Implementing PSO-ELM Model to Approximate Trolox Equivalent Antioxidant Capacity as One of the Most Important Biological Properties of Food

Marischa Elveny,1 Ravil Akhmadeev,2 Mina Dinari,3 Walid Kamal Abdelbasset,4,5 Dmitry O. Bokov,6,7 and Mohammad Mahdi Molla Jafari8

1DS & CI Research Group, Universitas Sumatera Utara, Medan, Indonesia
2Department of Accounting and Taxation, Plekhanov Russian University of Economics (PRUE), Stremyanny Lane 36, 117997 Moscow, Russia
3Department of Law, Faculty of Economics and Social Sciences, Shahid Chamran University of Ahwaz, Ahwaz, Iran
4Department of Health and Rehabilitation Sciences, College of Applied Medical Sciences, Prince Sattam Bin Abdulaziz University, Al-Kharj, Saudi Arabia
5Department of Physical Therapy, Kasr Al-Aini Hospital, Cairo University, Giza, Egypt
6Institute of Pharmacy, Sechenov First Moscow State Medical University, 8 Trubetskaya St., Bldg. 2, Moscow 119991, Russia
7Laboratory of Food Chemistry, Federal Research Center of Nutrition, Biotechnology and Food Safety, 2/14 Ustyinsky Pr., Moscow 109240, Russia
8Department of Petroleum Engineering, Ahwaz Faculty of Petroleum Engineering, Petroleum University of Technology (PUT), Ahwaz, Iran

Correspondence should be addressed to Marischa Elveny; marischaelveny@usu.ac.id and Mohammad Mahdi Molla Jafari; mohammad.molajafari@afp.put.ac.ir

Received 1 July 2021; Revised 10 July 2021; Accepted 16 July 2021; Published 3 August 2021

Academic Editor: Alireza Baghban

Copyright © 2021 Marischa Elveny et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

In this paper, the Trolox equivalent antioxidant capacity (TEAC) is estimated through a robust machine-learning algorithm known as the Particle Swarm Optimization-based Extreme Learning Machine (PSO-ELM) model. For this purpose, a large dataset from previously published reports was gathered. Various analyses were performed to evaluate the proposed model. The results of the statistical analysis showed that this model can predict the actual values with high accuracy, so that the calculated \( R^2 \) and RMSE values were equal to 0.973 and 3.56, respectively. Sensitivity analysis was also performed on the effective input parameters. The leverage technique was also performed to check the accuracy of real data, and the results showed that the majority of data are reliable. This simple yet accurate model can be very powerful in predicting the Trolox equivalent antioxidant capacity values and can be a good alternative to laboratory data.

1. Introduction

Many antioxidant compounds may be found in vegetable materials [1, 2]. To find possible sources of natural antioxidants for use in edible products, many plants have been researched and several substances have been identified [3, 4].

Biologists and physicians are also interested in antioxidants because of their use for guarding the human organs against the harm of reactive oxygen species (ROS) [5, 6]. Because of the alleged tight relationship between oxidative stress and illness, antioxidants are thought to be preventive agents against similar illnesses [7, 8].

There is a negative relationship between consuming fruits and vegetables, as the greatest sources of antioxidants, and the cancer risk, so that the risk is reduced by 30-50% [9, 10]. The antioxidant potential of a wide range of substances
is investigated by many researchers; the substances include different vegetables and fruits, drinks, green tea and black ones, plant parts (their leaves, the roots, and their bark), kernel, shell, and also industrial wastes and secondary products apart from the ones derived from plants [11]. The “antioxidant capacity” refers to the certain amount of free radicals being removed by a laboratory solution, without considering the properties of any antioxidant of the mix [12, 13].

The antioxidant capacity of the Trolox equivalent of nutritional excerpts is then measured by taking into account the aggregate activity of all antioxidants of the excerpts, together with their chain-breaking, cleansing, and chelating impacts, hence offering an inclusive factor instead of the measurable antioxidants’ simple sum [14]. So we can identify the antioxidants’ familiar and unfamiliar capacities, and their synergistic relationship, providing a method for identifying a diverse variety of nutrition for antioxidant characteristics [15, 16]. There are many kinds of research for measuring the antioxidant capacity of Trolox equivalent of nutritional products [17–21].

