Journal of Harbin Engineering University
ISSN: 1006-7043

Vol 47 No. 03
March 2026

Electricity Consumption Prediction Using a Multi-Model Machine Learning Framework

Mr. Kapil Jaiswal[l*1], Dr.Ashish Bansal [1,2], Dr.Puja Gupta[2]

[1,2]SoE, Avantika University, Ujjain[1]

[2]Department of Information Technology, Shri G.S. institute of Technology & Science, Indore

Abstract— Electricity demand prediction is crucial for efficient energy management in modern societies. In this research paper, we explore the efficacy of machine learning models in forecasting electricity consumption. Leveraging established ML models such as Long-Short term memory (LSTM), Random Forest, Support Vector Regression (SVR), we initially conduct predictions. However, to enhance predictive accuracy and address specific challenges inherent in prediction, we proposed modification in these traditional models. Our approach involves the development and utilization of novel ML architectures tailored to the electricity prediction task. We introduce three enhanced models: Deep LSTM, Random Forest combined with Neural Networks, and SVR integrated with Fuzzy Systems. These models are designed to capture intricate patterns, non-linear relationships, and uncertainties present in electricity consumption data. Our goal is to determine the best model for predicting power usage through empirical analysis and comparative review. Additionally, we provide insightful information on the advantages and disadvantages of each approach concerning energy management strategies and sustainable resource allocation.

Keywords— Machine Learning, Long Short-Term Memory, Deep LSTM, Random Forest, Neural Network, SVR

Introduction Precise forecasts of electricity consumption are critical for modern energy management systems. As societies worldwide shift toward more sustainable and efficient models, optimizing energy production, delivery, and consumption becomes paramount. However, power consumption patterns are complex and dynamic, posing challenges for traditional forecasting methods. To address this, sophisticated predictive modelling strategies powered by innovative machine learning algorithms have emerged as viable solutions. These algorithms offer improved prediction capabilities and adaptability across diverse data contexts.

Investigating Sophisticated Models

Our study focuses on several advanced machine learning models for power usage forecasting:

- **Support Vector Regression (SVR) with Neural Networks (NN):** This hybrid approach combines SVR’s robustness with NN’s ability to handle complex relationships.
- **Random Forest with Fuzzy Logic:** By merging Random Forest with fuzzy systems, we explore a versatile model capable of capturing intricate dependencies.
- **LSTM Networks with Deep LSTM Architectures:** These models excel at managing temporal dependencies, nonlinear relationships, and the intrinsic fuzziness of electricity consumption data.

Study Goals

Our primary objectives are as follows:

Thorough Evaluation: Assess each forecast model’s ability to effectively capture trends in both short- and long-term power usage.

- **Predictability, Efficiency, and Resilience:** Evaluate the predictability, computational effectiveness, and resilience of each model across various datasets and forecasting horizons.
- **Critical Elements:** Uncover critical factors that impact model performance, providing guidance on model selection and parameter adjustment.

Through the application of cutting-edge ML methodology, our research aims to support the advancement of more dependable and effective energy management systems. The results of our study have crucial ramifications for several stakeholders, such as customers, legislators, and utility companies. Equipped with precise forecasts of electricity use, utility firms can enhance operational efficiency and cost-effectiveness by optimizing resource allocation and scheduling energy output. Informed energy policies can be formulated by policymakers, and consumers can make informed decisions about their energy consumption to promote cost savings and energy conservation.

In the end, this research hopes to help make the shift to a future in energy that is more resilient and sustainable. Our objective is to enable stakeholders to proficiently handle the intricacies of electricity consumption dynamics by utilizing cutting-edge machine learning algorithms. We can all work together to create a more just and sustainable energy ecology, guaranteeing a better future for future generations, by making well-informed decisions and implementing proactive energy management techniques.

The structure of this document is as follows: Within Section 2, the literature on electricity consumption prediction is thoroughly reviewed, with an emphasis on methods, obstacles, and new developments. For each prediction model, Section 3 provides the theoretical framework and technique used, explaining the guiding ideas and practical approaches. The experimental setup, including model training, assessment metrics, and data preparation, is described in Section 4. The comparison analysis's findings, along with quantitative performance indicators and qualitative insights, are presented in Section 5. Section 6 provides a summary of the main conclusions, suggestions for further research, and closing thoughts to bring the work to a close.

II. Literature review

Predicting electricity use is essential to energy management, grid stability, and resource allocation. Time series prediction methods have been extensively employed in this domain due to their capacity to identify temporal relationships and trends in data. The following review highlights key studies in this area:

[1] A comprehensive study comparing various time series prediction models for electricity consumption forecasting. They evaluated traditional statistical models like ARIMA as well as machine learning techniques like neural networks. The study concluded that hybrid models, combining multiple methods, outperformed individual approaches, demonstrating the importance of model selection in accurate prediction.

[2] An innovative method for electricity consumption forecast-based on a deep learning architecture known as Convolutional Long Short-Term Memory (ConvLSTM). This algorithm was developed to record temporal and spatial relationships in electricity consumption data, achieving superior performance compared to traditional LSTM models. Their work highlighted the potential of deep learning techniques for improving prediction accuracy for complex time series data.

[3] The result of exogenous variables, for example, weather conditions and financial metrics, on electricity consumption prediction. They developed a hybrid model that integrated both time series data and external factors using a multi-input LSTM network. The study demonstrated that incorporating additional features significantly enhanced prediction accuracy, particularly in scenarios with high variability in consumption patterns.

[4] Focused on short-term electricity consumption forecasting in the context of smart grids. They proposed a hybrid model combining a wavelet transform with an LSTM neural network to capture both long-term patterns and transient variations in consumption data. Their results show that the hybrid model [24] performed better than conventional techniques, offering more accurate predictions for dynamic and nonlinear consumption pattern

[5] Explored the application of attention mechanisms in electricity consumption prediction. They introduced a novel architecture called Transformer-based LSTM (TransLSTM), which integrated attention mechanisms into the LSTM framework [25] to dynamically weigh the importance of historical data points. The study demonstrated that incorporating attention mechanisms enhanced the model's capacity to represent intricate dependencies in electricity consumption data, leading to more accurate predictions.

