



Research Article

Power Consumption Prediction in Urban Areas using Machine Learning as a Strategy towards Smart Cities

Emmanuel Gbenga DADA^{1*}, David Opeoluwa OYEWOLA² and Joseph Hurcha YAKUBU¹

¹Department of Mathematical Sciences, University of Maiduguri, Maiduguri, Nigeria

²Department of Mathematics and Computer Science, Federal University Kashere, Gombe, Nigeria

*Corresponding Author: gbengadada@unimaid.edu.ng

ARTICLE INFO:

Keyword:

Bagging (BAG),
Stochastic Gradient
Boosting (GBM),
Model Averaged Neural
Network (MAN),
K-Nearest Neighbours (KNN),
Power Consumption Prediction

ABSTRACT

The bedrock of every smart city is the continuous electricity supply. It is an important factor in the growth of smart cities. Smart transmission and distribution of power are crucial for the constant power supply in a smart city. The ability to correctly predict power consumption is a key factor in achieving demand-driven output through careful planning, and prompt usage of electricity generated. This paper applies Machine learning techniques to predict power consumption using historical data of power consumed in an urban area. Dataset obtained from Kaggle was used for the experiments in this paper. The paper used four machine learning models, namely, Bagging (BAG), Stochastic Gradient Boosting (GBM), Model Averaged Neural Network (MAN), and K-Nearest Neighbours (KNN). Experimental results showed that GBM outperforms other machine learning models used for the study with a prediction accuracy of 96%. This is an indication that GBM is a promising algorithm for predicting power consumption in a smart city if adopted.

Corresponding author: Emmanuel Gbenga DADA, Email: gbengadada@unimaid.edu.ng
Department of Mathematical Sciences, University of Maiduguri

INTRODUCTION

Power consumption prediction in Smart Buildings (SB) is very crucial for the implementation of Energy-efficient Control Systems. Most notably, power consumption prediction is a critical component of the innovative Smart Grids infrastructure, in which loads must be predicted and planned in real-time to account for the interconnected volatility in power consumption. As cities around the world have progressed technologically, electricity demand has risen to unimaginable peaks in recent years. If individual governments on a worldwide scale are unable to meet these energy needs, global warming challenges will worsen. A smart city can be defined as an innovative and beneficial metropolitan area that maximizes its resources to make sure a high standard of living is achieved. One of the most difficult problems in these cities is power conservation. Sun *et al.* (2016) defined smart communities as the amalgamation of many cities that incorporate and exploit these technologies to enhance a community's economic viability, sustainability, rejuvenation, and competitiveness.

In the last four decades, world energy consumption has risen by 100% (International Energy Agency, 2015). People have been forced to change their way of life and use resources more efficiently as a result of rising energy demand and the onset of an energy crisis. Urban area power efficiency has, without a doubt, become a big problem all over the world as the government of nations is making effort to attain smart cities that are characterised by power efficiency and greenhouse gas emission elimination. Due to the critical role, it plays in implementing city energy efficiency measures such as power generation control (Pedersen *et al.*, 2017), device fault identification and analysis (Li *et al.*, 2016), energy optimisation of buildings (Zhao and Magoulès, 2012), and smart building

evaluation and validation (Heo and Zavala, 2012), power consumption prediction is becoming increasingly important in improving efficiency in smart cities.

Engineering-based power generation simulation, which uses physical concepts to quantify the thermal equations and power consumption behaviours in smart cities (Wang *et al.*, 2018), is a common method to build power consumption prediction. Because of its ease and robustness, this technique aids users in constructing power efficient buildings in smart cities (Reeves *et al.*, 2012) and allows the use of hundreds of energy modeling tools around the world to be used in the smart city (Laroque *et al.*, 2012). Engineering-based power efficiency simulation, on the other hand, is frequently used during the design process because it examines all aspects of the buildings in smart cities.

Engineering-based power simulation in smart cities necessitates having the precise building details, such as geometrical parameters, component and part requirements, Heating, Ventilation, and Air Conditioning (HVAC) configuration files, and light system specifications, all of which are difficult to obtain from existing structures (Ryan and Sanquist, 2012). Furthermore, recent studies have revealed that this strategy occasionally leads to a significant difference between predicted and actual energy consumption in smart cities (Aydinalp *et al.*, 2002; Yalcintas and Ozturk, 2007).

Empirical simulation, which has been commonly used in forecasting power consumption for the past decade due to its dominance in experimental analysis and forecasting accuracy (Yu *et al.*, 2010; Ekici and Aksoy, 2009) is a much more appropriate technique for power consumption prediction. Empirical simulation employs machine learning (ML) algorithms like the Decision Tree

(DT) (Burkhart *et al.*, 2014), Artificial Neural Network (ANN) (Dong *et al.*, 2005), Gaussian process (Neto and Fiorelli, 2008), and Support Vector Machine (SVM) (Breiman, 1996) to extrapolate matching associations between input and output data. These ML methods are more effective than the engineering-based technique that has been applied to the power consumption prediction of smart cities. ML methods use data such as energy data, environmental data, and occupancy data for their predictions. This data can be collected from existing urban centres for use in smart cities.

