Hybrids ARIMA-ANN models for GDP forecasting in Nepal

Satish Chaudhary[1]
Dipika Uprety[2]

ABSTRACT

Forecasting Nepal's Gross Domestic Product (GDP) holds paramount importance for effective resource planning and allocation. In this research, Artificial Neural Networks (ANNs) have been introduced to predict the GDP time series, wherein the data have been dissected into linear and nonlinear components. The linear aspects have been handled by the ARIMA model, while the ANNs managed the nonlinear elements. Additionally, the study has delved into hybrid models, resulting in additive and multiplicative combinations of ARIMA and ANN. These hybrid models have aimed to enhance forecasting performance, minimize errors, and improve accuracy compared to standalone models. The findings revealed that both ANN and hybrid models surpassed other approaches in terms of prediction accuracy.

JEL Classification: B22, C45, C53

Key Words: ARIMA, ANN, Hybrid models, GDP forecasting

1 Satish Chaudhary, Assistant Director, Foreign Exchange Management Department, Nepal Rastra Bank, Email: satish.chaudhary@nrb.org.np

2 Dipika Uprety, Assistant Director, Nepal Rastra Bank, Biratnagar Office, Email: dipika.uprety@nrb.org.np

I. INTRODUCTION

Macroeconomic indicators like Gross Domestic Product (GDP) offer a comprehensive overview of a country's economic health, playing a crucial role in understanding economic globalization (Urasawa, 2014). Widely recognized as a valuable tool, GDP serves diverse stakeholders such as government officials, businesses, economists, and the public, aiding them in informed decision-making. Government officials and policymakers rely on GDP as a guiding compass for shaping economic policies and regulations (Fathia, 2021). In the business realm, GDP is a vital metric for strategic planning and assessing market conditions (Vatavu et al., 2022). Similarly, local and state governments heavily depend on GDP and similar metrics to formulate policies and evaluate the feasibility of public expenditure (Schaltegger & Torgler, 2005). Numerous studies over time have delved into GDP, offering valuable insights into the intricate mechanisms of the economy. Thus, making GDP an essential indicator for stakeholders, enabling them to navigate the complex landscape of economic globalization (Lumabao & Rosales, 2023; Maccarrone et al., 2021; Srinivasan et al., 2023; Yoon, 2021).

The drive behind undertaking this research lies in the intense impact that Gross Domestic Product (GDP) has on shaping Nepal's economic trajectory. With the nation facing persistent challenges such as poverty, inequality, trade imbalances, and internal conflicts, the urgency to formulate effective policies becomes imperative. Further, Nepal's ambitious goals of achieving sustainable development and transitioning to a middle-income status underscore the critical need for a deep understanding and precise forecasting of GDP. Given Nepal's classification as a least developed country by the United Nations[1], there is an inherent urgency to address these challenges. This research endeavors to play a pivotal role by exploring advanced forecasting models, particularly leveraging the combined strengths of Artificial Neural Networks (ANNs) and Autoregressive Integrated Moving Average (ARIMA) models. By doing so, the aim is to provide not just theoretical insights but practical solutions that enhance the accuracy of GDP forecasting.

One of common approaches for achieving accurate GDP prediction and developing economic strategies, time series models are employed. Time series forecasting, a method analyzing evenly spaced data points over time, is characterized by trends, seasonality, and randomness (Peixeiro, 2022). The value of time series analysis lies in its ability to provide accurate forecasts, enabling researchers/economists to make an informed decision, identify patterns, and gain crucial insights. This analytical approach is widely employed in various domains, including predicting stock markets, weather forecasts and economic indicators (Scott Armstrong, 1988), including economic forecasting.

Among various approaches applied for time series forecasting, Autoregressive Integrated Moving Average (ARIMA) has been in use since 1970's however, it has limitation. One significant drawback is its inability to capture nonlinear patterns in time series data (Zhu & Wei, 2013). Hence, to deal with nonlinearity artificial neurons and artificial neural networks (ANNs) has been used to identify intricate nonlinear relationships between dependent and independent variables (Tu, 1996). Moreover, ANNs are particularly adept at managing intricate relationships among diverse economic indicators. Given that GDP is influenced by a multitude of factors, ANNs excel at capturing the complex interdependencies among these variables. Similarly, economic data is dynamic and susceptible to sudden changes and external shocks. ANNs, with their capability to adapt to shifting patterns and variations in input, present a more flexible modeling framework compared to conventional methods. ANNs possess the ability to automatically extract pertinent features from the data, obviating the necessity for manual feature engineering. Additionally, they can be trained on real-time data, facilitating continuous learning and adaptation to evolving economic conditions.

As a result, artificial neural networks (ANNs) are widely used across numerous fields. Thus, this article is inspired by these developments and seeks to enhance the literature on GDP forecasting in Nepal by leveraging the capabilities of ANNs for better predictions. Additionally, the study will explore the hybrid models of ARIMA and ANNs, aiming to achieve better forecasting performance with low errors and increased accuracy compared to standalone models of ARIMA and ANN.

II. LITERATURE REVIEW

Despite the presence of both the theoretical and empirical criticisms regarding GDP as a comprehensive economic indicator for measure of social welfare and economic progress, its influence remains significant in the realms of economics (Bergh, 2009). While the alternative measures to GDP have proposed such as those suggested by Bloom et al. (2021) such as health and equality ofopportunity. However, per capita GDP has the virtues of being easy to interpret andto calculate with manageable data requirements. Against this backdrop, there is aneed for a measure of well-being that preserves the advantages of per capita GDP,but also includes health and equality. We propose a new parsimonious indicatorto fill this gap, and calculate it for 149 countries. This new indicator could beparticularly useful in complementing standard well-being indicators during theCOVID-19 pandemic. This is because (i, but they do not advocate for completely abandoning GDP as it continues to be a vital metric for evaluating economic strength of a country. Further, Giannetti et al. (2015) proposed variations of GDP including greening of GDP and inclusion of environmentally and socially oriented measures however these measures have not fully achieved progress towards sustainability. Similarly, Fan et al. (2018) summarized arguments against use of GDP, focusing its limited reflection for the value of better health and undervaluation of its true worth. But, no indicator is considered to be perfect. For alternative measures to gain popularity; repeated use, policy leadership and educational support are required for their diffusion. Therefore, despite the criticisms of GDP as a misaligned indicator, its prevalent usage as a prominent economic measure persists.

The need for precise economic predictions, particularly for GDP growth rate, cannot be ignored as it serves an important guide for policymakers, businesses, investors, and individuals, enabling them to manage risks and capitalize the opportunities effectively (Kitchen & Monaco, 2003). Consequently, GDP has emerged as a vital macroeconomic indicator for economic prediction across various economies. Studies by Golinelli & Parigi (2007) compared the forecasting performance of models for GDP growth in G7 countries, while Jiang et al. (2017) emphasized the importance of GDP growth targets for Chinese governments. Tacchella et al. (2018) argued that GDP growth forecasting models need not be to excessively complex, as even a simple forecasting scheme can outperform the accuracy of the International Monetary fund's (IMF) five year forecast by over 25% based on their analysis of 20 developing countries. Thus, GDP forecasting is of paramount significance for effective economic planning globally.

The Autoregressive Integrated Moving Average (ARIMA) model is a commonly utilized statistical method for GDP forecasting, which has capability to handle both stationary and non-stationary data (Peixeiro, 2022). In addition to ARIMA, regression analysis is also widely used for GDP prediction as evidenced by numerous studies (Samiyu, 2021; Stundziene, 2014). This analysis allows the discovery and measurement of relationship between GDP and other important variables considered in study. However, it is crucial to acknowledge the limitations of ARIMA model as it is susceptible to fat tails and volatility clustering (Petrica et al., 2016) and its linearity assumptions may not be applicable in certain complex situations (Peixeiro, 2022).

In addition, ARIMA models have been applied in various nations for GDP forecasting (Yang et al. (2016) focused on China and fitted the ARIMA (2, 4, 2) to model and predict the GDP while Rana (2019) employed ARIMA (0, 1, 2) for the prediction of GDP in Nepal. Eissa (2020) adopted ARIMA (1, 1, 2) for Egypt and ARIMA (1, 1, 1) for Saudi Arabia respectively to forecast GDP per capita. In Bangladesh, Salah Uddin & Tanzim (2021) used ARIMA (1, 2, 1) and Jamile et al. (2021) utilized ARIMA (0, 2, 1) for GDP forecasting in Qatar, Saudi Arabia and the UAE. Mohamed (2022) used ARIMA (5, 1, 2) to examine Somalia's GDP growth rate while Barbara et al. (2022) combined ARIMA (2, 2, 2) and nonlinear GARCH (1, 1) models for GDP forecasting in Ghana. Lastly, Muma & Karoki (2022) conducted a comprehensive review of applicability of ARIMA models in GDP studies. Thus, by utilizing the various ARIMA parameters combinations, researchers have been able to generate a reliable forecast for different economies.

