ISSN 2635-1676, Nanobiotechnology Reports, 2022, Vol. 17, No. 1, pp. 1–9. © Pleiades Publishing, Ltd., 2022.

REVIEWS

Artificial Intelligence for Nanostructured Materials

Z. M. Gadzhimagomedova[a,*], D. M. Pashkov[a], D. Yu. Kirsanova[a], S. A. Soldatov[a], M. A. Butakova[a], A. V. Chernov[a], and A. V. Soldatov[a]

[a] *The Smart Materials Research Institute, Southern Federal University, Rostov-on-Don, 344090 Russia*

** e-mail: zaira31may@gmail.com*

Received October 9, 2020; revised January 21, 2021; accepted April 11, 2021

Abstract—The current level of development of artificial intelligence (AI) technologies makes it possible to solve many complex problems just as well as a human does. Importance advances in AI are especially noticeable in machine learning, the methods and algorithms of which are successfully adapted and actively used to solve a wide range of problems, including those in the field of nanotechnology. In modern fields of nanotechnology, it is important to speed up the process of searching for the optimal synthesis parameters when creating new unique nanomaterials. The variety of approaches and techniques in both machine learning and nanotechnology makes it necessary to systematically review current data on the use of AI for solving problems in nanomaterials science at both the stage of computer design and of the chemical synthesis and diagnostics of the resulting nanomaterials. Particular attention is paid to the use of machine-learning technologies for studying the thermal and dynamic properties of nanofluids, the processes of sorption of nanocomposites, the catalytic activity of nanoparticles, and the toxicity of nanoparticles and for solving the problems of nanosensorics, as well as for processing the experimental data obtained during the diagnostics of various characteristics of nanomaterials.

DOI: 10.1134/S2635167622010049

TABLE OF CONTENTS

Introduction

1. Analysis of the Thermal and Dynamic Behavior of Nanofluids

2. Analysis of the Adsorption of Chemical Substances

3. Analysis of Various Spectra and Images of Nanoparticles

4. Problems of Nanosensorics

5. Determination of the Toxicity of Nanomaterials

6. Catalytic Activity of Nanoparticles

Conclusions

INTRODUCTION

Nanotechnologies have provided new opportunities for the development of industry, production, transportation, and telecommunications, leading to the emergence of a separate research area [1]. Studies in this area have made it possible to create entire classes of new materials with unique and practically important characteristics [2]. However, it should be noted that not all kinds of research and development activities in nanotechnology have positive aspects. For example, the production and widespread use of nanomaterials lead to problems associated, among other things, with environmental safety and make it necessary to develop and impose special normative regulations for their use [2]. Nevertheless, the benefits of introducing nanomaterials outweigh the possible risks of their use. At the same time, the development of innovative sectors of the economy reveals the constantly increasing need for the development of new nanomaterials that can solve the problems of developing environmentally friendly energy sources, for example, those based on solar energy [3]. New classes of photoactive nanocatalysts are increasingly being used [4]. Intensive research is being conducted in the development and use of two-dimensional nanomaterials, such as graphene-like structures [5]. In turn, they have a substantial effect on the development of sensors, nanoelectronics, and catalysis [6, 7]. The use of new nanomaterials in energy storage and conversion devices, including rechargeable electric batteries based on metal ions, is of no less importance [8]. The use of unique and often multistep novel technologies based on nanoporous materials from the family of metal-organic frameworks (MOFs) [9] makes it possible to synthesize nanomaterials for new generation supercapacitors [10, 11].

Increasing attention is being paid to green technologies for the synthesis of nanomaterials. For example, new nanomaterials for water purification are being synthesized with the aim of reducing the environmental impact. The production and use of such materials

are coordinated with the problems of environmental safety [12]. The use of new nanostructured materials for creating highly sensitive sensors for biomedical and food technologies is a matter of extreme urgency [13]. The synthesis of materials for nature-like technologies is a very promising new research area [14], which makes it possible to move closer, for example, to the creation of synthetic enzymes [15].

Turning to research in the field of artificial intelligence (AI), it should be noted that systematic studies in this field have been conducted for more than sixty years [16–23]. There are a significant number of AI methods that allow one to successfully solve various problems [24–35] that require substantial intellectual effort from humans. The use of AI approaches and methods in nanotechnologies gives rise to quite meaningful scientific and practical results and opens up new prospects for further progress in the area of speeding up the search for new nanomaterials. This paper presents a review of the most important examples of using AI to solve modern nanotechnology problems.

1. ANALYSIS OF THE THERMAL AND DYNAMIC BEHAVIOR OF NANOFLUIDS

In many studies, AI methods have been used to analyze the thermal behavior of nanofluids. For example, an artificial neural network (ANN) was used in [36] to study the thermal behavior of the Al_2O_3–H_2O nanofluid. The study of heat transfer in such channels is of crucial importance for the transport of fluids in the oil and gas industry. The Koo–Kleinstreuer–Li model was used to predict the behavior of a nanofluid, in which the influence of the Brownian motion of particles is taken into account. The heat-transfer rate was estimated using an ANN. The results of experiments with use of an ANN showed that heat transfer increases with an increase in the concentration of nanoparticles (NPs). At the same time, heat transfer decreases with an increase in the thermal-expansion coefficient of the liquid. Moreover, an increase in the rate of heat transfer in the flow upon the addition of NPs to the liquid was predicted.

In [37], an attempt was made to predict the heat-transfer coefficient during the boiling of a nanofluid based on Al_2O_3 by means of an ANN. As a result, it was found that the diameter of NPs, their mass concentration in the base liquid, the excess temperature (wall overheating), and the operating pressure are the best independent variables for assessing the parameter under consideration. The study results published in [37] show that a feedforward multilayer perceptron comprised of twelve hidden neurons (structure 4-12-1) is the best model for assessing the heat-transfer coefficient during the boiling of a nanofluid based on Al_2O_3. In this case, the coefficient of determination equals $R^2 = 0.9929$.

