



Deep Learning Approach for Prediction of Critical Temperature of Superconductor Materials Described by Chemical Formulas

Dmitry Viatkin^{1,2}, Begonya Garcia-Zapirain², Amaia Méndez-Zorrilla^{2*} and Maxim Zakharov¹

¹High School of Information Technologies and Automatic Systems, Northern Arctic Federal University, Arkhangelsk, Russia,

²eVIDA Research Group, Faculty of Engineering, University of Deusto, Bilbao, Spain

OPEN ACCESS

Edited by:

Zhenyu Li,

University of Science and Technology
of China, China

Reviewed by:

Yi Liu,

Shanghai University, China

Jun Jiang,

University of Science and Technology
of China, China

Youyong Li,

Soochow University, China

*Correspondence:

Amaia Méndez-Zorrilla
amaia.mendez@deusto.es

Specialty section:

This article was submitted to
Computational Materials Science,
a section of the journal
Frontiers in Materials

Received: 25 May 2021

Accepted: 29 July 2021

Published: 27 October 2021

Citation:

Viatkin D, Garcia-Zapirain B,
Méndez-Zorrilla A and Zakharov M
(2021) Deep Learning Approach for
Prediction of Critical Temperature of
Superconductor Materials Described
by Chemical Formulas.
Front. Mater. 8:714752.
doi: 10.3389/fmats.2021.714752

This paper proposes a novel neural network architecture and its ensembles to predict the critical superconductivity temperature of materials based on their chemical formula. The research describes the methods and processes of extracting data from the chemical formula and preparing these extracted data for use in neural network training using TensorFlow. In our approach, recurrent neural networks are used including long short-term memory layers and neural networks based on one-dimensional convolution layers for data analysis. The proposed model is an ensemble of pre-trained neural network architectures for the prediction of the critical temperature of superconductors based on their chemical formula. The architecture of seven pre-trained neural networks is based on the long short-term memory layers and convolution layers. In the final ensemble, six neural networks are used: one network based on LSTM and four based on convolutional neural networks, and one embedding ensemble of convolution neural networks. LSTM neural network and convolution neural network were trained in 300 epochs. Ensembles of models were trained in 20 epochs. All neural networks are trained in two stages. At both stages, the optimizer Adam was used. In the first stage, training was carried out by the function of losses Mean Absolute Error (MAE) with the value of optimizer learning rate equal to 0.001. In the second stage, the previously trained model was trained by the function of losses Mean Squared Error (MSE) with a learning rate equal to 0.0001. The final ensemble is trained with a learning rate equal to 0.00001. The final ensemble model has the following accuracy values: MAE is 4.068, MSE is 67.272, and the coefficient of determination (R²) is 0.923. The final model can predict the critical temperature for the chemistry formula with an accuracy of 4.068°.

Keywords: convolution neural network, superconductor, neural networks ensemble, critical temperature, LSTM neural network

INTRODUCTION

This paper presents a work deal with superconducting materials—materials that conduct current with zero resistance temperature equal to or below the critical temperature T_c (Hamidieh, 2018). Most of the known superconductors show the effect of superconductivity at extremely low temperatures below 100 K (Bonn, 2006; Flores-Livas et al., 2016; Nishiyama et al., 2017; Szeftel et al., 2018). However, despite the necessity of low temperatures for the appearance of the effect of superconductivity, superconductors are used in many areas.

The relevance of superconductors include societal challenges related to health and wellbeing. Superconductors are used in medicine mainly inside devices for CT scan and Magnetic Resonance Imaging, MRI, systems and in magnetometers for SQUID, Superconducting Quantum Interference Device (Alonso and Antaya, 2012). They are used for magnetoencephalography, magnetocardiography, and other processes for detection and mapping weak magnetic fields of the human body. The introduction of superconductors in such systems allows for conducting safe patient's research methods, to obtain a highly accurate three-dimensional picture of the state of the studied area of the human body.

Although the superconductivity effect can be used in many areas, the effect disappears when the temperature rises above critical (Suhl et al., 1959; Stewart, 2011; Flores-Livas et al., 2016). The need to maintain low temperatures to maintain the superconductivity effect is a complex and costly task. So, increasing the temperature of MRI (Sinanna et al., 2016) coils of the device above the critical temperature destroys the effect of superconductivity and triggers the process of destruction of the coils. This effect is used for emergency stops of MRI machines. The device is out of service and requires a long and expensive repair after an emergency stop. However, the use of high-temperature superconductors in the superconducting MRI coils would avoid such results of an emergency shutdown, making this procedure safe.

In addition, superconductors are often used in electrical, researchers, and other systems. Superconductors are used in superconducting fault current limiters, SCFCKs, for electrical current limitation (Noe and Steurer, 2007). Superconducting coils are used for generation and sustain high magnetic fields in the Large Hadron Collider at CERN (Hamidieh, 2018). Superconductors are used to create and support quantum states in quantum systems and quantum computers (Krinner et al., 2019). Superconductivity effect is used in the development of electric motors, unipolar machines, topological generators, rigid and flexible cables, switching and current-limiting devices, and magnetic separators, transport systems in the production of such coils for accelerators and for the creation of devices for measuring temperatures, costs, levels, and pressures.

At present, there are two main directions in the field of superconductivity application: in magnetic systems and in electric machines. Many types of research are underway to find new superconductors with high critical temperatures (Sleight et al., 1993; Flores-Livas et al., 2016; Si et al., 2016).

In the research (Hamidieh, 2018), models for predicting the properties of chemical compounds based on statistical algorithms XGBoost and multiple regression were developed. The algorithms were evaluated by RMSE and R2 metrics. The superconductor data comes from the Superconducting Material Database maintained by Japan's National Institute for Materials Science at http://supercon.nims.go.jp/index_en.html. After some data preprocessing, 21,263 superconductors are used. The developed model in the research based on the multiple regression method has RMSE value 17.6 K and R2 value 0.74. The model based on the XGBoost method has an RMSE value 9.5 K and R2 value 0.92.

In the research (Abdulkadir and Kemal, 2019), the model for predicting the properties of chemical compounds based on XGBoost statistical algorithm. In his article, he uses a statistical model to predict the superconducting critical temperature based on features extracted from the chemical formula of the superconductor. This article, like Kam Hamidiyeh's article, is based on the idea of analyzing the features of the chemical formula of the material and develops his ideas. The algorithms were evaluated by RMSE and R2 metrics. Developed XGBoost model has RMSE value 9.091 K and R2 value 0.928.

In the research (Li et al., 2019), the hybrid neural network that combines a convolution neural network and long short-term memory neural network is proposed to extract the characteristics of materials for critical temperature prediction of superconductors. The superconductor data comes from the Superconducting Material Database maintained by Japan's National Institute for Materials Science. The algorithms were evaluated by RMSE, MAE, and R2 metrics. The developed hybrid neural network has RMSE value 83.565 K and R2 value 0.899 and MAE value 5.023.

All these researches (Hamidieh, 2018; Abdulkadir and Kemal, 2019; Li et al., 2019) are using the same dataset.

