JOURNAL OF MECHANICS OF CONTINUA AND MATHEMATICAL SCIENCES
www.journalimcms.org

ISSN (Online) : 2454 -7190 Vol.-19, No.-5, May (2024) pp 116-135 ISSN (Print) 0973-8975

ANOMALY DETECTION IN SMART HOME ELECTRICAL APPLIANCES USING MACHINE LEARNING WITH STATISTICAL ALGORITHMS AND OPTIMIZED TIME SERIES ALGORITHMS

Basim Galeb[1], Haider Saad[2], Haitham Bashar[3], Kadhum Al-Majdi[4] Aqeel Al-Hilali[5]

[1] Department of Computer Technician Engineering, Al Hikma University College, Baghdad, Iraq

[2] Department of Cybersecurity Technology Engineering, Middle Technical University, Baghdad, Iraq

[3] Department of Accounting, Al-Esraa University, Baghdad, Iraq

[4] Department of Medical Instrumentation Engineering, Ashur University, Baghdad, Iraq

[5] Medical Instrumentation Engineering, Al-Farahidi University, Baghdad, Iraq

Email : [1]basim.ghalib@hiuc.edu.iq, [2]haidersaadct@mtu.edu.iq, [3]haitham@esraa.edu.iq, [4]dr.kadhum@au.edu.iq, [5]aqeel@uoalfarahidi.edu.iq

Corresponding Author: **Basim Galeb**

https://doi.org/10.26782/jmcms.2024.05.00008

(Received: March 04, 2024; Revised: April 17, 2024; Accepted: May 02, 2024)

Abstract

Over the last several years, there has been a significant increase in the amount of focus placed on the infrastructure development of smart cities. The primary issue that academics are attempting to address is the issue of energy efficiency. One of these issues was the identification of anomalies in energy usage, which was an essential component that needed to be taken into consideration when managing energy-saving systems that were efficient, hence lowering the total energy consumption and carbon emissions. Therefore, the proposal of a strong approach that is based on the Internet of Things (IoT) might provide more relevance for the identification of abnormal consumption in buildings and the provision of this information to customers and governments so that it can be handled in an appropriate manner to minimize payments. Consequently, the purpose of this work is to explore three different optimization methods, namely ADAM, AadMax, and Nadam, and to advocate for an optimization approach that makes use of the LSTM algorithm to identify anomalies. Statistical modelling techniques such as ARIMA and SARIMAX are used for the purpose of time series forecasting. The findings of the anomaly detection system reveal that the best results are obtained by using LSTM in *conjunction with Nadar. The MSE and RMSE values reached were 0.15348 and 0.02356 respectively. Additionally, the ARIMA model yields the best overall results, with the AIC value being 0.13859 and the MSE value being 300.94365 correspondingly. Confirmation of the suggested model's dependability and flexibility in optimizing anomaly detection is provided by this particular fact.*

Keywords: Anomaly Detection System, Abnormal Consumption, Energy-saving Systems, Statistical Modelling Techniques, Time Series Forecasting.

I. Introduction

Reducing greenhouse gas emissions and climate change is difficult. Nonrenewable resources including oil, coal, and nuclear reactors provide 85% of the world's energy [XXI]. Renewable energy is becoming essential. Optimizing energy utilization and reducing waste is crucial. Energy storage is difficult due to capacity and output issues. Energy efficiency may reduce the negative consequences of buildings on society, the economy, and the environment. Sustainable smart cities aim to balance energy production and demand.

Numerous countries worldwide are researching to create new energy-efficiency measures. Due to the rapid advancement of the Internet of Things (IoT) and the widespread use of Artificial Intelligence (AI) algorithms for data analysis, smart meters are widely used to gather detailed consumption patterns and detect electrical device abnormalities in real-time. Operations, energy usage, gadgets, and energy resources may be managed using the IoT. IoT ensures system stability, optimizes energy utilization, and prevents overload. IoT prevents system failures, injuries, and equipment damage [XI]. Device settings, illumination, home settings, and energy consumption optimization are essential IoT functionalities. Many difficulties emerge from malfunctioning equipment, outdated appliances, or defective system components, which IoT technology may fix.

Device usage, weather, and energy value affect smart home energy consumption. These traits may connect with energy use but not directly affect it. The traits significantly connected with energy utilization must be identified. Before predicting, data must be formatted, outlier identified, and aggregated to detect and manage erroneous data. Significant historical consumption data and parameters are needed to define the mathematical model that represents this system and offers forecasts [XXVI]. To anticipate future consumption, statistical and machine learning approaches analyze previous data and uncover probable links between consumption parameters. The identification of abnormalities might be monitored or unsupervised. Each method handles data differently: supervised with labelled datasets and unsupervised with unlabeled datasets. Power anomaly detection systems are difficult to create without annotated data sets and specific energy use criteria.

The LSTM algorithm, ARIMA and SARIMAX statistical models, and ADAM, AdaMax, and Nadam optimization algorithms are used in this research to manage electricity usage and identify abnormalities. Optimization will include a dropout value to reduce overfitting. The proposed system will be optimized by changing the learning rate. The article follows this structure: The second portion covers comparable work, whereas section 3 explains the neural network, statistical modelling, and optimization strategy. Section 4 covers the proposed system and dataset, while section 5 discusses the results. Section 5 retracts the conclusion.