In [38], the viscosity of the TiO_2–H_2O nanofluid was simulated using an ANN and established that ANNs make it possible to determine with high accuracy the changes in the behavior of the dynamic viscosity with changes in the temperature and the mass fraction of NPs. An ANN with one hidden layer and four neurons was used. As a result of simulation, a regression coefficient of 0.9998 was obtained, which indicates a very high prediction accuracy with a fairly simple ANN structure. Moreover, the nanofluid viscosity is predicted from data on the mass fraction and temperature. This correlation allows one to estimate the viscosity of the TiO_2–H_2O nanofluid in a wide range of mass fractions of NPs with a maximum error of 0.5%.

The study of the thermal conductivity of Fe_3O_4 magnetic nanofluids is of considerable interest [39]. Nanofluid samples were prepared using a two-stage method by dispersing Fe_3O_4 NPs in water with fractions of 0.1, 0.2, 0.4, 1, 2, and 3 vol % of solids. Thermal-conductivity measurements were carried out using a KD2 Pro thermal analyzer over a temperature range of 20 to 55°C. Taking into account the experimental data, a correlation was proposed to predict the thermal conductivity of a magnetic nanofluid. In the final stage, an optimal ANN was developed to predict the thermal conductivity of a magnetic nanofluid. The experiments showed that the maximum increase in the thermal conductivity of the nanofluid was about 90% with a fraction of 3 vol % of solids at a temperature of 55°C. Comparative results show that there are deviations of 5% from the experimental data in the case of using the proposed correlation and the deviation for an ANN is 1.5%.

In [40], the thermal conductivity of Al_2O_3 NPs in a water (40%)/ethylene glycol (60%) solution was determined. Using the results of temperature measurements and volume-fraction data, an empirical relationship has proposed. Further data analysis was conducted on a model that uses an ANN in the form of a feedforward multilayer perceptron. A structure with two hidden layers and five neurons in the first and second layers was chosen for data analysis. The results showed that ANNs can accurately assess the experimental thermal-conductivity data for Al_2O_3/water (40%)/ethylene glycol (60%) nanofluids.

The results of several experiments were used to determine the thermal conductivity of an aqueous nanofluid on multiwalled carbon nanotubes (MWNTs) containing the COOH functional group [41]. For this purpose, COOH MWCNT nanoparticles were dispersed in water by different methods. The thermal conductivity was measured for various concentrations of MWCNTs (up to 1%) at temperatures of 25–55°C. A multilayer ANN was used to simulate the thermal conductivity of this nanofluid. The temperature and volume fraction of solids were input network variables, and the thermal conductivity was an output

network variable. The root-mean-square (rms) error was 4.04×10^{-6}, which indicates a high degree of applicability of the ANN for predicting the thermal conductivity of the COOH MWCNT nanofluid.

The thermal conductivity of the ZnO–EG (ethylene glycol) nanofluid was also studied experimentally and with use of an AI [42]. For this purpose, zinc-oxide NPs with a nominal diameter of 18 nm were dispersed in EG in various volume fractions at different temperatures (24–50°C). Using the analysis of experimental data, an experimental model based on the temperature dependence of the solid concentration was proposed. Next, a feedforward multilayer ANN was used to simulate the thermal conductivity of the ZnO–EG nanofluid. Twenty eight of the forty measurements in total obtained during the experiments were selected for network training, and the remaining twelve measurements were used for network testing and validation. The results showed that the prediction made by an ANN is in good agreement with the experimental data.

In [43], the thermal conductivity of a hybrid nanofluid was studied experimentally. The nanofluid under study was obtained using a two-stage method by dispersing Cu and TiO_2 NPs with an average diameter of 70 and 40 nm in a water/EG (60 : 40) binary mixture. The properties of this nanofluid were measured for various concentrations of solids (0.1, 0.2, 0.4, 0.8, 1, 1.5, and 2 vol %) at temperatures in the range from 30 to 60°C. Two new correlations were then proposed to predict the thermal conductivity of the investigated hybrid nanofluid as a function of the concentration of solids and temperature. They use an ANN and are based on experimental data. The results showed that these two new models are capable of predicting the thermal conductivity and agree well with the experimental results.

Using AI methods, the entropy of nanofluids was also investigated. As was established in [44], entropy generation can be simulated by an ANN in terms of the particle fraction and heat flux. First of all, the characteristics of the first and second laws of thermodynamics, including the coefficient of convective heat transfer, the entropy-generation rate, and the Bejan number, were studied for a hybrid nanofluid containing graphene–platinum NPs. It was found that the heat-transfer coefficient and friction-entropy generation increase with an increase in the concentration of particles, while the thermal-entropy generation decreases. In addition, the rate of entropy generation due to heat transfer increases with an increase in the thermal load, while entropy generation due to friction decreases.

In [45], the thermal-conductivity coefficient and the dynamic-viscosity ratio of the Al_2O_3–H_2O nanofluid were accurately predicted using the methods of regression analysis of a Gaussian process. The temperature, volume fraction, and size of the NPs were used as input predictor variables. Two hundred and twenty two sets of experimental data were taken to predict the thermal-conductivity coefficient, dynamic-viscosity coefficient, and the effectiveness of predictor variables for predicting response variables. It was found that temperature is a decisive factor for improving the accuracy in determining the thermal-conductivity coefficient. It was also found that optimization of the regression predictor for a Gaussian process with the covariance function of the Matérn kernel shows very good agreement with experimental data on the thermal-conductivity coefficient with a rms error of 0.000126, and the square of the exponential kernel function shows good agreement with experimental data on the dynamic-viscosity coefficient with a rms error of 0.000045. The regression-coefficient value is 0.99 (closer to one) and, consequently, the predicted results are reliable. Thus, machine-learning methods enable highly accurate prediction of the thermodynamic characteristics of the behavior of nanofluids (which are very complex systems), such as thermal conductivity, viscosity, and entropy dynamics. The variety of types of nanofluids and the need to use them under various conditions make it difficult to use traditional and purely empirical methods for their study.