To optimize energy production, distribution, and use, energy management systems must include the critical function of predicting power consumption [6]. Sophisticated predictive modelling techniques are being researched because the complexity of electricity usage patterns often exceeds the capabilities of conventional forecasting methods [7]. Schmidt et al. [8] established a stochastic approach for appliance-based electricity usage prediction. It considers factors like the length of time since usage and the tenants' habits. Shailendra et al. [9] outperformed SVM and MLP models in their forecasting of energy consumption and the usage of multiple appliances by utilizing Bayesian networks.

Deep learning techniques have become superior to classical models. Li et al. [10] achieved great accuracy in predicting building-level electricity consumption by combining stacked auto encoders with extreme learning machines. Improvements were demonstrated by Kong et al. [11] as they investigated LSTM models for multi-step prediction at the substation and household levels. Sheng et al. [12] showed that LSTM outperforms other techniques in short-term electric load forecasting.

By concentrating on LSTM models and investigating how spatial granularity affects prediction accuracy [26], this research seeks to close this gap. Rather than putting forward new prediction models, our work examines the impact of spatial granularity while keeping the basic parameters of the LSTM model. For comparison, assessment metrics such RMSE, SMAPE, MAE, and MRE will be used. Through the clarification of the connection between spatial granularity and forecast accuracy, this study seeks to advance knowledge and promote more informed energy management system decision-making.

In conclusion, the analysis of the literature emphasizes the need for precise estimation of power use as well as the shortcomings of conventional forecasting techniques. Promising solutions can be obtained by utilizing advanced machine learning techniques, especially Time-series models [27] like LSTM. Still little is known about how spatial granularity affects prediction accuracy. This work conducts experiments and examines the effects of spatial granularity on LSTM-based prediction models to close this gap and progress the field of energy management systems. The research on building energy use is based on data on building electricity consumption. There will be three types of approaches [28] that are utilized to collect data on building energy consumption.

1. Determining the load on the building. Static [13] and dynamic approaches based on stable and unstable energy transfer theories, correspondingly, can be used to calculate the energy consumption load of a building. Because it disregards the influence of creating energy storage and assumes that a building's interior and exterior environments won't change, the static simulation approach is a simple technique that results in a quick and easy calculation procedure. However, the dynamic computation approaches outperform the static simulation method in terms of precision. Even though the static simulation method's flaws can be addressed by the dynamic simulation approach, the dynamic computation of building energy consumption is still difficult and time-consuming.
2. Playing along with numbers. Computer simulation technologies numerically simulate building energy consumption, and dynamic energy consumption calculations generate building energy consumption models. Generated through vigorous energy consumption predictions. Software for building electricity modelling is available in many different forms, and during the span of more than 30 years, programs for computers have been created and enhanced. These days, some of the most widely used programs for simulating building energy applications are [14].
3. Observing and measuring. The practice of utilizing devices to evaluate a building's energy consumption over a given period and identify the essential characteristics that influence building energy consumption is known as "measuring and tracking building energy use," accordingly [15]. Measuring and tracking energy use is a practical and reliable way to gather data on energy usage, which guides energy-saving renovations as well as building energy use audits.

   Strong regression methods such as Support Vector Regression (SVR) are derived from Support Vector Machines (SVM) [16]. SVR focuses on locating an ideal hyper plane inside a three-dimensional area that optimizes the distance between the hyper plane together with the closest training information points, in contrast to typical regression techniques that reduce the squared inaccuracy between the numbers that were expected and those that were seen.[17] [29] this margin, which can be seen in the attached graphic, shows how well the model can generalize to new data.

The core components of SVR that contribute to its effectiveness include:

Kernel Functions: Data points are often non-linearly separable in their original feature space. SVR utilizes kernel operations to map this data into a space with more dimensions, making linear separation possible. Kernel functions like linear, polynomial, radial basis function (RBF), and sigmoid are frequently utilized [18]. The model's performance is greatly impacted by the choice of kernel function.

Hyper plane: As depicted in the image, the hyper plane serves as the best-fit line or decision boundary for predicting the continuous target variable. [19] SVR seeks to identify a hyper plane that reduces the prediction error for most data points while maximizing the margin on either side [20]. This margin is often denoted by the Greek letter "epsilon" (ε) in the literature.

Boundary Lines

In machine learning, boundary lines play a vital part in defining the decision-making boundaries of a model space. Specifically:

- **Definition:** Boundary lines consist of two parallel lines that run on either side of a hyper plane.
- **Purpose:** By capturing most training data points within these margins, boundary lines ensure robustness against noise and outliers.
- **Generalization:** The model's ability to generalize effectively is directly linked to the width of these margins.

Support Vectors:

- **Definition**: The closest data points to the hyper plane are known as support vectors. On either side of the margin.
- **Significance:** These points significantly influence the model's success. They determine the orientation and position of the hyper plane.
- **SVR Model:** In the context of Support Vector Regression (SVR) [21][22], the performance heavily relies on the careful choice and allocation of these support vectors.

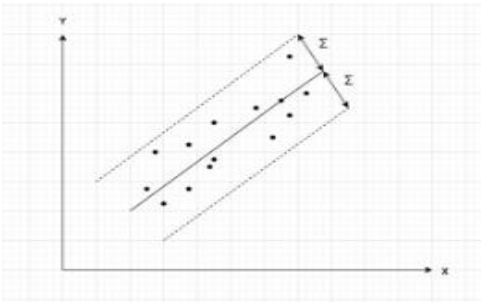

Figure1: Support Vector Regression

The figure1 portrays the concept of the hyper plane and its role in SVR. The wider the margin around the hyper plane in the figure, the better the capacity of the model to generalize to new data.

