International Research Journal of Innovations in Engineering and Technology (IRJIET)

ISSN (online): 2581-3048

Volume 7, Issue 8, pp 382-389, August-2023

https://doi.org/10.47001/IRJIET/2023.708057

Generating Reliable and Accurate Forecasts of Annual Neonatal Mortality Rate for Nepal Using the ARIMA Model

[1]Dr. Smartson. P. NYONI, [2]Thabani NYONI

[1]ZICHIRe Project, University of Zimbabwe, Harare, Zimbabwe
[2]Independent Researcher & Health Economist, Harare, Zimbabwe

Abstract **- This study uses annual time series data on neonatal mortality rate (NMR) for Nepal from 1960 to 2019 to predict future trends of NMR over the period 2020 to 2030. Unit root tests have shown that the series under consideration is an I (1) variable. The optimal model based on AIC is the ARIMA (4,1,3) model. The study findings indicate that neonatal mortality will gradually decline from around 19 in 2020 to approximately 14 deaths per 1000 live births by 2030. Therefore, we encourage the government of Nepal to design local policies to address various maternal and child health program challenges to keep neonatal deaths under control.**

Keywords: ARIMA, Forecasting, NMR.

I. INTRODUCTION

Maternal and child health issues have dominated many discussions on global health problems (UNICEF, 2019). The sustainable development goals were launched in 2015 and SDG-3 focuses on maternal and child health. SDG 3.1 focuses on the reduction of maternal mortality to less than 70 maternal deaths per 100 000 live births and target 3.2 aims to ensure substantial reduction in neonatal mortality rate to at least 12 deaths per 1000 live births and under five mortality to levels as low as 25 deaths per 1000 live births (UN, 2020; UNICEF, 2019; WHO, 2019; UNICEF, 2018; UN, 2015). The government of Nepal is committed to tackle the problem of mortality in neonates. The authorities pledged to reduce neonatal deaths to the set SDG 3 target by 2030 (Kc *et al.* 2019). Nepal reported a decline in under 5 mortality rate from 133 deaths per 1000 live births in 1991 to 39 per 1000 live births in 2016 (World Bank, 2019; Nepal MOHP, 2017). The main causes of mortality in neonates in Nepal are prematurity, asphyxia and neonatal sepsis (WHO, 2018). The main objective of this study is to model and forecast neonatal mortality rate for Nepal using the popular Box-Jenkins ARIMA model. ARIMA model has been found to be appropriate for modelling linear time series data (Nyoni, 2018; Box & Jenkins, 1970). The findings of this study are expected to inform neonatal policy, decision making and resource mobilization for maternal and child health programs in the country so that effective neonatal policies are implemented timeously to prevent and control mortality in neonates.

II. LITERATURE REVIEW

Li *et al*. (2021) examined the proportion of mothers with history of neonatal deaths using the most recent Demographic and Health Surveys from 56 low- and middle-income countries. Logistic regression models were used to assess the association between maternal history of neonatal death and subsequent neonatal mortality. The adjusted models controlled for socioeconomic, child, and pregnancy-related factors. Country-specific analyses were performed to assess heterogeneity in this association across countries. Study findings suggested that maternal history of neonatal death could be an effective early identifier of high-risk pregnancies in resource-poor countries. In another study by Khader *et al.* (2021) explored the healthcare professionals' perception about the usability of JSANDS.A descriptive qualitative approach, using focus group discussions, was adopted. A total of 5 focus groups including 23 focal points were conducted in five participating hospitals in Jordan. The study findings revealed that JSANDS was perceived positively by the current users. According to them, it provides a formative and comprehensive data on stillbirths and neonatal deaths and their causes. Nath *et al*. (2020) examined the effect of extreme prematurity and early neonatal deaths on infant mortality rates in England. Authors used aggregate data on all live births, stillbirths and linked infant deaths in England in 2006–2016 from the Office for National Statistic. Infant mortality decreased from 4.78 deaths/1000 live births in 2006 to 3.54/1000 in 2014 (annual decrease of 0.15/1000) and increased to 3.67/1000 in 2016 (annual increase of 0.07/1000). This rise was driven by increases in deaths at 0–6 days of life. A descriptive study was carried out by McNamara *et al*. (2018) to reveal