2. ANALYSIS OF THE ADSORPTION OF CHEMICAL SUBSTANCES

The study of the process of adsorption of certain NPs is another interesting application of machine learning in the field of nanotechnology. For example, in an experimental study [46], zinc-sulfide NPs deposited onto activated carbon (ZnS–NP–AC) were synthesized and used as a sorbent for the selective ultrasonic removal of brilliant green dye (BG) from aqueous solutions. Two machine-learning approaches were applied: an ANN with a radial basis function and the random forest method. These two approaches were evaluated on the basis of a quadratic model of the surface response to predict the maximum efficiency of BG removal from the ZnS–NP–AC aqueous medium. The good results achieved showed that this approach can be used in the future as a method for predicting the effectiveness of purification of contaminated water sources from other toxins as well.

In [47], the Box–Wilson method in combination with the desirability-function approach provided useful information on the simultaneous ultrasonic removal of BG and eosin B with ZnS–NP–AC nanoparticles. A multilayer ANN model optimized with the Levenberg–Marquardt algorithm was used to predict the effectiveness of the removal of BG and eosin B. The ANN could effectively predict the simultaneous removal of BZ and eosin B. The studies showed good agreement between the experimental data and the results obtained by the ANN. Being the best model for fitting experimental data, the Langmuir model reveals a fairly effective adsorption capacity

(258.7 and 222.2 mg/g) of the nanosystems under consideration, which confirms the applicability of the Langmuir model for selecting the best sorbents for wastewater treatment.

In [48], the properties of γ-Fe_2O_3 NPs deposited onto activated carbon were studied. The prepared nanomaterial was examined by field-emission scanning electron microscopy, energy dispersive X-ray spectroscopy, Fourier transform IR (FTIR) spectroscopy, and X-ray diffraction. Activated carbon with γ-Fe_2O_3 nanoparticles (γ-Fe_2O_3–NP–AC) was used as a novel adsorbent for the ultrasonic removal of methylene blue (MB) and malachite green (MG). The ANN and surface-response techniques were applied to simulate and optimize the adsorption of MB and MG in individual and binary solutions, and then to study the adsorption isotherm and kinetics. The individual effects of parameters, such as pH, the mass of the adsorbent, the sonication time, and the concentrations of MB and MG, were investigated in addition to the effect of possible interactions on the adsorption process. The pseudo-second order model proved to be applicable in studying adsorption kinetics. No degradation of the dyes was observed in the experiments without any adsorbent, which were performed to examine the possible decomposition of the dyes under conditions of ultrasonic treatment.

In [49], similar approaches were used to study MnO_2 NPs (MnO_2–NP–AC) deposited onto AC. This structure is an efficient, environmentally friendly, and cost-effective adsorbent. It has been synthesized and characterized using techniques, such as field-emission scanning electron microscopy, energy dispersive X-ray spectroscopy, Brunauer–Emmett–Teller (BET) surface analysis, single-crystal X-ray diffraction, and FTIR spectroscopy. The fast and simultaneous ultrasonic-assisted adsorption of BG, crystal violet (CV), and MB dyes with strong spectral overlap on the MnO_2–NP–AC system as a new and efficient adsorbent was investigated. An ANN was used to accurately predict the percentage of removal of dyes from their ternary solution by MnO_2–NP–AC adsorbent. Experimental equilibrium data were simulated using various isotherm models. The Langmuir model turned out to be most suitable for describing the experimental equilibrium data obtained under optimal conditions. A small amount of the MnO_2–NP–AC adsorbent (0.005 g) with a high adsorption capacity in a single-component system (206.20, 234.20, and 263.16 mg g^{-1} for BG, CV, and MB, respectively) was successfully used to remove dyes in a very short time (4 min). The study of the kinetics of the adsorption process showed the applicability of the second-order kinetic model.

Various AI methods such as those based on a multilayer perceptron, an adaptive network based on a fuzzy inference system optimized by a genetic algorithm, genetic programming, and a committee machine with intelligent inference, have been used to predict the sorption of gases, such as C_3H_8, H_2, CH_4, and CO_2, in nanocomposite membranes selective to hydrogen H_2 [50]. These matrices consisted of NPs of porous zeolite 4A as a dispersed phase and a polymeric (polydimethylsiloxane) matrix as a continuous phase. In this case, the sorption was evaluated considering the influence of the NP load, the critical temperature (gas characteristics), and the inlet pressure. The data obtained during the study were randomly divided into two parts: data sets for training (75% of experimental data) and for testing (25% of experimental data). As a result, it was established that the committee-machine method with intelligent inference gives more accurate results compared to other models.

A systematic analysis of the adsorption of CO on Pt nanoclusters with a size in the range of 0.2–1.5 nm was carried out in [51]. Such studies are necessary to identify trends in the effects of the size and morphology of clusters on an adsorbate binding and to develop models for predicting the adsorption of a substance at a certain site in the crystal lattice. It was shown that use of the genetic algorithm based on the empirical potential and simulation by the density-functional-theory method is associated with the existence of a size window of 40 to 70 atoms, in which Pt nanoclusters weakly bind CO and the binding energies are comparable to the binding energies at the (111) or (100) faces. In [51], machine-learning algorithms with several descriptors—in particular, the gradient-descent algorithm—were used. To train the algorithm and to ensure the accuracy of its predictions, the energies of adsorption of CO on different parts of the surface of Pt clusters were used as target data.