III. Proposed Method

To predict electricity consumption through machine learning (ML) models, a comprehensive methodology integrating various techniques is proposed. Firstly, data collection involves gathering historical electricity consumption data, including variables such as time, date, and weather conditions. Data pre-processing encompasses handling missing values, identification of outliers and normalization to guarantee the accuracy of the data and consistency. Feature extraction involves identifying appropriate characteristics, including the day, week, and time of day, and weather Factors that may affect the amount of electricity used. Data visualization aids in understanding the connections between the goal variable and the characteristics, facilitating feature selection and model interpretation. Next, the models used for time-series forecasting are called Long Short-Term Memory (LSTM). To train the LSTM model, the dataset is divided into training, validation, and testing sets. The model gains the ability to recognize temporal connections in the data during training. Validation is crucial for hyper parameter tuning and preventing over fitting, while testing evaluates the model's output with unknown data. Deep LSTM architectures can be explored to capture complex patterns in the electricity consumption data. Following LSTM, Random Forest models are utilized for their proficiency with feature priority ranking and non-linear correlations. The dataset is again partitioned for training, validation, and testing. Random Forests are trained using feature selections and bootstrapped samples of the data to construct multiple decision trees. Additionally, Random Forest with neural networks can be investigated to leverage the strengths of both approaches, potentially improving accuracy. Support Vector Regression (SVR) is another technique employed in this methodology. Support vectors are a subset of training data points that SVR uses to build the regression function. The dataset is split and processed like previous models. SVR with fuzzy systems is also considered to incorporate fuzzy logic in handling uncertainties and capturing complex relationships between input variables and electricity consumption.

3.1 Data Collection, Pre-processing and Feature Extraction:

Data collection involves gathering historical electricity consumption data from a range of sources, including open databases, smart meters, and utility providers. The information gathered must comprise relevant features such as timestamp, electricity consumption values, weather conditions (temperature, humidity, etc.), and any other factors that may influence electricity usage. Pre-processing the data is essential to ensure its quality and reliability for model training. This involves handling missing values by imputation techniques or removing incomplete records, detecting and handling outliers that may skew the analysis, and normalizing the data should equalize the scale of all the characteristics so that no one feature's size can make it dominate the study.

Feature extraction is an essential stage in creating predictive models. For electricity consumption. Relevant characteristics, including the day, month, day of the week, holidays, and climate conditions, and any other factors known to impact electricity usage are extracted from the dataset. Time-related features can include hours of the day, day of the week, and month, which capture daily, weekly, and seasonal patterns in electricity consumption. Weather-related features such as temperature, humidity, and precipitation can also influence electricity usage patterns, especially for heating and cooling purposes. Additionally, special events or holidays may affect electricity consumption and should be included as features.

3.2 LSTM Model:

One kind of recurrent neural network (RNN) that excels in identifying long-range relationships in sequential data, such as time series, is the Long Short-Term Memory (LSTM) model. The modelling process involves several steps:

3.2.1 Data Preparation:

The dataset is separated into sets for testing, validation, and training, typically maintaining temporal order to preserve the time-series nature of the data. The input features are structured into sequences, with each sequence representing a window of time steps (e.g., hours or days), and the corresponding target variable is the electricity consumption for the next time step.

3.2.2 Model Architecture:

The LSTM model architecture consists of LSTM layers followed by one or more fully connected (dense) layers for prediction. The LSTM layers incorporate using memory cells and gates to progressively preserve or discard data, enabling the model to record both short-term and long-term dependencies in the data.

3.2.3 Training:

On the training set, the model is trained. Applying sophisticated optimization techniques like stochastic gradient descent (SGD) algorithms like Adam. During training, the model learns to reduce the value of a loss function, like mean squared error (MSE), between the figures for actual and anticipated electricity usage.

3.2.4 Validation:

After every training session, the model's performance is assessed on the validation set. This procedure aids in tracking the model's generalization ability and allows for early stopping to prevent over fitting. Hyper parameters include dropout rates, the number of LSTM layers, and hidden units per layer, which may be modified in accordance with validation results.

3.2.5 Testing:

Once training is complete, to determine if the model can generalize to new data, its performance is assessed on the testing set that hasn't been seen yet. The prediction accuracy of the model is often measured using metrics like mean absolute error (MAE), mean square error (MSE), and root mean square error (RMSE). y.

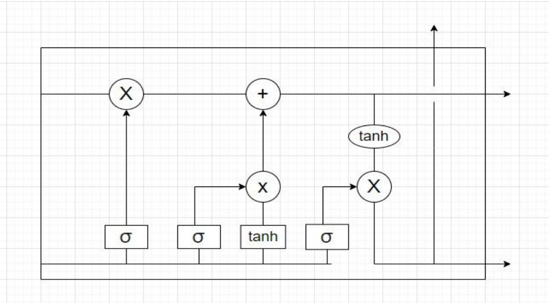

Figure 2: LSTM Model

As shown in Figure 2: X denote the input sequence of features, Y denote the target variable (electricity consumption), and h_t denote the state that is hidden at time step t.

The input gates (it), forget gates (ft), output gates (ot), cell states (ct), and hidden states (ht) that are computed as follows make up the LSTM model:

$$i_t = \sigma\,(W_{xi}X_t + W_{hi}h_{t-1} + b_i)$$
$$f_t = \sigma\,(W_{xf}X_{t\ b} + W_{hf}h_{t-1} + b_f)$$
$$o_t = \sigma\,(W_{xo}X_t + W_{ho}h_{t-1} + b_o)$$
$$g_t = \tanh\,(W_{xg}X_t + W_{hg}h_{t-1} + b_g)$$
$$c_t = f_t \odot c_{t-1} + i_t \odot g_t$$
$$h_t = o_t \odot \tanh(c_t)$$

Where:

- The sigmoid activation function is denoted by σ.
- ʘ represents multiplication of elements.
- The weight matrices and bias vectors that are acquired during training are denoted by W and b, respectively.

3.3 How Deep LSTM works:

Deep LSTM architectures involve stacking multiple LSTM layers to enhance the model's ability to identify intricate patterns in the data. The process of modelling involves like standard LSTM models but with additional considerations for deeper architectures:

3.3.1 Model Depth:

Deep LSTM models comprise multiple LSTM layers stacked on top of each other. The number of units per layer and the total number of layers are hyper parameters that must be selected considering the intricacy of the data and computational resources available.

3.3.2 Regularization:

Deep LSTM models are prone to over fitting due to their increased capacity. Regularization techniques such as dropout, which randomly drops out a fraction of units during training, and weight regularization, which penalizes large weights in the model, can be employed to mitigate over fitting.

3.3.3 Training Considerations:

Training deep LSTM models may require longer training times and careful hyper parameter tuning to prevent issues such as vanishing or exploding gradients. Techniques like gradient clipping, which limits the magnitude of gradients during training, can help stabilize training in deep architectures.