Therefore, this paper presents a novel technique for an accurate prediction of these cases.

Then the aim of our research is to develop a model for using the chemical formula of material and then to predict the critical temperature of superconductivity for this material. Our research considers and describes an approach based on the use of various neural network architectures and their combinations for chemical formulas analysis. This research considers the use of neural networks whose structure is based on the use of LSTM and convolution layers.

In the final neural network ensemble, six networks are used: one network based on LSTM and four based on convolutional neural networks, and one embedding ensemble of convolution neural networks. LSTM neural network and convolution neural network were trained in 300 epochs. Ensembles of models were trained in 20 epochs. All neural networks are trained in two stages. At both stages, the optimizer Adam was used. In the first stage, training was carried out by the function of losses Mean Absolute Error (MAE) with the value of optimizer learning rate equal to 0.001. In the second stage, the previously trained model was trained by the function of losses Mean Squared Error (MSE) with a learning rate equal to 0.0001. The final ensemble is trained with a learning rate equal to 0.00001.

This article is organized as follows: after presenting the introduction in *Introduction*, materials and methods used in this paper are described in *Materials and Methods*, including the preparation and standardization of materials data for neural network-based machine learning systems. *Process of Development and Training of Neural Network Model* describes the proposed neural network algorithms designed to process sequential data, detailing the features and parameters used in the ensemble of pre-trained neural networks with different architectures. Results and Discussion are presented in *Results and Discussion*. At the end of

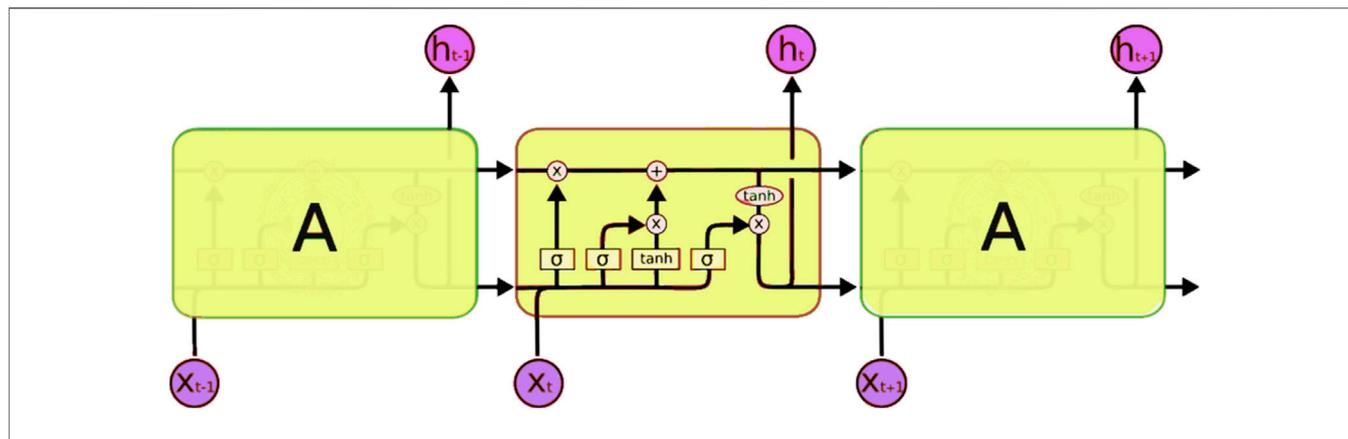


FIGURE 1 | Structure of LSTM neuron (Colah's blog, 2020).

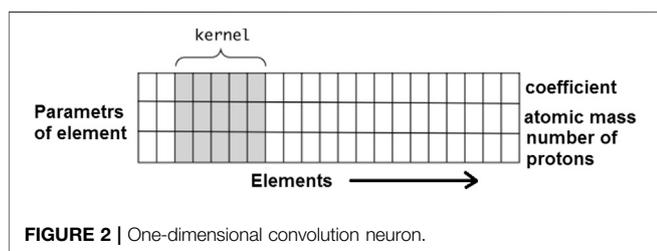


FIGURE 2 | One-dimensional convolution neuron.

the article, the conclusions of this research and future work are described.

MATERIALS AND METHODS

Materials

The research used a dataset from the research (Hamidieh, 2018). This dataset used in the research has 21,263 samples and is publicly available at the Center for Machine Learning and Intelligent Systems, Bren School of Information and Computer Science, University of California, Irvine (Center for Machine Learni, 2020). This set contains superconductor formulas and their parameters. The parameters of superconductor formulas were presented in a table with 21,263 rows by the number of superconductor formulas in the database. The formulas in the dataset contain from one to nine chemical elements. The dataset has 87 columns: columns 1 to 86 describe the chemical elements used in the formulas and column 87 contains the critical temperature T_c . The table has the coefficient value for each element of each formula. Values of element properties were obtained from the periodic table of elements in csv format (GitHub Gist, 2020)

The dataset used elements with atomic numbers up to 86, elements up to and including radon. For each element, 16 parameters were selected: atomic mass, number of neutrons, number of protons, period, atomic radius, electronegativity, first ionization, density, melting point, boiling point, number of shells, group, specific heat, is metal, is nonmetal, is metalloid.

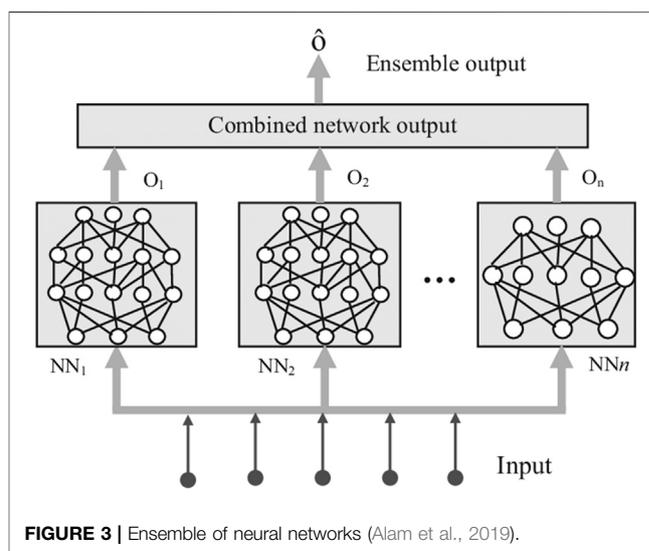


FIGURE 3 | Ensemble of neural networks (Alam et al., 2019).

The parameters “is metal”, “is nonmetal”, “is metalloid” were presented in 1 hot encoding format. These parameters have been selected because they provide a precise description for each of the 86 elements considered. After selecting these parameters, a table of 86 elements was standardized for these parameters.

Machine learning algorithms have low data accuracy with large differences in the size of input values. The parameters of chemical elements can vary by many orders of magnitude (Matthias, 1955; Montavon et al., 2013; Ramprasad et al., 2017; Meredig et al., 2018), so it was necessary to adapt the data for a quality learning process. For each of the 16 element parameters, an average value and an average deviation have been calculated based on the 86 elements used. Based on these data, a table of periodic elements has been standardized for better compatibility with machine learning algorithms.