A similar problem was attempted to be solved in [52] by studying the morphology of clusters with high symmetry. However, real NPs form low-symmetry structures rather than high-symmetry ones. This fact was taken into consideration in the study described above.

In [53], the best conditions were studied for the removal of benzene, toluene, ethylbenzene, and xylene (BTEX) compounds from aqueous solutions with use of magnetic nanosorbents, and relationships were established between the removal efficiency of these compounds and variations in the contact time, the initial concentration of the BTEX mixture, the adsorbent dose, stirring rate, pH, and temperature. Moreover, the ANN comprised of three layers (input, hidden, and output layers) was optimized to find a correlation between the studied parameters and the efficiency of BTEX removal. The network structure used for prediction was composed of six experimental parameters (temperature, stirring rate, initial BTEX concentration, contact time, pH, and adsorbent dose) at the input, ten neurons in the hidden layer, and one parameter at the output. The simulation results showed that the ANN with a mean absolute error of

0.6272% reliably describes the adsorption of BTEX at iron nanoparticles.

In [54], a three-layer ANN with the backpropagation algorithm, genetic algorithm, and the particle swarm optimization (PSO) method were used to optimize and predict optimal conditions and to obtain the maximum efficiency of removal of the CV dye from aqueous solutions with the use of bimetallic Fe/Ni nanoparticles. Moreover, it was shown that the ANN architecture with three neurons in the hidden layer turned out to be the optimal topology for machine learning and the value $R^2 = 0.9998$ obtained for the ANN model with the backpropagation algorithm points to a high accuracy of the prediction. The absolute errors between the predicted and experimental results were 5.6% for the genetic algorithm and 3.5% for the PSO method.

3. ANALYSIS OF VARIOUS SPECTRA AND IMAGES OF NANOPARTICLES

The identification and enumeration of individual NPs is an important part of nanotechnology research. In [55], deep-machine-learning methods showed a high efficiency in the automatic recognition of platinum NPs deposited onto highly oriented pyrolytic graphite from images obtained by scanning tunneling microscopy (STM). A neural network called CascadeRCNN was used. Training was conducted using a data set comprised of ten STM images of NPs. Five images containing 2052 nanoparticles were used for verification. As a result, the trained ANN recognized the nanoparticles in the control set with an accuracy of 50.8%, which is not very high. The nanoparticles had clear contours, which is necessary for further determination of their size parameters (width, height, etc.).

The development of powerful light microscopes capable of generating terabytes of high-resolution 2D and 3D video images in a single day has created a strong demand for automated image-analysis techniques. Tracking the movement of nanoscale particles (such as viruses, proteins, and synthetic-drug particles) is critical to understand how pathogens disrupt mucosal barriers, as well as to develop new drugs. Current methods for tracking such particles are based on a limited set of input parameters for identifying bright objects and are ill-suited to process the spectrum of spatial-temporal fluctuations due to a poor signal-to-noise ratio typically associated with submicrometer entities in complex biological environments. The optimization and implementation of tracking methods often requires the extensive involvement of users, which is not only inefficient, but also leads to subjective assessments. To create a fully automated tracking method, a convolutional neural network for the localization of particles from image data, which contained more than 6000 parameters, was developed in [56]. A neural-network tracker provides automation and a high accuracy with exceptionally low levels of false positive and false negative results for both 2D and 3D video simulations, and for experimental 2D videos with views that are difficult to track.

The properties of monometallic and bimetallic NPs can strongly depend on the compositional and structural (or geometric) characteristics and their dynamics. These parameters can be effectively described using the partial radial-distribution function (PRDF) of atoms in the material. For NPs with a size of several nanometers, finite size effects can play a substantial role in determining the crystal order, interatomic distances, and particle shape. Bimetallic NPs can also have a different distribution of the elemental composition compared to bulk materials. All the above factors make it difficult to determine the PRDF. Extended X-ray absorption fine structure (EXAFS) spectroscopy, molecular-dynamics simulation, and supervised machine learning were combined in [57] to extract the PRDF directly from experimental data. The application of this approach to several systems of Pt and PdAu nanoparticles demonstrated the effects of the finite size of NPs on the distribution of nearest neighbors, the dynamics of bonds, and distribution patterns of doping atoms in monometallic and bimetallic particles. In addition, the general applicability of ANNs for solving this class of problems was also shown.

Machine-learning methods were also used to analyze surface-enhanced Raman scattering (SERS) spectroscopy data [58]. In [58], a database was collected and analyzed for silicon-coated silver nanoparticles by the SERS method with the subsequent training of a deep neural network. As a proof of the concept, three types of representative tumor suppressor genes—namely, fragments p16, p21, and p53 that are easily distinguished without markers—were identified. The distinguishable and reproducible SERS spectra of these DNA molecules were collected and used as input data for learning and training a deep ANN, which made it possible to selectively recognize DNA targets. It was shown that the level of recognition accuracy of a particular DNA target in this approach reaches 90.28%.

Dynamic SERS spectroscopy, which was used to quickly detect acephate in rice in the range of 100.2–0.5 mg/L by means of simple height-adjustable gold nanorods, was also studied. Multivariate machine learning and deep-learning methods were used to develop regression models for automatic analysis of the quantitative level of acephate residues. Partial least squares regression made it possible to achieve optimal performance with a standard deviation of 5.4776 and a coefficient of determination of 0.9560.

