Prediction of reduced glass transition temperature of metallic alloys based on a neural network

Prediction of reduced glass transition temperature of metallic alloys based on a neural network

Dmitrii Viatkin[1,2], Maxim Zakharov[1] and Dmitrii Zhuro[1,2]

[1]Northern (Arctic) Federal University, Arkhangelsk, Russia
[2]University of Deusto, Bilbao, Spain

E-mail: d.vyatkin@narfu.ru, m.zakharov@narfu.ru, d.zhuro@narfu.ruhere

Abstract. The reduced glass transition temperature Trg is an important glass forming ability parameter. Trg describes the glass formation in materials and the behaviour of materials at the transition between solid and liquid states and is an important parameter for materials analysis, development, and production process. This article describes the process and results of research on the development of a system for prediction of the reduced glass transition temperature Trg of metallic alloys based on recurrent neural network algorithms. The developed system can predict the reduced glass transition temperature Trg of metallic alloys based on the analysis of its chemical formula with high accuracy. The accuracy was evaluated using the 3 metrics: MSE, RMSE, MAE. Obtained values are: MSE value is 0.000678, RMSE value is 0.0260, MAE value is 0.01835.

1. Introduction

Glass forming ability GFA describes glass formation in materials. GFA can be represented by many parameters [1]. The reduced glass transition temperature Trg is one of widely used GFA parameters and describes the behaviour of materials at the transition between solid and liquid states and is an important parameter for materials analysis, development, and production process. The reduced glass transition temperature Trg based on the glass transition temperature Tg and the liquidus temperature Tl. Trg calculated by formula 1.

$$Trg = Tg \,/\, Tl \tag{1}$$

The reduced glass transition temperature Trg depends on the composition of the material. Thus, an analysis of the chemical composition of the material can be used to calculate a given Trg.

This research is based on an open dataset of reduced glass transition temperature Trg for metallic alloys [2] and explores the ability to predict alloy reduced glass transition temperature Trg based on alloy chemical formula using recurrent neural networks. Creating a model to predict Trg for an unspecified alloy would allow these values to be used to estimate the GFA of the material directly or as inputs to another model to then predict the GFA of the material with higher accuracystyle).

2. Materials and methods

2.1. Data analysing and processing

This research based on open dataset of reduced glass transition temperature Trg for metallic alloys. The dataset has 585 samples with 23 columns of parameters.

The metallic glass dataset gives two columns with information about the material composition. The first column is the overall composition and describe alloy chemistry formula. This is input parameter for neural network model.

The second column is the highest chemistry formula element. This column just duplicates information from does not use in neural network model.

The reduced glass transition temperature Trg is in third column and it is target output parameter for neural network model. It is used as a rough predictor for GFA.

The columns from four to the 23 are the MAGPIE features that have been generated from the material composition column. These parameters give values such as properties averaged over the material composition as well as features that are only for the majority element in each alloy [3]. These parameters are used in neural network training process but are not used in final model.

Values of output Trg parameter are in the range 0.223 and 0.688. The largest number of alloys have a Trg value close to 0.6.

This dataset is divided to train, validation, and test datasets. Train dataset has 80% of full dataset, 468 samples. Test dataset has 10% of full dataset, 58 samples. Validation dataset has 10% of full dataset, 59 samples.

Source dataset has small number of samples and random division can have statistical anomalies in data distribution. Use of Kolmogorov–Smirnov test [4] help avoid problems with data distribution. Data was randomly divided and compared with source data distribution to obtain minimum value of Kolmogorov criteria in Kolmogorov–Smirnov test. The distribution with the minimum value of Kolmogorov criteria and most similar to the original distribution is shown in figure 1: (a) Data distribution in source dataset; (b) Data distribution in train dataset; (c) Data distribution in validation dataset; (d) Data distribution in test dataset.

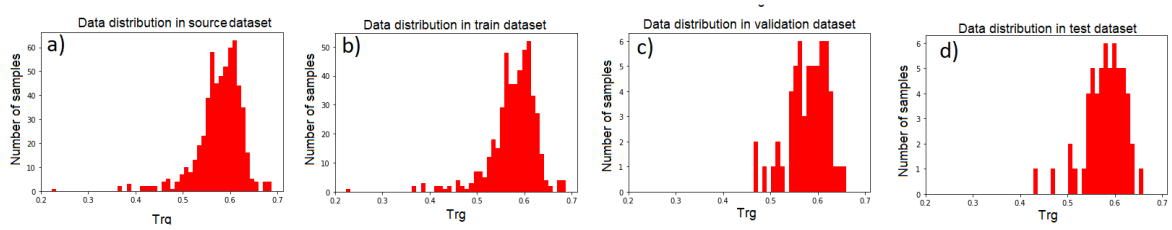

Figure 1. Reduced glass transition temperature distribution in source, train, test, and validation datasets.