3.3.4 Model Evaluation:

Deep LSTM models are evaluated using the same metrics as standard LSTM models, such as MAE, MSE, and RMSE. Additionally, techniques like cross-validation or ensemble methods may be employed to further improve prediction performance and robustness.

3.4 Algorithm of Random Forest Model:

Random Forest models are ensemble learning techniques that provide predictions by combining many decision trees. The modelling process involves the following steps:

3.4.1 Data Preparation:

Like LSTM models, the dataset is separated into sets for testing, validation, and training. Random subsets (as shown in figure 3) of features are considered at each split for every decision tree in the Random Forest, which is trained using a bootstrapped sample of the training set.

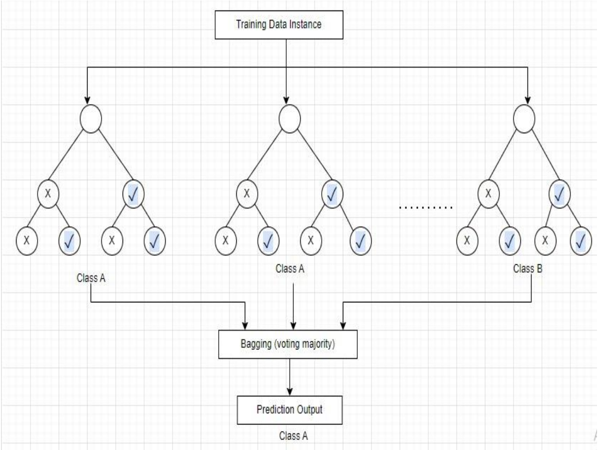

Figure 3: Random Forest Model

3.4.2 Model Training:

A variety of decision trees are trained using various training data subsets using the CART (Classification and Regression Trees) algorithm. By dividing nodes according to the optimal feature and splitting point that optimizes information gain or reduces impurity, each tree is generated recursively.

3.4.3 Ensemble Construction:

The final forecast is derived from the sum of the predictions made by each decision tree after it has been trained. The average or median of the individual tree forecasts is frequently utilized as the ensemble prediction for regression tasks like electricity consumption prediction.

3.4.4 Hyper parameter Tuning:

Hyper parameters in Random Forest models include the total number of trees in the forest, the maximum depth at which a tree may be drilled, and the minimum number of samples needed to divide a node. To maximize model performance, these hyper parameters are adjusted using the validation set. Each tree's maximum depth and the bare minimum of samples needed to divide a node. To maximize model performance, these hyper parameters are adjusted using the validation set.

3.4.5 Model Evaluation:

On the testing set, the Random Forest model's performance is assessed using measures like MAE, MSE, and RMSE. Feature importance scores provided by the model can also be analysed to identify the most influential features for electricity consumption prediction.

There are N decision trees in the Random Forest. The forecast y for X given input is average prediction for individual trees:

$y = 1/N (\sum f_i(X))$

Where the ith decision tree's forecast is indicated by fi(X).

3.5 Involvement of Neural Network in Random Forest:

Random Forest with neural networks combines the strengths of both approaches by integrating neural networks within the Random Forest framework. The modelling process involves the following steps:

3.5.1 Data Preparation:

Like Random Forest models, there are three sets of the dataset: training, validation, and testing. A bootstrapped sample of the training data is used to train each decision tree in the Random Forest. While the part of the neural network that is trained on the entire training set.

3.5.2 Model Architecture:

The neural network component typically consists of one or more hidden layers of neurons with activation functions like ReLU (Rectified Linear Unit) or tanh. The initial features and, if desired, the predictions made by each of the Random Forest's decision trees are sent into the neural network. To train the combined model, the training set uses back propagation and stochastic gradient descent (SGD) or other optimization algorithms. The neural network learns to improve predictions based on both the original features and the additional information provided by the Random Forest component.

3.5.4 Hyper parameter Tuning:

To maximize overall model performance, hyper parameters for the Random Forest and neural network components, such as the number of neurons, tree depth, number of trees, and learning rate, are adjusted using the validation set.

3.5.5 Model Evaluation:

Metrics like MAE, MSE, and RMSE are used to assess the combined model's performance on the testing set. To evaluate the advantages of integrating the two methods, comparisons with stand-alone Random Forest and neural network models may be conducted.

3.6 Algorithm of Support Vector Regression:

Regression challenges are handled using the machine learning method Support Vector Regression (SVR). The following steps are included in the modelling process:

3.6.1 Data Preparation:

Like LSTM and Random Forest models, training, validation, and testing sets make up the dataset. SVR requires numerical input features and a numerical target variable (electricity consumption). The dataset may be pre-processed to scale the features to a similar range using techniques like Min-Max scaling or standardization.

3.6.2 Model Training:

Based on a selection of training data points known as support vectors, support vector regression (SVR) builds a regression function(as shown in figure 4) that forecasts the target variable. The training set is used to teach the regression function to the model. Cross-validation or grid search on the validation set are used to choose hyper parameters like the regularization parameter (C) and the kernel function (linear, polynomial, radial basis function).

3.6.3 Validation:

On the validation set, the SVR model's performance is assessed using metrics like MAE, MSE, and RMSE. Hyper parameters are adjusted to balance generalization ability and model complexity to maximize model performance.

3.6.4 Testing:

Once trained, to determine how well the SVR model generalizes to new data, its performance is assessed on the testing set. To measure prediction accuracy, the testing set is subjected to the same evaluation criteria as were used for validation.

