MATHEMATICAL THEORY
OF PATTERN RECOGNITION

Prediction Based on the Solution of the Set of Classification Problems of Supervised Learning and Degrees of Membership

A. A. Lukanin[a],*, V. V. Ryazanov[a,b],, and N. N. Kiselyova[c],*****

[a] Moscow Institute of Physics and Technology (National Research University), Moscow, 115184 Russia
[b] Computer Science and Control, Federal Research Center, Russian Academy of Sciences, Moscow, 119333 Russia
[c] Baikov Institute of Metallurgy and Materials Science, Russian Academy of Sciences, Moscow, 119991 Russia
** e-mail: lukanin@phystech.edu*
*** e-mail: rvvccas@mail.ru*
****e-mail: kis@imet.ac.ru*

Abstract—It is proposed to use the degrees of membership of objects to each class in the process of recognition in the linear corrector model to solve the problem of restoring dependences from precedent samples. Two models of the algorithm for calculating estimates are used as classifiers. The work of the proposed model is compared with the original method and with the well-known data analysis methods. The dependence of the work of the linear corrector on its parameters is studied.

Keywords: regression, dependence, feature, classification algorithm, linear corrector, algorithm for calculating estimates, degree of membership

DOI: 10.1134/S1054661820010095

INTRODUCTION

We consider the problem of restoring the dependence between an object (a vector of independent variables (features) $\mathbf{x} = (x_1, x_2, \ldots, x_t)$, $x_j \in M_j$, where M_j is a set of valid descriptions of the jth feature) and the target value $y \in \mathbb{R}$ from the precedent sample $\{(\mathbf{x}_i, y_i)\}_{i=1}^m$. It is assumed that there exists a dependence $y = f(\mathbf{x})$ (regression problem).

There are numerous methods and approaches for restoring dependences from precedent samples. However, they all have their limitations in solving real problems.

In the case of a finite set of values of the target, the recognition or classification problem is solved. For this problem, there are various models and algorithms for which the properties of correctness and stability have been studied.

Therefore, with respect to regression problems, it was proposed [1, 2] to transfer the difficulties associated with comparing objects in the feature space (different types, different information content of the features, agreement of metrics for individual features) to classification problems and apply dependence restoration methods based on their solution with subsequent correction in the space of the target values.

In this paper, we propose to use the degrees of membership of objects to each class in the recognition process in the linear corrector model [3] based on the solution of the set of classification problems generated using the training sample, and subsequent correction in the space of the target values.

Objectives of the study:

1. Implementing the degrees of membership in the linear corrector model of dependence restoration.

2. Comparing two different methods for determining the proximity function in the algorithms for calculating estimates as classifiers of the linear corrector.

3. Comparing the results of the original linear corrector method and the proposed replacement method.

4. Studying the dependence of two variants of the linear corrector on its parameters.

5. Comparing the results of the proposed linear corrector model with the results obtained using the well-known data analysis methods.

Comparative experiments were carried out using samples with data on the parameters (a and c) of the tetragonal crystal lattice of known compounds of composition $A_2B_2DO_7$ with a melilite-type structure. Since these compounds can include cations A and B of different valence, two samples were analyzed. The first sample included the information on 37 melilites of composition $A^{+2}{}_2B_2^{+3}D^{+4}O_7$; and the second sample included the data on 28 compounds of composition $A^{+2}B_2^{+2}D_2^{+4}O_7$, where D is Si or Ge.

Received June 18, 2019; revised October 21, 2019;
accepted October 21, 2019

As features, we used the parameters of the chemical elements A, B, and D, as well as the properties of their simple oxides (AO, BO, B_2O_3, or DO_2) selected based on the experience of chemists. In addition, using a special software [4], the algebraic functions of these properties were selected, which were most important for separating the melilite-structure compounds from substances of the same composition but with a different crystal structure. The resulting training sample included 108 properties (all the initial parameters of elements A, B, and D and the most informative algebraic functions of these parameters).

