Forecasting Model for Quantification of Health Commodities in Emergency Response based on Demographic and Morbidity Data

Aung Chan Nyein
College of Informatics and Computing Science , Batangas State University
Alangilan, Batangas
24-05039@g.batstate-u.edu.ph

***Abstract*—The challenge of health commodity forecasting for effective quantification in emergencies is addressed in this paper. The paper introduces a model for forecasting by using the demographic and morbidity data of COVID-19 cases developed by John Hopkins University. Using the Orange Data Mining Tool with integrated Python and an advanced time series model for forecasting such as ARIMA, VAR, LSTM, and Facebook Prophet, the study has identified the trend of the confirmed cases based on historical data of India. It provides the best suitable model to conduct support data-driven decision-making to enhance emergency responses. The results indicated that the LSTM did a great job when it came to predictive accuracy as compared to other models; hence, it gives a strong framework for health logistics in crises. Supply chains for resource optimization integrate predictive analytics.**

***Index Terms*—Key terms related to this study include: ”forecasting algorithm,” ”healthcare commodities,” ”emergency response,” ”COVID-19 dataset,” ”Orange Data Mining,” ”quantification,” ”supply chain”, and ”logistics.”**

I. INTRODUCTION

Background of the Study

Seven viral pandemics, beginning in 1981, have infected the world over the last four decades and subsequently changed perceptions, practices, and policies worldwide. A quality health care service is not possible without a proper pharmaceutical management system with working LMIS and HMIS. Optimum stock status at each level of the supply pipeline of a Nation or Organization is of utmost importance. The study underlines the role of efficient logistics, especially during emergencies. [1] [4] It is hard to maintain an adequate stock level of health commodities even in normal situations without effective quantification and forecasting practices. To implement the quantification processes, gigantic amounts of LMIS and HMIS data need to be processed to produce better outcomes. Emergencies like disasters and pandemics trigger sudden changes in consumption for healthcare commodities that affect the capacity of the supply chain management system. Emergencies disrupt supply chains through lockdowns, travel restrictions, and transport constraints; hence, this affects resource availability. In most cases, such pharmaceutical supply chain management lags behind these delays in the effective conducting of procurement activities toward improved health outcomes.

Therefore, in emergencies, the need remains high to identify powerful algorithms that forecast health commodity requirements during emergencies, enhance resource mobilization, and set up an effective supply chain management system.

This paper discusses in detail a systematic review related to machine learning algorithms and their applications, specifically related to the forecasting of health commodities in emergency situations. The structure of the remainder of this paper is outlined as follows: Section 2 describes the literature review. Section 3 elaborates on the methodologies to be used for the forecasts. Section 4 addresses the results and discussion. Section 5 summarizes the conclusion and recommendations of the work.

Objectives of this Study

This study aims to review and analyze the trends in the morbidity of diseases in an emergency situation with the aim to strengthening the healthcare commodity management system. The objective of the paper is to develop a forecasting model using historical data that would optimize the quantification of health commodities for providing quality healthcare service to the community. This paper intends to achieve the following objectives in the research study:

Algorithm Development: A forecasting algorithm will be developed using the John Hopkins University ”COVID-19 dataset” and ”Orange Data Mining Tool” to achieve better healthcare logistics. [1] [18]

1. Disease Morbidity Pattern Analysis: Scrutinize the pattern of morbidity data emanating during the emergency, namely COVID-19 Positive Patients, to spot a weak link in the supply chain.

2. Model Evaluation: The assessment of the forecasting model for accuracy and reliability in predicting healthcare commodity needs.

3. Actionable Insights: Avail the output of forecasting to provide empirical insights for logistics optimization, decision-making processes, and emergency preparedness.

Problem Statement

The existing healthcare management system often limits the effective and efficient quantification of commodities during emergencies such as COVID-19, Zika Virus MERS, Ebola Outbreaks, etc.,

This is to say that to the unpredictable demand fluctuations because of factors like disease patterns, population vulnerability, emergency readiness, and government policies.

Supply chain constraints from lockdowns, restrictions, and transportation limitations increase the difficulty of providing essential health commodities. The insufficiency of practical, deployable models specifically designed for emergency settings and the prevalence of inaccurate and inconsistent data reporting delay effective forecasting and response efficacy. Consequently, healthcare systems struggle to ensure timely and adequate supplies, potentially impacting patient care and outcomes during emergencies.