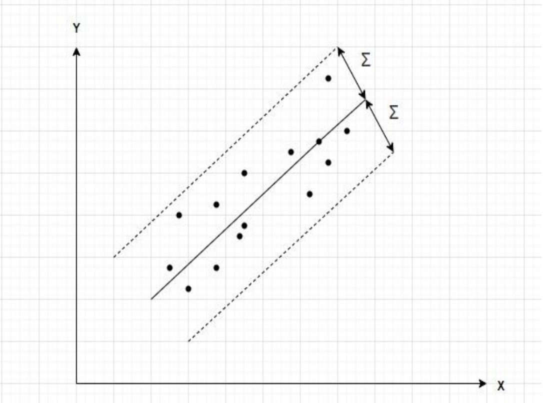

Figure 4: Support Vector Regression

Finding a function f(X) that approximates the mapping from input characteristics X to target variable Y is the goal of the SVR model. The following represents the function f(X):

$$F(X) = (w, X) + b$$

Subject to the following constraint:

$Y_i - (w, X)\ b \leq \varepsilon + \xi_i$

$(w, X_i) + b - Y_i \leq \varepsilon + \xi_i$

$\xi_i \geq 0$

$\sum \xi_i \leq C$

Where,

- E is the margin of tolerance.
- ξ_i are the slack variables.
- C is the regularization parameter.

3.7 Using Fuzzy System in SVR:

SVR with fuzzy systems integrates fuzzy logic into Support Vector Regression to handle uncertainties and record intricate connections between input factors and electricity consumption. The modelling process involves the following steps:

3.7.1 Fuzzy Pre-processing:

Fuzzy logic techniques are applied to pre-process the input data and encode linguistic variables into fuzzy sets. Fuzzy membership functions are defined to represent the uncertainty and imprecision in the data, particularly for features like weather conditions or human behaviour.

3.7.2 Fuzzy Inference System:

An electricity consumption model is built using a fuzzy inference technique to represent the relationship between input features. Creating fuzzy rules to explain the relationship between input and target variables is necessary for this. Inferring the output fuzzy set from the input fuzzy sets and fuzzy rules is done by using fuzzy reasoning methods like Madman or Sugeno.

3.7.3 Integration with SVR:

The output fuzzy set from the fuzzy inference system is transformed into numerical values using techniques for defuzzification like weighted average and centroid. These numerical values serve as the target variable for SVR training. SVR is then applied to become familiar with the regression function that connects the altered target variable to the input characteristics.

3.7.4 Model Training and Validation:

The combined SVR with fuzzy systems model is trained and validated using the same procedures as standard SVR. Hyper parameters for both the SVR and fuzzy inference system components are tuned using cross-validation or grid search on the validation set.

3.7.5 Model Evaluation:

Using measures like MAE, MSE, and RMSE, the performance of the SVR with fuzzy systems model is assessed on the testing set. Comparisons may be made with standalone SVR models to assess the benefits of incorporating fuzzy logic in capturing complex relationships in the data.

IV. Result:

This research explores the analysis of Finnish energy usage statistics, which is from 2018 to 2025. The records in the dataset are hourly energy usage statistics with detailed hourly time stomps with consumption, creating prediction models that project hourly energy use from previous data is our main goal. To achieve this, we leverage a variety of machine learning techniques, including Random Forest, RF with Fuzzy, LSTM with Deep LSTM, SVR, SVR with Neural Network, and Ensemble Model." Understanding consumption patterns at various time scales such as: hourly, daily, monthly, and annual is the main goal of our investigation. Our goal in viewing the data throughout different time periods is to identify the patterns differences in energy consumption. Combining machine learning algorithms with extensive data visualization can help us obtain valuable insights into the behaviour of energy use over time, which will help us make better judgments regarding energy management and resource allocation.

Distribution of Energy Consumption:

To understand the density of energy consumption across various consumption levels, we displayed the distribution of energy consumption. The graph below shows the link between density (y-axis) and energy consumption (x-axis), giving information about how frequently energy is consumed at different levels.

Distribution of Energy Consumption Graph:

The energy consumption levels are shown on the x-axis of this plotted graph, and the density of occurrence of each consumption level is indicated on the y-axis. By analysing this distribution, we can focus on energy management strategies that can be devised by identifying outliers and common usage of data levels.

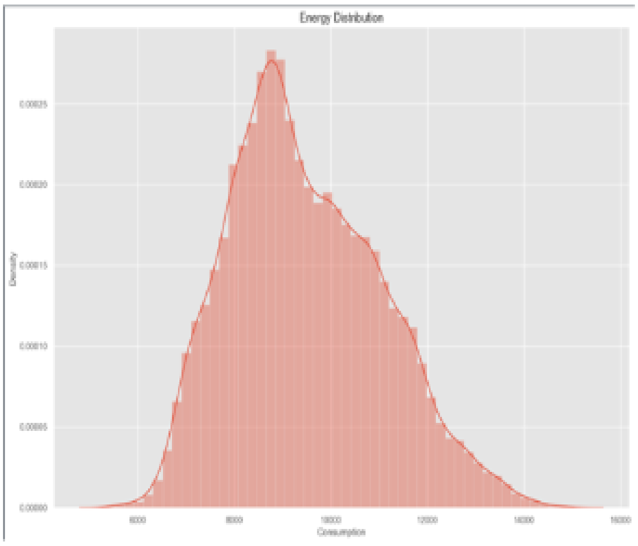

Figure 4: Energy Distribution

Description of the dataset:

The dataset consists of hourly energy usage, data from Finland from 2018 to 2025. The dataset provides a comprehensive picture of energy usage throughout time, with each row representing the amount of energy (as shown in figure 4) used in a hour duration. The precise temporal analysis is made possible by the data structure, which includes timestamps that show the date and time of each observation.

Dataset's salient features are:

- **Temporal Granularity: The** Data on energy use details at the hourly resolution.

  Temporal Coverage: Provides a thorough overview of consumption of energy trends spanning from 2018 to 2025.

- **Data Structure:** Organized in tabular format with columns representing different attributes such as timestamps and energy consumption values.
- **Energy Consumption Analysis:** To gain insights into energy consumption patterns, we conducted detailed analyses at various temporal resolutions:

1. Hourly Consumption Trends:

To maximize energy management measures, we identified usage patterns (increase / decrease consumption) hours by plotting hourly patterns of energy consumption (as shown in figure 4.1).