Elements from the standardized table of chemical elements were inserted into superconductor formulas from the dataset for the elements, respectively. Coefficients of chemical elements in standardization formulas were not exposed. If the number of

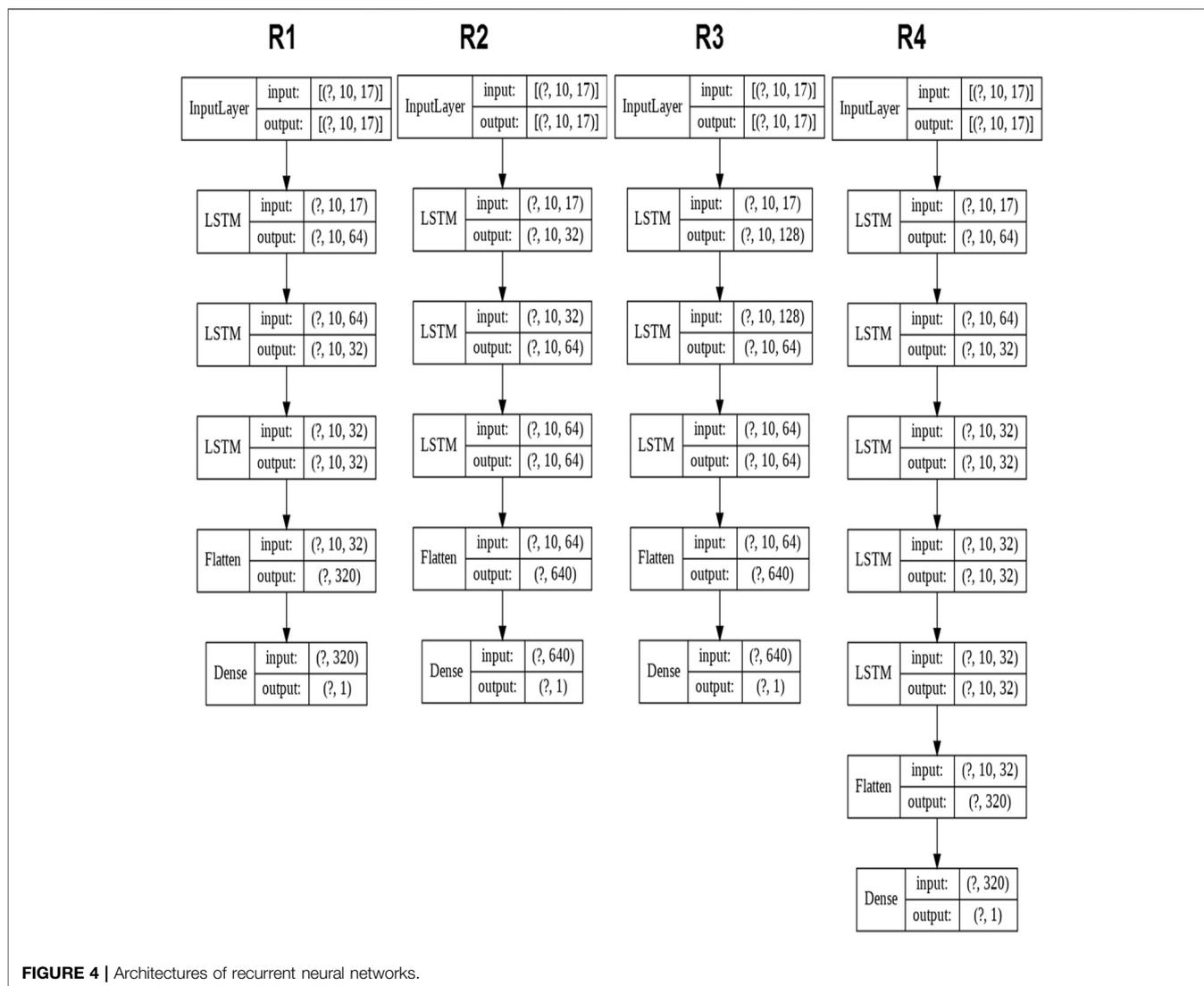


FIGURE 4 | Architectures of recurrent neural networks.

elements in a claim is less than 10, the claim is expanded to 10 and all parameters of all missing elements were set to zero (Montavon et al., 2013; Wei et al., 2016; Ramprasad et al., 2017). The result obtained was recorded in the processed dataset.

Two variants of chemical elements arranged in the processed dataset were considered: sorting by their indication in the chemical formula in the source dataset and sorting by their number in the periodic table of elements. The choice of sorting by number in the periodic table of elements was due to some differences in the order in which the chemical elements were arranged in the formulas in different fields of activity. In this variant of sorting the arrangement of elements in the formula view for the neural network, the specifics of recording the chemical formula do not affect the result of the neural network operation.

Methods

In this subsection, the details about the neural networks used in this research are described.

Long Short-Term Memory Based Neural Network

LSTM neural networks, Long Short-Term Memory neural networks, are a special kind of recurrent neural networks, capable of learning long-term dependencies. They were introduced by Sepp Hochreiter and Jurgen Schmidhuber (Hochreiter and Schmidhuber, 1997). LSTM networks are explicitly designed to avoid the long-term dependency problem. LSTM networks consist of LSTM neurons. **Figure 1** shows the principle of work of the LSTM neuron of a neural network as a whole and the structure of an individual LSTM layer in particular.

LSTM predictions are always based on past network input experience. However, as input data increase in size, the importance of data entering the neural network decreases at the beginning compared to data entering the neural network later (Sherstinsky, 2020). Therefore, the data at the beginning of the sequence have minimal impact on the result of the LSTM layer of the neural network, and the latest data have the maximum

TABLE 1 | Results of training of recurrent neural networks.

Name	Dropout	Activation	MAE (degrees*K)	MSE (degrees*K2)
R1	0	None	4.8908	92.1871
R1	0	Relu	5.1268	100.8792
R1	0.05	none	4.8328	93.8598
R1	0.05	Relu	5.0758	97.2897
R1	0.1	none	4.9218	94.8925
R1	0.1	Relu	5.1089	96.2879
R2	0	none	4.9705	100.8791
R2	0	Relu	5.1798	110.8971
R2	0.05	none	4.9518	99.8791
R2	0.05	Relu	5.1209	107.9871
R2	0.1	None	5.0129	97.5872
R2	0.1	Relu	5.1408	100.1879
R3	0	None	4.9287	96.7898
R3	0	Relu	5.1287	110.8987
R3	0.05	None	4.9019	95.1982
R3	0.05	Relu	5.0791	105.2847
R3	0.1	None	4.9271	95.0879
R3	0.1	Relu	5.0971	101.2878
R4	0	None	4.9833	99.7972
R4	0	Relu	5.0917	100.1972
R4	0.05	None	4.9613	96.4879
R4	0.05	Relu	5.0796	104.1975
R4	0.1	None	4.9486	95.1479
R4	0.1	Relu	5.0579	109.1871

impact on the result of the LSTM layer. To reduce this effect, the data dimensionality is added to the values used to operate the neural network. In this research, the value 0 is used to supplement the data to the standard size used by the neural network. These fillers are added before the data is processed so that the addition of data to the standard size taken by the LSTM layer for processing will have minimal impact on the extraction of information from the data (Montavon et al., 2013; Owolabi et al., 2014; Stanev et al., 2018). If we add these fillers after the data being processed, the neural network result will be less accurate because the filler values will have an effect on the result.

