Volume 19, Issue 2, 192541 (1-13)

Majlesi Journal of Electrical Engineering (MJEE)

https://doi.org/10.57647/j.mjee.2025.1902.41

Improving the criteria of electricity consumption forecasting in petrochemical industrial units based on deep learning

Ehsan Tavakoli Garmaserh[1], Mehran Emadi[2,*]

[1]*Department of Electrical Engineering, Majlesi Branch, Islamic Azad University, Majlesi, Isfahan, Iran.*
[2]*Department of Electrical Engineering, Mo.C., Islamic Azad University, Isfahan, Iran.*

[*]**Corresponding author:** mehranemadi49@iau.ac.ir

Original Research

Received:
1 March 2025
Revised:
2 April 2025
Accepted:
16 April 2025
Published online:
1 June 2025

Abstract:

Accurate forecasting of electricity consumption in petrochemical industrial units is essential for optimizing energy management and ensuring operational efficiency. This study presents a novel deep learning framework that integrates advanced feature engineering and Long Short-Term Memory (LSTM) networks to address the challenges posed by irregular seasonal trends and dynamic consumption patterns. Key innovations include the use of Fourier Transform-based feature extraction for enhanced data representation and a hybrid genetic-sparse matrix optimization technique for feature selection, ensuring high predictive performance. The proposed method effectively mitigates issues related to data irregularities through preprocessing techniques, resulting in improved accuracy and stability in both univariate and multivariate time series forecasting scenarios. Experimental evaluations using benchmark datasets demonstrate significant improvements, achieving a Root Mean Square Error (RMSE) of 0.0693 and a Mean Absolute Percentage Error (MAPE) reduction of over 15% compared to state-of-the-art methods. These results highlight the robustness and practical applicability of the proposed framework for industrial energy consumption forecasting and sustainable energy management.

Keywords: Electricity consumption forecasting; Petrochemical industrial units; Deep learning; Long short-term memory

Highlights

- Developed a deep learning-based framework to forecast electricity consumption in petrochemical industrial units.
- Implemented Long Short-Term Memory (LSTM) networks to enhance in multivariate time series predictions.
- Applied feature extraction using Fourier Transform to improve the input data quality.
- Introduced a novel feature selection method combining genetic algorithms and sparse matrix techniques.
- Achieved significant improvements in energy consumption forecasting through preprocessing and feature engineering.

1. Introduction

In recent years, the topic of social Internet of Things (IoT) has become an emerging IOT area of interest. The Internet of Things is now an established paradigm that supports a variety of applications and services [1]. Sectors that can benefit from this new paradigm and the connection of various smart objects to achieve comprehensive and effective answers to their problems include the sectors of transportation, water, energy, etc. [2]. The potential of energy efficiency management and consumption forecasting has been gradually recognized by governments and energy research institutions as an important part of sustainable development. With the increase in population and living standards of citizens, household energy consumption is steadily increasing [3]. Electricity production and distribution systems are always trying to maintain the balance between electricity supply and demand. According to the Annual Energy Outlook 2020 report, the annual growth of electricity demand is on average about 1% in the period 2019-2050 [4]. Industrial companies such as iron smelters, steel producing companies, automobile manufacturing and other cases including industrial uses in petrochemical industrial units are where electricity consumption forecasting is very important [5]. The rapid advancements in artificial intelligence and machine learning have revolutionized forecasting methods, particularly in energy-intensive industries. Traditional approaches such as Auto-Regressive Integrated Moving Average (ARIMA) or basic neural networks often fall short in addressing the complexities of modern industrial energy consumption, characterized by irregular seasonal trends, nonlinear relationships, and the impact of external variables like weather conditions and production cycles. Recent studies have shown that deep learning models, especially Long Short-Term Memory (LSTM) networks, offer superior capabilities for capturing long-term dependencies and dynamic patterns in multivariate time series data. However, the challenge lies in optimizing these models for industrial applications, where data preprocessing, feature extraction, and model tuning play a critical role in achieving accurate and reliable predictions. This research builds on these advancements to propose a comprehensive framework tailored for the unique demands of petrochemical industrial units. In the modern industrial age, oil and petrochemical industries have become a very important energy source and a strategic economic source [6]. The lack of oil and the sustainable development of the oil industry directly affect the development process of the national economy and defense security [7]. On the one hand, the world, especially Iran, is on the path of rapid industrialization and urbanization, and this fact undoubtedly accelerates the consumption of oil in secondary and tertiary industries such as petrochemicals [8]. In particular, it is observed that the ratio of oil consumption of secondary and tertiary industries basically dominates the total volume of oil consumption. Therefore, the development process of oil consumption will increase rapidly [9]. This increase in the consumption of oil and derived materials, like petrochemical production, requires the use of electrical energy for production [10–12]. In other words, electric energy consumption will increase in parallel. Increasing the consumption of electrical energy in industrial and petrochemical uses cannot lead to good results [13]. Those in charge should consider the necessary arrangements regarding the consumption of electrical energy. In this regard, the prediction of electric energy consumption can help these decisions [14]. One effective approach to intelligent forecasting is leveraging methods based on machine learning and deep learning [15]. Recurrent neural networks (RNN) are one of the deep learning methods. which has advantages such as accurate prediction, time series, high convergence speed and high adaptability [12]. In this network, the outputs of the hidden layers have feedback to themselves [16]. In other words, each neuron in the output layer has a feedback, and this feedback connection is made through a buffer layer. This feedback in the output layer makes the RNN learn better, recognize better and produce instant patterns better. Each hidden neuron is connected to only one recurrent neuron with a constant number of one [17]. But recurrent neural networks are unable to store information related to past inputs for a long time. In addition to the fact that this specification weakens the ability of this network to model long-term structures, this "forgetting" causes this type of network to be exposed to instability during sequence generation. Having a longer-term memory has a stabilizing effect because even if the network cannot get a correct understanding of its recent history, it is still able to complete its prediction by looking back [18]. Unlike the traditional recurrent neural network in which the content is rewritten at each time step, in an LSTM recurrent neural network, the network is able to decide to keep the current memory through introduced gates. Intuitively, if the LSTM unit recognizes an important feature in the input sequence in the initial steps, it can easily transmit this information along a long path, so it receives and maintains such possible long-term dependencies [19]. Therefore, an LSTM neural network can be used in the buffer units in a recurrent network. The combination of these two structures will improve memory in the LSTM-based recurrent neural network in learning selected thin features [20]. The performance of machine learning (ML) models mainly depends on the data representation [21]. While deep learning (DL)deals with non-linear transformation that provides high-level abstraction and ultimately greater benefit [22]. DL techniques have been widely used in various applications. RNN excels in natural language processing (NLP)tasks by storing sequential information. RNN also ensures that time series information can be preserved. LSTM, a type of RNN, is used to extract spatial and temporal features in combination with CNN [23]. The problem of predicting electric energy consumption is a time series problem [24]. Forecasting electric energy consumption in industrial uses in petrochemical industrial units is a multivariate time series problem that predicts electricity consumption. However, these irregular seasonal trends of electricity consumption make forecasting methods difficult to predict electric energy consumption. The dataset provided by the UCI repository consists of seven variables and energy consumption sampling for the years $2007-2011$ and is considered a benchmark dataset in time series forecasting [25]. These time series were collected from various social IoT devices such as smart meter readings. The output of the LSTM layer is fed to the fully connected layer which ultimately predicts the power demand [26]. In this work, the problem of multi-stage series electricity consumption is investigated, which is to estimate the expected electricity consumption for the next week using the recent consumption [27]. Then forecast the total active power each day to the next week using the forecasting model. The primary purpose of this article is to enhance the forecasting criteria of electricity consumption forecasting in petrochemical industrial units by leveraging advanced deep learning methodologies. The study introduces a novel framework that integrates LSTM networks with feature extraction techniques based on Fourier Transform and an optimized feature selection method using genetic algorithms and sparse matrix optimization. This combination addresses challenges such as irregular seasonal trends, data irregularities, and high dynamics in energy consumption patterns. The proposed approach not only improves predictive forecasting criteria but also provides a robust and efficient tool for industrial energy management, distinguishing itself through its innovative integration of preprocessing, feature engineering, and deep learning techniques. The main goal of this research is to predict the consumption of electrical energy in industrial uses in petrochemical industrial units based on deep recurrent neural network with the help of LSTM long-short-term memory networks [28]. In the proposed method, there are stages of pre-processing, feature extraction, dimensionality reduction and finally classification. In the proposed method, the consumption time series data is first pre-processed. Quality improvement and noise removal are also done at this stage. Then suitable features are extracted for pre-processing. The features extracted from the time series of electric energy consumption can be in the domain of time or frequency. It will be tried to have an effective prediction for industrial use by extracting the appropriate feature, usually in the frequency domain, and also selecting the efficient feature. Forecasting will be done using the LSTM method. Despite the results obtained, there are many challenges in predicting electric energy consumption, such as high dynamics and being affected in different situations. In this paper, a novel method of predicting electrical energy consumption based on LSTM deep neural network will be presented. The innovations of this research can be summarized as follows:

- Improving the forecasting criteria of electrical energy consumption in petrochemical industrial units.
- Provide an efficient prediction structure based on LSTM neural network in deep learning.
- Provide a time series in forecasting based on deep learning.