II. LITERATURE REVIEW

The study ”Predictive Model for Emergency Medical Services (EMS) using Machine Learning and Data Mining” by Ashish Dudhale, Dr. Akhilesh Mishra, and Dr. Priya Pise focuses on developing a predictive model for EMS demand to address the challenges of unbalanced service supply and demand. Utilizing data from healthcare sectors and various repositories, the researchers processed datasets related to COVID-19 cases in India and globally. They employed machine learning algorithms, including Random Forest, Linear Regression, and Support Vector Machine (SVM), for model development. While all models yielded accurate parameter predictions, SVM emerged as the most effective, despite challenges with Random Forest and Linear Regression. The study underscores the potential of machine learning in healthcare for enhancing EMS demand prediction, calling for further research to refine model accuracy. [9]

“Time Series Facebook Prophet Model and Python for COVID-19 Outbreak Prediction” by Mashael Khayyat, Kaouther Laabidi, Nada Almalki, and Maysoon Al-zahrani. This article was published in *Computer Modeling in Engineering & Sciences* in January 2021. It has been cited 19 times and read 6,707 times on ResearchGate. This study focused on using the Facebook Prophet model to analyze and predict the spread of COVID-19 in Saudi Arabia. The researchers concluded that the model was effective at predicting death cases but had a low ability to predict recovered cases. [5]

The authors, Seye Babatunde, Richard Oloruntoba, and Kingsley Agho, conducted a literature review on logistics models for humanitarian emergencies in Africa, focusing on healthcare commodities from 1990 to 2018. They searched 13 academic databases using keywords like ”humanitarian,” ”disaster,” and ”logistics model,” ultimately finding a scarcity of published articles in this area. The review highlighted the need for further research, particularly in the context of infectious diseases and pandemics, and proposed a new logistics model for policymakers and stakeholders. [6]

The authors, Liang Liu1, Gang Zhu, and Xinjie Zhao reviewed existing literature on inventory models for emergency medical supplies, highlighting factors influencing these models. They examined applications of machine learning and algorithms in medical supply inventory, noting that while some research has been conducted, there is a need for further investigation into testing and developing models for various categories of medical supplies. [10]

Chunyu Wang, Yue Deng, Ziheng Yuan, Chijun Zhang, Fan Zhang, Qing Cai, Chao Gao, and Jurgen Kurths examined the role of face masks in mitigating the spread of COVID-19 and analyzing global initiatives aimed at tracking the virus. They highlight the increased demand for ventilators and personal protective equipment (PPE) during the pandemic and explore the application of simulation models to better understand and optimize supply chains for medical emergency resources. This comprehensive review underscores the critical need for effective resource allocation during public health emergencies. [11]

The scoping review by Kehinde Olawale Ogunyemi, Eniola A Bamgboye, Adeola Fowotade, Fisayo Ogunwemimo, and David Oladimeji Alao emphasizes the critical role of mathematical modeling in epidemiology for enhancing public health emergency preparedness and response in low- and middle-income countries. A notable example discussed is the application of a forecasting model to predict the domestic and international spread of the COVID-19 outbreak, illustrating the utility of such models in managing public health crises. [12]

In their systematic review, Jahani, Jain, and Ivanov (2023) explore the methodologies employed in data science and big data analytics within the realm of supply chain and logistics research. They analyze a diverse range of studies, addressing applications, techniques, challenges, and future directions in the field. The authors highlight a significant limitation in many existing studies, noting that they often lack a structured approach and tend to focus narrowly on big data analytics, neglecting the broader scope of data science. Furthermore, they point out that some studies claiming to utilize big data fail to specify the techniques employed, underscoring the need for clearer definitions and methodologies in future research. [13]

In their 2020 publication in the Revista Científica Multidisciplinar Núcleo do Conhecimento, authors Rayanne Alves De Oliveira, Vinícius Mendes Lima, Lucielma Cavalcante De Jesus França, and Lucrécia Pereira Silva conducted a literature review focusing on hospital pharmacy management with an emphasis on supply chain logistics. They analyzed 25 articles published between 2005 and 2019, sourced from databases including LILACS, SPELL, Scielo, Medline, and the CAPES Journal Portal. The review highlights the critical role of logistics in optimizing supply chains within hospital pharmacies, underscoring the need for effective management practices to enhance pharmaceutical services. [14]

