A

Dissertation

Presentation on

Stacked Ensemble Model with Heap Based Optimization for Building Energy Consumption Prediction

Supervisors

Mr. Ashok Kumar Agarwal

(Associate Professor)

Dr. Prerna Jain

(Associate Professor)

Presented by

Puneet Pahadia

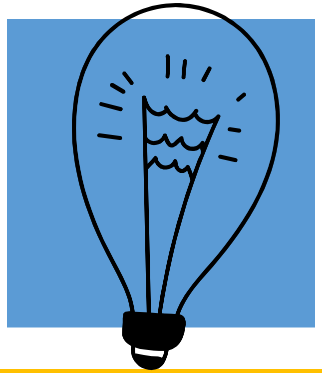
2021PSM5643

Department of Electrical Engineering

Malaviya National Institute of Technology Jaipur

2021-2023

Table of Contents



Introduction

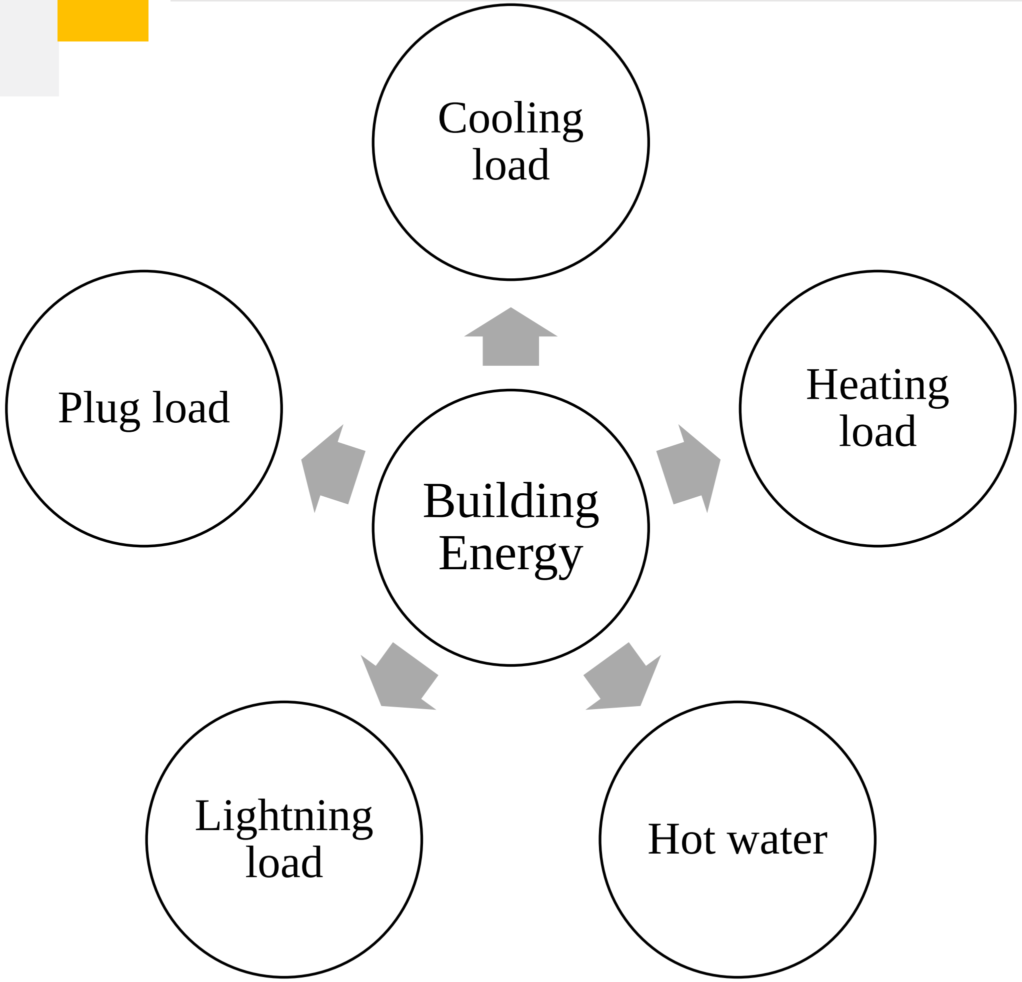


Fig.1 Different Kind of loads in Building

Source 5,6

- Building energy prediction is not only an important evaluation tool of energy-saving potential but also an essential component of smart buildings
- The term energy management usually refers to the saving of energy .
- In broad it means improving energy efficiency in power devices such as electrical equipment or vehicle and the development of renewable energy.

Energy Management Includes

- Planning and saving
- Operation of Energy Production
- Energy Consumption

Source 3

Motivation

- Building energy consumption accounts for a significant portion of global energy consumption and carbon emissions. About 40% of worldwide energy use and 30% of the world's carbon footprint are attributable to buildings.
- Traditionally, building energy prediction relies on energy simulation software, which requires detailed building models and significant computational resources. However, energy simulation software needs accurate inputs for the building model, which can be challenging to obtain.
- The computational requirements for energy simulations can be time-consuming and costly
- In order to resolve the above constraint and to use the resources efficiently there is a need to forecast accurate energy prediction which can help building owners and managers to optimize energy usage and lower their energy costs.
- The above problem can be solved by using machine learning techniques to forecast the building energy usage