Artificial Neural Networks (ANNs) were introduced to emulate the human brain functions and can change their internal structure based on the defined function objectives. This adaptability has made ANNs popular both from theoretical and practical applications (Z.R. Yang & Yang, 2014). ANNs application can be viewed in diverse real-time scenarios, such as regression analysis for function approximation, time series prediction, hand gesture control, pattern recognition, sequential decision making, media player control through hand motions, data processing (including filtering and clustering), scrolling webpages or eBooks by hand gestures, system identification and control, game playing and decision making, medical diagnosis, financial applications and data mining, among others (Maind & Wankar, 2014). Furthermore, ANNs have widespread use in economics field as well. Turan (2011) applied ANNs to study the dynamics of labor market in Turkey, while Ruiz-Real et al. (2020) examined the use of ANNs for business and economic research. Malte (2018) employed ANNs and found that ANN model could predict more precise GDP growth rate predictions than a comparable linear model. Consequently, ANN's models are highly useable in capturing the time series.

With comparative analysis of ARIMA and ANN, where each of models have been performed. However, in most cases, the results show that the ANN model performs better with high prediction accuracy than ARIMA models; but there are cases where ARIMA has outperformed ANN model (Safi & Sanusi, 2021; Sarvestani et al., 2022) Artificial Neural Network (ANN. Hence, to harness the ability of both models, hybrid models were introduced to capture both linear and nonlinear component for better prediction. Sahed et al. (2020) compared the adaptive network based fuzzy inference system (ANFIS) with ARIMA for quarterly GDP prediction in Algeria and found that ANFIS approach provided better accuracy than ARIMA. Similarly, Lu (2021) compared the forecast results of neural network and ARIMA model for price prediction revealing lower errors and higher accuracy. Further, Alsuwaylimi (2023), Li & Yang (2023), Pannakkong et al. (2019), Yusof & Samsudin (2018) compared ANN and ARIMA for time series forecasting and highlighted that hybrid models perform better compared to individuals models respectively. Finally, Zhang (2003)'s additive form and Wang et al. (2013)'s multiplicative form of hybrid model between ARIMA and ANN both resulted in better accuracy prediction for time series forecasting. Thus, to determine the appropriate GDP forecasting models, comprehensive experimental comparisons are conducted and the performance of the ARIMA, ANN and hybrid models are evaluated and compared.

III. TIME SERIES FORECASTING MODELS

Autoregressive Integrated Moving Average Model (ARIMA)

ARIMA stands as one the most popular used time series models in Box-Jenkins forecasting, due to its capability to depict various types of time series data. As a linear model, it bases its prediction based on the linear function of past observations, implying that the data into ARIMA needs to be both linear and stationary. The fundamental underlying principle is to estimate appropriate order of ARIMA process, which encompasses autoregressive (AR) model represented by ‘p’ and moving average (MA) model represented by ‘q’ models and differencing term, a data transformation element represented by ‘d’. However, this model linearity assumption presents limitations, which is hard to satisfy in real world scenarios (Das, 2019). Here, the formulation of ARIMA model is a special case of ARMA model as represented in the equation

$$Y_t = \mathrm{c} + \phi_1 Y_{(t-1)} + \phi_2 Y_{(t-2)} + \cdots + \phi_p Y_{(t-p)} + \epsilon_t - \theta_1 \epsilon_{(t-1)} - \theta_2 \epsilon_{(t-2)} - \cdots - \theta_q \epsilon_{t-q} \ldots (1)$$

$$Y_t = c + \Sigma^{p}_{i=1} \phi_i Y_{t-i} + \epsilon_t - \Sigma^{q}_{j=1} \theta_t \epsilon_{t-j} \ldots (2)$$

Here, ARMA model predicts time series by using lagged values of time series and errors ($Y_{t-1}, Y_{t-2}, \ldots, Y_{t-p}$) ($\epsilon_{t-1}, \epsilon_{t-2}, \ldots, \epsilon_{t-q}$). Further, $\epsilon_t \sim N(0, \sigma^2)$ and θ, ϕ are the coefficients of AR (p) and MA (q) respectively.

Using backward shift operator (B), where B is defined as $B^i Y_t = Y_{t-1}$. Reducing the equation (2) using backward shift operator ARMA model is formulated as follows

$$Y_t = \mathrm{c} + \Sigma^{p}_{i=1}\, \phi_i\, Y_t\, B^i + \epsilon_t - \Sigma^{q}_{j=1}\, \theta_j\, \epsilon_t\, B^j \qquad (3)$$

Rearranging equation (3), we obtain ARMA model as follows

$$(1 - \Sigma^{p}_{i=1}\, \phi_i\, B^i)\; Y_t = \mathrm{c} + (1 - \Sigma^{q}_{j=1}\, \theta_j\, B^j)\; \epsilon_t \qquad (4)$$

The compacted form

$$\phi_p\, (B)\; Y_t - c + \theta_q\, (B)\; \epsilon_t \qquad (5)$$

Where,

$$\phi_p\, (B) = 1 - \Sigma^{p}_{i=1}\, \phi_i\, B^i \qquad (6)$$

$$\theta_q\, (B) = 1 - \Sigma^{q}_{j=1}\, \theta_j\, B^j \qquad (7)$$

However, one limitation of ARMA model is the stationary condition, in such case the differencing is required to transform the non-stationary data into stationary time series by substituting $(1 - B)^{d}\, Y_t$ for Y_t, where d is degree of differencing. Thus, we obtain ARIMA model as follows

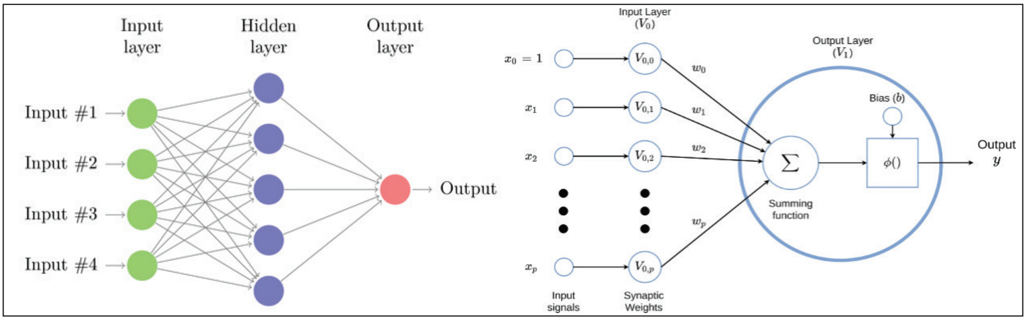
$$\phi_p\, (B)\;\; (1 - B)^{\mathrm{d}}\; Y_t = c + \theta_q\, (B)\; \epsilon_t \qquad (8)$$

Further, Peixeiro (2022) has described a detailed procedure for obtaining best fit parameters for ARIMA models and the final model is selected based on Akaike Information criteria (AIC), Bayesian Information criteria (BIC), Hannan Quinn information criteria (HIC), log-likelihood, and sigma square respectively.

Artificial Neural Networks (ANNs)

Artificial Neural Networks (ANNs) are an adaptive artificial system that has capability to mimic the human brain neurons. It has functionality to modify its internal structure with relation to target objective and because of this characteristic; ANNs are suitable for non-linear modeling technique. ANN fundamental components are node also known as processing elements and connections. Here, each node has its own input which it receives from other node and an output through which it communicates with other nodes. Further, each node has an activation function *f (x)* that transforms the input into output exhibiting excitatory connection with positive values and inhibitory connections with negative values (Grossi & Buscema, 2007).

Figure 1: Feed forward neural network (left) and node in neural network (right)



Source: (MontesinosLópez et al., 2022)

In general, Artificial Neural Networks (ANNs typically consist of a three layer; input layer, hidden layer and output layer. The structure typically involves one input layer followed by hidden layer and output layer. However, it is possible to have more than one hidden layer in ANNs architecture. The strength of ANNs lies in its ability to approximate any continuous function by adjusting the number of layers and nodes within each layer. The selection of layers and nodes plays a vital role in determining the forecasting performance. Having large number of layers and node can cause over fitting of data while too simple layers and nodes can cause under fitting of data and models might not learn from the given input. Unfortunately, there is no any predefined set of rules to decide these parameters, Thus, numerous experimentations are required for optimal configuration (Büyükşahin & Ertekin, 2019).