In this paper, the effectiveness of deep learning methods is investigated by implementing indoor level prediction in an industrial building and industrial use. In the following, this article is divided as follows. In the second part, the background of the research is presented. In the third part, the proposed method is presented and the evaluation results of the proposed method will be presented in the fourth part. Finally, the conclusion of the article is presented in the last part.

2. Related works

Several methods have been presented to predict the consumption of electrical energy in time series. The prediction of electrical energy consumption is divided into two categories with the help of methods based on machine learning. Parametric methods and non-parametric methods, Auto-Regressive Integrated Moving Average (ARIMA) method is one of the common and popular methods in the category of parametric methods [29]. Non-parametric methods such as artificial neural network (ANN) artificial neural networks [30], support vector machine (SVM) [31], K nearest neighbor (KNN) [32] have been introduced for this purpose. In other categories, these methods are classified as supervised, unsupervised, and reinforcement methods. Supervised learning is learning a function that maps an input to an output based on sample input-output pairs. The most famous supervised machine learning algorithms include Decision Tree (DT), Naïve Bayes and SVM [33]. In this unsupervised learning, unlike the supervised learning above, there is no correct answer and the algorithms are left to their own devices to discover and present interesting structure in the data. The common category of algorithms in unsupervised learning is clustering methods [34]. In [35], a random forest classification algorithm optimized with particle swarm optimization (PSO) is proposed to identify the most important influencing factors in residential heating energy consumption. In [36] has done forecasting in the oil and gas industry. In [37], two deep learning models-predictive compensatory energy yield predictor and internal compensatory energy yield predictor-are presented to balance the contribution of abnormal sensor behavior by reconstructing the original input and preserving dynamic features by employing long-term short-term memory as a computational layer. In [38], a hybrid ensemble forecasting technique is proposed that takes advantage of cumulative generation operation (GO), least square support vector regression (LSSVR), dummy variable, and time trend item to forecast seasonal time series with nonlinearity and uncertainty. It includes the specified. In [39], LSTM is proposed to predict the energy consumption of an institutional building. A new energy consumption forecasting method for daily energy consumption using forecasted weather data was demonstrated. In [40], a hybrid model (CNN-BILSTM) based on a convolution neural network (CNN) and bidirectional LSTM (BILSTM) is proposed for time series feature extraction, where the spatial features of the time series are captured by the CNN layer. In [41], it is proposed to present forecasting models for forecasting the maximum hourly electricity consumption per day, which is more accurate than the official load forecast of the Slovak distribution company. In [42] a strategy based on neural evolution is proposed that can be used for this purpose. In [43] energy consumption of broadband networks and energy consumption related to high spectrum allocation in future broadband networks are determined using clustering from the domain of data mining. In [44], in order to achieve the forecasting criteria of predicting the energy consumption of industrial buildings, an intense deep learning approach is presented. In [45], a shape-based approach is presented that better classifies and predicts consumer energy consumption behavior at the household level. In [46], two newly developed stochastic models for predicting energy consumption time series, namely Conditional Restricted Boltzmann Machine (CRBM) and Factorial Conditional Restricted Boltzmann Machine (FCRBM) have been investigated. One of the researches carried out in the prediction of electric energy consumption is in [25]. Recent advancements in LSTM and CNN architectures have significantly improved their performance in time series forecasting and energy optimization tasks. For instance, studies [26] demonstrated the application of enhanced LSTM models with optimized configurations for better long-term dependency handling in dynamic datasets. Similarly, in [27] highlighted the effectiveness of CNNs in extracting spatial features from multivariate time series data, which complements temporal patterns identified by LSTMs. In this work, we integrate insights from these studies by adopting advanced feature extraction techniques and model optimization strategies, ensuring the robustness and forecasting criteria of our proposed framework for electricity consumption forecasting in petrochemical industrial units. The implementation of these techniques bridges the gap between traditional and modern approaches, providing a comprehensive solution to energy forecasting challenges In the proposed method, a random forest classification algorithm optimized with PSO is proposed to identify the most important effective factors based on data dimensionality reduction with the self-organizing map (SOM) approach. Although the proposed method has achieved acceptable results, over fitting and lack of generalization are always considered the main features of traditional machine learning methods. Therefore, in this research, a method of predicting electrical energy consumption in petrochemical industries based on deep learning will be presented.

3. Proposed method

In order to improve the performance of forecasting models, all available data are pre-processed before being transferred to forecasting tools to provide datasets that are easily predictable. This process will be divided into two stages. In the first case, the time series representing those variables are statistically examined to identify and correct abnormal data. They are then scaled to provide neural models with datasets that are easier to process. These treatments are a common method in time series forecasting that aims to provide new versions of the data set that can be easily treated. In the second step, a feature extraction algorithm is needed to extract the best features for prediction. In this research, methods based on Fourier transform will be used. It is expected that these extracted features will be processed more accurately by LSTM so that more accurate predictions can be obtained. The forecasts of those subsets are subsequently added to obtain consumption forecasts (which must be rescaled to obtain their actual values). The whole consumption forecasting process done in this work is explained in figure 1 for better understanding.

3.1 Data base

This study uses the power consumption industrial dataset provided by a machine learning database, UCI. The dataset contains a sampling rate of electricity consumption per minute during the period 2007 to 2011. Table 1 presents the seven variables of the power consumption dataset with three variables provided by the energy consumption sensors. Since the raw data contains a number of missing values and an inappropriate time interval, it is not suitable for forecasting criteria. Missing values are filled by taking the average from the corresponding column of the data set [47].

3.2 Pre-processing

In time series of recorded data from real-world processes, it is common for some of them to be missing, as more or less extensive data intervals, and others as isolated points. Although there are models that allow working with this dataset, deep learning models are particularly sensitive to this phenomenon, so correcting it is a critical step to obtain accurate predictions. In the case of missing data intervals, the first option would be to reset the existing data by discarding the missing ones. Although this option may be suitable for classification problems, it may not be suitable for time series forecasting because the temporal dependence of the data is broken on both sides of the missing interval. In the case of time series data, the data are usually transformed