The formulas in the dataset are represented by lines containing formula descriptions in the classical chemical format like Cu47Ti33Zr11Ni2Sn6Si. Every formula split in two list. First list has formula elements. Second list has formula chemistry coefficients.

Every unique element from list 1 is analysed by tokenizer [5] and obtain unique integer label.

Chemistry coefficients in formula can have significantly different values. Neural network will not give good results for unprocessed chemistry coefficients values. These parameters are standardized [6] to avoid variation in values.

A chemical formula can have up to 9 elements in its composition. The neural network cannot analyse data with dynamic length. Every formula with a small number of elements is augmented with up to 9 by zeros at the beginning of the formula in both lists.

2.2. Neural network development and training

A neural network based on recurrent LSTM layers [7] is developed to analyse the data.

The neural network has 2 input layers, which receive data on the chemical composition of the alloy.

The 1st input layer receives data about the chemical elements of the alloy, tokenized by integer labels. They are analysed by the embedding layer [8] and each chemical element label is described by a tensor of 64 elements. The output of the embedding layer creates a two-dimensional tensors array of 9x64.

The 2nd input layer receives standardized coefficient values of chemical elements of the formula. It is merged with the two-dimensional tensors array of 9x64 by concatenate layer. The resulting 9x65 array is analysed by two successive LSTM recurrent layers. The output of the recurrence layers is a 9x128 tensors data array.

This 9x128 array is converted into a tensor with a size of 1152 and subjected to dropout regularization [9] to eliminate the effect of overfitting [10]. The resulting tensor is then sent to the last layer to the prediction of the target features. Last output layer does not use any activation function.

The neural network uses Mean Squared Error MSE as loss function and Mean Absolute Error MAE as additional metric. The training uses the Adam optimizer, which has a gradually decreasing learning rate. Learning rate start at 0.01 and is reduced until 0.0001.

Training is implemented in two stages.

At the 1st stage the neural network is trained to predict every available 21 features of the chemical formula: reduced glass transition temperature Trg and 20 MAGPIE features.

At the 2nd stage the neural network is trained to predict only reduced glass transition temperature Trg.

Such training allows to create a neural network on the 1st step of training the most complete understanding of the structure of the analysed data, which allows to obtain the best results of the target metrics on the 2nd step.

Architecture of neural network at the 1st stage of the training is shown in figure 2.

