African Journal of Mathematics and Statistics Studies
ISSN: 2689-5323
Volume 7, Issue 2, 2024 (pp. 262-275)
www.abjournals.org

SEASONALITY PREDICTION OF MENINGITIS USING ARTIFICIAL NEURAL NETWORK (ANN)

Aminu Titilope Funmilayo[1], Daniel Akindunjoye Akinboro[2], and Sanusi Akeem Olatoye[3]

[1&3]Department of Mathematics, Federal Polytechnic Offa, Kwara State.
Email: [1]titiaminu15@gmail.com; [3]olatotesanusi@fedpoffaonline.edu.ng

[2]Department of Statistics, University of Ilorin, Kwara State.
Email: danielakinboro10@gmail.com

Cite this article:

Aminu T. F., Daniel A. A., Sanusi A. O. (2024), Seasonality Prediction of Meningitis Using Artificial Neural Network (ANN). African Journal of Mathematics and Statistics Studies 7(2), 262-275. DOI: 10.52589/AJMSS-UTSMWUD8

Manuscript History

Received: 13 Apr 2024

Accepted: 1 Jun 2024

Published: 1 Jul 2024

ABSTRACT: *Communities are concerned about controlling, preventing, and handling infectious diseases due to recent epidemic outbreaks. Meningitis, an inflammation of the membranes surrounding the brain and spinal cord, is a significant risk in Nigeria. It can cause death within hours of infection, with an average case fatality rate of 10%. To prevent meningitis outbreaks, this paper focuses on using an Artificial Neural Network (ANN) to predict outbreaks based on climatological factors. Previous research has shown that climate plays a major role in these outbreaks. The study found that the Levenberg-Marquaralt ANN algorithm was the best model, with the lowest prediction error and fewer iterations. High temperature and low humidity were identified as major triggers for meningitis outbreaks. It is crucial to address these factors to prevent future outbreaks and protect communities.*

KEYWORDS: Artificial Neural Network, Meningitis, Prediction, Machine learning, Time steps.

Article DOI: 10.52589/AJMSS-UTSMWUD8
DOI URL: https://doi.org/10.52589/AJMSS-UTSMWUD8

African Journal of Mathematics and Statistics Studies
ISSN: 2689-5323
Volume 7, Issue 2, 2024 (pp. 262-275)

www.abjournals.org

INTRODUCTION

Meningitis is a condition where the membranes surrounding the brain and spinal cord become inflamed. These membranes, known as the meninges, are responsible for protecting the central nervous system in conjunction with the cerebrospinal fluid. Meningitis can be caused by various types of organisms, with bacteria and viruses being the most common culprits. When these organisms are present in the cerebrospinal fluid, inflammation occurs in the surrounding area. Bacterial meningitis is more prevalent than viral meningitis. [13], meningococcal disease is a contagious bacterial illness caused by the Neisseria meningitidis bacterium, which is a Gram-negative diplococcus bacterium.

There are three primary clinical forms of meningococcal disease. The most prevalent type is Meningococcal Meningitis (MM), particularly during epidemics. MM, also known as cerebrospinal meningitis, is the sole form of meningitis that leads to epidemics with high fatality rates. It spreads through respiratory droplets (throat secretions) of an infected person via person-to-person contact [12].

Epidemics can happen anywhere in the world, but the largest ones tend to occur in the semi-arid regions of sub-Saharan Africa, which is known as the African 'Meningitis belt' [8]. This belt includes countries such as Benin, Burkina Faso, Northern Cameroon, Chad, Gambia, Ghana, Mali, Niger, Northern Nigeria, and Sudan. Every year, these countries are heavily impacted by major MM disease outbreaks that occur in seasonal cycles between the end of December and the end of June, depending on the location and climate of the country. Meningitis outbreaks have been linked to climatic and environmental factors [15]. It is believed that low absolute humidity and high dust concentration can make it easier for meningococcal bacteria to invade by damaging the mucosa barrier directly, penetrating the epithelial lining of the upper respiratory tract or inhibiting mucosal immune defense [13].

[5] Published the first research that attempted to relate MM outbreaks with environmental conditions in Africa. Later on, other studies aimed to model the spatial distribution of MM risk based on various environmental factors such as rainfall, absolute humidity, aerosols, temperature, and dust concentrations [10]. Although the precise mechanism by which these climatic factors impact meningitis is still not fully understood, it is believed that when there are increased concentrations of dust, high winds, elevated temperature, and low humidity, there may be damage to the nasopharyngeal mucosa, which may increase the risk of meningitis [14].

Research has shown that the highest mean temperature is significantly correlated with the annual peak incidence of meningitis, while absolute humidity has an inverse correlation [3]. The Department of Invasive Disease supports the theory that prolonged dry seasons, which occur during the "cold dry" months (November to January), may increase the risk of precursory diseases that can lead to meningitis during the subsequent "hot dry" months [2].