The structure of ANN can be identified as follows

Input layer

The input layer of neural network consists of input nodes that represent input variables. In context of time series data, input can be categorized into two variables; technical and fundamental variables. Technical variables are lagged values (time series at t–1) or processed values. On the other hand, fundamental variables encompass other variables that are believed to have a relationship with the time series, such as inflation, interest rate, and other relevant factors (Montesinos López et al., 2022; Pannakkong et al., 2019).

Hidden layer

Hidden layer consists of neurons of network that isolated within the network i.e., it is the middle layer between the input and output layers which generates the output by aggregation of outputs from its subsequent layer. The actual learning process occurs in this layer as the information learned from the training data is stored in weights and biases. Further, it is crucial to highlight that hidden layer plays an important role in capturing intricate nonlinear pattern in data with greater efficiency (Montesinos López et al., 2022).

General form of hidden layer

$$h_j(x) = f\left(v_{j0} + \Sigma_{i=1}^{D} x_{ij} v_{ji}\right) \qquad (9)$$

Where j in index of hidden layer and D is number of inputs

Output layer

The layer responsible for processing the information from input to hidden layer and generating the result is output layer. This layer produces the answer or prediction based on the inputs provided. The output layer can take different layers such as binary, ordinal or in a continuous form.

General form of output layer

$$o_k(x) = g\left(w_{k0} + \Sigma_{j=1}^{j} h_j(x)\, w_{ki}\right) \qquad (10)$$

Where both in equation (9) and (10) represent; j is index of hidden units, k is the output units and D is the input units.

Hybrid models

In practice, time series data rarely exhibit either linear or nonlinear characteristics. Thus, to achieve more accurate predictions hybrid models of ARIMA-ANN methods are employed. This study considers two hybrid models as a potential approach for accurate prediction along with ARIMA and ANN models for prediction.

Zhang's hybrid ARIMA-ANN model

(Zhang, 2003) proposed a hybrid ARIMA-ANN model for time series forecasting, assumes any time series data is sum of two components, linear and nonlinear.

$$Y_t = L_t + N_t \quad (11)$$

Initially, an ARIMA is fit to the existing time series data and predictions $\widehat{L}_t$ is obtained using equation (1). The linear forecast from ARIMA $\widehat{L}_t$ are subtracted from original time series data the resulting value are residuals (error sequence) that consists of nonlinear variations of time series as ARIMA model only consider linear variations.

$$N_t = Y_t - \widehat{L}_t \quad (12)$$

The error sequences is fit using ANN and results are obtained as

$$\widehat{N}_t = f\,(n_1,\, n_{t-1},\, n_{t-2},\, \ldots,\, n_{t-n}) + v_t \quad (13)$$

Finally, the predictions obtained from both the linear model predicted using ARIMA and nonlinear model predicted using ANN are combined to obtained the final predictions.

$$\widehat{Y}_t = \widehat{L}_t + \widehat{N}_t \quad (14)$$

Wang hybrid ARIMA-ANN model

(Wang et al., 2013) proposed a hybrid model for time series forecasting which combines both linear and nonlinear component using multiplicative approach. The underlying assumption of this model considers both linear and nonlinear component of time series are represented by product rule.

$$Y_t = L_t \cdot N_t \quad (15)$$

Similar to (Zhang, 2003) assumptions, linear component of time series is measured using ARIMA model in equation (1). However, the nonlinear component in this hybrid model is represented as quotient of by the forecast of linear model

$$n_t = \frac{Y_t}{\widehat{L}_t} \quad \text{.......} \quad (16)$$

The nonlinear component N_t is fit using ANN

$$\widehat{N}_t = f\,(n_t,\ n_{t-1},\ n_{t-2},\ \ldots,\ n_{t-n}) + v_t \quad \text{.......} \quad (17)$$

The final predictions of model is obtained as product of linear and nonlinear component

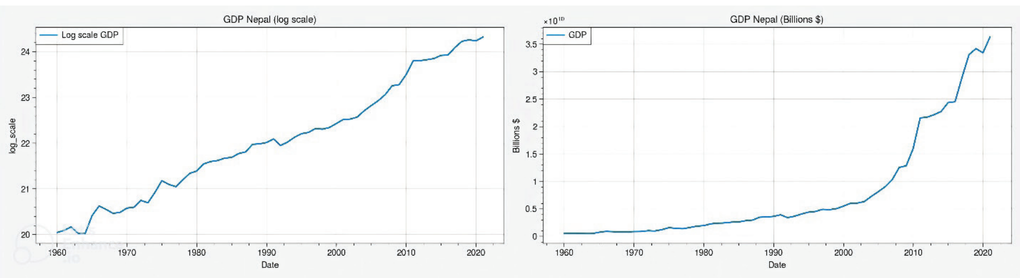
$$\widehat{Y}_t = \widehat{L}_t \cdot \widehat{N}_t \quad \text{.......} \quad (18)$$

IV. METHODOLOGY

Dataset and analysis software

The study uses univariate Box Jenkins method for modeling the linear component of time series data whereas ANN is used for nonlinear component. The dataset for GDP is sourced from World Bank official website titled as GDP (Current US$) Nepal; ranging from 1960 to 2021. The descriptive statistics of the model are reported in Table 1. From Figure 2, it is detected that the GDP of Nepal is in increasing trend and to facilitate the analysis, the data is transformed into logarithmic form. Further, the data analysis is conducted using open source Python programming language (ver-3.10.11).

Figure 2: GDP of Nepal in log scale (left) and in billions (right)



Source: Author's calculation

Table 1: Descriptive statistics of time series data ranging from 1960 to 2021

Statistics	Mean	Standard deviation	Minimum	Kurtosis	Skewness	Maximum
Original data (billions)	7.9368	9.9857	0.4960	1.2094	1.5746	36.288
Log form	22.04	1.27	20.02	-0.9800	0.1945	24.31

Source: Author's calculation

Accuracy measure

Accurate forecasting is significant when considering selection of models from various forecasting options. This study incorporates three metrics determining the forecasting errors: mean square error (MSE); calculates average of squared errors, root mean square error (RMSE); measures the average difference between statistical predicted values and actual values, and mean absolute percentage error (MAPE), measures the average of absolute percentage errors.

$$MSE= \frac{1}{N} \Sigma_{t=1}^{N} (Y_t - \acute{Y}_t)^{2} \quad (19)$$

$$RMSE= \sqrt{\frac{1}{N} \Sigma_{t=1}^{N} (Y_t - \acute{Y}_t)^2} \quad (20)$$

$$MAPE= \frac{1}{N} \Sigma_{t=1}^{N} \frac{|Y_t - \acute{Y}_t|}{|Y_t|} \quad (21)$$

ARIMA model identification, estimation and diagnostic checking

The optimal parameters for AR (p), I (d), MA (q) of ARIMA are determined using autocorrelation (ACF) functions and partial autocorrelation (PACF) functions and the stationary of data is tested using Augmented Dickey Fuller test (ADF) for identifying optimal differencing. The model parameters estimation is based on value maximum likelihood estimation (MLE) and the criteria mentioned in Table 2. Finally, the residuals of ARIMA are tested using Ljung Box Q test and roots are checked.

