Reaching Pandemic Milestones with Country Primary and Secondary Vaccination Inflection Points: An Assessment of Foundational and Hybrid Forecasting Methodologies

Marco M. Vlajnic & Steven J. Simske

Colorado State University

ABSTRACT

The devastating worldwide impact of the COVID-19 pandemic created a need to better understand the effects of vaccination on case fatality rates (CFR) in a pandemic setting. Foundational time series forecasting models (ARIMA, Prophet, LSTM) and novel hybrid models (SARIMA-Bidirectional LSTM and SARIMA-Prophet-Bidirectional LSTM) were compared for performance and accuracy to forecast vaccination inflection points for 26 countries. Correlation analyses demonstrated that stringency index, age 65 and older, life expectancy, and positive test rate, are factors correlating the most with the vaccination and case fatality rates. The primary vaccination inflection point was reached at 83.27 days (15-367 days), at the vaccination rate of 13.1% (0.1% - 50%), with 42% of countries seeing the initial impact in <50 days.

The secondary vaccination inflection point (SVIP) was reached at 339.31 days (161-560 days) at the cumulative vaccination rate of 67.8% (28% - 89%), with 23.1% of countries reaching it in < 300 days, 73% in the second half of 2021, and 27% in early 2022.

Index Terms: COVID-19, primary vaccination inflection point, secondary vaccination inflection point, ARIMA, prophet, LSTM, double hybrid, triple hybrid, SARIMA-bidirectional LSTM, SARIMA-prophet-bidirectional LSTM.

Classification: LCC Code: HB1-3840

Language: English

Great Britain Journals Press

LJP Copyright ID: 392941
Print ISSN: 2631-8474
Online ISSN: 2631-8482

London Journal of Engineering Research

Volume 24 | Issue 3 | Compilation 1.0

Reaching Pandemic Milestones with Country Primary and Secondary Vaccination Inflection Points: An Assessment of Foundational and Hybrid Forecasting Methodologies

Marco M. Vlajnic[α] & Steven J. Simske[σ]

ABSTRACT

The devastating worldwide impact of the COVID-19 pandemic created a need to better understand the effects of vaccination on case fatality rates (CFR) in a pandemic setting. Foundational time series forecasting models (ARIMA, Prophet, LSTM) and novel hybrid models (SARIMA-Bidirectional LSTM and SARIMA-Prophet-Bidirectional LSTM) were compared for performance and accuracy to forecast vaccination inflection points for 26 countries. Correlation analyses demonstrated that stringency index, age 65 and older, life expectancy, and positive test rate, are factors correlating the most with the vaccination and case fatality rates. The primary vaccination inflection point was reached at 83.27 days (15-367 days), at the vaccination rate of 13.1% (0.1% - 50%), with 42% of countries seeing the initial impact in <50 days.

The secondary vaccination inflection point (SVIP) was reached at 339.31 days (161-560 days) at the cumulative vaccination rate of 67.8% (28% - 89%), with 23.1% of countries reaching it in < 300 days, 73% in the second half of 2021, and 27% in early 2022. The highest vaccination rate was achieved in Portugal (89%) and the lowest in Bulgaria (28%). All assessed machine and deep learning methodologies performed with high accuracy relative to COVID-19 historical data, demonstrated strong forecasting value, and were validated by anomaly and volatility detection analyses. The novel triple hybrid model performed the best and had the highest accuracy across all performance metrics. Countries prioritizing the health of elderly and frail populations and utilizing AI technology will be better prepared for any future pandemic.

Index Terms: COVID-19, primary vaccination inflection point, secondary vaccination inflection point, ARIMA, prophet, LSTM, double hybrid, triple hybrid, SARIMA-bidirectional LSTM, SARIMA-prophet-bidirectional LSTM.

Author α, σ: Systems Engineering Department, Colorado State University, Engineering Building Suite 202, 400 Isotope Dr, Fort Collins, Colorado, 80523-6029 United States of America.

I. INTRODUCTION

COVID-19 is an infectious disease, caused by the Severe Acute Respiratory Syndrome Coronavirus 2 (SARS-CoV-2) characterized by high morbidity and mortality, and a significant burden on hospital systems and country economies. Over the last three years, the COVID-19 virus infected over 300 million people, caused death for approximately seven million people [1], and had a negative impact of $3.8 trillion on economies around the world. At the end of 2023, COVID-19 is still present with different virus mutations continuing to cause infections and deaths across the world [2, 3, 4].

The experience with the COVID-19 pandemic demonstrated the inadequate levels of preparedness across countries. Two thirds of world countries have a good capacity for public health threat surveillance and analytics in order to drive policy and planning. However, half of the countries have a limited capacity to systematically monitor care, including the impact of vaccination [5]. Both surveillance and monitoring are needed to adequately plan and prepare for possible infectious disease outbreaks with novel viruses [6, 7, 8, 9, 10, 11, 12, 13].

Vaccination is an effective way to obtain individual and herd immunity [14]. When the herd immunity threshold is reached, naturally or through vaccination, it creates an environment that is sufficient to control large outbreaks, reduce the number of infected individuals and possible deaths, protect vulnerable individuals in the society, and relax other public health measures [15]. There are many factors that influence a success of a vaccination campaign, such as availability of vaccine supply (e.g. speed of development, level of demand, difficulties with production and distribution of vaccines worldwide), vaccination strategy outlining priority groups for vaccination, and population acceptance of vaccination (e.g. anti-vaccination movement).

The vaccination efforts for COVID-19 started in December of 2020 for most of the countries in the world. There were several types of vaccines that were available: genetically engineered messenger RNA, viral vector, and protein subunit vaccines. The initial vaccinations from 2020 were followed with booster doses in 2021, 2022 and 2023, for a total of four booster doses, specifically in developed countries [16]. Understanding the impact of vaccination campaigns, the correlation of vaccination rates, incidence of COVID-19, and mortality rates was researched over the last two years. The results confirmed that successful vaccination efforts (e.g. availability of vaccines, public acceptance, strong government programs, etc.) can significantly reduce the negative effects of the COVID-19 pandemic, with a sharp decrease in the fatality rate [17, 18, 19]. Some researchers were able to define the vaccination threshold, identifying that a mean level of administering about 80 doses of vaccines per 100 inhabitants can sustain a reduction of confirmed cases and number of deaths [11], or when the mean cumulative vaccination rate reaches 29.06 doses per 100 people and 7.88 doses per 100 people, respectively, for spread and mortality [19]. Many researchers also looked at the sentiment around vaccination. Attitudes toward COVID-19 and vaccination, conspiracy beliefs, misconceptions, and complaints about COVID-19 control, were documented as dominant sentiments [21, 22, 23]. Researchers used data from different sources (local, national, and global registries) and different time frames (e.g., periods of 3 or 6 months post initial vaccination).

Diverse research methodologies were applied to increase sensitivity of analyses and achieve more accurate results, such as neural networks with cut effect [17], Augmented Artificial Neural Network Model for the COVID-19 Mortality Prediction relative to the vaccination rates [24]; Deep Learning Sequence Models for Forecasting COVID-19 Spread and Vaccinations with two recurrent neural network-based approaches, LSTM and GRU [25]; amalgamation of neural network with two powerful optimization algorithms, firefly algorithm and artificial bee colony based feed-forward neural networks to look at the effect of vaccinated population on the COVID-19 prediction [26]; and a multi-path long short term memory (LSTM) neural network for COVID-19 forecasting of new viral variants and vaccination [27]. Other researchers explored other models, structured and unstructured machine learning (ML) models [22], structural topic modeling [23], Latent Dirichlet Allocation (LDA) [28], deep learning and NLP [29, 30]. Cheng applied newly developed ARIMA models to improve the accuracy of weekly COVID-19 case growth rates and forecast COVID-19 spread according to protective behavior and vaccination [31]. Dhamodharavadhani and colleagues used hybrid models to forecast the vaccination rate, such as HARIMA, a hybrid of ARIMA and HGRNN, a hybrid of Generalized Regression Neural Network and the Gaussian Process Regression model [32]. Yi-Tui Chen and colleagues explored the effect of vaccination patterns and vaccination rates on the spread and mortality of the COVID-19 pandemic [19], and Kumar utilized the recurrent neural network (RNN) Convolutional Residual Network (RNNCON-Res) [33].

Nicholson and colleagues used both supervised and unsupervised methodologies to identify the critical county-level factors for studying COVID-

19 propagation prior to the widespread availability of a vaccine [40].

