Acta Materialia 159 (2018) 249–256

Contents lists available at ScienceDirect

Acta Materialia

journal homepage: www.elsevier.com/locate/actamat

Full length article

Predicting glass transition temperatures using neural networks

Daniel R. Cassar [a, *], André C.P.L.F. de Carvalho [b], Edgar D. Zanotto [a]

[a] Department of Materials Engineering, Center for Research, Technology and Education in Vitreous Materials, Federal University of São Carlos, São Carlos, SP, Brazil

[b] Institute of Mathematics and Computer Sciences, Center for Research in Mathematical Sciences Applied to Industry, University of São Paulo, São Carlos, SP, Brazil

ARTICLE INFO

Article history:
Received 12 June 2018
Received in revised form
8 August 2018
Accepted 10 August 2018
Available online 17 August 2018

Keywords:
Property prediction
Glass transition
Oxide glasses
Machine leaning
Artificial neural networks

ABSTRACT

The glass transition temperature (T_g) is a kinetic property of major importance for both fundamental and applied glass science. In this study, we designed and trained an artificial neural network to induce a model that can predict the T_g of multicomponent oxide glasses. To do this, we used a dataset containing more than 55,000 inorganic glass compositions and their respective experimental values of T_g. These compositions contain from 3 to 21 of the 45 chemical elements studied here. We implemented an optimization procedure to find artificial neural network hyperparameter values that were able to induce a model with high predictive performance. The resulting neural network model can correctly predict, with 95% accuracy, the published T_g value within less than ±9% error, whereas 90% of the data are predicted with a relative deviation lower than ±6%. This level of uncertainty is equivalent to the level present in the original dataset and allows a very satisfactory description of the T_g for multicomponent oxide glasses containing combinations of the 45 studied chemical elements. The prediction uncertainty does not depend on the number of elements in the glass composition. However, it is larger for glasses having very high T_g (above 1250 K). The most important aspect is the algorithm's ability to predict the T_g of glasses that are not included in the experimental dataset used for training, thus showing a high generalization ability. Besides, the procedure used here is general and can be easily extended to predict several other properties as a function of the glass composition. This handy feature will most probably help to develop new multicomponent glass compositions having remarkable properties.

1. Introduction

The glass transition temperature (T_g) is a defining parameter that is related to a dramatic change in the thermo-mechanical properties of non-crystalline materials [1]. It marks the temperature range below which the atoms of a supercooled liquid are temporarily frozen (without crystallizing) upon cooling. "Glass transition" is the second most cited keyword in the history of glass science [2], and is the guiding parameter used in important processes, such as annealing (to relieve spurious residual stresses that might break the product), thermal tempering (to strengthen glass products), ion-exchange (e.g. the Gorilla Glass® of smartphones), and crystallization (to control spontaneous devitrification). It is difficult to overstate its importance within the scope of Physics, Chemistry, and Materials Science and Technology [3].

Once glasses are continuous solutions of chemical elements, without fixed stoichiometry and long-range structural order, the possible number of vitreous compositions is gigantic. For instance, with 1% combinations of the 80 "friendly" elements of the periodic table, it is theoretically possible to synthesize 10^{52} compositions [4]. However, the number of inorganic glasses produced, tested and reported so far is "only" around 10^5 [5]. Therefore, there is a large window of opportunity for searching novel compositions, which can show very unusual and interesting properties.

Developing new glasses has been an empirical endeavor [6], mainly guided by educated guesses and the "cook and look" approach. This approach is expensive and time-consuming, less than ideal to support new technological demands in a fast-paced society. This fact has always been a major driving force to enhance our capacity to predict key properties of glasses and other materials with reasonable accuracy. Therefore, it is now timely to move from the traditional trial-and-error, "get lucky" method, to a

* Corresponding author.
E-mail addresses: daniel.r.cassar@gmail.com (D.R. Cassar), dedz@ufscar.br (E.D. Zanotto).
URL: http://www.cemeai.icmc.usp.br/, http://www.certev.ufscar.br

https://doi.org/10.1016/j.actamat.2018.08.022
1359-6454/

new era of predicting compositions that might lead to novel glasses having noteworthy, exotic and yet useful combinations of properties.

Within this context, some tools are available—ranging from theoretical to purely empirical— based on concepts from scientific domains such as the density functional theory, molecular dynamics simulations, the topological constraint theory, evolutionary algorithms, and big data analysis [6]. However, current theoretical tools are still limited to the simplest materials, with typically less than five different components due to their high computational demand. In this setting, machine learning in general [7–9] and artificial neural networks (ANNs), in particular, have been successfully used to identify patterns in large datasets [10].

When dealing with highly complex compositions, having more than five components, one has to resort to an empirical model. This study investigates the use of ANNs, which have been successfully employed to identify objects and facial expressions in pictures [10,11], recognize speech [12], discover new drugs [13,14] and play board games [15] and video games [16] at a level similar or superior to humans. More specifically, in Materials Science, ANNs have been used with moderate success to predict kinetic and mechanical properties of polymers [17–22], the glass forming ability of metallic alloys [23–25], and extract structural information from X-ray diffraction data [26], for example.

In the realm of *oxide glasses*, by far the most important family of commercial glasses, only two published studies have tested the applicability of ANNs to predict a property, chemical durability: the work of Brauer et al. on the solubility of phosphate glasses [27], and the recent work of Krishnan et al. on the solubility of silicate glasses [28]. However, the numbers of examples used in these two studies were quite small (31 and 299, respectively), which is not recommended [29] for building ANN models due to a strong possibility of overfitting. Nevertheless, both reported that the used ANNs reasonably predicted the chemical solubility of the investigated glasses, with R^2 values of 0.9996 and 0.982, respectively.

Two other related references [30,31] dealt with estimates of the *liquidus* temperature of oxide mixtures, T_{liq}. Although this is not a glass property *per se*, the knowledge of T_{liq} is critical for optimum glass melting operations to avoid undesirable, uncontrolled devitrification. Dreyfus & Dreyfus [30] tested four different datasets separated per composition system, each containing from 155 to 893 oxides. The tested systems contained 4 to 6 different chemical elements. The ANN used was able to predict the T_{liq} with a standard error of 20 K. Mauro et al. [31] studied 851 silicate glasses containing up to 9 different chemical elements. Their best ANN was able to predict the T_{liq} with an R^2 of 0.947, a high predictive performance for this particular application.

In this study, for the first time, we report a very successful application of ANNs to predict the glass transition temperature of multicomponent oxide glasses with a similar degree of accuracy as the original dataset used for training. The fingerprint used to predict the glass transition temperature was the composition of the glasses (in at%), for which we studied the influence of 45 different chemical elements using a dataset of 55,150 examples. More information on the data collection is available in the supplementary material (references [32–39] are cited only in the supplementary material).

