



Efficiency and accuracy comparison of machine learning algorithms for predicting US energy consumption across sectors

Seyed Matin Malakouti^{*}, Mohammad Bagher Menhaj, Amir Abolfazl Suratgar

Distributed and Intelligent Optimization Research Laboratory, Department of Electrical Engineering, Amirkabir University of Technology, Tehran, Iran

ARTICLE INFO

Keywords:

Energy consumption prediction
Machine learning algorithms
Ridge algorithm
Sector-wise analysis
Power outage mitigation

ABSTRACT

This study uses machine learning algorithms to predict energy consumption across the United States's critical sectors (commercial, residential, transportation, and industrial). The Ridge algorithm emerged as the most accurate and efficient in predicting energy consumption, outperforming other algorithms such as Lasso Regression, Elastic Net, Extra Tree, Random Forest, K Neighbors, and Orthogonal Matching Pursuit. The study employed data collection, feature engineering, model training, and evaluation based on metrics like MAPE, RMSLE, RMSE, MSE, and MAE, alongside speed evaluations. Key findings highlight the superior performance of the Ridge algorithm across residential, industrial, and commercial sectors in terms of accuracy (low MSE values) and computational efficiency. Additionally, the Orthogonal Matching Pursuit algorithm showed promise in predicting energy consumption in the transportation sector. These results provide valuable insights for energy management strategies, emphasizing the importance of accurate energy consumption predictions in effective power distribution and outage risk reduction.

Acronyms

AANNs	Artificial Neural Networks
SVR	Support Vector Regression
RF	Random Forest
M5Rule	M5Rules (a rule-based algorithm)
KNN	K Nearest Neighbors
ELM	Extreme Learning Machines
DT	Decision Trees
GP	Gaussian Processes
DT	Decision Trees
GBR	Gradient Boosting Regression
LR	Linear Regression
OMP	Orthogonal Matching Pursuit
ERT	Extra Trees or Extremely Randomized Trees

1. Introduction

One of the leading international initiatives is to reduce energy consumption in buildings. For instance, using renewable energy sources and saving energy have been the main goals of the European energy policy (D'Agostino et al., 2017). The effective use of energy helps improve the quality of life on Earth and prevents CO₂ emissions. Building buildings

use a significant share of the world's energy usage (J Naji et al., 2016; Chou and Tran, 2018; Chou et al., 2016). For instance, roughly a 40percent of the energy used in the United States is used by businesses (J Naji et al., 2016, Bokalders, 2010). Additionally, structures need power for approximately 5 Decades or longer throughout their operating phase. For a green economy, buildings must be efficiently constructed and operated in terms of energy use. Scientists have given energy conservation in structures their complete focus (Chou et al., 2016, Pham et al., 2020). Planning, managing, and conserving energy depends on anticipating energy usage in buildings (Zhong et al., 2019). Data-driven methods for the forecast of energy use in structures have been thoroughly examined by Amasyali (Amasyali and El-Gohary, 2018), Wei and Colleague (Wei et al., 2018), and Deb and Colleague (Deb et al., 2017). AANNs, SVR, gradient boosting, and extreme learning machines are just a few of the Machine learning that play a critical role in forecasting building energy trends (J Naji et al., 2016). Predictions of building energy use have been made using Machine learning, including support vector machines (Zhong et al., 2019), ELM (J Naji et al., 2016), artificial neural networks (Chaudhuri et al., 2019), and deep recurrent neural networks (Rahman et al., 2018, Tian et al., 2019). Gao and colleagues (Gao et al., 2019) compared the performance of sixteen Machine Learning algorithms, including RF, M5Rule, ANNs, and SVR methods, in

^{*} Corresponding author.

E-mail address: matin3475malakoti@gmail.com (S.M. Malakouti).

developing environmentally friendly home layouts. Their research showed that RF was highly predictive of HVAC loads in buildings. Wind power predictions were also made using RF, M5Rules, KNN, SVR, and ANN methods (Renani et al., 2016). Nasruddin and colleagues (Nasruddin et al., 2019) used artificial neural networks and evolutionary approaches to determine the most efficient energy use inside a structure. Predicting hotel energy use for HVAC systems required extensive roadway networks and more trees (Ahmad et al., 2018). Predictions of home HVAC loads were made using the XGBoost method (Al-Rakhami et al., 2019). The ensemble approach has been used to improve accuracy in solving prediction problems. For example, ensemble ML models have confirmed their power in predicting cooling loads to enhance buildings' energy-efficient design (Ngo, 2019). Araya et al. (Araya et al., 2017) developed an ensemble learning framework to identify anomalous consumption patterns in buildings. Their system can notify the building manager to perform appropriate energy-saving solutions. In addition, an ensemble approach has been applied successfully for the shear strength prediction of reinforced concrete beams (Chou et al., 2019) and for modeling the corrosion rate of steel structures (Chou et al., 2017). To forecast the total energy demand for electricity use in buildings, Ahmad and colleagues (Ahmad et al., 2018) used ML models. DT, GP, GPR, and LR are employed in their research. The values of long-term energy commodities were also predicted using Machine learning. Using monthly pricing, the scientists contrasted the forecasting abilities of Machine learning with conventional econometric methods (Herrera et al., 2019). Their findings confirmed the effectiveness of Machine Learning in creating a reliable prediction method to assist policymakers in the global energy market. Wang and colleagues reported an ensemble bagging tree algorithm for forecasting power usage in structures (Wang et al., 2018). In (Papadopoulos et al., 2018), tree-based ensemble learning methods predicted structural heating and cooling needs. GBR, ELM, LR, and SVR were all used by Huang and colleagues (Huang et al., 2019) to create an ensemble model to predict heating demand in apartment complexes. This research also illustrates how the quantities projected using the adaptive neuro-fuzzy inference technique, namely the hybrid-based model, might be used to forecast the efficiency of the ground-coupled heat pump system in a much more precise manner.

2. Literature review

The study by (Adebayo et al., 2023 Feb 1) investigates how country risks and renewable energy consumption impact environmental quality in Mexico, Indonesia, Nigeria, and Turkey (MINT) nations. CS-ARDL modeling shows that economic growth, political risk, urbanization, and trade openness increase the ecological footprint, while economic and financial risks, along with renewable energy use, improve environmental quality. Similarly, (Akram et al., 2023 Oct 15) emphasizes carbon neutrality as a crucial policy for sustainable development in G7 economies. It examines factors like natural resource dependence, eco-innovation, green energy, carbon tax, and environmental policy stringency in achieving carbon neutrality. Additionally, (Alola and Adebayo, 2023 Aug 20) focuses on Iceland's CAP 2020 and its targets for reducing greenhouse gas emissions, stressing the need for sustainable material use to achieve environmental goals. In China, (Adebayo and Ullah, 2023 Jul 1) explores how economic growth, financial development, nuclear energy, government stability, and socioeconomic conditions affect environmental quality, showing positive impacts of nuclear energy and government stability.

Moreover, (Ullah et al., 2023 Sep 15) highlights the positive effects of environment-related ICT innovations (EICT), financial development (FD), and human development (HD) on energy transition and GHG mitigation in G-7 economies, suggesting investments in clean technologies and education. In the United States, (Liu et al., 2023 Sep 10) studies the relationship between CO₂ emissions, coal efficiency, climate policy uncertainty, green energy, and green innovation to achieve CO₂ reduction goals. Lastly, (Hu et al., 2023 Dec 1) discusses the

interconnectedness of Canada's carbon dioxide emissions, financial development, renewable electricity, and economic growth.

These studies offer empirical evidence and policy recommendations for addressing environmental challenges and achieving sustainable development goals through targeted interventions and international cooperation, emphasizing the importance of holistic policies and investments in green technologies.

The body of this work is structured as follows: Section 2 contains details of the approaches used. Section 3 describes the data preparation steps. The results and comments are discussed in Section 4, and the article concludes in part 5.

3. Research algorithms

3.1. Lasso regression

Penalized regression technique is another name for Lasso regression. Machine learning often uses this technique to pick the subset of variables. It delivers more accurate predictions than other regression models. Lasso Regularization helps to boost model understanding. Lasso is a linear regression in which the model is penalized for the sum of the absolute weight values. Thus, the absolute weight values will (primarily) decrease, and a large proportion will likely be zero.

3.2. Elastic Net

Elastic Net uses L1 and L2 penalties to punish linear regression when training. Hyperparameters, such as "alpha," establish the relative relevance of the L1 and L2 penalties in "The Elements of Statistical Learning." To avoid overfitting, Lasso will remove certain features from your linear model. With the help of Ridge, you may reduce the impact of irrelevant characteristics on your y-value prediction. Elastic Net combines feature elimination from Lasso with feature coefficient reduction from the Ridge model to improve your model's predictive power.

3.3. Ridge Regression

Advantages. Ridge Regression overcomes the issue of overfitting, which occurs when ordinary squared error regression fails to distinguish less significant characteristics and utilizes them all, resulting in overfitting. To match the model to the actual values of the data, ridge regression incorporates a slight bias. Ridge regression reduces the complexity of a model but does not lower the number of variables since it never results in a coefficient being zero but minimizes it. Consequently, this model is not suitable for feature reduction. Ridge regression often decreases model overfitting and incorporates all model characteristics. It reduces the model's complexity by reducing the number of coefficients. Additionally, the Lasso regression reduces feature selection by lowering model overfitting.

3.4. Extra Tree

Extra Trees or Extremely Randomized Trees (ERT) is an ensemble approach for machine learning that aggregates the predictions of several decision trees. It is similar to the popular random forest method. Extremely Randomized Trees (ERT) resemble Random Forests in many ways. There are two significant differences: ERT does not resample data while creating a tree, and ERT does not use a random forest. They do not engage in bagging. The ERT does not use the "best split. Among machine learning algorithms, Random Forest is one of the most often used for Classification and Regression. The majority vote is used for classification, and the mean is used for regression to build decision trees from several samples.

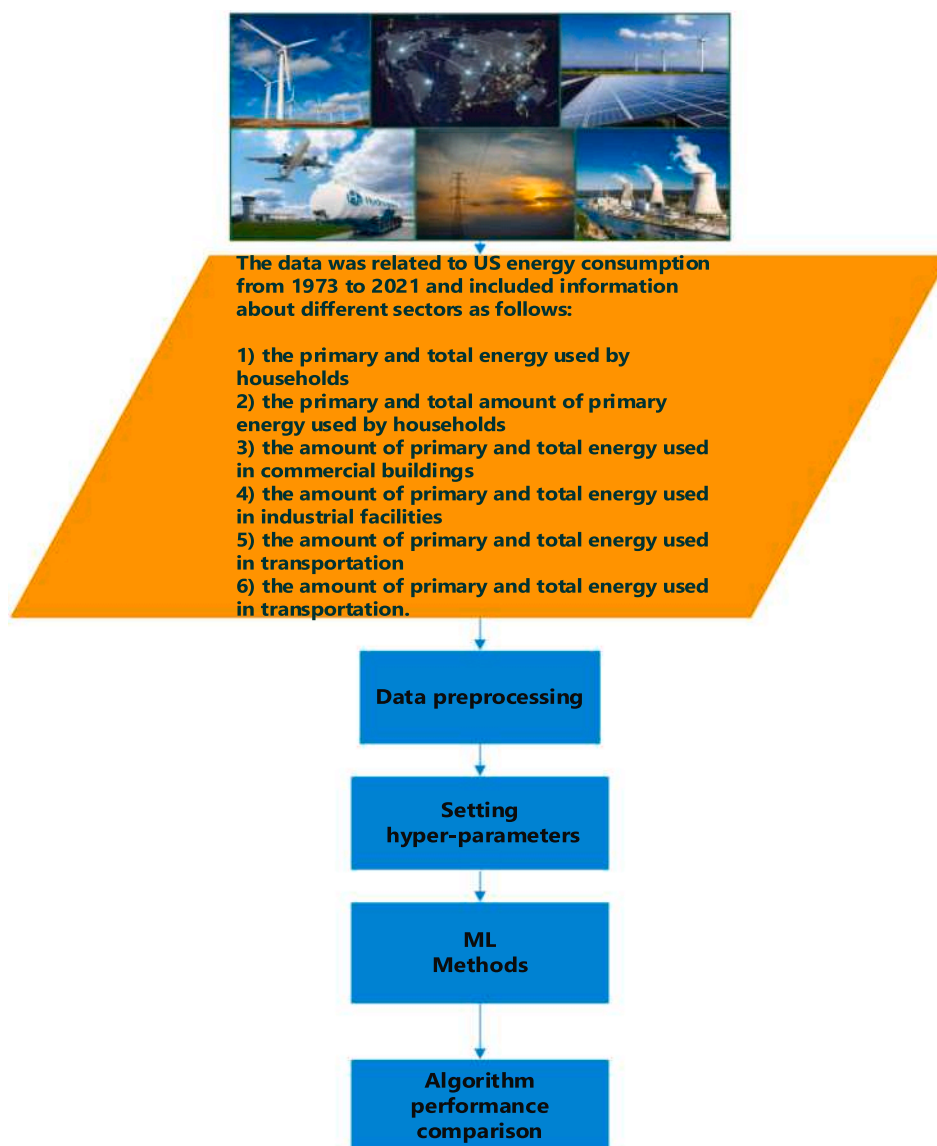


Fig. 1. Block diagram of how to predict energy consumption in America.

