A brief comparative study of the potentialities and limitations of machine-learning algorithms and statistical techniques

Li Eckart¹, Sven Eckart², and Margit Enke³

¹Department of Marketing and International Trade, Technische Universität Bergakademie Freiberg, Freiberg, Germany
²Institute of Thermal Engineering, Technische Universität Bergakademie Freiberg, Freiberg, Germany
³Department of Marketing and International Trade, Technische Universität Bergakademie Freiberg, Freiberg, Germany

Abstract: Machine learning is a popular way to find patterns and relationships in high complex datasets. With the nowadays advancements in storage and computational capabilities, some machine-learning techniques are becoming suitable for real-world applications. The aim of this work is to conduct a comparative analysis of machine learning algorithms and conventional statistical techniques. These methods have long been used for clustering large amounts of data and extracting knowledge in a wide variety of science fields. However, the central knowledge of the different methods and their specific requirements for the data set, as well as the limitations of the individual methods, are an obstacle for the correct use of these methods. New machine learning algorithms could be integrated even more strongly into the current evaluation if the right choice of methods were easier to make. In the present work, some different algorithms of machine learning are listed. Four methods (artificial neural network, regression method, self-organizing map, k-means al-algorithm) are compared in detail and possible selection criteria are pointed out. Finally, an estimation of the fields of work and application and possible limitations are provided, which should help to make choices for specific interdisciplinary analyses.

1 Introduction

“How to use the power of such big data” becomes a central topic across various research domains in the nowadays big data world. Machine-learning algorithms can provide practicable solutions to analyze and leverage the massive volume of available data. The roots of machine learning can be found in computer science, engineering, mathematics, physics, and astronomy. These technologies also find extensive practical use in the field of neuroscience and medicine [1-13]. In that field, artificial intelligence (AI) was used for very topical issues such as the spread of pandemics [14] and as an effective cancer prediction method [15]. Various problems were also investigated in the field of engineering sciences. Salehi et al. [16] use AI for the prediction of structures in the construction sector.
Whereas Tabor et al. [17] used these approaches to achieve a better synergy of different renewable energy areas and the chemical industry, Ali and Frimpong [18] applied these methods for the automation of mining processes and mining systems. In contrast, these methods are used in the social sciences to make risk analyses fairer [19] and to better understand learning processes and emotions [20]. However, merely a few studies discussed the application and perspectives of machine learning algorithms for the marketing research area [21,22].

Since the selection of the appropriate method represents a complex issue, this work attempts to introduce four novel machine-learning techniques: the artificial neural network (ANN), regression method, k-mean-algorithm, and self-organizing map (SOM), to compare them with similar approaches, as well as to point out the trends and perspectives for further research issues. Finally, an assessment of the fields of use and potential limitations will be provided to assist in the selection for specific interdisciplinary analyses.

Two main types of machine learning algorithms are considered in the present work: supervised and unsupervised learning. The supervised machine-learning approaches can solve prediction and regression problems, whereas the unsupervised learning process is involved in the explorative analysis with little or no prior knowledge, for instance, clustering problems. This work confines attention to ANN as representative of the supervised learning algorithms and to SOM as representative of the unsupervised learning algorithms. Furthermore, machine-learning approaches name the statistical terms differently from the conventional statistics. For instance, the input and output of machine learning algorithms equate respectively the independent and dependent variables in the conventional statistics. The target and error in machine learning can be viewed as the actual value and residuals of the conventional techniques. Moreover, the training cases or samples are applied to conduct the machine-learning process. Such training cases are called observations in the conventional framework.

