International Journal of Engineering Research and Development
e- ISSN: 2278-067X, p-ISSN: 2278-800X, www.ijerd.com
Volume 22, Issue 5 (May 2026), PP 418-435

Load Forecasting for Improved Performance of Rumuola 33/11 kV Distribution Network using the Artificial Neural Network

[1] Olie, Ijeoma Isreal. [2] Amadi, H.N [3] Igbogidi, O.N.
[1, 2, 3,] Department of Electrical/Electronic Engineering, Faculty of Engineering, Rivers State University, Nigeria

Abstract
To maintain economic activities and improve quality of life, reliable electricity supply plays an important role. In Nigeria, electricity distribution networks such as the Rumuola 33/11 kV feeders faces consistent challenges which involves power outages due to overloading. To tackle the problem, accurate load forecasting that includes the outage records becomes important in improving the grid system performance. This study developed and evaluated a short-term load forecasting of the Rumuola 33/11 kV network using Artificial Neural Network (ANN) model. The unique aspect of this study was the integration of outage indicators into the model. Another Artificial Neural Network model was developed using the same dataset but with the absence of the outage indicators. With the use of MATLAB, the two feedforward Artificial Neural Network models were developed with the aim of identifying the accuracy difference between both ANN models. Thirty-one days of hourly load demand were analyzed together with records of outage events to capture the influence of outage interruptions on load patterns. The ANN model having the outage indicators was benchmarked against linear regression that was trained with the same dataset. The aim was to show the superiority of the ANN model in capturing nonlinear behaviors. These models were checked using Standard error metrics such as Root Mean Square Error (RMSE), Mean Absolute Error (MAE) and coefficient of Determination (R^2). The results shows that the outage-sensitive ANN model had the best performance and achieved RMSE of 2.00017, MAE of 1.5284 and R^2 of 0.8240. It outperformed the ANN model without outage indicators which has RMSE = 4.1178, MAE = 3.0022 and R^2 = 0.2553. It also outperformed Linear regression model having RMSE = 2.2855, MAE = 1.8048 and R^2 = 0.7700. These findings confirms that for a country where outages are part of the grid's constant challenge, outage-sensitive ANN based forecasting can serve as a valuable tool to support decision making for the Nigerian electricity Distribution Companies.

Keywords: *Artificial Neural Network, Short-Term Load Forecasting, Distribution Network, Outage Indicators, Electricity Demand Prediction, MATLAB, Power System Performance, Rumuola Substation, Machine Learning, Load Demand Analysis.*

Date of Submission: 13-05-2026 Date of acceptance: 28-05-2026

I. INTRODUCTION

Load forecasting is a technique in the modern power system and planning. Its aims is to predict future demand based on the past or historical data such as load demand, weather patterns, population density, socio-economic factors such as price of electricity and financial capacity of the consumers. This helps the power system operators plan ahead of time to ensure stability between the load demanded and the load supplied (Arvanitidis *et al.,* 2021).

Multiple studies has discussed different aspects of load forecasting in the light of improving the performance of the power grid system. Akman *et al.* (2023) defined load forecasting as a prediction of future load demand using accepted intelligent models trained based on pattern learnt from historical data. Alahi *et al.* (2022) looked at load forecasting from an angle maximum peak demand. They further stated that load forecast is important for the prevention of overloads and helps in the support of grid infrastructural planning. On the other hand, Awad *et al.* (2020) discussed about the role of load forecasting in smart grid systems. They said that it enables proper optimal scheduling and efficient power management.

The accuracy of forecasting especially in short term is very critical because it supports generation scheduling and energy trading which helps in reducing cost of operation and improve the reliability of the system (Ahmad *et al.,* 2020).

Load forecasting and its complexities have long attracted researchers and especially in the aspect of its accuracy and reliability, because it plays a role in predicting, obtaining, and evaluating future electricity demand and consumption depending on the place, network, and regions (Hammad *et al.,* 2020).

Rodrigues *et al.* (2023) stated that one of the main challenges facing electricity power supply is the problem of a reliable electricity demand forecast. With the help of accurate forecasting, utility companies can predict future demands, thereby knowing how to distribute electrical power to meet the demand needed and satisfy the consumers. The benefits of load forecast range from efficient operation, maintenance to planning of the power grid.

The Rumuola electricity distribution network is a commercial environment with a large population of people due to its proximity to the Trans-Amadi industrial area. There are steady increase in network around the Rumuola distribution network because of the neighboring communities being linked to the national grid as a result of gradual development experienced in recent times (Igbogidi and Dahunsi, 2023). With its huge population density, it is important to acknowledge the importance of load forecasting in that area because wherever there are population, there are steady demand for electrical power, and it is then crucial to employ the use of load forecasting to predict the future load demand in that area to avoid overloading and system instability and faults (Luo *et al.*, 2023).

Load forecasting is needed because it gives future potential demand prediction thereby giving the utility companies the opportunity to plan for the power demand ahead of time.

Samuel *et al.* (2020) mentioned that the accuracy of the load forecast is dependent on the process, methods, and design used to carry out the forecast. For this reason, it is necessary to carry out studies like this to get to know what really works in forecasting electricity demand to overcome challenges faced by the power supply system in Nigeria. There is a need to carry out electricity load forecasting on the Rumuola Distribution network due to its increasing demand and electricity consumption to prevent future losses that can be avoided. And it also prevents undersupply of electrical power, thereby avoiding outages and faults.

II. EXTENT OF PAST WORKS

Awad *et al.* (2024) in their work, they used data both historical load and Temperature (Weather) of 2023, the highest electrical load of the month and data for Temperature per hour to build a Short term Load forecasting Model using ANN. They divided the time series into three groups in which, Learning group has 70%, Verification group has 15% and the testing group has 15%. They trained the network using Levenberg-Marquardt Learning Rule on MATLAB combined with the speed of Gauss-Newton approach and using the Mean Square Error (MSE) to validate the accuracy. It was discovered that the model developed on this research gave an impressive accuracy with estimated error of zero. The Historical Data was gotten from the Egyptian Electric Power Distribution Company.

Dinh *et al.* (2025) proposed the use of ANN for short term load forecasting on a wind farm in Gia Lai providing a single-layer, feed forward Artificial Network was used. The performance for selecting the training algorithm on MATLAB can be decided by the Mean Square Error (MSE). According to the article, they chose the Levenberg-Marquardt Algorithm. MAPE was used between output data curve and the predicted power and the results was a 3.25% while the MAPE error was at 3.07%.

Department of Electrical and Computer Engineering, Dilla University, Ethiopia *et al.* (2021) demonstrated an innovative use of short term load forecast on the Finnish electricity utility at Hawassa City. MAPE was gotten to be 1.529% in their model aster series of tests and tries. During their study historical load data, temperature and holiday season was used to carry out their study. On the result gotten, the actual and forecasted results had little differences due to minimal forecasting fluctuations. The model took account of the period of disconnected time of load. The results gotten were judged by calculating the absolute Error between the exact and predicted value and compared the result between 3 neural networks.

Kumar (2020) also proposed a forecasting model capable of predicting hourly load demand of the Victorian power system by the application of the neural network technique. He employed a series of reliable historical load data of the year 2006 obtained from the official website of the Australian national electricity market. Part of that was carefully used to train the neural network and while other portion was used to verify the level of accuracy of the resulting model. Interest, the model was found to be of a very high level of accuracy using MAPE for evaluation.

Dohen and Odiase (2021) carried out the forecasting of electric load demand within the Ugbowo campus of the University of Benin. Data used was for thirty days, starting from September 1, 2019 to September 30, 2019. The authors used artificial neural network (ANN) modelled in MATLAB in carrying out the work.

Ren *et al.* (2021) developed an approach for sensitivity analysis capable of analyzing the correlation existing between all input variables as well as detecting any variable that is important even though despite the fact that it may have duplicated values. The validation of the proposed approach was from a case study where the heat demand of a particular district network of several buildings within a university environment was carried out. It was found that the approach proposed was able to detect and eliminate unwanted variables, thereby helping the artificial neural network model developed to minimize about 20 percent training time.

Ganthia *et al.* (2023) explores the use of load forecasting on smart grids, electrical architecture. They studied the potential of various ANN transfer algorithm, transfer Function, Learning rate and moment in order to

find the suitable ANN model for developing an accurate grid. Different algorithms have been studied together with different types of neural networks. The study stated that the input and output vector normalization is important because it influences the back propagation's prediction performance.

