Arabian Journal of Chemistry (2022) **15**, 103843

King Saud University

Arabian Journal of Chemistry

www.ksu.edu.sa
www.sciencedirect.com

ORIGINAL ARTICLE

Development of multiple machine-learning computational techniques for optimization of heterogenous catalytic biodiesel production from waste vegetable oil

Walid Kamal Abdelbasset [a,b,*], Safaa M. Elkholi [c], Maria Jade Catalan Opulencia [d], Tazeddinova Diana [e,f], Chia-Hung Su [g,*], May Alashwal [h], Mohammed Zwawi [i], Mohammed Algarni [i], Anas Abdelrahman [j], Hoang Chinh Nguyen [k]

[a] *Department of Health and Rehabilitation Sciences, College of Applied Medical Sciences, Prince Sattam bin Abdulaziz University, P.O. Box. 173, Al-Kharj 11942, Saudi Arabia*
[b] *Department of Physical Therapy, Kasr Al-Aini Hospital, Cairo University, Giza 12613, Egypt*
[c] *Department of Rehabilitation Sciences, College of Health and Rehabilitation Sciences, Princess Nourah bint Abdulrahman University, P.O. Box. 84428, Riyadh 11671, Saudi Arabia*
[d] *College of Business Administration, Ajman University, Ajman, United Arab Emirates*
[e] *Department of Technology and Catering Organization, South Ural State University, Chelyabinsk, Russian Federation*
[f] *Zhangir Khan Agrarian Technical University, Uralsk, Kazakhstan*
[g] *Department of Chemical Engineering, Ming Chi University of Technology, New Taipei City, Taiwan*
[h] *Department of Computer Science, Jeddah International College, Jeddah, Saudi Arabia*
[i] *Mechanical Engineering Department, Faculty of Engineering, King Abdulaziz University, P.O. Box 344, Rabigh 21911, Saudi Arabia*
[j] *Department of Mechanical Engineering, Faculty of Engineering & Technology, Future University in Egypt, 11845 New Cairo, Egypt*
[k] *School of Life and Environmental Sciences, Deakin University, Geelong, Victoria 3216, Australia*

Received 6 December 2021; accepted 13 March 2022
Available online 16 March 2022

KEYWORDS

Biodiesel;
Esterification;
Renewable energy;

Abstract Multiple machine learning models were developed in this study to optimize biodiesel production from waste cooking oil in a heterogenous catalytic reaction mode. Several input parameters were considered for the model including reaction temperature, reaction time, catalyst loading, methanol/oil molar ratio, whereas the percent of biodiesel production yield was the only output.

* Corresponding authors.
E-mail addresses: walidkamal.wr@gmail.com (W. Kamal Abdelbasset), chsu@mail.mcut.edu.tw (C.-H. Su).
Peer review under responsibility of King Saud University.

Production and hosting by Elsevier

https://doi.org/10.1016/j.arabjc.2022.103843

Process optimization;
Machine learning

Three ensemble models were utilized in this study: Boosted Linear Regression, Boosted Multi-layer Perceptron, and Forest of Randomized Tree for optimization of the yield. We then found their optimized configurations for each model, namely hyper-parameters. This critical task is done by running more than 1000 combinations of hyper-parameters. Finally, The R^2-Scores for Boosted Linear Regression, Boosted Multi-layer Perceptron, and Forest of Randomized Tree, respectively, were 0.926, 0.998, and 0.992. MAPE criterion revealed that the error rates for boosted linear regression, boosted multi-layer perceptron, and Forest of Randomized Tree was 5.68×10^{-2}, 5.20×10^{-2}, and 9.83×10^{-2}, respectively. Furthermore, utilizing the input vector (X1 = 165, X2 = 5.72, X3 = 5.55, X4 = 13.0), the proposed technique produces an ideal output value of 96.7 % as the optimum yield in catalytic production of biodiesel from waste cooking oil.

1. Introduction

Recently, development of renewable energy sources as an alternative to conventional fossil fuels has attracted much attention worldwide in terms of economic profit as well as environmental protection (Yahya et al., 2020; Adebayo, 2022; Zhao, 2022; A.B.W, P. Computer Technology Simulation towards Power Generation Potential from Coproduced Fluids in South Lokichar Oil Fields., 2020; Shen, 2021; Mao, 2020; Ekramian and Etemad, 2014). Different techniques and bio-sources have been explored and studied for efficient development of biodiesel production as green fuel (Zhang, 2021; Hu, 2022; , xxxx; Johnson et al., 2022; Ibnou-Laaroussi et al., 2020; Rjoub, 2021; Wang, 2022; Lin, 2021; (Rikani, 2021)). The main aim of the most studies in this area is to enhance the biodiesel production yield for a given biomass source as feedstock (Fu, 2010; Nguyen, 2021). This goal can be achieved by experimental evaluation of the process, development of novel catalysts, and process optimization techniques such as response surface method (RSM) (Yahya et al., 2020; Jia et al., 2012). 2012.; , xxxx; Chen, 2021; Deng, 2019; Liu, 2021; Yin, 2022; Yin, 2022).

Fatty Acid Methyl Esters which are also known as FAME, are esters of fatty acids which are mainly known as biodiesel, are obtained from different sources such as vegetable oils. The FAME is recognized as the main biodiesel due to its similar structure and properties to the conventional fossil-based diesel fuel, however FAME is known as renewable alternative energy source. One of the hurdles toward development of biodiesel is limited availability of feedstock for biodiesel production as well as the price of feedstock (Nair et al., 2020; Huang, 2022; Adebayo and Rjoub, 2022). Various feedstocks have been employed for production of biodiesel such as palm oil (Kansedo et al., 2009), microalgae (Collotta and Basosi, 2019), rice bran oil (Zaidel, et al., 2019), waste cooking oil (WCO) (Mohadesi, 2019), etc.

Waste cooking oil (WCO) has been studied as one of the abundant and cheap sources for production of biodiesel. WCO masses contain fatty acids and triglyceride components which can be converted to biodiesel (FAME) using esterification or transesterification reactions in a batch or continuous reactor with or without catalysts (Yahya et al., 2020; Bakhshkandi and Ghoranneviss, 2019; Osanloo, 2019; Watandost et al., 2021; Andalib and Sarkar, 2021). These reactions are well proceeded in the presence of catalysts that can promote both the esterification and transesterification reactions (Zhao, 2022; Latif et al., 2021; Huang, 2021; Zhang et al., 2021). For the production of FAME from WCO, the process parameters need to be optimized in order to get the highest production yield with the lowest number of experiments. The latter can be implemented through process modeling and simulation, provided that the input and output parameters are well determined (Andalib and Sarkar, 2022; Ethier, 2021; Latif, 2021; Nourian, 2021; Sundaravadivelu Devarajan, 2020). The process parameters for production of FAME using heterogenous catalysts induce temperature of reaction mixture, time, amount of catalyst, and methanol to oil ratio. Also, the main output in the optimization is the production yield which needs to be defined as the objective to be maximized.