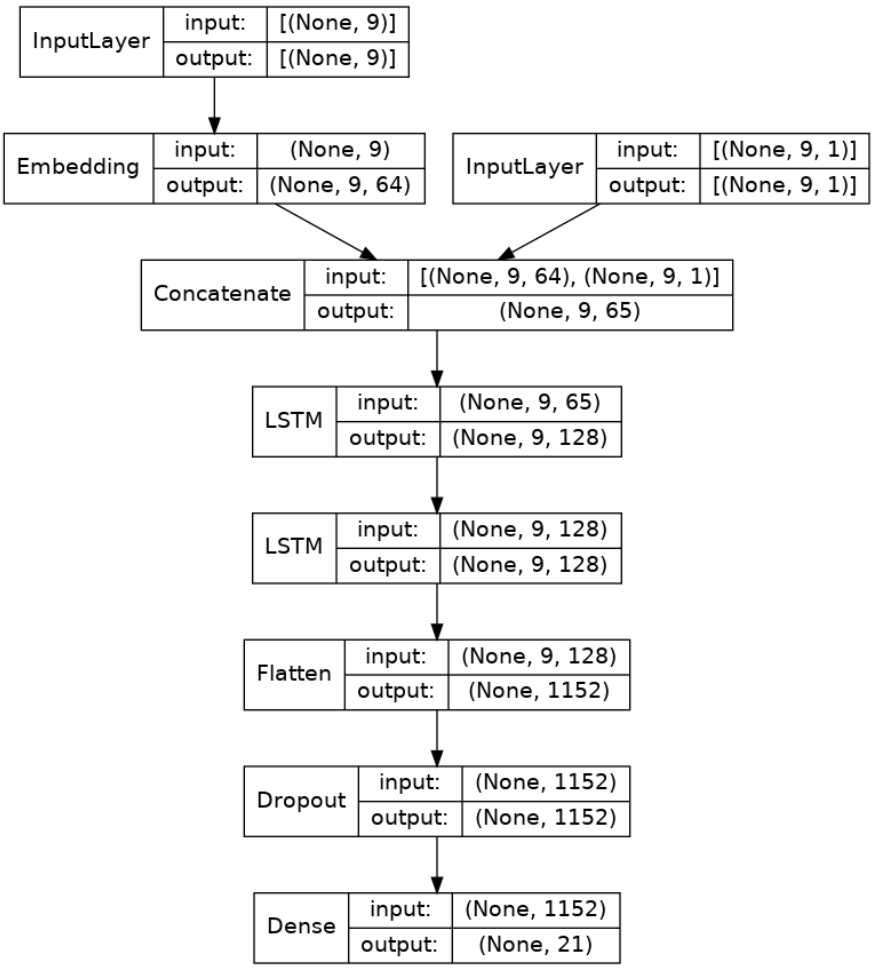

Figure 2. Architecture of neural network at the 1st stage of the training.

The training in Stage 1 is done in 1000 epochs. The process of the MSE loss function value changing is shown in figure 3.

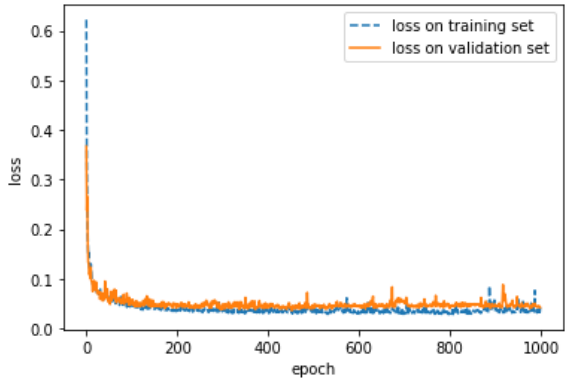

Figure 3. MSE loss function value changing in 1st stage of training process

The best obtained values of MSE loss function and MAE metrics for the 21 output features at 1st training stage are: MAE value is 0.0857, MSE value is 0.0348. This is sum of losses for every 21 output features. This model has 264.725 total parameters. This version is saved.

After 1st stage best version of neural network is restored. Output layer of restored model is removed. New output layer with one output value is added to the model at the 2nd stage. Training parameter for pretrained model is disabled for first step of 2nd stage because the new untrained layer can break pretrained neural network weights. In the first step of 2nd stage only last layer of model is trained. Model is trained 100 epochs with learning rate 0.01. At this step model has 241.665 total parameters and 1.153 trainable parameters.

Architecture of neural network at the 2nd stage of the training is shown in figure 4.

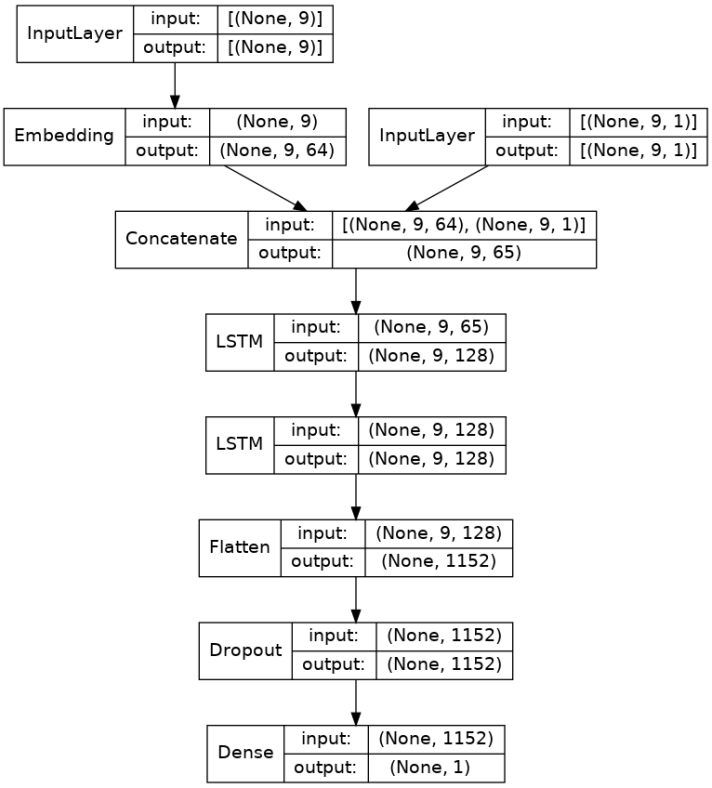

Figure 4. Architecture of neural network at the 2nd stage of the training.

