração de Na ₂ O na composição dos vidros . . .	30

LISTA DE ABREVIATURAS E SIGLAS

GFA	Habilidade de Formar Vidros
DTA	Análise Térmica Diferencial
DSC	Calorimetria Exploratória Diferencial
DRX	Difração de Raios-X
ML	Machine Learning
GS	Estabilidade do Vidro
Jezica	Modelo de Previsão da Habilidade de Formar Vidros
VFTH	Equação de Vogel-Fulcher-Tamman-Hess
SVM	Vetor de Suporte de Máquina
MLP	Perceptron Multi-Camadas
MAE	Erro Médio Absoluto
MSE	Erro Médio Quadrático
KNN	K-Vizinhos mais Próximos

LISTA DE SÍMBOLOS

$T_{liquidus}$	Temperatura liquidus
R_c	Valor Crítico de Resfriamento
η	Viscosidade
T_0	Temperatura absoluta
A	Constante do valor inicial de Viscosidade da Equação VFTH
B	Constante de ajuste da Equação VFTH
T_0	Temperatura de Vogel-Fulcher
T_g	Temperatura de transição vítrea
T_l	Temperatura liquidus

SUMÁRIO

1	INTRODUÇÃO	1
2	FUNDAMENTAÇÃO TEÓRICA	3
3	METODOLOGIA	9
3.1	CRIAÇÃO DAS BASES DE DADOS	9
3.2	SELEÇÃO DE DESCRITORES (COMPOSTOS)	10
3.3	OBTENÇÃO DA VISCOSIDADE EM $T_{liquidus}$	12
3.4	DESCRIÇÃO DO DATASET FINAL	13
3.5	<i>MACHINE LEARNING</i>	14
3.5.1	Seleção de Algoritmos	14
3.5.2	Hiperparâmetrização	15
3.5.3	Avaliação do Desempenho	15
3.5.4	Interpretação dos modelos	15
3.5.5	Extensão do modelo	15
4	RESULTADOS E DISCUSSÕES	17
5	CONSIDERAÇÕES FINAIS E PERSPECTIVAS	34
	REFERÊNCIAS	36

1 INTRODUÇÃO

Vidros podem ser formados por diversos métodos, tais como deposição de vapor, reações no estado sólido, técnicas sol-gel, dentre outras. No entanto, na maior parte dos casos, o fenômeno de vitrificação é verificado quando um líquido é submetido a um super-resfriamento até a temperatura de transição vítrea, na qual o aumento da viscosidade impede o rearranjo estrutural, conservando a estrutura amorfa. A princípio qualquer material pode ser vitrificado, desde que submetido à uma taxa de resfriamento suficientemente rápida para impedir a nucleação e o crescimento de cristais. Quanto mais lentamente um material puder ser resfriado até a temperatura de transição vítrea sem que ocorra cristalização, maior sua habilidade de formar vidros, ou *glass-forming ability*, GFA (NASCIMENTO et al., 2005). A GFA pode ser medida experimentalmente por meio de sucessivos ciclos de resfriamento, verificando a ocorrência ou não de cristalização por meio de caracterização térmica (DTA ou DSC) ou difração de raios-X (DRX) (SHELBY, 2005).

Diversos trabalhos têm se dedicado a buscar métodos indiretos para medir a GFA, evitando assim exaustivos e custosos testes experimentais, buscando correlacionar parâmetros mais fáceis de obter (NASCIMENTO et al., 2005), (JIUSTI et al., 2020). Recentemente, foi desenvolvido no LaMaV-UFSCar um modelo físico para prever a habilidade de formar vidros a partir da temperatura *liquidus* e da viscosidade nessa temperatura. Esses parâmetros contém implicitamente informações químicas e estruturais do material (JIUSTI et al., 2020). O modelo foi testado em mais de 30 vidros óxidos, corroborando o conceito de que líquidos com alta viscosidade e baixa $T_{liquidus}$ possuem uma boa GFA (JIUSTI et al., 2020). No entanto, as composições testadas correspondem a uma ínfima fração dos vidros que podem potencialmente ser produzidos (ZANOTTO; COUTINHO, 2004).

Nesse sentido, a utilização de modelos de *Machine Learning* (ML) para prever a $T_{liquidus}$, a viscosidade em $T_{liquidus}$ e a GFA são de extrema relevância no campo de pesquisa e desenvolvimento de novos materiais vítreos, permitindo criar vidros com propriedades desejadas e acelerando seu desenvolvimento. Nesse campo, a abordagem de ML é especialmente fértil, já que a maioria das **propriedades dos vidros são função, quase que exclusivamente, da composição química** (CASSAR et al., 2021b), (RAVINDER et al., 2021). Os modelos de *Machine Learning* apresentam elevada capacidade de identificar interações entre os dados (composição química e propriedades) que não são facilmente detectadas por outras metodologias. A integração de ML na área de materiais tem desempenhado um papel crucial no desenvolvimento de novos vidros, particularmente na abordagem utilizando algoritmos interpretáveis, cuja transparência

possibilita compreender e avaliar o impacto de cada componente em dada propriedade (CASSAR et al., 2021a).

No contexto deste trabalho propõe-se a integração de modelos físicos e de *Machine Learning* para a predição da GFA. Essa abordagem permite aliar o alto poder preditivo de modelos físicos com a eficiência de modelos de ML (JIRASEK; HASSE, 2023). **Este trabalho, original em nível internacional, pretende verificar a viabilidade da predição da habilidade de formar vidros óxidos com base apenas na composição química.** O esquema da Figura 1 ilustra dois caminhos possíveis para prever a GFA, de modo que o usuário precise fornecer apenas a composição química dos vidros como dados de entrada.

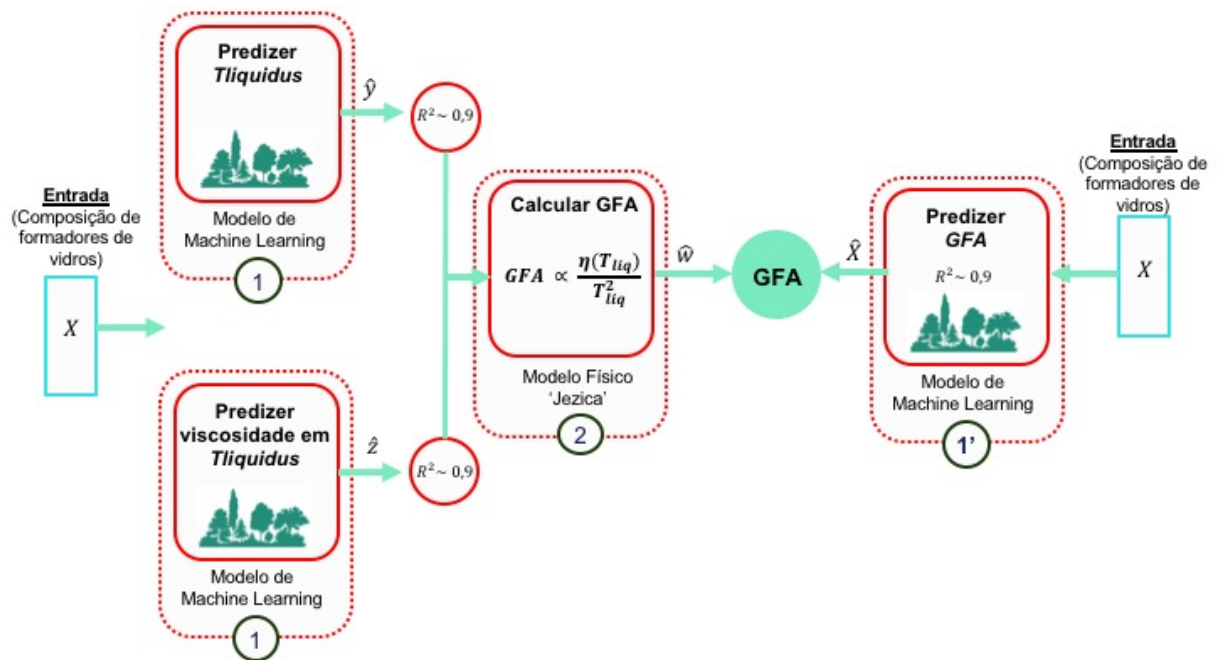


Figura 1.1 – Esquema para prever GFA utilizando a composição de vidros por duas vias.

Da esquerda para a direita está representado o caminho indireto, por meio do qual, os modelos preveem a $T_{liquidus}$ e viscosidade em $T_{liquidus}$, a partir das quais é possível calcular a GFA por meio do modelo físico ‘Jezica’. Da direita para a esquerda, o modelo encapsulado permite prever diretamente a GFA de um vidro a partir de sua composição nominal.

2 FUNDAMENTAÇÃO TEÓRICA

Vidros podem ser definidos como estruturas fora do equilíbrio termodinâmico, não cristalinas e que exibem o fenômeno de transição vítrea. A estrutura dos vidros é semelhante à dos líquidos super-resfriados, cuja tendência é cristalizar (ZANOTTO; MAURO, 2017). A capacidade dos vidros óxidos em transmitir a luz visível decorre do seu grande *band gap* e da ausência de contornos de grãos em sua microestrutura. A formação de vidros requer resfriar “rapidamente” um material a partir do estado fundido evitando a cristalização. Além disso, existem outros processos alternativos, como deposição de vapor e sol-gel, que possibilitam a produção de vidros, ampliando as opções tecnológicas para a fabricação desses materiais (SHELBY, 2005). Independentemente do processo adotado, para a produção de um vidro é imprescindível impedir a cristalização.

Sabe-se, contudo, que a fase cristalina é mais estável do ponto de vista termodinâmico que a fase líquida em temperaturas inferiores à temperatura de fusão. Desse modo, para formar vidros é preciso resfriar o líquido em um ritmo suficientemente rápido, para que este ultrapasse a região de transformação vítrea conservando a estrutura amorfa, mas com uma viscosidade tão elevada que seu escoamento e a cristalização não sejam mais viáveis em uma escala de tempo observável (ZANOTTO; MAURO, 2017). Alguns materiais tendem a formar vidros mais facilmente que outros, de modo que, inicialmente, as teorias ligadas à estrutura do material preponderaram na explicação dessa diferença, embora a maioria das teorias sejam empíricas, baseadas sobretudo nos vidros até então conhecidos, sobretudo os silicatos. Por exemplo, Smekal propôs que uma boa habilidade de formar vidros estaria ligada ao caráter misto das ligações químicas (OJOVAN, 2021). Por um lado, a presença de ligações covalentes favorece a formação de uma estrutura periódica, por seu caráter direcional; por outro, as ligações iônicas pela ausência de direcionalidade, dificultam a formação de uma rede (SHELBY, 2005). De fato, inúmeros exemplos práticos reforçam essa teoria.

Nesse mesmo sentido, Stanworth classificou os óxidos em três subgrupos, segundo a eletronegatividade do elemento ligado ao oxigênio e, portanto, ao caráter covalente-iônico das ligações químicas (SHELBY, 2005). O primeiro grupo apresenta um caráter covalente-iônico semelhante, os quais são denominados formadores de vidros, dentre os quais podemos destacar o SiO_2 , GeO_2 , B_2O_3 , P_2O_5 , entre outros. No outro extremo encontra-se um grupo composto por elementos de baixa eletronegatividade ligados ao oxigênio e, portanto, com caráter predominantemente iônico, sendo conhecidos como modificadores de rede, como, por exemplo, o Na_2O e o CaO (OJOVAN, 2021), (SHELBY, 2005). Entre esses dois grupos, há cátions que ligados ao oxigênio podem atuar como

formadores ou modificadores dependendo dos componentes presentes na composição do vidro, sendo assim denominados intermediários, tais como o Al_2O_3 e o SnO_2 . Embora as teorias ligadas a estrutura expliquem em parte a habilidade de formar vidros, a ênfase atual recai nas teorias cinéticas, segundo a qual o ritmo de resfriamento do fundido determinará a formação ou não de um vidro.

A habilidade de formar vidros (GFA) pode ser definida como a taxa crítica de resfriamento, acima da qual um líquido super-resfriado passa para o estado vítreo sem a formação de cristais. Uma fração máxima de cristais entre 0.1% e 0.0001% é por convenção considerada como suficiente para considerar o material como vítreo (NASCIMENTO et al., 2005). A princípio, vidros podem ser obtidos a partir de qualquer líquido, mesmo a partir de metais monoatômicos, desde que haja um resfriamento suficientemente rápido (ZHONG et al., 2014). Em teoria, basta aplicar um resfriamento superior a um valor crítico (R_c) para evitar a cristalização e vitrificar, sendo $\text{GFA} = 1/R_c$ e dependente do líquido a ser resfriado (JIUSTI et al., 2020).

Embora a compreensão do conceito de habilidade de formar vidros (GFA) seja relativamente simples, sua medida experimental é custosa, uma vez que implica fundir o material, mantendo-o em temperaturas superiores à de fusão por um determinado tempo, de modo que haja a completa dissolução dos cristais. Em seguida o material deve ser resfriado com diferentes taxas verificando se há ou não presença de cristais. Se for verificada a vitrificação o ensaio deve refeito, ou seja, o material deve ser reaquecido acima da temperatura de fusão e resfriado a uma taxa mais lenta, até que se verifique a formação de cristais (SHELBY, 2005). Após sucessivos ciclos de aquecimento/resfriamento poder-se-á verificar o menor ritmo de resfriamento necessário para a formação de vidro. Esse experimento para determinação do GFA é lento e ocorre, na maior parte das vezes, por meio de abordagens experimentais, como a que foi descrita anteriormente, nas quais as composições são selecionadas segundo o conhecimento do experimentador e utilizando metodologias de tentativa e erro (ALCOBAÇA et al., 2020), (RAVINDER et al., 2021). Experimentalmente a GFA pode ser obtida por duas metodologias: observando-se um pico de cristalização durante o resfriamento de um líquido utilizando análises térmicas ou conduzindo análises de difração de raios-X para detectar formação de cristais em líquidos resfriados à diferentes taxas (JIUSTI et al., 2020). Os autores ainda acrescentam: “Os erros nos valores de R_c podem diferir ordens de grandeza, quando determinados por diferentes autores”. (JIUSTI et al., 2020, p.10)

Dadas as limitações experimentais, no que se refere a tempo, custo e precisão, diversos trabalhos foram dedicados à obtenção do GFA por meios indiretos, ou seja, verificando como o GFA se correlaciona com outros parâmetros que podem ser mais facilmente medidos ou calculados (JIUSTI et al., 2020), como por exemplo, pela medida da

estabilidade do vidro (GS), ou seja, avaliando a resistência do vidro a cristalizar quando resquecido (NASCIMENTO et al., 2005). Sabe-se que à medida que um líquido é resfriado, sua viscosidade aumenta. A viscosidade eventualmente se torna tão grande que os átomos não têm mobilidade suficiente para se rearranjar em uma estrutura cristalina ordenada. A redução da temperatura, por outro lado, favorece a formação de cristais do ponto de vista termodinâmico (aumento de ΔG), uma vez que o arranjo cristalino diminui a energia livre do sistema. Desse modo, haverá duas variáveis concorrentes que explicam a formação de cristais: de um lado a variável termodinâmica que aumenta à medida que o líquido é resfriado. Por outro, o aumento da viscosidade torna o processo de cristalização mais lento, uma vez que reduz drasticamente a cinética de cristalização (OJOVAN, 2021). A barreira cinética para cristalização a partir de temperaturas inferiores a temperatura de transição vítrea se torna tão elevada que os átomos não conseguem mais se rearranjar (numa escala de tempo razoável) em virtude da alta viscosidade.

