A Novel Optimized Variant of Machine Learning Algorithm for Accurate Energy Demand Prediction for Tetouan City, Morocco

Natalie Rose[1]*, Otis Osbourne[2], Neil Williams[3], Syed Sajjad Hussain Rizvi[4]

[1,2,3,4]University of the Commonwealth Caribbean, Kingston, Jamaica
[4]SZABIST University, Karachi, Pakistan

ithod@ucc.edu.jm[1], itfaculty@ucc.edu.jm[2], itlecturer@ucc.edu.jm[3], sshussainr@gmail.com[4]

Abstract. Machine learning (ML) algorithms are an essential component of intelligent energy management systems. In the year 2021, a benchmark dataset of the power consumption of Tetouan city was published to train an ML algorithm for accurate energy demand prediction. However, parametric and empirical investigations for the best ML algorithm on this dataset are still undetermined. In this study, an exhaustive parametric evaluation of 26 ML variants is presented to advocate for the best algorithm for energy demand prediction in Tetouan city. After a thorough evaluation, the proposed Bayesian Fine Tree (BFT) outperforms the traditional Fine Tree algorithm. The simulation results provide strong evidence that the BFT is best at predicting the energy demand of Tetouan City.

Keywords: Machine Learning· Optimization· Hyper-Parameters· Data Driven Energy Management· Robust Energy Management· Power Consumption of Tetouan City Dataset· Bayesian Optimizer.

1. Introduction

Electricity is a critical and limited energy resource around the world [1]. However, the drastic increase in the demand for this limited resource has been observed around the world [2]. Researchers have proposed to handle this demand in two distinct ways. First, to increase the generation capacity of electricity and alternate energy resources [3]. Second, to have robust energy management methods [4]. In the literature, ML algorithms are reported as the best method in this case [5-9]. In recent literature, the benchmarked dataset that have been presented include REDD (Reference Energy Disaggregation Dataset) [10], UK-DALE (UK Domestic Appliance-Level Electricity) [11], Smart* Data Set [12], Pecan Street Data Set [13], AMPds (Almanac of Minutely Power dataset) [14], EcoSense [15], CEC (Commercial End-Use Survey) [16], Global Energy Forecasting Competition 2012 [17]. Existing studies have revealed that the scope of the major benchmarked dataset for energy demand prediction was limited to households, universities, buildings, or industries. In 2021, a city-level dataset was presented in the literature, which comprised of power consumption data of Tetouan city

C. Kiran Mai et al. (eds.), *Proceedings of the Fourth International Conference on Advances in Computer Engineering and Communication Systems (ICACECS 2023)*, Atlantis Highlights in Computer Sciences 18,
https://doi.org/10.2991/978-94-6463-314-6_7

[18](https://www.kaggle.com/datasets/gmkeshav/tetuan-city-power-consumption). The input attributes of the datasets included Date, Time, Temperature, Humidity, Wind Speed, General Diffuse Flows, and Diffuse Flows. Likewise, the output attributes were power consumption of zones 1, 2, and 3. Due to the recency of the dataset, limited research has been reported on this dataset, the Comparison of Machine Learning Algorithm for the Power Consumption Prediction - Case Study of Tetouan city [18], and Energy consumption prediction model with deep inception residual network inspiration and LSTM [19]. In [18], the researcher has proposed to employ only 5 machine learning algorithms with only one optimizer, the Grid-search optimizer. Moreover, only Root Mean Square Error (RMSE), and Mean Absolute Error (MAE) were taken as the measure of the performance parameter. This study reveals that the random forest is found to be the best from among the five given algorithms.

The scope of this study reveals that due to the limited number of machine learning algorithms employed for comparison, only two parameters were used as the measure of performance, one optimizer was employed, a limited graphical illustration of the results, and the non-availability to propose the new algorithm. These issues were considered and addressed in this study. In the recent literature is found to be deficient to present the comprehensive evaluation of machine learning algorithms and the best candidate of ML is presented in the literature. This has created a pressing need to first present a comprehensive parametric evaluation of ML algorithms and also introduce the best ML model for higher accuracy. The scope of this study is bi-folded. First, an exhaustive parametric evaluation of 26 variants of the machine learning algorithm is presented to advocate the best candidate of machine learning algorithm for energy demand prediction of Tetouan city. The performance was compared as the measure of RMSE, Mean Square Error (MSE), MAE, and R-squared for both training and testing. After thorough evaluation, the Fine Tree algorithm is found to be the best candidate for the said application. Subsequently, to improve the performance of the Fine Tree Algorithm with the hybridization including benchmark optimizers such as Bayesian optimizer, random search, and grid search.

