Full length article

Glass hardness: Predicting composition and load effects via symbolic reasoning-informed machine learning

Sajid Mannan [a], Mohd Zaki [a], Suresh Bishnoi [b], Daniel R. Cassar [c], Jeanini Jiusti [d], Julio Cesar Ferreira Faria [e], Johan F.S. Christensen [f], Nitya Nand Gosvami [g], Morten M. Smedskjaer [f], Edgar Dutra Zanotto [e,*], N.M. Anoop Krishnan [a,h,*]

[a] Department of Civil Engineering, Indian Institute of Technology Delhi, Hauz Khas, New Delhi 110016, India
[b] School of Interdisciplinary Research, Indian Institute of Technology Delhi, Hauz Khas, New Delhi 110016, India
[c] Ilum School of Science, Brazilian Center for Research in Energy and Materials (CNPEM), Zip Code 13083-970, Campinas, Sao Paulo, Brazil
[d] Laboratory of study and development of nuclear waste conditioning matrices, French Alternative Energies and Atomic Energy Commission, France
[e] Department of Materials Engineering, Federal University of São Carlos, Brazil
[f] Department of Chemistry and Bioscience, Aalborg University, 9220 Aalborg, Denmark
[g] Department of Materials Science and Engineering, Indian Institute of Technology Delhi, Hauz Khas, New Delhi 110016, India
[h] Yardi School of Artificial Intelligence, Indian Institute of Technology Delhi, Hauz Khas, New Delhi 110016, India

ARTICLE INFO

ABSTRACT

Glass hardness varies in a non-linear fashion with the chemical composition and applied load, a phenomenon known as the indentation size effect (ISE), which is challenging to predict quantitatively. Here, using a *curated* dataset of over 3,000 inorganic glasses from the literature comprising the composition, indentation load, and hardness, we develop machine learning (ML) models to predict the composition and load dependence of Vickers hardness. Interestingly, when tested on new glass compositions unseen during the training, the standard data-driven ML model failed to capture the ISE. To address this gap, we combined an empirical expression (Bernhardt's equation) to describe the ISE with ML to develop a framework that incorporates the symbolic equation representing the domain reasoning in ML, namely *Symbolic Reasoning-Informed ML Procedure* (SRIMP). We show that the resulting SRIMP outperforms the data-driven ML model in predicting the ISE. Finally, we interpret the SRIMP model to understand the contribution of the glass network formers and modifiers toward composition and load-dependent (ISE) and load-independent hardness. The deconvolution of the hardness into load-dependent and load-independent terms paves the way toward a holistic understanding of the composition effect and ISE in glasses, enabling efficient and accelerated discovery of new glass compositions with targeted hardness.

1. Introduction

Glasses are widely used in several applications involving interactions with other entities, namely, automobile windshields and other specialty windows, artificial gems, ballistic armors, smartphone and computer protective screens, biomedical implants, nuclear waste immobilization matrices, and artistic pieces [1,2]. The glass hardness is correlated with its ability to resist damage in such applications [3], which is typically determined by instrumented indentation experiments [4–7]. However, the hardness values obtained from such experiments are *not* an intrinsic glass property; it also depends on other parameters, such as the loading procedure, geometry of the indenter, and environmental conditions [8,

9]. Specifically, hardness monotonically decreases and saturates with increasing applied load—a phenomenon termed indentation size effect (ISE) [10–13]. This load-dependent plastic behavior of glass has been attributed to the stress concentration generated due to sharp contact loading, which causes localized structural changes in the glass network leading to permanent deformation [14]. The ISE prevents the comparison of hardness values obtained at different loading conditions. Predicting load-independent hardness values is thus crucial for comparing experiments conducted at varying loads.

Experimental studies have paved the way for several empirical laws capturing the load-dependence of glass hardness [10,15–18]. A widely used empirical law for predicting the ISE was provided by Bernhardt

[19]. Such relations are based on experimental observations fitted to simple polynomial-like functions. Thus, these laws are based on human observation/intervention, which is hard-coded in symbolic expressions by testing on a large family of glasses, an approach known as symbolic reasoning [20]. Symbolic reasoning has found extensive success in materials science, where several observations have been coded as empirical rules, enabling the prediction of material properties. However, these empirical laws require the fitting parameters to be calibrated for each family of glasses and typically are not transferable from one family to another. These laws are useful but require multiple experiments for each glass family to obtain the fitting parameters.

Recently, researchers have been successful in predicting glass properties using machine learning (ML) techniques [21–29]. For instance, Ravinder et al. [21] developed neural network models for predicting the physical, mechanical, optical, and thermal properties of oxide glasses. Since these models were trained on a limited number of components (~30 components) and properties (eight properties), there was a need for models which could predict properties for a broader range of components encompassing oxides, halides, and sulfides, to name a few [30]. To overcome this limitation, researchers [22] developed tree-based machine learning models for predicting 25 properties of oxide glasses for a larger dataset. In another work [23], researchers predicted the hardness of oxide glasses as a function of their composition, testing load, and annealing temperature, which demonstrated improved property prediction and strengthened the fact that the glass hardness is a function of composition, testing, and processing parameters [23]. However, these models are purely data-driven and unable to incorporate any empirical or physics-based knowledge that severely limits the generalizability of the model to compositions outside the training space.

Moreover, interpreting the compositional dependence of properties is also required for designing novel glasses with targeted properties, a challenging task for black-box ML models. To this end, Alcobaça et al. (2020) [24] used explainable machine learning models to predict the glass transition temperature of oxide glasses. This work was followed by others on predicting and explaining the properties of glasses using large databases, employing *explainable* ML models and a game theory-based approach known as Shapely additive explanations (SHAP) [22,31,32]. Similar attempts have been made to predict the properties of chalcogenide glasses along with their explanation of compositional effect using SHAP analysis [29,33].

Some recent studies have attempted to predict the Vickers hardness of oxide glasses using either composition only or composition and testing conditions, such as applied load [34,35]. However, these models use a classical data-driven model to predict the hardness that does not incorporate domain knowledge to ensure better generalization. Hence, these models cannot capture the load dependency or ISE, as demonstrated later in the present work. ML has also been used to predict other mechanical properties based on instrumented indentation [36]. An alternate approach incorporating domain knowledge is called Physics-Informed ML [37], where known physical laws governing the system are included as part of the loss function. Such methods have succeeded in predicting glass properties, such as viscosity, which exhibits a temperature-dependent variation based on the MYEGA or VFT equations [38]. Physics-informed ML has also been used to predict glass structure, where the physical information is provided by an independent statistical model [39]. Researchers have attempted to use data-driven knowledge and active machine learning approaches to model three-dimensional fracture problems from two-dimensional data [40, 41]. However, to the best of our knowledge, there have been no efforts to incorporate domain knowledge based on intuition and observation into ML, especially in the field of glass mechanics research. That is, no previous works combine symbolic reasoning and ML to improve the prediction and understanding of the mechanical properties of glasses. Also, no work has used a manually curated composition–load–hardness dataset for training an ML model.

Here, we propose a symbolic reasoning-informed ML procedure (SRIMP) framework, combining Bernhardt's law and ML, to understand the composition and load dependence of glass hardness. The contribution of this work is twofold: (i) we present the first highly *curated* dataset on composition–load–hardness values of glasses, which is made available publicly, and (ii) we develop a *SRIMP framework* that allows the definition of a load-independent hardness, while also providing insights into the indentation size effect for a given glass composition. To this end, we manually collected a large dataset (~3000 entries) of composition, load, and corresponding hardness values for different inorganic glasses from the literature, which was used to train traditional data-driven ML models. We demonstrate that even the ML models trained with the load as an input feature fail to capture the ISE. In contrast, the SRIMP model predicts the hardness in an improved fashion and exhibits an excellent capability to predict the load-dependent hardness variation. Further, we consider indentation experiments on three different glass compositions [42] to evaluate the transferability of the approach to new experiments performed in-house, on which the model has no information. Finally, the SRIMP provides interpretability to the ML model, thereby allowing the definition of a load-independent hardness. The Bernhardt equation implies that the hardness is inversely proportional to the applied load and converges to an asymptotic value at large loads. Thus, at any given load, hardness can be decomposed as the sum of its asymptotic value (H_∞) and the additional hardness value that varies as a function of load (a_{ISE}/L_D) (Fig.7). That is, $H = H_\infty + a_{ISE}/L_D$ (Eq. (5)). Thus, the load-independent hardness refers to the asymptotic value (shown in blue color in Fig. 7). Note that this asymptotic value might still depend on external factors such as humidity, surface quality, and even the indenter geometry. We demonstrate that the load-independent hardness corresponds to the asymptotic hardness at high loads and allows the hardness prediction at any load without performing an experiment. Possibly a modified Bernhardt equation could be developed using a symbolic regression framework in future research. However, additional work would be required to meaningfully interpret the modified equation in the context of indentation hardness.