In their 2018 paper published in Supply Chain Forum: An International Journal, authors Blandine Ageron, Smail Benzidia, and Michael Bourlakis provide a comprehensive overview of healthcare logistics and supply chain challenges. They introduce a special issue that includes various articles, highlighting a study by Valérie Belanger, Martin Beaulieu, and Sylvain Landry, which focuses on optimizing the location of medical supplies within nursing units to enhance replenishment system performance. Additionally, the authors reference other significant research addressing pharmaceutical supply chains and the implementation of lean management practices in hospitals, underscoring the evolving landscape of healthcare logistics and the need for innovative solutions. [15]

In their work, Mehmood and Graham (Year) examine the intersection of big data and healthcare logistics, particularly in the context of transport-sharing models. They reference a 2011 report by Audit Scotland, which highlights the importance of efficient transport for health and social care. Additionally, the authors discuss various articles that explore the implications of big data in smart cities and healthcare logistics, including insights from OR Insight on the redesign of healthcare logistics systems. This literature underscores the potential of big data to enhance operational efficiency and improve patient care within healthcare transport frameworks. [16]

Anthony Loritz. ”Prediction of COVID-19 Cases using Machine Learning and Data Analysis” (2022): This capstone project explores the use of machine learning, specifically a long short-term memory (LSTM) model, to predict the number of global COVID-19 cases. The project used COVID-19 data from GitHub repositories and involved training and testing the model on this data. The project also included an analysis of COVID-19 data, including global and country-specific statistics. The report discusses the design of the project, the evaluation metrics used, and key observations made. The report concludes by identifying open issues related to the quality and reporting of COVID-19 data. [17]

III. METHODOLOGY

Case Study Examination: Previous emergency situations, like the COVID-19 outbreak or certain natural disasters, serve as case studies to evaluate how well the model works in real life. Using the model on previous events will enable a comparison of its forecasts with real results, showing its ability to enhance response efforts. Table I provides the scope of this study;

Characteristics	Scope
Focus	Forecasting of Healthcare in Emergency
Goal	Develop & Evaluate Forecasting Model
Perspective	Data-driven Forecasting
Coverage	Algorithm Development & EDA
Organization	Conceptual & Applied Focus
Audience	Researchers, Data Scientists, & Healthcare Professionals

TABLE I
SCOPE OF THE STUDY

The methodology involves the following steps:

Data Collection: The COVID-19 dataset developed by the John Hopkins University has been used as the primary data source. [1] The dataset contains comprehensive information on confirmed cases, deaths, recovered, and other relevant variables from countries around the world. This dataset will be analyzed to identify trends and patterns that can inform resource allocation and healthcare strategies during emergencies. Furthermore, the WHO COVID-19 Essential Supplies Forecasting Tool (COVID-ESFT) will be used as the supplementary calculation for conducting the quantification of health provision and to enhance the accuracy of supply chain management. By integrating these data sources, this research aims to develop a robust framework that can predict future needs based on historical trends and current demand fluctuations. Figure 1 illustrates the process of this approach.

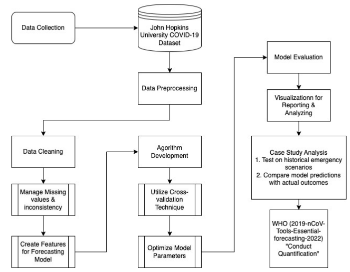

Fig. 1. Process Flow Diagram

Data Pre-processing: The Exploratory Data Analysis (EDA) is performed to ensure consistency and accuracy, involving steps such as cleaning, normalization, and transformation to eliminate any discrepancies or errors that could skew the analysis. This process will enhance the reliability of the findings and provide a solid foundation for subsequent modeling efforts, allowing for more precise forecasting and better decision-making in resource allocation during health emergencies. [2] [3] [18]

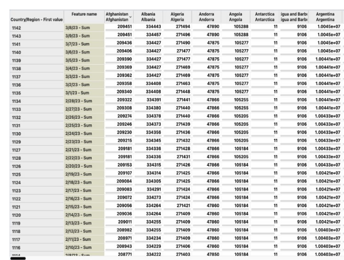

Fig. 2. Overview of the Dataset

The forecasting algorithms were developed using Orange Data Mining and Python. Firstly, the dataset was analyzed by using the geomap and timeslice features in order to visualize and report the significant trends and patterns present in the data. This visualization assists in identifying key variables that influence outcomes, facilitating a more informed approach to model selection and parameter tuning for improved accuracy in predictions. [1] [2] [3] [4] [18]

