Electricity Consumption Forecasting Using Facebook Prophet

Mohammed Abdel Hamid Musa Al-Shamoosy[1] and Suhad Ali Shaheed[2]
Department of Statistics, College of Administration and Economics, Al-Mustansiriyah University, 46167 Baghdad, Iraq
mohamad.abdalhamid@uomustansiriyah.edu.iq, dr.suhadali@uomustansiriyah.edu.iq

Keywords: Prediction, Electricity Consumption, Facebook Prophet, L-BFGS Algorithms.

Abstract: This research seeks for predicting electricity consumption in Baghdad Governorate using the Facebook Prophet Model, which was Proposed by Taylor and Latham in 2017 as a modern generalized additive Model specialized in tuning and forecasting time series data. This model includes non-linear functions that are combined into one model. The most important feature of this model is that it supports the presence of change points in the general trend, and uses the Fourier series to model seasonality, In addition, the model contains a function to represent the effect of special events (Holidays) in the time series. The electricity consumption data were analyzed and their estimated values were extracted according to the Facebook Prophet model using the Limited Memory Broyden, Fletcher, Goldfarb, and Shanno (L-BFGS) algorithm. After that, the estimated values were compared with the original values to measure the accuracy of the model`s performance. The results showed that the model provided relatively accurate prediction performance, which reflects the model`s ability to predict electricity consumption, and the possibility of using it as a tool to support energy demand management.

1 INTRODUCTION

Electricity is one of the essential elements for society development and the backbone of the economy and infrastructure, as it is relied upon in various industrial, service, domestic, and commercial activities. Demand for electric power increases significantly with population growth, growing economic activities, and climate change, which directly affects electricity consumption Patterns. As so, this requires the provision of highly accurate tools to predict this demand.

Forecasting electric power consumption is also an essential step in the strategic planning of energy production and distribution companies, as it contributes to ensuring the stability of the electricity grid and reducing resource waste. This issue is becoming increasingly important in Iraq, especially in Baghdad Governorate, as it is the largest urban and energy - consuming center in the country.

In this context, the Facebook Prophet model, discovered by Meta, emerges as a modern and effective generalized model for time series analysis and forecasting, as it has the ability to handle multi-seasonality, long-term trends, and missing and extreme values. There are many recent studies that have addressed the Facebook Prophet model in predicting time series data. The researchers Taylor and Latham were the first to discover this model in 2017 [1]. In 2021, the researcher Francesco Lomio used the Facebook Prophet model in his thesis to predict the number of arrivals the next day to the emergency department at the university of Tampere and the results showed that Facebook Prophet provided high prediction accuracy [2]. In 2023, the (Hasnain & et all) used Facebook Prophet to predict the ambient PM $_{2.5}$ concentration in China and the results indicated that there was a high match between the actual values and the estimated values [3].

This research aims to employ Facebook Prophet in analyzing electricity consumption data in Baghdad Governorate and predicting its future values.

2 FACEBOOK PROPHET MODEL

The Facebook Prophet model is considered one of the modern statistical tools for predicting time series. It was discovered by Taylor and Latham in 2017. It is considered a generalized additive model specialized in fitting time series data. It is characterized by its ability to deal with extreme values, and it can also deal well with missing values [4].

The work of the Facebook Prophet is based on the assumption that the time series can be described as a combination of several properties, such as trend, seasonality, in addition to special events (holidays) [5].

2.1 Model Structure and Components

The mathematical formula for the Facebook Prophet is written as follows:

$$Y(t) = g(t) + F(t) + h(t) + \varepsilon_t. \quad (1)$$

Where:

- $Y(t)$: the time series to be predicted;
- $g(t)$: a function that measures the nature of trends (linear or non-linear) and (non-periodic) in the data, and represents the growth function;
- $F(t)$: a function that measures seasonal trends and seasonal periodic changes;
- $h(t)$: a function that indicates any unusual changes over time (holiday seasons, national occasions, etc);
- ε_t: the random error term, which contains any pattern that the model does not recognize, and is assumed to be normally distributed $\varepsilon_t \sim N(0, \sigma^2)$ [1].

The growth function $g(t)$ in the Facebook Prophet model is of two types: the first is when the growth is non-linear, and the second is when the growth is linear.

