Early Detection and Forecasting of Influenza Epidemics Using a Hybrid ARIMA-GRU Model

Dr. Kabilan Annadurai[1], Dr. Aanandha Saravanan[2], Dr. S. Kayalvili[3],
Dr. Madhura K[4], Elangovan Muniyandy[5], Inakollu Aswani[6], Prof. Ts. Dr. Yousef A.Baker El-Ebiary[7]
Department of Public Health-School of Health Sciences, The Apollo University, Chittoor, Andhra Pradesh-517002, India[1]
Professor-Department of ECE, Vel Tech Rangarajan Dr. Sagunthala R&D Institute of Science and Technology, Chennai, India[2]
Associate Professor-Department of Artificial Intelligence, Kongu Engineering College, Perundurai – 638060, Tamilnadu, India[3]
Department of information Technology, Manipal Institute of Technology Bengaluru, Manipal Academy of Higher Education, Manipal, India[4]
Department of Biosciences, Saveetha School of Engineering, Saveetha Institute of Medical and Technical Sciences, Chennai - 602 105, India[5]
Applied Science Research Center, Applied Science Private University, Amman, Jordan[5]
Department of Computer Science and Engineering, Koneru Lakshmaiah Education Foundation, Vaddeswaram, Guntur Dist., Andhra Pradesh - 522302, India[6]
Faculty of Informatics and Computing, UniSZA University, Malaysia[7]

***Abstract*—Early diagnosis and accurate epidemic prediction are essential in limiting the public health impact of influenza epidemics because intervention on time can effectively curb both the spread of the disease and the strain on health services. Standard ARIMA models have proven their usefulness in short-term forecasting, particularly in stable contexts, but the fact that they cannot keep up with the complex and non-linear dynamics of disease spread makes them less capable of dealing with rapid-evolving outbreaks. This is especially the case when outbreaks are characterized by complicated seasonal trends and irregular peaks which are challenging for ARIMA to predict by itself. To fill this deficit, this study presents a hybrid model that marries ARIMA's statistical strength in dealing with short-term trends and the high-powered deep learning strengths of Gated Recurrent Units (GRU) that specialize in detecting long-term dependencies and non-linear relationships in data. The WHO Flu Net dataset, a trusted source of influenza surveillance, forms the foundation of training the model, with careful preprocessing operations conducted to normalize the data and eliminate any missing values, providing high-quality input to the model to make precise predictions. By combining ARIMA's linear prediction strengths with GRU's sophisticated pattern detection, the hybrid model delivers a powerful solution that is better than both regular ARIMA and other machine learning models, as evidenced by lower error rates on test metrics like MAE, RMSE and MAPE. The experimental findings validate that the ARIMA-GRU model not only enhances predictive performance but also increases the model's sensitivity to subtle trends, making it a valuable asset for early detection systems in public health. In the future, the incorporation of real-time environmental information such as temperature, humidity, and mobility patterns may further enhance the model's accuracy and responsiveness, providing more robust forecasting. Also, integrating healthcare infrastructure-related data, i.e., hospital capacity and availability of medical resources, would aid in developing a more complete epidemic management process. In total, the ARIMA-GRU hybridization is an effective and novel strategy for enhancing influenza surveillance, outbreak detection at the early stage, and epidemic control operations.**

Keywords*—Time-series analysis; gated recurrent unit; temporal patterns; influenza epidemic; auto regressive integrated moving average; early detection

I. INTRODUCTION

Early detection and prediction of influenza plays vital role in healthcare system to assign resources efficiently. Influenza is the first global epidemic to be tracked and it is a severe airway respiratory illness caused by the influenza virus. It is primarily spread by droplets, and most people in the population are found to be vulnerable. Influenza can cause epidemics and outbreaks; globally, the average yearly incidence of influenza in adults and children is roughly five percent and twenty percent respectively and seasonal influenza is responsible for between 290,000 and 650,000 deaths, leaving a massive disease burden [1]. Influenza is a serious and extensive problem that affects everything from socio economic stability to individual health issues. If the flu outbreak is not contained, it might have a catastrophic impact on society as a whole. For instance, the H1N1 pandemic struck the globe in 2009. The WHO reported that there are 1.3 million confirmed cases worldwide, with over 14,000 fatalities. This posed a serious threat to the global quarantine system [2]. Between 1976 and 2007, even the less serious seasonal epidemics contributed to over 24,000 deaths in the United States per year; the most recent data projects that the total number of flu-related deaths will reach 61,000 during the 2017–2018 flu season [3].

Forecasting the flu season is required for better influenza epidemic control in advance and to achieve optimum coverage for the administration of health care. Because of the upgrades in efficiency and model accuracy, deep learning methodologies have been implemented enormously in time series forecasting complications like stock exchange prediction, climatic predication, health supervision analysis, and traffic flow estimation[4]. DNNs, a form of ANNs, have been recommended to overcome the complex and non-stationary time series, where traditional and statistical learning fail in terms of prediction.

DNN techniques are capable to choose better features out of time series information in a way that helps them to discover more intricate and non-linear relations [5]. Use of AI and ML methods provides a unique opportunity to enhance existing Influenza early warning systems. Such innovative approaches of analysis can assess huge scales of information, identify correlations and patterns, and generate forecast models that may adapt to conditions. This could mean that while traditional models are developed to search for one distinct signal of influenza occurrence, it may also use machine learning algorithms to look for earlier indicators which may be faint [6], [7]. Although the best approach for defining these thresholds is still a global issue, further improvement and refinement of the influenza early warning system is possible by integrating more accurate and adaptive technologies and processes [8]. Influenza monitoring and primary cautioning systems permit for the timely detection of influenza epidemic trends and the provision of scientific backing for influenza management and prevention, both of which are crucial for public health. Currently, a wide range of techniques are used to predict infectious diseases, each with pros and cons of their own. These include the ARIMA, LR method, NN forecast model, grey prediction model, and infectious disease dynamic model [9] [10]. Recently, there has been a lot of influenza prediction due to a Deep learning method such as LSTM, which has produced great findings that are more reliable when compared to the of other methods [11]. ARIMA is a well-known linear model that is frequently employed for time series forecasting.

Although it has been around for a long time, it has been used in forecasting tasks in a unique perspective. They noted that even with little data, this model translates to a large gain in replicating more complex forms of time series that show patterns, seasonality's, and transient dependencies [12]. This model belongs to the time series methods which has been used broadly for flu predictions due to its ability in providing frequency, trend, and variation of data information. Since this model is developed on the basis of past data, it has a fairly good ability to predict the trends that are likely to happen in the future. Thereby, this model turns out to be very operative in accurately forecasting the incidence rate [13]. To the time series data, an ANN can also be applied in the forecasting process. The RNN is one of the neural network approaches often used to analyze time series data sets [15]. Nonetheless, while handling long sequences of data, RNN has a very crucial problem normally identified as the vanishing gradient problem in which the gradients of the function decrease sharply. This is undesirable as it reduces the capability of RNNs to discern long term dependencies used in prediction results. In a case, where the sequence is long, standard RNN procedures are known to suffer from the vanishing gradient issue [14], the GRU model has been designed to ensure that it does not happen. In other words, to capture the long-range dependency of the time series data, the model needs to learn important features from the earlier period phases and is also influenced by the vanishing gradient problem. GRU an RNN variant on the other hand is used to find long term dependencies in sequential data. This makes GRU models more effective than traditional deep learning models that rely on aggregated features. It proves useful in the case of non-linear time series such as influenza forecast. It is applicable to model non-stationary data and the temporal characteristic while other models such as statistical models like ARIMA cannot [15], [16]. ARIMA is a conventional statistical model and it is designed for identifying the short-term patterns while it has an inability to identify the long-term and non-linear characteristics. This can be fixed with the help of GRU that capture temporal long-term dependencies and more complex patterns within the data. In this study, a combined ARIMA-GRU model has been used to detect and predict the influenza in the early stage. The usage of both ARIMA and GRU strengthens the model and it is identified to achieve higher accuracy than the conventional methods and also it improves the early identification and prediction of the influenza.

