Journal of Nanomaterials & Molecular Nanotechnology

Research A SCITECHNOL JOURNAL

Cross-Validation and Cosine Similarity-based Deep Correlation Analysis of Nonlinear Properties in Transition Metal Clusters

Alireza Kokabi[1*], Zohreh Naghibi[2], Zahra Nasiri Mahd[1]

[1]Department of Electrical Engineering, Hamedan University of Technology, Hamedan, Iran

[2]Department of Computer Engineering, Hamedan University of Technology, Hamedan, Iran

***Corresponding author:** Alireza Kokabi, Department of Electrical Engineering, Hamedan University of Technology, Hamedan, Iran, E-mail: amjadqcm@gmail.com

Received date: 27 June, 2022, Manuscript No. JNMN-22-47753;

Editor assigned date: 29 June, 2022, PreQC No. JNMN-22-47753 (PQ);

Reviewed date: 12 July, 2022, QC No. JNMN-22-47753;

Revised date: 19 July, 2022, Manuscript No. JNMN-22-47753 (R);

Published date: 26 July, 2022, DOI: 10.4172/2324-8777.1000348

Abstract

A new approach for the rapid and accurate correlation study of the nonlinear properties in the Transition Metal (TM) clusters based on the Deep Leave-One-Out Cross-Validation (LOO-CV) method is presented. The method together with the cosine similarity delivers significant accuracy in the order of at most 10−9 for the prediction of total energy, lowest vibrational mode, binding energy and HOMO-LUMO energy gap of TM2, TM3, and TM4 nanoclusters. Based on the correlation errors, the most coupling TM clusters are ob- tained. In this regard, Mn and Ni clusters has the maximum and minimum amount of energy couplings with other transition metals, respectively. In general, energy prediction errors of TM2, TM3, and TM4 demonstrate comparable patterns while an even-odd behavior is observed for vibrational modes and binding energies. In addition, Ti, V and Co demonstrate maximum binding energy coupling to the TM2, TM3 and TM4 sets, respectively. For the case of the energy gap, Ni shows the maximum correlation in the smallest TM2 clusters while Cr dependence is highest for TM3 and TM4 sets. Finally, Zn has the highest error for HOMO-LUMO energy gap in all sets and consequently the maximum independent energy gap characteristics.

Keywords: Transition metal clusters; Density functional theory; Artificial neural network; Deep neural network.

Introduction

Cluster properties are significantly correlated with their corresponding size, composition and charge state. As a result, their numerous applications, specifically in heterogeneous catalysis and the various branches of nanoscience, are extensively related to their characteristics [1]. Analysis of similarities and differences between the characteristics of clusters and those from the corresponding bulk materials is com- mon [2]. Understanding the properties of the clusters will help to gain a deeper understanding and new insights into the behavior of the bulk. Small cluster characteristics are greatly different from bulk matter and are not scalable from those of bulk materials [3].

The atomic clusters of transition metals have been the topic of extensive investigations over the last decades. Due to the challenges of understanding transition metal nanocluster formation, stabilization, agglomeration, size and shape-dependent properties, there is still an enormous interest in transition metal nanoclusters [4]. Wide knowledge of these variations, especially for transition metal clusters, can open an extensive window of applications considering their structural, electronic, and magnetic properties [5,6]. These applications include catalysis, quantum computers, photochemistry, optics, nanoelectronics, chemical sensors, etc. An incomplete d shell plays a crucial role in the discrimination of transition metals properties in comparison with the clusters or solids of simple sp metals.

While many groups are empirically analyzing TM clusters, experiments may impose some limitations such as the exact control of the size and configuration. Atomic material modeling is now commonly and widely used, and includes a variety of techniques from exact quantum chemical methods to density functional theory (DFT) and semi-empirical quantum mechanics to analytical interatomic potentials. DFT is the key computational tool and most popular approach for explaining the chemical and physical properties of materials in different phases, forms, and degrees of aggregation. DFT calculations typically include a high computational cost and are sensitive to the employed exchange-correlation functional. Thus, its corresponding ability to simulate large-size systems and long-term scales is relatively limited [7].