1. LINEAR CORRECTOR AND PROPOSED REPLACEMENT

Consider the linear corrector model for restoring dependences [3]. In this model, the real line is first divided into n intervals: $y_1, \ldots, y_{m_1} \in \Delta_1$, $y_{m_1+1}, \ldots, y_{m_2} \in \Delta_2, \ldots, y_{m_{n-1}+1}, \ldots, y_{m_n} \in \Delta_n$. From the resulting partition $\boldsymbol{\Delta}$, N sets of l intervals $(2 \le l \le n)$ are obtained, based on which training samples of N classification problems for l classes are formed [3].

Let us express parameter N through n and l. To do this, we group the adjacent intervals of partition $\boldsymbol{\Delta}$ sequentially in all ways, without rearranging them along the real axis. The recursive formula for calculating the N value is as follows:

$$N(n,\ l) = \begin{cases} n-1, & \text{if} \quad l = 2, \\ 0, & \text{if} \quad n < l, \\ \sum\limits_{i=1}^{n-2} N(n-i,\ l-1), & \text{else.} \end{cases}$$

Now it will be easier to compare the work of methods at different values of pairs of parameters.

Further, we select and train classifiers.

Let $c_k = \dfrac{\sum_{i=m_{k-1}+1}^{m_k} y_i}{m_k - m_{k-1}}$ be the average value of the target in the interval Δ_k, $k = 1,\ldots,n$. Let $u_i^{(j)} = \dfrac{\sum_{k=k_i^{(j)}+1}^{k_i^{(j+1)}} c_k}{k_i^{(j+1)} - k_i^{(j)}}$, $i = 1,\ldots,N$, $j = 1,\ldots,l$, where $k_i^{(j)}$ is the index of the last partition from $\boldsymbol{\Delta}$ included in the jth interval in the ith classification problem. Classes are formed according to the following rule: object $\mathbf{x}$ belongs to class K_i^j if and only if $y = f(\mathbf{x}) \in \Delta_k$ and $k_i^{(j-1)} < k \le k_i^{(j)}$, $j = 1,\ldots,l$, $k_i^{(0)} = 1$, $k_i^{(l)} = n$.

Let us build a response vector $\mathbf{u}(\mathbf{x}) = (1, u_1(\mathbf{x}),\ldots,u_N(\mathbf{x}))$, where $u_i(\mathbf{x}) = u_i^{(j)}$ if classifier A_i assigned object $\mathbf{x}$ to class K_i^j. A linear corrector is a function f: $X \to \mathbb{R}$, $f(\mathbf{x}) = (\mathbf{u}(\mathbf{x}), \mathbf{w})$, where $\mathbf{w}$ is the vector of weights, $\mathbf{w} \in \mathbb{R}^{N+1}$ [3]. The vector of weights $\mathbf{w}$ of the linear corrector is obtained by solving the optimization problem

$$\sum_{i=1}^{m} (y_i - f(\mathbf{x}_i))^2 \to \min_{\mathbf{w}}$$

using the stochastic gradient descent.

It is proposed to replace each component $u_i(\mathbf{x})$ of the response vector $\mathbf{u}(\mathbf{x})$ with $\sum_{j=1}^{l} u_i^{(j)} m_i^j(\mathbf{x})$, $i = 1,\ldots,N$, where $m_i^j(\mathbf{x})$ is the degree (measure) of membership of object $\mathbf{x}$ to class K_i^j in the ith classification problem. These parameters are found as the result of optimization (training) for each partition. The following restrictions can be imposed:

1. $0 \le m_i^j \le 1$, $i = 1,\ldots,N$, $j = 1,\ldots,l$,
2. $\sum_{j=1}^{l} m_i^j = 1$, $i = 1,\ldots,N$.

Thus, we will consider two models of the linear corrector f: $X \to \mathbb{R}$, $f(\mathbf{x}) = (\mathbf{u}(\mathbf{x}), \mathbf{w})$, $\mathbf{u}(\mathbf{x}) = (1, u_1(\mathbf{x}),\ldots,u_N(\mathbf{x}))$, where

1. $u_i(\mathbf{x}) = u_i^{(j)}$ if classifier A_i assigned object $\mathbf{x}$ to class K_i^j $(i = 1,\ldots,N$, $j = 1,\ldots,l)$, and
2. $u_i(\mathbf{x}) = \sum_{j=1}^{l} u_i^{(j)} m_i^j(\mathbf{x})$ $(i = 1,\ldots,N)$.

