RESEARCH PAPER

Sm-Co ALLOYS COERCIVITY PREDICTION USING STACKING HETEROGENEOUS ENSEMBLE MODEL

Andrii Trostianchyn1*, Zoia Duriagina1,2, Ivan Izonin3, Roman Tkachenko1, Volodymyr Kulyk1, Olena Pavliuk4

1 Department of Materials Science and Engineering, Lviv Polytechnic National University, Lviv, S. Bandera str. 12, 79013, Ukraine
2 The John Paul II Catholic University of Lublin, Al. Racławickie 14, 20-950 Lublin, Poland
3 Department of Artificial Intelligence, Lviv Polytechnic National University, S. Bandera str. 12, Lviv, 79013, Ukraine
4 Department of Automated Control Systems, Lviv Polytechnic National University, S. Bandera str. 12, Lviv, 79013, Ukraine

*Corresponding author: andrii.m.trostianchyn@lpnu.ua, tel.: +380982149964, Department of Materials Science and Engineering, Lviv Polytechnic National University, S. Bandera str. 12, Lviv, 79013, Ukraine

Received: 15.09.2021
Accepted: 27.10.2021

ABSTRACT

The use of machine learning tools in modern materials science can significantly reduce the duration and cost of developing new materials and improving the properties of existing ones. This is especially true in studying expensive and strategic importance materials like alloys of rare earth metals, which are used to manufacture high-energy permanent magnets. At the same time, single machine learning algorithms do not always provide the accuracy required to solve a particular applied task. Therefore, the current paper aimed to develop an ensemble model for predicting the magnetic properties of Sm-Co system alloys with high accuracy. Based on literature data, we have collected the dataset of the relationship between phase composition, sample state, crystallographic orientation, microstructure, and magnetic properties. We have compared different machine learning algorithms. A stacking ensemble model was designed based on high-precision machine learning algorithms: Neural Networks, AdaBoost, Gradient Boosting, and Random Forest algorithm. The proposed ensemble scheme showed a significant increase in the accuracy for predicting the magnetic properties of Sm-Co alloys on the example of coercivity.

Keywords: Sm-Co alloys, magnetic properties, machine learning, ensemble model

INTRODUCTION

Due to their high magnetic properties at both room and elevated temperatures, permanent magnets based on Sm-Co system alloys are widely used in aerospace technology, various sensors, wind turbines, hybrid electric vehicles, etc., due to their high magnetic properties at both room and elevated temperatures [1,2]. Compared to the magnets of the Nd-Fe-B system, the main advantage of these magnets is the possibility of reliable operation at temperatures up to 500 °C. Because of the unceasing trend towards miniaturization of products, especially in high-tech industries, and the high cost and scarcity of the leading chemical elements of Sm-Co magnets, the primary efforts of researchers are currently aimed at developing new technologies to improve their properties. It is theoretically predicted [3] that the magnetic properties of nanocomposite magnetic materials consisting of nanoscale hard and soft magnetic phases can be twice as high as the properties of advanced permanent magnets. One of the promising methods for forming such a structure is hydrogen treatment using hydrogenation-disproportionation-desorption-recombination (HDDR) [4]. To date, there is a massive amount of experimental data on the phase transformations and the microstructure evolution in alloys based on SmCo5 and Sm2Co17 compounds depending on the modes of hydrogen treatment [5,6]. The magnetic properties of these alloys are susceptible to the: phase composition; microstructure features; the size of the structural components; the presence of crystallographic texture, etc. The experimental verification of the influence of these parameters on the magnetic properties is highly time-consuming and resource-intensive. This process can be significantly simplified, shortened, and reduced in price by using machine learning (ML), artificial intelligence, or neural network modeling for preliminary prediction of magnetic properties and subsequent experimental confirmation [7-9]. The availability of this approach was demonstrated by a wide range of examples, including materials for alkaline-ion batteries [9], green energy [10], medical destination [11], additive technologies [12], alloys with shape memory effect [13], magnetic materials [14-16], etc. For instance, the authors [14] showed the possibility of ML tools application to create new soft magnetic materials. In this case, the experimental data on the influence of chemical composition, modes of thermal treatment, and grain size on magnetic properties were used to predict saturation magnetization, coercivity, and magnetostriiction of different alloys using a random forest model. The domain structure of Sm (Co, Fe, Cu, Zr)17 alloy [15] was simulated depending on the heat treatment modes using FMRM software [17], developed based on a phenomenological approach to the analysis of reversal magnetization processes. Calculations based on the DFT method were used to predict the

DOI: 10.36547/ams.27.4.1173
effect of transition elements on the structure and magnetic properties of SmCo-based alloys [18].

