Arabian Journal for Science and Engineering (2021) 46:3155–3168
https://doi.org/10.1007/s13369-020-05069-2

RESEARCH ARTICLE-COMPUTER ENGINEERING AND COMPUTER SCIENCE

A Systematic Analysis for Energy Performance Predictions in Residential Buildings Using Ensemble Learning

Monika Goyal[1] · **Mrinal Pandey**[1]

Received: 10 June 2020 / Accepted: 23 October 2020 / Published online: 20 November 2020

Abstract
Energy being a precious resource needs to be mindfully utilized, so that efficiency is achieved and its wastage is curbed. Globally, multi-storeyed buildings are the biggest energy consumers. A large portion of energy within a building is consumed to maintain the desired temperature for the comfort of occupants. For this purpose, heating load and cooling load requirements of the building need to be met. These requirements should be minimized to reduce energy consumption and optimize energy usage. Some characteristics of buildings greatly affect the heating load and cooling load requirements. This paper presented a systematic approach for analysing various factors of a building playing a vital role in energy consumption, followed by the algorithmic approaches of traditional machine learning and modern ensemble learning for energy consumption prediction in residential buildings. The results revealed that ensemble techniques outperform machine learning techniques with an appreciable margin. The accuracy of predicting heating load and cooling load, respectively, with multiple linear regression was 88.59% and 85.26%, with support vector regression was 82.38% and 89.32%, with K-nearest neighbours was 91.91% and 94.47%. The accuracy achieved with ensemble techniques was comparatively better—99.74% and 94.79% with random forests, 99.73% and 96.22% with gradient boosting machines, 99.75% and 95.94% with extreme gradient boosting.

Keywords Machine learning · Energy optimization · Random forests · Multiple linear regression · Gradient boosting machines · Extreme gradient boosting · K-nearest neighbours · Support vector regression · Feature selection

1 Introduction

While designing smart buildings, optimal measures should be taken so that energy is used efficiently to safeguard the environment [1]. Studies done by researchers all over the world show that the highest percentage of energy is consumed by the multi-storey buildings [2–5]. Buildings consume about 40% of the total energy consumed in the world. After buildings, the second major energy consumer is industry which is reported for 32% energy consumption. The third major area is transport with 28% energy consumption. These studies motivated to devise solutions for energy optimization in buildings. Further studies show that within buildings, heating, ventilation and air conditioning (HVAC) system is one of the major energy consumers [6, 7]. HVAC consumes energy to maintain the desired temperature within a building and control humidity. It is responsible for meeting the heating load and cooling load of a building. Heating load can be defined as the amount of heat energy that is required to be added to a certain space for maintaining a desired temperature. Cooling load, on the other hand, is the amount of heat energy to be removed from a certain space to keep the temperature within desired limits. These two are related to the thermal load of the building. When the building is cold, the thermal load is converted into heating load and when the building is hot, the thermal load is converted into cooling load [8]. The heating and cooling loads of a building directly affect its energy performance. It requires analysis of factors that affect the heating and cooling loads. Studies reveal that various characteristics of a building and its structure affect heating and cooling loads to a major extent [9, 10]. Predicting energy consumption in buildings gives insight on the future demand of energy, and if more energy is being consumed than expected, appropriate measures can be adopted to stabilize energy use.

✉ Mrinal Pandey
mrinalpandey14@gmail.com

Monika Goyal
monikagoyal.er@gmail.com

[1] Computer Science and Technology, Manav Rachna University (Formerly, Manav Rachna College of Engineering), Faridabad, India

This paper focusses on several important features of buildings for, e.g. relative compactness, surface area, wall area, roof area, overall height, orientation, glazing area and glazing area distribution. The feature selections techniques were employed to derive the relevant features for predicting the heating load and cooling load. Further, three machine learning algorithms namely—multiple linear regression, K-nearest neighbours and support vector regression, and three ensemble learning algorithms namely random forests, gradient boosting machines, extreme gradient boosting were used for creating models. Additionally, these models have been evaluated using various performance measures for, e.g. RMSE, MSE, MAE, *R* squared and accuracy.

The organization structure of the paper is as follows: Sect. 1 introduces the problem.

Literature review is presented in Sect. 2. Section 3 describes the state-of-the-art machine learning techniques and ensemble techniques. The methodology adopted for the work is described in Sect. 4. The experiments performed and the results obtained are presented in Sect. 5. Finally, conclusions revealed in this research are mentioned in Sect. 6.

2 Literature Survey

Several researchers worldwide worked in the field of energy consumption and optimization in buildings. An important aspect in this domain is the analysis of energy performance gap. In this context, a study on German households [11] introduces prebound effect, in which the occupants actually consume 30% less energy than as calculated in the standard ratings. This gap can be due to incorrect assumptions made during energy ratings. In another study [12], common usages of the term rebound effect have been reviewed. Rebound effect is used in the context where actual energy consumption exceeds the calculated ratings. Some of the important researches in the field of energy using various individual machine learning and ensemble techniques are analysed and shown in Table 1.

3 Machine Learning

In this research, the applied machine learning techniques belong to two different categories.

3.1 Traditional Machine Learning Techniques

Three traditional machine learning techniques have been applied in our research work namely—multiple linear regression, K-nearest neighbours and support vector machines.

3.1.1 Multiple Linear Regression

It is although quite similar to linear regression but there is one significant difference. In MLR model, one response variable B is dependent upon multiple independent variables A1, A2, A3 … An. The relationship between predictor variable and response variable can be expressed in the form of conditional expectation as shown in Eq. (1).

$$E(Y|X) = \beta_0 + \beta_i X_i \tag{1}$$

β_i is the slope that depicts the change in response variable Y when the predictor variable j is varied by one unit and other predictors are kept constant. The complexity of results interpretation increases in this model as a result of the correlation between different independent variables [5]. The concept of MLR is graphically shown in Fig. 1 [35].

3.1.2 K-Nearest Neighbours

Also known as "Lazy Learner", this technique takes into consideration "k" number of closest instances in the training dataset to predict the value of the unknown instance. These k instances are found by applying a certain distance metric such that they are the k-nearest neighbours of the unknown instance [36]. The value returned is obtained after averaging the values of k-nearest neighbours [37].

If D is a dataset consisting of x_i training instances and the value of an unknown instance p is to be predicted, the distance between p and x_i can be obtained with Eq. (2).

$$d(p, x_i) = \sum_{f \in F} w_f \delta\left(p_f, x_{if}\right) \tag{2}$$

where $\delta\left(p_f, x_{if}\right) = \left|p_f - x_{if}\right|$ for continuous attribute.

Graphical representation of KNN regression can be seen in Fig. 2 [38]. In our experiments, the value of k is taken as 4.

3.1.3 Support Vector Regression

It aims at finding a function $f(x)$ which allows deviation to a certain extent ε from the obtained target values y_i in training data samples. It should also ensure maximum flatness. A linear function [39] can be described as:

$$f(x) = w,\ x + b \quad \text{with} \quad w \in \mathcal{X},\quad b \in \mathbb{R} \tag{3}$$

where $\mathcal{X}$ represents input pattern space such that $\mathcal{X} = \mathbb{R}^{\mathrm{d}}$.