Finally, a Bayesian Fine Tree (BFT) has been presented for the energy demand prediction of Tetouan city. The simulation results advocate that the BFT is performing better than the standard fine tree algorithm for energy demand prediction of Tetouan city. In summary, the literature review highlights the need for an optimized variant of a machine learning algorithm for accurate energy demand prediction in Tetouan City.

The novelty of this study, an exhaustive parametric evaluation of 26 ML variants is presented to advocate for the best algorithm for energy demand prediction of Tetouan city. After a thorough evaluation, the proposed a novel Bayesian Fine Tree (BFT) algorithm that outperforms among all ML algorithm. The simulation results provide strong evidence that the BFT is best in predicting the energy demand of Tetouan city.

2. Methodology

In the methodology to develop an optimized variant of a machine learning algorithm for accurate energy demand prediction for Tetouan City, Morocco; in its first phase, the dataset was pre-processed by filtering the missing values, sum up the power consumption of all three zones of Tetouan City, dividing the data set into 70% training set and 30% testing set, and segmenting the date and time stamp. In the second phase 26 variants of machine learning algorithm were trained and tested. The performance of each variant was recorded as the function of RMSE, MSE, R-squared, and MAE for both training and testing.

These measures support the best candidate of the machine learning algorithm. Subsequently, the best candidate was optimized using Bayesian optimizer, random search, and grid search. Finally, the proposed BFT is the major contribution of this research. The best implementation methodology for the energy demand prediction of Tetouan city is the Bayesian Fine Tree (BFT) algorithm. The study conducted an exhaustive parametric evaluation of 26 machine learning variants, and the BFT algorithm outperformed the traditional Fine Tree algorithm in predicting the energy demand of Tetouan city.

The dataset used in the study was collected in 2017 and has 52,416 energy consumption data points in 10-minute windows beginning on January 1, 2017, until December 30, 2017, from three different distribution networks of Tetouan city. The study aimed to provide a practical framework capable of forecasting electricity load patterns in smart cities, and the results showed that the BFT algorithm is the best algorithm for energy demand prediction of Tetouan city.

The data consists of 52,416 observations of energy consumption on a 10-minute window. Every observation is described by 9 feature columns, namely, Date Time, Temperature, Humidity, and Wind Speed.

3. Proposed Algorithm and Results

This section presents the complete layout of the proposed BFT algorithm. The proposed algorithm is the hybridization of Fine Tree algorithm and the Bayesian optimizer. The Bayesian optimizer in principle optimizes the hyper-parameters of Fine Tree to render the optimal performance for Accurate Energy Demand Prediction for Tetouan City, Morocco.

Input:
X: Training data
Y: Training data labels
Parameters: list hyperparameters
Number of iterations for optimization
Output:
Bayesian Fine Tree (BFT)
Process:
Model Training and Testing
def Fine_Tree(parameters):
  Initialize Fine Tree model with given hyperparameters
  Train model
  Evaluate performance parameters on training set
  Test .Fine Tree (X, Y)
  Evaluate performance parameters on testing set
 return performance parameters
Bayesian Optimizer
 optimizer = Bayesian Optimizer(parameters, Trained Fine Tree, iterations)
 for i in range(iterations):
  Select hyperparameters
  Train Fine Tree model with hyperparameters
  Report validation score back to optimizer
  Optimizer update (params, score)
Re-train Fine Tree with optimum hyperparameters
Return BFT

Table 1 illustrates the performance evaluation of 26 variants of machine learning algorithms for power consumption prediction for Tetouan City, Morocco. The performance evaluation utilizes the function of Root Mean Square Error (RMSE), Mean Absolute Error (MAE), Mean Square Error (MSE), and R-Squared for both training and testing [20-21].