Source 13,14

Objective

1

Obtain typical load profile of Building

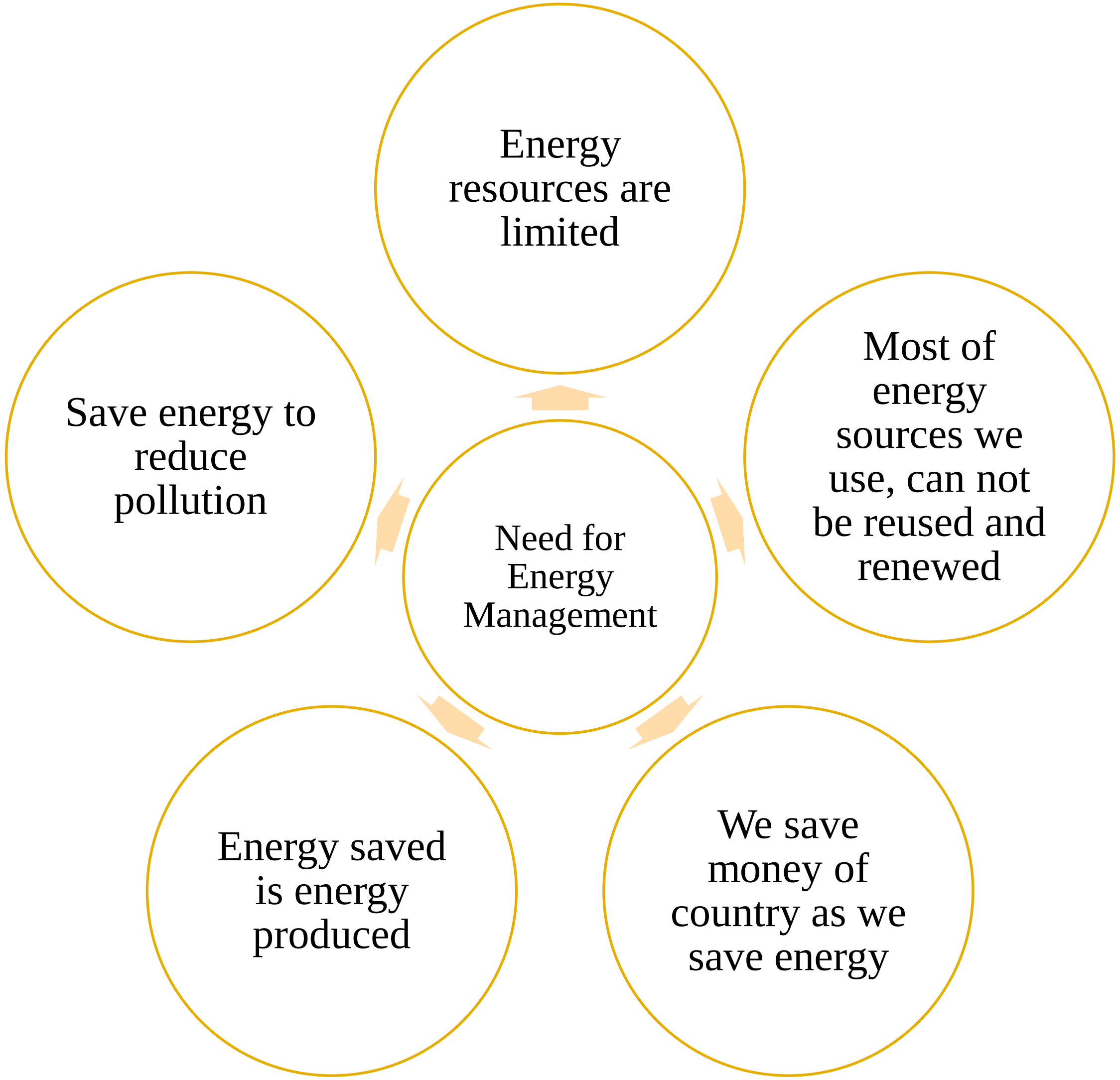
2

Predicting the Building Energy Consumption by Different Machine Learning Models

3

Hyperparameter tuning By Heap Based Optimization (HBO)

Need of Energy Management



Source 2

Fig.2 Need for Energy Management

Building Energy Management

Daily, many building electrical appliances are manufactured to make the building interesting, comfortable, and secure, as well as to facilitate living in the building

This issue has piqued the interest of the scientific society, necessitating the necessity of Building energy management systems, or BEMS.

Smart BEM, which was merged with a control based on events binary linear optimization to supply optimal electrical power to the domestic segment was introduced to develop the BEM idea in order to increase the system's level of performance.

Smart BEMS is provided with five primary features, including

- the capacity to monitor, which provides real-time data on energy use pattern,
- logging, which entails collecting and storing data

Source 4,11

ML in Energy Management Systems

Forecast Energy Bills

Identify Consumption Patterns

Operate Heat Ventilation and Illumination Systems

Analyze Big Amounts of Data

Help Decision Making

Obtain New Insights

Source 9

Motivation for Machine Learning

- Although Individual Machine learning models are also performing better for prediction, but the result can further be enhanced by using combination of model which is stacked ensemble model.
- In stacked ensemble model we combine two or more than two ML models in parallel and each model is trained using a subset of the data, each model produces its result and after this we select the meta model, one which is giving us the best result now we use this meta model for forecasting purpose
- And for input of the meta model we take the prediction results from level 0
- The level 1 model is trained on multiple predictions given by several base models, and then it optimally integrates the predictions of the base models on the testing data.

Source 11,12

Stacked Ensemble Model

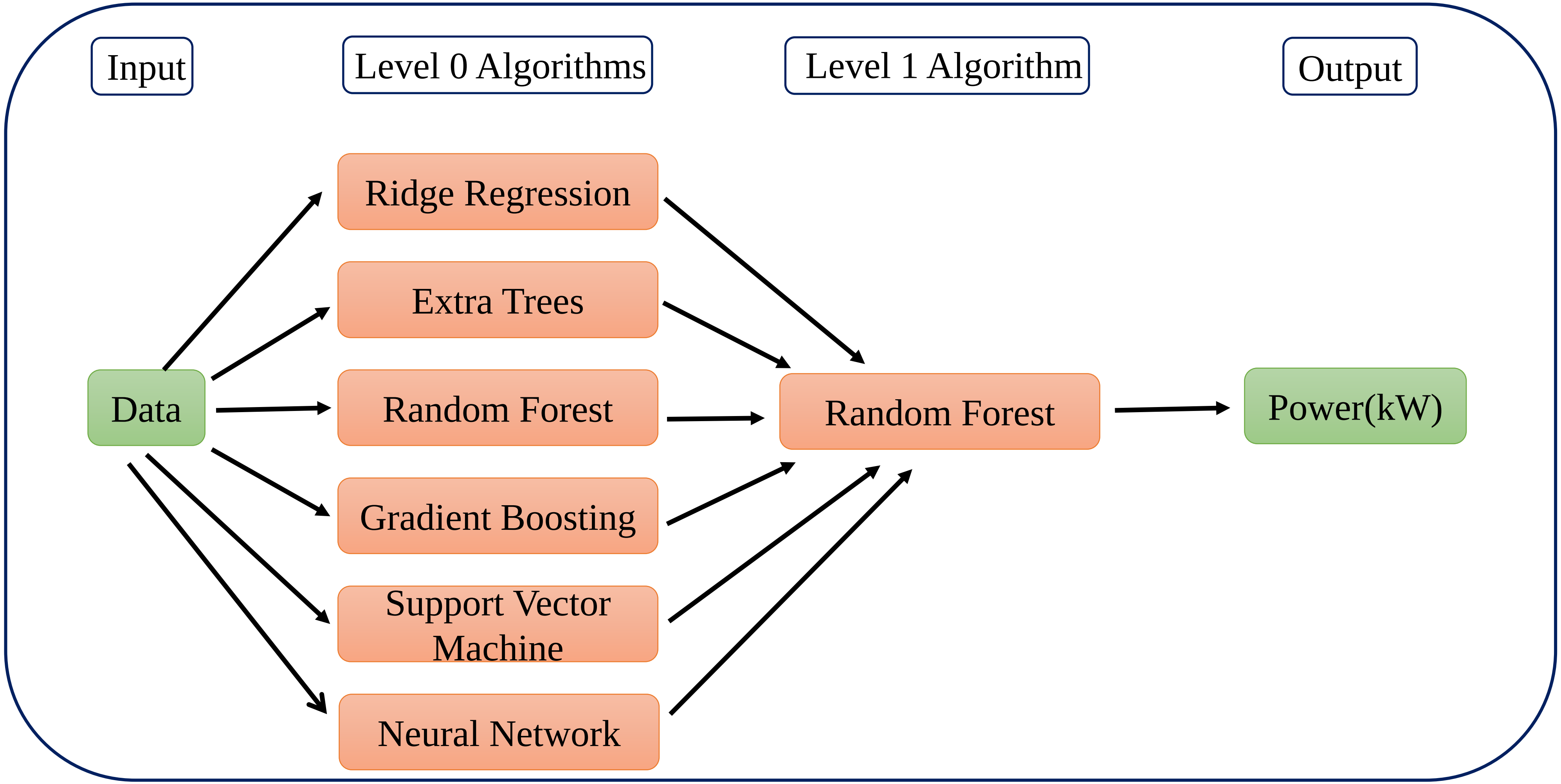


Fig.3 Stacked Ensemble Model

Hyperparameter Tuning

- Hyperparameters are parameters whose values control the learning process of the model and determines the performance of the model
- This process of choosing the optimal parameter for the ideal model architecture is called hyperparameter tuning
- In Ensemble Model, there is a large number of hyperparameters like max_depth, n_estimators, max_samples, etc.
- In order to get the best result from our existing model we perform hyperparameter tuning. So, that we can feed our model with the best parameters.