In material science, models based on different types of neural networks have shown excellent perfor- mance in material detection, material analysis, material design, and quantum chemistry (e.g., Convolutional Neural Network (CNN), Recurrent Neural Network (RNN), Deep Belief Network (DBN), and Deep Coding Network). In this area, the advancement of deep learning has made new advances in the application of data-driven approaches. Deep learning has been deployed to the integrated frontal poly-merization to adapt the chemistry to the desired manufacturing strategy. The model is applied to predict the cure kinetics parameters for a given set of front characteristics. The convolutional neural networks is also employed to design a predictive model for the electronic properties of the metal halide perovskites (MHPs). In addition, the tensorial ANN models of the product of atomic coordinate vectors and virtual NN outputs or their partial derivatives with respect to atomic coordinates have been previously developed. The proposed models accurately predict a large number of ab-initio data in a variety of benchmark systems. Furthermore, the ANNs are deployed to estimate the output characteristics of a multiscale model in the thin film formation obtained from the chemical vapor deposition method. This indicates that the ANN-based approach can acceptably estimate the behavior of the observables under uncertainty and mitigate a huge amount of undesired parasitic effects . Deep learning has been employed to estimate a variety of the electronic properties in the organic semiconductors such as their HOMO and LUMO energy levels and the excited states. In addition, the capability of the Deep Leave-One-Out Cross-Validation (LOO-CV) models has been shown in absorption spectra prediction, considerably reducing the number of quantum chemical calculations [8].

A deep Leave-One-Out Cross-Validation model is used to predict the TM properties in this paper. This model is trained to learn the nonlinear relationship within data without relying on the physical model. The Cross-Validation-based model has been used to investigate the correlated nonlinear properties of the transition metal small clusters. This network is utilized to predict the missing metal cluster properties based on the other properties of existing ones. Considering the similarity of transition metals, the probability of predicting the characteristics of any of these metal clusters with regard to the nine other transition metal ones in the fourth row of the periodic table is investigated. Moreover, the interdependence among these properties in the small transition metal clusters is reported which would enhance the understanding of the underlying factors. It is also of great significance to extract the constant patterns in the characteristics of TM small clusters as their size varies. The proposed model has been deployed to predict the interdependence of transition metal clusters based on their analogy.

Methods

ANN is an efficient method that can use complex algorithms and analyze data to investigate complex relations between features and predict the behavior of unknown systems. In this paper, the possibility of predicting the nonlinear properties of transition metals is examined using ANN-model. Since the properties of transition metal clusters share several similarities, it is expected that by having the properties of nine transition metals clusters, the properties of the tenth transition metal cluster are predictable.

A Deep Neural Network (DNN) model based on Feedforward Multilayer Perceptron (MLP) with Levenberg Marquart (LM) error Back-Propagation (BP) training algorithm is employed for modeling TM cluster nonlinear properties. The input layer, more than one hidden layer, and the output layer are three main components of DNNs. BP training algorithm is an iterative gradient designed to minimize the mean square error between the output of the neural network and the real values of output (target). Equation (1) refers to the rule of updating weights in the back-propagation algorithm [9].

TM2	TM3	TM4
Sc2	Sc3	Sc4
Ti2	Ti3	Ti4
V2	V3	V4
Cr2	Cr3	Cr4
Mn2	Mn3	Mn4
Fe2	Fe3	Fe4
Co2	Co3	Co4
Ni2	Ni3	Ni4
Cu2	Cu3	Cu4
Zn2	Zn3	Zn4

Table 1: Three datasets including TM2, TM3, and TM4 nanoclusters.

where W represents the weight matrix, E is the cost function which is here called mean squared error (MSE) and η is the learning rate. During the learning process, the weights and biases of the DNN are adjusted to reduce the error in the training data sets. Then, DNN can represent complex mappings and to address large nonlinear problems. In this work, the Tensorflow and Keras Python packages have been used to implement the DNN-based model.

Why Deep Neural Networks?

The recent advances in deep learning in various areas motivated us to pursue this approach in the nonlinear problems of TM clusters. Deep network structures are generally better generalized than shallow networks with the same number of parameters. Deep neural networks map the input to target via a deep sequence of simple data transformations and these data transformations are learned by exposure to examples. Finding the correct value for parameters and hyper-parameters of the deep neural network may seem like a challenging task, specifically given that modifying the value of one parameter will affect the behavior of all the others. Shallow neural network-based models cannot deal with the problems consisting of high complexity and nonlinearity. One of the reasons for the superiority of deep networks is that these networks apply a special form of composition in which by combining the features of one layer in different ways, more abstract features are created in the next layer. Deep architectures are more efficient and consequently more powerful in the modeling of the nonlinear relation between input and output data compared to the shallow neural ones. This has been attained by the same memory computational costs and a fixed amount of training data. DNNs can capture complex high-dimensional functions efficiently. On the other hand, DNNs have a strong ability and outstanding performance to approximate nonlinear functions due to the common saddle points in high dimensional spaces and their hierarchical structures.