II. Previous studies

It is essential, when analyzing actual power consumption data, to determine which factors stand out as unique in comparison to the rest. Anomalies are defined as such variables, and the objective of the data-driven collection of all such variables is to identify anomalous power use. Predictions of production patterns of anomalous power consumption can be made in this circumstance using aberrant data that has been acquired in the past. Numerous researchers and industries have devised numerous inventive algorithms to predict energy consumption anomalies in intelligent structures [XXXII, XXXVII]. As an illustration, a proposed methodology for anomaly detection consists of two stages: prediction of consumption and identification of an anomaly. The utilization of Autoregressive Integrated Moving Average (ARIMA) is employed for prediction purposes, whereas two signal rules are implemented for detection. A prediction accuracy ranging from 89.1% to 96.5 percent was achieved for the eight-week data window. Using such an antiquated approach, on the other hand, does not satisfy contemporary IoT-based detection techniques [XVI]. Furthermore, the authors in reference [XXIX] introduced the Days Exceeding Threshold (DET)-TOA method, which calculates temperature and energy consumption based on the difference in deviation between simulated and measured energy consumption. In contrast to conventional DET, analysis of results yields the highest overall performance as a consequence of its multiple sequential methods. Nonetheless, the robustness of the proposed procedure must be evaluated further using a variety of cases. A method for handling time series that is multivariate in nature was suggested in reference [XXX]. This method utilized a single class RNN and implemented Spatio Temporal-Long Short-Term Memory (TL-LSTM). The trust gates demonstrated improved performance, with accuracy rates of 95% and 100% for the dataset that was tested. An approach is proposed to learn features and identify anomalies by combining the Deep Learning (DL) techniques of Long Short-Term Memory (LSTM) and Multilayer Perceptron (MLP). The accuracy of the results generated by the proposed method is superior to that of other methods applied to historical and non-historical power data [XV]. A supervised learning approach was suggested to identify and categorize anomalous energy consumption. This was achieved by employing a Deep Neural Network (DNN) in conjunction with a dispersion of micro-moment classes. The accuracy and F1 score yield respective values of 99.58 and 97.85 percent. Nevertheless, additional refinement is required to accommodate various outdoor conditions [XXII]. A method based on electrical consumption was introduced for both centralized and decentralized cloud computing. However, further research is required to establish the efficacy of this method in terms of detection and to enhance its reliability [XVIII]. Additionally, [XXIII] presents two models that aim to detect anomalous consumption. One of these models employs a single-class Support Vector Machine (SVM) and does not require annotated data. Regarding the classification of the consumption fingerprint, the second model employed supervised micro-moment detection with K-nearest Neighbors (KNN). The investigation incorporated three datasets, and the proposed model achieves high accuracy and F1 score when k=5. Nonetheless, the model's performance accuracy has declined to 93.43 percent for a single dataset that is incapable of data classification and requires additional analysis. Recently, an ML-based framework for regulating energy consumption in smart homes has been developed. IoT devices provide the information utilized for time series analysis. The investigation of the framework incorporates three models: Vector Autoregressive (VAR), Light GBM, and Prophet. The performance of the final two models is most optimal with regard to future prediction and detection. The proposed procedure has a limitation concerning the quantity of the dataset and the sampling rate. Consequently, various approaches should be employed to enhance the efficacy of ML [XXXIII].

III. Methods and Techniques

This section will describe the major techniques and methods related to the topic presented in this paper as follows:

A- Time Series Forecasting

Various economic models are used to forecast demand. One of the core approaches in quantitative forecasting is time-series forecasting, which uses sequentially collected data whose stationarity may be represented by the mean, variance, and autocorrelation functions [XIII, XXIV]. Predictions made at time t must be used at t + l or at a lead time l that varies by property and problem to achieve the lowest probability accuracy between actual and anticipated values [XIII]. Time series forecasting is also used in supply chain management, inventory planning, manufacturing, and business and finance. Prediction difficulties often include time, requiring extrapolation or forecasting. Time series prediction is another important ML application that might use supervised learning. It may be used with random forests (RF), NN, SVM, and regression, among other machine-learning methods. Forecasting uses past data models to forecast future observations [X].

Moving averages, exponential smoothing, simple regression, and ARIMA in its many iterations have been used for forecasting. While these statistical models often use historical data, their efficacy depends on selecting appropriate parameters. Due to improved algorithm modelling, Artificial Neural Networks (ANN) have excelled at predicting across sectors. Prediction models capture seasonal and trend patterns. Stationarity, autocorrelation, and seasonality are investigated by the interrelated components to create a functional time series forecasting model employing the parameters. However, magnitude abnormalities that cause irregularities and impair the fitted model may also impact prediction models [XXIV, X]. Exogenous variables will be studied using seasonal SARIMAX and ARIMA statistical models.

B- ARIMA

ARIMA is the name of a statistical model utilized in time series analysis. ARIMA [LIII, XVII] is composed of two components: the Moving Average (MA) and the Autoregression (AR) models. Utilizing the dependent relationship between an observation and its latency, the AR model operates. Through the utilization of raw observation differencing, also known as the integrated term, the time series is rendered stationary. Equation (1) represents the computation of the AR model with respect to the order of P. The MA model utilizes the correlation between an observation and the residual error resulting from the application of a moving average to delayed data. The moving average hypothesis posits that the future value of a variable can be predicted by averaging its k prior values [LII]. In order to construct the moving average model [XXXIII], Equation (2) may be utilized.

$$y_t = c + \Phi 1 y_{t-1} + \Phi 2 y_{t-2} + \ldots\ldots + \Phi P y_{t-p} + \varepsilon_t \quad (1)$$

$$\hat{y_t} = \sum_{n=1}^{k} y_{t-n} \quad (2)$$

Where Φ is the parameters, ε_t is the white noise, $\hat{y_t}$ represents the predicted value, y_{t-n} is the variable value, and t represents the time. Using ARIMA provides some benefits, which are represented by being suitable for short-term forecasting, only requiring historical data, and providing models for nonstationary data. Figure 1, shows the components of ARIMA model,

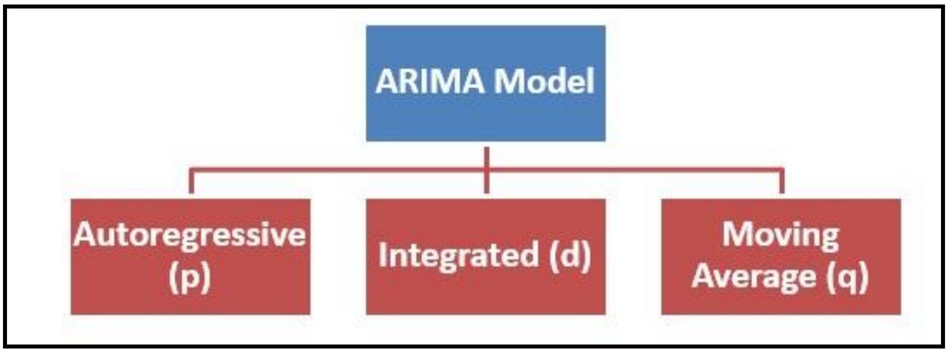

Fig. 1. Components of ARIMA model.

C- LSTM

A second-order recurrent neural network (RNN) architecture known as the Long Short-Term Memory (LSTM) is particularly effective in storing and retrieving consecutive short-term memories over a number of different time steps. The original LSTM training strategy provides essential spatial and temporal locality properties that other training techniques do not provide. However, this comes at the price of limiting its applicability to a limited number of network architectures [L, I]. The flow of input via the gates and the updating of the cell state are both controlled by LSTMs through the use of activation functions. Some examples of these functions are the sigmoid function and the hyperbolic tangent (tanh) function.

Furthermore, LSTM can be considered an extension of RNN because it employs a gating mechanism over an internal memory cell to learn and represent a better and more complex representation of the long-term dependencies among the input sequential data. This makes it suitable for feature learning over a sequence of temporal data [XXX]. LSTM is also capable of learning and representing a more complex representation of the long-term dependencies among serial data.

LSTM can learn to predict complicated temporal patterns in sequential data by successfully controlling information flow and sustaining long-term relationships. They have been a popular option for tasks like time series forecasting, machine translation, sentiment analysis, and other similar activities, where it is essential to comprehend and capture long-range relationships in order to achieve high performance [II].

