

Prediction of Unemployment Rates with Time Series and Machine Learning Techniques

Fa'iz Jibia Ibrahim

Department of Computer Science
Bayero University, Kano
Kano, Nigeria
fjibia@fptb.edu.ng

Hadiza Ali Umar

Department of Computer Science
Bayero University, Kano
Kano, Nigeria
haumar.cs@buk.edu.ng

Abdulkadir Shehu Bichi

Department of Computer Science
Bayero University, Kano
Kano, Nigeria
abdulkadirshehubichi@gmail.com

Ibrahim Said Ahmad

Department of Information Technology
Bayero University, Kano
Kano, Nigeria isahmad.it@buk.edu.ng

Nur Bala Rabi

Department of Computer Bayero
University, Kano
Kano, Nigeria
nurbashir2008@gmail.com

Aisha Mustapha Ahmad

Department of Computer Science
Bayero University, Kano
Kano, Nigeria
amahmad.cs@buk.edu.ng

Abstract— Unemployment represents a prevailing socio-economic challenge in every nation across the globe, exerting influence over individuals' living standards and a country's socio-economic position. Recent years have witnessed a surge in scholarly attention towards assessing unemployment rates. This study aims to introduce hybrid models for appraising Nigeria's unemployment rate. The accurate prediction of unemployment rates holds paramount significance, serving as a pivotal tool employed by experts and policymakers in sculpting a nation's economic and financial progress. In recent times, certain researchers have employed conventional analytical methodologies to model Nigeria's unemployment rate. However, it is crucial to note that unemployment datasets exhibit nonstationary behavior. Consequently, employing traditional techniques to model such time series frequently leads to arbitrary errors. To surmount the accuracy hurdles associated with conventional techniques, this inquiry advocates for an intelligent-based approach that dissects both the linear and nonlinear components of the time series. The study harnesses unemployment data obtained from the National Bureau of Statistics and the Macrotrend website to construct and evaluate the models. Precision of the models was assessed using MAE, MAPE, and RMSE evaluation metrics. Impressively, the hybrid ARIMA_ANN model displayed superior performance on the NBS dataset, while the hybrid ARIMA_SVM model outperformed the other models in the context of the Macrotrend dataset.

Keywords: Prediction, ARIMA, hybrid model, Time Series

I. INTRODUCTION

Time series prediction stands as a significant area of forecasting, involving the collection and analysis of past instances of a given variable to construct a model that represents its underlying connection. This model is then employed to extrapolate the time series into upcoming periods (Zhang, 2003) as cited in (Hewamalage, Bergmeir, Bandara, 2021). In recent decades, the research community

has invested substantial efforts to create more accurate models for forecasting, resulting in the introduction of several notable timeseries prediction models.

Unemployment constitutes a societal occurrence in which certain capable individuals seeking employment are unable to secure jobs (Begg, David KH and Fischer, Stanley and Dornbusch, Rudiger, 2007; Klos, Alicja, Tomaszewska, and Katarzyna, 2014; Kwiatkowski, Eugeniusz, 2002) as cited in (Gostkowski and Rokicki, 2021). Unemployment serves as a significant macroeconomic indicator that characterizes a nation's economic stability and influences social progress (Dritsakis & Klazoglou, 2018). Several factors, among others, have contributed to the escalation of unemployment in Nigeria. Firstly, structural elements involve the mismatch between educational institutions and labor market demands, the implementation of capital-intensive production techniques due to technological advancements, the composition of the workforce's skills, and the availability of job opportunities. Secondly, cyclical factors encompass the inadequate aggregate demand for goods and services, both domestically and internationally. Lastly, institutional factors include the influence of the World Bank and IMF during the 1980s, advising developing countries to downsize their public sector and governmental services (International Labour Organization, 2005).

The elevated level of unemployment has direct and ripple effects on poverty, crime, and socio-political issues, all of which are concurrently on the rise (Nikolaos et al., 2018). Unemployment is particularly rampant in Nigeria, with a significant presence among the youth population (Samuel, Samuel, John, Magdalene and Kajuru, 2021). According to the National Bureau of Statistics (NBS), as stated in its Q2 report in August 2020, Nigeria's unemployment rate has surged from 23.1% in 2018 to 27.1% in the second quarter of 2020. This indicates that approximately 21,764,617 (twenty-one million, seven hundred and sixty-four thousand, six hundred and seventeen) Nigerians are currently without jobs (National Bureau of Statistics, 2020).