Fig. 3. COVID-19 Confirmed Cases - World-map

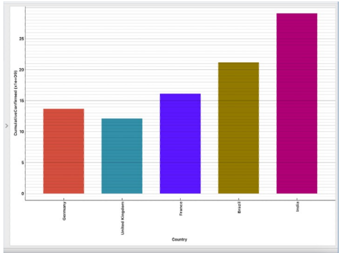

Fig. 4. Top Five Countries with Highest COVID-19 Confirmed Cases

Then, the Human Development Index (HDI) dataset was used to determine the correlation between the COVID-19 confirmed cases and factors such as healthcare infrastructure, demographics, epidemiological data, policy and governance interventions. This comprehensive analysis aims not only to enhance the understanding of the pandemic's impact but also to provide valuable insights for policymakers to devise targeted strategies aimed at mitigating future health crises. Then, the analysis undergoes by using the machine learning models, KNN (K Nearest Neighbor), SVM (Service Vector Machine), Linear Regression, Random Forest, Decision Tree, and Neural Network. Then, the dataset was split into test and train sets to utilize the time series forecasting techniques and measure the prediction performance of the different models, ARIMA, VAR, LSTM, and Prophet with Python integration. [2] [3] [5] [18]

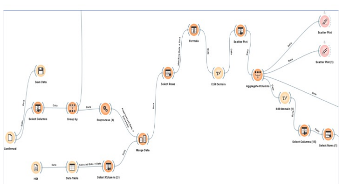

Fig. 5. Merged with HDI Dataset

With the support of the correlation widget and scatter plot, it can be seen that there is a strong correlation between median-aged people (x-axis) and cumulative confirmed cases (y-axis), as shown in Figure 6.

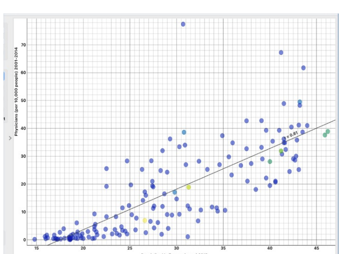

Fig. 6. Relation between Median-aged & Cumulative Confirmed Cases

For machine learning modeling purposes, data pre-processing such as selecting columns, ranking, and preprocess widgets are used, and then the dataset is split into training and testing subsets to enable model validation. A variety of machine learning algorithms, including Decision Trees, Neural Networks, Support Vector Machine (SVM), k-nearest Neighbors (kNN), Random Forest, and Linear Regression, are applied for model training purposes. To perform the model comparison, the "Test and Score" widget was utilized to conduct a comparative analysis of the performance of these algorithms by using metrics such as accuracy, precision, and recall.

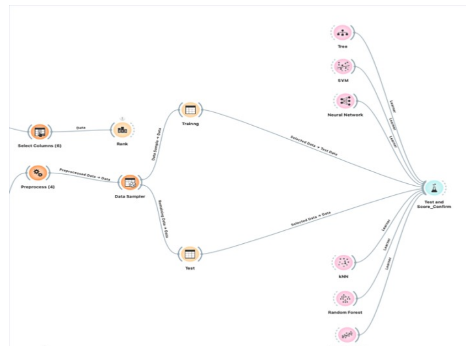

Fig. 7. Machine Learning Models

Time Series Model Training and Validation: The forecasting models were trained on a subset of the data and validated on a separate subset to assess its predictive accuracy. This involves splitting the data into training and testing sets, using cross-validation techniques to ensure model generalizability, and adjusting model parameters to optimize performance.

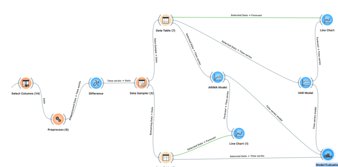

Fig. 8. Time Series Models

Visualization: Data trends, model results, and forecasting outcomes were displayed with charts, graphs, and maps. This helped to analyze and present the findings, providing important insights and patterns related to emergency preparedness and response. Orange Data Mining with Python integration offers various visualization tools for this purpose.

ARIMA models, autoregression (AR), integrated (I) with moving average (MA), are widely used in various fields including finance, economics, health, and engineering. These models can be used to forecast the future aspects of time series data by analyzing past observations. By leveraging historical data, ARIMA models can identify trends and seasonal patterns, enabling more accurate predictions that inform decision-making processes across different sectors. [2]

[3]

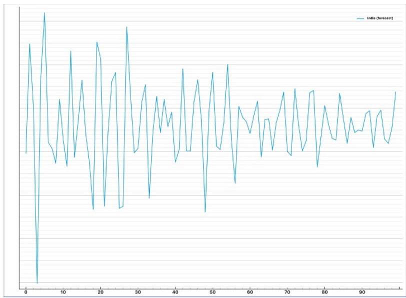

Fig. 9. Forecasting with ARIMA

Vector autoregressive (VAR) models are another powerful tool to model multivariate time series data by understanding the correlations between variables and forecasting the dependencies in order to provide valuable insights to support the decision-making processes. [2] [3]

