ORIGINAL PAPER

Multi-Term Electrical Load Forecasting of Smart Cities Using a New Hybrid Highly Accurate Neural Network-Based Predictive Model

Ashkan Safari[1] · Hamed Kharrati[1,2] · Afshin Rahimi[2]

Abstract

This paper presents FARHAN, a novel hybrid model designed to address the challenges of electrical load forecasting in smart grids. FARHAN combines descending neuron attention, long/short-term memory (LSTM), and Markov-simulated neural networks to optimize accuracy and analysis time for short-, mid-, and long-term smart grid planning decisions. FARHAN processes electricity load data efficiently by utilizing two LSTM blocks (LSTM.B1 & LSTM.B2) with attention layers, a 90% gain averager, and a Markov chain analyzer. The comparative analysis demonstrates FARHAN's superiority over traditional LSTM models and other methodologies, exhibiting remarkable Mean Absolute Percentage Errors (MAPEs) of 0.019162%, 0.0386%, and 0.039% for 14 years, annual, and monthly estimations, respectively. Root Mean Square Percentage Errors (RMSPEs) of 2.5%, 5.2%, and 1.2% and an overall R^2 of 1 validate its exceptional accuracy. FARHAN's innovative approach establishes it as a robust and intelligent tool for enhancing electrical load forecasting in smart grids and energy systems, promising significant advancements in the field.

Abbreviations

AI	Artificial Intelligence
ML	Machine Learning
NN	Neural Network
NB	Naïve Bayes
RF	Random Forest
DR	Demand Response
EV	Electrical Vehicle
ANN	Artificial Neural Network
QNN	Quantum Neural Network
CNN	Convolutional Neural Network
RNN	Recurrent Neural Network
KNN	K Nearest Neighbor
SVM	Support Vector Machine
DRM	Demand Response Management
DSM	Demand Side Management
RES	Renewable Energy Sources
ESS	Energy Storage System
GRU	Gate Recurrent Unit
KPI	Key Performance Indicator
LSTM	Long/Short Term Memory

Parameters & Subscripts

$Con.V$	Consumption Value
σ	Sigmoid Function
$\overline{[Con.V]}$	Mean of Initial Consumption Values
$S^2_{observed}$	Sample Variance of Observed Consumption Values
$\overline{h_{2,t}}$	Mean of Predicted Consumption Values
$S^2_{predicted}$	Sample Variance of Predicted Consumption Values
P	Probability
X_n	Random Variable
i_n	Possible Stages
i_{1t}	Input Gate of LSTM.B1
f_{1t}	Forget Gate of LSTM.B1
o_{1t}	Output Gate of LSTM.B1
$\widetilde{c}_{1t}$	Memory cell candidate of LSTM.B1

✉ Ashkan Safari
ashkansafari.unv@gmail.com

Hamed Kharrati
kharrati@tabrizu.ac.ir; kharrati@uwindsor.ca

Afshin Rahimi
arahimi@uwindsor.ca

1 Faculty of Electrical and Computer Engineering, University of Tabriz, Tabriz, Iran

2 Department of Mechanical, Automotive & Materials Engineering, University of Windsor, Windsor, ON, Canada

c_{1t}	Memory cell of LSTM.B1
b_1	Bias of LSTM.B1
h_{1t}	Output of LSTM.B1
X_{2t}	Input of LSTM.B2
i_{2t}	Input Gate of LSTM.B2
f_{2t}	Forget Gate of LSTM.B2
o_{2t}	Output Gate of LSTM.B2
$\widetilde{c}_{2t}$	Memory cell candidate of LSTM.B2
c_{2t}	Memory cell of LSTM.B2
b_2	Bias of LSTM.B2
h_{2t}	Output of LSTM.B2
R	Order of Markov Chain
Pr	Closing Price
H	Length of Price List
$\boldsymbol{P}$	Probability Distribution
N	Number of Distinction Price
t	Epoch indices for LSTM 1 & LSTM 2
n	Number indices for Markov Chain

Introduction

The field of electrical load forecasting in smart grids is a vital discipline that focuses on predicting electricity consumption patterns to ensure the efficient functioning of power systems. With the advent of smart grid technologies, accurate forecasting methods have become indispensable for energy providers and grid operators. These forecasts, derived from historical data and various influencing factors, facilitate optimal resource allocation, grid stability, and the integration of renewable energy sources. Researchers employ diverse techniques, ranging from traditional statistical models to advanced machine learning algorithms and artificial intelligence, to develop robust forecasting models essential for effective energy management and sustainable development in the evolving landscape of smart grids.

Motivation beyond Development

Recent technological advances, industrialization, and the number of electrical vehicles (EVs) resulting from electric transportation have increased electrification worldwide. Therefore, the electricity suppliers faced considerable demand in different residential, commercial, and industrial sectors.

Recent Advances and Literature Review

Hence, the significance of renewable energy resources, intelligent energy hubs [1], and intelligent demand response management (IDRM) are gaining recognition. Artificial intelligence-driven models have become prevalent across diverse sectors, including the financial domain [2, 3], smart electrical energy management, and its economic aspect. In intelligent DSM techniques, artificial intelligence-based load estimators, such as short-, medium-, and long-term, are the considerable methodologies utilized in different applications and scopes.

Yang et al. [4] forecasted the multiple time scales of a year and a week. Normalized Dynamic Time Warping (N-DTW), Density-Based Spatial Clustering of Applications with Noise (DBSCAN) for one year, and behavior similarity, Mutual Information (MI), as well as Principal Component Analysis (PCA) are used to have feature selection, and the reduction in data multi-dimensionality of the forecasting in a week, respectively. The presented model also contributes to Back Propagation Neural Network (BPNN), LSTM, Co-LSTM, and Extreme Gradient Boosting (XGBoost) to derive the prediction by analyzing the data gathered. This combined deep-learning model forecasted with an average MAAPE of 31.20%. To draw a short-term prediction, the Markov corrector model of day-ahead building load forecasting is expressed in [5] to be used in supervision control and data acquisition self-updating automation of the building system to execute the forecasting during the COVID-19, under influences of Omicron.

Relying on the nonlinearity in consumer usage behavior, Saeed et al. [6] represented a cross-channel-communication (C3)-enabled CNN-LSTM system to predict the load usage, aiming to help the experts make better decisions in optimal planning of smart grids. The C3 and CNN are utilized to extract the features from the gathered data to be predicted by LSTM with Leaky ReLu characteristics. The model derived the forecasting under a performance of 0.4560% MAPE. Chaianong et al. [7] combined the traffic data with other features to predict load usage in residential grid areas by random forest modeling. To be utilized by the proper operation of grid and electricity trading in the proposed market, a machine learning model of gated recurrent unit (GRU) and RF is implemented in [8]. In this model, GRU is the power load estimator. In contrast, RF reduces the dimension size of the input going to be analyzed by GRU investigating the effect of weekends on electrical power load performed by the model. Shaqour et al. [9] reviewed and evaluated five deep-learning forecasters to see short-term load forecasting criteria results. A deep neural network (DNN) assumed by a Bi-directional Gated recurrent unit with fully connected layers (Bi-GRU-FCL) presented the highest performance in sequence. The characteristics of thermal load over correlation and principal component analysis are investigated in [10].

Moreover, CCHP active operation optimization weekend and weekday system is implemented in heating, transition, and cooling systems, as well as dynamic matching optimization and evaluation. The MAPEs of 5.43% and 6.84% are carried out for heating load forecasting and cooling load prediction, respectively. Veeramsetty et al. [11] combined RNN and PCA to forecast hourly load in an electric power load 33/11 kV

MV substation. This algorithm has a dimensionality reduction in the data input size being analyzed. The mixup and transfer learning-based model is defined in [12]. The mixup option enhances data and forecasting quality by training sample distribution expansion. Transfer learning is used to check for similar data patterns to avoid overfitting during the process. As the next step, the model utilizes the maximal information coefficient (MIC) to measure the similarity between target load and source load. As final, the LSTM model performs the final time series prediction process on the expanded load data.

The outcomes in [13] present a review of AI-based electrical load estimating methodologies as the processing steps and forecasting strategy model. The review is completed on the multi-variance and single-variance types of time series data and the one-step and rolling forecasting methodologies. The work discussed the different intelligent forecasting models, such as artificial intelligence, fuzzy reasoning, and transfer learning approaches, for the last part. An STLF method based on empirical mode decomposition (EMD), BiLSTM, and attention mechanism is integrated in [14]. Firstly, EMD decomposes the load data series to IMFs to be analyzed by BiLSTM, predicting the tendencies of each IMF. For the final phase, the results of all IMFs contributed and combined for the final forecasting decision of IMFs optimum numbers with 3 or 4.

The work in [15] demonstrated a model combined with fuzzy cluster (FC) analysis, least-squares support vector machine (LSSVM), and a fireworks algorithm (FWA). A reduction is performed in the data features dimension to be optimized by FWA to cover up the model FC–FWA–LSSVM. A data generation method depending on the generative adversarial network (GAN) is exposed in [16] to perform prediction in EV loads. This research used a gating mechanism as a Mogrifier with LSTM to enhance performance. Kandilogiannakis et al. [17] used a recurrent neuro-fuzzy system for STLF as ReNFuzz-LF. This model is based on dynamic small-scale RNNs with a single hidden layer, as the recurrent characteristics of the model lead to the usage of inputs with minimal sets. An ensemble LSTM method retrieved by multi-source transfer learning as MTE-LSTM is imposed in [18]. Firstly, the model found similar buildings to the target building to be used for the LSTM estimator tuned by transfer learning, and fine-tune technology. Finally, the ensemble model weighted the output results dedicated to the forecasting results.

Tong et al. [19] derived a model by temporal inception convolutional network of multi-head attention (TICN-Att) to have ultra STLF. A Meta-learning tuned LSTM is in [20] to perform prediction by the LSTM model for the historical data and the meta-learning for the nonstationary load patterns of the grid. Using gradient descent algorithms, this model is trained by optimizing the base and error correction modules. Ullah et al. [21] intended a PLF methodology based on intelligent learning, in which the refined data pass through Conv.LSTM to generate feature maps. The feature maps are transformed into the deep GRU for providing the final PLF through the learning process.