2. Methodology

2.1. Dataset collection and preparation

The datasets used in this work were manually collected from published literature on glass, that is, research papers focusing on Vickers' indentation experiments in oxide glasses. Special care was taken to ensure that there was no bias in the manual selection process. Manuscripts were chosen based on an extensive search of the Web of Science and Scopus databases. Further, all the papers on oxide glasses obtained based on the search were used to extract data composition, load, and hardness. Those data were cross-checked against two glass databases–Interglad and SciGlass. The resulting dataset was normalized to ensure that the compositions added up to 100%. All the glass compositions were represented in the elemental form, with the value of each chemical taken in atomic mole percentage. Also, the hardness values were rounded off up to three decimal places. Duplicate entries were removed. Further, when multiple hardness values were available for a given composition, outliers were identified as those values lying beyond $\pm$ 2.5% of the mean of all hardness values for that particular composition. These outliers were removed, and the *mean* value of the remaining was considered as the hardness of that composition. Additionally, we considered only those elements present in at least *30 or more glass compositions* to ensure representative data in both training and test datasets. The above protocol resulted in hardness data for 3325 unique glasses with 52 features (51 different elements and load as the input features) for model training, testing, and further analysis. These data were used for developing the ML models using load and composition as the input.

However, for many of the selected compositions, the hardness values were available for only a single load value. However, hardness values

corresponding to at least three different loads are required to fit the parametric equation of ISE for any given glass composition. To this end, we further curated a smaller dataset of about approximately 150 glasses [15,41–45,45–47] from the entire dataset, for which the hardness values corresponding to three or more load values were available. These other data were used to train the SRIMP model.

2.2. Symbolic-reasoning informed machine learning procedure (SRIMP)

The hardness, H, from indentation experiments was computed as

$$H = \frac{2P\sin(\theta/2\ )}{L_D^2} \tag{1}$$

where, L_D is the diagonal length of the indent after the indenter is removed, θ is the tip angle of the indenter, and P is the applied load. The ISE using Bernhardt's model [48] is

$$\frac{P}{L_D} = a_1 + a_2 L_D \tag{2}$$

where a_1 and a_2 are fitting parameters. Substituting for L_D from Eq. (1) in Eq. (2), we get

$$P = a_1\sqrt{\frac{2P\sin(\theta/2)}{H}} + a_2\frac{2P\sin(\theta/2)}{H} \tag{3}$$

Solving this quadratic equation for hardness H, we get (see Supplementary materials for the detailed derivation)

$$H = \frac{C_1}{2P} + C_2 + \frac{\sqrt{C_1^2 + 4C_1 C_2 P}}{2P} \tag{4}$$

where $C_1\ =\ 2a_1^2\sin(\theta/2)$, and $C_2\ =\ 2a_2\sin(\theta/2)$.

Similarly, substituting for load P in Eq. (1) from Eq. (2), we get

$$H = 2a_2\sin(\theta/2) + \frac{2a_1\sin(\theta/2)}{L_D} = H_\infty + \frac{a_{ISE}}{L_D} \tag{5}$$

where the parameters $H_\infty = 2a_2\sin(\theta/2)$ and $a_{ISE} = 2a_1\sin(\theta/2)$ correspond to the load-independent hardness and the extent of the ISE, respectively. Although Eqs. (4) and (5) are equivalent, one is obtained by eliminating L_D, and the other is obtained by eliminating P. Thus, Eq. (5) requires L_D, a quantity that can only be determined experimentally, while Eq. (4) allows the prediction of H purely based on load.

Fig. 1 shows a graphical representation of the methodology employed in SRIMP. We train the ML model to predict the C_1 and C_2

values associated with each glass composition by minimizing the loss function,

$$Loss = \frac{1}{n}\sum_{i=1}^{n}\left(H_{SRIML}^i - H_{actual}^i\right)^2 \tag{6}$$

where H_{actual}^i represents the measured hardness of a certain glass at a given load and H_{SRIML}^i represents the predicted hardness using Eq. (4) for a total of **n** data points. The loss is then backpropagated through the multilayer perceptron (MLP) to learn the optimal weights that minimize it. In this process, the ML model learns the function relating to the composition and the parameters C_1 and C_2. Further, by comparing Eqs. (4) and (5), the relationship between C_1 and C_2 with a_{ISE} and H_∞, respectively, can be obtained as $C_2 = H_\infty$ and $C_1\ =\ a_{ISE}^2/(2\sin(\theta/2))$. This approach directly allows interpreting any glass composition's ISE and their inherent hardness.

2.3. Model training

To train the ML models, the cleaned datasets (both larger and smaller) were divided into 80:20 proportions for the training and testing sets, respectively. Fig. 1 shows the schematic of the training scheme with load (P) as an input. All the features were taken as input parameters, and the hardness was the output parameter for the ML model training. The mean and standard deviation of the training set features were used to scale the data before training by computing the z-score. [49] The training set was subjected to 10-fold cross-validation to enable hyperparameter search for an optimal fit. Since machine learning models have different hyperparameters, they need to be chosen appropriately to achieve generalized model performance. In the case of NN, these hyperparameters are the number of hidden layers, number of neurons in each hidden layer, learning rate, weight decay, dropout probability, activation function, and optimizer. The hyperparameters of the ML models were optimized using the Bayesian optimization-based library, Optuna [48], employing a previously reported approach [22]. The model with the best R^2 score on the validation fold was chosen as the final model. Since the test data were kept hidden in the training step to avoid data leakage [50], they were used only to evaluate the performance of the best model. The details of the bounds of each hyperparameter are provided in the supplementary material (Tables S1 and S2). MSE was chosen as the loss function for hyper-parameter optimization (see Eq. (6)). R^2 scores are also reported as it is more straightforward to interpret (closer to 1 means better training) than the MSE or RMSE values. The MAE and RMSE metrics along with the R^2 score are
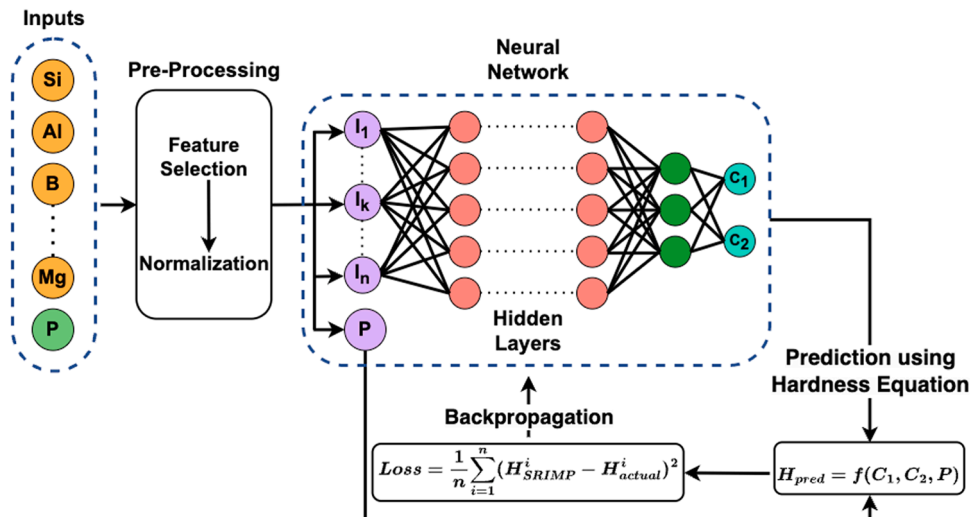

Fig. 1. Workflow diagram for the SRIMP model.

provided in supplementary materials (Table S3).

