# Energy Consumption Prediction in Smart Homes Using QIO-Enhanced Regression Models

Yuanjun Zhang
College of Art and Design, Luoyang Vocational College of Science and Technology, Luoyang 471822, China
E-mail: yzdavid09162@163.com

**Keywords:** smart homes, machine learning schemes, quadratic interpolation optimization

*Accurately forecasting household energy consumption remains a challenge due to the variability introduced by user behavior, appliance diversity, and environmental conditions. Smart homes have also become a key solution to the expanding global energy demands in the residential arena through the introduction of new strategies to maximize and control the use of energy. Smart systems comprise sophisticated sensors and networked appliances to provide accurate, affordable, as well as ecofriendly management of energy. Yet, forecasting the use of energy in such homes is challenging with the differences in end-user behavior, weather, and appliance efficiencies. In this work, the use of machine learning (ML) algorithms Extra Tree Regression (ETR), Naive Bayes Regression (NBR), and Elastic Net Regression (ENR) to predict the use of energy in a home is presented. These are also optimized with Quadratic Interpolation Optimization (QIO) to adjust their hyper parameters. Experiments were performed using the Kaggle Smart Home Energy Usage dataset, which provides comprehensive synthetic data across various features of occupancy, appliance usage, temperature, and timestamped consumption data. The developed hybrid schemes were tested using the criteria of $R^2$ and RMSE in the training, validation, and testing stages. Out of the varied modeled algorithms, the ETQI (ETR + QIO) model realized the highest accuracy of $R^2 = 0.985$ and RMSE = 0.245 while outperforming NBQI (NBR + QIO) with $R^2 = 0.974$ and RMSE = 0.233 and ENQI (ENR + QIO) with $R^2 = 0.952$ and RMSE = 0.319. These conclusions reflect the ability of optimized ML algorithms to drive more effective energy efficiency strategies in the smart home scenario.*

*Povzetek: Članek predstavi QIO-optimizirane regresijske modele (ETQI, NBQI, ENQI) za napoved porabe energije v pametnih domovih, kjer ETQI doseže najvišjo točnost in stabilnost.*

## 1 Introduction

Due to the unending demand for energy consumption in various parts of the world, especially in the residential sector, efforts towards enhancing energy management have been considered imperative [1]. Integrated homes with multiple devices and sensors are considered a significant step in energy matters [2]. Such dwellings can manage, regulate, and optimize energy intake in unique and valuable ways that preserve the environment while saving the homeowner's money [3]. However, predicting energy use in smart homes has unique difficulties since several variables affect energy use, including appliance efficiency, weather, and user behavior [4]. A smart home consists of two interactive elements: an intelligent network and a smart load comprising sensors, actuators, middleware, and a network [5], [6]. A bright house can focus on three primary goals: maximizing the percentage of home automation, making energy regulation more straightforward, and reducing environmental releases [7]. Today, smart buildings have become a phenomenon as they are overwhelming the world by developing brighter built spaces through the application of leading-edge computational and communication technologies [8]. Public awareness will relate to building automation, where operations are controlled through instruments and microcontrollers in an information exchange process [9]. Secondly, smart buildings utilize integrated systems that forecast energy consumption, hence triggering energy conservation and cost minimization [10]. More significantly, this is crucial given that it is proven that, within developed countries, buildings utilize approximately 40% of primary energy [11].

Energy regulation in the smart home relies on efficient appliances and increasing consumer engagement in demand-side management, energy management systems, and appliance design [12]. Many studies have demonstrated that appliance design and efficiency are paramount; outdated appliances and consumer ignorance of their efficiency are also significant concerns [13]. The authors continue to advocate stricter regulations over appliances, but the responsibility for energy efficiency lies with the consumer. Indeed, several studies have shown a high correlation between awareness of the amount of energy used and the actual consumption. Real-time feedback, through smartphones or fixed-site monitoring equipment, is encouraging, but there is still much interface opacity and technological issues. Energy efficiency extends beyond smart appliances into other areas like space heating, despite possible efficiencies from thermostat positioning. Managing smart home systems requires the customer's awareness and experience. Intertek

defines a smart home as a home with network-connected appliances that are managed remotely [14].

The current research seeks to (1) contrast the accuracy of the three machine learning regression algorithms Extra Tree Regression (ETR), Naive Bayes Regression (NBR), and Elastic Net Regression (ENR) in modeling energy usage in the home, (2) assess the impact of Quadratic Interpolation Optimization (QIO) on the improvement of the accuracy of these models, and (3) ascertain the most significant features responsible for the energy usage trend from the feature importance analysis.

## 1.1 Related works

This section investigates various research articles on appliances and socio-economic factors concerning energy consumption. Pratt et al. [15] presented data collection of hourly end-use from 140 commercial buildings and 454 residential homes from the Pacific Northwest. They mainly research the development of techniques to predict energy use profiles, with a special interest in refrigeration and freezer loads. A data acquisition system is presented that monitors 12 to 16 channels for energy consumption. Data from UK households was analyzed using a Markov Chain Monte Carlo approach. At the same time, Guo et al. utilized a Hidden Markov model to identify and estimate the energy consumption of individual appliances from aggregated power signals. Other techniques from machine learning have been used in various works: Ling et al. [16], Veit et al. [17], and Arghira et al. [18] analyzed neural networks (NNs), forecasting methods, and multiple regression techniques. Bharati et al. applied KNN, SVM, stochastic gradient descent, random forests, and neural networks for the classification attributes of the energy dataset [19]. They also utilized neuro-fuzzy inference for superconductor dataset feature prediction. Scott et al. [20] deployed a NARX Gaussian Process regression scheme and a linear regression scheme to predict energy loads from UK data, attaining a low prediction error of 0.26% [1], [2].

Other research filtered data from 1,628 households for electricity consumption, using the following criteria as a basis: weather, location, age of building, ownership, energy-efficient fixtures, and even income class [3], [4]. The analysis determined that weather, location, and floor area were the most significant factors in energy consumption. Nevertheless, daytime occupancy of the home was related to reduced appliance efficiency [21]. Overall, all literature underlines these different elements as an understanding of patterns in energy usage. Recent studies have introduced advanced forecasting and control models for smart energy systems. Razghandi et al. proposed a Seq2Seq model integrated with Q-learning for predictive control in smart homes [5]. Pan applied tree-based models optimized with War SO for accurate solar energy forecasting [6]. Nakıp et al. developed an rTPNN-based framework for joint forecasting and appliance scheduling in smart environments, achieving near-optimal scheduling and high predictive performance [7].

Table 1 gives a comparative overview of the significant studies in the area of predicting smart home energy usage. Comparison of the studies is in terms of the datasets, the machine learning algorithms used, the adopted feature engineering techniques, and model performances as measured by RMSE and $R^2$. The comparison highlights the novelty and strength of the adopted approach in this research, Quadratic Interpolation Optimization (QIO), in its ability to improve model performances in energy forecasting tasks.

Table 1: Comparative summary of key studies

| Study | Dataset | Algorithm | Feature Engineering | Results (RMSE, $R^2$) |
|---|---|---|---|---|
| Pratt et al. [15] | ELCAP, 140 buildings, 454 homes | Custom Regression Techniques | Manual Load Profiling | RMSE: N/A, $R^2$: N/A |
| Guo et al. [8] | UK Aggregated Load | Hidden Markov Model | Appliance-level Disaggregation | RMSE: ~0.32, $R^2$: ~0.91 |
| Ling et al. [16] | Synthetic Short-term Load | Neural Networks | Fuzzy Feature Selection | RMSE: 0.29, $R^2$: 0.94 |
| Bharati et al. [19] | FICTA Dataset | KNN, SVM, RF | Plot Analysis | RMSE: ~0.27, $R^2$: ~0.92 |
| Scott et al. [20] | UK Smart Meter | NARX + Gaussian Process | None Specified | RMSE: 0.26, $R^2$: 0.95 |
| Current Study | Kaggle Smart Home | ETR, NBR, ENR + QIO | Quadratic Interpolation Optimization | ETQI: RMSE: 0.245, $R^2$: 0.985 |

## 1.2 Objectives

Tree-based machine learning approaches were used in preference to deep learning solutions because of their reduced training time, improved interpretability, and good results on structured, table-like datasets such as the Smart Home Energy dataset. They use less data preprocessing and computational power as opposed to neural networks and are more apt for real-world, real-time applications in smart homes. This paper discusses a few Machines

Learning (ML) schemes that can be used for energy usage prediction in smart homes, mainly focusing on Extra Tree Regression (ETR), Naive Bayes Regression (NBR), and Elastic Net Regression (ENR). Each of these schemes has unique advantages when it comes to the estimation of energy consumption data. ETR has been widely

recognized for its efficiency in handling a large volume of data and its robustness to outliers, especially in the case of complex energy-use consumption patterns. In contrast to the above, NBR relies on a probability-based method that correctly addresses the uncertainty in the data, thereby leading to better prediction accuracy.

First, ENR embodies the characteristics of both Lasso and Ridge regressions and has stronger robustness for high correlations between features. Second, Quadratic Interpolation Optimization (QIO) is applied to improve the predictiveness of these schemes further. The QIO optimizer is essential in hyperparameter optimization, enabling it to find the best-fitted parameters for each model. The result is a holistic approach that leads to the development of predictive schemes with better performance, thus providing the most valuable insight into energy consumption efficiency in smart homes. This paper contributes to further understanding and optimally steering energy management within residential settings with the help of advanced ML techniques.

The remainder of this paper is organized as follows: Section 2 provides a detailed description of the dataset and preprocessing techniques. Section 3 explores the influence of occupancy status on energy consumption in smart homes. Section 4 presents the mathematical schemes, including the machine learning models and the QIO optimization process. Section 5 discusses the experimental results and model evaluations. Finally, Section 6 concludes the study and outlines future research directions.

# 2 Data description

## 2.1 Data preprocessing

Before model training, various steps in the preprocessing of the data were performed. Missing continuous variable values were replaced using mean substitution, but rows with missing categorical values were deleted as they were of low frequency. Continuous features like temperature and energy consumption were normalized with min–max scaling to have uniform contribution to the model during training. The day-of-week and season variables were encoded with one-hot encoding in order to maintain model consistency with non-numerical data. Timestamp data was also processed to its day, month, and hour components to enable more effective capturing of temporal trends in the energy behavior. The data was partitioned randomly to 70% for training, 15% for validation, and 15% for testing, with the aim of having a representative feature value distribution across phases. Five independent trials were conducted to ensure consistency of results.

Temporal characteristics were analyzed to augment their prediction value. The raw timestamp was analyzed to extract hour, day, and month numbers to identify periodic trends in energy use. The categorical characteristics "day of week" and "season" were encoded using one-hot encoding to preserve non-ordinal connections among the categories. This encoding enabled the models to leverage temporal context without creating erroneous hierarchies. No embedding layers were utilized, as the models are tree-based and proficiently manage one-hot encoded features.

## 2.2 Data description

All these details were obtained from the Smart Home Energy Usage Dataset from Kaggle, as noted at this link. The given dataset contains synthetic data, essential for further research on different scenarios related to the energy optimization subject within smart homes. Thus, the availability of data regarding electricity consumption, appliance usage, temperature settings, and occupancy status is critical for researchers and developers who want to enhance contemporary dwellings' energy efficiency and home automation. It is formatted in a manner that allows the detection of relevant consumption data, further facilitating the identification of potential energy-saving issues and optimization of home energy consumption to support automation. This dataset is useful in further research to analyze energy intensities to improve smart home operations and other automated processes that aim to improve home energy efficiency.

**Timestamp**: This column specifies the precise day and time the data was collected to allow comparison of customers' purchasing patterns at various time intervals.

**Home-id:** A unique number assigned to each house, allowing for comparisons and evaluations of the energy efficiency of various residences.

**Energy-consumption-kWh:** This column's first parameter, which allows for evaluating total energy consumption, is the energy usage measured in kilowatt-hours (kWh).

**Temperature-setting-C:** This feature displays the houses inside temperature in degrees Celsius, which is highly useful for estimating the energy a heater or air conditioner would use in a certain household.

**Occupancy status:** Indicates whether the house was occupied or unoccupied when the data was gathered, allowing the impact of presence on energy usage to be shown.

**Appliance:** It classifies the kind of appliance being used, including air conditioning units, washing machines, and dishwashers, helping to determine each appliance's contribution to the overall amount of energy used.

**Usage-duration-minutes:** The number of minutes an appliance was on helps distinguish between the quantity of energy used and the period the device was in use.

**Season:** The year when the data was collected since these aids in analyzing the patterns of energy consumption associated with each season, for instance, winter, spring, summer, or fall.