Published research has increased collective knowledge and has answered many questions. With limitations of every research, availability of more data and novel methodologies, there is a need and a responsibility to continue to expand the knowledge around pandemic vulnerability that can allow for better understanding of the dynamics of vaccination, infection rates and mortality.

This research was conducted to identify the vaccination inflection points and the time needed to reach the critical cumulative vaccination rate thresholds to observe continuous decrease of the case fatality rates. It was conducted both at an aggregate and at the country level. COVID-19 historical data was utilized to develop models that can be used for future pandemics. Applying advanced AI methodologies to forecast time to country specific vaccination inflection points, and assessing the vaccination rates relative to the case fatality rates, can provide another useful tool to guide countries in their pandemic risk preparedness.

II. MATERIALS AND METHODS

2.1 Data

This research utilized data from the Oxford University Our World in Data Covid 19 Dataset. This dataset contains data points collected on an ongoing basis from Johns Hopkins University, Center for Systems Science and Engineering COVID-19 data, European Centre for Disease Control, and OXFORD COVID-19 Government Response Tracker, from January 2020 to the present. The original dataset contains data from 207 countries and territories from which 26 countries were selected for this research: United States, Canada, Italy, Ireland, Finland, Iceland, Denmark, Belgium, Sweden, United Kingdom, Switzerland, Slovenia, Austria, Portugal, France, Netherlands, Luxembourg, Spain, Romania, Latvia, Cyprus, Estonia, Czechia, Slovakia, Serbia, and Bulgaria. Data for this research paper was accessed and downloaded on Dec 30, 2022 [35], and this longitudinal dataset was used from the period of December 2020, when most of the countries in the research dataset started vaccinating their population, to December 30, 2022.

The analyses in this research used 16 variables. Table 1 presents the 14 variables that represent the actual values from the research dataset. Two additional variables, case fatality rate and vaccination rate, were derived. The case fatality rate (CFR), an epidemiologic metric defined as the proportion of deaths within an observed population of interest [34], was calculated by dividing the respective values in the total deaths column by the total cases column of the dataset, for each of the 26 countries. The vaccination rate was calculated by dividing the number of people vaccinated (with at least one dose) by the total population of each country, for each of the 26 countries.

For a more meaningful interpretation of the data variables used to assess the correlation with the vaccination and CFR rates, data variables were organized into novel public health

Table 1: Public Health Indices definitions from the *Our World in Data* metadata file [20]

Population Health Index (PHI)	Pandemic Sensitivity Index (PSI)
cardiovasc death rate: death rate from the cardiovascular disease in 2017 (annual number of deaths per 100,000 people)	*stringency_index:* Government response stringency index: composite measure based on 9 response indicators including school and workplace closures, and travel bans.
diabetes prevalence: Diabetes prevalence (% of population aged 20 to 79) in 2017	*positive_rate:* The share of COVID-19 tests that are positive given as a rolling 7-day average
female smokers: Share of women who smoke, most recent years available	*hosp_ patients:* Number of COVID-19 patients in hospital on a given day
male smokers: Share of men who smoke, most recent years available	*icu_patients:* Number of COVID-19 patients in intensive care unit (ICUs) on a given day
life_expectancy: Life expectancy at birth in 2019	*reproduction_ rate:* Real time estimate of the effective reproduction rate of COVID-19
aged 65 or older: Share of the population that is 65 years or older, most recent years available	*total_cases:* Total confirmed cases of COVID-19
median age: Median age of the population, UN projection for 2020	*total_deaths:* Total deaths attributed to COVID-19

indices, the Population Health Index, PHI [35], and Pandemic Sensitivity Index, PSI (Table 1). The PHI contains the parameters that describe the health of the population such as: cardiovascular death rate, diabetes prevalence, female smokers, male smokers, life expectancy, age 65 and older, and median age. The PSI Index represents variables that are directly impacted by the pandemic, such as total COVID-19 cases and deaths, number of COVID-19 hospital and ICU admissions, Government response stringency index (a composite measure based on nine response indicators including school and workplace closures, and travel bans), reproduction rate of transmission of COVID-19, and positivity rate of COVID-19.

This research was conducted to identify the vaccination inflection points and the time needed to reach the critical cumulative vaccination rate thresholds to observe continuous decrease of the case fatality rates. It was conducted both at an aggregate and at the country level. To accommodate for the peaks and troughs of the case fatality rate curves, the vaccination inflection points were assessed at two different timepoints. The first vaccination inflection time point, primary vaccination inflection point (PVIP) was assessed from the vaccination start date to the date of the first CFR drop post vaccination. The secondary vaccination inflection point (SVIP) was assessed from the vaccination start date to the steepest, most significant CFR decline post vaccination. It represents the time point when the cumulative vaccination rate reached a critical threshold showing a continuous decrease of the case fatality rate, signaling the turnaround in the pandemic. Table 2 provides an overview of descriptions of critical variables used in this research relative to the vaccination inflection point. COVID-19 historical data was utilized to develop models that can be used for future pandemics.

In this research, it was assumed that all vaccines produced by different technologies and manufacturers have the same effectiveness. It was also assumed that distribution of different vaccines in different countries includes a combination of initial two-dose and single-dose vaccines and single dose booster vaccines over the two-year period (Dec 2020-Dec 2022). Since all vaccines require approximately two weeks to produce immunity, the effect of performance of vaccines on CFR was examined two weeks after the start of vaccination.

Several types of vaccines were available at the time of the initial vaccination: genetically engineered messenger RNA Pfizer/BioNTech and Moderna, viral vector vaccines (Janssen/Johnson

& Johnson and University of Oxford/AstraZeneca, Sputnik V), protein subunit vaccine (Novavax, Sinovac). The initial vaccinations in 2020 were delivered, in most cases, in sets of 2-doses, with a 3-week period in between (Pfizer/BioNTech, Moderna, Sinovac, Sputnik V). Some initial vaccines were delivered as a single dose vaccine (J&J, AZ/Oxford). Consequently, booster doses were delivered as single dose vaccines, starting in the third quarter of 2021 (Sep 2021 in the US, Oct/Nov 2021 in the EU) and continuing in 2022 (approved boosters in Mar and Sep 2022 in the US) and 2023 (approved in Sep 2023 in US and EU), for a total of four booster doses [16]. Today there are approximately 40 COVID-19 vaccines that were approved by regulatory agencies for full emergency use authorization. Of those 40, 16 have full authorization in only one country, 12 in ten or fewer countries, and 12 in more than 10 countries [36]. Emergence of new variants may be a challenge for the vaccines, reducing their protective power with the transmissibility of new variants substantially higher than the pre-existing SARS-CoV-2 variants. Booster dose vaccines were introduced to boost the protection power of vaccines and help the individuals with weakened immune systems. Efficacy of most vaccines range from 70-95%, mainly against symptomatic disease [37, 38]. All countries from this dataset (26 countries) are classified in three categories relative to their GDP per capita (>$50,000, $35,000-$50,000, and <$35,000) [20]. Table 3 summarizes the distribution of countries. This research was solely conducted by using publicly available data.

Table 2: Description of derived variables used for vaccination inflection point analyses

Variables	Description
vaccination start date	first documented date when vaccination started at the country level
CFR at vaccination start	Case fatality rate at the time on the 1st day of vaccination
CFR + 14 days	case fatality rate at the time when initial immunity from vaccination should be developed
vaccination rate at CFR +14 days	vaccination rate at the time of initial immunity
Primary vaccination inflection point (PVIP)	date when the first case fatality rate reduction is observed post vaccination, measured on the day of the 1st CFR peak post vaccination + one day
CFR at PVIP	case fatality rate at PVIP, measured on the day of the 1st CFR peak post-vaccination + one day
vaccination rate at PVIP	vaccination rate at the PVIP, measured as the vaccination rate on the day of the 1st CFR peak post vaccination + one day
Secondary vaccination inflection point (SVIP)	date when the most significant CFR reduction is observed post vaccination, measured on the day of the CFR peak that is followed by the most significant and continuous CFR reduction post vaccination + one day
CFR at SVIP	case fatality rate at the SVIP, measured as the CFR rate on the day of CFR peak that is followed by the most significant CFR reduction post vaccination + one day
vaccination rate at SVIP	vaccination rate at the SVIP, measured as the vaccination rate on the day of the CFR peak that is followed by the most significant CFR reduction post vaccination + one day

Table 3: Distribution of countries based on GDP per capita

GDP per Capita	Country Distribution
> 50,000	Ireland, Luxembourg, Switzerland, United States
35,000-50,000	Austria, Belgium, Canada, Denmark, Finland, France, Iceland, Italy, Netherlands, Sweden, United Kingdom.
< 35,000	Bulgaria, Cyprus, Czechia, Estonia, Latvia, Portugal, Romania, Serbia, Slovakia, Slovenia, Spain.