2. Data collection

We collected T_g data from the SciGlass 7.12 database, which in turn were gathered from academic literature and patents published up to May 2014 [5]. We imposed the following composition constraints in our query:

- An atomic fraction of at least 10% of oxygen must be present. This constraint was adopted to limit our dataset to the scope of this research, oxide glasses, by far the most abundant and commercially important;
- Each other chemical element must be present in at least 1% of the entries in the dataset. We wanted each atom in the analysis to have a minimum representation during the supervised training and validation;
- We only considered compositions that are made of at least two different chemical elements plus oxygen. Single oxides (such as SiO_2, B_2O_3, GeO_2, TeO_2, and others) are susceptible to considerable uncertainties because unavoidable (and unaccounted) impurities usually have a powerful impact on their T_g, which could jeopardize our analysis.

After applying these constraints, we obtained a dataset with 55,150 examples. Table 1 shows the chemical elements present in the studied compositions (45 in total), their maximum atomic percentage in any composition, and the number of compositions that contain this element. The most important oxide glass-forming

Table 1
Number of examples of each of the 45 studied chemical elements. The rightmost column is the maximum amount of a given element that is present in any glass in our dataset. The minimum amount is zero for all elements except oxygen, for which it is 10%.

Element	Compositions that contain the element	Maximum amount (at%)
Cd	565	25.0
Yb	620	16.7
Cs	654	45.6
N	707	43.8
Mn	825	23.1
S	981	58.8
Ce	985	15.8
Er	999	16.1
I	1038	46.3
Mo	1328	22.0
Cl	1364	63.2
As	1379	39.9
Ga	1410	33.4
Cu	1711	40.6
Sn	1998	28.3
Ag	2203	47.8
Ta	2306	28.3
Y	2340	18.4
Gd	2587	17.9
Ge	2620	33.3
V	3190	28.3
Fe	3322	31.6
W	3998	23.1
F	4230	64.4
Sb	4335	47.0
Sr	4525	40.9
Te	4564	32.9
Nb	4665	26.3
Bi	5352	37.6
La	6036	31.0
Pb	6361	47.6
Zr	7051	20.2
Ti	7614	27.3
Mg	9155	48.5
Ba	11,049	24.4
K	12,000	49.7
Ca	13,397	30.8
Zn	13,862	32.1
Li	14,326	58.4
P	14,493	28.1
Al	19,104	36.7
Na	19,731	56.7
B	25,186	40.1
Si	27,819	33.2
O	55,150	74.5

ions are included: silicon, boron, phosphorous, tellurium, germanium, arsenic, and antimony.

Fig. 1a and Fig. 1b complement the characterization of our dataset by showing the frequency of T_g values and the frequency of the number of different chemical elements per composition, respectively. The majority of the studied glasses have a T_g near 750 K, but the values span a range of 1150 K (from near 300 K to approximately 1450 K). Most oxide glasses are made with 4 different elements (including oxygen). Only about 10% of the compositions of the dataset contain 11 or more different elements.

Before they were fed to the ANNs, the values of T_g in the dataset were normalized by dividing them by the maximum T_g. This procedure is a common normalization practice for regression problems.

3. Designing the artificial neural networks

Detailed information about artificial neural networks, how they are built, trained, validated and used are available in various publications [10,16,40,41]. Briefly, these networks are made of artificial neurons that are interconnected by artificial synapses. These neurons receive numerical input data (either the raw data or data from other neurons) and apply a mathematical function whose output can either be the final network result (T_g, in our case) or be used as input for other neurons. The input values received by each neuron are weighted by an associated synapse. An ANN is usually trained by modifying these weights in order to improve the ANN prediction of the desired, correct output (in this case, to improve the prediction of the T_g values). The glossary in the supplementary material defines some relevant terms.

There are many types of ANNs and some are used for regression problems, as is the case of the network investigated in this article. After several experiments, we settled with the multilayer perceptron ANN [42,43], MLP, with two hidden layers, trained with the back-propagation learning algorithm. From now on, when mentioning ANNs, we mean MLP networks. One of the main challenges of using MLP is to avoid overfitting, which occurs when the network presents low generalization ability. We used two methods to reduce the occurrence of overfitting:

- we included a *dropout* routine [44] for the hidden layers. In this method, for each epoch of the training process, each neuron has a chance (dropout percentage) to be inactivated, not being updated and not propagating information during the epoch. Thus, the MLP network must be noise-robust;
- we used an *early stop* routine, instead of a fixed number of epochs, to train the MLP networks. In this case, the training is stopped when a certain degree of overfitting is identified, which is controlled by a hyperparameter called *patience*. This hyperparameter represents the number of epochs without improvement after which the training will stop. The improvement of the network predictive performance is measured by assessing how it performs in a validation set.

Machine learning algorithms have two types of variables that affect the performance of the induced model: *parameters* and *hyperparameters*. Parameters are model variables whose values can change during the model training. The synapse weights are examples of parameters of an artificial neural network. On the other hand, hyperparameters, also named free parameters, are variables whose values must be determined before training and do not change during training. The number and distribution of neurons of an ANN are examples of hyperparameters.

A typical ANN design challenge is to find a good set of hyperparameter values. It is possible to arbitrarily choose the hyperparameter values and check if they induce a model with high predictive accuracy. However, it is difficult to assess how optimized these hyperparameters are by only analyzing the trained network. In other words, the network itself cannot assess if its topology is optimized or if there is still room for improvement. In this framework, some combinations of hyperparameter values need to be identified and tested to search for the most suitable ANN topology.

Some hyperparameters for the investigated application are the *number of neurons* in the first and second hidden layers; the *dropout percentage* for the first and second hidden layers; the *batch size*; and the *patience*. We used optimization routines to search for a good combination of these hyperparameter values. Other hyperparameters of the ANN were set following a combination of conventional practices from the literature together with results from previous tests. For instance, we used the rectifier linear unit (ReLU) as the *activation function* for the hidden layers, and a linear activation function for the output layer. We also used the Huber *loss function* [45] because it is robust against outliers. We used the Adam (Adaptive Moment Estimation) gradient descent *optimization algorithm* [46] to adjust the network parameters during training.

Table 2 shows the search space for tuning the hyperparameters.

