Predictive Modeling of Energy Consumption: A Comparative Analysis of Machine Learning and Deep Learning Approaches

1. ABSTRACT

This study investigates energy consumption prediction in the PJM Interconnection LLC (AEP_hourly) dataset, spanning from December 31, 2004, to January 2, 2018, using a variety of machine learning (ML) and deep learning (DL) models. Key performance metrics, including Root Mean Square Error (RMSE), Mean Squared Error (MSE), and Mean Absolute Percentage Error (MAPE), are employed to assess the models' accuracy. The results reveal that XGBoost yields the most accurate predictions with the lowest MAPE. The study's findings provide valuable insights into model selection for enhancing energy consumption forecasting in the context of PJM Interconnection LLC.

2. INTRODUCTION

In an era marked by growing concerns about energy sustainability and efficiency, accurate energy consumption prediction is of paramount importance. This research paper addresses the crucial task of forecasting energy consumption, a key element in optimizing resource allocation, reducing costs, and promoting eco-friendly practices. The study centers on the PJM Interconnection LLC (AEP_hourly) dataset, covering the period from December 31, 2004, to January 2, 2018, characterized by its temporal nature. Our research aims to evaluate a range of machine learning (ML) and deep learning (DL) models, including traditional linear regression and advanced techniques like LSTM, to determine their efficacy in predicting energy consumption. By doing so, we contribute to the broader discourse on energy management and forecast accuracy, offering valuable insights into the most effective modeling approaches for this critical domain.

3. LITERATURE REVIEW

The importance of predicting energy usage in buildings for efficient energy management, lighting, equipment, and HVAC system optimization, as well as cooperation between facility managers and power suppliers, has been recognized. Various statistical and machine learning models have been applied to predict energy consumption patterns in buildings, including linear regression, random forest (RF), support vector regression (SVR), deep neural networks, and adaptive learning-based models. Short-term load prediction, crucial for HVAC system control, has been achieved through hybrid models that combine artificial neural networks (ANNs) with evolutionary optimization algorithms. Additionally, innovative approaches, such as non-intrusive occupant load monitoring (NIOLM) and generative adversarial networks (GANs), have been proposed to address data limitations and enhance the accuracy of energy consumption predictions.

Specifically, RF models have emerged as a powerful tool in solving regression problems, exhibiting enhanced predictive accuracy compared to other techniques like regression trees and SVR. However, while machine learning models have been used extensively for building energy forecasting, there is a need for further research to explore their effectiveness in predicting energy consumption across multiple buildings. This study aims to examine the potential of RF models for short-term energy consumption prediction in buildings.

4. Methodology

To prepare the dataset for analysis, we applied a Min-Max scaler to normalize the data, ensuring consistent scaling across features. Additionally, we performed feature engineering by extracting temporal information from the datetime data, resulting in features such as 'day of year,' 'hour,' 'day of week,' 'quarter,' 'month,' and 'year.' The target variable for our predictions was 'AEP_MW,' representing energy consumption.

In terms of the machine learning and deep learning algorithms employed, our toolkit encompassed a variety of models. This included traditional approaches such as Linear Regression and K-Nearest Neighbors (KNN), alongside advanced techniques like Artificial Neural Networks (ANN) and Long Short-Term Memory (LSTM) networks, ensuring a comprehensive exploration of predictive capabilities.

For evaluating the performance of these models, we utilized a suite of essential metrics:

Mean Squared Error (MSE), the mean squared error (MSE) or mean squared deviation (MSD) of an estimator measures the average of the squares of the errors

$$MSE = \frac{\sum(y_i - \hat{y}_i)^2}{n}$$

Mean Absolute Percentage Error (MAPE), assessing the percentage difference between predictions and true values, providing a relative error measure.

$$M = \frac{1}{n} \sum_{t=1}^{n} \left| \frac{A_t - F_t}{A_t} \right|$$

Root Mean Square Error (RMSE), quantifying the square root of the average squared differences between predicted and actual values.
These metrics were selected due to their ability to provide a well-rounded evaluation of model accuracy, capturing different aspects of prediction performance.

$$\mathrm{RMSE}_{fo} = [\sum_{i=1}^{N} (z_{f_i} - z_{o_i})^2 / N]^{1/2}$$

5. **Experimental Setup:**

The dataset was divided into two distinct sets for training and testing purposes, following a chronological split. Observations up to 'January 1, 2015,' were allocated to the training set, while data from and after this date formed the test set. This approach ensures that the models were trained on historical data and evaluated on more recent observations, replicating real-world forecasting scenarios.

