Khan, Shahzeb Ahmad; Rehman, Attique Ur; Arshad, Ammar; Alqahtani, Mohammed H.; Mahmoud, Karar; Lehtonen, Matti

Effective Voting-based Ensemble Learning for Segregated Load Forecasting with Low Sampling Data

Published in:
IEEE Access

DOI:
10.1109/ACCESS.2024.3413679

Published: 13/06/2024

Document Version
Publisher's PDF, also known as Version of record

Please cite the original version:
Khan, S. A., Rehman, A. U., Arshad, A., Alqahtani, M. H., Mahmoud, K., & Lehtonen, M. (2024). Effective Voting-based Ensemble Learning for Segregated Load Forecasting with Low Sampling Data. *IEEE Access*, *12*, 84074-84087. https://doi.org/10.1109/ACCESS.2024.3413679

IEEE Access
Multidisciplinary : Rapid Review : Open Access Journal

Received 24 May 2024, accepted 10 June 2024, date of publication 13 June 2024, date of current version 21 June 2024.

Digital Object Identifier 10.1109/ACCESS.2024.3413679

RESEARCH ARTICLE

Effective Voting-Based Ensemble Learning for Segregated Load Forecasting With Low Sampling Data

SHAHZEB AHMAD KHAN[1], ATTIQUE UR REHMAN[1], (Senior Member, IEEE), AMMAR ARSHAD[1], (Member, IEEE), MOHAMMED H. ALQAHTANI[2], (Member, IEEE), KARAR MAHMOUD[3], (Senior Member, IEEE), AND MATTI LEHTONEN[4]

[1]Faculty of Electrical Engineering, Ghulam Ishaq Khan Institute of Engineering Sciences and Technology, Topi 23640, Pakistan
[2]Department of Electrical Engineering, College of Engineering, Prince Sattam Bin Abdulaziz University, Al-Kharj 16278, Saudi Arabia
[3]Department of Electrical Engineering, Faculty of Engineering, Aswan University, Aswan 81528, Egypt
[4]Department of Electrical Engineering and Automation, School of Electrical Engineering, Aalto University, 02150 Espoo, Finland

Corresponding authors: Mohammed H. Alqahtani (mh.alqahtani@psau.edu.sa) and Matti Lehtonen (matti.lehtonen@aalto.fi)

This work was supported by Prince Sattam Bin Abdulaziz University under Grant PSAU/2024/01/99514.

ABSTRACT In power system planning and operation, load forecasting is an important task as it helps ensure a reliable and efficient electricity supply. For effective operation of the smart grid, load forecasting is also an important thing to keep balancing dispatch of power, load management, and load shifting. In this regard, this paper aims to propose an accurate load forecasting based on implementing and integrating different load forecasting models using standalone machine learning and ensemble machine learning models, particularly for segregated real-world load data. In the given context, machine learning models namely, k-nearest neighbor, random forest, decision tree, and voting ensemble regression, are used in this study. The time series load data for this research work was acquired from a real-world load database namely, Pecan Street Dataport. For performance evaluation, two statistical error matrices are used, i.e., mean absolute error (MAE) and mean squared error (MSE). For simulation purposes, Python along with different machine-learning libraries was employed. Moreover, for numerical data analysis and visualization, this research work utilizes different packages like NumPy, pandas, and matplotlib. The empirical study presents the comparative performance analysis of machine learning models for load forecasting utilizing low sampling load data, both at aggregated as well as at segregated levels. Standalone and ensemble machine learning algorithms yield very good forecasting results, and this research has revealed that machine learning models trained on segregated data exhibit superior performance compared to those trained on aggregated data. On segregated data, the proposed voting- based ensemble machine learning algorithm outperforms all the other models with MAE 0.05708, followed by k-nearest neighbors (with MAE 0.05879), random forest (with MAE 0.07069), and decision tree (with MAE 0.07361).

INDEX TERMS Load forecasting, segregated loads, low sampling data, machine learning, ensemble learning.

I. INTRODUCTION

Today, electricity plays a key role in any society and the daily lives of its citizens. Modern-day transportation, security, economy, and many other key sectors are greatly dependent on electricity. With the increasing population and the penetration of electronic gadgets in daily life, the electricity demand is consistently increasing, estimated to be an increase of 3.3% in the global demand for electricity by the year 2024 [1]. Modern-day transportation, security, economy, and many other key sectors are greatly dependent on electricity. Electricity has become one of the most fundamental human

The associate editor coordinating the review of this manuscript and approving it for publication was Akshay Kumar Saha.

necessities over the last century, with every machine, from a simple cell phone charger to a powerful water pump, all depending on electric power for their respective operations. Complex power networks have been developed over the years, with improvements still being made to provide uninterrupted electricity supply to millions of consumers around the world. While electricity cannot be easily stored, posing considerable issues for power generation, transmission, and distribution, consequently over access to power can result in wastage of energy while lack of power might result in power outages. Therefore, load forecasting is important to maintain a balance between power supply and demand. Moreover, maintaining a steady balance between energy production and consumption is crucial to prevent potential harm to the power grid [2].

The temporal scope of power system operational planning could be divided into various distinct forms, namely, short-term, mid-term, as well as long-term, each of which is dedicated to specific tasks. The short-term timeframe pertains to a duration of one day to one week, with a primary emphasis on the operational and security facets of the power system. The mid-term one, which spans from numerous weeks to various months, primarily concerns the effective management of production resources and the prevention of energy deficits in the context of current power plants. The long-term timeframe pertains to a period of several years to decades and is primarily concerned with determining the implementation of new power plants [2]. The primary focus of energy consumption management, from a grid standpoint, is to mitigate peak load during periods of elevated market costs or when the network's reliability is at risk. These actions yield advantages not only for the electricity supplier but also for the consumers. To effectively address the demand response (DR) issues, companies are implementing strategies that allow customers to engage in DR initiatives via their adjustable appliances as well as controllable loads. These strategies involve reducing energy consumption during designated events, shifting high-energy demand loads to non-peak hours, controlling programable demands in real time, and offering financial rewards [3].

The rapid pace of development within the electronics industry has resulted in a significant increase in energy demand over the past two decades. Utilities face challenges in maintaining the balance between supply and demand due to not only the unpredictability and uncertainty in energy consumption patterns but also due to the intermittent nature of renewables, e.g., solar, wind, etc. In this context, accurate energy prediction holds significant importance for modern power systems [4]. Consequently, the forecasting of electricity demand has received significant attention, particularly within the past decade [5]. Recent developments in AI and increased computational resources empower researchers to conduct load forecasting with greater precision. AI has the potential to be utilized in the power industry to predict load and supply [6]. Within AI, machine learning (ML) evolved significantly and recognized itself as an authoritative method in research and real-world application growth [3]. Therefore, much research has been conducted to solve the issues related to load forecasting, particularly with the development of machine learning [3]. In the given context, the development of big data and artificial intelligence (AI) has facilitated the implementation of novel machine learning techniques in the power sector, which necessitates the careful handling of large electricity data. AI has the potential to be utilized in the electricity industry to predict power demand and supply [6]. The significance of electricity forecasting and advanced techniques to address the increasing electricity demand in small and mid-income countries is pivotal [7]. Accurately predicting demand in the short term, ranging from one hour to a day, can aid energy providers in anticipating the optimal amount of power to generate to converge real-time user demand reliably and cost-effectively. Inaccurate estimation of demand can have significant consequences, such as power outages and unstable grid operation in the case of underestimation, and energy wastage in the case of overestimation. Consequently, a precise prediction can lead to improved energy administration and substantial reductions in supplier expenses.

In the available literature, the load forecasting models are primarily based on traditional statistical approaches, machine learning, and deep learning [8]. The common statistical techniques employed in literature comprise autoregressive moving average (AR-MA), exponential smoothing, the analogous day algorithm, and autoregressive integrated moving average (ARIMA) [8]. Nonetheless, such techniques are deficient in understanding the compound connections linking input and output vectors. In contrast, machine learning methods can address the limitations of statistically established methods as they have the capability to represent such compound nonlinear mapping among inputs and output sets, discover unseen forms in immense quantities of information, and propose effective scalability [8]. Besides, Artificial Neural Networks (ANNs), decision tree algorithms, support vector regression (SVR), random forest, and XGBoost are examples of artificial intelligence techniques that are commonly employed for load forecasting [9].