Source 13,14

Heap Based Optimization

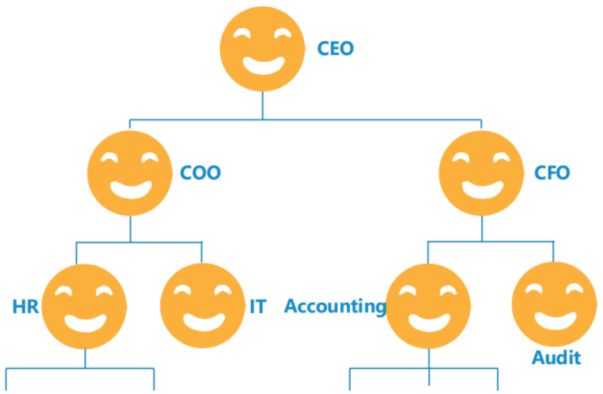


Fig.4 Corporate Rank Hierarchy

- Heap Based Optimization (HBO) is a meta-heuristic algorithm inspired by human social behavior..
- HBO minimizes memory allocations and deallocations by utilizing a heap data structure.
- It allocates a large chunk of memory at once and assigns smaller chunks as needed, reducing memory management overhead.

- In a corporation, HBO applies the heap structure to represent the Corporate Rank Hierarchy (CRH).
- CRH organizes search agents hierarchically based on fitness, resembling an organizational chart.
- The hierarchy follows a tree-like framework, with managers as parent nodes and subordinates as offspring.
- Each subordinate interacts with their direct supervisor.
- Example of a CRH is shown in Fig. 4, with the CEO as the primary node and other roles as subordinates.

Source 13,15

Heap Based Optimization

- **Mathematical Modeling of the Interaction with Immediate Boss**

$$X_i^k(t+1) = B^k + \left| \frac{\left(t \bmod \frac{T}{C} \right)}{\frac{T}{4C}} \right| (2r-1) \left| B^k - X_i^k(t) \right| \quad (1)$$

- **Mathematical Modeling of the Interaction Between Colleagues**

$$X_i^k(t+1) = \begin{cases} S_r^k + \gamma(2r-1)^k \left| S_r^k - X_i^k(t) \right|, \ f(S_r) < f(X_i) \\ X_i^k(t) + \gamma(2r-1)^k \left| S_r^k - X_i^k(t) \right|, \ f(S_r) \geq f(X_i) \end{cases} \quad (2)$$

- **Mathematical Modeling of the Employee Self-Contribution**

$$X_i^k(t+1) = X_i^k(t) \quad (3)$$

Nomenclature

$r \in [0,1]$ → randomized value,

B → optimal solution

t → present generation

k → component vector number

T → sum of number of the generations

C → random integer equal to T divided by 25

S_r → arbitrarily generated solution

X_i → updates its position using S_r

Heap Based Optimization

- **Overall Position Update**

$$X_i^k(t+1)=\begin{cases} X_i^k,\ p\le p_1 \\ B^k+y\lambda^k\left|B^k-X_i^k(t)\right|,\ p_1<p\le p_2 \\ S_r^k+\gamma\lambda^k\left|S_r^k-X_i^k(t)\right|,\ p_1<p\le p_3\ \&\ f(S_r)<f(X_i) \\ X_i^k(t)+\gamma\lambda^k\left|S_r^k-X_i^k(t)\right|,\ p_1<p\le p_3\ \&\ f(S_r)<f(X_i) \end{cases} \qquad (4)$$

$$p_1=1-\frac{t}{T},\ p_2=p_1+\frac{1-p_1}{2},\ p_3=p_1+\frac{1-p_1}{2} \qquad (5)$$

$$Parent_i=\left\lfloor\frac{i+1}{d}\right\rfloor \qquad (6)$$

(7)

child(i,j)=j-d+1+d*i

$$depth_i=\left\lceil log_d\left(d*i-i+1\right)\right\rceil-1 \qquad (8)$$

$$\left[\frac{d*d^{depth_i-1}}{d-1}+1,\frac{d*d^{depth_i-1}}{d-1}\right] \qquad (9)$$

Nomenclature

p → random number between 0 & 1

Proposed HBO-Stacked Ensemble Model

- ML models have multiple hyperparameters that require optimization for improved performance.
- Hyperparameters can be selected through testing different values and selecting the best performer or using default values.
- HBO algorithm is implemented to tune the hyperparameters of a stacked ensemble model.
- The selected fitness function for this research is the RMSE, mathematical representation of the given expression is as follows,

$$RMSE\left(\hat{y}\right)=\sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(y_i-\hat{y}\right)^2}$$

- The hyperparameters targeted for optimization are:
  - maximum depth
  - number of estimators
  - maximum number of samples
  - maximum number of features.
- The search space for these hyperparameters is defined with upper and lower limits ([10,100,0.1,0.1] and [50,400,1,1] respectively)
- The application of the heap-based optimization technique improves the accuracy of predicted energy consumption values.

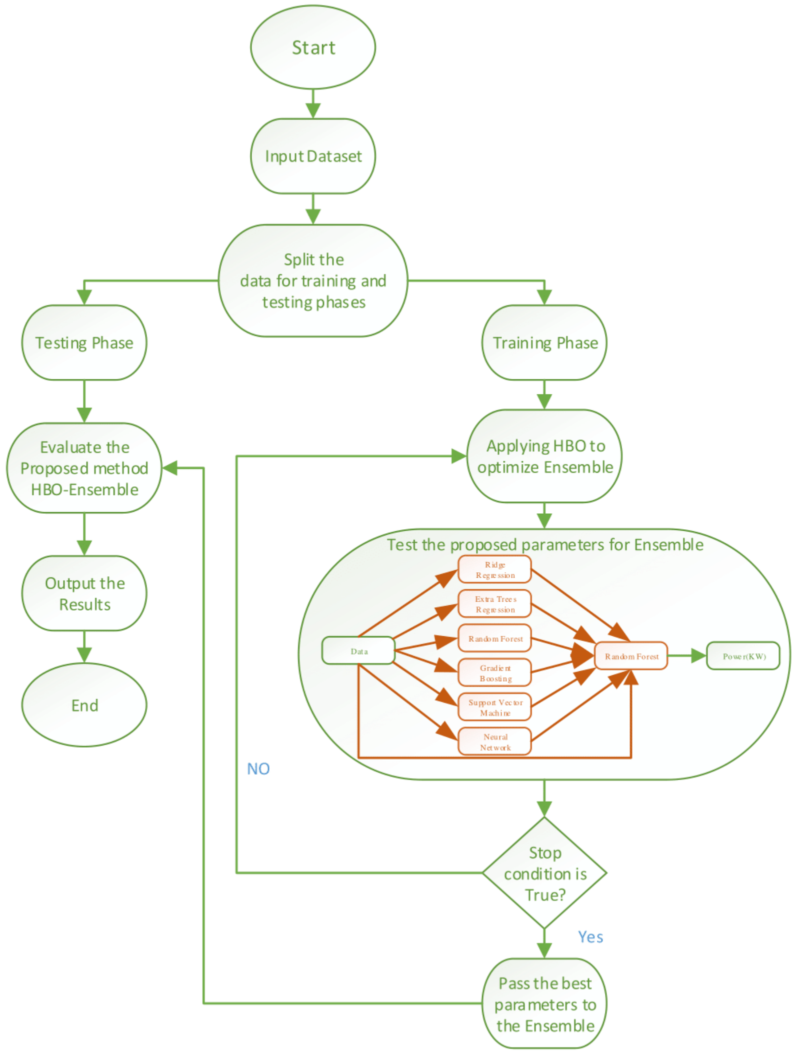


Fig.5 Flowchart Proposed HBO-Ensemble Model

Case Study

Data is collected directly from the PGEN database

→

Collected data is hourly for climatic data and electrical data with hourly and 1 minute time granularity respectively from the period of 01 Feb., 2018 to 13 Dec., 2019

→

To fetch the data, used SQL

→

Due to the fact that electrical and climatic data are maintained separately, both should be retrieved and preprocessed independently prior to integration.