Earlier studies have shown that empirical techniques such as ML algorithms when carefully selected and correctly used, can generate optimal predictive accuracy and surpass engineering-based power simulation approaches (Zekić-Sušac *et al.*, 2019; Tsanas and Xifara, 2012). However, some empirical simulation algorithms, such as DT and ANN are alleged to be unpredictable owing to their invariability problems (Candanedo *et al.*, 2017). Because of the instability of these ML algorithms, minor alterations in the input data can cause substantial deviations in the output (Candanedo *et al.*, 2017), causing the predicted outputs to vary significantly from the actual outputs. This will result in the abysmal performance of the prediction model. Since many energy-efficient measurements, such as system fault detection and energy standardization depend on the accuracy of smart city power consumption prediction outcomes, this restriction may prevent these algorithms from being used in the actual world. Furthermore, as smart cities' energy management technology advances, it now has an abundance of power consumption historical data.

To surmount the weaknesses associated with the existing ML algorithm in addition to enhancing prediction accuracy of power consumption in smart cities, this paper uses four the Bagging

(BAG), Stochastic Gradient Boosting (GBM), Model Averaged Neural Network (MAN), and K-Nearest Neighbours (KNN) ML models. These machine learning models generate prediction models that have high prediction accuracy.

After reviewing the relevant literature, it was discovered that the existing literature does not provide any work that uses BAG, GBM, MAN, and kNN ML techniques to solve the problem of power consumption prediction in a smart city. Moreover, the prediction accuracy of some of the existing work is rather low while some did not use state-of-the-art metrics to evaluate the performance of their work. Bearing in mind such inadequacy, the novelty of this work centers on the use of machine learning for power consumption prediction.

Related Works

This section presents a brief discussion on recent research done in the field of power or energy consumption prediction in smart cities. It has been explained in section one that ML algorithms have gained wide acceptance among data scientists and researchers for power consumption prediction. This is because of the efficacy of these algorithms in handling prediction problems. These ML models were chosen for use in this paper because they are highly nonlinear and can learn from past data. Zeki *et al.* (2019) demonstrated that power consumption in city areas does not conform to linearity hypotheses. BAG, GBM, MAN, and kNN ML models were used to create predictive models of power use in smart cities.

Tsanas and Xifara (2012) applied Random Forest (RF) algorithm to predict the power consumption of a building. They compared the performance of their proposed technique to that of the traditional linear regression method; the authors demonstrated that RF can provide more accurate heating and cooling load predictions. RF may also

be used to find an accurate practical association between input and output variables, according to the report. The downside of their approach is that the prediction accuracy is relatively low. RF was employed by Jurado *et al.* (2015) to predict the hourly power usage of multiple academic facilities. The study compares RF's prediction result to that of neural networks, fuzzy inductive reasoning (FIR), and the Auto-Regressive Integrated Moving Average (ARIMA). The findings demonstrated that for building energy prediction, FIR and RF outperform the other two approaches, and voting strategies can be used to merge various techniques and deliver more reliable predictions. The weakness of their work is the need to further reduce the prediction error value for the models used in the paper. Candanedo *et al.* (2017) utilised RF to estimate a building's short-term power use in Belgium. The output of RF was compared to that of three other prediction models: Multiple Linear Regression (MLR), Support Vector Regression (SVR), and GBM. The results revealed that RF and GBM outperformed MLR and SVR.

Ahmad *et al.* (2017) applied RF to estimate the hourly HVAC power usage in a hotel in Madrid, Spain. The authors contrasted the performance of RF and the commonly used ANN in terms of prediction. The results showed that both models are good at forecasting hourly HVAC power usage, and the RF model can be used to pick variables. However, there is a need to enhance the prediction accuracy of the proposed system. RF was used by Ma and Cheng (2016) to classify factors that influence the regional energy use intensity (EUI) of apartment blocks. To model the mean location EUI of residences in block groups, the authors included 171 influential features representing houses, economy, education, climate, households, neighborhood, and transportation. Premised on RF's out-of-bag (OOB) error, the 20

leading influential were recognised. As part of the effort towards the development of the smart city, Zekić-Sušac *et al.* (2021) addressed the question of how to fuse Big Data framework and ML into an intelligent framework for handling power consumption problems in the civil service. To construct a predictive model of power consumption of Croatian civil service facilities, deep neural networks, Rpart regression tree, and Random forest with variable reduction procedures were used. RF model generated the most accurate prediction result among the three approaches proposed in the paper. It was suggested that the models be incorporated into the MERIDA intelligent framework, which combines Big Data gathering and predictive models for managing power consumption for each power source in city areas. This will allow them to work effectively into a managing mechanism for enhancing power savings in a Big Data setting. The downside of their work is that it was just a framework and has not been experimented with to determine the performance of the proposed system.