Em estudos prévios, constatou-se que a viscosidade no ponto de fusão ($T_{liquidus}$) é um parâmetro significativo para prever a capacidade de formação de vidros (SUN et al., 2017). Por meio de modelagens, busca-se substituir ou, ao menos, orientar as análises experimentais dispendiosas necessárias para determinar a taxa crítica de resfriamento de um material vítreo. Recentemente, foi estabelecido um modelo correlacionando a GFA com a viscosidade em $T_{liquidus}$ (JIUSTI et al., 2020, p.2):

$$GFA = \frac{\eta(T_l)}{T_l^2} \quad (1)$$

Onde T_l se refere à temperatura *liquidus* e $\eta(T_l)$ à viscosidade na referida temperatura. Esse modelo, denominado ‘**Jezica**’ em homenagem aos autores (Jiusti, Zanotto, Cassar, Andreetta) foi testado e validado em mais de 30 vidros óxidos, estequiométricos ou não, incluindo vidros extremos com elevada e baixa $T_{liquidus}$, cobrindo materiais cuja GFA se estende por um intervalo de 9 ordens de grandeza, servindo de base para a proposição de uma classificação qualitativa quanto à GFA do material (JIUSTI et al., 2020). Sobre o modelo ‘Jezica’ os autores concluem:

Finalmente, este novo preditor de GFA corrobora o conceito amplamente difundido de que substâncias com alta viscosidade na temperatura *liquidus* e uma baixa *liquidus* ($T_{liquidus}$) são facilmente vitrificadas, fornecendo uma ferramenta preditiva poderosa para a busca e o design de novos vidros óxidos. (JIUSTI et al., 2020, p.10)

Dessa forma, decorre do exposto o interesse em caracterizar a viscosidade de distintas composições de formadores de vidros em $T_{liquidus}$. Contudo, a viscosidade de

vidros revela-se altamente sensível à temperatura, tornando necessárias extrapolações para determinar o valor da viscosidade em $T_{liquidus}$. Trabalhos anteriores indicam que as curvas de viscosidade em função da temperatura podem ser adequadamente modeladas utilizando a expressão empírica de equações do tipo Vogel-Fulcher-Tamman-Hesse (VFTH), as quais descrevem os movimentos moleculares cooperativos por meio da dependência da viscosidade com a temperatura (NASCIMENTO; APARICIO, 2007). A equação pode ser expressa como :

$$\log \eta(T) = A + \frac{B}{(T - T_0)} \quad (2)$$

Onde η é a viscosidade expressa em Pa.s e T a temperatura em Kelvin, sendo os parâmetros ajustáveis da equação A, B e T_0 . Diversas variantes dessa equação foram documentadas na literatura, demonstrando uma notável precisão na previsão das viscosidades, mesmo considerando que ela pode variar mais de 12 ordens de grandeza (HRMA, 2008) no intervalo de temperaturas entre T_g e $T_{liquidus}$ para diferentes composições (NASCIMENTO; APARICIO, 2007), podendo ser aplicada para vidros óxidos (NASCIMENTO; APARICIO, 2007), bem como para vidros metálicos (DUINE et al., 1991), permitindo assim obter a viscosidade em $T_{liquidus}$ para diferentes composições de formadores de vidros, cuja viscosidade foi medida em diferentes temperaturas. Além disso, foi apontada uma correlação entre os parâmetros A e T_0 , de modo que se constatou que manter A constante e igual a -5 resulta em um razoável ajuste da curva VFTH para diversas composições de vidros óxidos (NASCIMENTO; APARICIO, 2007).

As composições de vidros já estudadas e reportadas na literatura compõem uma pequena fração do total de vidros que teoricamente podem ser produzidos variando 80 elementos químicos em incrementos de 1%, as quais são estimados em mais de 10^{52} composições (ZANOTTO; COUTINHO, 2004), das quais estima-se que tenham sido produzidos 10^5 composições de vidros inorgânicos (CASSAR et al., 2018). Advém deste fato o interesse em utilizar ferramentas de *Machine Learning* para reconhecer padrões dos vidros já produzidos, com o intuito de inferir/predizer potenciais composições com propriedades desejáveis e ainda não estudadas. Essa abordagem é reforçada pelo fato de que as propriedades de materiais vítreos dependem, quase que exclusivamente de sua composição química nominal (CASSAR et al., 2018), não apresentando influência de fatores estruturais, tais como contornos de grãos e discordâncias em razão da sua estrutura desordenada (RAVINDER et al., 2021).

As aplicações de *Machine Learning* na área de vidros podem responder à diferentes questões, tais como o desenvolvimento, a aceleração e a redução de custos para o *design*

de novos vidros, além do desenvolvimento de novos conhecimentos impulsionados por algoritmos de *Machine Learning* (RAVINDER et al., 2021). Um destaque pode ser dado para o uso de ferramentas de ML para prever a viscosidade de vidros em $T_{liquidus}$, possibilitando um cálculo indireto de GFA para vidros ainda não explorados (JIUSTI et al., 2020).

As abordagens de modelagem por aprendizado de máquina supervisionado requerem a divisão dos dados em dois grupos, de treino e de teste. O conjunto de treino é utilizado para alimentar o modelo, de modo que ele tenha acesso tanto os dados de entrada (utilizados na predição), quanto aos dados de saída (os quais deseja-se prever), para que o modelo possa reconhecer padrões e relações complexas entre os dados de entrada que conduzam aos valores de saída. Em paralelo, o conjunto de teste, dos quais o modelo conhece apenas os dados de entrada, é utilizado para validar o modelo, ao verificar a habilidade do modelo de generalizar para dados previamente desconhecidos. O teste da capacidade de generalização do modelo é um requisito fundamental para garantir sua aplicabilidade para novos dados, para os quais deseja-se realizar previsões (MULLER; GUIDO, 2016).

Sabe-se que a presença de dados de boa qualidade e consistentes são fundamentais para a boa performance de modelos de *Machine Learning*, de modo que um desafio na aplicação na área de vidros consiste na elaboração de dados coletados “com os mesmos protocolos experimentais, incluindo a utilização das mesmas taxas de resfriamento, taxas de aquecimento, carregamento, agitação, cadinhos, forno, tempo de recozimento, entre outros”. (RAVINDER et al., 2021, p.4). O autor ainda acrescenta os desafios de reunir dados coerentes de $T_{liquidus}$ e de viscosidade, esses últimos dependendo da sensibilidade da técnica utilizada. Embora existam bases de dados relacionando a composição de vidros com diversas propriedades, ainda permanece um desafio detectar *outliers*, possibilitando assim predições mais acuradas (LAZAREVIC; KUMAR, 2005), (DOMINGUES et al., 2018). Diversas técnicas e algoritmos para identificação de anomalias foram mapeados, baseando-se no reconhecimento de padrões dos dados, podendo ser divididos em algoritmos que utilizam métricas de proximidade ou técnicas de conjunto (ZHAO et al., 2019).

Os modelos de *Machine Learning*, por meio dos dados de treinamento, conseguem capturar correlações, padrões e interações entre os dados que não são detectados por tradicionais análises de sensibilidade e por métodos de estatística descritiva. Em contrapartida, apresentam a desvantagem de funcionarem como uma ‘caixa preta’ com a funcionalidade de fazer previsões, mas cuja lógica escapa à compreensão humana. Recentemente, a área de modelos de ML explicáveis vem ganhando destaque ao tornar os modelos mais transparentes.

...há uma categoria de algoritmos que pode, em princípio, induzir modelos interpretáveis. Ao explicar como eles chegam às suas decisões, modelos explicáveis podem fornecer *insights* úteis, auxiliando na descoberta de conhecimentos novos e relevantes. (ALCOBAÇA et al., 2020, p.93).

Particularmente, o *Shapley Value*, originário da teoria dos jogos, é um conceito que pode ser aplicado ao contexto de ML para explicar a contribuição relativa de cada característica nas previsões realizadas pelo modelo. Essa metodologia tem se mostrado útil na aplicação em vidros óxidos (ALCOBAÇA et al., 2020), fornecendo valiosos *insights* do papel de formadores e modificadores em determinadas propriedades de materiais vítreos (RAVINDER et al., 2021), (CASSAR et al., 2021a).

Machine Learning tem desempenhado um papel crucial no desenvolvimento de novos vidros, particularmente na abordagem utilizando algoritmos interpretáveis. A transparência do processo decisório proporcionada por esses modelos permite a compreensão de fatores físicos/químicos que influenciam nas propriedades de vidros, possibilitando avaliar o impacto de cada componente em dada propriedade, sendo de grande relevância para área de pesquisa (CASSAR et al., 2021a). Estudos anteriores demonstram que os formadores de rede, modificadores e intermediários podem ter papéis diferentes a depender da propriedade avaliada (CASSAR et al., 2021a).

Além disso, a aplicação de modelos híbridos, que combinam modelos físicos com técnicas de ML, amplificam as capacidades preditivas. Integrando a física dos materiais com algoritmos de aprendizado de máquina, é possível prever comportamentos complexos em diferentes condições. Essa abordagem híbrida equilibra a precisão dos modelos físicos com a flexibilidade e eficiência dos modelos de ML (JIRASEK; HASSE, 2023).

Modelos híbridos têm o potencial de combinar o melhor dos dois mundos: do ML, a alta flexibilidade de algoritmos e estratégias avançadas de treinamento, e da física, a conformidade com as leis da física e o poder preditivo associado. (JIRASEK; HASSE, 2023, p.40).

3 METODOLOGIA

3.1 CRIAÇÃO DAS BASES DE DADOS

Os dados utilizados no trabalho foram obtidos das bases *SciGlass* e *Interglad*, das quais foram coletadas as composições de vidros, associando-as à temperatura *liquidus* e à viscosidade, medida em diferentes temperaturas. Posteriormente, foi realizada a preparação e limpeza dos dados, seguida de uma seleção dos descritores, conforme descrito nos parágrafos subsequentes. Inicialmente, foram obtidas bases de dados contendo a viscosidade de vidros medidas em uma única temperatura, a qual estava identificada no nome da base de dados, como por exemplo “Viscosity_73K”, sendo cada valor de viscosidade associado à composição do formador de vidro, por sua vez identificado por uma ID. As bases de dados iniciais foram reunidas em uma única base de dados, denominada doravante Base 1, como ilustrado no esquema da Figura 3.1.

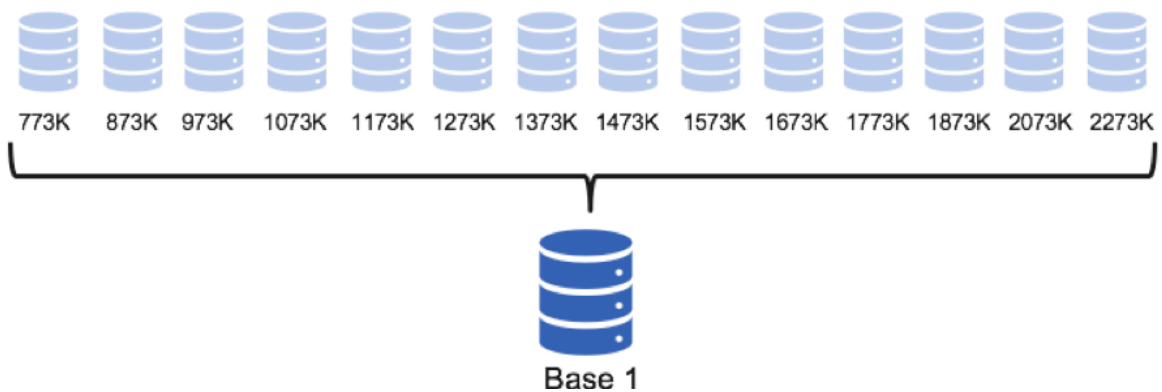


Figura 3.1 – Formação da base de dados contendo a composição e viscosidade em sua respectiva temperatura de medida.

Desse modo, a Base 1 passou a conter uma coluna com o valor da viscosidade, associada à uma coluna contendo a temperatura em que a viscosidade foi medida e associada à composição identificada por uma ID. Cabe ressaltar que uma ID pode estar duplicada, no caso de haver uma mesma composição de vidros, mas que foi medida em temperaturas diferentes, resultando, portanto, em viscosidades diferentes. Essa base de dados contém a composição de 83.388 formadores de vidros, das quais algumas tiveram sua viscosidade medidas em uma única temperatura, enquanto outras foram medidas em até 13 temperaturas diferentes.

A segunda base, doravante denominada Base 2 contém a composição de 45545 formadores de vidro (identificadas por uma ID), associada à sua temperatura $T_{liquidus}$.

Essa base foi combinada com a Base 1, formando a Base 3 como ilustrado no esquema da Figura 3.2.

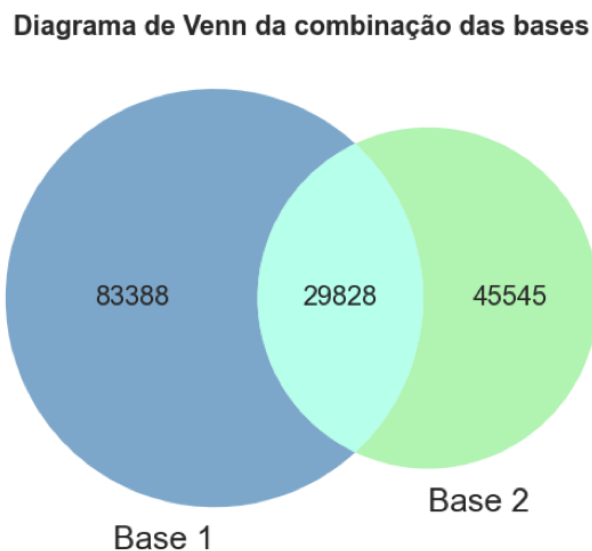


Figura 3.2 – Diagrama de Venn representando a junção das bases 1 e 2 para formar a base de dados 3.

Após a combinação da Base 1 contendo 83.388 e da Base 2 contendo 45545 foi formada a Base 3, contendo 29828 composições. A queda na quantidade de composições é explicada pelo fato que nem todos os dados presentes em um banco de dados tem seu respectivo valor no outro banco de dados, como pode ser observado no diagrama de Venn ilustrado no esquema da Figura 3.2.

Além disso, determinados vidros da base de dados apresentavam valores nulos para todos os compostos, evidentemente correspondendo a erros de digitação, sendo, portanto, retirados do banco de dados. De maneira análoga, o banco de dados apresentava compostos que não estão presentes em nenhuma das composições de vidro do banco de dados e foram, por essa razão, removidos da base de dados. Assim, a Base 3 passou a conter 72 óxidos que podem estar presentes ou não na composição de cada um dos vidros. As composições de vidros utilizadas para as análises subsequentes apresentam entre 1 e 14 compostos, de modo que a maioria das células são nulas. Por exemplo, um vidro contendo 100% SiO_2 em sua composição terá 71 valores nulos para os demais compostos. Cabe destacar que dos 29828 dados da Base 3, estão contidas 6710 composições diferentes.