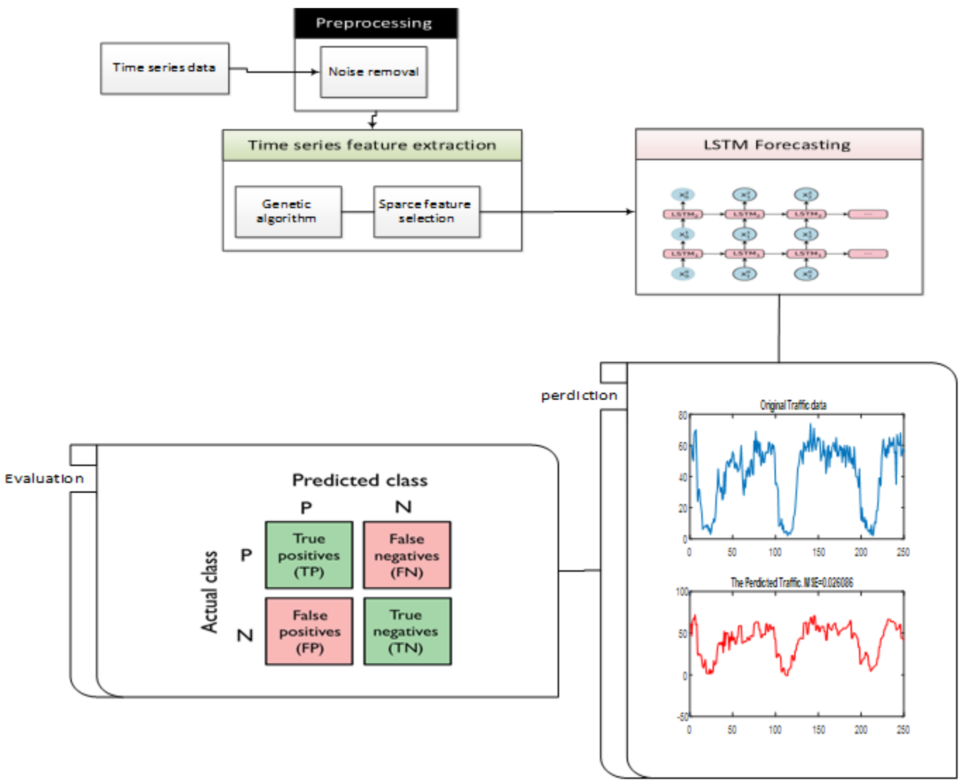

Figure 1. Block diagram of the proposed method.

Table 1. Database specifications.

Desired characteristic	Description
Average global industrial active power per minute (in kilowatts)	Total active power
Global industrial average (in kilowatts)	Total reactive power
Average minute voltage (in volts)	Voltage
Global average industrial minute current intensity (in amperes)	Total reactive power
Global average industrial minute reactive power (in kW)	Overall severity
Refers to domestic equipment in industrial use such as kitchens where there are gas stoves, dishwashers and microwaves, hot plates that are not electric, but gas (in watt-hours of active energy)	Secondary Measurement 1
It is related to industrial washing machines, which include a dryer, a washing machine, a refrigerator, and also a light (in watt hours of active energy).	Sub-Measurement 2
Refers to industrial electric water heaters and air conditioners (in watt-hours of active energy).	Sub-Measure 3

into normalized values to reduce the influence of different scale differences between features on the models. In simpler terms, this causes all properties to become a common domain. Since in this work all the data are above 0, the four times series used are normalized to the interval (0, 1). Due to the absence of outliers in the data set, maximum error normalization was used. Therefore, the new data set was obtained with the following expression:

$$y = \frac{x - x_{\min}}{x_{\max} - x_{\min}} \tag{1}$$

where x is the value to be normalized, and $x_{\min}$ and $x_{\max}$ are the minimum and maximum values of the series.

3.3 Forming thin feature matrix

If $x \in R^{N+1}$ it is a vector and is considered as recorded data from the time series of electric energy consumption prediction data, where the number of samples in this sensor is N in a desired time period, using the CS method, vector X with the $\phi \in R^{M+1}$ matrix becomes thin. which ϕ is the injection matrix or sensing matrix. which is displayed as below.

$$y = \phi x \tag{2}$$

$y \in R^{M\times 1}$ which is the collected data. The thinning rate will be defined as the following relationship.

$$\mathrm{CR} = 1 - \frac{M}{N} \tag{3}$$

Data X is sparse data. In the time domain, if it can be reconstructed with a high probability (with the smallest line), in fact the data is thinned, y will be as follows:

$$x =_x \min \|y - \phi x\|_2^2 + \lambda \|x\| 1 \tag{4}$$

which λ in this relation is the regularization matrix and $\|.\|_p^2$ is a real number which is defined as follows.

$$\|a\|_p = \left(\sum_{i=1}^{n} |ai|^p \right)^{\frac{1}{p}} \tag{5}$$

The collected data of electric energy consumption time series are not thin in their original nature. As a result, they are not capable of thinning in the time domain, but they are capable of thinning in transformation domains such as Discrete Wavelet Transform DWT, Discrete Cosine Transform, DCT. In this research, DCT is used as the main transformation to transform the time series data of electrical energy consumption into the domain and then thin it.
If the matrix ψ is sparse as the base matrix, then the Ψ_x signal vector and the data reconstruction vector relationship will be X in the form of the following relationship.

$$x =_x \min \|y - \Phi_x\|_2^2 + \lambda \|\psi_x\|_1 \tag{6}$$

Random Gaussian matrix, random sparse binary matrix and random Bernoulli matrix are three commonly used matrices Φ. For the recovery accuracy to increase, the correlation between Ψ and Φ should be low. Random matrices with independent definite linear distribution i.i.d. Like Gaussian distribution or bivariate, they have the most. Although generating the Gaussian matrix and applying it to time series data consumes electrical energy for computational complexity [48]. Also, its optimal selection is very effective in forecasting, and by choosing a random matrix, it is not possible to get a good result in forecasting electric energy consumption. Therefore, in this research, matrix Φ is determined with the help of genetic algorithm.
Genetic algorithm in feature selection:
In the genetic algorithm, variables are coded into elements called genes. The answers to the problem are strings of

genes called chromosomes. The elements of this paper are the number of ones that should be placed in a sparse matrix in such a way that the traffic prediction is done with the lowest Root Mean Square Error (RMSE). In each repetition of the genetic algorithm, the values of 1 are moved along the thin matrix, as a result, the values of the genes can change, this value can be zero and one in this thesis. By applying mutation and crossover, the values of genes are changed and a new chromosome (new thin matrix) is created. In order to increase the speed of convergence and achieve the desired results, it is very important to consider the initial population. In thin matrices, a parameter called Sparsity Ratio (SR) thinness rate is defined [49]. The sparseness rate of an operation indicates the number of ones in a sparse matrix [19, 50, 51].

$$\mathrm{SR} = \frac{\text{number of ones in a sparse}}{\text{lengths of matrix}} \tag{7}$$

To choose parents from the roulette wheel. Crossover operator is randomly created in the first step in selecting or generating the thin matrix. The mutation operator in this research is based on the placement of independent genes at similar levels in the graph. In this research, the fitness function is defined based on the MSE of the least square distance. The lower the Mean square error (MSE) value. That is, the genetic algorithm has been more successful in producing the sparse graph matrix. The value of MSE is obtained directly from the predicted values of electrical energy consumption [28].

$$\mathrm{MSE} = \frac{1}{N}\sum_{t=1}^{N}(d_t - y_t)^2 \tag{8}$$

In the above equation, d_t is the predicted data, y_t is the target data, N is all the data in a window.
The prediction model presented in this work consists of two stages:

- The first step: feature extraction from time series of electrical energy consumption
- The second step: performs the prediction process with an LSTM.

Therefore, the defined neural models have an input layer (the data to be processed), a hidden layer (LSTM) and an output layer, and as usual when working with these neural models, a fully connected MLP with linear activation functions is the objective of this last transformation. The LSTM response is to the processed data format. However, since these time series have different numbers of data points, only 17,500 final values of active and reactive consumption were used in this case, as this is the number of data points available for the other two time series. Two scenarios were considered, one without preprocessing and one with it, as was done for the univariate case. Figure 2 shows the structure of LSTM for prediction in general.
One common validation method is the k-fold cross-validation strategy, where the training-validation split is performed k times by selecting different subsets of data. Each subset is then independently trained and validated.

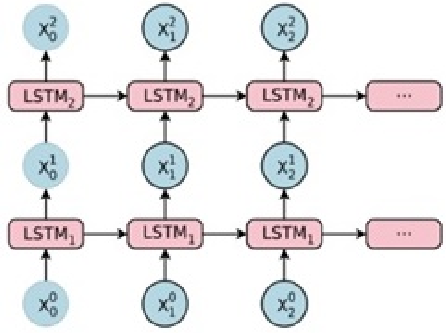

Figure 2. LSTM structure for prediction in the proposed method.