The non-linear growth function is used when there is an upper limit to the time series that cannot be exceeded, and it is represented by the logistic growth function, and it can be formulated according to the following equation:

$$g(t) = \frac{c}{1 + e^{(-k(t-m))}}. \quad (2)$$

Where c: the maximum limit of the time series, k the growth rate (logistic), t a continuous variable, and here represents time (days, years, ete), m offset Parameter.

The growth function has two assumptions, the first is that the maximum limit is not constant, as the larger the values of the observations of the time series, the maximum limit of the growth increases, and thus (c) is replaced with a time -varying capacity $c(t)$, i.e. it has a set of important parameters, and the second assumption is that the growth parameter (k) is also not constant, as the (change points), which are defined as moments in time at which the trend is allowed to change suddenly, either by an increase or decrease in the growth rate, can change the growth rate significantly. Therefore, the g(t) function must be able to incorporate a varying growth rate to fit the previously recorded data. (2) under the above two assumptions becomes as follows:

$$g(t) = \frac{c(t)}{1 + exp\left(-(k + a(t)^T\delta)\left(t - (m + a(t)^T\gamma)\right)\right)}. \quad (3)$$

Where:

- δ a vector representing the amount of increase or decrease in the growth rate at each change point;
- δ_j the amount of change in the growth rate that occurs at the change point A_j;
- γ a vector for corrections to the offset parameter;
- γ_j the amount of change in the offset parameter at the change point A_j;
- $a(t)$ a vector to determine whether time (t) falls before or after any specific change point and is calculated according to the following:

$$a(t) = \begin{cases} 1, if\ t \geq A_j \\ 0, otherwise \end{cases}. \quad (4)$$

Where: A_j the time in which the growth rate changes.

As for γ, it can be calculated as follows [6], [1]:

$$\gamma_j = \left(A_j - m - \sum_{l<j} \gamma_l\right)\left(1 - \frac{k + \sum_{l<j} \delta_l}{k + \sum_{l\leq j} \delta_l}\right). \quad (5)$$

Where: $\gamma_1 = 0$.

The linear growth function is used in the case of there being no maximum limit for the time series, and it is calculated according to the following (6):

$$g(t) = (k + a(t)^T\delta)t + (m + a(t)^T\gamma). \quad (6)$$

The seasonal function $F(t)$ is calculated according to the Fourier series, and its mathematical formula is as follows:

$$F(t) = \sum_{n=1}^{N} \left[\alpha_n \cos\left(\frac{2\pi nt}{s}\right) + b_n \sin\left(\frac{2\pi nt}{s}\right)\right]. \quad (7)$$

Where:

- s, the seasonal period of the data,
- N the number of terms in the Fourier series,
- t the time Period, $t = 1,2, \ldots, T$, b_n, α_n Fourier parameters [1].

The special events function $h(t)$ represents the events that affect the time series, and its mathematical formula is:

$$h(t) = Z(t)\,\Sigma. \quad (8)$$

Where:

- Σ_i a parameter given to each holiday,

- $Z(t)$ an indicator denoting whether the time (t) falls within the event period (holiday) and calculates as follows [2]:

$$Z(t) = [1\ (t \in Day_1), \ldots, 1\ (t \in Day_l)]. \quad (9)$$

2.2 Estimating the Model Parameters

The Facebook Prophet model is one of the models that deal with smoothing functions, as it includes several smoothing functions namely: trend smoothing functions, seasonal smoothing functions, and holiday smoothing functions. Estimating the model parameters is represented by achieving the Following loss function [3]:

$$Loss\ (\vartheta) = \sum_{t=1}^{T} (Y_t - \hat{Y}_t)^2 + \lambda \sum_{t=1}^{T} |\vartheta_T|. \quad (10)$$

Where:

- λ the amount of penalty imposed on the parameter to avoid overfitting,
- ϑ the vector of parameters for each of the smoothing functions $g(t)$, $F(t)$,$h(t)$, in addition to the variance parameter $\sigma^2{}_{\epsilon}$, where:

$$\vartheta = (k, m, \delta, \gamma, \beta, \Sigma, \sigma^2).$$

The loss function in (10) can be rewritten as follows:

$$\begin{aligned} Loss\ (\vartheta) = argmin \Big\{ &\sum (Y_t - \hat{g}(t) + \hat{F}(t) + \hat{h}(t))^2 \\ &+ \lambda_1 \|\delta\|^2 + \lambda_2 \|\gamma\|^2 \\ &+ \lambda_3 \sum \|\alpha_n + b_n\|^2 + \lambda_4 \|\Sigma\|^2 \Big\}, \end{aligned} \quad (11)$$

and in order to achieve the loss function in (11) we use the Limited Memory Broyden, Fletcher, Goldfarb, and Shanno (L-BFGS) algorithm, which gives us the optimal values for the model parameters.