3.2 SELEÇÃO DE DESCRITORES (COMPOSTOS)

A base de dados 3 antes da limpeza apresentava os seguintes componentes em ordem alfabética: Al_2O_3 , AlF_3 , As_2O_3 , BaF_2 , BaO , Bi_2O_3 , B_2O_3 , Br , CaCl_2 , CaF_2 ,

CaO , CdO , CeO_2 , Cl , Co_2O_3 , Co_3O_4 , CoO , Cr_2O_3 , Cs_2O , Cu_2O , CuO , Fe_2O_3 , FeO , F , F_2 , Ga_2O_3 , Gd_2O_3 , GeO_2 , H_2O , HfO_2 , In_2O_3 , K_2O , La_2O_3 , Li_2O , MgF_2 , MgO , $MnCl_2$, MnO , MnO_2 , MoO_3 , Na_2O , NaF , Nb_2O_3 , Nb_2O_5 , Nd_2O_3 , NiO , NO_2 , P_2O_5 , PbO , Pr_6O_{11} , Sc_2O_3 , SeO_2 , SiF_4 , SiO_2 , SnO_2 , SO_3 , SO_4 , SrF_2 , SrO , S , Ta_2O_5 , ThO_2 , TiO_2 , V_2O_5 , WO_3 , Y_2O_3 , YF_3 , ZnO e ZrO_2 .

No presente trabalho estamos interessados em avaliar apenas vidros óxidos, de modo que retiramos da base de dados os compostos fluorados, clorados, água e compostos voláteis: AlF_3 , BaF_2 , Br , $CaCl_2$, CaF_2 , Cl , F , F_2 , H_2O , MgF_2 , $MnCl_2$, NaF , S , SiF_4 , SO_3 , SO_4 , SrF_2 , S e YF_3 .

Em seguida, foi feita uma análise para verificar a frequência de cada um dos elementos restantes na composição dos vidros. O histograma da Figura 3.3 apresenta a contagem dos elementos presentes em ao menos 20 composições de vidros.

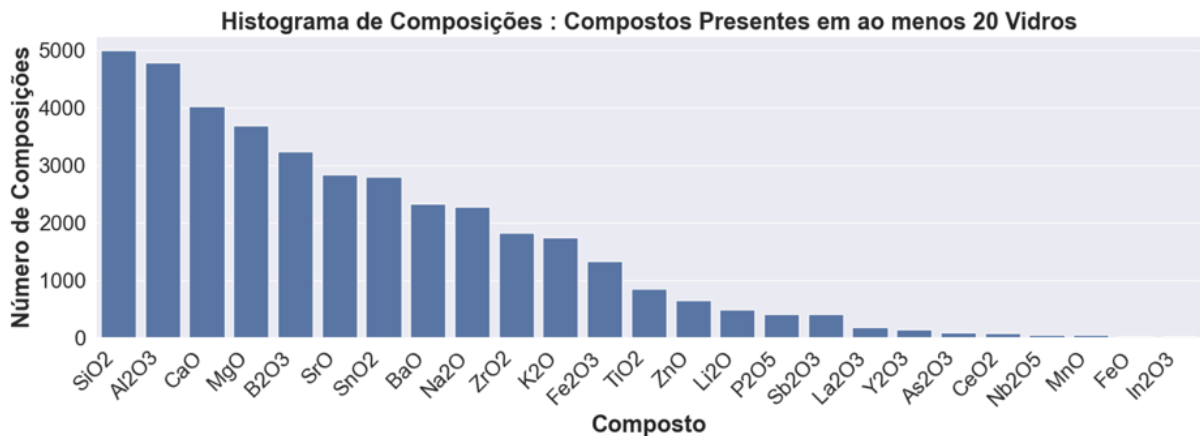


Figura 3.3 – Histograma contendo compostos que estão presentes em ao menos 20 composições de vidros.

É possível observar que a sílica (SiO_2) está presente em 4996 composições, seguida pela alumina (Al_2O_3) presente em 4788 e do óxido de cálcio presente em 4030. No outro extremo é possível observar o óxido de chumbo (In_2O_3) presente em 26 composições. No histograma da Figura 3.4 estão apresentadas as frequências de elementos que aparecem em menos de 20 composições de vidro e que, em razão de sua raridade optou-se por retirá-los da análise. O óxido de gadolínio (Gd_2O_3) aparece em apenas 13 composições.

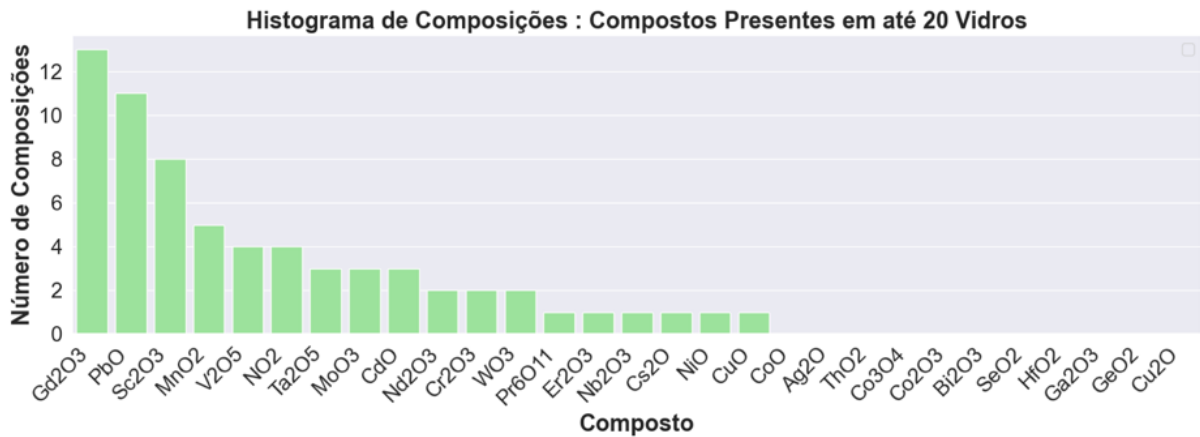


Figura 3.4 – Histograma contendo compostos que estão presentes em menos de 20 composições de vidros.

Os elementos retirados nessa etapa foram: Gd_2O_3 , PbO , Ta_2O_5 , CdO , V_2O_5 , MnO_2 , Sc_2O_3 , WO_3 , ThO_2 , Nd_2O_3 , NO_2 , Cu_2O , Co_3O_4 , MoO_3 , NiO , CuO , Cr_2O_3 , GeO_2 , Pr_6O_{11} , Ga_2O_3 , Nb_2O_3 , Cs_2O , Er_2O_3 , Bi_2O_3 , Co_2O_3 , HfO_2 , SeO_2 , Ag_2O , CoO . A razão dos elementos : CoO , Ag_2O , ThO_2 , Co_3O_4 , Co_2O_3 , Bi_2O_3 , SeO_2 , HfO_2 , Ga_2O_3 , GeO_2 , Cu_2O não terem sido filtrados anteriormente, embora apareçam em 0 composições de vidros pode ser explicada pelo fato de que inicialmente eles apareciam em algumas composições, mas na etapa de eliminação dos compostos fluorados, clorados, da água e dos compostos voláteis foram sendo removidos dos dados, por estarem em composições associadas com elementos não óxidos e que, portanto, não estão no escopo de análise do presente estudo. A eliminação de elementos raros tem por objetivo eliminar compostos pouco representativos e que, podem induzir erros indesejados na fase de treino de modelos de *Machine Learning*.

3.3 OBTENÇÃO DA VISCOSIDADE EM $T_{liquidus}$

Nessa etapa foram primeiramente selecionadas as composições de formadores de vidros cuja viscosidade foi medida em ao menos 4 temperaturas diferentes, resultando em 4550 composições de vidros diferentes para os quais foi encontrada a viscosidade em $T_{liquidus}$, utilizando a equação VFTH $\log \eta(T) = A + B/(T - T_0)$. Os parâmetros A, B e T_0 ótimos foram encontrados para cada composição, utilizando uma função de otimização com 2000 iterações, permitindo convergir para valores de A, B e T_0 que permitem um bom ajuste da função VFTH. Para certos casos, em que a convergência não foi possível mesmo após 2000 iterações, foi utilizada uma simplificação da equação VFTH, mantendo o parâmetro A constante e igual a -5, em conformidade com o que foi anteriormente explicado na sessão de fundamentos teóricos. Ainda nessa etapa, foi observado que 1713 composições apresentavam medidas de viscosidade em 2 ou 3 temperaturas diferentes. As

composições, cuja $T_{liquidus}$ distava até 100K de umas das temperaturas em que foi feita a medida foram ajustadas por uma função linear. A função linear permite aproximar o ajuste da função VFTH em um intervalo de temperaturas de aproximadamente 100K. Assim, foi possível calcular a viscosidade em $T_{liquidus}$ para 400 composições das 1713 contendo apenas 2 ou 3 medidas experimentais.

3.4 DESCRIÇÃO DO DATASET FINAL

O banco de dados utilizado para as análises subsequentes contém 4950 composições relacionando à $T_{liquidus}$ e a viscosidade em $T_{liquidus}$. Os compostos presentes estão exibidos no histograma da Figura 3.5, sendo o SiO_2 o mais presente (99,8% das composições) e o FeO o mais raro, presente em apenas 25 das 4950 composições (0.5%).

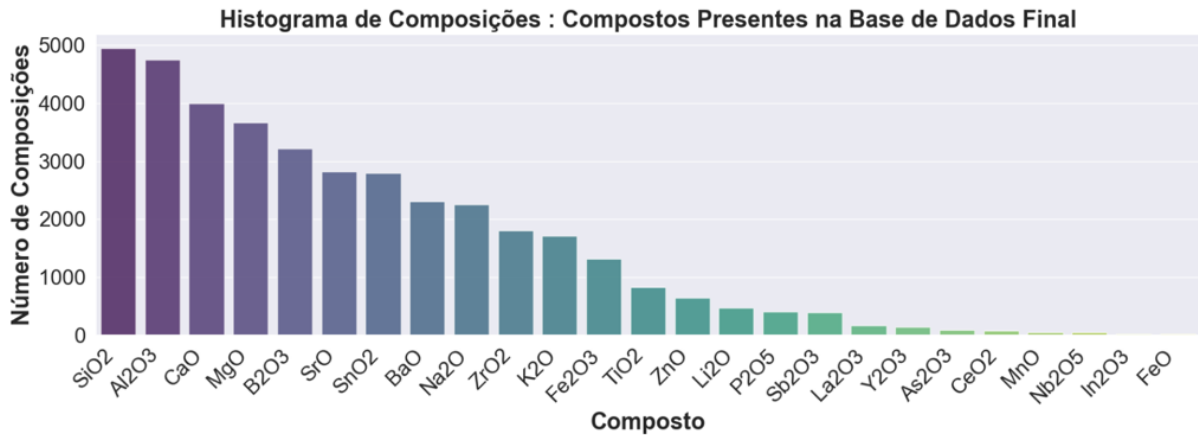


Figura 3.5 – Histograma da frequência dos 25 óxidos presentes nas composições de vidros.

Uma composição formadora de vidros presente na base de dados pode conter entre 2 e 14 compostos óxidos de um total de 25 possíveis como pode ser observado no histograma da Figura 3.5, de modo que a maior parte da tabela é preenchida com zeros. Como pode ser observado no histograma da Figura 3.6 a grande maioria dos vidros apresenta entre 5 e 11 compostos diferentes em sua composição, reforçando o interesse na utilização de algoritmos de ML para captar possíveis interações entre os componentes e seu efeito nas propriedades estudadas.

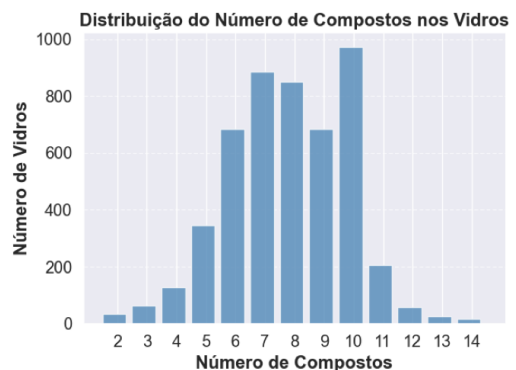


Figura 3.6 – Histograma do número de vidros com determinado número de compostos óxidos.

A descrição estatística dos valores de viscosidade em $T_{liquidus}$ e de $T_{liquidus}$ presentes no *dataset* estão dispostos na Tabela 3.1.

Tabela 3.1 – Estatística descritiva dos valores de viscosidade e temperatura *liquidus*.

	Viscosidade [Pa.s]	$T_{liquidus}$
Média	$4,08 \cdot 10^7$	1375,6
Desvio Padrão	$1,04 \cdot 10^9$	132,2
Mínimo	$1,35 \cdot 10^{-1}$	882,1
Mediana	$1,51 \cdot 10^4$	1383,1
Percentil 75	$4,35 \cdot 10^4$	1473,1
Máximo	$4,36 \cdot 10^{10}$	1873,1

Na Tabela 3.1 é possível observar um valor médio de viscosidade em $T_{liquidus}$ de $4,08 \cdot 10^7$ Pa.s e, portanto, bastante elevado quando comparado com o valor da mediana de $1,51 \cdot 10^4$ Pa.s. Essa discrepância é explicada pela sensibilidade da média à valores elevados, sobretudo para propriedades como a viscosidade que se estendem por várias ordens de grandeza. De todo modo, foi verificado que, das 4950 composições apenas 73 exibem viscosidade em $T_{liquidus}$ maior que 10^6 Pa.s (valor de referência de vidros de sílica pura).

3.5 MACHINE LEARNING

Inicialmente, técnicas de aprendizado não supervisionado foram empregadas para eliminar valores anômalos do *dataset*, sendo testados algoritmos que utilizam métrica de distância, de divisão de dados em subconjuntos calculando a probabilidade de cada instância ser um *outlier* ou a combinação de estratégias de eliminação, na qual cada uma contribui para pontuar o grau de anomalia de uma dada instância.

Em seguida, a metodologia adotada neste estudo visou a avaliação comparativa de quatro modelos de *Machine Learning*, a saber *Random Forest*, *XGBoost*, *Support Vector Machines* (SVM) e *Multilayer Perceptron* (MLP) para a predição da viscosidade em $T_{liquidus}$, $T_{liquidus}$ e GFA. O processo metodológico foi dividido em etapas distintas, conforme descrito a seguir:

3.5.1 Seleção de Algoritmos

Os algoritmos selecionados para análise supervisionada - *Random Forest*, *XGBoost*, SVM e MLP - foram escolhidos com base em suas características distintas e comprovada eficácia em problemas de regressão. Cada algoritmo foi considerado devido à sua capacidade de lidar com diferentes nuances do conjunto de dados, como complexidade não linear, alta dimensionalidade e com dados esparsos.