The following are the contributions made;

- Research proposed a Hybrid ARIMA-GRU Model that integrates the use of ARIMA for short-term prediction and GRU for catching long-term dependency and non-linear characteristics of epidemic data to achieve higher prediction precision for influenza outbreaks.
- Data processing enhancement occurs through the use of advanced normalization techniques which address missing value problems to optimize input quality. The performance assessment with three criteria MAE, RMSE and MAPE proves the gradient recurrence system's dominance compared to traditional forecasting models including ARIMA, SARIMA and LSTM.
- Work utilizes epidemiological data from the WHO Flu Net for model training which enhances practical use capabilities. The hybrid approach verifies its ability to perform real-time epidemic surveillance that supports public health decisions.
- Demonstration of the ARIMA-GRU model shows flexibility for using various infectious disease prediction applications. The forecasting capability of the system will improve by implementing future changes that incorporate real-time environmental data and mobility data for predictive modeling. This enhances epidemic preparedness strategies.

The construction of the study's remaining portion is as follows: Section II, review of relevant work is given. The problematic statement is in Section III. The suggested approach is explained in Section IV. The experimental findings are offered and compared in Section V. The conclusion and recommendations for further research are provided in Section VI.

II. RELATED WORKS

Three different feature spaces of EWARS data [17] from WHO worked in different models to predict weekly influenza rate of Syria. The initial approach involved the utilization of time series feature space and the application of seven distinct models. To predict the spreading of the devastating influenza pandemic using ML algorithm [18] an MSDII-FFNN, a forecasting model system for the influenza pandemic. It has the potential to identify the type of influenza causing pandemic using the proposed model. It can be utilized to control the harm and prevent its spread. Additionally, it can help the government manage the pandemic more effectively. Simulators are executed using MATLAB tools. The WHO is the source of the dataset. Two steps comprise this model, the model is revised on the cloud and trained using FFNN during the training phase. During the validation stage, the system's model is updated via cloud to anticipate the pandemic alarm each time an input is received by IoT devices. The dataset is split into 15% validation ratio and 85% training. It attains an output precision of about 90%. FFNNs are not designed to take temporal relationships in data. They are more expensive and challenging to train since they require more layers and neurons. Fractional SEIR model is for monitoring and predicting influenza transmission [19]. Also, the ARIMA model which predicts the yearly evolution of influenza epidemics. The analysis validates that the model with fractional orders agrees with empirical data and performs better than the ARIMA model. The fractional SEIR model was used to simulate the confirmed cases, while the ARIMA model was utilized to forecast the seasonal patterns of the influenza pandemic. The findings highlight the importance of developing numerical techniques with precise parameter values and applying fractional models to medical risk management. For this prediction, it is argued that improving the existing pandemic mathematical models and feasible measures to control influenza is crucial. Regarding the performance of the models, it has been found that the fractional SEIR model was better than the ARIMA (2, 0, 1) with a zero mean but a non-zero mean. The result indicates that the fractional SEIR model may provide maximum likelihood estimates for predicting the confirmed cases of flu. The high nonlinear interactions between variables as captured by the ARIMA models are expected to be linear thus restricting their effectiveness in designing models to capture non-linearity.

Utilization of XG Boost model to forecast the average monthly [20] detection of the influenza for the year 2020 and year 2021. The authors have also shown in this study how the ARIMA and the SARIMA have been compared. Forecasting techniques help in the tracking of the incidence of the influenza virus in minimizing the effects of the disease. The data was collected from Saudi Arabian Ministry of Health. From several calculations, MSE of 43791.75, MAE of 172.55, RMSE of 209.26, and a value of R^2 of 0.775 are deduced for the training set to determine that the effectiveness of the ARIMA models in forecasting the levels of influenza cases are found to be low. Concurrently, the result of “XG Boost” model was the peak with the $R^2 = 0.999$, RMSE = 1.94, MAE = 1.39, and MSE = 3.75. Based on these results, it is found XG Boost has more precision as compared to the other models. The nonlinear and complex interactions may be trapped by the XG Boost model as compared to the “ARIMA” and “SARIMA” models and may result in better accuracy in the predictions made. However, when it comes to sparse and unstructured data, XG Boost has a worse accuracy compared to other methods. The probability of mutations to occur in the upcoming flu season based on [21] previous glycoprotein hemagglutinin sequence data. Modeling and interpreting the outcome of the timing and dimensionality of successive influenza strains is one of the main concerns. A sound and effective Tempel for predicting influenza A viral mutations is presented here. Tempel takes into account past residue knowledge using RNN with attention processes. This research received datasets from H1NI, H3N2, and H5N1. Experimental results demonstrate that Tempel can greatly improve predictive performance over popular methods and shed new light on viral development and mutation dynamics. Moreover, the method precision can be enhanced by assigning proper attention weights. As the model rely on the past data for training it is unable to learn new features from the data. It needs additional data for training to improve the method.

The influenza provide, new deep neural network models like RNN and conventional autoregressive (AR) techniques [22]. Transformer models outperform RNN models in capturing long-range dependency. The model utilizes the ability of the Transformer to enhance predictability. To achieve integration of information from multiple sources and representation of author design a sources selection module that utilizes curve similarity measurement. It is compared with well-known AR models and “RNN-based” models using datasets in the USA and Japan. Six baseline models were compared in short- and long-term conditions to evaluate the methodology. The results indicate that the approach yields better long-term forecasting performance and approximation performance in short-term forecasting. When training separate models for different horizons, the proposed strategy is not flexible enough. The drawback of the traditional AR method is that it needs stationary data to provide better results. Various classifiers such as ML and DL have been employed in the prediction of influenza outbreaks. LSTM models were found to be more precise in identifying influenza patterns. The SEIR model also achieve better performance than the conventional ARIMA as the model understood non-linear interactions. In influenza prediction XG Boost surpassed both ARIMA and SARIMA by explaining intricate data structures and Transformer models excelled in long-term forecasting. In DL the methods like CNN is used by using X-rays, MRI. In pre-processing noise removal, scaling, augmentation, data partitioning and one hour encoding. The challenges are lack of reliable, large scale datasets during early pandemic. Data imbalance and limited labeled samples. No representative global datasets. Overfitting due to small training datasets. Limited generalizability of trained models. Develop more robust DL models for pandemic disease detection. Encourage dataset standardization and global sharing. Incorporate explainable AI for clinical decision support. Improve early detection systems and reduce dependency on lab tests. Still, there are some drawbacks in the models such as the model’s flexibility, the data on which the model depends, and capturing temporal features, which is quite hard to capture when employing certain models, including FFNN and other conventional AR approaches.

III. PROBLEM STATEMENT

The main emphasis of these studies is solving the problems of flu forecasting and control using different machine learning (ML) and deep learning (DL) methods. Nevertheless, conventional forecasting models such as ARIMA and SARIMA have limitations, especially in their capacity to capture the non-linear and intricate patterns that are characteristic of time series data [20]. A few approaches such as LSTM, XG Boost and transformer models have demonstrated satisfactory performance particularly when it comes to managing long-term dependencies. Nevertheless, difficulties associated with generalization and the capacity to handle rare and unstructured data are yet to be addressed. FFNN and RNN models have been found to fail in efficiently describing temporal dynamics and large datasets are required to improve predictive precision. Thus, there is a critical need for creating strong and flexible approaches which can respond to these issues [18]. Thus, it is crucial to design robust and adaptable techniques for forecasting utilized in influenza prevention and management. Hence, a hybrid ARIMA-GRU model have been proposed for influenza early detection and prediction. Non-linear and the complicated patterns in data can be extracted using GRU, a modified version of RNN and ARIMA. Time series technique is utilized for determining short term characteristics from data. It comprehends the data and forecast the upcoming points in time-series. Through the integration of the power of both statistical model and deep learning model, it will enhance the overall precision and provide improved results.