Galicia *et al.* (2019) proposed an ensemble technique for predicting big data time series based on ML approaches. The decision tree, gradient boosted trees, and RF were integrated using the ensemble technique. For each ensemble component's predictions, they used a weighted mean combination, with varying weights for each algorithm calculated based on its earlier results. The proposed model demonstrated to be promising based on its performance on 10-year Spanish power usage data estimated with a 10-minute rate of recurrence. The limitation of this work is that the values of MAE and RMSE are relatively high. Torres *et al.* (2018) applied deep learning for data analysis power consumption in Spain. The goal variable was power consumption in Spain from 2007 to 2016, with a ten-minute interval sampling frequency. The researchers concluded that deep

learning, along with decision trees, is the best approach concerning scalability in handling big data time series.

Following a detailed analysis of recently published scholarly papers, it can be observed that power consumption prediction in smart cities is a hot subject that needs to be explored further. ML and deep learning algorithms for predicting power consumption in smart cities have been used in a very limited number of papers. Furthermore, there have not been many ML approaches for predicting power consumption. This paper investigates the efficiency of machine learning techniques for predicting power usage in smart cities to overcome these issues. The next section explains how we used machine learning models to predict power consumption in smart cities.

Methodology

Dataset Description

The smart city power consumption dataset was collected from the Kaggle database. The dataset used in this paper contains 2075259 rows and 9 columns. This dataset provides measurements of electric power consumption in an urban area with a one-minute sampling rate over a period of almost 4 years. Various electrical quantities and some sub-metering values were used. The data is made up of timestamps such as Date, Time, Global Active Power, Global Reactive Power, Voltage, Global Intensity, Sub Metering-1, Sub Metering-2, Sub Metering-3, and Minutes. Table 1 comprises of terms and meaning of the dataset of smart city power consumption.

Table 1: Comprehensive Dataset of Smart City Power Consumption

Terms	Meaning	Type of Data
Date	Day of the month or year as specified by a number.	Discrete
Time	A time point in hours and minutes at 12 noon or midnight.	Discrete
Global Active Power	Electricity used by electrical devices and measured in watts.	Continuous
Global Reactive Power	Is the electricity between the loads moving back and forth? This power does not carry out any useful load function.	Continuous
Voltage	Is the pressure from the energy source of an electric circuit which pushes charged electrons through a conducting circuit?	Continuous
Global Intensity	Is the ratio of primary energy supply to Gross Domestic Product (GDP)?	Continuous
Sub Metering 1-3	Energy monitors that are attached downstream from electric meters.	Discrete

Table 2: Statistical measures of the Dataset

Power Consumption	Descriptive Statistics				
	Minimum	1 st Quartile	Median	Mean	Maximum
Global Active Power	0.1940	0.4720	1.5940	1.8450	9.1320
Global Reactive Power	0.0000	0.0000	0.1000	0.1187	0.7240
Minutes	0.0000	15.0000	30.0000	29.53	59.00
Voltage	228.20	239.10	241.90	241.70	251.70
Global Intensity	0.8000	2.2000	6.8000	7.7930	39.4000
Sub metering1	0.0000	0.0000	0.0000	1.2600	77.0000
Sub metering2	0.0000	0.0000	0.0000	2.3370	74.0000
Sub metering3	0.0000	0.0000	0.0000	7.5190	20.0000

Machine Learning Algorithms

Machine learning is a branch in mathematics, statistics and computer science that studies the design of algorithms that can learn. Four different machine learning were utilized in this study such as Bagging (BAG), K-Nearest Neighbor (KNN), Stochastic Gradient Boosting (GBM) and Moving Average Neural Network (MAN).

Bagging (BAG)

Bagging is an ensemble ML technique for integrating similar kinds of predictions. Bagging addresses the problems of over-fitting, reduces discrepancy, and unfairness in a model (Dada *et al.*, 2019). It is a uniformly weak learner with the capacity to learn from each other autonomously in parallel and combines them for determining the model average (Lemmens and Croux, 2006). In BAG, different learners' decisions can be merged into a single prediction. Voting is simply the process of integrating certain decisions in the case of classification. Both bagging and boosting employ this strategy. Individual models, on the other hand, are created by bagging and boosting in various ways. In bagging, the models are given the same weights, while in boosting, more efficient models are given more weighting because an executive may base alternative results on a variety

of experts' advice based on their previous accurate estimations. The experts are independent decision trees who are brought together by voting on each test. It is regarded right if one class receives more votes than the others. As there are more voters, predictions made by a larger number of votes are more accurate (Oyewola *et al.*, 2019). The following steps were utilized in the implementation of Bagging:

1. Take bootstrapped samples from the Kaggle database smart city electricity consumption dataset.
2. For each bootstrapped sample, create a decision tree.
3. Create a final model by averaging the forecasts of each tree.
4. Output the result of the predicted values.