2.2 Methodologies

Data utilized in this research was pre-processed by assigning the original time series dataset to training and testing datasets temporally. For each country, the training set included data from the beginning of the pandemic (March 1, 2020) until a few weeks post vaccination start. The testing set included the remaining data post vaccination until the end of the dataset (December 30, 2022).

Data cleaning was conducted by resolving the problem of missing and duplicate values, resolving data inconsistencies, removing outliers, and smoothing variables used for forecasting (*vaccination_rate* and *case_fatality_rate*), including all exogeneous variables (*stringency_index, aged_65_older, life_expectancy, and positive_rate)*. Smoothing was conducted by using a window of seven days to remove all noisy data. The current day value was calculated using the mean of the previous seven days for each variable. In this type of dataset, it is common that some data is missing, both at random and not at random. For this research, it was important that the data on the total number of cases and deaths was complete since it was used to derive the case fatality rates. This missing data was resolved by taking the mean values of the total number of cases and deaths from the previous day and the next day. Other missing data was managed in a similar manner. Data quality assessments (completeness, reliability, consistency, validity, and no redundancy) were also completed. Exploratory Data Analysis was conducted by exploring graphs and visuals in order to observe trends over time of the vaccination and case fatality rates for each country.

Three foundational forecasting methodologies were applied: Autoregressive Integrated Moving Average (ARIMA), Prophet, and Long-Short Term Memory (LSTM) models. These models were then enhanced and combined to develop novel double and triple hybrids, SARIMA-Bidirectional LSTM and SARIMA-Prophet-Bidirectional LSTM models. They were used to forecast the primary and secondary vaccination inflection points (PVIP and SVIP) relative to the case fatality rates, for each of the 26 countries. All machine learning and deep learning analyses were done using Python version 3.10.1 and the scikit-learn library version 1.2.0 [39]. In addition, the novel Vaccination Inflection Point Score was developed, and countries were classified according to the score.

2.2.1 Correlation Analysis

The correlation analysis was performed using Ordinary Least Squares Multifactor Regression Methodology to identify the top four variables that correlate the most with vaccination and case fatality rates for implementation into forecasting methodologies. These analyses were performed as an aggregate analysis of 14 variables that were assessed for correlation with vaccination and case fatality rates. All variables were used for the correlation assessment with the vaccination rate. Two variables, *total_cases* and *total_deaths* were not used in the assessment of the case fatality correlation since the CFR is a ratio of these two variables. In order to derive the list of the top four variables most correlated with both vaccination and case fatality rates together, the ranking order was assessed across both target variables (vaccination and case fatality rates).

2.2.2 Foundational Forecasting Methodologies

Baseline forecasting methodologies were selected based on literature search, model strengths and limitations.

A. Autoregressive Integrated Moving Average (Arima)

ARIMA (Autoregressive Integrated Moving Average) model is selected for its characteristics of being well-suited for forecasting time series data that exhibits trends and seasonality. It is deemed to be effective in forecasting a variety of real-world phenomena, which has good applicability for COVID-19, showing greater flexibility, accuracy, interpretability, and robustness. The parameters of the ARIMA model are defined as follows: p is the lag order, which represents the number of lag observations incorporated in the model, d is the degree of differencing, which denotes the number of times raw observations undergo differencing, and q is the order of the moving average, which indicates the size of the moving average window [41].

B. Prophet

The Facebook Prophet algorithm is an open-source software developed by Facebook's core Data Science Team. If the time series data has strong seasonal effects, this model works the best. It is a regression model for forecasting, specifically designed to forecast time series data that exhibits trends, seasonality, and coverage for holidays. It is also fast and scalable, and similar to ARIMA, this model is interpretable, robust, flexible, and accurate [42].

C. Long Short-Term Memory (LSTM)

LSTM Model is a neural network model that can learn long-term dependencies in time series data, handle nonstationary and noisy data, as well as leverage additional features. It is also accurate, flexible, and scalable [43, 83].

1) Double Hybrid Forecasting Model: Sarima-Bidirectional LSTM

Review of published literature showcases the use of different forecast models and enhancements in COVID-19 research, demonstrating better accuracy and performance in forecasting by hybrid models. For example, ARIMA-LSTM hybrid model was used to predict future COVID-19 transmissions in China where ARIMA-LSTM model was paralleled by weight of regression coefficient performing better than ARIMA alone [45]; the same group also looked at COVID-19 prediction using data from Germany and Japan and utilized three enhanced hybrid models: PSO-LSTM-ARIMA, MLR-LSTM-ARIMA, and BPNN-LSTM-ARIMA. The research showed that BPNN-LSTM-ARIMA had the best prediction accuracy [46]. Priya and colleagues compared time series forecasting models utilizing ARIMA, Facebook Prophet, Holt-Winters Model, and Hybrid ARIMA-ANN (to take advantage of the unique characteristics of ARIMA and ANN models in linear and nonlinear modeling). The Hybrid model showed better accuracy and root mean square error [47]; Morais looked at forecasting daily Covid-19 cases with a hybrid ARIMA and neural network model to capture the linear and non-linear structures of daily Covid-19 cases (MLP-ARIMA) [48]; and Nawi researched a hybrid ARIMA-SVM model [49]. Borges looked at COVID-19 ICU demand forecasting utilizing Prophet-LSTM approach vs a stand-alone approach in Brazil, confirming better performance of the hybrid model [50], and Long researched an efficient forecasting tool for Monkeypox outbreak in the US using ARIMA, Prophet, Neural Prophet, stacking model, and LSTM models. NeuralProphet achieved the optimal performance [51]. In addition, Guha, in his paper, presented two recurrent neural network-based approaches to predict the daily confirmed COVID-19 cases, daily total positive tests and total individuals vaccinated using LSTM and gated recurrent unit (GRU) [25]; Shastri looked at time series forecasting of Covid-19 using deep learning models: the recurrent neural network based variants of long-short term memory (LSTM) such as Stacked LSTM, Bi-directional LSTM and Convolutional [52]; Devaray utilized ARIMA, LSTM, Stacked LSTM (SLSTM) and Prophet approaches [53]; Zhenyu Li researched convolutional neural network combined with the stacked long-short-term-memory network model (CNN-Stack BiLSTM) [54]. The Stacked LSTM (SLSTM) model was also researched by Maaliw [55] and Ali, who also use the bidirectional enhancement to create a stacked Bi-directional long short-term memory (Stacked Bi-LSTM) network that forecasts COVID-19 more accurately [56]. Sah compared different COVID-19 forecasting models, Prophet, ARIMA, LSTM, and stacked LSTM-GRU models demonstrating better prediction results with the hybrid stacked LSTM-GRU model [57]. Other researchers looked at the Ensemble Empirical Mode Decomposition and Deep Learning creating an EEMD-LSTM hybrid model [58] and EEMD method with the Autoregressive Integrated Moving Average Exogenous inputs (ARIMAX) method, which they called EEMD-ARIMAX [59].

Hybrid models for this research were selected based on the literature search, strengths, and limitations of the individual components for forecasting performance, available enhancements to address limitations, and for their specific complementary characteristics that land them well for hybrid application. SARIMA-Bidirectional LSTM hybrid model combines the strengths of two powerful forecasting techniques, ARIMA enhanced with a seasonality component (the S) in SARIMA and enhancing the LSTM model to analyze data in both directions (Bidirectional component). This hybrid combines a linear and non-linear model, benefits from forecasting time series data that exhibits trends and seasonality and at the same time, an ability to learn long-term dependencies in time series data, as well as capture both forward and backward dependencies. SARIMA-Bidirectional LSTM complements the strength of each model and is expected to achieve better forecasting accuracy than either model individually [44].

2) Triple Hybrid Forecasting Model: Sarima-Prophet-Bidirectional LSTM

The triple hybrid SARIMA-Prophet-Bidirectional LSTM forecasting model enhances the previously mentioned hybrid model with a Facebook Prophet forecasting model that is specifically designed to forecast time series data that exhibits trends, seasonality, and holidays. The new triple hybrid combines the strengths of all three forecasting techniques with an ability to capture short-, medium-, and long-term dependencies, handle non-stationary and noisy data, and leverage additional features. Due to the complementary nature of the hybrid model components and a better fit for the data being researched, it would be expected that the new models would achieve better forecasting accuracy than either model individually.