Nonetheless, despite the increasing prevalence of unemployment in Nigeria, the utilization of the country's unemployment data has not been evaluated with the latest technological advancements, as evidenced by the examination undertaken by this researcher based on the available literature. Given this context, there exists a growing necessity to adopt a hybrid modeling approach to assess Nigeria's unemployment rate. Such an approach aims to establish a precise model that policymakers and other stakeholders can employ for effective forecasting and strategic planning to address the issue of unemployment in the nation. The primary contribution of this study revolves around the investigation of hybrid methodologies, which include a novel model designed to probe the interplay between the linear and nonlinear elements within the unemployment time series data. Consequently, a suitable model is identified for forecasting the unemployment rate in Nigeria.

The structure of this research is delineated as follows: Section 2 delves into the relevant prior research. In Section 3, the methodology employed in this study is expounded upon, encompassing an in-depth discussion of the hybrid modeling techniques developed. The ensuing section, Section 4, presents the outcomes and ensuing discussions. Finally, Section 5 concludes the report by summarizing key findings and outlining potential avenues for future research.

II. RELATED WORKS

The prediction of unemployment rate has been extensively investigated in various research studies, utilizing statistical models, machine-learning models, and hybrid models. This section presents an overview of the diverse methods employed to forecast the unemployment rate and other macroeconomic indicators.

A. Time Series Models

Adesegun, Mathew and Omotola (2020) conducted a study on Nigeria's annual unemployment rate using an ARIMA model. They assessed the trend through time plots and employed the KPSS and ADF unit root tests to test for stationarity. The order of the model was determined using ACF and PACF. The ARIMA (1, 1, 1) model was identified as the most suitable for the dataset, and a five-year forecast of the unemployment rate was generated. In a subsequent study, Samuel et al., (2021) proposed an ARIMA model for Nigeria's annual unemployment data, utilizing the Box-Jenkins' technique. They found that the ARIMA (2, 2, 2) model was the best fit for the time series data and used it to forecast the unemployment rate five years ahead. Model performance accuracy was assessed using RMSE, MAE, and MAPE. Similarly, Sharma and Soni (2021) applied ARIMA models with three exponential smoothing methods to India's youth unemployment rate data. Time series cross-validation was employed to determine the best model based on mean square error. The ARIMA (1, 1, 0) model outperformed other models in forecasting the values. In another study, Ghazo (2021) used the Box-Jenkins' methodology to analyze

Jordanian GDP and CPI data from 1976 to 2019. They selected ARIMA (3, 1, 1) and ARIMA (1, 1, 0) as the optimal models to forecast GDP and CPI three years ahead. Subsequently, Ismail, Nur Afiah and Ramzi, Nurin Alya and Wee Mah, Pauline Jin (2022) fitted an ARIMA (p, d, q) model and an ARFIMA (p, d, q) model to Malaysian monthly unemployment rate data from January 2010 to December 2020. The Sample ACF plot indicated the presence of long memory in the data. The ARIMA (2, 1, 2) model was identified as the best model, while ARFIMA (0, -0.2339, 0) was found to be the best ARFIMA model. Model comparison based on MAPE, MAE, and RMSE values showed that ARFIMA (0, -0.2339, 0) outperformed the ARIMA (2, 1, 2) model.

B. Machine Learning Approaches

Aiken (1996) as cited in (Mulaudzi and Ajoodha, 2020) was among the first researchers to showcase the applicability of neural networks for forecasting unemployment. His study aimed to address issues related to the accuracy and reliability of traditional macroeconomic methods used for predicting unemployment. One significant problem he highlighted was that conventional forecasting techniques required specific data conditions to be met, such as constant variance, constant mean, and no correlations among independent variables. In a different context, Comert and Yildiz (2022) proposed an improved Artificial Neural Network (ANN) model for forecasting Turkish electricity demand. The model comprised an input layer with 15 neurons (representing the year, 12 months, temperature, and unemployment rate), two hidden layers, and an output layer. Compared to previous models, their approach was notably successful in forecasting Turkey's electricity demand. In the same year, Karahan and Cetintas (2022) reported the effectiveness of another ANN model for predicting the future unemployment rate in Turkey. Their model demonstrated a degree of accuracy and provided reliable and consistent forecasts. It utilized five independent variables (inflation, employment rate, exchange rate, current deficit, industrial production) and one dependent variable, which was the monthly unemployment rate.