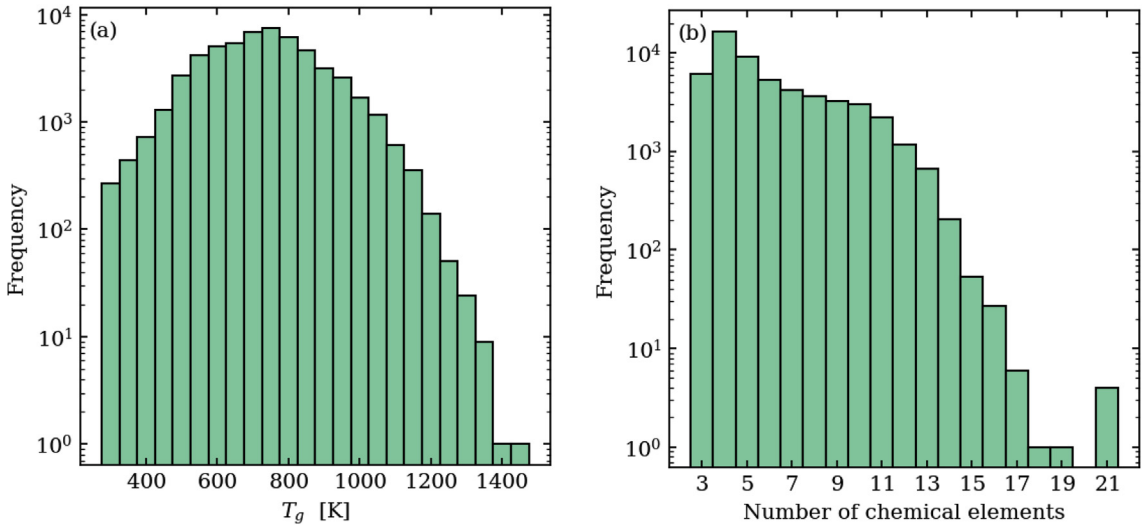

Fig. 1. Frequency of (**a**) T_g and (**b**) number of chemical elements in the glasses of the studied dataset. The y-axis is in logarithm scale.

Table 2
Lower and upper search space limits used during hyperparameter optimization and the best setting found.

Hyperparameter	Lower search space limit	Upper search space limit	Best topology
1st hidden layer			
Nº of neurons	60	500	350
Dropout (%)	0	25	6
2nd hidden layer			
Nº of neurons	50	500	500
Dropout (%)	0	50	11
Batch (2^n)	n = 5	n = 10	n = 7
Patience	7	15	15

The lower and upper values for the search space were obtained after various experiments. It is interesting to note that the ANN was allowed to have a zero dropout percentage, which means that no dropout routine would be present. Ultimately, choosing to use the dropout routine was left to the hyperparameter optimization algorithm. In the end, the best hyperparameter set had some degree of dropout.

Table 2 shows that the search space was far too ample to allow for an extensive grid search. Instead, the optimization algorithm progressed through the search space guided by a Tree-structured Parzen Estimator (TPE) algorithm [47,48]. The objective of the optimization algorithm was to minimize the mean validation loss in a stratified 10-fold cross-validation test, for which we used 8 folds for training, 1 fold for validation, and 1 fold for testing.

The learning algorithm stopped after 500 sets of the hyperparameter values were tested. From these, we selected the 30 sets that yielded the lowest mean loss. For each one of these selected sets, we computed the mean validation loss of a repeated stratified 10-fold cross-validation test (5 repetitions). Thus, we aimed to minimize any bias arising because of the random selection of data during the cross-validation.

Finally, we selected the topology with the hyperparameter values that yielded the lowest mean validation loss in the previous test. We trained a network with the selected topology, shown in Table 2, using (randomly chosen) 80% of all the data for training, 10% for validation and 10% for testing.

4. Testing the induced artificial neural network

The generalization capability of the induced ANN is shown in Fig. 2, with the predicted T_g versus the experimental T_g of the test dataset (containing 5515 examples). The R^2 of the prediction is 0.998. The network predicted 95% of the data with a relative deviation ϕ lower than ±9%, whereas 90% of the data were predicted with a ϕ lower than ±6%. This is a very relevant result since the scattering in the reported T_g in our dataset is of the same order.

To estimate the scatter of our dataset, we computed the relative neighborhood deviation (RND) for all the studied examples. Two examples are considered to be in the same neighborhood if the Euclidean distance of their composition vectors (in at%) is less than or equal to 1%. The reasoning for this is that it is not expected that T_g changes significantly for compositions in the same neighborhood. If $\check{T}_g$ is the set of all T_g values of compositions in a given neighborhood, then their RND is defined as:

$$\mathrm{RND} = 100\frac{\max\left(\check{T}_g\right) - \min\left(\check{T}_g\right)}{\max\left(\check{T}_g\right) + \min\left(\check{T}_g\right)}. \quad (1)$$

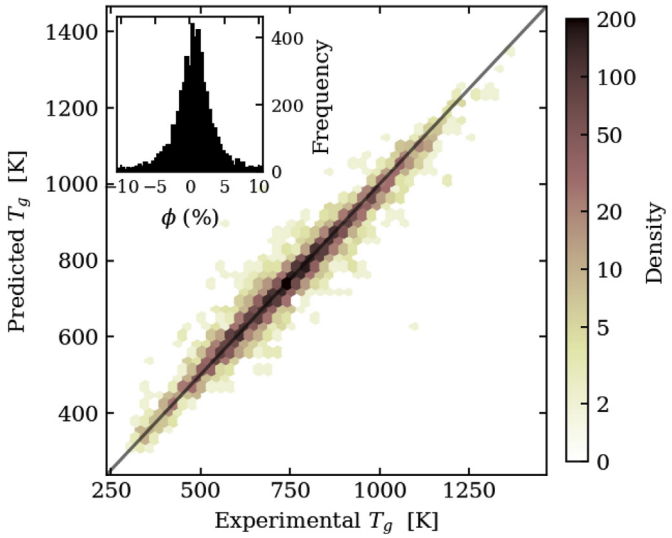

Fig. 2. Predicted T_g by the trained network *versus* reported values for the test dataset containing 5515 oxide glasses. The straight line is the identity where the prediction is equal to the reported value. The inset is the distribution of the relative deviation ϕ.

Fig. 3 shows the cumulative frequency of the RND—which is a measurement of the intrinsic data scatter—of all possible neighborhoods in our dataset. It should be mentioned that 95% of the data showed an RND of 6% or less. Therefore, it would not be possible to predict the data in this dataset with a better accuracy without overfitting or removing some data.

Another way of visualizing the generalization capability of the trained network is by looking at the residuals of the predicted T_g of the test set, shown in Fig. 4. For high values of T_g (above 1250 K), the prediction error is larger than expected, with a median residual of about 50 K, whereas all other ranges of T_g have a median residual not larger than 20 K. This occurrence is probably due to a combination of a small number of compositions (examples) having high values of T_g (only 86 examples with T_g above 1250 K in the dataset) combined with the usual high uncertainty associated to high-temperature measurement.