Panda *et al.* (2023) proposed the use of short term load forecasting for Artificial Neural Network (ANN) which is achieved using Back propagation (BP), cuockro search (CS) and Bat Algorithm. It was noticed that the use of the Bat Algorithm made the work Valuable as it proves a huge use in the research of electricity load forecasting.

Her *et al.* (2022) Conducted a detailed study on the best performance configuration for short term load forecasting using historical data gotten between 2016 to 2018 and they are trained and analyzed to be used to predict future load demands. Using a feed forward ANN for this study and a back propagation technique. they tested the model using MAPE and the results was used for more strict analysis to find out the best ANN configuration for the forecast they talked about the reduction of noise in the data which can lead to overfitting and it usually occurs on complex mode and it was gotten that the best ANN configuration is the Tanh activation function of performance over 1.38%.

Samuel *et al.* (2020) focused on the load forecasting on a 132/33kV transmission sustained in Port Harcourt using ANN for a week ahead forecast using MATLAB software. They obtained a very impressive result using the R-Value as they also compared their results the actual recorded value of the period and their result came really close to the actual data later gotten.

Fakhryza *et al.* (2024) looked into possibilities of ANN for load forecasting for a 150KV power plant in Indonesia, discussing the importance of variables like population, historical data on the accuracy of the forecast. Using Tensor Flow as a python library frame work to develop the ANN model with the Adam optimization algorithm use the weight adjustment. In their study, they collected historical data also power flow, voltage and phase angle from the power system. Now, using 70% of the data for training, 30% for testing and using a subtype of ANN such as a multilayer perceptron (MLP). Their results where compared with traditional methods such as ARIMA and Regression using MAPE and MSE for the comparison and it was seen and noticed that ANN handles data better . They went further to compare their study with other AI techniques such as sum, Random forest (RF) and gradient boosting machines (GBM). The results gotten shows that for short term load forecasting, parameters such as peak and minimum load has a huge direct impact on the forecast and dates from temperatures introduces more complexities and needs more complex training to handle.

Okelola *et al*. (2021) Using data obtained from national control center Osogba, Nigeria, they carried out a short term load forecasting by using Artificial Neural network. Using feed-forward, back-propagation network with not feedback system, they built an ANN system that compared the output with the expected output. It was noticed that the forecast gotten from ANN forecast model and actual load was almost identical. They achieved smoother curve from the ANN forecast which means better load transition. The Root mean square error was gotten to be 0.15% and this solidified the credibility of ANN for the forecast of electricity load.

Alabi *et al.* (2022) proposed long term forecasting using ANN on a 33kV substations in Abuja, Nigeria. Collecting historical data of direction of 4 years for this research on a wealthy demand bases. They analyzed the result using mean square (MSE). The ANN performed well due to the large data set fed to it was discovered that among the substations, the highest in demand is on the APO feeder and the lowest being Akwanga feeder. It was indicated that the paper can act as a guide for network expansion.

Tarmanini *et al.* (2023) developed an approach using ANN and ARIMA to forecast short term load. It was carried out using MATLAB. The historical data was gotten by recording 709 individual's household level during 12865h for 2009 to 2010 in Ireland. The result gotten between ANN and ARIMA was compared. It was discovered that the model (ANN and ARIMA) was dependent on season as the average error on both model was a higher in summer than winter. ANN got less error compared to ARIMA. ARIMA needs sufficient historical data to train for a better accuracy it was generally seen that ANN performed better that ARIMA.

Velasco *et al.* (2022) focused on conduction a performance analysis of 36 ANN models for hours ahead forecast. The models was made up of 9 input and 1 output neural. The historical data gotten was within the year of 2012 to 2014 in an urbanized city, a multilayer perceptron neural network and feed forward architecture during this study. Each layer on the ANN model has there corresponding duties and the neurons in the input layers where determine by conditions affecting, the next hour prediction. It was noticed and found that electricity load normalization alone isn't enough to produce a converging result. during this result and well when using it on an hour ahead due to the Gaussian Activating as errors arises when applying to resilient propagation, back propagation and quick propagation. It was noted that performance analysis proved to be effective due to its low value on MAPE.

III. MATERIALS AND METHOD

2.1 Materials Used

The data needed for this study includes the hourly peak load demand readings on the two 33kV feeder (Refinery Line 1 and Rumuola Feeder 3) of the Rumuola substation of the 31 days of August 2025. It was collected from the Authorities of the Port Harcourt Electricity Distribution Company (PHEDC) Rumuola Branch, Port Harcourt. The data collected was an hourly load consumption reading from the control room at the Rumuola substation. The data was made up of eight 11kV feeder Mega-Watts readings which was summed up to get the full consumer demand load value of the two incoming 33kV line. This was done because it gives the actual load demand of the 33kV line from the consumers, neglecting power loss on the 33kV lines and Step-down transformer loss.

The weather data collected are the hourly temperature and humidity value of the entire Port Harcourt metropolitan area. The data was gotten from a reliable weather source online: visualcrossing.com, and the site specializes in collecting data from multiple stations, and models. The reason behind this method was due to the unavailability of weather data within the areas Rumuola substation supply electrical power to (a limitation).

Software used involves:

Matrix Laboratory (MATLAB R2023a) software: Neural Network Toolbox and regression Leaner App.

Hardware used involved:

ASUS Vivo book Laptop, Core I3- 10110U, 20GB of ram, 265GB SSD

2.2 Method used

This study confirms to a data driven predictive (forecasting) model using an Artificial Neural Network (ANN) for load forecasting for the Rumuola 33/11kV distribution network. Here are the methods used for this study:

1. Artificial Neural Network model using Full and partial outage indicators.
2. Artificial Neural Network model without using outage indicators.
3. Linear Regression Model fed with the data containing outage indicators.
4. Performance Metrics (Root Mean Square Error, Mean Absolute Error and Coefficient of Determination).

2.2.1 Data Collection

(i) **Historical Load Data**: The Historical load Data was collected from the Port Harcourt Electricity Distribution Company (PHEDC), Rumuola substation. The importance of this data to this study cannot be underestimated as it plays a huge role in Predicting, training and testing of the ANN model. Within the data, there were abnormal event such as full outage events and Partial outage event. Partial outage events are those events that occurs when few but not all the 11kV feeder connected to one or both of the 33kV feeders goes offline due to circumstances such as emergency maintenance, load shedding, faults. This creates an instable value of power consumed by the customer. This affects the dataset negatively as the Load demand for electrical power during that time wasn't completely known.

Full outage event occurs when there is a complete blackout on the eight 11kV feeders that are connected to the two 33kV incoming Rumuola feeders. Within that time, no value of load demand was recording this created a gap in the dataset that cannot be ignored.

If ignored and used as it is, the ANN model might learn an inaccurate load demand pattern. The model will assume that during that time, power was not needed by the consumers or there was a drop of load demand from the consumers which is an inaccurate prediction.

To tackle the issue of Full outage event, other studies used the approach of cleaning the data by avoiding the outage zero values, some used an outage flag (Binary indictor) where there created a new column of data to track the outages. This way, their model learns from the outage event as well as the historical load demand.

This study will be looking into using both full outage flag and partial outage flag in my dataset together with day of the week flag and weather to predict the upcoming load demand.

The Outage flags/indicator will be removed and the model will be trained without the Outage indictor. The accuracy will be compared to prove the importance of Outage indictors in predicting a good model.

Historical Weather Data: The role of the weather data for this study isn't tied to any Equation or formulas. It is also one of the inputs fed to the Artificial Neural Network together with the day of the month and the days of the week. Their main purpose is aimed at helping the ANN model understand the patterns from the weather which influences human decisions. The data was collected from a reliable data site (visualcrossing.com). The site collects weather data from stations and Government archives around the world (it does not make its own weather values). Though the data was collected from a recognized historical weather site, it is still a limitation as specific weather dataset wasn't available for the locations Rumuola substation supply power to but the data gotten shows the weather value of the entire Port Harcourt city. The effect of Temperature and Humidity Weather value to the accuracy of the ANN model cannot be overlooked. It helps the model recognize weather patterns which affect the

consumer decision to consume electrical power. For example: during rainy time, consumer could decide to avoid inductive loads such as fans and Air conditional and stay indoors to enjoy the natural weather. While during Hot weather, the consumers can use AC and fans, they can use the refrigerator more during this time.