The ensemble deep Random Vector Functional Link (edRVFL) network is utilized in [22] to provide STLF results. The edRVFL generates the forecasting of outputs learning ensembles, and the raw data is decomposed by empirical wavelet transformation (EWT), feeding edRVFL to perform the final prediction. Abdolrezaei et al. [23] presented a knowledge-based methodology forecasting mid-term load forecasting (MTLF) based on preprocessing the data to refine them and perform the forecasting utilized by the linear equations estimator. The hybrid model, including SVM, BPNN, GRNN, and the genetic algorithm, is expressed in [24] to derive forecasting of individual residential loads. In this model, the genetic algorithm optimizes BPNN and SVM for enhancement. The DEM part of the model consists of BiLSTM networks optimized by the Bayesian algorithm (BA). Overall, the model separated the HAC and non-HAC load types through the RC model that predicts the indoor temperature. After that, the non-HAC load is divided into electric lighting and other loads. The DEM captures these loads, and the final prediction is performed by combining HAC and non-HAC separate results.

A smart grid consisting of a photovoltaic, wind turbine, battery energy storage system, electric vehicle charging stations, and intelligent estimator models of LSTM, group method of data handling, and adaptive neuro-fuzzy inference system are reviewed and analyzed in [25] considering the hardware requirements, and noisy conditions of the systems. As a result of the work, LSTM performed better accuracy than the other developed models. Azeem et al. [26] investigated the performance of three intelligent models of ARIMA, ANN, and LSTM, considering the input parameter change of real-world datasets of smart grids. The mentioned models have a 5%-15% change in accuracy by parameter varying. An adaptive framework to improve forecasting accuracy and quality was also introduced. A multi-task model as MultiDeT (Multiple-Decoder Transformer) is indicated in [27].

MultiDeT is adoptable of a one-encoder multiple decoder structure, in which all input data is encoded by the uniform encoder to be decoded by the multi-decoders. In sequence, the trained end-to-end with losses of each task is used to present the final forecasting result. In [28], the work presented a kernel-based Gaussian and Bayesian regression model to predict the day head residential load STLF analysis. The model is used in multivariate large datasets during winter. The Bayesian mixture density networks are also performed in [29] for STFL analysis. Given the importance of STLF [30], four models of STLF analysis by MLPs, K-means, and FC clustering algorithms are introduced. The first two models compare the data by the similarly developed modeling fed to MLPs. The other two models enhance the input

quality of MLPs to have better analysis, leading to a more accurate STLF. Moon et al. [31] expressed the Cubic forecasting model of daily peak load forecasting (DPLF) and total DPLF. The cubic learning-based model presented in this work performed with a MAPE of 10.06%. Focusing on decarbonization and zero greenhouse gas emissions, Vargas-Salgado et al. [32] described the utilization of renewable energy sources (RES), including PVs, and wind, considering their contribution to pumping storage and mega-batteries, as storage technologies to manage the variability. Forecasting the optimum demand response of Grand Canary Island, 3700 MW of PV, around 700 MW of offshore wind system, 607 MW of pump storage, and 2300 MW of EV battery capacity are considered as the maximum ability of demand as the optimal planning of demand management.

Dinh et al. [33] introduced a home energy management system (HEMS) considering demand response and customer usage behavior to control the energy storage system (EES) and renewable energy system (RES) to increase energy efficiency while lowering daily costs. This system ran under DNNs and MILP supervised learning strategy. MILPs perform the STFL analysis as a resource for DNNs to control and optimize EES and RES regarding the real-time environmental surroundings. Stanelyte et al. [34] reviewed DR services, including recent methodologies, IoT applications in adaptive monitoring of power electrical loads, concepts, and classifications of this field. Khan et al. [35] introduced a two-phase framework to perform STLF. Firstly, it preprocesses the data using deep residual CNN applied to the raw data. In the next phase, the featured datasets extracted by a deep residual CNN are analyzed by LSTM, learning the temporal information of the electricity data, leading to an RMSPE of 14.85%. Industrial demand response management techniques tuned by AI, and NN in smart sustainable cities (SSCs), and combined heat, and power (CHP) incorporated smart grids are reviewed in [36].

Wavelet transform-based ensemble forecasting model analyzing STLF is approached in [37], by which decomposing profile principles capture the portion of load daily profiles originated from variations. Safari and Ghavifekr [38, 39] defined and formulated quantum neural network (QNN)-based intelligent forecasters in smart grids and weather prediction of smart grids, respectively. Considering the optimal model selection, [40] presented a multi-space collaboration (MSC) framework and SVR model adopting the space separation strategy to perform the model selection on the subspace. Yazici et al. [41] conducted the one-dimensional CNN, LSTM, and GRU variants applying to the real-world simulations. They obtained the highest performance of one-dimensional CNN with a MAPE of 2.21% compared to the other variants.

Concerning the commercial buildings, a DNN model and LSTM-RNN are decomposed in [42]. Furthermore, Safari and Sabahi [43] presented a practical system of industrial data transfer structure that improves the data transfer rate and data sampling while lowering the data transfer time, which can be considered in the analysis of smart grid electricity load data. In the competitive electric energy market, accurately predicting electricity prices is crucial for effective planning and operations due to the unpredictability influenced by various factors, especially with the increasing use of wind energy. [44] have tackled this challenge by enhancing the Elman neural network with an improved Gorilla Troops Optimizer. This method optimizes wavelet decomposition and neural network architecture, providing efficient short-term electric power price predictions. Numerical testing using historical data from Chinese spot marketplaces demonstrated promising results, outperforming contemporary techniques.

A new intrusion detection method for the Internet of Vehicles, based on the Apache Spark framework [45], combines deep learning techniques (CNN and LSTM) to extract features and detect abnormal behaviors in large-scale car network data traffic. This approach achieves remarkable results, with a detection time of 20 units and an accuracy rate of 99.7%, outperforming existing models. Bahmanyar. et al. [45] present a Multi-Objective Arithmetic Optimization Algorithm (MOAOA) for a Home Energy Management System (HEMS) using IoT technology. HEMS optimizes appliance scheduling to reduce electricity costs, lower peak-to-average ratio, and enhance user comfort. The system outperforms other algorithms implemented on Raspberry Pi, demonstrating significant cost reductions and improved user comfort, especially when integrated with renewable energy sources.

An exergy assessment methodology for a power production system featuring a high-temperature proton exchange membrane fuel cell and an organic Rankine cycle for heat recovery is anticipated in [46]. The system variables are optimized to create an optimum balanced model. A new metaheuristic approach, the Fractional-order Coyote Optimization Algorithm, is applied to enhance the precision and accuracy of the results. Three cost functions are optimized: irreversibility, work, and exergy. The proposed method's simulation results are incorporated into a case study, and validation is performed by comparing them with experimental data, the original COA, and the Genetic Algorithm (GA) from existing literature. The final findings indicate that the experimental data provides the highest level of confirmation when employing the proposed algorithm.

Proposed Model

Despite its complexity, the FARHAN model's interpretability provides valuable insights into the factors steering electrical load forecasts. Experts can discern which input variables exert the most substantial influence on predictions by employing techniques such as feature importance analysis, offering a clear understanding of the critical factors shaping load forecasts. Furthermore, FARHAN's attention

mechanism enables experts to observe how the model weighs the importance of different elements within the input data, shedding light on the temporal patterns and correlations driving load predictions. This ability to discern the significance of specific features and time steps empowers experts to uncover the underlying dynamics of electrical load behavior, enhancing their capacity to make informed decisions and optimizations within smart grids and energy systems.

Integrating external factors, such as weather data or economic indicators, into the FARHAN model holds significant promise for augmenting its predictive capabilities in load forecasting in future works. Weather patterns, for instance, influence electricity consumption; incorporating meteorological data such as temperature, humidity, and precipitation can enable FARHAN to capture seasonal variations and sudden spikes in energy demand, thereby refining its predictions. Similarly, economic indicators such as GDP growth, industrial output, or employment rates can provide valuable context, especially in forecasting long-term load trends. By amalgamating these external factors with the existing dataset, FARHAN can discern intricate correlations, enhancing its ability to anticipate load variations during specific weather conditions or economic fluctuations. This integration not only strengthens the model's accuracy but also equips utility providers and grid operators with a more comprehensive understanding of the multifaceted factors shaping electricity demand, fostering smarter and more proactive decision-making in the management of energy resources.

Overall, the model performed in this work achieved considerably better performance with the MAPE and RMSPE of less than 0.02% and 2.5%, respectively. Furthermore, FARHAN has achieved R^2 of 1. Additionally, the contributions of the paper can be listed as:

- **Introduction of FARHAN:** FARHAN is a cutting-edge hybrid model developed for electrical load forecasting in smart grids, integrating descending neuron attention, LSTM, and Markov-simulated neural networks.
- **Enhanced Accuracy and Efficiency:** FARHAN overcomes challenges in accuracy and analysis time, crucial for optimal short-, mid-, and long-term smart grid planning, outperforming traditional LSTM models and other methodologies.
- **Model Components:** FARHAN comprises two LSTM blocks (LSTM.B1 & LSTM.B2) with attention layers, a 90% gain averager, and a Markov chain analyzer, ensuring comprehensive processing of electricity load data.
- **Comparative Performance:** Comparative analysis demonstrates FARHAN's superiority, with impressive MAPEs of 0.019162%, 0.0386%, and 0.039% for 14 years, annual and monthly estimations, respectively.
- **Validation and Accuracy:** FARHAN achieves remarkable RMSPEs of 2.5%, 5.2%, and 1.2%, along with an overall R^2 of 1, validating its exceptional accuracy and reliability in load forecasting.
- **Significance:** FARHAN's innovative approach establishes it as a robust and intelligent tool, promising significant advancements in electrical load forecasting within smart grids and energy systems.

The remainder of this paper is organized as follows:

Configuration principles and the operating framework of FARHAN are expressed in Sections "Configuration Principles" and "Operating Framework", respectively. Sections "Experiments & Results", "Comparison", and "Future Works" present the experimental results, comparative analysis, and future works. Finally, conclusions are drawn in Section "Conclusion".

Configuration Principles

In electrical load forecasting for smart grids, two fundamental configuration principles are paramount. Firstly, data integration and feature selection are essential. This involves integrating diverse data sources, such as historical consumption data and patterns while selecting relevant features that significantly impact electricity demand. Secondly, fine-tuning the model architecture and hyperparameters is crucial. Researchers experiment with different algorithms and neural network structures, adjusting hyperparameters to optimize the model's performance. These principles ensure that forecasting models are accurate, efficient, and capable of handling the complexities of smart grid data.