**Day-of-week:** This variable establishes the day of the week to differentiate daily energy use.

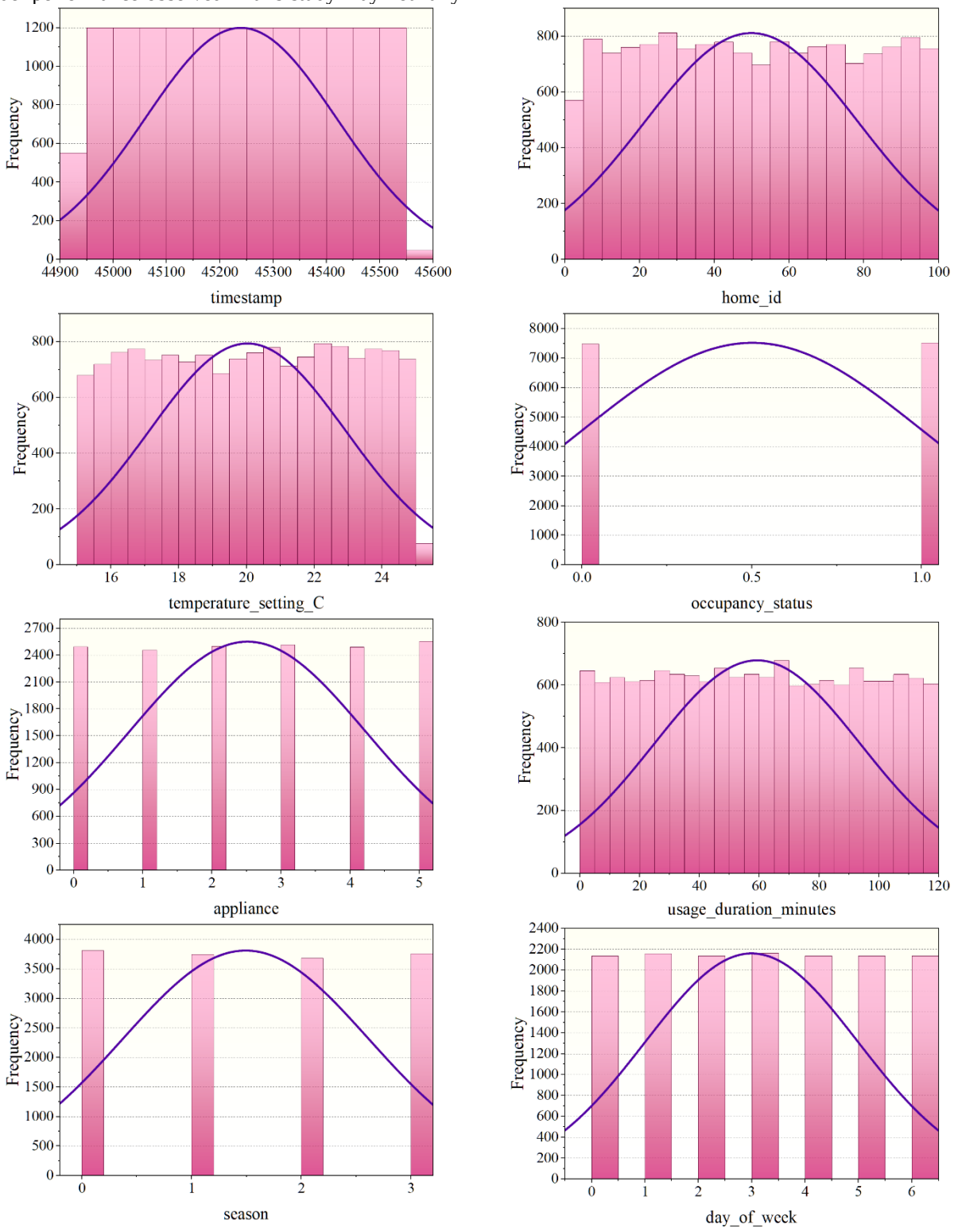
**Holiday:** A flag variable that, to filter out days that could vary due to the holiday, is equal to 1 if the day was a holiday and zero otherwise.

Therefore, it can be used in trend analysis, forecasting overall energy usage, enhancing intelligent automation at home, and, more importantly, developing automated energy-saving solutions. It allows for deeper and more

practical analysis of energy use in smart homes, thus making an environment even more effective and efficient.

The use of a synthetic dataset offers controlled conditions for model development and evaluation; however, it also presents limitations in terms of real-world applicability. Synthetic data may not capture the full complexity, variability, and noise present in actual household energy consumption patterns. Consequently, model performance observed in this study may not fully generalize to real-world deployments. Future work will incorporate real datasets, such as UK-DALE or REDD, to evaluate model robustness under realistic operational conditions and validate applicability in practical smart home environments.

Fig.1 presents the relationship between measured and predicted values, showing the model's capability to detect patterns and accuracy in prediction.



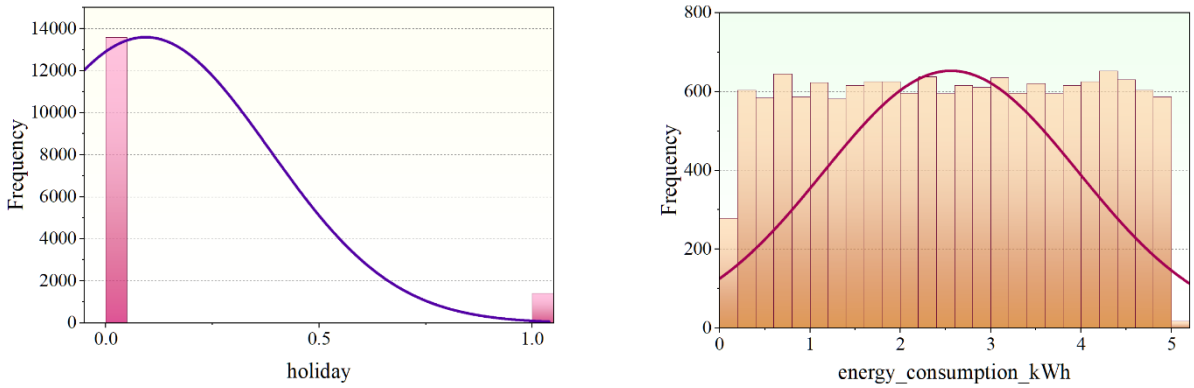



Figure 1: The links between input and output variables assessed utilizing a marginal histogram plot

# 3 Exploring the influence of occupancy status on energy consumption in smart homes

Energy consumption in smart homes is closely related to the occupancy status range, and it is primarily associated with HVAC. For instance, in the U.S., HVAC accounts for 43% of residential energy consumption, while it increases to 61% in colder regions like Canada and the U.K. Energy savings of up to 20-30% have also been estimated by switching these systems off. At the same time, the residents are asleep or away. However, these savings have often been hard to realize. Many residents do not usually change their thermostats, while programmable thermostats frequently result in higher energy use due to incorrect settings or user disinterest [22]. Occupancy-based control strategies have gained attention for improving energy efficiency by aligning HVAC operations with real-time user presence. Studies show that dynamic occupancy detection using sensors or predictive models can significantly reduce unnecessary energy consumption. Integrating occupancy data into energy prediction models enhances forecasting accuracy and enables proactive control, which is essential for smart home automation systems aiming to balance comfort, cost, and environmental sustainability.

# 4 Mathematical schemes

All experiments were implemented in Python version 3.13.3 using Visual Studio Code as the development environment. The machine learning models were developed using the scikit-learn library, which provided tools for model training, evaluation, and cross-validation.

The average runtime for base models was approximately 1.4 seconds for NBR, 3.1 seconds for ENR, and 4.2 seconds for ETR. When optimized with QIO, runtimes increased to 48.6 seconds for NBQI, 76.4 seconds for ENQI, and 101.8 seconds for ETQI. This added time reflects QIO's iterative tuning process, which enhances model performance and generalization.

## 4.1 Extra tree regression (ETR)

Geurts et al. [23] first introduced ETR, initially inspired by the Random Forest (RF) scheme. The ETR tactic generates regression trees, or sets of unpruned choices, in line with the conventional top-down methodology [23]. The RF scheme does the regression in two steps: bagging and bootstrapping. During the bootstrapping process, each tree develops to generate a collection of DTs using a sample from a stochastic training dataset. Upon arriving at the ensemble, the decision tree (DT) nodes undergo division by implementing a two-step bagging process. The first stage involves the selection of various random subsets of training data. The decision-making procedure is concluded by choosing the ideal subset and its value [24]. Breiman [25] thought of the RF scheme as being composed of several DTs where $G(x,\theta_r)$ displays the $G^{th}$ the unified individual distribution pattern assigned before the tree grows, denoted by θ in the prediction tree. Breiman is used to average and combine all of the trees, resulting in an ensemble of trees (a forest) of G(x). The procedure is shown in Eq (1):

$$G(x,\theta_1,\dots,\theta_r)=\frac{1}{R}\sum_{r=1}^{R}G(x,\theta_r) \quad (1)$$

There are two basic differentiations between RF and ETR. By randomly choosing a subset of all the cutting locations, the ETR separates nodes in the first place. It also cultivates trees by utilizing all of the learning samples to lessen bias. Two variables in the ETR technique control the splitting process: $k$ and $n_{min}$ .where the $n_{min}$ Variable indicates the least sample size needed to discriminate across nodes, and k r displays the number of characteristics the node randomly picks. Furthermore, $k$ and $n_{min}$, respectively, determine the medium result noise intensity and the feature extraction strength. These two elements reduce overfitting and boost the accuracy of the ETR scheme [26], [27].

## 4.2 Naive Bayes Regression

The example below looks at the challenge of anticipating a numerical goal value $Y$. Qualities in E $(X_1, X_2,\dots,X_m)$. are $m$. Features may be actual, quantitative, nominal, or a collection of unordered values [28]. To lower the predicted prediction error, $Y$ might be selected if the

probability density function $f(Y \mid E)$ of the target value was known. However, $f(Y \mid E)$ is usually unknown and necessitates data estimates. It is expected that the features. $X_1, X_2, \ldots, X_m$ are independent of the goal value $Y$ when Naive Bayes uses the Bayes theorem.

$$f(Y \mid E) = \frac{f(E,Y)}{\int f(E,Y)dY} = \frac{f(E \mid Y)f(Y)}{\int f(E \mid Y)f(Y)dY} \quad (2)$$

The Probability Density Function (PDF) of example E for a particular target value $Y$ is known as the like hood, or $f(E \mid Y)$, the pdf of the goal value before any examples are seen is known as the prior f(Y). Naive Bayes is represented in terms of attribute independence by Eq (2).

$$f(Y \mid E) = \frac{f(X_1 \mid Y)f(X_2 \mid Y) \cdots f(X_m \mid Y)f(Y)}{\int f(X_1 \mid Y)f(X_2 \mid Y) \cdots f(X_m \mid Y)f(Y)\,dY} \quad (3)$$

It is now possible to compute the individual PDFs $f(X_i \mid Y)$. instead of computing the overall pdf $f(E \mid Y)$. This reduction in dimensionality makes the learning task much more manageable. Since more data are needed to predict accurately $f(X_i \mid Y)$, this method is more trustworthy than estimating $f(E \mid Y)$.

## 4.3 Elastic Net Regression (ENR)

It is a high-performance linear regression technique that combines the best features of L1 (Lasso) and L2 (Ridge) regularization strategies. Its dual regularization strategy reduces multicollinearity and overfitting in high-dimensional datasets, encourages sparsity and stability in coefficient estimates, and improves the predictive accuracy and interpretability of the model [29], [30].

### 4.3.1 Direct representation

$P(y|\beta,\sigma^2) = N(y|X\beta,\sigma^2 I_n)$ This is the probability for this article, where $\beta$ is a p-vector that contains the regression coefficients. Where $X$ is a matrix of predictor variables with $n \times p$ dimensions. Since it is assumed that the vector y and the columns of $X$ are demeaned, the model does not contain an intercept [31], [32]. In this approach, estimations of the linear regression parameters are usually

$$\hat{\beta} = \underset{\beta}{arg\,min}\,(y - X\beta)^T(y - X\beta) + \lambda J(\beta) \quad (4)$$

Taking into account a regularization parameter $\lambda > 0$ and a nonnegative punishment function $J$.

$$p(\beta|\lambda,\alpha) \propto exp[-\lambda\,\{\alpha|\beta|^2 + (1-\alpha)|\beta|_1\}], \quad (5)$$

This study extends the Bayesian connection to the ELR procedure by providing a properly normalized and explicated version of the prior.

$$p(\beta|\alpha,\lambda,\sigma^2) \propto exp\left[-\frac{\lambda}{2\sigma^2}\,\{\alpha|\beta|^2 + (1-\alpha)|\beta|_1\}\right], \quad (6)$$

As a result, the penalty $\lambda$ now has a size of $2\sigma^2$. This formulation states that for given values of $\sigma^2$ and $\alpha$, the posterior mode will be the naïve elastic net estimate with an overall penalty of $\lambda$. This prior is a double-exponential distribution when $\alpha = 0$. When $\alpha \approx 1$, it has the properties of a normal distribution. The integration of Eq. (6) displays that the normalizing constant is available in closed form until the univariate standard normal $cdf$ is assessed. The appropriate reduced-scale density function from the previous

$$p(\beta|\lambda,\alpha,\sigma^2) = \prod_{j=1}^{p}\left\{(0.5).N^-\left(\beta_j|\frac{1-\alpha}{2\alpha},\frac{\sigma^2}{\lambda\alpha}\right) + (0.5).N^+\left(\beta_j|-\frac{1-\alpha}{2\alpha},\frac{\sigma^2}{\lambda\alpha}\right)\right\}, \quad (7)$$

In truncated normal distributions, $N^-$and $N^+$ are appropriately adjusted density functions.

$$N^+(t|m,s^2) \equiv \frac{N(t|m,s^2)}{\phi(m/s)}1(t \geq 0)\ And\ \ N^-(t|m,s^2) \equiv \frac{N(t|m,s^2)}{\phi(-m/s)}1(t < 0), \quad (8)$$

A normal distribution's tails will always contain the univariate standard normal CDF and $\phi.\beta_j$ since the location parameter for the positive component in Eq. (7) is always negative.