To optimize model performance, hyperparameter tuning and cross-validation were essential components of our experimental setup. GridSearchCV, a robust technique, was employed to systematically explore hyperparameter combinations and identify the most effective settings for each model. Furthermore, it's noteworthy that all model training was conducted on a single T4 GPU, ensuring consistency and comparability in the computational environment. This comprehensive experimental configuration aimed to produce accurate and robust energy consumption predictions while addressing potential overfitting concerns.

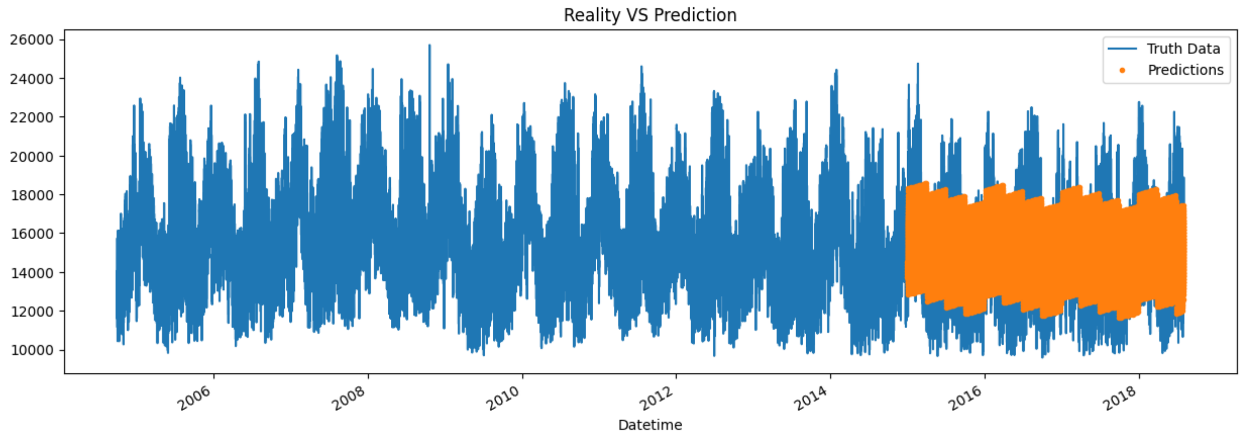
6. **Results**

The below are the following models used and their plots with predicted values on the test data.

Machine Learning Algorithms (ML):

Linear Regression: A fundamental ML approach that models a linear relationship between features and target, often used for its simplicity.