In a similar study, Kreiner and Duca (2019) utilized LASSO regression to predict unemployment rates in the USA. This regression model minimizes prediction error through regularization, effectively reducing overfitting by penalizing the reliance on specific variables for prediction. Researchers Smalter Hall and Cook (2017) and Katris (2019a) demonstrated the effectiveness of Support Vector Regression (SVR) with nonlinear decision boundaries in forecasting unemployment rates. SVRs outperformed linear regression in terms of accuracy. In another study, Diana, Budiantara, Purhadi and Darmesto (2014) employed a multivariate regression model to forecast Indonesia's unemployment. Their results were similar to those of Katris (2019b), showing an improvement in forecast accuracy. However, their implementation was semi-parametric, using Bayesian techniques to determine forecasting confidence intervals. They emphasized that semi-parametric multivariable regression models offer an advantage over standard non-parametric MARS approaches by not assuming a specific regression function form. Moreover, Sagheer & Kotb (2018)

proposed a deep long-short term memory (DLSTM) architecture, an extension of the traditional recurrent neural network, for forecasting petroleum production. They optimized DLSTM's architecture using a genetic algorithm. Two case studies from the petroleum industry domain validated the superiority of the proposed DLSTM model over standard approaches in forecasting performance.

C. Hybrid Modelling Approach

Lai, Khan, Thaljaoui, Chammam and Abbas (2021) developed three unique hybrid models to investigate the impact of COVID-19 on the unemployment rate in selected developed and developing countries in Asia. These hybrid models combined the widely used time series model, ARIMA, with either ANN, ARNN, or SVM models. The results indicated that in developing countries of Asia, the unemployment rate is projected to be three times higher compared to advanced countries in the upcoming years. Additionally, it will take double the time to recover from the impacts of the Coronavirus in developing countries than in developed countries in Asia. Similarly, Amin, Ahsan, Chowdhury, Amin, Islam (2021) analyzed the impact of the COVID-19 pandemic on unemployment in six European countries. They also employed three hybrid models: ARIMA-ARIMA, ARIMA-ANN, and ARIMA-SVM. The outcome revealed that the unemployment rate will continue to rise in the coming years as a consequence of the COVID-19 pandemic. It is expected to take a minimum of five years for these countries to overcome the effects of COVID-19 on unemployment. In a prior study, Chakraborty, Chakraborty, Biswas, Banerjee and Bhattacharya (2020) used another hybrid approach, ARIMA-ARNN, to forecast unemployment rates for seven countries (Canada, Germany, Japan, Netherlands, New Zealand, Sweden, and Switzerland). Furthermore, Shi, Khan and Tian (2022) proposed three hybrid models: ARIMA-ANN, ARIMA-ARNN, and ARIMA-SVM, to assess the impact of the COVID-19 pandemic on the unemployment rate in selected Asian countries. According to their findings, unemployment is expected to rise significantly in developing economies, with a particularly acute impact on the region's less-developed economies.

Ahmad, Khan, Jiang, Kazmi and Abbas (2021) proposed an intelligent hybrid modeling approach to predict the unemployment rate in selected European countries. They found that the hybrid ARIMA-ARNN model performed best for France, Belgium, Turkey, and Germany, while the hybrid ARIMA-SVM model outperformed others for Italy and Spain. In another study, Dave, Leonardo, Jeanice and Hanafiah (2021) used a hybrid ARIMA-LSTM model to predict Indonesia's future exports. The hybrid model integrated ARIMA and LSTM models, with LSTM handling the nonlinear components of the data and ARIMA handling the linear components. The (LSTM-ARIMA) hybrid model achieved the lowest error metrics among all evaluated models, outperforming other independent models with a MAPE value of 7.38%. Furthermore, Ji, Zou, He and Zhu, (2019) applied an ARIMA-CNN-LSTM model to forecast carbon futures prices. The ARIMA component captured linear features, the Convolutional

Neural Network (CNN) captured hierarchical data structures, and the Long Short-Term Memory network (LSTM) captured long-term dependencies in the data. The ARIMA-CNN-LSTM model demonstrated better prediction accuracy than benchmark models, as evidenced by improved Root Mean Square Error (RMSE) and Mean Absolute Percentage Error (MAPE) performance measures.

III. RESEARCH METHODOLOGY

On an economic level, the unemployment rate mirrors the extent of joblessness. It can be defined as the overall proportion of unemployed individuals within the labor force, presenting a comprehensive gauge. In this research, quarterly unemployment rates data in Nigeria were employed. Table 3.1 encapsulates the unemployment rate datasets for two distinct periods, while their historical depictions are visualized in Fig. 3.1 and Fig. 3.2. The graphical representation of the unemployment rate affirms the existence of nonstationary and nonlinear patterns within the data. To confirm the nonlinearity, a linearity test was employed, encompassing a broad range of nonlinearity (Chakraborty and Ghosh, 2020). This study concentrates on single-variable approaches for forecasting the unemployment rate. By amalgamating linear and nonlinear methods that account for specific data attributes, improved prediction accuracy was attained.

