

Using Machine Learning Algorithms in Home Energy Management Systems for Energy Cost Forecasting

Abderrauof Bouakkaz¹, Adel Lahsasna²

¹ Laboratoire de génie mécanique et matériaux, Université 20 aout 1955 Skikda. Bp26 route d'el Hadaik Skikda, Algeria

² Département d'informatique, Université 20 aout 1955 Skikda. Bp26 route d'el Hadaik Skikda, Algeria

*Corresponding author; Email: ab.bouakkaz@univ-skikda.dz

Article Info

Article history:

Received , 21/09/2024

Revised , 20/10/2024

Accepted , 25/11/2024

Keywords:

Machine learning
Energy management
Cost of energy
Energy consumption
Prediction

ABSTRACT

Facing global energy challenges, as energy demand continues to increase significantly, businesses and individuals are increasingly exploring ways to optimize energy use to reduce costs and environmental impacts. Artificial intelligence (AI) has become a key tool for accomplishing these goals, offering innovative ways to monitor, analyse and optimize energy consumption, thereby reducing expenses and promoting sustainability. In this study, energy cost prediction across different seasons is explored using various machine learning regression models, including Linear Regression (LR), Random Forest (RF), Support Vector Machine (SVM), and Artificial Neural Networks (ANN). The performance and accuracy of each model were evaluated, with all implementations conducted using Python version 3.11. The paper aims to predict the cost of energy consumption, ultimately minimizing the total energy cost. The assessment of performance and the accuracy of these models are evaluated using the MSE metric, showing how machine learning can help reduce energy consumption and costs.

I. Introduction

The energy sector is currently faced with a challenging situation as energy demand continues to rise, while supply undergoes a complex and uncertain transition [1]. Meanwhile, energy consumption costs pose a considerable challenge for households[2]. Consequently, the increase in the costs of energy consumption, besides the environmental concerns [3] and [4], and the need for efficient resource utilization, has spurred a paradigm shift toward employing advanced technologies like Artificial Intelligence (AI)[5] and [6]. In recent years, AI has emerged as an important tool for energy management, providing data-driven insights and predictive analytics to help streamline energy consumption and reduce costs [7], [8] and [9].

Many studies provide valuable insight into different approaches and methods used in energy management that use machine learning techniques. For instance, the authors of [10] have proposed an innovative approach (novel deep neural network) using machine learning, enabling them to accurately forecast hourly energy consumption in residential buildings. Artificial Neural Networks (ANN) and Genetic Algorithms are used to predict and plan energy consumption in smart buildings[11]. The proposed models are deployed in a real-world smart building. In addition, the models are trained and validated using real-world data collected from a PV installation along with smart building electrical appliances. In [12], authors propose a machine-learning approach using Support Vector Regression (SVR) to explore the correlation between various combinations of three smart information interventions: the current electricity consumption profile, the historical electricity consumption

profile, and a ranking of electricity savings compared to other participating households. Their findings reveal that the overall average electricity savings across the 14 households studied is 7.1%. An hour-ahead demand response (DR) based on supervised learning is proposed for behaviour-based home energy management systems[13]. Through learning the appliance usage habits of householders, the proposed strategy aims to control the energy storage system (ESS) and renewable energy system (RES) in order to reduce daily energy consumption costs.

These studies underscore the significance of energy forecasting, the challenges involved as well as the techniques and methodologies applied. In this paper, various machine-learning techniques have been used to predict the energy consumption costs for different seasons. With these different machine-learning algorithms, we aim to evaluate the accuracy of the proposed learning algorithms to accurately predict energy consumption cost, to optimize energy performance, thereby reducing the total cost of energy consumption.

The rest of the paper is organised as follow: the research methodology is presented in the second section; the section three the results and discussion; finally, the conclusion of the paper is given in the fourth section.

II. Research Method

A general overview of the proposed forecasting architecture is presented in Figure 1 which integrates machine learning models with an energy management system. The system includes weather forecasts, energy consumption forecasts, and energy prices, storing data in a logger (DB) for decision-making. It communicates with equipment (PV, grid, battery, loads) to measure parameters and provide instructions for efficient energy management. More details of the studied system and the methodology of forecasting are provided in the following section.

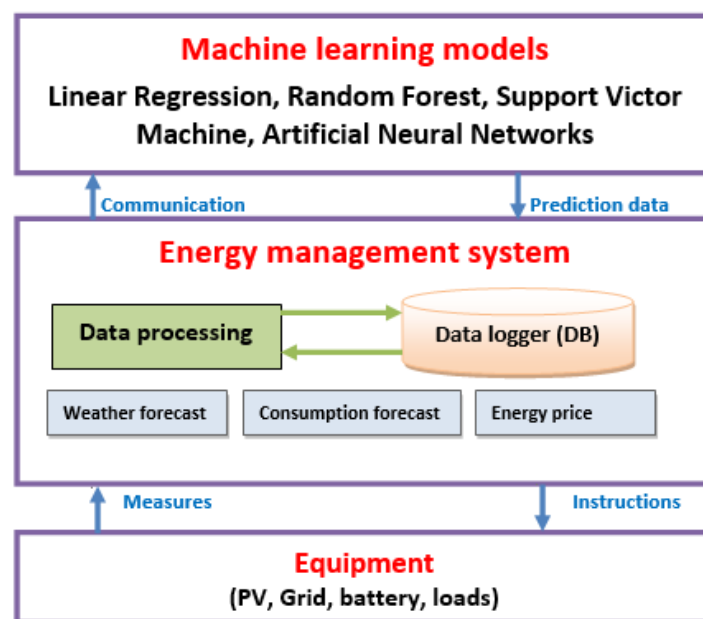


Figure 1. The general architecture of energy cost forecasting.