Features of Collected Data

Climatic Data

- Temperature (◦C)
- Atmospheric pressure (hPa)
- Precipitation (mm)
- Windspeed (m/sec)

Electrical Data

- Total Active Power(kW)
- Voltage (kV)
- Power Factor (dimensionless)

Kernel Density Estimation Plots

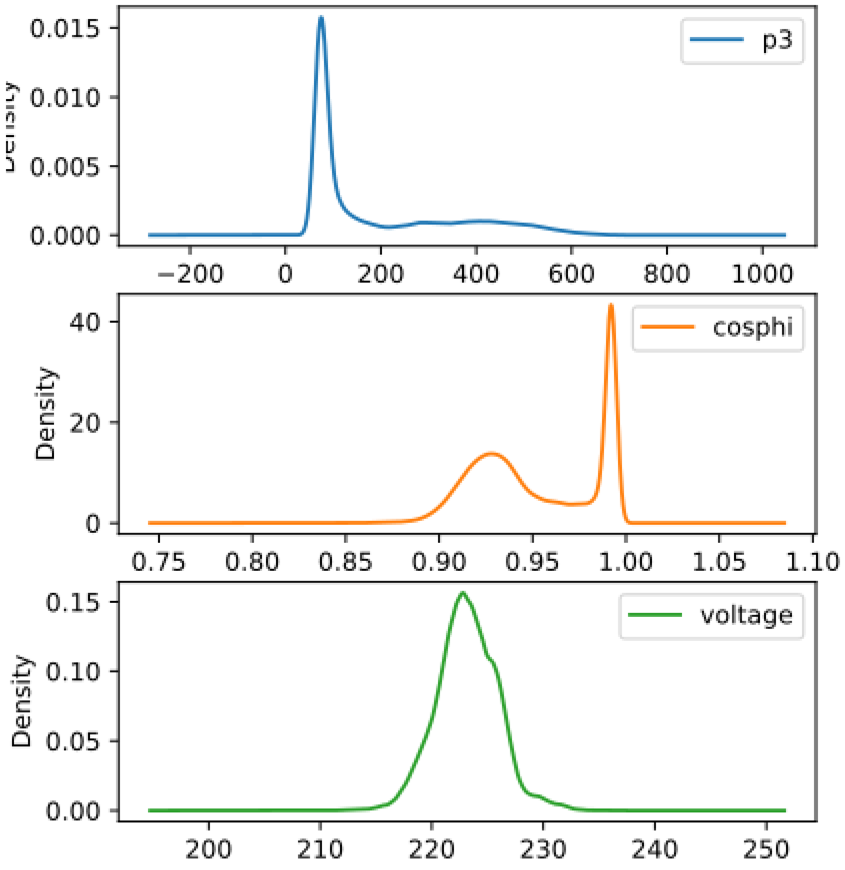


Fig. 6. KDE of electrical variables

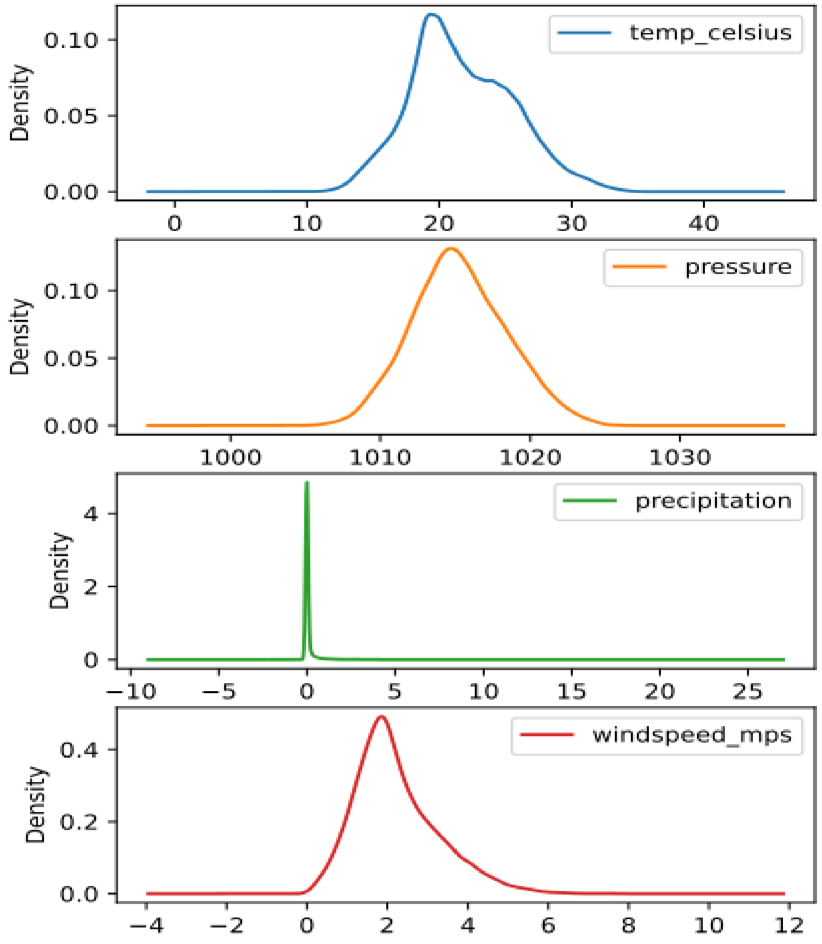


Fig. 7. KDE of climatic variables

Pre-processing and Data Integration

1. Probabilistic Density Calculation
2. Standardization of both the data
3. Load Factor Calculated
4. Separate Characteristics determined
5. Noisy Data Removed
6. Again data is Normalized
7. At end only active power is left unnormalized

$$Load\ Factor = \frac{Maximum\ Active\ Power\,(P_{max})}{Mean\ of\ all\ Powers\,(P_{mean})}$$

Correlation Plot

	Highest	Moderate	Weak or No
Positive	*cos_phi* (0.748)	*temperature* (0.413)	*hour* (0.327)
Negative	*pressure* (-0.320)	*cos_phi_std* (-0.421)	*voltage, load_factor, day_of_week, windspeed, month, and precipitation* (-0.03 to 0.10)

- The most significant variables associated with *p*3 are *cos_phi*, *temperature*, *pressure*, and *cos_phi_std*. These variables can be valuable in predicting *p*3 values and making informed decisions regarding power utilization and energy efficiency.