Table 2: Selection criteria for p, d, q of ARIMA model

Estimation parameters	Criteria
Log likelihood	Maximization
Akaike Information Criteria	Minimization
Bayesian Information Criteria	Minimization
Hanna Quinn Information Criteria	Minimization
Sigma Square	Minimization

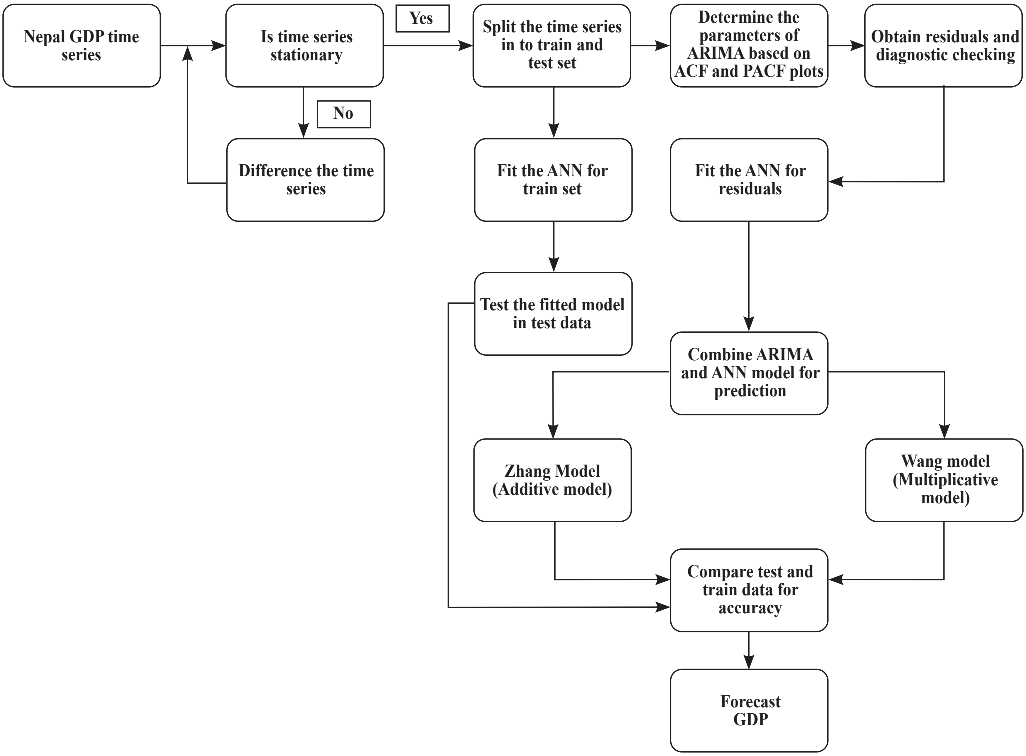
ANN model building, training and testing of models

Initially, GDP data undergo preprocessing for the Artificial Neural Network (ANN) whereby the window size of five is taken into consideration; this involves inputting the data of the previous five time periods ("t-5") into the model to predict the value for the current time period "t". Subsequently, the dataset is divided into training and testing sets to facilitate the training and testing of the model. After training the model, an optimal ANN architecture is employed to test the data, whereby 10 trials are conducted and the mean value of these trails are considered as the final output.

Table 3: Hyper parameters for ANN and hybrid models

Hype-parameters	ANN	ANN + ARIMA (Zhang, 2003)	ANN * ARIMA (Wang et al., 2013)
Architecture	35-35-35	5-10-5	25-20-25
Epochs	500	500	500
Learning rate	0.001	0.0001	0.0001
Regularization	L2 regularization	L2 regularization	L2 regularization
Activation function	Exponential linear unit	Exponential linear unit	Exponential linear unit
Optimizer	Adam	Adam	Adam
Loss	Mean squared error	Mean squared error	Mean squared error
Accuracy	Root mean squared error	Root mean squared error	Root mean squared error

Further, the detailed methodology is presented in figure



Forecasting: in sample forecasting and out of sample forecasting

In the study, the forecasting of the Gross Domestic Product (GDP) is undertaken among the models described in this paper to assess their predictive accuracy. The in-sample forecasting involves the application of models to predict GDP values for a specified period within the dataset, enabling the comprehensive evaluation of model performance over a 5 year horizon. This in-sample analysis serves as a benchmark allowing the comparison of model predictions against the actual observed data. Following the in-sample forecasting, the most accurate model will be identified, and subsequently employed for out-of-sample forecasting. The out-of-sample projection extends beyond the observed data, spanning a 10 year horizon.

V. EXPERIMENTAL RESULTS

ARIMA Model

ARIMA model was initially constructed using annual data from Nepal's GDP, with the time series from 1960 to 2016 serving as training data and the series from 2016 to 2021 as testing data. The raw log data represented a consistent upward trend, as depicted in Figure 2 (left). Now, the unit root test has been performed to assess the time series' stationarity. The results, at a significance level of 0.05, indicated that the original series became stationary after the first order difference and thus was taken into account. Hence, in this scenario, the first difference is sufficient for identifying the appropriate model for forecasting.

Table 4: Result of unit root

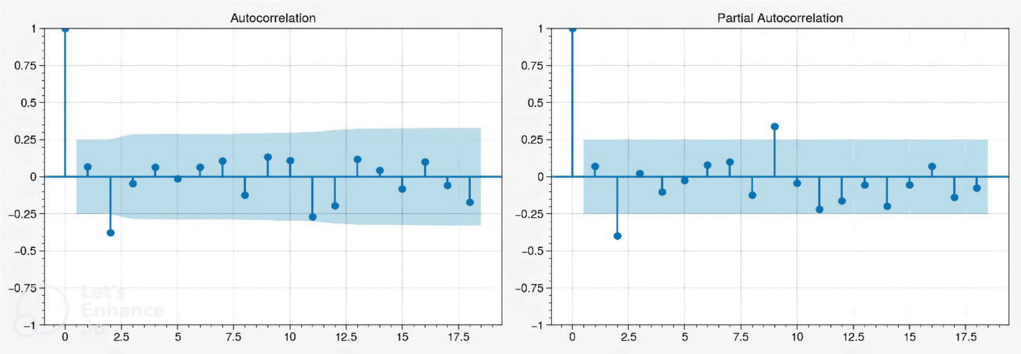
Variables	ADF test			
	Level		First level difference	
GDP time series (log form)	t – statistics	p-value	t – statistics	p-value
	0.3127	0.9779	-7.6864	1.457e-11

Source: Author's calculation

Model Identification

The next step, model identification, involves using autocorrelation (ACF) to determine the AR component 'p' and partial autocorrelation (PACF) to determine the MA component 'q'. The autocorrelation (ACF) and partial autocorrelation (PACF) function of the differenced GDP time series at first order is presented in Figure 3. Further, the ACF bar is found to be non-significant at lag 2, leading to the conclusion that the time series data originates from MA (q = 2). Similarly, the PACF graph shows the bar at lag 2 to be non-significant, indicating that the data originates from AR (p = 2). Hence, this gives us the initial ARIMA (2, 1, 2) model from the differenced series. In addition, to take a parsimonious model into account, several other models such as ARIMA (0, 1, 2) and ARIMA (2, 1, 0) are also developed.

Figure 3: Autocorrelation and Partial autocorrelation of GDP time series data



Source: Author's calculation

Following the ARIMA theoretical framework, several proposed ARIMA models are analyzed and based on the analysis, we selected the most suitable model by comparing the results based on the criteria mentioned in Table 2.

Table 5: Model selection

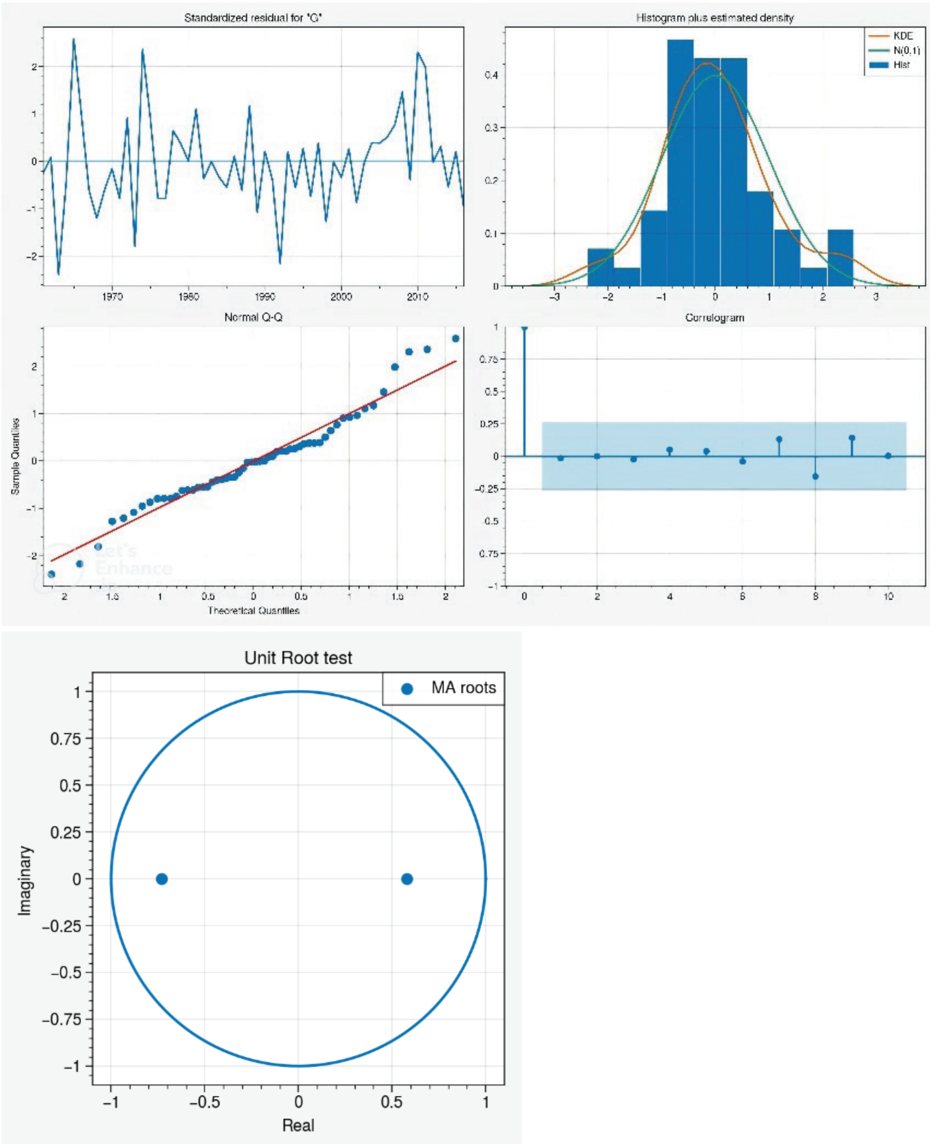
	Model selection process					
Models	**Log Likelihood**	**AIC**	**BIC**	**HQIC**	**Sigma square**	**Significance**
ARIMA (2, 1, 2)	54.205	-96.410	-84.258	-91.698	0.0083	2/5
ARIMA (0, 1, 2)	54.116	-100.23	-92.132	-97.092	0.0084	3/5*
ARIMA (2, 1, 0)	53.619	-99.239	-91.138	-96.098	0.0085	0/5