In the original method, we, in fact, multiply by vectors of zeros and one unit: the degree of membership is equal to one for the class to which the classifier assigned the object and zero for the other classes. In the proposed method, we will multiply by a vector of membership measures for each involved class of the classification problem.

2. TWO PROXIMITY FUNCTIONS IN MODELS OF ALGORITHMS FOR CALCULATING ESTIMATES

Algorithms for calculating estimates (ACE) are suited for finding degrees of membership of objects to each class in the process of recognition [5, 6]. Let us consider two ACE models: as a proximity function, we will use both the metric function and the function for arbitrary ordinal features [7].

The first model considers the set Ω of all possible support sets (these sets define the numbers of features by which parts of the reference objects and recognized objects are compared) with cardinality $\tilde{k}$ and the function of proximity of a recognized object $\mathbf{x}_i$ to some reference object $\mathbf{x}_\alpha$ of class K_ν^j, $\nu = 1,\ldots,N$, $j = 1,\ldots,l$, which appears as

$$B_\Omega(\mathbf{x}_i, \mathbf{x}_\alpha) = \begin{cases} 1, & |\mathrm{x}_{ip} - \mathrm{x}_{\alpha p}| \le \varepsilon_p, \quad \forall p \in \Omega, \\ 0, & \text{else} \end{cases}$$

$$\forall \mathbf{x}_\alpha \in K_\nu^j,$$

where $\boldsymbol{\varepsilon}$ is the vector of features' measurement accuracy. Thus, the distances between the features of the objects are calculated or, in other words, a metric is introduced.

In the second model, the following proximity function was considered for the same support sets:

$$\tilde{B}_\Omega(\mathbf{x}_\alpha, \mathbf{x}_i, \mathbf{x}_\beta)$$

$$= \begin{cases} 1, & (x_{\alpha p} \le x_{ip} \le x_{\beta p}) \vee (x_{\beta p} \le x_{ip} \le x_{\alpha p}), \\ \forall p \in \Omega, & \\ 0, & \text{else} \end{cases}$$

$$\forall \mathbf{x}_\alpha, \mathbf{x}_\beta \in K_\nu^j,$$

where $\mathbf{x}_i$ is the object being recognized; $\mathbf{x}_\alpha$ and $\mathbf{x}_\beta$ are the reference objects of class K_ν^j; and $\vee$ is a disjunction. This proximity function can already be used for any ordinal features (i.e., when the distance between the two values of the feature has no meaning), since it only examines whether the values of the features of the compared objects are within certain limits for a given informative fragment.

In the first case, we use an auxiliary function $d(\mathbf{x}_i, \mathbf{x}_\alpha) = |\{p : |x_{ip} - x_{\alpha p}| \le \varepsilon_p,\ p = 1, 2, \ldots, t\}|$. Then, the estimate for class K_ν^j looks as follows [5]:

$$\Gamma_\nu^j(\mathbf{x}_i) = \frac{1}{|K_\nu^j|} \sum_{\mathbf{x}_\alpha \in K_\nu^j} C_{d(\mathbf{x}_i, \mathbf{x}_\alpha)}^{\tilde{k}},$$

if $\mathbf{x}_i \notin K_\nu^j$ and

$$\Gamma_\nu^j(\mathbf{x}_i) = \frac{1}{|K_\nu^j| - 1} \sum_{\substack{\mathbf{x}_\alpha \in K_\nu^j \\ \alpha \ne i}} C_{d(\mathbf{x}_i, \mathbf{x}_\alpha)}^{\tilde{k}},$$

if $\mathbf{x}_i \in K_\nu^j$.