The L-BFGS algorithm is known as one of the memory - constrained Quasi – Newton approximation algorithms, and it has gained wide importance in optimizing objective functions due to its ability to deal with the complexities of nonlinear functions and deal with memory constraints. The basic idea of (L-BFGS) is to use the (curvature Information), which represents the second derivative (Second Derivative Information) of the most recent iterations only, to create a Hessian matrix approximation, as the curvature information from the previous iterations is ignored in order to save good storage space [7].

The Hessian matrix describes the curvatures of nonlinear functions, for numerical optimization, and the Hessian matrix is also called the information matrix of the observations, and the inverse of the Hessian matrix represents the variance-covariance of the estimators of the studied function, and is written in short as (H), and can be represented according to the following [8]:

$$H = \begin{bmatrix} \frac{\partial^2 f}{\partial x_1{}^2} & \frac{\partial^2 f}{\partial x_1 \partial x_2} & \cdots & \frac{\partial^2 f}{\partial x_1 \partial x_n} \\ \frac{\partial^2 f}{\partial x_2 \partial x_1} & \frac{\partial^2 f}{\partial x_2{}^2} & \cdots & \frac{\partial^2 f}{\partial x_2 \partial x_n} \\ \vdots & \vdots & \cdots & \vdots \\ \frac{\partial^2 f}{\partial x_n \partial x_1} & \frac{\partial^2 f}{\partial x_n \partial x_2} & \cdots & \frac{\partial^2 f}{\partial x_n{}^2} \end{bmatrix}_{n \times n}. \quad (12)$$

It can also be written at time r as follows [7]:

$$H_r = \nabla^2 l(\vartheta). \quad (13)$$

The estimation steps according to the L-BFGS algorithm are represented in the following form:

1) Gradient Computation, $l(\vartheta)$. It is a multi-parameter function, which is the vector of partial derivatives of the function at time point r that is to be minimized, since $i = 1,2, \ldots, r$, where:

$$\nabla l(\vartheta) = G(r). \quad (14)$$

2) Calculating the update of the approximate (Hessian) Matrix.

The Hessian matrix is updated according to the following relationship:

$$\begin{aligned} H_{r+1} = \left(I - \mathcal{P}_r S_r R_r{}^T\right) H_r \big(I \\ - \mathcal{P}_r R_r O_r{}^T\big). + \mathcal{P}_r S_r S_r{}^T \end{aligned} \quad (15)$$

Where:

- R_r the change in the gradient that calculated according to the following:

$$R_r = G_{r+1} - G_r. \quad (16)$$

- O_r: the change in the parameters that represented according to the following:

$$O_r = \vartheta_{r+1} - \vartheta_r. \quad (17)$$

- $\mathcal{P}_r$: the inverse of the product of the change in the parameters and the change in the gradient is calculated according to the following:

$$\mathcal{P}_r = \frac{1}{R_r{}^T O_r}. \quad (18)$$

The inverse of the approximation of the (Hessian) matrix is calculated through the following:

$$H_r = \mathcal{B}_r{}^{-1}. \quad (19)$$

- $\mathcal{B}_r{}^{-1}$ represents the variance-covariance matrix of the parameter estimates at the point (time r).

3) Calculating (search Direction).

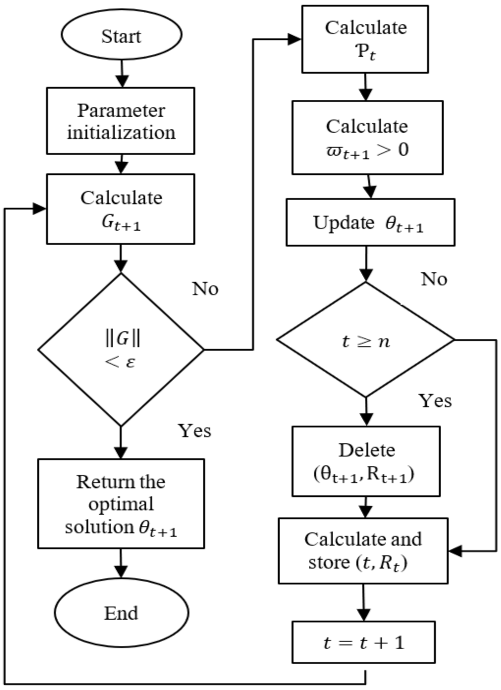

Figure 1: The flow chart of L-BFGS method.