Source: Author's calculation

The data in Table 5 indicates that the most appropriate model for forecasting Nepal's GDP is ARIMA (0, 1, 2). This model is chosen because it has the most relevant parameters with, the highest log likelihood, and the lowest Akaike, Bayesian, and Hannan Quinn information criteria. However, before using ARIMA (0, 1, 2) for forecasting, it must undergo diagnostic tests.

The diagnostics test results no heteroskedasticity for Ljung test, which tests for the absence of autocorrelations in residuals, where the test statistics is Q = 18.971 and p–value is 0.89902 which is larger than 0.05. Thus, we fail to reject the null hypothesis and conclude the residuals are independent at lag 28. Similarly, the correlogram of residuals is flat and Figure 3 (right) side figure represents inverse roots of MA polynomials for the stability of the models. The inverse root characteristics of AR/MA polynomials shows whether the estimated models is stationary or not. If all the roots falls within the unit circle the model is considered as stationary model. In the current case, MA inverse roots within the circle.

Figure 4: Diagonstic plot (left) and Inverse roots (right)



Source: Author's calculation

Table 6 indicates the estimated results of the chosen ARIMA (0, 1, 2) model with constant and MA components coefficients along with their associated p-values. Further, AIC, BIC, HQIC and log likelihood information are presented.

Table 6: ARIMA results

Variables	**Coefficients**	**Standard error**	**Z**	**P > |z|**	**0.025**	**0.975**
x1	0.0708	0.009	8.258	0.000	0.054	0.088
ma.L1	0.1136	0.103	1.098	0.272	-0.089	0.316
ma.L2	-0.4008	0.113	-3.548	0.000	-0.622	-0.179
sigma2	0.0081	0.001	6.174	0.000	0.006	0.011
Dependent variable		GDP				
Model		ARIMA (0,1,2)				
No. of observations		62				
Log likelihood		60.193				
AIC		-112.386				
BIC		-103.942				
HQIC		-109.076				

Source: Author's calculation

The final equation of ARIMA is represented as follows:

$$\delta GDP_t = 0.0708 + 0.1136\epsilon_{t-1} - 0.4008\epsilon_{t-2} \text{ where, } \epsilon_t \sim WN(0,0.0081) \quad ..(22)$$

Artificial Neural Networks (ANNs) and Hybrid models

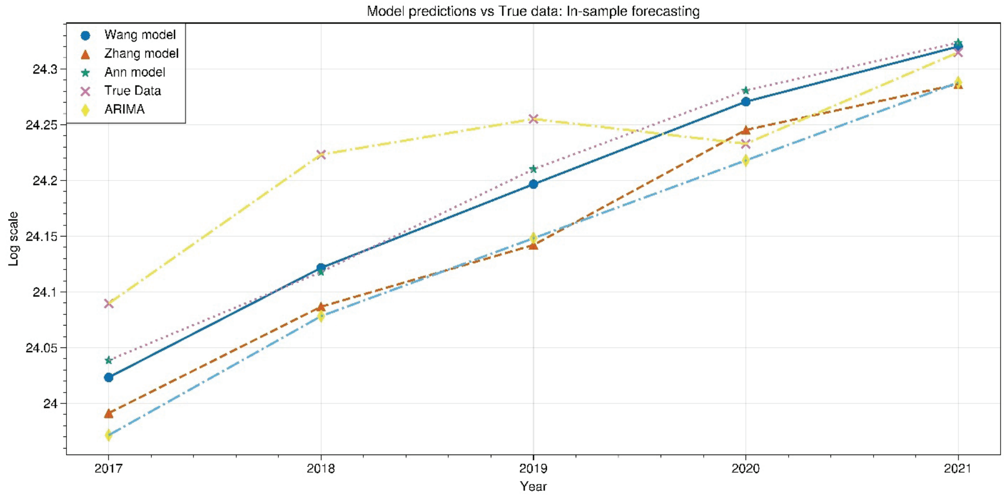
The process of experimenting with Artificial Neural Networks (ANNs) is carried out using a training set (80%) and a testing set (20%) of Nepal's Gross Domestic Product (GDP) time series data. After conducting numerous trials with three hidden layers, each containing a varying number of neurons, the architecture that demonstrated the most superior performance, as measured by the Mean Squared Error (MSE), Root Mean Squared Error (RMSE), and Mean Absolute Percentage Error (MAPE), consisted of 35 neurons in each layer. Consequently, this optimal architecture was used to fit and test the model over 10 trials, with the aim of comparing its performance against the benchmark set by the ARIMA model's accuracy measures as shown in Figure 6. The results indicated that the accuracy of the MSE, RMSE, and MAPE was superior to that of the ARIMA model, which is considered as a benchmark, when applied to the testing data.

For additive hybrid model analysis, once the most suitable ARIMA model for linear forecasting of Nepal's GDP time series data is identified, the linear modeling component is constructed using ARIMA orders (0, 1, 2). The nonlinear modeling component, on the other hand, is developed through trials to identify the most effective network structure. The optimal network structure, consisting of 5,10,5 neurons respectively in each of the three layers, is then applied to the residuals of the ARIMA results to capture the inherent nonlinearity in the time series data. The performance of additive hybrid model is shown in Appendix 1.

In contrast to the additive hybrid model, the multiplicative hybrid model assumes that the linear and nonlinear components of the time series are multiplicative in nature. In this approach, after identifying a suitable ARIMA model for linear modeling, the nonlinear residuals are trained through an optimal neural network. The best-fitted network 25,20,23 in each layers of network for the residuals of ARIMA is then multiplied with the results of the linear model (ARIMA) to make predictions. The performance of the multiplicative hybrid model on the training and testing data can be observed in Appendix 1.

Following a thorough analysis, the accuracy measures of the forecasting models are calculated and compared to determine the most suitable prediction model for Nepal's GDP. The MSE, RMSE, and MAPE of the models are calculated and evaluated as shown in figures in Appendix 1. The multiplicative model emerged as the best model for forecasting GDP in Nepal. Futher, the models performance on testing data are shown in Appendix 2.