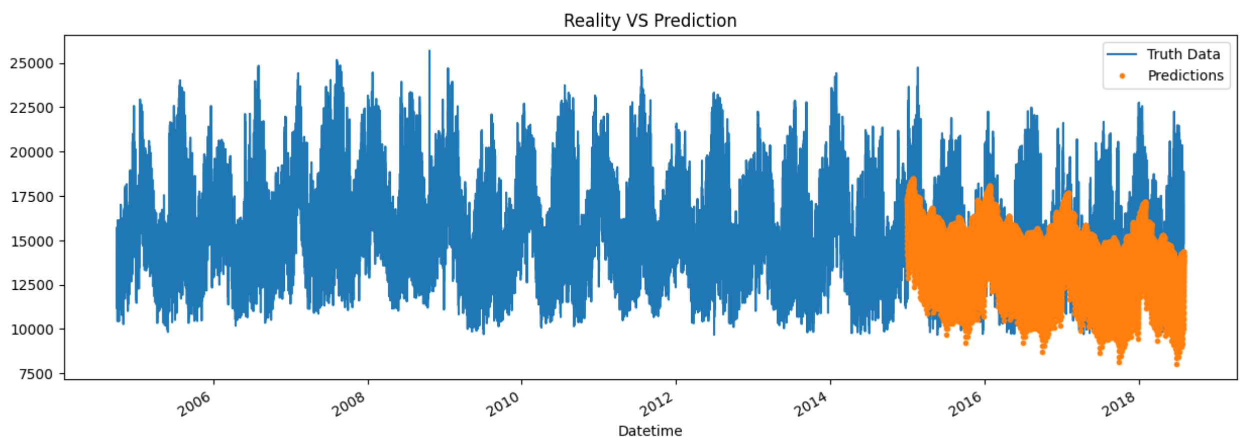
$$Y = \beta_0 + \beta_1 X_1$$



LINEAR REGRESSION

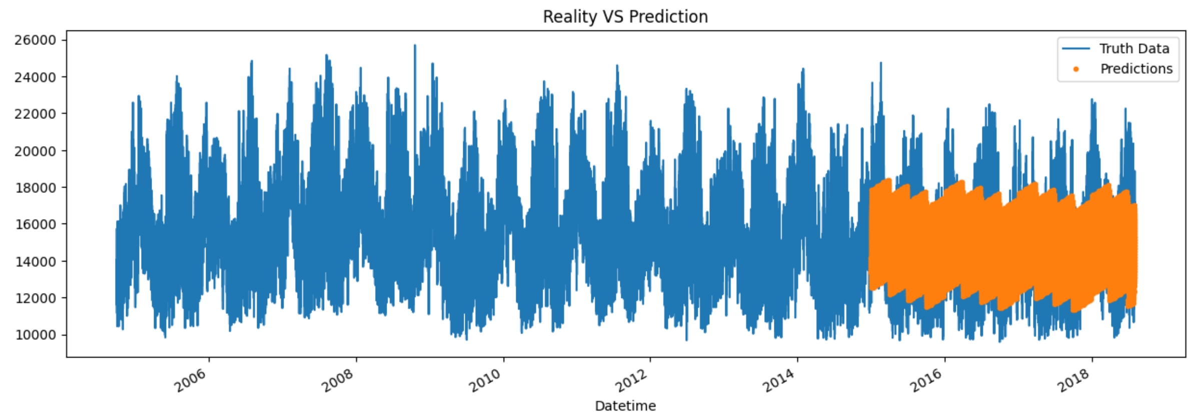
Polynomial Regression: An extension of linear regression that captures non-linear relationships by introducing polynomial terms.

$y= b_0+b_1x_1+ b_2x_1^2+ b_2x_1^3+...... b_nx_1^n$



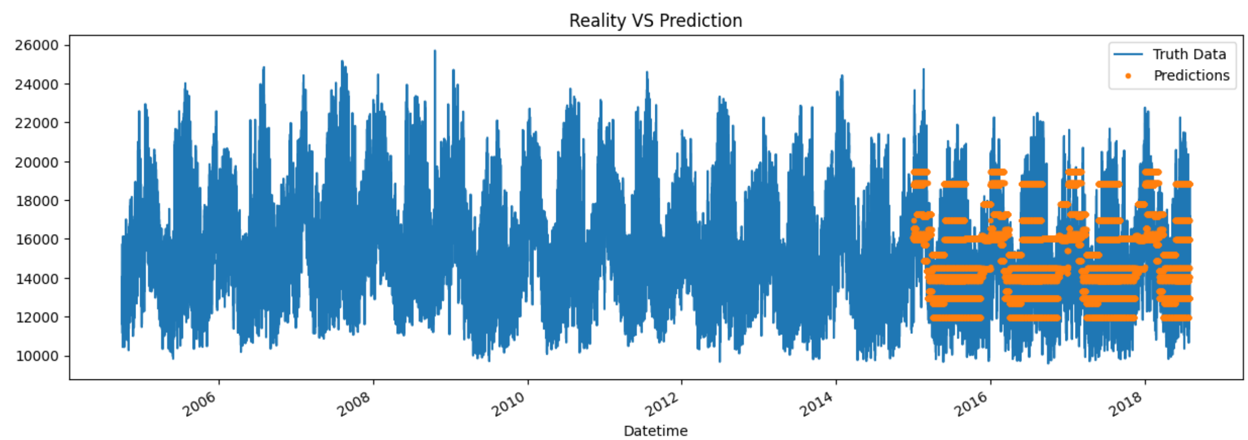
POLYNOMIAL REGRESSION

Linear SVR (Support Vector Regressor): A regression technique based on support vector machines, effective in handling both linear and non-linear data.