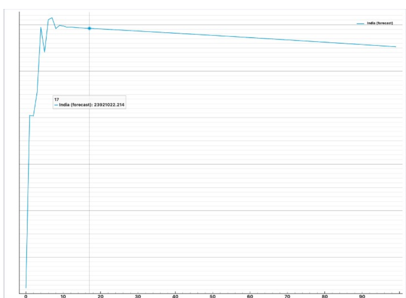

Fig. 10. Forecasting with VAR

Long-short-term memory (LSTM) models are one type of recurrent neural network (RNN) suitable for managing sequential data. They are particularly effective in capturing long-range dependencies, complex patterns, and trends in time series data. [2] [3]

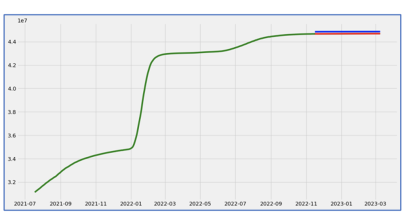

Fig. 11. Forecasting with LSTM

Prophet is a time series forecasting model developed by Facebook to handle time series data with strong seasonal effects and trends. This model allows users to exhibit the missing data, and outliers and make them suitable for real-world applications. The model decomposes time series into three main parts, a trend term, a seasonal term, and a residual term. It also incorporates the effect of holidays to accommodate real-world scenarios. The trend component can model non-linear growth using a piecewise linear or logistic growth model. The seasonal component captures repeating patterns like weekly or yearly seasonality using Fourier series. The residual term represents the remaining unexplained variation in the time series. [5]

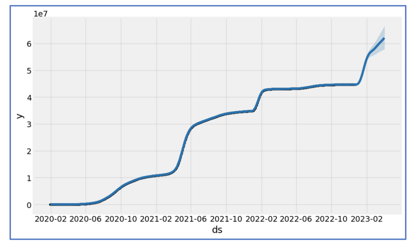

Fig. 12. Forecasting with Prophet (without Holidays)

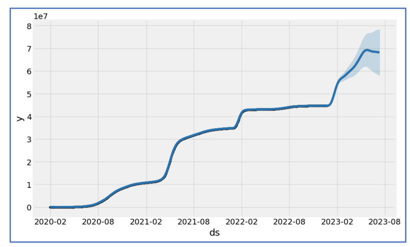

Fig. 13. Forecasting with Prophet (Holidays)

Model Evaluation: The models' effectiveness has been assessed using suitable metrics, including Mean Absolute Error (MAE), Mean Squared Error (MSE), Root Mean Squared Error (RMSE), and R-squared (coefficient of determination). These metrics were used to measure the model's prediction accuracy and error level. [2] [18]

IV. RESULTS AND DISCUSSION

Evaluation Metrics

The four metrics used to assess the accuracy of both machine learning and time series forecasting model regarding the models' performance, accuracy, and error rate: [2] [3] [5] [18]

1) RMSE (Root Mean Squared Error) indicates the error rate by taking the square root of MSE. The mathematical expression RMSE is:

$$RMSE = \sqrt{MSE} = \sqrt{\frac{1}{N}\sum_{t=1}^{N}(y_t - \hat{y})^2} \quad (1)$$

2) MAE (Mean Absolute Error) measures the difference between actual and predicted values by averaging the absolute differences across the dataset. The mathematical expression

for MAE is:

$$MAE = \frac{1}{N}\sum_{t=1}^{N}(y_t - \hat{y}) \tag{2}$$

3) MAPE (Mean Absolute Percentage Error) measures the prediction accuracy of a forecasting method by calculating the average percentage error between the actual and predicted values. It is expressed as a percentage, which makes it easier to interpret and compare across datasets with different scales. The mathematical expression for MAPE is:

$$MAPE = \frac{1}{N}\sum_{t=1}^{N}\left|\frac{y_t - \hat{y}_t}{y_t}\right| \times 100 \tag{3}$$

4) R-squared (Coefficient of Determination) shows how well the values align with the original values, with a range from 0 to 1 interpreted as a percentage. A higher value indicates a better model.