3.5. Random Forest

Random Forest Regression is an approach for supervised learning that uses ensemble learning for regression. Ensemble learning is a methodology that combines the predictions of numerous machine learning algorithms to provide more accurate forecasts than a single model. Random Forests may also be used for regression applications in addition to classification. The nonlinear structure of a Random Forest might offer it an advantage over linear algorithms, making it an excellent choice. Random Forest is more accurate than other algorithms. It features an efficient approach for guessing missing data and retains accuracy even when a significant amount is missing.

3.6. K neighbors

For example, KNN regression is a non-parametric method that approximates a link between independent variables and continuous outcomes by averaging data from the same neighborhood. The KNN method may be used to resolve problems in classification and regression. The KNN algorithm uses 'feature similarity' to forecast the values of new data points. The unique point receives a discount based on similarity to

other points in the training set. Classification and regression problems may be handled using the KNN, a fundamental supervised machine learning methodology.

3.7. Orthogonal Matching Pursuit

Orthogonal Matching Pursuit (OMP) is the most widely used greedy technique for locating a sparse solution vector to an underdetermined linear system of equations. OMP employs the projection method to determine the support indices of the light solution vector.

4. Methods

The information about the data is shown in Fig. 1. The energy consumption of 34 years was selected as training data, and the remaining 14 years were considered test data. First, the outliers were removed, and the data were normalized. Then, energy consumption was predicted using a 10-fold cross-validation method (Malakouti et al., 2022 Dec; Malakouti, 2023 Jun 1; Malakouti, 2023 Jul 1; Malakouti, 2023 Dec 1; Malakouti, 2023 Jun 1; Malakouti and Ghiasi, 2022 May 11; Malakouti et al., 2023 Aug 1; Malakouti et al., 2022 Jan 1; Malakouti, 2023 Mar;

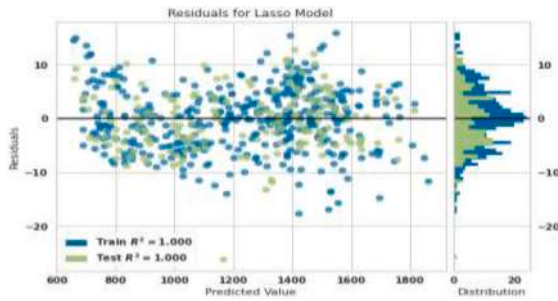


Fig. 2. Residual plot for Commercial with Lasso.

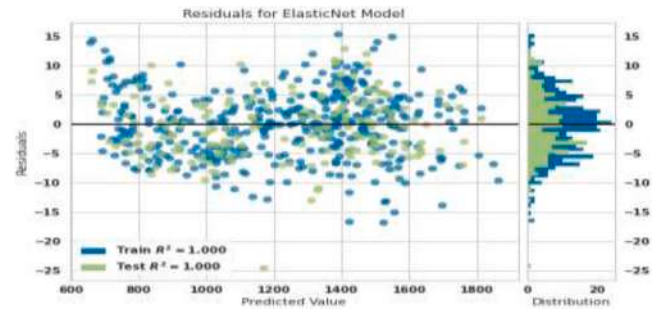


Fig. 3. residual plot for Commercial with elastic Net.

Malakouti, 2023 Mar 16; Malakouti, 2023 Sep 1; Malakouti, 2023; Malakouti et al., 2023 Dec 1; Malakouti, 2023 May 1; Malakouti et al., 2024 Mar 1) and hyper-parameter adjustment of Lasso Regression, Elastic Net, Ridge Regression, Extra Tree, Random Forest, K neighbors, and Orthogonal Matching Pursuit algorithms. Each algorithm can be used as a model for predicting energy consumption. Finally, the results of the algorithms in predicting energy consumption were compared with each other.

5. Results and discussion

The formula of MAPE, RMSLE, RMSE, MSE, and MAE evaluation criteria that was reviewed in the article is as follows:

N = number of data points

ϕ_i = actual values

$\hat{\phi}_i$ = predicted values

$$MSE = \frac{1}{N} \sum_{i=1}^N (\phi_i - \hat{\phi}_i)^2 \quad (1)$$

$$RMSE = \sqrt{\frac{\sum_{i=1}^N (\phi_i - \hat{\phi}_i)^2}{N}} \quad (2)$$

$$MAE = \frac{\sum_{i=1}^N |\phi_i - \hat{\phi}_i|}{N} \quad (3)$$

$$RMSLE = \sqrt{\frac{1}{N} \sum_{i=1}^N (\log(\phi_i) - \log(\hat{\phi}_i))^2} \quad (4)$$

$$MAPE = \frac{1}{N} \sum_{i=1}^N \left| \frac{\phi_i - \hat{\phi}_i}{\phi_i} \right| \quad (5)$$

A **residual plot** shows the difference between predicted and actual values. The ideal residual plot, also known as the null residual plot, depicts a random dispersion of points, creating a band of essentially constant width around the identity line. Figs. 4, 12, 20, and 32 are ideal residual plots because the residuals have a small range and are placed in the form of a narrow band around the identity line.

Fig. 2 shows the residual plot for the Commercial with Lasso. According to Fig. 2, the residuals range between -30 and 20 . With a bit of precision, it can be seen in the residuals of the training and test data that the residuals have the highest number in the range of -15 and 15 . Fig. (9-a) shows the lasso algorithm's commercial error for one year. As it is known, the error range of -6 to 13 has been obtained, and these error values are in the range of residuals in Fig. 2. This analysis showed that the lasso algorithm worked correctly.

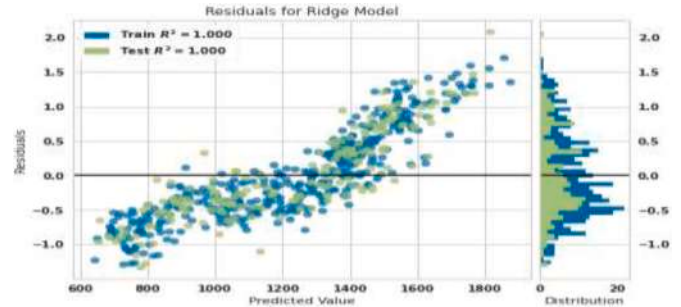


Fig. 4. Residual plot for Commercial with Ridge.

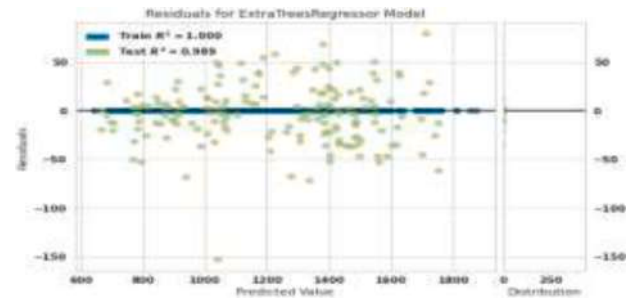


Fig. 5. Residual plot for Commercial with extra tree.

Fig. (9-a) shows the error for the Commercial with Lasso. February and June had the highest forecast error.

Fig. 3 shows the residual plot for Commercials with an elastic Net. According to Fig. 3, the residuals are in the range of -27 and 17 , and this Figure is very similar to Fig. 2. With a bit of precision, it can be seen in the residuals of the training and test data that the residuals have the highest number in the range of -15 and 15 . Fig. (9-b) shows the one-year commercial error with the elastic net algorithm. As it is known, the error range of -7 to 13 has been obtained, and these error values are in the range of residuals in Fig. 3. This analysis showed that the elastic net algorithm worked correctly. Elastic Net and lasso algorithms were very similar in the way they worked.

Fig. (9-b) shows the error for commercials with Elastic Net. February and June had the highest forecast errors, and October, July, and November had zero errors. This algorithm behaved similarly to the Lasso algorithm.

Fig. 4 shows the residual plot for Commercial with Ridge. According to Fig. 4, the residuals are between -1.5 and 2.25 . With a bit of precision, it can be seen in the residuals of the training and test data that the residuals have the highest number in the range of -1.5 and 2 . Fig. (9-c) shows the ridge algorithm's one-year commercial error. As it is known, the error range of -1 to 1.5 has been obtained, and this range of error values is in the range of residuals in Fig. 4. This analysis showed that the

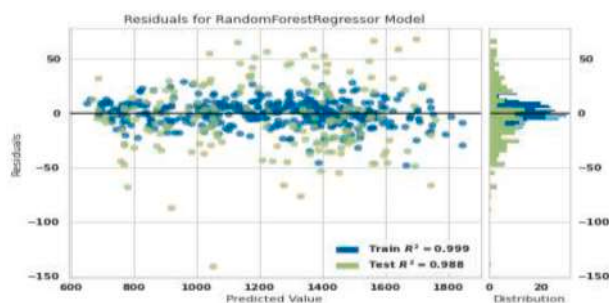


Fig. 6. Residual plot for Commercial with random forest.

ridge algorithm worked correctly. The ridge algorithm had the best performance.

Fig. (9-c) shows the error for the Commercial with Ridge. The Ridge algorithm should be used in the Commercial section.

Fig. 5 shows the residual plot for the Commercial with an extra tree. According to Fig. 5, the residuals range between -160 and 80 . With a bit of precision, it can be seen in the residuals of the training and test data that the residuals have the highest number in the range of -50 and 50 . Fig. (9-d) shows the one-year commercial error with the extra tree algorithm. As it is known, the error range of 40 to -50 has been obtained, and this range of error values is in the range of residuals in Fig. 5. This analysis showed that the extra tree algorithm worked correctly.

Fig. (9-d) shows the error for Commercial with et. June had the highest forecast error, and October and September had close to zero errors.

Fig. 6 shows the residual plot for Commercials with random forests. According to Fig. 6, the residuals are between -150 and 80 . With a bit of precision, it can be seen in the residuals of the training and test data that

the residuals have the highest number in the range of -70 and 80 . Fig. (9-e) shows a one-year commercial error with a random forest algorithm. As it is known, the error range of 80 to -60 has been obtained, and this range of error values is in the range of residuals in Fig. 6. This analysis showed that the random forest algorithm worked correctly.

Fig. (9-e) shows the error for commercials with rf. June had the highest forecast error, and October, November, September, and December had zero errors.

Fig. 7 shows the residual plot for Commercial with k neighbors. According to Fig. 7, the residuals range between -200 and 150 . With a bit of precision, it can be seen in the residuals of the training and test data that the residuals have the highest number in the range of -150 and 150 . Fig. (9-f) shows a one-year commercial error with the k neighbors algorithm. As it is known, the error range of -80 to 150 has been obtained, and this range of error values is in the range of residuals in Fig. 7. This analysis showed that the k neighbors algorithm worked correctly.

Fig. (9-f) Error for Commercial with knn. February had the highest forecast error. July, August, and September had zero errors.

Fig. 8 shows the residual plot for Commercial with OMP. According to Fig. 8, the residuals range between -140 and 100 . With a bit of precision, it can be seen in the residuals of the training and test data that the residuals have the highest number in the range of -140 and 100 . Fig. (9-g) shows a one-year commercial error with the OMP algorithm. As it is known, the error range of -60 to 75 has been obtained, and this range of error values is in the range of residuals in Fig. 8. This analysis showed that the OMP algorithm worked correctly.

Fig. (9-g) shows the error for Commercial with OMP. February had the highest forecast error. October and July had zero errors.