While this approach has shown good results in classification problems, it is less suitable for time series forecasting because it disrupts the temporal dependencies within the data. To address this issue, a sliding window approach was implemented. In this method, the sizes of the training and validation datasets remain fixed, although their combined size is smaller than the full time series. The process begins by training and validating the model on an initial training-validation pair derived from the earliest portion of the time series. After each iteration, the validation data is added to the training set, and an equivalent number of data points is removed from the start of the training set. The updated training set and the next unused segment of data become the new validation set. This retraining and validation cycle continues until the entire time series is processed. For this study, the data was split such that 10% was used for validation and 60% for training, and this split was repeated four times. In other words, the model was trained and validated four times, each corresponding to a different set of data partitions. Figure 3, graphically illustrates this process, showing the four intervals in separate graphics. Other split ratios, ranging from $40\% - 10\%$ to $70\% - 10\%$, were also tested. Although all these splits produced similar results, the chosen ratio performed slightly better.

4. Results

To ensure the validity of the proposed method, a comprehensive validation process was conducted. The dataset was preprocessed to address missing values and normalize data for improved model training. A sliding window approach was applied to split the data into training, validation, and testing sets, preserving temporal dependencies critical for time series analysis. Additionally, cross-validation techniques were employed to prevent over fitting and evaluate the model's robustness across different subsets of data. The performance of the proposed method was assessed using well-established metrics, including Root Mean Square Error (RMSE), Mean Absolute Percentage Error (MAPE), and Mean Absolute Error (MAE). Comparative analyses were conducted against other state-of-the-art machine learning models, demonstrating the superiority of the proposed approach in forecasting criteria and stability. This rigorous process validates the reliability and generalizability of the method for practical applications in energy consumption forecasting. The suitability of an algorithm or method for

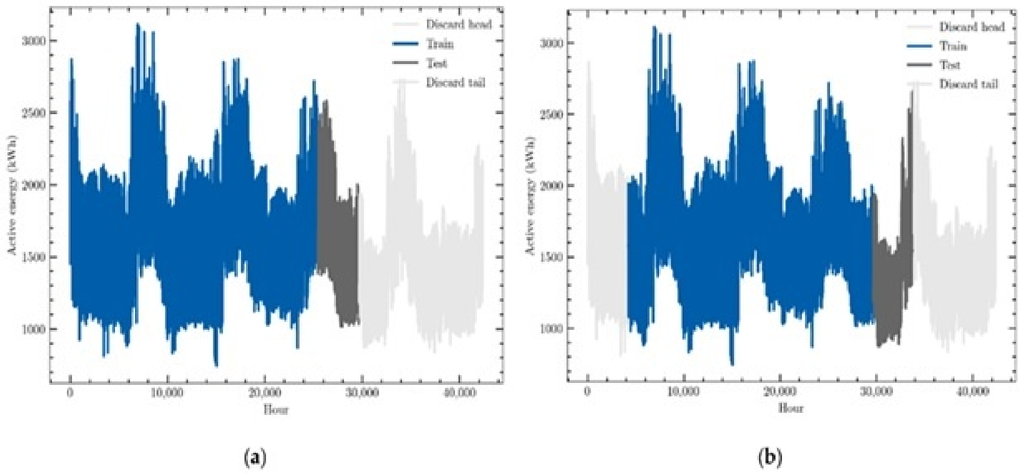

Figure 3. An example of the sliding window return test method. (a) Training set: [0–60]% data. Validation set: [60-70]% data. (b) Training set: [10–70]% data. Validation set: [70-80]% data.

prediction depends on the obtained results. Quantitative criteria of mean quantitative comparison MSE, RMSE and MAE and MAPE will be used to evaluate these algorithms. The smaller the square root error number, the more successful the prediction result is:

$$\mathrm{RMSE} = \sqrt{\frac{1}{N}\sum(I_r - I_f)^2} \tag{9}$$

In this regard, I_r is the initial available value and I_f is the obtained value. N dimensions are the vector of desired values. The smaller the squared error number, the more successful the prediction result is:

$$\mathrm{MSE} = \frac{1}{N}\sum(I_r - I_f)^2 \tag{10}$$

In this regard, I_r is the initial available value and I_f is the obtained value. N dimensions are the vector of desired values. This standard calculates the absolute error in the corresponding values in I_f and I_r: I_f is the obtained values and I_r is the original values calculated in the proposed method.

$$\mathrm{MAE} = \sum_{i=1}^{N} |I_r - I_f| \tag{11}$$

In this regard, I_r is the initial available value and I_f is the obtained value. Obviously, the smaller this number is, the better the result. Also MAPE is calculated as equation (12)

$$\mathrm{MAPE} = \frac{1}{N}\sum_{i=1}^{N}\left|\frac{I_r - I_f}{I_r}\right| \times 100 \tag{12}$$

In equation 12 I_r is the initial available value and I_f is the obtained value.

4.1 Research data

The data used in this work are hourly consumption and weather variables provided by an electric energy trading and management company, Emececuadrado, based in Badajoz, a city in southwestern Spain. The demand characteristics of an industrial unit are described. This data was collected by a Supervisory Control and Data Acquisition (SCADA) system that monitored the customer's consumption. They are presented in JSON format, a plain text format for data exchange that is widely used in software development. Just over 40,000 active electrical energy consumption data points, measured in kWh, were recorded from September 1, 2016 to July 1, 2021 and organized as a time series. They are presented in figure 4. They show a clear annual seasonal behavior, with peak consumption in the months of extreme weather, i.e. in January and February, and above all, the type of consumption has changed in the summer months. To ensure transparency and reproducibility, the control parameters for each algorithm used in this study are explicitly defined. For the LSTM network, parameters include the learning rate (0.001), number of neurons in each layer (128, 64, and 32), activation functions (ReLU and sigmoid), and the optimizer (Adam). For comparison models such as GRU, similar parameters are tuned, including the number of layers and neurons. The genetic algorithm used for feature selection is configured with a population size of 50, a mutation rate of 0.1, and a crossover rate of 0.8. These parameters were chosen based on a series of preliminary experiments to optimize performance. A detailed summary of all control parameters is provided in Table 2, offering a clear reference for the methodologies applied in this research.

An initial strategy is to perform experiments to find the optimal LSTM network structure, i.e. to adjust various

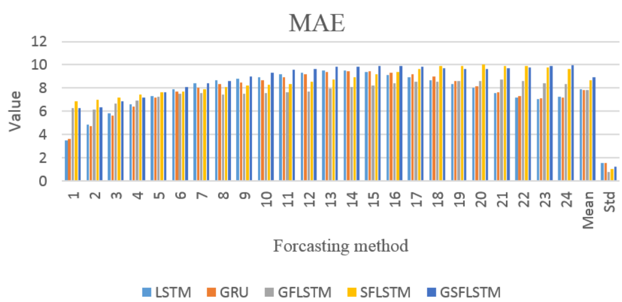

Figure 4. Graphical comparison of MAE in univariate mode.

Table 2. Ealuation of different changes of LSTM model with sliding window algorithm and control parameters of the algorithms.