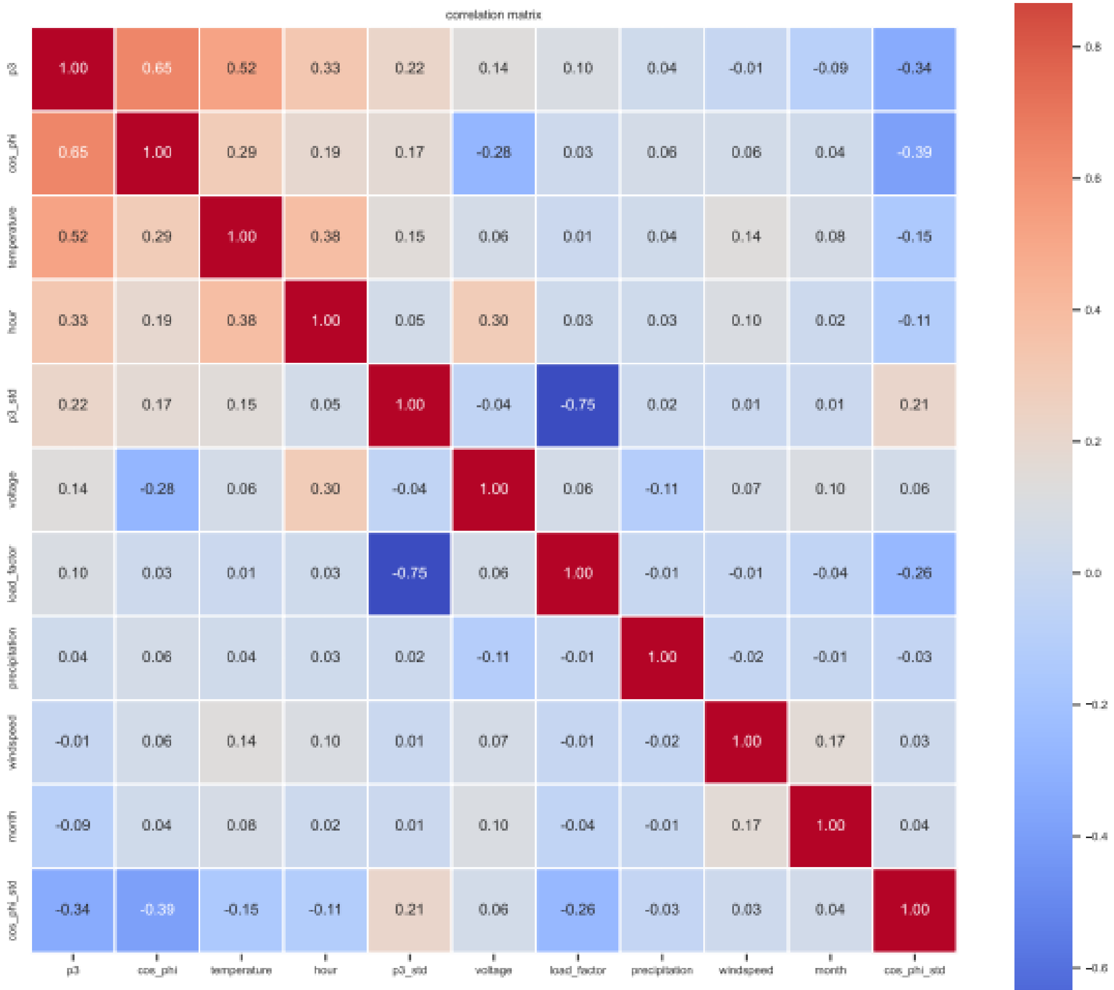


Fig.8 Correlation Plot

Model Training

- Used cross-Validation method: 70% of the data for training, 30% for tests i.e. out of 62509 data points 43756 is used for model training and 18753 is used for evaluating the errors
- Here training group is used to model the function
- While test group is used to evaluate the predictor error
- 6 Machine Learning techniques applied, chosen to have different strengths and weaknesses: Ridge Regression, Random Forest Regression, Extremely randomized trees, Gradient Tree Boosting, and Artificial Neural Network
- Each model is trained and evaluated separately
- Evaluation Metrics

$$RMSE(\hat{y}) = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y})^2} \qquad MAE(\hat{y}) = \frac{1}{n}\sum_{i=1}^{n}|y_i - \hat{y}_i|$$

$$Perc(\hat{y}) = \frac{MAE(\hat{y})}{Mean(\hat{y})}$$

y_i →output value of i^{th} energy

$\hat{y}$ →Value predicted by algorithm

n →No. of samples from test dataset

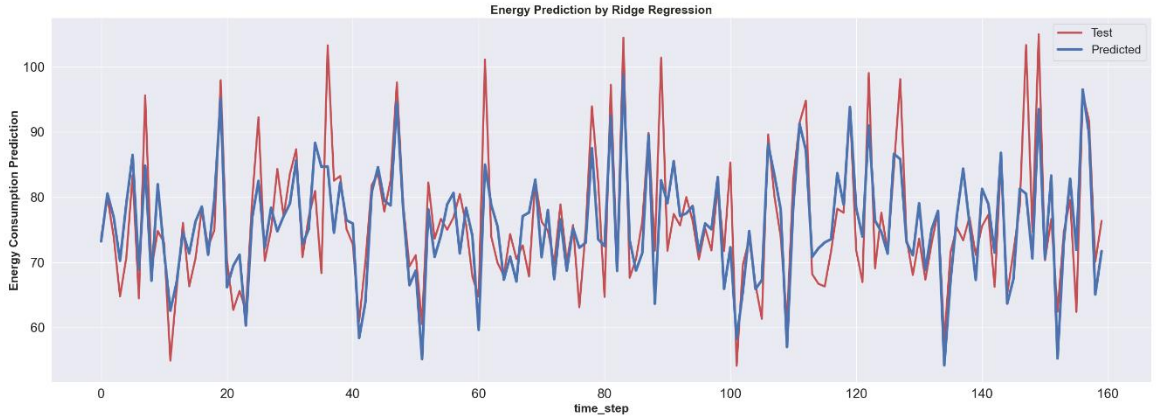
Evaluation of Models

Table-I Model Evaluation and Error Comparison

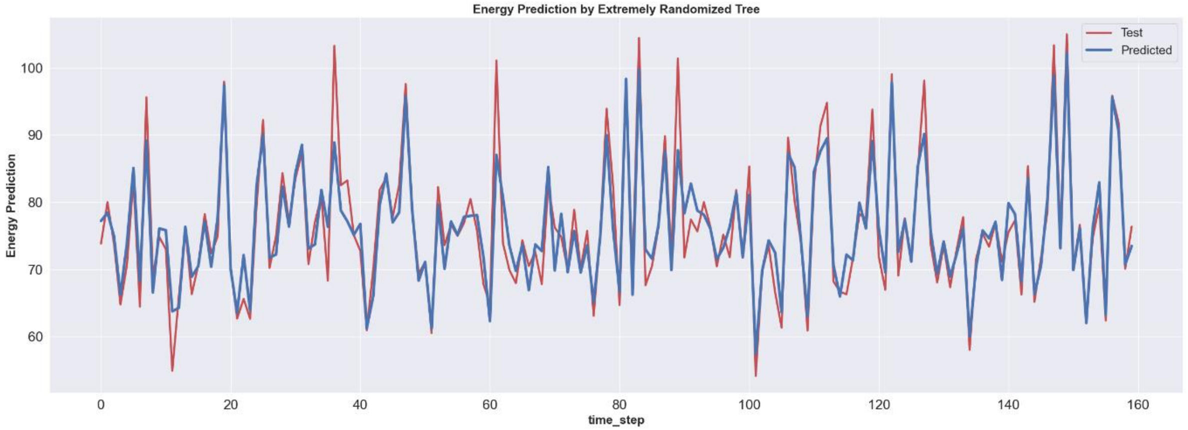
Algorithm	MAPE	RMSE	Percentual
Ridge Regression	0.058	5.76	5.82
Extremely Randomized Trees	0.029	3.20	3.02
Random Forest	0.029	3.20	2.99
Gradient Tree Boosting	0.029	3.10	2.99
Artificial Neural Networks	0.034	3.75	3.52
Support Vector Machine	0.053	5.11	5.24
Stacked Ensemble	0.028	3.06	2.86
Proposed HBO+Stacked Ensemble	0.027	3.05	2.85