II.1. Data description

This section involves the selection of a representative dataset that is used to train the machine learning models. This training dataset encompasses four different Quarter of data, each quarter present a season of the year (winter, spring, summer and autumn) with one-hour time step and covers different variables that affect energy consumption costs. In this study, three main parameters as features were considered (global horizontal irradiation, temperature and energy demand), where their data were obtained from National Renewable Energy Laboratory (NREL), which are shown in Figures 2, 3 and 4, respectively.

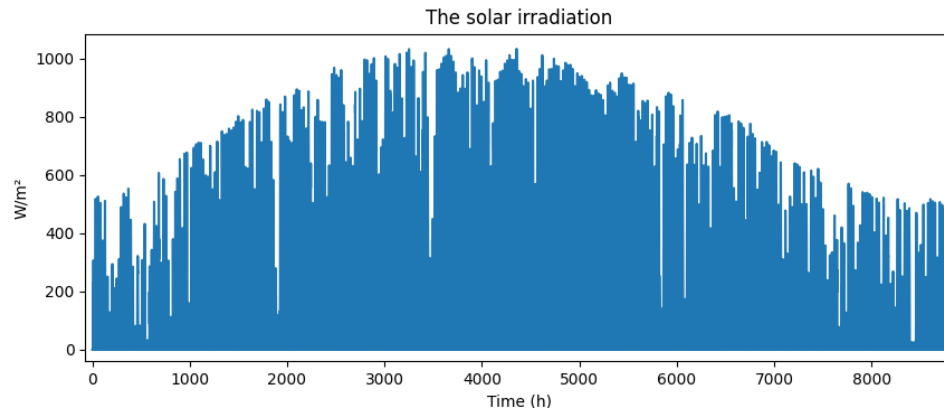


Figure 2. The solar irradiance of one year.

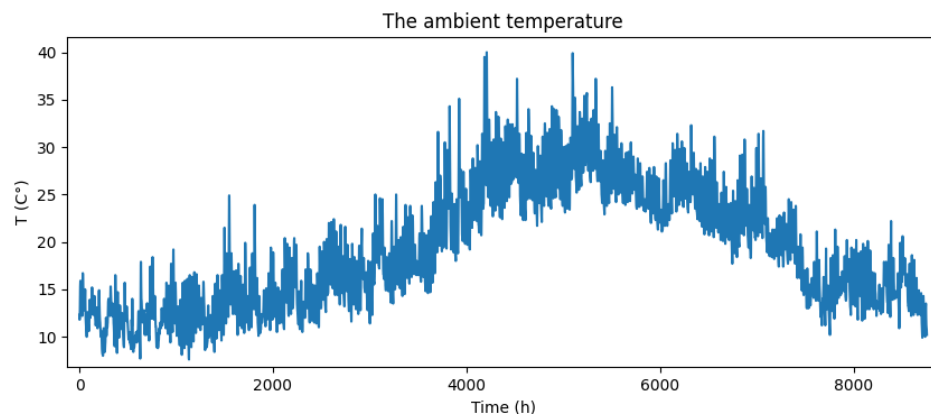


Figure 3. The ambient temperature of one year.

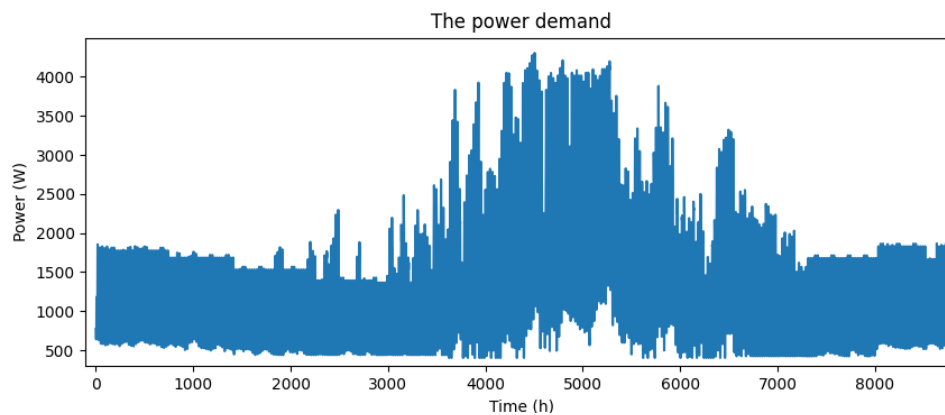


Figure 4. The energy demand of one year.

Table 1. The parameters of the studied system	
Variable	Value
Total PV power rate	2.5 kW
Total battery capacity	15 kWh

II.2. Mathematical models

This study focuses on energy forecasting in a typical house including several energy sources (grid, photovoltaic and energy storage system). The ultimate goal of this work is to use Machine-learning techniques to accurately predict the total energy consumption cost of energy used from the grid.

1) PV power model

The PV output power can be calculated using the solar irradiance and ambient temperature parameters as shown in the equation (1)[14].

$$P_{pv} = 250 \cdot \frac{G}{1000} \cdot [1 + k_t(T_{amb} + 0.0256G) - 25] \cdot N \quad (1)$$

where P_{pv} is the output power of PV panels, N is the number of PV panels, k_t is coefficient whose value equal to $3.7 \cdot 10^{-3} (1/^\circ\text{C})$, T_{amb} is the ambient temperature.

2) The energy from grid

The power used from the grid can be calculated as the difference between the power supplied by the PV and battery and the power demand by the loads, as shown in equation (2) while the energy is calculated by equation (3).

$$P_G = (P_{pv} + P_{B_disc}) \cdot \eta_{inv} - P_l \quad (2)$$

$$E_G(t) = P_G(t) \cdot \Delta T \quad (3)$$

where P_G , E_G are the output power and the energy from the grid, P_l is the power demanded from the household loads, P_{B_disc} is the power discharged from the battery, η_{inv} is the inverter efficiency, ΔT is the time step wish equal one hour in our case.