Previous studies have shown a potential correlation between climatic conditions and the incidence of meningitis, indicating the possibility of using environmental information to predict outbreaks. For accurate simulation and prediction of climate-related epidemics, it is essential to utilize climatic variables as explanatory factors collectively [9].

Research conducted in the fields of epidemiology and climatology since the mid-20th century has shown a correlation between meningitis epidemics and the dry season in terms of both location and timing [6]. The typical climate during the winter in subtropical areas creates an

Article DOI: 10.52589/AJMSS-UTSMWUD8
DOI URL: https://doi.org/10.52589/AJMSS-UTSMWUD8

environment conducive to meningitis development, while drought and strong winds can increase the risk of meningococcal invasion by damaging the mucous barrier or hindering the surface.

In Nigeria, the majority of meningitis cases and deaths occur in the Northern region due to the more severe dry season compared to the South [4]. It is widely known that climate greatly influences the spread of many infectious diseases, some of which are leading causes of death and illness in developing nations. These illnesses sometimes result in epidemics that can be sparked by changes in weather patterns that heighten transmission rates [14].

Meningitis has been a problem in Africa, especially in Nigeria, for the last 30 years. It has appeared in isolated cases, outbreaks, and even large epidemics. Until recently, the only way to control the disease was to conduct mass vaccination campaigns once a certain number of cases had been reported.

A model called the Artificial Neural Network (ANN) can predict the outbreak of meningitis before an epidemic happens. This can help improve surveillance and response monitoring to prevent the outbreak of meningitis, which can save lives and reduce costs while minimizing the negative impact on the economy. Machine learning (ML) is an artificial intelligence technique that teaches computers to learn from experience. It works by training with vast data sets and creating algorithms that can predict outcomes. These algorithms use mathematical methods, often derived from calculus and linear algebra, to identify trends and patterns in the data set and make accurate predictions.

In this paper, the analysis of meningitis climatology is conducted based on environmental parameters such as temperature, humidity, surface dust concentration, aerosol optical depth, and meridional wind. The Machine Learning approach is used specifically to establish predictive models for meningitis outbreaks. While many studies have shown a relationship between climatic factors and the occurrence of meningitis, few have focused on predicting the outbreaks of meningitis before an epidemic outbreak occurs. However, various predictive methods have been utilized by researchers to predict the outbreak of meningitis. There are several techniques used for disease prediction.

[1] Conducted a thorough study that highlighted the impact of certain weather conditions on the monthly occurrence of meningitis in Northwest Nigeria. The study also emphasized the significance of other risk factors that are not yet fully comprehended but could potentially be connected to societal and behavioral practices. The researchers concluded that by identifying and measuring both seasonal climate-related and non-climate-related risk factors through methods like improving disease surveillance, our understanding and prediction abilities can be significantly improved.

[11] Demonstrated a correlation ($r = 0.92$) between the onset of meningitis epidemics and the peak of winter in Mail, which occurs in the middle of February. This season is marked by the strongest harmattan winds and the lowest humidity levels.

[16] Investigated the "Comparative Study of Four Time Series Methods in Forecasting Typhoid Fever Incidence in China," A thorough analysis of various forecasting techniques based on monthly typhoid fever incidence rates was observed. They also compared three different models inspired by neural networks: backpropagation neural networks, radial basis function neural networks (RBFNN), and Elman recurrent neural networks. The study also examined the

Article DOI: 10.52589/AJMSS-UTSMWUD8
DOI URL: https://doi.org/10.52589/AJMSS-UTSMWUD8

theoretical and practical aspects of these models. The authors used data from 2005 to 2010 collected by the Centre for Disease Control and Prevention as the basis for their forecasting samples. The results showed that the neural network-based models were more effective than traditional statistical models.

Based on a study, it has been found that certain types of neural networks perform better than others. The focus of the research paper is on the Artificial Neural Network (ANN), which is widely used in Machine Learning (ML). In particular, the Multilayer Perceptron Artificial Neural Network (MLPANN) is being employed to comprehend the connection between MM outbreaks and various environmental factors. The ultimate goal is to create a warning system for MM outbreaks that integrates climatology forecasts.

METHODOLOGY

ANN is used to predict the future values of Meningitis outbreak given the past values. This model is developed using the following training algorithms: Levenberg-Marquardt, Bayesian Regularization and the Scaled Conjugate Gradient algorithm.

Levenberg-Marquardt Algorithm

The Levenberg-Marquardt algorithm combines the steepest descent method with the Gauss-Newton method and operates correctly in search for parameters both far from and close to the optimum one. In the former case the algorithm of the linear model of steepest descent is used, and in the latter one - squared convergence. Fast convergence is an additional advantage of the algorithm.