A literature review shows that modern materials science uses a wide arsenal of computer modeling methods, ranging from electronic structure calculations based on the theory of density functional (DFT method) [19], molecular dynamics [20], Monte Carlo [21], and the phase field [22] methods and ending with the macroscopic approach, when existing machine learning methods are used for analysis using computer programs such as the Toolkit for Modeling Materials for Machine Learning (MAST-ML) [23].

In general, machine learning tasks, most widely used in materials science, can be divided into regression, clustering, and classification. Besides, the problem of probability estimation is considered mainly for the discovery of new materials. At the same time, the tasks of regression, clustering, and classification in the vast majority of studies are used to predict the properties of materials at the macro and micro levels. In addition, machine learning methods are usually combined with various optimization algorithms, such as genetic algorithm (GA), annealing simulation algorithm (SAA), or particle swarm method (PSO), which are mainly used to optimize model parameters [7].

However, in addition to the apparent advantages, it is necessary to pay attention to a number of problems associated with predicting the properties of materials. First of all, it is the correlation dependence of properties on the microstructure, the parameters of which are pretty challenging to consider and describe for computer modeling. In addition, these relationships are individual for each system that requires an unconventional approach. When predicting the properties of materials, it is challenging to build a computational model that fully describes the complex relationships between many factors (chemical and phase composition, microstructure parameters, size of structural components, etc.) that affect the properties and is often unknown. Therefore, there is an urgent need to develop intelligent and highly efficient prediction models that can correctly predict the properties of materials at a short time and at a low cost. The basic idea of using machine learning methods to predict the properties of materials is to analyze and distinguish the relationships (primarily nonlinear) between different factors by extracting knowledge from existing empirical data. This approach allows us to fundamentally understand the relationship between the structure and properties of a particular material and their change during technological operations of production and processing. However, single machine learning algorithms do not always provide the accuracy required to solve a specific applied problem. Therefore, in recent years, the ensemble approach by combining several machine learning algorithms to increase the accuracy of prediction or classification has become quite common. Therefore, this paper aims to develop an ensemble model based on a set of heterogeneous ML-based regressors to solve the problem of Sm-Co alloy's magnetic properties prediction.

The main contribution of this paper can be summarized as follows:

- we have developed a stacking model for predicting coercivity of Sm-Co alloys based on a set of the most efficient heterogeneous machine learning algorithms; we have shown a significant increase in the accuracy of its operation compared to existing machine learning algorithms.

**MATERIAL AND METHODS**

**Data collection**

The formation of the original database was carried out based on literature data. Technological methods of production and processing of Sm-Co system alloys were not considered because, in the last case, the magnetic properties are determined by the formed microstructure and phase composition parameters. Each observation (so-called vector) entered into the database contains input information on the content of chemical elements in the alloy, main and impurity phases, the average size of structural components, the existence of crystallographic texture, basic microstructure characteristics, and magnetic properties as a result of interrelation between them (Fig. 1). It should be noted that a detailed analysis of more than 300 literature sources published over the past 10 years has shown that only about 30% of publications fully contain this information. Besides, data on coercivity are presented in all these publications, while information on other magnetic properties is given lesser. That is why we chose coercive force as a prediction characteristic. The main difficulties in data selection were related to the authors’ use of different measurement units and different approaches to describing the microstructure. In addition, a minimal number of publications contain numerical data on the proportion of phases in the studied materials and the effect of doping with such rare earth elements (REM) and 3d transition metals (TM) as La, Ce, Pr, Ti, and Ni. Accordingly, the features of the database are as follows. The content of chemical elements (the total content is for REM and Ti and Ni), the average size of the structural components, and magnetic properties are represented by numerical values. In this case, the numerical values were converted to the following units: the content of elements – at. %; the average size of the structural components – nm; coercive force \( H_c - \) kOe, saturation magnetization \( M_s \) and remanence \( M_r - \) emu/g; maximum energy product \( BH_{max} - \) MGOe. Data for all other parameters were transformed into a binary system as follows:

- phase composition: 1 corresponds to the presence of the phase, and 0 indicates its absence;
- state of material: the database presents the magnetic properties measured on powders or sintered magnets/ribbons. The form of investigated material is marked as 1 for "powder" and 0 for "sintered magnet/ribbon" in the case of powders and vice versa when otherwise;
- texture: 1 corresponds to the crystallographic orientation of a main ferromagnetic phase along the axis of easy magnetization (c-axis) while 0 indicates the absence of direction;
- microstructure: 1 means "yes" when 0 means "not," as described below.