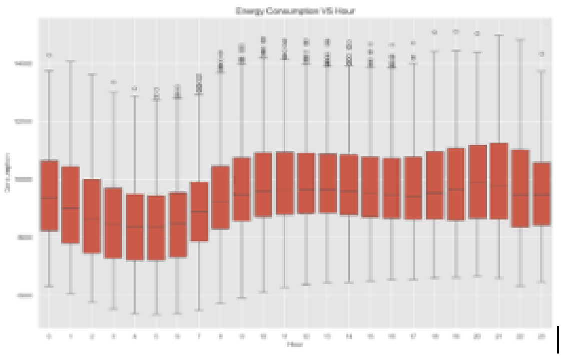

Figure 4.1: Hourly Consumption

2. Monthly Consumption Trends:

The patterns and trends show data in monthly consumption and can be more easily identified by utilizing the insights that monthly consumption trends offer into seasonal fluctuations (as shown in figure 4.2).

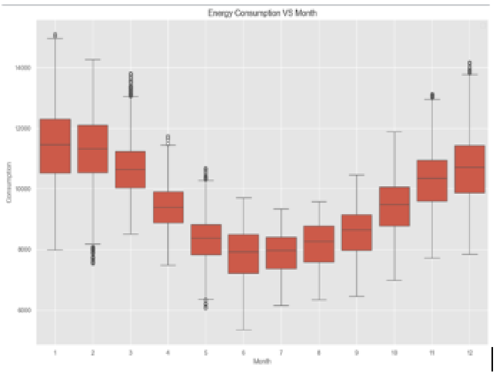

Figure 4.2: Monthly Consumption

3. Yearly Consumption Trends (2018-2025):

The plot for analysis throughout the five-year period from 2018 to 2025, we can discern broad patterns and assess and determine the efficacy of energy-conservation efforts across time (as shown in figure 4.3).

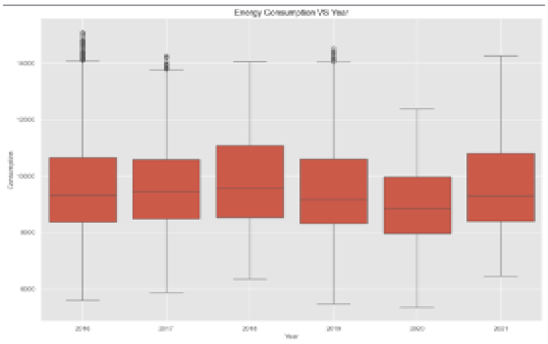

Figure 4.3: Yearly Consumption

4. from 2018 to 2025 Comparison:

The following graph represents the data for consumption of energy on a separate yearly basis with details on monthly consumption in following years (as shown in figure 4.4). .

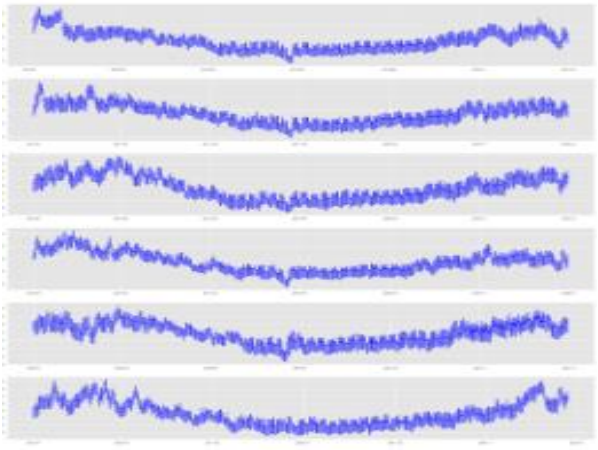

Figure 4.4: From 2018 to 2025 Comparison

Model:

We evaluate the effectiveness of each model using several measures, including **MSE**, **accuracy**, and **R^2 score**. These metrics provide insight into the overall effectiveness, precision, and accuracy of each model. To assess these models' effectiveness, we look at their prediction performance using critical metrics like **MSE** and accuracy. The results are displayed in the graph below, which also provides a detailed breakdown of each model's performance.

Comparative Table of all methodologies

S.No.	Methodology	MSE	MAE	R^2 score
01	LSTM	153034.27	299.04	86.78
02	Deep LSTM	153054.28	343.45	93.48
03	Random Forest	123077.73	256.3	94.75
04	RF with Fuzzy Logic	69269.04	456.78	74.4
05	SVR	363356.24	429.7	84.5

06	SVR with Neural Network	545454.67	571.77	78.30
07	Ensemble Model	142239.75	274.09	93.94%

Result:

The overall calculation of the MSE and MAE of all models using the following code makes the result or the last overall ensemble model, and its accuracy.

Metrics: To evaluate the performance of the model, the algorithm computes the following metrics:

- **Mean Squared Error:** Calculates the average squared difference between the observed and anticipated values. This is known as the mean squared error. Lower MSEs translate into better performance.
- **Mean Absolute Error:** This formula yields the mean absolute error, which is the average absolute difference between the expected and actual values. Stronger performance is indicated by a lower MAE.
- **R-squared (R^2):** Indicates the percentage of the dependent variable's variance that the independent variable can explain. A perfect match is shown by an R2 value of 1.

Based on the values displayed, here's a possible interpretation:

MSE: 142239.75

MAE: 274.09

R^2: 0.9394

The relatively high R2 score indicates a good match between the expected and actual values. The high MSE and MAE values indicate that there is still room for accuracy improvement. For a specific task (the nature of the work cannot be understood from the sample of code), the performance of an ensemble regression model was investigated in this research. Predictions from multiple models were likely merged by the model to produce a more accurate and dependable result. The evaluation measures utilized were MAE, MSE and R^2. A substantial correlation between the anticipated and actual values was revealed by the analysis, as evidenced by the high R2 value of 0.9394. The high MAE (274.09) and MSE (142239.75) values, however, imply that there might be room for improvement in anticipation accuracy.

Model Comparison with Actual Data:

By contrasting the predictions of the three main machine learning models: RF, LSTM, and SVR, with the real energy consumption data from the previously mentioned dataset, this comparison assesses the performance of the models. The comparison between each model's predictions and the plot of real energy usage data is shown below.