It should be noted that many methods have been developed to estimate the size and shape of NPs from transmission-electron-microscopy images. However, some of them have limiting characteristics, such as being difficult to use, expensive, or having steps that

make the results less reliable. For example, manual and automatic methods of analysis were compared in [59]. Gold nanoparticles were synthesized by the traditional Turkevich method using sodium citrate. Twenty-three images of Au NPs were used. The study results indicated that conglomerated particles could be excluded provided that a sufficient number of particles were included in the analysis. It has been shown that the CellProfiler software package can be very effective in studying the size and shape of NPs. These results demonstrate the benefits of using an open-source software package that can be applied to analyze data by machine learning and, consequently, make it possible to automate data processing when studying NPs.

In [60], the methods of principal component analysis and independent component analysis were used to analyze the energy-dispersive X-ray spectrum of core–shell NP clusters with the aim of determining the chemical composition of the samples. As a result of the study, the number of phases in the analyzed volume (core, shell, and substrate) was accurately determined and their spectral components were also identified. Moreover, the proposed approach divides the spectrum into three components, which precisely represent the isolated and unmixed X-ray signals coming from the carbon film, the iron-oxide shell, and the platinum–iron core. The results were also confirmed by comparing the calculated spectra from the bimetallic core and shell with the spectra obtained experimentally from these structures independently.

In [61], a set of electron-microscopy images containing individual NPs that were ordered on the surface in the direction of the formation of geometric patterns is described. This dataset covers the following three levels of nanoscale organization: individual NPs (1–5 nm) and NP arrays (5–20 nm); ordering effects (20–200 nm); complex patterns ranging in size from a few nanometers to a few micrometers. The described dataset provided, for the first time, an opportunity to develop machine-learning algorithms for studying the unique phenomena of the ordering of NPs and their hierarchical organization.

As was shown in [62], the evolution of proteins on the surface of NPs predicts the further biological fate of NPs in vivo. A library of the mass spectra of proteins served as input data and the clearance of blood and its accumulation in organs as output data were used to train a supervised deep neural network that predicted the biological fate of NPs. In a double-blind study, a network that predicted the accumulation of NPs in the spleen and liver with an accuracy of 94% was tested.

4. PROBLEMS OF NANOSENSORICS

There are a large number of biosensors that are used to detect various chemicals, including nanochemicals. For example, the air exhaled by a person contains more than 3 thousand volatile organic compounds (VOCs), at least 15 of which are directly or indirectly associated with internal biochemical processes in the body. Electronic sensors, such as the so-called "electronic nose" [63], already exist, which can play an important role in screening various respiratory and systemic diseases by analyzing exhaled components. The electronic nose combines a sensor array and an ANN that responds to specific VOCs and, consequently, can operate as a noninvasive technology for monitoring diseases. Lung cancer is one of the most feared types of cancer with a very high mortality rate. Standard methods, such as sputum cytology, chest X-ray, or computed tomography cannot be used in large-scale population screening. Breath products contain markers for lung cancer, diabetes, and markers for several other diseases, which can instantly signal the presence of a particular disease upon reaching the electronic nose. In the electronic-nose system, the output data of the sensor array is converted into an electrical signal. Each sensor provides a dynamic response in the form of an electrical signal through gravimetric measurements, standardization, and normalization, which make it suitable for statistical analysis. In general, analysis of the data pattern falls into the following three categories: graphical, multivariate, and neural-network analyses. Graphical analysis is among the simplest forms of data analysis, in which the response of sensors is analyzed using histograms or polar-diagram methods. These kinds of graphical analyses are useful for the visual interpretation of a chemical signature. However, the graphical analysis is useless when using high dimensional data. Multivariate data analysis is used to reduce the dimensionality of data that relate to the reactions of complex chemical compounds by eliminating redundant variables, such as the temperature and humidity of the environment. There are many methods for multivariate analysis, such as principal-component analysis (PCA), linear-discriminant analysis, discriminant-function analysis, canonical discriminant analysis, and partial least-squares regression, which are used in the electronic-nose system. In this system, PCA is frequently used to reduce the dimensionality. Principal-component analysis is an unsupervised learning method that allows one to identify chemicals without prior knowledge of the full list of substances. However, the main purpose of the electronic nose is the qualitative and quantitative determination of chemicals. Therefore, there is also a need for classification methods, such as the support-vector machine, the *k*-nearest neighbors algorithm, and the use of an ANN. If the results of an ANN and multivariate analysis show good agreement, then the precision of chemical identification improves.

Studies of cell-penetrating peptides (CPPs) that facilitate the transport of pharmacologically active molecules, such as plasmid DNA, short interfering RNAs, nanoparticles, and small peptides, are of no less interest. The precise determination of new and

unique CPPs is the first step to understanding their activity. Experiments can give a detailed idea about the ability of CPPs to penetrate cells. However, the synthesis and identification of CPPs in wet chemistry experiments require large investments of resources and time. Therefore, the development of an effective prediction tool is essential for identifying unique CPPs. With this aim in mind, a CPP prediction model was developed based on the KELM-CPPpred kernel of an extreme learning machine [64]. The main dataset used in [64] consisted of 408 CPPs and the same number of substances that are not CPPs. The input functions used to train the proposed predictive model included the composition of amino acids, the amino-acid composition of the dipeptide, the pseudo-amino-acid composition, and hybrid features based on motifs. Next, an independent data set was used to test the proposed model. Empirical tests have shown that the KELM-CPPpred kernel outperforms the existing approach to prediction based on random forests.