Figure 5: In sample prediction for next 5 year



Source: Author's calculation

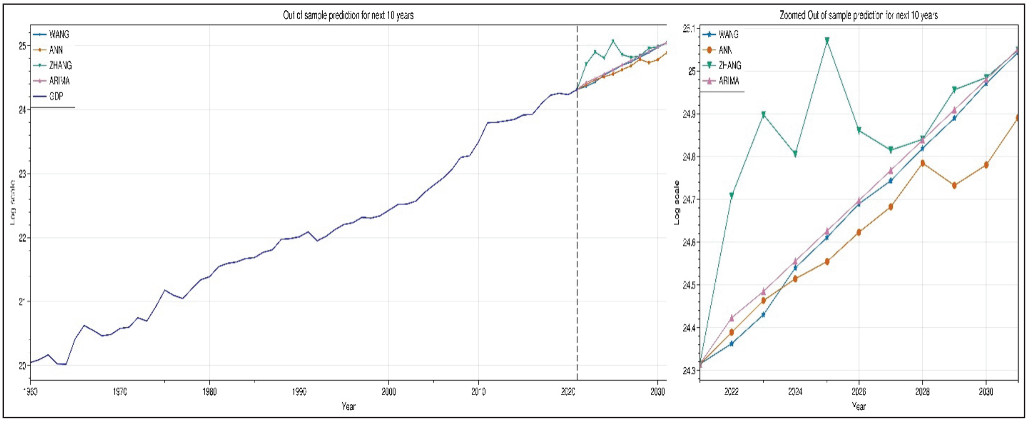
Table 7: Model performance based on training and testing data

Model/ Accuracy	MSE_train	RMSE_train	MAPE_train	MSE_test	RMSE_test	MAPE_test
ANN	0.00028	0.01655	0.01386	0.00456	0.06636	0.05575
Zhang	0.00889	0.09427	0.06525	0.00788	0.08875	0.07557
Wang	0.00002	0.00412	0.00362	0.00362	0.05677	0.04790
ARIMA	0.00846	0.09196	0.00340	0.00947	0.09730	0.00340

Source: Author's calculation

Building upon the evaluation of various forecasting models, the one that showcased superior performance during the in-sample analysis has been selected for the future predictions. As a result, models are used for out-of-sample prediction for next 10 years as show in Figure 6. These forecasts provide valuable insights into the potential trajectory of the Gross Domestic Product (GDP).

Figure 6: Out-of-sample prediction for next 10 year



Source: Author's calculation

VI. DISCUSSION

Several time series forecasting models, including ARIMA, ANN, additive hybrid, and multiplicative hybrid models, have been employed for predicting Nepal's GDP. The ARIMA parameters p, d, and q were found to be (0, 1, 2), aligning with (Rana, 2019) study. The ANN and hybrid models demonstrated superior predictive capabilities compared to the ARIMA models. The study concluded that multiplicative hybrid models outperformed all other models in predictive performance. The GDP data exhibited both linear and nonlinear characteristics, indicating that standalone models are insufficient for superior forecasting perfromance. Therefore, hybrid models were considered most suitable for accurate predictions (Alsuwaylimi, 2023; Li & Yang, 2023; Pannakkong et al., 2019; Wang et al., 2013; Zhang, 2003).

(National Planning Commission, 2020) projected an average economic growth of 10.3% by FY 2023/24, and the Government of Nepal has set optimistic goals for the economic growth rate accordingly. The Nepal SDG's economic growth target for 2025 is 5.4%, and the World Bank Development Update (April 2023) forecasts an average economic growth of 5.8% for FY 2024/25. The forecasting models predict that these targets are likely to be achieved within the specified timeframe, provided there are no external influences similar to the impact of

COVID-19, which can introduce uncertainty into forecasting, as demonstrated by (Safi & Sanusi, 2021) study.

Additionally, study's limitations include the scarcity of sectoral GDP data. While ANNs and the integration of ANNs with ARIMA are used for forecasting, newer techniques for more robust and high-performing models are under investigation. Future research is recommended to test hybrid models in the sectoral domain and explore various other hybrid models like support vector regression, long short term memory, recurrent neural network, and convolution neural network where such future studies could also compare these models and assess their robustness.

VII. CONCLUSION

GDP forecasting serves a multifaceted role, offering valuable insights across diverse economic dimensions, including planning, budgeting, monetary policy, business investment, foreign direct investment, and consumer spending. This study employed the ARIMA model to predict Nepal's GDP, optimizing parameters in alignment with (Rana, 2019). Subsequently, the introduction of artificial neural networks (ANNs) offered a novel approach, breaking down time series into linear and non-linear components, with the ARIMA model addressing the linear aspects and ANNs managing the non linear elements. The exploration further extends to hybrid model, with a focus in additive and multiplicative combinations. Notably, the findings revealed that the multiplicative hybrid models, demonstrated superior predictive accuracy compared to other models .

From economic prespective, the finding holds significant implications. The heightened accuracy of GDP forecasts, especially utilizing the hybrid approach of ARIMA-ANN, empowers the decision maker across the government, monetary policy and business sectors. Policymakers in Nepal can leverage these insights to refine the fiscal policies alinging them to the anticipated GDP trends. Similarly, the monetary authorities can adjust strategies to manage inflation effectively, while businesses gain valuable information for strategic planning, capacity expansion and market entry. Additionally, the outperformance of ANN based models, particulary multiplcative approach suggests a potential paradigm shift in forecasting methodologies in Nepal's GDP. And the study has also made a ten-year forecast, potraying an upward trajectory in Nepal's GDP calls for proactive long term planning. This sustained growth trend invites considerations for targeted investment in infrastructure, social programs, and policies that foster a conducive environment for economic development.

REFERENCES

Alsuwaylimi, A. (2023). Comparison of ARIMA, ANN and Hybrid ARIMA-ANN models for time series forecasting. *Information Sciences Letters*, *12*(2), 1003–1016. https://doi.org/10.18576/isl/120238

Armstrong, J. S. (1988). Research needs in forecasting. *International Journal of Forecasting, 4*(3), 449-465.http://dx.doi.org/10.1016/0169-2070(88)90111-2

Barbara, D., Li, C., Jing, Y., & Samuel, A. (2022). Modeling and forecast of Ghana's GDP using ARIMA-GARCH model. *OALib*, *09*(01), 1–16. https://doi.org/10.4236/oalib.1108335

Bergh, J. C. J. M. V. D. (2009). The GDP paradox. *Journal of Economic Psychology*, *30*(2), 117–135. https://doi.org/10.1016/j.joep.2008.12.001

Bloom, D. E., Fan, V. Y., Kufenko, V., Ogbuoji, O., Prettner, K., & Yamey, G. (2021). Going beyond GDP with a parsimonious indicator: Inequality-adjusted healthy lifetime income. *Vienna Yearbook of Population Research*, *19*. https://doi.org/10.1553/populationyearbook2021.res1.1

Büyükşahin, Ü. Ç., & Ertekin, Ş. (2019). Improving forecasting accuracy of time series data using a new ARIMA-ANN hybrid method and empirical mode decomposition. *Neurocomputing*, *361*, 151–163. https://doi.org/10.1016/j.neucom.2019.05.099

Das, P. (2019). *Econometrics in Theory and Practice: Analysis of Cross Section, Time Series and Panel Data with Stata 15.1*. Springer Singapore. https://doi.org/10.1007/978-981-32-9019-8

Eissa, N. (2020). Forecasting the GDP per capita for Egypt and Saudi Arabia using ARIMA models. *Research in World Economy*, *11*(1), 247. https://doi.org/10.5430/rwe.v11n1p247

Fan, V. Y., Bloom, D. E., Ogbuoji, O., Prettner, K., & Yamey, G. (2018). Valuing health as development: Going beyond gross domestic product. *BMJ*, k4371. https://doi.org/10.1136/bmj.k4371

Fathia, S. N. (2021). How good government governance affect the Economic Growth? An investigation on selected country around the World. *Asian Journal of Economics, Business and Accounting*, 93–98. https://doi.org/10.9734/ajeba/2021/v21i730405

Giannetti, B. F., Agostinho, F., Almeida, C. M. V. B., & Huisingh, D. (2015). A review of limitations of GDP and alternative indices to monitor human well-being and to manage eco-system functionality. *Journal of Cleaner Production*, *87*, 11–25. https://doi.org/10.1016/j.jclepro.2014.10.051

Golinelli, R., & Parigi, G. (2007). The use of monthly indicators to forecast quarterly GDP in the short run: An application to the G7 countries. *Journal of Forecasting*, *26*(2), 77–94. https://doi.org/10.1002/for.1007

Grossi, E., & Buscema, M. (2007). Introduction to artificial neural networks: *European Journal of Gastroenterology & Hepatology*, *19*(12), 1046–1054. https://doi.org/10.1097/MEG.0b013e3282f198a0

Jamile, Y., Nermeen, I., & Nour, F. (2021). GDP forecast of the Biggest GCC economies using ARIMA. *Munich Personal RePEc Archive*.

Jiang, Y., Guo, Y., & Zhang, Y. (2017). Forecasting China's GDP growth using dynamic factors and mixed-frequency data. *Economic Modelling*, *66*, 132–138. https://doi.org/10.1016/j.econmod.2017.06.005

Kitchen, J., & Monaco, R. (2003). Real-time forecasting in practice: The U.S. Treasury staff's Real-time GDP forecast system. *Business Economics*, *4*(38), 10–19.