3) Energy cost model

To calculate the cost of energy consumption, the price of the Algerian electrical operator (Sonelgaz) was considered. However, the cost of one kWh of energy consumption is affected by two different prices. The first one is the price C_1 related to the period of energy consumption in the day (Table 2) while the second is the price C_2 related to the amount of energy consumed in the season (Table 3).

$$C_{kWh} = C_1 \cdot C_2 \quad (4)$$

$$C_{total} = \sum_t^n (E_G(t) \cdot C_{kWh}) \quad (5)$$

where C_{kWh} , C_{total} are the cost of kWh and the total cost of energy used from the grid, n is total time step of dataset (which equal the number of hour for each season).

Table 2. The prices related to the energy consumption period

Period	Hours	Coefficient price
Night hours	11:00 p.m. to 6 a.m.	0.677
Off peak hours	6 a.m. - 5 p.m. and 9 p.m. -11.00 p.m.	1.216
Peak hours	5 p.m. -9 p.m.	4.562

Table 3. The prices related to the amount of energy consumption

Amount	Price (DZD)
0 to 125 kWh	1.778
125 to 250 kWh	4.179
250 to 1000 kWh	4.812
Up to 1000 kWh	5.479

II.3. Forecasting methodology

In this study, the prediction of energy consumption costs is carried out using various machine learning algorithms. The methodology of the forecasting considered in this study consists of the following key steps:

- Step 1: Data Selection and Preparation

The process begins with selecting a representative dataset for training the machine learning models. This dataset must cover a diverse range of scenarios and include variables that significantly influence energy consumption cost. In this study, three primary features were used: global horizontal irradiation, ambient temperature, and power demand.

- Step 2: Algorithm Selection

Based on the characteristics of the dataset and the prediction task, suitable regression-based machine learning algorithms are selected. Algorithms such as Linear Regression, Random Forest, Support Vector Machine, and Artificial Neural Networks are considered for their predictive capabilities and adaptability to different data types.

- Step 3: Performance Evaluation Using Cross-Validation

To assess the generalization performance of the selected models, the k-fold cross-validation technique is applied. By dividing the training data into 10 subsets ($k=10$), the model is iteratively trained and validated on different folds. This approach ensures a robust evaluation of the model's predictive accuracy and reduces the risk of overfitting.

- Step 4: Model Training and Error Calculation

The selected algorithms are trained using the prepared dataset. During the training process, the models learn underlying patterns and relationships between the input features and the target variable. The performance of each model is quantified by calculating the Mean Squared Error (MSE), which measures the average prediction error.

- Step 5: Prediction and Application

After training and validation, the finalized machine learning model is deployed to predict energy consumption costs for new, unseen data. By providing input variables such as global irradiation, temperature, and power demand, the model generates accurate cost predictions. These predictions enable residential users to optimize energy usage, make data-driven decisions, and reduce overall energy costs effectively.

III. Results and Discussion

In this section, the results obtained from forecasting energy consumption costs for various seasons using different models. The selected machine learning algorithms were executed by Python version 3.11. To predict the cost of energy consumption, three parameters were used as input to train the models (irradiation, ambient temperature and energy demand) and the cost of energy as output value for three months with a step time of one hour.

III.1. Quarter one of the year

Table 4 presents a summary of the energy cost forecasts for the first quarter (January, February and March), while Figure 5 illustrates the regression curves associated with each algorithm.

These forecasts highlight the variations in performance of the different machine learning models in estimating energy consumption costs. During this period, the total actual energy cost amounts to DZD 14,313.814, while the forecasts of the different algorithms are as follows:

Table 4. The energy cost prediction results for the first quarter

Predictor algorithm	Predicted energy cost (DZD)	Real energy cost (DZD)	Error (DZD)	Error (%)	MSE
LR	14275.374	14313.814	-38.44	0.269	48.968
RF	14258.875	14313.814	-26.94	0.188	3.851
SVM	5591.887	14313.814	-8721.928	60.934	137.604
ANN	14432.677	14313.814	118.863	0.830	28.479