3.5.2 Hiperparâmetrização

Os hiperparâmetros dos modelos foram ajustados com o intuito de otimizar o desempenho dos modelos. Para o *Random Forest*, foram ajustados parâmetros como o número de árvores e a profundidade máxima. No *XGBoost*, parâmetros como taxa de aprendizado e número de estimadores foram sintonizados. Para o SVM, a escolha do kernel do parâmetro gamma e C associados foi realizada. Já para o MLP, a estrutura da rede neural, número de camadas e unidades, foi objeto de otimização. Informações adicionais e detalhadas do impacto dos parâmetros nos modelos de *Machine Learning* estão amplamente discutidas na literatura, não sendo objeto desse trabalho (MULLER; GUIDO, 2016).

3.5.3 Avaliação do Desempenho

A avaliação do desempenho dos modelos foi conduzida por meio da divisão em dados de treino e teste, permitindo a avaliação da capacidade de generalização dos modelos. Foram utilizadas métricas padrão de regressão, tais como erro médio absoluto (MAE), erro médio quadrático (MSE) e coeficiente de determinação (R^2), para quantificar a acurácia e a precisão das predições.

3.5.4 Interpretação dos modelos

A interpretação dos modelos que apresentaram desempenho notável foi conduzida por meio da técnica de *Shapley Values*, oferecendo uma compreensão aprofundada das contribuições individuais das variáveis na predição da viscosidade de vidros. Essa análise proporcionou *insights* significativos sobre quais compostos exercem maior impacto nas predições.

3.5.5 Extensão do modelo

Os modelos foram estendidos para prever a GFA por duas metodologias. A primeira delas consistiu em, primeiramente, prever a $T_{liquidus}$ e a viscosidade em $T_{liquidus}$ para as composições de formadores de vidro e, em seguida, calcular o GFA, por meio da incorporação do modelo físico ‘Jezica’ tratado na sessão de fundamentação teórica.

A segunda metodologia consistiu em calcular o GFA utilizando a $T_{liquidus}$ e a viscosidades em $T_{liquidus}$, para em seguida dividir os dados em treino e teste e criar um modelo preditivo para a GFA. As duas metodologias estão resumidas no esquema da Figura 3.7. Cabe destacar que essa etapa está relacionada à criação dos modelos, tendo sido necessário empregar a equação VFTH para encontrar a viscosidade em $T_{liquidus}$. Uma vez criados, os modelos precisam ser alimentados apenas com a composição dos vidros para fazer as previsões.

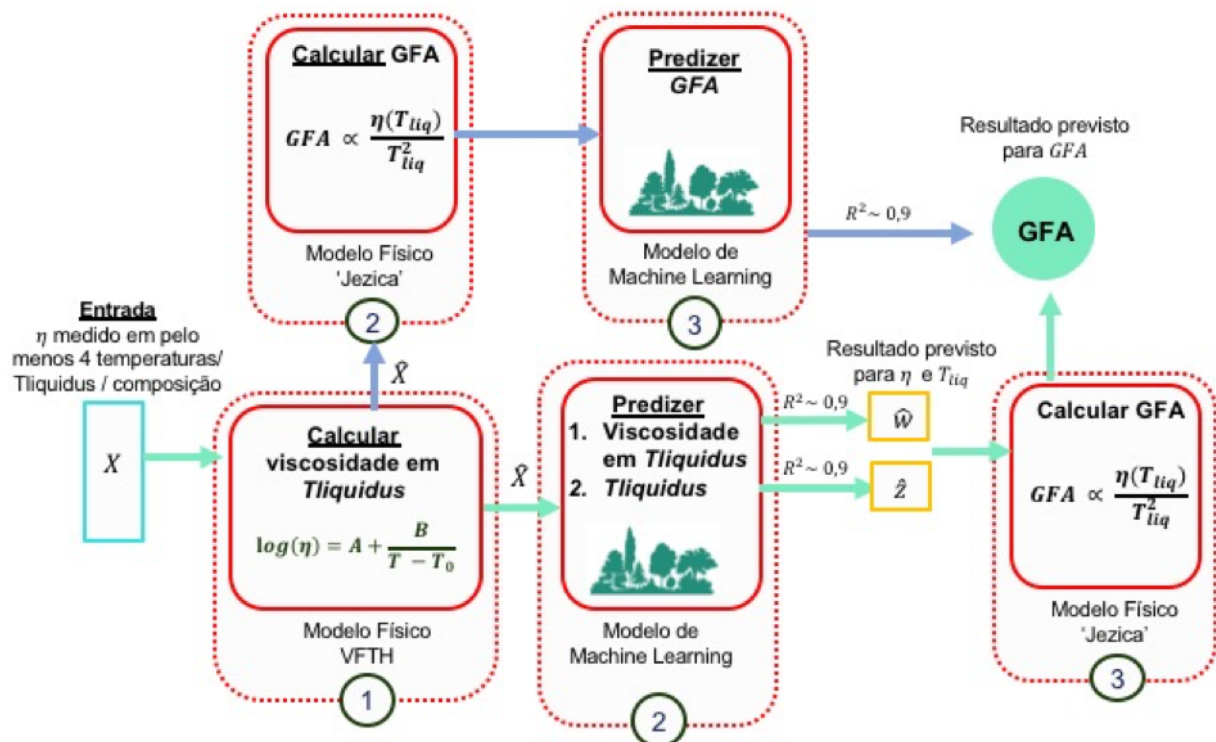


Figura 3.7 – Metodologia para criação dos modelos para prever GFA, por meio da incorporação de modelos físicos e de *Machine Learning*.

Adicionalmente, foi atribuída a cada composição um valor de GFA que possibilitou uma classificação da habilidade de formar vidros segundo o critério estabelecido no artigo '*Viscosity and liquidus-based predictor of glass-forming ability of oxide glasses*' e apresentado na Tabela 3.2:

Tabela 3.2 – Classificação da habilidade de formar vidros.

$\log(\text{GFA})$ [Pa.s/K ²]	Habilidade de vitrificação
>-1	Excelente
-1 a -4	Boa (vidros comerciais)
-4 a -7	Moderada à difícil
<-7	Extremamente difícil

4 RESULTADOS E DISCUSSÕES

A presente pesquisa foi conduzida com o objetivo de obter a viscosidade em $T_{liquidus}$ para diferentes composições de formadores de vidros, com o intuito de em seguida avaliar a capacidade de modelos de *Machine Learning* para prever a viscosidade e GFA para composições que não estão presentes na base de dados utilizada. A viscosidade em $T_{liquidus}$ é considerada um parâmetro relevante na predição da habilidade de formar vidros. Cabe ainda ressaltar que esse trabalho explora o fato de que as propriedades de materiais vítreos são função da sua composição nominal. Conforme explicitado anteriormente dois bancos de dados foram utilizados na realização das análises. O primeiro banco de dados contém a composição dos formadores de vidros, relacionando-as com a viscosidade medida em determinada temperatura. As temperaturas nas quais a viscosidade foi medida estão em intervalo de valores entre 773 e 2273 K, como exibido nos gráficos da Figura 4.1.

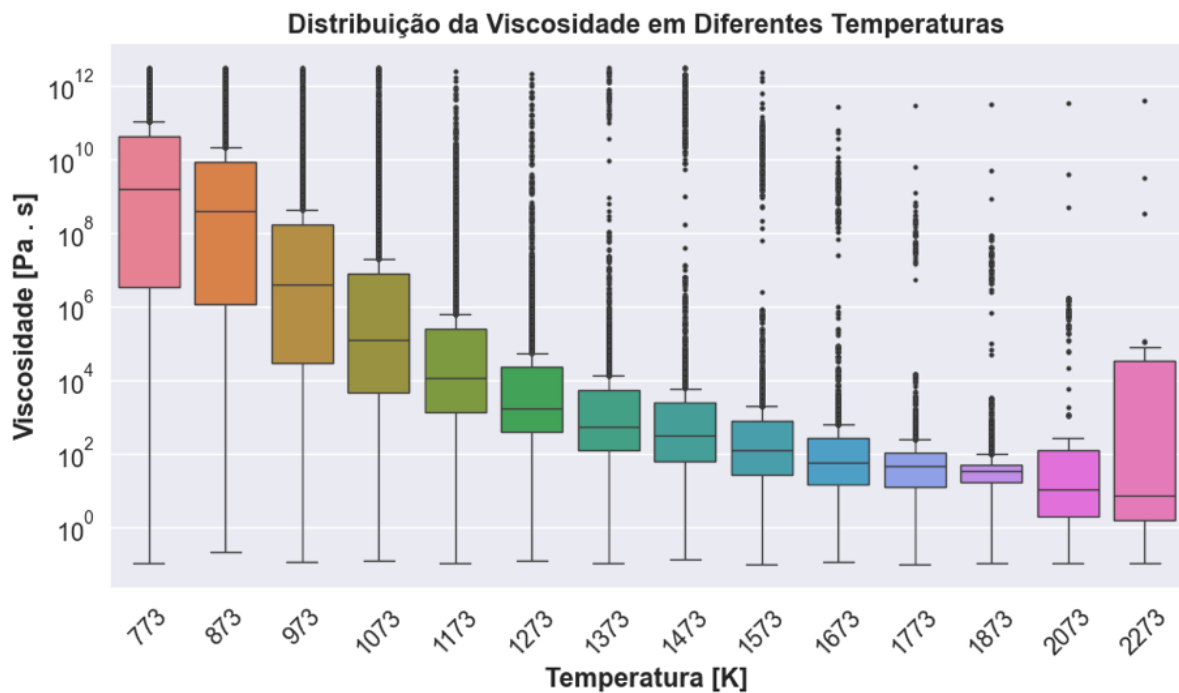


Figura 4.1 – *Box plot* da distribuição da viscosidade em diferentes temperaturas.

É possível observar que a viscosidade pode variar até mais de 12 ordens de grandeza em função da temperatura de medição e da composição dos formadores de vidros. Além disso, é possível observar uma tendência decrescente da viscosidade com o aumento da temperatura em que ela foi medida. Adicionalmente, se observamos para cada temperatura, a mediana da viscosidade (traço retilíneo no *box plot*) constatamos que a viscosidade decresce mais rapidamente, para variações de 100K, em temperaturas baixas. Em temperaturas

altas o decréscimo da viscosidade com a temperatura é mais lento. Essa base de dados contém a composição de 83.388 vidros, das quais algumas tiveram sua viscosidade medidas em uma única temperatura, enquanto outras foram medidas em até de 13 temperaturas diferentes.

A segunda base de dados contém a composição dos formadores de vidros e a temperatura $T_{liquidus}$ para 45545 composições. Na base de dados a viscosidade varia no intervalo de 302K a 3162K, tendo uma mediana de 1338K, como pode ser observado no gráfico da Figura 4.2.

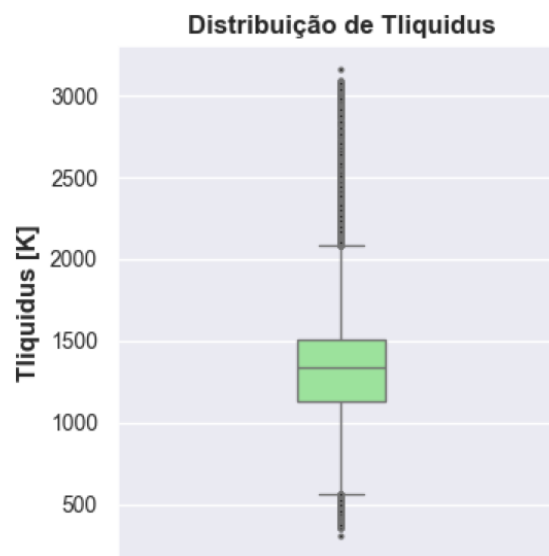


Figura 4.2 – *Box plot* da distribuição de $T_{liquidus}$.

Em uma etapa seguinte as duas bases de dados foram reunidas, correlacionando para cada caso a composição dos formadores de vidro, a viscosidade, a temperatura em que a viscosidade foi medida e a $T_{liquidus}$ do vidro. Cabe ressaltar que composições idênticas terão a mesma $T_{liquidus}$, muito embora o valor da viscosidade possa variar em função da temperatura em que foi feita a medição.

Após a união dos bancos de dados descritas na metodologia restaram 29828 composições de vidros, dos quais 6710 são composições únicas. O histograma da Figura 4.3 permite observar em quantas temperaturas diferentes as composições únicas dos formadores de vidros foi medida.

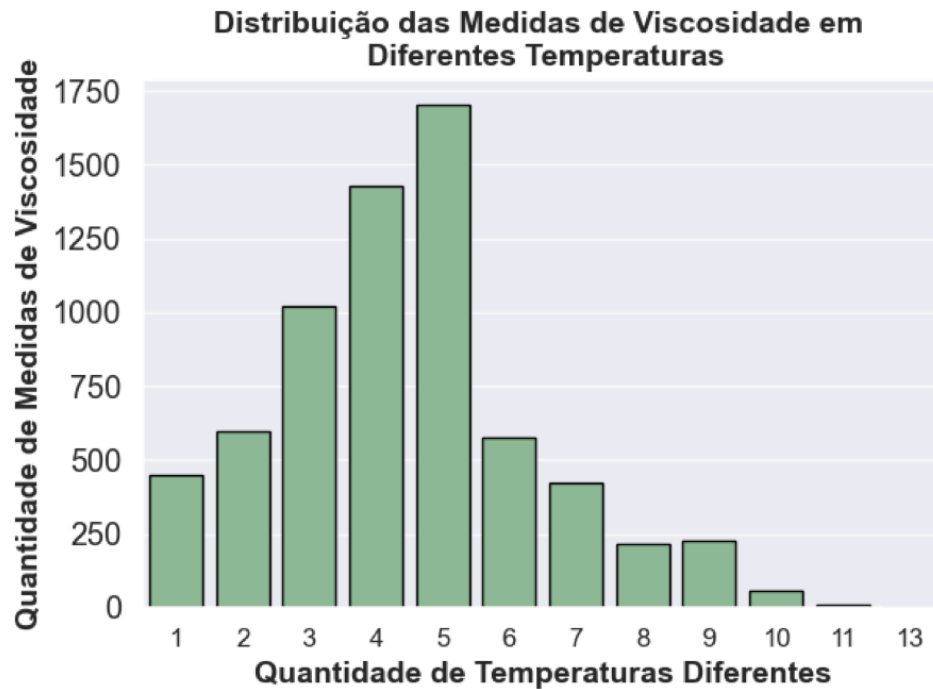


Figura 4.3 – Histograma da quantidade de medidas de viscosidade em função da quantidade de temperaturas de medida da viscosidade.

É possível observar que 447 composições tiveram sua viscosidade medida em uma única temperatura e que um total de 1705 composições tiveram sua viscosidade medida em 5 temperaturas diferentes. Essa análise descritiva é particularmente útil para a adoção de estratégias para encontrar a viscosidade dos formadores de vidros na temperatura *liquidus*. Cabe salientar que as temperaturas em que a viscosidade foi medida não correspondem na grande maioria dos casos à viscosidade em $T_{liquidus}$, de modo que foi necessário ajustar uma função aos dados experimentais que permitisse extrapolar à viscosidade para a $T_{liquidus}$. Para realizar esse ajuste de curva é preciso que a viscosidade de uma determinada composição tenha sido medida em um número mínimo de 4 pontos. Desse modo, não será possível encontrar a $T_{liquidus}$ para formadores de vidros que tiveram suas viscosidades medidas em 1, 2 ou 3 pontos apenas, o que totaliza 2067 composições únicas, de um total de 6710 dados. Com a eliminação dos elementos raros, a base de dados analisada foi composta de 4550 composições diferentes.