The classifier-regression mapping technique is proposed for short-term forecasting in [10]. Three classifiers namely DT, Ensemble, and SVM, and three regressors such as viz tree, neural networks, and Gaussian process regression are employed. Their hyperparameters are tuned using three different methods. It was found that DT as a classifier and neural networks as a regressor produce the best load forecasting results, and GridSearchCV turned the best hyperparameter tunning technique. The authors of ref. [11] studied different ML methods for assessing the electrical load in Cyprus (from 2016 to 2017) using both short- and long-term assessments. Economic, population, and climate variables were incorporated into the model to forecast the region's electricity demand. The study concluded that multiple linear regression was deemed inferior to ANN, and SVM methods [11].

In ref. [12] authors proposed an enhanced framework for ML based on SVM and extreme learning machines (ELM). Specifically, GridSearchCV was employed to optimize the hyperparameters and it is concluded that ELM provided quick training and accuracy under the given conditions. In recent years, ensemble learning has also been employed in different domains such as ref. [13] has proposed ensemble learning models for load monitoring and found performance improvement compared to their standalone learners. A detailed analysis of ensemble learning strategies, with a special emphasis on their use in deep learning is presented in [14]. Their findings highlighted the significance of ensemble approaches in boosting prediction accuracy and tackling challenges associated with deep learning models, such as hyperparameter tuning and model bias.

Numerous studies in the existing literature present novel approaches for load forecasting. However, from the available literature, it is noted that most of the studies have primarily focused on aggregated data for load forecasting. In-depth analysis and performance evaluation of load forecasting models particularly for segregated load data is lacking. Predicting segregated loads, i.e., forecasting appliance-level patterns/consumptions, can facilitate all the stakeholders of the power system and such feedback not only enhances the system reliability but also plays a key role in real-world energy efficiency applications like recommender systems. Moreover, it is also observed that the available research work mainly relies on aggregated load measurements that are acquired at an hourly sampling rate. The data acquisition at hourly intervals leads to the omission of much important information leading to the somehow inaccurate load forecasting. While this research utilizes appliances load data at minutely sampling rate, as a real-world energy efficiency application.

This research work aims to address the gaps along with better utilizing the load forecasting applications toward energy efficiency. Therefore, this research work focused on different aspects of load forecasting leading to a manifold contribution to the existing research, that is,

- Proposed a voting-based ensemble learning for aggregated and segregated forecasting applications.
- Employed real-world segregated load data for load forecasting in addition to conventional aggregated load measurements. Moreover, the research is carried out using a 1-minute (1/60Hz) interval load measurements compared to the existing research that is mostly carried out on hourly load data.
- Presented a comprehensive empirical analysis along with comparative performance evaluation of four different machine learning models, including ensemble and standalone learners, for segregated as well as conventional aggregated load forecasting.

The remainder of this paper is organized as follows, section II provides the detailed research methodology along with relevant literature, and section III presents the simulation, results, and analysis. The paper is concluded in section IV.

II. RESEARCH METHODOLOGY

This research work adopts a methodology that not only dives into a performance evaluation of standalone machine learning models for aggregated and segregated minutely load forecasting but also explores the domain of ensemble machine learning for load forecasting. Moreover, a comprehensive comparative performance analysis of the employed machine learning models for load forecasting (under the given conditions) is also carried out. In the given context, different regression base standalone machine learning models and ensemble learners, namely k nearest neighbor (k-NN), random forest (RF), decision tree (DT), as well as voting regressor are employed, respectively. An overall flow of the adopted research methodology is depicted in Figure 1.

A. DATA ACQUISITION AND PRE-PROCESSING

Targeting the real-world applications of load forecasting, this research work is carried out on real-world load measurements, acquired (via university license - free version) from a database namely, Pecan Street - Dataport [15]. The free version provides access to limited load data of limited number of houses. Overall, Dataport is comprised of power (in kilowatts) consumption profiles for hundreds of households in the United States of America (USA). Each household comprises load measurements (1/60 Hz sampling) at aggregated and segregated levels (having several different appliances including but not limited to the refrigerator, kitchen appliances, air conditioner, microwave, dishwasher, oven, and furnace). In this research work, data of 4 months duration, i.e., July-October 2014, is acquired from a household-based in California, USA. The acquired data is based on 1-minute intervals where both aggregated (total incoming power) and segregated (individual appliances power) are available. It is worth noting that during the data selection (for the mentioned timeframe), there is no missing value, and the data is consistent in terms of 1-minute interval reading. Moreover, the acquired data is further processed to remove any redundant information and outliers. In the given context, the visualization technique is also used to make sure that there is no outlier in the data [15], [16]. For the acquired load data, Figure 2 shows different types of loads that are under consideration in this research work.

B. MACHINE LEARNING MODELS

This research work addresses load forecasting using machine learning techniques. In the available literature, numerous ML techniques are employed for load forecasting, however from the literature review it is observed that no single technique is superior to others, as the overall performance of an ML model depends upon the given conditions. In the given context, to explore the domain of segregated load forecasting, four different ML models are employed, and the goal is to find the optimal model that offers the most accurate results under

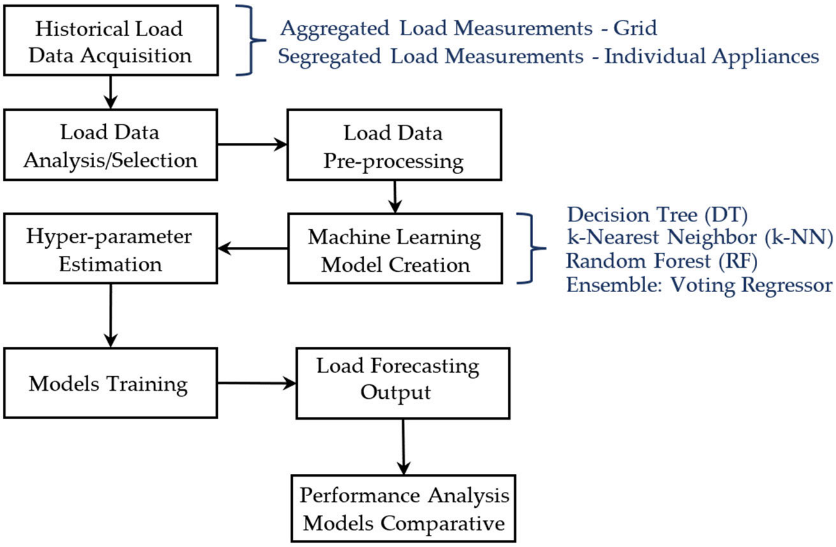


FIGURE 1. Flow of research methodology.

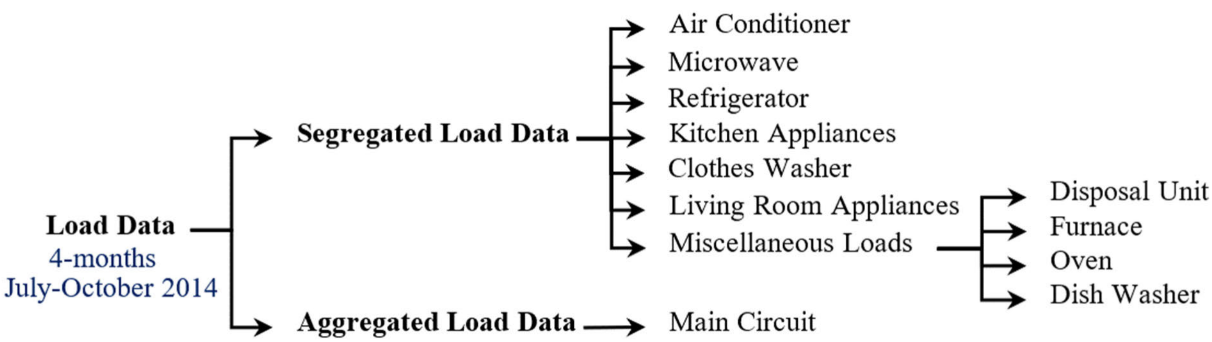


FIGURE 2. Data with both aggregated and segregated loads.

given conditions. For the said, the basic methodology, i.e., to execute and compare the different performance metrics of various ML models under the given condition and select the most accurate one. The ML models employed in this study are k-nearest neighbor, decision tree, random forest, and ensemble learner. These models were selected due to their diverse working principles and varying pros and cons provided in Table 1. The fundamental methodology involves executing and comparing different performance metrics of various machine learning models, with the objective of selecting the most accurate one. It should be noted that weather conditions and factors such as temperature, humidity, and seasonal variations are not currently considered in the employed ML models.