The Levenberg-Marquardt algorithm is an iterative method, in which the vector of unknown parameters are determined during step k +1 by the equation:

$$a_{k+1} = a_k^T - [J^T(a_k, t)J(a_k, t) + \mu_k I]^{-1} J^T(a_k, t)y(a_k, t) \qquad (2)$$

With the error:

$$I_2 = \int_0^T y^2\,(a_k, t)\,dt, \qquad (3)$$

Where:

$$y(a_k, t) = \int_0^t k(t-\tau)u(t)dt, \qquad (4)$$

Article DOI: 10.52589/AJMSS-UTSMWUD8
DOI URL: https://doi.org/10.52589/AJMSS-UTSMWUD8

African Journal of Mathematics and Statistics Studies
ISSN: 2689-5323
Volume 7, Issue 2, 2024 (pp. 262-275)

www.abjournals.org

$$J(a_k,t)=\begin{bmatrix}\frac{\delta y(a_k,t_1)}{\delta a_1} & \frac{\delta y(a_k,t_1)}{\delta a_2} & \cdots & \frac{\delta y(a_k,t_1)}{\delta a_m}\\ \frac{\delta y(a_k,t_2)}{\delta a_1} & \frac{\delta y(a_k,t_2)}{\delta a_2} & \cdots & \frac{\delta y(a_k,t_2)}{\delta a_m}\\ \vdots & \vdots & \ddots & \vdots\\ \frac{\delta y(a_k,t_n)}{\delta a_1} & \frac{\delta y(a_k,t_n)}{\delta a_2} & \cdots & \frac{\delta y(a_k,t_n)}{\delta a_m}\end{bmatrix} \quad (5)$$

The notations are as follows:

$k=1,2,\dots,p,\qquad p-\ number\ of\ iteration\ loops;$

$J_{(n\times m)}(a_k,t)-\ Jacobian\ matrix;$

$I_{(m\times m)}-\ unit\ matrix;$

$\mu_k-\ scalar, its\ value\ changes\ during\ iteration;$

$a=[a_1,a_2,\dots,a_m]-\ model\ parameters\ search\ for.$

The Levenberg-Marquardt algorithm is used for computation in the following stages:

Stage 1, for k=1

Assume the initial values of the parameters of vector a_k;

Assume the initial value of the coefficient μ_k (e.g. $\mu_k=0.1$);

Solve the matrix equation (5) and calculate (4);

Calculate the value of error (3);

Determine the parameters of vector a_{k+1}, a following (2).

Stage 2 and further steps, for $k=2,3,\dots,p$

Update the values of the parameters of vector a_k;

Solve the matrix equations (5), calculate (4) and (2);

Calculate the value of error (3);

Compare the values of error (3) for the step k and the step k - 1 If the result is $J(a_k,t)\geq J(a_{k-1},t)$, multiply μ_k by the specific value $\lambda\in\Re$ (e.g. λ =10) and return to step 2 of stage If the result is $J(a_k,t)<J(a_{k-1},t)$ divide μ_k by the value of λ and return to step of stage

The initial parameters of vector a are assumed in an arbitrary way, e.g $a=[1,1,\dots,1]$.

If in a consecutive stage the decrease in the value of error (3) is very small and insignificant, we then finish the iteration process. We fix $\mu_k=0$ and determine the final result for the parameters of vector a If the value of coefficient μ_k is high, it means that the solution is not

Article DOI: 10.52589/AJMSS-UTSMWUD8
DOI URL: https://doi.org/10.52589/AJMSS-UTSMWUD8

African Journal of Mathematics and Statistics Studies
ISSN: 2689-5323
Volume 7, Issue 2, 2024 (pp. 262-275)

www.abjournals.org

satisfactory$I_2 = max \int_0^T y^2\,(a_k, t)\,dt$. The values of parameters of the vector are not optimum ones, and the value of error (3) is not at the minimum level. At this point:

$$J^T(a_k, t)J(a_k, t) \ll \mu_k I, \qquad (6)$$

Can be assumed and this leads to the steepest descent method, for which we have:

$$a_{k+1} = a_k^T - \frac{1}{\mu_k} J^T(a_k, t)J(a_k, t) \qquad (7)$$

If the value of coefficient μ_k is small, it means that the values of the parameters of vector

a Are close to the optimum solution. Then:

$$J^T(a_k, t)J(a_k, t) \gg \mu_k I, \qquad (8)$$

Which means that the Levenberg-Marquardt algorithm is reduced to the Gauss-Newton method:

$$a_{k+1} = a_k^T - [J^T(a_k, t)J(a_k, t) + \mu_k I]^{-1} J^T(a_k, t)y(a_k, t) \qquad (9)$$

The selection of coefficient values μ and λ assumed as: μ = 1.0 and λ =10.