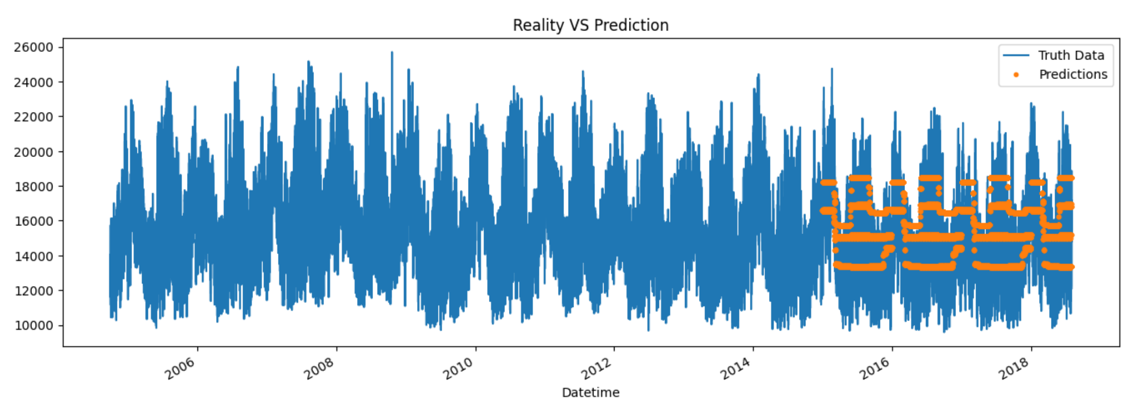
LINEAR SVR

Decision Tree: A tree-like model that makes decisions by splitting data into branches, widely used for its interpretability.



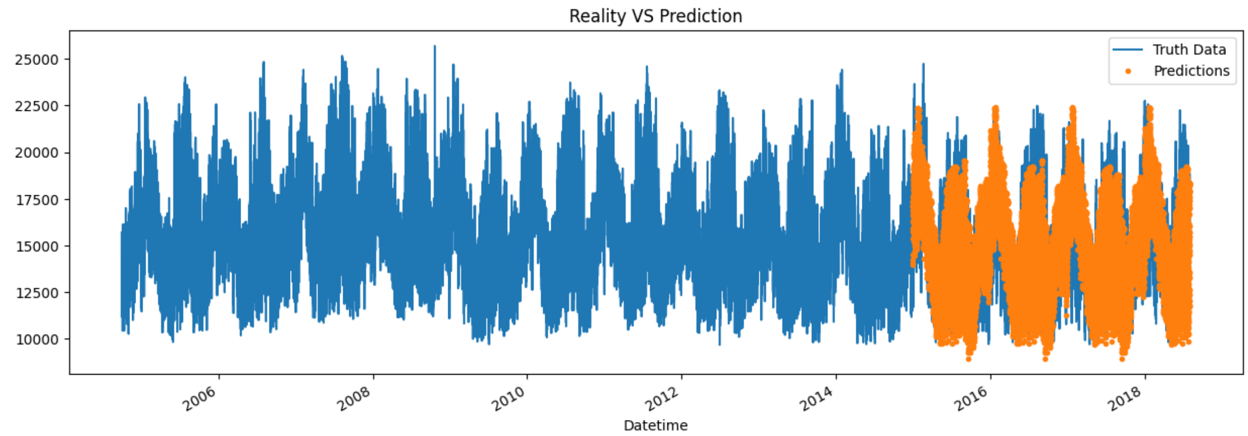
Decision Trees

Random Forest: An ensemble method employing multiple decision trees to enhance prediction accuracy and reduce overfitting.