Table 1. Parametric Comparison of Machine Learning Algorithm to Predict Electrical Energy Consumption

		Training				Testing			
	Benchmark ML Algorithm	RMSE e+3	MSE	R^2 e-1	MAE e+03	MAE e+03	MSE	RMSE e+03	R^2 e-1
Linear Regression	Linear	10.5	1.10E+08	6.27	8.29	8.13	1.06E+08	10.3	6.37
	Interaction	9.75	9.51E+07	6.77	7.58	7.47	9.26E+07	9.62	6.84
	Robust	10.5	1.11E+08	6.24	8.24	8.07	1.07E+08	10.3	6.35
	Stepwise	9.7	9.51E+07	6.77	7.58	7.47	9.25E+07	9.62	6.84
Tree	Fine	2.91	8.49E+06	9.71	1.84	1.69	7.17E+06	2.68	9.76
	Medium	3.23	1.05E+07	9.64	2.17	2.02	9.19E+06	3.03	9.69
	Coarse	3.92	1.53E+07	9.48	2.70	2.58	1.42E+07	3.77	9.52
SVM	Linear	10.6	1.12E+08	6.20	8.23	8.06	1.08E+08	10.4	6.31
	Quadratic	9.22	8.50E+07	7.11	6.76	6.69	8.33E+07	9.13	7.16
	Cubic	6.07	3.69E+07	8.75	4.39	4.37	3.67E+07	6.06	8.75
	Fine Gaussian	3.97	1.58E+07	9.46	2.90	2.71	1.37E+07	3.70	9.53
	M Gaussian	5.37	2.88E+07	9.02	3.82	3.74	2.80E+07	5.30	9.04
	C-Gaussian	9.16	8.39E+07	7.15	6.84	6.67	8.06E+07	8.98	7.25
Ense mble	Boosted Tree	6.16	3.79E+07	8.71	4.71	4.71	3.76E+07	6.13	8.72
	Bagged Tree	2.95	8.70E+06	9.70	2.09	1.95	7.58E+06	2.75	9.74
Gaussian Process Regression	Squared Exponential	4.47	2.00E+07	9.32	3.31	3.22	1.91E+07	4.37	9.35
	Matern 5/2	3.84	1.47E+07	9.50	2.81	2.85	1.50E+07	3.88	9.49
	Exponential	3.59	1.29E+07	9.56	2.61	2.47	1.15E+07	3.40	9.61
	Rationale Quadratic	3.95	1.56E+07	9.47	2.91	3.13	1.82E+07	4.26	9.38
Artificial Neural Networks	Narrow	8.17	6.68E+07	7.73	6.16	6.10	6.59E+07	8.12	7.75
	Medium	8.04	6.46E+07	7.80	6.02	5.96	6.36E+07	7.97	7.83
	Wide Neural Network	4.06	1.65E+07	9.44	3.03	3.16	1.81E+07	4.25	9.38
	Bi-layer	6.83	4.66E+07	8.42	5.11	4.88	4.27E+07	6.53	8.54
	Tri-layer	5.86	3.44E+07	8.83	4.31	4.13	3.36E+07	5.80	8.85
Kernel	SVM Kernel	6.16	1.10E+08	6.27	8.29	8.13	1.06E+08	10.3	6.37
	Least Square Regression	2.95	9.51E+07	6.77	7.58	7.47	9.26E+07	9.62	6.84
Optimiz ed Fine Tree	Bayesian	4.47	1.11E+08	6.24	8.24	8.07	1.07E+08	10.3	6.35
	Grid Search	3.84	9.51E+07	6.77	7.58	7.47	9.25E+07	9.62	6.84
	Random Search	3.59	8.49E+06	9.71	1.84	1.69	7.17E+06	2.68	9.76

It is evident in Table 1 that Fine Tree presents the minimum error measure or higher prediction accuracy for power consumption prediction as compared to the other benchmark ML algorithms. In this study the proposed algorithm is compared with 26 benchmark machine learning algorithms as the function of RMSE, MSE, MAE, and R-squared. This parametric comparison revels that the Fine Tree is outperforming as compared to other benchmark ML algorithm. The minimum error measure supports the nomination of Fine Tree as the best candidate for power consumption prediction for

Tetouan City. Figure 1, Figure 2, Figure 3, and Figure 4 illustrate that actual values are in close proximity to the approximate linearity of FT, BFT, GFT, and RFT algorithm.