An LSTM unit typically consists of an input gate i_t, a forget gate f_t, an output gate o_t, and an output state h_t, as well as an internal memory cell state c_t. The LSTM transition equations are written as follows [XXX]:

$$c_t = i_t \odot u_t + f_t \odot c_{t-1} \quad (4)$$

$$h_t = o_t \odot \tanh c_t \tag{5}$$

$$\begin{pmatrix} i_t & o_t \\ f_t & u_t \end{pmatrix} = \begin{pmatrix} \sigma & \sigma \\ \sigma & \tanh \end{pmatrix} \left(M \begin{pmatrix} x_t \\ h_{t-1} \end{pmatrix}\right) \tag{6}$$

Where $\odot$ is the wise product, x_t is the network input, u_t is the modulated input, σ sigmoid activation function, and M is the model transformation that includes the parameters. Figure 2 shows the LSTM.

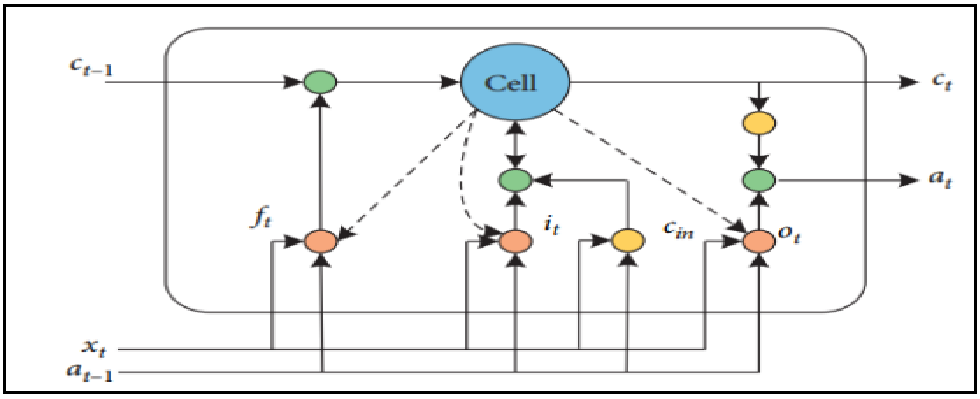

Fig. 2. LSTM concept of work.

D- Internet of Things (IoT)

Objects that are often used are connected to various devices, sensors, software, and networks. The following: Objects that are outfitted with sensors have the ability to collect data, send it over the network, and act together. The sensor of a refrigerator is responsible for gathering data on the temperature outside and adjusting the temperature inside to match the outside temperature. Critical to the fields of healthcare, energy, and environmental research is the Internet of Things. Internet of Things sensors installed on household appliances take temperature and energy readings in real-time. The Internet of Things platform is responsible for monitoring and controlling the energy use, devices, and sources of both commercial and residential properties. Optimizing the Internet of Things requires considerable analysis, which the majority of firms and people do not possess [I, II, VII, XXXIV, XIV].

E- Anomaly Detection

data analysis is the process of examining data in order to identify data points that do not conform to the pattern of conventional data. Based on the findings of the experiment, it is possible to identify points of change in the utilization pattern of energy consumption by using data on energy consumption. It has been shown and published via empirical research that the use of energy increases during months in which consumption is high and decreases during months in which consumption is anticipated to decrease. This is accomplished by tracking fluctuations in consumption patterns. It is possible to discover anomalies in time-series data using the moving average, which is the most straightforward method. Points that deviate from the moving average are referred to as anomalies [XIV, LX, LXI]. Figure 3 shows the different kinds of anomaly detection.

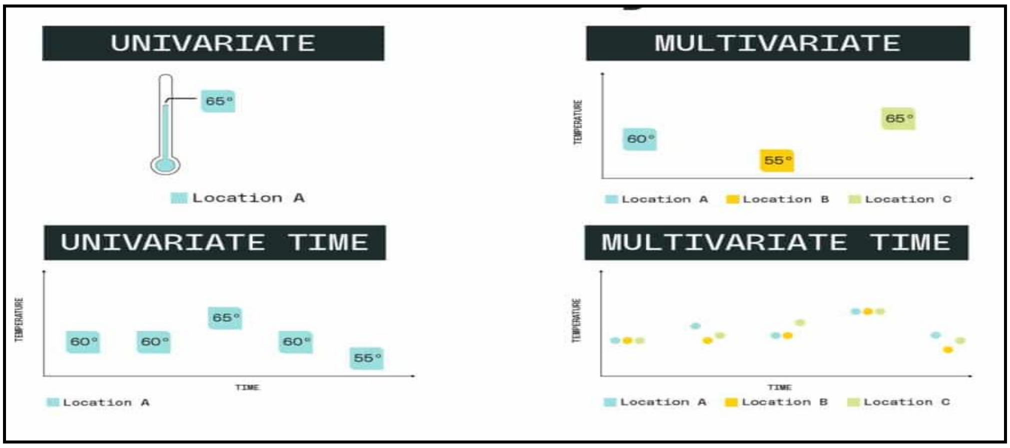

Fig. 3. Anomaly detection types [XXX].

IV. Optimization Algorithms

Optimization is a mathematical field that seeks to represent, analyze, and solve analytical or numerical problems involving the minimization or maximizing of a function on a given dataset. In this paper, the investigated optimization algorithms for the proposed model will be explained below:

1- *Adaptive Moment Algorithm (ADAM)*

Adam algorithm is a hybrid of Momentum and RMSProp. This approach also computes adaptive learning rates for each parameter. Adam, like AdaDelta and RMSprop, preserves an exponentially decaying average of the prior gradient squares *vt*, and Momentum keeps an exponentially decaying average. [XXX] The Adam algorithm equations are as follows [XXXVII, III]:

$$m \leftarrow \beta 1_{m} - (1 - \beta 1)\nabla_{\theta} J(\theta) \quad (7)$$

$$vt \leftarrow \beta 2_{vt} + (1 - \beta 1)\nabla_{\theta} J(\theta) \otimes \nabla_{\theta} J(\theta) \quad (8)$$

$$m \leftarrow m \otimes (1 - \beta_1^t) \quad (9)$$

$$vt = \frac{vt}{(1 - \beta_1^t)} \quad (10)$$

$$\theta = \theta + \frac{\eta m}{\sqrt{vt + U}} \quad (11)$$

The structure of the ADAM algorithm is expressed as seen in Figure 4 [14].