Figure 3 [40] shows the graphical representation of SVR. The legend in the figure represents the results of various SVR kernel functions applied on a sample dataset of 40 random numbers. Values on x-axis represent data points, values on

Table 1 Survey of literature

References	ML model used	Work done	Results
Tsanas and Xifara [13]	Iteratively regressive least squares, random forests	A ML framework developed to analyse the effect of different building parameters on heating load and cooling load	Results of random forests were better at revealing relationships between input and output variables
Fan et al. [14]	MLR, ARIMA, SVR, BT, RF, MLP, MARS, KNN	Ensemble models developed to predict next-day energy demand in buildings	Prediction accuracy of Ensemble models higher than individual models as evaluated by MAPE
Jain et al. [2]	Support vector regression	Developed a sensor-based model for forecasting energy consumption in multi-family residential buildings	Spatial granularity impact the prediction power of the model significantly
Wei et al. [15]	MLP ensemble	Model developed for HVAC energy optimization	Energy savings were more when internal air quality was taken into consideration
Park et al. [16]	Decision tree	Developed a new energy benchmark to improve the operational rating system of office buildings	Proposed benchmark better than conventional and baseline system
Candanedo et al. [17]	MLR, SVM, RF, GBM	Model developed for predicting energy usage by appliances in residential building	GBM outperformed other models; atmospheric pressure is an important predictor
Manjarres et al. [18]	Random forest	A framework developed to optimize HVAC energy consumption	Energy consumption reduced by 48% for heating and 39% for cooling
Peng et al. [19]	K-nearest neighbour	A model developed to optimize energy consumption in building space according to occupancy	Energy savings of 7–52% obtained
Gallagher et al. [20]	LSR, DT, KNN, ANN, SVM	ML algorithms used for measurement and verification of energy saved in industrial buildings	Error reduced by 51.09%
Deb et al. [21]	MLR, ANN	Prediction models developed for energy savings in HVAC in office buildings	ANN more accurate with MAPE of 14.8%
Nayak [22]	ARIMA, RBFNN, MLP, SVM, FLANN	Developed a new model which linearly combined the five ML models for better accuracy	Model developed was better in terms of feasibility and performance
Sethi and Mittal [23]	DT, Naïve Bayes, SVM, RF, LR, stacking ensemble, voting ensemble	ML techniques applied to predict accurate air quality index	Ensemble techniques outperformed others
Pallonetto et al. [24]	M5P regression algorithm	Demand response algorithms deployed for controlling an integrated heat pump and thermal storage system in Residential buildings	49% reduction in cost, 39% reduction in carbon footprint
Pham et al. [25]	RF, M5P, RT	ML algorithms applied on five building energy consumption datasets to predict the short-term energy consumption in buildings	RF was better at prediction accuracy as compared to other investigated algorithms. 49.21% better than RT and 49.95% better than M5P in terms of MAE
Walker et al. [26]	Boosted tree, RF, SVM, ANN	Algorithms employed on dataset obtained from 47 buildings–both at individual building level and aggregated level to predict the energy demand at hourly intervals	RF, boosted tree and ANN performed better than SVM when computation time and error metrics were considered
Xu et al. [27]	ANN	Social network analysis was integrated with ANN to predict energy use in a group of 17 buildings	90.28% accuracy achieved with the proposed method

Table 1 continued

References	ML model used	Work done	Results
Zhou et al. [28]	MLP	Two optimization techniques: artificial bee colony and particle swarm optimization were combined with mlp to predict heating load and cooling load of residential buildings	Coefficient of determination increased and MAE and RMSE decreased significantly when MLP combined with optimization techniques
Gao et al. [29]	EN, GPR, LMSR, MLR, MPR, MLP, RBFR, SMOR, functions XNV, Lazy K star, Lazy LWL, RDT, M5 Rules, AMT, DPC, RF	A correlation-based feature subset selection technique was applied to the original dataset consisting of 8 parameters that reduced it to 4 parameters. Then all ML algorithms were applied on the dataset	RF, Lazy K star, RDT, AMT outperformed other ML techniques in terms of reduction in RMSE and MAE
Seyedzadeh et al. [30]	ANN, SVM, GP, RF, GBRT	Sensitivity analysis performed to determine the importance of each feature. ML models applied on two datasets for predicting heating and cooling loads	40% improved RMSE than results obtained in previous studies
Roy et al. [31]	DNN, GBM, GPR, MPMR	Models compared for performance in predicting heating and cooling loads of residential buildings	Results obtained are 99.76% by DNN and 99.84% by GPR in terms of VAF
Iruela et al. [32]	ANN	A GPU-based parallel implementation of NSGA-II to train the ML model for predicting energy consumption in buildings	Improved computation time and errors
Das et al. [33]	Different types of ANN: Elman neural network, recurrent neural network and backpropagation network	The effect of various building parameters on heating load and cooling load of the building is studied using different algorithms	Backpropagation neural network was most accurate among all, with MAE 0.1 for HL and 0.1254 for CL prediction
Cozza et al. [34]	LASSO regression	Assessed the capability of existing energy certificates in calculating actual energy consumption and savings to be achieved post building retrofitting	An average of − 23% negative energy performance gap and 2% positive energy gap was found before and after building retrofit respectively

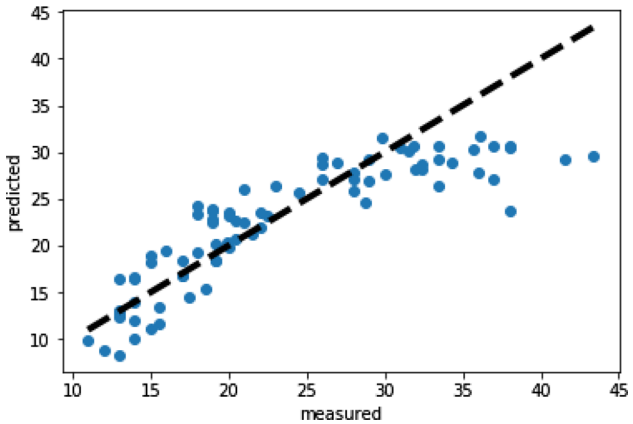

Fig. 1 Multiple linear regression

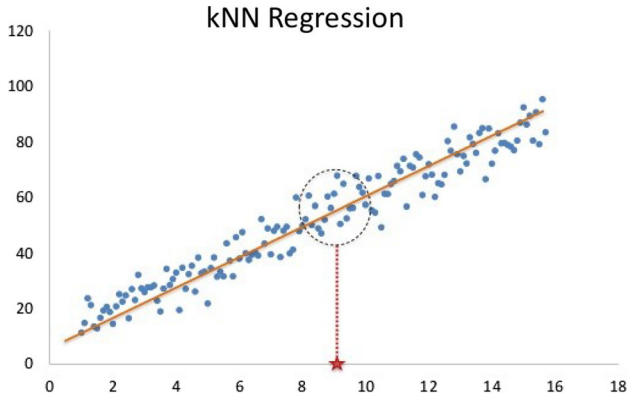

Fig. 2 K-nearest neighbours [38]

y-axis are target points. A radial basis function (RBF) kernel can be described as:

$$K(X_1, X_2) = \exp\left(-\gamma ||X_1 - X_2||^2\right) \quad (4)$$

where $||X_1 - X_2||$ is the Euclidean distance between points X_1 and X_2.

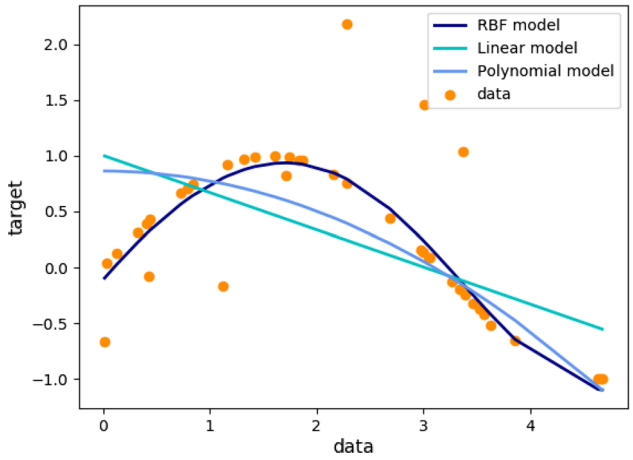

Fig. 3 Support vector regression [40]

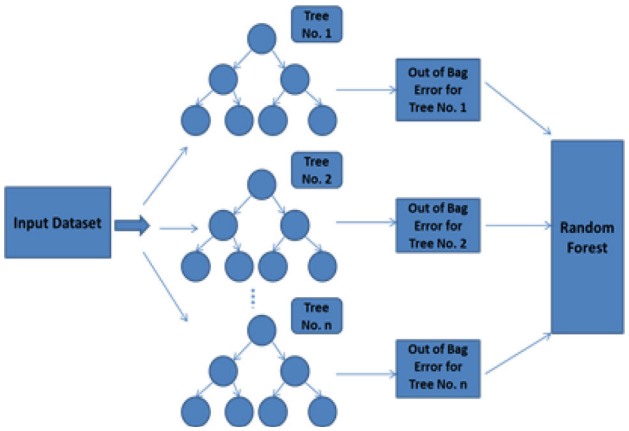

Fig. 4 Random forests

3.2 Ensemble Techniques

The basic Ensemble technique is to integrate the results of individual machine learning models, such that the prediction results exhibit improvement in terms of accuracy and robustness. Bagging and Boosting are two popular ensemble methods. The ensemble techniques used in this paper are explained as follows.