For simplicity, all possible types of microstructures were compiled to lamellar, flake, nanocrystalline, and cellular. In many cases, analyzed microstructures show a mixture of the parameters denoted by 1 in corresponding columns of the database. Absent types of microstructures were indicated by 0. In addition, the microstructure of each observation was characterized by such parameters as homogeneity (the uniform structure is denoted by 1) and regularity (the regular structure is represented by 1).
The resulting database [102] contains 419 vectors, each described by 31 input features collected from [24-101] (see Table in the Appendix).

**Investigated predictors**

Nowadays, there are many methods of machine learning that are successfully used to solve applied problems in material science. Moreover, many software packages, which implement a number of optimized versions of such methods with an intuitive and simple interface for the intellectual analysis of data, have been developed. One of them is Orange [103]. It contains a set of machine learning algorithms, in particular, to solve the regression task.

One of the simplest and most commonly used machine learning methods is a linear regression with Lasso or Ridge regularizations. It is designed to establish a linear relationship between a set of independent attributes and a dependent variable. The disadvantage of this method is low accuracy in the case of processing significantly nonlinear dependencies.

The timely and optimal implementation of all considered methods in Orange allows for their use in this study. The aim of their consideration is that some such heterogeneous machine learning methods will form the basis of the stacking ensemble model proposed in this article.

**Proposed model**

This work aims to build a prediction model that will provide the highest accuracy in solving the assigned task. In this case, instead of single-based regressors, which do not consistently demonstrate high efficiency, we use an ensemble approach. It is based on the use of a set of basic regressors, the results of which are summarized by a metaregressor. This will increase the accuracy compared to the use of single models that form such a model. The literature considers three main approaches to constructing ensemble models: boosting, bagging, and stacking.

The boosting approach involves the iterative step-by-step execution of the prediction algorithm. Its feature is that the selected ML-based regressor of each subsequent step must consider the regressor errors of the previous step. After performing all the necessary steps, the boosting algorithm combines all the answers into one resulting rule. The disadvantage of this approach is the large number of iterations needed for the effective operation of methods of this class, which requires considerable time for their work.

Bagging involves the use of only one machine learning method as a basic algorithm. However, its training needs a large number of data subsamples. These subsamples are formed by dividing the primary data set intended for model learning, where each subsample may contain a set of observations that are repeated in other ones. The disadvantage of this approach is the necessity to divide the total data set into small subsamples, which can reduce the generalizing properties of methods of this class.

The stacking approach uses a set of heterogeneous or homogeneous algorithms that can work in parallel. Each of them is trained on the whole data set. The results of all members of the
A. Trostianchyn et al. in Acta Metallurgica Slovaca

stacking model are weighed by a user-defined method. The disadvantage of this approach is the need for significant computing power to implement parallel training of all members that form the stacking model. However, modern hardware development eliminates this disadvantage and allows using a stacking approach for solving the various applied tasks.

In this work, we build a prediction model based on the stacking approach. The model assumes the presence of basic $N$-algorithms that will form a stacking ensemble. The meta-algorithm will weigh the results of their work. The work of the meta-algorithm will determine the impact of solving the stated task. A simplified flowchart of this approach is shown in Fig. 2.

The data set collected by us to solve the problem of predicting the magnetic properties of Sm-Co alloys contains many independent attributes. In addition, there are complex and nonlinear, unobvious and unexplored relationships between different features. It is evident that, in particular, many linear machine learning methods will not provide sufficient accuracy. If such algorithms are included in the general ensemble model, they will reduce the accuracy of their work. That is why we propose to perform a preliminary selection of basic algorithms that will form a stacking ensemble. It is based on initial modeling of machine learning algorithms and evaluation of their efficiency using some criteria, in particular:

![Fig. 2 A simplified flowchart of the stacking heterogeneous ensemble model](image)