Different techniques have been explored and implemented in order to optimize the biodiesel production such as CCD and RSM and provided great capability. Recently, the models based on machine learning (ML) techniques have attracted much attention in different fields of science and engineering for process understanding and process optimization as well (Wang, 2021; Panwar, 2021; Ghadiri, 2021; Pourtousi, 2021; (PUTRA, 2020)). This novel method requires experimental data for process to be used for training the model. Then the trained and validated model can be used in process prediction and optimization. The method has been successfully used in simulation of chemical processes (Shirazian, 2017; Ismail, 2019; Rezakazemi, 2018; Dashti, 2018; Pishnamazi, 2020; Marjani et al., 2020; Babanezhad, 2020; Babanezhad, 2020; Babanezhad et al., 2020; Babanezhad, 2020; Babanezhad, 2020; Nabipour, 2020; Tian, 2020). Machine learning (ML) is a general term for a bunch of AI (Artificial Intelligence) tools technique that enables computers to learn from data without being directly programmed. ML is focused on developing *meta*-programs that process experimental data and use it to train models (El Naqa and Murphy, 2015; Goodfellow et al., 2016). To cope with the non-linear, unpredictable, complex nature of biodiesel systems, data-driven ML technology provides a possible override to standard modeling methodologies (Aghbashlo, 2021; Gupta, et al., 2021; (Yosofvand, 2020)).

In this study, we used three different ensemble methods for simulation and optimization of FAME production using waste cooking oil (WCO) as feedstock via a heterogenous catalyst. These methods are a group of learning algorithms that use multiple base learners to make more robust models. Bagging and boosting are the most common approaches to make ensembles (Maclin and Opitz, 1997; Zhou, 2019). This study selected Forest of Decision Trees as a bagging method and Adaboost as a boosting method that combines with MLP and Linear Regression to make two other distinct models.

Forest of Decision Trees is a set of Decision Trees that are well-known as a method of knowledge representation, classifiers, and algorithms for solving diverse issues in optimization and other applications. The time complexity of trees and tree optimization algorithms have been thoroughly investigated for both finite and infinite sets of characteristics (Rokach and Maimon, 2007; Breiman, et al., 2017; Dr.s.srinivasareddy, d.y.v.n., dr.d.krishna, 2021). Also, Feed-forward neural networks are well-known and widely used techniques for dealing with nonlinear regression models. MLP models may be thought of as a parametric group of regression functions (White, 1992).

A few significant methodological characteristics characterize AdaBoost. First, unlike previous boosting algorithms, which trained multiple estimators using random sub-samples of data, AdaBoost trains multiple estimators with access to all available data points (Ferreira and Figueiredo, 2012). It improves the model by giving high weight to samples poorly estimated in former models.

2. Data Set

For this research, we have collected several experimental data from resources for development of optimization model of process. The data set of this research, which is identical to (Yahya et al., 2020), that is shown in Table 1. There are four input features and only one output and data containing 30 different sample vectors. As shown, the measured data have been obtained by variation of important process parameters including reaction temperature, reaction time, catalyst loading, and the ratio of methanol to oil, while the response variable is the biodiesel production efficiency. The experiments have been performed in batch operational mode, with the aid of a mineral catalyst, known as Montmorillonite *K*10 which is a heterogeneous catalyst for the biodiesel production. More details about the experimental measurements and description of parameters can be found elsewhere (Yahya et al., 2020). We developed the machine learning models for description of the process and finding the optimum point where the production yield is the maximum. The used machine learning models in this study will be explained in the next section.

3. Methodology

3.1. MLP model

The multilayer perceptron (MLP) model is a ML tool that is inspired by the structure of information processing in the human brain, and can be used as predictive tools for chemical/physical processes (Jain et al., 1996). MLP is used to handle a wide range of problems in various industries since it effectively predicts both discrete and numerical variables (Soltani Fesaghandis, 2017). The MLP is primarily made up of neurons, and the layers are formed up of clusters of neurons. In this method, the neurons in the previous layer get inputs from their counterparts in the next layer. Then process them with their activation function, and pass results into the following layer to use (Noriega, 2005).

The process begins with the input layer and progresses until units in the final layer provide some output type. Hidden layers are those that exist between the two input/output layers. The solver function, the activation functions, and the size of hidden layers are hyper-parameters that must be tuned in this algorithm to obtain the prediction accuracy of processes (Elmaz et al., 2020). The following is the output formulation for a Multi-layer Perceptron model with only 1 hidden layer and 1 output:

$$\widetilde{\mathbf{y}} = \delta_2\left(\sum_{i=1}^{m}\left(w_i^{(2)}\delta_1(\mathbf{X})\right) + \mathbf{b}^{(2)}\right)\mathbf{X} = \sum_{j=1}^{n}\left(\mathbf{x}_j\mathbf{w}_{xj}^{(1)}\right) + \mathbf{b}^{(1)}$$

where $\widetilde{\mathbf{y}}$ is the MLP model's prediction vector, m stands for the count of data point in the whole data, n is the size of data set features, and x_j is the j^{th} feature vector $w^{(2)}$ denotes the weights between the hidden and output layers, whereas $w^{(1)}$ denotes the weights of inputs linked to the hidden layer. δ_2 shows the output layer's activation (Zhou, 2018). Also, in

Table 1 The whole used dataset for simulation of biodiesel production used in this work (Yahya et al., 2020).

RUN	X1 = Temperature (^{0}C)	X2 = Reaction time (h)	X3 = Catalyst loading (weight%)	X4 = Methanol: oil molar ratio	Y = Actual yield (%)
1	125	4.5	3	10	62.93
2	175	4.5	3	10	66.13
3	125	7.5	3	10	65.36
4	175	7.5	3	10	66
5	125	4.5	5	10	63.72
6	175	4.5	5	10	68.92
7	125	7.5	5	10	68.81
8	175	7.5	5	10	70.25
9	125	4.5	3	14	64.34
10	175	4.5	3	14	60.25
11	125	7.5	3	14	65.2
12	175	7.5	3	14	58
13	125	4.5	5	14	74.52
14	175	4.5	5	14	75.31
15	125	7.5	5	14	78.24
16	175	7.5	5	14	73.13
17	100	6	4	12	62.02
18	200	6	4	12	64.7
19	150	3	4	12	60.24
20	150	9	4	12	68.12
21	150	6	2	12	58.59
22	150	6	6	12	79.79
23	150	6	4	8	71.44
24	150	6	4	16	73.15
25	150	6	4	12	97.53
26	150	6	4	12	95.75
27	150	6	4	12	95.26
28	150	6	4	12	96.41
29	150	6	4	12	96.23
30	150	6	4	12	97.78

the hidden layer the activation function of neurons is δ_1. The bias vectors in the output layer and all hidden layers are denoted by $b^{(2)}$ and $b^{(1)}$, respectively (Yang, 2008).

To improve the accuracy of predicting, the weights between each connection in a neural network are changed. The broadly useful back-propagation and batch gradient descent algorithms are employed in the learning step (Hecht-Nielsen, 1992).

3.2. Linear Regression

The other base learning algorithm in this study is linear regression, a fundamental regression approach. In a linear regression, the normality assumption is provided, and it pertains to the following equation:

$$y = \beta_0 + \beta_1 x + \varepsilon$$

In the above equation, y represents the output, the independent variable of the model is denoted by x. In fact, linear regression tries to minimize the sum of squares in this model (Pombeiro, 2017; Kim et al., 2020):

$$\sum_{k=1}^{n}\left(y_k - \bar{y}\right)^2 = \sum_{k=1}^{n}\left(y_k - \hat{y}_k\right)^2 + \sum_{k=1}^{n}\left(\hat{y}_k - \bar{y}_k\right)^2$$

Here, y_k denotes the observed value in actual data k, $\bar{y}_k$ is the average of y_k in all n data, and k shows the projected value of $\hat{y}_k$ for sample k.