Bayesian Regularization Backpropagation Neural Network

Bayesian regularization is an artificial neural network training algorithm which corrects the weight and refraction values based on the Levenberg-Marquardt optimization. This algorithm minimizes the combination of error squares and weights, then determine the correct combination so as to produce a good network. This process is called Bayesian regularization. Bayesian regularization neural network introduce network weights into the objective function of training. The objective function of the training is notated as follows

$$F(\omega) = \alpha E\omega + \beta ED \qquad (10)$$

Where $E\omega$ is the sum of squares of the network weight, and ED the sum of squares of network errors. The values α and β are parameters of the objective function. In the Bayesian process flow, network weights are seen as random variables, then the previous distribution of network weights and training is considered a Gaussian distribution follows. The following applies the Bayesian rule to optimize the objective function parameters α and β.

$$P(\alpha, \beta|D, M) = \frac{P(D|\alpha, \beta, M)P(\alpha, \beta|\, M)}{P(D|\, M)} \qquad (11)$$

Consider a neural network with training dataset D having n_t input and target vector pairs in the network model, i.e

$$D = \{(u_1, t_{o1}), (u_2, t_{o2}), \dots, (u_{n_t}, t_{on_t})\} \qquad (12)$$

For each input (u) to the network, the difference between target output (t_o) and predicted output (a_0) is computed as error e. In order to evaluate the performance of the network, i.e. how well the neural network is fitting the test data, a quantitative measure is needed.

Article DOI: 10.52589/AJMSS-UTSMWUD8
DOI URL: https://doi.org/10.52589/AJMSS-UTSMWUD8

African Journal of Mathematics and Statistics Studies
ISSN: 2689-5323
Volume 7, Issue 2, 2024 (pp. 262-275)

www.abjournals.org

This measure is called performance index of the network and is used to optimize the network parameters. The standard performance index $F(\bar{\omega})$ is governed by the sum of the squared errors (SSE).

$$F(\bar{\omega}) = E_D = \sum_{i=1}^{n_t} (e_i)^2 = \sum_{i=1}^{n_t} (t_{oi} - a_{0i})^T (t_{oi} - a_{0i}), \qquad (13)$$

Where $\bar{\omega}$ denotes the vector of size K containing all the weights and biases of the network. In order to generalize the neural network, the performance index of Eq. (13) is modified using a regularization method. A penalty term $(\mu/\nu)E_\omega$ is added to the performance index $F(\bar{\omega})$.

$$F(\bar{\omega}) = \mu\bar{\omega}^T\bar{\omega} + \nu E_D = \mu E_\omega + \nu E_D, \qquad (14)$$

Where μ and ν are the regularization parameters and E_ω represents the sum of the squared network weights (SSW).
Finding the optimum values for μ and ν is a challenging task, as their comparative values set up the basis for the training error. If $\mu << \nu$, smaller errors are generated, while $\mu >> \nu$, there should be reduced weight size at the cost of network errors. For the purpose of finding the optimum regularization parameters, a Bayesian regularization method is employed.

Considering the network weights $\bar{\omega}$ as random variables, the aim is to choose the weights that maximize the posterior probability distribution of the weights $P(\bar{\omega}|D, \mu, \nu, M_N)$ given a certain data D. According to Bayes' rule the posterior distribution of the weights depends on the likelihood function $P(D|\bar{\omega}, \nu, M_N)$, the prior density$P(\bar{\omega}|\mu, M_N)$, and the normalization factor $P(D|\mu, \nu, M_N)$ for a particular neural network model M_N and can be evaluated from

$$P(\bar{\omega}|D, \mu, \nu, M_N) = \frac{P(D|\bar{\omega}, \nu, M_N)P(\bar{\omega}|\mu, M_N)}{P(D|\mu, \nu, M_N)} \qquad (15)$$

Considering that the noise in the training set has a Gaussian distribution, the likelihood

function is given by

$$P(D|\bar{\omega}, \nu, M_N) = \frac{\exp(-\nu E_D)}{Z_D(\nu)}, \qquad (16)$$

Where $Z_D = (\pi/\nu)^{Q/2}$ and $Q = n_t \times N^{n_l}$.

Similarly, assuming a Gaussian distribution for the network weights, the prior probability

Density $P(\bar{\omega}|\mu, M_N)$ is given as

$$P(\bar{\omega}|\mu, M_N) = \frac{\exp(-\mu E_W)}{Z_W(\mu)} \qquad (17)$$

Where $Z_W = (\pi/\alpha)^{K/2}$.