The following python packages are used in this work: PyTorch, an open-source ML library [51], was used for developing a feed-forward neural network (NN). Batch normalization and dropout were used in the hidden layers to help in reducing the model overfitting to the data. NumPy, Pandas, and Matplotlib [52–54] were used for pre-processing and data visualization. Although we employed both tree-based and neural ML models, we obtained the best performance using an NN in terms of mean squared error and R^2 for both validation and training scores. The final hyperparameters used in the ML models are reproduced in Table 1. All the codes and data used for the present work are available at https://github.com/M3RG-IITD/SRIMP_Hardness

2.4. Shapley additive explanations

To understand the model's output, we used the SHAP [55], an interpretable machine learning algorithm, to understand the influence of individual features on the output. It is a unified game theoretic approach where players and the game can be considered as the feature values of instances and the model output, respectively. Shapely values tell us how fairly the prediction is distributed among the features. Mathematically, Shapley values are calculated as [22]

$$\phi_i = \sum_{S \subseteq F \setminus \{i\}} \frac{|S|!(|F| - |S| - 1)!|F|!}{|F|!} \left[f_{S \cup \{i\}} \left(x_{S \cup \{i\}} \right) - f_s(x_s) \right] \tag{7}$$

where F denotes the set of all features, S denotes a subset of features, $f_{S \cup \{i\}}$ denotes a model trained with the feature present, $f_{s\{i\}}$ denotes a model trained with the feature withheld, and x_s denotes input features in subset S. By comparing the prediction of these models on the current input, SHAP determines the significance of a feature. The feature importance is high if the prediction error is high; otherwise, it is low. Note that SHAP has several advantages compared to simply considering the derivatives of the output with respect to the input. For instance, (i) SHAP is a model agnostic approach that can be used for models that are non-differentiable, for instance, Classification and Regression Trees (CART), XGBoost, or Random Forest. (ii) For many properties, the output might depend on a combination of input variables and their cross-interaction. This is especially true in the case of glasses where one component can have a significant effect on the behavior of other components. In such cases, SHAP allows to capture this dependence through SHAP interaction value plots and SHAP dependency plots. SHAP has introduced various model approximations such as kernel, Deep, Linear, and Tree-based explainers. Based on the task and the model, any explainer can be opted to calculate the SHAP values for each feature. The explainer exploits the model's internal complexity and collapses it to a low-order polynomial complexity [22].

Various plot types, such as bar, violin, river flow, and bee swarm plots, can be used to visualize the SHAP values. In this study, we use bar and violin plots to evaluate the ML model. Bar plots give the average impact of each feature on the model output without showing the directionality of effects, whether it is influencing positively or

negatively. However, the violin plots represent the contribution of each feature towards the different model outputs as a function of the feature value. Thus, it is colored based on the feature value and represents the magnitude of impact on model output, whether it is impacting positively or negatively.

3. Results and discussion

3.1. Dataset visualization

First, we analyze the hardness and load data compiled from the literature. Fig. 2(a) shows the bar chart of glass compositions associated with each feature of the training and test sets. The training and test sets have 2656 and 665 glasses, respectively, of which 2555 and 632 are oxide glasses, and the others are oxynitrides, oxy-halides, or chalcogenide glasses. The inset of Fig. 2(a) shows the components present in a relatively small number of compositions. Fig. 2(b) shows the distribution of the number of chemical elements present in each glass composition in the training and test sets, that is, how many compositions present are two-, three-, four-component, etc. The most frequent value is for glasses having five elements, followed by six and four. The compositions with the most components in the dataset contain five elements. Fig. 2(c) shows the frequency of glass compositions in different hardness ranges. The glasses having hardness < 2 GPa are mostly phosphosilicates, and chalcogenides, while the most frequent hardness values are within 5–6 GPa. The glasses with high hardness are mostly oxynitrides with hardness values of 11–12 GPa. Thus, the dataset presented here covers a broad range of composition and hardness values for inorganic, non-metallic glasses.

3.2. Machine learning models

Figs. 3(a) and 3(b) show predicted hardness values with and without load as input features, respectively, compared to the experimentally measured values. These models will henceforth be referred to as ML hardness with load (MLHL) and ML hardness without load (MLH). We observe that the MLHL model, which takes both composition and load as input features, performs better than the MLH model, which takes only the composition as an input feature (training and test R^2 scores of 0.922 and 0.897, respectively, against 0.882 and 0.879, respectively). These results exemplify the importance of the applied load for hardness measurements and predictions. The insets represent the error distribution as a probability density function. The narrow region of the red-colored shaded portion indicates the 95% confidence limit, confirming a small error for most of the values. Altogether, we observe that the model with the load as an input feature exhibits reliable hardness predictions. It should be noted that although the model predictions are relatively good, it is unclear if it can capture the load dependence on the hardness, which will be discussed in detail later.

3.3. Symbolic reasoning informed ml

Now, we evaluate the performance of the induced SRIMP model to predict the hardness values. For the model to be trained effectively, we need at least three data points corresponding to each glass (composition, three load values, and hardness) since there are two fitting parameters, namely C_1 and C_2. To demonstrate this effect, we trained the SRIMP model with one load, two load and three load values. Our findings reveal that as the number of load values increases, the curve gets closer to actual data (see Supplementary materials Fig. S3). Thus, considering the trade-off between accuracy and available data points, we considered compositions for which hardness is available for at least three different loads. Thus, the dataset employed to train the SRIMP model is significantly smaller, with ~150 glass compositions (see Supplementary materials Fig. S1 for details). Fig. 4 shows the predicted hardness values using the SRIMP approach compared to the measured values. The model
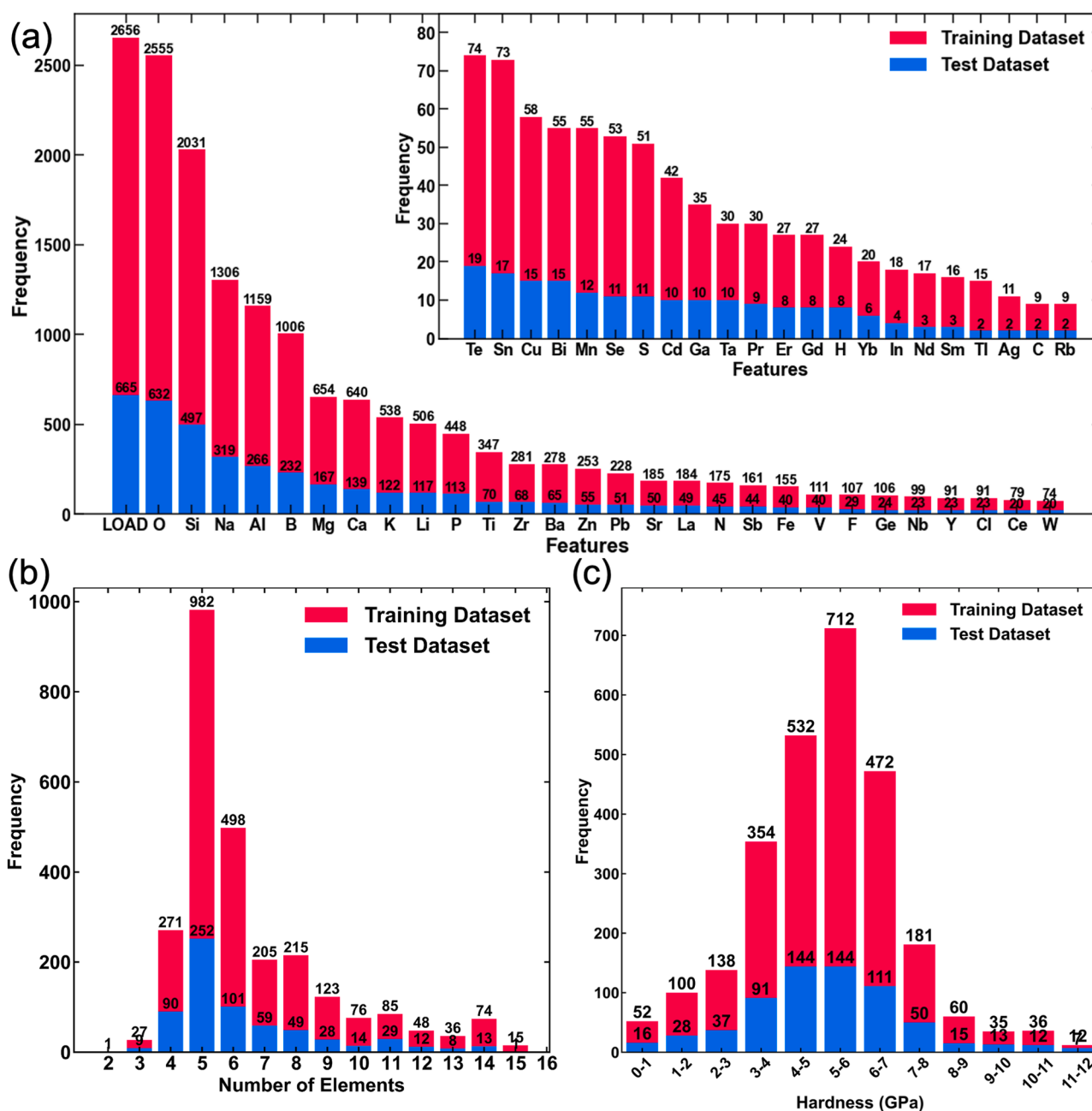
Table 1
Details of model hyperparameters.