A Methodology

In order to forecast the unemployment rate, this study utilized hybrid models that focus specifically on ARIMA, ANN, SVM and LSTM.

1. ARIMA Model

ARIMA stands for Autoregressive Integrated Moving Average, a linear model for stationary time series data capturing linear trends. Noted as ARIMA(p, d, q), where p and q signify autoregressive (AR) and moving average (MA) model orders, and d is differencing level. The model is expressed as follows:

$$y_t = \alpha_0 + \sum_{i=1}^p \theta_i x_{t-i} + \varepsilon_t + \sum_{j=1}^q \alpha_j \varepsilon_{t-j} \quad (1)$$

Where y_t demonstrates the veritable assessment of the variable feasible at point t, ε is the irregular mistake at point t [i, and α_j are the model's coefficients.

Optimal ARIMA orders are found via the Auto ARIMA function, which exhaustively explores parameter combinations within a range, choosing the model with the highest AIC (Akaike Information Criterion). Prior to applying Auto ARIMA, we decomposed data seasonally to detect patterns. Graphical analysis revealed clear seasonality, setting the seasonal parameter to true in Auto ARIMA. ARIMA(p, d, q) model fitting used the .fit() method in Jupyter Notebook, selecting the highest AIC model as the final one.

2. Artificial Neural Networks (ANNs)

The Artificial Neural Networks (ANNs) method has emerged as an alternative technique for time series prediction and has gained substantial recognition in recent times. The fundamentals of ANNs aim to replicate human brain intelligence in machines (Kihoro, Otieno, and Wafula, 2004 as cited in Adeyinka and Muhajarine, 2020). ANNs comprise interconnected artificial neural cells, which are mathematical renditions of biological neurons. These neural cells, termed neurons, consist of basic elements. Signals move through neuron connections, each with a weight value. Neurons generate network outputs, processing network inputs via an activation function (Hamzacebi, 2011). The model is composed of three layers of basic to-give units linked by a repeating connection. The interaction among the produce (z_t) and the inputs ($z_{t-1}, z_{t-2}, \dots, z_{t-p}$) has the following mathematical illustrations:

$$z = \alpha_0 + \sum_{j=1}^q \alpha_j g \left(\beta_{0j} + \sum_{i=1}^p \beta_{ij} x_{t-i} \right) + \epsilon_t \quad (2)$$

where j ($j = 0, 1, 2, \dots, q$) and β_{ij} ($i = 0, 1, 2, \dots, p; j = 0, 1, 2, \dots, q$), here p is the number of information hubs and q the number of hidden hubs. In other words, the strategic capacity is frequently utilized as the concealed layer move work, that is

$$f(x) = \frac{1}{1 + \exp(-x)}$$

Hence, the ANN model of Equation 2 performs a nonlinear functional mapping from the past observations ($z_{t-1}; z_{t-2}; \dots; z_{t-p}$) to the future value z_t , i.e.

$$z_t = g(z_{t-1}, z_{t-2}, \dots, z_{t-n}, w) + \epsilon_t$$

Where w is a vector of all parameters and f is a function defined by the configuration of the network and the weights of the connection. Thus, a nonlinear autoregressive model is identical to the neural network.

Table 1: Representation of Informational Indexes for the Unemployment Rate

	Unemployment Data (NBS)	Unemployment Data (Macrotrend)
Observations	244	124
Training	195	99
Testing	49	25
Max Value	33.3	9.81
Min Value	1.9	3.69

3. Support Vector Machines (SVMs)

Support Vector Machines (SVMs) stand as a supervised learning method, enabling the classification of non-linear data via a hyperplane (Sermpinis, Stasinakis, Theofilatos, and Karathanasopoulos, 2014). This approach assigns class labels to input data, effectively categorizing them into distinct classes and proving valuable for handling non-linear data. SVMs are divided into two types: support vector classifiers (SVC) and support vector regression (SVR). In regression tasks, SVR employs the ϵ -sensitive loss function to facilitate non-linear regression modeling (Katriss, 2019a).

$$f(x) = \sum_{i=1}^n (\alpha_i - \alpha_i^*) K(x_i, x) + b. \quad (3)$$

4. Long Short Term Memory (LSTM)

LSTM possesses a sophisticated architecture, primarily found in the hidden layer, known as the memory block. This block can consist of multiple memory cells, each housing a recurrent self-connected linear unit called the Constant Error Carousel (CEC). The CEC effectively addresses the vanishing gradient problem. Positioned alongside the RNN cell, the memory cell facilitates the accumulation and transfer of memories from previous time steps to the subsequent ones. This distinctive feature of LSTM makes it renowned for handling long-term dependencies, rendering it an ideal choice for time series data (Dave et al., 2021).