Another relevant information yielded by Fig. 4 is a subtle trend to underestimate the T_g value for compositions whose reported T_g is above 900 K. It is, however, essential to keep in mind that the

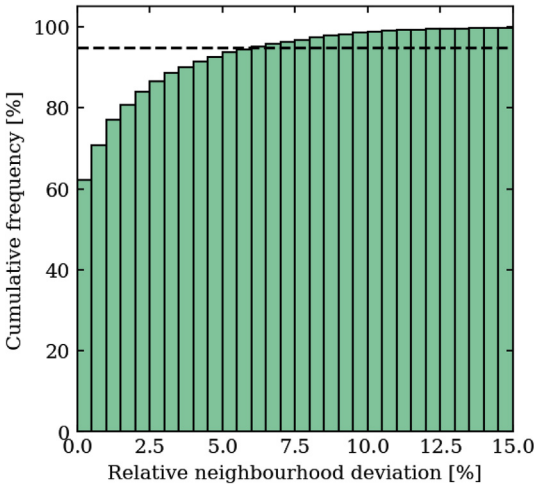

Fig. 3. Cumulative frequency of the relative neighborhood deviation (RND). The dashed black line shows that 95% of the data have an RND of 6% or less.

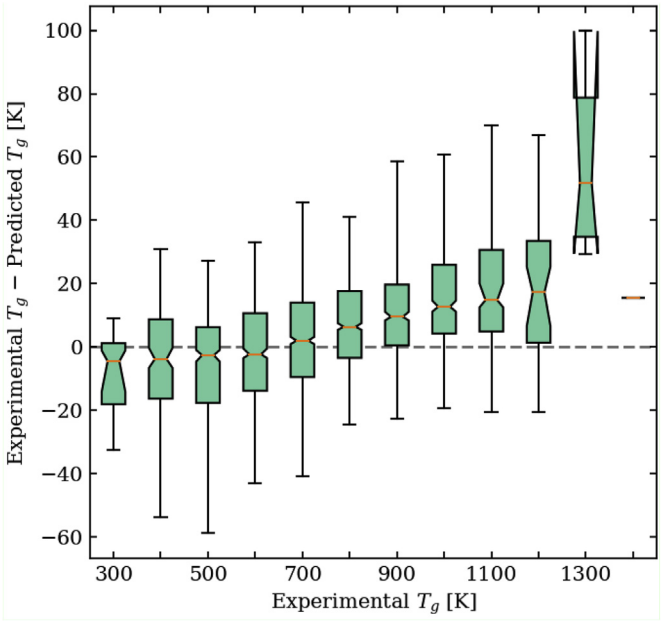

Fig. 4. Notched boxplot of the residuals of the test set predicted by the trained network as a function of the reported T_g. Each box spans a temperature range of 100 K in experimental T_g. The boxes are bound by the first and third quartiles (Q1 and Q3) and have caps that comprehend 90% of the data. The notch is the 95% confidence interval of the median considering 5000 bootstrap samples. Only one composition in this set had a reported T_g above 1350 K.

primary use of this particular ANN model is to help us predict the T_g of potentially exciting new glass compositions having desired properties. Therefore, the primary aim of its application is to make a fast *estimate* of the expected value of T_g (for any given composition) to guide actual experiments.

According to Fig. 2, the induced ANN model poorly describes some data points. We firmly believe this has to do with particularities of the glass transition temperature and its measurement, which sometimes lead to large errors and is likely the main reason for the intrinsic scatter in our dataset. It is also well known that the glass transition temperature of a given composition, being a kinetic property, depends on the time of the experiment. Thus, the rate at which temperature and pressure changes will affect the resulting glass transition temperature.

Moreover, the changes associated with the glass transition happen within a *range* of temperatures. This means that the *same* experimental data can yield *different* values of T_g if one researcher chooses the onset and another chooses the mid- or endpoint of the transition; hence researchers may disagree on the value of T_g when analyzing the same data (if, of course, no standard way of measurement is agreed beforehand). This lack of standardization alone can amount to several degrees. T_g also depends on the frozen non-crystalline structure of the glass, which means that a glass annealed at a low-temperature will have a different T_g than a fast-cooled fiber or thin film of the same composition. Additionally, T_g can be measured by various techniques, such as dilatometry, differential scanning calorimetry, and dynamic mechanical analysis, among others. Each of these techniques has its particularities (different heat transfer efficiency, heat conduction, temperature measurement apparatus, sensor type and placement, etc.) that may affect the result of the experiment. Finally, certain impurities, such as OH^-, have an enormous impact on T_g and are rarely accounted for in the chemical composition of glasses.

All these variables are natural sources of deviation in the published T_g data. On top of that, it is almost inevitable that in more than 55,000 reported T_g values, some are merely erroneous data. Some (hopefully a small fraction) are typos from the authors or the people that fed the database, whereas others deviate from the real value because of inaccurate analysis, incorrect chemical compositions, mistakes during sample handling, uncalibrated equipment, etc. Therefore, a fraction of the data significantly deviates from the mean value of T_g (for any given composition) due to one or more of the mentioned peculiarities. This readily explains why 5% of the whole dataset is poorly described by the best induced ANN model. Fortunately, these data with significant deviations from the mean did not pollute our analysis because they were few and because we used a robust loss function.

We also assessed the predictive power of the induced ANN model by closely studying part of the experimental data. For instance, Fig. 5a shows the data for all compositions having only sodium, silicon, and oxygen. It can be observed that the typical data scatter is very large. In this same plot, we set the transparency of

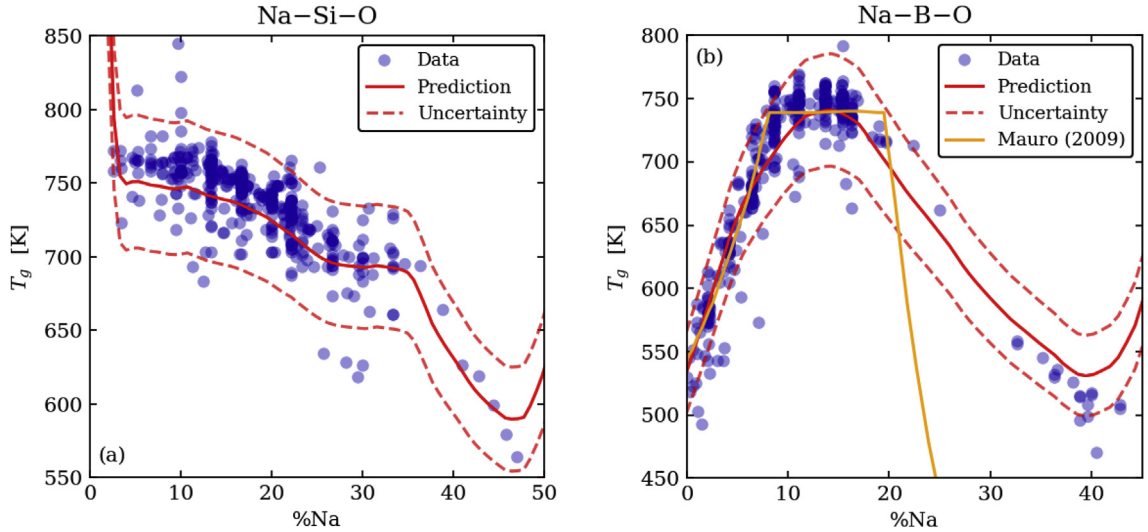

Fig. 5. Experimental T_g data and prediction bands for (**a**) sodium silicate and (**b**) sodium borate glasses. The orange line is the prediction of T_g by a topological model [50]. (For interpretation of the references to color in this figure, the reader is referred to the Web version of this article.)