The posterior probability with the network weights $\bar{\omega}$ can then be expressed as

Article DOI: 10.52589/AJMSS-UTSMWUD8
DOI URL: https://doi.org/10.52589/AJMSS-UTSMWUD8

African Journal of Mathematics and Statistics Studies
ISSN: 2689-5323
Volume 7, Issue 2, 2024 (pp. 262-275)

www.abjournals.org

$$P(\bar{\omega}|D,\mu,\nu,M_N) = \frac{\exp(-\mu E_W - \nu E_D)}{Z_F(\mu,\nu)} = \frac{\exp(-F(\bar{\omega}))}{Z_F(\mu,\nu)}, \qquad (18)$$

Where $Z_F(\mu,\nu) = Z_D(\nu)Z_W(\mu)$ is the normalization factor.

The complexity of the model M_N is governed by regularization parameters μ and ν, which need to be estimated from the data. Therefore, Bayes' rule is again applied to optimize them

From

$$P(\mu,\nu\,|D,M_N) = \frac{P(D|\mu,\ \nu,M_N)P(\mu,\ \nu|\ M_N)}{P(D|M_N)}, \qquad (19)$$

Where $P(\mu,\nu\,|M_N)$ denotes the assumed uniform prior density for the parameters μ and ν. From Eq. (19), it is evident that maximizing the likelihood function $P(D|\mu,\ \nu,M_N)$ eventually maximizes the posterior probability$P(\mu,\ \nu\,|D,M_N)$. Moreover, it can be noted

that the likelihood function in Eq. (19) is the normalization factor of Eq. (15). Therefore,

solving for the likelihood function $P(D|\mu,\ \nu,M_N)$ and expanding the objective function in Eq. (14) around the minimal point $\bar{\omega}^*$ via a Taylor series expansion, the optimum values of regularization parameters can be evaluated as follows

$$\mu^* = \frac{\gamma}{2E_W(\bar{\omega}^*)} \qquad \text{and} \qquad \nu^* = \frac{Q-\gamma}{2E_D(\bar{\omega}^*)} \qquad (20)$$

Where γ signifies the "number" of effective parameters exhausted in minimizing the error

Function $\gamma = K - \mu^* tr(H^*)^{-1}, for\ 0 \le \gamma \le K, \qquad (21)$

At the end of the training, a few checks regarding the number of effective parameters

are required for better performance of the network. The problem of computing the Hessian matrix at the minimal point $\bar{\omega}^*$ is implicitly solved in the Levenberg-Marquardt (LM) training algorithm while finding the minimum of $F(\bar{\omega})$. In the LM algorithm, the network weights and biases at the k^{th} iteration are adjusted according to

$$\bar{\omega}^{k+1} = \bar{\omega}^k - [J^T J + \lambda I]^{-1} J^T e,$$

and H^* is the Hessian matrix of the objective function evaluated at $\bar{\omega}^*$, which is calculated

using the Gauss-Newton approximation as

$$H^* \approx J^T J, \qquad (22)$$

Where J is the Jacobian matrix formed by the first derivatives of the network errors **e** with respect to network weights w_{ij}. In (21), tr(.) denotes the trace operator. The normalization factor $Z_F(\mu,\nu)$ can then be approximated as

$$Z_F(\mu,\nu) \approx (2\pi)^{K/2}(\det(H^*))^{-1/2}\exp\left(-F(\bar{\omega}^*)\right). \qquad (23)$$

Article DOI: 10.52589/AJMSS-UTSMWUD8
DOI URL: https://doi.org/10.52589/AJMSS-UTSMWUD8

African Journal of Mathematics and Statistics Studies
ISSN: 2689-5323
Volume 7, Issue 2, 2024 (pp. 262-275)

www.abjournals.org

Scaled Conjugate Gradient Algorithm

Conjugate gradient methods applied to the non-linear unconstrained minimization problem:

$$\min f(x),\ x \in R^n. \qquad (25)$$

Where $f: R^n \to R$ is continuously differentiable function and bounded below. A conjugate gradient method generates a sequence $x_k, k \geq 1$ starting from an initial guess $x_1 \in R^n$, using the recurrence

$$x_{k+1} = x_k + \alpha_k d_k \qquad (26)$$

Where the positive step size α_k is obtained by a line search, and the directions d_k are generated by the rule:

$$d_{k+1} = -g_{k+1} + \beta_k d_k,\ \ d_1 = -g_1 \qquad (27)$$

Where $g_k = \nabla f(x_k)$, and let $y_k = g_{k+1} - g_k$ and $s_k = x_{k+1} - x_k$, here β_k is the CG update parameter. Different CG methods corresponding to different choice for the parameter β_k. The first CG algorithm for non-convex problems was proposed by Fletcher and Revees (FR) in 1964, which defined as

$$\beta_k^{FR} = \frac{g_{k+1}^T g_{k+1}}{g_k^T g_k} \qquad (28)$$

We know that the other equivalents forms for β_k are Polack-Ribeir (PR) and Hestenes- Stiefel (HS) for example

$$\beta_k^{PR} = \frac{g_{k+1}^T y_{k+1}}{g_k^T g_k}, \quad and \quad \beta_k^{HS} = \frac{g_{k+1}^T y_{k+1}}{d_k^T g_k} \qquad (29)$$