One-Dimensional Convolution Neural Network

Deep convolution networks provide state-of-the-art classifications and regressions results over many high-dimensional problems (Mallat, 2016). In the convolution neural networks, a convolution operation is

used. The convolution operation is a mathematical operation for two functions, which results in a third function expressing a change in the shape of one function to another.

For different types of input data, there are different convolution variants defined by the parameters of the convolution kernel. For two-dimensional convolution, by the example of image analysis, the kernel is defined by resolution of image and depth, the number of color channels, of image. For one-dimensional convolution, the kernel parameters are the length of the input sequence, the number of elements in one sequence row, and depth, and the number of values per one sequence step.

A deep neural network consists of neuronal layers with set parameters. In one layer, all neurons have the same convolution parameters. An example of a neuron of a one-dimensional convolution neural network for the sequence analysis using the chemical formula is shown in **Figure 2**.

One-dimensional convolution layer is used for analyzing two-dimensional data sequences. This type of layer creates and uses a convolution kernel that is convolved with the layer input over a single spatial or temporal dimension to produce a tensor of outputs. In the case of the formula analysis in **Figure 2**, the neuron, depending on the size of the convolution core, analyzes the parameters of each group of elements, moving from left to right.

Neural Network Ensemble

Neural networks are flexible and scalable algorithms that can adapt to the data used in training. However, they are trained using a stochastic learning algorithm and adapt to the specifics of the learning data during training. So even the same architecture neural networks, trained on the same dataset, but started training with different parameters of weights, can find different variants of the optimal set of weights each time they are trained, which, in turn, leads to different forecasts (Hansen and Salamon, 1990).

To improve learning outcomes and reduce learning, a learning approach is used that is based on learning from the same data and then combines multiple neural networks with different architectures. This is called ensemble learning and not only reduces the variance of forecasts but can also lead to forecasts that are better than any single model. An example of an ensemble of neural networks is shown in **Figure 3**.

Accuracy Metrics of Neural Network

In the training process, different methods are used to assess the quality of training and neural network performance. Prediction of the critical superconductivity temperature value for a chemical formula is a regression problem. For the regression problem, metrics are used: mean absolute error, mean squared error, root mean square error, and coefficient of determination.

MAE is a measure of the error between paired observations expressing the same phenomenon. A special feature of MAE is its resistance to emissions in data. MAE calculated by the following equation:

$$MAE = \frac{\sum_{i=1}^n |x_i - y_i|}{n} \quad (1)$$

with actual value x_i , predicted value y_i , and number of values n .

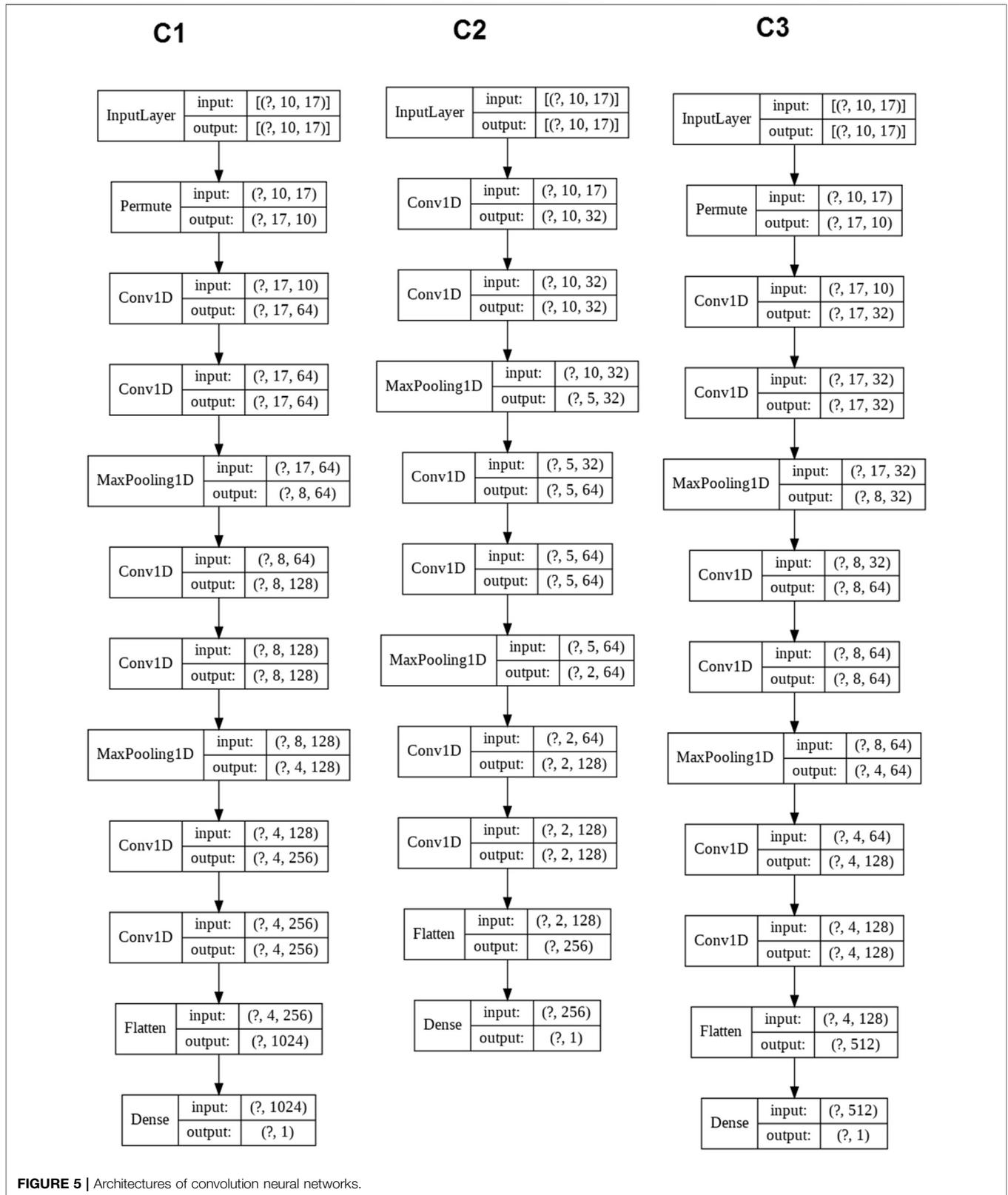


FIGURE 5 | Architectures of convolution neural networks.