As composições com ao menos 4 medidas de viscosidade foram ajustados pela equação VFTH previamente discutida e que, permite prever a viscosidade em função da temperatura, segundo a equação: $\log \eta(T) = A + B / (T - T_0)$. O gráfico da Figura 4.4 apresenta o ajuste da função VFTH para um vidro que teve sua viscosidade medida em 7 pontos experimentais, obtendo um ajuste muito preciso ($R_2 \approx 1,000$). Adicionalmente, os parâmetros A, B em Kelvin e T_0 em Kelvin da função VFTH foram obtidos para cada

uma das composições, constando para um determinado vidro analisado e representado na Figura 4.4.

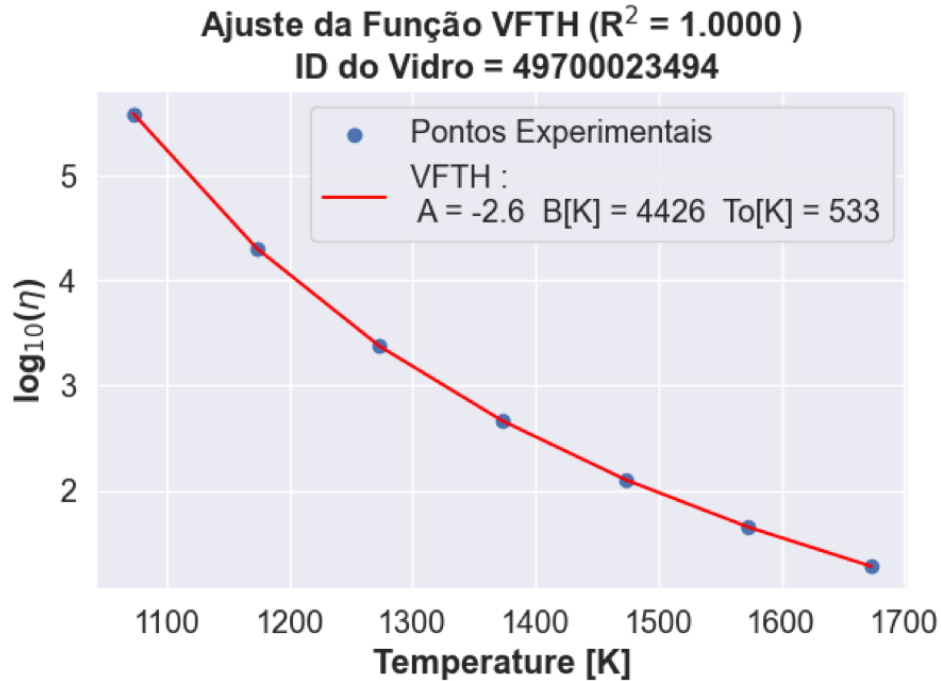
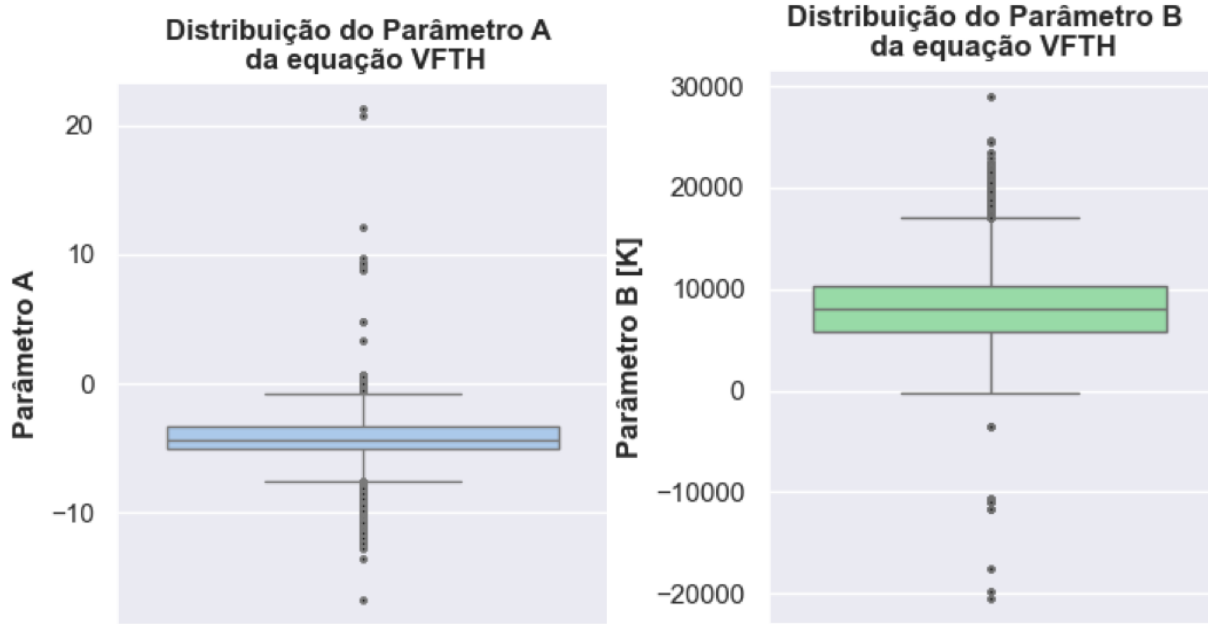


Figura 4.4 – Ajuste da função VFTH para um vidro.

A qualidade do ajuste permite obter elevada confiabilidade nos valores de viscosidade em $T_{liquidus}$ calculados pela extrapolação da função VFTH para a temperatura $T_{liquidus}$. A distribuição dos parâmetros A, B e T_0 estão apresentadas na Figura 4.5. A maior parte dos valores para o parâmetro A se encontram entre -3 e -6, o que está em conformidade com o que foi observado em outros trabalhos (BARRER, 1943), Angell, 1991), dos quais destaco um que usou a simplificação de adotar um valor de A fixo e igual a -5 (NASCIMENTO; APARICIO, 2007). Apenas para 0,3% dos dados o valor de A foi positivo. Por sua vez, a literatura indica uma maior variação do parâmetro B, entre 2000 e 22.000 K (NASCIMENTO; APARICIO, 2007); (DUINE et al., 1991) que está de acordo com a grande maioria (99%) dos vidros ajustados pela equação VFTH. Por fim, o parâmetro T_0 , relacionado com a temperatura na qual a viscosidade é infinita é superior a 0 K.



(a) Distribuição do parâmetro A da equação VFTH.

(b) Distribuição do parâmetro B da equação VFTH.

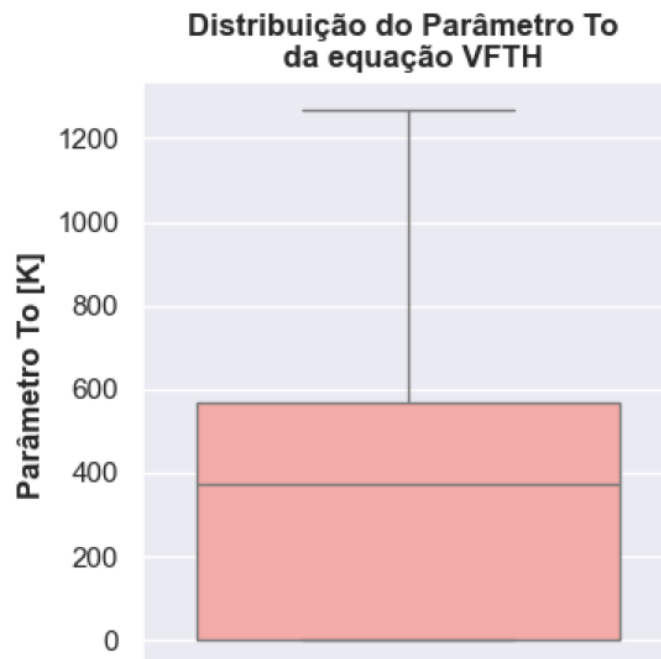
(c) Distribuição do parâmetro T_0 da equação VFTH.

Figura 4.5 – Distribuição dos parâmetros A, B e T_0 da equação VFTH obtidos para diversas composições de vidros.

Os valores discrepantes observados para A e B resultam do fato da função VFTH ter sido ajustada para pontos experimentais por vezes não representativos de uma curva de viscosidade completa, ou seja, que percorreriam pontos por uma vasta amplitude de temperaturas entre T_g e T_l . De todo modo, utilizando pontos experimentais próximos de

$T_{liquidus}$ a função será capaz de estimar com precisão a viscosidade nessa temperatura.

Cabe destacar que correlações entre os parâmetros A e T_0 foram encontradas em trabalhos anteriores sugerindo uma simplificação da equação para dois parâmetros (NASCIMENTO; APARICIO, 2007). A equação $\log \eta(T) = -5 + B/(T - T_0)$ foi utilizada para ajustar vidros para os quais não foi possível convergir para uma solução utilizando três parâmetros em até 2000 iterações. Embora o ajuste da equação simplificada não seja tão bom quanto utilizando três parâmetros, ele é justificável do ponto de vista da acurácia ($R^2 > 0,99$), evitando assim a redução da base de dados e preservando uma boa qualidade de ajuste.

Na Figura 4.6 estão representados dois exemplos de ajuste da equação para os casos limites, nos quais o número de pontos experimentais é elevado (11 pontos) e baixo (4 pontos).

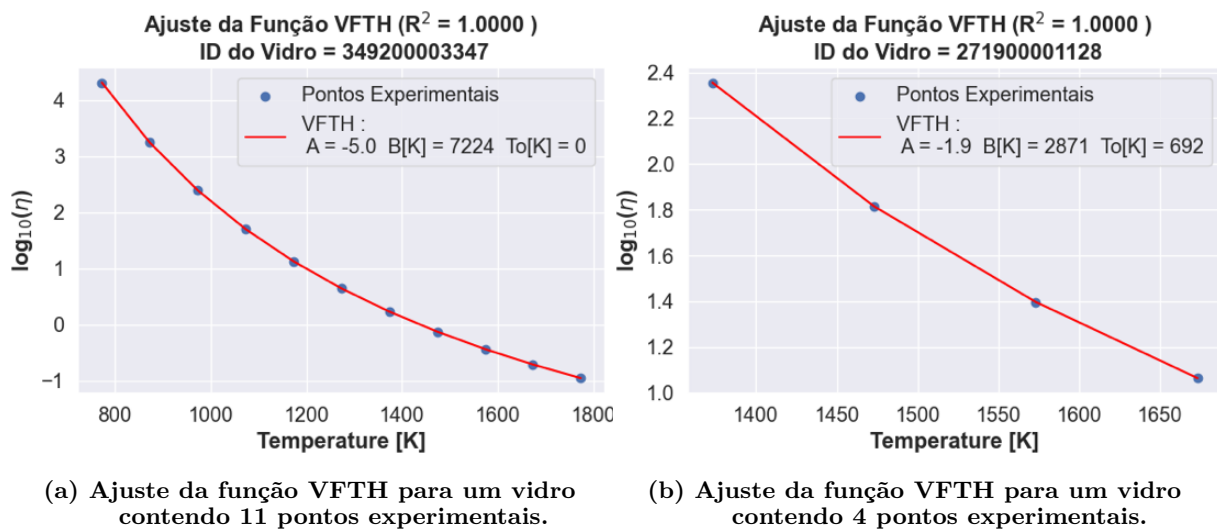


Figura 4.6 – Ajuste da função VFTH para vidros contendo diferentes quantidades de pontos de medida da viscosidade em função da temperatura.

O ajuste no caso em que há grande quantidade de pontos experimentais se mostrou eficiente, cobrindo um intervalo de temperatura de mais de 1000K e uma variação de viscosidade de mais de 5 ordens de grandeza. Nos casos em que há poucos pontos experimentais foi preservada uma boa qualidade do ajuste da equação VFTH. Além disso, os gráficos sugerem que para variações de temperaturas em uma faixa de 200K de intervalo a função apresenta um caráter aproximadamente linear, como pode ser observado entre 1400 e 1600K no gráfico com 4 pontos experimentais ou entre 1600 e 1800K no gráfico com 11 pontos experimentais.

Esse caráter aproximadamente linear da função VFTH para temperaturas com

diferença de até 200K, indica que pode haver boa precisão no ajuste de uma função linear para prever a viscosidade em $T_{liquidus}$. Assim, permitindo reincorporar à base de dados formadores de vidro cuja viscosidade foi medida em apenas 2 ou 3 temperaturas, para as quais não é possível ajustar a equação VFTH.

Os gráficos a seguir representam um exemplo de ajuste de uma reta para prever a viscosidade em $T_{liquidus}$ para um vidro cuja viscosidade foi medida em 2 e 3 pontos experimentais, respectivamente. Dos quais foram seleccionados apenas os vidros para os quais a $T_{liquidus}$ apresenta uma diferença de menos de 100K de um ponto experimental.

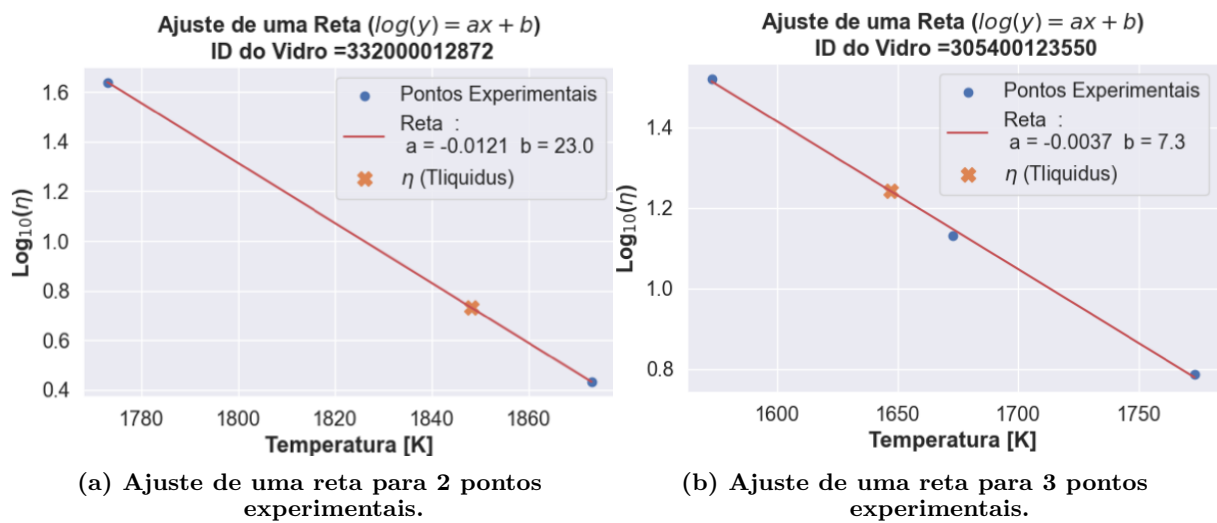


Figura 4.7 – Ajuste de uma função linear para prever a viscosidade de vidros cuja viscosidade foi medida em pontos próximo de $T_{liquidus}$.