Results

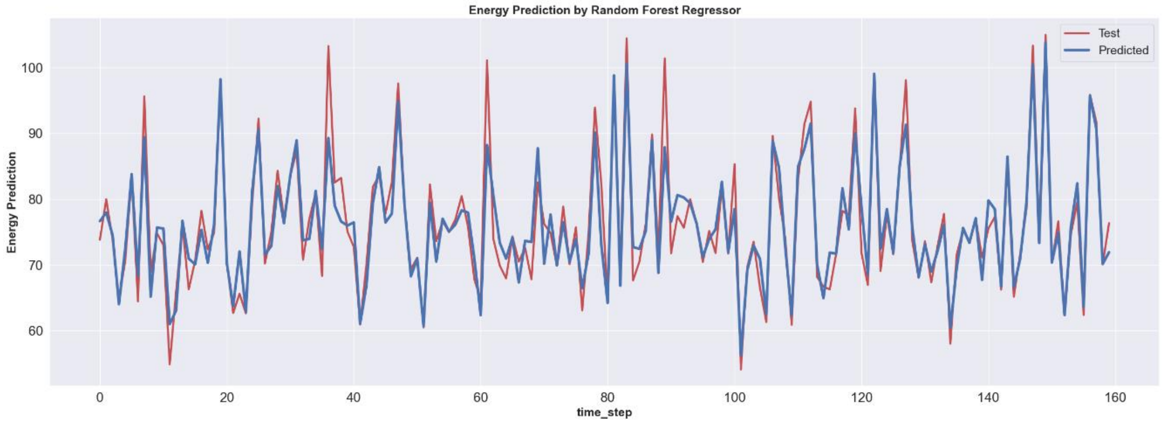
Ridge Regression



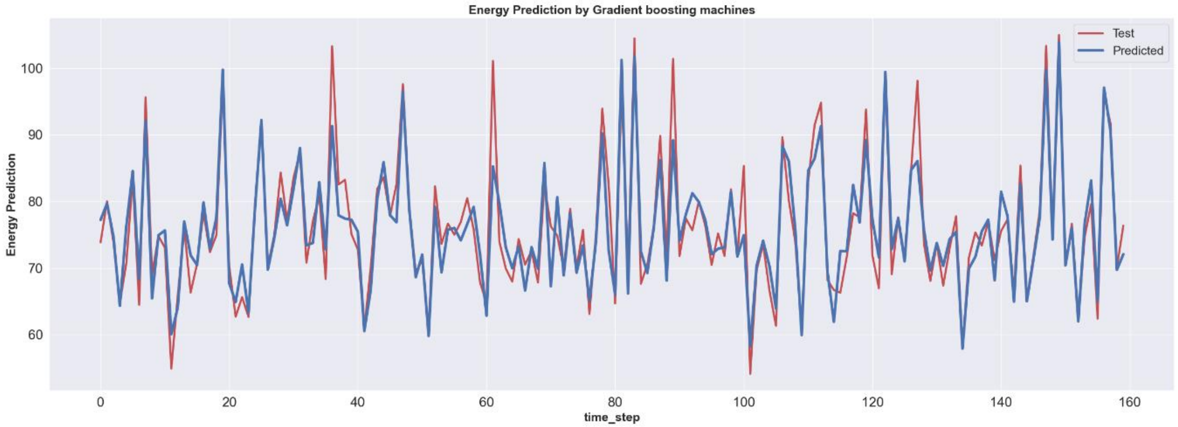
Extremely Randomized Trees



Random Forest



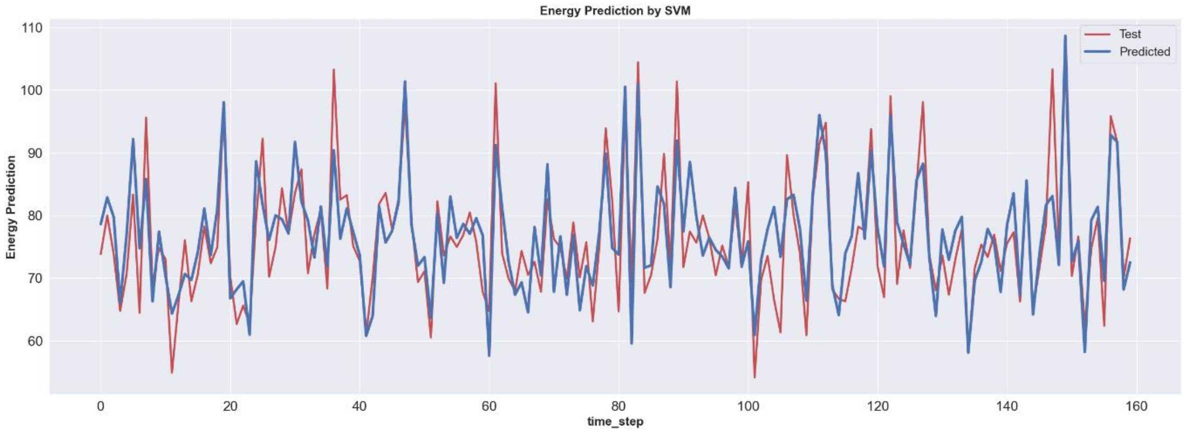
Gradient Boosting Regressor

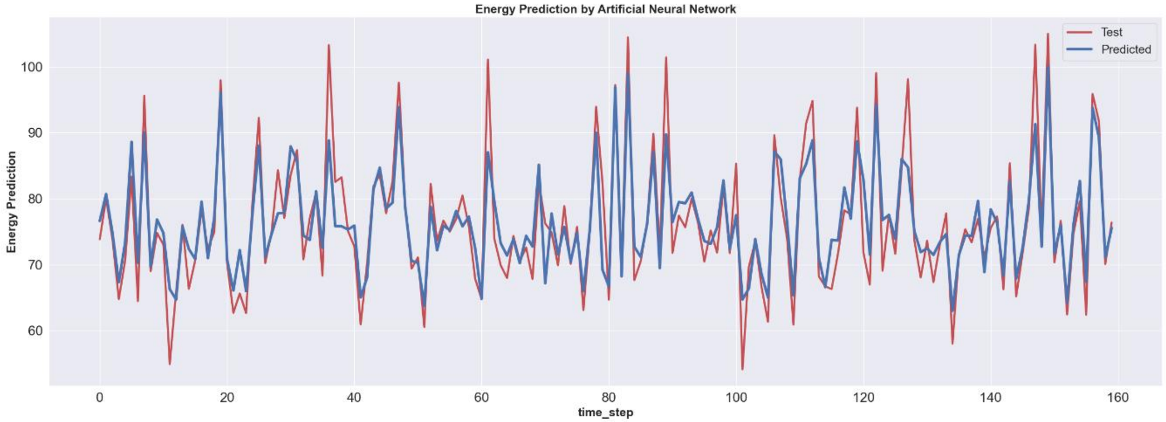


Results

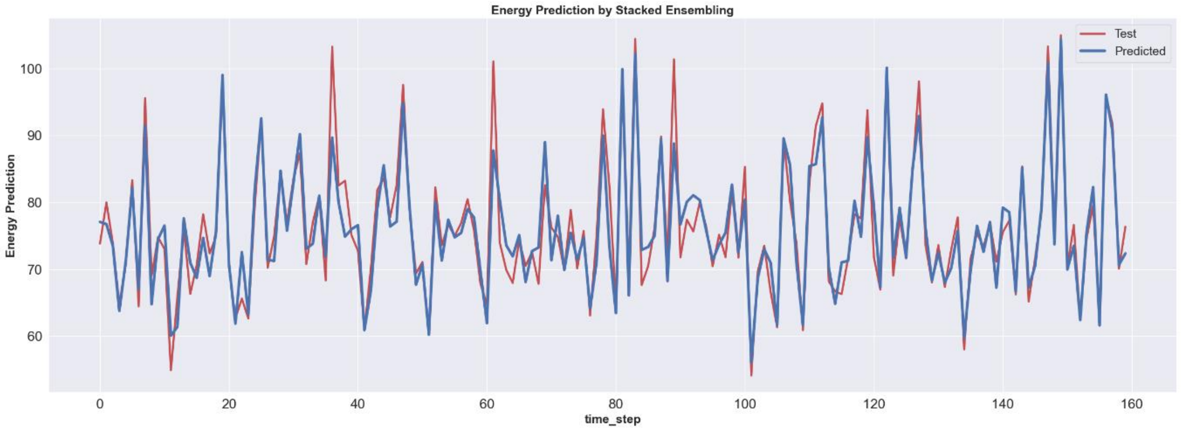
SVM

ANN



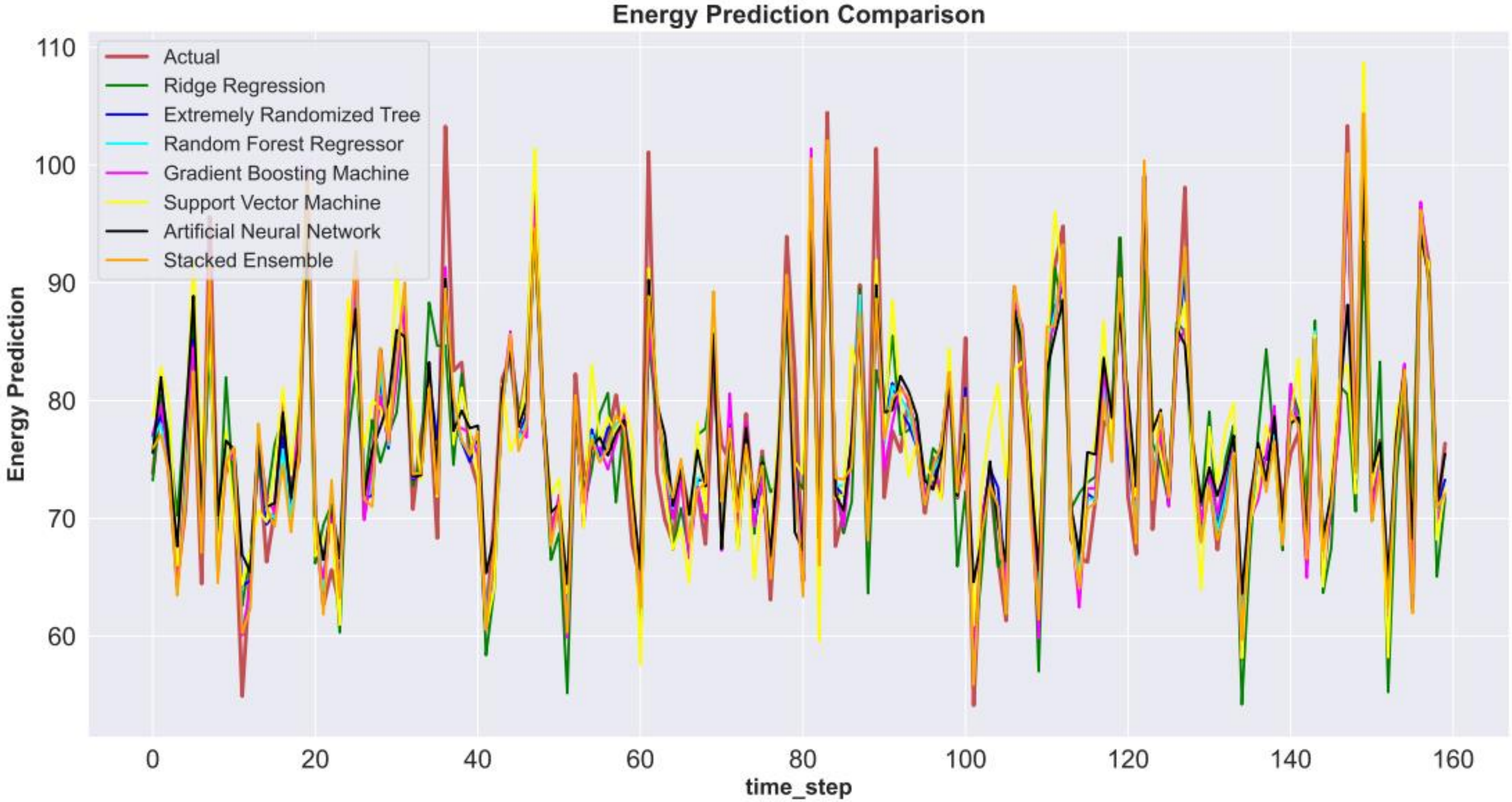


Stacked Ensemble



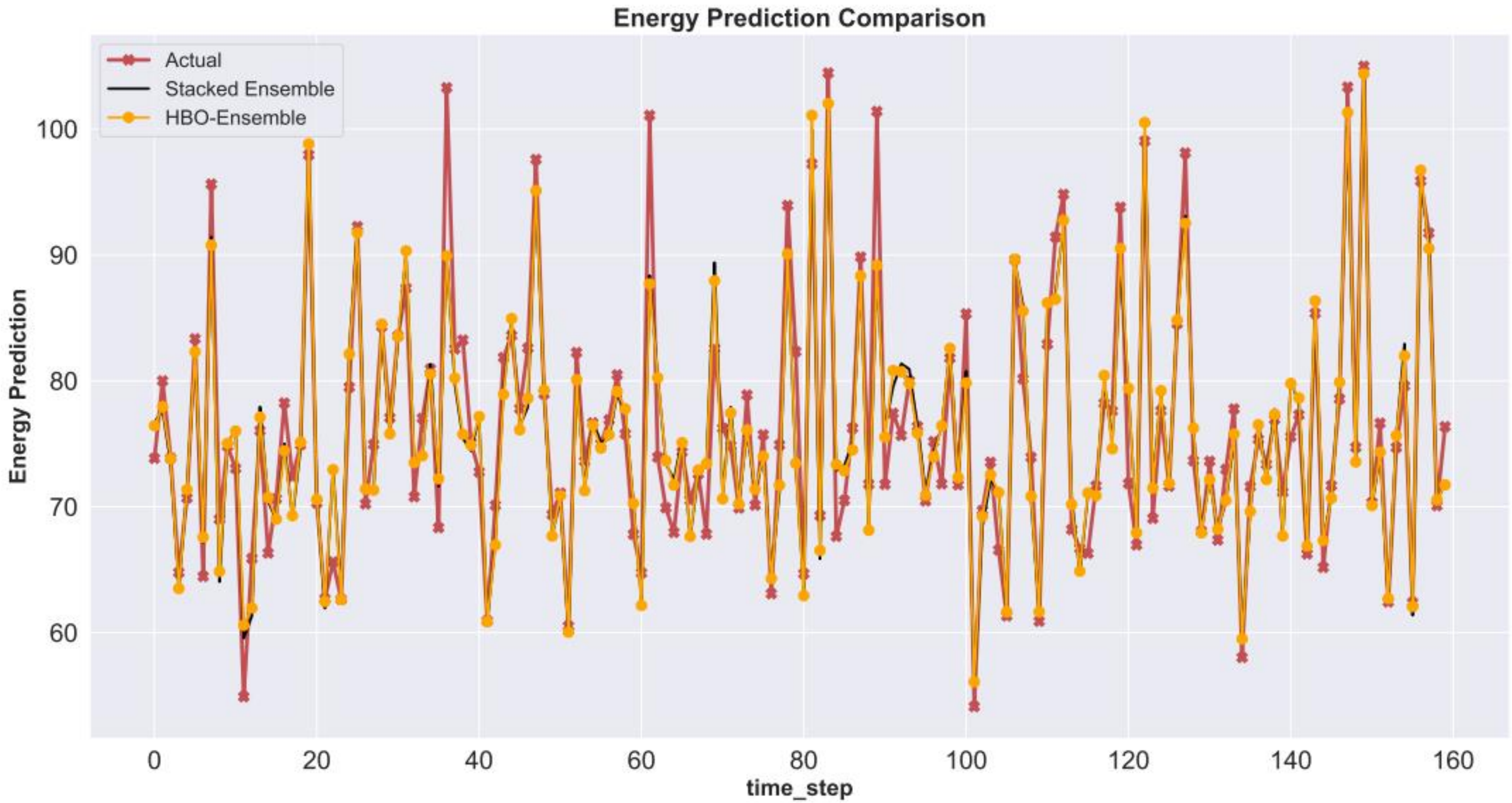
Results

Comparison of All Models with Ensemble Model



Results

Comparison of Stacked Ensemble Model with HBO-Ensemble Model



Conclusion

- Individual algorithms such as Extremely Randomized Trees, Random Forest, and Gradient Tree Boosting showed superior predictive performance with the lowest MAPE, RMSE, and Percentual values.
- The ensemble model demonstrated improved results compared to the individual algorithms, achieving a lower MAPE, RMSE, and Percentual.
- The proposed Heap-Based Optimization-Ensemble (HBO-Ensemble) algorithm outperformed all other approaches, including the ensemble model, by yielding the lowest MAPE, RMSE, and Percentual values.
- The novel optimization technique used in the HBO-Ensemble algorithm significantly enhanced the accuracy of energy consumption predictions.
- The study provides valuable insights into the application of machine learning algorithms for building energy consumption prediction.
- Ensemble methods, including the HBO-Ensemble algorithm, have the potential to achieve improved prediction accuracy.
- The findings of this study are essential for informing decision-making processes related to energy management and facilitating the development of energy-efficient strategies for buildings.

Future Scope

- Future research in this field could focus on expanding the dataset, incorporating additional relevant features, and exploring other optimization techniques to further improve prediction accuracy.