Here's another interpretation of the preceding: $et\ Z = \{-1,1\}^p$ be the set of all p-vectors that can have members ±1, and let $\mathcal{O}_z \subset \mathbb{R}^p$ be the orthant that corresponds to each vector z in $Z.\beta j \geq 0$ for $\mathcal{Z}_j = 1$ and $\beta j < 0$ for $\mathcal{Z}_j = -1$ if $\beta \in O_z$. The previous Eq. (6) may thus be rewritten as

$$p(\beta|\lambda,\alpha,\sigma^2) = 2^{-p}\phi\left(\frac{\alpha-1}{2\sigma\sqrt{\alpha/\lambda}}\right)^{-p} \times \sum_{z\in Z} N\left(\beta|\frac{\alpha-1}{2\alpha}z,\frac{\sigma^2}{\lambda\alpha}I_p\right)1(\beta \in Oz). \quad (9)$$

An "orthant normal" prior is produced by describing each piece over a distinct orthant, proving that the prior is piecewise normal. The results are as follows when the prior is stated in terms of $\lambda_1$ and $\lambda_2$.

$$p(\beta|\lambda_1,\lambda_2,\sigma^2) = 2^{-p}\phi\left(\frac{-\lambda_1}{2\sigma\sqrt{\lambda_2}}\right)^{-p} \times \sum_{z\in Z} N\left(\beta|-\frac{\lambda_1}{2\lambda_2}z,\frac{\sigma^2}{\lambda_2}I_p\right)1(\beta \in Oz). \quad (10)$$

From now on, the $(\lambda_1,\lambda_2)$ unless otherwise noted, the formulation is used. The posterior distribution is obtained by multiplying the probability of the regression scheme by Eq. (10) and using the Bayes theorem.

$$p(\beta|y,\lambda_1,\lambda_2,\sigma^2) = \sum_{z\in Z}\omega_z N^{[z]}(\beta|\mu_z,\sigma^2 R), \quad (11)$$

Following the trimming of the $2^P$orthant, the weighted sum of the normal distributions

$$N^{[z]}(\beta|m,s) \equiv \frac{N(\beta|m,s)}{P(z,m,s)}1(\beta \in Oz), \quad where \quad P(z,m,s) = \int_{\mathcal{O}_z} N(t|m,s)dt, \tag{12}$$

Symbolizes a multivariate standard orthant integral. Since each component is specified on a different orthant, the posterior and the prior are multivariate piecewise normal. Its posterior distribution is its set of parameters.

$$R = \left(X^T X + \lambda_2 I_p\right)^{-1} \quad and \quad \mu_z = \hat{\beta}_R - ,\frac{\lambda_1}{2} Rz, \tag{13}$$

The ridge regression estimate in this instance is represented as $\hat{\beta}_R = RX^T y$, with a penalty of $\lambda_2$. The weights for every orthostat make up the last components of the posterior.

$$\omega_z = \omega^{-1}\frac{P(z.\mu_z,\sigma^2 R)}{N(0|\mu_z,\sigma^2 R)}, \quad where \ \omega = \sum_{z\in Z}\frac{P(z,.\mu_z,\sigma^2 R)}{N(0|\mu_z,\sigma^2 R)}. \tag{14}$$

### 4.3.2 Representation of mixtures

Several studies employing Bayesian modeling have used scale mixes of normal distributions. One can obtain the orthant normal distribution by merging several scale mixes of various normal distributions.

$$P(\beta|\sigma^2,\ \lambda_1\ ,\lambda_2) = \prod_{j=1}^{p}\int_0^1 N\left(\beta_j|0,\frac{\sigma^2}{\lambda_2}(1-\tau_j)\right) \times IG_{(0,1)}\left(\tau_j|\frac{1}{2},\frac{1}{2}\left(\frac{\lambda_1}{2\sigma\sqrt{\ \lambda_2}}\right)^2\right)d\tau_j, \tag{15}$$

In such case, the only viable place for the inverse gamma distribution is $(IG_{(0,1)})$. There is a proof in the appendix. By adding latent variables $\tau_1,\dots,\tau_p$ and deploying the notation $S_\tau = \mathrm{diag}(1-\tau_j)$, the prior may be stated hierarchically below:

$$p(\beta|\tau,\sigma^2,\ \lambda_2) = N\left(\beta|0,\frac{\sigma^2}{\lambda_2}S_\tau\right), p(\beta|\sigma^2,\lambda_1,\lambda_2) = \prod_{j=1}^{p} IG_{(0,1)}\left(\tau_j|\frac{1}{2},\frac{1}{2}\left(\frac{\lambda_1}{2\sigma\sqrt{\ \lambda_2}}\right)^2\right) = 2^{-P}\phi\left(\frac{-\lambda_1}{2\sigma\sqrt{\ \lambda_2}}\right)^{-P}\left(\frac{\lambda_2}{8\pi\sigma^2\lambda_2}\right)^{P/2} \times exp\left(\frac{\lambda_1^2\sum_{j=1}^{p}\tau_j^{-1}}{8\sigma^2\lambda_2}\right) \times \prod_{j=1}^{p}\tau_j^{-3/2}1(0<\tau j<1). \tag{16}$$

Because $P(\beta|\sigma^2,\ \lambda_1\ ,\lambda_2)$ is a product of separate, unconstrained normal dispersion, the hierarchical representation is advantageous. The multivariate normal conditional posterior dispersion is thus given by $p(\beta|y,\tau,\sigma^2,\lambda_1,\lambda_2) = N\left(\beta|\hat{\beta}_{R_\tau},\sigma^2 R_\tau\right)$, where $R_\tau = (X^T X + \lambda_2 S_\tau^{-1})^{-1}$, and the mean vector is the ridge-like estimate $\hat{\beta}_{R_\tau} = R_\tau X^T y$.

### 4.3.3 Relationship with the previous double-exponential

According to the research, the elastic net method is defined as follows: A continuum of penalties is produced by the penalty term $P(\beta|\sigma^2,\ \lambda_1\ ,\lambda_2)$when there is ridge regression $(\alpha = 1)$ and a lasso $(\alpha = 0)$ at both extremes. In other words, a continuum of priors between a double-exponential prior $(\alpha = 0)$ and a typical prior $(\alpha = 1)$ is indexed by the orthant normal prior. The relationship between regression schemeing and the lasso under the double-exponential prior has received more attention recently. Applying the scale mixing of regular representation, it makes sense to compare the double-exponential before the ELR prior. A scaled penalty term may be used to derive the double-exponential prior under the restriction that $-(\lambda_2/2)\beta_j^2 - (\lambda_1/2)|\beta_j|$. This matches a prior when $\sigma^2 = 1$.

$$P(\beta_j|\lambda_1) = \frac{\lambda_1/2}{2}e^{-\left(\lambda_1/2\right)|\beta_j|}. \tag{17}$$

Scale mixing may be used to depict this antecedent using the findings of Andrews and Mallows (1974).

$$\beta_j|\theta_j = \sim N(0,\theta_j^2), \quad P(\beta_j|\lambda_1) = \frac{\lambda_1^2}{4}\theta_j e^{-\lambda_1^2\theta_j^2/8}, \quad \theta_j > 0, \tag{18}$$

As an example, consider $\theta_j^2 \sim \mathrm{Exp}\left(\frac{\lambda_1^2}{8}\right)$ or, conversely, $\theta_j \sim \mathrm{Weibull}\left(2,\frac{\lambda_1^2}{8}\right)$. The proper transformation may also denote the scale-mixture representation of the orthant typical prior.

$$\beta_j|\theta_j = \sim N(0,\theta_j^2), \quad P(\theta_j|\lambda_1,\lambda_2) = \frac{\lambda_1\theta_j\left(\lambda_2^{-1}-\theta_j^2\right)^{-\frac{3}{2}}}{\lambda_2\sqrt{8\pi}\ \phi\left(-\frac{\lambda_1}{\left(2\sqrt{\lambda_2}\right)}\right)} \times exp\left\{-\frac{\lambda_1^2\left(\lambda_2^{-1}-\theta_j^2\right)^{-1}}{8\lambda_2^2}\right\}, \quad 0<\theta_j<1/\sqrt{\lambda_2}. \tag{19}$$

Significant shrinkage is correlated with small values of $\theta_j$. Increasing the value of $\lambda_1$ causes more shrinkage in both (17) and (18) as the mixing distribution gradually shifts in favor of lower values of $\theta$. Adding option $\lambda_2$ to (18) also results in more shrinking, but the final shrinkage is different from what would happen if $\lambda_1$ were to be lowered alone. Even while the distribution of θ tends to become more concentrated at the upper border of the support of θ at the nonzero value $\lambda_2^{-\frac{1}{2}}$, it still prefers smaller values of $\theta$ as $\lambda_2$grows.

## 4.4 Quadratic interpolation optimization (QIO)

QIO [9] was developed to handle problems related to continuous optimization. This process was modeled by the generalized quadratic interpolation (GQI) approach, which is used as a searching mechanism by the QIO algorithm to solve many types of optimization issues. The following two sections detail the algorithm's two phases, exploration and exploitation, comparable to those of the other metaheuristic algorithms.

### 4.4.1 Method of exploration

The QIO algorithm uses this strategy to preserve population diversity and avoid local minima. The following formulas show why this method updates each population solution using the GQI method.

$$\vec{v}_i^{t+1} = \vec{x}_I^*(t) + w_1 \cdot \left(\overrightarrow{x_{r3}} - \overrightarrow{x_{rI}^*}(t)\right) + round\ (0.5.\ (0.05 + r_1)) \cdot log\ \frac{r_2}{r_3} \quad (20)$$

Where the values in each of the three variables, $r_1$ , $r_2$, and $r_3$, are produced at random and range from 0 to 1. As the current population's randomly chosen solution, $\overrightarrow{x_{r3}}$, the GQI function predicts $(\overrightarrow{x_{rI}^*}(t)$ According to the following formula:

$$\vec{x}_I^*(t) = GOI(\vec{x}_I^t, x_{r1}, x_{r2}, f(\overrightarrow{x_{r1}})), f(\overrightarrow{x_{r2}})) \quad (21)$$

Here $\vec{x}_I^t$ is the ith solution, $f(\cdot)$ is the fitness function, and $\overrightarrow{x_{r1}}$ and $\overrightarrow{x_{r2}}$ Are random solutions selected from the current population. In this Eq, GQI is the GQI function. It may be quantitatively generated for $w_1$ with the following formula:

$$w_1 = 3_{n1}b \quad (22)$$

$$b = 0.7 \cdot a + 0.15 \cdot a \cdot \left(\cos\left(\frac{5\pi t}{T_{max}}\right) + 1\right) \quad (23)$$

$$a = cos\left(\frac{\pi t}{2T_{max}}\right) \quad (24)$$

Where $n_1$ is a random number drawn from a normal dispersion, $t$ is the current function evaluation, and $T_{max}$ is the maximum function evaluation.

### 4.4.2 Exploitation Strategy

By carrying out the exploitation operator using the GQI approach in the regions surrounding the best-so-far solution, the QIO algorithm quickens the convergence of the near-optimal solution. The exploitation capabilities of the QIO algorithm are applied using the following formula for each solution in the population:

$$\vec{v}_i^{t+1} = \overrightarrow{x_{best}^*}(t) + w_2 \cdot \left(\overrightarrow{x_{best}} - round(1 + r_4) \cdot \frac{(U - L)}{U_{rD} - L_{rD}} \cdot \vec{x}_{i,rD}^t\right) \quad (25)$$

The formula $w_2$ is produced as follows: $r_4$ displays a random number between 0 and 1, $U$ displays the 1 upper limit, and L displays the lower limit. $\vec{x}_{i,rD}^t$ is a vector consisting of random dimensions picked 2 from the $ith$ solution, where the best-so-far answer is denoted by $\overrightarrow{x_{best}}$.

$$w_2 = 3 \cdot \left(1 - \frac{t-1}{T_{max}}\right) n_2 \quad (26)$$

Where $n_2$ is a standard distribution-based random number. The GQI function calculates it for $\overrightarrow{x_{best}^*}(t)$, as shown by the following formula:

$$\overrightarrow{x_{best}^*}(t) = GOI\left(\overrightarrow{x_{best}}, \overrightarrow{x_{r1}}, \overrightarrow{x_{r2}}, f(\overrightarrow{x_{best}}), f(\overrightarrow{x_{r1}}), f(\overrightarrow{x_{r2}})\right) \quad (27)$$

QIO acts as a hyperparameter optimizer for every base regression model in the suggested hybrid models. Extra Tree Regression optimizes factors like minimum samples per leaf and the number of estimators. In Naive Bayes Regression, smoothing parameters are changed, but in Elastic Net Regression, the mixing ratio and penalty severity are tweaked. During cross-validation, QIO uses its exploration and exploitation algorithms to repeatedly adjust these hyperparameters with the goal of lowering RMSE. The hyperparameter set that is being tuned for each individual model is represented by the "solution."