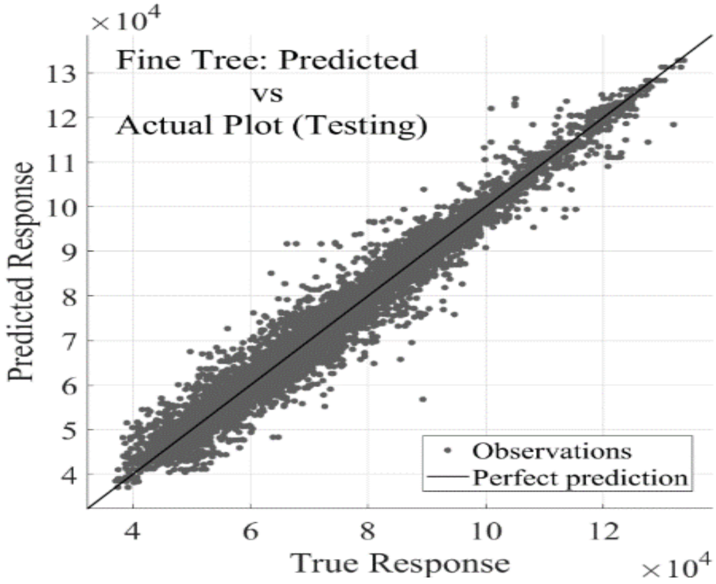

Fig. 1. Fine Tree: Predicted vs Actual Plot (Testing)

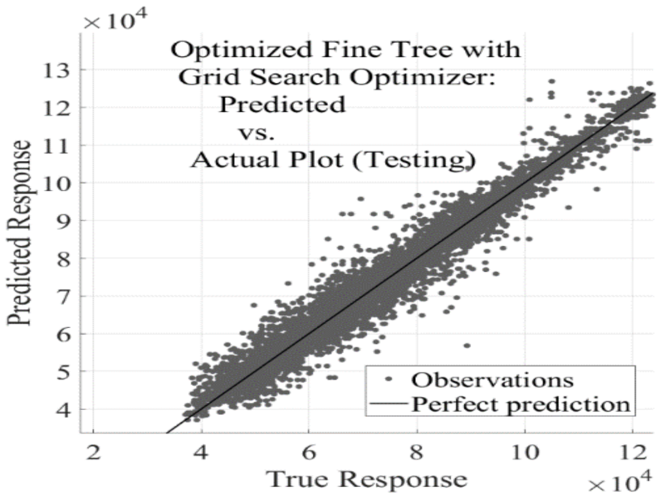

Fig. 2. Optimized Fine Tree with Grid Search Optimizer: Predicted vs Actual Plot (Testing)

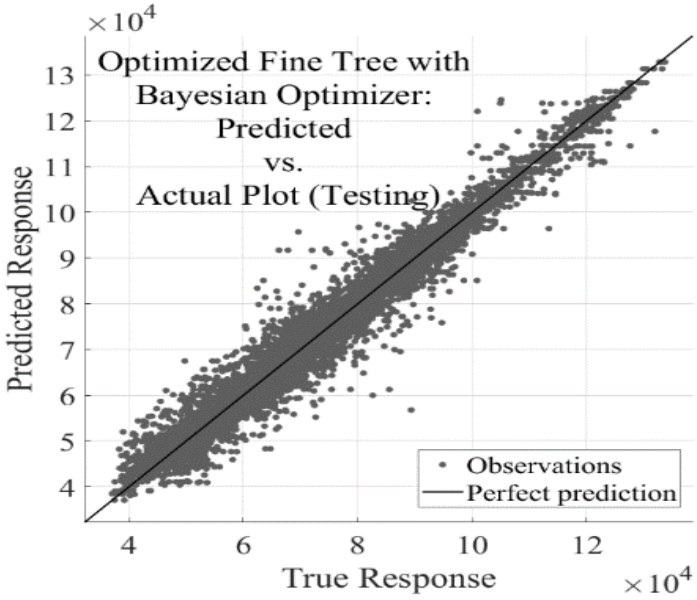

Fig. 3. Optimized Fine Tree with Bayesian Optimizer: Predicted vs Actual Plot (Testing)

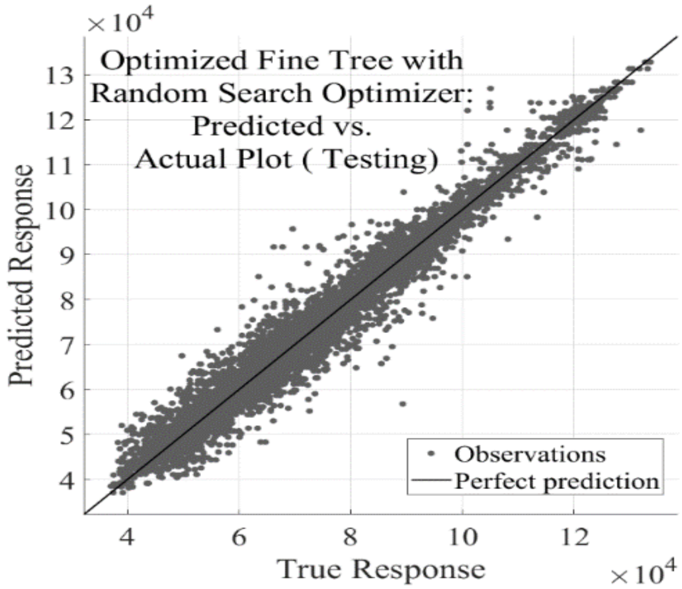

Fig. 4. Optimized Fine Tree with Random Search Optimizer: Predicted vs Actual Plot (Testing