A brief description of each employed model is provided in the below subsections, where Figure 3 presents a generalized framework of machine learning models for load forecasting.

1) K-NEAREST NEIGHBOR MODEL

k-NN, a supervised machine learning regressor [18], operates on the fundamental principle of loading a dataset containing independent values along with their corresponding dependent (target) values. The k-NN regressor predicts the target variable by computing the mean of the target profile of the k number of adjacent neighbors in the training set. The number of neighbors, k, is a hyperparameter that is tuned via the Grid-SearchCV technique. The nearest distance relating two data points is computed via a distance metric named Euclidean distance [19]. The details of the employed k-nearest neighbors' algorithm is outlined below:

2) DECISION TREE MODEL

A DT is an effective ML model that resembles a tree structure. It consists of root, decision, as well as leaf nodes. The topmost node of a DT is the root node, which leads to a sequence of decision nodes, each depicting specific decisions to be assembled. Extending from decision nodes are the leaf nodes, which denote the consequences of the decisions made. Every decision node signifies a split point, and the leaf nodes that stem from a decision node characterize the possible outcomes. A pseudocode of the employed DT algorithm is provided below.

TABLE 1. Comparison of employed machine learning models.

ML Model	Advantages	Disadvantages
k-NN	Easy implementation, works well with multi-class data, easy to interpret, computational time is less, versatility	Sensitive to the choice of *k*, the dimensionality issue, biased towards features with larger scales, vulnerable to noisy data, memory-intensive
DT	Handles non-linear relationships, applicable to both classification and regression, good generalization ability	Not suitable for unstructured data, error propagation issues, overfitting issues, instability
RF	Parallelizable training, noise robustness, high predictive accuracy, handles overfitting	Complexity, lack of interpretability, and more no of trees result in high computational time, memory usage
Ensemble Learning	Better Performance, robustness, increased stability	Complexity, computational cost, reduced interpretability, dependency on diverse models

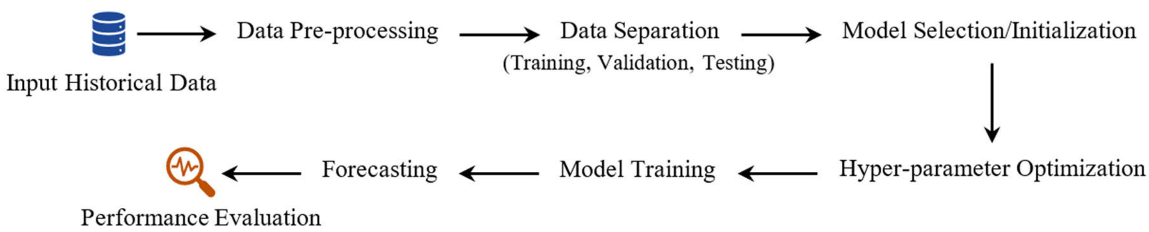


FIGURE 3. General framework of ML models for load forecasting.

Algorithm 1 k-NN Algorithm

Input: Training dataset $D = \{(x_1, y_1), (x_2, y_2) \ldots (x_N, y_N)\}$ consisting of N training instances each with m features, $k = 30$.
Output: Predicted value $\hat{y}_{test}$ for the test instance.
Algorithm:

i. Distance Calculation: For each training instance (x_i, y_i) in D, calculate the Euclidean distance d_i between x_i and x_{test} using,

$$d_i = \sqrt{\sum_{j=1}^{m} \left(x_{i,j} - x_{test,j}\right)^2}$$

ii. Sort Neighbors: On the basis of $d_{i,}$ sort the training instances in ascending order.

$$\{(x_1, y_1), (x_2, y_2) \ldots (x_N, y_N)\} \ni d_1 \leq d_2 \leq \ldots \leq d_N$$

iii. Selection of k nearest neighbors**:** Select top k=30 (hyperparameter) instances

$$\{(x_1, y_1), (x_2, y_2) \ldots (x_{30}, y_{30})\}$$

iv. Compute Prediction: Compute the $\hat{y}_{test}$ as,

$$\hat{y}test = \frac{1}{k} \sum_{i=1}^{k} y_i$$

v. Return Prediction: Return the calculated prediction, $\hat{y}_{test}$.

---

3) RANDOM FOREST MODEL

RF is an ML method that utilizes an ensemble of decision trees to improve its performance. Multiple decision trees are trained, and each tree casts a vote for its favored class. The final prediction is decided by choosing the class with the highest number of votes or averaging the base

---

Algorithm 2 DT Algorithm

Input:

- Training dataset $D = \{(x_1, y_1), (x_2, y_2) \ldots (x_N, y_N)\}$ consisting of N training instances each with m features.
- Maximum dept of the tree (max_depth) = 10
- Minimum number of samples required to split an internal node (min_samples_split) = 10
- Minimum number of samples required to be at a leaf node (min_samples_leaf) = 4

Output: Predicted value $\hat{y}_{test}$ for the test instance.
Algorithm:

i. Build Tree: Recursively split the data based on features until the maximum depth is reached or the number of samples is below a threshold.
ii. Stopping Criteria:
If depth = max_depth, **then** stop
If number of samples at node < min_samples_split, **then** stop.
If number of samples at node < min_samples_leaf, **then** stop.
iii. Prediction: Traverse the tree to a leaf node and return the average of target values in that node.
iv. Return Prediction: Return the calculated prediction, $\hat{y}_{test}$.

---

predictions [20], [21]. The Random Forest model exhibits both efficient training and robustness against overfitting, irrespective of how many trees are applied in combination [22]. The details of the employed random forest algorithm are given below.

4) PROPOSED ENSEMBLE LEARNING MODEL

Ensemble methods are the techniques used in machine learning that combine different standalone algorithms to produce improved results. One of the known ensemble techniques is

Algorithm 3 RF Algorithm

Input

- Training dataset $D = \{(x_1, y_1), (x_2, y_2) \ldots (x_N, y_N)\}$ consisting of N training instances each with m features.
- Maximum dept of the tree (max_depth) = 10
- Minimum number of samples required to split an internal node (min_samples_split) = 2
- Minimum number of samples required to be at a leaf node (min_samples_leaf) = 6
- Number of trees (n_estimator) = 200

Output: Predicted value $\hat{y}_{test}$ for the test instance.

Algorithm

i. Built Forest:
Initialize n_estimator = 200
For each tree t = 1 to n_estimator
Draw a bootstrap sample D_t from training dataset D.
Recursively split the data based on features until the maximum depth is reached or the number of samples is below a threshold.

ii. Stopping Criteria:
If depth = max_depth, **then** stop
If number of samples at node < min_samples_split, **then** stop.
If number of samples at node < min_samples_leaf, **then** stop.

iii. Prediction:
  a. For given text instance x_{test}:
  b. For each tree t:

Traverse the tree to a leaf node and return the average of target values in that node. Store the predicted value $\hat{y}t_{test}$ from each tree.

iv. Return Prediction: Return final predicted value $\hat{y}_{test}$ as average of all individual trees as,

$$\hat{y}_{test} = \frac{1}{\text{n\_estimator}} \sum_{t=1}^{\text{n\_estimator}} \hat{y}t_{test}$$

the voting ensemble, which is employed for both classification and regression tasks as a voting classifier and voting regressor, respectively. Voting is used by classification while averaging is used for regression. The voting classifier combines multiple base models, and the last forecast is dependent on the voting scheme, either hard or soft voting [23]. Hard voting is based on majority voting, whereas soft voting refers to average predicted probabilities. The voting regressor is based on a prediction average, which can be either a simple average or a weighted average. In simple average, for every instance of the test dataset, the average is calculated, whereas in the weighted average, the output of each model is first multiplied by some value, and then the average is calculated. The weighted average aims to give more weightage to the prediction of that good base model to improve performance to greater instant. In this research work, an ensemble learner is proposed and implemented using DT, RF, and k-NN models. Figure 4 depicts the layout of the proposed ensemble technique employed in this research work.