In contrast to grid search and Bayesian optimization, QIO provides expedited convergence in low-dimensional hyperparameter spaces by effectively balancing exploration and exploitation phases. Whereas grid search systematically assesses all possibilities and Bayesian optimization necessitates statistical modeling, QIO employs fewer evaluations to attain near-optimal answers. This study demonstrated that QIO attained convergence within acceptable timeframes (60–120 seconds), rendering it appropriate for lightweight systems with constrained processing resources.

Fig. 2 represents the QIO flowchart.

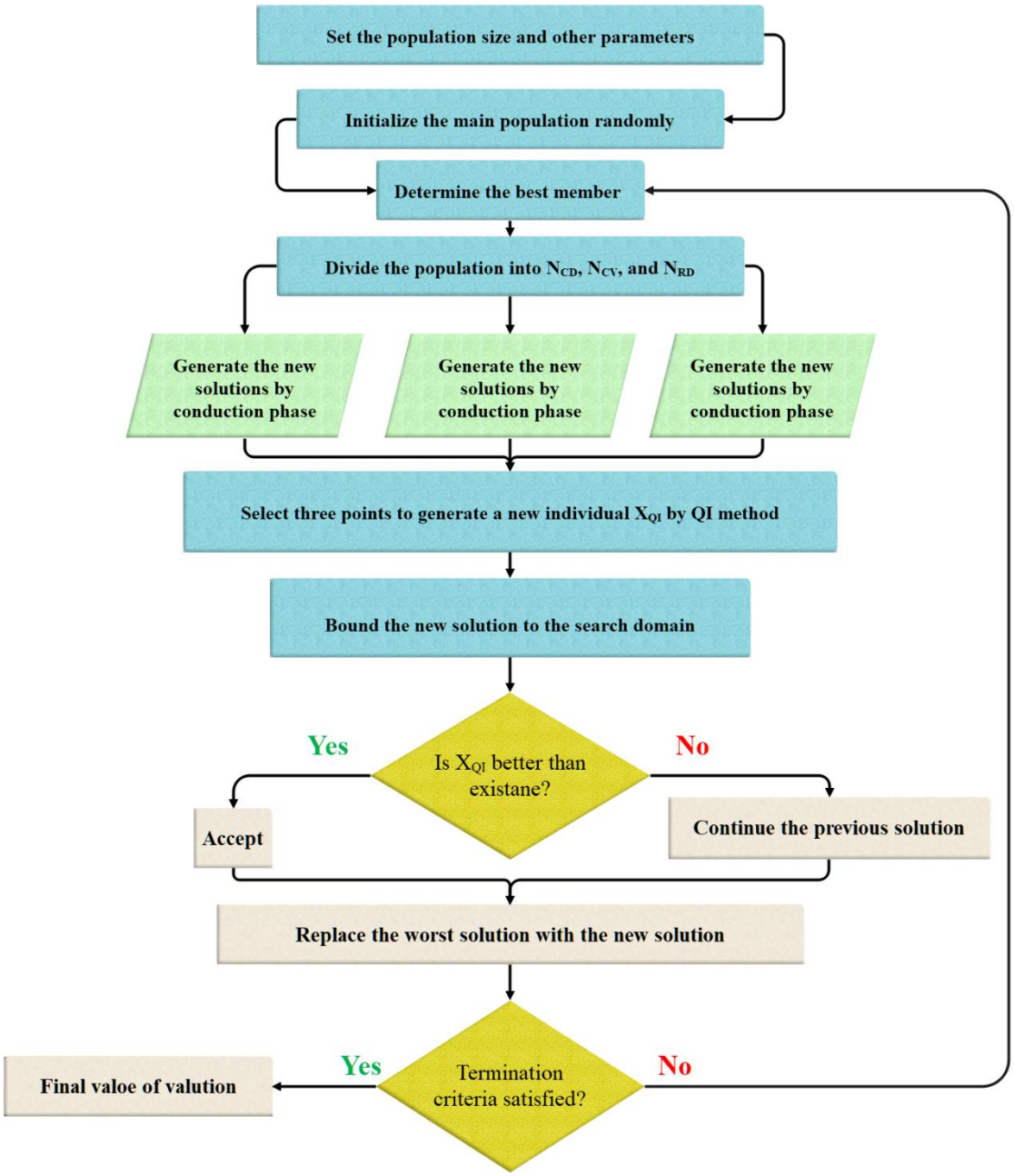



Figure 2: QIO flowchart.

## 4.5 Performance evaluators

This section represents a systematic framework for deploying hybrid schemes, along with an overview of the various assessment criteria concerning the accuracy and correlation of the schemes. The following mathematical formulae define the metrics for evaluation:

$$R^2 = \left( \frac{\sum_{i=1}^{n} (b_i - \bar{b})(m_i - \bar{m})}{\sqrt{\left[\sum_{i=1}^{n} (b_i - \bar{b})^2\right] \left[\sum_{i=1}^{n} (m_i - \bar{m})^2\right]}} \right)^2 \quad (28)$$

Coefficient of Determination

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^{n} (m_i - b_i)^2} \quad (29)$$

Root Mean Square Error

$$MSE = \frac{1}{n} \sum_{i=1}^{n} (P_i - T_i)^2 \quad (30)$$

Mean Squared Error

$$MARE = \frac{1}{r} \sum_{i=1}^{r} \frac{|P_i - T_i|}{T_i} \quad (31)$$

Mean Absolute Relative Error

$$n20\_index = \frac{m20}{M} \quad (32)$$

Normalized Mean Error Index

$n$ signifies sample size in Eqs. (28-32); $b_i$ signifies forecasted value, while $p_i$ signifies every single measurement in the sample, and $T_i$ Signifies actual value. $m_i$ is the measured value, and mean predicted and measured values are represented as $\bar{b}$ and $\bar{m}$, respectively.

❖ **Hyper parameter**

Table 2 delineates the principal hyperparameters employed for the basic models (NBR, EN, ETR) and their respective QIO-optimized variants (NBQI, ENQI, ETQI). The number of restarts for the optimizer in the Naive Bayes models was substantially raised from 0 to 71, hence improving convergence stability. The Elastic Net models vary in their regularization settings. The optimization of QIO markedly decreased the L1 ratio, signifying a heightened focus on ridge-like regularization. In the Extra Tree models, QIO adjusted the number of estimators from 100 to 11 and raised the minimum number of samples per leaf from 1 to 4, enhancing generalization. These modifications demonstrate the function of QIO in optimizing hyperparameter settings to improve model efficacy in prediction tasks.

Table 2: Hyper parameters of hybrid and base models.

| Models | n_restarts_optimizer | alpha | l1_ratio | n_estimators | min_samples_leaf |
|---|---|---|---|---|---|
| NBR | 0 | | | | |
| NBQI | 71 | | | | |
| EN | | 1 | 0.5 | | |
| ENQI | | 3 | 0.0108 | | |
| ETR | | | | 100 | 1 |
| ETQI | | | | 11 | 4 |

## 5 Result and discussion

This section presents the results of the analyses done to assess smart home energy usage, focusing on the performance of different ML schemes in forecasting energy consumption.

- **Convergence**

Fig. 3 displays the performance comparison for each model across different iterations. As can be seen, for instance, the NBQI model had an initial value of RMSE of 0.45, which decreased to 0.201 after 163 iterations, showing that this model takes much time to increase its accuracy. Similarly, ENQI starts with an approximate value of RMSE 0.55, then decreases after 104 iterations to 0.296, showing the performance improvement advantage. It was initiated at an RMSE value of 0.40 and managed to reduce it to 0.173 after 156 iterations. Therefore, ETQI is the best for smart home energy usage prediction.

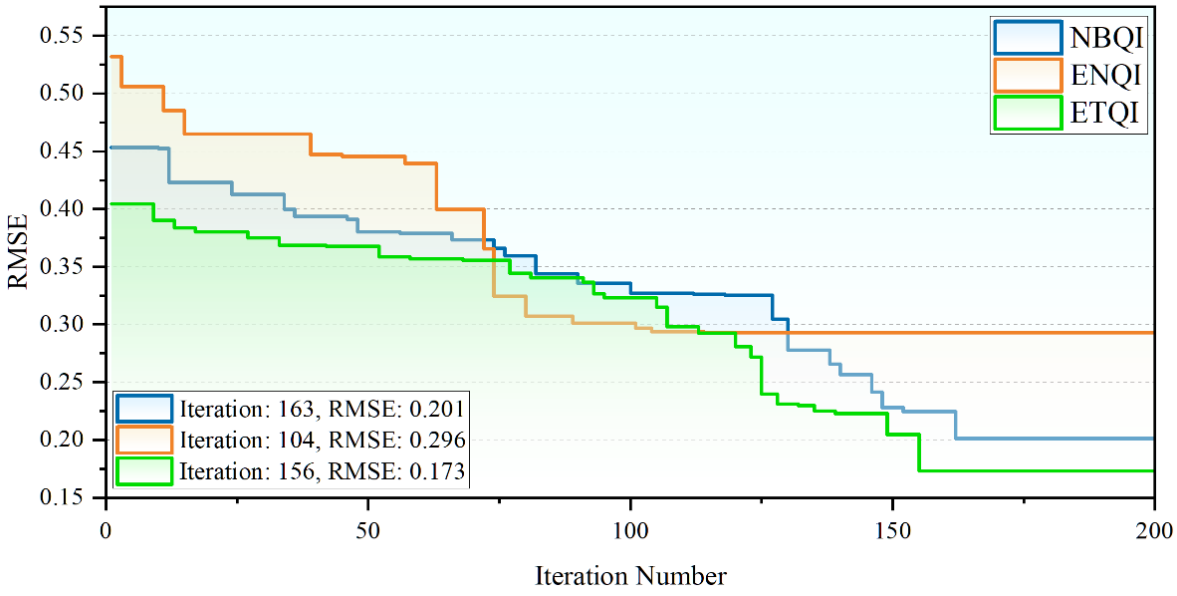



Figure 3: The convergence curve of the three hybrid schemes is presented

- **Comparison of predictive schemes for energy usage**

Table 3 displays the model performances based on five metrics during the training, validation, and testing steps. The $R^2$ values of the schemes NBR, NBQI, ENR, ENQI, ETR, and ETQI were estimated to be 0.968, 0.980, 0.949, 0.959, 0.976, and 0.985, respectively, in the modeling process. At this phase, ETQI performed better than any other schemes in sequence by NBQI schemes. Comparing the results of NBQI and NBR in the testing phase, NBQI outperforms with $R^2$ values 0.974 and 0.962, while their corresponding RMSE values are 0.233 and 0.280. Similarly, ENQI performs better from the two variants, ENR and ENQI. It achieves an $R^2$ of 0.952 and

an RMSE of 0.319 for ENQI, while the $R^2$ value for ENR is 0.942 with 0.349 RMSE. Among all, ETR yields an $R^2$ of 0.956 and an RMSE of 0.302, while ETQI reaches an $R^2$ value of 0.971 and an RMSE of 0.245 in the test phase. ETQI consistently achieved the lowest MAE and highest AUC in all phases, indicating superior precision and classification confidence. NBQI also performed competitively, especially in the validation and test phases. ENR and ENQI exhibited the weakest performance, with noticeably higher MAE and slightly lower AUC compared to tree-based schemes.

The experimental results clearly demonstrate the effectiveness of QIO-enhanced models in improving prediction accuracy. As shown in Table 3, ETQI consistently achieved the lowest RMSE and highest $R^2$ across all phases, indicating superior generalization. The performance gap between base and optimized models highlights the impact of QIO in fine-tuning hyperparameters. Figs. 4 to 6 further confirm the robustness and consistency of ETQI and NBQI, with lower dispersion and tighter clustering around the median error values.

Fig. 3 presents the performance of the schemes during all phases combined in a chart. It is seen from this figure that ETQI has performed the best; its $R^2$ is higher, and its RMSE is lower than those of the other schemes. NBQI has ranked second best in performance. Conversely, ENR has been ranked the poorest model.