3.2.1 Random Forests

It is a tree-based ensemble technique that can be applied for both classification as well as regression. Some of the features which make random forests appealing are: prediction efficiency, suitability for highly multi-dimensional problems, missing values handling, outlier removal, etc. [41, 42].

In regression using random forests, to predict a continuous variable, the trees are grown depending on θ in such a manner that $h\ (x, \theta)$ takes on numeric values.

Where θ : A random vector

$h(x,\ \theta)$: Tree predictor

The values of the response variable are numeric, and it is assumed that the training sample is drawn independently from the distribution X of random vector Y.

The RF predictor is created by taking the mean over k of the trees

$$\{h(x, \theta_k)\} \tag{5}$$

The mean square generalization error for a numeric predictor $h(x)$ is given by

$$E_{X,Y}(Y - h(X))^2 \tag{6}$$

For infinite numbers of trees in the forest, the RF predictor is defined as

$$E_{X,Y}(Y - av_k h(X, \theta_k))^2 \rightarrow E_{X,Y}(Y - E_\theta h(X, \theta))^2 \tag{7}$$

The schematic diagram of Random Forests is shown in Fig. 4.

3.2.2 Gradient Boosting Machines

GBM is also an ensemble learning technique, whose underlying structure is a decision tree. In GBM, additive regression models are created by iteratively fitting a simple base to currently updated pseudo residuals by calculating least squares at every continuous iteration [43]. In gradient boosting a function $F^*(x)$ is generated that maps x to y, so that when the joint distribution of all (y, x) values is taken, the expected value of $\Psi(y, F(x))$ is minimized, where $\Psi(y, F(x))$ is some specified loss function. This relation is depicted in Eq. (8).

$$F^*(x) = \arg\min E_{y,x}\Psi(y,\ F(x)) \tag{8}$$

where y: The random output or response variable, $x = \{x_1, x_2, \ldots x_n\}$: a set of random input variables.

Figure 5 shows the scheme behind gradient boosting machines.

3.2.3 Extreme Gradient Boosting

Apart from performance and speed as its key features, this technique has an added feature of Scalability. Several optimizations have been performed on the basic algorithm to ensure the scalability of the model [44]. Figure 6 shows the schematic diagram of XGBoost.

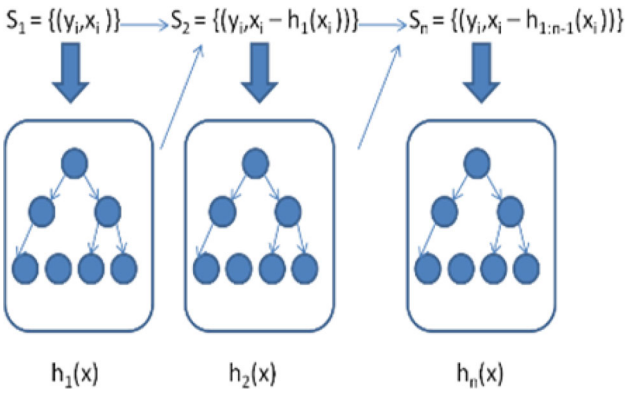

Fig. 5 Gradient boosting machines

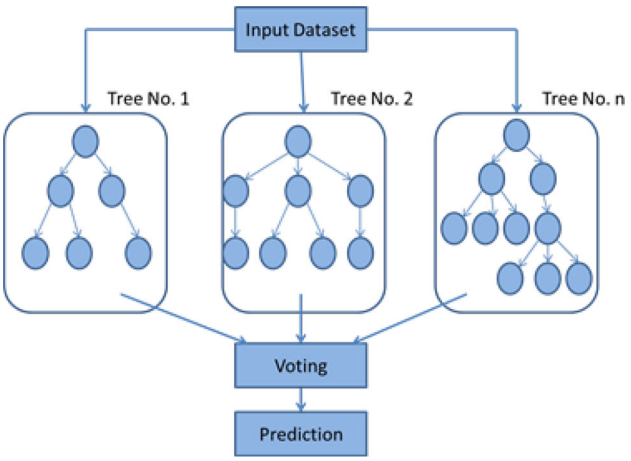

Fig. 6 Extreme gradient boosting

4 Method and Data

This section explains the workflow approach followed in this research. Figure 7 shows the steps of the methodology employed. The methodology starts with the data collections followed by data analysis and pre-processing, data partitioning and model constructions using various machine learning and ensemble learning algorithms. Finally, models have been evaluated on various parameters. Each phase in the process has been explained below.

4.1 Data Set Collection and Preparation

The dataset used in this research is a standard dataset that has been collected from the University of California, Irvine (UCI) repository [45].

This dataset is related to energy efficiency in buildings and consists of eight different characteristics of buildings which act as input variables (X_1,$X_2 \ldots X_8$) and heating load (Y_1) and cooling load (Y_2) of buildings as two output variables.

The detailed description of the parameters of the data used along with symbols and its respective type is given in Table 2.

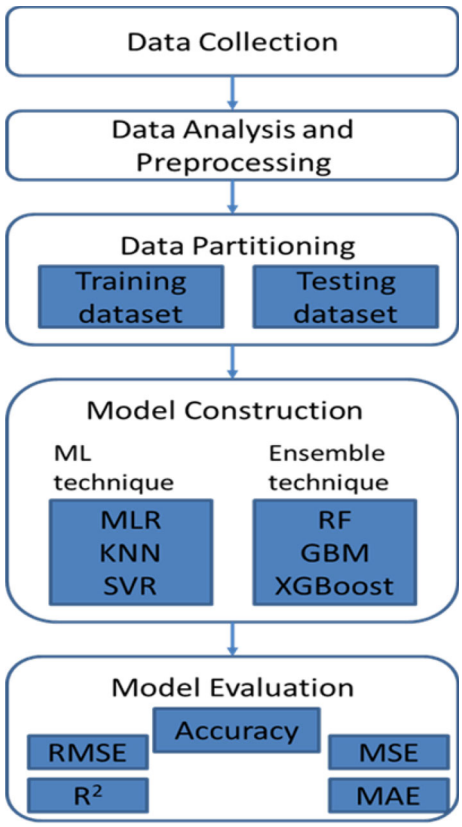

Fig. 7 Research approach

Table 2 Dataset parameter description

Parameter	Unit	Symbol	Type
Relative compactness	–	X_1	Input
Surface area	m^2	X_2	Input
Wall area	m^2	X_3	Input
Roof area	m^2	X_4	Input
Overall height	m	X_5	Input
Orientation	–	X_6	Input
Glazing area	m^2	X_7	Input
Glazing area distribution	–	X_8	Input
Heating load	KWh/m^2	Y_1	Output
Cooling load	KWh/m^2	Y_2	Output

4.2 Data Analysis and Pre-processing

Data pre-processing is a process that consists of checking the dataset for missing values and filling them with appropriate values, detecting and removing any outliers, converting it into a particular form suitable for applying algorithm, attribute selection, etc.