Input: *η, decay rate β_1 and β_2, small constant α (about 10^{-7}), initial θ.*
Initialize gradient accumulation variable $r \leftarrow 0$
$m_0 \leftarrow 0; v_0 \leftarrow 0; t \leftarrow 0$. *(Initialize 1st moment, 2nd moment and time step)*
While *θ_t not converged* ***do****:*
$t \leftarrow t+1$
Calculate gradients for step t: $g_t \leftarrow \nabla_\theta J_t(\theta_{t-1})$
Update biased 1st moment estimate: $m_t \leftarrow \beta_1 \cdot m_{t-1} + (1-\beta_1)\cdot g_t$
Update biased 2nd raw moment estimate: $v_t \leftarrow \beta_2 \cdot v_{t-1} + (1-\beta_2)\cdot g_t^2$
Calculate bias-corrected 1st moment estimate: $\widehat{m_t} \leftarrow m_t / (1-\beta_1^t)$
Calculate bias-corrected 2nd raw moment estimate: $\widehat{v_t} \leftarrow v_t / (1-\beta_2^t)$
Update parameters: $\theta_t \leftarrow \theta_{t-1} - \eta \cdot \widehat{m_t} / \left(\sqrt{\widehat{v_t}} + \alpha\right)$
EndWhile

Fig. 4. ADAM structure [14]

2- *AdaMax Algorithm*

The AdaMax algorithm is an infinite norm-based extension of the Adam algorithm. The factor vt in the update rule of the Adam algorithm adjusts the gradient inversely proportionate to the $l_{٢}$norm of previous gradients $v_{t-١}$, and the current gradient $|g_t|^{٢}$ [XXXVII, IV]. The equations that illustrate the AdaMax procedure can be seen below:

$$vt \leftarrow \beta_{٢\,vt-١} + (١ - \beta 2)|g_t|^{٢} \quad (12)$$

$$vt \leftarrow \beta^{p}_{٢\,vt-١} + (١ - \beta^{p}_{٢})|g_t|^{p} \quad (13)$$

$$ut \leftarrow \beta^{\infty}_{٢\,vt-١} + (١ - \beta^{\infty}_{٢})|g_t|^{\infty} \quad (14)$$

$$ut \leftarrow \max\,(\beta_{٢}.v_{t-١}, |g_t|) \quad (15)$$

The structure of the AdaMax algorithm is expressed as seen in Figure 5 [14].

Input: *η, decay rate β_1 and β_2, small constant α (about 10^{-7}), initial θ.*
$m_0 \leftarrow 0; v_0 \leftarrow 0; t \leftarrow 0$. *(Initialize 1st moment, 2nd moment and time step)*
While *θ_t not converged* ***do****:*
$t \leftarrow t+1$
Calculate gradients for step t: $g_t \leftarrow \nabla_\theta J_t(\theta_{t-1})$
Update biased 1st moment estimate: $m_t \leftarrow \beta_1 \cdot m_{t-1} + (1-\beta_1)\cdot g_t$
Update exponentially weighted infinity norm: $u_t \leftarrow \max\left(\beta_2 \cdot u_{t-1}, |g_t|\right)$
Update parameters: $\theta_t \leftarrow \theta_{t-1} - \left(\eta / \left(1-\beta_1^t\right)\right)\cdot m_t / u_t$
EndWhile

Fig. 5. AdaMax Structure [14].

3- *Nadam Algorithm*

The Nesterov-accelerated Adaptive Moment (Nadam) algorithm [18,40] combines Adam and Nesterov-accelerated algorithms. The objective behind this sort of method is to raise and decrease the decay factor over time. For clarity, a sequence of parameters β_1, β_1,..., β_t corresponding to steps 1, 2,..., t is performed. The momentum step in step *t*+1 is applied once by updating step t instead of *t*+1 as represented in the equations follows [XXXVII, XXVII]:

$$g_t \leftarrow \nabla_{\theta_{t-1}} j_t(\theta_{t-1}) \tag{16}$$

$$m_t \leftarrow \beta_t m_{t-1} + \eta g_t \tag{17}$$

$$\theta_t \leftarrow \theta_{t-1} - (\beta_{t+1} m_t + \eta g_t) \tag{18}$$

The steps related to the gradient and momentum depend on the current gradient as expressed in the following equations [XXXV, XXXVI, VI].

$$\theta_t \leftarrow \theta_{t-1} - \eta \left(\frac{\beta_t m_{t-1}}{1 - \prod_{i=1}^{t} \beta_t} + \frac{(1-\beta_t) g_t}{1 - \prod_{i=1}^{t} \beta_t} \right) \tag{19}$$

$$\theta_t \leftarrow \theta_{t-1} - \eta \left(\frac{\beta_{t+1} m_t}{1 - \prod_{i=1}^{t+1} \beta_t} + \frac{(1-\beta_t) g_t}{1 - \prod_{i=1}^{t} \beta_t} \right) \tag{20}$$

The flowchart of the Nadam algorithm is described in Figure 6.

Input: *learning rates $\eta_1, \eta_2, \ldots, \eta_t$; decay rates $\beta_1, \beta_2, \ldots, \beta_t$; small constant ε; initial θ;*
hyperparameter v.
$m_0 \leftarrow 0; n_0 \leftarrow 0$. *(Initialize 1st and 2nd moments)*
While *θ_t not converged* ***do****:*
Calculate gradients for step t: $g_t \leftarrow \nabla_{\theta_{t-1}} J_t(\theta_{t-1})$
Update biased 1st moment estimate: $m_t \leftarrow \beta_t m_{t-1} + (1-\beta_t) g_t$
Update biased 2nd moment estimate: $n_t \leftarrow v n_{t-1} + (1-v) g_t^2$
$\hat{m} \leftarrow \left(\beta_{t+1} m_t / \left(1 - \prod_{i=1}^{t+1} \beta_i\right)\right) + \left((1-\beta_t) g_t / \left(1 - \prod_{i=1}^{t} \beta_i\right)\right)$
$\hat{n} \leftarrow v n_t / (1 - v^t)$
Update parameters: $\theta_t \leftarrow \theta_{t-1} - \frac{\eta_t}{\sqrt{\widehat{n}_t} + \acute{U}} \widehat{m}_t$
EndWhile

Fig. 6. Nadam structure [14].

V. Proposed Model

The project's anomaly detection and consumption-based machine learning framework for smart homes will be shown here. The framework's stages are shown in Figure 7. The data comes from the Kaggle smart home dataset including weather information. Data from Internet of Things sensors and devices will be included in this dataset on appliance energy use. The selected appliance consumption dataset and its figures are shown in Table 1. From the needed dataset, environmental factors were selected for the suggested framework's performance evaluation. Table 2 displays these statistics. Thus, improving anomaly detection while reducing utilization. Preprocessing the data removes stages to keep the required column to ensure the right machine-learning technique. Extraction and feature selection modify data. Next, design the proposed framework using the LSTM algorithm with fifty layers and the Rectified Linear Activation Function (ReLU), which outputs positive input directly and zero otherwise. Since it's simpler to train and performs better, many neural networks utilize the default activation function. Two statistical models, ARIMA and SARIMAX, are used to predict time series with different orders. Dropout is one of the most important model performance optimizations researched for the LSTM method. Dropout reduces input and recurrent connections from activation and weight changes during training to regularize LSTM units. Probabilistic reasoning does this. Overfitting decreases, improving model performance. Three optimization strategies for performance evaluation are given. To define optimization algorithm step-taking, Adam, AdaMax, and Nadam are used with a learning rate of 0.001.