3.3. Tree-based Ensembles

Decision Tree (DT) is one of most used learners as weak learner for ensemble methods. A weak learner means a simple predictor that is just a bit more accurate than a random predictor. Tree-based ensemble methods are made up of many weak decision tree models that develop parallel to one another to minimize the model's variance and bias at the same time (Breiman, et al., 2017; Xue, 2020).

Random forest (RF) ensemble is a tree-based method that employs a voting process to improve the performance of several weak tree estimators like other ensemble learning methods (Jiang, 2009). In order to train a random forest, the original dataset is used for drawing N bootstrapped sample sets. Afterward, an unpruned regression tree (classification tree) will be grown using every single bootstrapped sample. In this step, instead of utilizing all of the existing predictors, a few and the predetermined number of κ predictors that are randomly sampled are chosen for playing the role of split candidates. This two-step procedure will then be repeated until C trees with the aforementioned properties are developed, and unseen data can be estimated by aggregating the predictions of these C trees. RF employs a bagging approach to increase tree diversity by creating trees from various training datasets, thereby lowering the model's overall variance (Rodriguez-Galiano, 2015). The following equation expresses an RF regression predictor:

$$\hat{f}_{RF}^{C}(x) = \frac{1}{C}\sum_{i=1}^{C} T_i(x)$$

In the above equation, C stands for the number of trees, x represents the data vector, and $T_i(x)$ represents a sole regression tree that is developed on based on bootstrapped samples and a subset consisting of input variables. RF can natively perform estimation of the out-of-bag errors during forest construction by utilizing the samples that are not chosen in the training of the ith tree during the bagging process. The subset that does not utilize an external data subset and can compute an impartial assessment of generalization error is known as out-of-bag (Breiman, et al., 2017; Zhang, 2022). To allocate relative significance score for every single input variable, an input variable is switched by RF while others are kept constant, and the mean reduction in estimation accuracy of the model is measured (Breiman, et al., 2017).

Extra Trees (Extremely randomized trees) is another ensemble approach based on trees, similar to the random forest. When splitting a tree node, it aggressively randomizes both the cut point decision and its characteristics. Extra Tree is useful for classification and regression tasks both (Geurts et al., 2006; Dutta et al., 2021).

Regarding the difference between these two models, they are similar in that both generate numerous trees and split nodes using random subsets of features. However, there are two significant differences: Extra trees do not make bootstrap observations, and nodes are divided on random rather than optimal splits.

3.4. Adaboost

The AdaBoost method is one of the essential ensemble methods. Because of its capabilities, this technique has become popular. As the name implies, basic models are adaptively boosted and used to solve complex problems in this technique. There are two approaches to complex problem solving: simple and complex models. Because of their simplicity of structure, simple models have excellent generalization properties. They are simple to implement in real-time problems, but they cannot solve complex problems due to high bias due to their structure (Freund and Schapire, 1997).

The use of complex models, on the other hand, increases the risk of over-fitting or implementation difficulties due to the complexity of the models (Buitinck, et al., 1309). Such problems can be resolved with the AdaBoost method. With this method, an unreliable base model (weak learner) is used as the starting point for a more reliable system to handle more challenging problems (Pedregosa and Scikit-learn, , 2011). This algorithm can be summarized in these steps:

- Initially
- Initialization:
  - o Make Decision on the number of estimators: M
  - o Set uniform example weights.

- For Each i ∈ [1,2, …, M}:
  - o Train a weak learner Li with a weighted sample.
  - o Test Li using whole dataset.
  - o Set weight for Li learner.
  - o Set sample weights.

4. Results and discussion

After tuning introduced models' hyper-parameters, two R^2 – score and MAPE metrics alongside fitting charts were used

to evaluate the performance of the selected models. These results are shown in Table 2 and Figs. 1-6. Figs. 1 and 5, which show the training stage in the boosted MLP models and the Forest, are the same. This fact shows that both models have crossed almost all points in the training data. Fig. 3 concludes that they had more proper training than the boosted linear regression model. However, when we compare the test figures specifically with Figs. 2 and 6, the Forest model has many

Table 2 Final Model Results.

Models	R^2	MAPE
Boosted MLP	0.998	5.20×10^{-2}
Boosted LR	0.926	5.68×10^{-2}
Forest of Decision Trees	0.992	9.83×10^{-2}

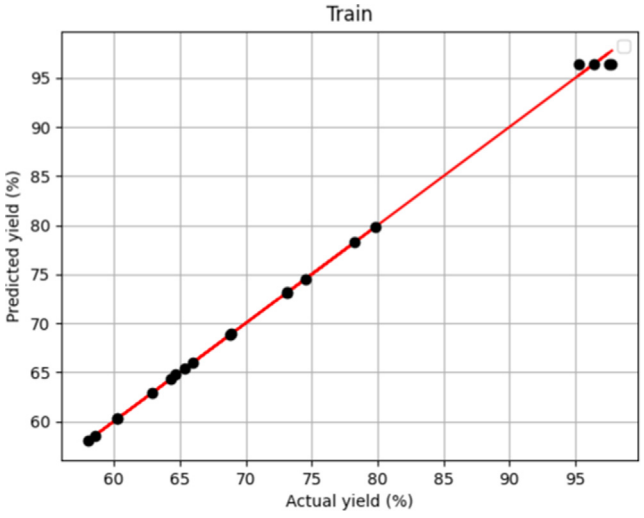

Fig. 1 Model Predicted and Actual yield in Boosted MLP – Train.

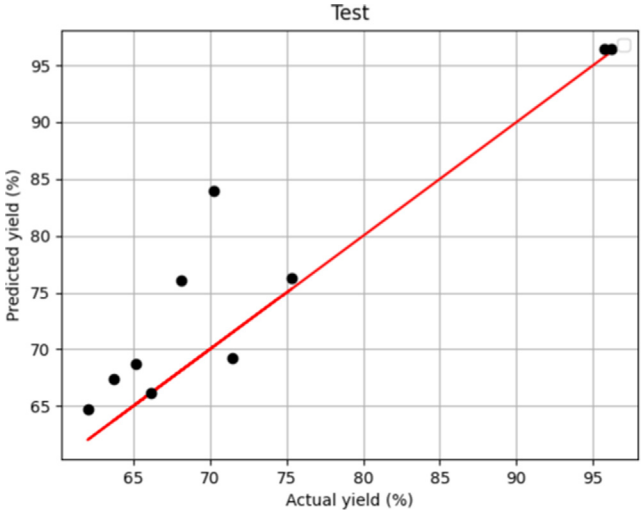

Fig. 2 Model Predicted and Actual yield in Boosted MLP –Test.

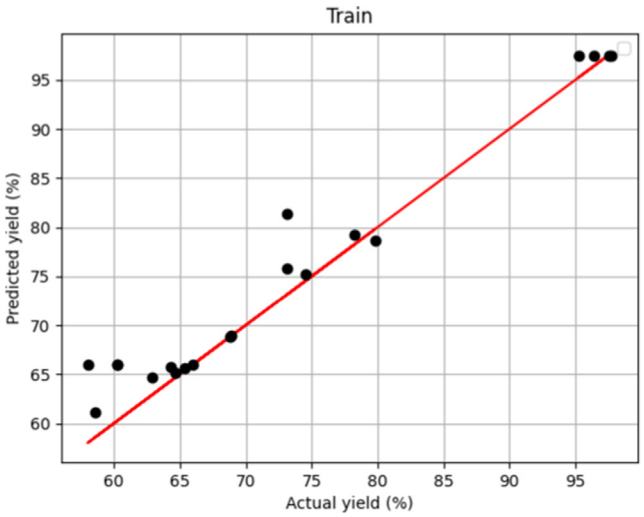

Fig. 3 Model Predicted and Actual yield in Boosted LR –Train.