Hyperparameter	Model with Load	Model without Load
Epoch	500	500
Batch Size	64	32
Number of layers	1	1
Neurons	92	97
Dropout	0.2	0.1
Activation	LeakyReLU	LeakyReLU
Optimizer	Adam	Adam
Learning rate	0.000832	0.000494
Weight Decay	0.000322	0.000416

Fig. 2. Visualizing the hardness dataset with load: (a) Number of elements and their frequency, (b) Frequency of glasses versus number of elements in their composition, (c) Distribution of hardness values in the datasets used for training and testing.

performance, as given by the R^2 values (training and test values of 0.979 and 0.948, respectively), is excellent in predicting glass hardness. To show the convergence of the training procedure, the loss curve for both training and test sets is provided in supplementary materials Fig S2.

3.4. Indentation size effect

To evaluate the ability of the models to generalize to unseen glass compositions and loads, we consider recent indentation experiments done on three new glass compositions synthesized in-house, that is, $25Na_2O$–$37.5B_2O_3$–$37.5SiO_2$ (NBS), $25Na_2O$–$75B_2O_3$ (NB), and $25Na_2O$–$75SiO_2$ (NS), compositions given in mol%. The hardness of these compositions was evaluated at multiple loads experimentally [42]. To quantify how different the new compositions are from the compositions in the training set, we use the Manhattan distance as a metric (see

Supplementary materials). Fig. 5 shows the minimum Manhattan distance for glass compositions used in the training and test set with their closest neighbor in the training dataset. We have also computed the minimum Manhattan distance for three selected glasses prepared and tested in-house: $25Na_2O$–$37.5B_2O_3$–$37.5SiO_2$ (NBS), $25Na_2O$–$75B_2O_3$ (NB), and $25Na_2O$–$75SiO_2$ (NS), respectively (Fig. 5). It is worth noting that no glass composition belonging to the NBS family exists in the training set. Most important is that Fig. 5 shows that the NBS composition is far away from all those used for model training, hence it provides a fair test for the extrapolation ability of the SRIMP.

To evaluate the ability of the ML models to capture the ISE, the hardness of these compositions was predicted using MLHL and SRIMP models. Note that the MLH model was not used as it cannot capture the load dependency. Figs. 6(a-c) show the hardness predicted by MLHL and SRIMP models, respectively. We observe that the data-driven MLHL
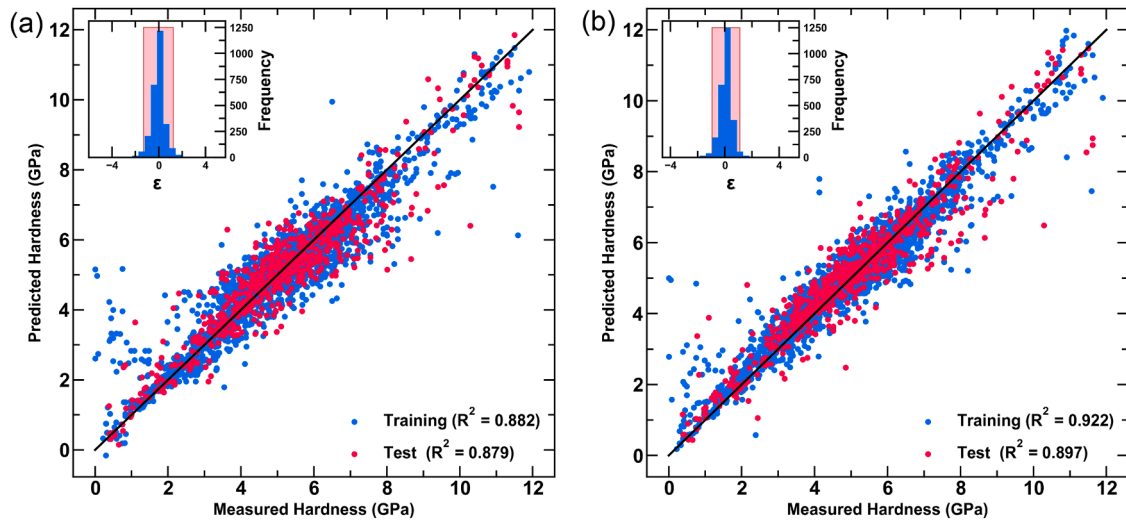
Fig. 3. Predicted values of hardness (a) without load and (b) with the load as input features using the ML models with respect to the experimental values for both the training and test datasets. Inset shows a histogram of errors, where the shaded region represents the 95% confidence interval. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)
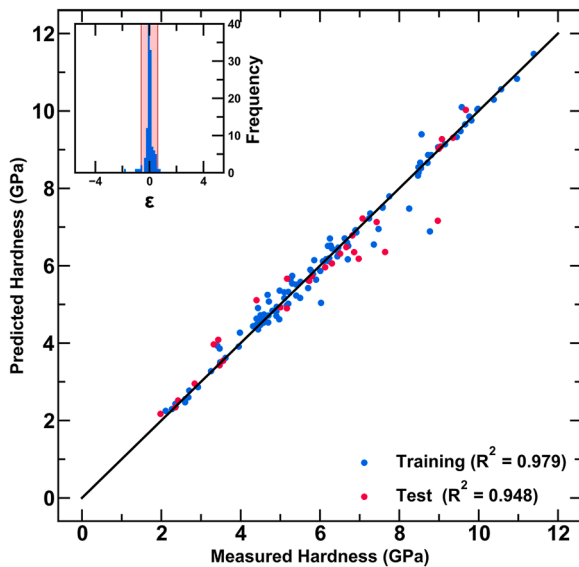

Fig. 4. Predicted hardness values using SRIMP with respect to the experimental values for both the training and test datasets. A histogram of errors is plotted in the inset, where the shaded region represents the 95% confidence interval. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

model cannot capture the load-dependent hardness variation of the glass compositions. Even the trend of the hardness predicted by the MLHL model is not representative of the ISE in glasses. Thus, MLHL performs poorly in extrapolating beyond trained compositions and loads and does not capture the ISE.

In contrast, the SRIMP model not only correctly predicts the ISE trend of hardness but also provides values reasonably close to the experimentally obtained ones that the model has no knowledge of. Thus, the SRIMP model captures the ISE, qualitatively and quantitatively, for completely unseen compositions and loads.

3.5. Load-independent hardness and indentation size effect

To interpret the formulation of the SRIMP, we now evaluate the load-dependent (a_{ISE}/L_D) and load-independent (H_∞) components of the
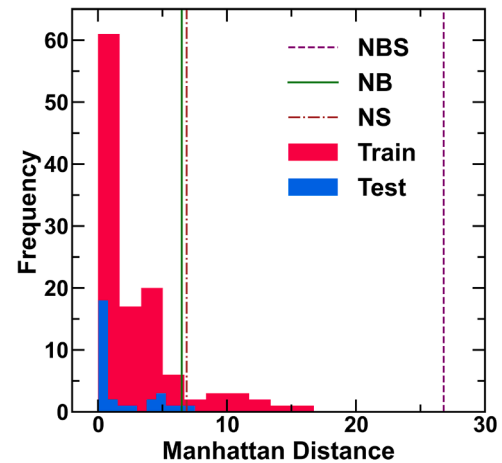
Fig. 5. Histogram of the Manhattan distance between a glass composition and its nearest neighbor from the training set along with three selected glass compositions prepared and tested in-house, $25Na_2O–37.5B_2O_3–37.5SiO_2$ (NBS), $25Na_2O–75B_2O_3$ (NB), and $25Na_2O–75SiO_2$ (NS). The compositions are in mol %. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