C. PERFORMANCE EVALUATION METRICS

Evaluation metrics play a key role in systematically assisting the performance of different models. Researchers employed numerous performance metrics, such as mean absolute error

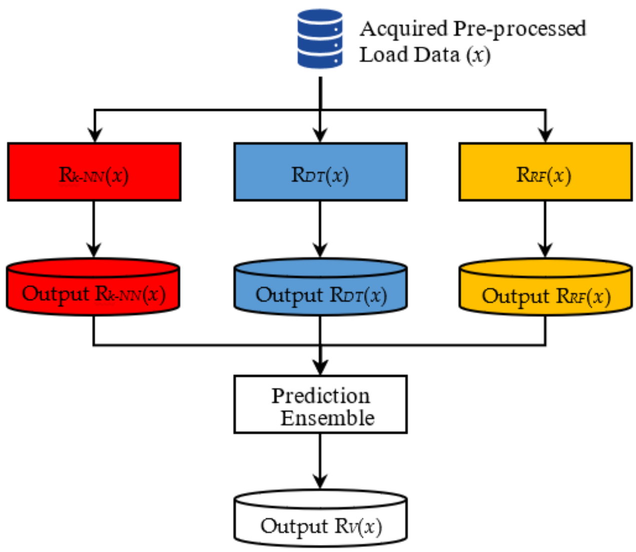


FIGURE 4. Proposed voting ensemble; $R_{k-NN}(x)$, $R_{DT}(x)$, $R_{RF}(x)$, and $R_V(x)$ represent the k-nearest neighbor, decision tree, random forest, and proposed voting regressor, respectively.

(MAE), mean absolute percentage error (MAPE), and mean squared error (MSE), for regression-based ML models. For comparative analysis and performance evaluation of employed ML models under the given conditions, the performance metrics of MAE and MSE are used, mathematically given in (1) and (2), respectively [24], [25], [26].

$$MAE = \frac{1}{n}\sum\nolimits_{i=1}^{n} |y_i - y_P| \tag{1}$$

$$MSE = \frac{1}{n}\sum\nolimits_{i=1}^{n} (y_i - y_p)^2 \tag{2}$$

where n is the number of times the summation iteration happens, y_i represents the actual value, and y_p represents the predicted/forecasted value. Moreover, MAPE, mathematically given in (3) [27], is not employed in this study as the segregated loads' (individual appliances) power is zero at intervals where the load is not in usage, i.e., turn-off state. This leads the MAPE performance metric to be infinite at that particular interval due to having an actual value (which is zero, i.e., the load is turned off) in the denominator, as shown in (3) [28].

$$MAPE = \frac{1}{n}\sum\nolimits_{i=1}^{n} \left|\frac{y_i - y_p}{y_i}\right| \tag{3}$$

III. SIMULATIONS AND RESULTS

As discussed in section II, four diverse machine learning models are employed to forecast load data at both aggregated and segregated levels (details of individual loads under consideration are depicted in Figure 2). For simulation purposes, Python programming language is used. The simulations are carried out using a notebook having a Core i7 5[th] Gen CPU (2.6 GHz clock) and 8 GB RAM. As mentioned earlier, a total of 4-month load data is under consideration, which is further split into around 3 months, i.e., 132,480 data samples, and

TABLE 2. ML model parameters.

Models	Hyper-parameters Details
k-NN	n_neighbors=30, p=2, weights= 'uniform'
DT	max_depth=10, min_samples_leaf=4, min_samples_split=10
RF	max_depth=10, min_samples_leaf=6, min_samples_split=2, n_estimators=200
Ensemble Learning	Voting (weighted averaging)

TABLE 3. K-NN performance for different loads.

Load Description	MSE	MAE
Air Conditioner	0.01258	0.01146
Clothes Washer	0.00005	0.00119
Kitchen Appliances	0.00172	0.00437
Living room Appliances	0.00063	0.00625
Microwave	0.00210	0.00344
Refrigerator	0.00509	0.02471
Miscellaneous Loads (Dish Washer, Disposal, Furnace, Oven)	0.00586	0.00737
Accumulated - Segregated Loads	**0.02803**	**0.05879**
Grid - Aggregated Load	**0.04678**	**0.08054**

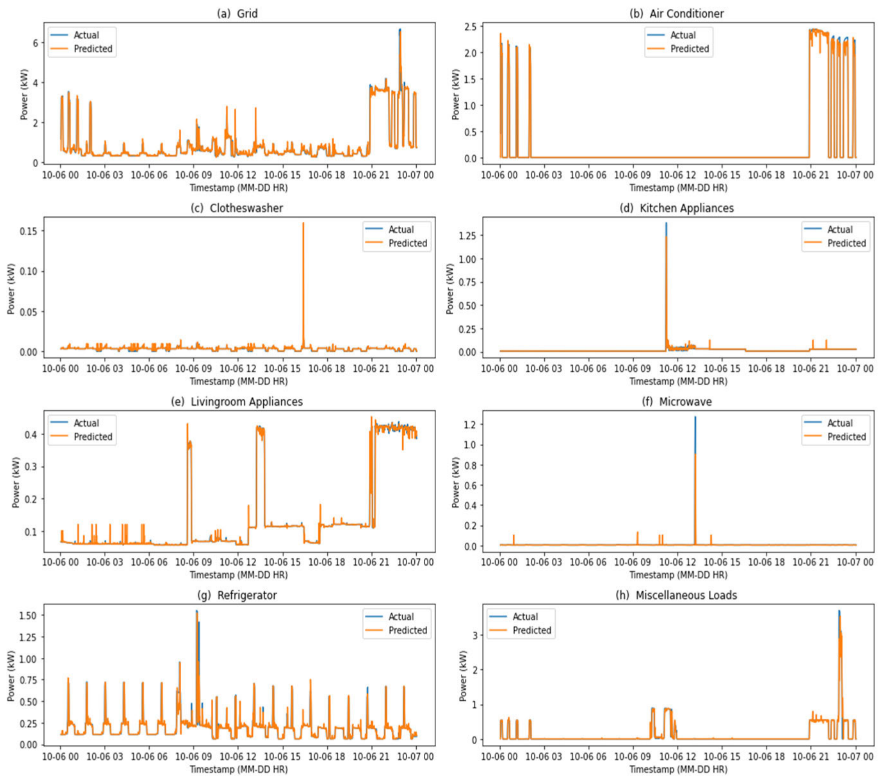


FIGURE 5. k-NN based forecasting results of aggregated and segregated loads.

1 month, i.e., 44,640 data samples, for training and testing purposes of each employed model, respectively. Moreover, during simulation, the hyper-parameters of each ML model are tuned using the Python GridSearchCV library [29]. In this technique range of parameters is defined, then each combination is tested, and the combination of parameters that results in optimum output are considered as tunned parameters. The optimum output selection is based on the objective function or scoring function, in this case, the MSE, of the GridSearchCV. The GridSearchCV is considered an exhaustive technique as it checks all the combinations. So, the computational time can be more, but the accuracy with this method can be high [29], [30]. Brief details of different hyperparameters of the employed ML models are presented in Table 2.