Neural Networks LSTM

Inspired by the human brain's neuron systems, artificial neural networks find applications in various fields. ANNs learn through a trial-and-error process to optimize their weight initialization, primarily excelling in forecasting tasks like electricity load prediction, stock market indices, and customer behavior analysis. They comprise three core components: the input layer, the processing layer (hidden, dense, and attention layers with varying sublayers and neurons), and the output layer.

Three central neural network types include ANNs, CNNs, and RNNs, with other models stemming from them. RNNs, a type of neural network, enable nodes to have feedback and cyclic processes, impacting subsequent inputs based on previous node outputs. LSTM neural networks, a variation of RNNs, possess the ability to remember or forget data for future processing phases. LSTMs excel in analyzing long-term dependencies, making them suitable for tasks involving time series analysis, model predictive control, and adaptive control systems. An overview of intelligent models is illustrated in Fig. 1, while Fig. 2 depicts the LSTM network's structure.

Markov Neural Network

In a probabilistic approach with feedback, the Markov neural network serves as an analytical system for assessing state transitions. It computes the probabilities of events occurring based on the previous state achieved. Specifically, in complex high-dimensional probability distributions, it systematically evaluates how much the next state depends on the current state. Subsequently, it conducts a conclusive analysis within the FARHAN framework, determining the predictive accuracy of future steps based on the predictive precision of the current state, which is influenced by the predictive accuracy of the preceding state. Markov NN is modeled and formulated as (3) in FARHAN:

$$P(X_n| = i_n ||X_n - 1 = |i_n - 1|) =|$$
$$P(X_n| = i_n ||X_0| = i_0|, X_1| = i_1|, \ldots, X_n - 1 = i_n - 1|) \quad (1)$$

where $P, X_n X_n$, and $i_n i_n$ present probability, random variables, and possible state, respectively. The developed model in this work implements Markov NN to improve future prediction phases regarding the current and previous forecasted values.

Operating Framework

This work's highly accurate predictive model stands for the Descending Neuron Coupled LSTM Averaged Markov Simulated Neural Network. FARHAN comprises two LSTM block, an avergizer, and a Markov Neural Network analyzer. The overall progress of the system workflow is shown in Fig. 3.

Based on Fig. 3, the system performs the prediction in eight steps. The user imports the electricity demand data, the DPU processes the data, and then it is reshaped by the feature set to

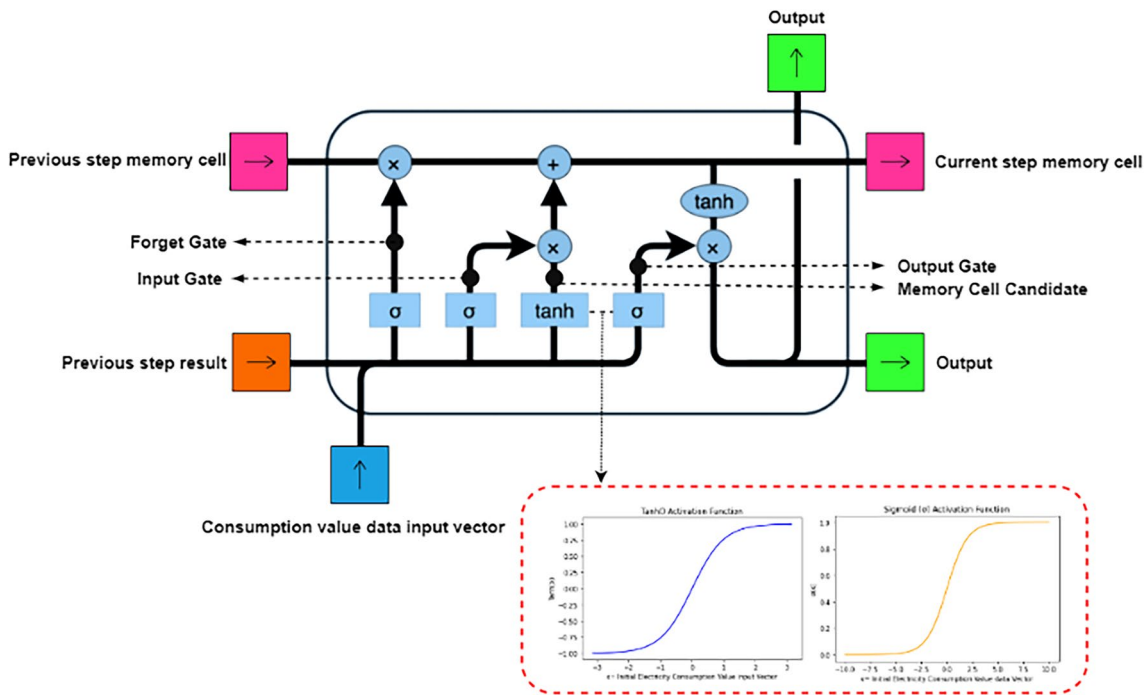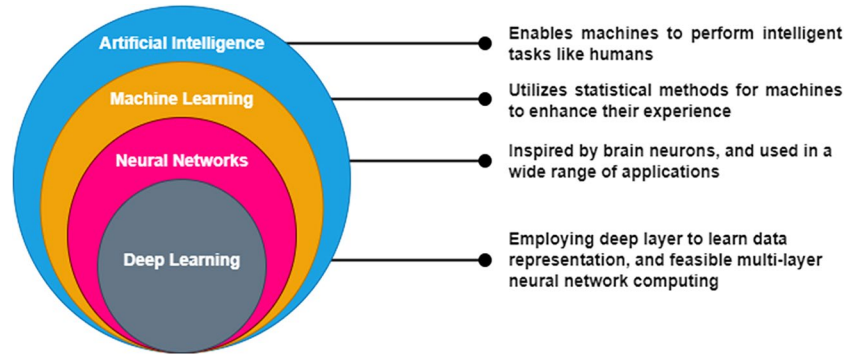**Fig. 1** The overall classification of AI models

Fig. 2 The structure of the LSTM network

be in the proper format analyzed by the intelligent analyzers of the model. The featured and reshaped data transferred to the first LSTM block. LSTM. B1 consists of two hidden layers with 100 neurons and a dense layer of 50 neurons. The output of LSTM. B1 is averaged by the averagizer with the 0.9 gain, and the new dataset is set up to be analyzed by the LSTM. B2. The gain of the averagizer is determined by system identification in which the system is in its best performance and draws -the most near-to-real prediction. As the next step, LSTM. B2 evaluates the new dataset resulting from the previous phase. Two hidden layers of 50 neurons and a dense layer with 25 neurons cover up the LSTM. B2 network structure. As the final evaluation, the results of LSTM. The Markov chain intelligence model analyzes B2 to perform the final decision and draw the most accurate forecasting to the user.

From Figs. 3 and 4, the LSTM neural network has three inputs of the previous step memory cell, results, and Therefore, output, and the current step memory cell set up the results of the network. The network utilized three types of forget, input, and output gates, as well as a memory cell candidate and two types of activation functions, $Tanh()$ and $\sigma()$ as formulated below:

$$Tanh(Con.V) = \frac{e^{Con.V} - e^{-Con.V}}{e^{Con.V} + e^{-Con.V}} \tag{2}$$

$$\sigma(Con.V) = \frac{e^{Con.V}}{1 + e^{Con.V}} \tag{3}$$

where $Con.V$ is the consumption value data input vector. The structure of FARHAN and the averagizer and Markov chain neural networks are described in Fig. 4. The neural modeling of FARHAN is also demonstrated in Fig. 5.

$Con.V$ From Figs. 4 and 5, the system starts to initialize the weights/bias of the LSTM. B1, including the neurons of input layer-first hidden layer (LSTM.B1.H1), first hidden layer-second hidden layer (LSTM.B1.H2), LSTM.B1.H2-dense layer, as well as the dense layer-output layer.

This process continues due to the time that the prediction tracks the real observed value with the least customized error rate. As the next, the output of LSTM.B1 averages by the 90% averagizer, and the new dataset moves as the input of LSTM. B2. The same weights/bias done by the LSTM.B2 to track the dataset imported from avergizer. For the final phase, the Markov chain performs the final analysis to illustrate the most accurate forecasting results to the user. The detailed structures of LSTM analyzers are represented in Table 1.

As shown in Fig. 3(C) and Eq. (1), R, Vc, and H are symbolized as Markov order, electricity consumption value as [kwh], and the length of the electricity consumption list. Moreover, P presents the probability distribution, and the distinct consumption values are noted by N, which predicts the further forecasting state prior to the results and data available in the current state.

The LSTM blocks and the averagizer of FARHAN are modeled and formulated as [47]:

$$i_{1t} = \sigma(W_{1i}h_{1t-1} + U_{1i}[Con.V]_{1t} + b_{1i}) \tag{4}$$

$$f_{1t} = \sigma(W_{1f}h_{1t-1} + U_{1f}[Con.V]_{1t} + b_{1f}) \tag{5}$$

$$O_{1t} = \sigma(W_{1o}h_{1t-1} + U_{1o}[Con.V]_{1t} + b_{1o}) \tag{6}$$

$$\widetilde{C}_{1t} = \sigma(W_1 h_{1t-1} + U_1[Con.V]_{1t} + b_1) \tag{7}$$

$$C_{1t} = (f_{1t} \odot C_{1t-1}) + (i_{1t} \odot \widetilde{C}_{1t}) \tag{8}$$

where $Con.V$, i_{1t}, and f_{1t} compensating the input consumption value vector, input gate, and the forget gate, respectively. O_{1t}, and $\widetilde{C}_{1t}$ are utilized as the output gate and memory cell candidate of the LSTM. B1 network. As the output vector of the LSTM.B1 h_{1t} is defined as the output of the network by:

$$h_{1t} = O_t \odot tanh(C_{1t}) \tag{9}$$

For the next step h_{1t} is averaged with $[Con.V]_{1t}$ by the predetermined gain as (8):