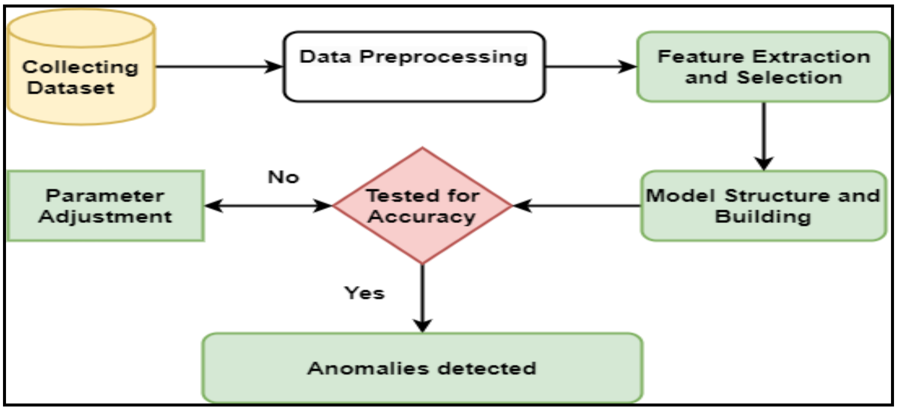

Fig. 7. Flowchart steps of the proposed framework.

Table 1: Sample dataset collected for the proposed framework.

time	House overall	Dishwasher	Home office	Fridge	Garage door	Barn	Well	Microwave	Living room	Furnace	Kitchen	Solar
2016-01-01 05:00:00	0.932833	0.000033	0.442633	0.12415	0.013083	0.03135	0.001017	0.004067	0.001517	0.020700	0.000417	0.003483
2016-01-01 05:01:00	0.934333	0.000000	0.444067	0.12400	0.013117	0.03150	0.001017	0.004067	0.001650	0.020717	0.000417	0.003467

Table 2: The utilized environmental data from the selected dataset.

temperature	humidity	visibility	apparentTemperature	pressure	windSpeed	cloudCover	windBearing	precipIntensity	dewPoint	precipProbability	month	day	weekday	hour	minute
36.14	0.62	10.0	29.26	1016.91	9.18	0.75	282.0	0.0	24.4	0.0	1	1	Friday	5	0
36.14	0.62	10.0	29.26	1016.91	9.18	0.75	282.0	0.0	24.4	0.0	1	1	Friday	5	1

VI. Results and Discussion

For the results of utilizing the LSTM algorithm with the three mentioned previously optimization methods for anomaly detection. The dataset studied uses testing and training stages which indicate simple overfitting occurs at the beginning of the curves, while when moving towards the curves gets more stability as seen in Figure 8, this may confirm the reliability of the proposed framework for testing and training stages. To quantify the achieved stability, the training loss is obtained as seen in Figure 9, which gives a measure of how well a deep learning model matches training data.

In addition, using the moving average for data trends can be defined as the simplest way to detect anomalies in time series forecasting by using the proposed framework. As a result, the LSTM with Nadar gives the best overall results for achieving the best results concerning the studied parameters as listed in Table III and shown in Figure 10.

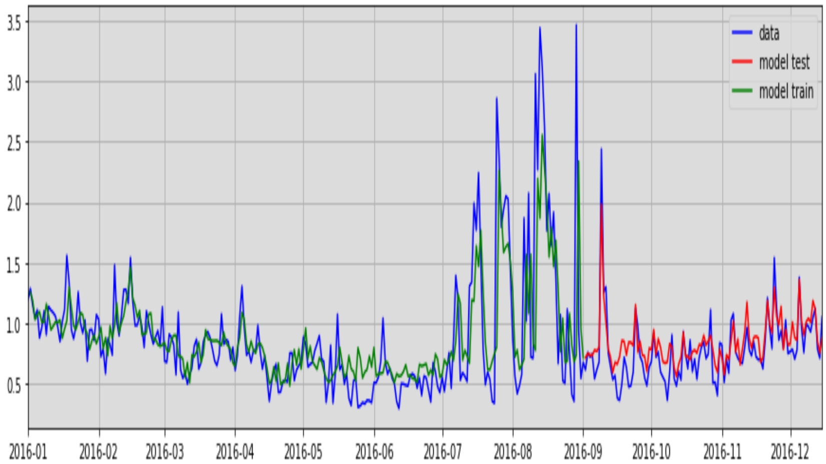

Fig. 8. Proposed framework data testing and training.

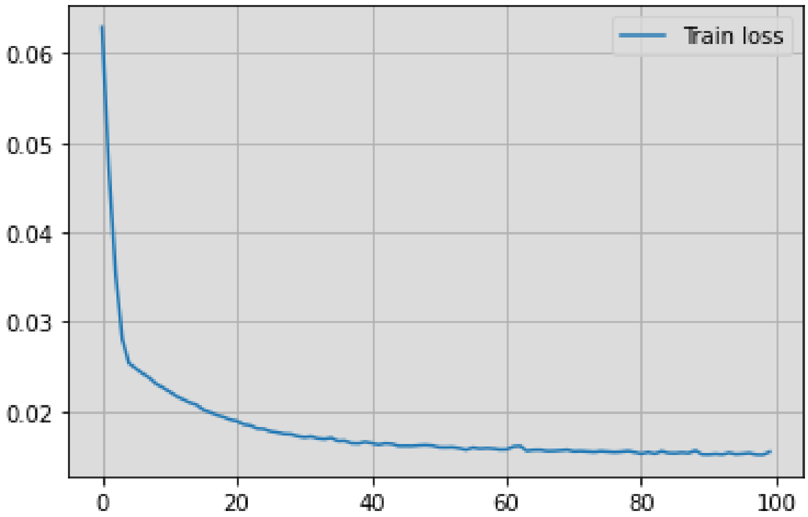

Fig. 9. Training loss parameter.

Table 3: best results of the proposed model.