3) Accuracy and Performance Assessment

Accuracy and performance assessment was conducted across all the models (foundational and hybrid models) evaluating vaccination and case fatality rates: Mean Absolute Error (MAE), Mean Squared Error (MSE), Root Mean Squared Error (RMSE) and Entropy, relative to the actual data. In addition, the accuracy of the forecasting results of each model was compared with actual historical data from the Our World in Data dataset, specifically, to the actual time needed to reach the vaccination inflection points for each country.

4) Anomaly and Volatility Analyses

Anomaly and Volatility analysis and assessments were conducted across all-time series analysis and forecasting models utilizing Isolation Forest and Generalized Autoregressive Conditional Heteroskedasticity (GARCH) models, both well-studied in this field. These methodologies were selected based on the review of published literature that showcase their good performance as well as being valuable algorithms for anomaly and volatility detection in the context of COVID-19 vaccination forecasting [60]. The results obtained upon performing anomaly and volatility detection were used to select the best performing model for forecasting the time to COVID-19 vaccination inflection point for each country.

A. Isolation Forest

The last part of the research was focused on the assessment of anomaly and volatility detection analysis across the time series analysis models. These analyses were conducted to identify unusual or unexpected patterns in data, to prevent overfitting, improve the accuracy, performance, and reliability of machine learning models and complex systems. It is often used in Systems Engineering to detect unusual activity in system logs, performance bottlenecks in systems, and anomalous patterns in system data and to improve overall reliability, efficiency, and security of complex systems. The first algorithm used in this research is Isolation Forest.

Isolation Forest can detect anomalies in an unsupervised manner. This model is used to compare the accuracy of different forecasting models and considered to be efficient, scalable, and robust to outliers. It works by randomly selecting features and splitting values to create partitions of the data. This process is repeated until isolation of the anomalies. It is particularly well-suited for high-dimensional data, which is the case with COVID-19 vaccination data, which includes features such as vaccination rate, case fatality rate, population density, and socio-economic factors. It is also relatively insensitive to outliers, which can be a problem for other anomaly detection algorithms. Isolation Forest can be used to detect anomalies in the vaccination and case fatality rates. This can be useful for identifying periods where the vaccination and CFR rate are significantly higher or lower than expected, adjusting, or improving the forecasts for the vaccination inflection point [61].

Isolation Forest measured three parameters: Precision, Recall, and F1-score. Precision measures the proportion of detected anomalies that are actually true anomalies, where high precision (closer to 1) is very accurate in its anomaly detections, with few false positives. A good threshold for Isolation Forest is 0.7 or higher. Recall measures the proportion of true anomalies that are correctly identified by the model, high recall (closer to 1) means the model is sensitive and can capture most anomalies. A good threshold for Isolation Forest is 0.7 or higher. F1-score combines precision and recall into a single metric, balancing their trade-off. A high F1-score (closer to 1) indicates a good balance between precision and recall, suggesting a reliable anomaly detector. Isolation Forest results at 0.7 or higher for all parameters are considered to be good results [61].

B. Generalized Autoregressive Conditional Heteroskedasticity (Garch)

The second algorithm used to compare the accuracy of different forecasting models is the Generalized Autoregressive Conditional Heteroskedasticity (GARCH) model. The GARCH model is a powerful tool employed to capture and model volatility patterns in the residuals. This model considers the conditional variance and accounts for the time-varying volatility and is especially well suited for time-series analysis, which is the case with COVID-19 vaccination and case fatality rate data.

The GARCH model was used to forecast the volatility of the COVID-19 vaccination and CFR rate. This helped to identify periods where the vaccination and CFR rates are likely to increase or decrease more rapidly than expected. If the model detects anomalies, this could indicate that the vaccination and CFR rates are not following the expected patterns [62].

Isolation Forest and GARCH models are both well-suited for anomaly and volatility detection, respectively, in the context of COVID-19 vaccination inflection point forecasting. They are both efficient, important for anomaly and volatility detection in large datasets, and robust to outliers. This can be a problem in COVID-19 vaccination data due to factors such as data entry errors and reporting delays. These models are also flexible, due to ease of adaptation to a variety of different anomaly detection tasks. The GARCH model also has several limitations, such as sensitivity to the choice of parameters, less robust performance for very short time series datasets, and the inability to capture all types of anomalies.

The GARCH model measures three parameters: Volatility Persistence, Relative Importance of ARCH Term, and Relative Importance of GARCH Term. Volatility Persistence represents the degree to which shocks to volatility persist over time, with an acceptable range between 0.7 and 1. Values below 1 are considered acceptable, ensuring stationarity of the volatility process. However, values closer to or exceeding 1, indicate stronger persistence, meaning shocks have longer-lasting impacts on volatility and might suggest issues like integrated volatility or model misspecification. The range that is typical and acceptable for Relative Importance of ARCH Term is 0 to 0.4. Relative Importance of GARCH Term captures the persistence of volatility shocks over time with an acceptable range of 0.3 to 0.9 [62].

1. Vaccination Inflection Point Score

Vaccination Inflection Point score was developed to categorize countries based on their actual time to achieving secondary vaccination inflection point, representing the time of the most significant CFR reduction post vaccination, and therefore, identifying the critical threshold signaling the turnaround in the pandemic. Countries were categorized into three groups corresponding to scores 1, 2, and 3, with a score of 1 indicating that the country needed the shortest amount of time to reach their secondary vaccination inflection point. This tool can help with the interpretation of changes in the pandemic dynamic, serve as a learning tool for the importance of the contribution of vaccination to achieving faster herd immunity, and improving the overall pandemic risk of countries.

III. RESULTS

3.1 Correlation Analysis Results

The correlation analysis was performed using Ordinary Least Squares Multifactor Regression Methodology. These analyses were performed as aggregate analysis with 14 variables. The correlation was assessed first with the vaccination rate as the target variable, followed by the case fatality rate. The top four variables most correlated with vaccination rate were: *stringency_index, life_expectancy, positive_rate, and total_deaths*. The top four variables for the case fatality rate were: *stringency_index, aged_65_older, life_expectancy, and positive_rate*. The top four variables that are the most correlated with both vaccination and case fatality rates together were derived by using the ranking order of variables across both vaccination and case fatality rates. The final ranking order of the four variables was: *stringency_index, aged_65_older, life_expectancy, and positive_ rate,* representing the exogeneous variables that were used in the primary and secondary vaccination inflection point forecasting analyses. The stringency index and positive rate were variables representing the PSI index and aged 65 and older and life expectancy represented the PHI index.

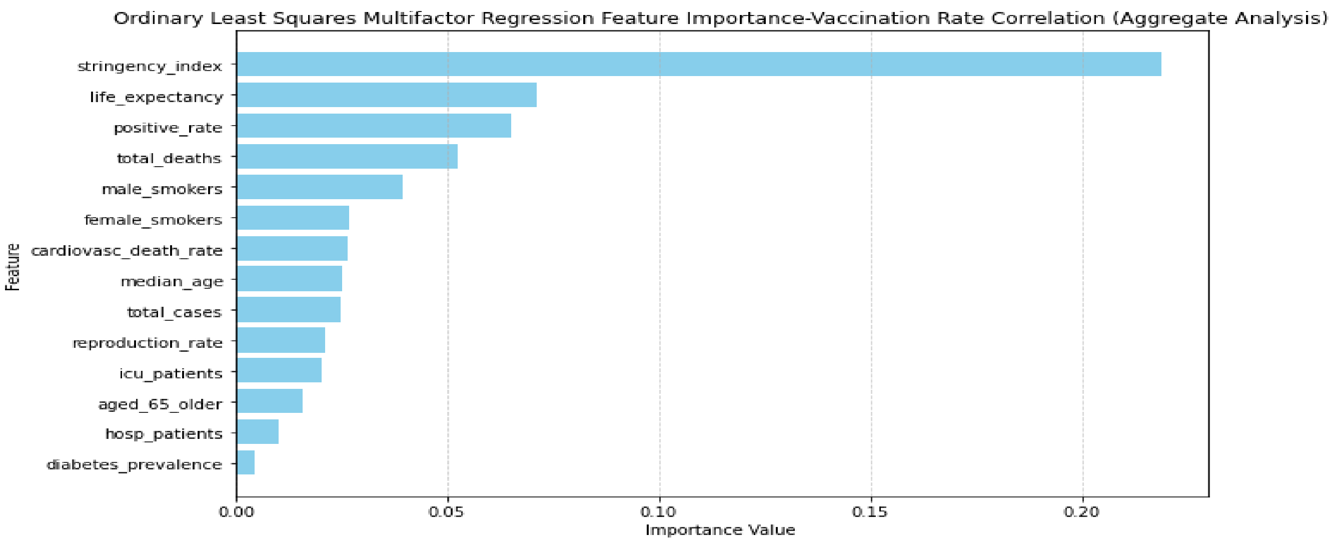

Figure 1: Correlation Analysis for Vaccination Rate (Aggregate Analysis)