$$[Con.V]_{2t} = \left(\frac{h_{1t} + [Con.V]_{1t}}{2}\right) \times 0.9 \tag{10}$$
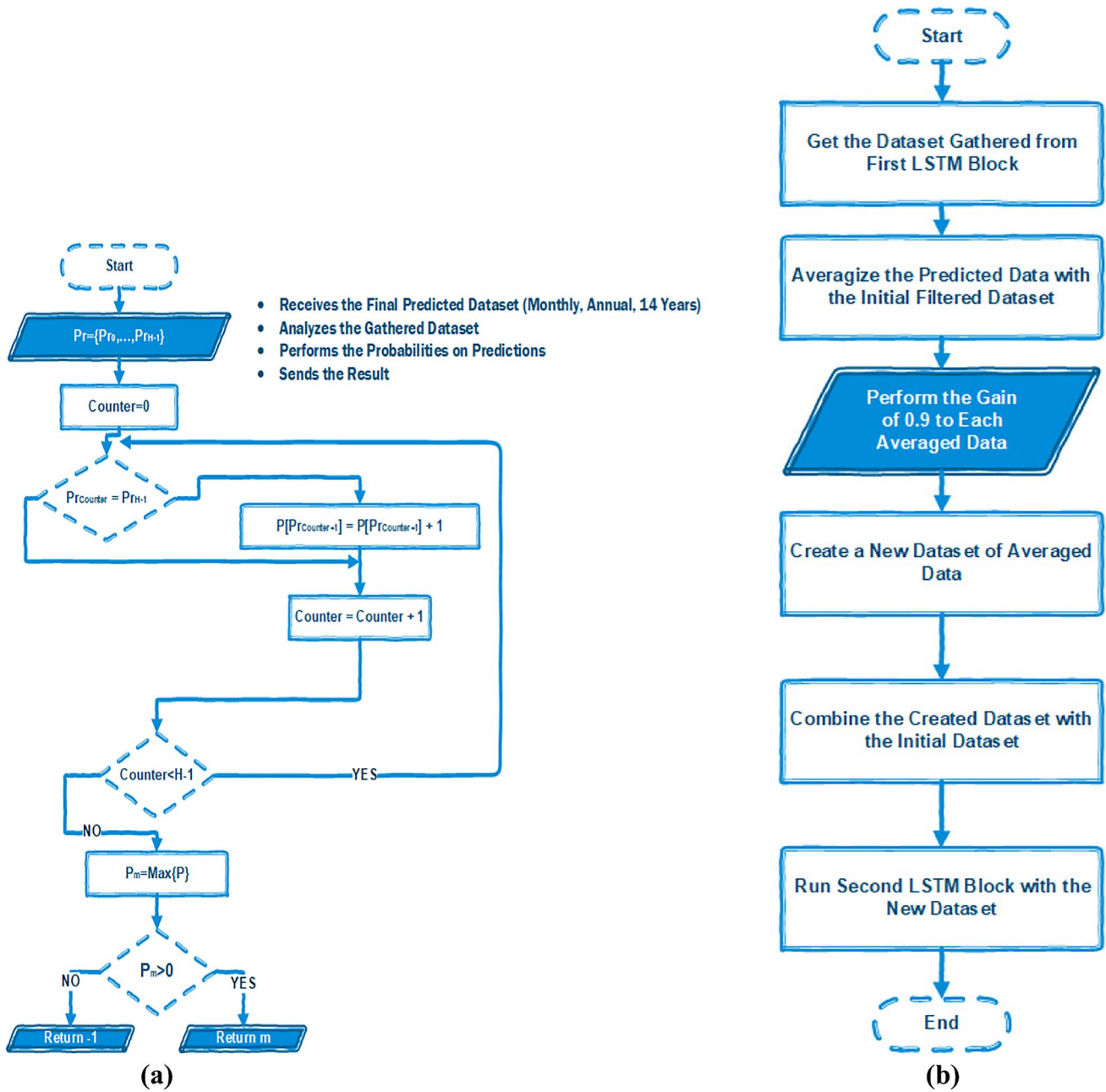
Fig. 3 The overall process performed by FARHAN
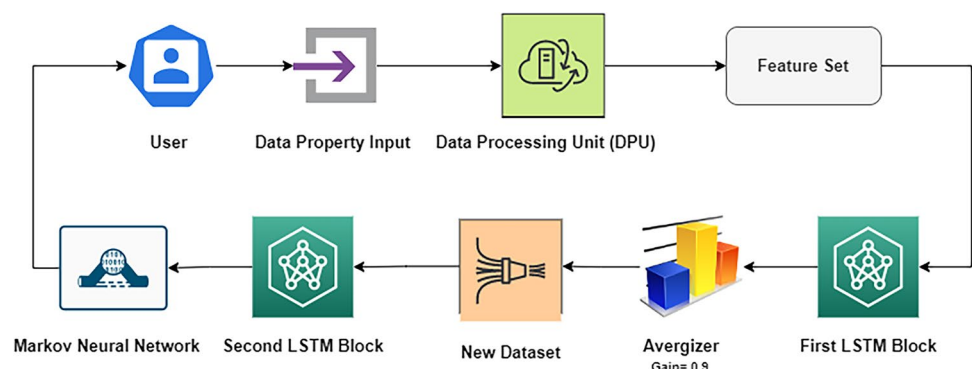
Fig. 4 The process of (**a**) MCNN and (**b**) Averagizer

where the averaged new dataset, as the input vector of LSTM.B2, is denoted by $[Con.V]_{2t}$. Conceptually, the results of the LSTM.B1, and avergizer are assumed as the input dataset of LSTM.B2, leading to a new dataset of $[Con.V]_{2t}$.
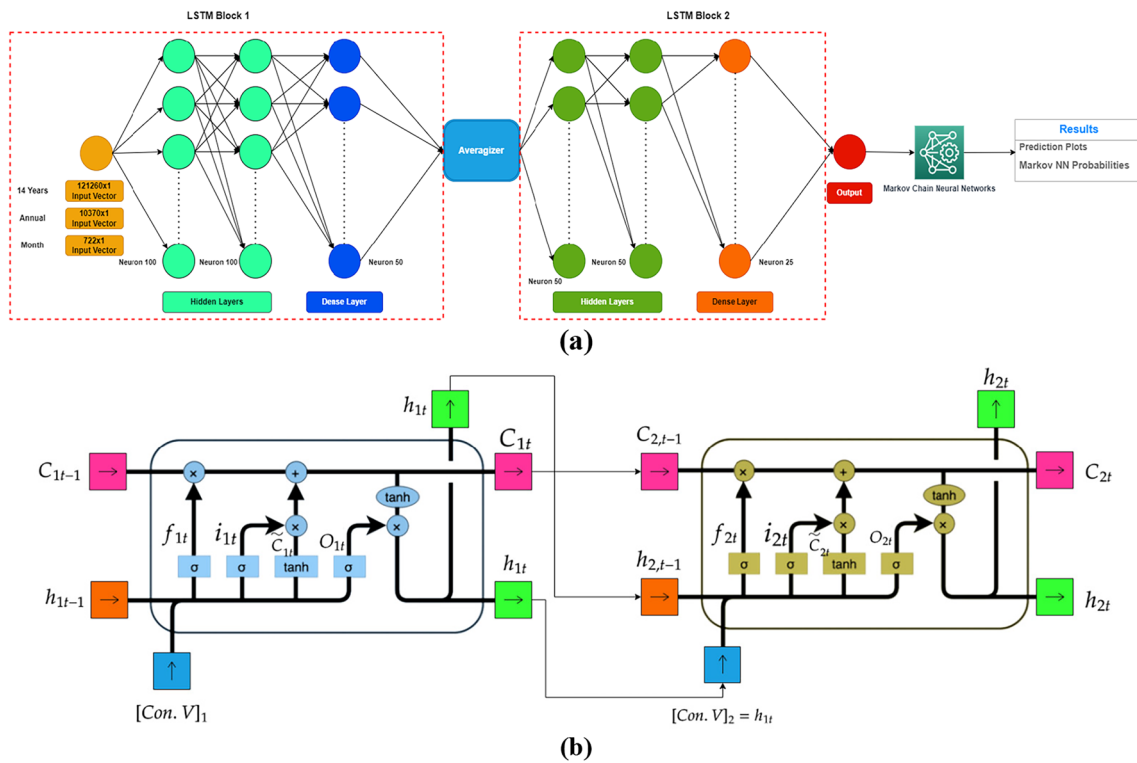
$$i_{2t} = \sigma(W_{2i}h_{2t-1} + U_{2i}((\frac{h_{1t} + [Con.V]_{1t}}{2}) \times 0.9) + b_{2i}) \quad (11)$$

$$f_{2t} = \sigma(W_{2f}h_{2t-1} + U_{2f}((\frac{h_{1t} + [Con.V]_{1t}}{2}) \times 0.9) + b_{2f}) \quad (12)$$

$$O_{2t} = \sigma(W_{2o}h_{2t-1} + U_{2o}((\frac{h_{1t} + [Con.V]_{1t}}{2}) \times 0.9) + b_{2o}) \quad (13)$$

$$\widetilde{C}_{2t} = \sigma(W_2 h_{2t-1} + U_2((\frac{h_{1t} + [Con.V]_{1t}}{2}) \times 0.9) + b_2) \quad (14)$$

$$C_{2t} = (f_{2t} \odot C_{2t-1}) + (i_{2t} \odot \widetilde{C}_{2t}) \quad (15)$$

$$h_{2t} = O_{2t} \odot tanh(C_{2t}) \quad (16)$$

Fig. 5 **a** Neural, and **b** cell modeling of FARHAN

Table 1 The properties of the LSTM block in FARHAN

LSTM Block	Properties – 31,901 Trainable Parameters							
	Input		Hidden		Dense		Output	
	Layers	Neurons	Layers	Neurons	Layers	Neurons	Layers	Neurons
LSTM. B1	1	2	2	50	1	25	1	1
LSTM. B2	1	2	2	100	1	50	1	1
Weights LSTM. B1	-		10,400 | 32.6%		1275 | 4.0%		-	
Weights LSTM. B2	-		20,200 | 63.3%		26 | 0.1%		-	

where $Con.V$, i_{2t}, and f_{2t} compensating the input consumption value vector, input gate, and the forget gate, respectively. O_{2t}, $\widetilde{C}_{2t}$, and h_{1t} are utilized as the output gate, memory cell candidate, and output vector of the LSTM. B2 network, respectively. FARHAN utilizes forget, and memory abilities to increase forecasting efficiency. Forget gate evaluates how much of the data should be omitted and the amount of data stored in memory cells for the subsequent phases. Forget gate is based on the Boolean logic, in which 0 means forgetting, and 1 means memorizing the data. PSUEDOCODE of FARHAN is presented in Algorithm 1.