hardness (see Eqs. (4) and (5)). Fig. 7(a-c) shows the actual versus predicted hardness for the NBS, NB, and NS compositions, where the pink and blue shaded regions represent the contribution of load-dependent (a_{ISE}/L_D) and independent hardness (H_∞), respectively, toward the hardness at any value of the load. As expected from the formulation, we observe that the H_∞ represents the asymptotically convergent hardness value. This value remains constant irrespective of the applied load. Thus, H_∞ can be used to describe the inherent hardness of the glass. In contrast, the load-dependent a_{ISE}/L_D decreases monotonically, asymptotically tending to zero with increasing applied load. Thus, the contribution of H_∞ towards the hardness is significantly more than the a_{ISE}/L_D, except for low load values. It should be noted that both H_∞ and a_{ISE} are computed directly from the C_1 and C_2 terms predicted by the SRIMP model. Thus, the SRIMP formulation provides a framework to decouple the contribution of ISE in hardness and directly predict them from the glass composition. However, due to the black-box nature of the NNs, it is not clear how each of the components in the composition controls the C_1 and C_2 terms, and in turn, a_{ISE} and H_∞. Therefore, we use
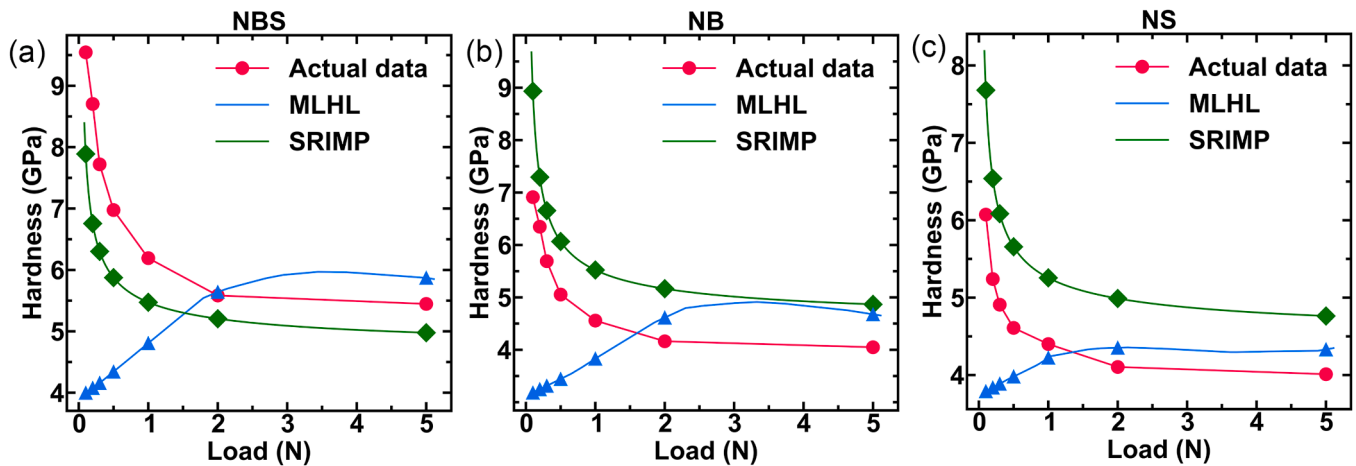
Fig. 6. Actual versus predicted hardness as a function of the applied load for three selected glass compositions prepared and tested in-house, (a) 25Na$_2$O–37.5B$_2$O$_3$–37.5SiO$_2$ (NBS), (b) 25Na$_2$O–75B$_2$O$_3$ (NB), and (c) 25Na$_2$O–75SiO$_2$ (NS), with the compositions in mol% using MLHL, and SRIML model. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)
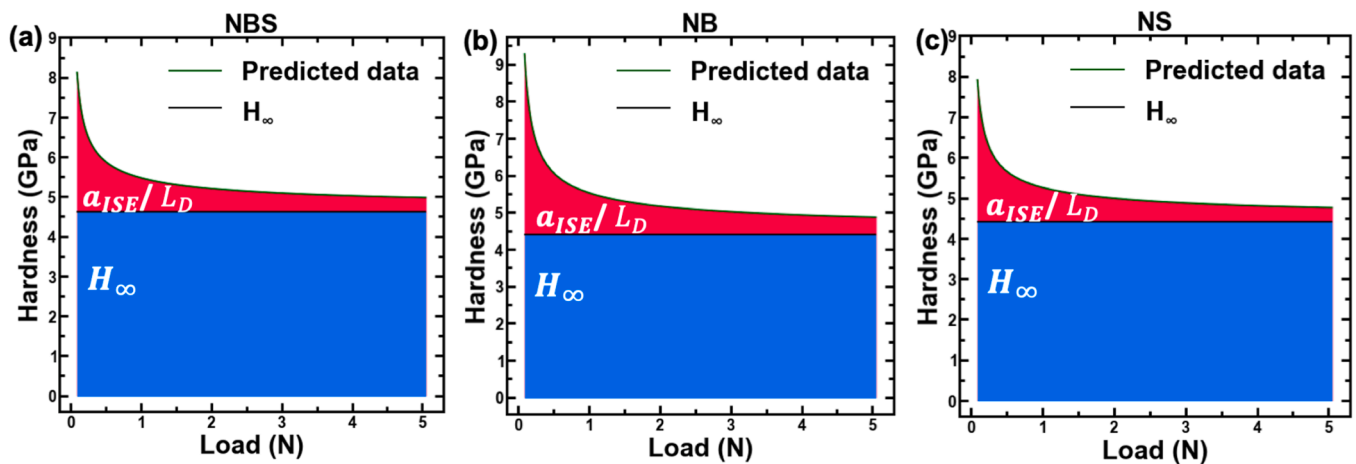

Fig. 7. Predicted hardness as a function of applied load for (a) NBS, (b) NB, and (c) NS glass compositions using the SRIMP model. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

the SHAP to interpret the effect of different elements present in the glass on the Vickers hardness.

### 3.6. Model interpretation using SHAP

SHAP is a model-agnostic post-hoc tool that can be used to interpret ML models. SHAP shows how each input feature increases or decreases the output value. Fig. 8(a-b) shows the SHAP bar and violin plots explaining the MLHL predictions. We analyze the MLHL predictions in this case, as they cover a larger compositional space than the SRIMP model. Moreover, here we focus on interpreting the role of individual elements toward the hardness glasses, not the ISE. Since the MLHL model provides good predictions for the compositions considered in the dataset, it is reasonable to apply the SHAP technique to interpret the MLHL model. While the bar plots reveal the mean absolute effect of a feature on the model predictions, violin plots indicate the directionality of the feature's impact on the model predictions, whether positively or negatively.

Fig. 8(a) shows that the top ten chemical elements governing the hardness of inorganic glasses are nitrogen (N), silicon (Si), sodium (Na), lithium (Li), phosphorous (P), and aluminum (Al), magnesium (Mg), boron (B), calcium (Ca), and lead (Pb). However, while some of these features impact the hardness positively, others have a negative impact.

To this end, we analyze the role of these elements through the violin plots (Fig. 8(b)), where the x-axis represents the SHAP value in GPa (that is, the contribution toward hardness by a given datapoint), and color represents the normalized feature value. Among these top ten elements, while N, Si, Ca, and Mg positively affect the hardness (they increase it), Na, P, and Pb have negative effects. Other elements that increase the hardness are La, Zr, Ti, and Zn. Interestingly, Al, B, and Li have mixed effects, that is, the effect of these elements on hardness depends on the other species present in the system. Note that elements such as B and Al are well-known to have anomalous effects due to their unique structure characterized by the Loewenstein rule and the boron anomaly [9,56,57].

The property dependence on glass composition is highly non-linear (especially for multicomponent glasses), reflecting that glass properties depend on more than a single physical-chemical feature of each element (e.g., field strength, atomic packing density, dissociation energy, polarizability, etc.). To address these non-linear dependencies, we have opted for a ML model agnostic approach (SHAP) to unveil composition-property relationships.