)

In this study, the performance of the Fine Tree algorithm is further improved by creating three variants of Fine Tree. These variants are the hybridization of Bayesian optimizer, Grid search, and Random search optimizer. The variants are termed as Bayesian-Fine Tree (BFT), Grid-Fine Tree (GFT), and Random-Fine Tree (RFT) in this study. In this study, the performance of these three variants were evaluated as the function of RMS, MSE, MAE, and R-squared for accurate power consumption prediction for Tetouan City, Morocco. Table 1 reflects that the BFT is outperforming when compared to GFT and RFT. As a result, the performance measures of BFT report the least value of all the error measures as compared to Fine Tree, GFT, and RFT. Figure 5 shows the Minimum MSE of BFT with optimum point of hyper-parameters. It can be seen from the figure that the best point hyperparameter was recorded at the 7th iteration whereas for GFT in Figure 6 was recorded at the 9th iteration. Likewise, for RFT in Figure 7, it was found at the 11th iteration. This evidence advocates that the BFT has a higher convergence than GFT and RFT. Finally, the dependence plot of BFT, GFT, and RFT in respect to Figure 8, Figure 9, and Figure 10 has shown the competitive edge of BFT over GFT, and RFT. A partial dependence plot (PDP) is a visualization tool that shows the marginal effect one or two features have on the predicted outcome of a machine learning model. It can help us understand how different values of a particular feature impact a model's predictions.