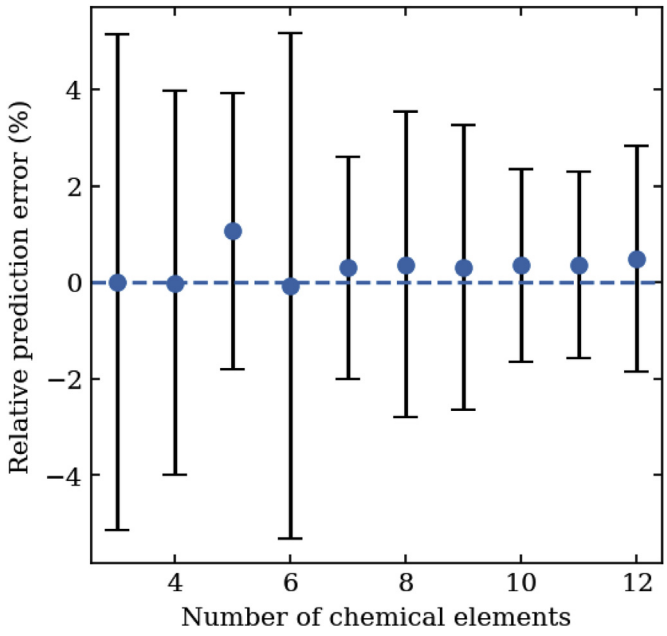

Fig. 6. Mean and standard deviation of the relative prediction error in T_g versus the number of chemical elements in the composition. For each given number of chemical elements, 50 different examples were randomly drawn from the test set.

Table 3
Test of the T_g prediction capability of the induced network for glass compositions containing only two chemical elements showing the mean and standard deviation, minimum, and maximum reported values of T_g, as well as the prediction from the trained model. All values are in Kelvin.

Composition	T_g (mean and std)	Min. T_g	Max. T_g	Predicted T_g
SiO_2	1440 (60)	1293	1495	1210
B_2O_3	540 (20)	483	580	534
GeO_2	800 (30)	743	883	753
V_2O_5	484 (3)	482	492	497
TeO_2	580 (40)	498	658	573
P_2O_5	610 (50)	536	673	609
Sb_2O_3	550 (50)	500	618	547
As_2O_3	441 (7)	433	453	438

the scatted points in such a way that is possible to check the regions that are denser on data points. The predictions of the induced ANN model are reasonably close to these denser areas. Fig. 5b shows a similar plot for compositions having only sodium, boron, and oxygen. In particular, the well-known boron anomaly [3,49] was successfully captured and described by the network. In this plot, we also compare our prediction with those from a topological model proposed by Mauro and co-authors [50]. The agreement between them is reasonably good up to 20% of atomic fraction of sodium.

Another assessment of the predictive performance of the proposed network was carried out by randomly sampling compositions from the test set having different number of chemical elements. We randomly sampled 50 different compositions for glasses made from 3 to 13 different chemical elements. Fig. 6 summarizes the experimental results, i.e., the average predicted T_g and the respective standard deviation as a function of the number of chemical elements in the glass. No trend is visible in these results. The reasoning for this test was to check if the uncertainty increased with the complexity of the glass composition. Fortunately, this was not the case; the prediction error in T_g does *not* depend on the number of elements in the glass! This result is very significant for the development of new multicomponent glasses composed of a large number of chemical elements.

Another assessment of the model was how well it predicts the T_g of glass compositions made with two chemical elements that were not used to build our dataset (the single oxides mentioned in Section 2). We hypothesized that the trained network would be able to extrapolate the behavior of complex glasses to a reasonable value of T_g for single oxide glass-formers. Table 3 shows that this was indeed the case for most of these compositions, with the notable exception of silica (SiO_2), for which the predicted T_g was lower than expected. This result, however, is in agreement with what we observed with glasses having T_g values above 1250 K: the network tends to underestimate the glass transition temperature in these cases.

As a final test, we randomly sampled 15 glass compositions published in a specialized journal on glasses (Journal of Non-Crystalline Solids) between 2016 and 2018, which were not part of the original dataset. With these compositions, we hoped to evaluate how the induced ANN behaves in actual and recent glass development applications. Table 4 shows the results from this test. These 15 compositions were manufactured from 28 chemical elements; each composition contains from 4 to 10 different species and shows a wide range of T_g, from 474 to 1118 K.

The results are appealing. The overall relative deviation, ϕ, is very reasonable and corroborates the use of our ANN for the main objective of this work: assisting the search and development of new glass compositions. The average level of uncertainty is low and much more preferable than the classical “cook and look” approach.

The induced ANN model, as of now, is an empirical model. We

Table 4
Test of the T_g predicting capability of the induced network for 15 glass compositions that were not included in the original dataset. T_g in Kelvin and ϕ in %.

Composition	Reported T_g	Predicted T_g	ϕ	Ref.
$Ag_{14.3}V_{14.3}Te_{9.5}O_{61.9}$	474	475	−0.2	[51]
$Pb_{9.7}Ti_{1.7}Te_{16.9}B_{10}O_{61.7}$	563[a]	613	−8.9	[52]
$Mo_{11.4}W_{2.9}Te_{14.3}O_{71.4}$	589	626	−6.2	[53]
$Bi_{11.1}Mo_{10.3}Te_{10.3}O_{68.3}$	610	619	−1.4	[54]
$Cu_{7.3}Na_{17.1}P_{17.1}O_{58.5}$	623	612	1.8	[55]
$Bi_{14.3}Ti_{2.4}Te_{7.1}B_{14.3}O_{61.9}$	683	695	−1.7	[56]
$Ga_{9.7}Zn_{4.9}P_{18.9}O_{66.5}$	762	716	6.1	[57]
$Ba_{4.7}Cu_{0.3}K_{5.2}Zn_{3.1}B_{30.5}O_{56.2}$	775	757	2.4	[58]
$Cu_{12.8}Ti_{3.8}P_{17.9}O_{65.5}$	777	811	−4.3	[59]
$Ca_{2.9}Fe_{2.1}Mg_{4.2}Na_{0.4}P_{16.8}B_{8.4}O_{65.2}$	839	836	0.3	[60]
$Ca_{5.1}Fe_{0.1}K_{8.1}Mg_{3.9}Mn_{0.3}Na_{1.2}Al_{10.4}P_{1.4}Si_{19.8}O_{59.7}$	866	904	−4.4	[61]
$Ca_{15.8}Na_{2.6}Nb_{2.6}P_{15.8}O_{63.2}$	885	836	5.5	[62]
$Li_{1.9}Mg_{0.3}Zn_{0.4}Zr_{0.5}Ti_{0.5}Al_{10.2}P_{3.7}B_{2.7}Si_{15.8}O_{64}$	890	951	−6.9	[63]
$Ca_{1.9}Mg_{2.5}Sr_{1.9}Al_{6.9}B_5Si_{19.2}O_{62.6}$	940	984	−4.7	[64]
$La_{7.3}Al_{16.6}P_{6.7}Si_{5.7}O_{63.7}$	1118	1062	5.0	[65]

[a] Value obtained from Fig. 7 of ref. [52].

are studying methods to extract information from its internal parameters that could be used to build simpler and easily interpretable models, such as decision trees [66] or physical models (Occam's razor [67]).