Table 1 shows MAPE, RMSLE, RMSE, MSE, and MAE evaluation criteria for the Lasso Regression, elastic Net, Ridge, extra tree, random forest, k neighbors, and Orthogonal Matching Pursuit algorithms presented in the Commercial section. The ridge algorithm was the most

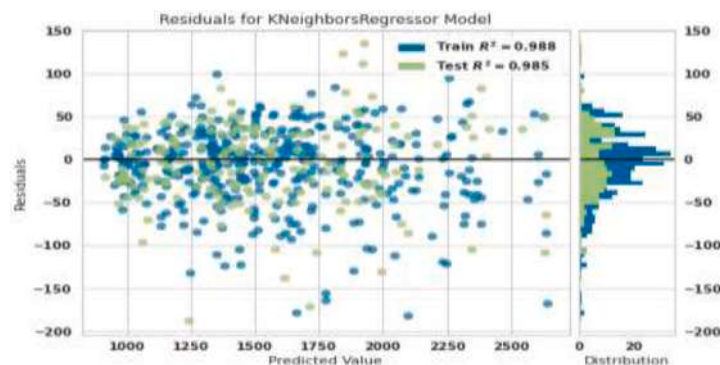


Fig. 7. Residual plot for Commercial with k neighbors.

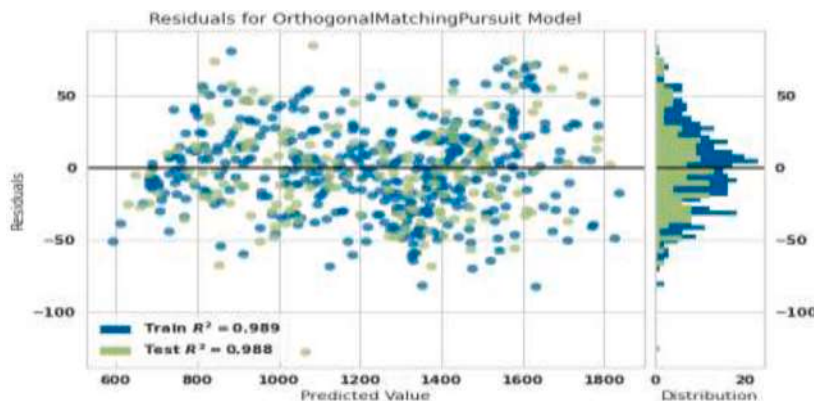


Fig. 8. Residual plot for Commercial with OMP.

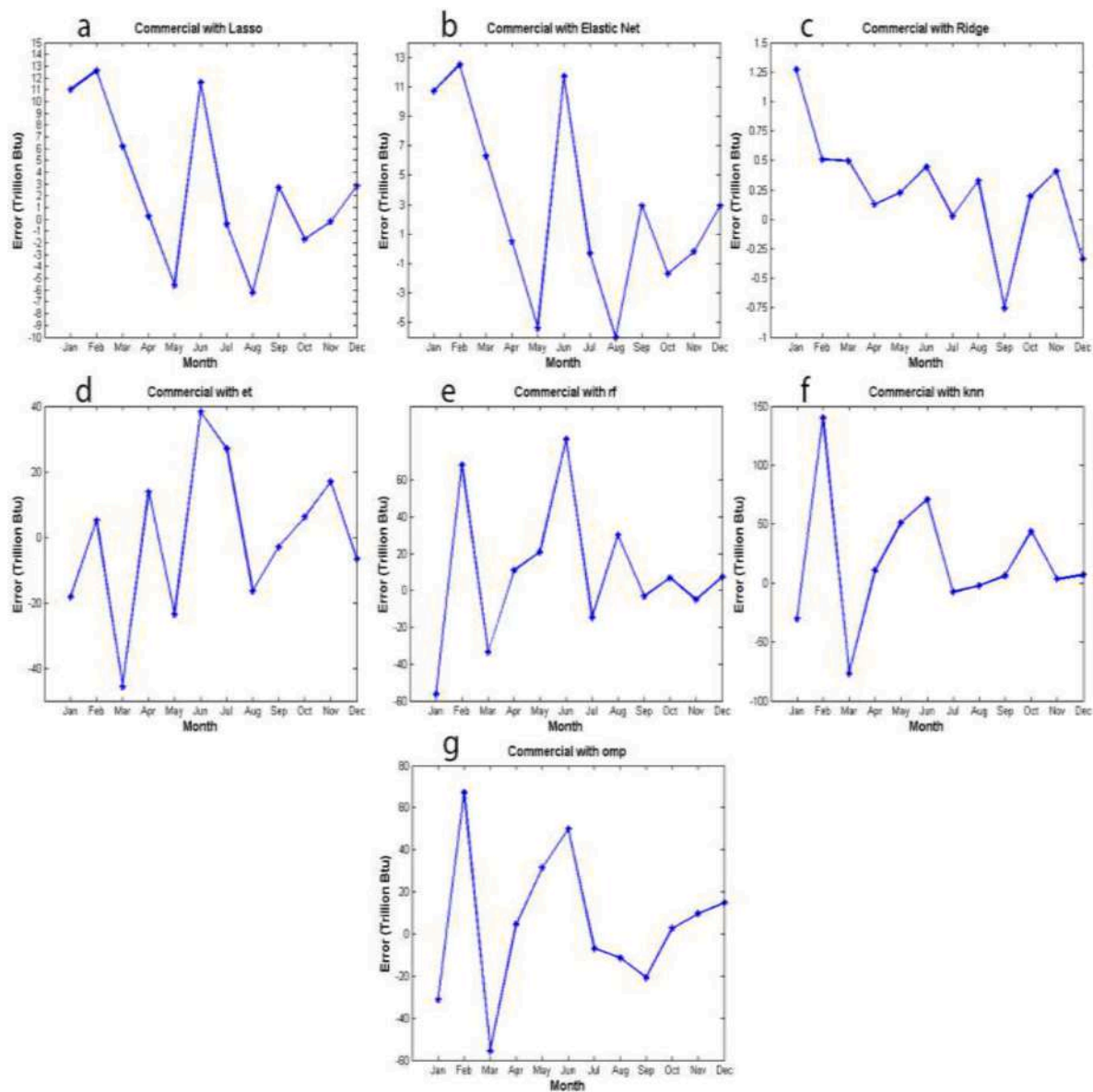


Fig. 9. Error for Commercial with (9-a) lasso, (9-b) elastic net, (9-c) ridge, (9-d) extra tree, (9-e) random forest, (9-f) k neighbors, (9 g) OMP for one year.

Table 1						
MAPE, RMSLE, RMSE, MSE, and MAE evaluation criteria for Commercial.						
Sector	MAE	MSE	RMSE	RMSLE	MAPE	Time (Sec)
lasso	3.59	22	4.66	0.0039	0.003	3
elastic Net	3.6	22.1	4.66	0.0039	0.003	2.8
ridge	1.12	2.61	1.33	0.0013	0.001	2
extra tree	21.6	754.6	27.36	0.0231	0.0183	9.36
random forest	24.91	989.42	31.24	0.0268	0.0213	4
k neighbors	27.78	1261.3	35.28	0.0308	0.024	3
Orthogonal Matching Pursuit	25.17	965.87	30.97	0.0271	0.0217	1.5

powerful for predicting power consumption in the Commercial area. Iran’s residential and commercial energy consumption in this research was anticipated using three distinct ML techniques: LR, logarithmic multiple linear regression, and nonlinear autoregressive with exogenous input artificial neural networks. The outcomes determined the best R2 value to be 0.99902 (Liu et al., 2023 Sep 10). The range of residuals in the Lasso algorithm is between –30 and 20

(Fig. 2). Also, the evaluation criterion of R2=1 for the test data was obtained by this algorithm. The range of residuals in the Ridge algorithm is between –1.5 and 2.5 (Fig. 4). Also, the evaluation criterion of R2=1 for the test data was obtained by this algorithm. Maybe this question arises as to why there is this difference in the range of residuals of Lasso and Ridge algorithms if R2=1 is the same for both algorithms. This is because the unit of energy under consideration is Trillion BTU, which is natural. After all, it is a very large-scale unit. Carefully, in Table 1, the MSE evaluation criterion for the lasso algorithm was assigned a value of 22, and the MSE evaluation criterion for the ridge algorithm was a value of 2.61. As can be seen, the MSE evaluation criterion for the lasso algorithm is about ten times the MSE evaluation criterion for the ridge algorithm. It can be found carefully in the residual range of the ridge and lasso algorithm that the residual range is about ten times the residual range of the ridge algorithm, which indicates the validity of the obtained results. The range of residuals in the random forest algorithm is between –150 and 90 (Fig. 6). Also, this algorithm obtained the evaluation criterion of R2=0.988 for the test data. Despite the significant criterion of Trillion BTU, despite the range of residuals between –150 and 90, R2 =

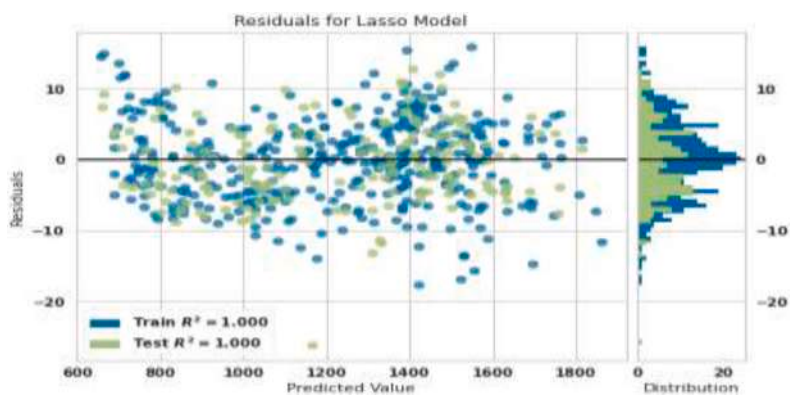


Fig. 10. Residual plot for industrial with Lasso.

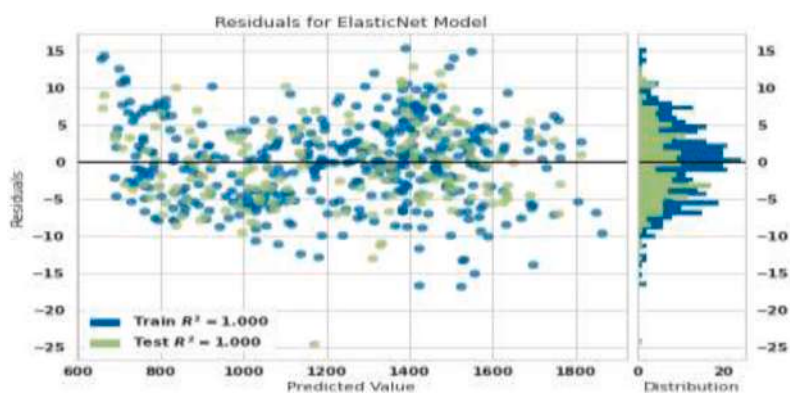


Fig. 11. Residual plot for industrial with elastic Net.

0.988 has been obtained.

As can be seen, in the random forest algorithm, $R^2 = 0.988$ was obtained, and the range of residuals was between -150 and 90 . Comparing the results of this algorithm with the Ridge algorithm, it can be seen that the R^2 evaluation criterion in the Ridge algorithm is only 0.012 higher than the R^2 evaluation criterion in the random forest algorithm. Still, the range of the majority of the residuals of the random forest algorithm is about 25 times that of the majority of the residuals of the Ridge algorithm. This claim can be seen by comparing Figs. 4 and 6 with each other and Figs. (9-c) and (9-e). This claim is proven carefully in the results of random forest and Ridge algorithms in Table 1.

Fig. 10 shows the residual plot for industrial with Lasso. According to Fig. 10, the residuals range between -30 and 20 . With a bit of precision, it can be seen in the residuals of the training and test data that the residuals have the highest number in the range of -10 and 10 . Fig. (17-a)

shows the lasso algorithm's industrial error for one year. As it is known, the error range of -6 to about three has been obtained, and this range of error values is in the range of residuals in Fig. 10. This analysis showed that the lasso algorithm worked correctly.

Fig. (17-a) shows errors for industrial with Lasso. January had the highest forecast error, and March had a near-zero error.