The MSE is calculated as the mean square difference between the predicted and actual values. The result is always positive regardless of the sign of the predicted and actual values, and the ideal value is 0.0. A

square value means that larger errors result in more errors than smaller errors, which means that the model is penalized for larger errors. MSE calculated by the following equation:

TABLE 2 | Results of training of convolutional neural networks with different kernels.

Name	Kernel	MAE (degrees*K)	MSE (degrees*K2)
C1	2	4.7972	85.8867
C1	3	4.7287	85.4572
C1	4	4.7553	83.9879
C1	5	4.7843	89.4839
C2	2	4.8138	88.3587
C2	3	4.7628	89.7028
C2	4	4.8575	86.9126
C2	5	4.8052	85.8867
C3	2	4.8975	86.0501
C3	3	4.7559	85.7088
C3	4	4.8641	85.3812
C3	5	4.8384	88.4810

$$MSE = \frac{1}{n} \sum_{i=1}^n (x_i - y_i)^2 \quad (2)$$

with actual value x_i , predicted value y_i , and number of values n .

RMSE is a measure of the difference between the values predicted by the model or evaluator and the observed values. It is calculated as the square root of MSE and big value of errors have a disproportionate impact on RMSE. RMSE is calculated by the following equation:

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (x_i - y_i)^2}{n}} \quad (3)$$

with actual value x_i , predicted value y_i , and number of values n .

The coefficient of determination, R^2 , is the proportion of dispersion in the dependent variable, which is predictable from the independent variable or variables. It provides an estimation of how well the observed outcomes are reproduced by the model based on the proportion of common variations in outcomes explained by the model. The coefficient of determination is calculated by the following equation:

$$R^2 = 1 - \frac{\sum_{i=1}^n (x_i - y_i)^2}{\sum_{i=1}^n (x_i - \bar{y})^2} \quad (4)$$

with actual value x_i , predicted value y_i , mean value of actual values \bar{y} , and number of values n .

PROCESS OF DEVELOPMENT AND TRAINING OF NEURAL NETWORK MODEL

The processed dataset was shuffled and randomly divided into three parts: 80% - training set, 10% - test set, and 10% - validation set. Each value within each subset was duplicated 5 times, after which the

TABLE 3 | Results of training of neural networks with cross validation.

Name	MAE (degrees*K)	MSE (degrees*K2)
R1	4.8009	79.1871
C1	4.6608	78.0079
C2	4.7025	76.4987
C3	4.6989	79.7912

subset was mixed. This operation was performed so that the training did not focus on the features of individual batches. Since formulas had a different number of elements in their composition, they were added to 10 elements by an element, all parameters of which and the coefficient in the formula have the value 0.

After pre-processing, the original set containing 21,263 formulas was divided into a training set of 17,010 unique formulas, a test set of 2,126 formulas, and a validation set of 2,127 formulas. Each formula in each subset was repeated 5 times, after which the subset was mixed. As a result, the training set had 85,050 formulas, the test set had 10,630 formulas, and the validation set had 10,635 formulas.

After dividing the dataset into subsets, they were trained in different neural network architectures. Since the formulas were a sequence of elements with predefined parameters, the focus was on neural network architectures based on LSTM and one-dimensional convolution layers. The input of the neural network was supplied a processed formula added to 10 elements, each of which has 17 parameters: the coefficient of the element in the formula and 16 parameters of the chemical element from the periodic table of elements.

Models were trained in two stages. At both stages, the optimizer Adam was used. In the first stage, training was carried out by the function of losses Mean Absolute Error, MAE with the value of learning rate equal to 0.001. In the second stage, the previously trained model was trained by the function of losses Mean Squared Error, MSE, with a low learning rate equal to 0.0001.

LSTM Based Neural Network Models

Various architectures based on LSTM layers were considered during the research. The architectures and their names, which gave the most accurate results, are shown in **Figure 4**. For these networks, various activation functions and dropout values have been investigated. The results of training for these architectures are shown in **Table 1**.

According to the analysis of these results, the best results were obtained by R1 neural network of three LSTM layers, without the function of the activation of neurons on the recurrent layers, with a value of dropout on each of the layers equal to 5%. The best results were obtained with the number of epochs equal to 300 and the size of the batch equal to 200.

One-Dimensional Convolution Based Neural Network Models

Also, in the process of research were considered the architecture of convolution networks based on one-dimensional convolution layers. L2 regularization with a value of 0.0001 was applied for each layer for reducing the effect of overfitting. Architectures and their indexes, which gave the most accurate results, are shown in **Figure 5**. For the