Mean Square Error:

$$\text{MSE} = \frac{1}{n} \sum_{i=1}^{n} (y_i - y_i^{\text{pred}})^2,$$

Root Mean Square Error:

$$\text{RMSE} = \sqrt{\frac{\sum_{i=1}^{n} (y_i^{\text{pred}} - y_i)^2}{n}},$$

Mean Absolute Error:

$$\text{MAE} = \frac{1}{n} \sum_{i=1}^{n} |y_i^{\text{pred}} - y_i|,$$

Coefficient of determination:

$$R^2 = 1 - \frac{\sum_{i=1}^{n} (y_i - y_i^{\text{pred}})^2}{\sum_{i=1}^{n} (y_i - \bar{y})^2},$$

where: $y_i$ - true value, $y_i^{\text{pred}}$ - predicted value, $n$ - number of observations, $\bar{y} = \frac{1}{n} \sum_{i=1}^{n} y_i$.

It is known that the values of the coefficient of determination ($R^2$) are in the range of 0 to 1. High values of this indicator show a clear relationship between the dependent and the set of independent attributes. Accordingly, values close to 0 indicate the absence of such a relationship; negative values of $R^2$ show the complete inadequacy of the chosen machine learning model.

In this work, we propose including a machine learning algorithm in the stacking model based on the efficiency of its position $R^2$. Since a value of 0.5 indicates the adequacy of the selected model to solve the prediction task, the criterion for including a machine learning algorithm in the stacking ensemble model will be as follows:

$$R^2 \geq 0.5$$

Thus, the proposed stacking model will form a set of basic machine learning algorithms that provide satisfactory results. However, their combined use by weighing their output signals by a certain meta-algorithm should increase the prediction accuracy.

RESULTS AND DISCUSSION

The simulation of all the methods studied in this work was based on the data set collected by us. It contains 419 vectors, each of which is characterized by 31 independent attributes. The value of coercivity was the output-dependent attribute. The data set was randomly divided into training (80%) and test (20%) samples (at each start of one or another method). Each experiment was repeated ten times. As a result, this article presents the average results (after ten runs) of all studied techniques.

The experimental modeling was performed using the Orange data analysis and visualization tool [102]. This visual programming tool contains a set of all the necessary machine learning methods, tools for building stacking models, and powerful tools for visualizing the work results. It was chosen to solve the assigned task due to simplicity, clarity, and ease of use.

Performance evaluation of the investigated ML-based algorithms

In this work, the experimental comparison of the efficiency of eight existing machine learning methods using the criteria (1) - (4) is carried out. A flowchart of this process created using Orange tools is shown in Fig. 3.
The purpose of this experiment is to evaluate the accuracy of the existing regressors and select the best of them based on (4) according to criterion (5) to build a stacking model. All the studied methods based on criteria (1) - (4) are summarized in Table 1.

As can be seen from Table 1, linear regression, SGD and SVR show the lowest accuracy by all indicators. Somewhat better results were obtained when using the Tree algorithm. However, the accuracy of its work is not sufficient. Ensemble methods and an artificial neural network are algorithms that satisfy criterion (5). They will be the basis of the proposed stacking model of heterogeneous regressors.

Table 1 The results of prediction based on criteria (1) - (4) using all the studied machine learning algorithms

<table>
<thead>
<tr>
<th>Machine learning method</th>
<th>Performance indicators</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tree</td>
<td>MSE: 71.317, RMSE: 8.445, MAE: 5.822, R²: 0.413</td>
</tr>
<tr>
<td>SVR</td>
<td>MSE: 95.180, RMSE: 9.756, MAE: 6.540, R²: 0.217</td>
</tr>
<tr>
<td>SGD</td>
<td>MSE: 76.139, RMSE: 8.726, MAE: 5.782, R²: 0.374</td>
</tr>
<tr>
<td>Random Forest</td>
<td>MSE: 49.713, RMSE: 7.051, MAE: 4.473, R²: 0.591</td>
</tr>
<tr>
<td>Neural Network</td>
<td>MSE: 55.655, RMSE: 7.460, MAE: 4.995, R²: 0.542</td>
</tr>
<tr>
<td>Linear regression</td>
<td>MSE: 75.630, RMSE: 8.697, MAE: 5.951, R²: 0.378</td>
</tr>
<tr>
<td>Gradient Boosting</td>
<td>MSE: 49.084, RMSE: 7.006, MAE: 4.539, R²: 0.596</td>
</tr>
</tbody>
</table>

Modeling of the proposed model

In this work, the stacking ensemble was developed on the basis of a set of heterogeneous machine learning algorithms:
- Random Forest;
- AdaBoost;
- Gradient Boosting;
- Neural Network.