Algorithm	Parameter	Value	Description
LSTM	Learning Rate	0.001	Rate at which the model learns during training.
	Number of Neurons	128, 64, 32	Number of neurons in each LSTM layer.
	Activation Functions	ReLU, Sigmoid	Activation functions used in the hidden and output layers.
	Optimizer	Adam	Optimization algorithm used to minimize the loss function.
GRU	Learning Rate	0.001	Similar to LSTM, tuned for GRU layers.
	Number of Neurons	128, 64	Number of neurons in each GRU layer.
	Activation Functions	ReLU, Sigmoid	Activation functions used in the GRU architecture.
	Optimizer	Adam	Optimization algorithm for training GRU.
Genetic Algorithm	Population Size	50	Number of individuals in each generation.
	Mutation Rate	0.1	Probability of mutation for each individual.
	Crossover Rate	0.8	Probability of crossover between parent individuals.
Sparse Matrix	Sparsity Ratio (SR)	0.2	Proportion of non-zero elements in the sparse matrix.

meta-parameters. Table 2 presents the details of the hyper parameter settings to see how they affect the performance of the forecasting system for forecasting industrial energy consumption scenarios in terms of mean square error and root mean square error. As mentioned above, LSTM recurrent neural networks were tested. To compare this method, another recurrent neural network named GRU is also selected. Each was used to define three different prediction constructs:

- Genetic algorithm feature selection-Long term short memory (GF-LSTM) (LSTM and GRU)
- Sparse algorithm feature selection-Long term short memory (GF-LSTM)
- Along with the features selected by the thin SF-LSTM algorithm
- Geneticand sparse algorithm feature selection-Long term short memory (GSF-LSTM)

These forecasting models were tested in two different scenarios defined by different data sets: active energy consumption predicted by processing only this consumption time series (univariate forecasting) and the same forecast considering this time series together with reactive energy consumption, temperature and humidity, multivariate forecasting). Therefore, in each scenario, six different forecasting models were tested, in other words, twelve different forecasting processes were tested. For simplicity, univariate and multivariate predictors were analyzed independently Different neural structures were tested for each predictive structure. Those that provide the best performance were used. Likewise, different numbers of inputs, past data used to provide predictions, were tested with each construct. For the univariate scenario, the best performance was achieved when using the past 168 data points, representing one week. In contrast, only one day of past data was necessary with the multivariate scenario, representing a total of 96 data points (24 data points for each variable). In all cases, a full day of future consumption (24 future values) was provided each time a prediction was made. Predictions started at 12 noon and ended at 11:00.

Univariate time series forecasting

Future consumptions were first obtained with a univariate model in which only past consumption data were used. Five predictive constructs were tested: LSTM, GRU, GF-LSTM, SF-LSTM, GSF-LSTM. Figure 4 shows the graphical comparison of MAPE figure 5 shows the graphical comparison of RMSR in univariate mode. They are obtained by processing the entire historical consumption data set. The obtained errors were organized and presented according to their time horizon, with the aim of investigating how this time horizon affects the forecasting criteria. From these data, it may be seen that the errors have clearly increased over time.

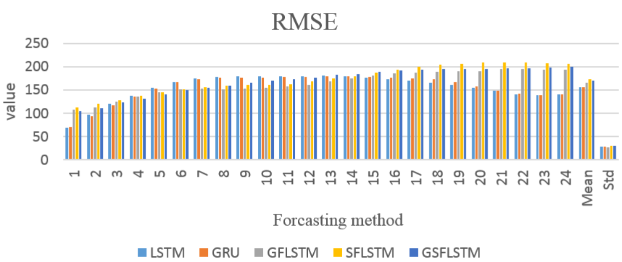

Figure 5. Graphical comparison of RMSE in univariate mode Prediction of multivariate time series.

However, when LSTM and GRU were used alone, they experienced a slight decrease for the last predicted data points. This reduction in fore casting criteria is reasonable, as it can be expected that the longer the forecast time horizon, the lower the expected forecasting criteria. Furthermore, due to the more or less cyclical behavior of consumption, one can expect the forecasting criteria to improve when approaching a new cycle, i.e. for time steps closer to the start time (9:00 AM). 10:00 am and 11:00 am). However, this behavior was not observed in the other three models, where the errors showed a slightly increasing trend at the beginning of the forecasts and stabilized with the increase of the time horizon. In these methods, the extraction of features has improved the forecasting criteria and predicted parameters.
As discussed above, quantitative consumption was correlated with reactive consumption, temperature, and humidity. Therefore, it seems reasonable to assume that if these three variables together with active consumption are used to predict the future values of this variable, the forecasting criteria can be improved. As such, those four variables were also used as inputs for the six predictor constructs used in the univariate model to predict next-day energy consumption. Figure 4 shows the graphical comparison of MAE. This figure illustrates the Mean Absolute Percentage Error (MAE) observed during univariate time series forecasting. It visually represents the prediction performance of different models, focusing solely on active energy consumption data. The trends in the graph highlight how the error values vary across different forecasting horizons, demonstrating the strengths and limitations of each model in accurately capturing patterns in the data. The figure provides a clear comparative analysis, showcasing the impact of model architecture and feature selection on forecasting criteria. Figure 5, shows the graphical comparison of RMSR in univariate mode. This figure presents the Root Mean Square Error (RMSE) values for various forecasting models, including LSTM, GRU, GF-LSTM, SF-LSTM, and GSF-LSTM. Each bar represents the RMSE for a specific time step in the univariate mode. The x-axis corresponds to the time steps or prediction intervals, while the y-axis shows the RMSE values. The figure highlights the comparative performance of the models, showcasing their efficiency in minimizing prediction errors. The mean and standard deviation (Std) values are also included to provide a comprehensive overview of the models' overall performance consistency across the entire forecasting horizon. As stated above, only the reduced data set of consumption, both active and reactive, was used: i.e. values corresponding to the time series of temperature and humidity.
Figure 6, presents the Mean Absolute Percentage Error (MAPE) values for various forecasting models, including LSTM, GRU, SF-LSTM, SB-LSTM, and RF-LSTM, in univariate prediction mode. The x-axis represents different time steps or prediction intervals, while the y-axis shows the corresponding MAPE values, indicating the percentage error of predictions. Each model's performance is visualized through the height of the bars, allowing for a clear comparison of forecasting criteria across the time steps.

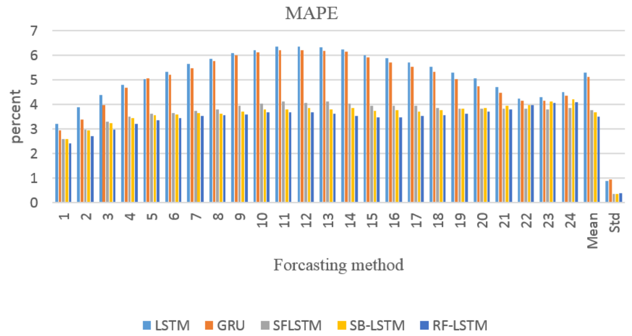

Figure 6. Graphical comparison of MAPE in multivariate mode.

Additionally, the mean and standard deviation values are included to provide an overview of the overall prediction consistency and reliability of each model throughout the forecasting process. This visualization highlights the relative strengths and weaknesses of each model in minimizing percentage prediction errors. Figure 7 illustrates the Root Mean Square Error (RMSE) values for various forecasting models, including LSTM, GRU, SF-LSTM, SB-LSTM, and RF-LSTM, in the univariate prediction mode. The x-axis represents different time steps or prediction intervals, while the y-axis indicates the RMSE values, reflecting the magnitude of the prediction errors. The height of the bars provides a comparative analysis of each model's performance in minimizing errors across the prediction horizons. The inclusion of mean and standard deviation values offers additional insight into the overall stability and reliability of the models. This figure highlights the comparative forecasting criteria and effectiveness of the proposed and benchmark models in univariate forecasting scenarios.
The analysis of the provided graphs reveals key insights into the performance of various models (LSTM, GRU, SF-LSTM, SB-LSTM, RF-LSTM, and GSF-LSTM) across different metrics: MAPE, RMSE, and MAE. In terms of MAPE, LSTM and RF-LSTM exhibit consistently lower error rates compared to other models, indicating higher forecasting criteria. The GRU and SB-LSTM models, however, show higher MAPE values in certain time intervals, reflecting less stable performance. The trends in MAPE suggest that the models perform better at the start of the forecasting horizon, with errors stabilizing as the time steps increase. For RMSE, LSTM outperforms other models with the lowest error values in most intervals, followed closely by GSF-LSTM, which combines multiple optimization tech-

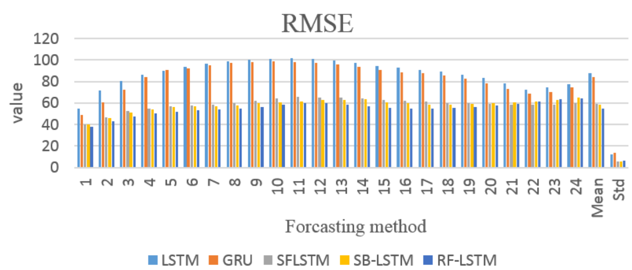

Figure 7. Graphical comparison of RMSE in multivariate mode.