The mathematical expression for R-squared is:

$$R^2 = 1 - \frac{\sum(y_t - \hat{y})^2}{\sum(y_t - \bar{y})^2} \tag{4}$$

Where:

$$y = \text{actual value of } y$$

$$\hat{y} = \text{predicted value of } y$$

$$\bar{y} = \text{mean value of } y$$

Machine Learning Models

Fatality Rate & HDI

Tree-based models (especially Tree and Random Forest) yield the most dependable outcomes in forecasting the link between Fatality Rate and HDI, achieving the least errors and superior R² values. Linear Regression is a valid option, while kNN faces some challenges. More advanced models like Neural Networks and SVM did not perform well, likely due to issues with data limitations or model tuning.

Model	MSE	RMSE	MAE	MAPE	R2
Tree (1)	0.000	0.012	0.008	2.606	0.245
Linear Regression (1)	0.000	0.013	0.008	2.900	0.183
Random Forest (1)	0.000	0.013	0.008	2.091	0.142
kNN (1)	0.000	0.014	0.009	3.303	0.051
Neural Network (2)	0.002	0.046	0.040	9.514	-9.494
SVM (2)	0.007	0.083	0.082	22.791	-33.208

Fig. 14. Fatility & HDI

Recovery Rate & HDI

Linear Regression offers the most trustworthy results in predicting the connection between Recovery Rate and HDI, with the lowest error metrics and the highest R² value (0.274). kNN and SVM show moderate performance, while more complex models like Random Forest, Tree, and Neural Networks do not deliver significant outcomes, probably due to overfitting, underfitting, or data constraints.

Model	MSE	RMSE	MAE	MAPE	R2
Linear Regression (2)	0.002	0.047	0.036	32.104	0.274
kNN (2)	0.003	0.052	0.042	22.410	0.085
SVM (1)	0.003	0.054	0.044	29.600	0.032
Random Forest (2)	0.003	0.057	0.044	32.707	-0.081
Tree (2)	0.004	0.062	0.049	18.165	-0.289
Neural Network (1)	0.006	0.078	0.068	28.672	-1.038

Fig. 15. Recovery Rate & HDI

Cumulative Cases & HDI

The k-Nearest Neighbors (kNN) model excels among the models assessed, with the least errors and the best R² value (0.052). While still modest, kNN slightly identifies the relationship between Cumulative Confirmed Cases and HDI. Other models such as Linear Regression and Random Forest show poor performance, while Tree, SVM, and Neural Network models struggle to fit the data, as indicated by their low or negative R² values.

Model	MSE	RMSE	MAE	MAPE	R2
kNN	13668...	3697137...	16081...	19.479	0.052
Linear Regression	13858...	372267...	171524...	35.305	0.039
Random Forest	14645...	3826911...	16378...	18.243	-0.016
Tree	15396...	392389...	18576...	19.812	-0.068
SVM	15649...	395598...	13692...	11.143	-0.085
Neural Network	16448...	405562...	14243...	1.000	-0.141

Fig. 16. Cumulative Cases & HDI

Time Series Forecasting Models

Table II provides the comparison of performance metrics among ARIMA, VAR, LSTM, and Facebook Prophet Models:

Model	RMSE	MAE	MAPE	R^2
ARIMA (11, 1, 11)	1.72	1.71	0.75	0.022
ARIMA (11,1,11) (in-sample)	1.62	1.61	0.70	0.014
VAR(1, ctt)	2.11	1.89	0.89	-0.47
VAR(1, ctt) (in-sample)	1.24	9.72	0.49	0.42
LSTM (Training)	0.0	0.24	0.00	1.0
LSTM (Test)	0.07	0.05	0.22	0.96
Prophet	9.32E+06	7.82E+06	0.174922	-9.61E+06
Prophet (Holiday)	9.14E+06	7.68E+06	0.171832	-9.25E+06

TABLE II
MODEL PERFORMANCE METRICS

A. ARIMA Model The ARIMA model's performance is modest, with a small R² value near zero, indicating it does not explain much variance in the data. There is only a slight difference between in-sample and test performance, suggesting the model does not overfit. High RMSE and MAE values compared to the LSTM suggest ARIMA may not be the best option for the dataset.

B. VAR Model The negative R² in the test set indicates the VAR model performs worse than a simple mean-based baseline. In-sample metrics are considerably better, with a positive R², but the very high MAE (9.72) suggests possible issues with scaling or errors in predictions. VAR does not generalize well to the test set, indicating it might not be a suitable choice.

C. LSTM Model The LSTM model shows exceptional performance on both the training and test sets. Training RMSE = 0.0 and R^2 = 1.0 suggest a perfect fit to the training data. However, this could indicate the overfitting of the model. Test metrics (RMSE = 0.07, R^2 = 0.96) demonstrate strong generalization with low errors and high predictive power.