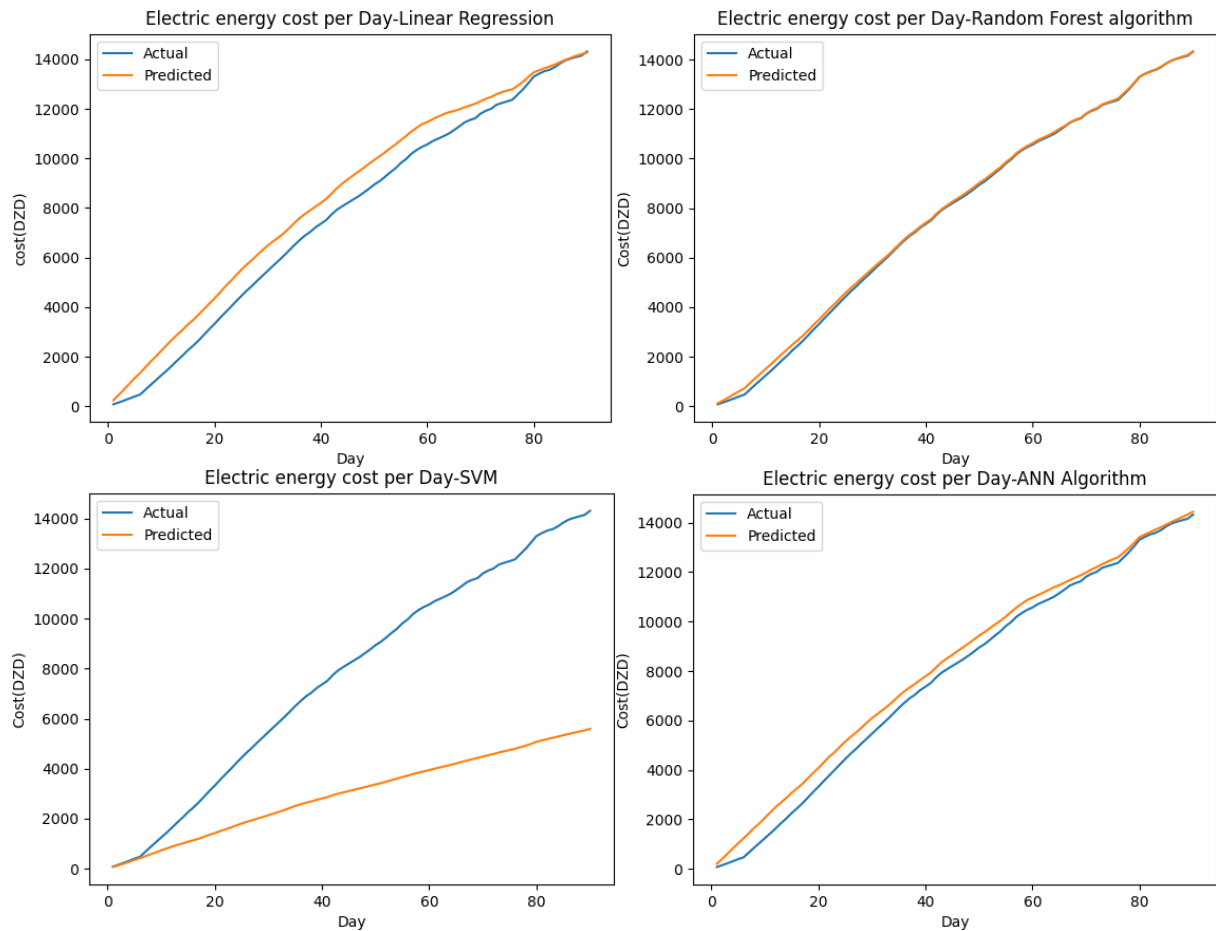


Figure 5. Energy cost prediction for the first quarter.

- **Linear Regression (LR):** The LR model performed well, with a total predicted cost of 14275.374 DZD, which is very close to the actual cost of 14313.814 DZD. The absolute error is -38.44 DZD, corresponding to a relative error of 0.269%, making it one of the most accurate models in this context.
- **Random Forest (RF):** The RF algorithm achieved the best performance among the models, predicting a total cost of 14258.875 DZD, with an absolute error of -26.94 DZD and a relative error of 0.188%. This indicates the RF model's robustness and its ability to capture the underlying patterns in the data effectively.
- **Support Vector Machine (SVM):** The SVM model significantly underperformed, with a predicted cost of 5591.887 DZD, resulting in a large absolute error of -8721.928 DZD and a relative error of 60.934%. This substantial deviation suggests that the SVM algorithm struggled to model the relationship between the input features and energy consumption costs, potentially due to its sensitivity to data scaling or the complexity of the dataset.
- **Artificial Neural Networks (ANN):** The ANN model also performed well, with a predicted cost of 14432.677 DZD, slightly overestimating the actual cost. The absolute error of 118.863 DZD corresponds to a relative error of 0.830%, which, while higher than LR and RF, is still within an acceptable range for practical applications.

III.2. Quarter two of the year

Table 5 provides an overview of the energy cost predictions for the second quarter (April, May, and June), with Figure 6 displaying the regression curves corresponding to each algorithm. The results emphasize the ability of machine learning models to forecast energy consumption costs during this period.

The total actual energy cost for the second quarter is 15,592.805 DZD, while the predicted costs for each algorithm are as follows:

Table 5. The energy cost prediction results for the second quarter

Predictor algorithm	Predicted energy cost (DZD)	Real energy cost (DZD)	Error (DZD)	Error (%)	MSE
LR	15564.629	15592.805	-28.176	0.181	109.218
RF	15645.207	15592.805	52.403	0.336	4.974
SVM	2792.304	15592.805	-12800.501	82.092	255.306
ANN	15335.168	15592.805	-257.637	1.652	33.248