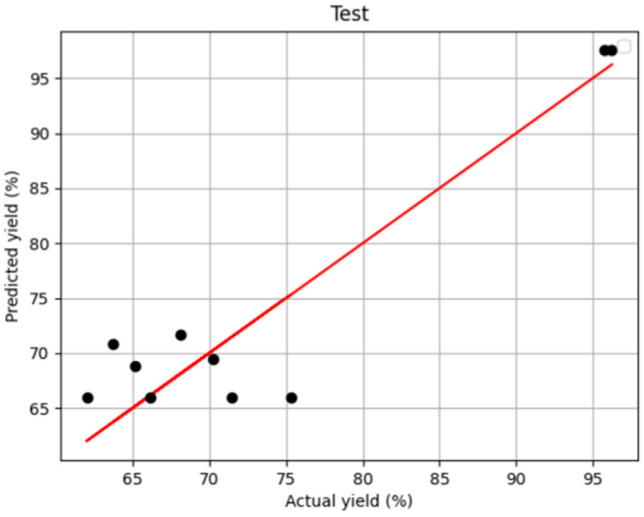

Fig. 4 Model Predicted and Actual yield in Boosted LR –Test.

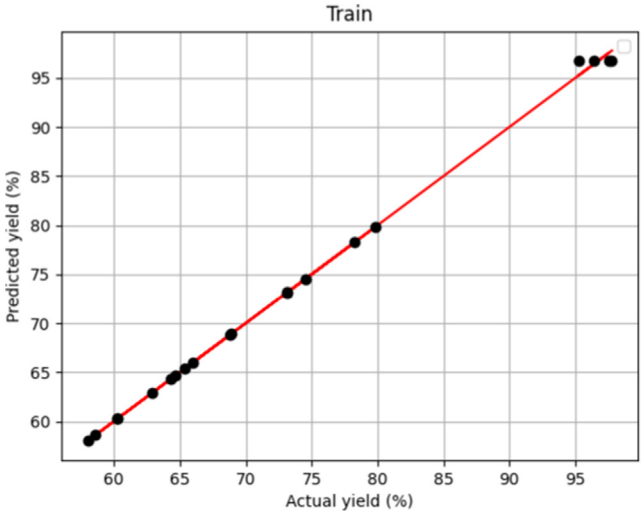

Fig. 5 Model Predicted and Actual yield in Forest of Decision Trees –Train.

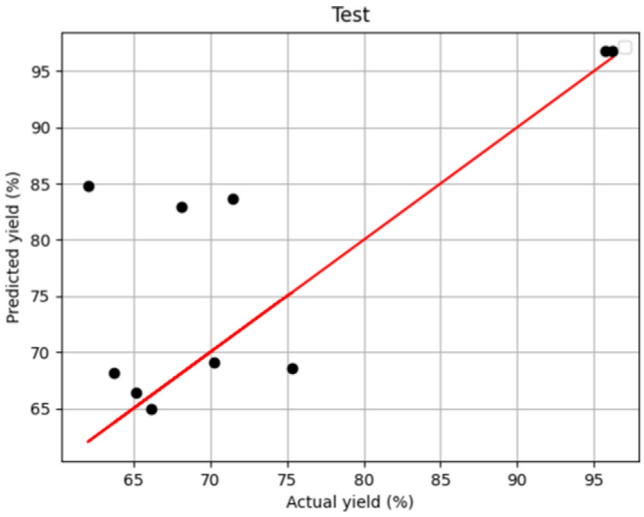

Fig. 6 Model Predicted and Actual yield in Forest of Decision Trees –Test.

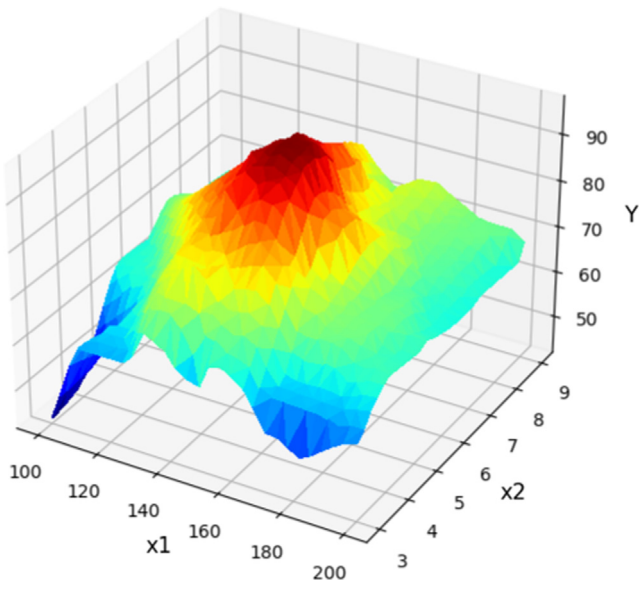

Fig. 7 X1 and X2 projection with prediction surface in final Boosted GPR model. X3 = 4 and X4 = 12 considered constant. Optimum value is y = 95.7 for x1 = 153 x2 = 5.57.

points where the actual values are significantly different that is not the case in boosted MLP. Therefore, we accept the boosted MLP model as the most common model in prediction of this process.

The optimal values of the input parameters at the optimum point are also listed in Table 3, where the highest value of yield is attained as 96.7 % which is consistent with the reported value in literature (Yahya et al., 2020). Furthermore, the 3D surface plots of predicted yield for the combination of various inputs are indicated in Figs. 7-12. In these figures, the red points indicate the optimum values which are specified for each case. Also, the 2D plots of optimization are illustrated in Figs. 13-16 to exactly spot the optimum parameter for each input parameter.

It is indicated that longer reaction time (e.g., 7 h) and higher reaction temperature (greater than170 °C) would degrade the biodiesel production yield to lower values which could be attributed to the rate of reaction and also accumulation of water as by-product of the WCO conversion to FAME. Indeed, the reaction time and temperature must be kept at an optimum value in order to obtain the highest yield of biodiesel production.

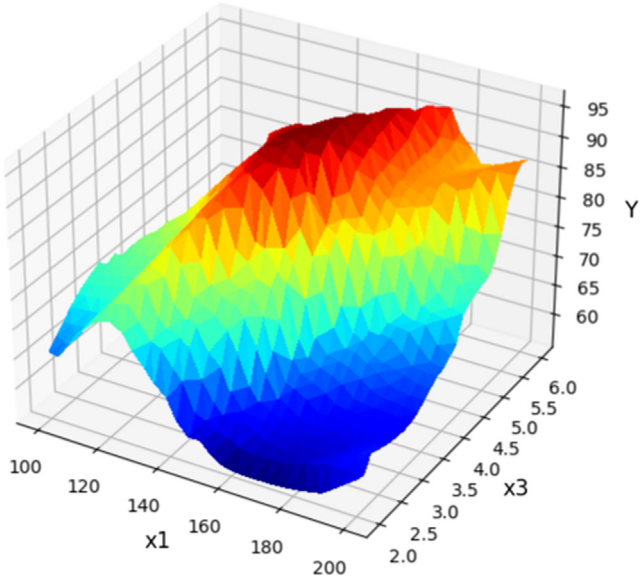

Fig. 8 X1 and X3 projection with prediction surface in final Boosted GPR model. X2 = 6 and X4 = 12 considered constant. Optimum value is y = 96.66 for x1 = 153 x3 = 4.