IV. PROPOSED ARIMA-GRU MODEL FOR EARLY DETECTION AND PREDICTION OF INFLUENZA

The Proposed ARIMA-GRU Model for Early Detection and Prediction of Influenza aims to exploit the strengths of two robust techniques comprises ARIMA (Auto Regressive Integrated Moving Average) and GRU (Gated Recurrent Unit) deep learning models to enhance influenza forecasting accuracy and reliability. The ARIMA technique is utilized for capturing short-run trends and linear patterns in epidemic data, and GRU for modeling intricate non-linear dependencies as well as long-run trends. Through the combination of both methods, the hybrid model seeks to offer more accurate early warning and prediction of influenza outbreaks, enabling the public health authorities to initiate prompt preventive actions and mitigate the effects of the disease.

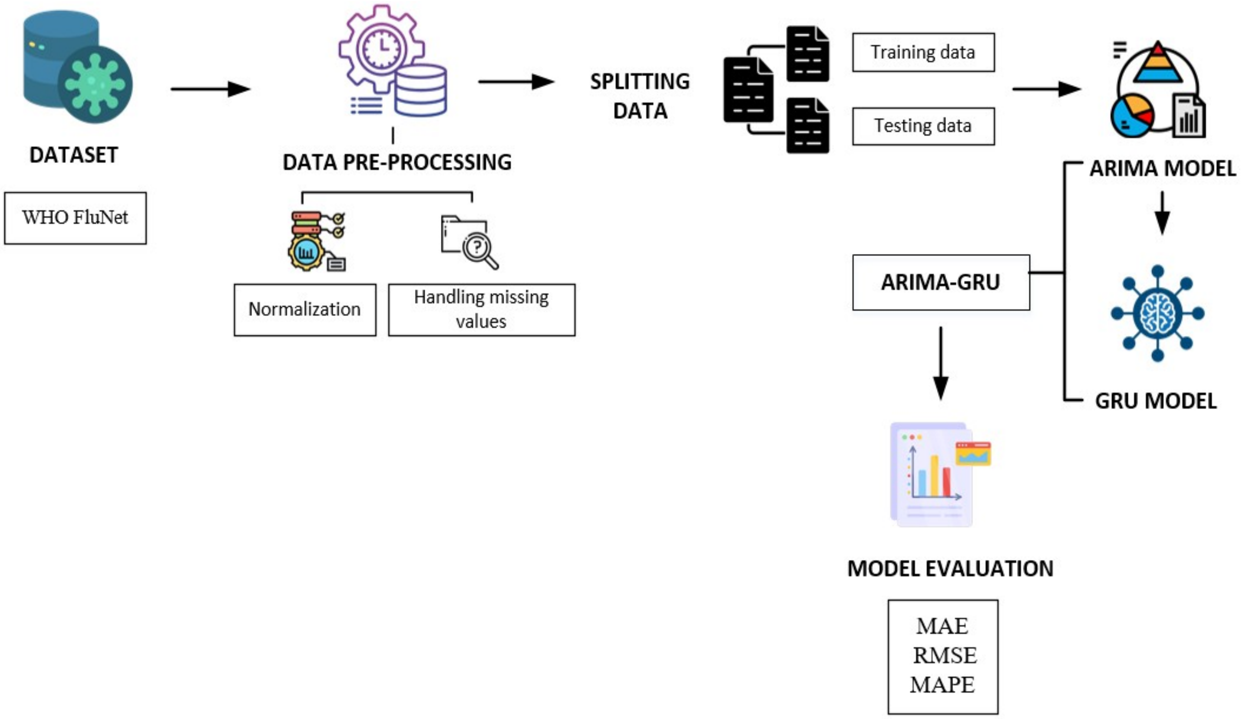


Fig. 1. Workflow of proposed method.

Fig. 1 presents a dual-panel illustration that encapsulates both the data handling pipeline and the performance evaluation of different machine learning models. The left panel is a detailed flowchart depicting the process starting with the WHO Flu Net dataset symbolized by a database icon paired with a virus motif followed by essential preprocessing steps such as data normalization and addressing missing values. The cleaned data is then split into training and testing sets, with the training portion used to develop an ARIMA model that is integrated with a GRU network, culminating in an evaluation phase that utilizes metrics like MAE, RMSE and MAPE. In contrast, the right panel is a bar chart titled "Model Accuracy Comparison" that visually contrasts the accuracy percentages of several models: Logistic Regression (91.2%), Random Forest (93.5%), Support Vector Machine (94.1%), Long Short-Term Memory (96.0%) and the Proposed CNN + BiLSTM (99.5%). Together, these panels present a comprehensive narrative of the project by linking the methodical progression of data transformation and model construction with a clear comparative assessment of model performance.

A. Data Collection

Data obtained in the influenza dataset is based on weekly reports on flu submitted to the WHO Flu Net by 167 countries for several years starting from 2004. Therefore, this dataset is informative of influenza trends worldwide, but it gives only a glimpse at the number of infections in a population. In the data, the cases are not standardized across the years or even between countries for a number of reasons including, differences in the reporting systems, healthcare centers, and diagnostic tests.

Table I displays weekly data for a specific variable (likely related to health, such as influenza cases or another measure) across four countries: Australia, Algeria, Norway and Spain. Over the span of six weeks, all four countries show a steady upward trend in the values. Australia begins with 750 in week 1 and reaches 950 by week 6. Algeria starts at 800 and rises to 950, while Norway shows a similar increase, from 800 in week 1 to 980 in week 6. Spain exhibits the highest values, starting at 900 and increasing to 1000 by week 6. This data highlights the gradual growth of the metric in each country over time.

TABLE I. WEEKLY INFLUENZA CASES IN SELECTED COUNTRIES

Week	Australia	Algeria	Norway	Spain
1	750	800	800	900
2	780	820	830	920
3	820	850	870	940
4	860	890	910	960
5	910	920	950	980
6	950	950	980	1000

The dataset is in CSV format contains important epidemiological data which can help in model development and forecasting of flu outbreaks. However, the set of data provides a very exhaustive picture of the flu trends and could be an informative data input for epidemic prognosis and early warning models [23].

B. Data Pre-processing

It is the procedure of preparing unprocessed data for deep learning model training and it represents the first phase of the development of the model. The deep learning models cannot be taught just feeding it raw data. The most critical and significant factor influencing the model's ability to generalize is data pre-processing. In order to identify and eliminate inaccurate or noisy data from the dataset, this method involves data cleansing. It usually functions to detect and replace any noisy, inaccurate, incomplete, or irrelevant data and records. Pre-processing data plays an important part in artificial intelligence by improving the accuracy of the models. So, for the proposed model ARIMA, it is essential to check whether data is stationary or not, which means that the data's variance and mean must remain constant. There are numerous techniques for converting data into a stationary state. Log-scale transformation and time-shifting transformation were used in this instance. Time-series data with non-constant variance can be stabilized using the log-scale transformation to give more normal distribution. It is done by taking the log of numerical values in the given dataset [26].