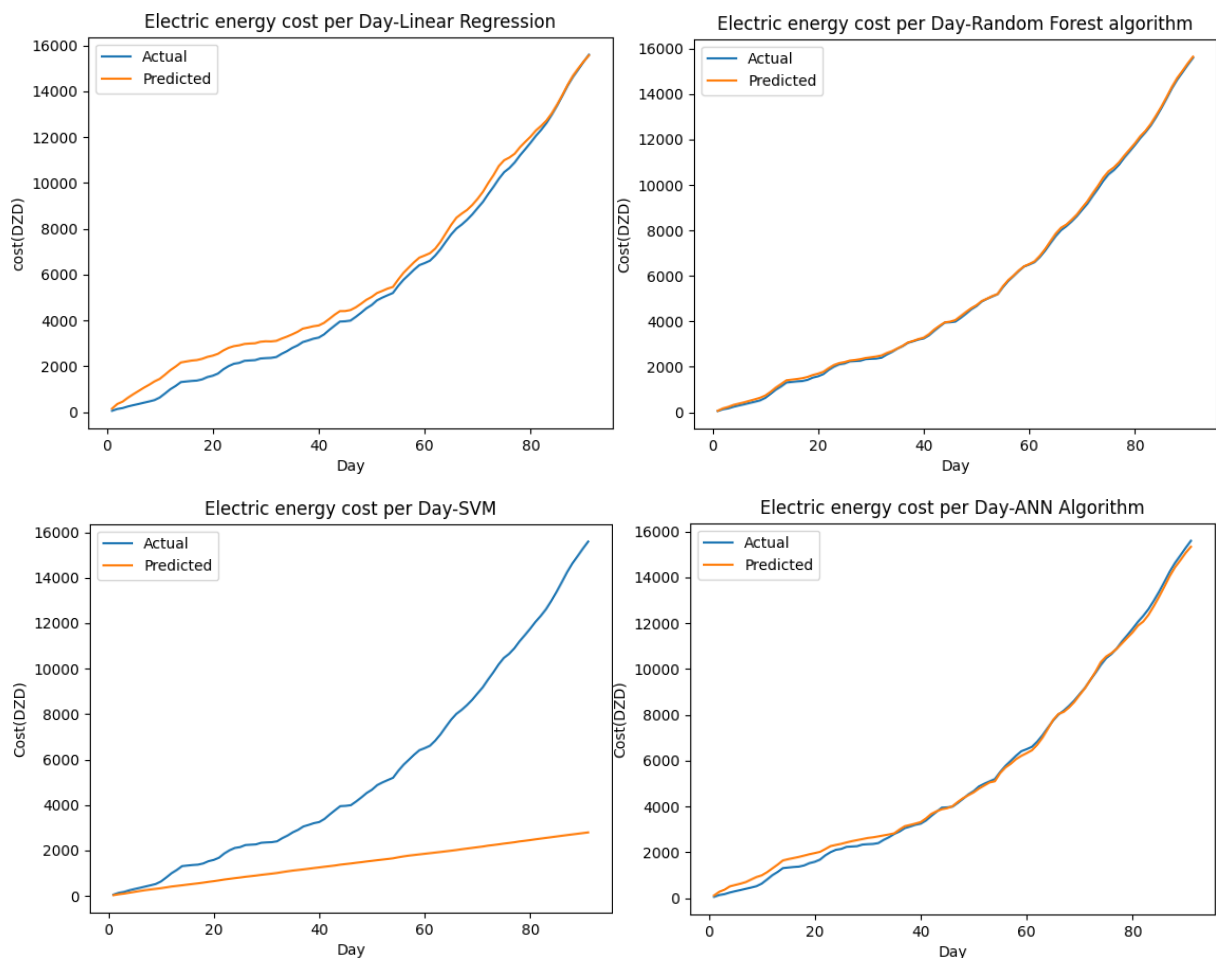


Figure 6. Energy cost prediction for the second quarter.

- Linear Regression (LR): The LR model performed exceptionally well, predicting a total cost of 15564.629 DZD, which is very close to the actual cost of 15592.805 DZD. The absolute error is -28.176 DZD, corresponding to a relative error of 0.181%, making LR the most accurate algorithm for this quarter.
- Random Forest (RF): The RF algorithm also delivered strong performance, with a predicted cost of 15645.207 DZD and an absolute error of 52.403 DZD. This corresponds to a relative error of 0.336%, which is slightly higher than LR. RF continues to demonstrate its reliability in capturing complex relationships in the data.
- Support Vector Machine (SVM): Similar to the first quarter, the SVM model performed poorly, predicting a total cost of only 2792.304 DZD, resulting in a massive absolute error of -12800.501 DZD and a relative error of 82.092%. This significant underestimation underscores SVM's unsuitability for this task, likely due to the same issues noted previously.
- Artificial Neural Networks (ANN): The ANN model showed moderate performance, predicting a cost of 15335.168 DZD. The absolute error was -257.637 DZD, with a relative error of 1.652%. While its predictions were less accurate compared to LR and RF, ANN still provided reasonably acceptable results.

III.3. Quarter three of the year

Table 6 summarizes the energy cost predictions for the third quarter (July, August, and September), while Figure 7 illustrates the regression curves for each algorithm. The prediction results for this quarter further reveal the strengths and weaknesses of machine learning algorithms in estimating energy consumption costs.

During this period, the total actual energy cost is 32,884.313 DZD, with the predicted costs for each algorithm as follows:

- Linear Regression (LR): The LR model provided highly accurate predictions, with a predicted cost of 32836.864 DZD, closely matching the actual cost of 32884.313 DZD. The absolute error was -47.449 DZD, corresponding to a relative error of 0.144%, making LR the most accurate algorithm for this quarter.
- Random Forest (RF): The RF model performed well, predicting a total cost of 32954.907 DZD, with an absolute error of 70.594 DZD and a relative error of 0.215%. While slightly less accurate than LR, RF maintained strong predictive performance and demonstrated its robustness.
- Support Vector Machine (SVM): The SVM model continued to underperform significantly, predicting a total cost of only 9006.306 DZD. This resulted in a substantial absolute error of -23878.008 DZD and a relative error of 72.612%, indicating that SVM is not suitable for this prediction task without significant adjustments.
- Artificial Neural Networks (ANN): The ANN model showed moderate accuracy, predicting a cost of 33420.703 DZD. The absolute error was 536.39 DZD, corresponding to a relative error of 1.631%, which is higher than both LR and RF but still within an acceptable range for some applications.