Fig. 11 shows the residual plot for industrial with elastic net. According to Fig. 11, the residuals range between -28 and 18 . With a bit of precision, it can be seen in the residuals of the training and test data that the residuals have the highest number in the range of -10 and 10 . Fig. (17-b) shows the industrial error for one year with the elastic net algorithm. As it is known, the error range of -6 to 3 has been obtained, and this range of error values is in the range of residuals in Fig. 11. This analysis showed that the elastic net algorithm worked correctly. Elastic Net and lasso algorithms had very similar performance.

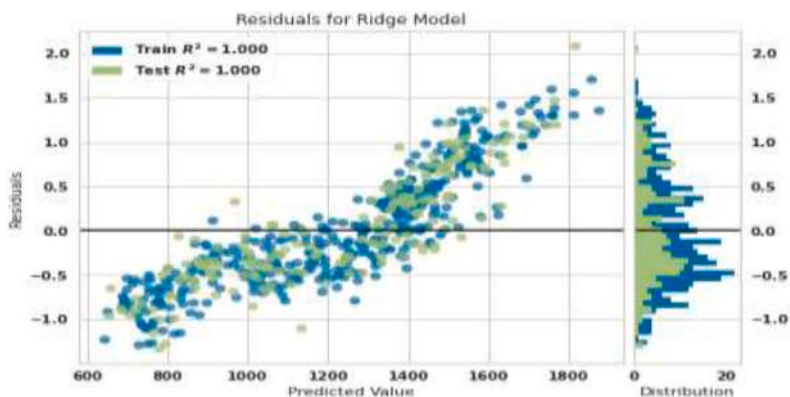


Fig. 12. Residual plot for industrial with Ridge.

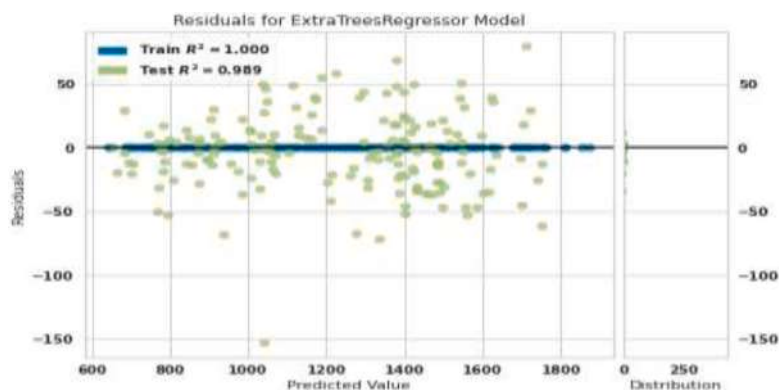


Fig. 13. Residual plot for industrial with extra tree.

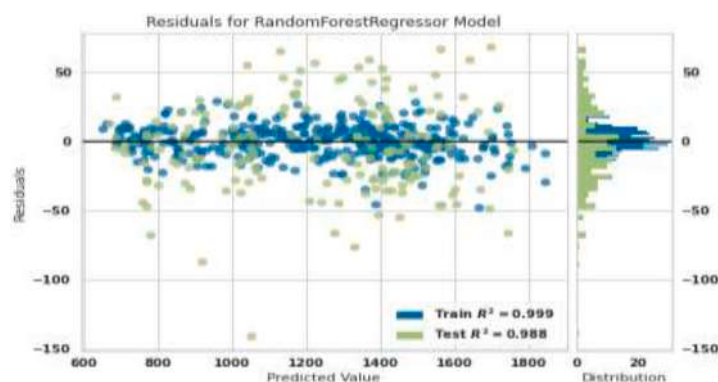


Fig. 14. Residual plot for industrial with random forest.

Fig. (17-b) shows the error for industrial with Elastic Net. January had the highest forecast error, and March had a near-zero error. This algorithm's behavior was similar to the Lasso algorithm. This issue was realized by comparing this Figure with Fig. (17-a).

Fig. 12 shows the residual plot for industrial with Ridge. According to Fig. 12, the residuals are between -2 and 2.5 . With a bit of precision, it can be seen in the residuals of the training and test data that the residuals have the highest number in the range of -1 and 1.5 . Fig. (17-c) shows the ridge algorithm's industrial error for one year. As it is known, the error range of -0.5 to 0.6 has been obtained, and this range of error values is in the range of residuals in Fig. 12. This analysis showed that the ridge algorithm worked correctly.

Fig. (17-c) shows the error for industrial with Ridge. The lowest error was 0 , and the highest error was about 0.6 .

Fig. 13 shows the residual plot for industrial with the extra tree. According to Fig. 13, the residuals range between -170 and 90 . With a

bit of precision, it can be seen in the residuals of the training and test data that the residuals have the highest number in the range of -70 and 70 . Fig. (17-d) shows industrial error for one year with an extra tree algorithm. As it is known, the error range of -70 to 50 has been obtained, and this range of error values is in the range of residuals in Fig. 13. This analysis showed that the extra tree algorithm worked correctly.

Fig. (17-d) shows errors for industrial with et. September had the highest forecast error.

Fig. 14 shows the residual plot for industrial with random forest. According to Fig. 14, the residuals range between -150 and 75 . With a bit of precision, it can be seen in the residuals of the training and test data that the residuals have the highest number in the range of -150 and 75 . Fig. (17-e) shows a random forest algorithm's industrial error for one year. As it is known, the error range of -80 to 60 has been obtained, and this range of error values is in the range of residuals in Fig. 14. This

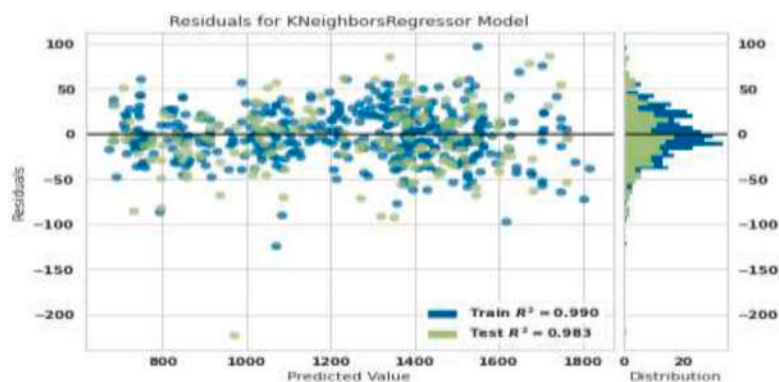


Fig. 15. Residual plot for industrial with k neighbors.

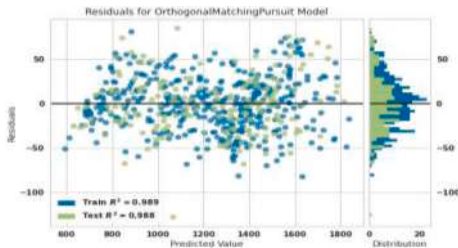


Fig. 16. Residual plot for industrial with OMP.

analysis showed that the random forest algorithm worked correctly.

Fig. (17-e) shows the error for industrial with rf. September had the highest forecast error, and June had a near-zero error.

Fig. 15 shows the residual plot for industrial with k neighbors. According to Fig. 15, the residuals range between –225 and 110. With a bit of precision, it can be seen in the residuals of the training and test data that the residuals have the highest number in the range of –150 and 100. Fig. (17-f) shows industrial error for one year with the k neighbors algorithm. As it is known, the error range of –150 to 100 has been obtained, and this range of error values is in the range of residuals in Fig. 15. This analysis showed that the k neighbors algorithm worked correctly.

Fig. (17-f) shows the error for industrial with knn. We made mistakes in almost all months of the year, and the maximum mistakes were

around 140. June had the highest forecast error, and August had a near-zero error.

Fig. 16 shows the residual plot for industrial with OMP. According to Fig. 16, the residuals range between –125 and 100. With a bit of precision, it can be seen in the residuals of the training and test data that the residuals have the highest number in the range of –100 and 100. Fig. (17-g) shows the industrial error for one year with the OMP algorithm. As it is known, the error range of 30 to –100 has been obtained, and this range of error values is in the range of residuals in Fig. 16. This analysis showed that the OMP algorithm worked correctly.

Fig. (17-g) shows errors for industrial with OMP. October had the highest forecast error, and April and February had near-zero errors.

Table 2 shows MAPE, RMSLE, RMSE, MSE, and MAE evaluation

Table 2 MAPE, RMSLE, RMSE, MSE, and MAE evaluation criteria for industrial.						
Sector	MAE	MSE	RMSE	RMSLE	MAPE	Time (Sec)
lasso	4.287	30.5	5.47	0.0020	0.0016	2.7
elastic Net	4.21	29.71	5.39	0.002	0.0016	2.8
ridge	1.286	3.5	1.51	0.0006	0.0005	1.2
extra tree	35.99	2140.4	45.83	0.0173	0.0137	8.3
random forest	36.07	2083.9	45.21	0.017	0.0137	6
k neighbors	63.85	7011	82.55	0.0778	0.0242	3.2
Orthogonal Matching Pursuit	33.87	1875.6	43	0.0164	0.0128	2

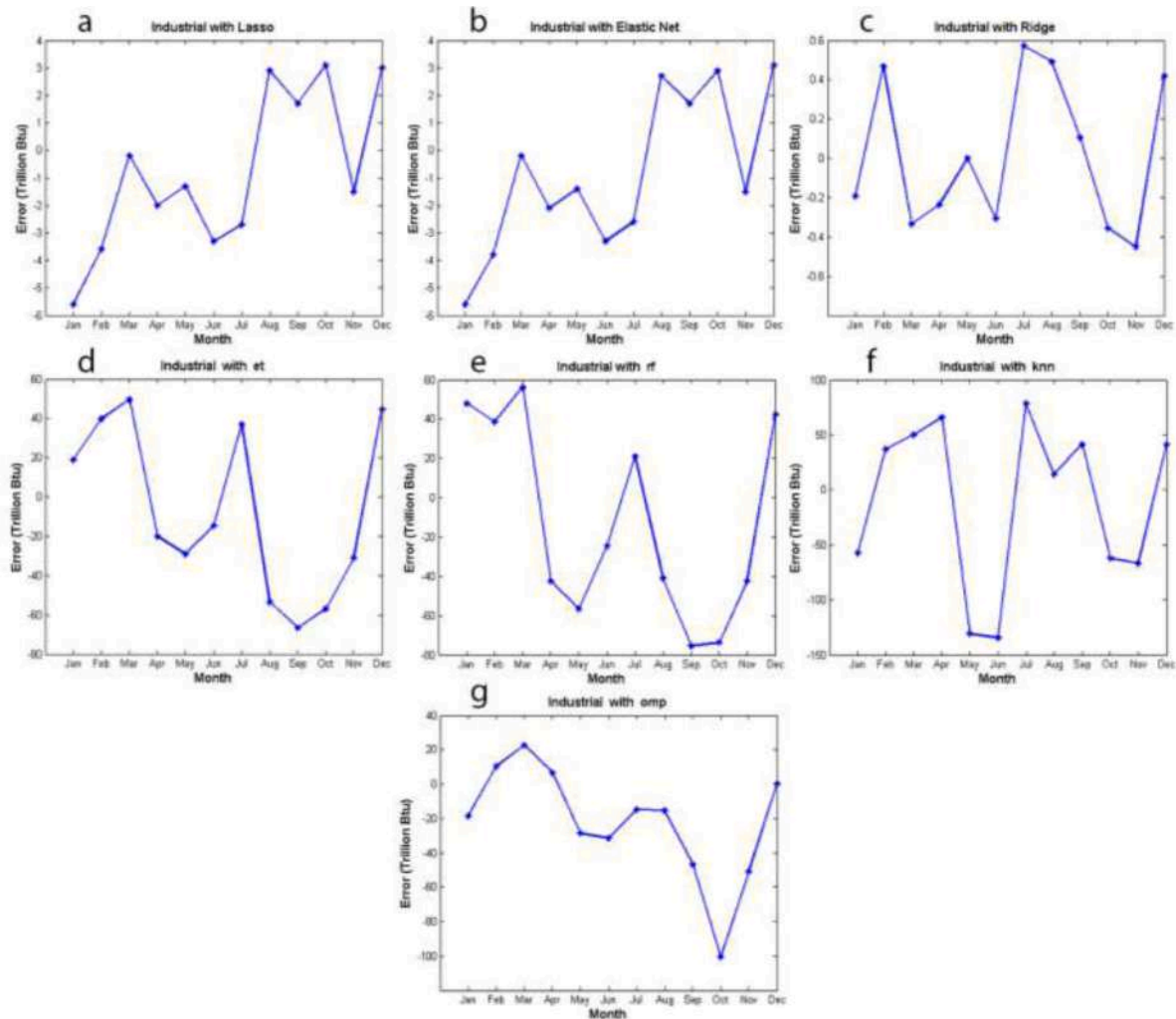


Fig. 17. error for industrial with (17-a) lasso, (17-b) elastic net, (17-c) ridge, (17-d) extra tree, (17-e) random forest, (17-f) k neighbors, (17-g) OMP for one year.