1) Missing values: Deep learning, Missing Values (MVs) are the data attributes that may be absent from a dataset as a result of faults that may occur due to inaccurate measurements or failure of the device. Insufficient data in the collection can lead to poor accuracy in the mathematical model produced from the data. The accuracy of the model can be additionally impacted by missing attributes. Missed attributes, for instance, might lead to uneven length in the environment of decision-tree induction. They can also cause uneven feature allocation and divide the dataset into testing and training sets. The data must be removed if over twenty percent of the data utilized is missing. In a dataset, it is common for certain columns to have missing values. Data collection or data validation guidelines may be the cause of the problem. However, missing values might lead to the elimination of a model's feature, it is essential to take them into account. Simple methods of interpolation can fill in the gaps left by a fair number of missing values. The most popular approach to handling it is to employ model feature mean, median, or mode values [24]. To find out if there is a correlation among factors in a dataset, a missing value algorithm is utilized. For example, M contains a dataset (a, b), where N is a random variable, 'b' is M's missing value and 'a' is M's observed value. Assuming that N = 1 holds true independent of whether M's have been detected or missing values, the observed value can be found by expressing M = 0 as a model Q (N/M, ¸), where Ø denotes the missing value. N's reliance on the variables in the dataset forms the basis of the method for filling in the missing values [25].

2) Normalization: Normalization is necessary to bring the characteristics of numerous features in the same scale or to avoid getting poor outcomes because each feature may have its own scale. It consists of "decimal scaling", "z-score normalization", and "min-max normalization"[24]**.** Estimating the mean and standard deviation are the process involved in normalization. The raw EEG data is normalized using the min-max approach. The data will be prepared for additional extraction and classification steps following this preprocessing technique [26]. One of the normalization technique that can be applied is min-max scaling, which works with features that have a linear distribution and feature values that fall between 0 and 1 or (-1) and 1 [27]. The method for min-max normalization is given in Eq. (1),

$$E_{norm} = \frac{E - E_{min}}{E_{max} - E_{min}} \tag{1}$$

where, normalized data is represented as E_{norm} . E is the value of raw feature data, minimum and maximum feature values are indicated as E_{min} and E_{max} respectively.

C. ARIMA for Feature Extraction

Different time series modeling methods have been created, with the most popular method being ARIMA [12]. This can be used to model univariate data that is trended, seasonal, and has a cyclical pattern. A variable is forecasted based on its past values by the ARIMA (a, b, c) model. The integration (I), moving average (MA), and autoregression (AR) constitute the three components of the ARIMA model. Autoregressive, or AR, models compare the single period's pattern to previous time periods of the same. Moving averages, or MAs, used the errors of a previous time-step to forecast the variable in a future process [28]. The method of producing the forecast by reversing the differencing process is referred to as integration (I). Three parameters constitute the ARIMA model: a, b, c. Autoregressive term of the relation among the present value and the preceding values is accounted for by parameter a. Number of transformations via differencing steps taken to achieve a stationary form of the time series is described by parameter b. The term moving average, or parameter c, to strip off the randomness fluctuations. a -order autoregressive model is represented by the AR (a) model. c-moving average epresented by the MA (c) model. The generalized equations for the AR (a) model are given in Eq. (2) and Eq. (3),

$$x_t = \emptyset_1 x_{t-1} + \emptyset_2 x_{t-2} + \cdots + \emptyset_a x_{t-a} + \varepsilon_t \qquad (2)$$

$$x_t = \varepsilon_t - \theta_1 \varepsilon_{t-1} - \theta_2 \varepsilon_{t-2} - \cdots - \theta_a \varepsilon_{t-c} \qquad (3)$$

where, $\emptyset_i$ $(i = 1,2 \ldots a)$ is the moving average parameter at ith time-stamp, $\theta_i (i = 1,2 \ldots a)$ is the auto-regressive parameter, and ε_t is a zero-mean white noise series. The AR (a) model and the MA (c) model can alternatively be represented using the backshift operator; that is represented in Eq. (4) and Eq. (5),

$$\begin{cases} \emptyset(B) x_t = \varepsilon_t \\ \emptyset(B) = 1 - \emptyset_1 B - \cdots - \emptyset_a B^a \end{cases} \qquad (4)$$

$$\begin{cases} x_t = \theta(B)\, \varepsilon_t \\ \theta(B) = 1 - \theta_1 B - \cdots - \theta_c B^c \end{cases} \qquad (5)$$

where, $\emptyset(B)$ is the p-order auto-regressive polynomial, $\theta(B)$ is the c -order moving average polynomial, B is the backshift operator which is represented in Eq. (6),

$$B^m x_t = x_{t-m} \qquad (6)$$

where, B^m represents an operator or matrix applied m times to x_t. x_t is the value of the variable (or data point) at time t. It could be a signal, a sequence, or some time-dependent data point. x_{t-m} is the value of the variable at time $t - m$, meaning m steps before t. The differencing process can alternatively be represented using the backshift operator as average polynomial, which is given in Eq. (7),

$$y_t = (1 - B)^b\, x_t \qquad (7)$$

where, y_t is represented as fixed time series and x_t is a non-stationary time series. The equation of the ARIMA (a, b, c) model can be obtained by incorporating the Eq. (4), (5), and (7). It can be stated as Eq. (8),

$$\emptyset(B)(1 - B)^b\, x_t = \theta(B) \varepsilon_t \qquad (8)$$

where, $\emptyset(B)$ a polynomial in the backward shift operator B usually describing an autoregressive (AR) part of a model, B is typically a lag operator, $(1 - B)^b$ represents the differencing operator applied to the time series datax_t, $\theta(B)$ is a polynomial in the backward shift operator, ε_t represents the error term.

ARIMA is a data-driven linear approach that modifies the parameters data. Consequently, nonlinearity in the data significantly affects how well the ARIMA model performs. This is an ARIMA model limitation because significant non-linear data patterns may reduce the ARIMA model's applicability.

D. GRU

GRU is the enhanced method of LSTM and RNN, the gated recurrent unit can efficiently retain the relevant information and relationships between input sequences while removing irrelevant information to save processing time and memory use. GRU is frequently employed in predicting sequential data and shortens processing times because of its uniqueness. To decrease the delayed execution in the neural network, GRU is modified from LSTM. The LSTM's structure is simplified into the GRU, which has two gates but no independent memory cell. To quickly analyze the current output state, a single update gate is built in GRU, which took the role of the input gate and the forget gate in LSTM. In order to remove irrelevant information from the previous hidden state, the reset gate was added to GRU. The vanishing and expanding gradient problem, which arises from constant multiplication during Backpropagation Through Time, is the core difficulty with RNNs. As shown in Fig. 2, GRU uses the update gate and reset gate to solve this issue.

Fig. 2 offers a concise dual-panel view of the forecasting process using WHO Flu Net data. The left panel illustrates GRU's core mechanism that emphasize its update and reset gates and manage the influence of past information. The right panel outlines the data flow: raw WHO Flu Net data is preprocessed (with normalization and missing value handling), split into training and testing sets, and modeled using an integrated ARIMA and GRU approach with performance measured via MAE, RMSE and MAPE.