ISSN (online): 2581-3048
Volume 7, Issue 8, pp 382-389, August-2023
https://doi.org/10.47001/IRJIET/2023.708057

intrapartum fetal deaths and unexpected neonatal deaths in Ireland from 2011 to 2014. Anonymised data pertaining to all intrapartum fetal deaths and unexpected neonatal deaths for the study time period was obtained from the national perinatal epidemiology centre. The findings of the study indicated that the corrected intrapartum fetal death rate was 0.16 per 1000 births and the overall unexpected neonatal death rate was 0.17 per 1000 live births. Boulos *et al.* (2017) investigated the etiology of severe bacterial infections in neonates. Researchers conducted a secondary retrospective analysis of a de-identified database from the Neonatal Intensive Care Unit (NICU) at Nos Petit Fre`res et Soeurs-St. Damien Hospital (NPFS-SDH). Records from 1292 neonates admitted to the NICU at NPFS-SDH in Port-au-Prince Haiti from 2013 to 2015 were reviewed. Sepsis accounted for 708 of 1292 (54.8%) of all admissions to the NICU. The most common organism cultured was Streptococcus agalactiae, followed by Klebsiella pneumoniae, Pseudomonas aeroginusa, Enterobacter aerogenes, Staphylococcus aureus and Proteus mirabillis.

III. METHODOLOGY

The Autoregressive (AR) Model

A process X_t (NMR at time t) is an autoregressive process of order p, that is, AR (p) if it is a weighted sum of the past p values plus a random shock (Z_t) such that:

$$X_t = \emptyset_1 X_{t-1} + \emptyset_2 X_{t-2} + \emptyset_3 X_{t-3} + \cdots + \emptyset_p X_{t-p} + Z_t \ldots\ldots [1]$$

Using the backward shift operator, B, such that $BX_t = X_{t-1}$, the AR (p) model can be expressed as in equation [2] below:

$$Z_t = \emptyset(B) X_t \ldots\ldots [2]$$

where $\emptyset(B) = 1 - \emptyset_1 B - \emptyset_2 B^2 - \emptyset_3 B^3 - \cdots - \emptyset_p B^p$

The 1st order AR (p) process, AR (1) may be expressed as shown below:

$$X_t = \emptyset X_{t-1} + Z_t \ldots\ldots [3]$$

Given $\emptyset = 1$, then equation [3] becomes a random walk model. When $|\emptyset| > 1$, then the series is reffered to as explosive, and thus non-stationary. Generally, most time series are explosive. In the case where $|\emptyset| < 1$, the series is said to be stationary and therefore its ACF (autocorrelation function) decreases exponentially.

The Moving Average (MA) Model

A process is reffered to as a moving average process of order q, MA (q) if it is a weighted sum of the last random shocks, that is:

$$X_t = Z_t + \theta_1 Z_{t-1} + \theta_2 Z_{t-2} + \cdots + \theta_q Z_{t-q} \ldots\ldots [4]$$

Using the backward shift operator, B, equation [4] can be expressed as follows:

$$X_t = \theta(B) Z_t \ldots\ldots [5]$$

where $\theta(B) = 1 + \theta_1 B + \theta_2 B^2 + \cdots + \theta_q B^q$

Equation [4] can also be expressed as follows:

$$X_t - \sum_{j \le 1} \pi_j X_{t-j} = Z_t \ldots\ldots [6]$$

for some constant π_j such that:

$$\sum_{j \le 1} |\pi_j| < \infty$$

International Research Journal of Innovations in Engineering and Technology (IRJIET)

ISSN (online): 2581-3048

Volume 7, Issue 8, pp 382-389, August-2023

https://doi.org/10.47001/IRJIET/2023.708057

This implies that it is possible to invert the function taking the Z_t sequence to the X_t sequence and recover Z_t from present and past values of X_t by a convergent sum.

The Autoregressive Moving Average (ARMA) Model

While the above models are good, a more parsimonious model is the ARMA model. The AR, MA and ARMA models are applied on stationary time series only. The ARMA model is just a mixture of AR (p) and MA (q) terms, hence the name ARMA (p, q). This can be expressed as follows:

$$\emptyset(B)M_t = \theta(B)Z_t \ldots \ldots [7]$$

Thus:

$$X_t\left(1 - \emptyset_1 B - \emptyset_2 B^2 - \cdots - \emptyset_p B^p\right) = Z_t\left(1 + \theta_1 B + \theta_2 B^2 + \cdots + \theta_q B^q\right) \ldots \ldots [8]$$

where $\emptyset(B)$ and $\theta(B)$ are polynomials in B of finite order p, q respectively.