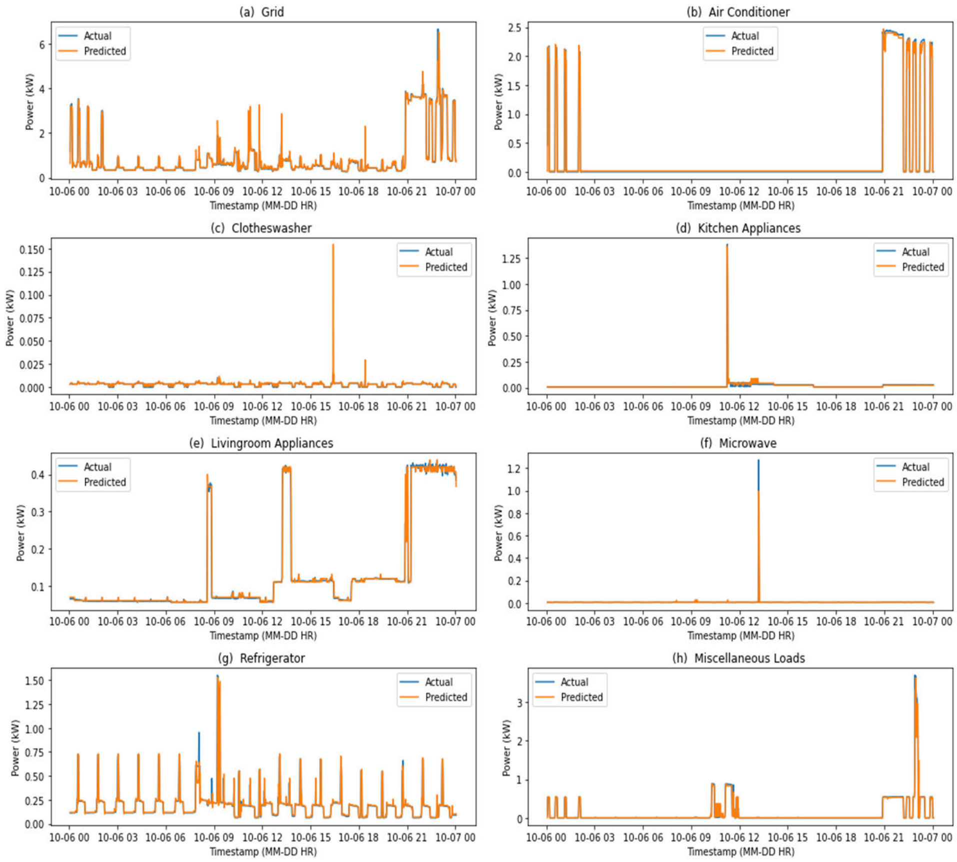


FIGURE 6. **DT-based forecasting results of aggregated and segregated loads.**

A. INTRA-MODEL PERFORMANCE EVALUATION

Based on the presented methodology (Figure 1) and hyper-parameters (Table 2), digital simulations are carried out for each ML model using aggregated as well as segregated load data. In the given context, the performance metrics of MAE and MSE are extracted for each load under consideration. Table 3 presents the details of the obtained results for the k-NN model.

It is evident from the presented performance results, in Table 3, that the k-NN model performs well for both aggregated as well as segregated loads. However, it is worth noting that at the segregated level, even at the accumulated level, forecasting performance, both MSE and MAE, is better compared to the forecasting results obtained for the aggregated load (bold text in Table 3).

In the given context, for segregated loads (accumulated) the overall improvement of 40% and 27% has been recorded in terms of MSE and MAE, respectively, in comparison to the MSE and MAE results obtained for the aggregated load. The acquired performance can be further analyzed graphically in Figure 5, depicting the employed model, k-NN, performance in terms of forecasted/predicted vs. actual load. It is also evident from the presented results, Figure 5, that the k-NN model precisely forecasted the aggregated as well as segregated loads.

Similarly, Figures 6-8 present a graphical depiction of forecasting performance for DT, RF, and Ensemble learners, respectively. It is evident from the presented results that all the tuned models perform well and precisely forecasted the load including aggregated and segregated load, under consideration. From the presented results (Figures 5-8), a slight variation in the forecasting performance has been observed for similar loads, however, it is expected because even though the load data is similar, the employed models are diverse, having different working principles with varying pros and cons. Furthermore, likewise the k-NN model, the other employed models also followed the trend of performance edge for segregated loads compared to forecasting results obtained for the aggregated load. This is anticipated, as the segregated loads have a more deterministic usage pattern which is easy to predict.

B. INTER-MODEL PERFORMANCE EVALUATION

To evaluate the inter-model performance for the given load data, a comparative analysis is carried out in terms of mean square error and mean absolute error performance metrics. Table 4 presents the employed ML models' performance, in terms of MSE metric, for all the loads under consideration.

The presented results can be further analyzed using the graphical illustration presented in Figure 9. It is evident from

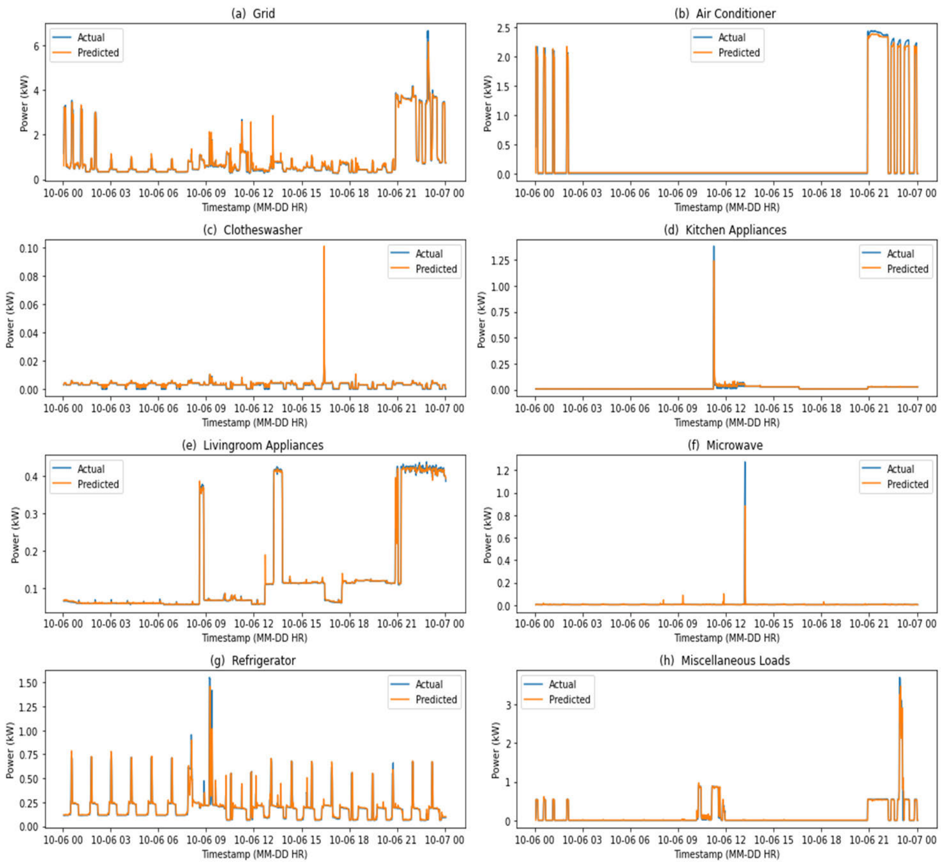


FIGURE 7. RF-based forecasting results of aggregated and segregated loads.

TABLE 4. Inter-model performance - mean square error.

Loads	DT	k-NN	RF	Ensemble
Air Conditioner	0.01250	0.01258	0.01256	0.01248
Clothes Washer	0.00007	0.00005	0.00005	0.00005
Kitchen Appliances	0.00184	0.00172	0.00165	0.00164
Living room Appliances	0.00067	0.00063	0.00059	0.00059
Microwave	0.00186	0.00210	0.00190	0.00183
Refrigerator	0.00519	0.00509	0.00460	0.00457
Miscellaneous Loads	0.00660	0.00586	0.00510	0.00520
Segregated Loads (Acc.)	**0.02873**	**0.02803**	**0.02645**	**0.02636**
Grid - Aggregated Load	**0.05209**	**0.04678**	**0.04461**	**0.04599**

the presented results that among all the employed models, our proposed ensemble learner performs optimally, with a smaller MSE value, in comparison to the other standalone learners. As discussed earlier, it is also evident from the MSE performance that the employed models perform well for segregated loads compared to the aggregated load, bold in Table 4 and highlighted in Figure 9.

Likewise, Table 5 presents the extracted results for employed models in terms of mean absolute error metric. The trend of the given results can be graphically analyzed in Figure 10.

As obvious from the obtainable result in Table 5, all utilized models perform well for both aggregated as well as segregated data. However, for segregated loads, the proposed ensemble learner has an edge over the rest of the employed models. On the other hand, for aggregated load data, the RF model has a slight edge over the proposed ensemble learner. Moreover, under the given conditions, the DT model had the poorest performance compared to the other employed ML models, both in terms of MSE and MAE.