	Optimizer	MSE	RMSE	MAE	MAPE	R^2
LSTM	Adam	0.025	0.16	0.129	0.197	0.655
	AdaMax	0.024	0.15	0.122	0.181	0.677
	Nadam	0.023	0.15	0.122	0.187	0.685

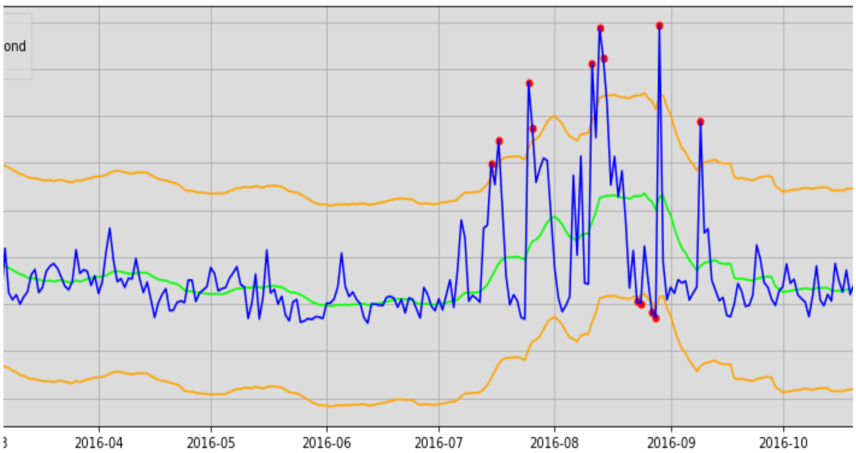

Fig. 10. Anomaly detection with moving average

For the results of utilizing the statistical models that were included with the proposed model for performance evaluation. The results using ARIMA give the best results as compared to the utilized SARIMAX model. These results were based on the evaluation of the AIC and MSE for different orders as seen in Table 4 and Figure 11.

Table 4 : Results of the proposed statistical models

Algorithm	Order (p, d, q)	AIC	MSE
ARIMA	(1,1,2)	300.944	0.13859
ARIMA	(2,1,1)	300.98	0.13861
SARIMAX	(2,1,1)	305.247	0.13536
SARIMAX	(1,1,2)	307.032	0.13615

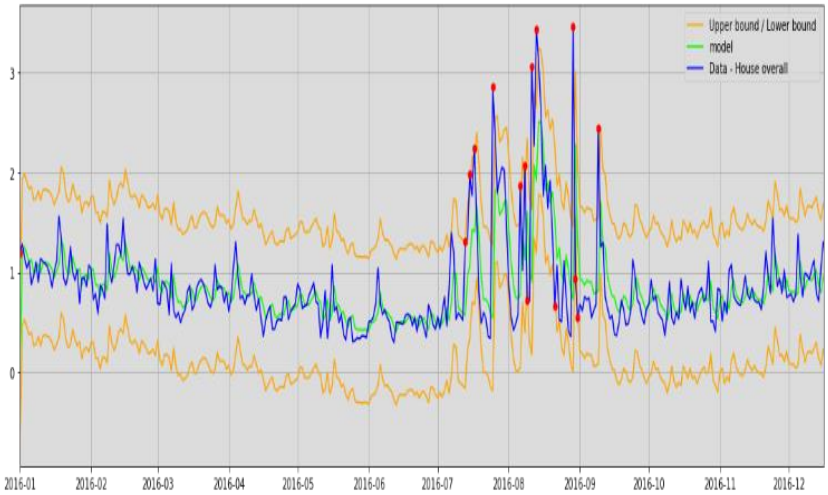

Fig. 11. ARIMA statistical model with moving average.

This study can be enhanced more and more using efficient microstrip RF devices, cloud, e-government, and fog computing as future trends.

VII. Conclusion

This study proposes an optimization framework for energy consumption anomaly detection for smart home IoT sensors and devices. The suggested model uses the strong LSTM algorithm for time series forecasting and two statistical models, ARIMA and SARIMAX, to create the best model. Three optimization strategies were chosen for performance assessment. Based on the analysis parameter, LSTM with Nadam optimization yielded the best results, with MSE and RMSE of 0.15348 and 0.02356. With AIC and MSE values of 0.13859 and 300.94365, the

ARIMA model yields the best results. This proves the model's stability and flexibility in optimizing anomaly detection and lowering power costs.

Conflict of Interest:

There was no relevant conflict of interest regarding this paper.

References

I. Abdulwahid, M. M., Al-Ani, O. A. S., Mosleh, M. F., & Abd-Alhameed, R. A. : 'Investigation of millimeter-wave indoor propagation at different frequencies'. *In 2019 4th Scientific International Conference Najaf (SICN)*. (2019, April). (pp. 25-30). IEEE.

II. Abdulwahid, M. M., Al-Ani, O. A. S., Mosleh, M. F., & Abd-Alhameed, R. A. : 'A Comparison between Different C-band and mmWave band Frequencies for Indoor Communication'. *J. Commun.*, 14(10), (2019). pp. 892-899.

III. Abdulwahid, M. M., Al-Hakeem, M. S., Mosleh, M. F., & Abd-alhmeed, R. A. : 'Investigation and optimization method for wireless AP deployment based indoor network'. *In IOP Conference Series: Materials Science and Engineering*. Vol. 745, No. 1. (2020, February). pp. 012031. IOP Publishing.

IV. Abdulwahid, M. M., & Kurnaz, S. : 'The channel WDM system incorporates of Optical Wireless Communication (OWC) hybrid MDM-PDM for higher capacity (LEO-GEO) inter-satellite link'. Optik. VoL. 273 (2023). 170449. doi.org/10.1016/j.ijleo.2022.170449

V. Abd-Alhameed, R. A., Abdulwahid, M. M., & Mosleh, M. F. : 'Effects of Antenna Directivity and Polarization on Indoor Multipath Propagation Characteristics for different mmWave frequencies'. Informatica 2(1). 2021. pp. 20-28.

VI. Ali, O. M. A., Kareem, S. W., & Mohammed, A. S. : 'Evaluation of Electrocardiogram Signals Classification Using CNN, SVM, and LSTM Algorithm: A review'. *In 2022 8th International Engineering Conference on Sustainable Technology and Development (IEC)*, pp. 185-191. IEEE, 2022.

VII. Almetwali, A. S., Bayat, O., Abdulwahid, M. M., & Mohamadwasel, N. B. : 'Design and Analysis of 50 Channel by 40 Gbps DWDM-RoF System for 5G Communication Based on Fronthaul Scenario'. *In Proceedings of Third Doctoral Symposium on Computational Intelligence*. (2023). (pp. 109-122). Springer, Singapore.

VIII. Alhamadani, N. B., & Abdelwahid, M. M. : 'Implementation of microstrip patch antenna using MATLAB. Informatica'. *Journal of Applied Machines Electrical Electronics Computer Science and Communication Systems*. 2(1). (2021). pp. 29-35.

IX. Alsalemi A, Himeur Y, Bensaali F, et al. : 'Achieving domestic energy efficiency using micro-moments and intelligent recommendations. *IEEE Access*. Vol.8. 2020. pp.15047-15055,

X. Ampountolas, A. : 'Modeling and Forecasting Daily Hotel Demand: A Comparison Based on SARIMAX, Neural Networks, and GARCH Models'. *Forecasting*, Vol.3, No. 3, pp.580-595., 2021

XI. A. Mosavi, A. Bahmani. : 'Energy consumption prediction using machine learning: A review'. *Preprints 2019*, doi.org/10.20944/preprints201903.0131.v1, 2019.