Cabe destacar que, diferentemente dos gráficos da Figura 4.7 apresentados acima, não necessariamente a $T_{liquidus}$ se encontra entre dois pontos experimentais, como pode ser visto no gráfico da Figura 4.8.

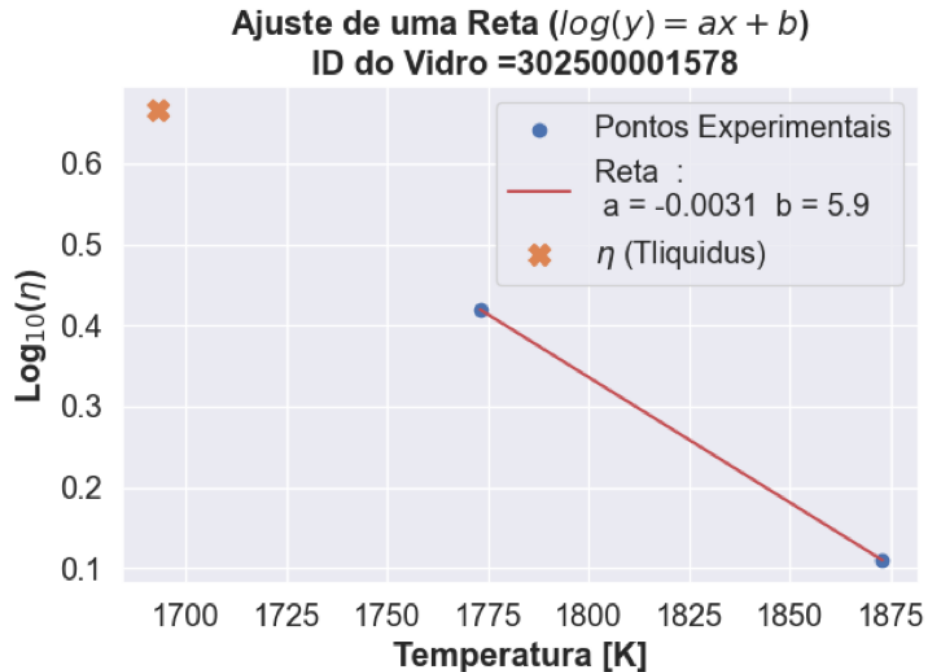


Figura 4.8 – Ajuste de uma função linear para prever a viscosidade de vidros cuja viscosidade foi medida em pontos próximo de $T_{liquidus}$.

No caso de haver dois pontos experimentais a acurácia R^2 não é representativa da qualidade dos ajustes, uma vez que dois pontos experimentais resultam necessariamente em uma reta perfeita. No caso de haver 3 pontos experimentais o R^2 variou de 0,98 à 0,99. Essa predição da viscosidade permitiu recuperar para a base de dados 400 composições de vidros, que adicionadas as 4550 composições precedentes totalizam 4950.

Adicionalmente, as 4950 composições foram avaliadas segundo a sua habilidade de formar vidros, utilizando a equação $GFA = \eta(T_l) / T_l^2$. Os formadores de vidros foram classificados segundo sua GFA em 4 grupos, a saber: excelentes, bons, moderado à difícil, e extremamente difícil, segundo explicado na sessão metodologia. Os 4950 formadores da base de dados foram agrupados segundo a Tabela 4.1.

Tabela 4.1 – Classificação das composições segundo a habilidade de formar vidros.

Habilidade de formar vidros	Quantidade de vidros
Excelente	450
Boa (vidros comerciais)	4195
Moderada à difícil	300
Extremamente difícil	5

Na Tabela 4.1 observa-se que os grupos excelente, bom e moderado à difícil apresentam boa representatividade nos dados. Assim, em um primeiro momento o presente

trabalho estende a classificação de vidros óxidos em relação à sua habilidade de formar vidros para 4950 composições. A base de dados, em teoria, contém as composições de vidros inorgânicos que foram obtidos por resfriamento rápido e tiveram suas propriedades medidas de acordo com a metodologia de cada artigo. Por essa razão, a presença de formadores de vidros classificados como extremamente difíceis de serem formados não eram esperados na base de dados, de modo que sua origem foi rastreada. De fato, as 5 composições classificadas como extremamente difíceis estão reportadas no mesmo artigo (SEGERS et al., 1979) apresentando alta $T_{liquidus}$ e baixa viscosidade em $T_{liquidus}$, o que implicaria em uma baixa GFA. Cabe destacar que no artigo mencionado as viscosidades foram medidas nos líquidos formadores de vidros em diferentes temperaturas, mas esses líquidos não foram vitrificados. Desse modo o estudo visou identificar a viscosidade de misturas líquidas contendo CaO , SiO_2 e MnO . Essa observação exemplifica a importância de buscar metodologias que possibilitem identificar possíveis valores anômalos na base de dados.

Assim, as 4950 composições foram testadas quanto a presença de valores anômalos no *dataset*, utilizando algoritmos não supervisionados para detecção de *outliers* (ZHAO et al., 2019). Dos algoritmos testados o mais eficiente foi o K-Vizinhos mais próximos, ou *K-Nearest Neighbors(KNN)*. O algoritmo KNN mede a distância de uma instância até seus vizinhos mais próximos. Instâncias que têm vizinhos mais distantes são consideradas mais prováveis de serem *outliers*. Muitos dos valores identificados pelo algoritmo como potenciais *outliers* foram também identificados como valores anômalos baseando-se no conhecimento da viscosidade em $T_{liquidus}$ de algumas composições, destacando-se aqui vidros com viscosidade em $T_{liquidus}$ acima de 10^8 Pa.s, as quais estão muito acima de valores reportados na literatura. A proporção de possíveis *outliers* foi considerada como aproximadamente 10% dos valores, sendo eliminadas 439 composições nessa etapa.

Na etapa subsequente, as 4511 composições de vidros foram divididas em dois grupos: um deles contendo 3383 composições (75% dos dados) foi utilizado para treinar o modelo. O restante dos dados (25%) foi utilizado para validação do modelo, segundo apresentado na sessão metodologia. As quantidades dos compostos (entre 0 e 100%) para cada formador de vidro foram utilizadas como descritores para prever a viscosidade em $T_{liquidus}$. Cabe destacar que os hiperparâmetros dos modelos foram previamente ajustados para otimizá-lo.

O gráfico da Figura 4.9 permite comparar os valores de viscosidade $T_{liquidus}$ das composições reservadas para teste (25% dos dados) com as previsões geradas por um modelo de *Random Forest*.

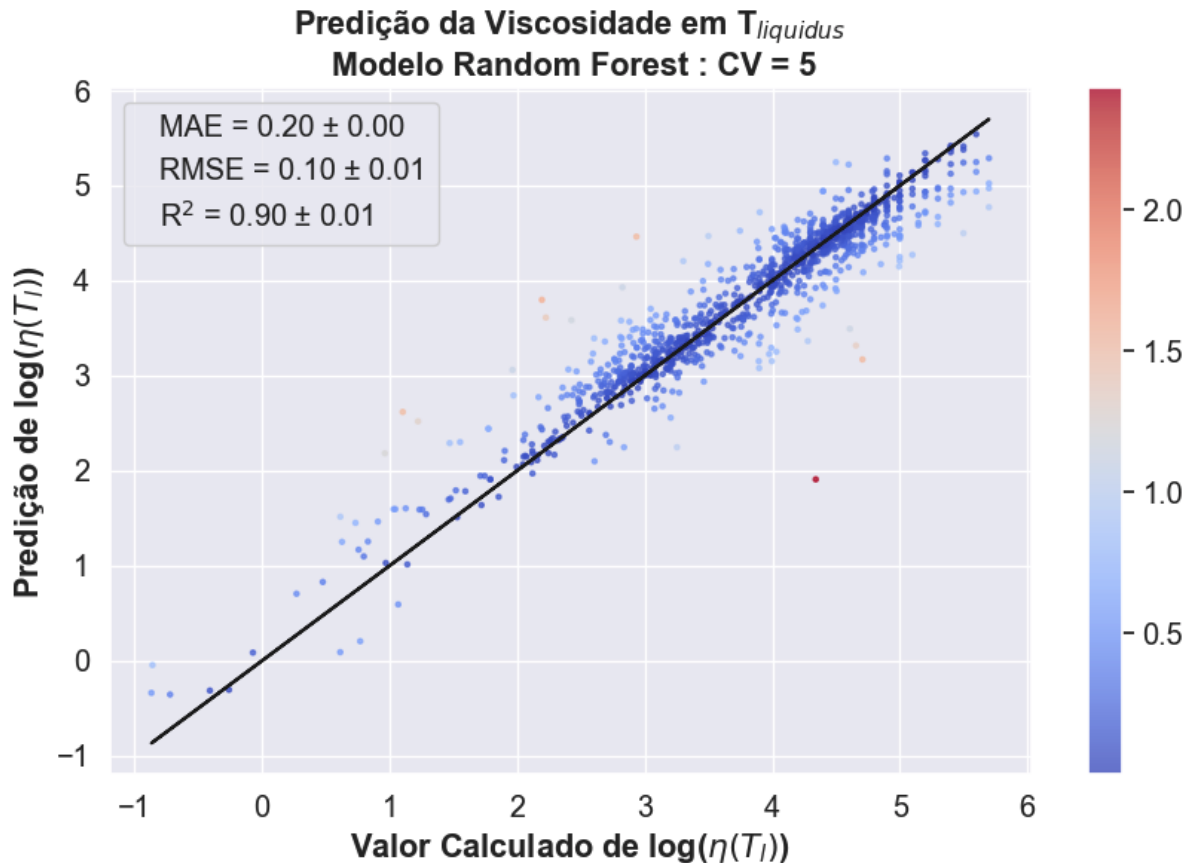


Figura 4.9 – Gráfico dos valores previstos X valores reais de viscosidade em $T_{liquidus}$ utilizando o modelo *Random Forest*.

A linha diagonal (primeira bissetriz) representa uma correspondência perfeita entre as previsões e os valores reais. Desvios da linha diagonal indicam discrepâncias entre as previsões do modelo e os valores verdadeiros, sendo os pontos mais distantes do valor verdadeiro representados na cor vermelha e os mais próximos em azul escuro. A métrica de $R^2 = 0,90 \pm 0,1$ indica uma boa capacidade preditiva do modelo. Os valores médios das métricas de avaliação, exibidos na Figura 4.9, foram obtidos utilizando 5 diferentes grupos de treino e teste, sempre na proporção de 75 e 25% respectivamente, o que permite obter resultados menos dependentes de uma única divisão dos dados em treino e teste. Além disso, é possível observar que o modelo conserva sua precisão para valores de viscosidades em $T_{liquidus}$ extremos, sejam muito baixos ou muito elevados. Esse fator indica uma boa sensibilidade do modelo a variações na composição e, em seu impacto na viscosidade.

O modelo *XGBoost* também foi avaliado quanto à sua capacidade de predição da viscosidade em função dos hiperparâmetros de profundidade máxima e taxa de aprendizado como pode ser visto na Figura 4.10.

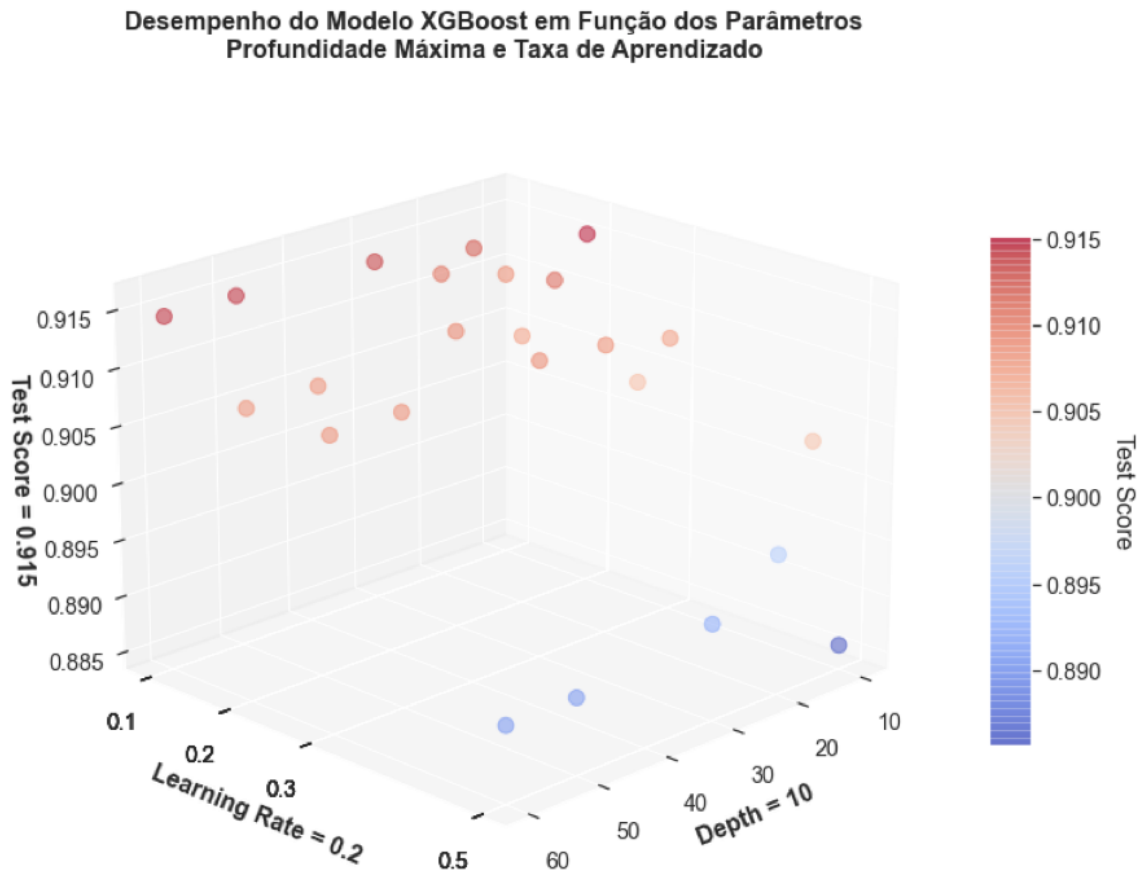


Figura 4.10 – Performance do modelo *XGBoost* para prever a viscosidade em $T_{liquidus}$ em função dos hiperparâmetros selecionados.

Pode ser observado que a performance do modelo oscila entre 0,89 e 0,92 dependendo dos hiperparâmetros utilizados. De todo modo, essa análise é sensível ao modo como foi feita a divisão dos dados entre treino e teste, uma vez que não foi realizado um processo de validação cruzada. O intuito é apenas demonstrar que os modelos *Random Forest* e *XGBoost* apresentam performances comparáveis na predição da viscosidade em função da composição. Os demais modelos testados, a saber SVM ($R^2 = 0,75$) e o MLP ($R^2 = 0,60$) tiveram performances muito inferiores quando comparados aos modelos anteriores.

Em seguida, foi feita uma avaliação para verificar quais dos compostos impactaram mais na previsão da viscosidade utilizando a metodologia *Shapley* detalhada na sessão de fundamentos teóricos. A seguir está apresentada na Figura 4.11 os valores absolutos do impacto de cada composto da predição da viscosidade, em relação ao valor médio das previsões.