2.2.2 Data Description & Preprocessing

This Part of the study focuses on the process of preparing the dataset before it can be fed to the Model, and the justification behind the choices of Data properties used to build the ANN model.

2.2.2.1 The Dataset

The data was arranged and saved using the Microsoft Excel software.

The Dataset is made up of 13 columns. They are sets of values arranged in such a way that will make the ANN model understand what pattern to look for.

Table 3.1 show the few rows out of the 744 rows or samples that makes up the full Dataset. The Full dataset is presented in the Appendix section of this study.

The dataset is made up of temperature and humidity reading during the time of the load demand. The load demand column is the electrical power consumed during the specific hour corresponding to the hour and day of the week. Other columns includes:

i. **Day of the Week/Hourly Indicator:** The ANN model isn't able to recognize text/words because it was only built to work with numbers and find patterns with numbers. So, days of the week such as Monday and Tuesday won't be recognized. The day of the week indicator was employed in the dataset to solve the problem. It allows the ANN model understand the Dataset and find specific weekly patterns that affects the historical load data. The Day of the week indicator was implemented by using the same binary system used for the Outage indicator, i.e. (0, 1). This study used 7 columns to identify the 7 days of the week, from Monday to Sunday. This method was used because the month of August consists of 4 weeks and the model needs to track weekly load consumption patterns in other to predict a more accurate value. For each of the columns, '1' was used to tell the model that a specific load consumed belong to a specific day of the month. While '0' tells the model that specific load demand consumed doesn't belong to a specific day of the week. This study also used weekend indicators, predicting load based on weekend load consumed. Because it was noticed that the load patterns spotted during the week is different from the weekend load pattern.

ii. **Outage Indicator (Full and Partial):**

   a. **Full outage indicator:** These are indicators used to indicate when there is a complete outage on the entire eight 11 Kv feeders of the Rumuola station. A complete blackout. Using a binary indicator where (0) indicates that there is no outage and (1) indicates a complete outage.

   b. **Partial Outage Indicator:** These are indicators used to represent partial outages on the feeders. It is used to show when 1 or more 11Kv feeders connected to the Rumuola substation is experiencing an outage event. It still uses the binary indicators to show when a partial outage occurs.

From the dataset sample shown in Table 3.1, the outage indicators (Full Outage and Partial Outage) columns were used to train the ANN model to understand the outage patterns. Without these columns, the model will only predict the load based on the hourly, temperature humidity, and weekly patterns, which can affect the accuracy of the model. For this study, we compared the accuracy level of the ANN model with and without the Outage indicators. This will further prove the relevance of adding the outage indicator to the improved performance of the ANN model.

Table 1 Arranged Dataset Structure

	Hour	Temp	Humidity	Mon	Tues	Wed	Thurs	Fri	Sat	Sun	Full_Outage	Partial_Outage	Load
0	24	94.15	0	0	0	0	1	0	0	0	0	17.6
1	24	94.15	0	0	0	0	1	0	0	0	0	15.7
2	24	94.15	0	0	0	0	1	0	0	0	0	14.6
3	24	100	0	0	0	0	1	0	0	0	0	15.5
4	23.9	94.17	0	0	0	0	1	0	0	0	1	13.5
5	24	94.15	0	0	0	0	1	0	0	0	1	12.7
6	24	94.15	0	0	0	0	1	0	0	0	1	10.6
7	24	94.15	0	0	0	0	1	0	0	0	0	19.2

8	23.9	99.98	0	0	0	0	1	0	0	0	0	19.3
9	25	94.19	0	0	0	0	1	0	0	0	0	19.1
10	25	94.17	0	0	0	0	1	0	0	0	1	3.5
11	26	88.76	0	0	0	0	1	0	0	0	1	11.5
12	27	83.67	0	0	0	0	1	0	0	0	0	18.8
13	27	83.66	0	0	0	0	1	0	0	0	0	18
14	27	83.67	0	0	0	0	1	0	0	0	0	17.8
15	27	83.67	0	0	0	0	1	0	0	0	0	21.9
16	27	78.79	0	0	0	0	1	0	0	0	0	21.5
17	26	83.56	0	0	0	0	1	0	0	0	0	18.1
18	25.9	88.77	0	0	0	0	1	0	0	0	1	11.4
19	25	94.18	0	0	0	0	1	0	0	0	1	11.9
20	24	94.15	0	0	0	0	1	0	0	0	1	11.5
21	24	94.15	0	0	0	0	1	0	0	0	1	10.4
22	24.1	93.28	0	0	0	0	1	0	0	0	1	12.8
23	24	94.15	0	0	0	0	1	0	0	0	0	14.7
0	24	94.15	0	0	0	0	0	1	0	0	0	13.8
1	23.9	94.14	0	0	0	0	0	1	0	0	0	16.7
2	24	94.15	0	0	0	0	0	1	0	0	0	15.4
3	24.1	95.87	0	0	0	0	0	1	0	0	0	15.2
4	24	95.87	0	0	0	0	0	1	0	0	0	13.5
5	24.1	94.72	0	0	0	0	0	1	0	0	1	9
6	23	94.11	0	0	0	0	0	1	0	0	1	12.8
7	23.9	94.2	0	0	0	0	0	1	0	0	1	10.2
8	24	99.99	0	0	0	0	0	1	0	1	1	0
9	24	99.97	0	0	0	0	0	1	0	0	1	11.8
10	26.4	87	0	0	0	0	0	1	0	0	1	12.2
11	27.8	84.77	0	0	0	0	0	1	0	0	1	12.6
12	29	74.46	0	0	0	0	0	1	0	0	1	12.1
13	29.9	70.35	0	0	0	0	0	1	0	0	1	12.9
14	29	74.47	0	0	0	0	0	1	0	1	1	0
15	31	66.37	0	0	0	0	0	1	0	1	1	0
16	29	79.05	0	0	0	0	0	1	0	0	1	17
17	27	83.67	0	0	0	0	0	1	0	0	1	10.7
18	26	94.21	0	0	0	0	0	1	0	0	0	17.8
19	26.1	92.01	0	0	0	0	0	1	0	0	1	17.3
20	26	94.23	0	0	0	0	0	1	0	0	1	14.3
21	25.3	91.97	0	0	0	0	0	1	0	0	1	12.7
22	25	94.19	0	0	0	0	0	1	0	0	1	12.5
23	25	94.19	0	0	0	0	0	1	0	0	0	17.1

Table 2 briefly describes each columns (Variable Name) of the data set. It shows the nature of the data they represent, their units and where they were gotten from. It shows the full description of the dataset.

Table 2: The Description of the Dataset Variable

Variable Name	Variable Type	Source
Load Demand	Numerical	Rumuola Distribution Station
Temperature	Numerical	Online-Archives (visualcrossing.com)
Humidity	Numerical	Online-Archives (visualcrossing.com)
Partial Outage	Binary Value (0, 1)	Rumuola load demand record
Full Outage	Binary Value (0, 1)	Rumuola load demand record
Day of week	Binary Value (0, 1)	Load demand timestamp
Hourly	Numerical	Load dataset

2.2.2.2 Predictor and Response

Figure 1 illustrates the division of the dataset into the predictors and the responses.

i. **Predictors:** These are the Independent variable on the ANN model. In other studies, they are referred to as INPUTS. They are like a clues the model use to predict the future load demand. For example, the model developed in this study uses clues from hours, temperature, humidity levels, days of the week, full and partial outage data of past dates to try to predict future load demand.

ii. **Responses:** These are the dependent variables which values are what we are trying to predict. Other studies use the preferred to it as TARGETS. It's important is classified as most significant because with without it, the model becomes blind.