Traditionally, glass properties are analyzed and understood in terms of the network forming and network modifying species present in the glass structure. To interpret the MLHL model similarly, we performed a SHAP analysis separately for the network forming and modifying elements (see Fig. 9). We observed that most of the network formers, except
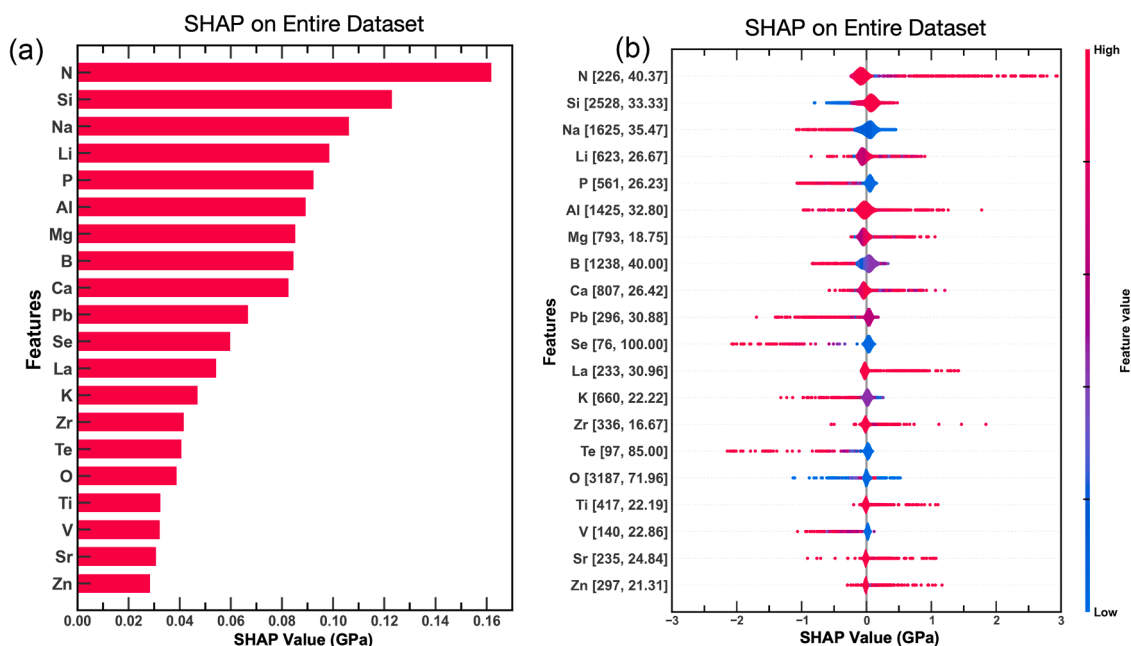
Fig. 8. SHAP plot for the (a) absolute impact of the chemical elements on hardness and (b) relative impact of the elements on hardness. In (b), the color of points represents the feature value, that is, in terms of the mol% of the oxide components in the glasses, with red representing high contents and blue representing low contents in the glass composition. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)
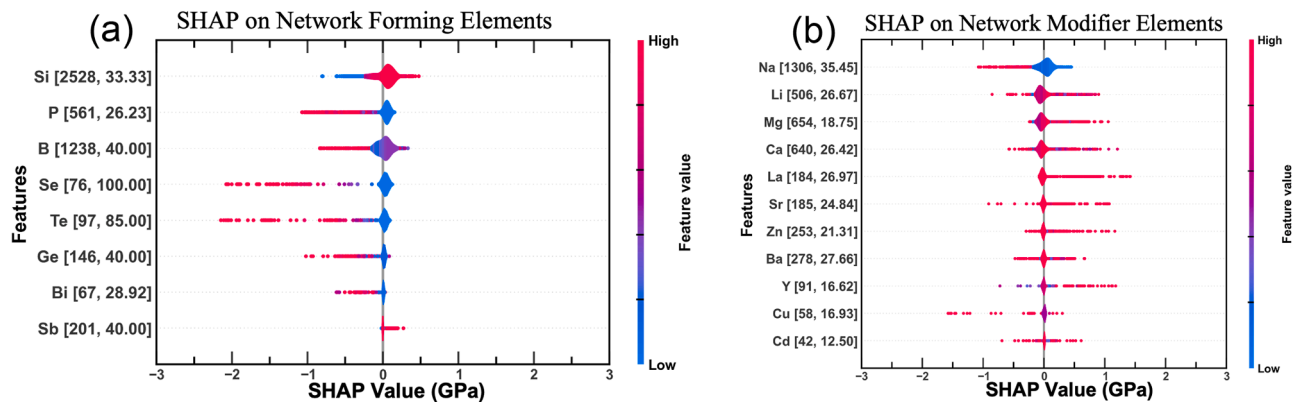
Fig. 9. SHAP values for (a) the glass network forming elements and (b) the glass network modifying elements. The color of points represents the magnitude of the mol% of the elements.

Si and B, decrease the hardness with increasing concentrations. To understand this, we can consider an existing model of glass hardness based on topological constraint theory [58]. Predicting hardness depends on the density of rigid bond constraints. That is, both the connectivity of the network and the atomic packing density govern glass hardness. For example, this explains why pure SiO_2 glass is relatively hard but is not among the hardest known oxide glasses. Neglecting some defects, SiO_2 glass has a fully polymerized structure with four bridging oxygens per tetrahedron, but this connectivity is counterbalanced by its very low packing density (open structure) [59]. Furthermore, network formers, such as P, have a double bonded terminal oxygen, which is not considered a part of the network. Therefore, a fully polymerized phosphorous atom is effectively three-fold coordinated, providing fewer bond constraints per atom than silicon [60].

Fig. 9(b) indicates that Na decreases the hardness. In contrast, other network modifiers either increase the hardness or exhibit mixed effects, that is, they exhibit different behavior depending on the other elements in the glass composition. This complex behavior can be understood as follows. While the network modifiers, such as Na, will generally tend to

decrease the network connectivity by breaking bridging oxygen bonds, they can also increase the atomic packing density (e.g., in the case of silicate glasses), thus creating competing effects on glass hardness, as discussed above. Moreover, in the case of, e.g., borate-containing glasses, modifiers can also increase the connectivity of network formers by converting boron from three- to four-fold coordination. Indeed, the SHAP values have proven to be insightful in determining the contribution of each feature (chemical elements) towards property value, and several agree with existing knowledge. For example, the Loewenstein rule and Boron anomaly are well-known and have been observed experimentally and predicted by SHAP in previous articles for other properties and here for hardness [25,31,61]. For convenience, we summarize some of these important results in Table 2 with suitable references corresponding to each effect predicted by our SHAP analysis. Note that these predictions should be considered as general guidelines, as there will be exceptions depending on the chemical environment.

Altogether, the analysis above presented can be useful for identifying the components that, on average, affect hardness positively or negatively, thereby enabling experimentalists to design new glasses in a

Table 2

Effect of increasing concentration of different elements on glass hardness.

Element	SHAP prediction for Hardness	Remarks
Na	Decreases	Depolymerize the network due to the formation of NBO [62].
Si	Increases	Increases the network connectivity [63].
B	Mixed effects	Boron anomaly [64].
P	Decreases	Depolymerize the network connectivity due to the lone pair of electrons [65].
Al	Mixed effects	Al-tetrahedra requires charge balancing Na-ions, which reduce the NBO. Further, Al tetrahedra tends to avoid the formation of Al–O–Al bonds tends to preferential bonding with other network formers present in the glass structure. This can further affect the network connectivity and packing of the glass structure, thus leading to mixed effects [66].

rational and effective fashion.

4. Conclusions

We used ML and symbolic reasoning to describe, predict, and interpret the chemical composition and load dependency of the hardness of inorganic glasses. To this end, we constructed a carefully curated dataset of glasses. We then employed a NN model for predicting the glass hardness, which exhibits good performance but is unable to capture the ISE. To address this challenge, we proposed and successfully tested a SRIMP framework combining symbolic reasoning Bernhardt's equation and ML. We demonstrate that the resulting SRIMP model can capture the ISE for completely unseen compositions and load values.

Further, the framework allows the deconvolution of hardness into load-independent and load-dependent terms. The load-independent hardness proposed here can be used as an inherent material property, which along with the load-dependent term, can predict the hardness corresponding to any load. Such a framework allows for easy and rational comparison of the glass hardness measured at different loads. Finally, employing the SHAP analysis, we interpreted the black-box ML models to identify the role of individual elements in governing the hardness. Such an approach enables the accelerated design of glasses with targeted hardness.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.actamat.2023.119046.

References

[1] M.I. Ojovan, W.E. Lee, Glassy wasteforms for nuclear waste immobilization, Metall. Mater. Trans. A 42 (2011) 837–851.

[2] R. R, J. Kaewkhao, Glass material and their advanced applications, KnE. Social. Sci (2019), https://doi.org/10.18502/kss.v3i18.4769.

[3] S. Yoshida, H. Tanaka, T. Hayashi, J. Matsuoka, N. Soga, Scratch Resistance of Sodium Borosilicate Glass, J. Ceram. Soc. Japan. 109 (2001) 511–515, https://doi.org/10.2109/jcersj.109.1270_511.

[4] T. Rouxel, J. Jang, U. Ramamurty, Indentation of glasses, Prog. Mater. Sci 121 (2021), 100834, https://doi.org/10.1016/j.pmatsci.2021.100834.

[5] S. Yoshida, Indentation deformation and cracking in oxide glass –toward understanding of crack nucleation, J. Non-Cryst. Solids:. X 1 (2019), 100009, https://doi.org/10.1016/j.nocx.2019.100009.

[6] K. Januchta, M.M. Smedskjaer, Indentation deformation in oxide glasses: quantification, structural changes, and relation to cracking, J. Non-Cryst. Solids:. X 1 (2019), 100007, https://doi.org/10.1016/j.nocx.2018.100007.