The basis of each technique is producing various free radicals via a range of methods, followed by measuring a variety of endpoints at a defining moment or above the limit [22]. A spectrophotometric test of the antioxidant capacity of the Trolox equivalent is a frequent technique, which has relied on the antioxidants’ relative capacities that exist in nutritional products to remove the radical cation of ABTS*+ (2,2′-azinobis(3-ethyl-benzothiazoline-6-sulfonic acid)), comparing to the capacity of the antioxidant of 6-hydroxy-2,5,7,8-tetramethylchroman-2-carboxylic acid (Trolox) standard amounts [17, 23]. We previously evaluated the substances of physiologically active chemicals and the cruciferous seeds’ Trolox equivalent antioxidant capacity during sprouting, and the findings on ascorbic acid levels were previously reported [24].

The present paper is aimed at developing a PSO-ELM model for forecasting the antioxidant capacity of the Trolox equivalent of various sprouting cruciferous seeds dependent upon the entire phenolic composites, isonitole hexahexaphosphate, glucosinolates, soluble proteins, ascorbic acid, and the entire tocopherol content (as an add-up of α-T, β-T, γ-T, and δ-T labeled as T_m) and the antioxidant capacity of the Trolox equivalent of sprouted cruciferous seeds, as established experimentally. After the model construction stage, various analyses are used to evaluate its accuracy. Sensitivity analysis is also used to determine the effect of each of the input parameters on the target values.

2. Description of Modeling

2.1. PSO. One of the techniques for stochastic optimization is PSO which was presented by Eberhart and Kennedy [25]. The application and manual for this procedure are introduced in [26–28]. An abstract of this method is summarized in six steps which will be mentioned as follows [29].

\[ x_i^k(t + 1) = x_i^k(t) + \alpha(t) \left( p_i^k(t) - x_i^k(t) \right) + \beta_i(t) \left( g_i^k(t) - x_i^k(t) \right) \]

\[ x_i^k(t + 1) = x_i^k(t) + v_i^k(t + 1) \]

Step 1: a majority of stochastic solutions is formed as the searching space. Assume that the searching space has two parameters: dimension (D) and particle number (N). Each possible solution is dedicated to two attributes: position and velocity of the \( x_i^k(t) \) particle in iteration \( k \). These particles are then “flown” through the search space of possible solutions which are indicated as follows.

Step 2: measure the fitness between each particle in the swarm.

Step 3: for each iteration, evaluate the particle’s fitness with the best fitness acquired in the previous ones. If this value is better than the best acquired previous ones, replace the amount and location of the previous one with the current value and location, respectively.

Step 4: evaluate some fragments together and update the finest place with the best fitness \( g_i^k \).

Step 5: the pace of each fragment is increased towards its \( (g_i^k) \) and \( (p_i^k) \). This speed or acceleration is valued by an accidental term.

Step 6: start again from step 2 based on favorable factors until a new convergence is reached.

2.2. ELM. An ELM can be introduced as a least square-based single hidden layer (HL) feed-forward neural network (SLFN) for two problems: regression and classification. Huang et al., for the design of an ELM, apply the kernel function instead of the HL with a huge amount of nodes [30]. Hung et al. and Pal and Deswal [31] both suggested techniques; the abstract of these methods is as follows.

Hidden neurons \( H \), ELM for the training set \( N \), and activation function \( f(x) \) can be described as follows:

\[ e_j = \sum_{i=1}^{H} \alpha_i f(w_{ij}, x_i) = 1 \cdots N. \]

\[ \alpha_i \text{ and } \omega_i \text{ are } H \text{-output layers and weight vectors (WVs) of the connecting input HLs (input Ws)}, \text{respectively.} \]

\[ x_j \text{ indicates input variables. } C_i \text{ indicates the } H \text{ bias for the } i \text{th } H \text{ neuron, and } e_j \text{ is the output of ELM for multiple data points (j).} \]

The process of generating input Ws is random and is based on consecutive distribution. Via a linear function, the output and result of Ws are calculated which are as follows:

\[ \beta = A \dagger Y. \]