D. Prophet Model This prediction model, developed by Facebook, is designed to effectively manage seasonality and holidays within the data. However, the results indicate that the model may not be suitable for the given dataset. The Root Mean Square Error (RMSE) values are notably high, at 9.32E+06 and 9.14E+06, which suggests that the predictions are significantly off from the actual values. Additionally, the R^2 values are alarmingly low, recorded at -9.61E+06 and -9.25E+06. These negative values imply that the model is failing to capture any of the variance present in the data, highlighting its inadequacy in providing reliable forecasts. These findings point to a need for a more effective modeling approach to better understand, and predict the underlying trends.

Overall Observation

LSTM clearly outperforms ARIMA and VAR in terms of all metrics, especially on the test set. The high R^2 and low error metrics suggest that it captures the underlying patterns effectively. ARIMA is relatively weak, with low R^2 and higher errors. While it provides stable results, it does not perform as well as LSTM. VAR and Prophet Models struggled, particularly on the test set. Its negative R^2 and higher RMSE suggest it fails to model the relationships in the dataset effectively. Table III, the comparison between the actual confirmed cases and the predicted number of populations shown. Healthcare professionals and data analysts could review and utilize these results as an inputs into WHO COVID-19 Essential Supplies Forecasting Tool (COVID-ESFT) to conduct the quantification of health provision and to enhance the accuracy of supply chain management.

Date	Actual Cases	LSTM Prediction	Prophet	Prophet with Holiday
28/02/2023	44,687,837	43,913,130	56,625,904	39,516,542
01/03/2023	44,688,105	43,913,196	56,692,286	39,561,388
02/03/2023	44,688,388	43,913,268	56,758,668	39,606,233
03/03/2023	44,688,722	43,913,340	56,825,050	39,651,079
04/03/2023	44,689,046	43,913,416	56,891,433	39,695,924
05/03/2023	44,689,327	43,913,492	56,957,815	39,740,770
06/03/2023	44,689,593	43,913,580	57,024,197	39,785,616
07/03/2023	44,689,919	43,913,668	57,090,579	39,830,461
08/03/2023	44,690,298	43,913,760	57,156,962	39,875,307
09/03/2023	44,690,738	43,913,868	57,223,344	39,920,152

TABLE III
COMPARISON OF ACTUAL CASES, LSTM PREDICTION, PROPHET, AND PROPHET WITH HOLIDAY

V. CONCLUSION

This study aimed to develop a forecasting model to enhance the quantification of healthcare goods for emergency scenarios. The paper examined data science research, concentrating on the forecasting of essential healthcare commodities during emergencies. The research utilized the John Hopkins University COVID-19 dataset along with the Orange Data Mining Tool with Python integration to identify trends and forecast future consumption. The findings showed that the LSTM model greatly surpassed the ARIMA and VAR models in predicting the needs for healthcare commodities. The LSTM model provided outstanding results on both the training and testing datasets. The paper suggests deploying the LSTM model for making forecasts but it also emphasizes the importance of continuous monitoring and updating of the model to adapt to changing circumstances and emerging data patterns in real time to avoid the overfitting of the model that could reduce the accuracy of the forecast. To conclude, leveraging machine learning and predictive analysis to optimize the forecasting models can lead to more efficient resource allocation and strengthen data-driven decision-making in healthcare information systems, especially for emergency situations. The prediction results from time series forecasting models can be used as an input for the WHO COVID-19 Essential Supplies Forecasting Tool (COVID-ESFT) to conduct the quantification and forecasting of health commodities.

REFERENCES

[1] E. Dong, H. Du, and L. Gardner, "An interactive web-based dashboard to track COVID-19 in real time," *Lancet Infect. Dis.*, vol. 20, no. 5, pp. 533–534, May 2020, doi: 10.1016/S1473-3099(20)30120-1.

[2] R. H. Shumway and D. S. Stoffer, "Time Series Analysis and Its Applications".

[3] S. J. Taylor and B. Letham, "Forecasting at Scale," *Am. Stat.*, vol. 72, no. 1, pp. 37–45, Jan. 2018, doi: 10.1080/00031305.2017.1380080.

[4] A. Desale, B. Taye, G. Belay, and A. Nigatu, "Assessment of laboratory logistics management information system practice for HIV/AIDS and tuberculosis laboratory commodities in selected public health facilities in Addis Ababa, Ethiopia," *Pan Afr. Med. J.*, vol. 15, 2013, doi: 10.11604/pamj.2013.15.46.1969.