Data Collection

Three datasets are defined, which are clusters of first-order transition metals from groups IIIB to IIB (Sc to Zn) with n=2 to 4 where, n represents the number of atoms in each cluster. Here, datasets are marked with TM2, TM3 and TM4. These three datasets are shown in Table 1

Calculations have been performed by Density Functional Theory (DFT) within B3LYP functional and 6-311++G basis set. For each cluster of Table 1 various spin multiplicities have been calculated. Further-more, for each cluster, the structure with the lowest total energy is considered as the final structure. The calculations have been performed by using Gaussian 09W. Figure 1 shows the configurations obtained for TMn clusters, where TM stands for Sc, Ti, V, Cr, Mn, Fe, Co, Ni, Cu, Zn and n is in the range of 2 − 4. Physical and electronic properties of these clusters including group number in the periodic table, bond length between atoms in each cluster, mass, total energy, binding energy, the energy gap between HOMO and LUMO, HOMO, magnetic moment and low-frequency vibrational modes, are extracted from DFT calculations.

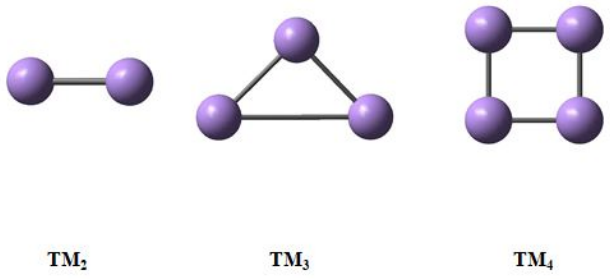

Figure 1: TM2, TM3 and TM4 nanocluster configurations of first-order transition metals.

Data Prepration and Feature Selection DFT calclations are pe

Cluster	Magnetic	Mean Bond Length (A˚)	Binding Energy (eV)	HOMO-LUMO Energy Gap (eV)	Lowest Vibrational Mode (cm-1)
Sc2	4	2.6013	-0.2471	1.3872	262.6104
Ti2	2	1.9115	-0.554	1.8503	492.1252
V2	2	1.7482	-0.4688	2.336	677.2642
Cr2	10	2.7803	-0.2731	1.8922	187.5621
Mn2	10	2.627	0.0861	1.5447	204.9463
Fe2	6	1.994	-0.4249	1.1978	435.5859
Co2	6	2.0882	-0.189	0.7284	347.2473
Ni2	0	2.0672	-0.0203	2.0403	346.837
Cu2	0	2.2788	-0.8925	4.2381	241.5759
Zn2	0	3.8794	0.0004	4.9658	15.9489
Sc3	3	2.9353	-0.5469	1.4968	125.1783
Ti3	4	2.4342	-0.711	1.3839	126.5158
V3	5	2.4795	-0.8372	1.8245	112.1235
Cr3	6	2.9091	-0.6089	1.9439	38.8681
Mn3	5	2.9798	-0.1795	1.8032	109.3713
Fe3	10	3.1762	-0.9844	2.0854	29.8353
Co3	7	3.0517	-0.915	2.1807	25.4997
Ni3	4	2.9358	-0.9688	1.7507	47.8888
Cu3	1	2.4699	-0.8807	1.3036	66.8161
Zn3	0	3.3824	-0.008	4.4534	32.1097
Sc4	4	3.1954	-0.6667	1.2672	188.7053
Ti4	8	2.7074	-0.7984	1.0443	111.7593

V4	6	3.0379	-0.8273	1.4011	130.9942
Cr4	2	2.9381	-0.5864	0.4636	139.6622
Mn4	2	2.9665	-0.058	0.8963	163.5971
Fe4	6	2.7789	-0.8481	0.4911	160.4016
Co4	10	2.657	-1.0123	1.5213	203.1061
Ni4	4	2.7053	-1.0312	0.9589	191.1343
Cu4	0	2.7451	-1.1691	1.9219	142.5269
Zn4	0	4.374	0.0003	4.2302	14.1698

Table 2: Cluster geometric parameters, lowest vibrational modes, binding energies and HOMO-LUMO energy gaps.

Results and Discussion

In this section, the shallow and deep structures of neural networks are employed to investigate the non- linear correlation between TM cluster properties. Different combinations of features were opted to predict the outputs. The total energy, lowest vibrational mode, binding energy and HOMO-LUMO energy gap are considered as the model targets for the DNN approach used for prediction. Many hyper-parameters must be tuned to achieve the proper performance of DNN. The number of hidden layers and neurons is deter- mined to an extent to mitigate the training and test errors. In addition, Swish and Selu have been picked as the activation functions. Moreover, the optimization process is performed through the Adam and RMSProp algorithms. Regarding the importance of the activation function and optimization algorithm for the desired prediction accuracy, the selection of these functions has been performed attentively through the sweeping process.