The Autoregressive Integrated Moving Average (ARIMA) Model

The AR, MA and ARMA processes are usually not applied empirically because in most cases many time series data are not stationary; hence the need for differencing until stationarity is achieved.

$$\left.\begin{array}{c} The\ first\ difference\ is\ given\ by: \\ X_t - X_{t-1} = X_t - BX_t \\ The\ second\ difference\ is\ given\ by: \\ X_t(1-B) - X_{t-1}(1-B) = X_t(1-B) - BX_t(1-B) = X_t(1-B)(1-B) = X_t(1-B)^2 \\ The\ third\ difference\ is\ given\ by: \\ X_t(1-B)^2 - X_{t-1}(1-B)^2 = X_t(1-B)^2 - BX_t(1-B)^2 = X_t(1-B)^2(1-B) = X_t(1-B)^3 \\ The\ d^{th}\ difference\ is\ given\ by: \\ X_t(1-B)^d \end{array}\right\} \ldots [9]$$

Given the basic algebraic manipulations above, it can be inferred that when the actual data series is differenced "d" times before fitting an ARMA (p, q) process, then the model for the actual undifferenced series is called an ARIMA (p, d, q) model. Thus equation [7] is now generalized as follows:

$$\emptyset(B)(1-B)^d X_t = \theta(B)Z_t \ldots \ldots [10]$$

Therefore, in the case of modeling and forecasting NMR, equation [10] can be written as follows:

$$\emptyset(B)(1-B)^d X_t = \theta(B)Z_t \ldots \ldots [11]$$

The Box – Jenkins Approach

The first step towards model selection is to difference the series in order to achieve stationarity. Once this process is over, the researcher will then examine the correlogram in order to decide on the appropriate orders of the AR and MA components. It is important to highlight the fact that this procedure (of choosing the AR and MA components) is biased towards the use of personal judgement because there are no clear – cut rules on how to decide on the appropriate AR and MA components. Therefore, experience plays a pivotal role in this regard. The next step is the estimation of the tentative model, after which diagnostic testing shall follow. Diagnostic checking is usually done by generating the set of residuals and testing whether they satisfy the characteristics of a white noise process. If not, there would be need for model re – specification and repetition of the same process; this time from the second stage. The process may go on and on until an appropriate model is identified (Nyoni, 2018). The Box – Jenkins technique was proposed by Box & Jenkins (1970) and is widely used in many forecasting contexts, including health sector. In this paper, hinged on this technique; the researcher will use automatic ARIMA modeling for estimating equation [10].

International Research Journal of Innovations in Engineering and Technology (IRJIET)
ISSN (online): 2581-3048
Volume 7, Issue 8, pp 382-389, August-2023
https://doi.org/10.47001/IRJIET/2023.708057

Data Issues

This study is based on annual NMR in Nepal for the period 1960 to 2019. The out-of-sample forecast covers the period 2020 to 2030. All the data employed in this research paper was gathered from the World Bank online database.

Evaluation of ARIMA Models

Criteria Table

Table 2: Criteria Table

Model Selection Criteria Table
Dependent Variable: D(X)
Date: 01/29/22 Time: 10:25
Sample: 1960 2019
Included observations: 59