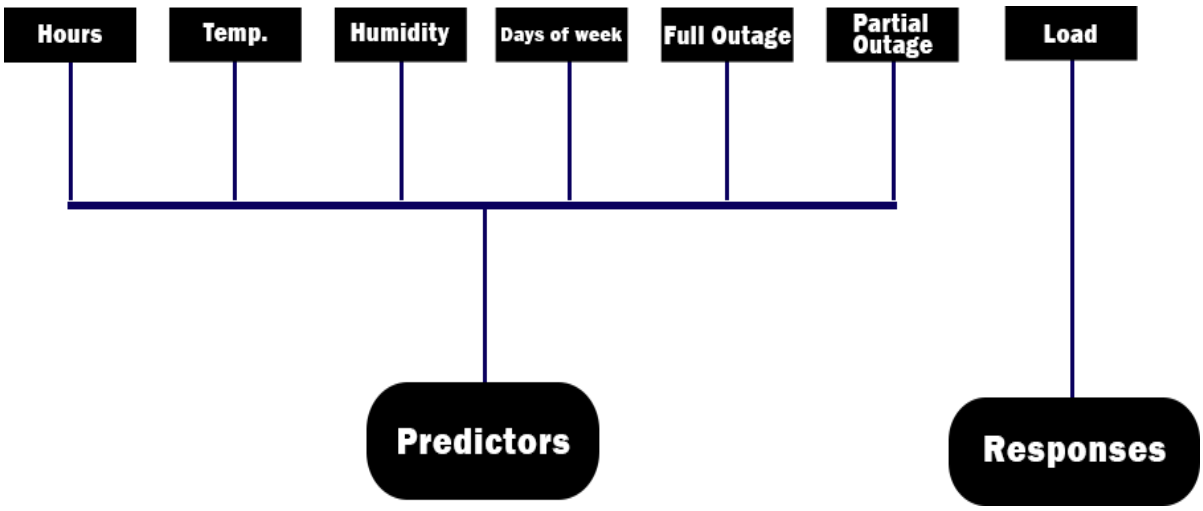


Figure 1 Dataset – Predictor and Response

2.2.3 Model Development – Artificial Neural Network

This study confirms to a data driven predictive (forecasting) model using an Artificial Neural Network (ANN) technique for load forecasting for the Rumuola 33/11kV distribution network.

The concept of Artificial Neural Network was gotten from the idea of developing a system or model that can learn like the human brain.

The regression problem such as the load forecasting that needs a continuous Value for its output cannot be easily resolved through the help of conventional mathematical models but with the help of a model such as the Artificial Neural network. The ability of the model to solve problems lies in its ability to learn (Training).

The training of the Artificial Neural Network cannot be underestimated as it is necessary for the development of a strong and reliable system. Without Training, the model is like an empty box, but with the right training and the right data, the model can solve complex mathematical problems, especially a continuous problem such as the electricity load forecast.

The major benefit of using Artificial Neural Network is its ability to cope with different types of situations (Noise), and handling of Errors.

The Artificial Neural Network contains Neurons just as the Neurons found in the Human Brain and it's grouped into Layers.

Figure 2 illustrates the basic representation of a Neural Network contain basic layers, which are:

i. Input Layer: this Layer contains nodes that moves input signals (x1,x2,x3,x4) and (w1,w2,w3,w4)

ii. Hidden layer: it's the layer between the input and output layer. It's the layer where the complex calculation takes place because it contains values of artificial neurons that input signals, process them and manipulates it using weights and bias and activation function. It is also responsible for the handling nonlinear relationships.

iii. Output layer: it is tasked with collecting data processed by the hidden layer and gives an output value and for electricity load forecasting, the value is usually continuous.

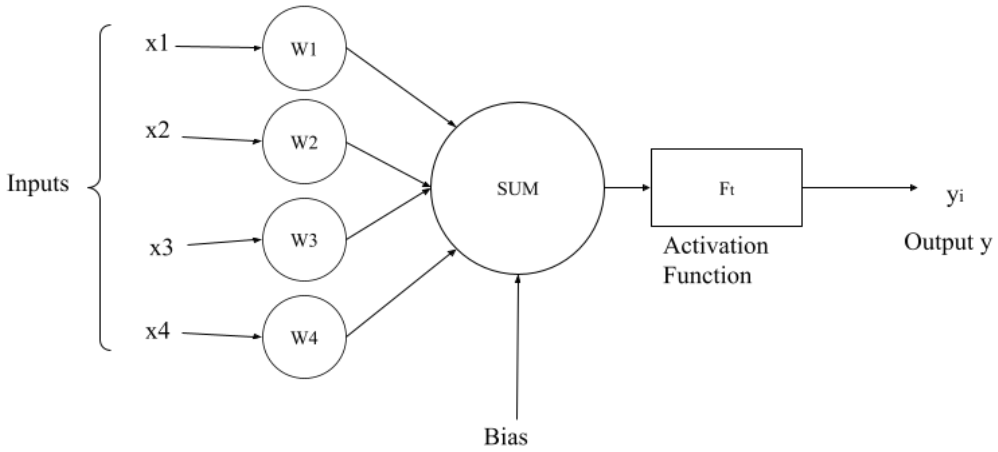


Figure 2 Basic Representation of a Neural Network

2.2.3.1 Training of the Artificial Neural Network

ANN can be trained by modifying its internal parameters such as the weight and biases with the goal of increasing its accuracy in prediction, minimizing errors and performs well on unseen data. The Network learns from data because it contains patterns, relationships and trends. When a neural network is trained properly, for a particular task, the training process is usually carried out by identifying the differences between the network's processed output which is known as the predicted value, and a targeted output.

The benefits of training a Neural Network ranges from continuous learning without being supervised, handling of complex and non-linear relationships and improved performance over time because of its adaptability feature.

2.2.3.2 Mathematical Model for ANN

STEP 1: Arrange the input data as Predictor and Response (Input Layer)

The input sent to the Neuron:

$$Input\ (x) = \begin{bmatrix} x1 \\ x2 \\ x3 \\ \dots \\ xn \end{bmatrix} \quad (1)$$

Where:

X1, x2, x3, xn are the input received by the neurons (Past load values, Temperatures, Humidity, Day of the Week, full outage indicator and Partial Load indicator).

Input (x) is the total input

STEP 2: Setup the weight and Bias (Hidden Layer)

The weight and biases of the ANN structure is given as:

$$w = [w1\ \ w2\ \ w3], [b] \quad (2)$$

Where:

w1, w2, w3 are the assigned weights matrix

w being the total weight

b is the bias vector or the Neuron Threshold.

The neuron sums al the signals it receives, with each signal being multiplied by its associated weight on the connection. The weight sum of the output node is giving as:

$$F = \left([w1\ \ w2\ \ w3] \times \begin{bmatrix} x1 \\ x2 \\ x3 \end{bmatrix} \right) + b \quad (3)$$

$$F = \big((w1 \times x1) + (w2 \times x2) + (w3 \times x3)\big) + b \quad (4)$$

$$F = (WnXn + \mathrm{b}) \quad (5)$$

Introducing the activation Function, we will get:

$H = \emptyset(WnXn + \mathrm{b})$ (6)

$\mathrm{H} = \emptyset(F)$ (7)

Where:

Wn is the total number of weight
Xn is the total number of inputs feed to the neuron
$\emptyset$ Is the introduced activation function
F is the weighted sum of the output node
H is the output

STEP 3: Calculating the Output (Output Layer)

From Eqn. (6) and (7) we can expand the equation by using Tan Sigmoid Function as the activation function and we will have:

$\emptyset(\mathrm{H}) = \left(\frac{1}{1+e^{-y}}\right)$ (8)

$\emptyset(\mathrm{H}) = \left(\frac{1}{1+e^{-((w_1 \times x_1)+(w_2 \times x_2)+(w_3 \times x_3)+b)}}\right)$ (9)

Where:

E is the error in node
y is the output that passes through a sigmoid transfer function that is normally non-linear to give the final output.

2.2.3.3 Neural Network Fitting App

The ANN model was built using the Neural Network Fitting App (nftool) on MATLAB software. The ANN model was used for this study due to its ability to handle complex patterns (non-linear) as well as linear patterns, unlike the traditional model that struggles with complex patterns. It has the ability to learn and comprehend complex data, provided the full data is numerical. The benefits of using ANN for load forecasting lies with its ability to learn and memorize data which it was initially fed with, without the operator providing instructions. It uses back propagation for its pattern and relationship learning because it continuously adjust its weight once it learns something new and this is beneficial because it reduces errors in the process.

The features of the ANN model that makes it stand out from other model is its ability to work with very noisy, incomplete or massive data. It still can provide an output with what it was fed it, but the quality of the output will depend on the quality of data it was fed with.

2.2.3.4 Splitting the data

To split the data, we trained the models with 70% of the data, because previous studies has shown that training the model with majority of the dataset helps the model adjust the weight and Bias of the network. For validating the data, we used 15% as its purpose is to tune evaluate the performance of the model based on what it has learnt. We used the remaining 15% of the data for testing. The testing is important as it gives the overall performance of the model after the training is complete.