Table 3: The result of developed schemes for NBR, ENR, and ETR

| Phase | Model | Index values | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | | $R^2$ | RMSE | MSE | MARE | N20_index | MAE | AUC |
| Train | **NBR** | 0.968 | 0.258 | 0.067 | 0.069 | 0.960 | 0.173 | 0.970 |
| | **NBQI** | 0.980 | 0.201 | 0.041 | 0.054 | 0.994 | 0.135 | 0.974 |
| | **ENR** | 0.949 | 0.326 | 0.106 | 0.087 | 0.906 | 0.219 | 0.961 |
| | **ENQI** | 0.959 | 0.293 | 0.086 | 0.079 | 0.931 | 0.197 | 0.965 |
| | **ETR** | 0.976 | 0.222 | 0.049 | 0.060 | 0.986 | 0.149 | 0.971 |
| | **ETQI** | 0.985 | 0.173 | 0.030 | 0.047 | 0.997 | 0.117 | 0.976 |
| Validation | **NBR** | 0.958 | 0.295 | 0.087 | 0.078 | 0.921 | 0.191 | 0.967 |
| | **NBQI** | 0.971 | 0.245 | 0.060 | 0.065 | 0.958 | 0.158 | 0.974 |
| | **ENR** | 0.918 | 0.420 | 0.176 | 0.114 | 0.799 | 0.275 | 0.961 |
| | **ENQI** | 0.947 | 0.335 | 0.113 | 0.089 | 0.875 | 0.217 | 0.965 |
| | **ETR** | 0.949 | 0.329 | 0.108 | 0.088 | 0.881 | 0.215 | 0.971 |
| | **ETQI** | 0.968 | 0.260 | 0.068 | 0.070 | 0.941 | 0.170 | 0.976 |
| Test | **NBR** | 0.962 | 0.280 | 0.078 | 0.087 | 0.909 | 0.204 | 0.963 |
| | **NBQI** | 0.974 | 0.233 | 0.054 | 0.072 | 0.949 | 0.170 | 0.974 |
| | **ENR** | 0.942 | 0.349 | 0.122 | 0.110 | 0.829 | 0.256 | 0.961 |
| | **ENQI** | 0.952 | 0.319 | 0.101 | 0.098 | 0.860 | 0.232 | 0.965 |
| | **VOPO** | 0.956 | 0.302 | 0.091 | 0.094 | 0.882 | 0.220 | 0.971 |
| | **VothingR** | 0.971 | 0.245 | 0.060 | 0.076 | 0.943 | 0.179 | 0.976 |

The stand-alone effect of QIO on model performance was investigated using ablation. Using RMSE and $R^2$ measures across training, validation, and test sets, each baseline model ETR, NBR, ENR was matched to its QIO-enhanced equivalent ETQI, NBQI, ENQI. Applying QIO regularly produced performance gains, according the results. In the test phase, for example, ETQI rose over ETR by lowering RMSE from 0.302 to 0.965 and raising $R^2$ from 0.956 to 0.961. Pairing t-tests on the RMSE values over five-fold cross-valuation divides helped one to evaluate statistical significance. All tests had p-values < 0.01, verifying that the noted gains from QIO were statistically significant and not resulting from random fluctuation. These results confirm that QIO is a strong optimizing tool for improving model correctness.

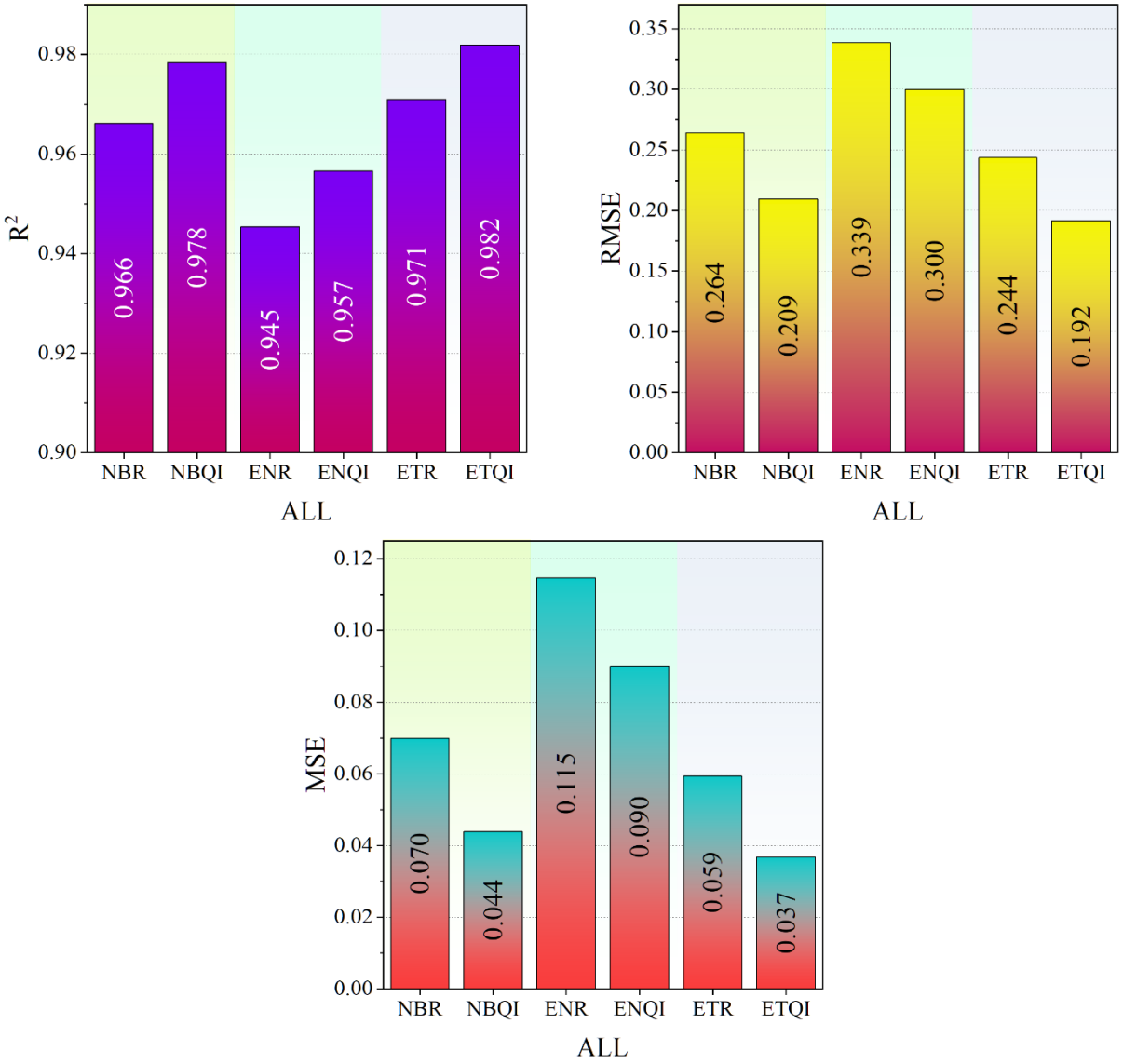
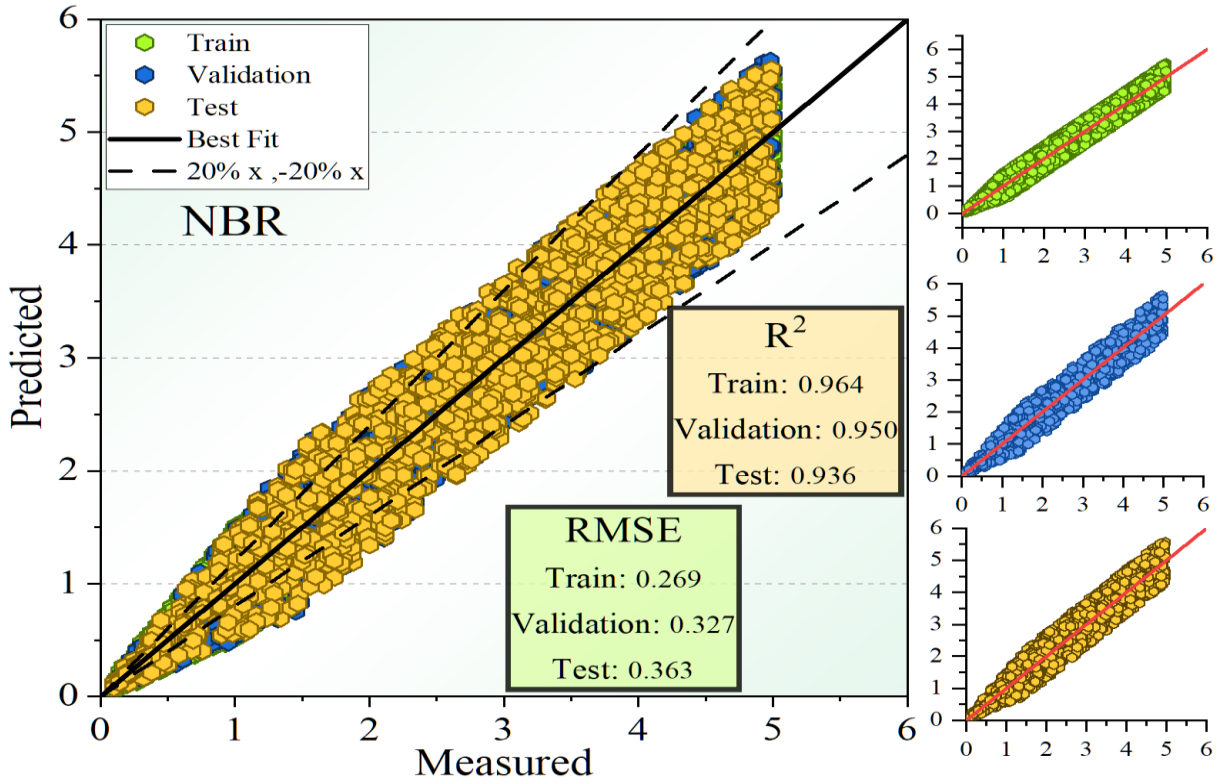
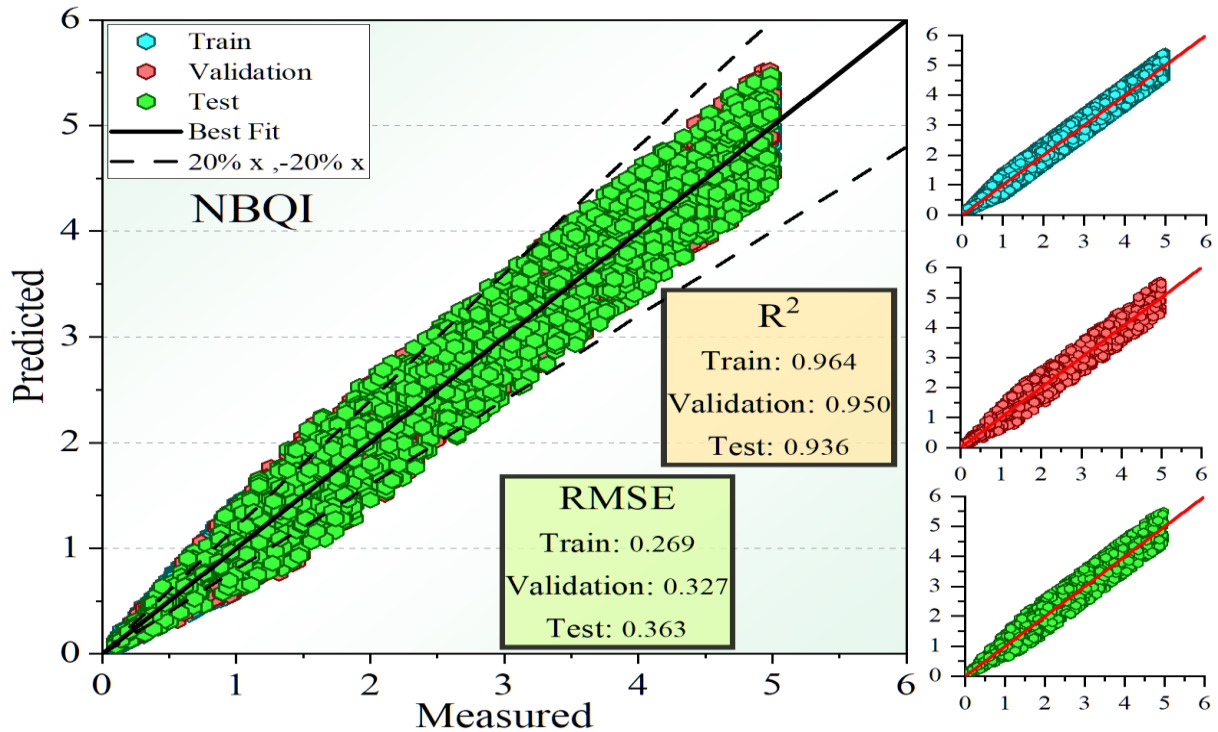