Algorithm 1   Pseudocode of FARHAN

```
Input: [Con. V]₁ = [(Con. V)₁₁, (Con. V)₁₂, ..., (Con. V)₁Testset_Length]
Given Parameters: W₁f, U₁f, b₁f, W₁c̄, U₁c̄, b₁c̄, W₁i, U₁i, b₁i, W₁o, U₁o, b₁o
Initialize W₁o, C₁o = 0⃗
for t=1,..., Testset_Length do
            Calculate f₁t, C̄₁t, i₁t
            Update Cell State C₁t
            Calculate O₁t, □₁t
end for
Output: h₁ = [□₁₁, □₁₂, ..., □₁Testset_Length], □₁i ∈ Rⁿ
for t=1,..., Test_Length do
            Calculate [Con. V]₂ = h₁[t]×(Con.V)₁[t]/2 (0.9)
end for


Input: [Con. V]₂ = [(Con. V)₂₁, (Con. V)₂₂, ..., (Con. V)₂Testset_Length]
Given Parameters: W₂f, U₂f, b₂f, W₂c̄, U₂c̄, b₂c̄, W₂i, U₂i, b₂i, W₂o, U₂o, b₂o
Initialize W₂o, C₂o = 0⃗
for t=1,..., Testset_Length do
            Calculate f₂t, C̄₂t, i₂t
            Update Cell State C₂t
            Calculate O₂t, □₂t
end for
Output: h₂ = [□₂₁, □₂₂, ..., □₂Testset_Length], □₂i ∈ Rⁿ
for Counter := 0 to R-1 do
            C[Counter]:= Pr[H-R+Counter]
Endfor

for h := R to H-1 do
            IS_CONTEXT:= TRUE
            for Counter := 0 to R-1 do
                        if E[h-R+Coounter] != C[Counter] then
                        IS_CONTEXT := false
                        break
                        end if
            end for

   if IS_CONTEXT then
            P[Pr[h]] := P[Pr[h]] + 1
   end if
end for

PREDICTION := 0
MAX := P[0]
for i := 1 to N-1 do
            if P[i] > MAX then
                        MAX := P[i]
                        PREDICTION := i
            endif
end for

if MAX>0 then
            return PREDICTION
end if
return -1
end
```

Experiments & Results

In order to use and evaluate the performance of FARHAN, a consumption dataset including 121,260 data of power consumption, from 12/31/2004, 1:00:00 AM to 1/2/2018, 12:00:00 AM, is provided. In this process, the model performs prediction in three different ranges: monthly, annual, and for 14 years. Sincerely, the model analyzed 722, 10,370, and 121,260 data in the monthly, annual, and 14-year range,

respectively. The detailed properties of the utilized dataset, the number of trained and tested data in each range, are expressed in Table 2.

The data/time step window of FARHAN is drawn in Fig. 6.

Figure 6 shows the gathered electricity consumption demand data imported to the model by the LSTM. B1, the dataset is split into the train set, from 12/31/2004 1:00 AM to 10/26/2006 4:00 AM, and the test set, from 10/26/2006 4:00 AM to 01/02/2018 12:00 AM. The LSTM.B1 analyzes the data and sends the resulting dataset to the averagizer to be averaged with a gain of 0.9. The analysis in LSTM.B1 was performed by two hidden layers of 100 neurons each and a dense layer of 50 neurons. The new averaged dataset was imported to the LSTM. B2 to perform the final technical forecasting with two hidden layers of 50 neurons each and a 25 neurons dense layer. As the final step, the output of LSTM.B2 was analyzed by the Markov model to present the most near-to-value prediction. The results of the model are illustrated in Fig. 7.

From Fig. 7, the data index between 12/31/2004, 1:00:00 AM—1/2/2018, 12:00:00 AM (a), 1/2/2017, 1:00:00 AM—1/2/2018, 1:00:00 AM (b), and 2/1/2018, 12:00:00 AM—1/2/2018, 12:00:00 AM (c) are analyzed by FARHAN for 14 years, a year, and a month, respectively. The gap between the value graph and the predictions graph can also be considered as the forecasting error of FARHAN. The model presents a noticeable forecasting performance with considerable improvements in the least prediction error. Based on the system's progress, the resulting dataset of LSTM.B2 imported to Markov chain neural networks that the Markov parameters yield, and described in Table 3, below regarding the analyzed dataset.

Accordingly, the Markov model exports the following states:

1- If FARHAN predicts an upside in price, its probability will be 0.586326
2- If FARHAN predicts a downside in price, it will occur, with the probability of 0.557607

In addition, the mean and variance of the dataset forecasted by FARHAN and the observed ones are noted in Table 4, based on (15–19).

$$\overline{[Con.V]} = \frac{\sum [Con.V]_i}{n} \tag{17}$$

$$S^2_{observed} = \frac{\sum ([Con.V]_i - \overline{[Con.V]})^2}{n-1} \tag{18}$$

$$\overline{h_{2,t}} = \frac{\sum h_{2,t}}{n_{predicted}} \tag{19}$$

Table 2 The details of the dataset used in the experiment

Total Number of Data			
14 Years	121,260		
A Year	10,370		
A Month	722		
Date			
Initial Date	Month	Year	14 Years
Final Date	2/1/2018, 12:00:00 AM—1/2/2018, 12:00:00 AM	1/2/2017, 1:00:00 AM—1/2/2018, 1:00:00 AM	12/31/2004, 1:00:00 AM—1/2/2018, 12:00:00 AM
Duration	31 Days	365 Days	5110 Days
Train & Test Properties			
Number of Trained Data	Month	Year	14 Years
	595	5204	108,707
Number of Tested Data	127	5166	12,553

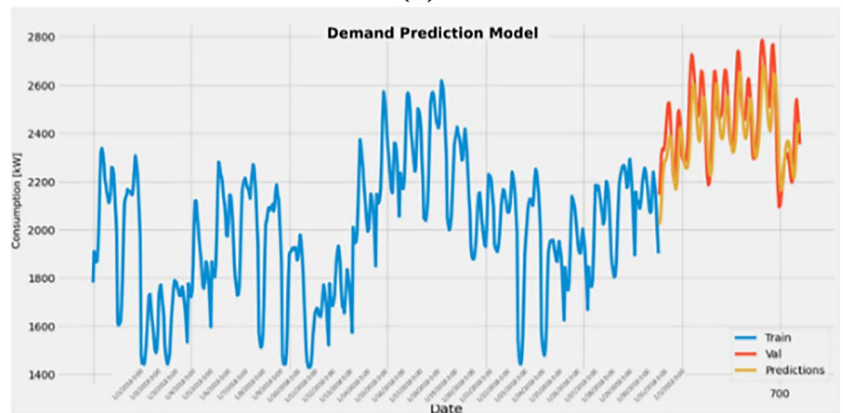**Fig. 6** The data/time step performed by the model
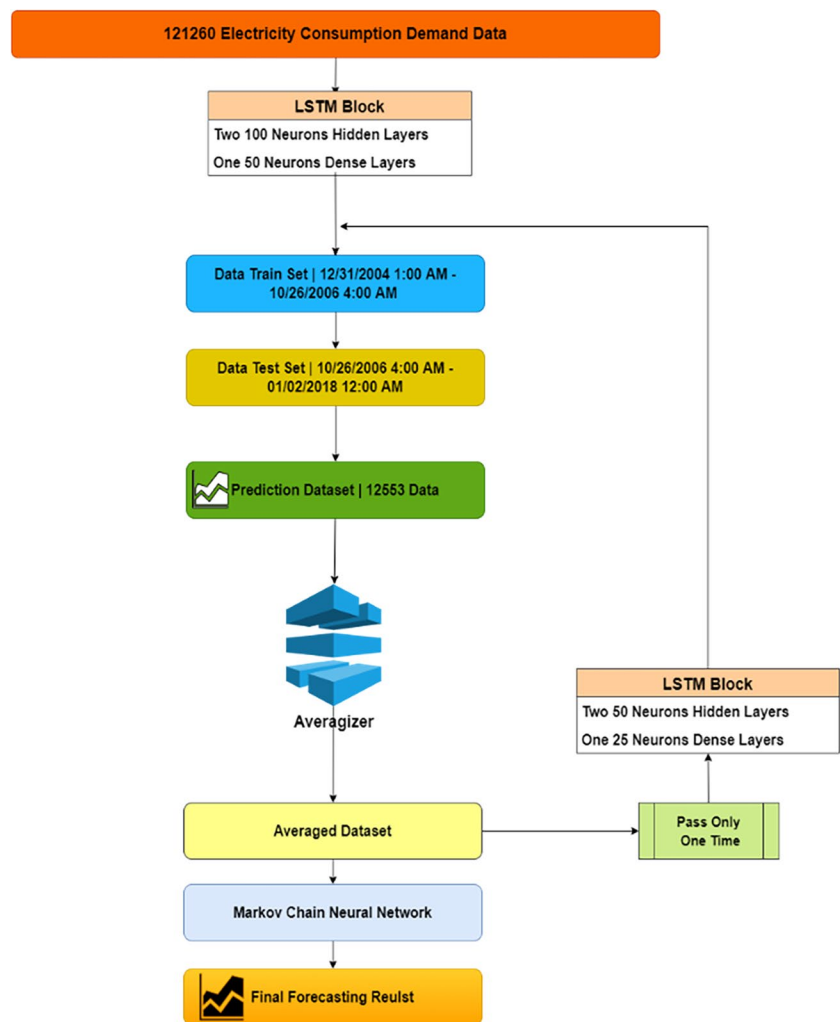
Fig. 7 The (**a**) 14 years, (**b**) Annual, and (**c**) Monthly load forecasting analysis
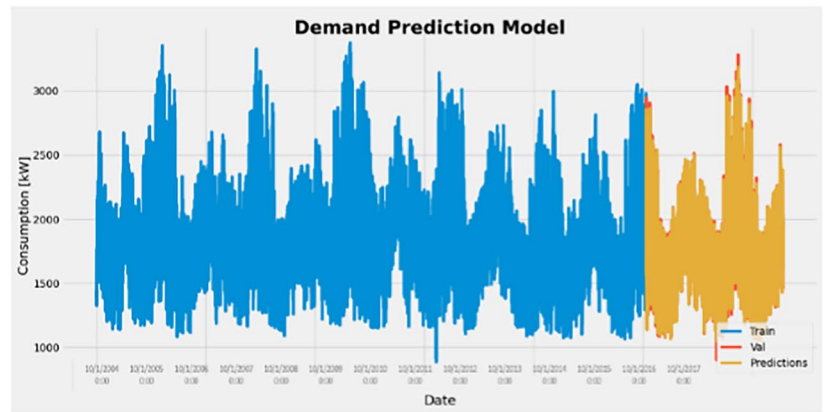
(a)
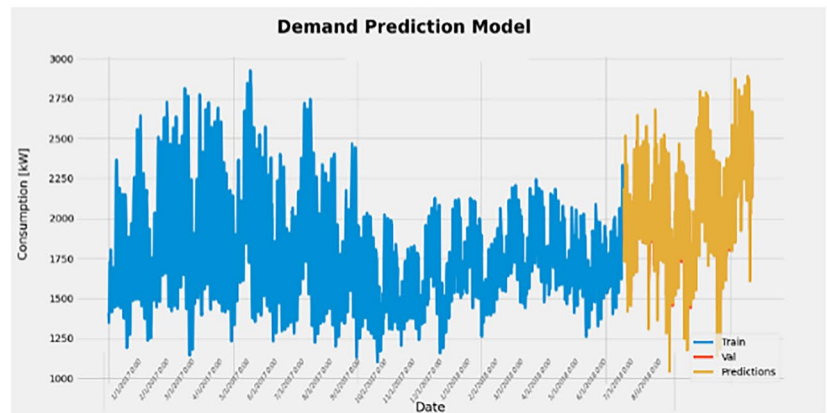
(b)