2.2.3.5 Architectural Design

Figure 3 illustrate the workflow of the two Artificial Neural Network forecasting models and the decisions taken. It starts by collecting Previous Load data (1st August, 2025 to 31st August, 2025) from the Rumuola distribution substation and the weather data specific to the Port Harcourt City. Now, the data collected was arranged, creating a dataset of specific columns aimed to improve the accuracy of the model.
The ANN model optimization technique used was set to Levenberg-Marquardt with the number of hidden layers and data was shared into Training, Validation and Testing.

Two ANN models will be trained to get the best possible accuracy. The first Model will be trained with dataset that contain the full and partial outage Indicators/flags, and the next model will be trained without those outage indicators. As said previously, it is used to study the importance outage indictors for the improved performance of the ANN model.

The two model will be evaluated using the Mean Absolute Error (MAE), coefficient of Determination R^2, and Root Mean Square Error (RMSE) and if the Model with outage flag performs better, it will be compare with Linear Regression model as it will prove the usefulness of Artificial Neural Network model in Load forecasting compared to other traditional methods. The ANN model will be used to predict the Next 24 hours of Load demand.

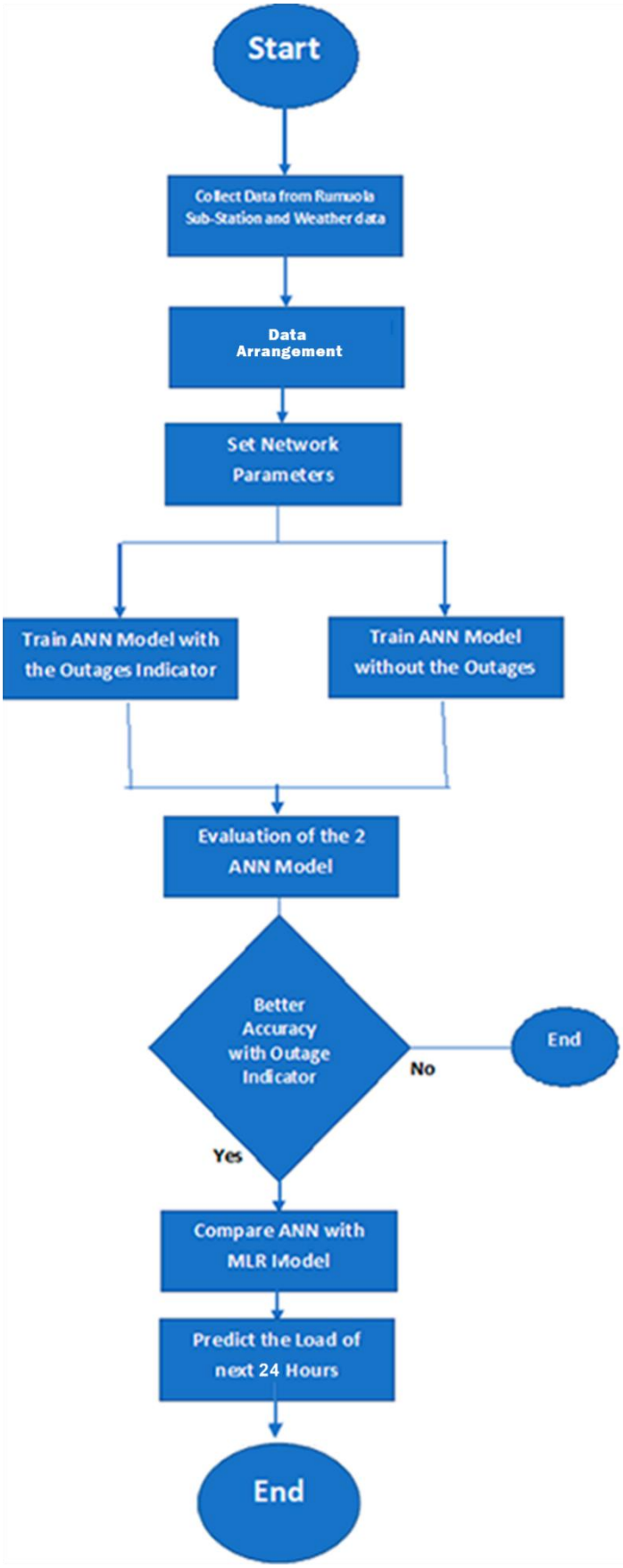


Figure 3 The Workflow of the Two ANN models

2.2.3.6 The Model Training

The training was conducted using the MATLAB Neural Network Fitting App. The App, also called 'nftool' is a no code app that helps to build as well as validating and visualizing regression related problem such as load forecasting.

How it works consist of taking the input data, allows the users to set the number of hidden layers for feedforward network and for the optimization methods it was designed to train, it only allows the user choose from 3 types of optimizers. They are Levenberg-Marquardt optimizer, Bayesian Regularization and Scaled Conjugate Gradient. The user has full control over the data is to be spilt (Training, Validation and Testing). The main benefit of using this app lies in its quick and efficient way of deciding if the dataset it was given was a good one or not. It gives you script as the training the model. This scripts can be modified later to enhance the ANN model. The down side to this app, it hides a lot of what is really going on in the background and the risk of overfitting.

Figure 4 explains the ANN network. It shows that the model was fed with 12 column inputs which is made up of the values from the hours, temperature, humidity, and the 7 days of a week, full outage and partial outage. The model is made up of one hidden layer with 10 Neurons, while the output part contains the predicted Load.

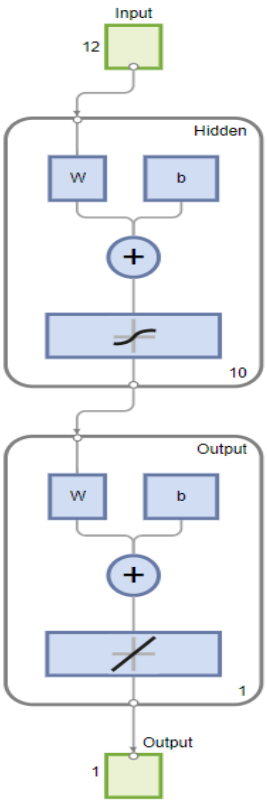


Figure 4 The ANN Network

2.2.3.7 The Optimizer used

The Levengberg-Marquardt optimizer is an algorithm used on the fitting app for the purpose of solving non-linear least square related problems and reducing the Mean Square Error (MSE) in the network. It is the fastest optimizers among the 3 optimizers that the fitting app uses. It is specially designed to solve regression problem such as load forecasting. Its main work is to adjust the weight and bias of the ANN model to get the best possible result, automatically reducing errors.

2.2.3.8 The Neurons

The neurons used for this study are 10 Neurons. The decisions that lead to this choice are:

(i) 10 Neurons, 15 Neurons, 20 Neurons and 25 Neurons were used to test the model but the difference in accuracy wasn't noticeable.

(ii) 10 Neurons has been used in past studies because it is usually the default number of neural per hidden layer, it works well with Levenberg-Marquardt optimization algorithm and it usually lead to good fitting

performance. Its training is usually faster compared to using larger numbers of neurons which can over complicate the ANN model.

(iii) With just 10 Neurons, (744 Rows and 13 columns) 12 inputs and 1 output of data, it balances the ANN model, and detect the existing nonlinear relationship, avoiding under fitting due to fewer Neurons and Overfitting due to excessive number of Neurons per hidden layer.

2.2.3.9 Activation Function

The Activation Function used by the hidden layer is the Hyperbolic Tangent Sigmoid. It is the default Activation function on the MATLAB Neural Network Fitting App. It helps in exploring nonlinear relationships of the model. For the Output layer, Purelin was used as an Activation function.

Table 3 summarized the configuration details of the ANN models. The ANN model with the outage indicators and the ANN model without outage indicator used the same configurations.

Table 3 Configuration Details of the Model

Configuration of the Model	Details
Optimizer/Trainer	Levenberg-Marquardt
Number of Neurons (Hidden Layer)	10
Number of Hidden Layers	1
Activation Function (Hidden Layers)	Tansig (Hyperbolic Tangent Sigmoid)
Activation Function (Output Layer)	Purelin
Training Data	70%
Validation Data	15%
Testing Data	15%

2.2.4 Evaluation of the Model

Evaluating the model is an important part of this study as it helps in determining the performance and reliability of the ANN model for predicting Load. For this study, Mean Square Error (MSE), Root Mean Square Error (RMSE) and Coefficient of Regression where used to evaluate the ANN model. These evaluation Metrics where in this study to show the difference in accuracy between the ANN model and the Multi-Linear Regression Technique and to show the importance of adding outage values for the improvement of the ANN model.