XII. Alsalemi A, Himeur Y, Bensaali F, et al. : 'Achieving domestic energy efficiency using micro-moments and intelligent recommendations. *IEEE Access*. Vol.8. 2020. pp.15047-15055.

XIII. Box, G. E., and Jenkins, G. M. : 'Time series analysis: forecasting and control'. *Holden dsy. Inc. California*, 1976.

XIV. Burhan, I. M., Al-Hakeem, M. S., Abdulwahid, M. M., & Mosleh, M. F. : 'Investigating the Access Point height for an indoor IOT services'. *In IOP Conference Series: Materials Science and Engineering*. Vol. 881, No. 1, (2020, July). pp. 012116. IOP Publishing.

XV. Buzau, M. M., Tejedor-Aguilera, J., et al. : 'Hybrid deep neural networks for detection of non-technical losses in electricity smart meters. *IEEE Transactions on Power Systems*, Vol. 35, No. 2, 2019. pp.1254-1263.

XVI. Chou, J. S., & Telaga, A. S. : 'Real-time detection of anomalous power consumption'. *Renewable and Sustainable Energy Reviews*, Vol. 33. 2014. pp.400-411.,

XVII. Farsi, M., Hosahalli, D., Manjunatha, B. R., Gad, I., et al. : 'Parallel genetic algorithms for optimizing the SARIMA model for better forecasting of the NCDC weather data. *Alexandria Engineering Journal*, Vol.60, No. 1, (2021). pp. 1299-1316.

XVIII. Feng, L., Xu, S., Zhang, L., Wu, J., Zhang, J., et al. : 'Anomaly detection for electricity consumption in cloud computing: framework, methods, applications, and challenges'. *EURASIP Journal on Wireless Communications and Networking*, Vol. 1. (2020). pp.1-12,

XIX. F. Abayaje, S. A. Hashem, H. S. Obaid, Y. S. Mezaal, and S. K. Khaleel. : 'A miniaturization of the UWB monopole antenna for wireless baseband transmission'. *Periodicals of Engineering and Natural Sciences*, vol. 8. no. 1. (2020). pp. 256-262.

XX. G. Zhao, L. Xing. 'Reliability analysis of IoT systems with competitions from cascading probabilistic function dependence. *Reliab. Eng. Syst. Saf.*, Vol. 198. (2020). pp.106812. https://doi.org/10.1016/j.ress.2020.106812

XXI. Himeur Y, Elsalemi A, Bensaali F, Amira A. : 'Robust event-based non-intrusive appliance recognition using multi-scale wavelet packet tree and ensemble bagging tree. *Appl Energy.* Val. 267. (2020). pp.114877.

XXII. Himeur, Y., Alsalemi, A., Bensaali, et al. : 'A novel approach for detecting anomalous energy consumption based on micro-moments and deep neural networks'. *Cognitive Computation*, Vol. 12. No. 6. (2020). pp. 1381-1401.

XXIII. Himeur, Y., Alsalemi, A., Bensaali, F., at al. : 'Smart power consumption abnormality detection in buildings using micromoments and improved K-nearest neighbors. *International Journal of Intelligent Systems*, Vol. 36, No. 6. (2012). pp. 2865-2894.

XXIV. Hyndman, R. J., and Athanasopoulos, G. : 'Forecasting: principles and practice. *OTexts*. 2018.

XXV. H. A. Hussein, Y. S. Mezaal, and B. M. Alameri. : 'Miniaturized microstrip diplexer based on FR4 substrate for wireless communications'. *Elektron. Ir Elektrotech*. Vol. 27 No. 5. (2021) . doi.org/10.5755/j02.eie.28942

XXVI. I. O. Essiet, Y. Sun, Z. Wang. 'Optimized energy consumption model for smart home using improved differential evolution algorithm'. *Energy* Vol. 172. (2019). pp.354–365. https://doi.org/10.1016/j.

XXVII. Jamal, S. A., Ibrahim, A. A., Abdulwahid, M. M., & Wasel, N. B. M. (2020). 'Design and implementation of multilevel security-based home management system'. *International Journal of Advanced Trends in Computer Science and Engineering*. Vol. 9(4). (2020). pp. 5716-5720. doi.org/10.30534/ijatcse/2020/224942020

XXVIII. J. Ali and Y. Miz'el. 'A new miniature Peano fractal-based bandpass filter design with 2nd harmonic suppression 3rd IEEE International Symposium on Microwave.' Antenna, Propagation and EMC Technologies for Wireless Communications, Beijing, China, 2009.

XXIX. Lin, G., & Claridge, D. E. : 'A temperature-based approach to detect abnormal building energy consumption'. *Energy and Buildings*, Vol. 93. (2015). pp. 110-118.

XXX. Liu, J., Shahroudy, A., Xu, D., et al. : 'Spatio-temporal lstm with trust gates for 3d human action recognition'. *In European conference on computer vision*. (2016). pp. 816-833. Springer, Cham.,

XXXI. Liu Y, Geng G, Gao S, Xu W. : Non-intrusive energy use monitoring for a group of electrical appliances'. *IEEE Trans Smart Grid*, Vol.9, No. 4. (2018). pp. 3801-3810,

XXXII. Luo, Y., Li, W., & Qiu, S. : 'Anomaly detection based latency-aware energy consumption optimization for IoT data-flow services'. *Sensors*, Vol.20. No. 1. (2019). pp.122.

XXXIII. Malki, A., Atlam, E. S., and Gad, I. : 'Machine learning approach of detecting anomalies and forecasting time-series of IoT devices'. *Alexandria Engineering Journal*, Vol. 61 No. 11. (2022). pp. 8973-8986.

XXXIV. Mohsen, D. E., Abbas, E. M., & Abdulwahid, M. M. : 'Design and Implementation of DWDM-FSO system for Tbps data rates with different atmospheric Attenuation'. *In 2022 International Congress on Human-Computer Interaction, Optimization and Robotic Applications* (HORA) (2022, June). (pp. 1-7). IEEE. 10.1109/HORA55278.2022.9799974

XXXV. Mohsen, D. E., Abbas, E. M., & Abdulwahid, M. M. : 'Performance Analysis of OWC System based (S-2-S) Connection with Different Modulation Encoding'. *International Journal of Intelligent Systems and Applications in Engineering*. 11(4s). (2023). pp. 400-408.

XXXVI. Monner, D., and Reggia, J. A. : 'A generalized LSTM-like training algorithm for second-order recurrent neural networks. *Neural Networks*, Vol. 25. (2012). pp.70-83.