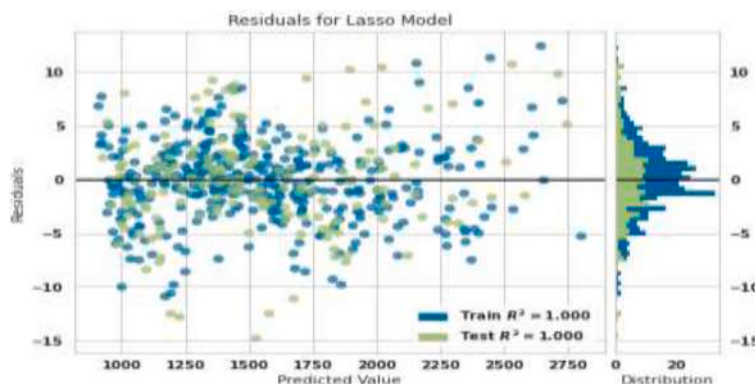


Fig. 18. Residual plot for residential with Lasso.

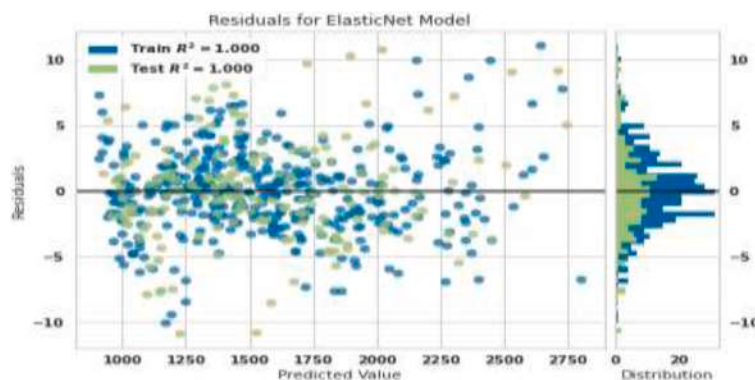


Fig. 19. Residual plot for residential with elastic Net.

criteria for the Lasso Regression, elastic Net, Ridge, extra tree, random forest, k neighbors, and Orthogonal Matching Pursuit algorithms presented in the industrial section. The ridge algorithm was the most powerful for predicting power consumption in the industrial area.

In (Ungureanu et al., 2019 Sep 3), the amount of energy consumption in the Industrial sector was predicted with Regression Simple NN, Classical regression, LSTM, and Random Forest algorithms. The best evaluation criterion of MAPE was 17.10 %. Comparing the results of the article (Ungureanu et al., 2019 Sep 3) with our results, the superiority of the results obtained by all the used algorithms is clear.

Fig. 18 shows the residual plot for residential with Lasso. According to Fig. 18, the residuals range between -16 and 15 . With a bit of precision, it can be seen in the residuals of the training and test data that the residuals have the highest number in the range of -10 and 10 . Fig. (25-a) shows the lasso algorithm's one-year residential error. As it is known, the error range of 3.5 to -1 has been obtained, and this range of error

values is in the range of residuals in Fig. 18. This analysis showed that the lasso algorithm worked correctly.

Fig. (25-a) shows that the maximum error value for residential with Lasso was approximately three, and the minimum error value was 0. February, April, and June had the highest prediction error, and August and September had nearly zero errors.

Fig. 19 shows the residual plot for residential with elastic net. According to Fig. 19, the residuals range between -12 and 12 . With a bit of precision, it can be seen in the residuals of the training and test data that the residuals have the highest number in the range of -5 and 5 . Fig. (25-b) shows the one-year residential error with the elastic net algorithm. As it is known, the error range of 3.5 to -0.7 has been obtained, and this range of error values is in the range of residuals in Fig. 19. This analysis showed that the elastic net algorithm worked correctly.

Fig. (25-b) shows the error for residential with Elastic Net. February, April, and June had the highest prediction error, and August and

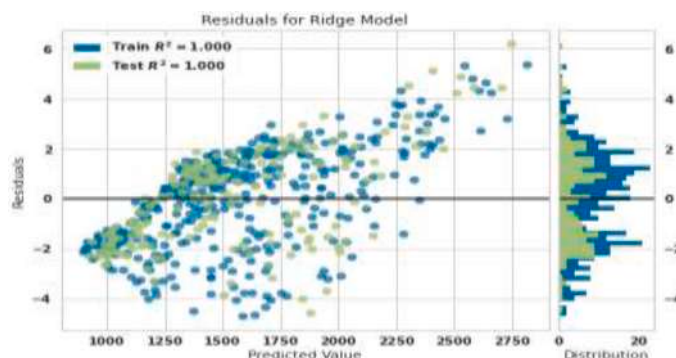


Fig. 20. Residual plot for residential with Ridge.

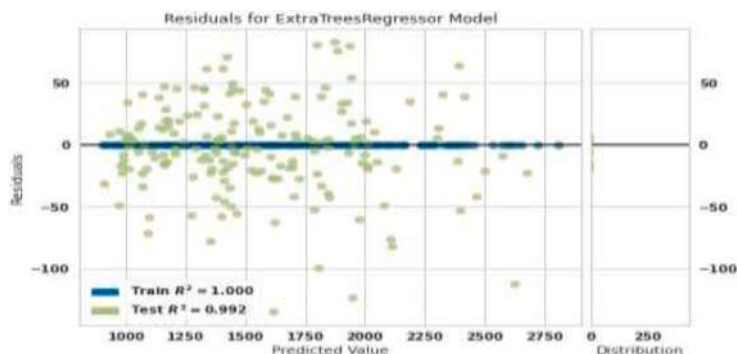


Fig. 21. Residual plot for residential with extra tree.

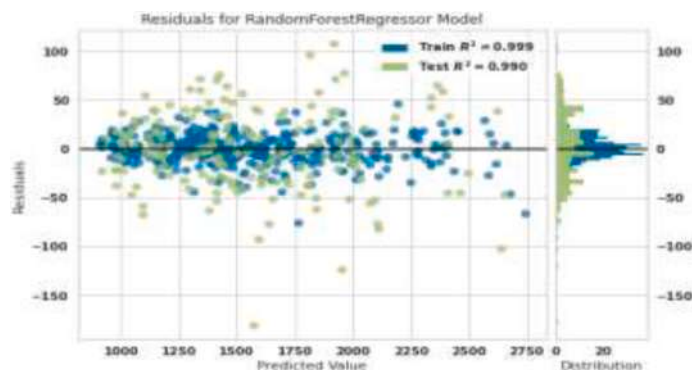


Fig. 22. Residual plot for residential with random forest.

September had nearly zero errors. This algorithm's behavior was similar to the Lasso algorithm. This issue was realized by comparing this Figure with Fig. (25-a).

Fig. 20 shows the residual plot for residential with Ridge. According to Fig. 20, the residuals range between -5 and 6.5 . With a bit of precision, it can be seen in the residuals of the training and test data that the residuals have the highest number in the range of -4 and 4 . Fig. (25-c) shows the ridge algorithm's one-year residential error. As it is known, the error range of 2 to -0.5 has been obtained, and this range of error values is in the range of residuals in Fig. 20. This analysis showed that the ridge algorithm worked correctly.

Fig. (25-c) shows the error for residential with Ridge. February, April, and December had the highest prediction error, and May, June, and August had nearly zero error.

Fig. 21 shows the residual plot for residential with the extra tree. According to Fig. 21, the residuals range between -150 and 100 . With a bit of precision, it can be seen in the residuals of the training and test

data that the residuals have the highest number in the range of -50 and 50 . Fig. (25-d) shows the one-year residential error with the extra tree algorithm. As it is known, the error range of about -45 to about 48 has been obtained, and this range of error values is in the residual range of Fig. 21. This analysis showed that the extra tree algorithm worked correctly.

Fig. (25-d) shows the error for residential with et. February, January, and June had the highest prediction error, and September had close to zero error.

Fig. 22 shows the residual plot for residential with random forests. According to Fig. 22, the residuals range between -200 and 120 . With a bit of precision, it can be seen in the residuals of the training and test data that the residuals have the highest number in the range of -100 and 50 . Fig. (25-e) shows a random forest algorithm's one-year residential error. As it is known, the error range of about 40 to -60 has been obtained, and this range of error values is in the range of residuals in Fig. 22. This analysis showed that the random forest algorithm worked

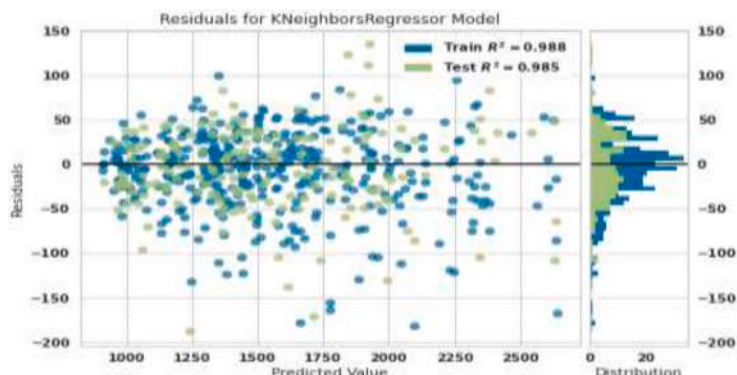


Fig. 23. Residual plot for residential with k neighbors.

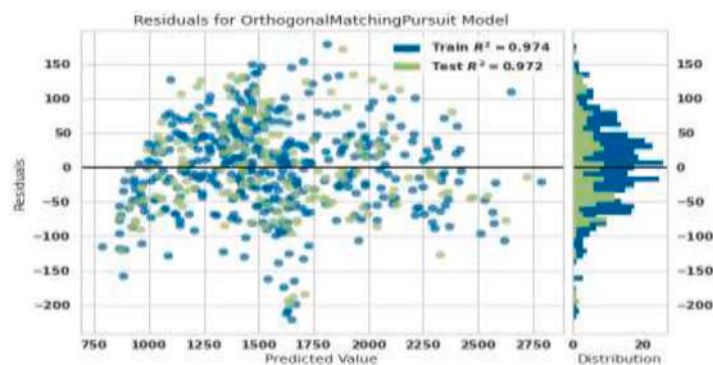


Fig. 24. Residual plot for residential with omp.