[5] M. Khayyat, K. Laabidi, N. Almalki, and M. Al-zahrani, "Time Series Facebook Prophet Model and Python for COVID-19 Outbreak Prediction," *Comput. Mater. Contin.*, vol. 67, no. 3, pp. 3781–3793, 2021, doi: 10.32604/cmc.2021.014918.

[6] S. Babatunde, R. Oloruntoba, and K. Agho, "Healthcare commodities for emergencies in Africa: review of logistics models, suggested model and research agenda," *J. Humanit. Logist. Supply Chain Manag.*, vol. 10, no. 3, pp. 371–390, Jul. 2020, doi: 10.1108/JHLSCM-09-2019-0064.

[7] D. Božić, D. Šego, R. Stanković, and M. Šafran, "Logistics in healthcare: a selected review of literature from 2010 to 2022," *Transp. Res. Procedia*, vol. 64, pp. 288–298, 2022, doi: 10.1016/j.trpro.2022.09.033.

[8] N. M. Mkumbwa, G. A. B. Kagashe, E. M. Mlugu, and R. Mwakalukwa, "Factors Affecting the Use of Electronic Logistics Management Information System (eLMIS) Data in Bottom-up Quantification of Health Commodities in Public Health Facilities in Coast Region, Tanzania: A mixed-methods study," Mar. 29, 2023, *In Review*. doi: 10.21203/rs.3.rs-2714700/v1.

[9] A. Dudhale, A. Mishra, and P. Pise, "Predictive Model for Emergency Medical Services (EMS) using Machine Learning and Data Mining," Jan. 23, 2023, *In Review*. doi: 10.21203/rs.3.rs-2457360/v1.

[10] L. Liu, G. Zhu, and X. Zhao, "Application of medical supply inventory model based on deep learning and big data," *Int. J. Syst. Assur. Eng. Manag.*, vol. 13, no. S3, pp. 1216–1227, Dec. 2022, doi: 10.1007/s13198-022-01669-3.

[11] C. Wang *et al.*, "How to Optimize the Supply and Allocation of Medical Emergency Resources During Public Health Emergencies," *Front. Phys.*, vol. 8, p. 383, Oct. 2020, doi: 10.3389/fphy.2020.00383.

[12] K. O. Ogunyemi, E. A. Bamgboye, A. Fowotade, F. Ogunwemimo, and D. O. Alao, "Developing forecasting capacity for public health emergency management in Africa using syndemics approach: lessons from the COVID-19 pandemic," *BMJ Glob. Health*, vol. 7, no. 8, p. e010148, Aug. 2022, doi: 10.1136/bmjgh-2022-010148.

[13] H. Jahani, R. Jain, and D. Ivanov, "Data science and big data analytics: a systematic review of methodologies used in the supply chain and logistics research," *Ann. Oper. Res.*, Jul. 2023, doi: 10.1007/s10479-023-05390-7.

[14] R. A. D. Oliveira, V. M. Lima, L. C. D. J. França, and L. P. Silva, "Hospital pharmacy management: Logistics focus on the supply chain," *Rev. Científica Multidiscip. Núcleo Conhecimento*, pp. 87–98, Sep. 2020, doi: 10.32749/nucleodoconhecimento.com.br/health/hospital-pharmacy.

[15] B. Ageron, S. Benzidia, and M. Bourlakis, "Healthcare logistics and supply chain – issues and future challenges," *Supply Chain Forum Int. J.*, vol. 19, no. 1, pp. 1–3, Jan. 2018, doi: 10.1080/16258312.2018.1433353.

[16] R. Mehmood and G. Graham, "Big Data Logistics: A health-care Transport Capacity Sharing Model," *Procedia Comput. Sci.*, vol. 64, pp. 1107–1114, 2015, doi: 10.1016/j.procs.2015.08.566.

[17] A. Loritz, "Prediction of Covid-19 Cases using Machine Learning and Data Analysis".

[18] J. Demšar, T. Curk, A. Erjavec, Č. Gorup, T. Hočevar, M. Milutinović, M. Možina, M. Polajnar, M. Toplak, A. Starič, M. Štajdohar, N. Umek, Z. Žagar, J. Žbontar, M. Zupan, and B. Zupan, "Orange: Data Mining Toolbox in Python," *Journal of Machine Learning Research*, vol. 14, pp. 2349–2353, 2013. [Online]. Available: https://orangedatamining.com