It is a vector that determines the optimal direction for updating the parameters in the optimization process and is calculated using the inverse approximation of the (Hessian) matrix based on the concept of (Limited Memory) and is calculated according to the following formula:

$$\text{Ƥ}_r = -\mathcal{B}_r \nabla l(\vartheta_r). \qquad (20)$$

- Ƥ_r: provides us with information about any parameter that needs to be modified, and it is calculated without the need to store the matrix H_r completely.

4) Updating Parameter ϑ.

The last step is to update the vector of parameters (ϑ) through the following formula:

$$\vartheta_{r+1} = \vartheta_r + \varpi_r \text{Ƥ}_r. \qquad (21)$$

Where: ϖ_r the value of the (Line Search) and is used to determine the optimal step size (step Size) when updating the parameters in the gradient calculation $L(\vartheta)$).

The appropriate λ value of (regulatory limits) is chosen through the cross- validation (C.V) in a way that achieves the restrictions imposed on each function of the Facebook Prophet model, and these steps continue until the loss function in (10) is achieved [7], [9], [10].

The Flow Chart of the (L-BFGS) method is shown in Figure 1 as follows [10]:

3 THE PRACTICAL APPLICATION

The Prophet model was used in the practical application of the data of the studied phenomenon according to a program written in the R language, the study data represents the consumed quantities of electrical energy in Baghdad Governorate measured in (Kw/h) taken on a daily basis for the period from 1/1/2024 to 12/31/2024. This data was obtained from the Ministry of Electricity/Operation and Control Department, and through the following Figure 2 we show the plot of the study data.

When analyzing the time series according to the studied model, its components become clear to us, such as trend, seasonality, and special events. The following Figure 3 shows this.

Figure 3 shows the presence of a general trend that changes with time, in addition to the clarity of the effect of holidays on the time series data, as we notice the presence of sharp peaks and valleys, and it is also clear to us the presence of the weekly seasonal component, as it appears that there is a pattern that changes with the change of days of the week, and the presence of the annual component in the data is also clear, as we notice the repetition of the pattern during the months of the year, with a sharp rise in the summer months and a sharp decline in the winter months.

When applying the L-BFGS algorithm, we obtained the parameter values shown in Table 1:

Table 1: The estimation of the parameters of the Facebook Prophet.

Parameter	Coefficient
Growth rate k	0.2411
offset parameter m	1.0271
δ represents the amount of increase or decrease in the growth rate	[1,1:30] -8.001, 0.6781, -0.2427, …
β represents the vector of parameters of the Fourier Series	[1,1:74] -0.5451, -0.1083, 0.0047, …

The residuals of the time Series are shown in Figure 4.

The predicted values are Presented in Table 2.

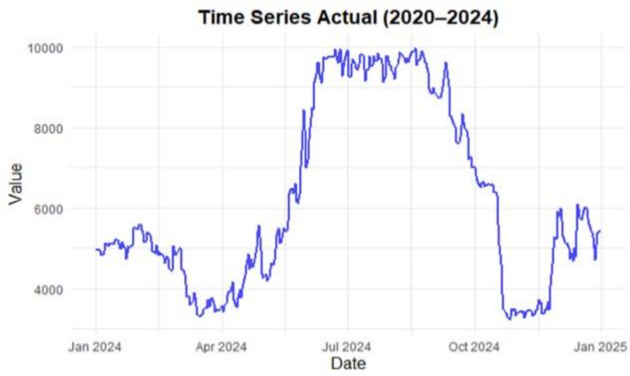

Figure 2: Represents the time series graphically.