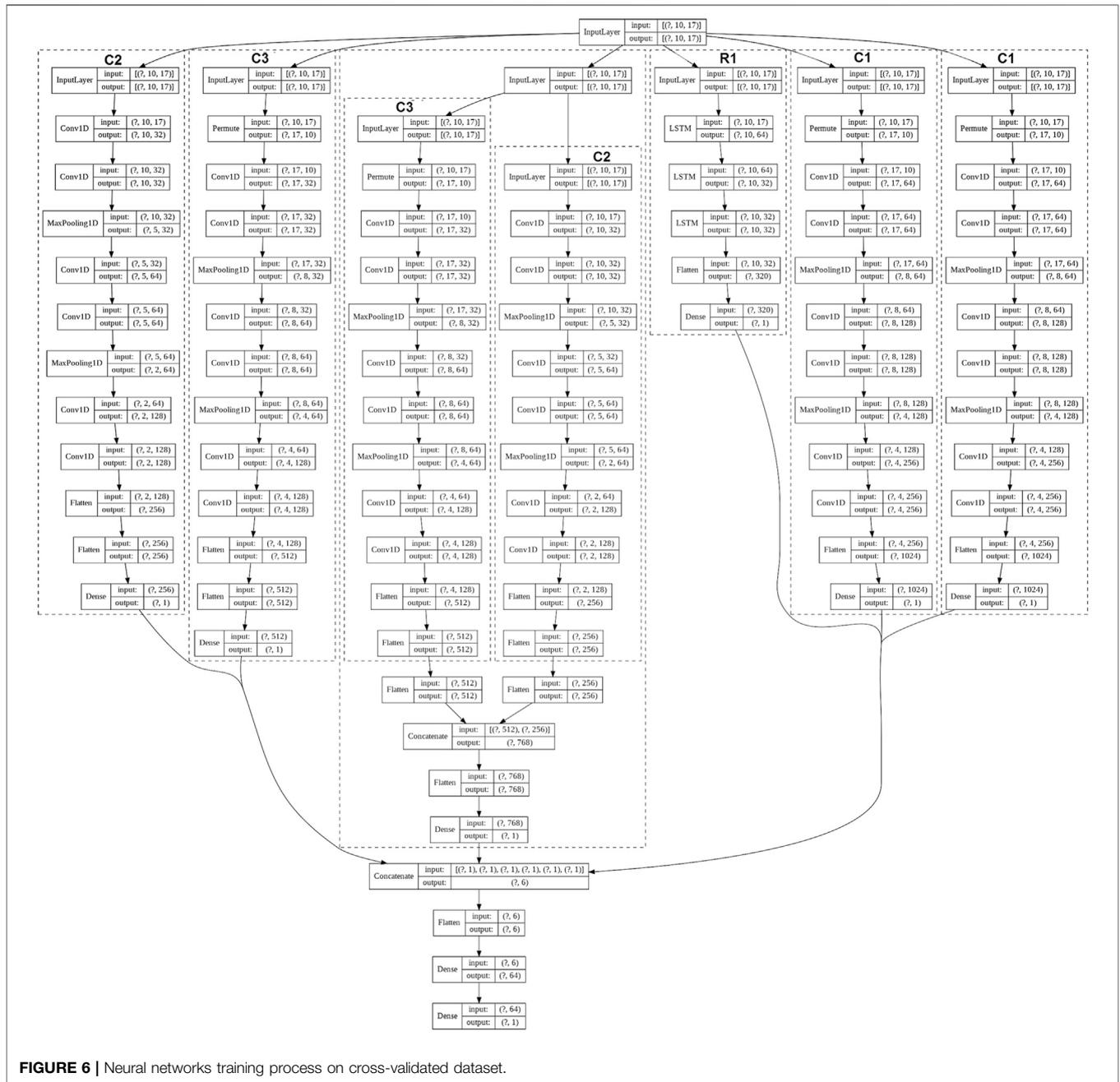


FIGURE 6 | Neural networks training process on cross-validated dataset.

convolution neural networks, training at different kernel sizes was provided. The size of the convolution kernel was changed from 2 to 5. The training results for these architectures are shown in Table 2.

The best result for convolutional neural networks was obtained with the kernel size equal to three. The best results were obtained with the number of epochs equal to 300 and the size of the batch equal to 200.

Cross-Validation Approach for Training Neural Network Models

For the variants of architectures, cross-validation training was used. The source dataset of 21,263 values was divided 10 times

into 90% train subset and 10% test subset. In addition, each value within each subset was duplicated 5 times, after which the subset was mixed. The cross-validation train subset had 95,685 formulas and the test dataset had 10,630 formulas. The best results of training with the use of cross-validation are presented in Table 3. All architectures of the neural network were trained on the best cross-validated dataset. All modules have two options: adapted for MAE and adapted for MSE. The training process of neural networks for MAE is shown in Figure 6.

The cross-validation variant of the dataset, which gave the best training results and the neural networks trained on this data, were saved and used to create an ensemble of neural networks.

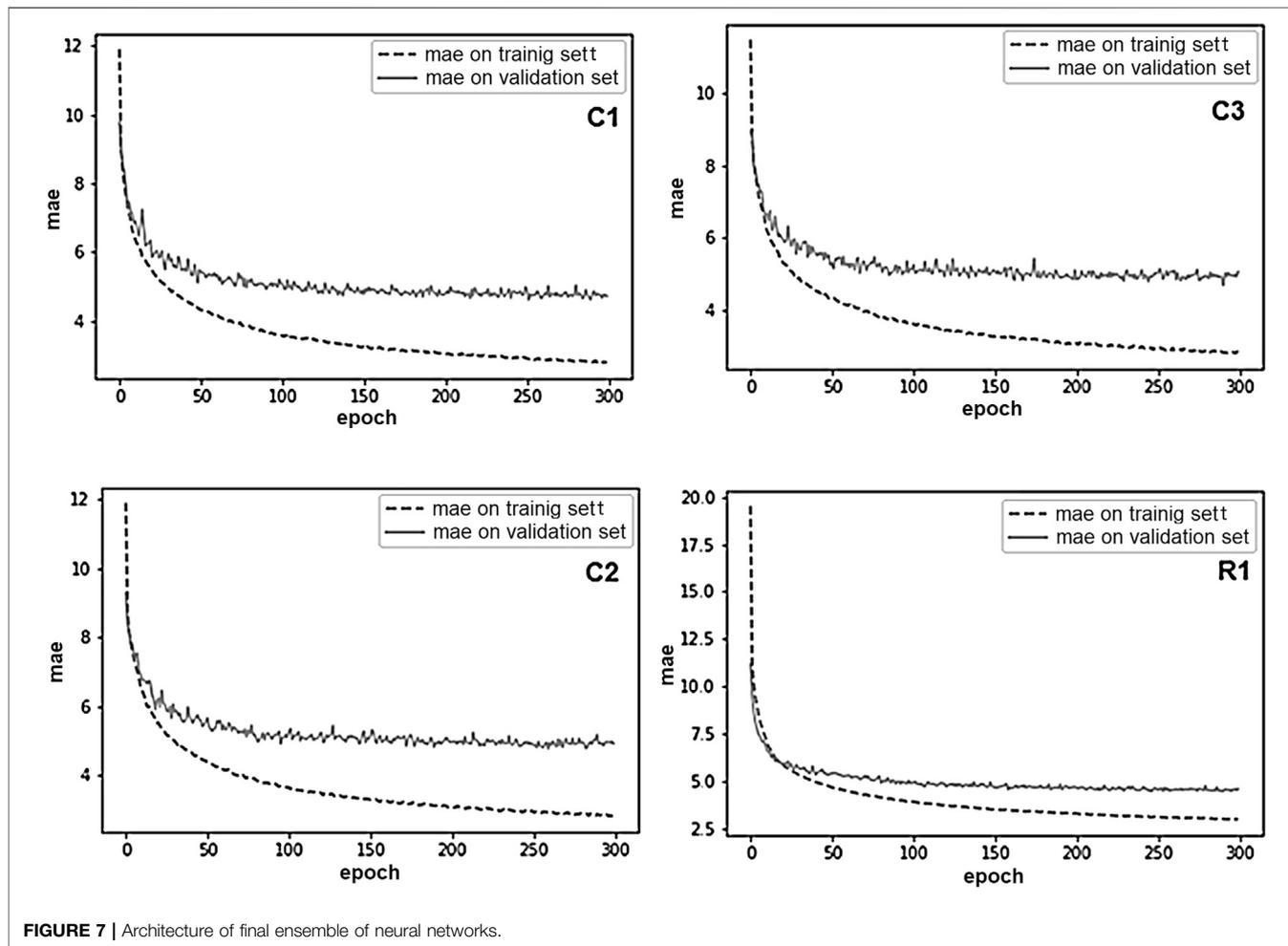


FIGURE 7 | Architecture of final ensemble of neural networks.

Ensemble of Neural Network Models

Since the C2 and C3 neural networks have similar architecture, they were adapted to minimize MSE loss and combined into an ensemble. After training, an ensemble of C2 and C3 neural networks was inserted into the final ensemble. Model C1 has been added in two variants: adapted to minimize MAE loss and adapted to minimize MSE loss. The final ensemble architecture from pre-trained neural networks is shown in **Figure 7**. In the ensemble, all outputs of all pre-trained models were combined and analyzed by a fully connected layer. In the process of training, weights of pre-trained models were frozen. Only the last fully connected layer of 64 neurons was trained. Output values of 6 pre-trained networks were combined and sent to the full layer of 64 neurons with “ReLU” activation function. The result of the full layer is used to estimate the critical superconductor temperature.