Although all the above formulas are equivalent for convex quadratic functions, but they have different performance for non-quadratic functions, the performance of a non-linear CG algorithm strongly depends on coefficient β_k. Dai and Yuan (DY) in proposed a non-linear CG method (28) and (29) with β_k defined as

$$\beta_k^{DY} = \frac{g_{k+1}^T g_{k+1}}{d_k^T y_k} \qquad (30)$$

Which generates a descent search directions

$$d_k^T g_k < 0 \qquad (31)$$

At every iteration k and convergence globally to the solution if the following Wolfe conditions are used to accept the step-size α_k:

$$f(x_k + \alpha_k d_k) \leq f(x_k) + c_1 \alpha_k g_k^T d_k \qquad (32)$$

$$g(x_k + \alpha_k d_k)^T d_k \geq c_2 g_k^T d_k \qquad (33)$$

Where $0 < c_1 < c_2 < 0$. Condition (32) stipulates a decrease of f and d_k if (33) satisfied. Condition (33) is called the curvature condition and it's role is to force α_k to be sufficiently far

Article DOI: 10.52589/AJMSS-UTSMWUD8
DOI URL: https://doi.org/10.52589/AJMSS-UTSMWUD8

African Journal of Mathematics and Statistics Studies
ISSN: 2689-5323
Volume 7, Issue 2, 2024 (pp. 262-275)

www.abjournals.org

away from zero. Which could happen if only condition (32) were to be used. Conditions (32) and (33) are called standard Wolfe conditions (SDWC). Notice that if equation (32) is satisfied then always there exists $\bar{\alpha} > 0$ such that for any $\alpha_k \in [0, \bar{\alpha}]$ the conditions (32) and (33) will be satisfied according to the theorem (1) given later. If we wish to find a point α_k, which is closer to a solution of the one dimensional problem

$$\min_{\alpha>0} \emptyset(\alpha) = \min_{\alpha>0} f(X_k + \alpha d_k) \qquad (34)$$

Than a point satisfying (32) and (33) we can impose on α_k the strong Wolfe conditions (STWC):

$$f(x_k + \alpha_k d_k) \leq f(x_k) + c_1 \alpha_k g_k^T d_k \qquad (35)$$

$$|g(x_k + \alpha_k d_k)^T d_k| \leq c_2 |g_k^T d_k| \qquad (36)$$

Where $0 < c_1 < c_2 < 0$. In contrast to (SDWC) $g_{k+1}^T d_k$ cannot be arbitrarily large. The (STWC) with the sufficient descent property

$$d_k^T g_k < c\|g_k\|., c \in (0,1) \qquad (37)$$

Widely used in the convergence analysis for the CG methods.

DATA PRESENTATION

The following data consists of a sample of 500 elements that shows the number of meningitis cases based on various environmental factors. The input data is in the form of a 500x5 matrix, representing dynamic data for 500 timesteps of 5 elements. The target data, on the other hand, is a 500x1 matrix, representing dynamic data for 500 timesteps of 1 element.

The environmental factors used to train the algorithms for ANN were acquired from the Meteorological State Agency of Spain (AEMET) and various collaborating bodies. These factors were obtained from the Journal of International Research Institute for Climate and Society at Columbia University (IRI) for meningitis cases. Additionally, data from the Banizoumbou AERONET station, dust reanalysis from two different sources, and satellite-based information from the World Data Centre for Remote Sensing of the Atmosphere, WDC-RSAT, and the Giovanni web-based application from NASA were also used. Furthermore, additional simulations were performed using the MINITAB application package.

The model inputs consist of 5 variables for 500 weeks, which is approximately 10 years. These variables include temperature ranging from 16 to 36 degrees Celsius, absolute humidity ranging from 3.0 to 20.0gm-3, dust concentration ranging from 150 to 650ugm-3, aerosol optical depth ranging from 0.3 to 1.5nm, and meridional wind ranging from -2.0 to 2.5ms-1. The model output includes 1 sample of 500 elements, which represents the weekly number of meningitis cases. To train the ANN, the final data used was a matrix of 500 timesteps or weeks multiplied by 5 and a vector of 500 timesteps or weeks multiplied by 1 response variable as output. MATLAB was used on the Neural Network fitting to compare the performance of the MLPANN using Bayesian Regularization, Levenberg-Marquardt Algorithm, and Scaled

Article DOI: 10.52589/AJMSS-UTSMWUD8
DOI URL: https://doi.org/10.52589/AJMSS-UTSMWUD8

Conjugate Gradient Algorithm (SCG). The goal was to minimize the prediction error by considering the error terms in the training of the data.