5. Hybrid Modelling Approach

Hybrid modeling is a two-step process designed to tackle nonstationary time series data in two stages. Firstly, traditional ARIMA (p, d, q) captures linear trends in the time series. In the second step, advanced techniques like ANN, SVM, or LSTM model residuals for nonlinear aspects. Both step results combine for final estimate. In the modern era of big data, the application of intelligent-based hybrid approaches for large-scale data is a popular research area (Li et al. 2019). Utilizing ARIMA with Machine Learning algorithms (ANN, SVM, and LSTM) efficiently enhances prediction accuracy. This study employs hybrid ARIMA-ANN, hybrid ARIMA-SVM, and hybrid ARIMA-LSTM models to evaluate the unemployment rate for Nigeria. Below is the outline of the hybrid modelling algorithm.

Hybrid Modelling Algorithm for the Prediction of Unemployment rate

Step 1: Start

- Divide the data in to two groups: Training set and Testing set

Step 2 Accomplished: Determined optimal ARIMA (p, d, q) by means of training data

- ✓ Parameters p, d & q of ARIMA are determined by using `auto_arma` function in python.

- ✓ Predictions of ARIMA (p,d,q) are obtained using training data
- ✓ Obtain residuals of the ARIMA prediction
- ✓ Acquire best ANN (p, k)

Step 3 Performed: Obtain the best ANN (p,k) model from the training dataset residuals

- ✓ Performed lag selection on the residuals of training data and then apply ANN model with p selected lagged input from residuals and k hidden units.
- ✓ Acquire the residuals prediction using ANN model

Step 4: Final Prediction of Unemployment rate: Combine both the predictions of ARIMA and ANN models to get the final prediction.

Step-5: Replication: repeat step 3–4 by initiating SVM and LSTM.

B. Data Acquisition and Preprocessing

The data employed in this study was collected from two main sources. The initial dataset was collected and prepared using labor force survey bulletins and additional economic indicators released by the National Bureau of Statistics (NBS). The second dataset was acquired from the popular economic statistics platform, *macrotrends.net*. The experimentation was conducted within Jupyter Notebook, an interactive computational environment accessible via the Anaconda Package. The visual representation of the two datasets is depicted in Figures 3.1 and 3.2 below.

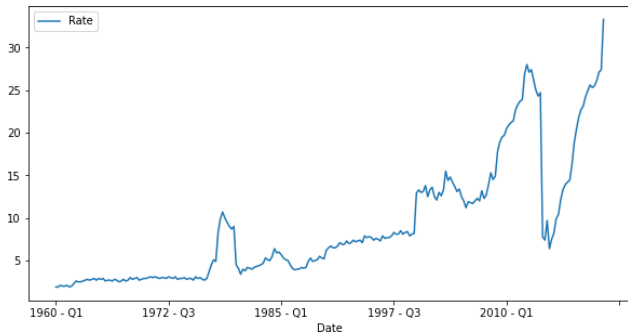


Fig 1: Visualization of the observed unemployment rate (NBS)

As illustrated in **Fig. 1** above, Nigeria's unemployment rate trajectory (sourced from the National Bureau of Statistics) spanned from Q1 1960 to Q4 2020, comprising 244 data points. On the graph, the vertical axis signifies unemployment rates, while the horizontal axis represents dates. Initially, figures remained relatively stable until approximately Q4 1972, when they began to ascend, followed by an immediate decline in the following quarter. The rates then alternated between increases and decreases until the last data point, which marks the highest point. By observation and verified through the ADF test, the time series data appears non-stationary.