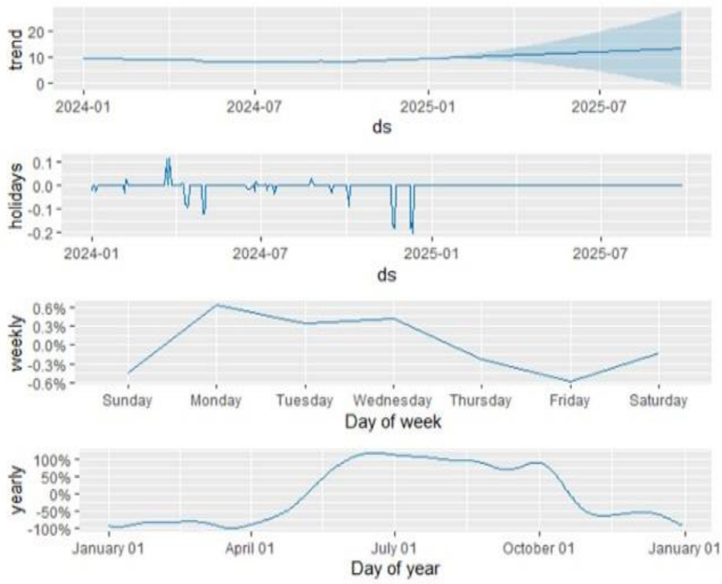

Figure 3: The components of time.

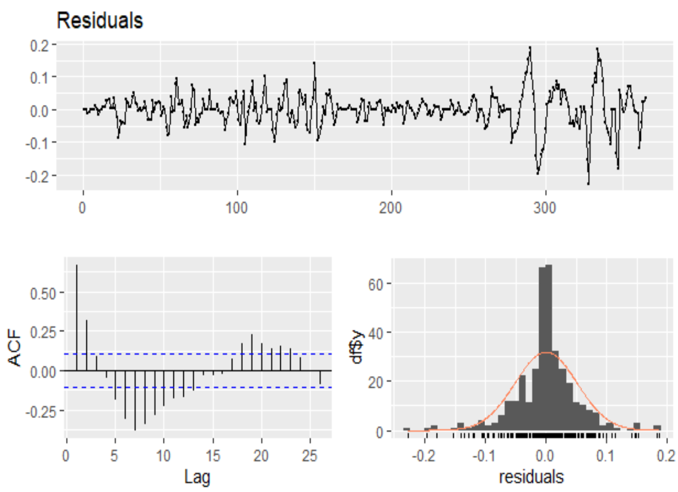

Figure 4: Represents the residuals of the time series, its autocorrelation function, and the shape of its distribution.

Table 2: represents the predictive values according to the FB Prophet model.

Date	Forecast	Date	Forecast
1/1/2025	5285	2/15/2025	11320
1/2/2025	5266	2/16/2025	11476
1/3/2025	5275	2/17/2025	11796
1/4/2025	5337	2/18/2025	11959
1/5/2025	5371	2/19/2025	12166
1/6/2025	5493	2/20/2025	12279
1/7/2025	5553	2/21/2025	12425
1/8/2025	5646	2/22/2025	12662
1/9/2025	5709	2/23/2025	12794
1/10/2025	5801	2/24/2025	13096
1/11/2025	5950	2/25/2025	13209
1/12/2025	6064	2/26/2025	13358
1/13/2025	6273	2/27/2025	13393
1/14/2025	6408	2/28/2025	13454
1/15/2025	6573	3/1/2025	13605
1/16/2025	6695	3/2/2025	13639
1/17/2025	6842	3/3/2025	13850
1/18/2025	7046	3/4/2025	13860
1/19/2025	7198	3/5/2025	13911
1/20/2025	7453	3/6/2025	13851
1/21/2025	7607	3/7/2025	13828
1/22/2025	7787	3/8/2025	13909
1/23/2025	7906	3/9/2025	13885
1/24/2025	8044	3/10/2025	14057
1/25/2025	8239	3/11/2025	14044
1/26/2025	8368	3/12/2025	14092
1/27/2025	8608	3/13/2025	14047
1/28/2025	8728	3/14/2025	14061
1/29/2025	8875	3/15/2025	14202
1/30/2025	8952	3/16/2025	14255
1/31/2025	9052	3/17/2025	14530
2/1/2025	9219	3/18/2025	14633
2/2/2025	9316	3/19/2025	14816
2/3/2025	9544	3/20/2025	14917
2/4/2025	9644	3/21/2025	15092
2/5/2025	9782	3/22/2025	15414
2/6/2025	9850	3/23/2025	15651
2/7/2025	9953	3/24/2025	16139
2/8/2025	10139	3/25/2025	16443
2/9/2025	10255	3/26/2025	16838
2/10/2025	10522	3/27/2025	17137
2/11/2025	10655	3/28/2025	17516
2/12/2025	10834	3/29/2025	18061
2/13/2025	10939	3/30/2025	18497
2/14/2025	11084	3/31/2025	19222

Figure 5 shows the predictions that were reached through the model with the confidence intervals.