K-Nearest Neighbor (KNN)

The kNN is a classic example of a promising classification algorithm based on the principle of nearest learning examples in the feature space for solving classification problems (Oniyide *et al.*, 2020; Dada *et al.*, 2021). Because it generalizes the data after a query, the kNN algorithm is known as the "lazy learning" algorithm. After the

classification process, any calculation involved in the classification process takes place. Despite the need for a training dataset, no definite learning or model formation occurs during the training stage (Dada *et al.*, 2017). The dataset is used by the kNN to fill a sample of the search space with only instances of a known class. For this reason, this algorithm is referred to as a "lazy learning" algorithm (Saini *et al.*, 2013). Since the kNN algorithm is a slow learner, it does not have a training stage, and when it does, it completes it in a short amount of time. The testing stage, on the other hand, is time and memory intensive (Thirumuruganathan, 2010). The following steps were utilized in the implementation of K-Nearest Neighbor:

1. Decide on the K neighbor number.
2. Select the K Nearest Neighbor samples from the Kaggle database smart city power consumption dataset.
3. Count how many data points each category has among the K-Neighbours.
4. Assign the new data point to the category in which the most neighbours were counted.
5. Output the result of the predicted values.

Stochastic Gradient Boosting (GBM)

Breiman (1996) pioneered the idea of introducing arbitrariness into function estimation procedures to increase their accuracy. Random sampling was used in early AdaBoost developments (Freund and Schapire, 1996), but it was regarded as an approximation to deterministic weighting when the base learner execution did not endorse observation weights, instead of an integral part. Breiman (1999) lately presented a hybrid bagging

boosting technique (referred to as "adaptive bagging") for least squares appropriate incremental extensions. It swaps "out of bag" residuals for ordinary residuals in every boosting phase, replacing the base learner in normal boosting procedures with the equivalent bagged base learner. Gradient boosting was modified slightly to provide arbitrariness like an essential component of the technique, as inspired by Breiman (1999). Principally, a sample population of the training data is haphazardly selected (without substitution) from the complete training data collection for each iteration. This arbitrarily picked sample population is then used to match the base learner and calculate the model adjustment for the present iteration, rather than the entire sample. The following steps were utilized in the implementation of Stochastic Gradient Boosting:

1. Subsample rows of samples from the smart city power consumption dataset which was collected from the Kaggle database before creating each tree.
2. Subsample columns of samples from the smart city power consumption dataset which was collected from the Kaggle database before creating each tree.
3. Subsample columns of samples from the smart city power consumption dataset which was collected from the Kaggle database before considering each split in each tree.
4. Output the result of the predicted values.

Model Averaged Neural Network (MAN)

Model Averaged Neural Network (MAN) is a model where the same neural network model is fitted using different random number seeds. All the resulting models are used for prediction. In

regression, the output from each network are averaged while in classification, the model scores are first averaged, then translated to predicted classes (Ripley, 1996). The following steps were utilized in the implementation of Model Averaged Neural Network:

1. Combining Neural Networks obtained from prediction of data collected from Kaggle database.
2. Averaging their parameters instead of combining outputs based on knowledge of theoretical asymptotic weight distribution.
3. Output the result of the predicted values.

Proposed System Architecture

This paper used four machine learning models for power consumption prediction as depicted in Figure 1. The proposed architecture uses time-series data preprocessed to extract spatial features using different machine learning models. The spatial features refer to the shape, size and location of the features such as Date, Time, Global Active Power, Global Reactive Power, Voltage, Global Intensity, Sub Metering-1, Sub Metering-2, and Sub Metering-3. The dataset used in this work contains 40,000 data of household power consumed in an urban area. The dataset was partitioned into a training set (80%) and a test set (20%). Afterwards, selected features of the data were modeled using BAG, GBM, MAN, and kNN algorithms.