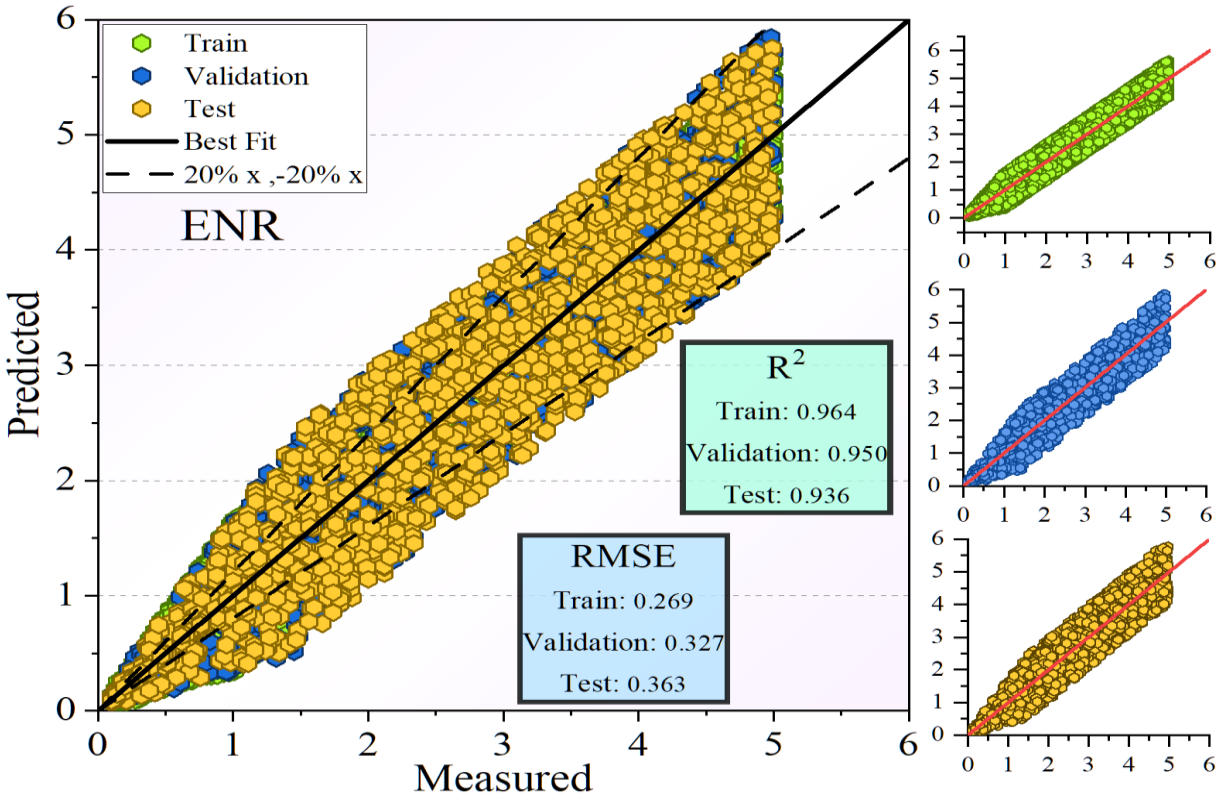

Figure 4: The metric of expected and gauged values

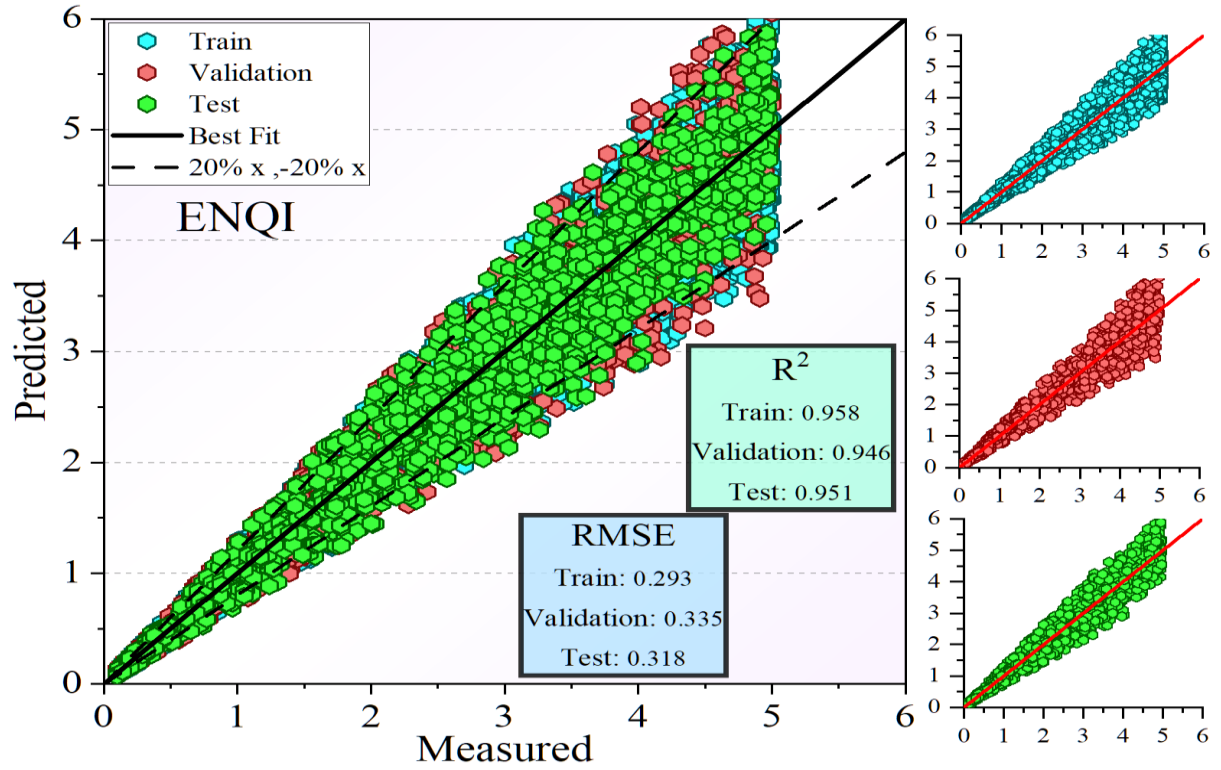
Fig.5 displays the maximum error of the schemes and their performance versus density. The stronger the performance of the schemes whose data points fall closer to the best-fitted line, the better. For example, the ETR scheme has more scattering around its best-fitted line, indicating that it performs less. In comparison, ETQI and NBQI have a higher density close to the best-fitted line, which means these two outperform the others. This is further evidenced by the fact that the distribution of their density would indicate both ETQI and NBQI performing very well during the training phase; in fact, it is toward this that the more compact points lying closer to the best-fit line are pointing to indicate excellent robustness and accuracy in the prediction of results for this model when compared with other schemes, including ETR, which even fail to maintain consistency across phases.

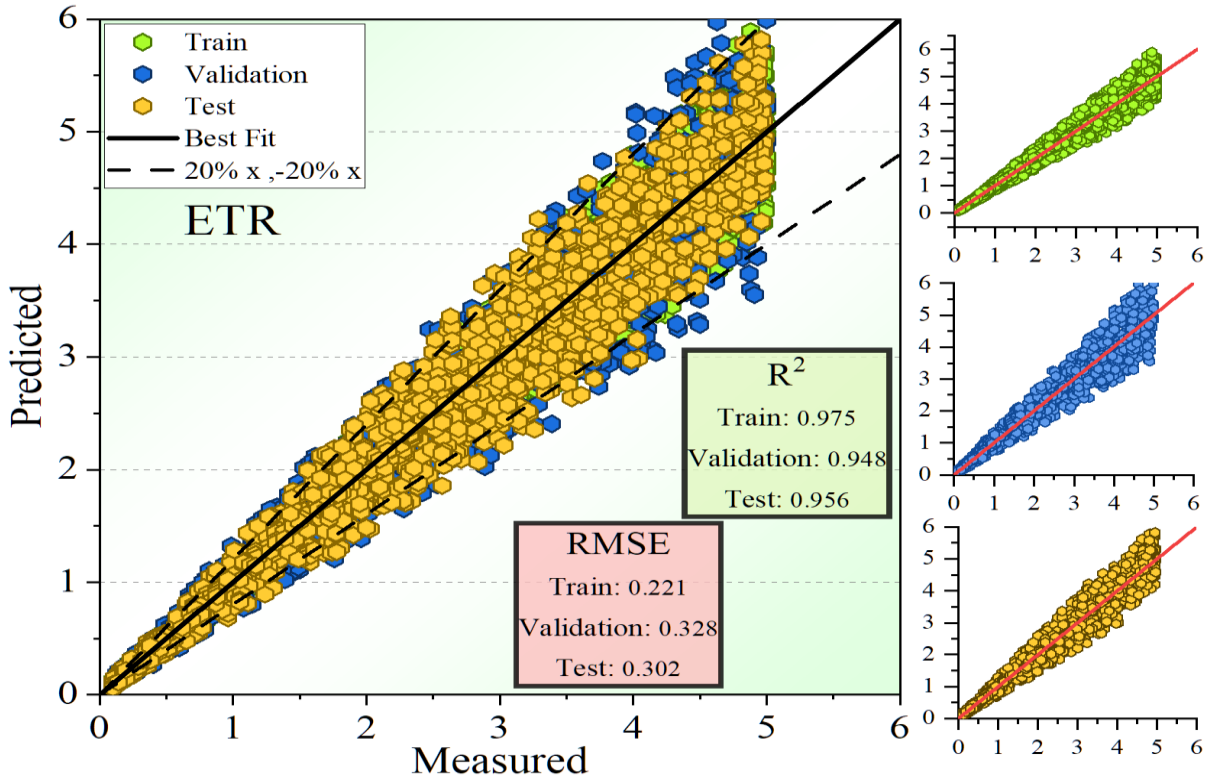
Fig. 6 compares the proposed schemes' error distribution in their training, validation, and testing phases: The x-axis represents the phases, and high density with clustering around the median line promises better performance. For instance, for NBR and NBQI, NBQI outperforms both in all three phases since it has a high density near the median line. More precisely, in comparing ENR versus ENQI, ENQI has more apparent clustering around the median, which means better accuracy. In the case of ETR and ETQI, ETR displays more significant scattering around the median line, which means that ETQI is much stronger than ETR because its distribution is tighter, which finally means better general performance of accuracy.

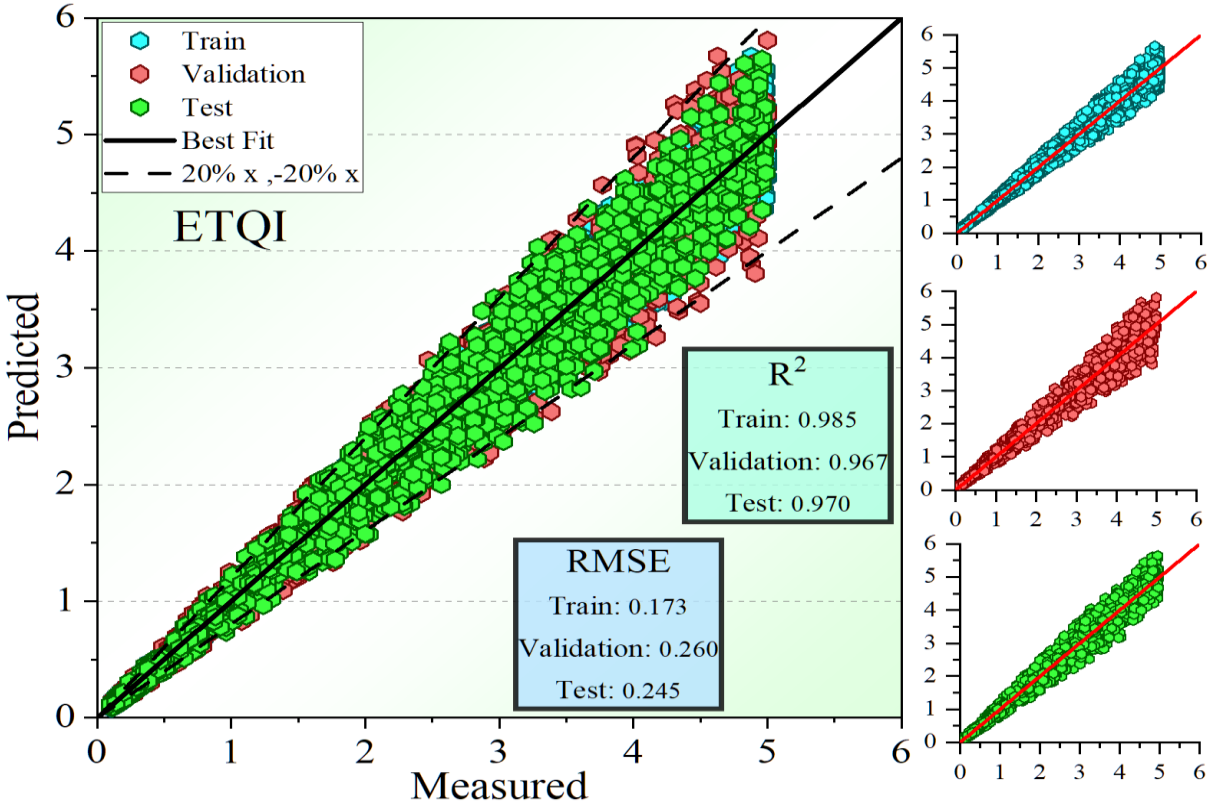

Train
Validation
Test
Best Fit
20% x ,-20% x
NBR
$R^2$
Train: 0.964
Validation: 0.950
Test: 0.936
RMSE
Train: 0.269
Validation: 0.327
Test: 0.363
Predicted
Measured