To summarize our analysis, our induced ANN was able to reasonably predict the T_g for all combinations of the 45 elements used here, and will very likely help researchers to develop new glass compositions. Finally, the procedure developed in this study can be easily extended to build new networks to predict other properties of glasses.

5. Final remarks

This study comprehended a very substantial portion of all the available T_g data for oxide glasses. For such, we collected approximately 55,000 reported glass compositions and their respective values of glass transition temperatures for the network training and testing. These compositions contain from 3 to 21 different chemical elements, from a total of 45 that were studied. We then designed, implemented, trained, and evaluated an artificial neural network that can predict the T_g of oxide glasses within a very reasonable uncertainty, which is akin to the intrinsic scatter of the original dataset. The prediction uncertainty is larger for glasses with very high T_g (above 1250 K), fortunately, however, it does not depend on the number of elements in the glass composition. To the best of our knowledge, this is the first time a refined technique for ANN hyperparameter optimization has been used to predict the T_g of multicomponent glasses.

Besides, the research protocol developed here can be readily adapted to predict several other glass properties. This feature may lead to new insights and applications of ANNs helping to select and develop new glasses having noteworthy properties for advanced applications. This predictive algorithm will be soon available to the community as a web application.

Acknowledgments

We are thankful to the São Paulo Research Foundation (FAPESP) contract #2013/07793-6 (EDZ), #2013/07375-0 (ACPLFC), and #2017/12491-0 (DRC), as well as to the Nippon Sheet Glass overseas research grant for funding this research (DRC). DRC is grateful to Prof. Ivan N. da Silva for his lectures on artificial neural networks taught at the University of São Paulo (São Carlos campus) in 2017 and grateful for the educational series on neural networks by Welch Labs, which increased his interest in the topic. The educational discussions with Dr. Tiago Botari are also much appreciated.

Appendix A. Supplementary data

Supplementary data related to this article can be found at https://doi.org/10.1016/j.actamat.2018.08.022.