A shows the output of the HL matrix (equation (5)), A \( \dagger \) indicates the inverse of A when using the Moore-Penrose method, and Y represents the values that ELM tries to reach. The compact and simplified form of equation (4) is \( Aa = Y \). A is the matrix of HL in the neural network (NN) and Y is the vectors of the output variable. The three matrixes, \( A, \alpha \) , and \( Y \), can be represented as follows:
\[
A = \begin{bmatrix}
    h(x_i) \\
    \vdots \\
    h(x_N)
\end{bmatrix} = \begin{bmatrix}
    f(w_1, c_1, x_1) & \cdots & f(w_H, c_H, x_1) \\
    \vdots & \cdots & \vdots \\
    f(w_1, c_1, x_j) & \cdots & f(w_H, c_H, x_j)
\end{bmatrix}, \alpha
\]

\[
\begin{bmatrix}
    \alpha_1^T \\
    \vdots \\
    \alpha_H^T
\end{bmatrix},
\begin{bmatrix}
    y_1^T \\
    \vdots \\
    y_N^T
\end{bmatrix} = \begin{bmatrix}
    f(w_1, c_1, x_j) & \cdots & f(w_H, c_H, x_j)
\end{bmatrix}
\]

in which \(h(x)\) represents the feature mapping of the HL. Matrix \(A\) strongly assigns the result of the ELM algorithm. One of the traditional solutions is using NNs in the HLs in which this type of solution uses a gradient descent algorithm as represented in [31]. In order to solve ELM, the kernel function is applied, and on the other hand, for solving the kernel matrix, feature mapping is used as can be seen in the following [30]:

\[
k(x_i, x_j) = h(x_i) \cdot h(x_j). \quad (6)
\]

In this article, we attempt to comprehend the efficiency of the kernel on the Mr prediction via KELM application and integrating ELM method with PSO for designing a fresh model for predicting Mr. In the point of view of the learning rate, predictive performance, and generalization capability, ELM has better performance in comparison with normal NNs. ELM via the Moore-Penrose generalized reverse method distinguishes the Ws of the output and input layers and produces some random values for \(H\) biases and input Ws [32, 33]. The general structure of a NN known as SLFN (single HL feed-forward NN) included the output and input layer neurons, \(m\) and \(n\), respectively, and also HL neurons. For instance, suppose \(\{X_i, Y_i\}\) is a training dataset, then it can be understood that the input dataset is \(X_i = [X_{i1}, X_{i2}, \ldots, X_{in}]\) and the output dataset is \(Y_i = [Y_{i1}, Y_{i2}, \ldots, Y_{im}]\) for \(i = 1, 2, \ldots, n, m\) is the number of training samples.

### 3. Data Gathering

There are 172 data points in the database utilized in the present study [34], with two train (129 data points (about 75%)) and test (43 data points (about 25%)) datasets for training and testing the efficiency of proposed models, respectively. To boost the efficiency of the study models, the data points were normalized between −1 and +1.

### 4. Results and Discussion

The efficiency analysis needs to be carried out for evaluating the capability of the model. Accordingly, various statistical analyses were carried out between the actual values and the model outputs, including standard deviations (STD), mean relative errors (MRE), root mean square error (RMSE), mean squared error (MSE), and \(R^2\)-squared \((R^2)\) to evaluate the capability of the study model [35–38].

Figure 1 represents the actual values versus model outputs for the output data at the train and test phases. The target is accurately estimated by the study model with a decent agreement between the real information and model yields, highlighting their capability in output prediction.

Also, the model results from regression analysis represented in Figure 2 at the train and test phases. Based on related literature, the \(R^2\) value is an eminent statistic indicating the model output-actual value relationship. The basic objective was to conduct a comparative analysis between model yields and real values. The accuracy of the model’s accuracy is improved when the fitted line approaches the bisector line [39, 40]. A remarkable linear correlation is achieved between the model outputs and actual values for \(R^2 = 1\), which gets weaker when the \(R^2\) value approaches to ...
zero [41, 42]. Accuracy is represented by a close-fitting of data points around the 45° line for the prediction models. As shown in this figure, this model shows a high ability to predict target values in different phases.