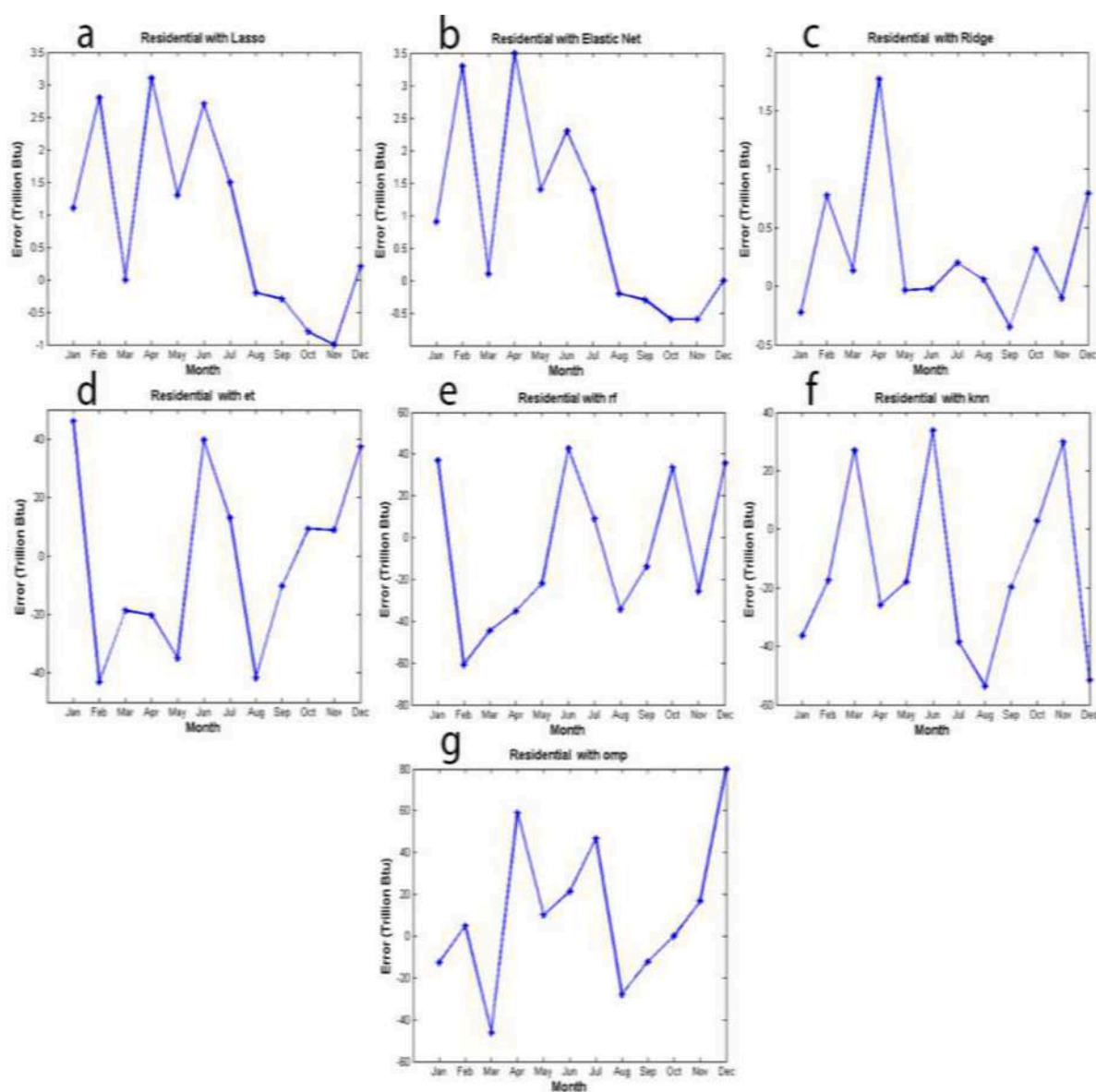


Fig. 25. error for residential with (25-a) lasso, (25-b) elastic net, (25-c) ridge, (25-d) extra tree, (25-e) random forest, (25-f) k neighbors, (25-g) omp for one year.

correctly.

Fig. (25-e) shows the error for residential with rf. February and June had the highest prediction error, and September had nearly zero errors.

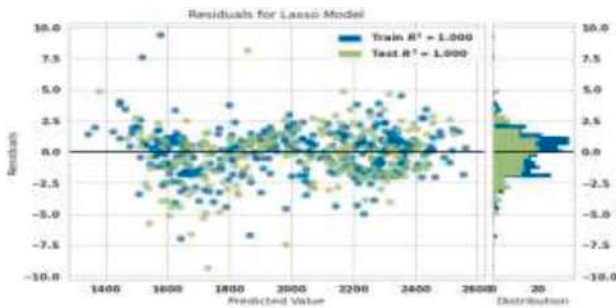
Fig. 23 shows the residual plot for residential with k neighbors.

According to Fig. 23, the residuals range between -200 and 150 . With a bit of precision, it can be seen in the residuals of the training and test data that the residuals have the highest number in the range of -50 and 50 . Fig. (25-f) shows the one-year residential error with the k neighbors

Table 3

MAPE, RMSLE, RMSE, MSE, and MAE evaluation criteria for residential.

Sector	MAE	MSE	RMSE	RMSLE	MAPE	Time (Sec)
lasso	1.65	4.626	2.12	0.0014	0.0011	2.8
elastic Net	1.49	3.92	1.95	0.0013	0.0010	2.4
ridge	0.59	0.597	0.735	0.0005	0.0004	1.3
extra tree	26.56	1247.2	34.7	0.0219	0.0172	6.66
random forest	31.79	1680.1	40.7	0.0259	0.0207	4
k neighbors	45.74	3828.2	61.21	0.0374	0.0289	2.9
Orthogonal Matching Pursuit	45.25	3618.5	58.61	0.0396	0.0303	0.58

**Fig. 26.** Residual plot for transportation with Lasso.

algorithm. As it is known, the error range of 40 to –50 has been obtained, and this range of error values is in the range of residuals in Fig. 23. This analysis showed that the k neighbors algorithm worked correctly. Fig. (25-f) shows the error for residential with knn.

Fig. 24 shows the residual plot for residential with OMP. According to Fig. 24, the residuals range between –250 and 200. With a bit of precision, it can be seen in the residuals of the training and test data that the residuals have the highest number in the range of –100 and 150. Fig. (25-g) shows the one-year residential error with the OMP algorithm. As it is known, the error range of –50 to 80 has been obtained, and this range of error values is in the range of residuals in Fig. 24. This analysis showed that the OMP algorithm worked correctly.

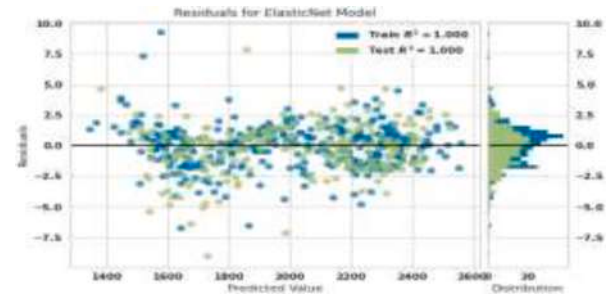
Fig. (25-g) shows the error for residential with OMP. We had errors in all months of the year. The maximum error was about 80, and the minimum was about 5.

Table 3 shows MAPE, RMSLE, RMSE, MSE, and MAE evaluation criteria for the Lasso Regression, elastic Net, Ridge, extra tree, random forest, k neighbors, and Orthogonal Matching Pursuit algorithms presented in the residential section. The ridge algorithm was the most powerful for predicting power consumption in residential areas.

The provided approach was used on a benchmark data set of power use for a single residential customer. For the same dataset, the findings from the Convolutional neural networks were compared to those from LSTM S2S, factoring restricted Boltzmann Machines, Artificial neural networks, and SVM. According to experimental results, the CN beat SVR while delivering outcomes that were on par with those of Artificial neural networks and deep learning techniques. Further research is necessary to compare the abilities of various deep learning architectures in load forecasting.

The RMSE evaluation criterion of the Convolutional neural networks method was 0.677, and the RMSE evaluation criterion for the LSTM-S2S method was 0.625. The RMSE evaluation criteria for FCRBM, SVM, and ANN methods were 0.663, 0.814, and 0.691, respectively. In this article (Amarasinghe et al., 2017 Jun 19), a residential complex was tested, and the reported results were obtained. In our article, the energy consumption of the United States in the residential sector was accepted with much more favorable results than (Amarasinghe et al., 2017 Jun 19).

Fig. 26 shows the residual plot for transportation with Lasso. According to Fig. 26, the residuals range between –10 and 10. Fig. (33-a)

**Fig. 27.** Residual plot for transportation with elastic Net.

shows the lasso algorithm's transportation error for one year. As it is known, the error range of 2 and –10 has been obtained, and this range of error values is in the residual range of Fig. 26. This analysis showed that the lasso algorithm worked correctly.

Fig. (33-a) shows Error for Transportation with Lasso. March had the highest prediction error, and August had zero error.

Fig. 27 shows the residual plot for transportation with elastic net. According to Fig. 27, the residuals range between –10 and 10. Fig. (33-b) shows the transportation error for one year with the elastic net algorithm. As it is known, the error range of 2 and –10 has been obtained, and this range of error values is in the residual range of Fig. 27. This analysis showed that the elastic net algorithm worked correctly. The elastic net and lasso algorithms had the same results.

Fig. (33-b) shows the Error for Transportation with Elastic Net. March had the highest prediction error, and August had a near-zero error. The Elastic Net algorithm behaved like Lasso. This issue was realized by considering Fig. (33-a) and comparing it with Fig. (33-b).

Fig. 28 shows the residual plot for transportation with Ridge. According to the Fig., 28 residuals are between –25 and 15. Fig. (33-c) shows the ridge algorithm's transportation error for one year. As it is known, the error range of 8 and –8 has been obtained, and this range of error values is in the range of residuals in Fig. 28. This analysis showed that the ridge algorithm worked correctly.

Fig. 29 shows the residual plot for transportation with an extra tree. According to Fig. 29, the residuals range between –7 and 8. Fig. (33-d) shows the transportation error for one year with the extra tree algorithm. As it is known, the error range of 3 and –5 has been obtained, and this range of error values is in the residual range of Fig. 29. This analysis showed that the extra tree algorithm worked correctly.

Fig. (33-d) shows the Error for Transportation with et. March had the highest prediction error, and August had a near-zero error.

Fig. 30 shows the residual plot for transportation with random forest. According to Fig. 30, the residuals range between –45 and 70. Fig. (33-e) shows the transportation error for one year with a random forest algorithm. As it is known, the error range of 5 and –35 has been obtained, and this range of error values is in the range of residuals in Fig. 30. This analysis showed that the random forest algorithm worked correctly.

Fig. 31 shows the residual plot for transportation with k neighbors. According to Fig. 31, the residuals range between –300 and 150. Fig. (33-f) shows the transportation error for one year with the k neighbors algorithm. As it is known, the error range of 100 and –250 has been obtained, and this range of error values is in the range of residuals in Fig. 31. This analysis showed that the k neighbors algorithm worked correctly.

Fig. (33-f) shows Error for Transportation with knn. We made significant mistakes throughout the year. This algorithm had worse results than Lasso, Extra Tree, Random Forest, and Elastic Net algorithms.

Fig. 32 shows the residual plot for transportation with OMP. According to Fig. 32, the residuals range between –1.6 and 1.8. Fig. (33-g) shows the transportation error for one year with the OMP algorithm. As it is known, the error range of –0.8 and 0.8 has been obtained, and this range of error values is in the range of residuals in Fig. 32. This analysis

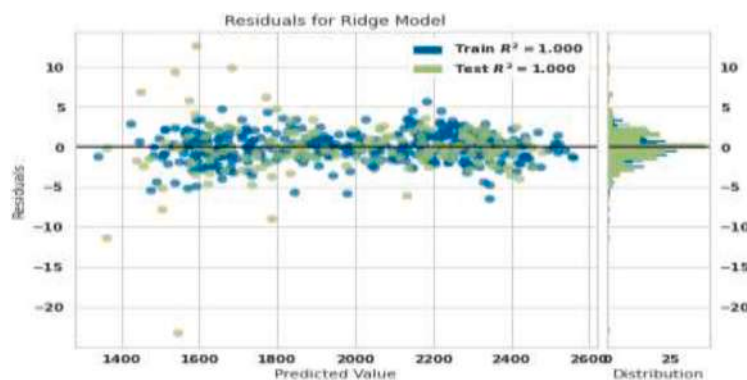


Fig. 28. Residual plot for transportation with Ridge.

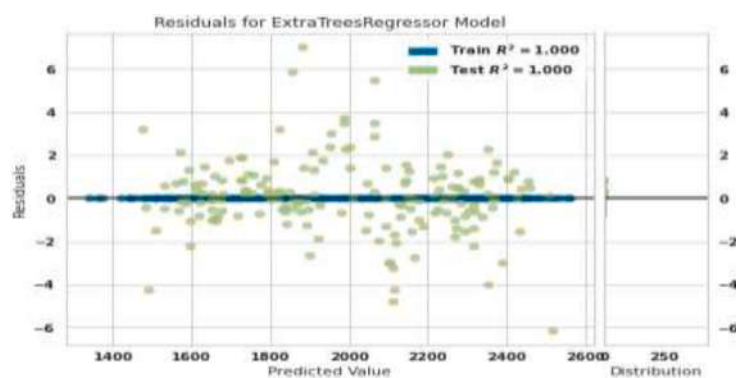


Fig. 29. Residual plot for transportation with extra tree.