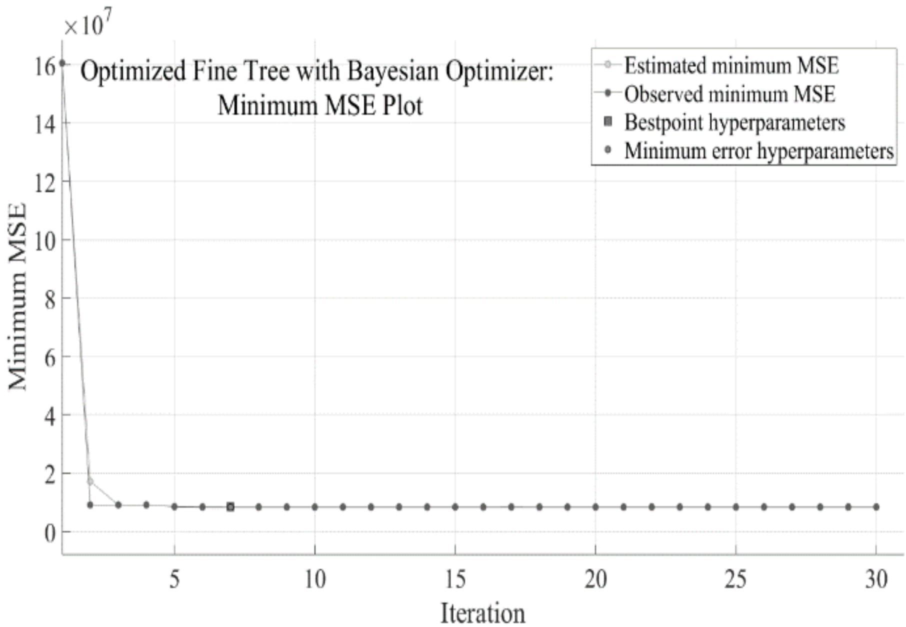

Fig. 5. Optimized Fine Tree Bayesian Optimizer: Minimum MSE Plot

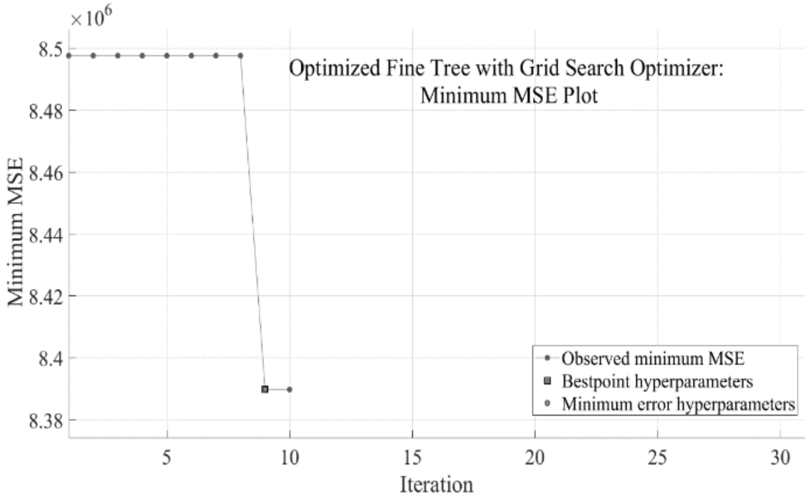

Fig. 6. Optimized Fine Tree with Grid Search Optimizer: Minimum MSE Plot

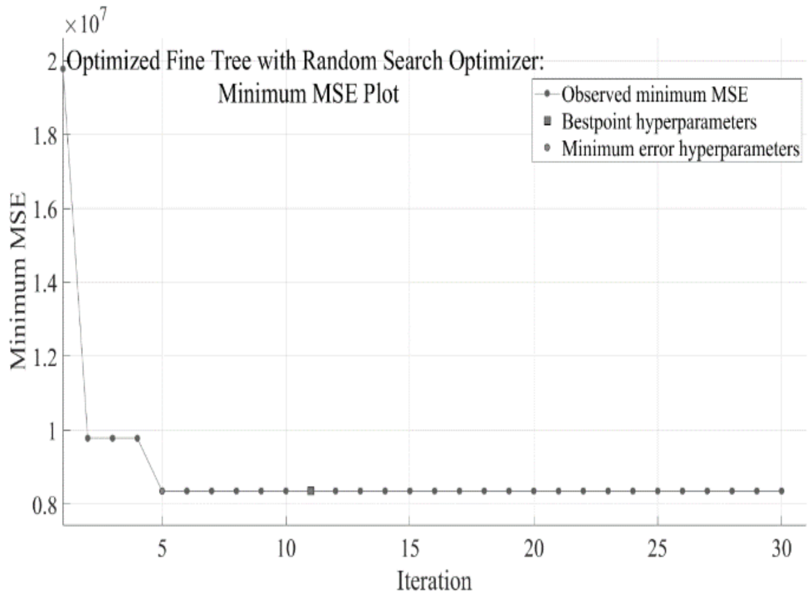

Fig. 7. Optimized Fine Tree with Random Search Optimizer: Minimum MSE Plot

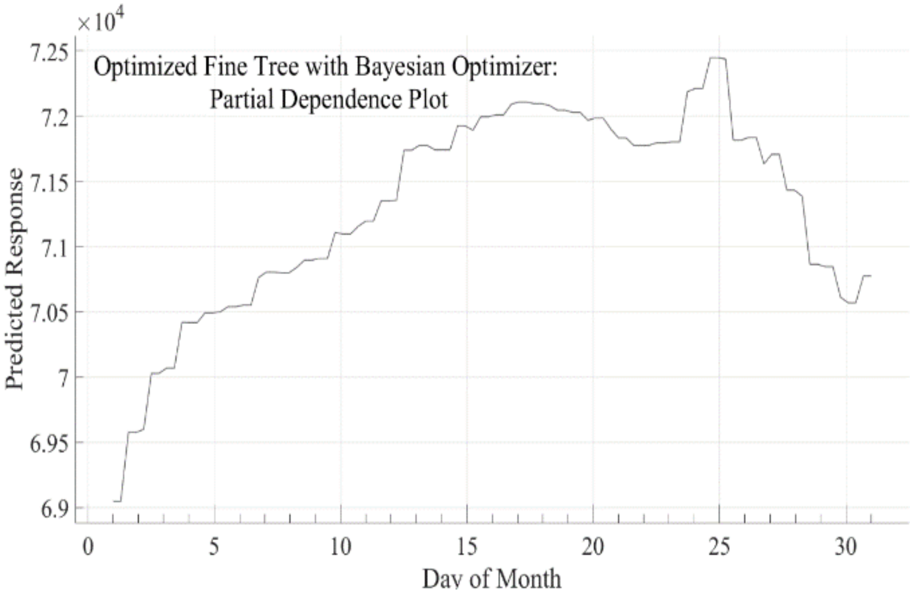

Fig. 8. Optimized Fine Tree with Bayesian Optimizer: Partial Dependence Plot

4. Conclusion and Future Work

This study has presented an optimized variants BFT for accurate energy demand prediction. The rationale of selecting the FT for optimization is its minimum error measure as compared to other variants of machine learning algorithms. The comparison is established after the systematics and