- Furthermore, investigating the generalizability of the proposed approach across different geographical regions and building types would provide a comprehensive understanding of its applicability

Publication

- P.Pahadia, P.Jain, A.K.Agarwal, A. Prajesh and A.Harit, ‘Stacked Ensemble Approach for Building Energy Consumption Prediction’ In 2023 14th International Conference on Computing, Communication and Networking Technologies (ICCCNT) [ACCEPTED]
- P.Pahadia, P.Jain, A.K.Agarwal, A. Prajesh and A.Harit. "HBO-Ensemble: Based Building Energy Consumption Prediction." Electric Power Systems Research. [DRAFTED]

References

[1] Z. Wang and R. S. Srinivasan, “A review of artificial intelligence based building energy prediction with a focus on ensemble prediction models,” in 2015 Winter Simulation Conference (WSC). IEEE, 2015, pp. 3438–3448.

[2] R. Olu-Ajayi, H. Alaka, I. Sulaimon, F. Sunmola, and S. Ajayi, “Building energy consumption prediction for residential buildings using deep learning and other machine learning techniques,” Journal of Building Engineering, vol. 45, p. 103406, 2022.

[3] L. S. B. Pereira, R. N. Rodrigues, and E. A. C. A. Neto, “Modeling of energy management systems using artificial intelligence,” in 2020 IEEE International Systems Conference (SysCon). IEEE, 2020, pp. 1–6..

[4] Doukas, K. D. Patlitzianas, K. Iatropoulos, and J. Psarras, “Intelligent building energy management system using rule sets,” Building and environment, vol. 42, no. 10, pp. 3562–3569, 2007.

[5] B. Mahapatra and A. Nayyar, “Home energy management system (hems): Concept, architecture, infrastructure, challenges and energy management schemes,” Energy Systems, vol. 13, no. 3, pp. 643–669, 2022.

References (Contd.)

[6] D. Amarnath and S. Sujatha, “Internet-of-things-aided energy management in smart grid environment,” The Journal of Supercomputing, vol. 76, no. 4, pp. 2302–2314, 2020