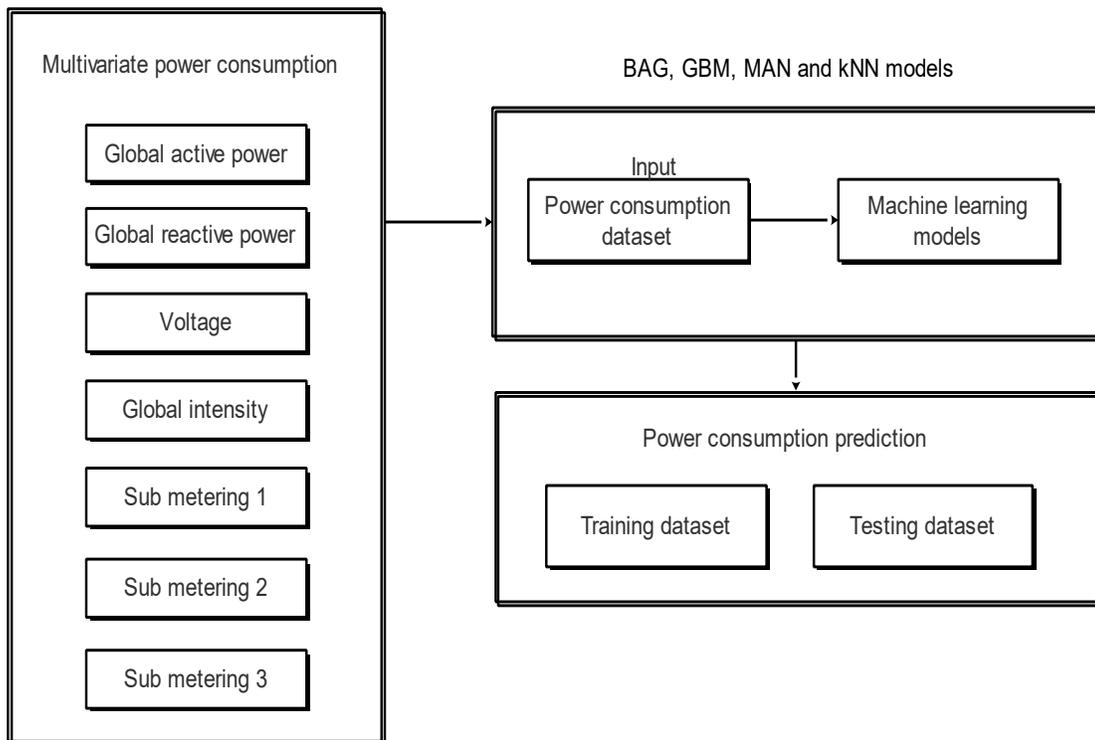


Figure 1: Proposed power consumption prediction architecture

Performance Measurements

The prediction performance of the proposed system was evaluated using three metrics: Mean Square Error (MSE), Root Mean Square Error (RMSE), and Mean Absolute Scaled Error (MASE).

MSE is defined as the average of the squares of the error difference between predicted and actual values in the test.

$$MSE = \frac{1}{n} \sum_{n=1}^n (I_n - \hat{I}_n)^2 \quad (1)$$

RMSE is defined as the standard deviation of prediction errors in a test.

$$RMSE = \sqrt{\frac{1}{n} \sum_{n=1}^n (y_t^n - \hat{y}_t^n)^2} \quad (2)$$

MASE is defined as a measure of the accuracy of predictions that represents it as a percentage in comparison to a standard mean error.

Figure 2. The voltage is measured in volts between 230-250 volts.

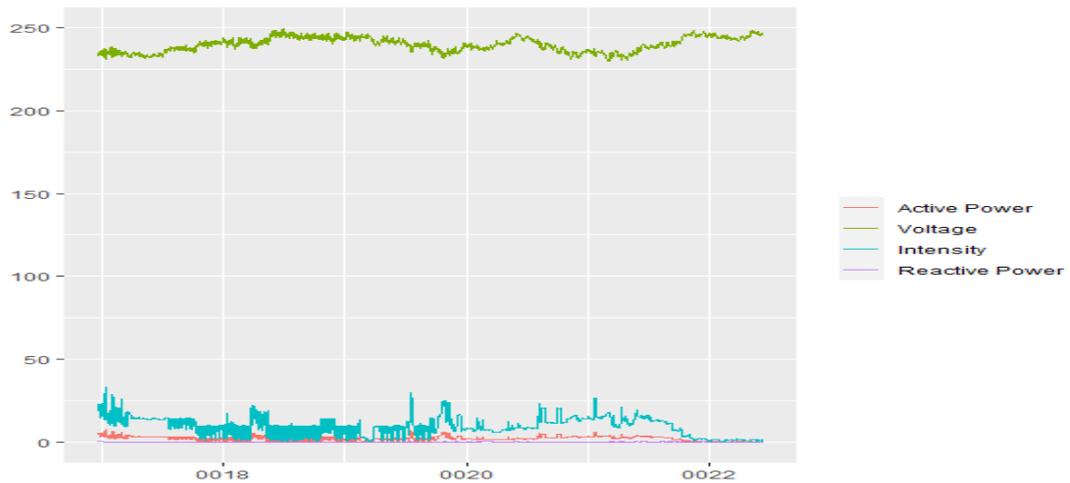


Figure 2: Smart City Power Consumption of Active Power, Voltage, Global Intensity and Global Reactive Power

$$MASE = \frac{1}{n} \sum_{n=1}^n \left(\frac{|y_t^n - \hat{y}_t^n|}{\frac{1}{n-m} \sum_{n=m+1}^n |y_t^n - y_{t-m}^n|} \right) \quad (3)$$

Results and Discussion

Figure Statistical results of the Smart City Power Consumption are presented in Table 2. The Minimum, 1st Quartile, Median, Mean and Maximum were also analyzed. Voltage has the highest value while sub-metering 1 has the lowest value followed by sub-metering 2. The Global Active Power is within the range of 0-9 while Global Reactive Power is within the range of 0-0.8. The consumption of current varies between 0-59 minutes. This indicates that smart city consumes current daily. Smart City Power Consumption of Active Power, Voltage, Global Intensity and Global Reactive Power is shown in

In Figure 2 high global intensity leads to an increase in global active power consumed daily. Also, the amount of electricity consumed at a time leads to an increase in Global Active Power. This means that usage of electricity can lead to an

increase of Global Active Power. The performance of the actual global active power and predicted active power is measured using three performance metrics: RMSE, MSE and MASE.