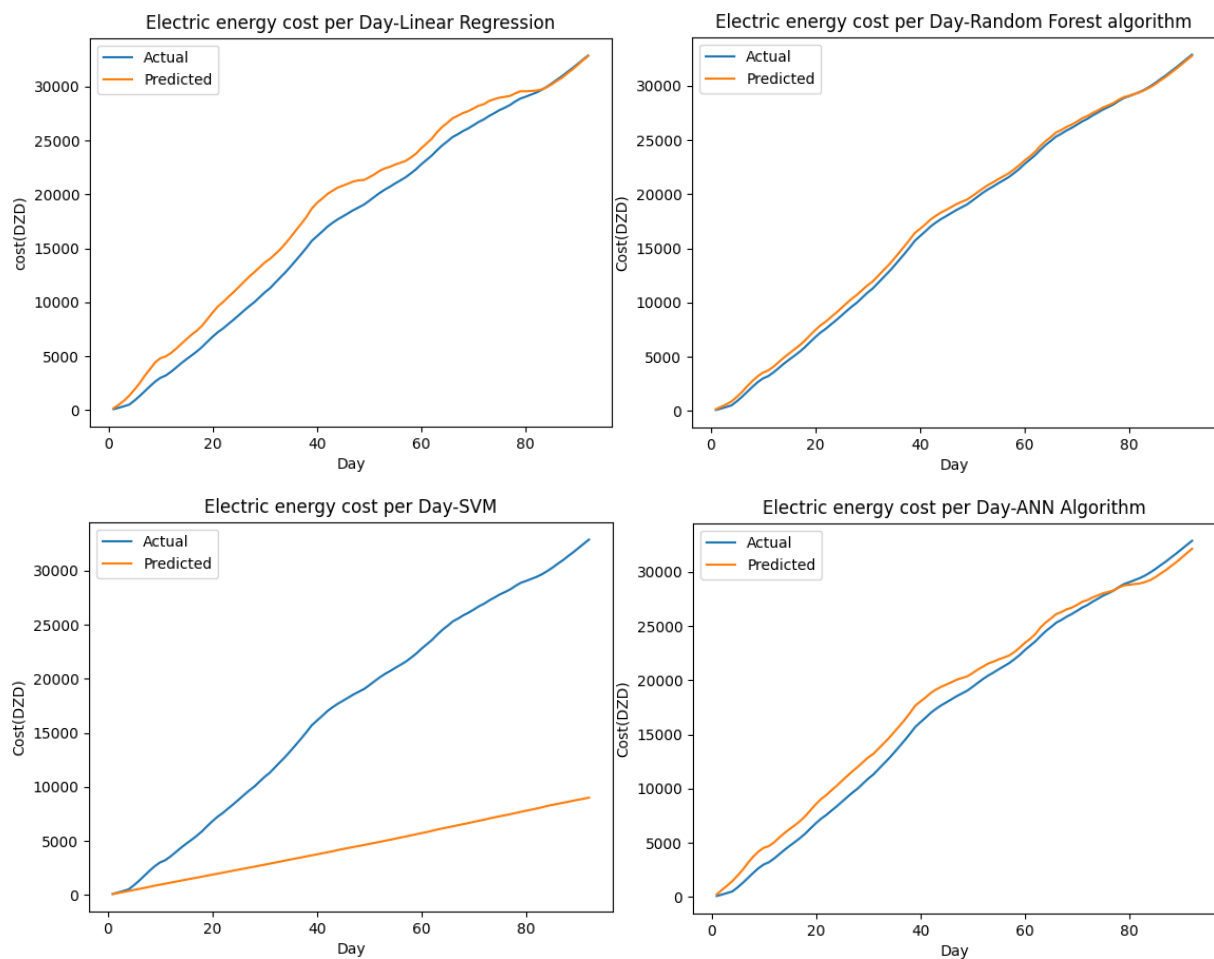


Figure 7. Energy cost prediction for the third quarter.

Table 6. The energy cost prediction results for the third quarter

Predictor algorithm	Predicted energy cost (DZD)	Real energy cost (DZD)	Error (DZD)	Error (%)	MSE
LR	32836.864	32884.313	-47.449	0.144	218.758
RF	32954.907	32884.313	70.594	0.215	16.513
SVM	9006.306	32884.313	-23878.008	72.612	652.963
ANN	33420.703	32884.313	536.39	1.631	110.376

III.4. Quarter four of the year

Table 7 provides a summary of the energy cost predictions for the fourth quarter (October, November, and December), with Figure 8 showcasing the regression curves for each algorithm. These results highlight the varying performance levels of machine learning algorithms in forecasting energy consumption costs during this period.

Table 7. The energy cost prediction results for the fourth quarter

Predictor algorithm	Predicted energy cost (DZD)	Real energy cost (DZD)	Error (DZD)	Error (%)	MSE
LR	19127.688	18770.027	357.661	1.905	70.679
RF	18761.427	18770.027	-8.599	0.046	1.709
SVM	6597.788	18770.027	-12172.238	64.849	200.913
ANN	19194.615	18770.027	424.589	2.262	21.997