Table 3 Parameter statistics

Statistic	Parameter							
	X_1	X_2	X_3	X_4	X_5	X_6	X_7	X_8
Min	0.62	514.5	245	110.25	3.5	2	0	0
Max	0.98	808.5	416.5	220.5	7	5	0.4	5
Mean	0.76	671.7	318.5	176.6	5.25	3.5	0.23	2.8
SD	0.1	88	43.59	45.13	1.75	1.11	0.13	1.54
Variance	0.01	7749.06	1900.79	2037.3	3.06	1.25	0.01	2.40

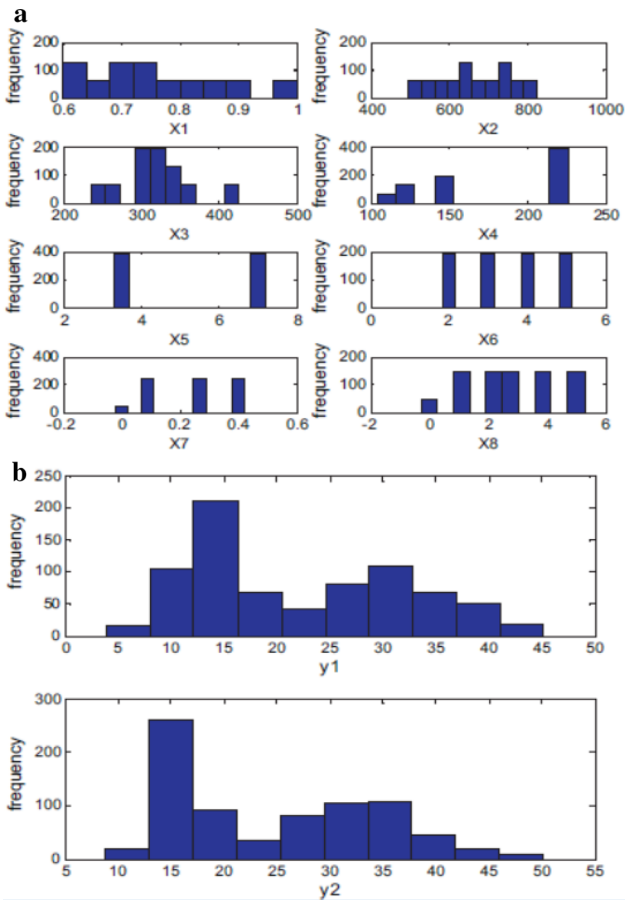

Fig. 8 Probability density estimates [12]

4.2.1 Statistical Analysis

The statistics of input parameters like minimum value, maximum value, mean, standard deviation, and variance were derived and are described in Table 3.

Probability distribution of all input variables, X_1 to X_8 and both the output variables, Y_1 and Y_2 using histograms is shown in Fig. 8. The distribution graphs show that none of the input and output variables follow Normal distribution.

4.2.2 Feature Selection

Feature selection is an important step in the process of predicting results using machine learning because all the features are generally not equally important for predicting the response value. Some features carry more weights than others for deriving a particular value and are thus more important, whereas some are very less or not at all important in the derivation of results. Such irrelevant features need to be excluded from the input to save training time and computation time. Additionally, applying the algorithm on only important and relevant features may result in more accurate prediction, reducing over fitting. In this research, feature selection has been performed in the following two ways:

Filter Feature Selection It is a univariate method in which statistical techniques are used to derive the relationship between each input variable and the target variable. The features which are strongly related to the response variable are selected as input for algorithm application and the features which are weakly related to the response variable can be eliminated. In our research, Spearman correlation coefficient was calculated to derive the strength of the relationship of several independent variables of the dataset with each of the response variables. Spearman's method for computing correlation was employed as the distribution of dataset used is non-Gaussian. A zero value for the correlation coefficient means the variables are not correlated, i.e., they are independent. A value closer to 1 indicates a high correlation among variables [46]. High correlation among two variables means one variable varies in accordance with the other; if one increases the other also increases. Similarly, reduction in one variable tends to reduce the other. The values of correlation coefficient between independent variables X_1–X_8 and Y_1 are represented in Fig. 9 and the same with Y_2 are represented in Fig. 10.

Feature Importance As mentioned earlier, features play a very important role in prediction and some features tend to be more important than others. In this context, feature importance graphs were generated to obtain the degree of effectiveness of each of the independent parameters, so that the contribution of each feature in prediction can be known and accordingly selection can be made. Figures 11 and 12 show the feature importance graph generated using random forests and gradient boosting machines, respectively. These techniques showed similar importance of features for both response variables, Y_1 and Y_2. According to random forests, overall height has maximum importance, followed by relative compactness, then surface area, wall area, glazing area, roof area, glazing area distribution and orientation. The sequence of features as derived by gradient boosting machines in the decreasing order of importance is—relative compactness,

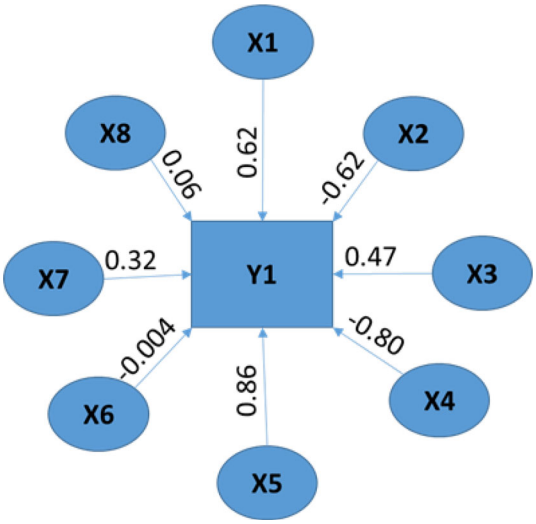

Fig. 9 Correlation of input variables with Y_1

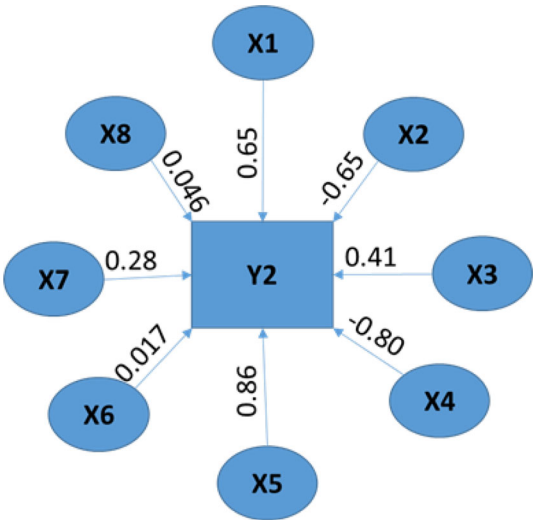

Fig. 10 Correlation of input variables with Y_2

Surface area, roof area, overall height, glazing area, wall area, orientation and glazing area distribution.

Figures 13 and 14 represent the graphs generated by applying LASSO technique for feature importance for Y_1 and Y_2, respectively. According to LASSO, relative compactness, overall height, glazing area and glazing area distribution are more important for predicting the heating load of a building. Furthermore, for predicting cooling load, the parameters—relative compactness, surface area, overall height, orientation and glazing area—are more important than others. Therefore, for performing experiments, X_1, X_5, X_7 and X_8 have been selected as input parameters for predicting Y_1. Likewise, X_1, X_2, X_5, X_6 and X_7 have been selected for the prediction of Y_2.

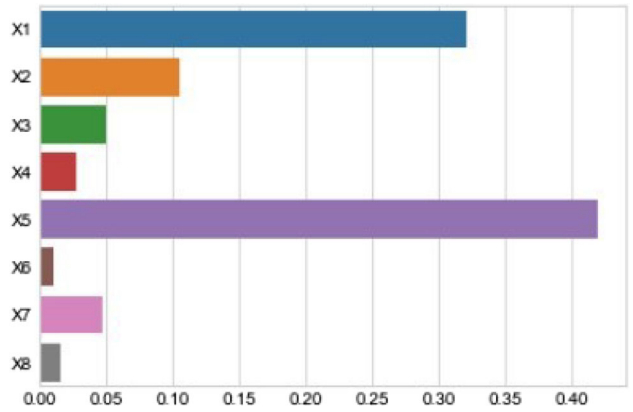

Fig. 11 Feature importance using RF

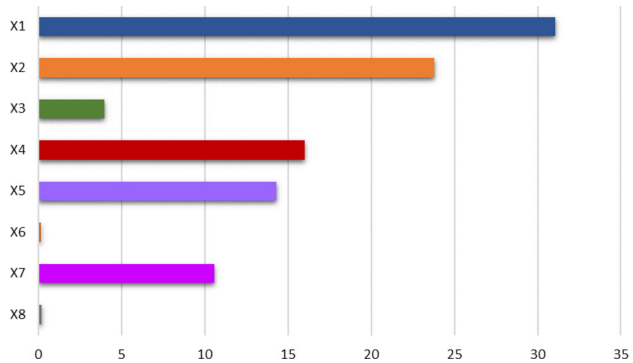

Fig. 12 Feature importance using GBM

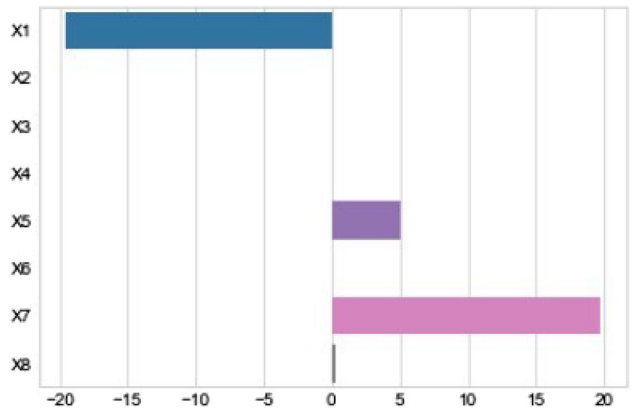

Fig. 13 Feature importance for Y_1 using LASSO

4.3 Data Analysis and Pre-processing

Dataset was partitioned according to 70–30% rule into two subsets: training dataset and testing dataset. For partitioning, random sampling without replacement was applied which resulted in 70% training data, on which the algorithms were applied, and the remaining 30% was used for testing the algorithms.