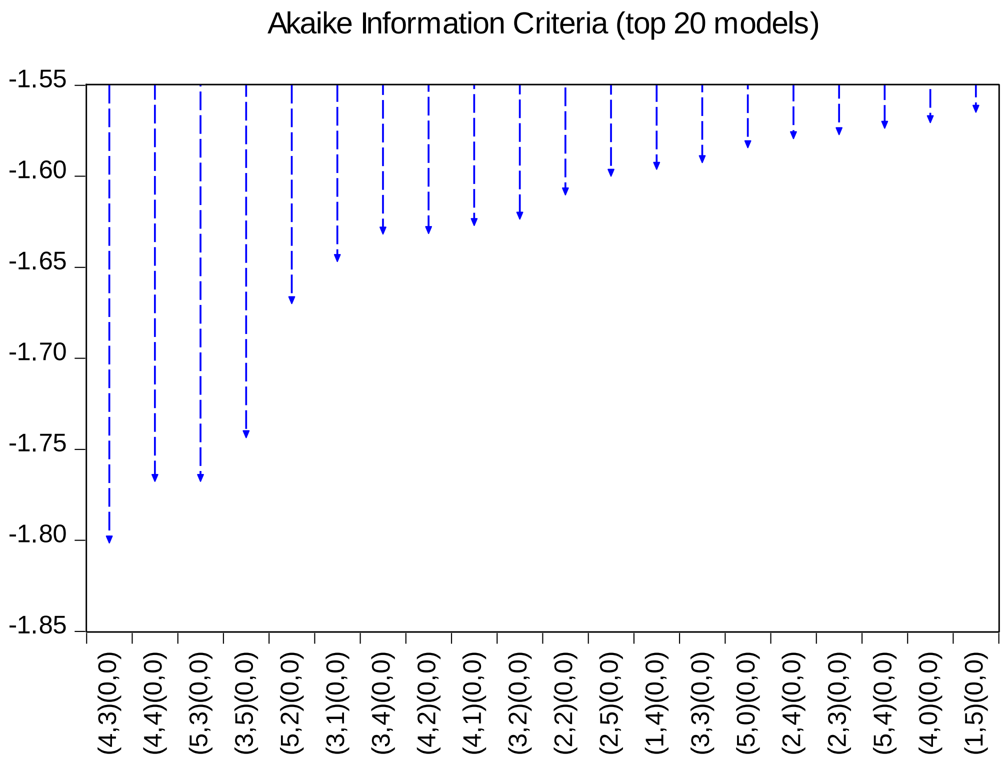
Model	LogL	AIC*	BIC	HQ
(4,3)(0,0)	62.077924	-1.799252	-1.482339	-1.675542
(4,4)(0,0)	62.080336	-1.765435	-1.413310	-1.627980
(5,3)(0,0)	62.079947	-1.765422	-1.413297	-1.627967
(3,5)(0,0)	61.370594	-1.741376	-1.389251	-1.603921
(5,2)(0,0)	58.200986	-1.667830	-1.350918	-1.544120
(3,1)(0,0)	54.515789	-1.644603	-1.433328	-1.562130
(3,4)(0,0)	57.073204	-1.629600	-1.312688	-1.505890
(4,2)(0,0)	56.063763	-1.629280	-1.347580	-1.519316
(4,1)(0,0)	54.934803	-1.624909	-1.378421	-1.528690
(3,2)(0,0)	54.830112	-1.621360	-1.374872	-1.525141
(2,2)(0,0)	53.438505	-1.608085	-1.396810	-1.525612
(2,5)(0,0)	56.134128	-1.597767	-1.280855	-1.474057
(1,4)(0,0)	54.023474	-1.594016	-1.347529	-1.497797
(3,3)(0,0)	54.916404	-1.590387	-1.308687	-1.480422
(5,0)(0,0)	53.673190	-1.582142	-1.335655	-1.485923
(2,4)(0,0)	54.525507	-1.577136	-1.295436	-1.467172
(2,3)(0,0)	53.462104	-1.574987	-1.328499	-1.478768
(5,4)(0,0)	57.359335	-1.571503	-1.184165	-1.420302
(4,0)(0,0)	52.267026	-1.568374	-1.357099	-1.485901
(1,5)(0,0)	54.098575	-1.562664	-1.280964	-1.452699
(3,0)(0,0)	50.806123	-1.552750	-1.376687	-1.484022
(5,1)(0,0)	53.745048	-1.550680	-1.268980	-1.440715
(1,3)(0,0)	51.567993	-1.544678	-1.333403	-1.462205
(1,2)(0,0)	50.351083	-1.537325	-1.361262	-1.468597
(1,0)(0,0)	46.466028	-1.473425	-1.367787	-1.432188
(5,5)(0,0)	55.068371	-1.459945	-1.037395	-1.294998
(2,1)(0,0)	48.027069	-1.458545	-1.282482	-1.389817
(2,0)(0,0)	46.969607	-1.456597	-1.315747	-1.401615
(1,1)(0,0)	46.750899	-1.449183	-1.308333	-1.394201
(4,5)(0,0)	53.611128	-1.444445	-1.057108	-1.293244
(0,5)(0,0)	40.022832	-1.119418	-0.872931	-1.023199
(0,4)(0,0)	34.743987	-0.974372	-0.763097	-0.891899
(0,3)(0,0)	27.158727	-0.751143	-0.575081	-0.682416
(0,2)(0,0)	17.259160	-0.449463	-0.308613	-0.394481
(0,1)(0,0)	0.120880	0.097597	0.203235	0.138834