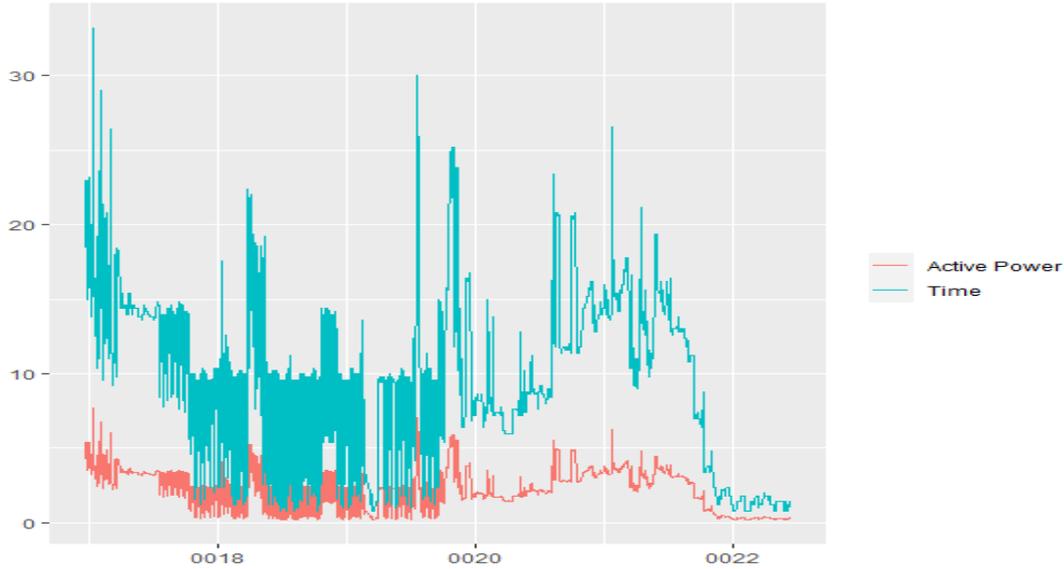


Figure 3: Smart City Power Consumption of Global Active Power and amount of electricity consumed referred to as time.

Figure 3 is the power consumption and amount of electricity consumed per hour. It represents the power consumption cycle throughout the day. The Figure displays the rise and fall of power consumed daily. This variation in power

consumed follows the daily (time) patterns of energy use by households and businesses. This shows that as more time is spent using the electricity by turning on the lights, use hot water, and others, it affects the daily power consumption.

Table 3: Performance Evaluation of BAG, GBM, MAN and KNN

MODEL	RMSE	MSE	MASE
BAG	0.2554	0.0652	0.6643
GBM	0.0747	0.0055	0.1769
MAN	1.6444	2.7041	4.3200
KNN	0.1907	0.0363	0.3199

In Table 3, MAN and BAG performed poorly while GBM perform excellently. This shows that GBM has better predicting power than KNN in

predicting Global active power. In all the performance evaluation used in this paper, GBM has a very low-value which is an indication of

good performance. In Figure 4 BAG and MAN perform poorly while GBM performs excellently well in all the algorithms considered in this paper.

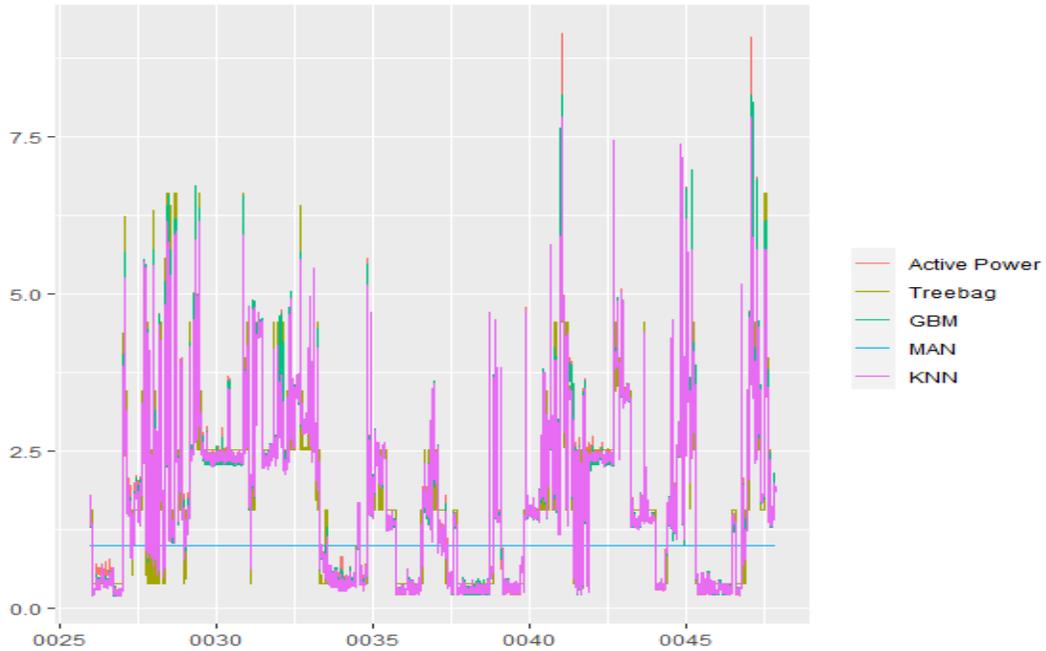


Figure 4: Smart City Power Consumption of Actual (Active Power) and Predicted (Treebag, GBM, MAN and KNN).