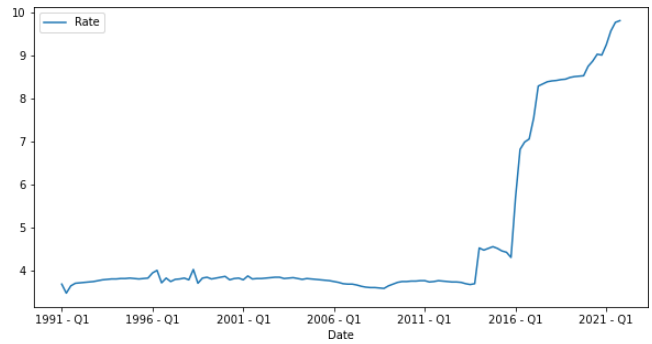


Fig 2: Visualization of the observed unemployment rate (Macrotrend)

As shown in **Fig. 2** above, Nigeria's unemployment rate trend (as gathered from *macrotrend.com*) spanned from Q1 1991 to Q4 2021, encompassing 124 data points. The vertical axis signifies unemployment rates, while the horizontal axis represents dates. Initially, the time series data showed relative stability from Q1 1991 to approximately Q4 2014. From there, it started to increase and maintained this trajectory until Q4 2021, marking the highest recorded point. By observation and validated by the ADF statistical test, the time series data appears to exhibit non-stationary behavior.

a. Performance Evaluation Metrics

Forecasting precision acts as a yardstick to assess forecasting model efficacy. It evaluates predictive capability and data fit. In this research, models' predictive ability was assessed using three error metrics: Root Mean Squared Error (RMSE), Mean Absolute Error (MAE), and Mean Absolute Percent Error (MAPE). A model with lower error metrics indicates better forecasting performance.

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2}; \quad (4)$$

$$MAE = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i|; \quad (5)$$

$$MAPE = \frac{1}{n} \sum_{i=1}^n \left| \frac{y_i - \hat{y}_i}{y_i} \right|; \quad (6)$$

RMSE quantifies predicted vs. actual series discrepancies, providing a measure for gauging predictive strength. MAE on the other hand calculates average absolute errors, assessing proximity of predictions to actual data. MAPE expresses accuracy as error percentage. Thus, a superior forecasting model demonstrates diminished RMSE, MAE, and MAPE values.

IV. RESULTS AND DISCUSSIONS

The underlined time series exhibit both non-stationary and non-linear characteristics, as confirmed. Figures 4.1 and 4.2 illustrate the nonlinearity within these datasets.

This research utilized the hybrid ARIMA-ANN model introduced by (Lai et al., 2021), the hybrid ARIMA-SVM model proposed by (Shi et al., 2022), and the novel hybrid ARIMA-LSTM model to analyze and contrast the chosen study data. Initially, the conventional ARIMA (p, d, q) model was implemented using the .fit () method within Jupyter Notebook. To determine p, d, and q, the auto_arima function of the pmdarima package was utilized. The top ARIMA model with the highest AIC was chosen. Additionally, residual errors were generated using predicted training results. The training and testing dataset split was 80% and 20%, respectively.

In the subsequent stage, ARIMA residuals were incorporated into ANN (p,k) models, employing one hidden layer with 'kn' neurons for simplified calculations. ARIMA (1, 0, 0) and ARIMA (2, 0, 0) had the highest AIC values of 160.763 and 158.488 for the two datasets, respectively. Residuals were integrated into non-linear models. Hybrid ARIMA-ANN, ARIMA-SVM, and ARIMA-LSTM models were then developed for future prediction using testing observations. Model accuracy was assessed using MAE, MAPE, and RMSE metrics. The outcomes of proposed models on the distinct datasets are summarized in **Table 2** and **Table 3**.

Table 2: Result of NBS Dataset

Models	MAE	RMSE	MAPE
ARIMA_ANN	5.52	6.99	0.62
ARIMA_SVM	5.58	8.09	0.82
ARIMA_LSTM	7.62	8.46	0.41

As shown in Table 4.1, the hybrid models demonstrated generally satisfactory performance. However, the foundational ARIMA_ANN and ARIMA_SVM models marginally surpassed the novel model in most evaluation measures, except for Mean Average Percentage Error (MAPE), in which the new model outperformed the two baseline models.

Table 3: Result of Macrotrend Dataset

Models	MAE	RMSE	MAPE
ARIMA_ANN	1.41	1.88	0.30
ARIMA_SVM	0.36	0.95	0.09
ARIMA_LSTM	4.69	4.77	0.55

As per the outcomes in Table 4.3 for the Macrotrend dataset, the baseline models exhibited superior performance compared to the new model, with a noticeable difference. This could be

attributed to the lesser number of observations in the dataset compared to the other.