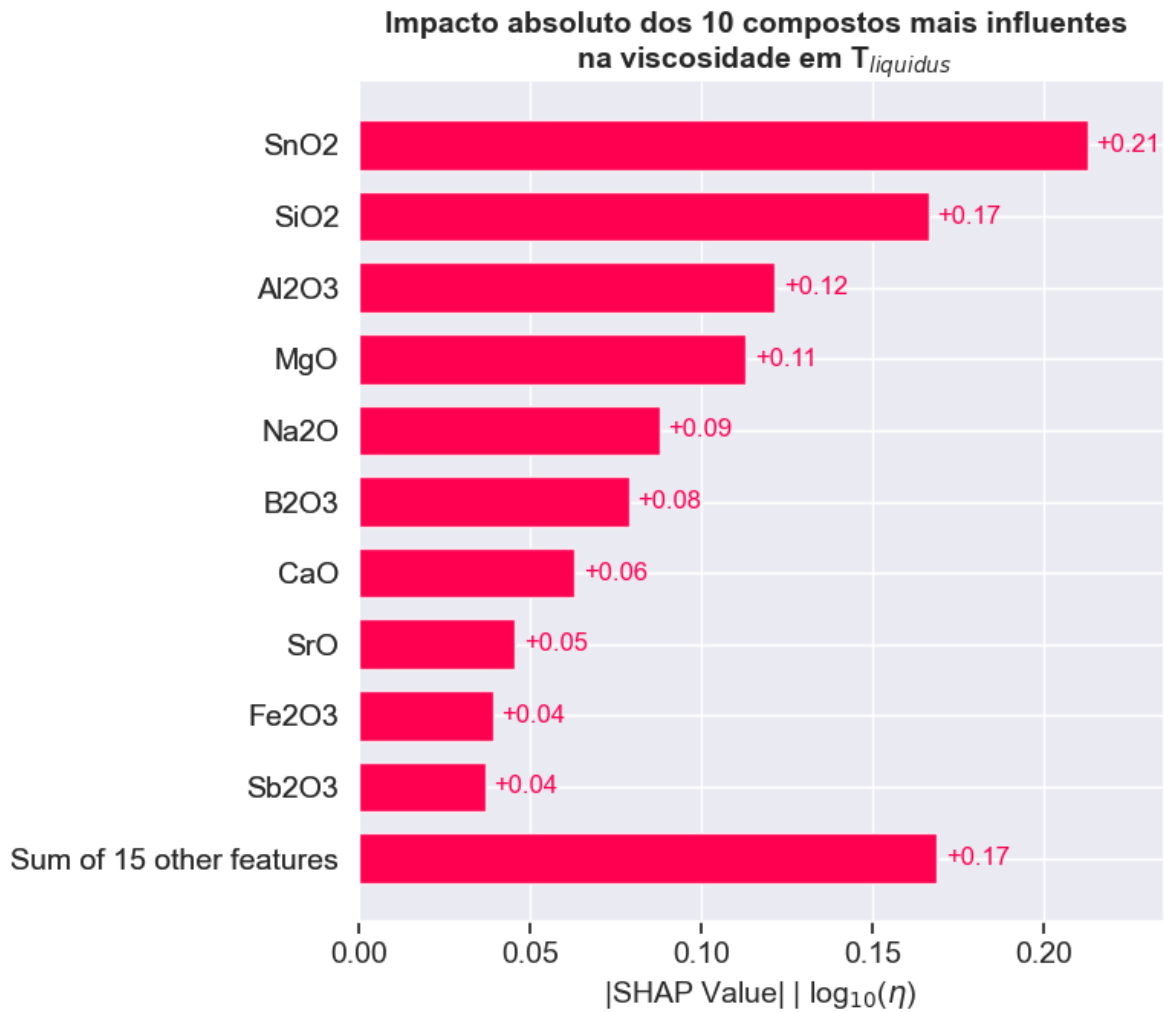


Figura 4.11 – Impacto absoluto dos compostos na viscosidade em $T_{liquidus}$ avaliado pela metodologia *Shapley*.

Pode ser verificado que dentre os compostos mais influentes temos formadores de vidro (SiO_2 , B_2O_3 , Sb_2O_3), intermediários (SnO_2 e Al_2O_3) e modificadores (Na_2O , MgO , CaO , SrO e Fe_2O_3). No entanto, o gráfico da Figura 4.11 informa pouco sobre como cada composto impacta na previsão da viscosidade, uma vez que se baseia no valor absoluto. O gráfico da Figura 4.12 permite uma melhor visualização do impacto de cada composto para a previsão da viscosidade.

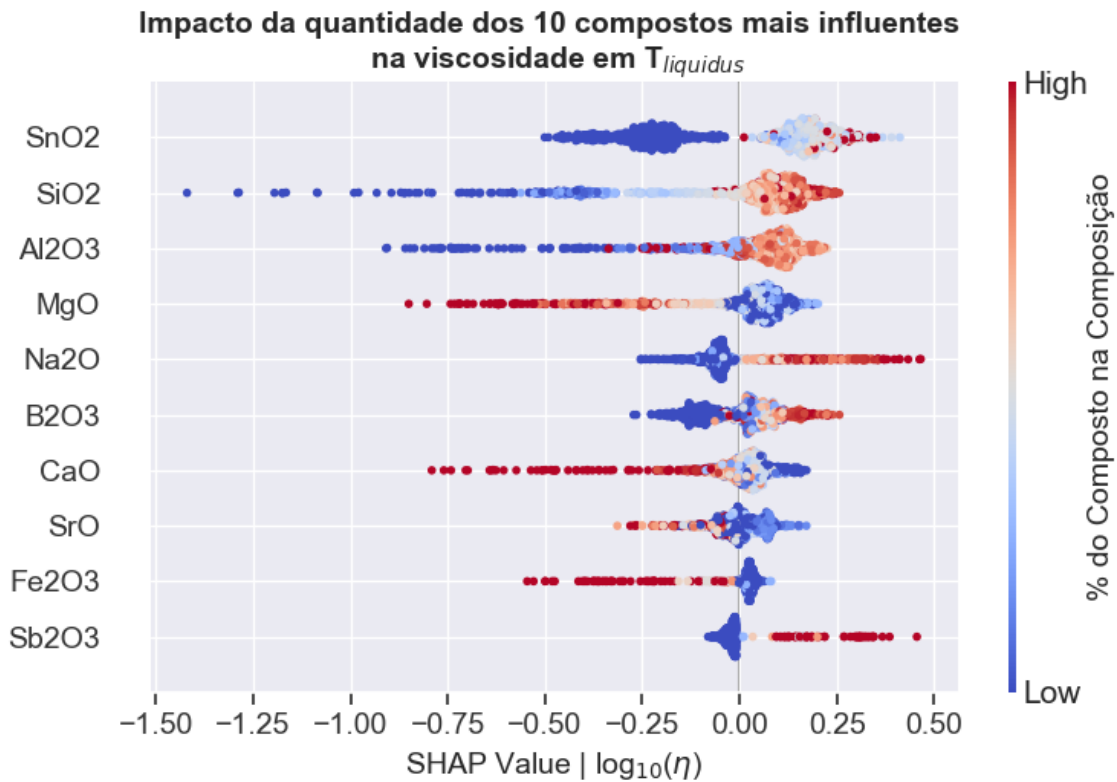


Figura 4.12 – Impacto relativo de cada composto na viscosidade em $T_{liquidus}$ avaliado pela metodologia *Shapley*.

Inicialmente observamos que para os formadores de vidros (SiO_2 , B_2O_3 e Sb_2O_3), de maneira geral, altos valores de composição resultam em previsões de viscosidades acima da média (cuja referência é a linha vertical preta). No caso de vidros borossilicatos o B_2O_3 pode atuar como modificador de rede, explicando a presença de altos valores desse elemento por vezes contribuir para reduzir a viscosidade. Para os modificadores de rede (MgO , CaO , SrO e Fe_2O_3) observa-se que alta fração desses compostos está relacionada com baixos valores de viscosidade e que baixas frações desses compostos tendem a gerar valores de viscosidade acima da média, esse comportamento está bem nítido para o MgO e o CaO . Por outro lado, no caso do Na_2O o comportamento é justamente o inverso do esperado se consideramos seu papel como modificador de rede. Por essa razão seu caso particular será discutido com mais detalhamento adiante.

No caso dos compostos intermediários (SnO_2 e Al_2O_3) é possível observar um comportamento ambíguo, segundo o qual tanto altas como baixas frequências do composto podem resultar em previsões acima da média. Esse caráter ambíguo, pode ser explicado pela própria natureza dos compostos intermediários, os quais podem atuar como formadores ou modificadores em função dos demais compostos presentes na composição. De todo o modo, baixos valores de SnO_2 tendem a resultar em vidros com viscosidade significativamente

abaixo da média, enquanto alto valores de Al_2O_3 tendem a contribuir para o aumento da viscosidade.

De todo modo, cabe destacar que nenhuma análise de explicação de complexos modelos de *Machine Learning* pode ser tomada em absoluto, pois a influência de um elemento na viscosidade pode na verdade estar sendo influenciada pela presença de outro elemento, pela fração em que ele está presente nos compostos avaliados, de modo que certas correlações podem ser ignoradas por explicações simplistas. Um exemplo disso é o comportamento inesperado obtido para a influência do Na_2O na viscosidade dos formadores de vidros, o qual exige uma análise mais aprofundada. Foi reportado na literatura que a adição de óxidos da forma R_2O em vidros silicatos resulta na formação de oxigênios não ligantes, reduzindo as conectividades na estrutura, ocasionando um decréscimo da viscosidade com pequenas adições (SHELBY, 2005). No entanto, foi observado, por exemplo, que adições de óxidos de fórmula R_2O (óxidos de sódio, óxido de lítio, óxido de potássio etc), proporcionam aumento na viscosidade de vidros gálio-silicatos se adicionados em uma fração molar acima de 10-20% (SHELBY, 2005). Cabe ainda destacar que o Na_2O quando em presença de óxidos da família 3A, tais como Al_2O_3 e B_2O_3 pode atuar como compensador de carga e não como modificador de rede. De modo, que é possível que a metodologia *Shapley* esteja atribuindo ao Na_2O um efeito na viscosidade que, na verdade, resulta da interação com esses outros compostos.

Uma análise descritiva da fração da presença do dióxido de sódio na composição dos formadores de vidros pode ser observada na Tabela 4.2. Cabe ressaltar que essa análise é relativa à 2256 composições, desconsiderando as que não contém o Na_2O .

Tabela 4.2 – Estatística descritiva da fração de Na_2O na composição dos vidros.

	Fração de Na_2O (%)
Média	10,30
Desvio Padrão	6,07
Mínimo	0,01
Percentil 25	5,05
Mediana	10,85
Percentil 75	15,2
Máximo	40,0

Observa-se que aproximadamente 25% dos formadores de vidros apresentam uma fração de mais de 15% de Na_2O . Desse modo, a Figura 4.12 contendo o a análise *Shapley* é coerente para o Na_2O , uma vez que em pequenas frações observa-se um decréscimo da viscosidade, mas por outro lado, altas frações resultam em aumento da viscosidade.

A metodologia *Shapley* utilizada para explicar o modelo de *Machine Learning*,

baseado no *Random Forest*, é extremamente poderosa e, permite *insights* importantes como os evidenciados acima. De todo modo, o modelo de ML é capaz de ‘perceber’ interações, efeitos sinérgicos e padrões que podem escapar a esses modelos explicativos. De todo o modo, essa ferramenta permite um ‘*sanity check*’ do modelo ao evidenciar que os resultados obtidos para os 10 elementos que mais influenciam nas previsões se encontram em conformidade com o que está reportado na literatura.

Dessa forma, os modelos *Random Forest* e *XGBoost* exibiram boas capacidades preditivas e boa habilidade para generalizar para dados desconhecidos do modelo, conjunto reservado para treinamento. Adicionalmente, foi evidenciada coerência entre as análises explicativas utilizando a metodologia *Shapley* e as encontradas na literatura, ressaltando-se a capacidade do modelo de absorver informações estruturais implícitas na composição dos formadores de vidros, traduzidas na adequação das explicações às teorias estruturais que destacam o papel dos óxidos formadores, intermediários e modificadores.

Em seguida os dados 4511 dados foram utilizados para prever a $T_{liquidus}$ dos formadores de vidros, utilizando a mesma metodologia empregada para prever a viscosidade. Os resultados dos valores preditos em função dos valores reais (obtidos experimentalmente) estão apresentados na Figura 4.13.

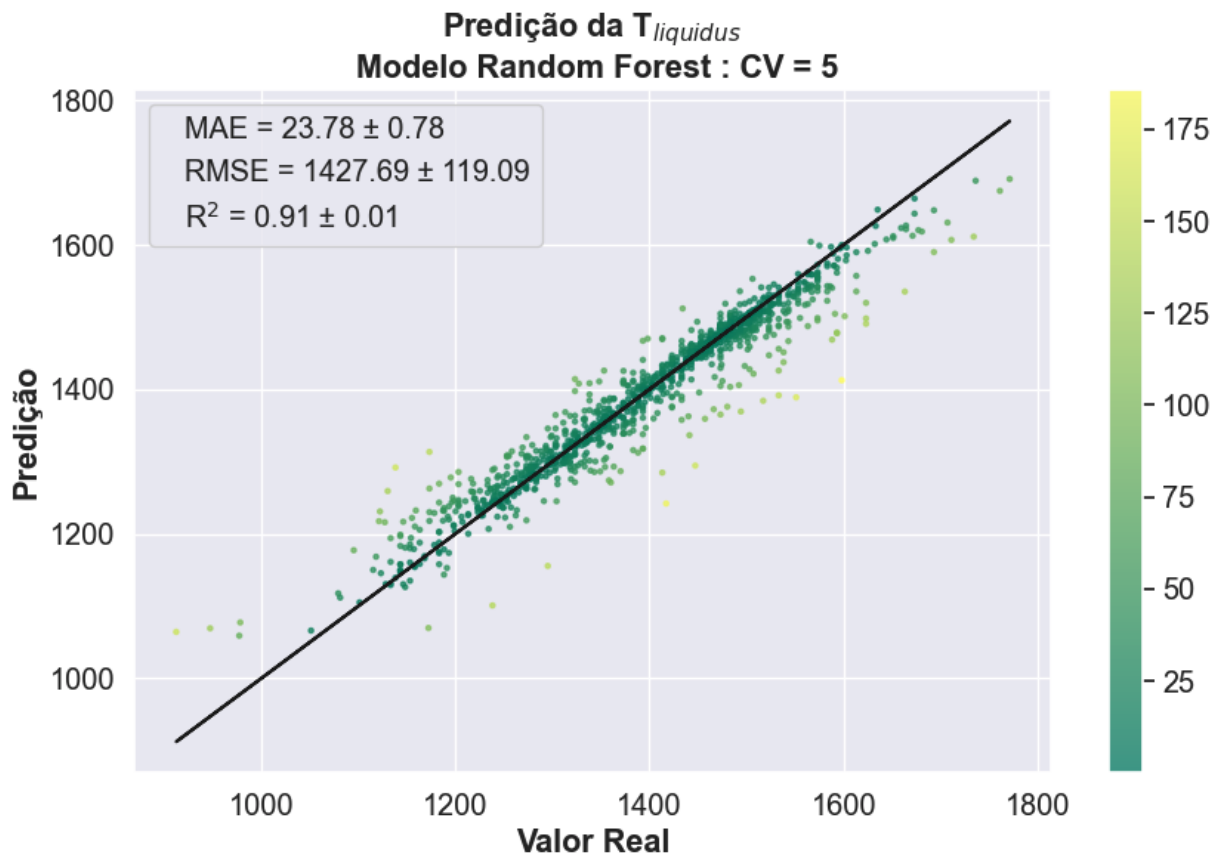


Figura 4.13 – Previsão de $T_{liquidus}$ utilizando modelo *Random Forest*.