Table 1 presents the results from statistical analyses of the study model based on the RMSE, MSE, STD, MRE, and $R^2$ parameters [43, 44].

In a study with similar input data, Buciński et al. predicted TAEC values using the artificial neural network (ANN) method [34]. Their model showed an accuracy of $R^2 = 0.931$ in estimating the output data in the testing phase, which was weaker than the model proposed in this paper.

Furthermore, Figure 3 represents the absolute relative deviation between the actual values and model yields of output anticipated utilizing the examination model. William’s plot was utilized for determining the outliers of the model [45, 46]. Figure 4 represents the standardized residuals versus hat values. This figure clearly shows three
limited boundaries: leverage limit, upper limit, and down suspected limit [47]. Outliers are data with higher standardized residual values > 3 or <-3, and the data with hat > hat* (referred to as the warning leverage value) are beyond the applicability domain of the study model [48]. As can be seen from this figure, among all the data points, only three suspicious points are seen.

Finally, sensitivity analysis was used to determine the effect of different input parameters on the target parameter. More details about this analysis are given elsewhere [49, 50]. According to Figure 5, it was found that T_{tot} has the most direct effect on the target parameter, which corresponds to the relevancy factor (r) equal to +0.26, while other input parameters showed an inverse effect on the target parameter so that GLS showed the most negative effect with r equal to -0.85.

5. Conclusion

This study was aimed at seeing how well a statistical learning-based model could predict the antioxidant capacity of cruciferous sprouts. To this end, the PSO was implemented in the ELM model. When it came to setting the tuning parameters, the PSO algorithm showed good performance. Estimates were found to be quite accurate when compared to actual data points. The efficiency of the proposed techniques was verified by an excellent agreement achieved between the model outputs and the actual values in assessing the model.
during the train and test phases, as demonstrated by results from statistical analyses. The models’ accuracy was confirmed as predicted by a comparison which was made between the proposed models’ outcomes and another reported correlation. The proposed strategy to predict the antioxidant capacity of cruciferous sprouts is user-friendly so that they can be considered a useful tool for researchers, particularly in related fields, unlike the sophisticated mathematical techniques developed for this output prediction.

Data Availability

The data used to support the findings of this study are provided within the article.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

References

[1] H. Anwar, G. Hussain, and I. Mustafa, “Antioxidants from natural sources,” Antioxidants in Foods and its Applications, 2018.

[2] S. C. Lourenço, M. Moldão-Martins, and V. D. Alves, “Antioxidants of natural plant origins: from sources to food industry applications,” Molecules, vol. 24, no. 22, p. 4132, 2019.

[3] A. E. Hagerman, K. M. Riedl, G. A. Jones et al., “High molecular weight plant polyphenolics (tannins) as biological antioxidants,” Journal of Agricultural and Food Chemistry, vol. 46, no. 5, pp. 1887–1892, 1998.

[4] I. Gulcin, “Antioxidants and antioxidant methods: an updated overview,” Archives of Toxicology, vol. 94, no. 3, pp. 651–715, 2020.

[5] N. J. Temple, “Antioxidants and disease: more questions than answers,” Nutrition Research, vol. 20, no. 3, pp. 449–459, 2000.

[6] H. Zielinski, “Low molecular weight in the cereal grains—a review,” Polish Journal of Food and Nutrition Sciences (Poland), vol. 52, pp. 3–6, 2002.

[7] S. Asgary, A. Rastqar, and M. Keshvari, “Functional food and cardiovascular disease prevention and treatment: a review,” Journal of the American College of Nutrition, vol. 37, no. 5, pp. 429–455, 2018.

[8] J. W. Ballway and B.-J. Song, “Translational approaches with antioxidant phytochemicals against alcohol-mediated oxidative stress, gut dysbiosis, intestinal barrier dysfunction and fatty liver disease,” Antioxidants, vol. 10, no. 3, p. 384, 2021.

[9] G. Block, B. Patterson, and A. Subar, “Fruit, vegetables, and cancer prevention: a review of the epidemiological evidence,” Nutrition and Cancer, vol. 18, no. 1, pp. 1–29, 1992.