[7] H. Jo and Y. I. Yoon, “Intelligent smart home energy efficiency model using artificial tensorflow engine,” Human-centric Computing and Information Sciences, vol. 8, no. 1, pp. 1–18, 2018.

[8] Y. Zhou, Y. Chen, G. Xu, Q. Zhang, and L. Krundel, “Home energy management with pso in smart grid,” in 2014 IEEE 23rd International Symposium on Industrial Electronics (ISIE). IEEE, 2014, pp. 1666–1670.

[9] M. S. Ahmed, A. Mohamed, R. Z. Homod, and H. Shareef, “A home energy management algorithm in demand response events for household peak load reduction,” PrzeglAd Elektrotechniczny, vol. 93, no. 3, p. 2017, 2017.

[10] M. H. Elkazaz, A. Hoballah, and A. M. Azmy, “Artificial intelligent-based optimization of automated home energy management systems,” International Transactions on Electrical Energy Systems, vol. 26, no. 9, pp. 2038–2056, 2016.

References (Contd.)

[11] Lu, Chujie, Sihui Li, and Zhengjun Lu. "Building energy prediction using artificial neural networks: A literature survey." *Energy and Buildings* (2021): 111718.

[12] Shao, Minglei, Xin Wang, Zhen Bu, Xiaobo Chen, and Yuqing Wang. "Prediction of energy consumption in hotel buildings via support vector machines." Sustainable Cities and Society 57 (2020): 102128.

[13] Askari, Qamar, Mehreen Saeed, and Irfan Younas. "Heap-based optimizer inspired by corporate rank hierarchy for global optimization." Expert Systems with Applications 161 (2020): 113702.

[14] A. Parizad and C. Hatziadoniu, "Deep Learning Algorithms and Parallel Distributed Computing Techniques for High-Resolution Load Forecasting Applying Hyperparameter Optimization," in IEEE Systems Journal, vol. 16, no. 3, pp. 3758-3769, Sept. 2022, doi: 10.1109/JSYST.2021.3130080.

[15] Ewees, Ahmed A., Mohammed AA Al-qaness, Laith Abualigah, and Mohamed Abd Elaziz. "HBO-LSTM: Optimized long short term memory with heap-based optimizer for wind power forecasting." Energy Conversion and Management 268 (2022): 116022.

Thank You