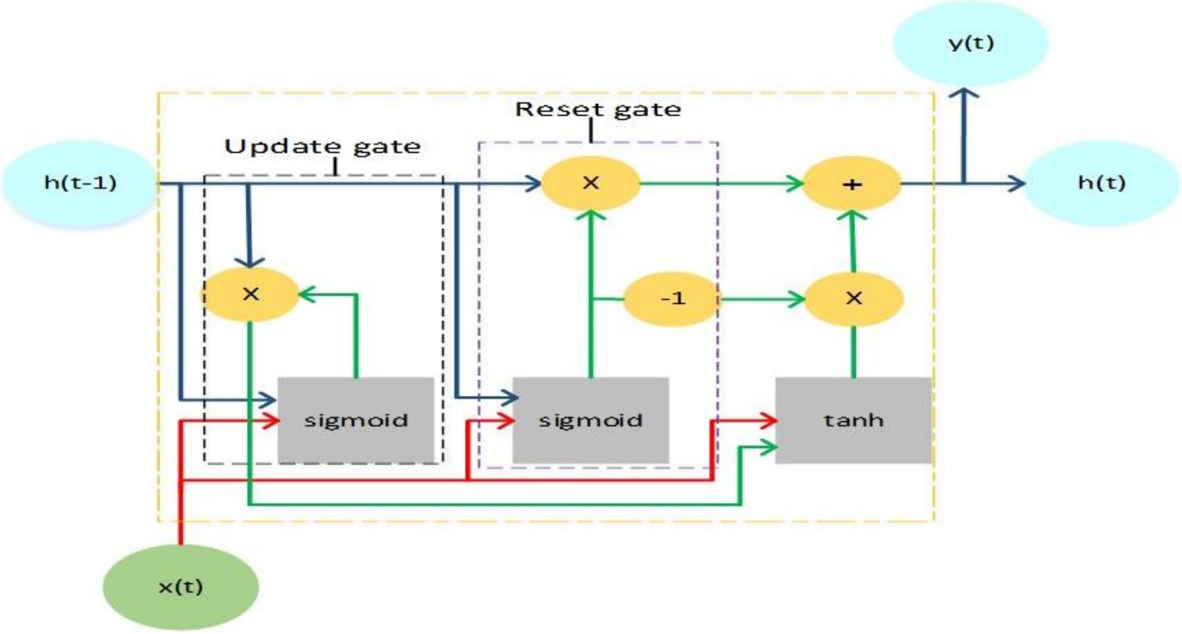


Fig. 2. Structures of gated recurrent units.

The first step is to use Eq. (9) to compute the update gate (z_t), which indicates the amount of past information that needs to be stored.

$$z_t = \sigma(w^{(t)}x_t + u^{(t)}h_{t-1} + \text{b}) \tag{9}$$

where, b is the bias, w and u were denoted as weights, x_t input, and h_{t-1} is the hidden state. Eq. (10) is then used to compute the reset gate (r_t), indicating that the amount of previous data should be removed and how to incorporate the new input with the previously stored data. Here is the formula for the reset gate:

$$r_t = \sigma\,(w^{(r)}x_t + u^{(r)}h_{t-1} + \text{b} \tag{10}$$

After that, determine the candidate hidden state ($h't$), that the reset gate will employ to preserve significant historical data. The candidate hidden state is expressed by Eq. (11)

$$h't = \tan h\,(wx_t + r_t \odot uh_{t-1} \tag{11}$$

where, $\odot$ is represented as Hadamard product. The last step is utilizing Eq. (12) to compute the hidden state (ht). The output (y_t) is this:

$$h_t = z_t \odot h_{t-1}\,(1 - z_t) \odot h't \tag{12}$$

Numerous hyperparameters, including batch size, learning rate drop, amount of hidden layers, are involved in GRU which has an impact on the prediction results. Batch size specifies how often the weights are changed, learning rate drop tells how many iterations were used to determine the learning rate, and the number of hidden layers indicates the extent of the training process.

E. Integration of ARIMA-GRU Model

The suggested approach, ARIMA-GRU incorporates the best aspects learned from ARIMA and GRU to improve influenza forecasts. Automated, effective, and strong, ARIMA is one of the most accurate ways of modelling linear trend level, seasonality and cycles in a time series. It does this by breaking down the time series into three pieces: “ARIMA” which is an abbreviation of autoregressive (AR), integrated (I), and moving average (MA) to forecast future values based on past values. But it is challenging for the model to operate in nonlinear cases, something that occurs regularly in real-world circumstances like infection rates of the actually nonlinear Previous.

The Gated Recurrent Unit serves as another special neural network design which specializes in solving progressive dataset vanishing gradient problems. The memory control of GRU helps the network remember significant past inputs while discarding trivial information making it an ideal solution for processing temporal patterns with strong dependencies. The circuit utilizes two control mechanisms termed “update gate” and “reset gate” to administer information flow throughout its structure. The update gate detects the previous data to conserve and the reset gate determines forgotten data proportions. ARIMA-GRU allows modeling of cyclic and seasonal patterns in influenza dataset predictions through its prediction component. The residuals from ARIMA receive input into a GRU system for finding non-linear relationships that cannot be detected through ARIMA. The combined approach allows the model to regulate both linear and non-linear time series components so it achieves better predictive performance. Flu Net receives data cleaning followed by transformation into stationary format using differencing in preparation for an ARIMA model used for flu prediction. The linear trends are estimated or modeled and the forecasts are created through the ARIMA component. The non-linear patterns visible in the residuals from the ARIMA model are fed to the GRU which enhances the predictions by providing a learning of the temporal patterns that is not amenable to learning by the ARIMA model, as it is designed to do. Thus, combining these two approaches is optimal in terms of performance when it comes to foreseeing factors related to influenza outbreaks, thus ensuring timely measures for prevention and intervention.

V. RESULT AND DISCUSSION

The evaluation of the proposed hybrid GRU-ARIMA model through predictive performance indicators revealed better accuracy than standard ARIMA models as well as other machine learning methods. The testing results demonstrated that the GRU-ARIMA model generated adjusted MAPE, RMSE and MAE scores that were lower than other testing samples indicating accurate forecasting while monitoring short-term fluctuations and long-term patterns in influenza diseases. This means that hybrid models are superior in measuring complex temporal patterns and doing early forecasting about the onset of epidemics compared to other methods, making them the best available tool for forecasting influenza outbreaks. These results prove the model's versatility in the area of early epidemic detection and the potential use of such detected patterns for decision-making and public health interventions. All these results demonstrate the versatility of the model in early epidemic pattern identification and the usefulness of the revealed patterns for decision making and public health intervention.

A. Evaluation Metrices

Model performance evaluation is the process of monitoring a model to determine how well it performs the specific task for which it was developed. There are several ways in which model evaluation can be done when it comes to model monitoring. The performance of the proposed model is evaluated using “MAE”, “RMSE”, “MAPE”, and other statistics.

1) Root mean squared error: Standard deviation residuals or variance among the estimated and real values is referred to as RMSE. It is measured by RMSE, indicates the degree of the spread of these residuals whereas the residuals alone provide a measure of how much away these data points are from the regression line. It refers to how focused the data is on the best-fit line. In regards to the experiments on climatology, forecasting, and regression analysis, it is typical to employ root mean square error, as given in Eq. (13),

$$\text{RMSE} = \left(\frac{1}{X}\right) \sum_{x}^{(i=1)} (z_i - \cap z_i)^2 \tag{13}$$

where, $\cap z_i$ is the model's prediction, X is the sample size, and z_i represents the actual expected output [29].

2) Mean absolute percentage error: The prediction precision is measured using MAPE. It can be used to get the MAPE: [12], that is given in Eq. (14),

$$\text{MAPE} = \frac{1}{n}\sum_{t=1}^{n}\left|\frac{e_t}{y_t}\times 100\right| \tag{14}$$

where, MAPE Mean Absolute Percentage Error, $\frac{1}{n}$ is the average over the number of observations, $\sum_{t=1}^{n}$ represents the summation (sum) over all time periods, e_t is the forecast error at time, y_t is the actual value at time

3) Mean absolute error: A statistic method called MAE finds the mean proportion of the absolute errors among expected and actual values. It is calculated using Eq. (15),

$$\text{MAE} = \frac{1}{T}\ \frac{\sum_{I=1}^{T}|X_1 - \hat{X}|}{T} \tag{15}$$

where, MAE represents Mean Absolute Error, $\frac{1}{T}$ represents the average over the total number of observations, $\left|X_1 - \hat{X}\right|$ is the absolute error for the $i-th$ observation, T is the total number of observations.