2.2.4.1 Mean Absolute Error

This study made use of the Mean Absolute error (MSE) as one of its evaluation metrics because of its ability to work with any data size. It is not reluctant with errors. It collects errors by measuring the difference between the actual values and the predicted values, then it finds the mean of the total errors it collects. It's important because it helps us determine the average error the model has made (Kachalla *et al.* 2024).

$$\mathbf{MAE} = \frac{1}{N}\sum_{i=1}^{N}(\boldsymbol{Actual\ Values - Predicted\ Values}) \quad (10)$$

Where: N: Number of samples the model predicted

2.2.4.2 Root Mean Square Error

The Root Mean Square Error (RMSE) is useful because it helps to show how a model differs in performance from another as well as show how reliable a model really is. It is a standard metrics for researches because of its simplicity. It takes the difference between the actual values from the target and the predicted value just like the Mean Absolute error, followed by squaring each error values before finding the mean of the value. The square root of the mean is taken at the end (Roman-Portabales *et al.,* 2021).

$$\boldsymbol{RMSE} = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(\boldsymbol{Actual\ Values - predicted\ Values})^2} \quad (11)$$

2.2.4.3 Coefficient of Determination

This matric is used to show how well the model can predict the actual value of the load. It takes the error of the model value and divides it by the average guess error. Its resultant value ranges from 0 to 1. When the error of determination (R^2) gives a value close to 1, it shows that the model is really good at predicting but when it shows a value close to zero, it indicates that the model isn't impressive when it comes to predicting (Saglam *et al.,* 2022).

$$R2 = 1 - \frac{Error\ of\ the\ model}{Average\ Guess\ Error} \quad (12)$$

2.5 Rumuola Distribution Substation

The Rumuola distribution Substation is one of the other substations owned by the Port Harcourt Distribution Company (PHEDC) and it's supplied by two 33kV lines, which are the Refinery Line 1 and Rumuola Feeder 3 directly from the Port Harcourt Main sub Transmission Station. All four power Transformer inside the Rumuola Substation are rated at 15MVA, 33/11kV each and uses a 33kV auto reclose breaker. using eight 11kV load feeder feeding New GRA, Barracks, Bori Camp, Shell Industrial, Rumuokwuta, Omerelu, Rumuomoi and Waterlines. Figure 3.5 shows the line diagram of the Rumuola substation 33/11kV network.

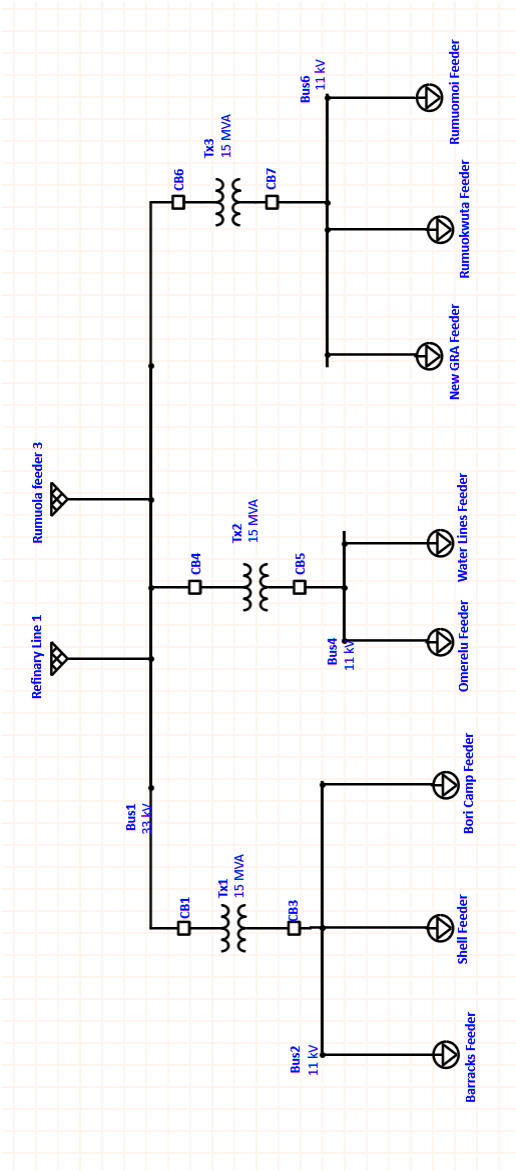


Figure 5: Single Line Diagram of Rumuola Station

IV. RESULTS AND DISCUSSION

3.1.1 Performance Results

Table 4.1 presents the performance results of the model using the Mean Absolute Error (MAE), Root Mean Square Error (RMSE), and Coefficient of determination. It shows the performance of the training, validation and testing. It also presents the overall model performance.

From the table it was noticed that for the R^2, the model do explain much of the variability as it has a high value of 0.8240 which indicates that it was able to explain only 82% of the variance of the actual historical load data. The RMSE has an overall error difference of 2.0017 MW which is good a result because the model forecasted

over 744 hours of load. The prediction was imperfect with a difference of 1.5284 units according to the MAE evaluation.

Table 4 Performance Metrics of the Model (Without Outage Indictor)

	MAE	RMSE	R^2
Training	1.4575	1.8907	0.8484
Validation	1.6470	2.0228	0.7521
Testing	1.7387	2.4345	0.7512
Overall Model	1.5284	2.0017	0.8240

Figure 4 illustrates the Bar Chart visualization of the Evaluation Metric and from the chart, it was seen that the RMSE for the model testing proves to have the highest error while the model training has the lowest RMSE value. This shows that the model performed better on the training than on the testing. The same is to be said for the R^2 as the Training shows better results than the testing. The Bar chat illustrates the table 4.1 as it shows the performance of each part of the model.

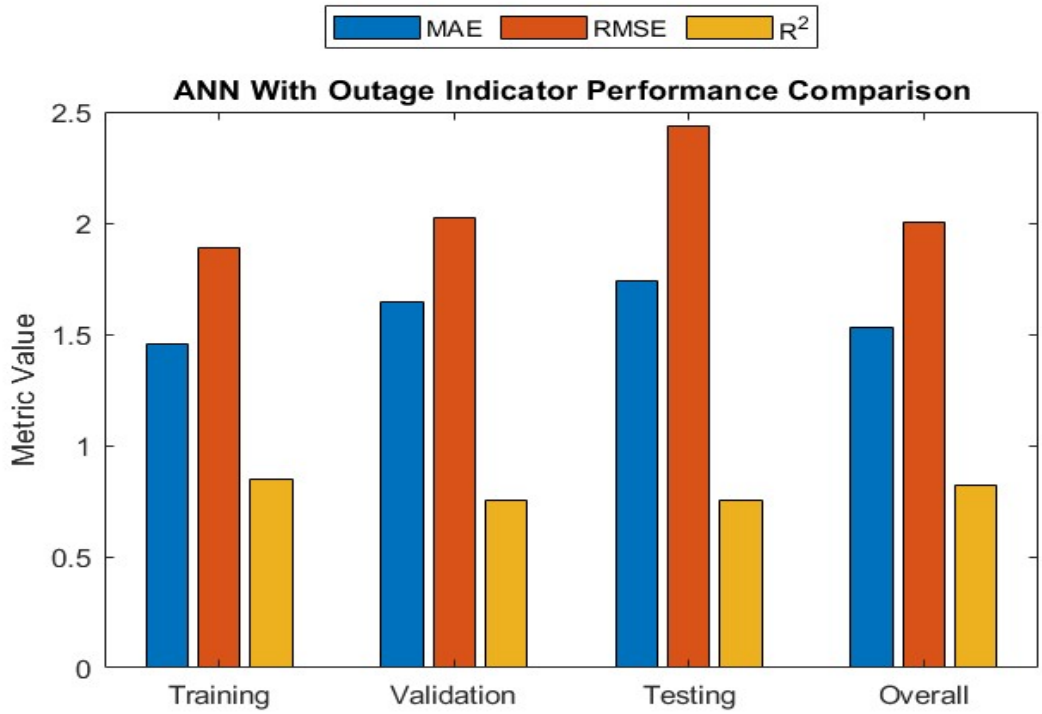


Figure 6 Bar Chart of Evaluation Metric

The Figure 6 illustrates the scatter plot of the general performance of the model and it can be seen that training, validation and testing values are almost cluster at 45 degree which indicates that they follow a clear trend and the model was not overfitting. This shows that as the output follows the patterns of the historical load data (Target). This shows that this model understands the relationship of the data it was fed with.