V. CONCLUSION AND FUTURE WORK

The aim of this research is to develop hybrid models to evaluate the unemployment rate in Nigeria. The findings of the study demonstrated that, the hybrid ARIMA_ANN model is the most suitable for forecasting Nigeria's unemployment rate. The study also revealed that the two datasets utilized exhibited clear characteristics of non-stationary and nonlinear time series.

This study's significance for economic policy formulation is substantial. Accurate unemployment rate forecasts empower the government to effectively address unemployment issues and plan for the well-being of jobless individuals, particularly youth who influence the nation's future.

In terms of academic contributions, this study evaluated Nigeria's unemployment rate using three distinct hybrid models: ARIMA_ANN, ARIMA_SVM, and ARIMA_LSTM.

Future Work

In spite of the significance of the results attained in this research, which overcome the challenges of model selection between linear and nonlinear models in the forecast Nigeria's unemployment rate. Thus, this research opens an avenue of directions for future work to explore on new models that can give better forecast accuracy. Likewise, the performance of the models can be improved using optimization algorithms (like PSO, crow search, etc.).

REFERENCES

- [1] Adesegun, O. I., Mathew, O. O., & Omotola, S. B. (2020). Modeling Unemployment Rates in Nigeria using Time Series Approach.
- [2] Adeyinka, D. A., & Muhajarine, N. (2020). Time series prediction of under-five mortality rates for Nigeria: comparative analysis of artificial neural networks, Holt-Winters exponential smoothing and autoregressive integrated moving average models. *BMC medical research methodology*, 20, 1-11.
- [3] Ahmad, M. (2021). The impact of COVID-19 on unemployment rate : An intelligent based unemployment rate prediction in selected countries of Europe. November 2020, 1–16. <https://doi.org/10.1002/ijfe.2434>.
- [4] Aiken, M. (1996). A neural network to predict civilian unemployment rates. *Journal of International Information Management*, 5(1). <https://scholarworks.lib.csusb.edu/jiim/vol5/iss1/3>
- [5] Begg, D. K., Fischer, S., & Dornbusch, R. (2007). *Mikroekonomia*. Polskie Wydawnictwo Ekonomiczne.
- [6] Boccaletti Stefano, Ditto William, Mindlin Gabriel, A. A. (2020). Chaos , Solitons and Fractals Modeling and forecasting of epidemic spreading : The case of Covid-19 and beyond. 135.
- [7] Castillo, O., & Melin, P. (2021). A Novel Method for a COVID-19 Classification of Countries Based on an Intelligent Fuzzy Fractal Approach.
- [8] Chakraborty, T., & Ghosh, I. (2020). Real-time forecasts and risk assessment of novel coronavirus (COVID-19) cases: A data-driven analysis. 1–18.
- [9] Chakraborty, T., Kumar, A., & Munmun, C. (2020). Unemployment Rate Forecasting: A Hybrid Approach. *Computational Economics*, 0123456789. <https://doi.org/10.1007/s10614-020-10040-2>
- [10] Comert, M., & Yildiz, A. (2022). A novel artificial neural network model for forecasting electricity demand enhanced with population-weighted