Total Energy Prediction

The DNN-based model has been employed to predict the Total energy of the TM clusters. Inputs are applied to the neural network with respect to the features is presented in Table Various sets of features were examined to find a combination that precisely predicts the outputs. The corresponding results are given in Table for the TM2, TM3 and TM4 nanoclusters. As can be seen from Table by variation of the number of layers and neurons, training and test error will be swept. The appropriate number of layers and neurons must be chosen to avoid the under-fitting and over-fitting issues. In the first step, the ANN-based model with a shallow structure has been deployed for the prediction. This model includes three layers which consist of an input, a hidden, and an output layer. In the next step, the ANN-based model with a deep structure has been used. This deep structure is formed by the same number of neurons used in the shallow one presented in Table The shallow structure 1, 25, 1 in Table consists of three layers that the input and output layers contain one neuron and the hidden layer includes 25 neurons. Also, the deep structure 1, 5, 5, 5, 10, 1 which is shown in Table consists of six layers that the input layer contains one neuron, the first, second, third and fourth hidden layers contain 5, 5, 5, and 10 neurons, respectively. Finally, the output layer also contains one neuron. As can be seen, the total number of hidden layer neurons is equal in both shallow and deep structures. The results reveal that the values of MSE were significantly lower for the ANN-based model with deep structure compared to that of the ANN-based model with a shallow structure for all nanoclusters.

For the prediction of the total energy, among the different mixture of features, the highest correlation is obtained with the group number. Figure illustrates the correlation between features and targets corresponding to the total energy prediction. The test error value in predicting total energy based on the group

Conclusion

In this paper, the nonlinear correlation between the properties of transition metal nanoclusters was stud- ied using the LOO-CV method. The method together with the cosine similarity delivers significant accuracy for the prediction of the total energy, lowest vibrational mode, binding energy and HOMO-LUMO energy gap. These parameters have been analyzed for all TM2, TM3, and TM4 nanoclusters. The correlation errors demonstrate some characteristic behavior of the elements. For instance, the total energy of the vanadium has similar errors in all three sets. On the other hand, Mn cluster has the lowest and Ni has the highest errors. This means that Mn and Ni clusters have the maximum and minimum amount of energy correlations with other transition metals, respectively. In general, TM2, TM3, and TM4 seem to follow comparable pat- terns. In contrast, the mentioned pattern discriminates TM3 clusters from the TM2 and TM4 sets in the case of vibrational modes. Therefore, for these parameters, an even-odd behavior is observed for the considered range. In the analysis of the binding energy, Ti, V and Co demonstrate maximum coupling to the TM2, TM3 and TM4 sets, respectively. In this regard, the binding energy similarity is highest for TM4 and lowest for TM3 clusters. Again, the discrimination error shows the even-odd pattern for this parameter. Finally, for the energy gap, Ni shows the maximum coupling in the smallest TM2 clusters while Cr dependence is highest for TM3 and TM4 sets. In this regard, Zn has the highest error for HOMO-LUMO energy gap in all sets and consequently the maximum independent energy gap characteristics. Speedup comparison between the deep and shallow structure of neural networks presents better performance for DNNs at the expense of higher computation cost. However, the DNN-based model is still much faster than conventional compu- tational approaches such as DFT. In this regard, the DNN-based model is shown to be more than 29,000 times faster than the conventional DFT calculations.

References

1. Arnold MA, Yanik SC, Suryadevara AC (2019) Septal fractures predict poor outcomes after closed nasal reduction: retrospective review and survey. Laryngoscope; 129: 1784-1790.
2. Li K, Moubayed SP, Spataro E, Most SP (2018) Risk factors for corrective septorhinoplasty associated with initial treatment of isolated nasal fracture. JAMA Facial Plast Surg 20: 460-467.
3. Baykal B, Erdim I, Guvey A, Oghan F, Kayhan FT. Caudal septal stabilization suturing technique to treat crooked noses. J Craniofac Surg 2016; 27 (07) 1830-1833.
4. Ozyazgan I. Septal deviation treatment using bone or cartilage grafts fixed with cyanoacrylate tissue adhesive. Aesthetic Plast Surg 2017; 41 (03) 618-627.
5. Foda HM (2008) The caudal septum replacement graft. Arch Facial Plast Surg 10: 152-157.
6. Loyo M, Markey JD, Gerecci D, Rassi EE, Li Rj, et al. (2018) Technical refinements and outcomes of the modified anterior septal transplant. JAMA Facial Plast Surg 20: 31-36
7. Gubisch W. Twenty-five years experience with extracorporeal septoplasty. Facial Plast Surg 2006; 22 (04) 230-239.
8. Spataro EA, Saltychev M, Kandathil CK, Most SP. Outcomes of extracorporeal septoplasty and its modifications in treatment of severe L-strut septal deviation: a systematic review and meta-analysis. JAMA Facial Plast Surg 2019.
9. Choi JW, Kim MJ, Jeong WS. One-piece nasal osteotomy for the correction of a centrally deviated nose. Aesthetic Plast Surg 2018; 42 (06) 1625-1634.