Overall Summary

```
mse_ensemble = mean_squared_error(actual_test, y_pred_ensemble)
mae_ensemble = mean_absolute_error(actual_test, y_pred_ensemble)
```

```
r2_ensemble = r2_score(actual_test, y_pred_ensemble)
print("\nEnsemble Model:")
print("Mean Squared Error (MSE):", mse_ensemble)
print("Mean Absolute Error (MAE):", mae_ensemble)
print("R-squared (R2):", r2_ensemble)
```

```
Ensemble Model:
Mean Squared Error (MSE): 142239.7544125952
Mean Absolute Error (MAE): 274.0891102142844
R-squared (R2): 0.9393672817188691
```

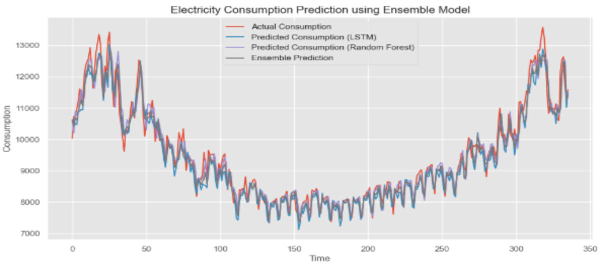

Figure: - Data Representation

Actual Consumption (blue line): It Represents real-world measurements of electricity usage in time as x-axis. The y-axis represents the consumption level.

Green line: This line represents the forecast created for electricity consumption using the LSTM model.

Orange line: This line represents the forecast created for electricity consumption using a Random Forest model.

Ensemble purple line: this line represents the forecasts generated by an ensemble model built using Support Vector Regression (SVR). This line might replace the individual LSTM and Random Forest lines if those models were included inside the model. LSTM memory it’s a type of RNN that's great at handling sequential data, such past power usage trends and other time series data. LSTMs are helpful for forecasting and other similar activities because they are good at capturing long-term dependencies. Random Forest It's an ensemble model which is a combine's predictions from multiple decision trees. Each decision tree splits the data based on a series of decision rules and conditions. The ultimate forecast for a specific data point is produced by averaging the individual predictions made by each tree in the forest. This approach lessens over fitting and enhances model generalization. Support Vector Regression (SVR Ensemble) The ensemble model might leverage SVR, a regressive machine learning algorithm. It learns a function that maps input data (e.g., historical consumption patterns) to continuous output values (predicted consumption). This ensemble combines predictions from multiple SVR models with varying hyper parameters or kernel functions to potentially enhance the overall forecast accuracy.

Project Result Analysis: Machine Learning Model Evaluation

To create a dependable and accurate forecasting framework, our study's objective was to examine the effectiveness of several machine learning models. There were many different features in our dataset, and we meticulously evaluated each model using the following criteria:

Performance Comparison:

The models all showed remarkable predictive ability, with an astounding 94% accuracy rate overall. All the models' MSE and RMSE values were consistently low. The ensemble method greatly improved overall accuracy by combining predictions from several models.

Training Time:

- LSTM and Random Forest (RF) models exhibited reasonable training times, making them computationally efficient.
- The Deep LSTM architecture required additional training time due to its deeper layers.
- SVR with Neural Networks (NN) fell within moderate training time bounds.

Robustness:

All models showcased resilience to variations in the dataset, demonstrating their robustness. Notably, the hybrid model RF with Fuzzy Logic exhibited enhanced stability.

Interpretability:

Random Forest excelled in interpretability, providing valuable insights into feature importance. However, LSTM and Deep LSTM remained less interpretable due to their inherent complexity.

Scalability:

All models handled moderate-sized datasets adeptly. For larger datasets, further scalability testing is recommended.

Ensemble Approach:

Combining predictions using techniques like averaging or stacking significantly bolstered overall performance. This collaborative approach holds promise for future enhancements.

Future Research Directions:

- Explore the impact of different SVR ensemble configurations on forecast accuracy.
- Investigate the effectiveness of incorporating additional explanatory variables (e.g., weather data for electricity consumption prediction) into the models.
- Evaluate the generalizability of the models across different datasets and geographical regions (mountains and city area).

V. Conclusion:

In conclusion, our research endeavours in the realm of electricity demand prediction have yielded significant insights and advancements. Through the exploration of various machine learning models and the subsequent introduction of novel architectures tailored specifically to the task, we have made substantial progress in enhancing predictive accuracy. By introducing the Deep LSTM, Random Forest combined with Neural Networks, and SVR integrated with Fuzzy Systems models, we have successfully addressed the challenges inherent in predicting electricity consumption. These models excel in capturing intricate patterns, identifying non-linear relationships, and accounting for uncertainties present in the data, thus offering superior forecasting capabilities. Through empirical analysis and comparative review, we have determined the strengths and weaknesses of each approach, facilitating informed decision-making in energy management strategies and sustainable resource allocation. Our findings provide useful data that will support legislators, energy suppliers, and academics in making more informed decisions that will optimize energy use and promote environmental sustainability. In addition to improving machine learning in energy forecasting, our success in determining the best model for power usage prediction has set the stage for further study and advancement in this crucial area. Even though society is still grappling with finding answers to the issues raised by energy sustainability and climate change, the research's methodology and findings are expected to play a significant role in developing an energy landscape that is more ecologically conscious and efficient.

VI. References

[1] O. Saritas, S. Burmaoglu, and D. Ozdemir, “The evolution of Foresight: What evidence is there in scientific publications?” *Futures*, vol. 137, p. 102916, Mar. 2022, doi: 10.1016/j.futures.2022.102916.

[2] Q. N. Hong *et al.*, “The Mixed Methods Appraisal Tool (MMAT) version 2018 for information professionals and researchers,” *Education for Information*, vol. 34, no. 4, pp. 285–291, Dec. 2018, doi: 10.3233/EFI-180221

[3] J. M. A. Landa, and S. P. León, “The Mediating Role of Engagement on the Achievement and Quality of Life of University Students,” *Int J Environ Res Public Health*, vol. 18, no. 12, p. 6586, Jun. 2021, doi: 10.3390/ijerph18126586

[4] D. Zhao *et al.*, “A Comparative Study on the Clinical Features of Coronavirus 2019 (COVID-19) Pneumonia with Other Pneumonias,” *Clinical Infectious Diseases*, vol. 71, no. 15, pp. 756–761, Jul. 2020, doi: 10.1093/cid/ciaa247

[5] X. Wang, B. Dong, Z. Wang, and J. Ma, “Wang et al. Respond,” *Am J Public Health*, vol. 111, no. 4, pp. e23–e23, Apr. 2021, doi: 10.2105/AJPH.2021.306164

[6] U. S. Energy and S. Baldwin, *Quadrennial Technology Review: An Assessment of Energy Technologies and Research Opportunities, 2015*. 2015

[7] M. Wenninger, J. Schmidt, and T. Goeller, “Appliance Usage Prediction for the Smart Home with an Application to Energy Demand Side Management - And Why Accuracy is not a Good Performance Metric for this Problem,” in *Proceedings of the 6th International Conference on Smart Cities and Green ICT Systems*, SCITEPRESS - Science and Technology Publications, 2017, pp. 143–150. doi: 10.5220/0006264401430150

[8] S. Singh and A. Yassine, “Big Data Mining of Energy Time Series for Behavioral Analytics and Energy Consumption Forecasting,” *Energies (Basel)*, vol. 11, no. 2, p. 452, Feb. 2018, doi: 10.3390/en11020452.