References

[1] E.D. Zanotto, J.C. Mauro, The glassy state of matter: its definition and ultimate fate, J. Non-Cryst. Solids 471 (2017) 490–495, https://doi.org/10.1016/j.jnoncrysol.2017.05.019.
[2] J.C. Mauro, E.D. Zanotto, Two centuries of glass research: historical trends, current status, and grand challenges for the future, Int. J. Appl. Glass Sci. 5 (2014) 313–327, https://doi.org/10.1111/ijag.12087.
[3] A.K. Varshneya, Fundamentals of Inorganic Glasses, second ed., Society of Glass Technology, Sheffield, 2006.
[4] E.D. Zanotto, F.A.B. Coutinho, How many non-crystalline solids can be made from all the elements of the periodic table? J. Non-Cryst. Solids 347 (2004) 285–288, https://doi.org/10.1016/j.jnoncrysol.2004.07.081.
[5] O.V. Mazurin, A.I. Priven, SciGlass - Glass Information System - Glass Database - Glass Properties, ITC, Inc., Newton, MA, USA, 2017. http://www.sciglass.info/.
[6] J.C. Mauro, Decoding the glass genome, Curr. Opin. Solid State Mater. Sci. 22 (2018) 58–64, https://doi.org/10.1016/j.cossms.2017.09.001.
[7] G. Pilania, C. Wang, X. Jiang, S. Rajasekaran, R. Ramprasad, Accelerating materials property predictions using machine learning, Sci. Rep. 3 (2013) 2810, https://doi.org/10.1038/srep02810.
[8] L.M. Ghiringhelli, J. Vybiral, S.V. Levchenko, C. Draxl, M. Scheffler, Big data of materials science: critical role of the descriptor, Phys. Rev. Lett. 114 (2015) 105503, https://doi.org/10.1103/PhysRevLett.114.105503.
[9] A. Mannodi-Kanakkithodi, G. Pilania, T.D. Huan, T. Lookman, R. Ramprasad, Machine learning strategy for accelerated design of polymer dielectrics, Sci. Rep. 6 (2016) 20952, https://doi.org/10.1038/srep20952.
[10] Y. LeCun, Y. Bengio, G. Hinton, Deep learning, Nature. 521 (2015) 436, https://doi.org/10.1038/nature14539.
[11] M.N. Dailey, G.W. Cottrell, C. Padgett, R. Adolphs, EMPATH: a neural network that categorizes facial expressions, J. Cognit. Neurosci. 14 (2002) 1158–1173.
[12] G. Hinton, L. Deng, D. Yu, G.E. Dahl, A. r Mohamed, N. Jaitly, A. Senior, V. Vanhoucke, P. Nguyen, T.N. Sainath, B. Kingsbury, Deep neural networks for acoustic modeling in speech recognition: the shared views of four research groups, IEEE Signal Process. Mag. 29 (2012) 82–97, https://doi.org/10.1109/MSP.2012.2205597.
[13] A. Mendyk, R. Jachowicz, Unified methodology of neural analysis in decision support systems built for pharmaceutical technology, Expert Syst. Appl. 32 (2007) 1124–1131.
[14] O. Deeb, Correlation ranking and stepwise regression procedures in principal components artificial neural networks modeling with application to predict toxic activity and human serum albumin binding affinity, Chemometr. Intell. Lab. Syst. 104 (2010) 181–194.
[15] D. Silver, A. Huang, C.J. Maddison, A. Guez, L. Sifre, G. van den Driessche, J. Schrittwieser, I. Antonoglou, V. Panneershelvam, M. Lanctot, S. Dieleman, D. Grewe, J. Nham, N. Kalchbrenner, I. Sutskever, T. Lillicrap, M. Leach, K. Kavukcuoglu, T. Graepel, D. Hassabis, Mastering the game of Go with deep neural networks and tree search, Nature 529 (2016) 484, https://doi.org/10.1038/nature16961.
[16] V. Mnih, K. Kavukcuoglu, D. Silver, A.A. Rusu, J. Veness, M.G. Bellemare, A. Graves, M. Riedmiller, A.K. Fidjeland, G. Ostrovski, S. Petersen, C. Beattie, A. Sadik, I. Antonoglou, H. King, D. Kumaran, D. Wierstra, S. Legg, D. Hassabis, Human-level control through deep reinforcement learning, Nature 518 (2015) 529, https://doi.org/10.1038/nature14236.
[17] Z. Zhang, K. Friedrich, Artificial neural networks applied to polymer composites: a review, Compos. Sci. Technol. 63 (2003) 2029–2044, https://doi.org/10.1016/S0266-3538(03)00106-4.
[18] N.Y. Steiner, D. Hissel, P. Moçotéguy, D. Candusso, Diagnosis of polymer electrolyte fuel cells failure modes (flooding & drying out) by neural networks modeling, Int. J. Hydrogen Energy 36 (2011) 3067–3075.
[19] X. Chen, L. Sztandera, H.M. Cartwright, A neural network approach to prediction of glass transition temperature of polymers, Int. J. Intell. Syst. 23 (2008) 22–32, https://doi.org/10.1002/int.20256.
[20] W. Liu, C. Cao, Artificial neural network prediction of glass transition temperature of polymers, Colloid Polym. Sci. 287 (2009) 811–818, https://doi.org/10.1007/s00396-009-2035-y.
[21] S.J. Joyce, D.J. Osguthorpe, J.A. Padgett, G.J. Price, Neural network prediction of glass-transition temperatures from monomer structure, J. Chem. Soc., Faraday Trans. 91 (1995) 2491, https://doi.org/10.1039/ft9959102491.
[22] A. Afantitis, G. Melagraki, K. Makridima, A. Alexandridis, H. Sarimveis, O. Iglessi-Markopoulou, Prediction of high weight polymers glass transition temperature using RBF neural networks, J. Mol. Struct. THEOCHEM 716 (2005) 193–198, https://doi.org/10.1016/j.theochem.2004.11.021.
[23] A. Cai, X. Xiong, Y. Liu, W. An, J. Tan, Y. Luo, Artificial neural network modeling for undercooled liquid region of glass forming alloys, Comput. Mater. Sci. 48 (2010) 109–114, https://doi.org/10.1016/j.commatsci.2009.12.012.
[24] A.H. Cai, Y. Liu, W.K. An, G.J. Zhou, Y. Luo, T.L. Li, X.S. Li, X.F. Tan, Prediction of critical cooling rate for glass forming alloys by artificial neural network, Mater. Des. 52 (2013) 671–676, https://doi.org/10.1016/j.matdes.2013.06.012, 1980-2015.
[25] F. Ren, L. Ward, T. Williams, K.J. Laws, C. Wolverton, J. Hattrick-Simpers, A. Mehta, Accelerated discovery of metallic glasses through iteration of machine learning and high-throughput experiments, Science Advances 4 (2018), https://doi.org/10.1126/sciadv.aaq1566 eaaq1566.
[26] A. Ziletti, D. Kumar, M. Scheffler, L.M. Ghiringhelli, The Face of Crystals: Insightful Classification Using Deep Learning, 2017. ArXiv:1709.02298 [Cond-Mat], http://arxiv.org/abs/1709.02298. (Accessed 13 January 2018).
[27] D.S. Brauer, C. Rüssel, J. Kraft, Solubility of glasses in the system P_2O_5–CaO–MgO–Na_2O–TiO_2: experimental and modeling using artificial neural networks, J. Non-Cryst. Solids 353 (2007) 263–270, https://doi.org/10.1016/j.jnoncrysol.2006.12.005.
[28] N.M. Anoop Krishnan, S. Mangalathu, M.M. Smedskjaer, A. Tandia, H. Burton, M. Bauchy, Predicting the dissolution kinetics of silicate glasses using machine learning, J. Non-Cryst. Solids 487 (2018) 37–45, https://doi.org/10.1016/j.jnoncrysol.2018.02.023.
[29] W. Sha, K.L. Edwards, The use of artificial neural networks in materials science based research, Mater. Des. 28 (2007) 1747–1752, https://doi.org/10.1016/j.matdes.2007.02.009.
[30] C. Dreyfus, G. Dreyfus, A machine learning approach to the estimation of the liquidus temperature of glass-forming oxide blends, J. Non-Cryst. Solids 318 (2003) 63–78, https://doi.org/10.1016/S0022-3093(02)01859-8.
[31] J.C. Mauro, A. Tandia, K.D. Vargheese, Y.Z. Mauro, M.M. Smedskjaer,