5. Conclusion

Multiple machine learning regression models were developed in this work in order to optimize production of biodiesel from a feedstock. Waste cooking oil (WCO) was considered as the feedstock for the reaction, and heterogenous catalyst was considered in the process. Four different input parameters including reaction temperature, reaction

Table 3 Optimal values of the paramours for maximum response.

X1 = Temperature (°C)	X2 = Reaction time (h)	X3 = Catalyst loading (wt%)	X4 = Methanol: oil molar ratio	Y = Actual yield (%)
165	5.72	5.55	13.0	96.7

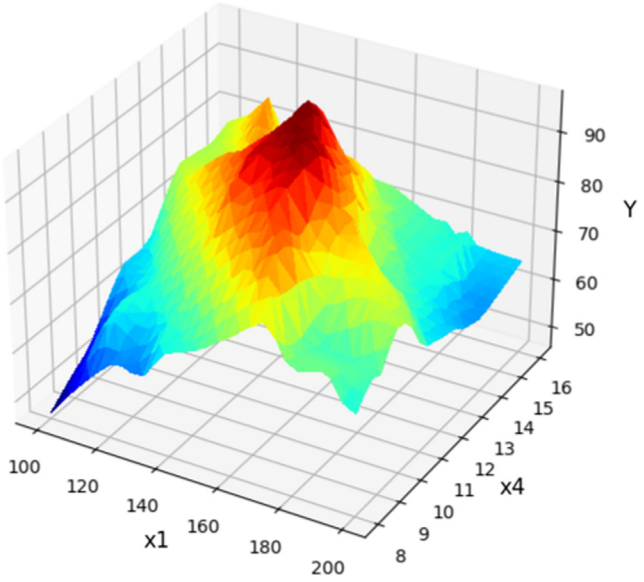

Fig. 9 X1 and X4 projection with prediction surface in final Boosted GPR model. X2 = 6 and X3 = 4 considered constant. Optimum value is y = 96.52 for x1 = 146 x4 = 13.09.

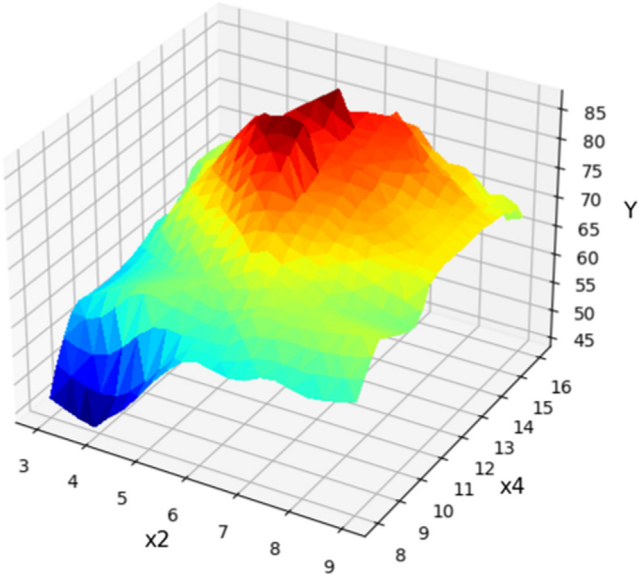

Fig. 11 X2 and X4 projection with prediction surface in final Boosted GPR model. X1 = 125 and X3 = 4 considered constant. Optimum value is y = 87.22 for x2 = 6, x4 = 14.3.

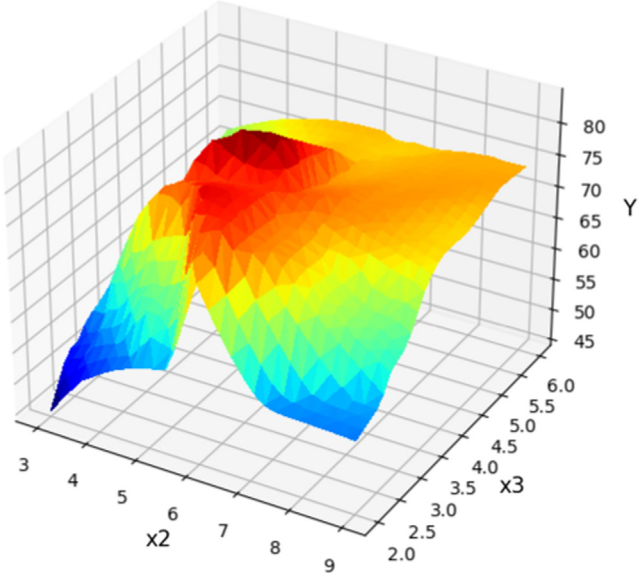

Fig. 10 X2 and X3 projection with prediction surface in final Boosted GPR model. X1 = 125 and X4 = 12 considered constant. Optimum value is y = 84.48 for x2 = 5.35, x3 = 3.69.

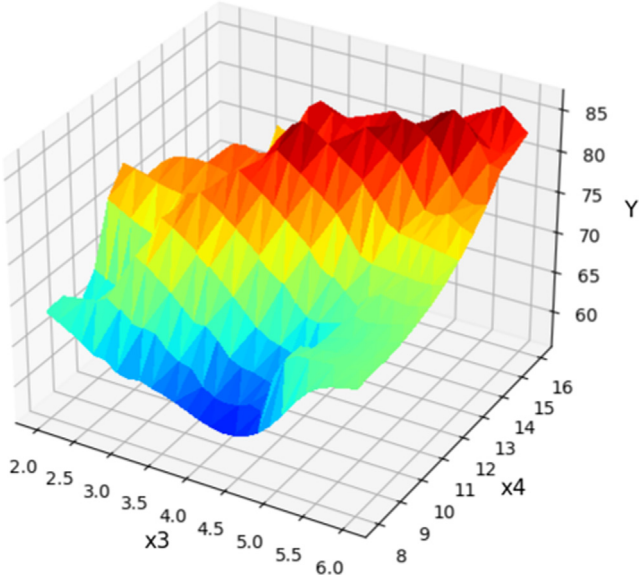

Fig. 12 X3 and X4 projection with prediction surface in final Boosted GPR model. X1 = 125 and X2 = 4 considered constant. Optimum value is y = 86.75 for x3 = 5.38, x4 = 14.66.

time, catalyst loading, and MeOH:oil ratio were considered in the model, while the only predicted output was the production yield (%). An attempt was made to find the optimum values of the input parameters to achieve the highest production yield. In the modeling of process, three ensemble models were used: boosted linear regression, boosted multi-layer perceptron, and forest of randomized tree. We then determined their optimum configurations, or hyper-parameters, for each model. This crucial search is carried out by executing over 1000 different combinations of hyper-parameters. Finally, R^2-Scores for Boosted Linear Regression, Boosted Multi-layer Perceptron, and Forest of Randomized Tree were 0.926, 0.998, and 0.992, respectively. The error rates for boosted linear regression, boosted multi-layer perceptron, and Forest of Randomized Tree were 5.6810^{-2}, 5.2010^{-2}, and 9.8310^{-2}, respectively, according to MAPE. It was indicated that boosted multi-layer perceptron was the best model among other models in terms of predictive accuracy.