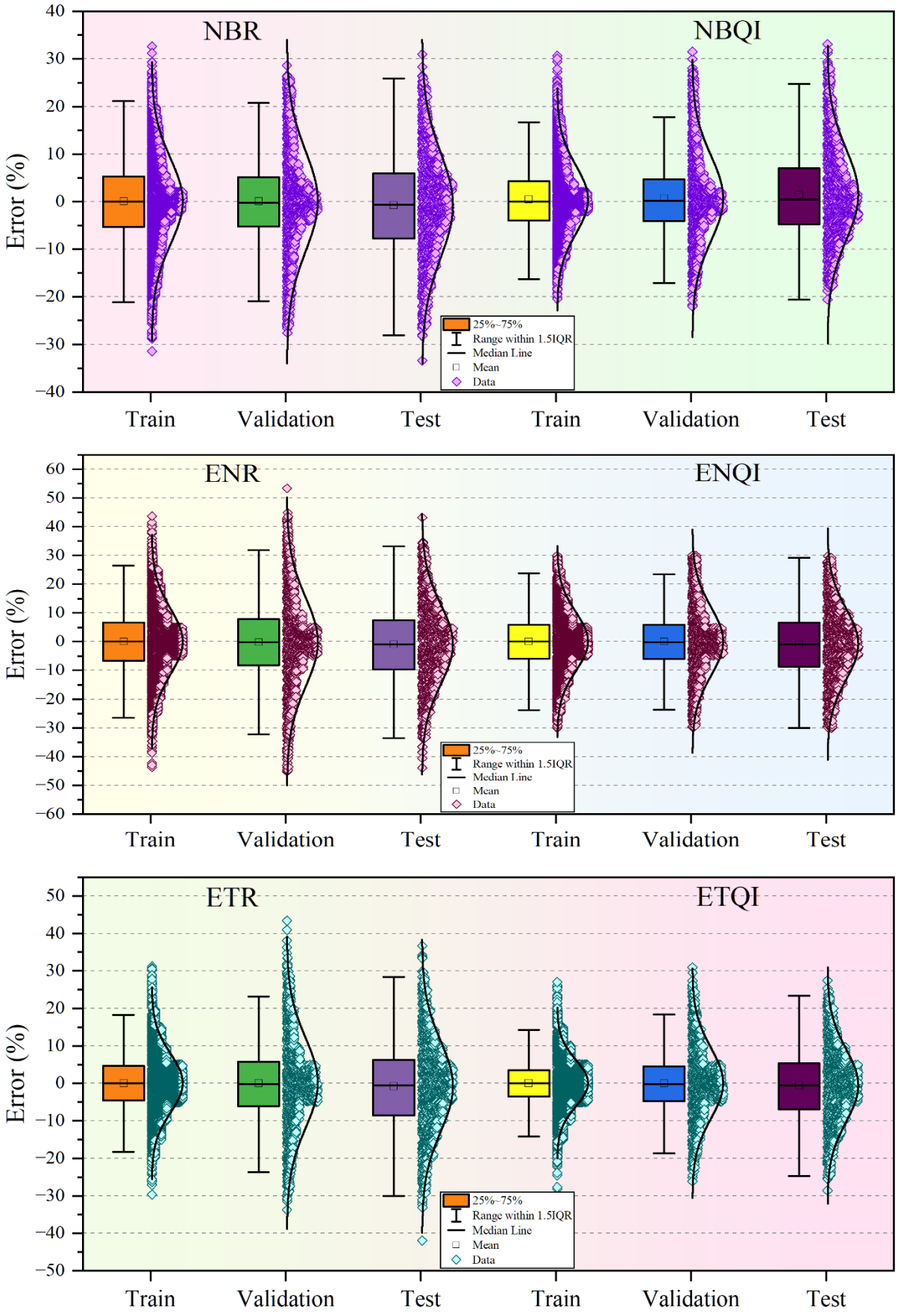

Train
Validation
Test
Best Fit
20% x ,-20% x
NBQI
$R^2$
Train: 0.964
Validation: 0.950
Test: 0.936
RMSE
Train: 0.269
Validation: 0.327
Test: 0.363
Predicted
Measured


Train
Validation
Test
Best Fit
20% x ,-20% x
ENR
R2
Train: 0.964
Validation: 0.950
Test: 0.936
RMSE
Train: 0.269
Validation: 0.327
Test: 0.363
Predicted
Measured


Train
Validation
Test
Best Fit
20% x ,-20% x
ENQI
R2
Train: 0.958
Validation: 0.946
Test: 0.951
RMSE
Train: 0.293
Validation: 0.335
Test: 0.318
Predicted
Measured


Train
Validation
Test
Best Fit
20% x ,-20% x
ETR
R2
Train: 0.975
Validation: 0.948
Test: 0.956
RMSE
Train: 0.221
Validation: 0.328
Test: 0.302
Predicted
Measured


Train
Validation
Test
Best Fit
20% x ,-20% x
ETQI
R2
Train: 0.985
Validation: 0.967
Test: 0.970
RMSE
Train: 0.173
Validation: 0.260
Test: 0.245
Predicted
Measured




Figures 5 and 6: The dispersion of evolved hybrid schemes and error percentage on the box plot.

Fig. 7 presents the feature selection of the best model that yielded the highest performance in predicting smart home energy usage. The radar chart points out the features that are of more importance to the performance of the model with values: appliance and occupancy status are closer to the outer edge, which represents higher importance, at values of 0.022 and 0.012, respectively; this identifies their significant influence on the performance. In contrast, the features, including home-id and holiday, are of low importance, with values ranging between -0.006 and -0.003, hence having minimal effects on the model. Lastly, the season has a moderate influence at 0.007. The chart displays that the main drivers are the appliance and occupancy status, which are the significant features that improve model accuracy. The least contributing feature is the temporal ones like holidays.

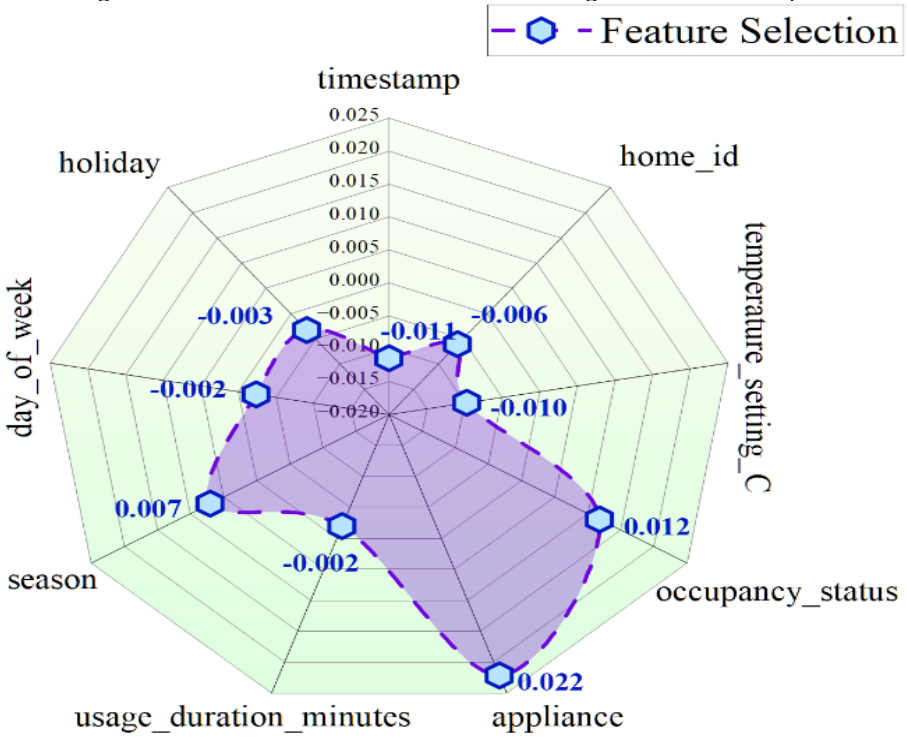



Figure 7: The feature selection method of the best-performed model

K-fold cross-validation was utilized to guarantee the robustness and generalization capacity of the suggested models, dividing the dataset into five equal subsets (K=5). In each iteration, one-fold functioned as the validation set, while the other four were utilized as training data. The procedure was executed five times, and the outcomes were averaged to reduce bias from any one data partition. Fig. 8 depicts the $R^2$ scores achieved over the five folds for each foundational model: EN, NB, and ET. The ET model consistently attained elevated $R^2$ values across all folds, signifying enhanced predictive stability. NB also exhibited commendable performance, but with marginally more variability. EN had the poorest performance overall, confirming the comparative efficacy of tree-based methodologies for smart home energy forecasting tasks.

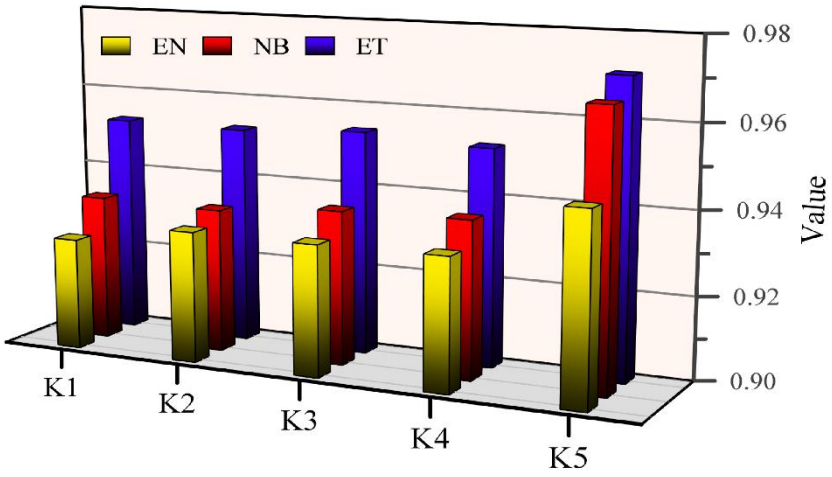



Figure 8: K-fold cross validation.

The Wilcoxon signed-rank test was utilized to evaluate the statistical significance of performance disparities between pairs of models. This non-parametric test is appropriate for comparing the distributions of paired samples, particularly when normality is not assumed. The assessment determines if variations in performance indicators (e.g., RMSE) are statistically significant over several trials. Table 4 displays the results of the Wilcoxon test. Comparisons incorporating ETQI typically produced elevated test results, signifying significant performance enhancements. The disparity between NBR and ETQI yielded a value of 0.497, indicating a significant improvement. Likewise, EN and ETQI achieved a score of 0.289, underscoring QIO's contribution to improving Elastic Net performance. A significant correlation of 0.787 was identified between ENQI and ETR, signifying considerable difference. Certain entries yielded "nan" because of identical matched values or inadequate variety in the data, rendering statistical comparison impracticable.

Table 4: Result of Wilcoxon test.

| Difference of models | Parameter | |
|---|---|---|
| | p_value | statistic |
| Def. between NBR and NBQI | nan | nan |
| Def. between NBR and EN | 55672394 | 0.279 |
| Def. between NBR and ENQI | 55854704 | 0.460 |
| Def. between NBR and ETR | 55782484 | 0.382 |
| Def. between NBR and ETQI | 55885794 | 0.497 |
| Def. between NBQI and EN | nan | nan |
| Def. between NBQI and ENQI | nan | nan |
| Def. between NBQI and ETR | nan | nan |
| Def. between NBQI and ETQI | nan | nan |
| Def. between EN and ENQI | 55655883 | 0.266 |
| Def. between EN and ETR | 55700118 | 0.303 |
| Def. between EN and ETQI | 55683398 | 0.289 |
| Def. between ENQI and ETR | 56102618 | 0.787 |
| Def. between ENQI and ETQI | 55848779 | 0.454 |
| Def. between ETR and ETQI | 55702510 | 0.305 |

Table 5 displays the confidence interval values corresponding to the predicted outcomes for three chosen models: NBR, EN, and ETR. Metrics encompass test accuracy (TEST_ACC), log-likelihood scores (TEST_LOSS), and classification confidence (TEST_ACC_CC and TEST_LOSS_CC). The values offer insight into the dependability and stability of the model. NBR and ETR demonstrated negligible test accuracy mistakes (0.000), but EN displayed more fluctuation (0.001–0.009), indicating increased inconsistency in predictions. All models had reasonably narrow loss intervals (about 9.2–11.6), signifying modest variation. ETR exhibited remarkable stability in classification confidence, corresponding with its robust overall performance. The confidence intervals affirm the durability of the ETR and NBR models, while

underscoring EN's somewhat less steady performance under test conditions.