In the second case, similarly to the first, we introduce the function $d(\mathbf{x}_\alpha, \mathbf{x}_i, \mathbf{x}_\beta) = |\{p : (x_{\alpha p} \le x_{ip} \le x_{\beta p}) \vee (x_{\beta p} \le x_{ip} \le x_{\alpha p}),\ p = 1, 2, \ldots, t\}|$ and obtain

$$\Gamma_\nu^j(\mathbf{x}_i) = \frac{2}{|K_\nu^j|\left(|K_\nu^j| - 1\right)} \sum_{\mathbf{x}_\alpha, \mathbf{x}_\beta \in K_\nu^j, \alpha < \beta} C_{d(\mathbf{x}_\alpha, \mathbf{x}_i, \mathbf{x}_\beta)}^{\tilde{k}},$$

if $\mathbf{x}_i \notin K_\nu^j$ and

$$\Gamma_\nu^j(\mathbf{x}_i) = \frac{2}{\left(|K_\nu^j| - 1\right)\left(|K_\nu^j| - 2\right)} \sum_{\substack{\mathbf{x}_\alpha, \mathbf{x}_\beta \in K_\nu^j,\ \alpha < \beta \\ \alpha \ne i,\ \beta \ne i}} C_{d(\mathbf{x}_\alpha, \mathbf{x}_i, \mathbf{x}_\beta)}^{\tilde{k}},$$

if $\mathbf{x}_i \in K_\nu^j$.

Let us write out the process of calculating estimates in the cross-validation mode. We assume that the test sample coincides with the training sample. In the case of the first model, it is necessary to calculate the matrix $D_1 = \|d_{ij}\|_{m \times m}$, where $d_{ij} = C_{d(\mathbf{x}_i, \mathbf{x}_j)}^{\tilde{k}}$ (the parameter value $\tilde{k}$ is chosen by ourselves). Then, $\Gamma_\nu^j(\mathbf{x}_i) = \frac{1}{|K_\nu^j|} \sum_{\mathbf{x}_\alpha \in K_\nu^j} d_{i\alpha}$, if $\mathbf{x}_i \notin K_\nu^j$, and $\Gamma_\nu^j(\mathbf{x}_i) = \frac{1}{|K_\nu^j| - 1} \sum_{\substack{\mathbf{x}_\alpha \in K_\nu^j \\ \alpha \ne i}} d_{i\alpha}$, if $\mathbf{x}_i \in K_\nu^j$. In the case of the second model, it is necessary to calculate the matrix $D_2 = \|d_{\alpha i \beta}\|_{m \times m \times m}$, where $d_{\alpha i \beta} = C_{d(\mathbf{x}_\alpha, \mathbf{x}_i, \mathbf{x}_\beta)}^{\tilde{k}}$ and use it in the calculation of estimates.

For both ACE models, in order to obtain the degree of membership of an object to each class, it is necessary to normalize estimates for classes using the ℓ_1-norm: $\|\mathbf{x}\|_1 = \sum_i |x_i|$. This is necessary to satisfy the limitations of the proposed replacement in the linear corrector.

In the original linear corrector model, i.e., when all degrees of membership except one are zero, it is also possible to use ACE. However, in this case, it is necessary to determine the following decision rule: the object will belong to the class with the maximum calculated estimate.

Let us give a description of finding the parameters ε_i $(i = 1, \ldots, t)$, i.e., the components of the vector of the features' measurement accuracy, $\boldsymbol{\varepsilon}$, which is used for the first ACE model. First, we obtain all the unique target values of the training sample. To do this, we set the number of digits for rounding the targets: to the right of the decimal point if this number is positive and to the left of the decimal point if the number is negative. Those targets whose rounded values are equal up to this order of magnitude will be considered nonunique. We find the minimum and maximum values of each feature corresponding to each unique target value and calculate their difference. Further, we obtain the features' measurement accuracies as the averages of these differences over all the unique values of the target.

3. STUDY OF THE PROPOSED MODEL

The regression problem of estimating the parameters of a melilite crystal lattice will be solved using the original linear corrector model and the proposed replacement model with the two described ACE mod-