From the presented results, it is obvious that almost all the employed ML models showed superior forecasting performance on segregated loads data as compared to aggregated data. This is anticipated as segregated loads of data offer more granularity in terms of not only exclusivity but also usage patterns. These attributes of segregated

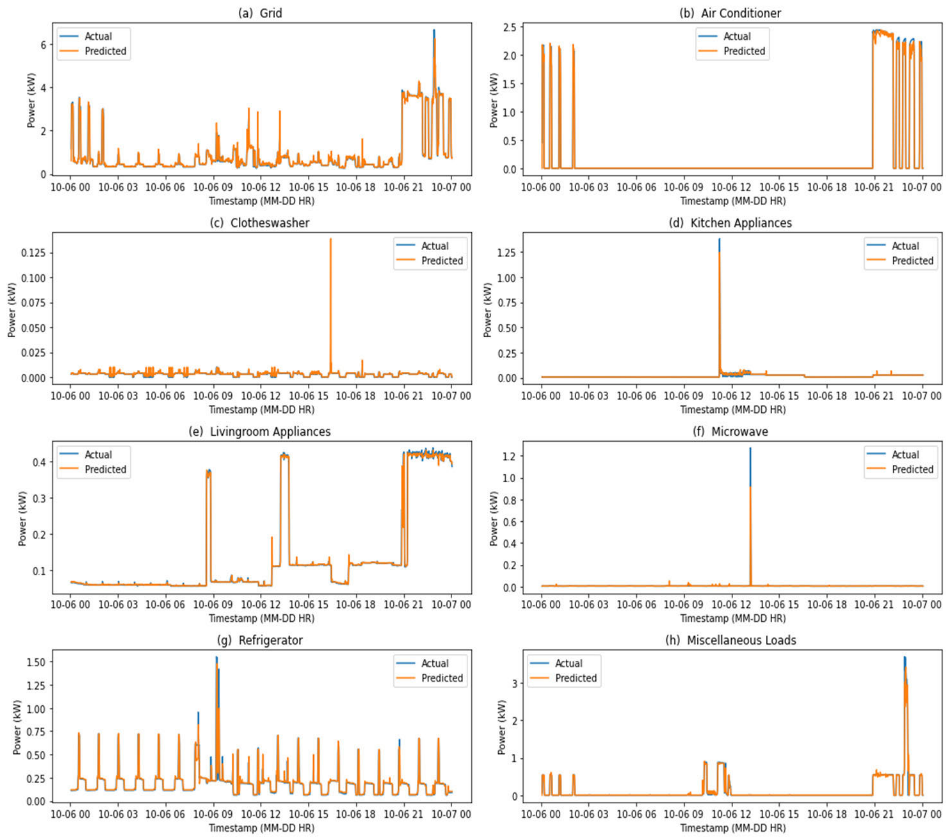


FIGURE 8. Ensemble learner forecasting results of aggregated and segregated loads.

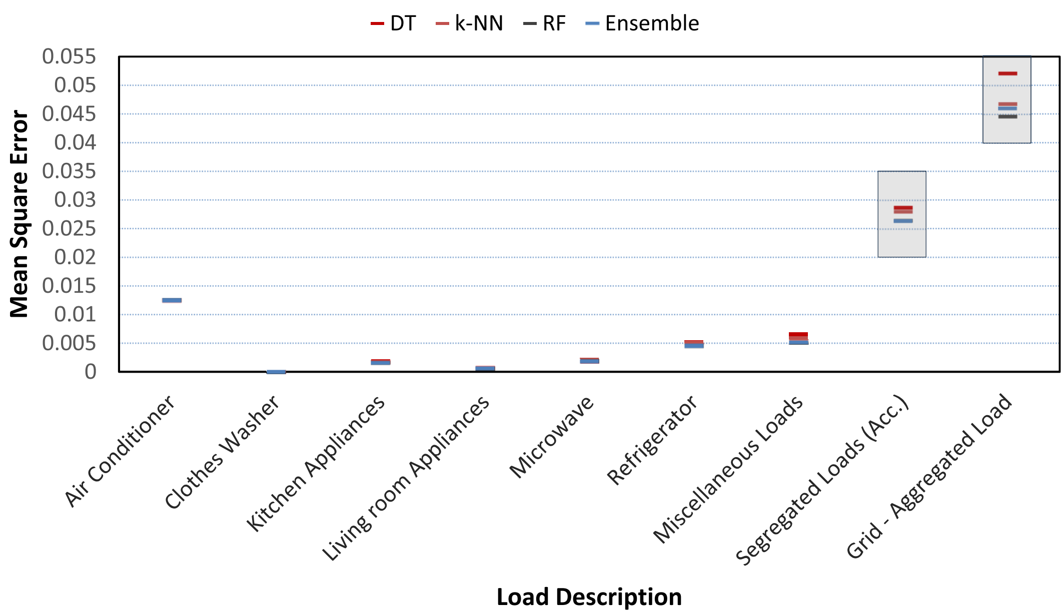


FIGURE 9. Inter-model performance - mean square error.

loads enable a more accurate portrayal of load patterns and fluctuations, consequently leading to a more precise forecasting outcome. Aggregated load, on the other hand, may mask these exclusive variations, resulting in less precise predictions. In a nutshell, the increased granularity of segregated data allows ML models to incorporate load-specific

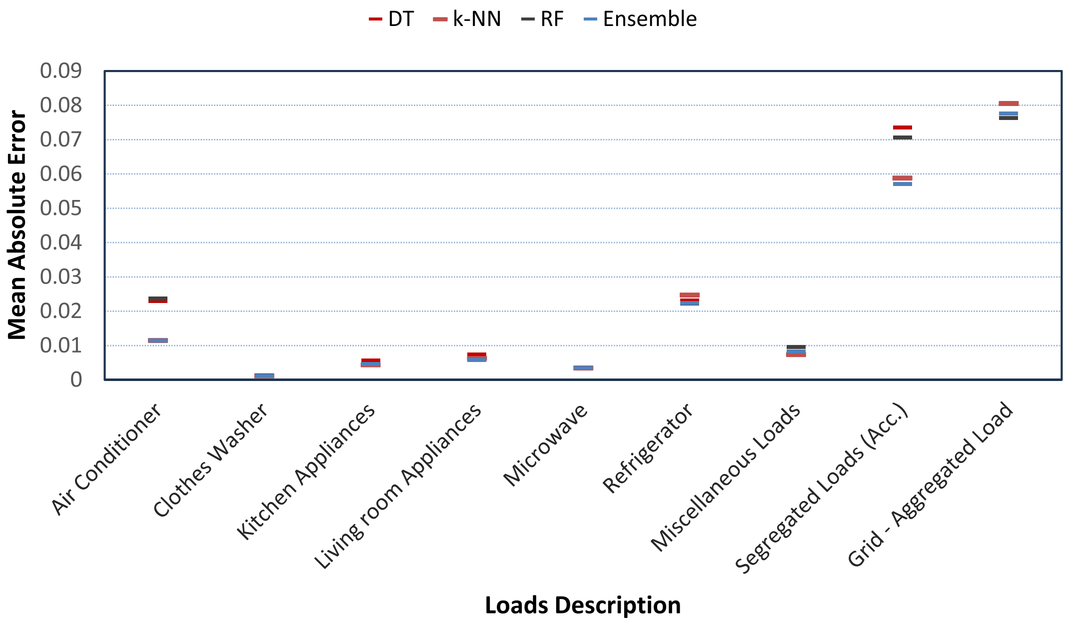


FIGURE 10. Inter-model performance - mean absolute error.

TABLE 5. Inter-model performance - mean absolute error.

Loads	DT	k-NN	RF	Ensemble
Air Conditioner	0.02294	0.01146	0.02366	0.01152
Clothes Washer	0.00131	0.00119	0.00127	0.00117
Kitchen Appliances	0.00558	0.00437	0.00471	0.00456
Living room Appliances	0.00735	0.00625	0.00576	0.00575
Microwave	0.00366	0.00344	0.00343	0.00339
Refrigerator	0.02315	0.02471	0.02227	0.02241
Miscellaneous Loads	0.00962	0.00737	0.00959	0.00828
Segregated Loads (Acc.)	**0.07361**	**0.05879**	**0.07069**	**0.05708**
Grid - Aggregated Load	**0.08064**	**0.08054**	**0.07626**	**0.07764**

properties and variations, resulting in better prediction performance.

In addition to performance accuracy, load forecasting at a segregated level can further facilitate many energy efficiency applications such as demand response, load shifting, and energy saving [32]. The detailed insights gained from segregated load forecasting enable more precise demand response actions, efficient load-shifting schemes, and targeted energy-saving measures. Therefore, the advantages extend beyond performance accuracy, contributing to a more nuanced and impactful implementation of real-world energy efficiency applications.

It is also worth noting that under the given conditions, the ensemble learners outperform the standalone learners. For aggregated data the RF, being an ensemble learning model, outperforms the rest of the models by successfully managing the complicated relations of the acquired aggregated load data. RF model effectively captures the non-linear correlations among the acquired aggregated load data and adapts to various load patterns, consequently leading to its improved performance. On the other side, for segregated loads, our proposed ensemble model outperforms the rest of the employed models including the RF model. In the given context, the proposed ensemble learner utilizes individual models and takes the average of individual predictions. Under the given conditions, the employed weighted average technique for the proposed ensemble learner improves the performance, as the individual regressors who performed well on individual loads were given more weightage while averaging the predictions.