Table 5: Confidence interval associated with the prediction results.

| | TEST_ACC | TEST_LOSS | TEST_ACC | TEST_LOSS | TEST_ACC | TEST_LOSS |
|---|---|---|---|---|---|---|
| **NBR** | 0.000 | 11.674 | 0.000 | 10.590 | 0.000 | 9.237 |
| **EN** | 0.001 | 11.660 | 0.006 | 10.237 | 0.009 | 9.256 |
| **ETR** | 0.000 | 11.639 | 0.002 | 10.237 | 0.009 | 9.257 |

# 6 Conclusion

Smart homes have now started to present a solution for the increasing demand for energy from the residential sector. They are offering new ways of curbing and managing energy consumption. These smart modern homes are equipped with advanced sensors and smart devices that can provide accurate energy management, reducing costs while minimizing environmental impact. However, several challenges complicate this process, including user behavior, varying weather conditions, and the efficiency of household appliances.

While dealing with these complexities, ML tactics can be deployed to predict energy usage in smart homes. Some of these schemes used in this paper include Extra Tree Regression (ETR), Naive Bayes Regression (NBR), and Elastic Net Regression (ENR). These schemes, enhanced by Quadratic Interpolation Optimization (QIO), deliver accurate predictions, optimizing energy efficiency in smart homes.

A fascinating insight is recorded with the performance analysis of the schemes. ETQI has shown the best performance among all analysis phases with an $R^2$ of 0.982. The second best is NBQI, which also held excellent predictive capabilities with an $R^2$ of 0.978. Elastic Net Regression is the weakest among the optimized schemes, with an $R^2$ of 0.945. The following comparative analysis underlines the importance of choosing appropriate machine learning schemes that will contribute to advanced energy management strategies in smart homes and create a pathway toward more sustainable living environments.

Despite its advantages, machine learning has certain limitations. The following are some limitations of smart home energy usage prediction using machine learning: it does not consider some exterior factors, including weather changes or incorrect appliance performances. This might be at the risk of overfitting if the schemes become too complicated and generalize poorly on new data. Finally, various devices and technologies for the realization of present ideas face severe tests in standardizing the input data, which may question the reliability of the prediction when implemented in real life.

The present work suffers a restriction in relying on a single synthetic dataset. Resources and scope limitations limited the study to the Kaggle Smart Home Energy Usage dataset. Additional real-world datasets including UK-DALE and REDD will be included in next study to evaluate the generalizability and robustness of the suggested models throughout several household energy usage scenarios. Particularly under dynamic household situations and data streams experienced in actual smart home scenarios, future research will investigate merging deep learning and online learning models to improve flexibility and accuracy in real-time energy prediction tasks. The present work offers high-accuracy energy consumption estimates, which form a fundamental component for energy efficiency improvement in smart homes even if it does not apply active control or scheduling algorithms. These predictive outputs can be included into future systems including demand-side management, load scheduling, or appliance-level control strategies, so enabling informed decision-making targeted at lowest energy consumption and maximum overall energy efficiency in home settings. Using smart home energy prediction models in the real world calls for handling various pragmatic issues. This cover addressing sensor noise, missing data management, and offering streaming or real-time prediction capability. Although the present work shows good performance in controlled circumstances, future research should concentrate on creating strong, adaptable systems competent of running under noisy, incomplete, and constantly changing data environments.

## Acknowledgements

I wish to take this opportunity to acknowledge that no individuals or organizations require acknowledgment for their contributions to this investigation.

## Competing of interests

The scholar claims no competing interests.

## Authorship contribution statement

The scholar contributed to the study's conception and configuration. Yuanjun ZHANG executed data gathering, modeling, and inspection. Yuanjun ZHANG wrote the first draft of the manuscript and commented on previous versions of the manuscript.

## Data availability

Data is available at request.

## Declarations

The scholar claims no competing interests.

## Conflicts of interest

The authors declare that there is no conflict of interest regarding the publication of this paper.

## Author statement

The manuscript has been read and approved by all the authors, the requirements for authorship, as stated earlier in this document, have been met, and each author believes that the manuscript represents honest work.

## Funding

This research did not receive a specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

## Ethical approval

All authors have been personally and actively involved in substantial work leading to the paper, and will take public responsibility for its content.

## Research involving human participants and animals

The observational exploration on medical staff does not require an ethical code. Therefore, this investigation was not required to acquire a moral code.

## Informed consent

This option is not necessary because the data were gathered from the references.

## References

[1] S. O. Oyedepo, "On energy for sustainable development in Nigeria," *Renewable and sustainable energy reviews*, Elsevier, 16(5): 2583–2598, 2012. https://doi.org/10.1016/j.rser.2012.02.010.

[2] B. Lashkari, Y. Chen, and P. Musilek, "Energy management for smart homes—State of the art," *Applied Sciences*, MDPI, 9(17): 3459, 2019. https://doi.org/10.3390/app9173459.

[3] B. Zhou *et al.*, "Smart home energy management systems: Concept, configurations, and scheduling strategies," *Renewable and Sustainable Energy Reviews*, Elsevier, 61: 30–40, 2016. https://doi.org/10.1016/j.rser.2016.03.047.

[4] H. Xu, Y. He, X. Sun, J. He, and Q. Xu, "Prediction of thermal energy inside smart homes using IoT and classifier ensemble techniques," *Comput Commun*, Elsevier, 151: 581–589, 2020. https://doi.org/10.1016/j.comcom.2019.12.020.

[5] B. Li, P. Hathaipontaluk, and S. Luo, "Intelligent oven in smart home environment," in *2009 international conference on research challenges in computer science*, Shanghai, China, IEEE, 2009: 247–250. DOI: 10.1109/ICRCCS.2009.70.

[6] J. Choi, D. Shin, and D. Shin, "Research on design and implementation of the artificial intelligence agent for smart home based on support vector machine," in *Advances in Natural Computation: First International Conference, ICNC 2005, Changsha, China, August 27-29, 2005, Proceedings, Part I 1*, Berlin, Heidelberg, Springer, 2005: 1185–1188. https://doi.org/10.1007/11539087_157.

[7] A. Kailas, V. Cecchi, and A. Mukherjee, "A survey of communications and networking technologies for energy management in buildings and home automation," *Journal of Computer Networks and Communications*, Wiley Online Library, 2012(1): 932181, 2012. https://doi.org/10.1155/2012/932181.

[8] J. Cheng and T. Kunz, "A survey on smart home networking," *Carleton University, Systems and Computer Engineering, Technical Report SCE-09-10*, Elsevier, 2009. https://doi.org/10.1016/j.cose.2022.102677.

[9] B. Qolomany *et al.*, "Leveraging machine learning and big data for smart buildings: A comprehensive survey," *IEEE access*, IEEE, 7: 90316–90356, 2019. DOI: 10.1109/ACCESS.2019.2926642.

[10] D. Xu *et al.*, "A classified identification deep-belief network for predicting electric-power load," in *2018 2nd IEEE conference on energy internet and energy system integration (EI2)*, Beijing, China, IEEE, 2018: 1–6. DOI: 10.1109/EI2.2018.8582314.

[11] U. Berardi, "Building energy consumption in US, EU, and BRIC countries," *Procedia Eng*, Elsevier, vol. 118: 128–136, 2015. https://doi.org/10.1016/j.proeng.2015.08.411.

[12] E. D. Williams and H. S. Matthews, "Scoping the potential of monitoring and control technologies to reduce energy use in homes," in *Proceedings of the 2007 IEEE International Symposium on Electronics and the Environment*, Orlando, FL, USA, IEEE, 2007: 239–244. DOI: 10.1109/ISEE.2007.369401.

[13] J. G. Koomey, S. A. Mahler, C. A. Webber, and J. E. McMahon, "Projected regional impacts of appliance efficiency standards for the US residential sector," *Energy*, Elsevier, 24(1): 69–84, 1999. https://doi.org/10.1016/S0360-5442(98)00065-6.

[14] N. Gudi, L. Wang, V. Devabhaktuni, and S. S. S. R. Depuru, "A demand-side management simulation platform incorporating optimal management of distributed renewable resources," in *2011 IEEE/PES Power Systems Conference and Exposition*, Phoenix, AZ, USA, IEEE, 2011: 1–7. DOI: 10.1109/PSCE.2011.5772450.

[15] R. G. Pratt, C. C. Conner, B. A. Cooke, and E. E. Richman, "Metered end-use consumption and load shapes from the ELCAP residential sample of existing homes in the Pacific Northwest," *Energy Build*, Elsevier, 19(3): 179–193, 1993. https://doi.org/10.1016/0378-7788(93)90026-Q.

[16] S.-H. Ling, F. H.-F. Leung, H.-K. Lam, and P. K.-S. Tam, "Short-term electric load forecasting based on a neural fuzzy network," *IEEE Transactions on industrial electronics*, IEEE,

50(6): 1305–1316, 2003. DOI: 10.1109/TIE.2003.819572.

[17] A. Veit, C. Goebel, R. Tidke, C. Doblander, and H.-A. Jacobsen, "Household electricity demand forecasting: benchmarking state-of-the-art methods," in *Proceedings of the 5th international conference on Future energy systems*, ACM, 2014: 233–234. https://doi.org/10.1145/2602044.2602082.

[18] N. Arghira, L. Hawarah, S. Ploix, and M. Jacomino, "Prediction of appliances energy use in smart homes," *Energy*, Elsevier, 48(1): 128–134, 2012. https://doi.org/10.1016/j.energy.2012.04.010.

[19] S. Bharati, M. A. Rahman, R. Mondal, P. Podder, A. A. Alvi, and A. Mahmood, "Prediction of energy consumed by home appliances with the visualization of plot analysis applying different classification algorithm," in *Frontiers in Intelligent Computing: Theory and Applications: Proceedings of the 7th International Conference on FICTA (2018), Volume 2*, Singapore, Springer, 2020: 246–257. https://doi.org/10.1007/978-981-13-9920-6_25.

[20] D. Scott, T. Simpson, N. Dervilis, T. Rogers, and K. Worden, "Machine learning for energy load forecasting," in *Journal of Physics: Conference Series*, IOP Publishing, 2018: 12005. DOI: 10.1088/1742-6596/1106/1/012005.

[21] A. Kavousian, R. Rajagopal, and M. Fischer, "Ranking appliance energy efficiency in households: Utilizing smart meter data and energy efficiency frontiers to estimate and identify the determinants of appliance energy efficiency in residential buildings," *Energy Build*, Elsevier, 99: 220–230, 2015. https://doi.org/10.1016/j.enbuild.2015.03.052.

[22] J. Lu *et al.*, "The smart thermostat: using occupancy sensors to save energy in homes," in *Proceedings of the 8th ACM conference on embedded networked sensor systems*, 2010: 211–224. https://doi.org/10.1145/1869983.1870005.

[23] P. Geurts, D. Ernst, and L. Wehenkel, "Extremely randomized trees," *Mach Learn*, Springer, 63: 3–42, 2006. https://doi.org/10.1007/s10994-006-6226-1.

[24] A. Sharafati, S. B. H. S. Asadollah, and M. Hosseinzadeh, "The potential of new ensemble machine learning models for effluent quality parameters prediction and related uncertainty," *Process Safety and Environmental Protection*, Elsevier, 140: 68–78, 2020. https://doi.org/10.1016/j.psep.2020.04.045.

[25] L. Breiman, "Random forests," *Mach Learn*, Springer, 45: 5–32, 2001. https://doi.org/10.1023/A:1010933404324.

[26] V. John, Z. Liu, C. Guo, S. Mita, and K. Kidono, "Real-time Lane estimation using deep features and extra trees regression," in *Image and Video Technology: 7th Pacific-Rim Symposium, PSIVT 2015, Auckland, New Zealand, November 25-27, 2015, Revised Selected Papers 7*, Cham, Springer, 2016, : 721–733. https://doi.org/10.1007/978-3-319-29451-3_57.

[27] G. Mishra, D. Sehgal, and J. K. Valadi, "Quantitative structure activity relationship study of the anti-hepatitis peptides employing random forests and extra-trees regressors," *Bioinformation*, National Library of Machine, 13(3): 60, 2017. DOI: 10.6026/97320630013060.

[28] E. Frank, L. Trigg, G. Holmes, and I. H. Witten, "Naive Bayes for regression," *Mach Learn*, Springer, 41: 5–25, 2000. https://doi.org/10.1023/A:1007670802811.

[29] H. Zou and T. Hastie, "Regularization and variable selection via the elastic net," *J R Stat Soc Series B Stat Methodol*, Oxford Academic, 67(2): 301–320, 2005. https://doi.org/10.1111/j.1467-9868.2005.00503.x.

[30] C. Hans, "Elastic net regression modeling with the orthant normal prior," *J Am Stat Assoc*, Taylor & Francis, 106(496):1383–1393, 2011. https://doi.org/10.1198/jasa.2011.tm09241.

[31] Q. F. Xu, X. H. Ding, C. X. Jiang, K. M. Yu, and L. Shi, "An elastic-net penalized expectile regression with applications," *J Appl Stat*, Taylor & Francis, 48(12): 2205–2230, 2021. https://doi.org/10.1080/02664763.2020.1787355.

[32] Z. Zhang, Z. Lai, Y. Xu, L. Shao, J. Wu, and G.-S. Xie, "Discriminative elastic-net regularized linear regression," *IEEE Transactions on Image Processing*, IEEE, 26(3): 1466–1481, 2017. DOI: 10.1109/TIP.2017.2651396.