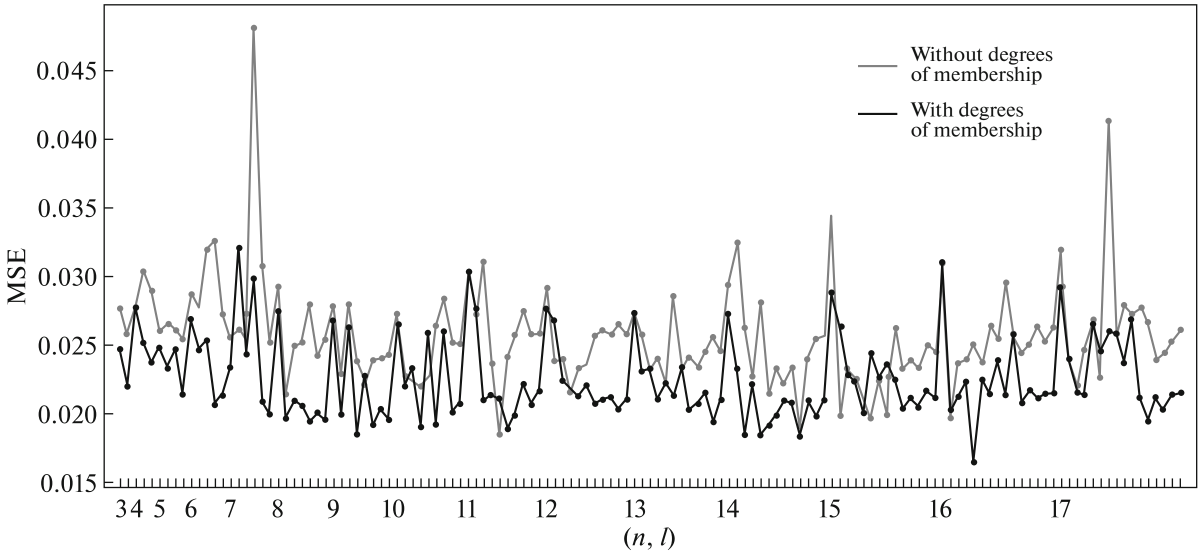

Fig. 1. Mean square error of the prediction as a function of parameters n and l of two linear correctors that use the first ACE model.

els as classifiers. We will apply cross-validation to the data and compare the averaged mean square errors of the prediction on the test samples.

3.1. Comparison of the Linear Corrector Versions

In this part of the study, we compare the results of the two versions of the linear corrector: the original method and the method with the proposed replacement.

The following ranges of values were considered as hyperparameters n and l of the algorithms: $3 \leq n \leq 17$ and $2 \leq l \leq n$.

Figure 1 shows the chart of the mean square error of the prediction of the linear corrector algorithms that use the first ACE model versus hyperparameters n and l. The x axis shows the values of parameter n; the corresponding values of parameter l ($2 \leq l \leq n$) are indicated by strokes for compactness. In other words, the value of parameter n marked on the x axis corresponds to the pair of the algorithm parameters $(n,2)$, the stroke to the right of it corresponds to $(n,3)$, etc., until (n,n). The y axis shows the mean square error (MSE) of the prediction. The gray markers indicate the prediction errors of the original linear corrector method without using the degrees of membership, i.e., when the object belongs to the class with the maximum obtained estimate. The black markers show the same for the method with the proposed replacement, taking into account the degrees of membership of objects to classes. The markers are connected by lines for convenience.

It can be seen from this chart that the method with the proposed replacement almost always works better than the original; i.e., it gives a smaller error for the problem under consideration.

Figure 2 shows a similar chart for the second ACE model.

We can find out from Fig. 2 that the proposed method with the second model (with the proximity function for arbitrary ordinal features) also often gives smaller mean square errors.

This is easier to see in the following two charts. We average the mean square errors of the prediction over all values of parameter l for fixed values of n. Figures 3a and 3b show this dependence for both correctors using the first and second ACE models, respectively.

It is worth noting that the linear corrector with the proposed replacement gives the smallest errors in the case of both the first (0.016) and the second (0.0205) ACE model.

3.2. Dependence of the Linear Corrector on Its Hyperparameters

Let us note the dependence of the errors of the two versions of the linear correctors on its parameters n and l.

Figure 1 clearly shows small fluctuations of the error of the proposed method for almost all the pairs of n and l: if we take non-extreme values of l, all the MSEs are approximately 0.02. Figure 2 shows that for fixed n, the errors are smaller when l is close to $\frac{n}{2}$: they are in the neighbourhood of 0.025.