Figure 4 depicts a graph that visualizes the predicted results using the Treebag (BAG), GBM, MAN and KNN models. It was confirmed that power consumption prediction in the power consumption dataset was very fruitful. Results of prediction show that GBM performs excellently well when compared with the real value. In addition, brilliant prediction results are attained even in circumstances where the periodicity is undetected. Visual confirmation proved that the GBM models perform excellently well in power consumption prediction. Moreover, it also substantiates the fact that GBM models predict local features well in power consumption prediction.

Conclusion

Machine learning models for predicting power consumption in a smart city have been proposed in this paper. The Bagging (BAG), Stochastic Gradient Boosting (GBM), Model Averaged Neural Network (MAN), and K-Nearest Neighbours (KNN) Machine Learning models were applied for power consumption prediction in a smart city. Simulation results demonstrate the supremacy of GBM over other models applied for power consumption prediction in this study. The findings of this paper show that GMB is a promising algorithm that can be adopted for empirical modeling-based power usage prediction in smart cities. In the future, our work will focus on integrating GBM with Softcomputing

algorithms such as Moth Flame Optimization (MFO) algorithm, Harris Hawks Optimization (HHO), League Championship Algorithm (LCA), Firefly algorithm (FA), and Dragonfly algorithm (DA) to improve the prediction accuracy of GBM and extending the applications of the proposed hybrid model to the power consumption prediction of different types of buildings (school, offices, hospitals, train station) in a smart city. The summary of our contribution is as follows:

- i. This paper used four (4) machine-learning techniques for power consumption prediction in a smart city. Machine learning models have the capacity to reduce variability between training and test datasets. It also reduces bias during the model's training and testing phases. As a result, the proposed system solves the limitations associated with the existing

techniques used for the prediction of power consumption.

- ii. The BAG, GBM, MAN and kNN ML models for solving the problem of power consumption prediction in the smart city was presented in this paper.
- iii. The proposed models were evaluated using different performance metrics such as Root Mean Square Error (RMSE), Mean Average Square Error (MASE), and Mean Square Error (MASE) for the GBM model was attained.

Conflict of Interest

The authors declare no conflict of interest.

References

- Ahmad, M. W., Mourshed, M., & Rezgui, Y. (2017). Trees vs Neurons: Comparison between random forest and ANN for high-resolution prediction of building energy consumption. *Energy and buildings*, 147, 77-89.
- Aydinalp, M., Ugursal, V. I., & Fung, A. S. (2002). Modeling of the appliance, lighting, and space-cooling energy consumptions in the residential sector using neural networks. *Applied energy*, 71(2), 87-110.
- Breiman, L. (1996). Heuristics of instability and stabilization in model selection. *The annals of statistics*, 24(6), 2350-2383.
- Breiman, L., 1996. Bagging predictors. *Mach. Learning* 26, pp. 123-140.
- Breiman, L. (1999). *Using adaptive bagging to debias regressions* (p. 16). Technical Report 547, Statistics Dept. UCB.
- Burkhardt, M. C., Heo, Y., & Zavala, V. M. (2014). Measurement and verification of building systems under uncertain data: A Gaussian process modeling approach. *Energy and Buildings*, 75, 189-198.
- Candanedo, L. M., Feldheim, V., & Deramaix, D. (2017). Data driven prediction models of energy use of appliances in a low-energy house. *Energy and buildings*, 140, 81-97.
- Dada, E. G., Bassi, J. S., & Adekunle, O. O. (2017). An Investigation into the Effectiveness of Machine Learning Techniques for Intrusion Detection. *Arid Zone Journal of Engineering, Technology and Environment*, 13(6), 764.
- Dada, E. G., Bassi, J. S., Chiroma, H., Adetunmbi, A. O., & Ajibuwa, O. E. (2019). Machine learning for email spam filtering: review, approaches and open research problems. *Heliyon*, 5(6), e01802.
- Dada, E. G., Oyewola, D. O., Joseph, S. B., & Dauda, A. B. (2021). Ensemble Machine Learning Model for Software Defect Prediction, *Advances in Machine Learning & Artificial Intelligence* 2 (1), 11-21