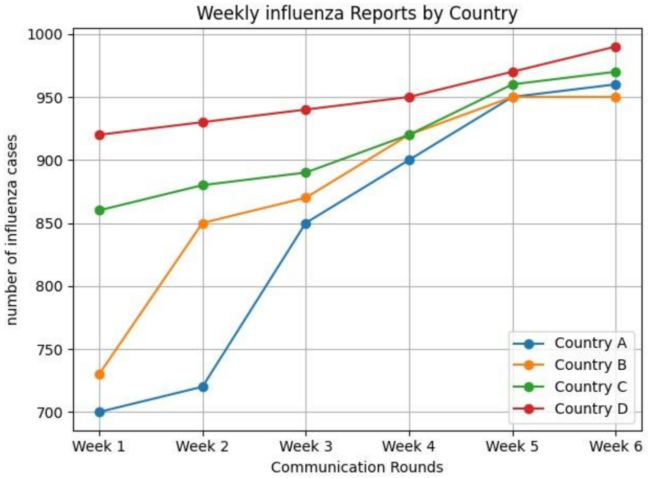


Fig. 3. Influenza reports of 4 countries.

Fig. 3 combines two panels into one succinct visualization. The left panel is a line graph charting weekly influenza cases over six weeks for four different countries, showing gradual increases in case numbers. The right panel presents a simple schematic of a GRU cell, highlighting its update and reset gates that manage how previous information is retained or forgotten. Together, these elements illustrate both the observed temporal trends in flu cases and the neural mechanism used for processing sequential data.

Country A (Australia) began with about 750 cases and concluded near to 950, which points to a consistent upward trend. Approximately 800 cases began in Country B (Algeria) around Week 3, escalating to 950 by Week 6, much like Australia's trend. Norway's journey started much like Algeria's, but concluded a bit higher than both, maintaining a steady upward trend. Country D (Spain) started nearly 900 cases and ended at nearly 1000, which indicates a stable increase throughout the time. Overall, all countries observed an upward trend in reported cases of influenza, stressing the need for powerful public health initiatives to deal with rising infection figures. The comparison of data reveals distinctive trends that may guide intended interventions and resource allocation.

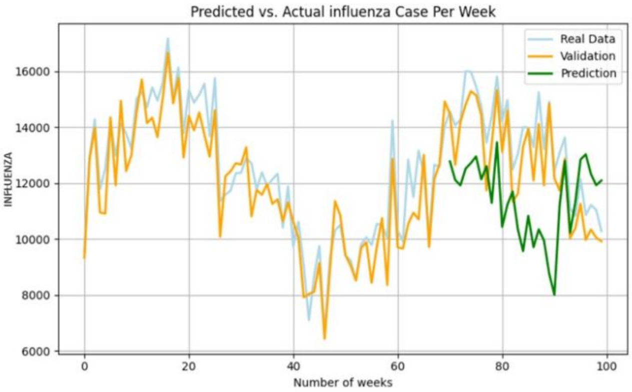


Fig. 4. Comparisons of predicted and actual cases of influenza per week.

Fig. 4 is the comparison of actual case data for influenza with model predictions and validation. The blue line on the graph gives the number of actual recorded cases per week of flu which are real incidence as there is reasonable fluctuation from week to week. The model performance evaluation scale using the orange line from a validation dataset closely following the real data. In same way the green line shows the predicted ARIMA model values and also it follows with real data trend. It is a method of getting an idea about how well the ARIMA model can predict future influenza cases which gives us some indication of how closely the model matches reality and hence how trust we can have in the predictions when doing public health planning.

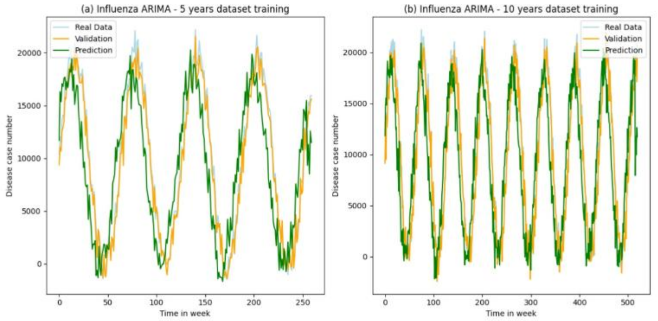


Fig. 5. Comparison of ARIMA model predictions with real data for influenza.

Fig. 5 is the Assessment of Forecasts with Real Statistics for assessing how well an ARIMA model predict the cases of influenza. Actual recorded cases (blue line), which vary quite a bit from week to week due to real Wattage fluctuations. On the other hand, we have an orange line (mean) using it from validation dataset which is real data. The green line above is ARIMA model prediction, and it well fits with the trend of real data. This comparison illustrates the precision in estimating influenza cases.

Fig. 6 is the Training and Validation Accuracy of model that illustrates the learning progress of the model over time. The blue line represents training accuracy, beginning around 0.70 at epoch 0 and gradually increasing to approximately 0.95, The orange line, representing validation accuracy, tracks the training accuracy with minor fluctuations. The close alignment between training and validation accuracy shows minimal overfitting, indicating that the version specifies fit to hidden data.

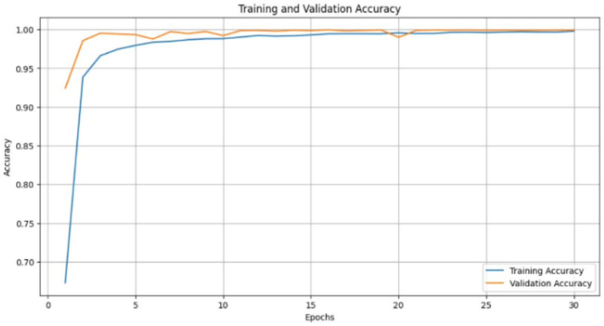


Fig. 6. Training and validation accuracy of GRU.