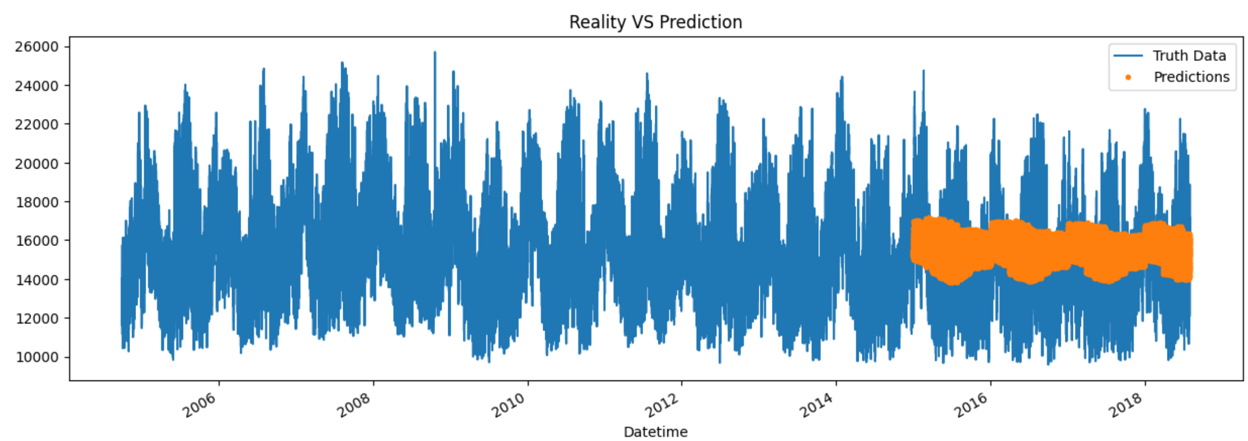
RANDOM FOREST

Gradient Boosting Regressor: A boosting algorithm that combines multiple weak models to create a strong predictive model.

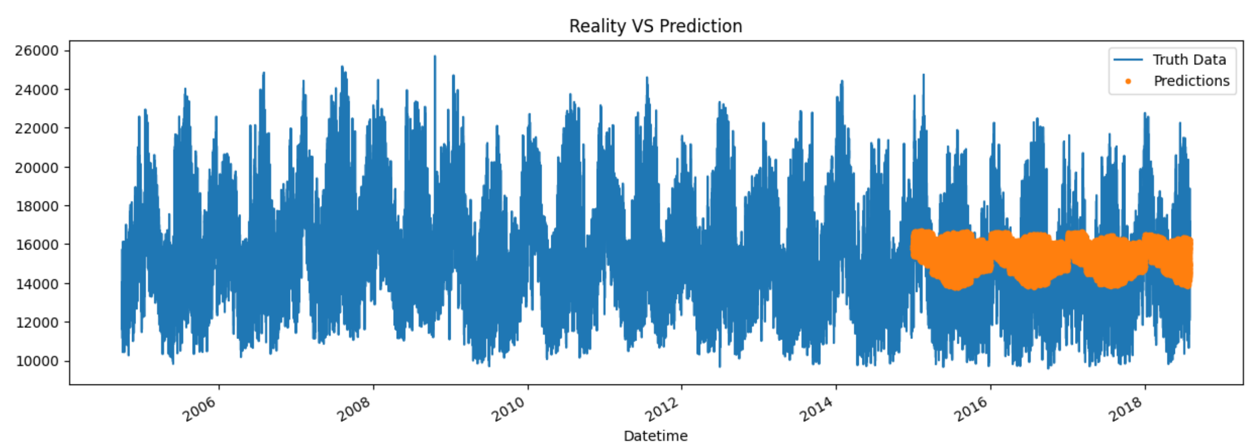


GRADIENT BOOSTING REGRESSOR

Bagging and Pasting with K-Nearest Neighbors (KNN): Techniques that use K-Nearest Neighbors for predicting values by aggregating neighboring data points.