In order to allow a general comparison of the four selected forecast methods, publications since 1980 were evaluated and shown in Figure 1. All publications examined were in the scientific database Scopus. Keywords, abstracts, and titles were included in the clustering and filtered according to the four methods used in this work. Subsequently, the publications were assigned to the research areas and were investigated in detail. It can be seen that the first rise in usage can be observed, in particular, after the release of the world wide web and its commercial use around 1990. The second rise appeared in the mid-2000s, where the electronic data sum quadrupled in only four years. [23]. An additional rapid increase in machine learning, which can be related to the topics of smart city [24,25], smart home [26], industry 4.0 [27,28] and digitization [29,30], can be observed in the last few years after 2014 [31]. Overall, the ANN method is mentioned and applied most often, followed by K-means, regression method, and SOM.
The total number of publications was spread mainly among nine scientific fields, covering over 60% of all papers on this topic. This data set is shown in Figure 2. The four fields (computer science, engineering, mathematics, and medicine) account for at least 40% of each method. In the ANN, SOM, and K-means, the fields of engineering and computer science are represented by more than 20%. The usual disciplines play a rather small role,
sometimes less than 1%. It can be shown that the dissemination of the method could be significantly expanded, especially in the business, management, and economic sectors. Furthermore, the scientific areas in which the methods are already being used have been identified. They are shown in Figure 3. Specific fields of science are not uniformly distributed. While regression methods are much more commonly used in medicine and economics, neural networks are mainly used in engineering, physics, and environmental sciences. The K means method is being used increasingly in the decision sciences, as well as in mathematics. The SOM method is also a common approach in computational science and mathematics.

2 Methods

In the present work, four different mathematical approaches were used, namely Artificial Neural Network (ANN), Self-Organizing Map (SOM), Regression Method, and K-mean method. A short description of the classifiers used and their current applications can be found in the following section.

2.1 Artificial neural network

As a special type of supervised machine learning, the artificial neural network (ANN) is inspired by the human neural systems, first introduced in the 1950s. Due to the massive scale of big data and novel supervised learning algorithms, this approach has lately resurfaced.

An ANN is viewed as a system of connected directed artificial neurons. In an ANN, the process of a biological neural network will be reproduced with connected artificial neurons. Each network layer consists of artificial neurons and bias. The network signals will pass through various network layers and finally arrive at the output layers. Mathematical functions (propagation function and activation function) represent the strengths of neural connections and activation levels of neurons. The connecting weights will be stored in the propagation function. The received signal will be transformed into network input using the activation function, also called the transfer function. Bias can be seen as a weight, of which input always equals one. Adding bias into an the ANN enables the model to perform an affine transformation (nonlinear) instead of a linear transformation. The application of bias prevents the problem that a neuron without a bias will always transfer zero when the network inputs are zero.

The magnitude and direction of the neural connections show the different contributions of inputs referring to the output. Inputs with larger connection weights possess greater intensities in the signal transfer process, compared to inputs with smaller connection weights. The connection directions can be inhibitory (negative) or excitatory (positive). The inhibitory connections reduce the intensity of the received signal and the excitatory connections, oppositely, intense the received signal. It indicates that the connection weights can provide insight into the relative importance of input variables. Therefore, the information about the weights can be used to assess the robustness of the observed system.

Based on the universal approximation theorem [32], an ANN with one hidden layer can approximate any continuous function arbitrarily well, regardless of the activation function and the dimension of the input space [33]. An ANN with bias, a nonlinear, and a linear output layer is capable of approximating any function with a finite number of discontinuities without any previous justification. This shows the high predictive ability of ANNs and gives the theoretical support for the great performance of ANNs in nonlinear statistical modeling.
Given input and target data, the weights of the ANN can be trained by various learning algorithms, for instance, backpropagation. In each learning step of the backpropagation, the weights in the propagation function will be adapted to minimize learning errors (differences between outputs and targets). The adaptive changes of weights follow the learning rule determined by learning algorithms. The knowledge of the trained ANN stored in weights and bias can be used for further prediction.

Among different structures of ANNs, a multilayer feed-forward neural network (MFNN) consisting of three layers (input, hidden, and output layer) is favored among the scholars (see Figure 4). An ANN with \( n \) \((x_i, i = 1, \cdots, n)\) inputs is considered. \( w_{i,j} \) implies the weights of the connections between the input layer and the hidden layer, which consists of \( m \) neurons \((j = 1, \cdots, m)\). \( b_j \) describes the bias of the neuron \( j \). The network signal \( net_j \) is represented by the propagation function in terms of inputs, weights, and bias

\[
net_j = \sum (x_i \cdot w_{i,j} + b_j)
\]  

(1)

Depending on the activation level of neurons, the network signal is transmitted with the help of the activation function

\[
a_j = f(net_j) = f(\sum (x_i \cdot w_{i,j} + b_j)).
\]  

(2)

After the activation process, the signal \( a_j \) is passed to the output layer. For simplicity, it is assumed that merely one neuron is contained in the output layer. The output value of the neuron in the output layer is

\[
Y = \sum w_{j,r} \cdot f(\sum (x_i \cdot w_{i,j} + b_j)) + b_r
\]  

where \( b_r \) is the bias of the output neuron. The frequently employed learning goal includes squared error \( \sum (Y - y)^2 \) and absolute error \( |Y - y| \) between the output \( Y \) and the target \( y \).[13]