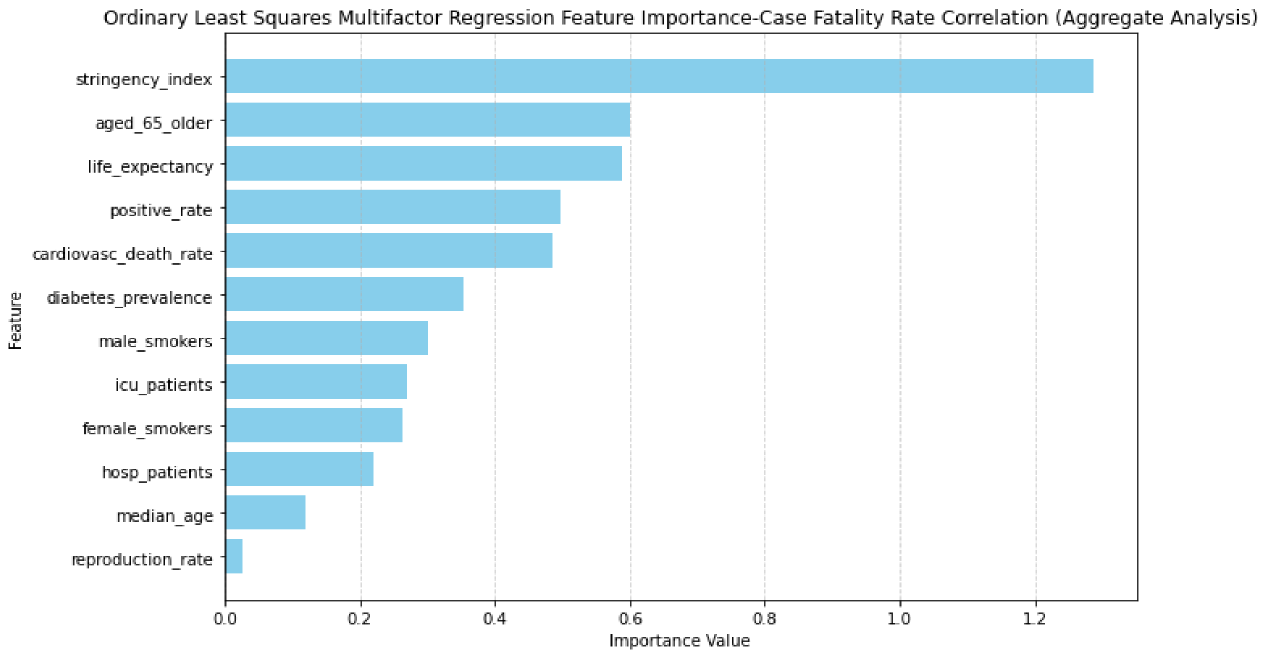

Figure 2: Correlation Analysis for Case Fatality Rate (Aggregate Analysis)

3.2 Forecasting Analysis Results

The summary of the conducted analyses is presented in Table 4, first as an aggregate and then per GDP per capita category (>$50,000, $35,000-$50,000, and <$35,000). Overall, all countries started their vaccination campaigns within the 43 days, starting with Latvia on December 4, 2020, and ending with the UK starting on January 10, 2021. Countries with the higher GDP initiated their vaccination efforts faster than the other countries (15 days vs 26 and 35 days), however, countries with the mid-range GDPs reached the PVIP and SVIP faster than the other two groups, with high and low GDP per capita. PVIP was reached in 37.5 days vs 76 and 131.7 days, and SVIP in 299.2 days vs 336.5 and 380.4 days.

Similar results were observed when median numbers were used, with the mid-range GDP countries again performing better, with the shortest time needed to reach both PVIP (34 days vs 80.5 and 82 days) and SVIP (316 days vs 343.5 and 365 days), and with the highest achieved vaccination rate (74.7% vs 70.6% and 63.3%), for GDP mid-range, high-range, and low-range respectively. Overall, all countries reached an average vaccination rate of 67.8% (mean) and 71.25% (median) at the time they observed the significant CFR drop post-vaccination (SVIP). The highest vaccination rate was achieved in Portugal (89%) and the lowest in Bulgaria (28%).

Analysis of vaccinations by age group in Our World in Data (except for three countries) showed similar distribution by age [20]. The elderly population (60-70, 70-80, and 80+ years of age) achieved the highest vaccination rates in all, but three countries (Latvia, Romania, and Bulgaria), followed by the middle age group (18-24, 25-59). The smallest vaccination rates were observed in the youngest age group (0-17).

The data for the US and UK were not available in the Our World in Data dataset, however, data from official government sites demonstrated the same patterns observed with the rest of the countries [81, 82], supplement Tables S10, S11, and S12. There were no official records available for Serbia at the time of this research. This confirms earlier statements that most countries prioritize elderly and frail populations in their vaccination campaigns. Looking at the countries based on their GDP per capita grouping, the mid-range group on average achieved higher vaccination rates of the elderly population than the countries with higher and lower GDP per capita. These findings support the better performance of the countries in the mid-range

GDP group, demonstrating the importance of prioritizing the needs of the elderly population (age 65 and older and life expectancy) in a pandemic setting. It should be assumed that other factors, such as acceptance and robustness of the vaccination campaign and vaccination mandates imposed by governments played a significant role as well [71].

Table 5 presents the ranking order of the countries based on the time to reach SVIP. The UK observed the SVIP in the shortest amount of time, at 161 days (CFR 3.3%, vaccination rate 63.8%), while Romania reached the same point in 560 days (CFR 2.2%, vaccination rate 41.6%).

Supplemental Tables (Table S1, S2A-B, S3A-B) present all results of all forecasting models, the three foundational (ARIMA, PROPHET, LSTM) and the two hybrid forecasting models (double hybrid: SARIMA-Bidirectional LSTM, and triple hybrid: SARIMA-Prophet-Bidirectional LSTM). The baseline data for each country, as well as the actual historical data from the COVID-19 pandemic are also documented in these supplemental tables.

Table 4: Summary of Results Across 26 Countries

	Aggregate data for 26 countries	Countries with GDP per capita >$50,000	Countries with GDP per capita $35,000-$50,000	Countries with GDP per capita < $35,000
		4 countries (15.4%)	11 countries (42.3%)	11 countries (42.3%)
vaccination start	43 days (Dec 8, 2020 - Jan 10, 2021	15 days (Dec 13 - Dec 28, 2020)	26 days (Dec 8, 2020 - Jan 3, 2021)	35 days (Dec 4, 2020 - Jan 8, 2021)
	mean (range)			
time to reach PVIP*	83.27 days (15-367)	76 days (49-94)	37.5 days (15-75)	131.7 days (16-367)
vaccination rate at PVIP	13.1% (0.1-50)	9.7% (0.9-24.8)	5% (0.4-33.1)	18.6% (1.5-50.1)
time to reach SVIP**	339.31 days (161-560)	336.5 days (296-363)	299.2 days (161-371)	380.4 days (319-560)
vaccination rate at SVIP	67.8% (28-89)	71% (66.4-76.5)	74.2% (63.8-81.8)	60.3% (28-89.1)
	median (range)			
time to reach PVIP	57.5 days (15-367)	80.5 days (49-94)	34 days (15-75)	82 days (16-367)
vaccination rate at PVIP	6.05% (0.1-50)	6.6% (0.9-24.8)	2.4% (0.4-33.1)	9.5% (1.5-50.1)
time to reach SVIP	355.5 days (161-560)	343.5 days (296-363)	316 days (161-371)	365 days (319-560)
vaccination rate at SVIP	71.25% (28-89)	70.6% (66.4-76.5)	74.7% (63.8-81.8)	63.3% (28-89.1)