ISSN (online): 2581-3048
Volume 7, Issue 8, pp 382-389, August-2023
https://doi.org/10.47001/IRJIET/2023.708057

(0,0)(0,0)	-29.798883	1.077928	1.148353	1.105419

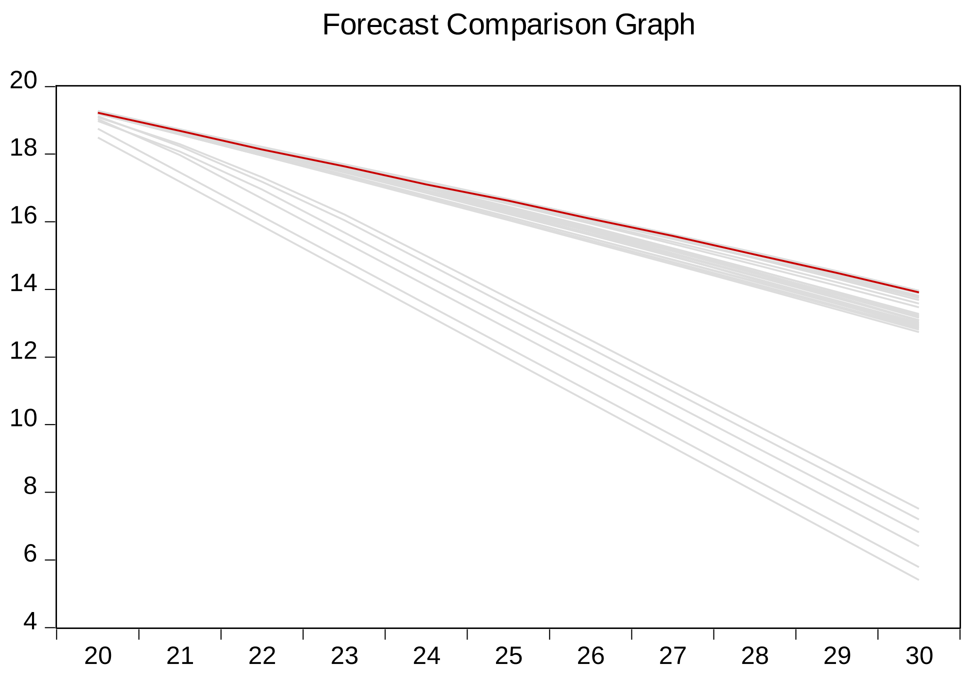
Criteria Graph

Figure 1: Criteria Graph



Forecast Comparison Graph

Figure 2: Forecast Comparison Graph



International Research Journal of Innovations in Engineering and Technology (IRJIET)
ISSN (online): 2581-3048
Volume 7, Issue 8, pp 382-389, August-2023
https://doi.org/10.47001/IRJIET/2023.708057

Table 2 and Figure 1 indicate that the optimal model is the ARIMA (4,1,3) model. Figure 2 is a combined forecast comparison graph showing the out-of-sample forecasts of the top 25 models evaluated based on the AIC criterion. The red line shows the forecast line graph of the optimal model, the ARIMA (4,1,3) model.

IV. RESULTS

Summary of the Selected ARIMA () Model

Table 3: Summary of the Optimal Model

Automatic ARIMA Forecasting
Selected dependent variable: D(X)
Date: 01/29/22 Time: 10:25
Sample: 1960 2019
Included observations: 59
Forecast length: 11

Number of estimated ARMA models: 36
Number of non-converged estimations: 0
Selected ARMA model: (4,3)(0,0)
AIC value: -1.79925165146

Main Results of the Selected ARIMA () Model

Table 4: Main Results of the Optimal Model

Dependent Variable: D(X)
Method: ARMA Maximum Likelihood (BFGS)
Date: 01/29/22 Time: 10:25
Sample: 1961 2019
Included observations: 59
Failure to improve objective (non-zero gradients) after 86 iterations
Coefficient covariance computed using outer product of gradients