Li, G., & Yang, N. (2023). A hybrid SARIMA-LSTM model for air temperature forecasting. *Advanced Theory and Simulations*, *6*(2), 2200502. https://doi.org/10.1002/adts.202200502

Lu, S. (2021). Research on GDP forecast analysis combining BP neural network and ARIMA model. *Computational Intelligence and Neuroscience*, *2021*, 1–10. https://doi.org/10.1155/2021/1026978

Lumabao, M. K., & Rosales, J. F. (2023). Determinants of GDP growth in the Philippines: 1970-2020. *Journal of Economics, Finance and Accounting Studies*, *5*(1), 73–97. https://doi.org/10.32996/jefas.2023.5.1.6

Maccarrone, G., Morelli, G., & Spadaccini, S. (2021). GDP forecasting: Machine learning, linear or autoregression? *Frontiers in Artificial Intelligence*, *4*, 757864. https://doi.org/10.3389/frai.2021.757864

Maind, S., & Wankar, P. (2014). Research paper on basic of artifical neural network. *International Journal on Recent Innovation Trends in Computing and Communication*, *2*(1), 96–100.

Malte, J. (2018). Artificial neural network regression models: Predicting GDP growth. *HWWI Research Paper, No. 185*.

Mohamed, A. O. (2022). Modeling and forecasting Somali economic growth using ARIMA models. *Forecasting*, *4*(4), 1038–1050. https://doi.org/10.3390/forecast4040056

Montesinos López, O. A., Montesinos López, A., & Crossa, J. (2022). *Multivariate Statistical Machine Learning Methods for Genomic Prediction*. Springer International Publishing. https://doi.org/10.1007/978-3-030-89010-0

Muma, B., & Karoki, A. (2022). Modeling GDP using autoregressive integrated moving average (ARIMA) model: A systematic review. *OALib*, *09*(04), 1–8. https://doi.org/10.4236/oalib.1108355

Pannakkong, W., Huynh, V.-N., & Sriboonchitta, S. (2019). A novel hybrid autoregressive integrated moving average and artificial neural network model for Cassava export forecasting: *International Journal of Computational Intelligence Systems*, *12*(2), 1047. https://doi.org/10.2991/ijcis.d.190909.001

Peixeiro, M. (2022). *Time series forecasting in Python*. Manning Publications Co.

Petrica, A.-C., Stancu, S., & Tindeche, A. (2016). Limitation of ARIMA models in financial and monetary economics. *Theoretical and Applied Economics*, *23*(4).

Rana, S. B. (2019). Forecasting GDP movements in Nepal using Autoregressive Integrated Moving Average (ARIMA) modelling Process. *Journal of Business and Social Sciences Research*, *4*(2), 1–20. https://doi.org/10.3126/jbssr.v4i2.29480

Ruiz-Real, J. L., Uribe-Toril, J., Torres, J. A., & De Pablo, J. (2020). Artificial intelligence in business and economics research: Trends and future. *Journal of Business Economics and Management*, *22*(1), 98–117. https://doi.org/10.3846/jbem.2020.13641

Safi, S. K., & Sanusi, O. I. (2021). A hybrid of artificial neural network, exponential smoothing, and ARIMA models for COVID-19 time series forecasting. *Model Assisted Statistics and Applications*, *16*(1), 25–35. https://doi.org/10.3233/MAS-210512

Sahed, A., Kahoui, H., & Mekidiche, M. (2020). Forecasting Algerian GDP using adaptive neuro fuzzy inference system During the Period 1990-2019. *Journal of Smart Economic Growth*, *5*(2).

Salah Uddin, K. M., & Tanzim, N. (2021). The Role of internal auditors in achieving the social responsibility of the commercial banks operating in Jordan: A field study. *International Journal of Business and Management*, *16*(6), 56. https://doi.org/10.5539/ijbm.v16n6p56

Samiyu, M. (2021). Multiple regression model for predicting GDP using macroeconomic variables (Part 1). *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.3895177

Sarvestani, S. E., Hatam, N., Seif, M., Kasraian, L., Lari, F. S., & Bayati, M. (2022). Forecasting blood demand for different blood groups in Shiraz using auto regressive integrated moving average (ARIMA) and artificial neural network (ANN) and a hybrid approaches. *Scientific Reports*, *12*(1), 22031. https://doi.org/10.1038/s41598-022-26461-y

Schaltegger, C. A., & Torgler, B. (2005). Growth effects of public expenditure on the state and local level: Evidence from a sample of rich governments. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.663685

Srinivasan, N., M, K., V, N., S M, K., Kumar, S., & R, S. (2023). Predicting Indian GDP with machine learning: A comparison of regression models. *2023 9th International Conference on Advanced Computing and Communication Systems (ICACCS)*, 1855–1858. https://doi.org/10.1109/ICACCS57279.2023.10113035

Stundziene, A. (2014). Prediction of Lithuanian GDP: Are regression models or time series models better? *Economics And Management*, *18*(4), 721–734. https://doi.org/10.5755/j01.em.18.4.5041

Tacchella, A., Mazzilli, D., & Pietronero, L. (2018). A dynamical systems approach to gross domestic product forecasting. *Nature Physics*, *14*(8), 861–865. https://doi.org/10.1038/s41567-018-0204-y

Tu, J. V. (1996). Advantages and disadvantages of using artificial neural networks versus logistic regression for predicting medical outcomes. *Journal of Clinical Epidemiology*, *49*(11), 1225–1231. https://doi.org/10.1016/S0895-4356(96)00002-9

Turan, G. (2011). Artificial Neuro Network (ANN) applications in economics: A survey of emprical literature and its using on economic Studies. *1st International Symposium on Computing in Informatics and Mathematics (ISCIM 2011)*. https://www.researchgate.net/publication/301560273_Artificial_Neuro_Network_ANN_Applications_in_Economics_A_Survey_of_Emprical_Literature_and_Its_Using_on_Economic_Studies

Urasawa, S. (2014). Real-time GDP forecasting for Japan: A dynamic factor model approach. *Journal of the Japanese and International Economies*, *34*, 116–134. https://doi.org/10.1016/j.jjie.2014.05.005

Vatavu, S., Dogaru, M., Moldovan, N.-C., & Lobont, O.-R. (2022). The impact of entrepreneurship on economic development through government policies and citizens' attitudes. *Economic Research-Ekonomska Istraživanja*, *35*(1), 1604–1617. https://doi.org/10.1080/1331677X.2021.1985566

Wang, L., Zou, H., Su, J., Li, L., & Chaudhry, S. (2013). An ARIMA-ANN hybrid model for time series forecasting: An ARIMA-ANN hybrid model. *Systems Research and Behavioral Science*, *30*(3), 244–259. https://doi.org/10.1002/sres.2179

Yang, B., Li, C., Li, M., Pan, K., & Wang, D. (2016). Application of ARIMA Model in the prediction of the Gross Domestic Product. *Advances in Intelligent Systems Research*, *130*.

Yang, Z. R., & Yang, Z. (2014). Artificial neural networks. In *Comprehensive Biomedical Physics* (pp. 1–17). Elsevier. https://doi.org/10.1016/B978-0-444-53632-7.01101-1

Yoon, J. (2021). Forecasting of real GDP growth using machine learning models: Gradient boosting and random forest Approach. *Computational Economics*, *57*(1), 247–265. https://doi.org/10.1007/s10614-020-10054-w

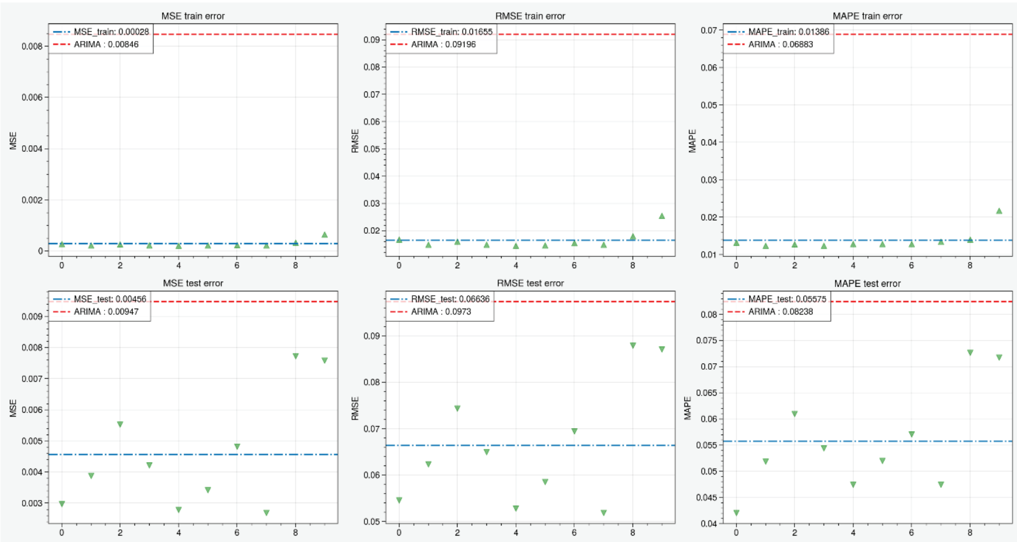
Yusof, S., & Samsudin, R. (2018). Comparison of artificial neural network (ANN) and autoregressive integrated moving average (ARIMA) models for WTI crude oil price forecasting. *Innovations in Computing Technology and Applications*, *3*.