**PVIP: Primary vaccination inflection point*
***SVIP: Secondary vaccination inflection point*

Table 5: Ranking of the Countries based on the Time to Reach SVIP

Rank	Country	Time (days) to reach SVIP	Vaccination start date	Date SVIP reached	Vaccination rate at SVIP
1	United Kingdom	161 days	Jan 10 2021	Jun 20 2021	63.88%
2	Iceland	201 days	Dec 30 2020	Jul 19 2021	71.64%
3	Denmark	274 days	Dec 8 2020	Sep 8 2021	73.99%
4	Belgium	292 days	Dec 28 2020	Oct 16 2021	74.77%
5	Netherlands	293 days	Jan 8 2021	Oct 28 2021	69.99%
6	Ireland	296 days	Dec 28 2020	Oct 20 2021	76.58%
7	Italy	316 days	Dec 27 2020	Nov 8 2021	79.32%
8	Portugal	319 days	Jan 1 2021	Nov 16 2021	89.10%
9	France	323 days	Dec 27 2020	Nov 15 2021	76.88%
10	Switzerland	329 days	Dec 21 2020	Nov 15 2021	66.42%
11	Finland	333 days	Jan 3 2021	Dec 2 2021	77.16%
12	Spain	337 days	Jan 4 2021	Dec 7 2021	80.84%
13	Cyprus	353 days	Jan 6 2021	Dec 25 2021	71.53%
14	Sweden	358 days	Jan 3 2021	Dec 27 2021	72.39%
15	Luxembourg	358 days	Dec 28 2020	Dec 21 2021	70.99%
16	Serbia	361 days	Jan 8 2021	Jan 4 2022	48.20%
17	Estonia	362 days	Dec 27 2020	Dec 24 2021	63.29%
18	United States	363 days	Dec 13 2020	Dec 11 2021	70.30%

19	Bulgaria	365 days	Dec 29 2020	Dec 29 2021	28.08%
20	Canada	370 days	Dec 14 2020	Dec 19 2021	81.80%
21	Austria	371 days	Dec 27 2020	Jan 2 2022	75.10%
22	Czechia	372 days	Dec 27 2020	Jan 3 2022	65.12%
23	Slovenia	374 days	Dec 27 2020	Jan 5 2022	59.07%
24	Slovakia	386 days	Jan 3 2021	Jan 24 2022	45.73%
25	Latvia	395 days	Dec 4 2020	Jan 3 2022	71.06%
26	Romania	560 days	Dec 27 2020	Jul 10 2022	41.64%

In the dataset used for this research, 65% of countries started their vaccination efforts in December 2020, and 35% started in January 2021. The primary vaccination inflection point representing the first observed reduction in the CFR post vaccination was reached at 83.27 days (mean, range 15-367 days), with 42% of countries seeing the initial impact in less than 50 days, 38.4% in 50-100 days, and 19.2% above 100 days (Figure 3). This reduction was achieved with the initial vaccination rate of 31.1% (mean, range 0.1% to 50%), with 27% of countries reaching the vaccination rate of >25%, 15.3% reaching the rate between 11-25%, and 57.7% reaching the rate of <10% (Figure 4). Finland observed the fastest PVIP in only 15 days (CFR 1.6%, vaccination rate of 1.1%), while Romania had the longest wait to first reduction at 367 days (CFR 3.2%, vaccination rate 27.8%).

The secondary vaccination inflection point (SVIP), representing the most significant reduction in CFR post vaccination, signaling the start of the continuous CFR reduction and turnaround in the pandemic, was reached at 339.31 days (mean, range 161-560 days), with 23.1% of countries observing this impact in less than 300 days, 53.8% from 300-370 days, and 23.1% in more than 370 days (Figure 5). This reduction was achieved with the cumulative vaccination rate of 67.8% (mean, range 8%-89%), with 50% of countries reaching the vaccination rate between 50-75% (Figure 6). Most of the countries reached a significant drop in the CFR in 2021 (73%), out of which 61.5% reached it in the 4th quarter of 2021, 11.5% in the 3rd quarter of 2021, and 27% in early 2022. The highest vaccination rate at this inflection point was achieved in Portugal (89%) on November 16, 2021.

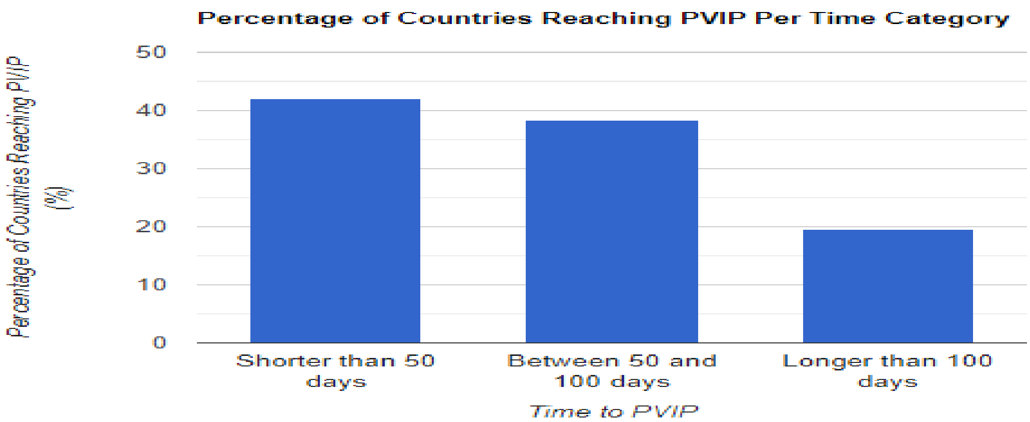

Figure 3: Percentage of Countries Reaching PVIP Per Time Category

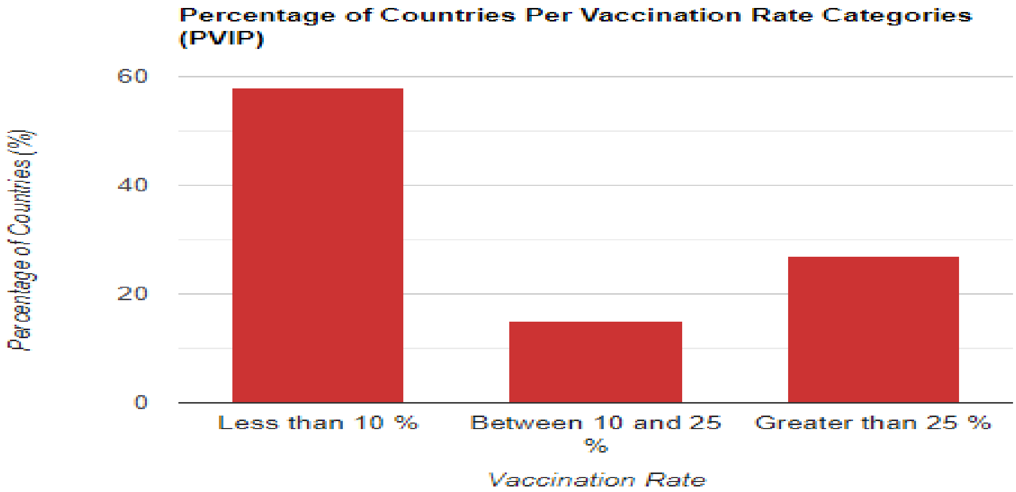

Figure 4: Percentage of Countries Per Vaccination Rate Categories (PVIP)