5. DETERMINATION OF THE TOXICITY OF NANOMATERIALS

It is well known that NPs can cause harmful effects on various biological systems and their ecosystems. Toxicological analysis is a very important step in evaluating the potential risks associated with NPs, but experimental tests are often very expensive and usually too slow to determine the amount of NPs that may cause harmful effects. In-silico models based on quantitative structure–activity/toxicity relationships are alternative tools that have become valuable supports to risk assessment, which rationalize the search for safer NPs. In [65], a unified model of the quantitative structure–property relation (QSPR) was developed on the basis of an ANN. It aims to simultaneously predict the overall toxicity profiles of NPs under different experimental conditions. The model uses 54371 pairs of NPs generated by applying perturbation theory to a set of 260 unique NPs and has shown an accuracy of over 97% for both training and validation sets. Next, the QSPR perturbation model was used to predict the toxic effects of several NPs not included in the original data set. The theoretical results obtained for this independent set fully agree with the experimental data; therefore, one can assume that the present QSPR perturbation model can be considered as a promising and reliable computational tool for studying the toxicity of NPs. It should be noted that the development of such quantitative structure–property relationship models for nanomaterials (nano-QSPR) requires meticulous data collection and processing. For example, a list of open-source nano-QSPR models that were developed on the basis of an online chemical modeling environment called OCHEM is given in [66]. A variety of data on the toxicity of NPs to different living organisms were collected from academic papers and uploaded to the OCHEM database. The main characteristics of NPs, such as the chemical composition of NPs, the mean size, shape, and surface charge of particles, and information about the tested biological samples were used as descriptors for the development of QSPR models. Random forests and associative neural networks were used as models in the QSPR.

6. CATALYTIC ACTIVITY OF NANOPARTICLES

A good example of studying the catalytic activity of nanocomposites is published in [67], in which a machine-learning model for comparing the catalytic efficiency of Pt nanocrystals with structural features of NPs, such as the diameter, surface area, sphericity, face configuration, and type of surface defects, is described. A theoretical dataset of more than three hundred thousand NPs was used. In this case, the most important combinations of only two or three functions (molar catalytic activity, selectivity, and thermodynamic stability) that have an effect on the catalytic efficiency were investigated using a decision tree and an ANN. It was established that extremely accurate predictions are obtained when all functions are simultaneously used to train ANN models, and the models constructed for nonsymmetrical NPs predict catalytic efficiency and stability with an accuracy of better than 0.93. Moreover, the effect of temperature on the catalytic efficiency is also predicted quite accurately for more than 300 thousand samples by adding an additional input neuron to the ANN architecture.

The development of active sites of the catalyst is the key to creating high performance heterogeneous catalysts. The rate of a catalytic reaction can be easily predicted by simulating the arrangement of surface atoms with well-defined single-crystal surfaces [68]. However, this method has limitations in the case of highly inhomogeneous atomic configurations, such as NPs of alloys with defects at the atomic scale, at which the structure cannot be decomposed into single crystals. In [69], a universal machine-learning scheme based on the local similarity kernel method is published. The proposed scheme makes it possible to study catalytic activity using local atomic configurations. The algorithm consists of the following two steps: studying the density functional theory (DFT) data from the single-crystal surface and extrapolating them to the entire NP. The DFT calculations were compiled into a training set. The data consisted of geometric information about the unrelaxed regions of the single-crystal surface. Next, this approach was applied to study the direct decomposition of NO on $Rh_{(1-x)}Au_x$ nanoparticles. It made it possible to effectively predict the thermochemistry of catalytic reactions using the DFT data of single crystals, and the combination with kinetics analysis was able to provide detailed information about the structures of active sites and the catalytic activity, which depends on the size and composition of nanocatalysts.

CONCLUSIONS

Artificial-intelligence technologies, especially machine learning, allow us to analyze the hidden relationships between the structure and properties of materials at the nanoscale. They represent a fast, highly efficient, and resource-saving computational tool for studying the parameters of nanoparticles, as well as for predicting the possible characteristics of nanomaterials prior to their synthesis. Machine learning opens up new opportunities for solving the problems of the thermal and dynamic behavior of nanofluids, the adsorption of chemicals, diagnostics of the results of NP imaging, and nanosensorics. Artificial neural networks are successfully used to classify large data on the spectra and images of NPs. Evaluation of the toxicity of nanomaterials, prediction of the behavior of NPs in vivo, and determination of the NP-surface chemical compositions optimal for their introduction into the body are performed faster and more precisely. Important characteristics, such as the heat-transfer coefficient, thermal conductivity, and the dynamic-viscosity ratio of the Al_2O_3–H_2O nanofluid are predicted quite successfully. Moreover, the adsorption values of substances on ZnS–NP–AC, γ-Fe_2O_3–NP–AC, and MnO_2–NP–AC are accurately predicted. The study results can be used to develop methods for purifying polluted water sources from various toxins. The use of ANNs in solving the problems of determining the catalytic activity of NPs gives substantial results in the research field of structures with defects and inhomogeneities.

FUNDING

This study was supported by the Ministry of Science and Higher Education of the Russian Federation within State assignment no. 0852-2020-0019 in the field of scientific activity.

CONFLICT OF INTEREST

We declare that we have no conflicts of interest.