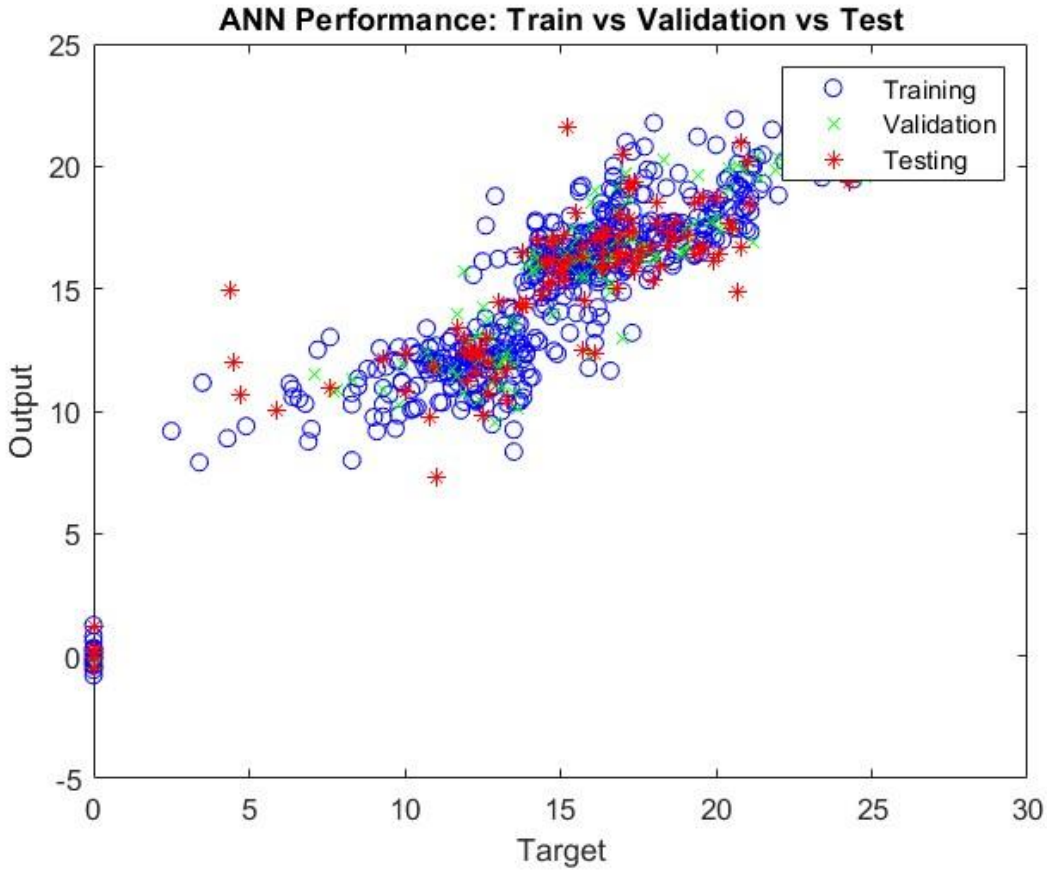


Figure 7 Scatter Plot of the ANN Performance

Comparative Analysis of Model Performance

This subsections talks about the evaluations of the results between the ANN model (with the outage indicator) and ANN model (without outage indicator). It also evaluates the results between the ANN model (with outage indicator) and Linear Regression (with outage flags). This is to show how ANN handles load forecasting better when using outage indicator compared with ANN model that was trained without outage indicators. It also shows how complex the model is in handling non-linear relationships compared with linear regression.

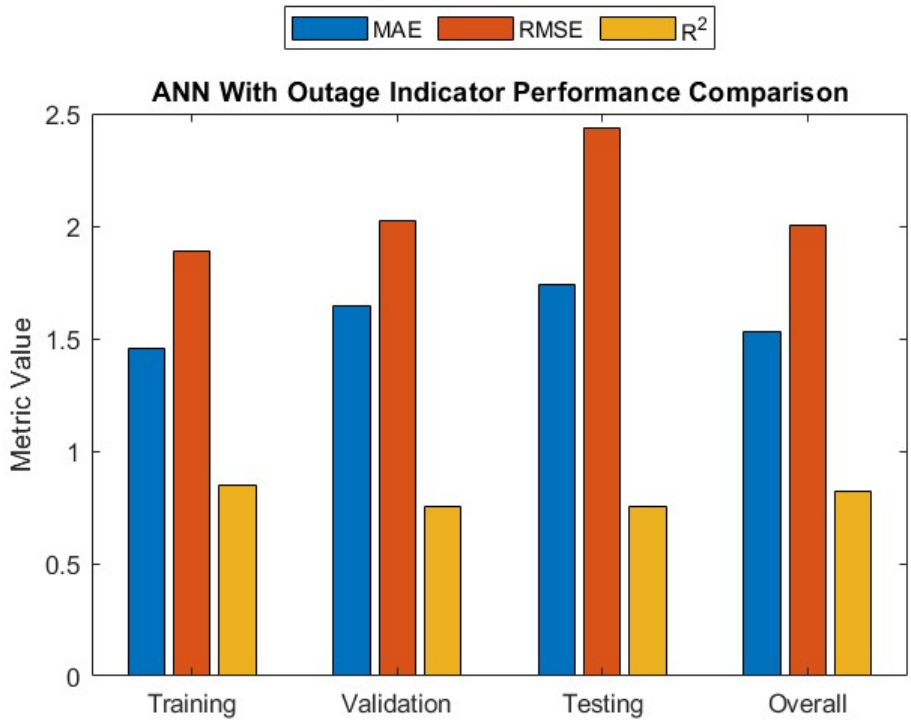
3.2.1 The ANN Models (with outage indicators Vs. without Outage indicators)

Both ANN models where configured the same, they both contain 1 hidden layers with 10 Neurons. Their training functions are also alike as 70% of the data was used for training, 15% of the data was used for Validation and the remaining 15% was used for testing of the model. The table 4.4 shows compared performance results between the two ANN models. This is to show the significant improvement of the model when outage indicators are added as an input that is to be fed to the model. It was seen that the model with the outage indicator performed a lot better. It was noticed the model without the outage flag has twice as much error compared to the model with the outage flag when using the Root mean Square Error (RMSE) and the Mean Absolute Error (MAE). For the coefficient of determination (R^2), due to the presence of the outage indicator, the model was able to explain 82% of the variance while without the model, it can only explain just 25% which is very poor performance. They both had the same configuration but absolutely different result due to the difference in outage indicator.

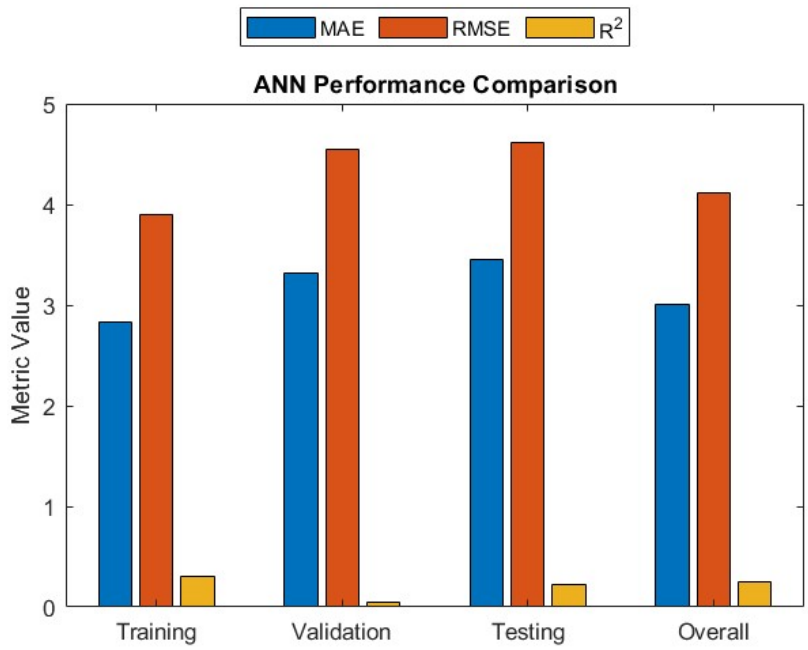
Table 5 Performance Matric difference between ANN Models (with and without outage indicators)

Model ID	Hidden Layers/ Neurons	Training Function	MAE	RMSE	R^2
ANN (with outage indicator)	1 layer with (10 Neurons)	Levenberg-Marquardt	1.5284	2.0017 MW	0.8240
ANN (without outage indicator)	1 layer with (10 Neurons)	Levenberg-Marquardt	3.0022	4.1178 MW	0.2553

Figure 7 illustrates the performance further as it presents the performance of the training, validation and testing together with the overall performance of the model. The Training of the model has the best results for the model with the outage indicator.