![Fig. 4. Schematic illustration of ANNs.](image)

### 2.2 Regression method

Similar attempts to examine the relationship between the independent and dependent variables were also identified by the idea of the regression. Wold’s decomposition theorem [34] contributes to the foundation of the linear forecasting issue of Kolmogorov in 1941. In this paper, the regression method can be subdivided into generalized linear models (MLM) and support vector regression (SVR).
MLM employs the linear function to describe the relationship, considering parallelly the
error distribution. Thus, the predicted value can be given by
\[ Y_i = g(\mathbf{x}_i^T, \beta_j) + E_i. \]

\( E_i \) represents the error term and \( g \) is the linear regression function, which must be
predefined. Previous studies have shown the easy usage and the efficiency of the MLM.
However, the main limitation of the MLM, its incompetence to handle nonlinear problems
has also been pointed out.

In contrast to the MLM, SVR can be applied to a nonlinear context. Additionally, SVR
defines the acceptable area in the model to fit the data in a better way. The input is mapped
through a nonlinear function (so-called kernel function) [35]. To predict the dependent
variable, a linear regression function is computed in high-dimensional feature space [36].
The determination of the kernel function is crucial for the result accuracy of the SVR.
However, there are still no common guidelines for the selection of the kernel function.

SVR optimizes the prediction based on the principle of structural risk minimization.
Derived from this principle, the upper bound of the general error which includes the
training error will be minimized [37].

<table>
<thead>
<tr>
<th>Table 1. Comparison of advantages and disadvantages of the SVR regression method.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Advantage</td>
</tr>
<tr>
<td>---</td>
</tr>
<tr>
<td>● Consideration of the general error, not only the training error</td>
</tr>
<tr>
<td>● Better balance between the result accuracy and computational cost</td>
</tr>
</tbody>
</table>

2.3 Self-organizing map

As part of unsupervised learning, the self-organizing map (SOM) acts as the cluster
approach in machine learning. Kohonen[38] first introduced the SOM. The unsupervised
learning rules refer to the machine-learning rules of the artificial neural networks by giving
the input but without revealing the output, and it assumes no special data structure. Hence,
this approach is especially appropriate for explorative data analysis in terms of identifying
hidden patterns in unknown structured data sets. Indeed, the SOM is a single-layer neural
network with a pre-specified number of neurons spread on a two-dimensional lattice. The
purpose of this approach is to classify objects and to reduce the complexity of high-
dimensional problems.

The input data will be transformed into a low-dimensional discrete map by topological
mapping. The nonlinear statistical relationships between high-dimensional data can be
converted into geometric relations on a low-dimensional display [38]. The SOM follows
the “winner-takes-all” rule. Its process generally comprises four steps: initialization,
competition, cooperation, and adaption.

1) In the first step of initialization, all the connection weights are initialized with
small random values.

2) For a given input vector \( \mathbf{x}_i, i = 1, \cdots, n \), its similarity to the neurons \( \mathbf{w}_j, i = 1, \cdots, n \) can be depicted as the Euclidean distance between them. The closest neuron \( \mathbf{w}_k \) to
the input vector is the so-called winning neuron, mathematically: \( \| \mathbf{x}_i - \mathbf{w}_k \| = \min\| \mathbf{x}_i - \mathbf{w}_j \|, \ i = 1, \cdots, n \). The values computed by the discrimination function (e.g., the Euclidean
distance function) provide the basis for the next step.

6
3) In the last step, the connection weights of the winning neuron and its connected neighborhood neurons will also be adapted according to the rule proposed by Kohonen [38,39]:

\[ w_i(t+1) = w_i(t) + \alpha(t) \cdot g_i(t) \cdot [x_i(t) - w_i(t)]. \]

\( t \) presents the number of iteration steps. The connection weights in the \( t + 1 \) iteration step depend on the previous outcome \( w_i(t) \) and the learning rate \( \alpha(t) \), as well as the neighborhood function \( g_i(t) \). \( g_i(t) \) determines how fast the connection weights of the neighborhood are allowed to be adapted. One of the most commonly appropriated neighborhood functions of the SOM approach is the Gaussian function. Based on the Gaussian function, the neurons near the winning neuron will “learn” faster than those far away. For more details about this approach, see the work of Kohonen [38,39].