(c)

Table 3 Results of the Markov Neural Network Model

Prior Sate	State	
	Downside	Upside
Downside	0.586326	0.413674
Upside	0.442393	0.557607

Table 4 Statistical Results of FARHAN

Duration	$\overline{h_{2,t}}$	$\overline{[Con.V]}$	$S^2_{predicted}$	$S^2_{observed}$
14 Years	1779.841	1783.634	115,144.9	127,534
1 Year	2288.992	2285.486	954,911.63	92,879.41
1 Month	2797.302	2658.236	37,525.21	35,860.39

$$S^2_{predicted} = \frac{\sum(h_{2,t} - \overline{h_2})^2}{n_{predicted} - 1} \qquad (20)$$

where, $\overline{[Con.V]}$, $S^2_{observed}$, and n present the mean, sample variance, the size of the dataset, respectively. In the other hand, the mean, sample variance, and the prediction dataset size are denoted by $\overline{h_{2,t}}$, $S^2_{predicted}$, and $n_{predicted}$.

In order to evaluate the performance of the developed model in this work, MAPE (%), RMSPE (%), and R^2 KPIs are defined, formulated, and used as forecasting performance indicators during the process. Mean absolute percentage error is a predictive model's key statistical error calculation over its forecasting regarding the observable values. RMSPE is among the forecasting KPIs that standardly deviate the residuals as the prediction errors. MAPE, RMSPE, and R^2 in FARHAN are utilized as [48]:

$$MAPE = \frac{1}{n} \sum_{i=1}^{n} |\frac{[Con.V]_{i,actual} - h_{i,predicted}}{[Con.V]_{i,actual}}| \qquad (21)$$
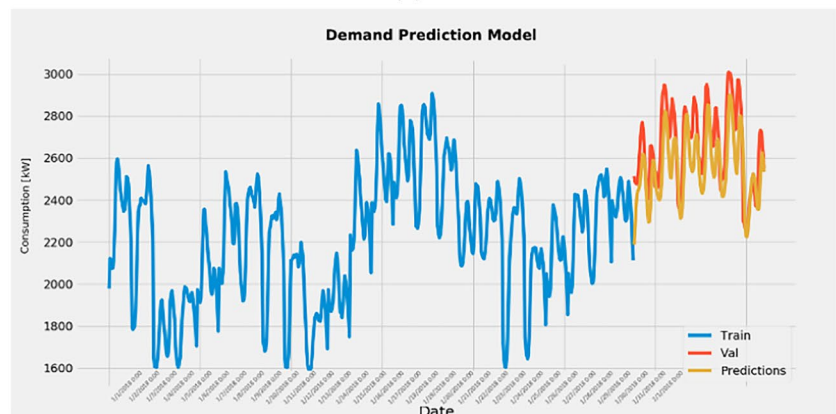
Fig. 8 The (**a**) 14 years, (**b**) Annual, and (**c**) Monthly load forecasting analysis by LSTM
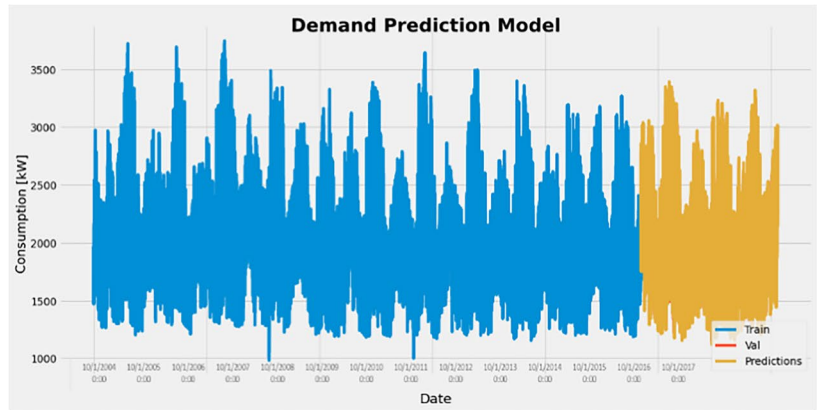
(a)
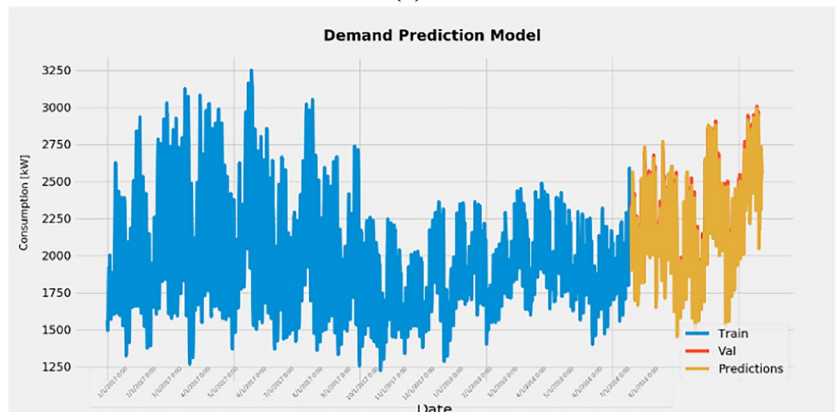
(b)

(c)

$$RMSPE = \sqrt{\frac{\sum_{i=1}^{n}([Con.V]_{i,actual} - h_{i,predicted})^2}{n \times [Con.V]_{i,actual}}} \qquad (22)$$

The R^2, also known as the coefficient of determination, is a statistical measure representing the proportion of the variance in the dependent variable (target variable) that is predictable from the independent variables (features) in a regression model. It ranges from 0 to 1, where 0 indicates that the model does not explain the variability of the target variable, and 1 indicates that the model perfectly predicts the target variable, as defined below:

$$R^2 = 1 - \left(\frac{\sum([Con.V]_{iObserved} - [Con.V]_{iPredicted})^2}{\sum([Con.V]_{iObserved} - [Con.V]_{iMean})^2} \right) \qquad (23)$$

where $[Con.V]_{i,actual}$, and $h_{i,predicted}$ are symbolized as the predicted, and observed consumption values, in [kW], respectively. Additionally, n denotes the number of imported data. From (21) and (22), the forecasting accuracy of the model depends on the parameters of n and the difference between the predicted and observed consumption values. As a result, the MAPE and RMSPE of FARHAN tend to zero, and in their least value, that demonstrates the forecasting accuracy of the model. Therefore, during the experiments, MAPE (%) and RMSPE (%) are calculated as 0.019162 and 2.5, respectively. As the mentioned KPIs indicate, the observed values are nearly fully tracked by the forecasted values performed by the presented model (Fig. 7). Additionally, an LSTM model analyzes the same dataset, and the results are presented in Fig. 8.

From Fig. 8, the observed values of the same dataset are not fully tracked by LSTM forecasted values, which leads to higher MAPE and RMSPE, indicating a higher amount of forecasting error regarding FARHAN. The details about the KPI results of FARHAN and LSTM are also available in Table 5.

Comparison

Performing the comparison between FARHAN and other intelligent methodologies, the MAPE KPI is the comparison metric and the process performed among the models. Consequently, FARHAN performed with the lowest MAPE forecasting error rate and has the highest and nearest accuracy prediction rate among the other intelligent models. The comparison is represented in Table 6, and Fig. 9.

From Fig. 9, FARHAN has showcased exceptional forecasting capabilities across various timeframes, including Yearly, Long-Term, Monthly, and Annual predictions. In the Yearly forecast, FARHAN displayed impressive accuracy with a low MAE of approximately 75.08, indicating minimal prediction errors. In the Long-Term forecast, FARHAN significantly

Table 5 KPI Results of FARHAN & LSTM

Duration	MAPE (%)		RMSPE (%)	
	Proposed	LSTM	Proposed	LSTM
14 Years	0.019162	0.0210	2.5	6.3
1 Year	0.0386	0.0310	5.5	9.5
1 Month	0.039	0.0454	1.2	1.5

improved its performance, reducing the MAE to around 37.15, highlighting its enhanced predictive accuracy for extended periods. Even in Monthly forecasts, FARHAN maintained a competitive edge with an MAE of approximately 93.14, demonstrating its ability to handle shorter timeframes effectively. Moreover, in both Long-Term and Monthly scenarios, FARHAN achieved a perfect score in Accuracy, Precision, and F1 Score, emphasizing its precision and reliability. Additionally, the Annual forecast further solidified FARHAN's capabilities, showcasing its ability to achieve near-perfect predictions with minimal errors across diverse timeframes. FARHAN's consistent high accuracy, precision, and low error rates underscore its exceptional performance in forecasting tasks.