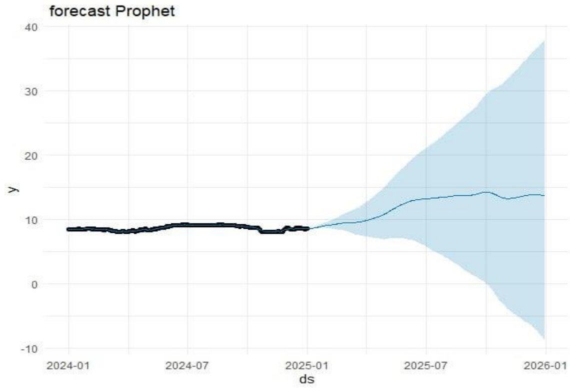

Figure 5: Represents the predictive values of the Facebook Prophet model.

4 CONCLUSIONS

The results confirm that the Facebook Prophet model provides a reliable and effective framework for forecasting electricity consumption in Baghdad Governorate. The model successfully captured the long-term trend, seasonal variations, and the influence of special events, demonstrating its flexibility in handling complex time series data. The close agreement between the predicted and actual values reflects the robustness of the model and its ability to generate accurate forecasts. These findings highlight the potential of Prophet as a valuable analytical tool for supporting electricity demand management and planning. Its structure, which accommodates change points and multiple seasonal patterns, enhances both interpretability and predictive accuracy. Future work may focus on improving the model's performance by integrating Prophet with other statistical or machine learning techniques and incorporating additional explanatory factors to further refine forecasting precision and applicability in real-world energy analysis.

REFERENCES

[1] S. J. Taylor and B. Letham, "Forecasting at scale," PeerJ Preprints, Sep. 27, 2017, doi: 10.7287/peerj.preprints.3190v2.

[2] A.-J. Mäkipää, "Forecasting emergency department arrivals with Facebook Prophet library," Electrical Engineering, 2021.

[3] A. Hasnain, M. Z. Hashmi, B. Nadeem, M. M. Nizamani, and S. U. Bazai, "Ambient PM2.5 prediction based on Prophet forecasting model in Anhui Province, China," in Proceedings of the International Conference on Information Technology and Applications, vol. 614, S. Anwar, A. Ullah, Á. Rocha, and M. J. Sousa, Eds., Lecture Notes in Networks and Systems, vol. 614, Singapore: Springer Nature Singapore, 2023, pp. 27–34, doi: 10.1007/978-981-19-9331-2_3.

[4] Z. Luo et al., "A combined model of SARIMA and Prophet models in forecasting AIDS incidence in Henan Province, China," International Journal of Environmental Research and Public Health, vol. 19, no. 10, p. 5910, May 2022, doi: 10.3390/ijerph19105910.

[5] S. F. Stefenon, L. O. Seman, V. C. Mariani, and L. D. S. Coelho, "Aggregating Prophet and seasonal trend decomposition for time series forecasting of Italian electricity spot prices," Energies, vol. 16, no. 3, p. 1371, Jan. 2023, doi: 10.3390/en16031371.

[6] P. Ma et al., "Multiscale superpixelwise Prophet model for noise-robust feature extraction in hyperspectral images," IEEE Transactions on Geoscience and Remote Sensing, vol. 61, pp. 1–12, 2023, doi: 10.1109/TGRS.2023.3260634.

[7] J. Nocedal and S. J. Wright, Numerical Optimization, 2nd ed., New York, NY: Springer, 2006.

[8] T. Wästerlid, "Application of L-BFGS to a large-scale Poisson MAP estimation," 2012.

[9] D. R. S. Saputro and P. Widyaningsih, "Limited memory Broyden-Fletcher-Goldfarb-Shanno (L-BFGS) method for the parameter estimation on geographically weighted ordinal logistic regression model (GWOLR)," in AIP Conference Proceedings, Yogyakarta, Indonesia, 2017, doi: 10.1063/1.4995124.

[10] X. Lu et al., "Improved reconstruction algorithm of wireless sensor network based on BFGS quasi-Newton method," Electronics, vol. 12, no. 6, p. 1267, Mar. 2023, doi: 10.3390/electronics12061267.