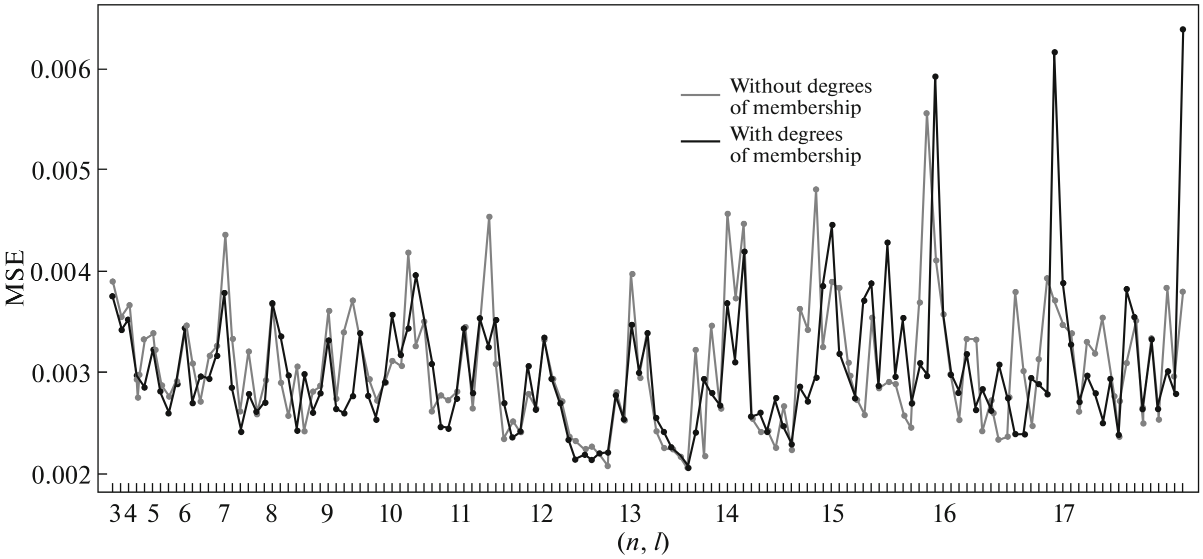

Fig. 2. Mean square error of the prediction as a function of parameters n and l of two linear correctors that use the second ACE model.

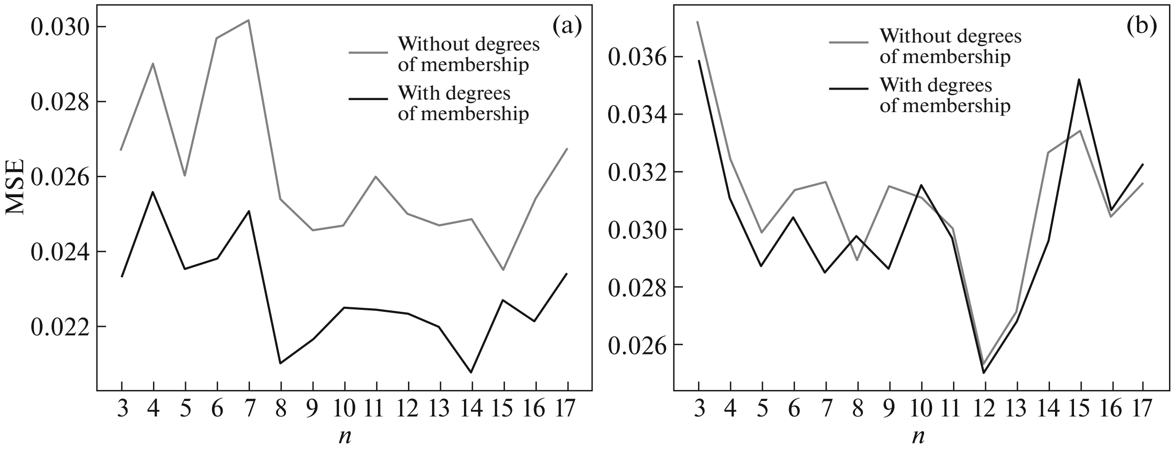

Fig. 3. Mean square error of the prediction averaged over parameter l as a function of parameter n of the two linear correctors that use the (a) first and (b) second ACE models.

Figures 4a and 4b show the sections of charts *1* and *2* for $n = 16$ and $n = 13$, respectively; they refer to the cases where the lowest values of the mean square error of the prediction were obtained. Since n is fixed, the x axis of these charts shows the values of parameter l.