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-1.163480	0.183055	-6.355885	0.0000
AR(1)	0.145146	0.065795	2.206051	0.0320
AR(2)	1.723696	0.089517	19.25558	0.0000
AR(3)	0.010831	0.064696	0.167421	0.8677
AR(4)	-0.908739	0.068778	-13.21272	0.0000
MA(1)	0.866561	427.4365	0.002027	0.9984
MA(2)	-0.999977	271.0160	-0.003690	0.9971
MA(3)	-0.866540	596.8523	-0.001452	0.9988
SIGMASQ	0.005771	0.395105	0.014605	0.9884

R-squared	0.964108	Mean dependent var	-1.308475
Adjusted R-squared	0.958365	S.D. dependent var	0.404410
S.E. of regression	0.082519	Akaike info criterion	-1.799252
Sum squared resid	0.340467	Schwarz criterion	-1.482340
Log likelihood	62.07794	Hannan-Quinn criter.	-1.675542
F-statistic	167.8813	Durbin-Watson stat	2.055163
Prob(F-statistic)	0.000000		

ISSN (online): 2581-3048
Volume 7, Issue 8, pp 382-389, August-2023
https://doi.org/10.47001/IRJIET/2023.708057

Inverted AR Roots	.99+.09i	.99-.09i	-.92+.28i	-.92-.28i
Inverted MA Roots	1.00	-.87	-1.00	

ARIMA () Model Forecast

Tabulated Out of Sample Forecasts

Table 5: Tabulated Out of Sample Forecasts

2020	19.22291711154258
2021	18.68940112851589
2022	18.13756922763609
2023	17.643029559047
2024	17.10488072160438
2025	16.61936728198398
2026	16.08358961344497
2027	15.57870769608321
2028	15.03186936898864
2029	14.4838198343136
2030	13.90928610887946

Table 2 clearly indicates that neonatal mortality will gradually decline from around 19 in 2020 to approximately 14 deaths per 1000 live births by 2030.

V. POLICY IMPLICATION & CONCLUSION

Surveillance mechanisms are essential in public health programming as they facilitate planning, decision making and allocation of resources. In this study we proposed the ARIMA model to predict future trends of NMR for Nepal and the findings indicate that neonatal mortality will gradually decline from around 19 in 2020 to approximately 14 deaths per 1000 live births by 2030. Therefore, we encourage the government of Nepal to design local policies to address various maternal and child health program challenges to keep neonatal deaths under control.

REFERENCES

[1] Box, D. E., and Jenkins, G. M. (1970). Time Series Analysis, Forecasting and Control, Holden Day, London.
[2] Nyoni, T. (2018). Box-Jenkins ARIMA Approach to Predicting net FDI Inflows in Zimbabwe, *University Library of Munich*, MPRA Paper No. 87737.
[3] UNICEF. (2019). Levels and trends in child mortality: report 2019. Estimates developed by the UN Inter-agency Group for child mortality estimation. New York: UNICEF.
[4] United Nations. (2015). transforming our world: The 2030 agenda for sustainable development, A/RES/70/1. New York: UN General Assembly.
[5] Kc A., Jha A. K., Shrestha M. P., Zhou H., Gurung A ., Thapa J., Budhathoki S.S (2020). Trends for Neonatal Deaths in Nepal (2001–2016) to Project Progress towards the SDG Target in 2030, and Risk Factor Analyses to Focus Action. Maternal and Child Health Journal, 2020, 24 (Suppl 1), S5–S14.
[6] Ministry of Health and Population (MOHP). Nepal demographic and health survey. Kathmandu, 2017.
[7] World Bank (2019). Mortality rate, under 5.
[8] WHO MCEE. MCEE-WHO, 2018.
[9] UN (2020) sustainable development goals. https://www.un.org/sustainabl development/development-agenda
[10] UNICEF (2018). Every Child alive. New York: UNICEF
[11] World Health Organization (WHO) (2019). SDG 3: Ensure healthy lives and promote wellbeing for all at all ages.

International Research Journal of Innovations in Engineering and Technology (IRJIET)

ISSN (online): 2581-3048

Volume 7, Issue 8, pp 382-389, August-2023

https://doi.org/10.47001/IRJIET/2023.708057

Citation of this Article:

Dr. Smartson. P. NYONI, Thabani NYONI, "Generating Reliable and Accurate Forecasts of Annual Neonatal Mortality Rate for Nepal Using the ARIMA Model" Published in *International Research Journal of Innovations in Engineering and Technology - IRJIET*, Volume 7, Issue 8, pp 382-389, August 2023. Article DOI https://doi.org/10.47001/IRJIET/2023.708057

****