4) Repeat steps 2 and 3 until the convergence criteria are fulfilled.

As a result, the information of the initial high-dimensional problems will be visualized by different color shades on a 2-dimensional lattice in SOM weight planes. After the SOM learning process, each training sample is connected to all neurons on the lattice and assigned to a neuron according to connection weights (similarities). It implies that all the input samples in a certain neuron share similar characteristics. The following illustration (Figure 5) shows an exemplary visualization of the SOM using MATLAB:

![Visualization of a self-organizing map](image)

Fig. 5. Visualization of a self-organizing map.

2.4 K-means algorithm

Like the SOM, the conventional approach k-means algorithm belongs to the squared error-based (vector quantization) clustering algorithms. Except for the squared error-based approaches, hierarchical, density-based, matrix optimization-based and mixture algorithms have also been reviewed in the previous studies [40,41]; see the following table.

<table>
<thead>
<tr>
<th>Category</th>
<th>Clustering algorithm</th>
</tr>
</thead>
<tbody>
<tr>
<td>Square error-based</td>
<td>• k-means [42]</td>
</tr>
<tr>
<td>algorithms</td>
<td>• self-organizing analysis [43,44]</td>
</tr>
<tr>
<td></td>
<td>• genetic k-means algorithm [45]</td>
</tr>
<tr>
<td></td>
<td>• PAM (Partitioning Around Medoids) [46]</td>
</tr>
</tbody>
</table>
Due to the limited scope of this paper, the k-means algorithm is selected to conduct the comparison with the SOM. The cross-categorical observations could be interesting fields for future research.

The K-means algorithm is classified as a non-hierarchical approach. Therefore, the number of clusters should be predefined. The aim of this method is to assign objects with similar characteristics to the same cluster, and different clusters should be heterogeneous. The K-means algorithm is a greedy algorithm. Due to the characteristics of the greedy algorithm, the best solution can converge into a local minimum, and the global optimum cannot be guaranteed. Besides, the initial centroid choice influences the accuracy of the result of the k-means algorithm.

Based on the comparison of the Euclidean distances between one object and the centroid of each cluster, the object is assigned to the closest cluster (with minimal Euclidean distance). After the update of cluster members, the centroids will be recalculated. The stop criteria of this algorithm are 1) the marginal modification of the centroids between the iteration steps and 2) every object is assigned to one cluster [58].

The basis iteration steps of the k-means algorithm:

1) Determine the Cluster number (K)
2) Initialization of the centroids (c_j, j = 1,⋯,K)
3) Allocation of the training data to each cluster based on the minimal Euclidean distance criterion: ||x_i - c_l|| = min ||x_i - c_j||, j = 1,⋯,K, 1 ≤ l ≤ K. x_i is the i-th sample of input data set and is assigned to the l-th cluster.
4) The regarded centroid will be recalculated after the assignment of x_i
   \[ c_{l,new} = \frac{\sum_{x_i \in N_{l,new}} x_i}{N_{l,new}} \]
   \( c_{l,new} \) is the updated centroid considering the new cluster member \( x_i \). \( N_{l,new} \) is the number of the cluster members including \( x_i \).
5) Repeat steps 3 and 4 till convergence occurs.

3 Discussion and classification

In the following, the methods described above are examined and validated for different possible applications. Based on this, concrete recommendations for their application are derived. Two of the four methods are compared with each other for a specific application area. First, the use of the ANN and regression is investigated, and then the SOM and k-means are compared in detail.
### 3.1 Comparison between artificial neural network and regression

In this section, two approaches predicting the relationship between the independent and the dependent variables are discussed. While the ANN algorithm has been successfully applied as a powerful prediction tool to various research disciplines [3,6,59,60], the regression method is viewed as one of the standard tools to deal with prediction problems [35,59,36]. The ANN algorithm and the regression method are selected to conduct the comparative analysis owing to the following considerations: 1) the ANN algorithm, as well as the regression method, is proposed to examine the relationship between the independent and dependent variables; 2) both techniques can be used with respect to minimizing the mean squared error between the predicted value and the actual value.