XXXVII. Mustapha, A., Mohamed, L., and Ali, K. : 'Comparative study of optimization techniques in deep learning: Application in the ophthalmology field'. *In Journal of Physics: Conference Series*, Vol. 1743, No. 1. (2021) pp. 012002. IOP Publishing.

XXXVIII. M. Kh. : 'Cybercrime Challenges in Iraqi Academia: Creating Digital Awareness for Preventing Cybercrimes.' *International Journal of Cyber Criminology*. vol. 16, no. 2. (2022). pp. 1–14.

XXXIX. M. Q. Mohammed. : 'HARNESSING CLOUD OF THING AND FOG COMPUTING IN IRAQ: ADMINISTRATIVE INFORMATICS SUSTAINABILITY'. *Journal of Mechanics of Continua and Mathematical Sciences*. vol. 19, no. 2. (2024) pp. 66–78. 10.26782/jmcms.2024.02.00004

XL. M. S. Jameel, Y. S. Mezaal, and D. C. Atilla. : 'Miniaturized coplanar waveguide-fed UWB Antenna for wireless applications'. *Symmetry*. vol. 15, no. 3. (2023). pp. 633. doi.org/10.3390/sym15030633

XLI. M. S. Shareef et al. : 'Cloud of Things and fog computing in Iraq: Potential applications and sustainability.' : *Heritage and Sustainable Development*. vol. 5, no. 2. (2023). pp. 339–350.

XLII. M. Zekic´-Susˇac, S. Mitrovic´, A. Has. : 'Machine learning based system for managing energy efficiency of public sector as an approach towards smart cities'. *Int. J. Inf. Manage*. Vol 58. (2021) doi.org/10.1016/j.ijinfomgt.2020.102074

XLIII. Sardianos C, Varlamis I, Chronis C, et al. : 'A model for predicting room occupancy based on motion sensor data'. *In: 2020 IEEE International Conference on Informatics, IoT, and Enabling Technologies (ICIoT)*. (2020). pp. 394-399.

XLIV. Sardianos C, Varlamis I, Dimitrakopoulos G, et al. : 'REHAB-C: recommendations for energy HABits change'. *Future Gener Comput Syst*. Vol. 112. (2020). pp. 394-407. 10.1016/j.future.2020.05.041

XLV. Sohani, A., Sayyaadi, H., Cornaro, C., Shahverdian, et. al. : 'Using machine learning in photovoltaics to create smarter and cleaner energy generation systems: A comprehensive review'. *Journal of Cleaner Production*. Vol 364 (2022). doi.org/10.1016/j.jclepro.2022.132701

XLVI. S. A. Abdulameer. : 'Security Readiness in Iraq: Role of the Human Rights Activists'. *International Journal of Cyber Criminology*. vol. 16. no. 2. (2022) pp. 1–14.

XLVII. S. I. Yahya et al.. 'A New Design Method for Class-E Power Amplifiers Using Artificial Intelligence Modeling for Wireless Power Transfer Applications'. *Electronics*. vol. 11, no. 21 (2022). pp. 3608.

XLVIII. S. Roshani et al. : 'Design of a compact quad-channel microstrip diplexer for L and S band applications'. *Micromachines* (Basel), vol. 14, no. 3. (2023). doi.org/10.3390/mi14030553

XLIX. T. Abd, Y. S. Mezaal, M. S. Shareef, S. K. Khaleel, H. H. Madhi, and S. F. Abdulkareem. : 'Iraqi e-government and cloud computing development based on unified citizen identification'. *Periodicals of Engineering and Natural Sciences*. vol. 7. no. 4. (2019) pp. 1776–1793.

L. Vagropoulos, S. I., Chouliaras, G. I., Kardakos, E. G., et al. : 'Comparison of SARIMAX, SARIMA, modified SARIMA and ANN-based models for short-term PV generation forecasting'. *In 2016 IEEE International Energy Conference (ENERGYCON)*, pp. 1-6. IEEE. 10.1109/ENERGYCON.2016.7514029

LI. V. Marinakis, H. Doukas. : ‘An advanced IoT-based system for intelligent energy management in buildings’. *Sensors*, Vol.18, No. 2, (2018). pp.610. https://doi.org/10.3390/s18020610

LII. Wu, D., Zhang, L., and Lin, L. : ‘Based on the moving average and target motion information for detection of weak small target’. *In 2018 International Conference on Intelligent Transportation, Big Data & Smart City (ICITBS)*, pp. 641-644. IEEE, 2018.

LIII. Yang, L., & Yang, H. : ‘A Combined ARIMA-PPR Model for Short-Term Load Forecasting’. *In 2019 IEEE Innovative Smart Grid Technologies-Asia (ISGT Asia)*. IEEE. (2019). pp. 3363-3367.

LIV. Y. S. Mezaal and H. T. Eyyuboglu. : ‘A new narrow band dual-mode microstrip slotted patch bandpass filter design based on fractal geometry’. *In 2012 7th International Conference on Computing and Convergence Technology (ICCCT)*. IEEE. (2012). pp. 1180-1184.

LV. Y. S. Mezaal, and J. K. Ali. : ‘A new design of dual band microstrip bandpass filter based on Peano fractal geometry: Design and simulation results’. *Presented at the 2013 13th Mediterranean Microwave Symposium (MMS),* IEEE. 2013. 10.1109/MMS.2013.6663140

LVI. Y. S. Mezaal and S. F. Abdulkareem. : ‘New microstrip antenna based on quasi-fractal geometry for recent wireless systems’. *In 2018 26th Signal Processing and Communications Applications Conference (SIU)*. IEEE. (2018) pp. 1-4.

LVII. Y. S. Mezaal et al. : ‘Cloud computing investigation for cloud computer networks using cloudanalyst’. *Journal of Theoretical and Applied Information Technology*, vol. 96. no. 20. (2018)

LVIII. Y.S. Mezaal et al. : ‘Cloud computing investigation for cloud computer networks using cloudanalyst’. *Journal of Theoretical and Applied Information Technology*. vol. 96, no. 20, (2018)

LIX. Y. S. Mezaal, H. H. Saleh, and H. Al-Saedi. ‘New compact microstrip filters based on quasi fractal resonator’. *Advanced Electromagnetics*. vol. 7. no. 4. (2018) pp. 93-102.

LX. Y. S.Mezaal, Eyyuboglu, H. T., &Ali, J. K. Wide bandpass and narrow bandstop microstrip filters based onHilbert fractal geometry: design and simulation results. PloS one, 9(12),e115412, 2014.

LXI. Zaal, R. M., Mustafa, F. M., Abbas, E. I., Mosleh, M. F., & Abdulwahid, M. M. : ‘Real measurement of optimal access point localizations’. *In IOP Conference Series: Materials Science and Engineering*. Vol. 881. No. 1. (2020). pp. 012119). IOP Publishing.