C. COMPUTATIONAL ANALYSIS

The computational analysis of the employed models, namely k-NN, DT, RF, and ensemble learner, was conducted using the time library in Python. All models were trained on a minutely interval three-month data, and their respective training times were recorded. Among the employed models, k-NN exhibited the shortest training time, while the proposed ensemble learner had the longest training time. The results regarding training time for each model are detailed in Table 6.

TABLE 6. Computational analysis of employed machine learning models.

Models	Training Time (ms)
K-Nearest Neighbors	527
Decision Tree	903
Random Forest	43668
Ensemble Learning	58225

The computational times for random forest and proposed ensemble learning are longest among all the employed models because both are ensemble leaners, however these models perform optimally (evident from Tables 4, 5). Here, it is important to mention that the performance enhancement made possible by the ensemble model comes at a cost to model complexity and processing time. As a result, there is a trade-off between performance and computing power. Therefore, it is up to the user's preference and the sensitivity of the problem to favor performance above computational efficiency and vice versa.

IV. CONCLUSION

This research work explored the domain of load forecasting for real-world energy efficiency applications. Unlike the existing studies, this research considered not only aggregated load data but also explored segregated load data. For said purpose, real-world load measurements, at a sampling rate of 1/60 Hz, were acquired from a household in California, USA. Moreover, this research study not only employed four diverse machine learning models but also carried out a comprehensive performance analysis of the employed ML models.

It is also worth noting that the employed ML models were not only evaluated in a standalone configuration but also assessed in an ensemble configuration. In the given context, in addition to the RF ensemble, another voting-based ensemble model is proposed. All the employed ML models were trained and tested on the acquired real-world load measurements. The corresponding results and analysis showed that overall, the employed ML models performed best on aggregated as well as segregated load data. However, the forecasting results for the segregated loads (at the accumulated level) were better compared to the forecasting performance at the aggregated load. The k-NN performs better on segregated loads with MAE, MSE 0.05879, and 0.02803 than aggregated loads with MAE and MSE of 0.08054, and 0.04678 respectively. Similarly, other models such as DT, RF, and voting ensemble followed the same trend and DT performed superior at segregated load data (of MAE, MSE 0.07361, 0.02873) than aggregated load data (of MAE, MSE 0.08064, 0.05209). For RF at segregated load data MAE, MSE is 0.07069, and 0.02645 and at aggregated load data MAE and MSE are 0.07626, and 0.04461 respectively. The proposed voting ensemble also has a lower MAE, and MSE (of 0.05708, 0.02636) at segregated load data than aggregated load data with MAE and MSE (0.07764, 0.04599). Therefore, it is suggested that considering the individual (segregated) load consumption behaviors yields more precise predictions. Further, segregated load forecasting can facilitate many real-world energy efficiency applications, e.g., demand response, load shifting, effective tariff design, etc.

In terms of inter models' comparison while considering aggregated data, the RF gave best forecasting accuracy with MAE of 0.07626, followed by the voting ensemble (of MAE 0.07764), k-NN (of MAE 0.08054) and DT (of MAE 0.08064). For segregated load data, the proposed voting ensemble outperforms all other models with an MAE of 0.05708, followed by k-NN (of MAE 0.05879), RF (of MAE 0.07069), and DT (of MAE 0.07361). It is worth noting that for both aggregated as well as segregated loads, the ensemble machine-learning techniques outperform the standalone models. But here it is worth mentioning that an ensemble learner has more complexity and computational cost than standalone models. This leads to a trade-off between performance and computational expense. Therefore, the end-user has the flexibility to prioritize either performance or computational efficiency according to their needs or the sensitivity of the problem.

The study concludes that it is crucial to consider segregated loads when utilizing machine learning models for load forecasting. Moreover, it has been determined that ensemble learning can facilitate load forecasting performance. In the future, the objective is to expand the scope by exploring deep learning techniques for segregated load forecasting. A key aspect of the future work involves broadening the study by incorporating a larger number of households. Additionally, the intention is to introduce a recommender system as a practical application for real-world energy efficiency.

REFERENCES

[1] (2024). *Global Electricity Demand to Grow By 3.3 Percent in 2024—Forecast*. Accessed: Dec. 4, 2023. [Online]. Available: https://gmk.center/en/news/global-electricity-demand-to-grow-by-3-3-in-2024/

[2] E. Aguilar Madrid and N. Antonio, "Short-term electricity load forecasting with machine learning," *Information*, vol. 12, no. 2, p. 50, Jan. 2021.

[3] A. Bellahsen and H. Dagdougui, "Aggregated short-term load forecasting for heterogeneous buildings using machine learning with peak estimation," *Energy Buildings*, vol. 237, Apr. 2021, Art. no. 110742, doi: 10.1016/j.enbuild.2021.110742.

[4] A. U. Rehman, T. T. Lie, B. Vallès, and S. R. Tito, "Comparative evaluation of machine learning models and input feature space for non-intrusive load monitoring," *J. Modern Power Syst. Clean Energy*, vol. 9, no. 5, pp. 1161–1171, Sep. 2021, doi: 10.35833/MPCE.2020.000741.

[5] S. R. Khan, I. A. Hayder, M. A. Habib, M. Ahmad, S. M. Mohsin, F. A. Khan, and K. Mustafa, "Enhanced machine-learning techniques for medium-term and short-term electric-load forecasting in smart grids," *Energies*, vol. 16, no. 1, p. 276, 2023, doi: 10.3390/en16010276.

[6] P.-H. Kuo and C.-J. Huang, "A high precision artificial neural networks model for short-term energy load forecasting," *Energies*, vol. 11, no. 1, p. 213, Jan. 2018, doi: 10.3390/en11010213.

[7] A. A. Mir, M. Alghassab, K. Ullah, Z. A. Khan, Y. Lu, and M. Imran, "A review of electricity demand forecasting in low and middle income countries: The demand determinants and horizons," *Sustainability*, vol. 12, no. 15, p. 5931, Jul. 2020, doi: 10.3390/su12155931.

[8] B. Ibrahim, L. Rabelo, E. Gutierrez-Franco, and N. Clavijo-Buritica, "Machine learning for short-term load forecasting in smart grids," *Energies*, vol. 15, no. 21, p. 8079, Oct. 2022, doi: 10.3390/en15218079.

[9] M. Cordeiro-Costas, D. Villanueva, P. Eguía-Oller, M. Martínez-Comesaña, and S. Ramos, "Load forecasting with machine learning and deep learning methods," *Appl. Sci.*, vol. 13, no. 13, p. 7933, Jul. 2023, doi: 10.3390/app13137933.

[10] B. V. Surya Vardhan, M. Khedkar, and V. Suresh, "Hyper-parameter tuned short term load forecasting using stochastic classifier-regression mapping for power system operator," in *Proc. IEEE PES 14th Asia–Pacific Power Energy Eng. Conf. (APPEEC)*, Nov. 2022, pp. 1–6, doi: 10.1109/APPEEC53445.2022.10072174.

[11] D. Solyali, "A comparative analysis of machine learning approaches for short-/long-term electricity load forecasting in Cyprus," *Sustainability*, vol. 12, no. 9, p. 3612, Apr. 2020, doi: 10.3390/su12093612.

[12] W. Ahmad, N. Ayub, T. Ali, M. Irfan, M. Awais, M. Shiraz, and A. Glowacz, "Towards short term electricity load forecasting using improved support vector machine and extreme learning machine," *Energies*, vol. 13, no. 11, p. 2907, Jun. 2020, doi: 10.3390/en13112907.

[13] A. U. Rehman, T. T. Lie, B. Vallès, and S. R. Tito, "Non-intrusive load monitoring of residential water-heating circuit using ensemble machine learning techniques," *Inventions*, vol. 5, no. 4, p. 57, Nov. 2020, doi: 10.3390/inventions5040057.

[14] A. Mohammed and R. Kora, "A comprehensive review on ensemble deep learning: Opportunities and challenges," *J. King Saud Univ. Comput. Inf. Sci.*, vol. 35, no. 2, pp. 757–774, Feb. 2023, doi: 10.1016/j.jksuci.2023.01.014.