The final version of the ensemble was trained in two stages. In the first stage, training was carried out by the function of Mean Squared Error loss, MSE, with the value of learning rate equal to 0.000001. In the second stage, the previously trained model was trained by the function of

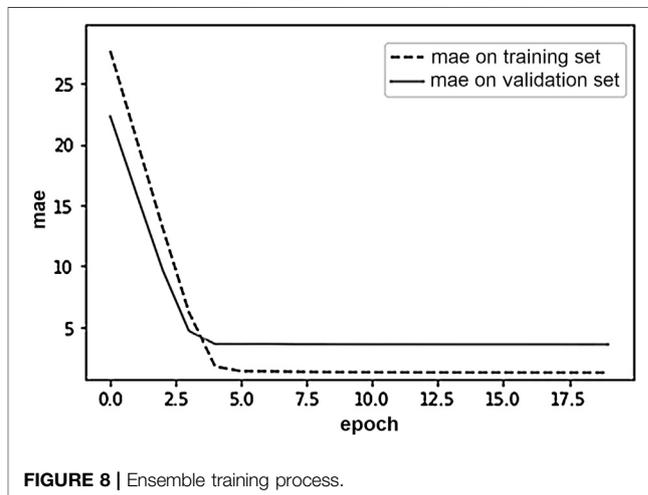


FIGURE 8 | Ensemble training process.

Mean Squared Error loss, MSE, with a low learning rate equal to 0.0000001. The low rate of training is discussed by the fact that the ensemble consists of already pre-trained models. The best results were obtained with the number of

TABLE 4 | Accuracy metrics of ensemble.

Neural network name	RMSE (degrees°K)	R2	MAE (degrees°K)
Ensemble of R1, C1, C2, C3 networks	8.202	0.923	4.068
R1	8.899	0.892	4.801
C1	8.832	0.897	4.661
C2	8.749	0.906	4.703
C3	8.932	0.890	4.699

TABLE 5 | Compilation different algorithms.

Algorithm	Author	Year	RMSE (degrees°K)	R2	MAE (degrees°K)
Multiple regression Hamidieh (2018)	Kam Hamidieh	2018	17.6	0.74	—
XGBoost Hamidieh (2018)	Kam Hamidieh	2018	9.5	0.92	—
XGBoost Abdulkadir and Kemal (2019)	Abdulkadir Karacı	2019	9.091	0.928	—
Hybrid neural network Li et al. (2019)	Shaobo Li	2020	9,141 ^a	0.899	5.023
Proposed model	Authors of this paper	2020	8.202	0.923	4.068

^aOriginal value was 83.565. In the article (Abdulkadir and Kemal, 2019), the MSE value was printed instead of the RMSE value.

epochs equal to 20 and the size of the batch equal to 200. The process of ensemble training for the minimization of MAE loss is shown in **Figure 8**.

The error of this ensemble after training is 4.068 for MAE loss and 67.272 for MSE loss. After minimization, MSE, loss coefficient of determination, R2, was calculated. The coefficient of determination, R2, is 0.923. Also, after minimizing MSE loss, root mean square error, RMSE, was calculated. RMSE is 8.202. The final ensemble of neural networks has 1,330,247 trainable parameters.

RESULTS AND DISCUSSION

In this research, the application of neural networks to the analysis of superconductor tempers was considered. The analysis was based on the properties of chemical elements and their coefficients in the chemical formula. As a result, neural networks with convolution and recurrence architecture were trained. The ensemble of neural networks was created as a combination of the best variants of architectures of pre-trained neural architectures.

The accuracy metrics of ensemble after training are shown in **Table 4**:

The final ensemble of neural networks has 1,330,247 trainable parameters.

The results of the prediction of this ensemble in comparison with previously developed algorithms are presented in **Table 5**.

The root mean square error of the developed ensemble is smaller than the minimum value of the RMSE error of previous algorithms. RMSE of the ensemble from this article was decreased by 0.889°.

The coefficient of determination is more than the value of previous algorithms. The mean absolute error is smaller than the

MAE of previous algorithms. R2 metric of the ensemble from this article was decreased by 0.005.

Mean absolute error of the developed ensemble is smaller than the minimum value of MAE error of previous algorithms. Mean absolute error of ensemble from this article was decreased by 0.937°.

Although the neural network ensemble has lost exactly the coefficient of determination, MAE and RMSE values have decreased. Comparison of MAE and RMSE value changes with the coefficient of determination value allows considering the decrease in the value by the coefficient of determination as insignificant in comparison with improved accuracy and decrease in MAE and RMSE values.

The developed neural network ensemble algorithm, as well as earlier algorithms presented in this article, is based on the same dataset. This dataset includes only the chemical formula of the superconductor and its critical temperature. However, many substances may change their internal structure depending on many factors. Adding to the model of atomic structure information of investigated material could considerably increase the quality of work of models intended for analysis of chemical compounds.

CONCLUSION

In this paper, an ensemble of neural networks was developed to predict the critical temperature of superconductors. Input data for this neural network model are only the chemical formula of the material. Sorting chemical elements in the formula by the element number in the periodic table of chemical elements allowed the neural network to concentrate more actively on real parameters of chemical elements rather than on features of their representation in the chemical formula.

As the chemical formula represents a sequence of parameters, recurrence and convolution algorithms of neural networks were used in this neural network model. Combined use of these algorithms helped to give high accuracy calculation of the target parameter, critical temperature for chemical formula. Our proposed method based on a neural network model can be useful for searching high-temperature superconductors.

Nowadays, superconductors are used in various fields. However, the necessity to provide low temperatures for the use of the effect of superconductivity makes these devices with the use of superconductors expensive and difficult to service. These devices are large with a relatively small size of the used superconductor. Many efforts and materials are used to simply maintain the temperature enough to support the superconductivity effect. Also, low temperatures are dangerous for humans, where the superconductivity effect is manifested.

Due to the increase of using superconductor material in different devices, which solves problems related to societal challenges, the proposed research line could be implemented in real time in the near future.