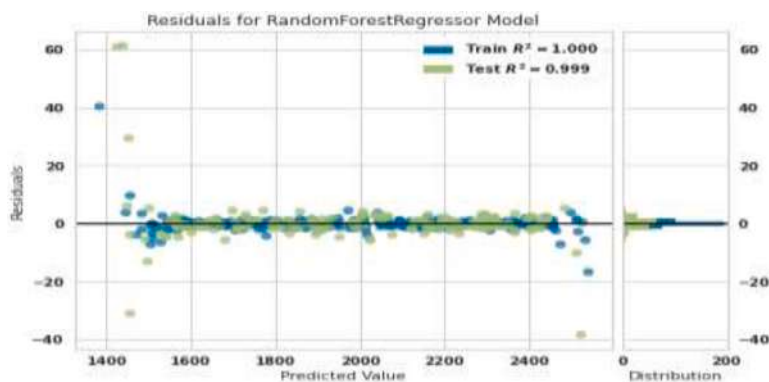


Fig. 30. Residual plot for transportation with random forest.

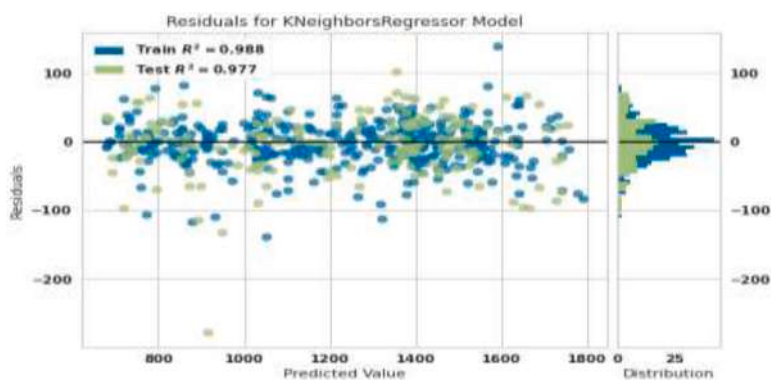


Fig. 31. Residual plot for transportation with k neighbors.

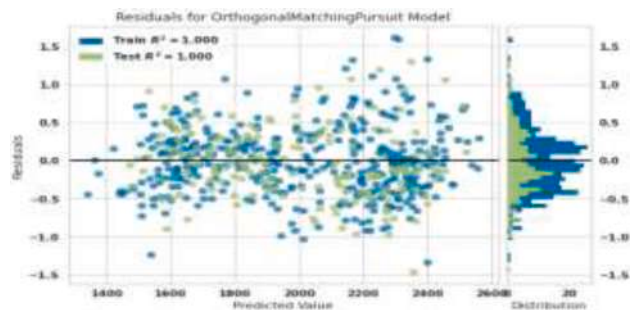


Fig. 32. Residual plot for transportation with omp.

showed that the OMP algorithm worked correctly. The OMP algorithm had the best results in the transportation section.

Table 4 shows MAPE, RMSLE, RMSE, MSE, and MAE evaluation criteria for the Lasso Regression, elastic Net, Ridge, extra tree, random forest, k neighbors, and Orthogonal Matching Pursuit algorithms presented in the transportation section. The Orthogonal Matching Pursuit algorithm was the most powerful for predicting power consumption in the transportation section.

Nowadays (Ağbulut, 2022 Jan 1), more than a 99percent of transportation is driven by fossil-based fuels, and air pollution-related illnesses kill nearly 6.5 million people globally each year. Several ml algorithms DL, SVM, and ANN, are utilized in this system to anticipate Turkey’s transportation-based CO2 emissions and energy consumption.

To provide a more accurate comparison, the outcomes of these algorithms are described in conjunction with six commonly used statistical indicators (R2, RMSE, MAPE, MBE, rRMSE, and MABE). R2 ratings for all ML techniques range between 0.8639 and 0.9235.

By employing the empirical Kernel-Based Regularized Least Squares (KRLS) approach and quantile regression to address non-normality issues in variables, the research reveals (Özkan et al., 2023 Aug 1) compelling insights. It highlights that improving nonrenewable energy efficiency leads to a more environmentally sustainable outcome when compared to intensifying renewable energy utilization. However, the environmental performance derived from intensive renewable energy use exceeds that of environmental-related technologies. Notably, the EU showcases superior environmental sustainability achievements across these measured metrics compared to the USA. The high R2 values of

Table 4
MAPE, RMSLE, RMSE, MSE, and MAE evaluation criteria for transportation.

Sector	MAE	MSE	RMSE	RMSLE	MAPE	Time (Sec)
lasso	1.38	3.499	1.836	0.001	0.0007	2.7
elastic Net	1.35	3.35	1.799	0.0010	0.0007	3
ridge	1.249	3.941	1.544	0.0008	0.0007	1.6
extra tree	1.43	15.89	2.99	0.0018	0.0008	4.2
random forest	3.09	54.14	5.85	0.0036	0.0017	3.8
k neighbors	42.21	2988.8	54.34	0.0284	0.0218	2.34
Orthogonal Matching Pursuit	0.37	0.23	0.479	0.0002	0.0002	1.34

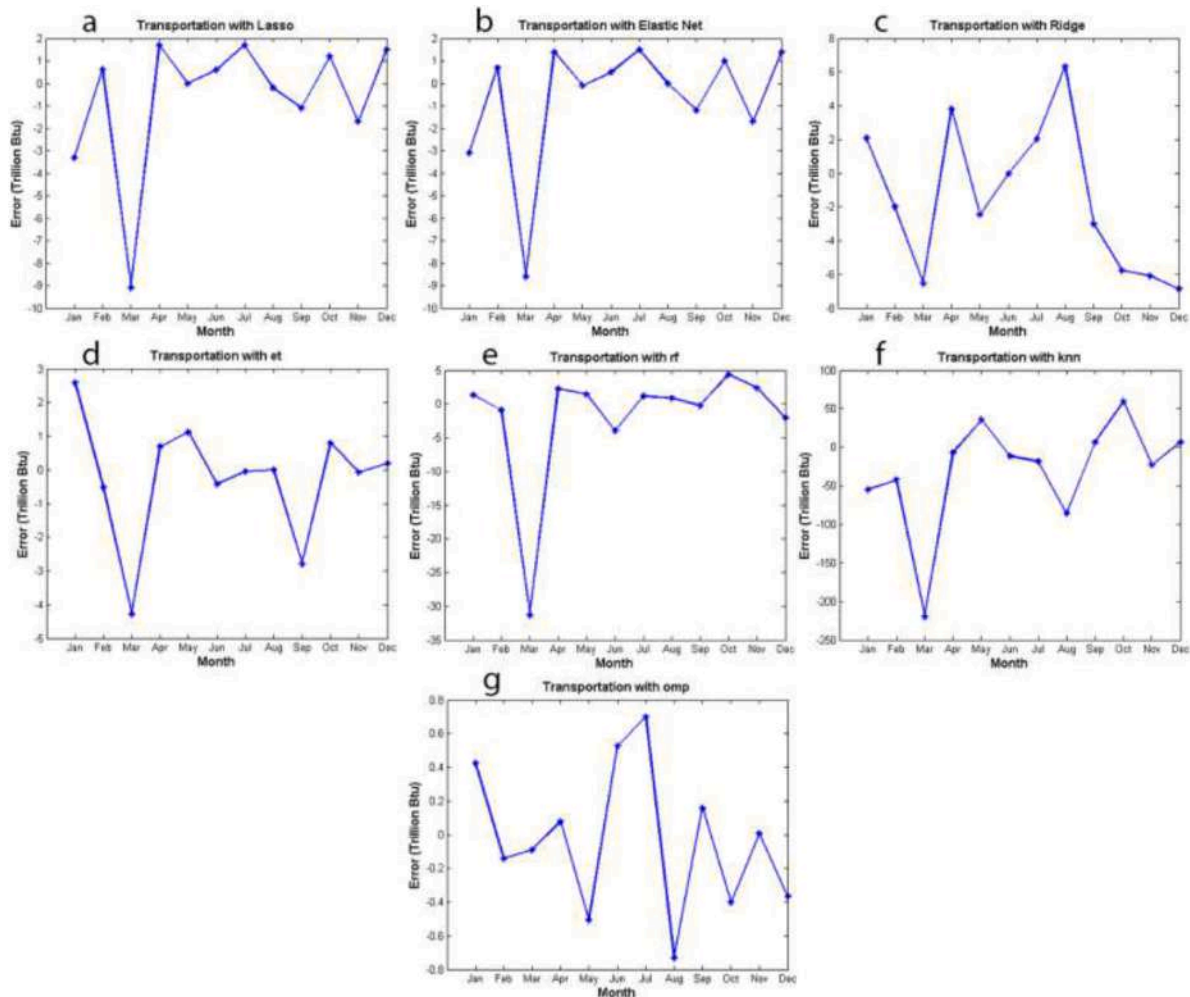


Fig. 33. error for transportation with (33-a) lasso, (33-b) elastic net, (33-c) ridge, (33-d) extra tree, (33-e) random forest, (33-f) k neighbors, (33-g) omp for one year.



Fig. 34. Time for algorithms.

0.994 for the USA and 0.977 for the EU signify explanatory solid power, indicating that the considered regressors can account for a significant portion of the variations in GHG emissions (99 % for the USA and 97 % for the EU). These regressors include nonrenewable energy efficiency (NREE), renewable energy intensity (REI), environmental technologies (ET), natural resources (NR), and urbanization (UR).

Examining the average marginal effects sheds light on how a 1 % increase in each regressor impacts GHG emissions. Higher levels of NREE, REI, and ET in the USA result in reduced GHG emissions, while increases in NR and UR lead to higher emissions. Conversely, similar trends are observed in the EU, albeit with varying magnitudes for each regressor.

In Figs. 26 to 32, R2 evaluation criteria in the transportation sector of the United States were calculated using Lasso Regression, elastic Net, Ridge, extra tree, random forest, k neighbors, and Orthogonal Matching Pursuit algorithm. The R2 evaluation criterion for the training data was assigned a value of 1, and for the test data, the R2 evaluation criterion was obtained between 0.977 and 1.

6. Conclusion

The energy consumption of the United States was predicted by machine learning in four sectors: industry, housing, transportation, and trade. The Ridge algorithm was the most powerful in predicting energy consumption. The K Neighbors algorithm had the worst results of the MAPE, RMSLE, RMSE, MSE, and MAE evaluation criteria and performed poorly in predicting energy consumption.

The MSE evaluation criterion for the Ridge algorithm in the commercial sector was 2.61. In the residential sector, it was 0.597; in the industrial sector, it was 3.5. These values were much lower than those of other algorithms. Therefore, the Ridge algorithm was considered the most accurate in predicting energy consumption. Also, the MSE evaluation criterion for the Orthogonal Matching Pursuit algorithm in the transportation section obtained a value of 0.23, the lowest value compared to other algorithms.

Fig. 34 shows the execution time of the proposed algorithms lasso, elastic Net, Ridge, extra tree, random forest, k neighbors, and Orthogonal Matching Pursuit. By comparing the results of Tables 1 to 4, it was found that the ridge algorithm had the best results in predicting the energy consumption of residential, industrial, and commercial sectors, and the Orthogonal Matching Pursuit algorithm had the best results in predicting energy consumption in the transportation sector. By comparing the execution times of the algorithms in Fig. 34, it was found that the ridge algorithm, in addition to its high accuracy in predicting energy consumption, also takes a little time to execute, and it should be noted that the Orthogonal Matching Pursuit algorithm takes a little time to execute after the ridge algorithm.

CRediT authorship contribution statement

Seyed Matin Malakouti: Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Project administration, Resources, Software, Validation, Visualization, Writing – original draft, Writing – review & editing. **Mohammad Bagher Menhaj:** Supervision, Validation. **Amir Abolfazl Suratgar:** Validation, Visualization.

Declaration of competing interest

None.