[15] (2024). *Dataport—Pecan Street Inc*. Accessed: May 17, 2024. [Online]. Available: https://www.pecanstreet.org/dataport/

[16] D. Crain and C. Lysy. (2020). *Outlier Analyses Step-by-Step Guide*. [Online]. Available: https://www.ideadata.org/

[17] (2024). *What is Outlier Analysis and How Can it Improve Analysis*. Accessed: May 17, 2024. [Online]. Available: https://www.smarten.com/blog/outlier-analysis-improve-analysis/

[18] J. Zhang, Y. Li, Z. Wu, C. Rong, T. Wang, Z. Zhang, and S. Zhou, "Deep-reinforcement-learning-based two-timescale voltage control for distribution systems," *Energies*, vol. 14, no. 12, p. 3540, Jun. 2021, doi: 10.3390/en14123540.

[19] D. A. Khan, A. Arshad, and Z. Ali, "Performance analysis of machine learning techniques for load forecasting," in *Proc. 16th Int. Conf. Emerg. Technol. (ICET)*, Dec. 2021, pp. 1–6, doi: 10.1109/ICET54505.2021.9689903.

[20] A. Nadi and H. Moradi, "Increasing the views and reducing the depth in random forest," *Expert Syst. Appl.*, vol. 138, Dec. 2019, Art. no. 112801, doi: 10.1016/j.eswa.2019.07.018.

[21] B. Ait Hammou, A. Ait Lahcen, and S. Mouline, "An effective distributed predictive model with matrix factorization and random forest for big data recommendation systems," *Expert Syst. Appl.*, vol. 137, pp. 253–265, Dec. 2019, doi: 10.1016/j.eswa.2019.06.046.

[22] K. Fawagreh, M. M. Gaber, and E. Elyan, "Random forests: From early developments to recent advancements," *Syst. Sci. Control Eng.*, vol. 2, no. 1, pp. 602–609, Dec. 2014, doi: 10.1080/21642583.2014.956265.

[23] A. U. Rehman, "Load disaggregation: Towards energy efficient systems," Ph.D. dissertation, Dept. Elect. Electron. Eng., School Eng., Comput. Math. Sci., Auckland Univ. Technol., Auckland, New Zealand, 2021.

[24] R. F. Berriel, A. T. Lopes, A. Rodrigues, F. M. Varejão, and T. Oliveira-Santos, "Monthly energy consumption forecast: A deep learning approach," in *Proc. Int. Joint Conf. Neural Netw.*, May 2017, pp. 4283–4290, doi: 10.1109/IJCNN.2017.7966398.

[25] T.-J. Hsieh, H.-F. Hsiao, and W.-C. Yeh, "Forecasting stock markets using wavelet transforms and recurrent neural networks: An integrated system based on artificial bee colony algorithm," *Appl. Soft Comput.*, vol. 11, no. 2, pp. 2510–2525, Mar. 2011, doi: 10.1016/j.asoc.2010.09.007.

[26] C. Willmott and K. Matsuura, "Advantages of the mean absolute error (MAE) over the root mean square error (RMSE) in assessing average model performance," *Climate Res.*, vol. 30, pp. 79–82, May 2005, doi: 10.3354/cr030079.

[27] S. Kim and H. Kim, "A new metric of absolute percentage error for intermittent demand forecasts," *Int. J. Forecasting*, vol. 32, no. 3, pp. 669–679, Jul. 2016, doi: 10.1016/j.ijforecast.2015.12.003.

[28] (2023). *Accuracy—What Are the Shortcomings of the Mean Absolute Percentage Error (MAPE)—Cross Validated*. Accessed: May 29, 2023. [Online]. Available: https://stats.stackexchange.com/questions/299712/what-are-the-shortcomings-of-the-mean-absolute-percentage-error-mape

[29] (2023). *Sklearn Model Selection GridSearchCV—scikit-Learn*. Accessed: Dec. 15, 2023. [Online]. Available: https://scikit-learn.org/stable/modules/generated/sklearn.model_selection.GridSearchCV.html

[30] T. Yu and H. Zhu, "Hyper-parameter optimization: A review of algorithms and applications," 2020, *arXiv:2003.05689*.

[31] B. V. S. Vardhan, M. Khedkar, I. Srivastava, P. Thakre, and N. D. Bokde, "A comparative analysis of hyperparameter tuned stochastic short term load forecasting for power system operator," *Energies*, vol. 16, no. 3, p. 1243, Jan. 2023, doi: 10.3390/en16031243.

[32] R. Dharani, M. Balasubramonian, T. S. Babu, and B. Nastasi, "Load shifting and peak clipping for reducing energy consumption in an Indian University campus," *Energies*, vol. 14, no. 3, p. 558, Jan. 2021, doi: 10.3390/en14030558.

SHAHZEB AHMAD KHAN received the B.Sc. degree in electrical engineering from Namal College, Mianwali, affiliated with the University of Engineering and Technology, Lahore, Pakistan, in 2021, and the M.Sc. degree (Hons.) from the Ghulam Ishaq Khan Institute of Engineering Sciences and Technology, Pakistan, in 2024. His research interests include smart grids, machine learning, optimization, load forecasting, and demand-side management.

ATTIQUE UR REHMAN (Senior Member, IEEE) received the B.E. degree in electrical engineering from Air University, Islamabad, Pakistan, in 2009, the M.Sc. degree from RWTH Aachen University, Germany, in 2013, and the Ph.D. degree from Auckland University of Technology, New Zealand, in 2021.

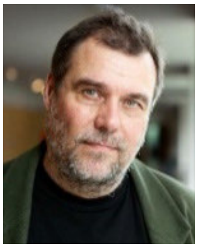
He is currently an Assistant Professor with the Faculty of Electrical Engineering, Ghulam Ishaq Khan Institute of Engineering Sciences and Technology, Pakistan. His research interests include smart grids, applied machine learning, and energy-efficient systems.

AMMAR ARSHAD (Member, IEEE) received the B.Sc. degree in electrical engineering from the University of Engineering and Technology, Lahore, Pakistan, in 2013, and the M.Sc. and D.Sc. degrees from Aalto University, Finland, in 2016 and 2019, respectively.

He has been with the Ghulam Ishaq Khan Institute of Engineering Sciences and Technology, since 2020. His main research interests include smart grid technologies' utilization for PV integration in distribution networks, distributed voltage control, enhancement of PV hosting capacity, and demand response.

MOHAMMED H. ALQAHTANI (Member, IEEE) was born in October 1989. He received the B.S. degree in electrical engineering from Prince Sattam Bin Abdulaziz University, Saudi Arabia, in 2012, the M.S. degree from the University of South Florida (USF), in 2016, and the Ph.D. degree from the Smart Grid Power Systems Laboratory, USF, in 2021. He is currently an Assistant Professor with Prince Sattam Bin Abdulaziz University. His research interests include power system computing and modeling.

KARAR MAHMOUD (Senior Member, IEEE) received the Ph.D. degree from the Electric Power and Energy System Laboratory (EPESL), Graduate School of Engineering, Hiroshima University, Hiroshima, Japan, in 2016. In 2019, he has been with Aalto University, Finland, as a Research Fellow with the Group of Prof. M. Lehtonen, Department of EEA, School of Electrical Engineering. Since 2010, he has been with Aswan University, where he is currently an Associate Professor with the Department of Electrical Engineering, Faculty of Engineering. He has published many advanced scientific papers, including more than 100 scientific papers, in several high-level publishing houses, including scientific journals, international conferences, and scientific book chapters. His research interests include power systems, energy storage, forecasting, electrical vehicles, renewable energy, smart grids, and applied machine learning. He awarded the Prestigious Egyptian State Encouragement Award in the field of engineering sciences, in 2021, and was honored by Egypt's Academy of Scientific Research and Technology (ASRT).



MATTI LEHTONEN received the master's and Licentiate degrees in electrical engineering from Helsinki University of Technology, in 1984 and 1989, respectively, and the Doctor of Technology degree from Tampere University of Technology, in 1992. He was with VTT Energy, Espoo, Finland, from 1987 to 2003. Since 1999, he has been a Full Professor and the Head of the Power Systems and High Voltage Engineering Group, Aalto University, Espoo. His research interests include power system planning and assets management and power system protection, including earth fault problems, harmonic-related issues, high-voltage systems, power cable insulation, and polymer nanocomposites. He is an Editor and a Special Issue Editor of *IET Generation, Transmission and Distribution journal* and an Associate Editor of *Electric Power Systems Research* journal.