empirical performance evaluation of FT with the benchmark machine learning algorithm as the function of RMSE, MSE, MAE, and R-squared for both training and testing. Furthermore, the performance of BFT is also compared with GFT, and RFT. All the indicator strongly advocated for the accuracy of BFT for accurate energy demand prediction for Tetouan City, Morocco. As the future recommendation the deep learning model can be applied to test and compare the performance. Moreover, the scope of explainable AI can be evaluated on said problem.

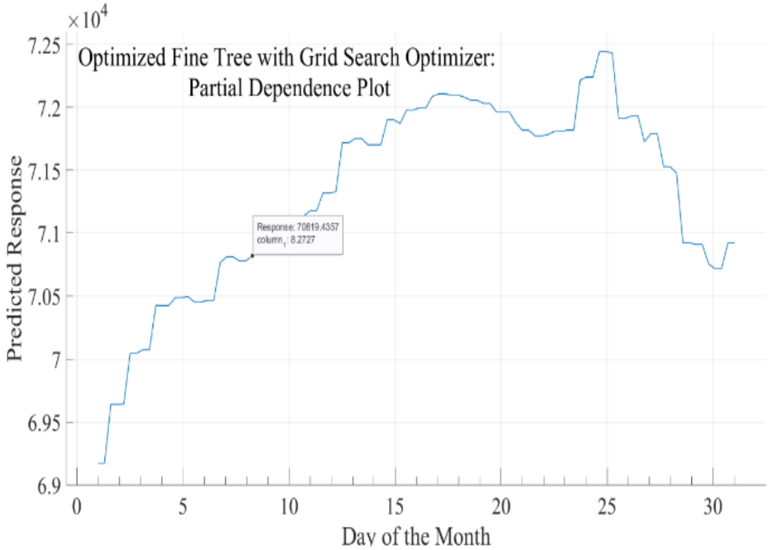

Fig. 9. Optimized Fine Tree with Grid Search Optimizer: Partial Dependence Plot

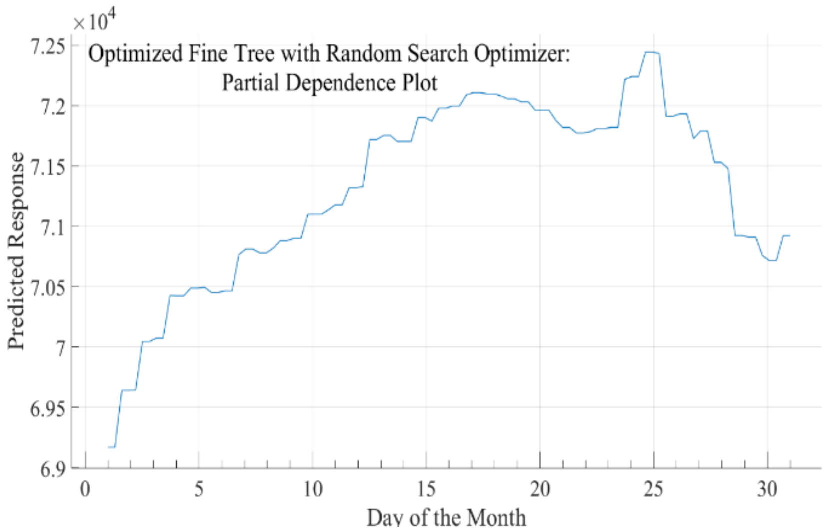

Fig. 10. Optimized Fine Tree with Random Search Optimizer: Partial Dependence Plot