These charts confirm these conditions for obtaining the smallest errors with a linear corrector with the proposed replacement. Thus, in the case of using the metric proximity function, there is no need to experiment with the hyperparameters of the linear corrector and we can simply take any n and the corresponding non-extreme value of l. And in the case of the proximity function for arbitrary ordinal features, it is sufficient to consider various values of parameter n and the corresponding values of l close to $n/2$.

Let us now study how the work of the two described versions of the linear corrector depends on the main parameter n. This dependence was shown in Figs. 3a and 3b, which illustrate nonmonotonic functions. However, we can clearly see the minima that can be found by dividing the set of targets of the training sample into different numbers of intervals.

3.3. Comparison of the Proposed Model with the Well-Known Methods

Let us compare the results of the work of the two versions of linear correctors with the results obtained using the well-known methods and approaches of data analysis. Table 1 shows various mean square errors of the prediction.

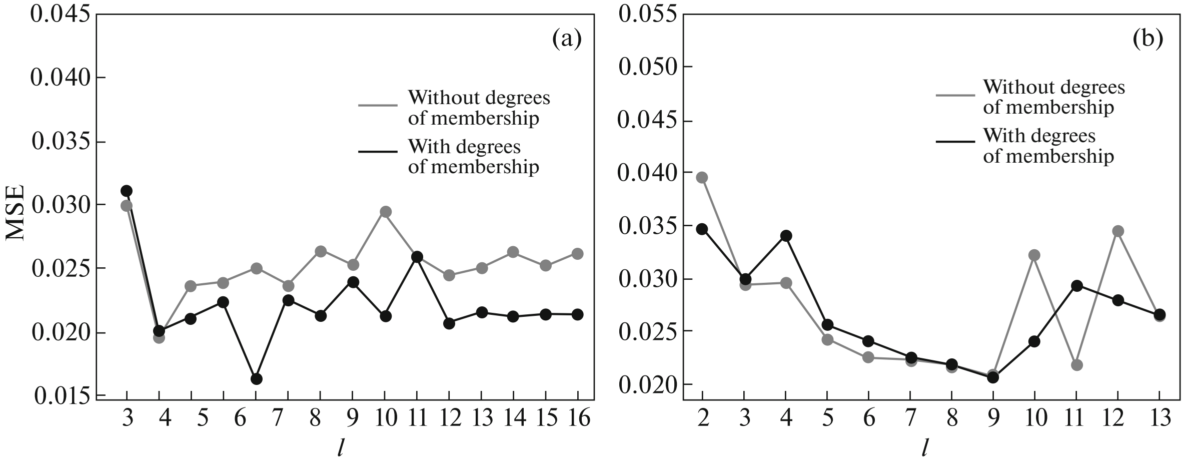

Fig. 4. Mean square error of the prediction as a function of parameter l of the two linear correctors that use the (a) first ACE model at $n = 16$ and the (b) second one at $n = 13$.

The table shows that the proposed version of the linear corrector is in second place after linear regression, giving a better result than the rest of the well-known regression models.

CONCLUSIONS

The main results of this study are as follows:

1. The problem of restoring the dependence from precedent samples is considered.

2. The degrees of membership are implemented in the linear corrector model of restoring dependences.

3. Two different ways of determining the proximity function in the algorithms for calculating estimates as classifiers are compared: the model with the metric proximity function works slightly better than the model with the proximity function for arbitrary ordinal features.

4. The results of the original linear corrector method and the method with the proposed replacement are compared: the proposed model works better.

5. The dependence of the two versions of the linear corrector on its hyperparameters is studied: the smallest mean square error of the prediction of the algorithms is achieved at non-extreme values of parameter l.

6. The results of the proposed linear corrector model are compared with the results obtained using the well-known data analysis methods: both versions of the linear corrector work better than most of the regression models.

Table 1. Mean square errors of the prediction obtained by linear correctors and various methods of data analysis

Method	MSE
Linear corrector with proposed replacement + first ACE model	0.016
Original linear corrector model + first ACE model	0.018
Linear corrector with proposed replacement + second ACE model	0.0205
Original linear corrector model + second ACE model	0.0207
Linear regression	0.006
Theil-Sen Estimator regressor	0.02
Random Forest regressor	0.019
Decision Tree regressor	0.026

FUNDING

The study was supported in part by the Russian Foundation for Basic Research, project nos. 18-01-00557-a and 18-07-00080-a.