Accelerating the design of functional glasses through modeling, Chem. Mater. 28 (2016) 4267–4277, https://doi.org/10.1021/acs.chemmater.6b01054.
[32] T.E. Oliphant, Python for scientific computing, Comput. Sci. Eng. 9 (2007) 10–20, https://doi.org/10.1109/MCSE.2007.58.
[33] Python Data Analysis Library — Pandas, 2014. http://pandas.pydata.org/. (Accessed 4 April 2014).
[34] F. Chollet, Keras, (2015). https://github.com/fchollet/keras (accessed January 13, 2017).
[35] J.D. Hunter, Matplotlib: a 2D graphics environment, Comput. Sci. Eng. 9 (2007) 90–95, https://doi.org/10.1109/MCSE.2007.55.
[36] M. Abadi, P. Barham, J. Chen, Z. Chen, A. Davis, J. Dean, M. Devin, S. Ghemawat, G. Irving, M. Isard, TensorFlow: a system for large-scale machine learning, OSDI (2016) 265–283.
[37] S.C. Larson, The shrinkage of the coefficient of multiple correlation, J. Educ. Psychol. 22 (1931) 45.
[38] F. Mosteller, J.W. Tukey, Data analysis, including statistics, Handbook of Social Psychology 2 (1968) 80–203.
[39] M. Stone, Cross-validatory choice and assessment of statistical predictions, J. Roy. Stat. Soc. B (1974) 111–147.
[40] I.N. da Silva, D. Hernane Spatti, R. Andrade Flauzino, L.H.B. Liboni, S.F. dos Reis Alves, Artificial Neural Networks: a Practical Course, Springer International Publishing, Cham, 2017, https://doi.org/10.1007/978-3-319-43162-8.
[41] H.B. Demuth, M.H. Beale, O. De Jess, M.T. Hagan, Neural Network Design, Martin Hagan, 2014.
[42] D.E. Rumelhart, J.L. McClelland, P.R. Group, Parallel Distributed Processing, MIT press, Cambridge, MA, USA, 1987.
[43] F. Rosenblatt, Principles of Neurodynamics. Perceptrons and the Theory of Brain Mechanisms, Cornell Aeronautical Lab Inc Buffalo NY, Buffalo, 1961. http://www.dtic.mil/docs/citations/AD0256582.
[44] N. Srivastava, G. Hinton, A. Krizhevsky, I. Sutskever, R. Salakhutdinov, Dropout: a simple way to prevent neural networks from overfitting, J. Mach. Learn. Res. 15 (2014) 1929–1958.
[45] P.J. Huber, Robust estimation of a location parameter, Ann. Math. Stat. 35 (1964) 73–101.
[46] D.P. Kingma, J. Ba, Adam: a Method for Stochastic Optimization, 2014. ArXiv: 1412.6980 [Cs], http://arxiv.org/abs/1412.6980. (Accessed 9 January 2018).
[47] J.S. Bergstra, R. Bardenet, Y. Bengio, B. Kégl, Algorithms for hyper-parameter optimization, in: Advances in Neural Information Processing Systems, 2011, pp. 2546–2554.
[48] J. Bergstra, D. Yamins, D. Cox, Making a science of model search: hyper-parameter optimization in hundreds of dimensions for vision architectures, in: International Conference on Machine Learning, 2013, pp. 115–123.
[49] J. Krogh-Moe, On the structure of boron oxide and alkali borate glasses, Phys. Chem. Glasses 1 (1960) 26.
[50] J.C. Mauro, P.K. Gupta, R.J. Loucks, Composition dependence of glass transition temperature and fragility. II. A topological model of alkali borate liquids, J. Chem. Phys. 130 (2009) 234503, https://doi.org/10.1063/1.3152432.
[51] D. Souri, Physical and thermal characterization and glass stability criteria of amorphous silver-vanadate-tellurate system at different heating rates: inducing critical Ag_2O/V_2O_5 ratio, J. Non-Cryst. Solids 475 (2017) 136–143, https://doi.org/10.1016/j.jnoncrysol.2017.09.008.
[52] Q. Chen, M. Zhang, H. Wang, Q. Wang, Q. Ma, J. Li, Structures and magneto optical property of diamagnetic TiO_2-TeO_2-PbO-B_2O_3 glass, J. Non-Cryst. Solids 468 (2017) 58–66, https://doi.org/10.1016/j.jnoncrysol.2017.04.020.
[53] N. Gedikoğlu, A.E. Ersundu, S. Aydin, M. Çelikbilek Ersundu, Crystallization behavior of WO_3-MoO_3-TeO_2 glasses, J. Non-Cryst. Solids (2017), https://doi.org/10.1016/j.jnoncrysol.2017.11.030.
[54] O.A. Zamyatin, A.D. Plekhovich, E.V. Zamyatina, A.A. Sibirkin, Glass-forming region and physical properties of the glasses in the TeO_2–MoO_3–Bi_2O_3 system, J. Non-Cryst. Solids 452 (2016) 130–135, https://doi.org/10.1016/j.jnoncrysol.2016.08.027.
[55] R.A. Mouss, S. Krimi, B. Glorieux, I. Khattech, M. Couzi, T. Cardinal, A. El Jazouli, Structural characterization and calorimetric dissolution behavior of Na2O-CuOP2O5 glasses, J. Non-Cryst. Solids 452 (2016) 144–152, https://doi.org/10.1016/j.jnoncrysol.2016.08.029.
[56] N. Gupta, A. Kaur, A. Khanna, F. Gonzàlez, C. Pesquera, R. Iordanova, B. Chen, Structure-property correlations in TiO_2-Bi_2O_3-B_2O_3-TeO_2 glasses, J. Non-Cryst. Solids 470 (2017) 168–177, https://doi.org/10.1016/j.jnoncrysol.2017.05.021.
[57] J. Holubová, Z. Černošek, E. Černošková, Structural investigation and physical properties ofGa_2O_3–ZnO–P_2O_5 glasses, J. Non-Cryst. Solids 454 (2016) 31–38, https://doi.org/10.1016/j.jnoncrysol.2016.10.022.
[58] G.N. Devde, G. Upender, V. Chandra Mouli, L.S. Ravangave, Structure, thermal and spectroscopic properties of Cu^{2+} ions doped $59B_2O_3$–$10K_2O$–(30-x) ZnO–BaO ($0 \leq x \leq 30$), J. Non-Cryst. Solids 432 (2016) 319–324, https://doi.org/10.1016/j.jnoncrysol.2015.10.022.
[59] L. Zhang, S. Liu, Structure and crystallization behavior of 50CuO-xTiO_2-(50-x) P_2O_5 glasses, J. Non-Cryst. Solids 473 (2017) 108–113, https://doi.org/10.1016/j.jnoncrysol.2017.08.003.
[60] C. Zhu, I. Ahmed, A. Parsons, K.Z. Hossain, C. Rudd, J. Liu, X. Liu, Structural, thermal, in vitro degradation and cytocompatibility properties of P_2O_5-B_2O_3-CaO-MgO-Na_2O-Fe_2O_3 glasses, J. Non-Cryst. Solids 457 (2017) 77–85, https://doi.org/10.1016/j.jnoncrysol.2016.11.001.
[61] O. Gulbiten, J.C. Mauro, X. Guo, O.N. Boratav, Viscous flow of medieval cathedral glass, J. Am. Ceram. Soc. 101 (2018) 5–11, https://doi.org/10.1111/jace.15092.
[62] H. Maeda, S. Lee, T. Miyajima, A. Obata, K. Ueda, T. Narushima, T. Kasuga, Structure and physicochemical properties of CaO–P_2O_5–Nb_2O_5–Na_2O, J. Non-Cryst. Solids 432 (2016) 60–64, https://doi.org/10.1016/j.jnoncrysol.2015.06.003.
[63] R. Mishra, M. Goswami, A. Dixit, M. Krishnan, Study on thermophysical properties and phase evolution in Nd doped Li_2O-Al_2O_3-SiO_2 glass nucleated by multiple nucleating agents, J. Non-Cryst. Solids 447 (2016) 66–73, https://doi.org/10.1016/j.jnoncrysol.2016.05.035.
[64] J. Xie, H. Tang, J. Wang, M. Wu, J. Han, C. Liu, Network connectivity and properties of non-alkali aluminoborosilicate glasses, J. Non-Cryst. Solids 481 (2018) 403–408, https://doi.org/10.1016/j.jnoncrysol.2017.11.023.
[65] S. Guo-Malloy, P.F. McMillan, W.T. Petuskey, Glass formation and characterization in the 3Al2O3·2SiO2-LaPO4 system, J. Non-Cryst. Solids 451 (2016) 77–83, https://doi.org/10.1016/j.jnoncrysol.2016.03.017.
[66] M.W. Craven, Extracting Comprehensible Models from Trained Neural Networks, University of Wisconsin, Madison, 1996.
[67] P. Domingos, The role of Occam's razor in knowledge discovery, Data Min. Knowl. Discov. 3 (1999) 409–425.