Future Works

In the electrical load forecasting for smart grids, future research is expected to focus on several key areas. Firstly, advancements in machine learning techniques, particularly in deep learning

Table 6 The KPI Comparisons of the Models

| Forecasting Period	Model	MAPE %	RMSPE %	R^2
Long-Term	GRU	0.61	1.40	0.99
	RNN	0.80	1.99	0.99
	KNN	0.01	0.06	1.00
	DT	0.03	0.14	1.00
	RF	0.01	0.07	1.00
	FARHAN	0.03	4.70	1.0
Monthly	GRU	99.94	99.94	-41.19
	RNN	99.76	99.76	-41.03
	KNN	0.15	0.29	1.00
	DT	0.20	0.53	1.00
	RF	0.12	0.26	1.00
	FARHAN	0.02	3.09	1.0
Annual	GRU	7.69	9.17	0.64
	RNN	24.41	24.74	-1.23
	KNN	0.05	0.11	1.00
	DT	0.08	0.28	1.00
	RF	0.04	0.16	1.00
	FARHAN	0.03	5.53	1.0

Fig. 9 The KPI comparison results of (**a**) Long-term, (**b**) Monthly, and (**c**) Annual analysis
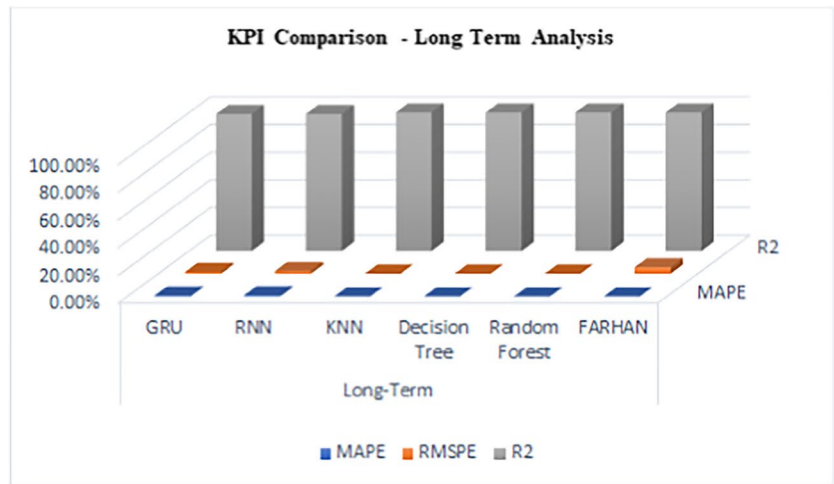
(a)
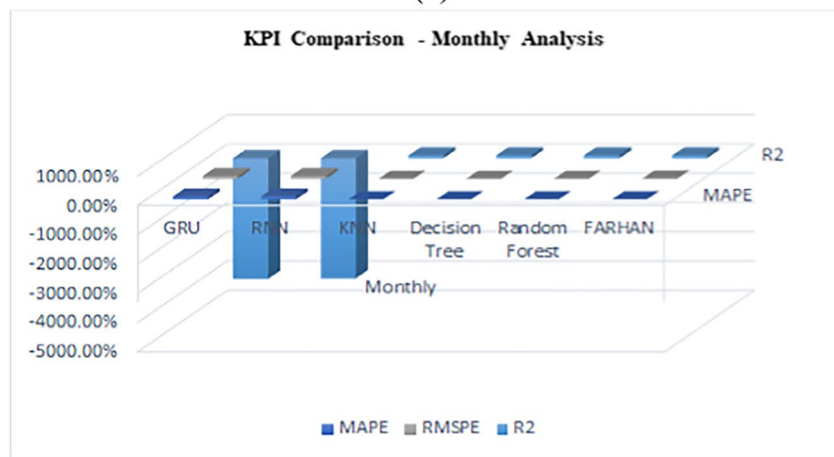
(b)
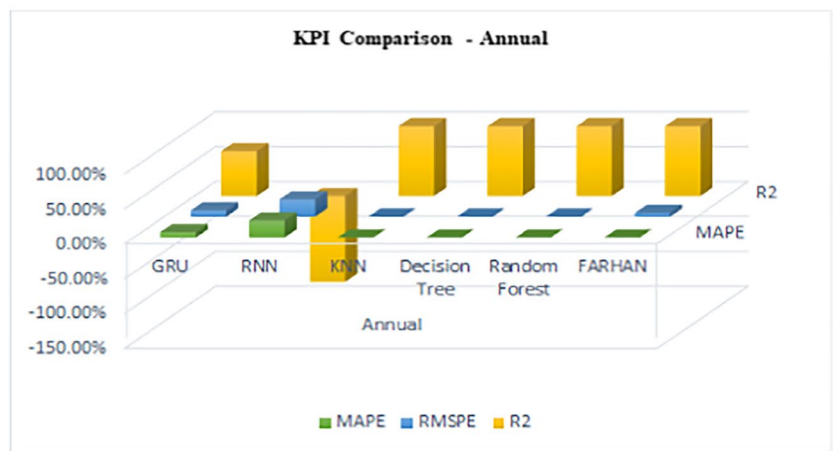
(c)

and reinforcement learning, will likely lead to the development of more sophisticated forecasting models capable of handling complex, high-dimensional data. Additionally, integrating real-time data streams from IoT devices and sensors within smart grids will enhance the accuracy and responsiveness of forecasting models. Furthermore, there will be a continued emphasis on improving the interpretability of these advanced models, ensuring that their predictions are understandable and trustworthy for decision-makers. Collaborative efforts between academia, industry, and policymakers are anticipated to drive

research towards creating more resilient and adaptive forecasting systems, capable of accommodating the dynamic nature of modern energy systems, thereby contributing significantly to the sustainability and efficiency of smart grids. Moreover, the integration of renewable energy sources and the development of forecasting techniques specific to microgrids are likely to be prominent areas of future exploration, aligning with the global push towards greener energy solutions and grid decentralization.

Challenges and Considerations

Implementing FARHAN in practical scenarios presents several challenges and limitations. Firstly, dealing with diverse and complex datasets could impact the model's accuracy, necessitating preprocessing and feature engineering. Additionally, substantial computational resources are required for efficient model training and prediction, particularly when handling large-scale real-world datasets. Developing expertise in configuring and fine-tuning FARHAN demands a deep understanding of neural networks and machine learning algorithms. Furthermore, the model's complexity might hinder interpretability, making it challenging to explain predictions to non-experts. Finally, the efficiency of FARHAN in real data analysis depends on the quality of the input data; inaccurate or incomplete data could lead to unreliable predictions, emphasizing the importance of data quality in the implementation process. Despite these challenges, it's worth noting that FARHAN has demonstrated efficiency in real data analysis, underscoring its potential in advancing the field of electrical load forecasting within smart grids and energy systems.

Candidate Solutions

In addition to the achievements of traditional intelligence models and their global impact, these models often face limitations when dealing with vast quantities of data. Consequently, their performance can become less efficient when processing substantial big data, resulting in reduced accuracy and longer execution times. To address this challenge, it is crucial to develop intelligent systems with the capacity to process extensive big data, providing highly accurate predictions and significantly reducing execution time. Quantum technology emerges as a promising solution, leveraging the principles of quantum mechanics, including superposition, entanglement, and qubits, to enable faster computing. Therefore, intelligent quantum systems like Quantum Neural Networks (QNN), Industrial Quantum Internet of Things (IQIoT) [49], Quantum Fuzzy Logic (QFL), and other related innovations hold the potential to enhance the analysis of intelligent models. In future research endeavors, a broad spectrum of investigations and simulations will focus on developing quantum AI-based intelligent systems for more precise and efficient forecasting and intelligent management of smart grids.

Conclusion

In this research, a novel intelligent estimator system named FARHAN (Descending Neuron Coupled LSTM Averaged Markov Simulated Neural Networks) was developed and applied. An electricity consumption dataset was used to assess FARHAN's performance in the context of smart grids and smart, sustainable cities. FARHAN analyzed a dataset containing 121,260 instances of electricity consumption and conducted mid-to-long-term forecasting. The key performance indicators (KPIs) obtained were MAPE (%) = 0.019162, RMSPE (%) = 2.5, and $R^2 = 1$, indicating an exceptional level of accuracy. Comparing these results with classical and other intelligent models confirms FARHAN's accuracy. The contributions of this paper are as follows: 1) Efficiently integrating various intelligent methodologies. 2) Introducing an exclusive model structure developed for this research. 3) Achieving a remarkable forecasting accuracy of 99.980838%, with minimal error rates (MAPE (%) and RMSPE (%) of 0.019162 and 2.5, respectively). 4) Implementing a system identification-based averaging mechanism that optimizes FARHAN's performance. 5) Demonstrating the ability to analyze extensive volumes of big data.

Permission to Reproduce Materials from other Sources

None.

Author Contribution   *A.S*–prepared the simulations, wrote, and provided the manuscript. *H.K* with *A.R* – rechecked scientific formulations and reviewed the simulations and the manuscript.

Declarations

Ethical Approval   This article does not contain any studies with human participants or animals performed by any of the authors.

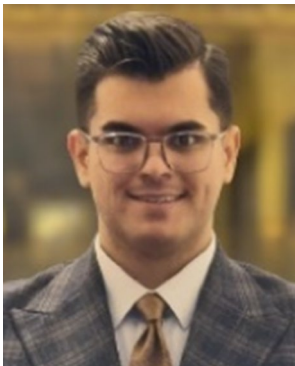**Conflict of Interest**   None.

References

1. Daneshvar M, Mohammadi-Ivatloo B, Zare K (2022) A novel transactive energy trading model for modernizing energy hubs in the coupled heat and electricity network. J Clean Prod 344:131024
2. Safari A, Ghavifekr AA (2021) International stock index prediction using artificial neural network (ANN) and Python programming. In *2021 7th International Conference on Control, Instrumentation and Automation (ICCIA)*, IEEE, pp. 1–7

3. Safari A (2022) Data Driven Artificial Neural Network LSTM Hybrid Predictive Model Applied for International Stock Index Prediction. In *2022 8th International Conference on Web Research (ICWR)*, IEEE, pp. 115–120

4. Yang W, Shi J, Li S, Song Z, Zhang Z, Chen Z (2022) A combined deep learning load forecasting model of single household resident user considering multi-time scale electricity consumption behavior. Appl Energy 307:118197

5. Besanger Y (2022) Short term Markov corrector for building load forecasting system–Concept and case study of day-ahead load forecasting under the impact of the COVID-19 pandemic. Energy Build 270:112286

6. Saeed F, Paul A, Seo H (2022) A Hybrid Channel-Communication-Enabled CNN-LSTM Model for Electricity Load Forecasting. Energies 15(6):2263

7. Chaianong A, Winzer C, Gellrich M (2022) Impacts of traffic data on short-term residential load forecasting before and during the COVID-19 pandemic. Energ Strat Rev 43:100895

8. Veeramsetty V, Reddy KR, Santhosh M, Mohnot A, Singal G (2022) Short-term electric power load forecasting using random forest and gated recurrent unit. Electr Eng 104(1):307–329