[7] G. Gehlhoff, M. Thomas, Die physikalischen Eigenschaften der Gläser in Abhängigkeit von der Zusammensetzung, Z. Techn. Physik. 6 (1926).

[8] S. Kasimuthumaniyan, N.N. Gosvami, N.M.A. Krishnan, Towards understanding the scratchability in functional glasses, Ceram. Int 47 (2021) 20821–20843, https://doi.org/10.1016/j.ceramint.2021.04.233.

[9] A.K. Varshneya, J.C. Mauro, Fundamentals of Inorganic Glasses, Elsevier, 2019.

[10] S. Shim, H. Bei, E.P. George, G.M. Pharr, A different type of indentation size effect, Scr. Mater 59 (2008) 1095–1098.

[11] N. Van Steenberge, J. Sort, A. Concustell, J. Das, S. Scudino, S. Suriñach, J. Eckert, M. Baró, Dynamic softening and indentation size effect in a Zr-based bulk glass-forming alloy, Scr. Mater 56 (2007) 605–608.

[12] G. Ryge, D.E. Foley, C.W. Fairhurst, Micro-indentation hardness, J. Dent. Res 40 (1961) 1116–1126.

[13] S. Sebastian, M.A. Khadar, Microhardness indentation size effect studies in 60B2O3-(40-x)PbO-xMCl2 and 50B2O3(50-x)PbO-xMCl2 (M = Pb, Cd) glasses, J. Mater. Sci 40 (2005) 1655–1659, https://doi.org/10.1007/s10853-005-0666-z.

[14] K. Redmond, Oxide glass exhibits plasticity without fracture at room temperature, MRS. Bull. 45 (2020) 78–79, https://doi.org/10.1557/mrs.2020.48.

[15] H. Li, R.C. Bradt, The indentation load/size effect and the measurement of the hardness of vitreous silica, J. Non. Cryst. Solids 146 (1992) 197–212, https://doi.org/10.1016/S0022-3093(05)80492-2.

[16] Y. Cao, M. Kazembeyki, L. Tang, N.M.A. Krishnan, M.M. Smedskjaer, C.G. Hoover, M. Bauchy, Modeling the nanoindentation response of silicate glasses by peridynamic simulations, J. Am. Ceram. Soc. 104 (2021) 3531–3544, https://doi.org/10.1111/jace.17720.

[17] S. Kasimuthumaniyan, A.A. Reddy, N.M.A. Krishnan, N.N. Gosvami, Understanding the role of post-indentation recovery on the hardness of glasses: case of silica, borate, and borosilicate glasses, J. Non. Cryst. Solids 534 (2020), 119955, https://doi.org/10.1016/j.jnoncrysol.2020.119955.

[18] M. Kazembeyki, M. Bauchy, C.G. Hoover, New insights into the indentation size effect in silicate glasses, J. Non. Cryst. Solids 521 (2019), 119494, https://doi.org/10.1016/j.jnoncrysol.2019.119494.

[19] E.O. Bernhardt, Über die Mikrohärte der Feststoffe im Grenzbereich des Kick'schen Ähnlichkeitssatzes, Zeitschrift. Für. Metallkunde 33 (1941) 135–144.

[20] N. Prentzas, A. Nicolaides, E. Kyriacou, A. Kakas, C. Pattichis, Integrating machine learning with symbolic reasoning to build an explainable AI model for stroke prediction, in: 2019 IEEE 19th International Conference on Bioinformatics and Bioengineering (BIBE), IEEE, Athens, Greece, 2019, pp. 817–821, https://doi.org/10.1109/BIBE.2019.00152.

[21] R. Ravinder, K.H. Sridhara, S. Bishnoi, H. Singh Grover, M. Bauchy, Jayadeva, H. Kodamana, N.M.A. Krishnan, Deep learning aided rational design of oxide glasses, Mater. Horiz (2020), https://doi.org/10.1039/D0MH00162G.

[22] R. Bhattoo, S. Bishnoi, M. Zaki, N.M.A. Krishnan, Understanding the compositional control on electrical, mechanical, optical, and physical properties of inorganic glasses with interpretable machine learning, Acta. Mater (2022), 118439, https://doi.org/10.1016/j.actamat.2022.118439.

[23] M. Zaki, N.M.A. Krishnan, Jayadeva, Extracting processing and testing parameters from materials science literature for improved property prediction of glasses, Chem. Engin. Proces (2021), 108607, https://doi.org/10.1016/j.cep.2021.108607.

[24] E. Alcobaça, S.M. Mastelini, T. Botari, B.A. Pimentel, D.R. Cassar, A.C.P. de L.F. de Carvalho, E.D. Zanotto, Explainable machine learning algorithms for predicting glass transition temperatures, Acta. Mater 188 (2020) 92–100, https://doi.org/10.1016/j.actamat.2020.01.047.

[25] D.R. Cassar, S.M. Mastelini, T. Botari, E. Alcobaça, A.C.P.L.F. de Carvalho, E.D. Zanotto, Predicting and interpreting oxide glass properties by machine learning using large datasets, Ceram. Int 47 (2021) 23958–23972, https://doi.org/10.1016/j.ceramint.2021.05.105.

[26] D.R. Cassar, G.G. Santos, E.D. Zanotto, Designing optical glasses by machine learning coupled with a genetic algorithm, Ceram. Int 47 (2021) 10555–10564, https://doi.org/10.1016/j.ceramint.2020.12.167.

[27] S.M. Mastelini, D.R. Cassar, E. Alcobaça, T. Botari, A.C.P.L.F. de Carvalho, E.D. Zanotto, Machine learning unveils composition-property relationships in

chalcogenide glasses, Acta. Mater 240 (2022), 118302, https://doi.org/10.1016/j.actamat.2022.118302.

[28] S. Singla, S. Mannan, M Zaki, N.A. Krishnan, Accelerated design of chalcogenide glasses through interpretable machine learning for composition–property relationships, J.Phys. Materials 6 (2) (2023), 024003, https://doi.org/10.1088/2515-7639/acc6f2.

[29] S.M. Mastelini, D.R. Cassar, E. Alcobaca, T. Botari, A.C.P.L.F. de Carvalho, E.D. Zanotto, Machine learning unveils composition-property relationships in chalcogenide glasses, Acta. Mater 240 (2022), 118302, https://doi.org/10.1016/j.actamat.2022.118302.

[30] V.Venugopal Ravinder, S. Bishnoi, S. Singh, M. Zaki, H.S. Grover, M. Bauchy, M. Agarwal, N.M.A. Krishnan, Artificial intelligence and machine learning in glass science and technology: 21 challenges for the 21st century, Int. J. Appl. Glass. Sci 12 (2021) 277–292, https://doi.org/10.1111/ijag.15881.

[31] M. Zaki, V. Venugopal, R. Bhattoo, S. Bishnoi, S.K. Singh, A.R. Allu, N.M.A. Krishnan Jayadeva, Interpreting the optical properties of oxide glasses with machine learning and Shapely additive explanations, J. Am. Ceram. Soc. 105 (2022) 4046–4057, https://doi.org/10.1111/jace.18345.

[32] D.R. Cassar, S.M. Mastelini, T. Botari, E. Alcobaça, A.C.P.L.F. de Carvalho, E.D. Zanotto, Predicting and interpreting oxide glass properties by machine learning using large datasets, Ceram. Int (2021), https://doi.org/10.1016/j.ceramint.2021.05.105.

[33] S. Singla, S. Mannan, M. Zaki, N. Krishnan, ArXiv Preprint, in: ArXiv Preprint, 2211, ArXiv, 2022, p. 00691.

[34] N. Khatavkar, S. Swetlana, A.K. Singh, Accelerated prediction of Vickers hardness of Co- and Ni-based superalloys from microstructure and composition using advanced image processing techniques and machine learning, Acta. Mater 196 (2020) 295–303, https://doi.org/10.1016/j.actamat.2020.06.042.

[35] X. Liu, Z. Long, L. Peng, Prediction of Vickers hardness of amorphous alloys based on interpretable machine learning, J. Non. Cryst. Solids 602 (2023), 122095, https://doi.org/10.1016/j.jnoncrysol.2022.122095.

[36] L. Lu, M. Dao, P. Kumar, U. Ramamurty, G.E. Karniadakis, S. Suresh, Extraction of mechanical properties of materials through deep learning from instrumented indentation, Proc. Natl. Acad. Sci. U.S.A. 117 (2020) 7052–7062, https://doi.org/10.1073/pnas.1922210117.