- Dong, B., Cao, C., & Lee, S. E. (2005). Applying support vector machines to predict building energy consumption in tropical region. *Energy and Buildings*, 37(5), 545-553.
- Ekici, B. B., & Aksoy, U. T. (2009). Prediction of building energy consumption by using artificial neural networks. *Advances in Engineering Software*, 40(5), 356-362.
- Freund, Y. and Schapire, R. (1996). Experiments with a new boosting algorithm. *Machine Learning: Proceedings of the 13th International Conference*, pp. 148–156.
- Galicia, A., Talavera-Llames, R., Troncoso, A., Koprinska, I. and Martínez-Álvarez, F. (2019). Multi-step forecasting for big data time series based on ensemble learning. *Knowledge-Based Systems*, 163, 830–841.
- Heo, Y. and Zavala, V.M. (2012). Gaussian process modeling for measurement and verification of building energy savings. *Energy and Buildings*, 53, pp. 7-18.
- International Energy Agency, (2015) Key World Energy Statistics," International Energy Agency.
- Jurado, S., Nebot, À., Mugica, F. and Avellana, N. (2015). Hybrid methodologies for electricity load forecasting: Entropy-based feature selection with machine learning and soft computing techniques, *Energy*, vol. 86, pp. 276–291.
- Laroque, C., Himmelspach, J., Pasupathy, R., Rose, O. and Uhrmacher, A.M. (2012). Proceedings of the 2012 Winter Simulation Conference (WSC 2012), December 9-12, Berlin, Germany.
- Lemmens, A. and Croux, C. (2006) Bagging and boosting classification trees to predict churn. *Journal of Marketing Research*, 43(2), pp. 276-286.
- Li, D., Hu, G. and Spanos, C. J. (2016) A data-driven strategy for detection and diagnosis of building chiller faults using linear discriminant analysis. *Energy and Buildings*, 128, pp. 519-529.
- Ma, J. and Cheng, J. C. (2016) Identifying the influential features on the regional energy use intensity of residential buildings based on Random Forests," *Applied Energy*, vol. 183, pp. 193-201.
- Neto, A.H. and Fiorelli, F.A.S. (2008) Comparison between detailed model simulation and artificial neural network for forecasting building energy consumption. *Energy and buildings*, 40(12), pp. 2169-2176.
- Oniyide, M. K., Kolawole, A. O., & Dada, E. G. (2020). Performance Comparison Of Computer-Based Test And Paper-Pencil Test Using Support Vector Machine. *FUDMA Journal Of Sciences-ISSN: 2616-1370*, 3(4), 478-487.
- Oyewola, D. O., Dada, E. G., Omotehinwa, T. O. and Ibrahim, A. I. (2019). Comparative Analysis of Linear, Non Linear and Ensemble Machine Learning Algorithms for Credit Worthiness of Consumers. *Computational Intelligence & Wireless Sensor Networks*, vol. 1 (1) pp. 1-11
- Pedersen, T. H., Hedegaard, R. E., & Petersen, S. (2017). Space heating demand response potential of retrofitted residential apartment blocks. *Energy and Buildings*, 141, 158-166.
- Reeves, T., Olbina, S., & Issa, R. (2012, December). Validation of building energy modeling tools: ecotect™, green building studio™ and IES< VE>™. In *Proceedings of the 2012 Winter Simulation Conference (WSC)* (pp. 1-12). IEEE.
- Ripley, B. D. (1996). *Pattern Recognition Neural Networks*. Cambridge.
- Ryan, E. M., & Sanquist, T. F. (2012). Validation of building energy modeling tools under idealized and realistic conditions. *Energy and buildings*, 47, 375-382.
- Saini, I., Singh, D., & Khosla, A. (2013). QRS detection using K-Nearest Neighbor algorithm (KNN) and evaluation on standard ECG databases. *Journal of advanced research*, 4(4), 331-344.
- Sun, Y., Song, H., Jara, A. J., & Bie, R. (2016). Internet of things and big data analytics for smart and connected communities. *IEEE access*, 4, 766-773.

- Thirumuruganathan, S. (2010) A detailed introduction to K-nearest neighbor (KNN) algorithm. Retrieved on May 15th, 2020.
- Torres, J. F., Galicia, A., Troncoso, A. and Martínez-Álvarez, F. (2018) A scalable approach based on deep learning for big data time series forecasting. *Integrated Computer-Aided Engineering*, 25(5), 1–14.
- Tsanas, A. and Xifara, A. (2012) Accurate quantitative estimation of energy performance of residential buildings using statistical machine learning tools. *Energy and Buildings*, 49, pp. 560-567.
- Wang, Z., Wang, Y., Zeng, R., Srinivasan, R.S. and Ahrentzen, S. (2018) Random Forest based hourly building energy prediction. *Energy and Buildings*, 171, pp. 11-25.
- Yalcintas, M. and Aytun Ozturk, U. (2007) An energy benchmarking model based on artificial neural network method utilizing US Commercial Buildings Energy Consumption Survey (CBECS) database. *International Journal of Energy Research*, 31(4), pp. 412-421.
- Yu, Z., Haghghat, F., Fung, B.C. and Yoshino, H. (2010) A decision tree method for building energy demand modeling. *Energy and Buildings*, 42(10), pp. 1637-1646.
- Zekić-Sušac, M., Knežević, M. and Scitovski, R. (2019, May) Deep learning in modeling energy cost of buildings in the public sector. In *International Workshop on Soft Computing Models in Industrial and Environmental Applications* (pp. 101-110). Springer, Cham.
- Zekić-Sušac, M., Mitrović, S. and Has, A. (2021) Machine learning based system for managing energy efficiency of public sector as an approach towards smart cities. *International journal of information management*, 58, 102074.
- Zhao, H.X. and Magoulès, F. (2012) A review on the prediction of building energy consumption. *Renewable and Sustainable Energy Reviews*, 16(6), pp. 3586-3592.