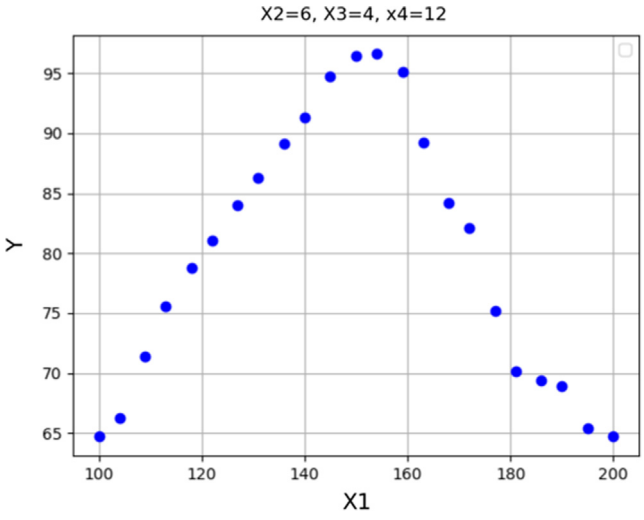

Fig. 13 Response trend for X1.

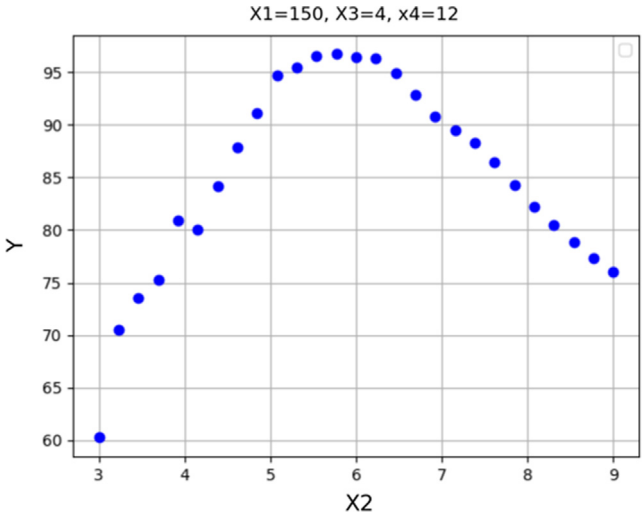

Fig. 14 Response trend for X2.

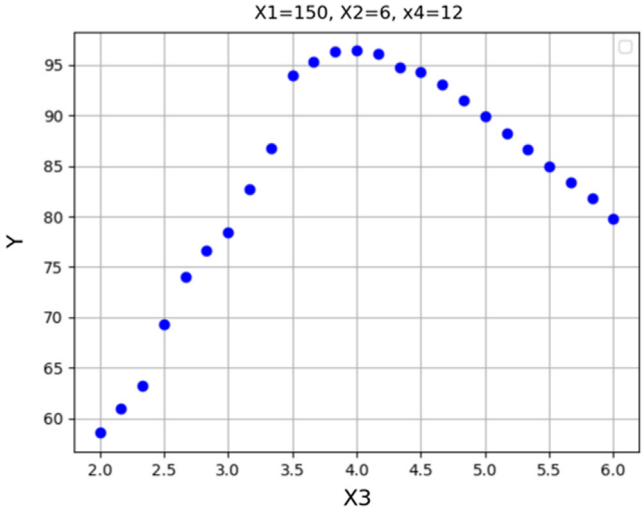

Fig. 15 Response trend for X3.

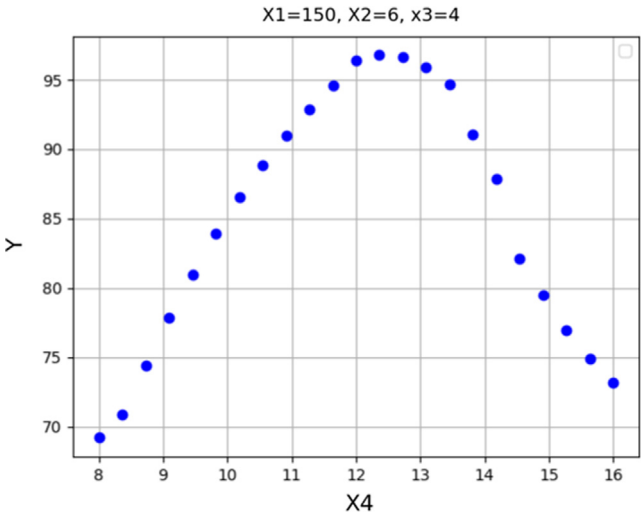

Fig. 16 Response trend for X4.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

This research was funded by Princess Nourah bint Abdulrahman University Researchers Supporting Project number (PNURSP2022R145), Princess Nourah bint Abdulrahman University, Riyadh, Saudi Arabia.

Tazeddinova Diana is thankful to the South Ural State University (Russian Federation) and Zhangir Khan Agrarian Technical University (Republic of Kazakhstan) for their help and support in completing this work.

References

Yahya, S., Muhamad Wahab, S.K., Harun, F.W., 2020. Optimization of biodiesel production from waste cooking oil using Fe-Montmorillonite K10 by response surface methodology. Renew. Energy 157, 164–172.

Adebayo, T.S. et al, 2022. The asymmetric effects of renewable energy consumption and trade openness on carbon emissions in Sweden: new evidence from quantile-on-quantile regression approach. Environ Sci Pollut Res Int 29 (2), 1875–1886.

Zhao, C. et al, 2022. Computational design of BC3N2 based single atom catalyst for dramatic activation of inert CO2 and CH4 gases into CH3COOH with ultralow CH4 dissociation barrier. Chin. Chem. Lett.

A.B.W, P. *Computer Technology Simulation towards Power Generation Potential from Coproduced Fluids in South Lokichar Oil Fields*. 2020.

Shen, Z. et al, 2021. A critical review of plant-based insulating fluids for transformer: 30-year development. Renew. Sustain. Energy Rev. 141, 110783.

Mao, J., et al., *The Effect of Marine Power Generation Technology on the Evolution of Energy Demand for New Energy Vehicles*. J. Coast. Res. 2020. 103(sp1): p. 1006–1009, 4.

Ekramian, E., S.G. Etemad, and M. Haghshenasfard. *Numerical Analysis of Heat Transfer Performance of Flat Plate Solar Collectors*. 2014.

Yosofvand, Mohammad et al, 2020. AdipoGauge software for analysis of biological microscopic images. Adipocyte 9 (1), 360–373. https://doi.org/10.1080/21623945.2020.1787583.

Zhang, S. et al, 2021. Prediction of energy photovoltaic power generation based on artificial intelligence algorithm. Neural Comput. Appl. 33.

Hu, S. et al, 2022. A preliminary study on the eco-environmental geological issue of in-situ oil shale mining by a physical model. Chemosphere 287, 131987.

Hu, Y., et al., *Recent Technologies for the Extraction and Separation of Polyphenols in Different Plants: A Review*. Journal of Renewable Materials.

Johnson, J., Shanmugam, R., Lakshmi, T., 2022. A review on plant-mediated selenium nanoparticles and its applications. J. Popul. Ther. Clin. Pharmacol. 28 (2), e29–e40.

Ibnou-Laaroussi, S., Rjoub, H., Wong, W.-K., 2020. Sustainability of Green Tourism among International Tourists and Its Influence on the Achievement of Green Environment: Evidence from North Cyprus. Sustainability 12 (14).