REFERENCES

1. B. Bhushan, *Encyclopedia of Nanotechnology* (Springer, Netherlands, 2012).
2. J. Jeevanandam, A. Barhoum, Y. S. Chan, et al., Beilstein J. Nanotechnol. **9**, 1050 (2018). https://doi.org/10.3762/bjnano.9.98
3. S. Kundu and A. Patra, Chem. Rev. **117**, 712 (2017). https://doi.org/10.1021/acs.chemrev.6b00036
4. E. A. Konstantinova, M. P. Kushnikov, V. B. Zaitsev, V. G. Kytin, A. V. Marikutsa, G. V. Trusov, A. S. Sedegov, and P. K. Kashkarov, Nanotechnol. Russ. **14**, 190 (2019). https://doi.org/10.1134/S1995078019030078
5. C. Cheng, S. Li, A. Thomas, et al., Chem. Rev. **117**, 1826 (2017). https://doi.org/10.1021/acs.chemrev.6b00520
6. C. Tan, X. Cao, X. J. Wu, et al., Chem. Rev. **117**, 6225 (2017). https://doi.org/10.1021/acs.chemrev.6b00558
7. H. Jin, C. Guo, X. Liu, et al., Chem. Rev. **118**, 6337 (2018). https://doi.org/10.1021/acs.chemrev.7b00689
8. Y. Zhao, L. P. Wang, M. T. Sougrati, et al., Adv. Energy Mater. **7**, 1901260 (2017). https://doi.org/10.1002/ente.201901260
9. V. V. Butova, M. A. Soldatov, A. A. Guda, K. A. Lomachenko, and C. Lamberti, Russ. Chem. Rev. **85**, 280 (2016). https://doi.org/10.1070/RCR4554
10. S. Dang, Q.-L. Zhu, and Q. Xu, Nat. Rev. Mater. **3**, 17075 (2018). https://doi.org/10.1038/natrevmats.2017.75
11. R. R. Salunkhe, Y. T. Kaneti, and Y. Yamauchi, ACS Nano **11**, 5293 (2017). https://doi.org/10.1021/acsnano.7b02796
12. M. Nasrollahzadeh, M. Sajjadi, S. Iravani, and R. S. Varma, J. Hazard. Mater. **401**, 123401 (2021). https://doi.org/10.1016/j.jhazmat.2020.123401
13. R. Gupta, N. Raza, S. K. Bhardwaj, et al., J. Hazard. Mater. **401**, 123379 (2021). https://doi.org/10.1016/j.jhazmat.2020.123379
14. R. R. Nasaruddin, T. Chen, Q. Yao, et al., Coord. Chem. Rev. **426**, 213540 (2021). https://doi.org/10.1016/j.ccr.2020.213540
15. J. Wu, X. Wang, Q. Wang, et al., Chem. Soc. Rev. **48**, 1004 (2019). https://doi.org/10.1039/D1TB00964H
16. E. A. Patrick and F. P. Fisher, Inf. Control. **16**, 128 (1970).
17. J. E. Laird, Ch. Lebiere, and P. S. Rosenbloom, AAAI **1**, 26 (2017). https://doi.org/10.1609/aimag.v38i4.2744
18. R. Calegari, G. Ciatto, E. Denti, and A. Omicini, Information **11**, 1 (2020). https://doi.org/10.3390/info11030167
19. L. Perez Cruz and D. Treisman, **1**, 144 (2018). doi
20. V. N. Vagin, *Knowledge and Conviction in Data Mining* (Fizmatlit, Moscow, 2019) [in Russian].
21. M. van Gerven, Front. Comput. Neurosci. **11**, 112 (2017). https://doi.org/10.1101/166785
22. A. K. Pogodaev, Vestn. Voronezh. Tekh. Univ. **11**, 1 (2015).
23. M. I. Anchekov, V. V. Bova, O. V. Nagoeva, et al., Izv. Kab.-Balk. Nauch. Tsentra RAN **5**, 24 (2015).
24. J. R. Quinlan, Mach. Learn. **1**, 81 (1986).
25. H. Drucker, C. J. C. Burges, L. Kaufman, et al., in *Proceedings of the Conference on Support Vector Regression Machines, NIPS, 1996,* Ed. by M. C. Mozer et al.
26. C. Cortes and V. Vapnik, Mach. Learn. **20**, 273 (1995).
27. P. Langley, W. Iba, and K. Thompson, in *An Analysis of Bayesian Classifiers, Proceedings of the 10th National Conference on Artificial Intelligence* (1992), p. 223.
28. W. Pitts, Bull. Math. Biophys. **5**, 115 (1943).