References

1. Cebullaa, F., Naeglera, T., & Pohla, M.A. (2017). Electrical energy storage in highly renewable European energy systems: capacity requirements, spatial distribution, and storage dispatch.
2. Prada, R.B., & Souza, L.J. (1998). Voltage stability and thermal limit: constraints on the maximum loading of electrical energy distribution feeders.
3. Bongers, T., Kellermann, J., Franz, M.L., & Moser, A. (2016). Impact of curtailment of renewable energy sources on high voltage network expansion planning. 2016 Power Systems Computation Conference (PSCC), 1-8.
4. Chang, X., Xu, Y., Gu, W., Sun, H., Chow, M., & Yi, Z. (2021). Accelerated Distributed Hybrid Stochastic/Robust Energy Management of Smart Grids. IEEE Transactions on Industrial Informatics, 17, 5335-5347.
5. Charytanowicz, M., Olwert, A., Radziszewska, W., Jarnicka, J., Gajowniczek, K., Zabkowski, T.S., Brożyna, J., Mentel, G., & Matejko, G. (2022). Machine Learning for Energy Load Prediction and its Interpretation. 2022 IEEE 11th International Conference on Intelligent Systems (IS), 1-8.
6. Fan, L., Li, J., & Zhang, X. (2020). Machine learning load prediction methods for home energy management systems based on human behavior patterns recognition. CSEE Journal of Power and Energy Systems, 6, 563-571.
7. Alomoush, W., Khan, T.A., Nadeem, M., Janjua, J.I., Saeed, A., & Athar, A. (2022). Residential Power Load Prediction in Smart Cities using Machine Learning Approaches. 2022 International Conference on Business Analytics for Technology and Security (ICBATS), 1-8.
8. Gaboitaolelwe, J., Zungeru, A.M., Yahya, A., & Lebekwe, C.K. (2022). Electricity Load Prediction Using Machine Learning. 2022 International Conference on Smart Applications, Communications and Networking (SmartNets), 01-04.
9. Chaganti, R., Rustam, F., Daghriri, T., Díez, I. T., Mazón, J. L. V., Rodríguez, C. L., & Ashraf, I. (2022). Building Heating and Cooling Load Prediction Using Ensemble Machine Learning Model. Sensors (Basel, Switzerland), 22(19), 7692.
10. Kolter, J. Z., & Johnson, M. J. (2011, August). REDD: A public data set for energy disaggregation research. In Workshop on data mining applications in sustainability (SIGKDD), San Diego, CA (Vol. 25, No. Citeseer, pp. 59-62).
11. Kelly, J., & Knottenbelt, W. (2015). The UK-DALE dataset, domestic appliance-level electricity demand and whole-house demand from five UK homes. Scientific data, 2(1), 1-14.
12. Barker, S., Mishra, A., Irwin, D., Cecchet, E., Shenoy, P., & Albrecht, J. (2012). Smart*: An open data set and tools for enabling research in sustainable homes. SustKDD,
13. Haskell, B. Pecan Street, Inc.: Smart Grid, Smart Meter, Big Data.
14. Makonin, S., Popowich, F., Bartram, L., Gill, B., & Bajić, I. V. (2013, August). AMPds: A public dataset for load disaggregation and eco-feedback research. In 2013 IEEE electrical power & energy conference (pp. 1-6). IEEE.
15. Díaz, P., Olivares, T., Galindo, R., Ortiz, A., Royo, F., & Clemente, T. (2011). The ecosense project: an intelligent energy management system with a wireless sensor and actor network. In Sustainability in Energy and Buildings: Results of the Second International Conference on Sustainability in Energy and Buildings (SEB'10) (pp. 237-245). Springer Berlin Heidelberg.
16. Placet, M., Katipamula, S., Liu, B., Dirks, J., Xie, Y., Sullivan, G., ... & Williamson, I. R. (2010). Energy end-use patterns in full-service hotels: A case study. ACEEE summer study on energy efficiency in buildings. American council for an energy-efficient economy
17. Hong, T., Pinson, P., & Fan, S. (2014). Global energy forecasting competition 2012. International Journal of Forecasting, 30(2), 357-363.

18. Salam, A., & El Hibaoui, A. (2018, December). Comparison of machine learning algorithms for the power consumption prediction:-case study of tetouan city–. In 2018 6th International Renewable and Sustainable Energy Conference (IRSEC) (pp. 1-5). IEEE.
19. Salam, A., & El Hibaoui, A. (2021). Energy consumption prediction model with deep inception residual network inspiration and LSTM. Mathematics and Computers in Simulation,
20. S. Kedia and M. Bhushan, "Prediction of Mortality from Heart Failure using Machine Learning," 2022 2nd International Conference on Emerging Frontiers in Electrical and Electronic Technologies (ICEFEET), Patna, India, 2022, pp. 1-6, doi: 10.1109/ICEFEET51821.2022.9848348.
21. U. Verma, C. Garg, M. Bhushan, P. Samant, A. Kumar and A. Negi, "Prediction of students' academic performance using Machine Learning Techniques," 2022 International Mobile and Embedded Technology Conference (MECON), Noida, India, 2022, pp. 151-156, doi: 10.1109/MECON53876.2022.9751956.