Validation and Test Data

There are three (3) divisions of the dataset

1) Training: These are presented to the network during training and the network is adjusted according to its error. 70% of the data is used for training i.e. 350 target timesteps.

2) Validation: These are used to measure network generalization, and to halt training when generalization stops improving. 15% of the data is used for validation i.e. 75 target timesteps.

3) Testing: These have no effect on training and so provide an independent measure of network performance during and after training. 15% of the data is used for testing i.e. 75 target timesteps.

Network Architecture This network is created and trained in open loop form. Open loop (single-step) is more efficient than closed loop (multi-step) training. Open loop allows to supply the network with correct past outputs as the model is trained to produce the correct current outputs.

Several numbers of Hidden Neurons and Number of neurons or delays d is used to evaluate the performances of the Neural Network using the Levenberg-Marquardt algorithm, Bayesian Regularization algorithm and the Scaled Conjugate Gradient algorithm. Note, training automatically stops when generalization stops improving, as indicated by an increase in the mean square error of the validation.

The Mean Squared Error is the average squared difference between outputs and targets. Lower values are better. Zero means no error. This is used to evaluate the performance of the Neural Network.

The Regression R values measure the correlation between outputs and targets. An R value 1 means a close relationship, 0 a random relationship.

Multilayer Perceptron

Network Information

Input Layer	Factors	1	Temperature
		2	Humidity
		3	Dust_ Concentration
	Covariates	1	Aerosol_Optical_Depth
		2	Meridional Wind
	Number of Units[a]		416
	Rescaling Method for Covariates		Standardized
Hidden Layer(s)	Number of Hidden Layers		2

Article DOI: 10.52589/AJMSS-UTSMWUD8
DOI URL: https://doi.org/10.52589/AJMSS-UTSMWUD8

African Journal of Mathematics and Statistics Studies
ISSN: 2689-5323
Volume 7, Issue 2, 2024 (pp. 262-275)

	Number of Units in Hidden Layer 1[a]	10
	Activation Function	Hyperbolic tangent
Output Layer	Dependent Variables 1	Number of Cases
	Number of Units	1
	Rescaling Method for Scale Dependents	Standardized
	Activation Function	Identity
	Error Function	Sum of Squares

FIGURE 1

Model Summary

Training	Sum of Squares Error	344.158
	Relative Error	.983
	Stopping Rule Used	1 consecutive step(s) with no decrease in error[a]
	Training Time	0:00:01.42
Testing	Sum of Squares Error	106.190
	Relative Error	.987

FIGURE 2

Dependent Variable: Number of Cases

Multilayer Perceptron Algorithms for LM, BR and SCG Model Summary

Algorithms	Number of Hidden Neurons	Number of Epoch (Iterations)	TIME	MSE		R
Levenberg-Marquardt	10	8	01.42secs	Training	0.999529	0.0971883
				Validation	1.00351	0.0403117
				Testing	1.07120	-0.12291
Bayesian Regularization	10	41	03secs	Training	1.01129	0.0620949
				Validation	0.00000	0.00000
				Testing	0.922408	0.0169559
Scaled Conjugate Gradient	10	23	02secs	Training	0.998475	0.106632
				Validation	0.955837	-0.02630
				Testing	1.05843	-0.09360
Levenberg-Marquardt	5	9	01secs	Training	0.983776	00.159546
				Testing	1.00237	0.0643188
				Validation	1.04327	-0.11750

Article DOI: 10.52589/AJMSS-UTSMWUD8
DOI URL: https://doi.org/10.52589/AJMSS-UTSMWUD8

Bayesian Regularization	5	1000	70secs	Training	1.00411	0.0547818
				Validation	0.00000	0.00000
				Testing	0.922408	00.0169559
Scaled Conjugate Gradient	5	12	02 secs	Training	1.04684	-0.00642879
				Validation	0.963054	00.122583
				Testing	1.04196	-0.0879877

i. In Levenberg-Marquardt algorithm, training automatically stops when generalization stops improving, as indicated by an increase in the mean square error of the validation sample.

ii. In Bayesian Regularization algorithm, this algorithm typically takes more time, but can result in good generalization for difficult, small or noisy datasets. Training stops according to adaptive weight minimization (regularization).

iii. In Scaled Conjugate Gradient algorithm, this algorithm takes less memory. Training automatically stops when generalization stops improving, as indicated by an increase in the mean square error of the validation samples.