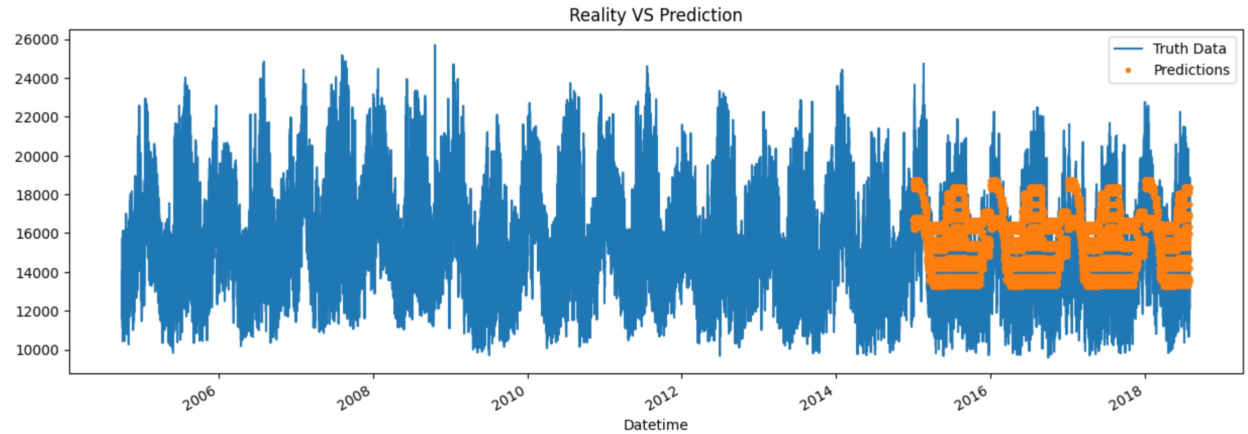


PASTING WITH KNN



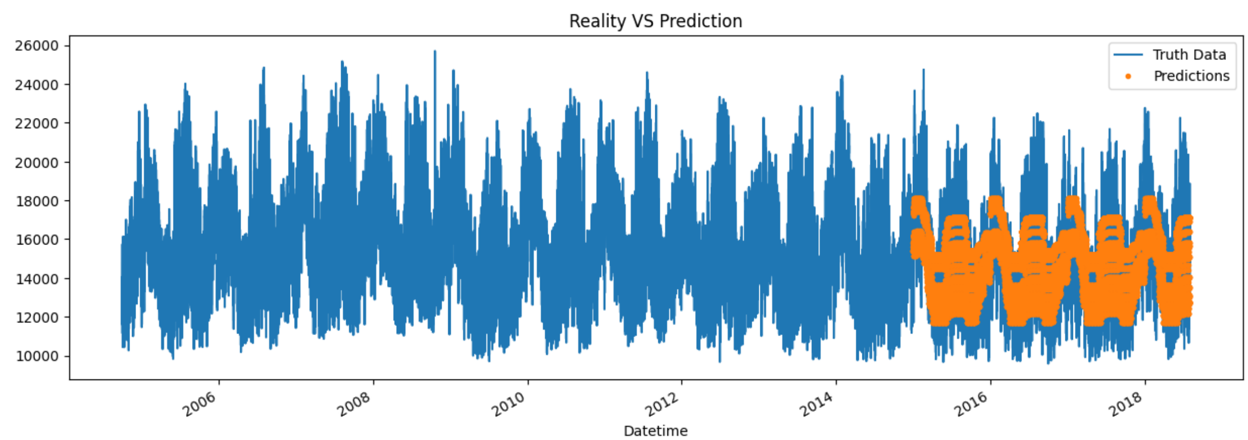
BAGGING WITH KNN

Adaptive Boosting with Random Forest Regressor: A boosting approach that improves the performance of Random Forest models through iterative learning.



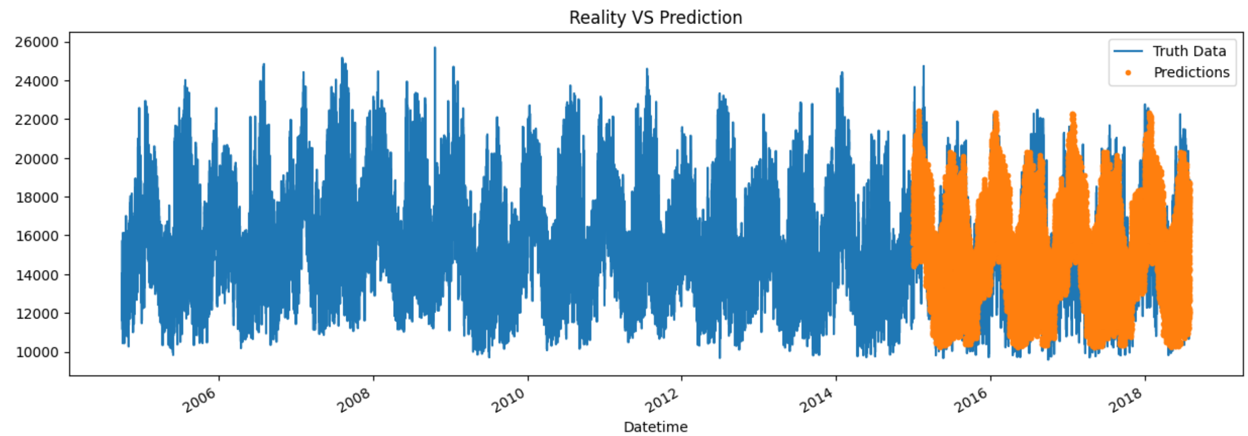
ADAPTIVE BOOSTING WITH RANDOM FOREST REGRESSOR

XGBoost: A popular gradient boosting framework known for its speed and accuracy in predictive modeling.