The first step of training in Stage 2 is done in 100 epochs. The process of the MSE loss function value changing is shown in figure 5.

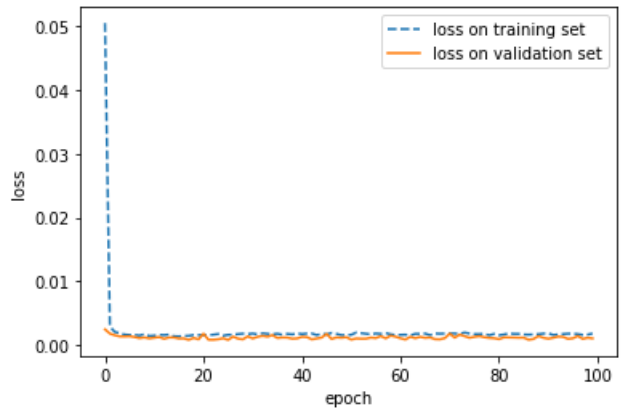

Figure 5. MSE loss function value changing in 1st step of 2nd stage of training process.

The best obtained values of MSE loss function and MAE metrics for the 1 output features at 1st step of 2nd stage of training process are: MAE value is 0.02098, MSE value is 0.000799. This version is saved.

After 1st step of 2nd stage best version of neural network is restored. Training parameter for full model is enabled. Model is trained 100 epoch with low learning rate value 0.0001. At this step model has 241.665 total trainable parameters.

The best obtained values of MSE loss function and MAE metrics for the 1 output features at 2st step of 2nd stage of training process is: MAE value is 0.01835, MSE value is 0.000678. This version is saved. This is final version of model.

3. Results

This dataset has a small size and its dividing into test, validation, and training samples randomly creates anomalies in the data distribution. Using a dataset with a large amount of data can significantly improve the performance of the neural network.

Also, the quality of neural network training can be improved by using additional data augmentation.

Another way to get the most stable model can be additional balancing of the data by clustering it. This approach, however, will reduce the quality of target metrics of the model, but will allow it to give the same quality on the target metrics on any distribution of data.

4. Discussion

The research developed a neural network model based on recurrent layers to analyse the chemical formulas of alloys and predict their reduced glass transition temperature Trg. Training was performed in 2 stages. In the 1st stage of training, the model obtained the fullest available information about the chemical compound. At stage 2, the model trained at stage 1 was upgraded to predict only the target parameter, reduced glass transition temperature Trg.

The model works with the processed data. The names of chemical elements were subjected to tokenization. The coefficients of chemical elements were subjected to standardization.

As a result of combining data processing algorithms and neural network training algorithms, the best metric values were obtained. The best obtained values of MSE loss function and MAE metrics are: MAE value is 0.01835, MSE value is 0.000678. Calculated value of Root Mean Squared Error RMSE is 0.0260.

5. Conclusion

This article shows the process of development of a neural network based on recurrent layers to analyse the chemical formulas of alloys for prediction its reduced glass transition temperature Trg. As a result, a final neural network model has metric values: MAE value equal 0.01835, MSE value equal 0.000678, RMSE value equal 0.0260.

The possibilities and features of using neural networks for the analysis of chemical formulas of dynamic size with a wide variation of their coefficients are demonstrated.

Since the analysed dataset has a small size, the paper gives examples of data anomalies elimination and data processing.

References

[1] Lu Z P, Li Y and Ng S C 2000 *Journal of Non-Crystalline Solids* **270(1-3)** 103–114
[2] Wu H, Mayeshiba T and Morgan D 2016 *Scientific Data* **3**
[3] Ward L, Agrawal A, Choudhary A and Wolverton C 2015 A *Npj Computational Materials* 1–7
[4] Daniel W Wayne 1990 *Applied Nonparametric Statistics* (Kolmogorov: Smirnov one-sample test) pp 319–30
[5] Ramasubramanian K and Singh A 2018 *Deep Learning Using Keras and Tensor Flow*
[6] Shanker M and Hu M 1996 *Omega* **24(4)** 385–97
[7] Liu G and Guo J 2019 *Neurocomputing*
[8] Vo N and Hays J 2019 Generalization in Metric Learning: Should the Embedding Layer Be Embedding Layer? *IEEE Winter Conference on Applications of Computer Vision (WACV)*
[9] Yarin G, Jiri H and Kendall A 2017 *Concrete Dropout*
[10] Hawkins D 2004 *Journal of Chemical Information and Computer Sciences* **44(1)** 1–12