Zhang, G. P. (2003). Time series forecasting using a hybrid ARIMA and neural network model. *Neurocomputing*, *50*, 159–175. https://doi.org/10.1016/S0925-2312(01)00702-0

Zhu, B., & Wei, Y. (2013). Carbon price forecasting with a novel hybrid ARIMA and least squares support vector machines methodology. *Omega*, *41*(3), 517–524. https://doi.org/10.1016/j.omega.2012.06.005

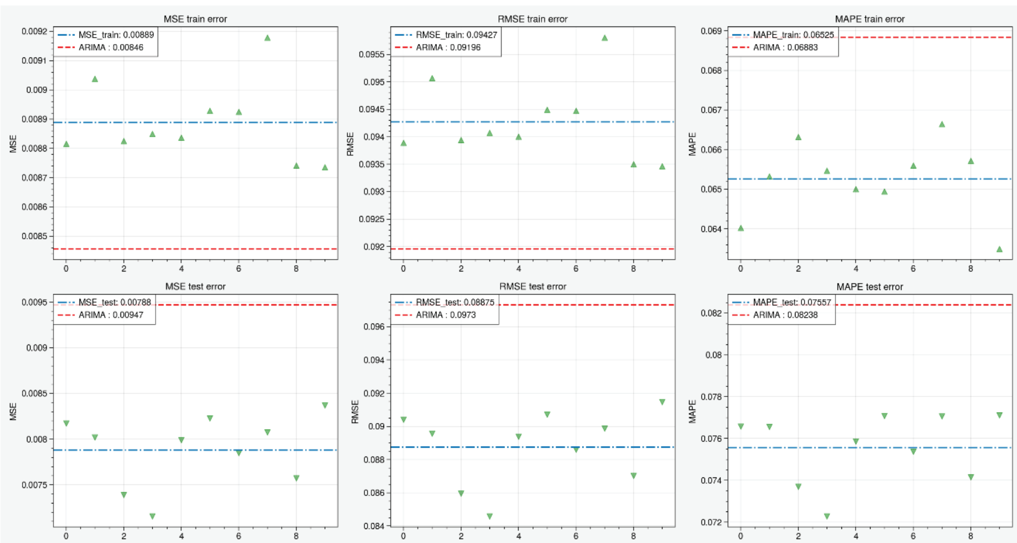
APPENDIX 1

A: ANN training and testing performance



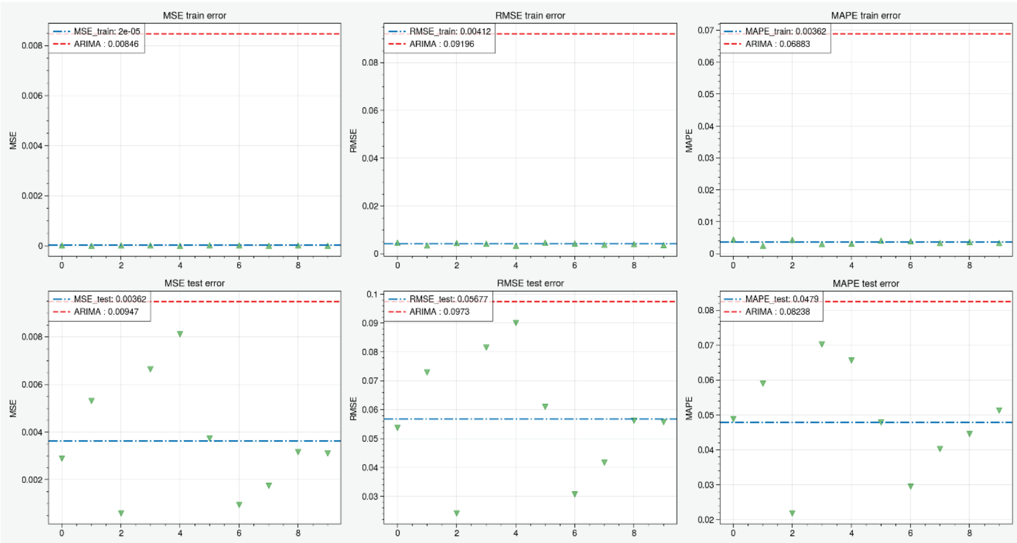
Source: Author's calculation

B: Zhang hybrid training and testing performance



Source: Author's calculation

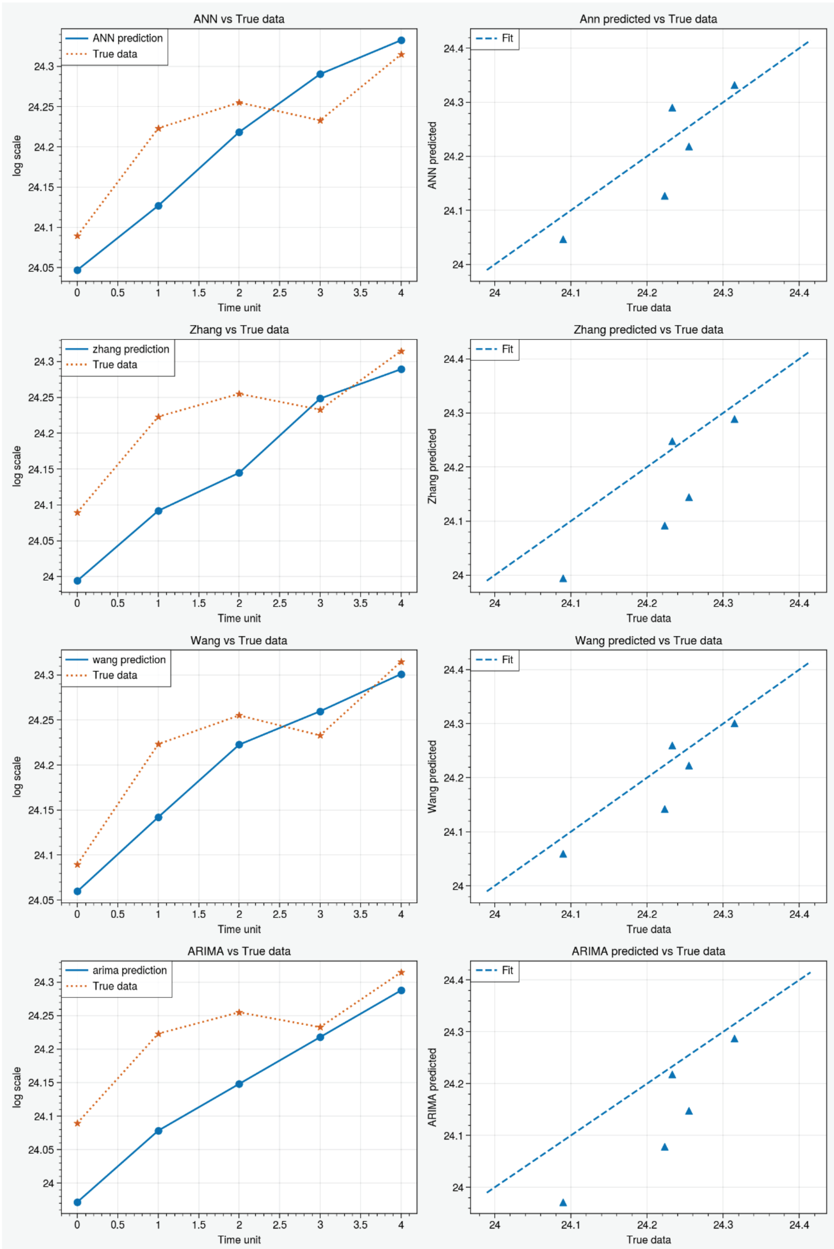
C: Wang hybrid model training and testing performance



Source: Author's calculation

APPENDIX 2

True value plotted against predicted value



Source: Author's calculation