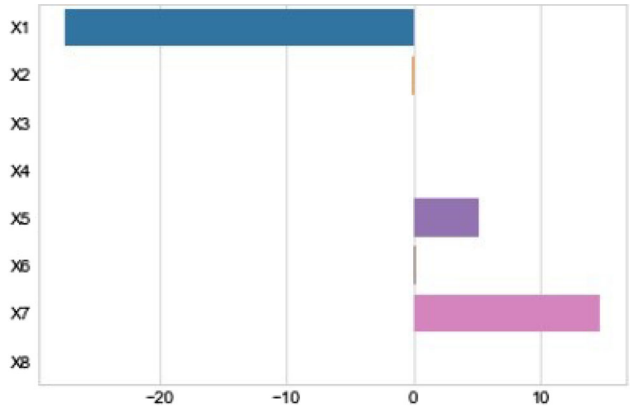

Fig. 14 Feature importance for Y_2 using LASSO

4.4 Model Construction

The model was constructed by applying three machine learning techniques namely—multiple linear regression, K-nearest neighbours and support vector regression. Three ensemble techniques were applied namely—random forests, gradient boosting machines, and extreme gradient boosting. The models were applied on the training dataset and tested using the testing dataset.

4.5 Model Evaluation

The evaluation of the results obtained after applying algorithms was done using five well-known performance measures namely root mean square error, mean square error, mean absolute error, R squared and accuracy. These measures can be calculated by applying following formulae, where Y_i: is the observed value for the ith observation, $\hat{Y}_i$: is the predicted value, N: is sample size.

The original dataset consisting of 768 instances has been partitioned into two subsets—70% training dataset and 30% testing dataset, by random sampling. So $N = 538$ for model construction on training dataset and $N = 230$ for testing purpose.

Root mean square error Following equation defines the formula for RMSE:

$$\text{RMSE} = \sqrt[2]{\sum_{i=1}^{N} \frac{\left(\hat{Y}_i - Y_i\right)^2}{N}} \tag{9}$$

Mean square error Following equation defines the formula for MSE:

$$\text{MSE} = \frac{1}{N} \sum_{i=1}^{N} \left(Y_i - \hat{Y}_i\right)^2 \tag{10}$$

Table 4 Results of ML algorithms

Performance metric	ML algorithms					
	MLR		KNN		SVR	
	Y_1	Y_2	Y_1	Y_2	Y_1	Y_2
RMSE	3.4	3.68	2.86	2.25	4.22	3.13
MSE	11.56	13.54	8.2	5.07	17.85	9.8
MAE	2.61	2.62	1.96	1.54	3.19	2.25
R Squared	0.87	0.83	0.90	0.94	0.76	0.84
Accuracy (%)	88.59	85.26	91.91	94.47	82.38	89.32

Mean absolute error MAE can be defined by the following equation:

$$\text{MAE} = \sum_{i=1}^{N} \frac{\left|Y_i - \hat{Y}_i\right|}{N} \tag{11}$$

R Squared R squared can be defined by the following equation:

$$R^2 = 1 - \frac{\sum\left(Y_i - \hat{Y}_i\right)^2}{\sum\left(Y_i - \bar{Y}\right)^2} \tag{12}$$

Accuracy Accuracy of a model can be calculated using the following formula:

$$\text{Accuracy} = \frac{|V_\text{A} - V_\text{O}|}{V_\text{A}} * 100 \tag{13}$$

where V_A: actual value, V_O: obtained value.

Sample calculations performed on the dataset using the above equations are shown in "Appendix".

5 Results

All the experiments of the research were performed using Python programming language. Three machine learning algorithms namely MLR, KNN and SVR and three ensemble techniques namely, RF, GBM, and XGBoost have been experimented on the collected dataset. The results of the ML and Ensemble experiments are described in Tables 4 and 5 respectively.

5.1 Results of Classical ML Techniques

The results obtained after applying all the three aforementioned classical machine learning algorithms on the dataset are summarized in Table 4. These results are based on the performance measures. The values of RMSE range between

Table 5 Results of ensemble algorithms

Performance metric	Ensemble algorithms					
	RF		GBM		XGBoost	
	Y_1	Y_2	Y_1	Y_2	Y_1	Y_2
RMSE	0.50	2.18	0.52	1.86	0.50	1.93
MSE	0.25	4.78	0.27	3.47	0.25	3.72
MAE	0.36	1.39	0.38	1.25	0.37	1.27
R Squared	0.99	0.94	0.99	0.96	0.99	0.95
Accuracy (%)	99.74	94.79	99.73	96.22	99.75	95.94

3.13 and 4.22 for both output variables Y_1 and Y_2 after applying MLR and SVR, whereas it is lower, 2.86 for Y_1 and 2.25 for Y_2 when KNN is applied. Correspondingly MSE values for KNN are also lower than MLR and SVR. Similarly, MAE values range between 2.25 and 3.19 using MLR and SVR, and the values are 1.96 and 1.54 using KNN. *R* Squared values are better in KNN (0.90 and 0.94), as compared to MLR (0.87 and 0.83) and SVR (0.76 and 0.84). KNN results are better than the other two algorithms in terms of accuracy also.

5.2 Results of Ensemble Techniques

Table 5 summarizes the results of the experiments performed by applying Ensemble techniques. As per the results, RMSE is 0.50 for output variable Y_1 for RF and XGBoost and 0.52 for GBM. Y_2 value varies slightly, 2.18 for RF, 1.86 for GBM and 1.93 for XGBoost. Correspondingly MSE value is also same 0.25 for Y_1 with RF and XGBoost and 0.27 with GBM. For Y_2, the values of MSE are 4.78, 3.47 and 3.72 with RF, GBM and XGBoost, respectively. MAE value varies slightly for Y_1 in the range 0.36–0.38 for all three algorithms, whereas the range for Y_2 is 1.25–1.39. *R* Squared values for Y_1 for all three algorithms are same, 0.99 and for Y_2; they vary from 0.94 to 0.96. Accuracy is also same, above 99% for Y_1 with all three algorithms, whereas accuracy percentage for Y_2 is 94.79%, 96.22% and 95.94% when RF, GBM and XGBoost are applied, respectively.

Figures 15 and 16 show the graphs plotted for results obtained in Table 4 for ML algorithms and Table 5 for ensemble algorithms, respectively. Figure 17 represents combined results for all six algorithms (ML and Ensemble) for both response variables Y_1 and Y_2.

5.3 Comparative Analysis of Machine Learning and Ensemble Learning Algorithms

In this section, the results of experiments are represented graphically based on various performance measures used for results evaluation. The graphs for RMSE, MAE, *R* squared and accuracy are shown in Figs. 18, 19, 20 and 21, respectively.