CONFLICT OF INTEREST

The authors declare that they have no conflict of interest.

REFERENCES

1. V. V. Ryazanov and Yu. I. Tkachev, “Solution of the dependence estimation problem using groups of recognition algorithms,” Dokl. Math. **81** (3), 500–504 (2010).
2. V. Ryazanov, “Regression via logic supervised classification,” in *Pattern Recognition, Proc. 5th Mexican Conference, MCPR 2013*, Ed. by J. A. Carrasco-Ochoa, J. F. Martínez-Trinidad, J. S. Rodríguez, and G. S. di Baja, Lecture Notes in Computer Science (Springer, Berlin, Heidelberg, 2013), Vol. 7914, pp. 242–253.
3. Yu. I. Tkachev, *Methods for Solving the Dependence Estimation Problem Using Groups of Recognition Algo-*

rithms, Dissertation for the Candidate of Sciences Degree in Physics and Mathematics (Dorodnicyn Computing Centre, Russian Academy of Sciences, Moscow, 2013) [in Russian].

4. O. V. Senko, "An optimal ensemble of predictors in convex correcting procedures," Pattern Recogn. Image Anal. **19** (3), 465–468 (2009).
5. Yu. I. Zhuravlev, *Selected Scientific Works* (Magistr, Moscow, 1998) [in Russian].
6. Yu. I. Zhuravlev, V. V. Ryazanov, and O. V. Sen'ko, *Recognition*: *Mathematical Methods. Program System. Practical Applications* (FAZIS, Moscow, 2005) [in Russian].
7. A. A. Lukanin and V. V. Ryazanov, "Prediction based on the solution of a set of classification problems of supervised learning," in *Applied Mathematics and Computer Science*: *Proc. 60th Scientific Conference of MIPT* (Moscow, Russia, 20–26 November 2017) (Moscow Institute of Physics and Technology, Moscow, 2017), pp. 63–65.

Translated by M. Chubarova

Artem Alexandrovich Lukanin. Born 1996. Graduated with a bachelor's degree from the Department of Control and Applied Mathematics of the Moscow Institute of Physics and Technology (MIPT) in 2017. Currently finishing a master's degree at the MIPT and is entering postgraduate education. Scientific interests: methods for optimization of recognition models, algorithms for searching and processing logical regularities of classes according to precedents, mathematical recognition models based on voting on sets of logical regularities of classes, committee synthesis of collective clustering and construction of stable solutions in clustering problems, restoring data gaps, restoring regressions from sets of recognizing algorithms, creating software classification systems, solving practical problems in medicine, technology, chemistry, and other areas.

Vladimir Vasil'evich Ryazanov. Born 1950. Graduated from the Moscow Institute of Physics and Technology in 1973. Received candidate's degree in 1977 and doctoral degree in 1994. Full member of the Russian Academy of Natural Sciences since 1998 and Professor since 2008. Has been working at the Computing Center of the Russian Academy of Sciences since 1976. Currently the head of the department of classification methods and data analysis of the Dorodnitsyn Computing Center of the Computer Science and Management Federal Research Center of the Russian Academy of Sciences. Author of 208 papers. Scientific interests: methods for optimization of recognition models, algorithms for searching and processing logical regularities of classes according to precedents, mathematical recognition models based on voting on sets of logical regularities of classes, committee synthesis of collective clustering and construction of stable solutions in clustering problems, restoring data gaps, restoring regressions from sets of recognizing algorithms, creating software classification systems, solving practical problems in medicine, technology, chemistry, and other areas.

Nadezhda Nikolaevna Kiselyova. Born 1949. Graduated from the Faculty of Chemistry of Moscow State University in 1971 and postgraduate program at the same faculty in 1974. Received candidate's degree in 1975 and doctoral degree in 2004. Currently the head of the laboratory of semiconductor materials at the Baikov Institute of Metallurgy and Materials Science of the Russian Academy of Sciences. Scientific interests: computer support for design of inorganic compounds, databases of properties of inorganic substances and materials, electronic materials. Author of more than 150 papers and two monographs.