Rikani, A.S., 2021. Numerical analysis of free heat transfer properties of flat panel solar collectors with different geometries. Journal of Research in Science, Engineering and Technology 9 (1), 95–116.

Rjoub, H. et al, 2021. Sustainability of the Moderating Role of Financial Development in the Determinants of Environmental Degradation: Evidence from Turkey. Sustainability 13 (4).

Wang, B. et al, 2022. Synergistic effect on the co-gasification of petroleum coke and carbon-based feedstocks: A state-of-the-art review. J. Energy Inst. 102, 1–13.

Lin, W. et al, 2021. Understanding the effects of different residual lignin fractions in acid-pretreated bamboo residues on its enzymatic digestibility. Biotechnol. Biofuels 14 (1), 143.

Fu, C.-C. et al, 2010. Hydrolysis of microalgae cell walls for production of reducing sugar and lipid extraction. Bioresource Technology 101 (22), 8750–8754.

Nguyen, H.C. et al, 2021. Bio-Derived Catalysts: A Current Trend of Catalysts Used in Biodiesel Production. Catalysts 11 (7), 812.

Jia, N., J.S. Wang, and N. Li. Application of data mining in intelligent power consumption. in International Conference on Automatic Control and Artificial Intelligence (ACAI 2012). 2012.

Aslanova, F., *A comparative study of the hardness and force analysis methods used in truss optimization with metaheuristic algorithms and under dynamic loading*. Journal of Research in Science, Engineering and Technology. 8(1), 25–33.

Chen, T.-C. et al, 2021. Engineering of Novel Fe-Based Bulk Metallic Glasses Using a Machine Learning-Based Approach. Arab. J. Sci. Eng. 46 (12), 12417–12425.

Deng, W. et al, 2019. Evolution of Aromatic Structures during the Low-Temperature Electrochemical Upgrading of Bio-oil. Energy Fuels 33 (11), 11292–11301.

Liu, J. et al, 2021. Highly efficient photocatalytic degradation of oil pollutants by oxygen deficient SnO2 quantum dots for water remediation. Chem. Eng. J. 404, 127146.

Yin, G. et al, 2022. Machine learning method for simulation of adsorption separation: Comparisons of model's performance in predicting equilibrium concentrations. Arabian J. Chem. 15, (3) 103612.

Yin, G. et al, 2022. Multiple machine learning models for prediction of CO2 solubility in potassium and sodium based amino acid salt solutions. Arabian J. Chem. 15, (3) 103608.

Nair, K.G.S., Velmurugan, R., Sukumaran, S.K., 2020. Influence of polylactic acid and polycaprolactone on dissolution characteristics of ansamycin-loaded polymeric nanoparticles: An unsatisfied attempt for drug release profile. J. Pharm. Negative Results 11, 23–29.

Huang, C. et al, 2022. Lignin-enzyme interaction: A roadblock for efficient enzymatic hydrolysis of lignocellulosics. Renew. Sustain. Energy Rev. 154, 111822.

Adebayo, T.S., Rjoub, H., 2022. A new perspective into the impact of renewable and nonrenewable energy consumption on environmental degradation in Argentina: a time-frequency analysis. Environ. Sci. Pollut. Res. Int. 29 (11), 16028–16044.

Kansedo, J., Lee, K.T., Bhatia, S., 2009. Biodiesel production from palm oil via heterogeneous transesterification. Biomass Bioenergy 33 (2), 271–276.

Collotta, M. et al, 2019. Life Cycle Analysis of the Production of Biodiesel from Microalgae. In: Basosi, R. (Ed.), Life Cycle Assessment of Energy Systems and Sustainable Energy Technologies: The Italian Experience. Springer International Publishing, Cham, pp. 155–169.

Zaidel, D.N.A., et al., *18 - Production of biodiesel from rice bran oil*, in *Biomass, Biopolymer-Based Materials, and Bioenergy*, D. Verma, et al., Editors. 2019, Woodhead Publishing. p. 409–447.

Mohadesi, M. et al, 2019. Production of biodiesel from waste cooking oil using a homogeneous catalyst: Study of semi-industrial pilot of microreactor. Renew. Energy 136, 677–682.

Bakhshkandi, R. and M. Ghoranneviss. *Investigating the synthesis and growth of titanium dioxide nanoparticles on a cobalt catalyst*. 2019.

Nourian, P. et al, 2021. Implementation of active probe rheology simulation technique for determining the viscoelastic moduli of soft matter. Journal of Rheology 65 (4), 617–632.

Osanloo, M. et al, 2019. Larvicidal activity of chemically synthesized silver nanoparticles against Anopheles stephensi. J. Pharm. Negative Results 10, 69–72.

Watandost, H., Achak, J., Haqmal, A., 2021. Oxidation of hydrogels based of sodium alginate and MnO2 as catalyst. Int. J. Innov. Res. Sci. Stud. 4 (4), 191–199.

Andalib, V., Sarkar, J., 2021. A Repairable System Supported by Two Spare Units and Serviced by Two Types of Repairers. J. Stat. Theory Appl. 20, 180.

Zhao, C. et al, 2022. Electro-reduction of N2 on nanostructured materials and the design strategies of advanced catalysts based on descriptors. Mater. Today Phys. 22, 100609.

Latif, M. et al, 2021. Flowing atmospheric-pressure afterglow drift tube ion mobility spectrometry. Analytica Chimica Acta 1163, 338507.

Latif, M., Zhang, D., Gamez, G., 2021. Flowing atmospheric-pressure afterglow drift tube ion mobility spectrometry. Analytica Chimica Acta 1163, 338507.

Huang, C. et al, 2021. Gold Nanoparticles-Loaded Polyvinylpyrrolidone/Ethylcellulose Coaxial Electrospun Nanofibers with Enhanced Osteogenic Capability for Bone Tissue Regeneration. Mater. Des. 212, 110240.

Zhang, D., Latif, M., Gamez, G., 2021. Instantaneous Differentiation of Functional Isomers via Reactive Flowing Atmospheric Pressure Afterglow Mass Spectrometry. Anal. Chem. 93 (29), 9986–9994.

Wang, J. et al, 2021. Interruptible load scheduling model based on an improved chicken swarm optimization algorithm. CSEE J. Power Energy Syst 7 (2), 232–240.

Panwar, A. et al, 2021. Comprehensive modelling of pharmaceutical solvation energy in different solvents. J. Mol. Liq. 341, 117390.

Ghadiri, M. et al, 2021. An insight into the estimation of relative humidity of air using artificial intelligence schemes. Environ. Dev. Sustain. 23 (7), 10194–10222.

Pourtousi, Z. et al, 2021. Ability of neural network cells in learning teacher motivation scale and prediction of motivation with fuzzy logic system. Sci. Rep. 11 (1), 9721.

Shirazian, S. et al, 2017. Artificial neural network modelling of continuous wet granulation using a twin-screw extruder. Int. J. Pharm. 521 (1–2), 102–109.