Based on the previous considerations, it can be summarized:

**Table 3.** Comparison between artificial neural network and regression.

<table>
<thead>
<tr>
<th>Advantages of ANN</th>
<th>Disadvantages of ANN</th>
</tr>
</thead>
<tbody>
<tr>
<td>1) The ANN algorithm is a data-driven approach. The underlying relationship between inputs and outputs is based on the given data sets. To develop a regression model, this underlying relationship is predefined by a (linear or nonlinear) regression function. The connection weights between the neurons are similar to the beta coefficients in a least-squares regression model. However, the effect of an independent variable on the dependent variable is depicted by different neurons on different layers instead of by only one beta coefficient in the regression approach.</td>
<td>Due to lack of transparency, the ANN model has been criticized as a “black box” model. Based on data sets, an ANN is trained to learn the relationship between the independent and dependent variables by itself. There still are no standard guidelines to interpret the connection weights of ANNs, even if prior scholars attempted to reveal more insight into the connection weights of the ANN system [64-66].</td>
</tr>
<tr>
<td>2) The ANN algorithm does not require the absence of multicollinearity. Since multicollinearity impacts the performance of the regression method, in particular, the relevant information in highly collinear data sets may not be discovered by the regression approach [61]. The ANN can outperform the regression method in the case of multicollinearity.</td>
<td>Compared with regression methods, the computational cost of ANNs is relatively high. The development of the ANN model is an intensive process, which requires more computational resources than the conventional regression method.</td>
</tr>
<tr>
<td>3) The ANN algorithm can investigate implicitly any complex nonlinear relationship between independent and dependent variables [62,32,63]. Compared with ANNs, the conventional regression approach employs merely one layer to describe the relationship between independent and dependent variables. The multilayer structure enables ANNs to model more complex nonlinear relationships [12]. Hence, ANNs can be tailored to the idiosyncratic marketing research problems owing to various potential learning algorithms and activation functions.</td>
<td>The ANN requires a large number of training samples.</td>
</tr>
<tr>
<td>4) An ANN including hidden layers considers all the interplays between the independent variables. Due to the particular structure of ANNs, all the interactions between the independent variables are taken into account in the learning process of the ANN algorithm, whereas those interactions should be explicitly added to the regression model.</td>
<td>4) The ANN algorithm is susceptible to overfitting. The hidden layer structure of an ANN enables the depiction of all possible interactions of</td>
</tr>
</tbody>
</table>
independent variables. However, the superfluous high-order consideration can lead to overfitting [12]. To solve this problem, prior research suggested to reduce the number of hidden nodes. Adding a penalty term or limiting the amount of training, as well as using cross-validation, may also help avoid the overfitting problem in ANNs [67].

The ANN is a powerful and versatile tool to conduct data analysis and prediction. This data-driven approach uses the given historical or real-time data to perform a non-linear mapping between the input and output arrays. Despite the potential overfitting problem, the ANN as a universal approximator is preferable to the conventional techniques. In particular, the ANN outperforms other methods in terms of the result accuracy by dealing with complex and nonlinear relationships, for instance, in the area of medical imaging analysis [68,69], fundamental flame parameters [70], times series forecasting [13], and demand analysis [71]. To ensure high accuracy, the ANN requires usually a huge amount of data sets. Consequently, the application of the ANN is connected with high computational costs. Therefore, the ANN is less attractive for small-sized samples with simple relationships. In this case, the conventional regression method is beneficial for forecasting, for instance, in the business administration research domain [72,73].

3.2 Comparison between the self-organizing map and k-means algorithm

In this part of the work, the focus is put on the approaches dealing with the classification problems. Such problems are the main concerns of the pattern and language recognitions, image data compression [74], and robotics [75-80]. SOM and k-means algorithms are both acknowledged to handle the clustering tasks well. For both techniques, the Euclidean distance plays an important role in the clustering process. Based on the previous consideration, it can be summarized:

**Table 4. Comparison between the self-organizing map and k-means algorithm.**

<table>
<thead>
<tr>
<th>Advantage of SOM:</th>
<th>Disadvantages of SOM:</th>
</tr>
</thead>
<tbody>
<tr>
<td>1) The SOM does not predefine the number of clusters, whereas the number of clusters should be determined in advance in the k-means algorithm. The initial settings (i.e., the choices of the initial centroids) will impact the accuracy of the result of the k-means approach [81]. The result of the k-means approach is sensitive to the previous model setting.</td>
<td>1) No standard guidance of the interpretation of the SOM visualization results.</td>
</tr>
<tr>
<td>2) The SOM takes into consideration the neighborhood of the winning neuron. In the k-means algorithm, neighborhood samples are not taken into account in the clustering process.</td>
<td>2) Overfitting problem of the SOM. Due to the structure, the SOM is also prone to overfitting problems, like the ANN algorithm [60].</td>
</tr>
<tr>
<td>3) The method provides new visualization possibilities.</td>
<td>3) The theoretical groundwork in terms of network parameter settings is still incomplete. The decision about the structure of ANNs, such as the number of layers and neurons should be made by the scholars themselves [80].</td>
</tr>
<tr>
<td>4) Additionally, the SOM can project a high-dimensional data space into a low-dimensional data space [80].</td>
<td>4) The disadvantage of the k-means: this approach is based on the Euclidean distance and inclines to generate hyperspherical formed</td>
</tr>
</tbody>
</table>
clustering. For other geometric forms, the k-means has low efficiency.

The superiority of the SOM in the treatment of multiple optima problems is shown. In principle, the SOM offers the possibility of early exploration of the search space, and as the search progresses, it gradually narrows the search. At the end of the search process (if the neighborhood radius decreases to zero), the SOM is exactly equal to the k-mean, which allows minimizing the distances between the observations and the cluster centers. Therefore, it can be concluded that self-organizing maps are a time-efficient method for data sets with both high numbers of samples and high dimensionality and could be further used in the fields of, for example, customer behavior, energy consumptions diagnostics, additive manufacturing, or investigations of high-speed diffusion flames.

4 Conclusion

This paper provides a brief overview of various approaches to conducting the data analysis, ranging from conventional statistics to machine-learning algorithms, evolving from different research disciplines. The focus of this work was on the novel machine-learning algorithms and comparing them with other similar approaches, including their basic ideas, strengths, and weaknesses. In particular, the ANN vs. the regression methods and the SOM vs. the k-means algorithms are compared. The central knowledge of the different methods, their specific requirements, as well as the limitations of the individual methods, can often be an obstacle for the correct use of the methods.

As a data-driven prediction tool, ANNs are already widely applied for a large number of research areas ranging from engineering and environmental science to physics and astronomy. The flexibility of ANNs as universal approximation makes them preferable to the regression method. The challenges in this approach are to determine the appropriate structure and learning algorithm. Compared with the regression approach, ANNs can provide more accurate predictions, especially for the short-term prediction. Moreover, the regression method will easily run into its limits, if a wide range of predictors should be considered. As a data-driven cluster tool, the SOM can transfer the high-dimensional problems into a one- or two-dimensional topographic map [85]. Such dimension-reduced results can be employed for further cluster analysis using other algorithms. It has been noted that a good SOM result requires the appropriate data pretreatment [86]. Compared with the k-means algorithm, the SOM does not necessitate a priori specifications of the cluster number and the initial positions of the cluster centers.

As the computational cost of both machine-learning algorithms will increase rapidly with the increasing data size, the tradeoff between the result accuracy and the computational effort should be always taken into account. In general, data-driven techniques may pose the risk of over fitting and case-dependence. To address this issue, additional research is required. Further, future research should extend the application areas of the machine-learning algorithms, which are evidently underrepresented, such as marketing, sociology, and political science.

The results of the comparisons indicate that there is no universal approach to solve all research problems. Thus, the selection of the appropriate technique is based on the specific applications and experiment conditions. The new technology enables scholars to deal with high-complexity and high-dimensional problems. However, there remain questions referring to the efficiency of the novel algorithms:
1) How to conduct data analysis with high-dimensional features within the scope of the acceptable computational cost?
2) How to handle the outliers in a better way?
3) How to provide more insight into the analysis process of black-box models?

In summary, the aforementioned methods of machine-learning discipline enable the science to communicate to analyze complex data sets of enormous size and solve challenging big data problems on applicable time scales. Despite the complicacy, it is widely distributed and implemented due to the lack of clear and reliable guidelines for model settings and the burden of interpreting results. Nevertheless, to pursue such state-of-the-art methods, scientists must be motivated by a deeper understanding of the problems in their areas of research.

References


77. S.Kuramoto, H.Sawada, P.Hartono, Visualization of the topographical internal representation of learning robots. *International Joint Conference on Neural Networks*: 1–7(2020)