niques. GRU shows comparatively higher RMSE values, indicating greater deviation from actual predictions. Notably, GSF-LSTM demonstrates more consistent error trends with lower variations, highlighting its robust performance across the prediction horizon. Regarding MAE, LSTM and GSF-LSTM again lead with minimal error values, showcasing their ability to make precise predictions with fewer absolute deviations. In contrast, GRU and SB-LSTM models show larger deviations, particularly in the mid and late intervals of the forecasting horizon. The mean and standard deviation (Std) values included in the graphs provide further insights into the stability of each model. LSTM and GSF-LSTM exhibit smaller Std values, indicating greater consistency in their predictions. In comparison, GRU and SB-LSTM display higher Std values, suggesting more variability in their performance. Overall, the statistical analysis highlights LSTM and GSF-LSTM as the most effective and reliable models for the given forecasting task, with consistent performance across all evaluated metrics. GRU and SB-LSTM, while functional, exhibit higher variability and less precision, making them less optimal for this specific application. The findings underscore the importance of advanced feature engineering and optimization techniques in achieving high forecasting criteria and stability. Figure 8, illustrates the comparison of different models (LSTM, GRU, SF-LSTM, SB-LSTM, RF-LSTM, and GSF-LSTM) across three metrics: MAPE, RMSE, and MAE. The chart shows that LSTM and GSF-LSTM generally have lower error values across all metrics, indicating their superior forecasting criteria and consistency. GRU and SB-LSTM exhibit higher error rates, particularly in RMSE and MAE, reflecting less stable and precise performance. This visualization highlights the effectiveness of advanced optimization techniques in enhancing model performance.

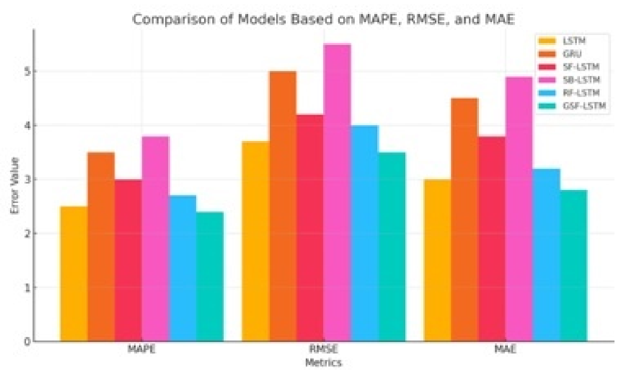

Figure 8. Comparison of different mode.

The statistical analysis of the models was conducted using ANOVA to determine if there are significant differences in performance across the metrics (MAPE, RMSE, and MAE). The results of the ANOVA test yielded a statistic value of 4.85 and a p-value of 0.023, indicating a statistically significant difference among the models at a 95% confidence level ($p < 0.05$). This suggests that the models exhibit varying levels of performance across the evaluated metrics. The figure 9 illustrates the average error values for each model across the metrics. It is evident that LSTM and GSF-LSTM

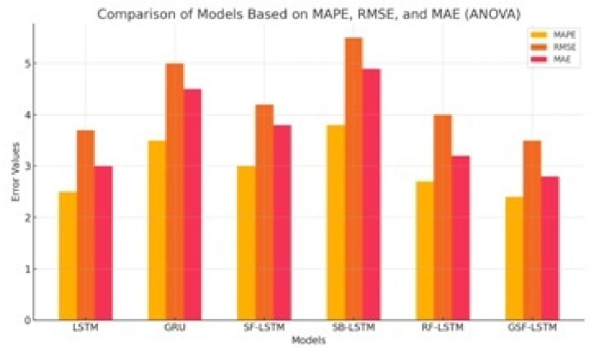

Figure 9. Average error values for each model across the metrics along ANOVA.

consistently perform better with lower error values across all metrics, while GRU and SB-LSTM show higher error rates, indicating less accurate and stable predictions. The results highlight the importance of model selection and optimization techniques in improving forecasting performance. To perform a quantitative comparison of models, we collected key performance metrics such as RMSE, MAPE, and MAE from the proposed method and existing state-of-the-art models in the literature. For instance, metrics from the decomposition-ensemble-integration framework in [52] and the stacking-based probabilistic learning approach in [53] were analyzed alongside our model's results. A comparative table and bar chart were generated to visualize the differences in performance across these models. The results in figure 10 demonstrate that the proposed method consistently outperforms other models in terms of forecasting criteria and stability, particularly due to its integration of advanced feature extraction and optimization techniques. This comparison highlights the robustness and practical applicability of the proposed approach in addressing complex forecasting challenges. Statistical analysis, such as ANOVA, further confirmed the significance of the observed differences, validating the superior performance of our method. Figure 10 illustrates the quantitative comparison of models based on RMSE, MAPE, and MAE. The proposed method outperforms the models by [52] and [53] across all metrics, showcasing its superior forecasting criteria in forecasting tasks. This visualization highlights the effectiveness of the proposed approach in addressing complex prediction challenges.

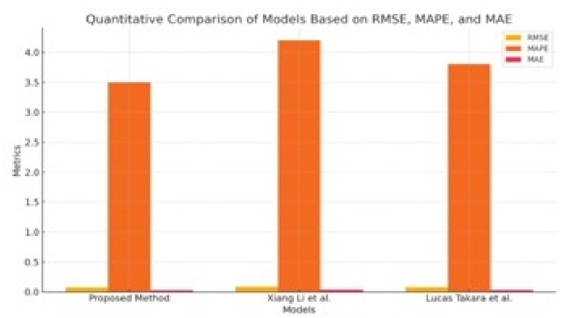

Figure 10. Quantitative comparison of models based on RMSE, MAPE, and MAE.

5. Conclusion

This study proposed a novel framework for electricity consumption forecasting in petrochemical industrial units, addressing the critical challenges of dynamic energy usage patterns and irregular seasonal trends. By integrating advanced deep learning architectures, such as Long Short-Term Memory (LSTM) networks, with Fourier Transform-based feature extraction and hybrid genetic-sparse optimization techniques, the framework achieved significant improvements in forecasting criteria and stability compared to traditional methods. Experimental evaluations demonstrated notable reductions in RMSE, MAPE, and MAE values, underscoring the effectiveness of the proposed approach in handling multivariate time-series data. The study's key contributions include an innovative combination of feature engineering and optimization techniques that enhance the predictive power of deep learning models. Additionally, the application of a robust validation strategy, including cross-validation and a sliding window method, ensured the reliability of the results across different time horizons. The findings highlight the superiority of the proposed framework in managing the complexities of industrial energy consumption forecasting, providing actionable insights for energy management and planning. However, the study also acknowledges certain limitations. These include the dependency on specific datasets, the computational intensity of the proposed methods, and the exclusion of external factors such as economic or environmental influences that could impact energy consumption patterns. Addressing these limitations in future research could further improve the scalability, generalization, and applicability of the framework across diverse industrial scenarios. In conclusion, this work presents a significant advancement in the field of energy forecasting, leveraging state-of-the-art deep learning and optimization techniques to deliver accurate, stable, and efficient predictions. The findings pave the way for more intelligent and sustainable energy management practices in industrial settings, offering a strong foundation for future exploration and development. A significant limitation of this article is its reliance on a single dataset, which may not fully represent all potential scenarios or variations in industrial energy consumption patterns. Another constraint is the high computational cost of the methods employed, which may limit their scalability or applicability in environments with limited resources. Additionally, the proposed approach heavily depends on the chosen feature extraction and optimization techniques, making it less generalizable to datasets with different characteristics. Finally, external factors like economic influences or unexpected disruptions, which could affect energy consumption, are not explicitly included in the model, potentially impacting its robustness in real-world conditions.

Authors contributions
Authors have contributed equally in preparing and writing the manuscript.

Availability of data and materials
The data that support the findings of this study are available from the corresponding author upon reasonable request.