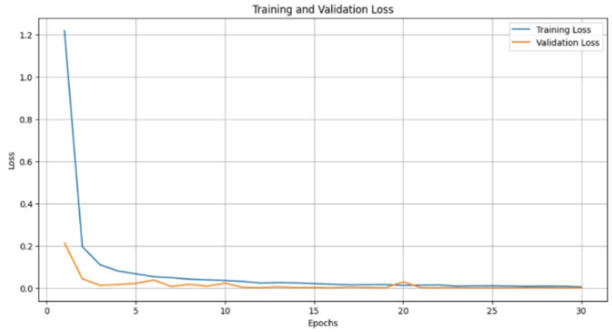


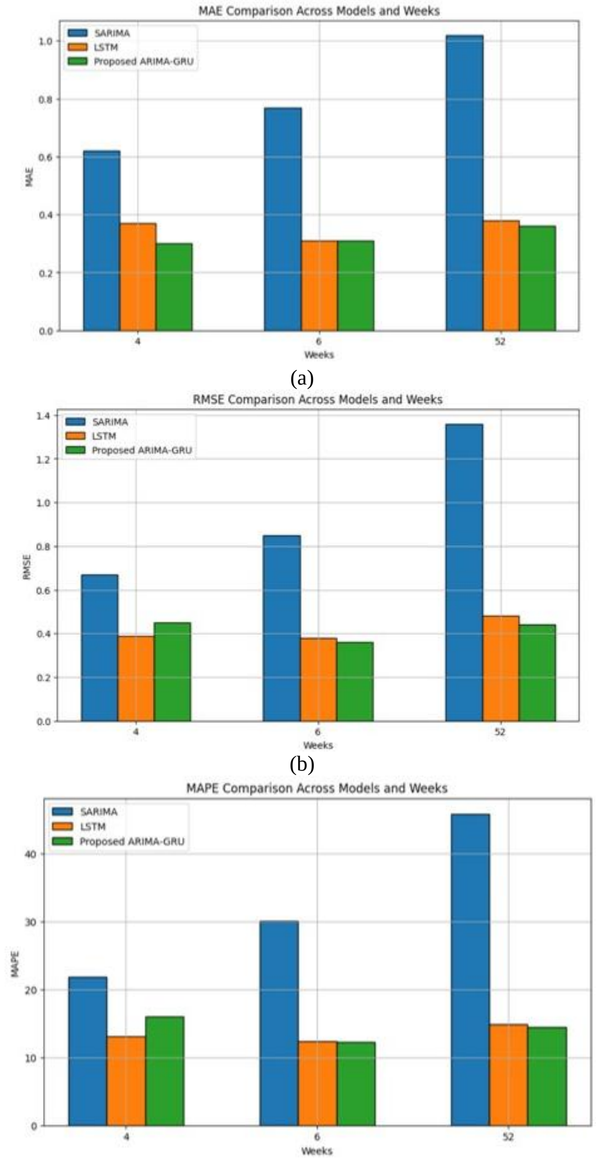
Fig. 7. Training and validation loss for GRU.

Fig. 7 shows the graph of loss in a GRU model. It represents the progression in training and validation losses across epochs. At epoch 0, the training loss initiates just past 1 and quickly drops in the initial few epochs, finally stabilizing near zero around epoch 15, meaning enhancements in model fit. Simultaneously, the validation loss starts somewhat above 1, proceeds at a slower rate, hitting its lowest state between epochs 5 and 10, and then a tiny increase before stabilizing around 0.2. The stabilization of both losses points that more training does not offer much improvement, require a balance between learning and avoiding overfitting. The close connection observed between the two loss curves indicates that the model is adapting well to unseen data without a substantial degree of overfitting. Doing this helps to guarantee that the model's performance remains sound and it does not overfit as training moves onward.

Table II demonstrates the performance of the “SARIMA”, “LSTM” and proposed “ARIMA-GRU” methodologies across different forecasting periods (4, 6, 52 weeks), where the hybrid ARIMA-GRU model proves superior. ARIMA-GRU achieved superior performance than SARIMA through results with lower MAE, RMSE and MAPE measures. At week 4 ARIMA-GRU produced an MAE of 0.30 as well as RMSE of 0.45 and MAPE of 15.98 whereas SARIMA yielded substantially greater assessment errors (MAE: 0.62 and RMSE: 0.67 and MAPE: 21.87). During the six weeks’ timeframe ARIMA-GRU demonstrated remarkable assessment results (MAE 0.31 RMSE 0.36 and MAPE 12.28) which outperformed both SARIMA and matched the LSTM model's performance. The 52-week forecasting results of ARIMA-GRU yielded the most accurate prediction values of MAE (0.36), RMSE (0.44) and MAPE (14.48) [30], while surpassing LSTM and surpassing greatly the performance of SARIMA. The ARIMA-GRU model demonstrates ability to handle short-term and long-term patterns through effective tracking with superior precision outcomes. Fig. 8 displays the model comparison.

TABLE II. COMPARISON WITH VARIOUS MODELS

MODEL	WEEKS	MAE	RMSE	MAPE
SARIMA	4	0.62	0.67	21.87
	6	0.77	0.85	30.04
	52	1.02	1.36	45.84
LSTM	4	0.37	0.39	13.14
	6	0.31	0.35	12.37
	52	0.38	0.48	14.86
Proposed ARIMA-GRU	4	0.30	0.45	15.98
	6	0.31	0.36	12.28
	52	0.36	0.44	14.48



(a)

(b)

(c)

Fig. 8. (a) MAE comparison across models and weeks; (b) RMSE comparison across models and weeks; (c) MAPE comparison across models and week.

Fig. 8(a) illustrates the Mean Absolute Error (MAE) for three forecasting models such as SARIMA, LSTM and the Proposed ARIMA-GRU evaluated over 4, 6 and 52-week periods. The visualization clearly shows that SARIMA consistently exhibits the highest MAE values, while the Proposed ARIMA-GRU achieves the lowest MAE across all time frames.

Fig. 8(b) presents a side-by-side bar chart comparing the Root Mean Square Error (RMSE) of the same three models (SARIMA, LSTM, and Proposed ARIMA-GRU) over the same weekly intervals (4, 6, and 52 weeks). Consistent with the MAE results, SARIMA registers the highest RMSE, whereas the Proposed ARIMA-GRU demonstrates superior performance by attaining the lowest RMSE values.

Fig. 8(c) focuses on the Mean Absolute Percentage Error (MAPE) for SARIMA, LSTM, and the Proposed ARIMA-GRU models, again with evaluations at 4, 6 and 52 weeks. The chart makes it evident that SARIMA leads to the highest MAPE, while the Proposed ARIMA-GRU delivers the best performance evidenced by the lowest MAPE values.

B. *Discussion*

Coupling the ARIMA model with GRU proposes a sophisticated hybrid forecasting method that enhances the forecasting accuracy of influenza epidemics. Standard ARIMA models have been popular choices for time-series forecasting because of their efficiency in identifying short-term patterns. Their linear nature restricts their capacity to express complex, non-linear dependencies in epidemiological data, especially in dynamically changing outbreaks. Conversely, GRU, a specialized type of RNN, is best suited to analyze oscillating trends of diseases. By combining ARIMA and GRU, this hybrid model is able to take the best out of both techniques—ARIMA improves short-term prediction, while GRU improves the capacity to identify long-term trends and complex dependencies in time-series data. This merge leads to increased prediction accuracy, which enables earlier diagnoses of epidemic outbursts. The model is trained on the WHO Flu Net dataset obtained from Kaggle, with its real-world practicability assured. To build prediction credibility, data preprocessing methods, which involve normalization, scaling, and imputation of missing values, are utilized. The performance of the ARIMA-GRU model is verified through performance measures like MAE, RMSE, and MAPE, which reflect higher accuracy than conventional techniques. The ability to specify well on unseen data renders it a strong tool for real-world epidemic forecasting. By learning on past data and validating on fresh data, it effectively forecasts influenza trends, aiding proactive decision-making in public health. Finally, this hybrid model yields a strong, flexible, and interpretable prediction model, providing public health officials with a useful instrument for resource planning and timely intervention measures. Future work may involve real-time integration of data, including environmental and population mobility data, to enhance predictions further. Some of the demerits are data dependency, complex integration, High computational cost, and limited interpretability, sensitivity to parameter tuning, struggles with real-time external shocks, needs regular retraining and data pre-processing overhead. The ARIMA-GRU model therefore offers a scalable, high-performance solution to epidemic surveillance and control, enabling improved preparedness and response to influenza outbreaks.

VI. CONCLUSION AND FUTURE WORKS

The ARIMA-GRU hybrid model surpasses conventional baseline models by successfully integrating statistical and neural network methods. ARIMA is superior in capturing short-term variations, whereas the GRU model is capable of identifying long-term temporal patterns with high accuracy, resulting in better forecasting results. Tests based on metrics like MAE, RMSE, and MAPE validate the model's higher accuracy in forecasting influenza cases across various forecasting horizons, from 4 to 52 weeks. This model turns out to be extremely useful for public health organizations, yielding actionable information for the early identification of influenza patterns, which can assist in the timely control of epidemics.