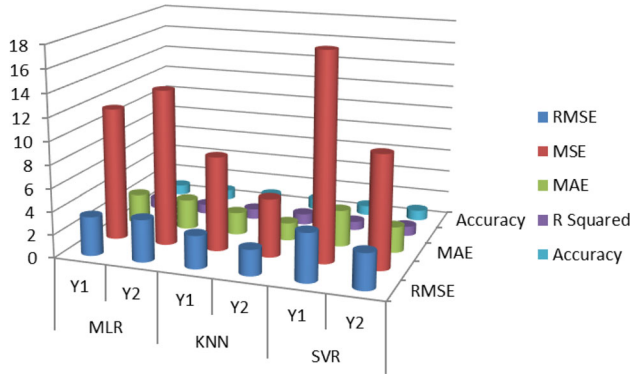

Fig. 15 Graphical representation of Table 4 results

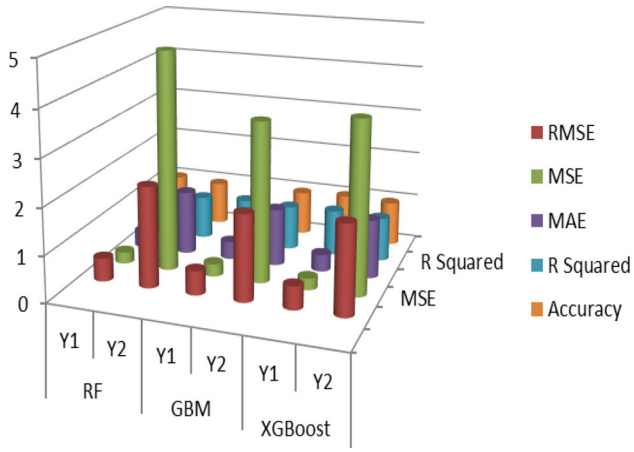

Fig. 16 Graphical representation of Table 5 results

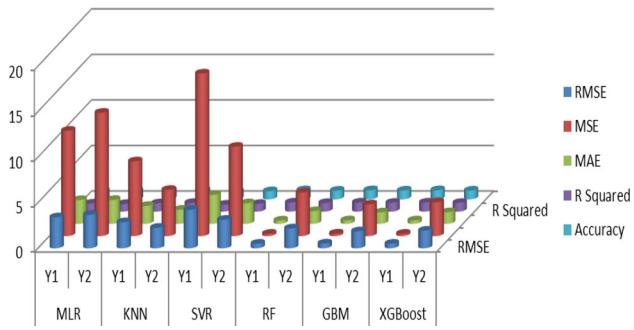

Fig. 17 Combined graph for all six algorithm results

It can be observed from Fig. 18 that the RMSE value is high (more than 3.0) for SVR and MLR algorithms for both the output variables Y_1 and Y_2, comparatively lower (approximately 2.0) for KNN, whereas the error values are extremely low (below 0.5) with all ensemble algorithms—RF, GBM and XGBoost for Y_1 and between 1.8 and 1.9 for Y_2.

Similar pattern can be observed from Fig. 19 for MAE. The values for MLR and SVR algorithms range

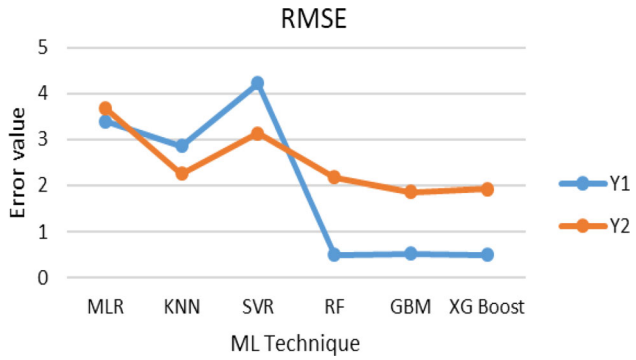

Fig. 18 Graph for RMSE for Y_1 and Y_2

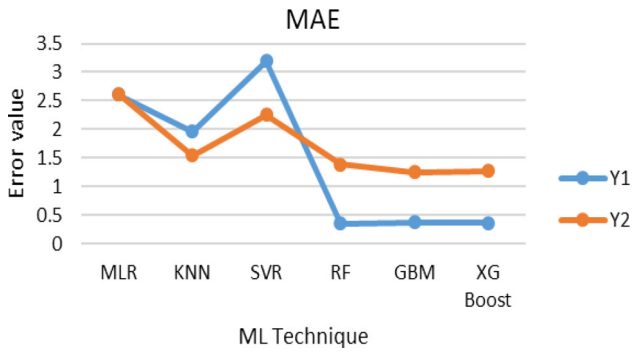

Fig. 19 Graph for MAE for Y_1 and Y_2

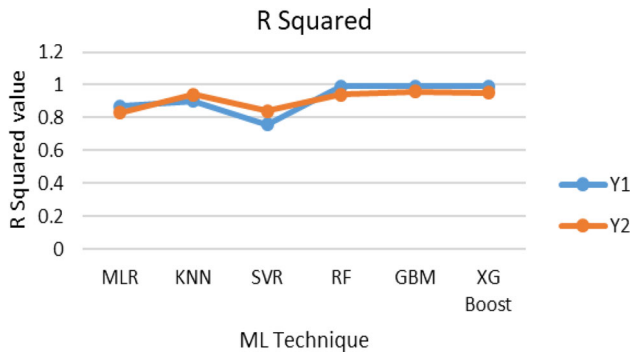

Fig. 20 Graph for R squared for Y_1 and Y_2

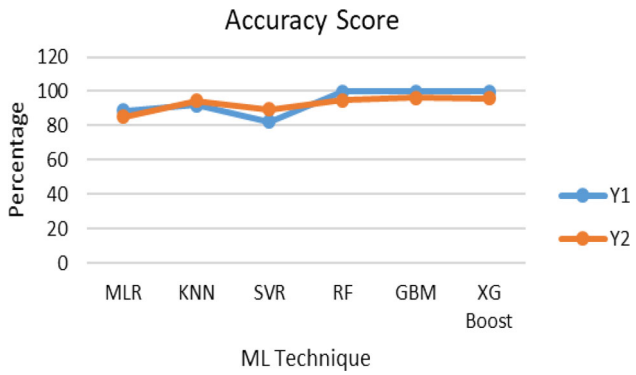

Fig. 21 Graph for accuracy for Y_1 and Y_2

between 2.32 and 2.63 for both Y_1 and Y_2. The values are lower when KNN is applied, 1.50 for Y_1 and 1.35 for Y_2. The results are better with ensemble algorithms with MAE between 1.19 and 1.27 for Y_2 and even lower values (0.35–0.36) are obtained for Y_1.

For R squared (Fig. 20), higher values, i.e., the values approaching 1 are considered better. In this context, again ensemble techniques have outperformed traditional ML techniques. The lowest values for R squared are obtained for SVR, 0.82 and 0.79 for Y_1 and Y_2, respectively. Slightly higher values are obtained for MLR, 0.88 and 0.83, and even higher for KNN with 0.94 and 0.96. R Squared results obtained with ensemble algorithms are significantly better than traditional algorithms with values ranging from 0.94 and going up to 0.99.

In Fig. 21, the plot for accuracy score also concludes that ensemble techniques perform significantly better than traditional algorithms with accuracy ranging between 96 and 99.76%. Among classical ML techniques, KNN performs better with an accuracy score of 95.12% and 96.47%, while the accuracy score for MLR and SVR ranges between 85.75 and 89.63%.

Ensemble techniques became popular from last two decades in the area of classification and prediction. The idea behind ensemble methods is that it can be compared to situations in real life, such as when critical decisions has to be taken, often opinions of several experts are taken into account rather than relying on a single judgment. Ensembles have shown to be more accurate in many cases than the individual models. Ideal ensembles consist of models with high accuracy which differ as much as possible. If each model makes different mistakes, then the total error will be reduced, if the models are identical, then a combination is useless since the results remain unchanged. It is evident from the survey of literature performed in Table 1 [13, 14, 17, 18, 23, 25, 26, 29] that ensemble techniques are far better in terms of performance prediction as compared to traditional machine learning algorithms. On similar terms, the results of experiments performed in this research also conclude that the predictions done by Ensemble models resulted in much lower error values RMSE, MSE and MAE, better R squared values and improved accuracy as compared to the traditional machine learning models used.