- temperature mean and the unemployment rate. 6(2), 178–189. <https://doi.org/10.31127/tuje.903876>
- [11] Dave, E., Leonardo, A., Jeanice, M., & Hanafiah, N. (2021). Forecasting Indonesia Exports using a Hybrid Model ARIMA-LSTM. *Procedia Computer Science*, 179(2020), 480–487. <https://doi.org/10.1016/j.procs.2021.01.031>
- [12] Diana, R., Budiantara, I. N., & Darmesto, S. (2014). Statistical modeling for unemployment rate using smoothing spline in semiparametric multivariable regression model with Bayesian approach. 9, 159–166. <https://doi.org/10.3233/MAS-130287>
- [13] Dritsakis, N., & Klazoglou, P. (2018). Forecasting Unemployment Rates in USA Using Box-Jenkins Methodology. *International Journal of Economics and Financial Issues*, 8(1), 9–20. <http://www.econjournals.com>
- [14] Ghazo, A. (2021). Applying the ARIMA Model to the Process of Forecasting GDP and CPI in the Jordanian Economy. 12(3), 70–77. <https://doi.org/10.5430/ijfr.v12n3p70>
- [15] Hamzaçebi, C., & Pekkaya, M. (2011). Determining of stock investments with grey relational analysis. *Expert Systems with Applications*, 38(8), 9186–9195.
- [16] Hewamalage, H., Bergmeir, C., & Bandara, K. (2021). Recurrent Neural Networks for Time Series Forecasting: Current status and future directions. *International Journal of Forecasting*, 37(1), 388–427. <https://doi.org/10.1016/j.ijforecast.2020.06.008>
- [17] I. Sozen, & M. B. T. (2021). Forecasting Unemployment Rate in the Aftermath of the Covid-19 Pandemic: The Turkish Case. 0–3. <https://doi.org/10.24988/ije.202136312>
- [18] Ji, L., Zou, Y., He, K., & Zhu, B. (2019). Carbon futures price forecasting based with ARIMA-CNN-LSTM model. *Procedia Computer Science*, 162(Itqm 2019), 33–38. <https://doi.org/10.1016/j.procs.2019.11.254>
- [19] Kabir, A. (2021). An Intellectual Based Unemployment Rate Forecast in Preferred Countries of An Intellectual Based Unemployment Rate Forecast in Preferred Countries of Europe. February.
- [20] Katris, C. (2019a). Forecasting the Unemployment of Med Counties using Time Series and Neural Network models. 8(2), 37–49.
- [21] Katris, C. (2019b). Prediction of Unemployment Rates with Time Series and Machine Learning Techniques. *Computational Economics*. <https://doi.org/10.1007/s10614-019-09908-9>
- [22] Kihoro, J., Otieno, R. O., & Wafula, C. (2004). Seasonal time series forecasting: A comparative study of ARIMA and ANN models.
- [23] Kłós, A., & Tomaszewska, K. (2014). Bezrobocie jako zjawisko społeczne stanowiące zagrożenie dla kształtowania się rynku pracy w Polsce. *Społeczeństwo. Edukacja. Język*, 2.
- [24] Kreiner, A., & Duca, J. V. (2019). Can machine learning on economic data better forecast the unemployment rate? *Applied Economics Letters*, 00(00), 1–4. <https://doi.org/10.1080/13504851.2019.1688237>
- [25] Kwiatkowski, E., Kucharski, L., & Tokarski, T. (2002). Bezrobocie i zatrudnienie a PKB w Polsce w latach 1993-2001. *Ekonomista*, (3), 329-346.
- [26] Lai, H. (2021). COVID-19 pandemic and unemployment rate : A hybrid unemployment rate prediction approach for developed and developing countries of Asia. *Soft Computing*, 6. <https://doi.org/10.1007/s00500-021-05871-6>
- [27] Mehmet KARAHAN, F. Ç. (2022). FORECASTING OF TURKEY ' S UNEMPLOYMENT RATE FOR FUTURE In world economies , in being able to achieve high national income level , employment has an important effect . Therefore , it is necessary for unemployment to be highly low . Labor force structure of. 6 2, 163–184. <https://doi.org/10.18070/erciyesiibd.1056618>
- [28] Mulaudzi, R. (n.d.). Application of Deep Learning to Forecast the South African Unemployment Rate : A Multivariate Approach.
- [29] Newsletter, M. (2020). National Bureau of Statistics. January.
- [30] Organization, I. L. (2005). International Labour Organization.
- [31] Ozili, P. K. (2021). COVID-19 pandemic and economic crisis: the Nigerian experience and structural causes. *Journal of Economic and Administrative Sciences*, 37(4), 401–418. <https://doi.org/10.1108/jeas-05-2020-0074>
- [32] Rokicki, T. (2021). Forecasting the Unemployment Rate : Application of Selected Prediction Methods. XXIV(3), 985–1000.
- [33] Sagheer, A., & Kotb, M. (2018). ARTICLE IN PRESS Time series forecasting of petroleum production using deep LSTM. xxxx.
- [34] Samuel, A. N., Samuel, B. S., John, D. I., Magdalene, P., & Kajuru, Y. J. (2021). *Modeling Nigeria ' s Unemployment Rates : Box -Jenkins Approach*. 2(1), 28–35.
- [35] Sharma, S., & Soni, H. K. (2021). An Unemployment Prediction Rate for Indian Youth Through Time Series.
- [36] Shi, L., Ali, Y., Id, K., & Tian, M. (2022). COVID-19 pandemic and unemployment rate prediction for developing countries of Asia : A hybrid approach. 1–17. <https://doi.org/10.1371/journal.pone.0275422>
- [37] Smalter Hall, A., & Cook, T. R. (2017). Macroeconomic Indicator Forecasting with Deep Neural Networks. *SSRN Electronic Journal*, September. <https://doi.org/10.2139/ssrn.3046657>
- [38] Zhang, G. P. (2003). Time series forecasting using a hybrid ARIMA and neural network model. *Neurocomputing*, 50, 159–175. www.elsevier.com/locate/neucom