[37] G.E. Karniadakis, I.G. Kevrekidis, L. Lu, P. Perdikaris, S. Wang, L. Yang, Physics-informed machine learning, Nat. Rev. Phys 3 (2021) 422–440, https://doi.org/10.1038/s42254-021-00314-5.

[38] D.R. Cassar, ViscNet: neural network for predicting the fragility index and the temperature-dependency of viscosity, Acta. Mater 206 (2021), 116602, https://doi.org/10.1016/j.actamat.2020.116602.

[39] M.L. Bødker, M. Bauchy, T. Du, J.C. Mauro, M.M. Smedskjaer, Predicting glass structure by physics-informed machine learning, NPJ. Comput. Mater 8 (2022) 192, https://doi.org/10.1038/s41524-022-00882-9.

[40] X. Liu, C.E. Athanasiou, N.P. Padture, B.W. Sheldon, H. Gao, Knowledge extraction and transfer in data-driven fracture mechanics, Proc. Natl. Acad. Sci 118 (23) (2021), https://doi.org/10.1073/pnas.2104765118 e2104765118.

[41] F. Zhu, J. Leng, J.W. Jiang, T. Chang, T. Zhang, H. Gao, Thermal-fluctuation gradient induced tangential entropic forces in layered two-dimensional materials, J. Mech. Phys. Solids 163 (2022), 104871, https://doi.org/10.1016/j.jmps.2022.104871.

[42] J.F. Christensen, N.A. Krishnan, M. Bauchy, M.M. Smedskjaer, Indenting glasses with indenters of varying stiffness and sharpness, J. Non. Cryst. Solids 603 (2023) 122111, https://doi.org/10.1016/j.jnoncrysol.2022.122111.

[43] S. Striepe, M.M. Smedskjaer, J. Deubener, U. Bauer, H. Behrens, M. Potuzak, R.E. Youngman, J.C. Mauro, Y. Yue, Elastic and micromechanical properties of isostatically compressed soda–lime–borate glasses, J. Non. Cryst. Solids 364 (2013) 44–52, https://doi.org/10.1016/j.jnoncrysol.2013.01.009.

[44] A.S. Hakeem, J. Grins, S. Esmaeilzadeh, La-Si-O-N glasses, Part II: vickers hardness and refractive index, J. Eur. Ceram. Soc 27 (2007) 4783–4787, https://doi.org/10.1016/j.jeurceramsoc.2007.04.003.

[45] A.K. Dutta, N. Narasaiah, A.B. Chattopadhyaya, K.K. Ray, The load dependence of hardness in alumina–silver composites, Ceram. Int 27 (2001) 407–413, https://doi.org/10.1016/S0272-8842(00)00095-X.

[46] Z.H. Li, J.H. Gong, Z.J. Peng, H.Z. Miao, Indentation size effect in the nanohardness of soda-lime glass, KEM. 280–283 (2007) 1659–1662, https://doi.org/10.4028/www.scientific.net/KEM.280-283.1659.

[47] A.R. Alao, M.H. Danish Bujang, Load effect on the mechanical behaviour of zirconia-reinforced lithium silicate glass ceramics, Ceram. Int 47 (2021) 1353–1363, https://doi.org/10.1016/j.ceramint.2020.08.257.

[48] E.O. Bernhardt, Über die Mikrohärte der Feststoffe im Grenzbereich des Kick'schen Ähnlichkeitssatzes, Int. J. Mater. Res. 33 (1941) 135–144, https://doi.org/10.1515/ijmr-1941-330305.

[49] sklearn.preprocessing.StandardScaler, Scikit-Learn. (n.d.). https://scikit-learn.org/stable/modules/generated/sklearn.preprocessing.StandardScaler.html (accessed January 5, 2023).

[50] S. Kaufman, S. Rosset, C. Perlich, O. Stitelman, Leakage in data mining: formulation, detection, and avoidance, ACM. Trans. Knowl. Discov. Data 6 (2012), https://doi.org/10.1145/2382577.2382579, 15:1-15:21.

[51] PyTorch, (n.d.). https://www.pytorch.org (accessed October 14, 2022).

[52] C.R. Harris, K.J. Millman, S.J. van der Walt, R. Gommers, P. Virtanen, D. Cournapeau, E. Wieser, J. Taylor, S. Berg, N.J. Smith, R. Kern, M. Picus, S. Hoyer, M.H. van Kerkwijk, M. Brett, A. Haldane, J.F. del Río, M. Wiebe, P. Peterson, P. Gérard-Marchant, K. Sheppard, T. Reddy, W. Weckesser, H. Abbasi, C. Gohlke, T.E. Oliphant, Array programming with NumPy, Nature. 585 (2020) 357–362, https://doi.org/10.1038/s41586-020-2649-2.

[53] J. Reback, W. McKinney, jbrockmendel, J.V. den Bossche, T. Augspurger, P. Cloud, gfyoung, Sinhrks, S. Hawkins, M. Roeschke, A. Klein, T. Petersen, J. Tratner, C. She, W. Ayd, S. Naveh, M. Garcia, J. Schendel, A. Hayden, D. Saxton, V. Jancauskas, A. McMaster, P. Battiston, patrick, S. Seabold, K. Dong, chris-b1, h-vetinari, S. Hoyer, M. Gorelli, pandas-dev/pandas: pandas 1.2.0rc0, (2020). https://doi.org/10.5281/zenodo.4311557.

[54] J.D. Hunter, Matplotlib: a 2D Graphics Environment, Comput. Sci. Eng 9 (2007) 90–95, https://doi.org/10.1109/MCSE.2007.55.

[55] S.M. Lundberg, S.I. Lee, A unified approach to interpreting model predictions, Adv. Neural. Inf. Process. Syst (2017) 30.

[56] X. Lu, L. Deng, J. Du, J.D. Vienna, Predicting boron coordination in multicomponent borate and borosilicate glasses using analytical models and machine learning, J. Non. Cryst. Solids (2020), 120490, https://doi.org/10.1016/j.jnoncrysol.2020.120490.

[57] W. Loewenstein, The distribution of aluminum in the tetrahedra of silicates and aluminates, Am. Mineral. 39 (1954) 92–96.

[58] Q. Zheng, Y. Yue, J.C. Mauro, Density of topological constraints as a metric for predicting glass hardness, Appl. Phys. Lett 111 (2017), 011907.

[59] T. Rouxel, Elastic Properties and Short-to Medium-Range Order in Glasses, J. Am. Ceram. Soc. 90 (2007) 3019–3039, https://doi.org/10.1111/j.1551-2916.2007.01945.x.

[60] C. Hermansen, X. Guo, R.E. Youngman, J.C. Mauro, M.M. Smedskjaer, Y. Yue, Structure-topology-property correlations of sodium phosphosilicate glasses, J. Chem. Phys. 143 (2015) 064510, https://doi.org/10.1063/1.4928330.

[61] R. Bhattoo, S. Bishnoi, M. Zaki, N.M.A. Krishnan, Understanding the compositional control on electrical, mechanical, optical, and physical properties of inorganic glasses with interpretable machine learning, Acta. Mater 242 (2023), 118439, https://doi.org/10.1016/j.actamat.2022.118439.

[62] Y. B. Stevensson, M. Edén, Structural Role of Sodium in Borosilicate, Phosphosilicate, and Borophosphosilicate Glasses Unveiled by Solid-State NMR and MD Simulations, J. Phys. Chem 123 (2019) 25816–25832, https://doi.org/10.1021/acs.jpcc.9b06823.

[63] Q. Zheng, Y. Yue, J.C. Mauro, Density of topological constraints as a metric for predicting glass hardness, Appl. Phys. Lett 111 (2017), 011907, https://doi.org/10.1063/1.4991971.

[64] X. Lu, L. Deng, J. Du, J.D. Vienna, Predicting boron coordination in multicomponent borate and borosilicate glasses using analytical models and machine learning, J. Non. Cryst. Solids 553 (2021), 120490, https://doi.org/10.1016/j.jnoncrysol.2020.120490.

[65] C. Hermansen, X. Guo, R.E. Youngman, J.C. Mauro, M.M. Smedskjaer, Y. Yue, Structure-topology-property correlations of sodium phosphosilicate glasses, J. Chem. Phys 143 (2015), 064510, https://doi.org/10.1063/1.4928330.

[66] W. Loewenstein, The distribution of aluminum in the tetrahedra of silicates and aluminates, Am. Mineral. (1954). https://www.semanticscholar.org/paper/The-distribution-of-aluminum-in-the-tetrahedra-of-Loewenstein/bcc1e1d4b8a1e8bdebfddcca0d329f5dc93dd32e. accessed April 5, 2023.