XGBoost

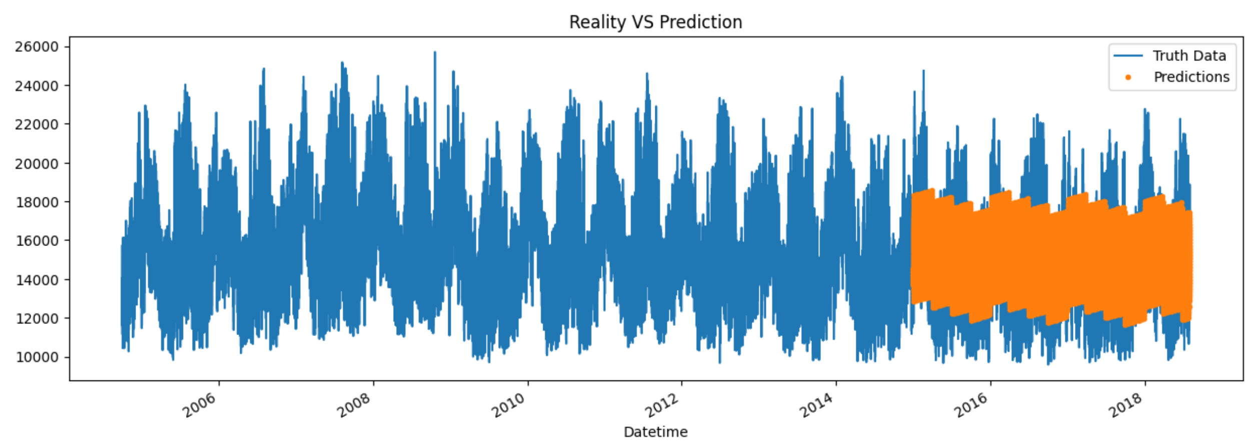
Stacking Regressor (Meta-regressor as Support Vector Regressor): An ensemble technique that combines the predictions of multiple models, with a Support Vector Regressor as the meta-learner.



STACKING REGRESSOR

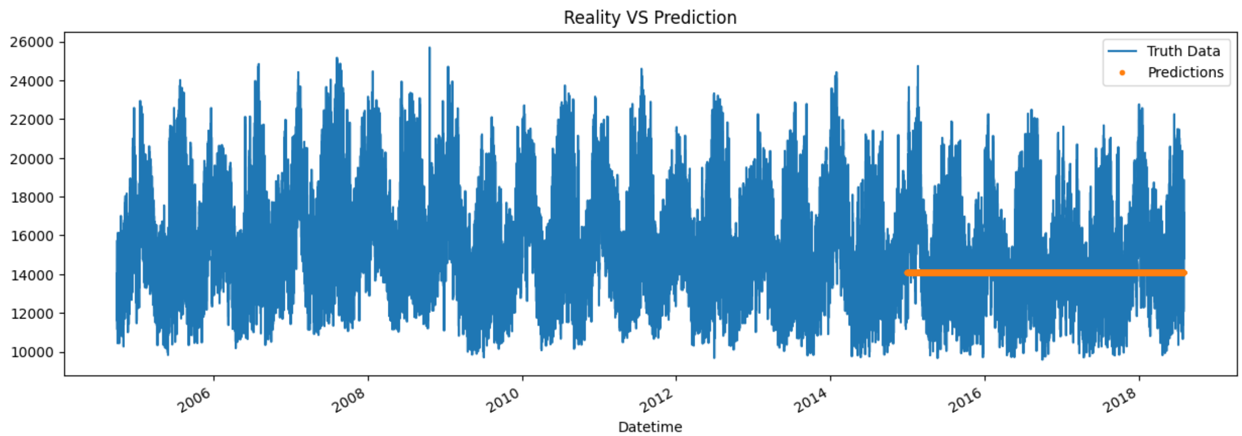
Deep Learning Algorithms (DL):

Artificial Neural Network (ANN): A deep learning model inspired by the human brain, consisting of interconnected layers of neurons for complex pattern recognition.