6 Conclusion and Future Scope

The issue of energy consumption at a fast pace and in large amounts demands solutions in this area, which can help in using the energy efficiently. Globally, buildings are the largest energy consumers, accounting for nearly

40% energy consumption. Therefore, the analysis of various energy-consuming components of a building reveal that the HVAC system consumes a large percentage of the building's total energy. HVAC needs energy for operation so that it can meet the heating load and cooling load requirements of the building. Heating and cooling loads are largely affected by various attributes of a building. This research shows that relative compactness, surface area, overall height, orientation and glazing area are more important in predicting heating load and cooling load of the buildings. Furthermore, the results of experiments prove that ensemble techniques perform better than traditional machine learning techniques.

In this research, only one dataset is used. In future, we can apply experiments on multiple datasets with large number of instances to better prove the accuracy of models. Apart from the models applied in this research, more advanced models like stacking and voting can be applied for better analysis.

Authors' Contribution Mrinal Pandey and Monika Goyal conducted the research and analyze the data. Monika Goyal performed the literature survey and experiments. Statistical Analysis is done by Mrinal Pandey. The research article is written by Mrinal Pandey and Monika Goyal.

Data Availability Yes, Data are available.

Code Availability Yes, code is available.

Appendix

Sample Calculations for Model Evaluation

The sample calculations using formulae in Eqs. 9–13 are described here. Table 6 contains the predicted values, observed values of response variables Y_1 and Y_2 from the dataset and predicted values after applying KNN algorithms. The calculations for model evaluation on the basis of values given in Table 6 have been performed manually on 20 and 100 sample size, respectively, which has been selected in respective order from 1–10 and 1–100.

Referring to Eqs. 9–13, applying the formulae on observed values and values predicted using KNN, For Y_1 calculated results for samples of initial 20 records,

RMSE = 12.01
MSE = 144.29
MAE = 10.2
R Squared = −5.2
Accuracy = 43.82%

Referring to Eqs. 9–13, applying the formulae on observed values and values predicted using KNN, For Y_1 calculated results for samples of initial 100 records,

Table 6 Sample dataset showing all predictor values and predicted values using KNN

X_1	X_2	X_3	X_4	X_5	X_6	X_7	X_8	Y_1 (observed)	Y_1 (predicted using KNN)	Y_2 (observed)	Y_2 (predicted using KNN)
0.98	514.5	294	110.25	7	2	0	0	15.55	13.325	21.33	14.595
0.98	514.5	294	110.25	7	3	0	0	15.55	29.1475	21.33	31.155
0.98	514.5	294	110.25	7	4	0	0	15.55	34.46	21.33	35.505
0.98	514.5	294	110.25	7	5	0	0	15.55	39.0375	21.33	41.9675
0.9	563.5	318.5	122.5	7	2	0	0	20.84	32.55	28.28	37.0325
0.9	563.5	318.5	122.5	7	3	0	0	21.46	16.4125	25.38	19.75
0.9	563.5	318.5	122.5	7	4	0	0	20.71	16.635	25.16	19.99
0.9	563.5	318.5	122.5	7	5	0	0	19.68	28.4525	29.6	29.7375
0.86	588	294	147	7	2	0	0	19.5	13.63	27.3	16.3875
0.86	588	294	147	7	3	0	0	19.95	24.61	21.97	30.33
0.86	588	294	147	7	4	0	0	19.34	13.1225	23.49	16.095
0.86	588	294	147	7	5	0	0	18.31	31.775	27.87	32.715
0.82	612.5	318.5	147	7	2	0	0	17.05	25.4975	23.77	28.755
0.82	612.5	318.5	147	7	3	0	0	17.41	14.505	21.46	17.6525
0.82	612.5	318.5	147	7	4	0	0	16.95	12.265	21.16	15.03
0.82	612.5	318.5	147	7	5	0	0	15.98	38.89	24.93	45.29
0.79	637	343	147	7	2	0	0	28.52	36.775	37.73	39.5125
0.79	637	343	147	7	3	0	0	29.9	13.65	31.27	16.8475
0.79	637	343	147	7	4	0	0	29.63	13.9125	30.93	14.63
0.79	637	343	147	7	5	0	0	28.75	35.7175	39.44	36.4425

Table 7 Sample dataset showing all predictor values and predicted values using XGBoost

X_1	X_2	X_3	X_4	X_5	X_6	X_7	X_8	Y_1 (observed)	Y_1 (predicted using XGBoost)	Y_2 (observed)	Y_2 (predicted using XGBoost)
0.98	514.5	294	110.25	7	2	0	0	15.55	12.8211	21.33	14.1680
0.98	514.5	294	110.25	7	3	0	0	15.55	29.0209	21.33	31.4201
0.98	514.5	294	110.25	7	4	0	0	15.55	35.6678	21.33	35.8985
0.98	514.5	294	110.25	7	5	0	0	15.55	39.0577	21.33	39.6939
0.9	563.5	318.5	122.5	7	2	0	0	20.84	35.3105	28.28	36.2770
0.9	563.5	318.5	122.5	7	3	0	0	21.46	16.6586	25.38	19.9970
0.9	563.5	318.5	122.5	7	4	0	0	20.71	15.2877	25.16	19.2294
0.9	563.5	318.5	122.5	7	5	0	0	19.68	28.5264	29.6	29.7966
0.86	588	294	147	7	2	0	0	19.5	14.4582	27.3	17.7255
0.86	588	294	147	7	3	0	0	19.95	26.1724	21.97	28.9496
0.86	588	294	147	7	4	0	0	19.34	14.6530	23.49	17.1983
0.86	588	294	147	7	5	0	0	18.31	32.4698	27.87	34.1773
0.82	612.5	318.5	147	7	2	0	0	17.05	26.0209	23.77	28.6788
0.82	612.5	318.5	147	7	3	0	0	17.41	15.1882	21.46	17.9761
0.82	612.5	318.5	147	7	4	0	0	16.95	12.6591	21.16	15.8440
0.82	612.5	318.5	147	7	5	0	0	15.98	38.9452	24.93	40.7834
0.79	637	343	147	7	2	0	0	28.52	28.8261	37.73	32.9217
0.79	637	343	147	7	3	0	0	29.9	14.9532	31.27	17.8712
0.79	637	343	147	7	4	0	0	29.63	14.5771	30.93	15.2797

RMSE = 14.02

MSE = 196.6

MAE = 10.9

R Squared = −1.55

Accuracy = 48.46%

The sample calculations using formulae in Eqs. 9–13 are described here. Table 7 contains the predicted values, observed values of response variables Y_1 and Y_2 from the dataset, and predicted values after applying XGBoost algorithms. The calculations for model evaluation on the basis of values given in Table 7 have been performed manually on 20 and 100 sample size, respectively, which has been selected in respective order from 1–20 and 1–100.

Applying the formulae on the values predicted using XGBoost, For Y1 the calculated results for samples of initial 20 records,

RMSE = 11.98

MSE = 143.59

MAE = 9.82

R Squared = −5.17

Accuracy = 40.68

Applying the formulae on the values predicted using XGBoost, For Y_1 the calculated results for samples of initial 100 records,