29. D. Opitz and R. Maclin, J. Artif. Intell. Res. **11**, 169 (1999). https://doi.org/10.1613/jair.614
30. L. Brieman, Mach. Learn. **45**, 5 (2001). https://doi.org/10.1023/A:1010933404324
31. P. Geurts, D. Ernst, and L. Wehenkel, Mach. Learn. **63**, 3 (2006). https://doi.org/10.1007/s10994-006-6226-1
32. R. E. Schapire, in *Proceedings of the 16th International Joint Conference on Artificial Intelligence* (1999), Vol. 11, p. 169.
33. T. Chen and C. Guestrin, "XGBoost: A scalable tree boosting system," in *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, Series XGBoost* (2016), p. 785.
34. W. A. Barbakh, Y. Wu, and C. Fyfe, *Non-Standard Parameter Adaptation for Exploratory Data Analysis,* Vol. 249 of *Studies in Computational Intelligence* (Springer, Berlin, 2009), p. 7. https://doi.org/10.1007/978-3-642-04005-4
35. J. MacQueen, in *Proceedings of the 5th Berkeley Symposium on Mathematical Statistics and Probability* (1967), Vol. 1, p. 281.
36. M. Sheikholeslami, M. B. Gerdroodbary, R. Moradi, et al., Comput. Method Appl. Mech. Eng. **344**, 1 (2019). https://doi.org/10.1177/0954408919878984
37. M. Hassanpour, B. Vaferi, and M. E. Masoumi, Appl. Therm. Eng. **128**, 1208 (2018). https://doi.org/10.1615/JEnhHeatTransf.2019031660
38. M. Hemmat Esfe, M. R. Hassani Ahangar, M. Rejvani, et al., Int. Commun. Heat Mass. **75**, 192 (2016). https://doi.org/10.1016/j.icheatmasstransfer.2016.04.002
39. M. Afrand, D. Toghraie, and N. Sina, Int. Commun. Heat Mass. **75**, 262 (2016).
40. M. Hemmat Esfe, W.-M. Yan, M. Afrand, et al., Int. Commun. Heat Mass. **74**, 125 (2016).
41. M. Hemmat Esfe, A. Naderi, M. Akbari, et al., J. Therm. Anal. Calorim. **121**, 1273 (2015).
42. M. Hemmat Esfe, S. Saedodin, A. Naderi, et al., Int. Commun. Heat Mass. **63**, 35 (2015). https://doi.org/10.1007/s10973-017-6744-z
43. M. Hemmat Esfe, S. Wongwises, A. Naderi, et al., Int. Commun. Heat Mass. **66**, 100 (2015). https://doi.org/10.1016/j.icheatmasstransfer.2015.05.014
44. R. Khosravi, S. Rabiei, M. Bahiraei, and A. R. Teymourtash, Int. Commun. Heat Mass. **109**, 104351 (2019). https://doi.org/10.1016/j.icheatmasstransfer.2019.104351
45. P. C. Mukesh Kumar and R. Kavitha, Heliyon **6**, e03966 (2020). https://doi.org/10.1016/j.heliyon.2020.e03966
46. Ahmadi M. H. Azqhandi, M. Ghaedi, F. Yousefi, and M. Jamshidi, J. Colloid Interface Sci. **505**, 278 (2017). https://doi.org/10.1016/j.jcis.2017.05.098
47. M. Jamshidi, M. Ghaedi, K. Dashtian, et al., Spectrochim. Acta, A **153**, 257 (2016). https://doi.org/10.1016/j.saa.2015.08.024
48. A. Asfaram, M. Ghaedi, S. Hajati, and A. Goudarzi, Ultrason. Sonochem. **32**, 418 (2016). https://doi.org/10.1016/j.ultsonch.2016.04.011
49. A. Asfaram, M. Ghaedi, S. Hajati, and A. Goudarzi, RSC Adv. **5**, 72300 (2015). https://doi.org/10.1039/C5RA10815B
50. A. Dashti, H. R. Harami, and M. Rezakazemi, Int. J. Hydrogen Energy **43**, 6614 (2018). https://doi.org/10.1007/s11814-019-0330-y
51. R. Gasper, H. Shi, and A. Ramasubramaniam, J. Phys. Chem. C **121**, 5612 (2017). https://doi.org/10.1021/acs.jpcc.7b08686
52. L. Li, A. H. Larsen, N. A. Romero, et al., J. Phys. Chem. Lett. **4**, 222 (2013). https://doi.org/10.1021/jz3018286
53. S. A. Abdel-Gawad, M. K. Mostafa, and A. S. Mahmoud, Water Supply **18**, 1650 (2018). https://doi.org/10.1177/11786221211028162
54. W. Ruan, J. Hu, J. Qi, et al., Materials (Basel) **11**, 865 (2018). https://doi.org/10.3390/ma11050865
55. A. G. Okunev, A. V. Nartova, and A. V. Matveev, in *Proceedings of the International Multi-Conference on Engineering, Computer and Information Sciences SIBIRCON, 2019,* p. 940.
56. J. M. Newby, A. M. Schaefer, P. T. Lee, et al., Proc. Natl. Acad. Sci. U. S. A. **115**, 9026 (2018). https://doi.org/10.1073/pnas.1804420115
57. J. Timoshenko, C. J. Wrasman, M. Luneau, et al., Nano Lett. **19**, 520 (2019). https://doi.org/10.1021/acs.nanolett.8b04461
58. H. Shi, H. Wang, X. Meng, et al., Anal. Chem. **90**, 14216 (2018).
59. G. P. Ribeiro, R. S. Valotto, J. P. de Oliveira, et al., Chem. Papers **74**, 2821 (2020). https://doi.org/10.1007/s11696-020-01123-3
60. D. Rossouw, P. Burdet, F. de la Pena, et al., Nano Lett. **15**, 2716 (2015). https://doi.org/10.1021/acs.nanolett.5b00449
61. D. A. Boiko, E. O. Pentsak, V. A. Cherepanova, and V. P. Ananikov, Sci. Data **7**, 101 (2020). https://doi.org/10.6084/m9.figshare.11926929
62. J. Lazarovits, S. Sindhwani, A. J. Tavares, et al., ACS Nano **13**, 8023 (2019). https://doi.org/10.1021/acsnano.9b02774
63. B. Behera, R. Joshi, G. K. Anil Vishnu, et al., J. Breath Res. **13**, 024001 (2019). https://doi.org/10.1088/1752-7163/aafc77
64. P. Pandey, V. Patel, N. V. George, and S. S. Mallajosyula, J. Proteome Res. **17**, 3214 (2018). https://doi.org/10.1021/acs.jproteome.8b00322
65. R. Concu, V. V. Kleandrova, A. Speck-Planche, and M. Cordeiro, Nanotoxicology **11**, 891 (2017). https://doi.org/10.1080/17435390.2017.1379567
66. V. Kovalishyn, N. Abramenko, I. Kopernyk, et al., Food Chem. Toxicol. **112**, 507 (2018). https://doi.org/10.1016/j.fct.2017.08.008
67. M. Fernandez, H. Barron, and A. S. Barnard, RSC Adv. **7**, 48962 (2017). https://doi.org/10.1039/C7RA06622H
68. X. Ma, Z. Li, L. E. K. Achenie, and H. Xin, J. Phys. Chem. Lett. **6**, 3528 (2015). https://doi.org/10.1021/acs.jpclett.5b01660
69. R. Jinnouchi and R. Asahi, J. Phys. Chem. Lett. **8**, 4279 (2017). https://doi.org/10.1021/acs.jpclett.7b02010

Translated by O. Kadkin