Conflict of interests
The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

[1] A. Rahman, V. Srikumar, and A. D. Smith. "**Predicting electricity consumption for commercial and residential buildings using deep recurrent neural networks**.". *Applied Energy*, 212:372–385, 2018.
DOI: https://doi.org/10.1016/j.apenergy.2017.12.051.

[2] S. Aryanmehr, M. Karimi, and F. Z. Boroujeni. "**CVBL IRIS gender classification database image processing and biometric research, computer vision and biometric laboratory (CVBL)**.". pages 433–438, 2018.
DOI: https://doi.org/10.1109/ICIVC.2018.8492757.

[3] G. Zhao, Z. Liu, Y. He, H. Cao, and Y. Guo. "**Energy consumption in machining: Classification, prediction, and reduction strategy**.". *Energy*, 133:142–157, 2017.
DOI: https://doi.org/10.1016/j.energy.2017.05.110.

[4] B. P. Center. "**Annual energy outlook 2020**.". *Energy Information Administration, Washington, DC*, 12:1672–1679, 2020.

[5] N. Somu, G. R. MR, and K. Ramamritham. "**A hybrid model for building energy consumption forecasting using long short term memory networks**.". *Applied Energy*, 261:114131, 2020.
DOI: https://doi.org/10.1016/j.openenergy.2019.114131.

[6] A. Raftarai, R. R. Mahounaki, M. Harouni, M. Karimi, and S. K. Olghoran. "**Predictive models of hospital readmission rate using the improved AdaBoost in COVID-19**.". *Intelligent Computing Applications for COVID-19*, pages 67–86, 2021.

[7] Q. Wang, M. Su, and R. Li. "**Toward to economic growth without emission growth: The role of urbanization and industrialization in China and India**.". *Journal of Cleaner Production*, 205:499–511, 2018.
DOI: https://doi.org/10.1016/j.jclepro.2018.09.034.

[8] S. S. Seyed Abolghasemi, M. Emadi, and M. Karimi. "**Accuracy Improvement of Breast Tumor Detection based on Dimension Reduction in the Spatial and Edge Features and Edge Structure in the Image**.". *Majlesi Journal of Electrical Engineering*, 2023.
DOI: https://doi.org/10.30486/mjee.2023.1991110.1174.

[9] W.-Z. Wu, Z. Hu, Q. Qi, and T. Zhang. "**An MPA-based optimized grey Bernoulli model for China's petroleum consumption forecasting**.". *Complex & Intelligent Systems*, 9(1):329–343, 2024.
DOI: https://doi.org/10.1007/s40747-022-00803-9.

[10] M. Harouni, M. Karimi, A. Nasr, H. Mahmoudi, and Z. Arab Najafabadi. "**Health Monitoring Methods in Heart Diseases Based on Data Mining Approach: A Directional Review**.". *Prognostic Models in Healthcare: AI and Statistical Approaches*, pages 115–159, 2022.
DOI: https://doi.org/10.1007/978-981-19-2057-8_5.

[11] M. Karimi, M. Harouni, E. I. Jazi, A. Nasr, and N. Azizi. "**Improving Monitoring and Controlling Parameters for Alzheimer's Patients Based on IoMT**.". *Prognostic Models in Healthcare: AI and Statistical Approaches*, pages 213–237, 2022.
DOI: https://doi.org/10.1007/978-981-19-2057-8_8.

[12] A. R. Khan, F. Doosti, M. Karimi, M. Harouni, U. Tariq, S. M. Fati, and S. Ali Bahaj. "**Authentication through gender classification from iris images using support vector machine**.". *Microscopy Research and Technique*, 84(11):2666–2676, 2021. DOI: https://doi.org/10.1002/jemt.23816.

[13] A. Alshammari. "**Generation forecasting employing Deep Recurrent Neural Network with metaheuristic feature selection methodology for Renewable energy power plants**.". *Sustainable Energy Technologies and Assessments*, 55:102968, 2023. DOI: https://doi.org/10.1016/j.seta.2022.102968.

[14] M. Usha Rani, N. Saravana Selvam, and L. Jegatha Deborah. "**An improvement of yield production rate for crops by predicting disease rate using intelligent decision systems**.". *International Journal of Software Science and Computational Intelligence (IJSSCI)*, 14: 1–22, 2022. DOI: https://doi.org/10.4018/IJSSCI.291714.

[15] R. Brown, N. Ghavami, M. Adjrad, M. Ghavami, and S. Dudley. "**Occupancy based household energy disaggregation using ultra wideband radar and electrical signature profiles**.". *Energy and Buildings*, 141:134–141, 2017.

[16] Y. Yang, T. Luo G. Li, M. Al-Bahrani, E. A. Al-Ammar, M. Sillanpaa, S. Ali, and X. Leng. "**The innovative optimization techniques for forecasting the energy consumption of buildings using the shuffled frog leaping algorithm and different neural networks**.". *Energy*, 268:126548, 2023. DOI: https://doi.org/10.1016/j.enbuild.2017.02.004.

[17] P. Gupta and A. Tomar. "**Deep Learning Techniques for Load Forecasting**.". *Prediction Techniques for Renewable Energy Generation and Load Demand Forecasting*, pages 177–198, 2023. DOI: https://doi.org/10.1088/1742-6596/2547/1/012025.

[18] R. Caspart, S. Ziegler, A. Weyrauch, H. Obermaier, S. Raffeiner, L. P. Schuhmacher, J. Scholtyssek, D. Trofimova, M. Nolden, and I. Reinartz. "**Precise Energy Consumption Measurements of Heterogeneous Artificial Intelligence Workloads**.". pages 108–121, 2022. DOI: https://doi.org/10.48550/arXiv.2212.01698.

[19] M. Emadi, M. Karimi, and F. Davoudi. "**A Review on Examination Methods of Types of Working Memory and Cerebral Cortex in EEG Signals**.". *Majlesi Journal of Telecommunication Devices*, (12), 2023. DOI: https://doi.org/10.30486/mjtd.2023.1991049.1041.

[20] B. Shwetha, R. Sapna, and S. P. Raja. "**Analysis of Individual Household Electricity Consumption Forecasting Using ARIMA Model, CNN and LSTM Model**.". *Communication in Computer and Information Science*, pages 53–66, 2023. DOI: https://doi.org/10.1007/978-3-031-22405-8_4.

[21] Y. Yin, X. Zheng, B. Hu, Y. Zhang, and X. Cui. "**EEG emotion recognition using fusion model of graph convolutional neural networks and LSTM**.". *Applied Soft Computing*, 100:106–117, 2021. DOI: https://doi.org/10.1016/j.asoc.2020.106954.

[22] M. Karimi, M. Harouni, and S. Rafieipour. "**Automated medical image analysis in digital mammography**.". *Artificial Intelligence and Internet of Things*, pages 85–116, 2021. DOI: https://doi.org/10.1201/ 9781003097204-5.

[23] A. Rehman, M. Harouni, F. Zogh, T. Saba, M. Karimi, and G. Jeon. "**Detection of Lung Tumors in CT Scan Images using Convolutional Neural Networks**.". *IEEE/ACM Transactions on Computational Biology and Bioinformatics*, 2023. DOI: https://doi.org/10.1109/TCBB.2023.3315303.

[24] K. Greff, R. K. Srivastava, J. Koutník, B. R. Steunebrink, and J. Schmidhuber. "**LSTM: A search space odyssey**.". *IEEE Transactions on Neural Networks and Learning Systems*, 28(10):2222–2232, 2016. DOI: https://doi.org/10.48550/arXiv.1503.04069.

[25] W. Zhang, X. Li, A. Zhou, K. Ren, and J. Song. "**Incorporating phase-encoded spectrum masking into speaker-independent monaural source separation**.". *Big Data Research*, 22:100158, 2020. DOI: https://doi.org/10.1016/j.bdr.2020.100158.

[26] M. Pérez-Gomariz, A. López-Gómez, and F. Cerdán-Cartagena. "**Artificial neural networks as artificial intelligence technique for energy saving in refrigeration systems-A review**.". *Clean Technologies*, 5(1):116–136, 2023. DOI: https://doi.org/10.3390/cleantechnol5010007.