References

- D'Agostino, D., Cuniberti, B., Bertoldi, P., 2017. Energy consumption and efficiency technology measures in European non-residential buildings. *Energy Build.* 153, 72–86. <https://doi.org/10.1016/j.enbuild.2017.07.062>.
- J. Naji, S., Keivani, A., Shamshirband, S., Alengaram, U.J., Jumaat, M.Z., Mansor, Z., Lee, M., 2016. Estimating building energy consumption using extreme learning machine method. *Energy* 97, 506–516. <https://doi.org/10.1016/j.energy.2015.11.037>.
- Chou, J.-S., Tran, D.-S., 2018. Forecasting energy consumption time series using machine learning techniques based on usage patterns of residential householders. *Energy* 165, 709–726. <https://doi.org/10.1016/j.energy.2018.09.144>.
- Chou, J.S., Ngo, N.T., Chong, W.K., Gibson, G.E., 2016. 16: big data analytics and cloud computing for sustainable building energy efficiency. In: Pacheco-Torgal, F., Rasmussen, E., Granqvist, C.-G., Ivanov, V., Kaklauskas, A., Makonin, S. (Eds.), *Start-Up Creation*. Woodhead Publishing, pp. 397–412.
- Bokalders, V.M.B., 2010. *The Whole Building Handbook-How to Design Healthy, Efficient and Sustainable Buildings*. Earthscan, London.
- Pham, A.-D., Ngo, N.-T., Ha Truong, T.T., Huynh, N.-T., Truong, N.-S., 2020. Predicting energy consumption in multiple buildings using machine learning for improving energy efficiency and sustainability. *J. Clean. Prod.* 260, 121082. <https://doi.org/10.1016/j.jclepro.2020.121082>.
- Zhong, H., Wang, J., Jia, H., Mu, Y., Lv, S., 2019. Vector field-based support vector regression for building energy consumption prediction. *Appl. Energy* 242, 403–414. <https://doi.org/10.1016/j.apenergy.2019.03.078>.
- Amasyali, K., El-Gohary, N.M., 2018. A review of data-driven building energy consumption prediction studies. *Renew. Sustain. Energy Rev.* 81, 1192–1205.
- Wei, Y., Zhang, X., Shi, Y., Xia, L., Pan, S., Wu, J., Han, M., Zhao, X., 2018. A review of data-driven approaches for prediction and classification of building energy consumption. *Renew. Sustain. Energy Rev.* 82, 1027–1047. <https://doi.org/10.1016/j.rser.2017.09.108>.
- Deb, C., Zhang, F., Yang, J., Lee, S.E., Shah, K.W., 2017. A review on time series forecasting techniques for building energy consumption. *Renew. Sustain. Energy Rev.* 74, 902–924.
- Chaudhuri, T., Soh, Y.C., Li, H., Xie, L., 2019. A feedforward neural network based indoor-climate control framework for thermal comfort and energy saving in buildings. *Appl. Energy* 248, 44–53.
- Rahman, A., Srikumar, V., Smith, A.D., 2018. Predicting electricity consumption for commercial and residential buildings using deep recurrent neural networks. *Appl. Energy* 212, 372–385.
- Tian, C., Li, C., Zhang, G., Lv, Y., 2019. Data driven parallel prediction of building energy consumption using generative adversarial nets. *Energy Build.* 186, 230–243.
- Gao, W., Alsarraf, J., Moayedi, H., Shahsavari, A., Nguyen, H., 2019. Comprehensive preference learning and feature validity for designing energy-efficient residential buildings using machine learning paradigms. *Appl. Soft. Comput.* 84, 105748.

- Renani, E.T., Elias, M.F.M., Rahim, N.A., 2016. Using data-driven approach for wind power prediction: a comparative study. *Energy Convers. Manag.* 118, 193–203.
- Nasruddin, S., Satrio, P., Mahlia, T.M.I., Giannetti, N., Saito, K., 2019. Optimization of HVAC system energy consumption in a building using artificial neural network and multi-objective genetic algorithm. *Sustain. Energy Technol. Assess.* 35, 48–57.
- Ahmad, M.W., Mouraud, A., Rezgui, Y., Mourshed, M., 2018. deep highway networks and tree-based ensemble for predicting shortterm building energy consumption. *Energies (Basel)* 11 (12), 3408.
- Al-Rakhami, M., Gumaei, A., Alsanad, A., Alamri, A., Hassan, M.M., 2019. An ensemble learning approach for accurate energy load prediction in residential buildings. *IEEE Access*. 7, 48328–48338. <https://doi.org/10.1109/ACCESS.2019.2909470>.
- Ngo, N.-T., 2019. Early predicting cooling loads for energy-efficient design in office buildings by machine learning. *Energy Build.* 182, 264–273.
- Araya, D.B., Grolinger, K., ElYamany, H.F., Capretz, M.A.M., Bitsuamlak, G., 2017. An ensemble learning framework for anomaly detection in building energy consumption. *Energy Build.* 144, 191–206.
- Chou, J.-S., Pham, T.-P.-T., Nguyen, T.-K., Pham, A.-D., Ngo, N.-T., 2019. Shear strength prediction of reinforced concrete beams by baseline, ensemble, and hybrid machine learning models. *Soft. Comput.*
- Chou, J.-S., Ngo, N.-T., Chong, W.K., 2017. The use of artificial intelligence combiners for modeling steel pitting risk and corrosion rate. *Eng. Appl. Artif. Intell.* 65, 471–483.
- Ahmad, T., Chen, H., Huang, R., Yabin, G., Wang, J., Shair, J., Azeem Akram, H.M., Hassnain Mohsan, S.A., Kazim, M., 2018. Supervised based machine learning models for short, medium and longterm energy prediction in distinct building environment. *Energy* 158, 17–32.
- Herrera, G.P., Constantino, M., Tabak, B.M., Pistori, H., Su, J.-J., Naranpanawa, A., 2019. Long-term forecast of energy commodities price using machine learning. *Energy* 179, 214–221.
- Wang, Z., Wang, Y., Srinivasan, R.S., 2018. A novel ensemble learning approach to support building energy use prediction. *Energy Build.* 159, 109–122.
- Papadopoulos, S., Azar, E., Woon, W.-L., Kontokosta, C.E., 2018. Evaluation of tree-based ensemble learning algorithms for building energy performance estimation. *J. Build. Perform. Simul.* 11 (3), 322–332.
- Huang, Y., Yuan, Y., Chen, H., Wang, J., Guo, Y., Ahmad, T., 2019. A novel energy demand prediction strategy for residential buildings based on ensemble learning. *Energy Proc.* 158, 3411–3416.
- Adebayo, T.S., Kartal, M.T., Ağa, M., MA, Al-Faryan, 2023 Feb 1. Role of country risks and renewable energy consumption on environmental quality: evidence from MINT countries. *J. Environ. Manag.* 327, 116884.
- Akram, R., Ibrahim, R.L., Wang, Z., Adebayo, T.S., Irfan, M., 2023 Oct 15. Neutralizing the surging emissions amidst natural resource dependence, eco-innovation, and green energy in G7 countries: insights for global environmental sustainability. *J. Environ. Manag.* 344, 118560.
- Alola, A.A., Adebayo, T.S., 2023 Aug 20. Analysing the waste management, industrial and agriculture greenhouse gas emissions of biomass, fossil fuel, and metallic ores utilization in Iceland. *Sci. Total Environ.* 887, 164115.
- Adebayo, T.S., Ullah, S., 2023 Jul 1. Formulating sustainable development policies for China within the framework of socioeconomic conditions and government stability. *Environ. Pollut.* 328, 121673.
- Ullah, S., Adebayo, T.S., Irfan, M., Abbas, S., 2023 Sep 15. Environmental quality and energy transition prospects for G-7 economies: the prominence of environment-related ICT innovations, financial and human development. *J. Environ. Manag.* 342, 118120.
- Liu, X., Adebayo, T.S., Ramzan, M., Ullah, S., Abbas, S., VO, Olanrewaju, 2023 Sep 10. Do coal efficiency, climate policy uncertainty and green energy consumption promote environmental sustainability in the United States? An application of novel wavelet tools. *J. Clean. Prod.* 417, 137851.
- Hu, B., Alola, A.A., Tauni, M.Z., Adebayo, T.S., Abbas, S., 2023 Dec 1. Pathway to cleaner environment: how effective are renewable electricity and financial development approaches? *Struct. Chang. Econ. Dyn.* 67, 277–292.
- Özkan, O., Alola, A.A., Adebayo, T.S., 2023 Aug 1. Environmental benefits of nonrenewable energy efficiency and renewable energy intensity in the USA and EU: examining the role of clean technologies. *Sustain. Energy Technol. Assess.* 58, 103315.
- Ungureanu, S., Țopa, V., Czikier, A., 2019 Sep 3. Industrial load forecasting using machine learning in the context of smart grid. In: 2019 54th International Universities Power Engineering Conference (UPEC). IEEE, pp. 1–6.
- Amarasinghe, K., Marino, D.L., Manic, M., 2017 Jun 19. Deep neural networks for energy load forecasting. In: 2017 IEEE 26th international symposium on industrial electronics (ISIE). IEEE, pp. 1483–1488.
- Ağbulut, Ü., 2022 Jan 1. Forecasting of transportation-related energy demand and CO₂ emissions in Turkey with different machine learning algorithms. *Sustain. Prod. Consum.* 29, 141–157.
- Malakouti, S.M., Ghiasi, A.R., Ghavifekr, A.A., Emami, P., 2022 Dec. Predicting wind power generation using machine learning and CNN-LSTM approaches. *Wind Engineering* 46 (6), 1853–1869.
- Malakouti, S.M., 2023 Jun 1. Estimating the output power and wind speed with ML methods: a case study in Texas. *Case Stud. Chem. Environ. Eng.* 7, 100324.
- Malakouti, S.M., 2023 Jul 1. Heart disease classification based on ECG using machine learning models. *Biomed. Signal. Process. Control* 84, 104796.
- Malakouti, S.M., 2023 Dec 1. Improving the prediction of wind speed and power production of SCADA system with ensemble method and 10-fold cross-validation. *Case Stud. Chem. Environ. Eng.* 8, 100351.
- Malakouti, S.M., 2023 Jun 1. Utilizing time series data from 1961 to 2019 recorded around the world and machine learning to create a Global temperature change prediction Model. *Case Stud. Chem. Environ. Eng.* 7, 100312.
- Malakouti, S.M., Ghiasi, A.R., 2022 May 11. Evaluation of the application of computational model machine learning methods to simulate wind speed in predicting the production capacity of the Swiss basel wind farm. In: 2022 26th International Electrical Power Distribution Conference (EPDC). IEEE, pp. 31–36.
- Malakouti, S.M., Menhaj, M.B., Suratgar, A.A., 2023 Aug 1. The usage of 10-fold cross-validation and grid search to enhance ML methods performance in solar farm power generation prediction. *Clean. Eng. Technol.* 15, 100664.
- Malakouti, S.M., Ghiasi, A.R., Ghavifekr, A.A., 2022 Jan 1. AERO2022-flying danger reduction for quadcopters by using machine learning to estimate current, voltage, and flight area. *e-Prime-Adv. Electric. Eng., Electron. Energy* 2, 100084.
- Malakouti, S.M., 2023 Mar. Use machine learning algorithms to predict turbine power generation to replace renewable energy with fossil fuels. *Energy Explor. Exploit.* 41 (2), 836–857.
- Malakouti, S.M., 2023 Mar 16. Discriminate primary gammas (signal) from the images of hadronic showers by cosmic rays in the upper atmosphere (background) with machine learning. *Phys. Scr.* 98 (4), 045506.
- Malakouti, S.M., 2023 Sep 1. Babysitting hyperparameter optimization and 10-fold-cross-validation to enhance the performance of ML methods in predicting wind speed and energy generation. *Intell. Syst. Appl.* 19, 200248.
- Malakouti, S.M., 2023. Prediction of wind speed and power with LightGBM and grid search: case study based on Scada system in Turkey. *Int. J. Energy Prod. Manag.* 8 (1), 35–40. [https://doi.org/10.1016/j.ies.2023.8\(1\)](https://doi.org/10.1016/j.ies.2023.8(1)).
- Malakouti, S.M., Menhaj, M.B., Suratgar, A.A., 2023 Dec 1. Machine learning techniques for classifying dangerous asteroids. *MethodsX*. 11, 102337.
- Malakouti, S.M., 2023 May 1. Cancer risk assessment based on Family history and smoking habits. *Syst. Rev. Pharm.* 14 (6).
- Malakouti, S.M., Menhaj, M.B., Suratgar, A.A., 2024 Mar 1. ML: early breast cancer diagnosis. *Curr. Probl. Cancer: Case Rep.* 13, 100278.