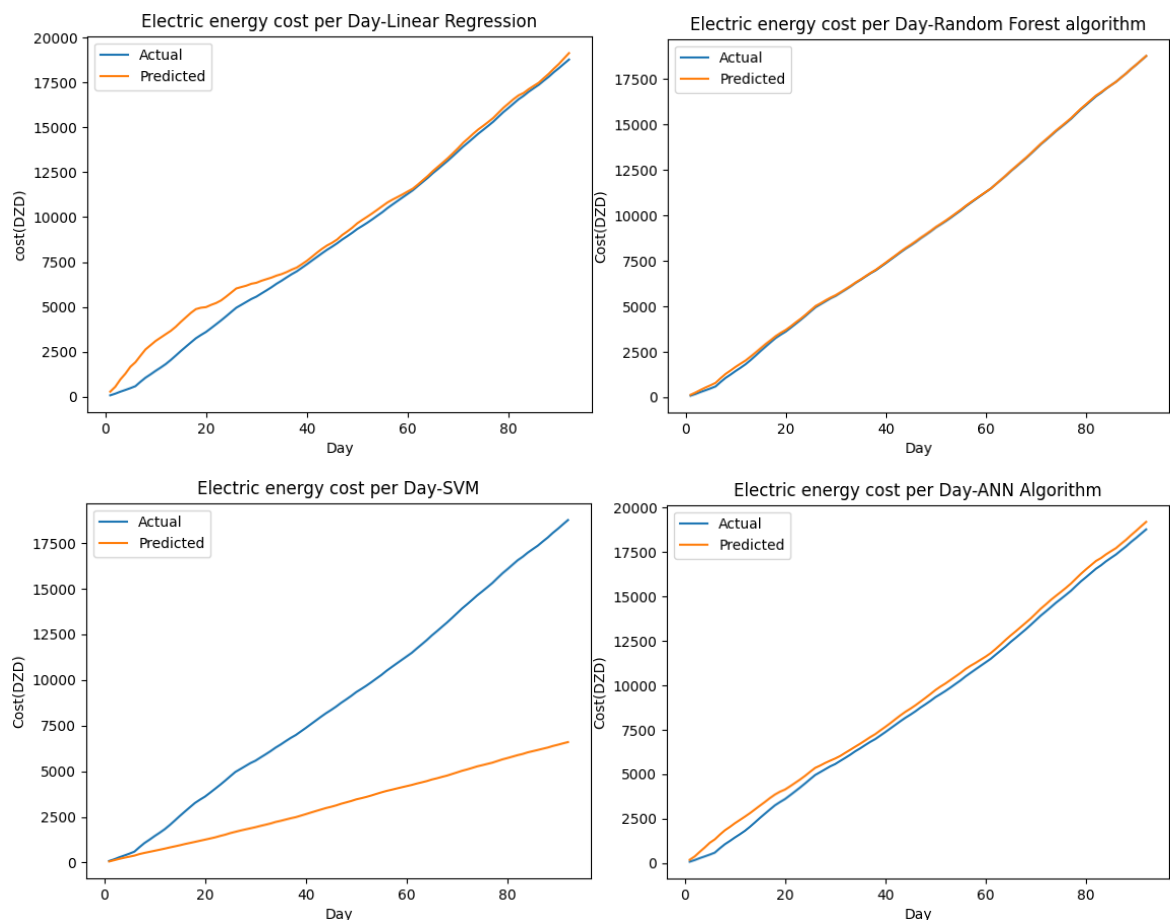


Figure 8. Energy cost prediction for the fourth quarter.

For the fourth quarter, the total real energy cost equal 18770.027 DZD, while the predicted energy consumption for each algorithm is:

- Linear Regression (LR): The LR model predicted a total energy cost of 19127.688 DZD, resulting in an absolute error of 357.661 DZD and a relative error of 1.905%. While the prediction was close, it was less precise compared to the Random Forest model.
- Random Forest (RF): The RF model delivered the most accurate prediction, with a total cost of 18761.427 DZD. The absolute error was only -8.599 DZD, corresponding to a relative error of 0.046%, making RF the best-performing algorithm for this quarter.
- Support Vector Machine (SVM): The SVM model again performed poorly, predicting a total cost of 6597.788 DZD, which is significantly lower than the actual cost. This resulted in a massive absolute error of 12172.238 DZD and a relative error of 64.849%, highlighting its continued unsuitability for this prediction task.
- Artificial Neural Networks (ANN): The ANN model predicted a total cost of 19194.615 DZD, with an absolute error of 424.589 DZD and a relative error of 2.262%. While its predictions were relatively close to the actual values, the ANN model was outperformed by both RF and LR.

Figure 9 visualizes the Mean Squared Error (MSE) for energy cost predictions across four quarters for each machine learning algorithms: Linear Regression (LR), Random Forest (RF), Support Vector Machine (SVM), and Artificial Neural Networks (ANN).

- First Quarter: The MSE values are 48.968 for LR, 3.851 for RF, 137.604 for SVM, and 28.479 for ANN.
- Second Quarter: The MSE values are 109.218 for LR, 4.974 for RF, 255.306 for SVM, and 33.248 for ANN.
- Third Quarter: The MSE values are 218.758 for LR, 16.513 for RF, 652.963 for SVM, and 110.376 for ANN.
- Fourth Quarter: The MSE values are 70.679 for LR, 1.709 for RF, 200.913 for SVM, and 21.997 for ANN.

Based on the previous results, the Random Forest (RF) algorithm demonstrated the highest accuracy in forecasting energy consumption costs, with the lowest total error in energy consumption predictions. It was followed by Linear Regression (LR) and Artificial Neural Networks (ANN), while the Support Vector Machine (SVM) showed the poorest performance in terms of prediction errors.

Regarding the Mean Squared Error (MSE), the Random Forest (RF) also achieved the best performance (the best fit) with the lowest MSE, followed by the ANN and LR algorithms, whereas the SVM recorded the highest MSE, indicating the weakest performance.