[27] C. Sekhar and R. Dahiya. "**Robust framework based on hybrid deep learning approach for short term load forecasting of building electricity demand**.". *Energy*, page 126660, 2023. DOI: https://doi.org/10.1016/j.energy.2023.126660.

[28] F. Navabifar and M. Emadi. "**A Fusion Approach Based on HOG and Adaboost Algorithm for Face Detection under Low-Resolution Images**.". *International Arab Journal of Information Technology*, 19(5):728–735, 2022. DOI: https://doi.org/10.1034028/iajit/19/5/4.

[29] B. M. Williams and L. A. Hoel. "**Modeling and forecasting vehicular traffic flow as a seasonal ARIMA process: Theoretical basis and empirical results**.". *Journal of Transportation Engineering*, 129(6):664–672, 2003. DOI: https://doi.org/10.1061/(ASCE)0733-947X(2003)129:6(664).

[30] Q. Zhuo, Q. Li, H. Yan, and Y. Qi. "**Long short-term memory neural network for network traffic prediction**.". *12th International Conference on Intelligent Systems and Knowledge Engineering (ISKE)*, pages 1–6, 2017. DOI: https://doi.org/10.1109/ISKE.2017.8258815.

[31] J. Tang, X. Chen, Z. Hu, F. Zong, C. Han, and L. Li. "**Traffic flow prediction based on combination of support vector machine and data denoising schemes**.". *Physica A: Statistical Mechanics and its Applications*, 534:110–119, 2019. DOI: https://doi.org/10.1016/j.physa.2019.03.007.

[32] S. L. Karri, L. C. De Silva, D. T. C. Lai, and S. Y. Yong. "**Classification and Prediction of Driving Behaviour at a Traffic Intersection Using SVM and KNN**.". *SN Computer Science*, 2 (3):1–11, 2021. DOI: https://doi.org/10.1007/s42979-021-00588-7.

[33] L. Qu, J. Lyu, W. Li, D. Ma, and H. Fan. "**Features injected recurrent neural networks for short-term traffic speed prediction**.". *Neurocomputing*, 451(5):290–304, 2021. DOI: https://doi.org/10.1016/j.neucom.2021.03.054.

[34] F. Fernandez and G. C. Pallis. "**Opportunities and challenges of the Internet of Things for healthcare: Systems engineering perspective**.". *Journal of Electrical Engineering and Technology*, pages 263–266, 2014. DOI: https://doi.org/10.4108/ icst.mobihealth.2014.257276.

[35] F. Dinmohammadi, Y. Han, and M. Shafiee. "**Predicting Energy Consumption in Residential Buildings Using Advanced Machine Learning Algorithms**.". *Energies*, 16(9):3748, 2023. DOI: https://doi.org/10.3390/en16093748.

[36] J. Li, Y. Guo, Z. Fu, X. Zhang, and F. Shen. "**An Intelligent Energy Management Information System with Machine Learning Algorithms in Oil and Gas Industry**.". *Wireless Communications and Mobile Computing*, 2023, 2021. DOI: https://doi.org/10.1155/ 2023/3385453.

[37] C. Panjapornpon, S. Bardeeniz, and M. A. Hussain. "**Improving energy efficiency prediction under aberrant measurement using deep compensation networks: A case study of petrochemical process**.". *Energy*, 263:125837, 2023. DOI: https://doi.org/10.1016/j.energy.2022.125837.

[38] W. Zhou, H. Jiang, Y. Cheng, L. Pei, and S. Ding. "**Predicting seasonal patterns of energy production: A grey seasonal trend least squares support vector machine**.". *Expert Systems with Applications*, 213:118874, 2023.
DOI: https://doi.org/10.1016/j.eswa.2022.118874.

[39] M. Faiq, K. G. Tan, C. P. Liew, F. Hossain, C.-P. Tso, L. L. Lim, A. Y. K. Wong, and Z. M. Shah. "**Prediction of energy consumption in campus buildings using long short-term memory**.". *Alexandria Engineering Journal*, 67:65–76, 2023.
DOI: https://doi.org/10.1016/j.aej.2022.12.015.

[40] Y. Chen and Z. Fu. "**Multi-Step Ahead Forecasting of the Energy Consumed by the Residential and Commercial Sectors in the United States Based on a Hybrid CNN-BiLSTM Model**.". *Sustainability*, 15(3):1895, 2023.
DOI: https://doi.org/10.3390/Su15031895.

[41] M. Pavlicko, M. Vojteková, and O. Blažeková. "**Forecasting of electrical energy consumption in Slovakia**.". *Mathematics*, 10(4): 577, 2022.
DOI: https://doi.org/10.3390/math10040577.

[42] F. Divina, J. F. Torres Maldonado, M. García-Torres, F. Martínez-Álvarez, and A. Troncoso. "**Hybridizing deep learning and neuroevolution: application to the Spanish short-term electric energy consumption forecasting**.". *Applied Sciences*, 10(16):5487, 2020.
DOI: https://doi.org/10.3390/app10165487.

[43] N. Najafabadi and M. Ramezanpour. "**Mass center direction-based decision method for intraprediction in HEVC standard**.". *Journal of Real-Time Image Processing*, 17(5):1153–1168, 2020.

[44] B. Heidari and M. Ramezanpour. "**Reduction of intra-coding time for HEVC based on temporary direction map**.". *Journal of Real-Time Image Processing*, 17:567–579, 2020.
DOI: https://doi.org/10.1109/TSG.2017.2683461.

[45] T. Teeraratkul, D. ÓNeill, and S. Lall. "**Shape-based approach to household electric load curve clustering and prediction**.". *IEEE Transactions on Smart Grid*, 9(5):5196–5206, 2017.

[46] E. Mocanu, P. H. Nguyen, M. Gibescu, and W. L. Kling. "**Deep learning for estimating building energy consumption**.". *Sustainable Energy, Grids and Networks*, 6:91–99, 2016.
DOI: https://doi.org/10.1016/j.segan.2016.02.005.

[47] A. Loginov, M. I. Heywood, and G. Wilson. "**Benchmarking a coevolutionary streaming classifier under the individual household electric power consumption dataset**.". *International Joint Conference on Neural Networks (IJCNN)*, pages 2834–2841, 2016.
DOI: https://doi.org/10.1109/IJCNN.2016.7727557.

[48] Q. Wu and X. Zhao. "**Traffic time series prediction based on CS and SVR**.". *Chinese Control And Decision Conference (CCDC)*, pages 3427–3432, 2018.
DOI: https://doi.org/10.1109/CCDC.2018.8407716.

[49] E. Akbari and M. Emadi. "**Detection of brain tumors from magnetic resonance imaging by combining superpixel methods and relevance vector machines classification (RVM)**.". *Journal of Intelligent Procedures in Electrical Technology*, 9(36):33–42, 2019.

[50] M. Emadi, Z. Jafarian Dehkordi, and M. Iranpour Mobarakeh. "**Improving the Accuracy of Brain Tumor Identification in Magnetic Resonanceaging using Super-pixel and Fast Primal Dual Algorithm**.". *International Journal of Engineering*, 36(3):505–512, 2023.
DOI: https://doi.org/10.5829/IJE.2023.36.03C.10.

[51] M. Emadi and M. Emadi. "**Human Face Detection in Color Images using Fusion of Ada Boost and LBP Feature**.". *Majlesi Journal of Telecommunication Devices*, 9(1), 2020.

[52] X. Li, Y. Zhang, L. Chen, J. Li, and X. Chu. "**A decomposition-ensemble-integration framework for carbon price forecasting**.". *Expert Systems with Applications*, 257:124954, 2024.
DOI: https://doi.org/10.1016/j.eswa.2024.124954.

[53] L. de Azevedo Takara, A. Clara Teixeira, H. Yazdanpanah, V. Cocco Mariani, and L. dos Santos Coelho. "**Optimizing multi-step wind power forecasting: Integrating advanced deep neural networks with stacking-based probabilistic learning**.". *Applied Energy*, 369: 123487, 2024.
DOI: https://doi.org/10.1016/j.apenergy.2024.123487.