ANN

Long Short-Term Memory (LSTM): A type of recurrent neural network (RNN) designed for sequential data, ideal for capturing temporal dependencies in time series forecasting.



LSTM

The below table displays the RMSE, MSE and MAPE values for the above mentioned models.

	RMSE	MSE	MAPE
1. Linear Regression	2225.74	4953933.50	12.55
2. Polynomial Regression	2309.80	5335203.37	11.34
3. Linear SVR	2209.80	4883249.54	12.01
4. Decision Tree	1887.93	3564305.9	10.82
5. Random Forest	2035.89	4144866.36	12.25
6. Gradient Boosting Regressor	1803.87	3253952.88	9.36
7. Bagging with KNN	2303.72	5307155.73	13.72

8.Pasting with KNN	2326.71	5413582.65	13.82
8. Adaptive Boosting with RF	2078.15	4318731.05	12.85
9. XGBoost	**1649.42**	**2720589.36**	**8.95**
10. Stacking Regressor	1800.47	3241711.71	9.2
11. ANN	2226.20	4955984.50	12.56
12. LSTM	2617.36	6850604.38	13.31

In our analysis, it was observed that XGBoost exhibited superior performance in terms of Mean Absolute Percentage Error (MAPE) when compared to the Deep Learning (DL) models. This implies that the XGBoost, yielded more accurate predictions of energy consumption. This outcome demonstrates the effectiveness of an ensemble approach in enhancing predictive accuracy and highlights the practical advantages of leveraging a diverse set of machine learning models for energy consumption forecasting.

7. Conclusion:

In conclusion, this study has shed light on the effectiveness of a diverse range of machine learning algorithms in predicting energy consumption, with the XGBoost emerging as a standout performer, surpassing even the deep learning models. This underscores the value of combining the strengths of various ML techniques to enhance predictive accuracy, making it a promising avenue for future energy forecasting applications.

Looking ahead, the incorporation of more advanced deep learning models, such as LSTM+CNN architectures, presents an exciting opportunity to further improve forecasting precision. Additionally, the augmentation of the dataset through Generative Adversarial Networks (GANs) and Variational Autoencoders (VAEs) can be explored to expand the dataset's size, potentially unlocking even more accurate and robust energy consumption predictions. The future of energy forecasting holds great promise, and continued research in these areas can contribute significantly to sustainable energy management and efficiency.

8. References

[1] Abdelkader, S.S., Grolinger, K., Capretz, M.A.M., 2015. Predicting energy demand peak using M5 model trees. In: 2015 IEEE 14th International Conference on Machine Learning and Applications (ICMLA), pp. 509e514.
MAE
[2]Allouhi, A., El Fouih, Y., Kousksou, T., Jamil, A., Zeraouli, Y., Mourad, Y., 2015. Energy
consumption and efficiency in buildings: current status and future trends.
J. Clean. Prod. 109, 118e130.

[3]Amasyali, K., El-Gohary, N.M., 2018. A review of data-driven building energy consumption
prediction studies. Renew. Sustain. Energy Rev. 81, 1192e1205.
Breiman, L.J.M.L., 2001. Random Forests, 45 (1), 5e32.

[4]Chan, J.C.-W., Paelinckx, D., 2008. Evaluation of Random Forest and Adaboost treebased
ensemble classification and spectral band selection for ecotope mapping
using airborne hyperspectral imagery. Rem. Sens. Environ. 112 (6), 2999e3011.

[5]Chou, J.-S., Bui, D.-K., 2014. Modeling heating and cooling loads by artificial intelligence
for energy-efficient building design. Energy Build. 82, 437e446.

[6]Fan, C., Xiao, F., Zhao, Y., 2017. A short-term building cooling load prediction method using deep learning algorithms. Appl. Energy 195, 222e233.

[7]Foucquier, A., Robert, S., Suard, F., Stephan, L., Jay, A., 2013. State of the art in building modelling and energy performances prediction: a review. Renew. Sustain. Energy Rev. 23, 272e288.

[8]Goudarzi, S., Anisi, M.H., Kama, N., Doctor, F., Soleymani, S.A., Sangaiah, A.K., 2019.
Predictive modelling of building energy consumption based on a hybrid natureinspired optimization algorithm. Energy Build. 196, 83e93.