[10] K. A. Steinmetz and J. D. Potter, “Vegetables, fruit, and cancer prevention: a review,” Journal of the American Dietetic Association, vol. 96, no. 10, pp. 1027–1039, 1996.

[11] N. Chhikara, R. Kaur, S. Jaglan, P. Sharma, Y. Gat, and A. Panghal, “Bioactive compounds and pharmaceutical and food applications of Syzygium cumini—a review,” Food & Function, vol. 9, no. 12, pp. 6096–6115, 2018.

[12] A. Ghiselli, M. Serafini, F. Natella, and C. Scaccini, “Total antioxidant capacity as a tool to assess redox status: critical view and experimental data,” Free Radical Biology and Medicine, vol. 29, no. 11, pp. 1106–1114, 2000.

[13] H. Zielinski, “Peroxyl radical-trapping capacity of germinated legume seeds,” Food/Nahrung, vol. 46, no. 2, pp. 100–104, 2002.

[14] S. Ortega-Requena, S. Rebouillat, and F. Pla, “Paving the highway to sustainable, value adding open-innovation integrating bigger-data challenges: three examples from bio-ingredients to robust durable applications of electrochemical impacts,” Journal of Biomaterials and Nanobiotechnology, vol. 9, no. 2, pp. 117–188, 2018.

[15] G. F. Biggi, Investigating the molecular genetic basis of antioxidants in Lactuca sativa for the enhancement of its nutritional qualities, University of Southampton, 2010.

[16] K. O. Chu, The study of feasibility of green tea treatment on fetus: from chemistry to treatment, The Chinese University of Hong Kong, Hong Kong, 2005.

[17] N. Miller and C. Rice-Evans, “Spectrophotometric determination of antioxidant activity,” Redox Report, vol. 2, no. 3, pp. 161–171, 1996.

[18] C. Rice-Evans and N. J. Miller, “Antioxidants the case for fruit and vegetables in the diet,” British Food Journal, vol. 97, no. 9, pp. 35–40, 1995.

[19] N. Salah, N. J. Miller, G. Paganga, L. Tijburg, G. P. Bolwell, and C. Riceevans, “Polyphenolic Flavanols as Scavengers of Aqueous Phase Radicals and as Chain- Breaking Antioxidants,” Archives of Biochemistry and Biophysics, vol. 322, no. 2, pp. 339–346, 1995.

[20] H. Wang, G. Cao, and R. L. Prior, “Total antioxidant capacity of fruits,” Journal of Agricultural and Food Chemistry, vol. 44, no. 3, pp. 701–705, 1996.

[21] H. Zielinski and H. Kozlowska, “Antioxidant activity and total phenolics in selected cereal grains and their different morphological fractions,” Journal of Agricultural and Food Chemistry, vol. 48, no. 6, pp. 2008–2016, 2000.

[22] R. Apak, M. Ozyurek, K. Guçlü, and E. Çapanoğlu, “Antioxidant activity/capacity measurement. I. Classification, physicochemical principles, mechanisms, and electron transfer (ET)-based assays,” Journal of agricultural and food chemistry, vol. 64, no. 5, pp. 997–1027, 2016.

[23] K. Guclu, M. Altn, M. Ozyurek, E. Saliba, and R. A. Karademir, “Antioxidant capacity of fresh, sun- and sulphited-dried Malatya apricot (Prunus armeniaca) assayed by CUPRAC, ABTS/TEAC and folin methods,” International Journal of Food Science & Technology, vol. 41, no. 1, pp. 76–85, 2006.

[24] H. Zielinski, A. Bucinski, and H. Kozlowska, “Monitoring of the vitamin C content in germinating cruciferous seeds by HPLC,” Polish Journal of Food and Nutrition Sciences, vol. 11, pp. 142–146, 2002.

[25] R. Eberhart and J. Kennedy, “A new optimizer using particle swarm theory In MHS’95,” in Proceedings of the Sixth International Symposium on Micro Machine and Human Science, Ieee, 1995.

[26] P. Wilson and H. A. Mantooth, Model-based engineering for complex electronic systems, Newnes, 2013.

[27] A. M. Sharaf and A. A. Elgammal, Novel AI-based soft computing applications in motor drives, in Power Electronics Handbook, Elsevier, 2018.