To further improve prediction accuracy, subsequent research must involve incorporating new deep learning methodologies like transformers and CNN-based recurrent networks into the current ARIMA-GRU framework. By using real-time data like weather patterns and population mobility patterns would greatly enhance the model's applicability and relevance in ever-changing environments. Scaling the model's applicability to predict infectious diseases in different regions will also be crucial. Additionally, computational efficiency of the ARIMA-GRU model is necessary for its real-time application to respond quickly to new global health issues.

REFERENCES

[1] S. Yang and Y. Bao, “Comprehensive learning particle swarm optimization enabled modeling framework for multi-step-ahead influenza prediction,” Appl. Soft Comput., vol. 113, p. 107994, 2021.

[2] C.-T. Yang et al., “Influenza-like illness prediction using a long short-term memory deep learning model with multiple open data sources,” J. Supercomput., vol. 76, pp. 9303–9329, 2020.

[3] Y.-L. Xia, W. Li, Y. Li, X.-L. Ji, Y.-X. Fu, and S.-Q. Liu, “A deep learning approach for predicting antigenic variation of influenza A H3N2,” Comput. Math. Methods Med., vol. 2021, no. 1, p. 9997669, 2021.

[4] S. O. Olukanmi, F. V. Nelwamondo, and N. I. Nwulu, “Utilizing Google Search Data with deep learning, machine learning and time series modeling to forecast influenza-like illnesses in South Africa,” IEEE Access, vol. 9, pp. 126822–126836, 2021.

[5] A. Kara, “Multi-step influenza outbreak forecasting using deep LSTM network and genetic algorithm,” Expert Syst. Appl., vol. 180, p. 115153, 2021.

[6] H. Ge et al., “How to determine the early warning threshold value of meteorological factors on influenza through big data analysis and machine learning,” Comput. Math. Methods Med., vol. 2020, no. 1, p. 8845459, 2020.

[7] S. Mishra, R. Kumar, S. K. Tiwari, and P. Ranjan, “Machine learning approaches in the diagnosis of infectious diseases: a review,” Bull. Electr. Eng. Inform., vol. 11, no. 6, pp. 3509–3520, 2022.

[8] L. Yang et al., “Enhancing infectious diseases early warning: A deep learning approach for influenza surveillance in China,” Prev. Med. Rep., vol. 43, p. 102761, 2024.

[9] P. Meng, J. Huang, and D. Kong, “[Retracted] Prediction of Incidence Trend of Influenza-Like Illness in Wuhan Based on ARIMA Model,” Comput. Math. Methods Med., vol. 2022, no. 1, p. 6322350, 2022.

[10] R. Kumar, S. Maheshwari, A. Sharma, S. Linda, S. Kumar, and I. Chatterjee, “Ensemble learning-based early detection of influenza disease,” Multimed. Tools Appl., vol. 83, no. 2, pp. 5723–5743, 2024.

[11] S. Kandula and J. Shaman, “Near-term forecasts of influenza-like illness: An evaluation of autoregressive time series approaches,” Epidemics, vol. 27, pp. 41–51, 2019.

[12] M. Ridwan, K. Sadik, and F. M. Afendi, “Comparison of ARIMA and GRU Models for High-Frequency Time Series Forecasting.,” Sci. J. Inform., vol. 10, no. 3, pp. 389–400, 2023.

[13] S. Zhou, C. Song, J. Zhang, W. Chang, W. Hou, and L. Yang, “A hybrid prediction framework for water quality with integrated W-ARIMA-GRU and LightGBM methods,” Water, vol. 14, no. 9, p. 1322, 2022.

[14] C. B. Vennerød, A. Kjærran, and E. S. Bugge, “Long short-term memory RNN,” ArXiv Prepr. ArXiv210506756, 2021.

[15] N. Wu, B. Green, X. Ben, and S. O’Banion, “Deep transformer models for time series forecasting: The influenza prevalence case,” ArXiv Prepr. ArXiv200108317, 2020.

[16] A. A. Tyndall et al., “Quantifying the impact of avian influenza on the northern gannet colony of Bass Rock using ultra-high-resolution drone imagery and deep learning,” Drones, vol. 8, no. 2, p. 40, 2024.

[17] A. Darwish, Y. Rahhal, and A. Jafar, “A comparative study on predicting influenza outbreaks using different feature spaces: application of influenza-like illness data from Early Warning Alert and Response System in Syria,” BMC Res. Notes, vol. 13, no. 1, p. 33, 2020.

[18] M. A. Khan et al., “Forecast the influenza pandemic using machine learning,” Comput. Mater. Contin., vol. 66, no. 1, pp. 331–340, 2020.

[19] S. M. Alzahrani, R. Saadeh, M. A. Abdoon, A. Qazza, F. Guma, and M. Berir, “Numerical simulation of an influenza epidemic: Prediction with fractional SEIR and the ARIMA model,” Appl Math, vol. 18, no. 1, pp. 1–12, 2024.

[20] S. M. Alzahrani and F. E. Guma, “Improving Seasonal Influenza Forecasting Using Time Series Machine Learning Techniques,” J. Inf. Syst. Eng. Manag., vol. 9, no. 4, p. 30195, Sep. 2024, doi: 10.55267/iadt.07.15132.

[21] R. Yin, E. Luusua, J. Dabrowski, Y. Zhang, and C. K. Kwoh, “Tempel: time-series mutation prediction of influenza A viruses via attention-based recurrent neural networks,” Bioinformatics, vol. 36, no. 9, pp. 2697–2704, 2020.

[22] L. Li, Y. Jiang, and B. Huang, “Long-term prediction for temporal propagation of seasonal influenza using Transformer-based model,” J. Biomed. Inform., vol. 122, p. 103894, 2021.

[23] Alexander Lachmann, “Weekly Influenza Reports by Country.” Accessed: Sep. 21, 2024. [Online]. Available: https://www.kaggle.com/datasets/lachmann12/weekly-influenza-reports-by-country

[24] K. Maharana, S. Mondal, and B. Nemade, “A review: Data pre-processing and data augmentation techniques,” Glob. Transit. Proc., vol. 3, no. 1, pp. 91–99, 2022.

[25] A. E. Karrar, “The effect of using data pre-processing by imputations in handling missing values,” Indones. J. Electr. Eng. Inform. IJEEI, vol. 10, no. 2, pp. 375–384, 2022.

[26] S. Mekruksavanich and A. Jitpattanakul, “Effective Detection of Epileptic Seizures through EEG Signals Using Deep Learning Approaches,” Mach. Learn. Knowl. Extr., vol. 5, no. 4, pp. 1937–1952, Dec. 2023, doi: 10.3390/make5040094.

[27] A. Omar and T. Abd El-Hafeez, “Optimizing epileptic seizure recognition performance with feature scaling and dropout layers,” Neural Comput. Appl., vol. 36, no. 6, pp. 2835–2852, Feb. 2024, doi: 10.1007/s00521-023-09204-6.

[28] Y. Jian et al., “ARIMA model for predicting chronic kidney disease and estimating its economic burden in China,” BMC Public Health, vol. 22, no. 1, p. 2456, 2022.

[29] Y.-T. Tsan, D.-Y. Chen, P.-Y. Liu, E. Kristiani, K. L. P. Nguyen, and C.-T. Yang, “The prediction of influenza-like illness and respiratory disease using LSTM and ARIMA,” Int. J. Environ. Res. Public. Health, vol. 19, no. 3, p. 1858, 2022.

[30] G. Li et al., “Forecasting and analyzing influenza activity in Hebei Province, China, using a CNN-LSTM hybrid model,” BMC Public Health, vol. 24, no. 1, p. 2171, 2024.