Ismail, H.Y. et al, 2019. Developing ANN-Kriging hybrid model based on process parameters for prediction of mean residence time

distribution in twin-screw wet granulation. Powder Technol. 343, 568–577.
PUTRA, A.B.W, 2020. Computer Technology Simulation towards Power Generation Potential from Coproduced Fluids in South Lokichar Oil Fields. International Journal of Communication and Computer Technologies 8 (2), 9–12. https://doi.org/10.31838/ijccts/08.02.03.
Rezakazemi, M. et al, 2018. Development of hybrid models for prediction of gas permeation through FS/POSS/PDMS nanocomposite membranes. Int. J. Hydrogen Energy 43 (36), 17283–17294.
Dashti, A. et al, 2018. Estimating CH4 and CO2 solubilities in ionic liquids using computational intelligence approaches. J. Mol. Liq. 271, 661–669.
Pishnamazi, M. et al, 2020. ANFIS grid partition framework with difference between two sigmoidal membership functions structure for validation of nanofluid flow. Sci. Rep. 10 (1), 1–11.
Marjani, A., Babanezhad, M., Shirazian, S., 2020. *Application of adaptive network-based fuzzy inference system (ANFIS) in the numerical investigation of Cu/water nanofluid convective flow*. Case Studies. Therm. Eng. 100793
Andalib, V., Sarkar, J., 2022. A System with Two Spare Units, Two Repair Facilities, and Two Types of Repairers. Mathematics 10, (6). https://doi.org/10.3390/math10060852 852.
Babanezhad, M. et al, 2020. Artificial intelligence simulation of suspended sediment load with different membership functions of ANFIS. Neural Comput. Appl., 1–15
Babanezhad, M. et al, 2020. Bubbly flow prediction with randomized neural cells artificial learning and fuzzy systems based on k–ε turbulence and Eulerian model data set. Sci. Rep. 10 (1).
Babanezhad, M., Nakhjiri, A.T., Shirazian, S., 2020. Changes in the Number of Membership Functions for Predicting the Gas Volume Fraction in Two-Phase Flow Using Grid Partition Clustering of the ANFIS Method. ACS Omega 5 (26), 16284–16291.
Babanezhad, M. et al, 2020. *gbell Learning function along with Fuzzy Mechanism in Prediction of Two-Phase Flow*. ACS Omega.
Babanezhad, M. et al, 2020. Pattern recognition of the fluid flow in a 3D domain by combination of Lattice Boltzmann and ANFIS methods. Sci. Rep. 10 (1), 1–13.
Nabipour, N. et al, 2020. Prediction of Nanofluid Temperature Inside the Cavity by Integration of Grid Partition Clustering Categorization of a Learning Structure with the Fuzzy System. Acs Omega 5 (7), 3571–3578.
Sundaravadivelu Devarajan, D. et al, 2020. Molecular simulation of nanocolloid rheology: Viscosity, viscoelasticity, and time-concentration superposition. Journal of Rheology 64 (3), 529–543.
Tian, E. et al, 2020. Simulation of a Bubble-Column Reactor by Three-Dimensional CFD: Multidimension- and Function-Adaptive Network-Based Fuzzy Inference System. Int. J. Fuzzy Syst. 22 (2), 477–490.
El Naqa, I., Murphy, M.J., 2015. What is machine learning? In: machine learning in radiation oncology. Springer, pp. 3–11.
Goodfellow, I., Bengio, Y., Courville, A., 2016. Machine learning basics. Deep learning 1 (7), 98–164.
Aghbashlo, M. et al, 2021. Machine learning technology in biodiesel research: A review. Prog. Energy Combust. Sci. 85, 100904.
Gupta, K., et al., *Machine Learning-Based Predictive Modelling of Biodiesel Production—A Comparative Perspective. Energies 2021, 14, 1122*. 2021, s Note: MDPI stays neutral with regard to jurisdictional claims in published ….
Maclin, R., Opitz, D., 1997. An empirical evaluation of bagging and boosting. AAAI/IAAI 1997, 546–551.
Zhou, Z.-H., *Ensemble methods: foundations and algorithms*. 2019: Chapman and Hall/CRC.
Rokach, L. and O.Z. Maimon, *Data mining with decision trees: theory and applications*. Vol. 69. 2007: World scientific.
Breiman, L., et al., *Classification and regression trees*. 2017: Routledge.
DR.S.SRINIVASAREDDY, D.Y.V.N., DR.D.KRISHNA, *Sector Beam Synthesis in Linear Antenna Arrays using Social Group Optimization Algorithm*. National Journal Of Antennas And Propagation, 2021. 3 (2), 6–9.
White, H., 1992. Artificial neural networks. Blackwell Cambridge, Mass.
Ethier, J.G. et al, 2021. Microrheology analysis in molecular dynamics simulations: Finite box size correction. Journal of Rheology 65 (6), 1255–1267.
Ferreira, A.J., Figueiredo, M.A., 2012. Boosting algorithms: A review of methods, theory, and applications. Ensemble Mach. Learn., 35–85
Jain, A.K., Mao, J., Mohiuddin, K.M., 1996. Artificial neural networks: A tutorial. Computer 29 (3), 31–44.
Soltani Fesaghandis, G. et al, 2017. Comparison of Multilayer Perceptron and Radial Basis Function in Predicting Success of New Product Development. Eng. Technol. Appl. Sci. Res. 7.
Noriega, L., *Multilayer perceptron tutorial*. School of Computing. Staffordshire University, 2005.
Elmaz, F., Yücel, Ö., Mutlu, A.Y., 2020. Predictive modeling of biomass gasification with machine learning-based regression methods. Energy 191, 116541.
Zhou, X. et al, 2018. Feature extraction and physical interpretation of melt pressure during injection molding process. J. Mater. Process. Technol. 261, 50–60.
Yang, J. et al, 2008. Computation of two-layer perceptron networks' sensitivity to input perturbation. IEEE.
Hecht-Nielsen, R., 1992. Theory of the backpropagation neural network. In: Neural networks for perception. Elsevier, pp. 65–93.
Pombeiro, H. et al, 2017. Comparative assessment of low-complexity models to predict electricity consumption in an institutional building: Linear regression vs. fuzzy modeling vs. neural networks. Energy Build. 146, 141–151.
Kim, M.K., Kim, Y.-S., Srebric, J., 2020. Predictions of electricity consumption in a campus building using occupant rates and weather elements with sensitivity analysis: Artificial neural network vs. linear regression. Sustain. Cities Soc. 62, 102385.
Xue, M. et al, 2020. Identification of potential type II diabetes in a large-scale Chinese population using a systematic machine learning framework. J. Diabetes Res. 2020.
Jiang, R. et al, 2009. A random forest approach to the detection of epistatic interactions in case-control studies. BMC Bioinf. 10 (1), 1–12.
Rodriguez-Galiano, V. et al, 2015. Machine learning predictive models for mineral prospectivity: An evaluation of neural networks, random forest, regression trees and support vector machines. Ore Geol. Rev. 71, 804–818.
Zhang, J. et al, 2022. Rapid evaluation of texture parameters of Tan mutton using hyperspectral imaging with optimization algorithms. Food Control 108815.
Geurts, P., Ernst, D., Wehenkel, L., 2006. Extremely randomized trees. Mach. Learn. 63 (1), 3–42.
Dutta, S., Mukherjee, U., Bandyopadhyay, S.K., 2021. Pharmacy Impact on Covid-19 Vaccination Progress Using Machine Learning Approach. J. Pharm. Res. Int., 202–217
Freund, Y., Schapire, R.E., 1997. A decision-theoretic generalization of on-line learning and an application to boosting. J. Comput. Syst. Sci. 55 (1), 119–139.
Buitinck, L., et al., *API design for machine learning software: experiences from the scikit-learn project*. arXiv preprint arXiv:1309.0238, 2013.
Pedregosa, F., et al., Scikit-learn: Machine learning in Python. J. Mach. Learn. Res. 2011. 12: p. 2825–2830.