[28] F. Han, H.-F. Yao, and Q.-H. Ling, “An improved evolutionary extreme learning machine based on particle swarm optimization,” Neurocomputing, vol. 116, pp. 87–93, 2013.

[29] H. Guo, B. Li, W. Li, F. Qiao, X. Rong, and Y. Li, “Local coupled extreme learning machine based on particle swarm optimization,” Algorithms, vol. 11, no. 11, p. 174, 2018.
[30] G.-B. Huang, H. Zhou, X. Ding, and R. Zhang, “Extreme learning machine for regression and multiclass classification,” IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics), vol. 42, no. 2, pp. 513–529, 2011.

[31] M. Pal and S. Deswal, “Extreme learning machine based modeling of resilient modulus of subgrade soils,” Geotechnical and Geological Engineering, vol. 32, no. 2, pp. 287–296, 2014.

[32] J. Cao, Z. Lin, and G.-B. Huang, “Self-adaptive evolutionary extreme learning machine,” Neural Processing Letters, vol. 36, no. 3, pp. 285–305, 2012.

[33] G.-B. Huang, Q.-Y. Zhu, and C.-K. Siew, “Extreme learning machine: theory and applications,” Neurocomputing, vol. 70, no. 1-3, pp. 489–501, 2006.

[34] A. Buciński, H. Zieliński, and H. Kozłowska, “Artificial neural networks for prediction of antioxidant capacity of cruciferous sprouts,” Trends in Food Science & Technology, vol. 15, no. 3-4, pp. 161–169, 2004.

[35] M. Mahdaviara, A. Rostami, F. Keivanimehr, and K. Shahbazi, “Accurate determination of permeability in carbonate reservoirs using Gaussian process regression,” Journal of Petroleum Science and Engineering, vol. 196, p. 107807, 2021.

[36] E. Khamehchi and A. Bemani, “Prediction of pressure in different two-phase flow conditions: machine learning applications,” Measurement, vol. 173, p. 108665, 2021.

[37] X. Zhou, F. Zhou, and M. Naseri, “An insight into the estimation of frost thermal conductivity on parallel surface channels using kernel based GPR strategy,” Scientific Reports, vol. 11, no. 1, 2021.

[38] M. R. Kalooop, A. Bardhan, N. Kardani, P. Samui, J. W. Hu, and A. Ramzy, “Novel application of adaptive swarm intelligence techniques coupled with adaptive network-based fuzzy inference system in predicting photovoltaic power,” Renewable and Sustainable Energy Reviews, vol. 148, p. 111315, 2021.

[39] A. Hemmati-Sarapardeh, M. Fathinasab, and A. H. Mohammadi, “Determination of minimum miscibility pressure in N2-crude oil system: A robust compositional model,” Fuel, vol. 182, pp. 402–410, 2016.

[40] H. Mokarizadeh, S. Atashrouz, H. Mirshekar, A. Hemmati-Sarapardeh, and A. M. Pour, “Comparison of LSSVM model results with artificial neural network model for determination of the solubility of SO2 in ionic liquids,” Journal of Molecular Liquids, vol. 304, p. 112771, 2020.

[41] A. Ghanbari, M. N. Kardani, A. M. Goodarzi, M. J. Larche, and A. Baghban, “Neural computing approach for estimation of natural gas dew point temperature in glycol dehydration plant,” International Journal of Ambient Energy, vol. 41, no. 7, pp. 775–782, 2020.

[42] N. Kardani, A. Bardhan, D. Kim, P. Samui, and A. Zhou, “Modelling the energy performance of residential buildings using advanced computational frameworks based on RVM, GMDH, ANFIS-BBO and ANFIS-IPSO,” Journal of Building Engineering, vol. 35, 2021.

[43] M. Ahmadi, M. Ghazvini, A. Baghban et al., “Soft computing approaches for thermal conductivity estimation of CNT/water nanofluid,” Revue des Composites et des Matériaux Avancés, vol. 29, no. 2, 2019.

[44] X. Zhou, F. Zhou, and M. Naseri, “An insight into the estimation of frost thermal conductivity on parallel surface channels using kernel based GPR strategy,” Scientific Reports, vol. 11, no. 1, 2021.