REFERENCES

- Abdulkadir, K., and Kemal, A. (2019). Predicting the Critical Temperature of Superconductors with XGBoost. *ICAIAME 2019*, 74–76. Conference paper.
- Alam, K. M. R., Siddique, N., and Adeli, H. (2019). A dynamic ensemble learning algorithm for neural networks. *Neural Comput. Applic* 32, 8675–8690. doi:10.1007/s00521-019-04359-7
- Alonso, J. R., and Antaya, T. A. (2012). Superconductivity in Medicine. *Rev. Accl. Sci. Tech.* 05, 227–263. doi:10.1142/s1793626812300095
- Bonn, D. A. (2006). Are high-temperature superconductors exotic? *Nat. Phys.* 2 (3), 159–168. doi:10.1038/nphys248
- Center for Machine Learning and Intelligent Systems (2020). *Bren School of Information and Computer Science*. Irvine: University of California. <https://archive.ics.uci.edu/ml/datasets/Superconductivity+Data>.
- Colah's blog (2020). *Understanding LSTM Networks*. <http://colah.github.io/posts/2015-08-Understanding-LSTMs/>.
- Flores-Livas, J. A., Sanna, A., and Gross, E. K. U. (2016). High temperature superconductivity in sulfur and selenium hydrides at high pressure. *Eur. Phys. J. B* 89 (3). doi:10.1140/epjb/e2016-70020-0
- GitHub Gist (2020). *Periodic Table of Elements*. <https://gist.github.com/GoodmanSciences/c2dd862cd38f21b0ad36b8f96b4bf1ee>.
- Hamidieh, K. (2018). A data-driven statistical model for predicting the critical temperature of a superconductor. *Comput. Mater. Sci.* 154, 346–354. doi:10.1016/j.commatsci.2018.07.052
- Hansen, L. K., and Salamon, P. (1990). Neural network ensembles. *IEEE Trans. Pattern Anal. Machine Intell.* 12 (10), 993–1001. doi:10.1109/34.58871
- Hochreiter, S., and Schmidhuber, J. (1997). Long Short-Term Memory. *Neural Comput.* 9 (8), 1735–1780. doi:10.1162/neco.1997.9.8.1735
- Krinner, S., Storz, S., Kurpiers, P., Magnard, P., Heinsoo, J., Keller, R., et al. (2019). Engineering cryogenic setups for 100-qubit scale superconducting circuit systems. *EPJ Quan. Technol.* 6 (1). doi:10.1140/epjqt/s40507-019-0072-0
- Li, S., Dan, Y., Li, X., Hu, T., Dong, R., Cao, Z., et al. (2019). Critical Temperature Prediction of Superconductors Based on Atomic Vectors and Deep Learning. *Symmetry* 12, 262. doi:10.3390/sym12020262
- Mallat, S. (2016). Understanding deep convolutional networks. *Phil. Trans. R. Soc. A.* 374 (2065), 20150203. doi:10.1098/rsta.2015.0203
- Matthias, B. T. (1955). Empirical Relation between Superconductivity and the Number of Valence Electrons per Atom. *Phys. Rev.* 97 (1), 74–76. doi:10.1103/physrev.97.74

DATA AVAILABILITY STATEMENT

The raw data supporting the conclusions of this article will be made available by the authors without undue reservation.

AUTHOR CONTRIBUTIONS

Funding acquisition: BG-Z and AM-Z; investigation: DV, BG-Z, and AM-Z; methodology: BG-Z; project administration: BG-Z and AM-Z; resources: BG-Z; software: DV; supervision: AM-Z; writing—original draft: DV; writing—review and editing: BG-Z and AM-Z.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fmats.2021.714752/full#supplementary-material>

- Meredig, B., Antono, E., Church, C., Hutchinson, M., Ling, J., Paradiso, S., et al. (2018). Can machine learning identify the next high-temperature superconductor? Examining extrapolation performance for materials discovery. *Mol. Syst. Des. Eng.* 3, 819–825. doi:10.1039/c8me00012c
- Montavon, G., Rupp, M., Gobre, V., Vazquez-Mayagoitia, A., Hansen, K., Tkatchenko, A., et al. (2013). Machine learning of molecular electronic properties in chemical compound space. *New J. Phys.* 15 (9), 095003. doi:10.1088/1367-2630/15/9/095003
- Nishiyama, S., Fujita, H., Hoshi, M., Miao, X., Terao, T., Yang, X., et al. (2017). Preparation and characterization of a new graphite superconductor: Ca_{0.5}Sr_{0.5}C₆. *Sci. Rep.* 7 (1). doi:10.1038/s41598-017-07763-y
- Noe, M., and Steurer, M. (2007). High-temperature superconductor fault current limiters: concepts, applications, and development status. *Supercond. Sci. Technol.* 20 (3), R15–R29. doi:10.1088/0953-2048/20/3/r01
- Owolabi, T. O., Akande, K. O., and Olatunji, S. O. (2014). Estimation of Superconducting Transition Temperature T_C for Superconductors of the Doped MgB₂ System from the Crystal Lattice Parameters Using Support Vector Regression. *J. Supercond. Nov. Magn.* 28 (1), 75–81. doi:10.1007/s10948-014-2891-7
- Ramprasad, R., Batra, R., Pilania, G., Mannodi-Kanakkithodi, A., and Kim, C. (2017). Machine learning in materials informatics: recent applications and prospects. *Npj Comput. Mater.* 3 (1). doi:10.1038/s41524-017-0056-5
- Sherstinsky, A. (2020). Fundamentals of Recurrent Neural Network (RNN) and Long Short-Term Memory (LSTM) network. *Physica D: Nonlinear Phenomena* 404, 132306. doi:10.1016/j.physd.2019.132306
- Si, Q., Yu, R., and Abrahams, E. (2016). High-temperature superconductivity in iron pnictides and chalcogenides. *Nat. Rev. Mater.* 1 (4). doi:10.1038/natrevmats.2016.17
- Sinanna, A., Belorgey, J., Bredy, P., Donati, A., Dubois, O., Guihard, Q., et al. (2016). High Reliability and Availability of the Iseult/Inumac MRI Magnet Facility. *IEEE Trans. Appl. Supercond.* 26 (3), 1–5. doi:10.1109/tasc.2016.2516914
- Sleight, A. W., Gillson, J. L., and Bierstedt, P. E. (1993). High-temperature superconductivity in the BaPb_{1-x}BixO₃ system. *Solid State. Commun.* 88 (11–12), 841–842. doi:10.1016/0038-1098(93)90253-j
- Stanev, V., Oses, C., Kusne, A. G., Rodriguez, E., Paglione, J., Curtarolo, S., et al. (2018). Machine learning modeling of superconducting critical temperature. *Npj Comput. Mater.* 4 (1). doi:10.1038/s41524-018-0085-8
- Stewart, G. R. (2011). Superconductivity in iron compounds. *Rev. Mod. Phys.* 83 (4), 1589–1652. doi:10.1103/revmodphys.83.1589

- Suhl, H., Matthias, B. T., and Walker, L. R. (1959). Bardeen-Cooper-Schrieffer Theory of Superconductivity in the Case of Overlapping Bands. *Phys. Rev. Lett.* 3 (12), 552–554. doi:10.1103/physrevlett.3.552
- Szefel, J., Sandeau, N., and Khater, A. (2018). Comparative study of the meissner and skin effects in superconductors. *Pier M* 69, 69–76. doi:10.2528/pierm18012805
- Wei, J. N., Duvenaud, D., and Aspuru-Guzik, A. (2016). Neural Networks for the Prediction of Organic Chemistry Reactions. *ACS Cent. Sci.* 2 (10), 725–732. doi:10.1021/acscentsci.6b00219

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors, and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2021 Viatkin, Garcia-Zapirain, Méndez-Zorrilla and Zakharov. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.