9. Shaqour A, Ono T, Hagishima A, Farzaneh H (2022) Electrical demand aggregation effects on the performance of deep learning-based short-term load forecasting of a residential building. Energy AI 8:100141

10. Kang L et al (2022) Feed-forward active operation optimization for CCHP system considering thermal load forecasting. Energy p. 124234

11. Veeramsetty V, Chandra DR, Grimaccia F, Mussetta M (2022) Short Term Electric Power Load Forecasting Using Principal Component Analysis and Recurrent Neural Networks. Forecasting 4(1):149–164

12. Lu Y, Wang G, Huang S (2022) A short-term load forecasting model based on mixup and transfer learning. Electric Power Syst Res 207:107837

13. Hou H et al (2022) Review of load forecasting based on artificial intelligence methodologies, models, and challenges. Electric Power Syst Res 210:108067

14. Meng Z, Xie Y, Sun J (2022) Short-term load forecasting using neural attention model based on EMD. Electr Eng 104(3):1857–1866

15. Zhao X, Shen B, Lin L, Liu D, Yan M, Li G (2022) Residential Electricity Load Forecasting Based on Fuzzy Cluster Analysis and LSSVM with Optimization by the Fireworks Algorithm. Sustainability 14(3):1312

16. Shen X, Zhao H, Xiang Y, Lan P, Liu J (2022) Short-term electric vehicles charging load forecasting based on deep learning in low-quality data environments. Electric Power Syst Res 212:108247

17. Kandilogiannakis G, Mastorocostas P, Voulodimos A (2022) ReNFuzz-LF: A Recurrent Neurofuzzy System for Short-Term Load Forecasting. Energies 15(10):3637

18. Peng C, Tao Y, Chen Z, Zhang Y, Sun X (2022) Multi-source transfer learning guided ensemble LSTM for building multi-load forecasting. Expert Syst Appl 202:117194

19. Tong C, Zhang L, Li H, Ding Y (2022) Temporal inception convolutional network based on multi-head attention for ultra-short-term load forecasting. IET Gener Transm Distrib 16(8):1680–1696

20. Xiao X, Mo H, Zhang Y, Shan G (2022) Meta-ANN–A dynamic artificial neural network refined by meta-learning for Short-Term Load Forecasting. Energy 246:123418

21. Ullah FU, Ullah A, Khan N, Lee MY, Rho S, Baik SW (2022) Deep learning-assisted short-term power load forecasting using deep convolutional LSTM and stacked GRU. Complexity 26:2022

22. Gao R, Du L, Suganthan PN, Zhou Q, Yuen KF (2022) Random vector functional link neural network based ensemble deep learning for short-term load forecasting. Expert Syst Appl 206:117784

23. Abdolrezaei H, Siahkali H, Olamaei J (2022) Substation mid-term electric load forecasting by knowledge-based method. Energy Ecol Environ 7(1):26–36

24. Lin X, Zamora R, Baguley CA, Srivastava AK (2022) A hybrid short-term load forecasting approach for individual residential customer. IEEE Trans Power Delivery 38(1):26–37

25. Salehimehr S, Taheri B, Sedighizadeh M (2022) Short-term load forecasting in smart grids using artificial intelligence methods: a survey. J Eng 2022(12):1133–1142

26. Azeem A, Ismail I, Jameel SM, Romlie F, Danyaro KU, Shukla S (2022) Deterioration of Electrical Load Forecasting Models in a Smart Grid Environment. Sensors 22(12):4363

27. Wang C, Wang Y, Ding Z, Zheng T, Hu J, Zhang K (2022) A Transformer-Based Method of Multienergy Load Forecasting in Integrated Energy System. IEEE Trans Smart Grid 13(4):2703–2714

28. Dab K, Agbossou K, Henao N, Dubé Y, Kelouwani S, Hosseini SS (2022) A compositional kernel based gaussian process approach to day-ahead residential load forecasting. Energy Build 254:111459

29. Brusaferri A, Matteucci M, Spinelli S, Vitali A (2022) Probabilistic electric load forecasting through Bayesian Mixture Density Networks. Appl Energy 309:118341

30. Arvanitidis AI, Bargiotas D, Daskalopulu A, Kontogiannis D, Panapakidis IP, Tsoukalas LH (2022) Clustering Informed MLP Models for Fast and Accurate Short-Term Load Forecasting. Energies 15(4):1295

31. Moon J, Park S, Rho S, Hwang E (2022) Interpretable short-term electrical load forecasting scheme using cubist. Comput Intell Neurosci 8:2022

32. Vargas-Salgado C, Berna-Escriche C, Escrivá-Castells A, Díaz-Bello D (2022) Optimization of All-Renewable Generation Mix According to Different Demand Response Scenarios to Cover All the Electricity Demand Forecast by 2040: The Case of the Grand Canary Island. Sustainability 14(3):1738

33. Dinh HT, Lee K-H, Kim D (2022) Supervised-learning-based hour-ahead demand response for a behavior-based home energy management system approximating MILP optimization. Appl Energy 321:119382

34. Stanelyte D, Radziukyniene N, Radziukynas V (2022) Overview of Demand-Response Services: A Review. Energies 15(5):1659

35. Khan ZA et al (2022) Efficient short-term electricity load forecasting for effective energy management. Sustain Energy Technol Assess 53:102337

36. Safari A, Ghavifekr AA (2022) Use case of artificial intelligence and neural networks in energy consumption markets and industrial demand response. Industrial demand response: methods, best practices, case studies, and applications 4(4):379

37. Kondaiah V, Saravanan B (2022) Short-Term Load Forecasting with a Novel Wavelet-Based Ensemble Method. Energies 15(14):5299

38. Safari A, Ghavifekr AA (2022) Quantum Technology & Quantum Neural Networks in Smart Grids Control: Premier Perspectives. In: 2022 8th International Conference on Control, Instrumentation and Automation (ICCIA), IEEE, pp. 1–6

39. Safari A, Ghavifekr AA (2021) QNN Application in Weather Prediction of Smart Grids. In: 2021 11th Smart Grid Conference (SGC), IEEE, pp. 1–6

40. Xian H, Che J (2022) Multi-space collaboration framework based optimal model selection for power load forecasting. Appl Energy 314:118937

41. Yazici I, Beyca OF, Delen D (2022) Deep-learning-based short-term electricity load forecasting: A real case application. Eng Appl Artif Intell 109:104645

42. Haque A, Rahman S (2022) Short-term electrical load forecasting through heuristic configuration of regularized deep neural network. Appl Soft Comput 122:108877

43. Safari A, Sabahi M (2021) Practical data connection between MATLAB and microcontrollers using virtual serial port and MicroPython Pyboard: A survey. IET Circuits Devices Syst 15(5):485–492

44. Zhang H, Razmjooy N (2023) Optimal Elman Neural Network based on Improved Gorilla Troops Optimizer for Short-Term Electricity Price Prediction. J Electr Eng Technol 21:1–5

45. Alferaidi A, Yadav K, Alharbi Y, Razmjooy N, Viriyasitavat W, Gulati K, Kautish S, Dhiman G (2022) Distributed deep CNN-LSTM model for intrusion detection method in IoT-based vehicles. Math Probl Eng 19:2022

46. Sun L, Han XF, Xu YP, Razmjooy N (2021) Exergy analysis of a fuel cell power system and optimizing it with fractional-order coyote optimization algorithm. Energy Rep 1(7):7424–7433

47. Safari A, KheirandishGharehbagh H, Nazari-Heris M, Oshnoei A (2023) DeepResTrade: A Peer-to-Peer LSTM-Decision Tree-Based Price Prediction and Blockchain-Enhanced Trading System for Renewable Energy Decentralized Markets. Front Energy Res 11:1275686

48. Safari A, KheirandishGharehbagh H, Nazari HM (2023) Deep-VELOX: INVELOX Wind Turbine Intelligent Power Forecasting Using Hybrid GWO–GBR Algorithm. Energies 16(19):6889

49. Safari A, Kharrati H (2023) Application of Optical Wireless Communications in IoT Devices of Smart Grids within Smart Sustainable Cities: With Hybrid Perspectives to Metaverse & Quantum IoT. In: 2023 8th International Conference on Technology and Energy Management (ICTEM). IEEE, pp. 1–7

Hamed Kharrati was born in Tehran, Iran. He received his Ph.D. degree in Electrical Engineering from the Faculty of Electrical and Computer Engineering, University of Tabriz, Iran in 2013. He is now an Associate Professor in the Faculty of Electrical and Computer Engineering at the University of Tabriz. His research interests include Model Predictive Control, Optimal & Nonlinear Control, Fault-tolerant Control, Robotics, and Artificial Intelligence.
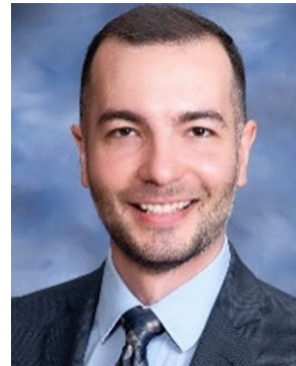
Afshin Rahimi (*Senior MemberIEEE*) is an Associate Professor at the University of Windsor. He received his B.Sc., M.Sc., and Ph.D. degrees in Aerospace Engineering from KNTU in Iran and TMU in Canada in 2010, 2012, and 2017, respectively. He worked at Pratt & Whitney Canada from 2017 to 2018. Since 2010, he has been involved in various industrial research, technology development, and systems engineering projects/contracts related to the control & diagnostics of satellites, UAVs, and commercial aircraft subsystems. His research primarily focuses on model-based and data-driven fault detection, diagnostics, and prognosis; machine learning and intelligent systems; linear & nonlinear controller/observer design; and avionics, sensors, and measurement technologies.

Ashkan Safari (*Student Member, IEEE*) is a distinguished %1 technology-oriented student researcher specializing in Engineering and Intelligent Finance. With a proven track record of delivering high-quality research and management, he has presented his technical findings at numerous international journals and conferences. His research interests span a variety of topics, including Electrical Engineering, Optimization & Control Intelligent Methodologies, Smart Sustainable Cities, Smart Grids, Finance, Management, and Investment in Engineering Science and Technology. He is a prolific author, having contributed to papers and peer reviews. He is a respected professional member of more than 50 committees and is also an esteemed international professional member of the Institute of Electrical and Electronics Engineers (IEEE), a professional organization dedicated to the advancement of technology. His technical skills are demonstrated by his successful completion of more than 250 certifications from prestigious institutions such as top universities, scientific institutes, and giant tech corporations.