RMSE = 14.25

MSE = 203.1

MAE = 11.13

R Squared = −1.63

Accuracy = 49.57

References

1. Lam, J.C.; Wan, K.K.; Tsang, C.L.; Yang, L.: Building energy efficiency in different climates. Energy Convers. Manag. **49**(8), 2354–2366 (2008)
2. Ahmad, M.W.; Mourshed, M.; Rezgui, Y.: Trees vs neurons: comparison between random forest and ANN for high-resolution prediction of building energy consumption. Energy Build. **147**, 77–89 (2017)
3. Chou, J.S.; Bui, D.K.: Modeling heating and cooling loads by artificial intelligence for energy-efficient building design. Energy Build. **82**, 437–446 (2014)
4. Jain, R.K.; Smith, K.M.; Culligan, P.J.; Taylor, J.E.: Forecasting energy consumption of multi-family residential buildings using support vector regression: investigating the impact of temporal and spatial monitoring granularity on performance accuracy. Appl. Energy **123**, 168–178 (2014)
5. Krzywinski, M.; Altman, N.: Multiple linear regression: when multiple variables are associated with a response, the interpretation

of a prediction equation is seldom simple. Nat. Methods **12**(12), 1103–1105 (2015)
6. Carreira, P.; Costa, A.A.; Mansu, V.; Arsénio, A.: Can HVAC Really Learn from Users? A Simulation-Based Study on the Effectiveness of Voting for Comfort and Energy Use Optimization. Sustain. Cities Soc. **41**, 275–285 (2018)
7. Drgoňa, J.; Picard, D.; Kvasnica, M.; Helsen, L.: Approximate model predictive building control via machine learning. Appl. Energy **218**, 199–216 (2018)
8. Roy, S.S.; Roy, R.; Balas, V.E.: Estimating heating load in buildings using multivariate adaptive regression splines, extreme learning machine, a hybrid model of MARS and ELM. Renew. Sustain. Energy Rev. **82**, 4256–4268 (2018)
9. Kumar, S.; Pal, S.K.; Singh, R.P.: A novel method based on extreme learning machine to predict heating and cooling load through design and structural attributes. Energy Build. **176**, 275–286 (2018)
10. Ngo, N.T.: Early predicting cooling loads for energy-efficient design in office buildings by machine learning. Energy Build. **182**, 264–273 (2019)
11. Sunikka-Blank, M.; Galvin, R.: Introducing the prebound effect: the gap between performance and actual energy consumption. Build. Res. Inf. **40**(3), 260–273 (2012)
12. Galvin, R.: Making the 'rebound effect'more useful for performance evaluation of thermal retrofits of existing homes: defining the 'energy savings deficit'and the 'energy performance gap'. Energy Build. **69**, 515–524 (2014)
13. Tsanas, A.; Xifara, A.: Accurate quantitative estimation of energy performance of residential buildings using statistical machine learning tools. Energy Build. **49**, 560–567 (2012)
14. Fan, C.; Xiao, F.; Wang, S.: Development of prediction models for next-day building energy consumption and peak power demand using data mining techniques. Appl. Energy **127**, 1–10 (2014)
15. Wei, X.; Kusiak, A.; Li, M.; Tang, F.; Zeng, Y.: Multi-objective optimization of the HVAC (heating, ventilation, and air conditioning) system performance. Energy **83**, 294–306 (2015)
16. Park, H.S.; Lee, M.; Kang, H.; Hong, T.; Jeong, J.: Development of a new energy benchmark for improving the operational rating system of office buildings using various data-mining techniques. Appl. Energy **173**, 225–237 (2016)
17. Candanedo, L.M.; Feldheim, V.; Deramaix, D.: Data driven prediction models of energy use of appliances in a low-energy house. Energy Build. **140**, 81–97 (2017)
18. Manjarres, D.; Mera, A.; Perea, E.; Lejarazu, A.; Gil-Lopez, S.: An energy-efficient predictive control for HVAC systems applied to tertiary buildings based on regression techniques. Energy Build. **152**, 409–417 (2017)
19. Peng, Y.; Rysanek, A.; Nagy, Z.; Schlüter, A.: Using machine learning techniques for occupancy-prediction-based cooling control in office buildings. Appl. Energy **211**, 1343–1358 (2018)
20. Gallagher, C.V.; Bruton, K.; Leahy, K.; O'Sullivan, D.T.: The suitability of machine learning to minimise uncertainty in the measurement and verification of energy savings. Energy Build. **158**, 647–655 (2018)
21. Deb, C.; Lee, S.E.; Santamouris, M.: Using artificial neural networks to assess HVAC related energy saving in retrofitted office buildings. Sol. Energy **163**, 32–44 (2018)
22. Nayak, S.C.: Escalation of forecasting accuracy through linear combiners of predictive models. EAI Endorsed Trans. Scalable Inf. Syst. **6**(22), 1–14 (2019)
23. Sethi, J.S.; Mittal, M.: Ambient air quality estimation using supervised learning techniques. EAI Endorsed Trans. Scalable Inf. Syst. **6**(22) (2019)
24. Pallonetto, F.; De Rosa, M.; Milano, F.; Finn, D.P.: Demand response algorithms for smart-grid ready residential buildings using machine learning models. Appl. Energy **239**, 1265–1282 (2019)
25. Pham, A.D.; Ngo, N.T.; Truong, T.T.H.; Huynh, N.T.; Truong, N.S.: Predicting energy consumption in multiple buildings using machine learning for improving energy efficiency and sustainability. J. Clean. Prod. **260**, 121082 (2020)
26. Walker, S.; Khan, W.; Katic, K.; Maassen, W.; Zeiler, W.: Accuracy of different machine learning algorithms and added-value of predicting aggregated-level energy performance of commercial buildings. Energy Build. **209**, 109705 (2020)
27. Xu, X.; Wang, W.; Hong, T.; Chen, J.: Incorporating machine learning with building network analysis to predict multi-building energy use. Energy Build. **186**, 80–97 (2019)
28. Zhou, G.; Moayedi, H.; Bahiraei, M.; Lyu, Z.: Employing artificial bee colony and particle swarm techniques for optimizing a neural network in prediction of heating and cooling loads of residential buildings. J. Clean. Prod. **254**, 120082 (2020)
29. Gao, W.; Alsarraf, J.; Moayedi, H.; Shahsavar, A.; Nguyen, H.: Comprehensive preference learning and feature validity for designing energy-efficient residential buildings using machine learning paradigms. Appl. Soft Comput. **84**, 105748 (2019)
30. Seyedzadeh, S.; Rahimian, F.P.; Rastogi, P.; Glesk, I.: Tuning machine learning models for prediction of building energy loads. Sustain. Cities Soc. **47**, 101484 (2019)
31. Roy, S.S.; Samui, P.; Nagtode, I.; Jain, H.; Shivaramakrishnan, V.; Mohammadi-Ivatloo, B.: Forecasting heating and cooling loads of buildings: a comparative performance analysis. J. Ambient Intell. Humaniz. Comput. **11**(3), 1253–1264 (2020)
32. Iruela, J.R.S.; Ruiz, L.G.B.; Pegalajar, M.C.; Capel, M.I.: A parallel solution with GPU technology to predict energy consumption in spatially distributed buildings using evolutionary optimization and artificial neural networks. Energy Convers. Manag. **207**, 112535 (2020)
33. Das, S.; Swetapadma, A.; Panigrahi, C.; Abdelaziz, A.Y.: Improved method for approximation of heating and cooling load in urban buildings for energy performance enhancement. Electr. Power Compon. Syst. **48**, 1–11 (2020)
34. Cozza, S.; Chambers, J.; Deb, C.; Scartezzini, J.L.; Schlüter, A.; Patel, M.K.: Do energy performance certificates allow reliable predictions of actual energy consumption and savings? Learning from the Swiss national database. Energy Build. **224**, 110235 (2020)
35. https://sweetcode.io/simple-multiple-linear-regression-python-scikit/
36. Cunningham, P.; Delany, S.J.: k-Nearest neighbour classifiers. Multiple Classif. Syst. **34**(8), 1–17 (2007)
37. Martínez, F.; Frías, M.P.; Pérez, M.D.; Rivera, A.J.: A methodology for applying k-nearest neighbor to time series forecasting. Artif. Intell. Rev. **52**(3), 2019–2037 (2019)
38. https://www.slideshare.net/amirudind/k-nearest-neighbor-presentation
39. Smola, A.J.; Schölkopf, B.: A tutorial on support vector regression. Stat. Comput. **14**(3), 199–222 (2004)
40. https://scikit-learn.org/0.18/auto_examples/svm/plot_svm_regression.html
41. Breiman, L.: Random forests. Mach. Learn. **45**(1), 5–32 (2001)
42. Cutler, A.; Cutler, D.R.; Stevens, J.R.: Random forests. In: Ensemble Machine Learning, pp. 157–175. Springer, Boston, MA (2012)
43. Friedman, J.H.: Stochastic gradient boosting. Comput. Stat. Data Anal. **38**(4), 367–378 (2002)
44. Chen, T.; Guestrin, C.: Xgboost: a scalable tree boosting system. In: Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, pp. 785–794 (2016)
45. https://archive.ics.uci.edu/ml/datasets/Energy+efficiency
46. Myers, L.; Sirois, M.J.: Spearman correlation coefficients, differences between. Encycl. Stat. Sci. (2004)