Experimental evaluation of their work showed that they are adequate models for solving the problem of predicting the coercivity of Sm-Co alloys based on the collected data set. Three of them are ensemble methods, two of which are based on the boosting approach. In addition, an element of the stacking model is also a multilayer perceptron.

Various machine learning methods can be used as a meta-algorithm. However, logistic regression (for classification task) or ridge regression (for prediction task) are the most commonly used methods. Since the Stacking widget in Orange by default involves the use of the latter, we decided to focus on it.

Flowchart of the proposed ensemble model using Orange tools is presented in Fig. 4.

Fig. 4 Flowchart of the proposed stacking model for Sm-Co alloys coercivity prediction

The results of the developed stacking model for predicting the coercivity of Sm-Co alloys based on (1) - (4) are summarized in Table 2.

Table 2 Prediction results using a stacking model

<table>
<thead>
<tr>
<th>Proposed ensemble model</th>
<th>Performance indicators</th>
</tr>
</thead>
<tbody>
<tr>
<td>Proposed model</td>
<td>MSE: 44.223, RMSE: 6.650, MAE: 4.201, R²: 0.636</td>
</tr>
</tbody>
</table>

As can be seen from Table 2, the proposed approach provides high accuracy.

Comparison and discussion

To confirm the effectiveness of the proposed model, we compared the results of its work with the results of existing methods. In particular, the comparison was made using basic regressors that form the proposed ensemble based on (1) and (4). SVR, SGD, Linear regression, and Tree algorithms were not taken into account. This is because the value of the indicator (4) for these methods is lower than 0.5. The comparison results are summarized in Fig. 5 and Fig. 6.

Fig. 5 MSE values for the proposed ensemble and its single components

As shown from Fig. 5 and Fig. 6, the multilayer perceptron demonstrates the lowest performance accuracy compared to other ensemble members. The Random Forest method, despite the optimal value of the number of trees, does not exceed the results of algorithms based on boosting. AdaBoost and Gradient Boosting demonstrate minor errors compared to all other components of the ensemble. However, the highest accuracy in all four performance indicators was obtained using the devel-
The proposed ensemble scheme showed a significant increase in the prediction accuracy of the magnetic properties of alloys of the Sm-Co system on the example of coercivity. It allows us to use it when solving applied problems of materials science.

CONCLUSION

The development of ML-based methods for predicting the properties of engineering materials is a critical task for modern materials science. Such an approach allows using of computer modeling instead of protracted, expensive, and resource-consuming experimental studies. The properties of materials are determined by many factors, which are characterized by complex nonlinear relations that have their own features for each system. It requires an individual approach to developing the ML model, which provides the best accuracy in each case. In this article, based on the collected data set, an experimental comparison of eight existing machine learning methods was performed when solving the problem of predicting the coercivity of Sm-Co alloys. Based on the collected data set, an experimental comparison of eight existing machine learning methods was performed when solving the problem of predicting the coercivity of Sm-Co alloys. The accuracy of their work was estimated using MSE, RMSE, MAE, and $R^2$ indicators. This modeling aimed to select the optimal machine learning method based on the proposed criterion for building an ensemble model. This approach should provide a significant increase in the accuracy of solving the assigned task. The authors built a stacking ensemble model using heterogeneous elements, Neural Networks, AdaBoost, Gradient Boosting, and Random Forest algorithm were chosen as components of the ensemble. This choice is due to the high value of the coefficient of determination provided by the selected method. The stacking ensemble was modeled using the Orange data analysis and visualization software. Experimental studies have shown a significant increase in the accuracy of the proposed scheme compared with single-based algorithms forming it and other machine learning methods (SVR, SGD, Linear regression, and Random Forest) based on the proposed criterion for building an ensemble model. This approach should provide a significant increase in the accuracy of solving the assigned task.

The stacking ensemble was modeled using the Orange data analysis and visualization software. Experimental studies have shown a significant increase in the accuracy of the proposed scheme compared with single-based algorithms forming it and other machine learning methods (SVR, SGD, Linear regression, and Random Forest) based on the proposed criterion for building an ensemble model. This approach should provide a significant increase in the accuracy of solving the assigned task.

REFERENCES


Fig. 6 $R^2$ values for the proposed ensemble and its single components