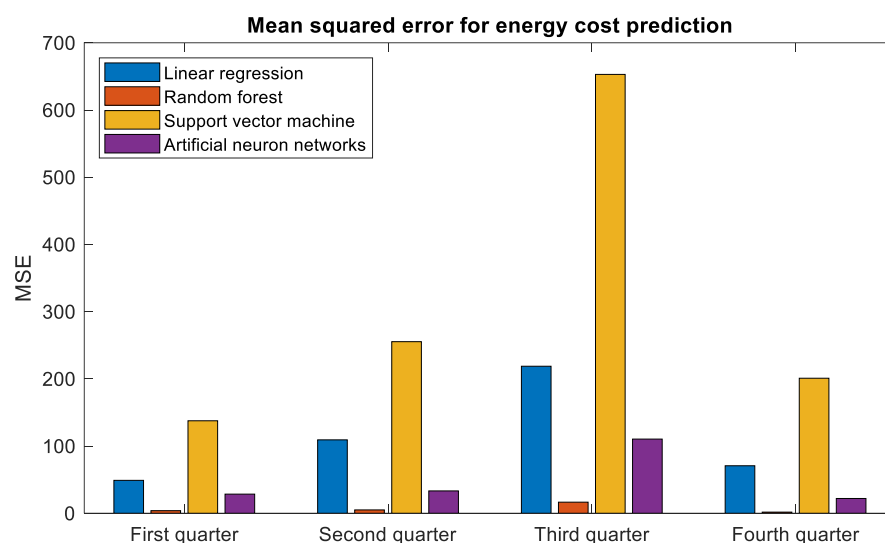


Figure 9. The MSE for energy cost prediction.

IV. Conclusion

In this paper, the prediction of energy consumption costs using various machine learning algorithms, including Linear Regression (LR), Random Forest (RF), Support Vector Machine (SVM), and Artificial Neural Networks (ANN), is presented. The algorithms were trained on four seasonal datasets representing winter, spring, summer, and autumn.

The results demonstrate that the Random Forest (RF) model outperformed the other algorithms achieving the highest accuracy in predicting energy consumption costs, with the lowest Mean Squared Error (MSE) values of 3.851, 4.974, 16.513, and 1.709 for the first, second, third, and fourth quarters, respectively. These forecasts provide householders with actionable insights to optimize their energy usage, helping them reduce their overall energy costs effectively.

References

- [1] IEA, International Energy Agency (IEA) World Energy Outlook 2022, 2022. <https://www.iea.org/reports/world-energy-outlook-2022>.
- [2] K. Steemers, G.Y. Yun, Household energy consumption: A study of the role of occupants, *Build. Res. Inf.* 37 (2009) 625–637. <https://doi.org/10.1080/09613210903186661>.
- [3] K. Holmberg, A. Erdemir, Influence of tribology on global energy consumption, costs and emissions, *Friction*. 5 (2017) 263–284. <https://doi.org/10.1007/s40544-017-0183-5>.
- [4] V. San, T. Sriv, V. Spoann, S. Var, S. Seak, Economic and environmental costs of rural household energy consumption structures in Sameakki Meanchey district, Kampong Chhnang Province, Cambodia, *Energy*. 48 (2012) 484–491. <https://doi.org/https://doi.org/10.1016/j.energy.2012.10.017>.
- [5] S.U. Khan, N. Khan, F.U.M. Ullah, M.J. Kim, M.Y. Lee, S.W. Baik, Towards intelligent building energy management: AI-based framework for power consumption and generation forecasting, *Energy Build.* 279 (2023). <https://doi.org/10.1016/j.enbuild.2022.112705>.
- [6] Asem Alzoubi, Machine Learning for Intelligent Energy Consumption in Smart Homes, *Int. J. Comput. Inf. Manuf.* 2 (2022) 62–75. <https://doi.org/10.54489/ijcim.v2i1.75>.
- [7] Y. Himeur, K. Ghanem, A. Alsalemi, F. Bensaali, A. Amira, Artificial intelligence based anomaly detection of energy consumption in buildings: A review, current trends and new perspectives, *Appl. Energy*. 287 (2021). <https://doi.org/10.1016/j.apenergy.2021.116601>.
- [8] I. Antonopoulos, V. Robu, B. Couraud, D. Kirli, S. Norbu, A. Kiprakis, D. Flynn, S. Elizondo-Gonzalez, S. Wattam, Artificial intelligence and machine learning approaches to energy demand-side response: A systematic review, *Renew. Sustain. Energy Rev.* 130 (2020). <https://doi.org/10.1016/j.rser.2020.109899>.
- [9] K.T. Chui, M.D. Lytras, A. Visvizi, Energy sustainability in smart cities: Artificial intelligence, smart monitoring, and optimization of energy consumption, *Energies*. 11 (2018) 1–20. <https://doi.org/10.3390/en11112869>.
- [10] L.H.M. Truong, K.H.K. Chow, R. Luevisadpaibul, G.S. Thirunavukkarasu, M. Seyedmahmoudian, B. Horan, S. Mekhilef, A. Stojcevski, Accurate prediction of hourly energy consumption in a residential building based on the occupancy rate using machine learning approaches, *Appl. Sci.* 11 (2021) 1–19. <https://doi.org/10.3390/app11052229>.
- [11] S. Bourhnane, M.R. Abid, R. Lghoul, K. Zine-Dine, N. Elkamoun, D. Benhaddou, Machine learning for energy consumption prediction and scheduling in smart buildings, *SN Appl. Sci.* 2 (2020). <https://doi.org/10.1007/s42452-020-2024-9>.
- [12] A. Wang, J.C.K. Lam, S. Song, V.O.K. Li, P. Guo, Can smart energy information interventions help householders save electricity? A SVR machine learning approach, *Environ. Sci. Policy*. 112 (2020) 381–393. <https://doi.org/10.1016/j.envsci.2020.07.003>.
- [13] H.T. Dinh, K. haeng Lee, D. Kim, Supervised-learning-based hour-ahead demand response for a behavior-based home energy management system approximating MILP optimization, *Appl. Energy*. 321 (2022) 119382. <https://doi.org/10.1016/j.apenergy.2022.119382>.
- [14] A.-K. Daud, M.S. Ismail, Design of isolated hybrid systems minimizing costs and pollutant emissions, *Renew. Energy*. 44 (2012) 215–224. <https://doi.org/https://doi.org/10.1016/j.renene.2012.01.011>.