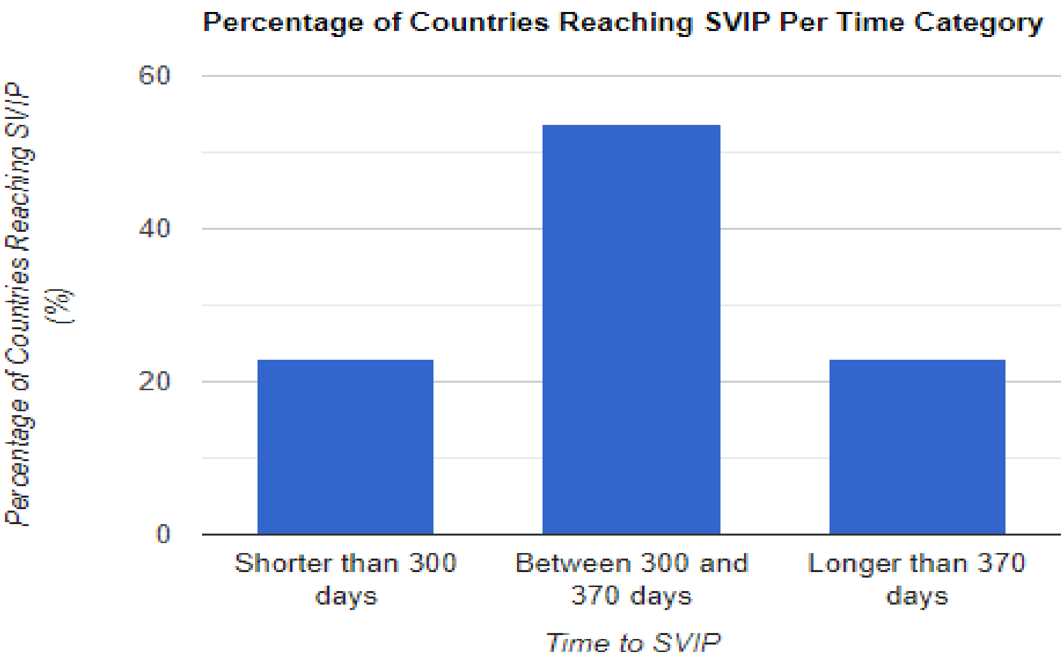

Figure 5: Percentage of Countries Reaching SVIP Per Time Category

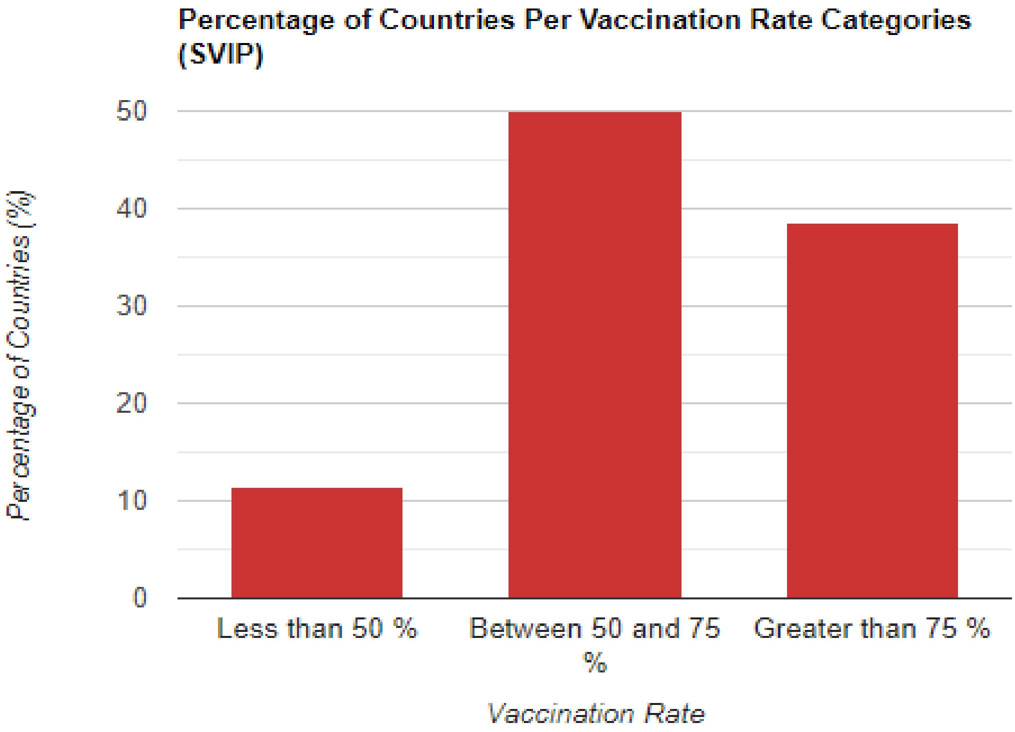

Figure 6: Percentage of Countries Per Vaccination Rate Categories (SVIP)

Overall, at the time of the SVIP, all countries with the exception of three, showed a reduction in the CFRs relative to the CFRs at the beginning of the vaccination. The highest CFR at the time of the SVIP was in Bulgaria (4.1%), followed by the UK (3.32%), and Italy (2.75%). Belgium and Romania had the CFRs that were > 2.0%, and the remaining countries had the CFRs <2.0%. The lowest CFRs were documented in Cyprus (0.42%) and Iceland (0.45%). Bulgaria, Latvia and Slovakia had the CFRs at the SVIP that were higher than the CFR at the vaccination start date, however, all three countries showed a reduction in the CFRs from the PVIP to the SVIP, indicating a positive impact of the vaccination.

3.3 Accuracy and Performance Assessment

Accuracy and performance assessment was conducted across all the models (foundational and hybrid models) evaluating vaccination and case fatality rates: Mean Absolute Error (MAE), Mean Squared Error (MSE), Root Mean Squared Error (RMSE) and Entropy, relative to the actual data. Tables 6-7 showcase the mean and median results for all calculated metrics indicating the superior performance of the triple hybrid model SARIMA-Prophet- Bidirectional LSTM.

3.4 Anomaly and Volatility Analysis Results

Anomaly and Volatility analysis and assessments were conducted across all time-series-analysis and forecasting models utilizing Isolation Forest and Generalized Autoregressive Conditional Heteroskedasticity (GARCH) models. In the Isolation Forest model, precision, recall, and F1-score values above 0.7 indicate good performance. As presented in Tables 8 and 9, both mean and median values were above 0.7, indicating that all forecasting methodologies are performing well and accurately, validating performance of all forecasting models. In the GARCH model, Volatility Performance between 0.7-1, Relative Importance ARCH Term between 0-0.4, and Relative Importance of GARCH Term between 0.3 - 0.9, indicate good performance.

Tables 10 and 11 presented that both mean and median values are within typical and acceptable ranges for all three indicators, suggesting that all forecasting methodologies are performing well and accurately, validating performance of all forecasting models.

Table 6: Vaccination Rate Forecasting Metrics

Metric	Country	ARIMA	Prophet	LSTM	SARIMA-Bidirectional LSTM Double Hybrid	SARIMA-Prophet-Bidirectional LSTM Triple Hybrid
Mean Absolute Error (MAE)	Mean:	0.273440346	0.269496	0.246185113	0.197624417	0.04688918
	Median:	0.2801870395	0.264832	0.190341017	0.091510384	0.008675794
Mean Squared Error (MSE)	Mean:	0.147081	0.185847	0.184565242	0.12681021	0.022887863
	Median:	0.104444003	0.161433	0.076258508	0.0723305005	0.0000841
Root Mean Squared Error (RMSE)	Mean:	0.333789454	0.321616	0.298212928	0.210208282	0.053976834
	Median:	0.327567389	0.322921	0.31088121	0.187344609	0.009161993
Entropy	Mean:	0.197672158	0.172476	0.093079818	0.111602614	0.10330825
	Median:	0.157320015	0.181081	0.02683001	0.0245333975	0.02033074

Table 7: Case Fatality Rate Forecasting Metrics

Metric	Country	ARIMA	Prophet	LSTM	SARIMA-Bidirectional LSTM Double Hybrid	SARIMA-Prophet-Bidirectional LSTM Triple Hybrid
Mean Absolute Error (MAE)	Mean:	0.423921554	0.240282	0.243192654	0.240685426	0.059632661461538
	Median:	0.271977023	0.211509	0.206126957	0.163648243	0.033323647
Mean Squared Error (MSE)	Mean:	0.430524623	0.225626	0.165527034	0.147206818	0.008526044
	Median:	0.106648124	0.213829	0.0766076	0.062206746	0.001274246
Root Mean Squared Error (RMSE)	Mean:	0.500104124	0.275272	0.271744391	0.303519309	0.063672814
	Median:	0.326569872	0.273384	0.2643577545	0.258133644	0.037807365
Entropy	Mean:	0.199711931	0.19532	0.083534021	0.042455902	0.034290933
	Median:	0.217112239	0.193974	0.0327498355	0.009875701	0.0095256035

Table 8: Isolation Forest: Anomaly Detection for Vaccination Rate

Isolation Forest-Anomaly Detection Results (Vaccination Rate Forecasting)			
Country	Precision	Recall	F1 Score
United States	0.957	0.739	0.8007
Austria	0.9117	0.8159	0.9772
Serbia	0.7469	0.9786	0.8216
Canada	0.9969	0.8121	0.7191
Belgium	0.7331	0.9941	0.8332
Bulgaria	0.7079	0.8811	0.7067
Czechia	0.9397	0.8437	0.7551
Denmark	0.9514	0.8081	0.7378
Estonia	0.7033	0.9454	0.9022
Finland	0.8061	0.8942	0.7076
France	0.7569	0.9209	0.7978
Iceland	0.7113	0.7931	0.7926