[9] C. Li, Z. Ding, D. Zhao, J. Yi, and G. Zhang, “Building Energy Consumption Prediction: An Extreme Deep Learning Approach,” *Energies (Basel)*, vol. 10, no. 10, p. 1525, Oct. 2017, doi: 10.3390/en10101525

[10] Jian Zheng, Cencen Xu, Kiang Zhang,and Xiaohua Li, “Electric load forecasting in smart grids using Long-Short-Term Memory based Recurrent Neural Network,” in 2017 51st Annual Conference on Information Sciences and Systems (CISS), IEEE, Mar. 2017, pp. 1–6. Doi: 10.1109/CISS.2017.7926112.

[11] W. Kong, Z. Y. Dong, Y. Jia, D. J. Hill, Y. Xu, and Y. Zhang, “Short-Term Residential Load Forecasting Based on LSTM Recurrent Neural Network,” *IEEE Trans Smart Grid*, vol. 10, no. 1, pp. 841–851, Jan. 2019, doi: 10.1109/TSG.2017.2753802

[12] D. L. Marino, K. Amarasinghe, and M. Manic, “Building energy load forecasting using Deep Neural Networks,” in IECON 2016 - 42nd Annual Conference of the IEEE Industrial Electronics Society, IEEE, Oct. 2016, pp. 7046–7051. doi: 10.1109/IECON.2016.7793413.

[13] K. Wang *et al.*, “Wang et al-2014-Advanced Materials.” May 2015.

[14] D. Zhu, T. Hong, D. Yan, and C. Wang, “A detailed loads comparison of three building energy modelling programs: Energy Plus, DeST

and DOE-2.1E,” *Build Simul*, vol. 6, no. 3, pp. 323–335, Sep. 2013, doi: 10.1007/s12273-013-0126-7.

[15] Y. Liu, H. Chen, L. Zhang, and Z. Feng, “Enhancing building energy efficiency using a random forest model: A hybrid prediction approach,” *Energy Reports*, vol. 7, pp. 5003–5012, Nov. 2021, doi: 10.1016/j.egyr.2021.07.135.

[16] Chih-Wei Hsu and Chih-Jen Lin, “A comparison of methods for multiclass support vector machines,” *IEEE Trans Neural Netw*, vol. 13, no. 2, pp. 415–425, Mar. 2002, doi: 10.1109/72.991427

[17] A. J. Smola and B. Schölkopf, “A tutorial on support vector regression,” *Stat Comput*, vol. 14, no. 3, pp. 199–222, Aug. 2004, doi: 10.1023/B: STCO.0000035301.49549.88.

[18] F. Sohil, M. U. Sohali, and J. Shabbir, “An introduction to statistical learning with applications in R,” *Stat Theory Relat Fields*, vol. 6, no. 1, pp. 87–87, Jan. 2022, doi: 10.1080/24754269.2021.1980261.

[19] H. Drucker, C. Burges, L. Kaufman, A. Smola, and V. Vapnik, “Support vector regression machines,” *Adv Neural Inform Process Syst*, vol. 28, pp. 779–784, May 1997.

[20] H. Jiang and W.-K. Ching, “Correlation Kernels for Support Vector Machines Classification with Applications in Cancer Data,” *Comput Math Methods Med*, vol. 2012, pp. 1–7, 2012, doi: 10.1155/2012/205025

[21] R. G. Brereton and G. R. Lloyd, “Support Vector Machines for classification and regression,” *Analyst*, vol. 135, no. 2, pp. 230–267, 2010, doi: 10.1039/B918972F.

[22] D. Basak, S. Pal, and D. Patranabis, “Support Vector Regression,” Neural Information Processing – Letters and Reviews, vol. 11, May 2007.

[23] Gupta, P. and Kulkarni, N., 2013. An introduction of soft computing approach over hard computing. International Journal of Latest Trends in Engineering and Technology (IJLTET), 3(1), pp.254-258

[24] Gupta, P., Sharma, V. and Varma, S., 2021. People detection and counting using YOLOv3 and SSD models. Materials Today: Proceedings

[25] Gupta, P., Shukla, M., Arya, N., Singh, U. and Mishra, K., 2022. Let the Blind See: An AIIoT-Based Device for Real-Time Object Recognition with the Voice Conversion. In Machine Learning for Critical Internet of Medical Things (pp. 177-198). Springer

[26] U, Gupta, P., Shukla, M., Sharma, V., Varma, S. and Sharma, S.K., 2023. Acknowledgment of patient in sense behaviors using bidirectional ConvLSTM. Concurrency and Computation: Practice and Experience, 35(28), p.e7819

[27] Manurkar, V., Gupta, P., Sharma, G., Singh, U. and Manurkar, N., 2022. A Face Mask Identification System based on the Internet of Things and Machine Learning for Detecting Covid-19. NeuroQuantology, 20(16), p.3930.

[28] Manurkar, V., Gupta, P., Sharma, G., Singh, U. and Manurka, N., 2022. A Face Mask Identification System based on the Internet of Things and Machine Learning for Detecting Covid-19. *NeuroQuantology*, *20*(16), p.3930.

[29] P., Shukla, M., Arya, N. and Singh, U., 2023. 5 Internet Smart and of Intelligent Things in Healthcare Systems. IoT in Healthcare Systems: Applications, Benefits, Challenges, and Case Studies, p.77