Em virtude dos resultados validando as análises, realizou-se uma ampliação dos modelos de *Machine Learning* que passaram a ser treinados utilizando todo o conjunto de dados, o que aumenta sua robustez e as composições abrangidas. Aqui destaco, que com base nas composições de um vidro que se deseja estudar é possível obter uma predição de $T_{liquidus}$ e da viscosidade em $T_{liquidus}$. Com esses dois valores é possível calcular a GFA, utilizando o modelo "Jezica". Os modelos foram encapsulados e estão disponíveis no repositório *Git Hub*: [glassgit](https://github.com/glassgit), podendo ser usados para prever a $T_{liquidus}$, a viscosidade em $T_{liquidus}$ e o GFA para formadores de vidros que não estão contidos nas bases de dados com base exclusivamente na fração dos compostos óxidos presentes nos formadores de vidros (entre 0 e 100%).

Em uma segunda etapa, os 4511 dados contendo a composição dos vidros associados ao valor do logaritmo de GFA, calculados segundo o modelo 'Jezica', foram divididos em dois grupos: um contendo 75% dos dados foi utilizado para treinar o modelo *Random Forest* e o outro contendo 25% dos dados foi usado para testar o modelo, de maneira análoga ao que foi feito para prever a viscosidade. Os resultados dos valores preditos em função dos calculados está apresentado na Figura 4.14.

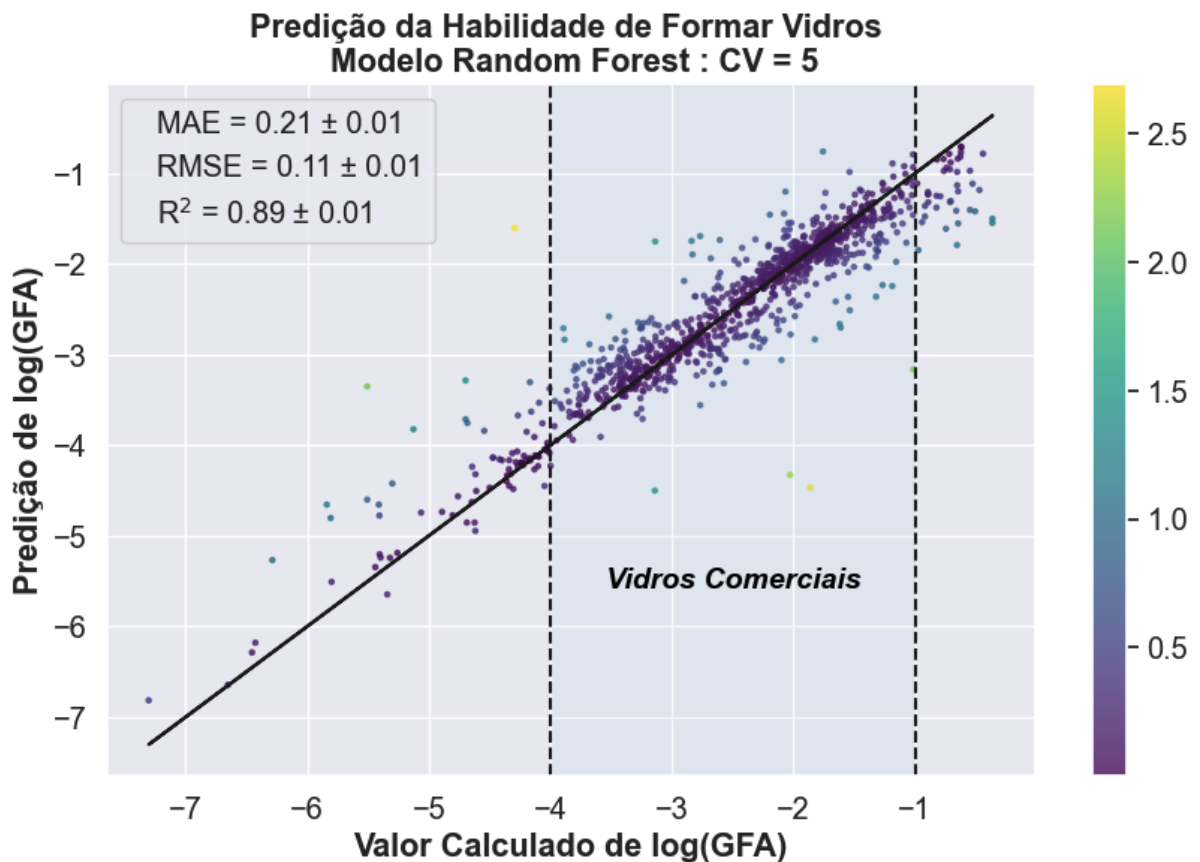


Figura 4.14 – Previsão de GFA utilizando o modelo *Random Forest*.

Além de verificar que o modelo apresenta boa capacidade de prever a GFA, é

possível observar que a performance do modelo é similar ao do modelo utilizado para prever a viscosidade, de modo que se constata que a etapa intermediária de passagem do cálculo de viscosidade em $T_{liquidus}$, utilizando o modelo ‘Jezica’, não interfere na habilidade preditiva dos resultados. Observa-se na Figura 4.14 que o modelo é bastante acurado na predição de vidros comerciais, os quais tem uma GFA intermediária. Para os vidros com baixa GFA, portanto, a esquerda dos vidros comerciais a GFA tende a ser superestimada. No outro extremo, para os considerados excelentes formadores a GFA tende a ser subestimada. De qualquer forma, o modelo exibe capacidade de generalizar para vidros em um grande extensão de valores de GFA.

Uma vez validado nos dados de teste, o modelo foi treinado utilizando todo o conjunto de dados, aumentando sua robustez e as composições abrangidas. O modelo está disponível no repositório *Git Hub* para utilização dos usuários : [glassgit](#). Assim, com esse trabalho cria-se a possibilidade de prever a habilidade de formar vidros diretamente a partir da composição de vidros óxidos. Desse modo, minimizando a necessidade de obter experimentalmente a GFA, a qual é custosa em termos de tempo e recursos.

5 CONSIDERAÇÕES FINAIS E PERSPECTIVAS

Este estudo buscou criar modelos de *Machine Learning* para prever a habilidade de formar vidros (GFA) de formadores de vidros compostos por óxidos. Para este fim, os formadores de vidros, cuja viscosidade foi medida em diferentes temperaturas, foram ajustados pela função VFTH, com o intuito de obter a viscosidade em $T_{liquidus}$, um parâmetro chave para a predição de GFA. O ajuste da função VFTH foi testado para 4950 composições de vidro, demonstrando a alta capacidade preditiva associada a esse modelo, mesmo considerando uma grande variedade de composições, temperaturas e um intervalo de viscosidades de 12 ordens de grandeza.

A viscosidade em $T_{liquidus}$ é considerada um fator chave para prever a habilidade de formar vidros, de modo que foi avaliada a capacidade de modelos de ML na predição dessa propriedade, utilizando a fração dos compostos como descritores do modelo. Constatou-se uma boa capacidade de prever a viscosidade com informações da composição, destacando-se o alinhamento dos *insights* obtidos com as teorias estruturais, segundo as quais os óxidos são agrupados como formadores de rede, intermediários e modificadores. Paralelamente, foram criados modelos para prever a $T_{liquidus}$ para as diferentes composições, a qual associada a viscosidade em $T_{liquidus}$ permite obter a GFA. Essa via, oferece uma transparência na predição de GFA, uma vez que utiliza dois modelos de ML intermediários, os quais fornecem explicitamente os valores de $T_{liquidus}$ e a viscosidade nessa temperatura.

Paralelamente, os dados de $T_{liquidus}$ experimentais e os dados de viscosidade de $T_{liquidus}$ foram utilizados para calcular a GFA por meio do modelo ‘Jezica’. Nesse caso, os modelos de *Machine Learning* foram testados na sua capacidade de prever diretamente GFA a partir da composição dos vidros, obtendo uma boa acurácia. Esses resultados têm implicações significativas no desenvolvimento de novos vidros óxidos, possibilitando a obtenção de GFA com base apenas na composição e, portanto, prescindindo de análises experimentais dispendiosas. Esse estudo pode potencialmente ser expandido para vidros não óxidos e para vidros metálicos. Cabe destacar que os modelos estão restritos à composições que contenham unicamente os 25 óxidos utilizados no treinamento do modelo, ressaltando-se portanto, o potencial de estender os modelos para contemplar outros compostos.

Em última análise, esse trabalho sublinha o poder associativo da interação de modelos físicos e de *Machine Learning*, oferecendo uma contribuição para direcionar e/ou minimizar a necessidade de testes laboratoriais, estendendo a análise para uma quantidade grande de vidros e fornecendo uma capacidade de prever, em princípio, a GFA para qualquer

nova composição de vidro que contenha apenas os óxidos utilizados no treinamento do modelo.

REFERÊNCIAS

- ALCOBAÇA, E.; MASTELINI, S. M.; BOTARI, T.; PIMENTEL, B. A.; CASSAR, D. R.; FERREIRA, A. C. P. de L.; ZANOTTO, E. D. et al. Explainable machine learning algorithms for predicting glass transition temperatures. **Acta Materialia**, Elsevier, v. 188, p. 92–100, 2020.
- BARRER, R. M. The viscosity of pure liquids. ii. polymerised ionic melts. **Transactions of the Faraday Society**, v. 39, p. 59, 1943.
- CASSAR, D. R.; CARVALHO, A. C. de; ZANOTTO, E. D. Predicting glass transition temperatures using neural networks. **Acta Materialia**, Elsevier, v. 159, p. 249–256, 2018.
- CASSAR, D. R.; MASTELINI, S. M.; BOTARI, T.; ALCOBAÇA, E.; CARVALHO, A. C. de; ZANOTTO, E. D. Predicting and interpreting oxide glass properties by machine learning using large datasets. **Ceramics international**, Elsevier, v. 47, n. 17, p. 23958–23972, 2021.
- CASSAR, D. R.; SANTOS, G. G.; ZANOTTO, E. D. Designing optical glasses by machine learning coupled with a genetic algorithm. **Ceramics international**, Elsevier, v. 47, n. 8, p. 10555–10564, 2021.
- DOMINGUES, R.; FILIPPONE, M.; MICHIARDI, P.; ZOUAOUI, J. A comparative evaluation of outlier detection algorithms: Experiments and analyses. **Pattern recognition**, Elsevier, v. 74, p. 406–421, 2018.
- DUINE, P.; SIETSMA, J.; BEUKEL, A. van den. Determining the parameters of the fulcher-vogel equation from equilibrium viscosity data. **Materials Science and Engineering: A**, Elsevier, v. 133, p. 649–652, 1991.
- HRMA, P. Arrhenius model for high-temperature glass-viscosity with a constant pre-exponential factor. **Journal of Non-Crystalline Solids**, Elsevier, v. 354, n. 18, p. 1962–1968, 2008.
- JIRASEK, F.; HASSE, H. Combining machine learning with physical knowledge in thermodynamic modeling of fluid mixtures. **Annual Review of Chemical and Biomolecular Engineering**, Annual Reviews, v. 14, p. 31–51, 2023.
- JIUSTI, J.; ZANOTTO, E. D.; CASSAR, D. R.; ANDREETA, M. R. Viscosity and liquidus-based predictor of glass-forming ability of oxide glasses. **Journal of the american ceramic society**, Wiley Online Library, v. 103, n. 2, p. 921–932, 2020.
- LAZAREVIC, A.; KUMAR, V. Feature bagging for outlier detection. In: **Proceedings of the eleventh ACM SIGKDD international conference on Knowledge discovery in data mining**. [S.l.: s.n.], 2005. p. 157–166.
- MULLER, A.; GUIDO, S. **Introduction to Machine Learning with Python**. [S.l.]: O'Reilly Media, Incorporated, 2016. ISBN 144936988X, 9781449369880.

NASCIMENTO, M. L.; SOUZA, L. A.; FERREIRA, E. B.; ZANOTTO, E. D. Can glass stability parameters infer glass forming ability? **Journal of Non-Crystalline Solids**, Elsevier, v. 351, n. 40-42, p. 3296–3308, 2005.

NASCIMENTO, M. L. F.; APARICIO, C. Data classification with the vogel–fulcher–tammann–hesse viscosity equation using correspondence analysis. **Physica B: Condensed Matter**, Elsevier, v. 398, n. 1, p. 71–77, 2007.

OJOVAN, M. I. Glass formation. **Encyclopedia of Glass Science, Technology, History, and Culture**, Wiley Online Library, v. 1, p. 249–259, 2021.

RAVINDER; VENUGOPAL, V.; BISHNOI, S.; SINGH, S.; ZAKI, M.; GROVER, H. S.; BAUCHY, M.; AGARWAL, M.; KRISHNAN, N. A. Artificial intelligence and machine learning in glass science and technology: 21 challenges for the 21st century. **International journal of applied glass science**, Wiley Online Library, v. 12, n. 3, p. 277–292, 2021.

SEGERS, L.; FONTANA, A.; WINAND, R. Viscosite de melanges de silicates fondus du systeme cao sio2 mno. **Electrochimica acta**, Elsevier, v. 24, n. 2, p. 213–218, 1979.

SHELBY, J. E. **Introduction to Glass Science and Technology**. Illustrated. [S.l.]: Royal Society of Chemistry, 2005. 291 p. ISBN 0854046399, 9780854046393.

SUN, Y.-T.; BAI, H.-Y.; LI, M.-Z.; WANG, W.-H. Machine learning approach for prediction and understanding of glass-forming ability. **The journal of physical chemistry letters**, ACS Publications, v. 8, n. 14, p. 3434–3439, 2017.

ZANOTTO, E.; COUTINHO, F. How many non-crystalline solids can be made from all the elements of the periodic table? **Journal of non-crystalline solids**, Elsevier, v. 347, n. 1-3, p. 285–288, 2004.

ZANOTTO, E. D.; MAURO, J. C. The glassy state of matter: Its definition and ultimate fate. **Journal of Non-Crystalline Solids**, Elsevier, v. 471, p. 490–495, 2017.

ZHAO, Y.; NASRULLAH, Z.; LI, Z. Pyod: A python toolbox for scalable outlier detection. **arXiv preprint arXiv:1901.01588**, 2019.

ZHONG, L.; WANG, J.; SHENG, H.; ZHANG, Z.; MAO, S. X. Formation of monatomic metallic glasses through ultrafast liquid quenching. **Nature**, Nature Publishing Group UK London, v. 512, n. 7513, p. 177–180, 2014.