(A) Bar chart of Evaluation Metric
(with outage indicator)



(B) Bar chart of Evaluation Metric
(with outage indicator)

Figure 8 Bar Chart of Evaluation Matric between The 2 ANN Models (with & without outage indication)

V. CONCLUSION

This paper focuses on applying Artificial Neural Network ANN for a short term load forecasting of the Rumuola 33/11kV network. The models were built using the MATLAB R2023B Neural Networking fitting app and Regression learner app.

The objectives of the research focuses on collecting and analyzing the Historical data gotten from the Rumuola substation, develop an ANN model using MATLAB having a dataset containing temperature, humidity, days of the week indicators, full and partial outage indicators. Then develop another ANN model without the outage indicators in the dataset, compare, evaluate and test the ANN models for best performance. Use the ANN model with the best performance to forecast a 24 Hours load demand. Compare the best ANN model with a traditional technique (linear regression) to show the superiority of the ANN model.

Contribution to knowledge

(i) This research has helped to close multiple gaps on existing research. The major contribution of this research was showing the importance of including outage indicators when forecasting Rumuola substation short term Electricity load demand as it was noticed that outages are events that frequently occur due to overloading in the Nigerian electricity power system unlike other developed countries that treats outages as rare events and because of this, outage data are usually ignored or interpolated. This research proved that by adding the outage indicators to the dataset, it can significantly increase the forecasting accuracy, especially in a network that contains a lot of partial and full outages. The indicators helps the model learn certain patterns and relationships between the outages and the load demands.

(ii) Another gap that was covered in this research was the limited comparison between traditional models with more advanced methods. By comparing Artificial Neural Network with the Linear Regression method under the same conditions (both model contain outage indicators). It proves that neural networks are better in capturing non-linear behaviors of load forecasting than linear methods using a dataset filled with load demand and outages.

REFERENCE

[1]. Aderibigbe, A. O., Ani, E. C., Ohenhen, P. E., Ohalete, N. C., & Daraojimba, D. O. (2023). Enhancing energy efficiency with AI: A review of machine learning models in electricity demand forecasting. *Engineering Science & Technology Journal, 4*(6), 341–356.

[2]. Ahmad, W., Ayub, N., Ali, T., Irfan, M., Awais, M., Shiraz, M., & Glowacz, A. (2020). Towards short-term electricity load forecasting using improved support vector machine and extreme learning machine. *Energies, 13*(11), 2907. https://doi.org/10.3390/en13112907

[3]. Akman, T., Yılmaz, C., & Sönmez, Y. (2023). Short-term electric energy load forecasting of Ankara region using artificial intelligence methods. *Politeknik Dergisi, 26*(4), 1517–1531.

[4]. Aıroboman, A. (2022). Forecasting of maximum peak load demand of Abuja region interconnected power network using artificial neural network. *International Journal of Engineering Science and Application, 6*(1), 12–28.

[5]. Arvanitidis, A. I., Bargiotas, D., Daskalopulu, A., Laitsos, V. M., & Tsoukalas, L. H. (2021). Enhanced short-term load forecasting using artificial neural networks. *Energies, 14*(22), 7788. https://doi.org/10.3390/en14227788

[6]. Aslam, J., Kamran, M., Arif, A., Wasif, M., Irshad, J., & Hussain, I. (2022, November 16). Short-term load forecasting under extensive power outages using domestic energy meter load profile: A case study. *Journal of Engineering Research*, ICEPE Special Issue. https://doi.org/10.36909/jer.ICEPE.19559

[7]. Awad, H., Mahmoud, H. M., Allam, E., Abd Alfatah, S., Didane, D. H., Amin, A. A., & El-Sebah, M. I. A. (2019). Short-term electrical load forecasting using artificial neural networks. *International Journal of Electrical and Computer Engineering, 9*(6), 4590–4598. *(Year inferred and completed reference example.)*

[8]. Brahim, R., Ahlam, L., Hamza, A., & Güler, İ. (2025). Power forecasting using ANN and ELM: A comparative evaluation of machine learning approaches. *Mathematical Modelling of Engineering Problems, 12*(1).

[9]. Briggs, K. C., Idoniboyeobu, D. C., & Braide, S. Study Of Transmission/Distribution Line Interface For Improved Grid Stability Using Moment Distribution Technique.

[10]. Chen, S., Ren, Y., Friedrich, D., Yu, Z., & Yu, J. (2020). Sensitivity analysis to reduce duplicated features in ANN training for district heat demand prediction. *Energy and AI, 2*, 100028. https://doi.org/10.1016/j.egyai.2020.100028

[11]. Chikezie, O., Dikio Clifford Idoniboyeobu, D. R., Braide, U., & Okpara, K. (2022). Enhancement of 11kV Distribution Network for Power Quality Improvement Using Artificial Neural Network Based DVR. *Iconic Research and Engineering Journals*, *5*(10), 102-111.

[12]. Ciechulski, T., & Osowski, S. (2021). High-precision LSTM model for short-time load forecasting in power systems. *Energies, 14*(11), 2983. https://doi.org/10.3390/en14112983

[13]. Cordeiro-Costas, M., Villanueva, D., Eguía-Oller, P., Martínez-Comesaña, M., & Ramos, S. (2023). Load forecasting with machine learning and deep learning methods. *Applied Sciences, 13*(13), 7933. https://doi.org/10.3390/app13137933

[14]. Chane, K., Gebru, F. M., & Khan, B. (2021). Short-term load forecasting of distribution feeder using artificial neural network technique. *Journal of Informatics Electrical and Electronics Engineering, 2*(2), 1–22.

[15]. Dinh, T. H., Tran, T. T., & Do, N. D. P. (2025). Short-term active power forecasting for wind farms using artificial neural networks. *CTU Journal of Innovation and Sustainable Development, 17*(Special Issue: ETMD), 44–52.

[16]. Edohen, O. M., & Odiase, O. F. (2021). Short-term prediction of electric load demand of University of Benin using artificial neural network. *NIPES Journal of Science and Technology Research, 3*(3), 1–10.

[17]. Ezugwu, C. O., Opata, C. E., Dimeze, A. I., Ogbuka, C. U., & Nwosu, C. M. (2020). Development of a hybrid ARIMA–ANN model for enhanced electric load forecasting accuracy. *International Journal of Energy Applications and Technologies, 7*(2), 89–96.*

[18]. Fakhryza, K. A., Budisusila, E. N., & Nugroho, A. A. (2024). Application of artificial neural network for peak load forecasting in 150 kV Semarang power system. *Journal of Electrical Systems and Information Technology, 11*(1), 40. https://doi.org/10.1186/s43067-024-00140-5

[19]. Ganthia, B. P., Mohanty, M., Mohapatra, S. S., Pradhan, R., Satapathy, S., Patra, S., & Pahadasingh, S. P. (2023). Artificial neural network optimized load forecasting of smart grid using MATLAB. *Control Systems and Optimization Letters, 1*(1), 46–51.

[20]. Grandón, T. G., Schwenzer, J., Steens, T., & Breuing, J. (2024). Electricity demand forecasting with hybrid classical statistical and machine learning algorithms: Case study of Ukraine. *Applied Energy, 355*, 122249. https://doi.org/10.1016/j.apenergy.2023.122249

[21]. Hammad, M. A., Jereb, B., Rosi, B., & Dragan, D. (2020). Methods and models for electric load forecasting: A comprehensive review. *Logistics, Supply Chain, Sustainability and Global Challenges, 11*(1), 51–76.

[22]. Her, O. Y., Mahmud, M. S. A., Abidin, M. S. Z., Ayop, R., & Buyamin, S. (2022). Artificial neural network-based short-term electrical load forecasting. *International Journal of Power Electronics and Drive Systems, 13*(1), 586–594.