DISCUSSION OF RESULTS

In order to minimize prediction error, the Multilayer Perceptron is utilized with five independent variables. Each variable is trained using 500 observations to generate a predicted number of meningitis cases. Figure 1 displays the model information, while Figure 2 shows the model summary of the Multilayer Perceptron algorithm. The Levenberg-Merquardt (LM) algorithm with 5 hidden neurons and 9 iterations of 0.01 seconds has a 99% correct prediction rate for training, testing, and validation. The Bayesian Regularization algorithm (BR) with 5 hidden neurons is also reliable, but requires a larger number of iterations (1000) and takes 70 seconds to run. However, it has a 99% correct prediction rate for training and testing and a 100% correct prediction rate for validation. The Scaled Conjugate Gradient algorithm with 5 hidden neurons and 12 iterations of 0.02 seconds has an approximate 99% correct prediction rate for training, testing, and validation. Increasing the number of hidden neurons to 10 results in a small mean square error of approximately 1%. Levenberg-Marquardt offers the fastest convergence rate. Figures 3, 4, and 5 display the histogram of errors for the three methods used in training, testing, and validation.

CONCLUSION

This article presents an Artificial Neural Network Model that predicts meningitis outbreaks based on environmental factors. The model uses multilayer perceptron to minimize prediction errors. The study found that temperature is the most significant factor that affects the occurrence of meningitis epidemics, along with low absolute humidity and high dust concentration. When the atmospheric dust concentration exceeds 320ugm, the humidity drops below 5m\s, and the temperature remains below 20°C in December, January, and February, a

Article DOI: 10.52589/AJMSS-UTSMWUD8
DOI URL: https://doi.org/10.52589/AJMSS-UTSMWUD8

sudden increase in temperature can trigger a meningitis outbreak. This finding is a strong signal for the surveillance and response monitoring team to create an alert and be watchful. The model also considers categorical variables, denoted by C.

REFERENCES

[1] Abdussalam, A.F., Monaghan, A.J., Dukic, V.M., Hopson, T.M., Hayden, M.H., Leckebusch, G.C. & Thornes, J.E. (2013). Climate Influences on Meningitis Incidence in Northwest Nigeria.
[2] CartWright K.A.V, (1995): Meningococcal Carriage and disease.
[3] Dukic, V. & coauthor (2012) The role of weather in meningitis outbreaks in Navrongo, Ghana; A generalized additive modelling approach Journal Agric Biological Environmental statistics, 17, 422-460.
[4] Hodgson A., Smith T., Gagneu X.S,. Adjuilk M., Pluschke G., Mensah N.K. & Genton B. (2001) Risk factors for meningococcal meningitis in northern Ghana Trans. Roy Society Tropical Medicine Hyg. 95, 477-480.
[5] Lapeyssonie L (1963). La meningite cerebro-spinale en Afrique, bull OMS, 28 (suppl).1-100.
[6] Molesworth A, Cuevas L.E, Morse A.P, Heiman J. R. & Thomson M. C. (2002). Dust clouds and spread of infections, lancer 359, 81-82.
[7] Molesworth A. M., Cuevas L. E., Connor S. J., Morse A.P. & Thomson M.C.(2003) Environmental risk and meningitis epidemics in Africa, Emerging infections disease 1287-1293.
[8] Molesworth, A. M., Thomson, M. C., Connor, S. I., Cresswell M. P., Morser A.P. (2002). Where is the meningitis Belt? Defining an area at risk of epidemics meningitis in Africa. Trans Response Society Medicine Hyg. 96, 242-248.
[9] Palmgreen H. (2009). Meningococcal disease and climate. Global Health Action 2, Doi: 10:3402/gha.v2io,2061.
[10] Savory E.C, Cuevas L.E. Yassin M.A. Hart C.A. Molesworth A.M. & Thompson M. C. (2006) Evaluation of the meningitis epidemic risk model in Africa. Epidemiology and infection 134. 1047-1051.
[11] Sultan, B., Labadi, K., Guegan, J. & Janicot, S. (2005). Climate Drives the Meningitis Epidemics Onset in West Africa. https://doi.org/10.1371/journal.pmed.0020006.
[12] WHO (2000). Meningitis.
[13] WHO (2009). Meningitis.
[14] WHO (2012). Meningitis.
[15] Yaka, P., Sultan, B., Broutin, H., Janicot, S., Philippon, S. & Fourquet, N. (2008). Relationships between climate and year-to-year variability in meningitis outbreaks: A case study in Burkina Faso and Niger. International Journal of Health Geographics, Article number: 34.
[16] Zhang, X., Liu, Y., Yang, M., Zhang, T., Young, A.A. & Li, X. (2013). Comparative Study of Four Time Series Methods in Forecasting Typhoid Fever Incidence in China. https://doi.org/10.1371/journal.pone.0063116.

Article DOI: 10.52589/AJMSS-UTSMWUD8
DOI URL: https://doi.org/10.52589/AJMSS-UTSMWUD8