Ireland	0.8135	0.8411	0.7034
Italy	0.7614	0.8705	0.8592
Latvia	0.7289	0.8352	0.786
Luxembourg	0.7961	0.7753	0.8538
Netherlands	0.7212	0.7771	0.9907
Portugal	0.8517	0.9156	0.8336
Romania	0.9704	0.7877	0.7137
Slovakia	0.7616	0.904	0.9632
Slovenia	0.8737	0.7902	0.7825
Spain	0.8443	0.868	0.9747
Sweden	0.8305	0.7749	0.973
Switzerland	0.9417	0.7178	0.7002
United Kingdom	0.915	0.8303	0.8497
Cyprus	0.8723	0.9244	0.7774
Mean:	0.8309	0.8476	0.8197
Median:	0.822	0.8382	0.7993

Table 9: Isolation Forest: Anomaly Detection for Case Fatality Rate

Isolation Forest-Anomaly Detection Results (Case Fatality Rate Forecasting)			
Country	Precision	Recall	F1 Score
United States	0.8875	0.7503	0.857
Austria	0.926	0.7301	0.8791
Serbia	0.9929	0.8165	0.8089
Canada	0.797	0.9341	0.7126
Belgium	0.9709	0.7439	0.8172
Bulgaria	0.7817	0.9051	0.9914
Czechia	0.9896	0.9788	0.7884
Denmark	0.8408	0.9157	0.7323
Estonia	0.8036	0.8121	0.8191
Finland	0.9038	0.8677	0.8376
France	0.8538	0.9745	0.7094
Iceland	0.9427	0.9422	0.9627
Ireland	0.8745	0.7822	0.769
Italy	0.9391	0.907	0.9095
Latvia	0.9576	0.9067	0.7518
Luxembourg	0.8179	0.7254	0.996
Netherlands	0.747	0.8287	0.8018
Portugal	0.7876	0.918	0.8372
Romania	0.8899	0.752	0.7196
Slovakia	0.9771	0.9905	0.9455
Slovenia	0.921	0.9264	0.7131
Spain	0.7366	0.9903	0.7089
Sweden	0.7791	0.9988	0.8754
Switzerland	0.8953	0.7339	0.7227
United Kingdom	0.7583	0.7996	0.8798
Cyprus	0.9981	0.7523	0.7227
Mean:	0.8757	0.8609	0.818
Median:	0.8887	0.8864	0.8131

Table 10: GARCH: Volatility Detection for Vaccination Rate

GARCH-Volatility Detection Results (Vaccination Rate Forecasting)			
Country	Volatility Persistence	Relative Importance of ARCH Term	Relative Importance of GARCH Term
United States	0.8282	0.2504	0.852
Austria	0.7063	0.1871	0.6404
Serbia	0.8652	0.3208	0.8104
Canada	0.8244	0.026	0.8845
Belgium	0.714	0.0902	0.4021
Bulgaria	0.7451	0.1982	0.3019
Czechia	0.8241	0.1447	0.5649
Denmark	0.7312	0.0737	0.3504
Estonia	0.8505	0.069	0.6422
Finland	0.7691	0.2895	0.4325
France	0.7524	0.1025	0.4491
Iceland	0.8171	0.063	0.63
Ireland	0.7553	0.0197	0.8928
Italy	0.7087	0.2291	0.3192
Latvia	0.7707	0.1126	0.7709
Luxembourg	0.7432	0.1899	0.4111
Netherlands	0.7774	0.0394	0.5235
Portugal	0.75	0.0377	0.7828
Romania	0.8011	0.0934	0.3291
Slovakia	0.7535	0.0961	0.6822
Slovenia	0.7218	0.2805	0.3256
Spain	0.7524	0.3012	0.4183
Sweden	0.7409	0.1956	0.8429
Switzerland	0.7716	0.1138	0.4444
United Kingdom	0.8017	0.0097	0.7839
Cyprus	0.7527	0.3958	0.6299
Mean:	0.7703	0.1511	0.5814
Median:	0.7544	0.1132	0.5974

Table 11: GARCH: Volatility Detection for Case Fatality Rate

GARCH-Volatility Detection Results (Case Fatality Rate Forecasting)			
Country	Volatility Persistence	Relative Importance of ARCH Term	Relative Importance of GARCH Term
United States	0.7459	0.055	0.5546
Austria	0.7973	0.295	0.6177
Serbia	0.7003	0.1989	0.4148
Canada	0.7955	0.0907	0.5704
Belgium	0.837	0.1485	0.5992
Bulgaria	0.7926	0.1119	0.4248
Czechia	0.8525	0.2854	0.815
Denmark	0.8412	0.0142	0.5472
Estonia	0.7639	0.2216	0.5843
Finland	0.7499	0.1335	0.3002
France	0.8511	0.2473	0.8511

Iceland	0.7068	0.2207	0.3123
Ireland	0.7093	0.3246	0.6198
Italy	0.8087	0.0608	0.8999
Latvia	0.7109	0.0828	0.3823
Luxembourg	0.8555	0.3423	0.6214
Netherlands	0.8424	0.1738	0.3842
Portugal	0.809	0.216	0.7016
Romania	0.7274	0.1294	0.3674
Slovakia	0.8668	0.1789	0.4596
Slovenia	0.846	0.3062	0.7968
Spain	0.8044	0.3552	0.3832
Sweden	0.7273	0.3213	0.602
Switzerland	0.7126	0.3562	0.6668
United Kingdom	0.8057	0.3061	0.5669
Cyprus	0.7119	0.356	0.3568
Mean:	0.7835	0.2128	0.5539
Median:	0.7964	0.2184	0.5687

3.5 *Vaccination Inflection Point Score Results*

Vaccination Inflection Point score was developed to categorize countries based on their actual time to achieving secondary vaccination inflection point, representing the time of the most significant CFR reduction post vaccination. Countries were categorized into three groups with scores 1, 2, and 3, with a score of 1 indicating the country needing the shortest amount of time to reach their secondary vaccination inflection point.

Table 12 and Figure 7 present the distribution of countries per VIP score. This data indicates that the majority of countries (53.8%) reached the SVIP between 300-370 days (score 2). While there is a broad range in the achieved vaccination rates across different countries, the median values show numerically higher vaccination rates in the countries with the shortest time to SVIP, score 1 (72.75%), over score 2 (71.75%), and score 3 countries (62%).

Table 12: Distribution of Countries per SVIP Score

SVIP Score	Days to SVIP	Distribution of Countries	% of Countries	Vaccination Rate Range
1	< 300	Denmark, Belgium, Iceland, Ireland, Netherlands, UK (6)	23.10%	63.8-76%. median 72.75%
2	300-370	US, Serbia, Canada, Bulgaria, Estonia, Finland, France, Italy, Luxemburg, Portugal, Spain, Sweden, Switzerland, Cyprus (14)	53.80%	28-89%. median 71.75%
3	>370	Austria, Czechia, Latvia, Romania, Slovakia, Slovenia (6)	23.10%	41.6-75.1%. median 62%

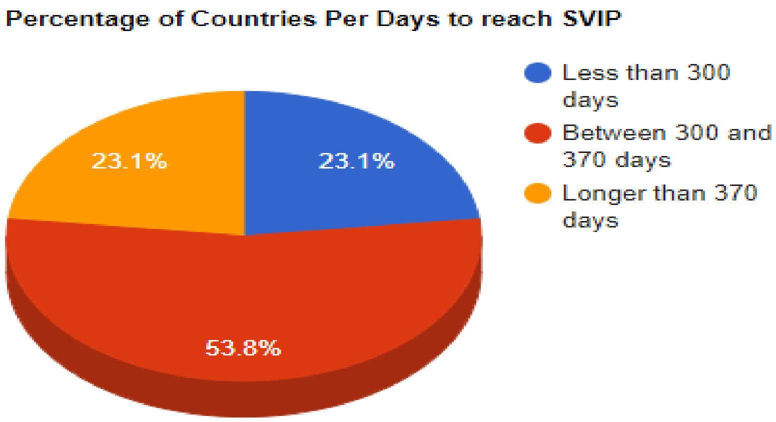

Figure 7: Percentage of Countries per days to reach SVIP

IV. DISCUSSION

The significant negative impact of the COVID-19 pandemic highlighted the need for better preparedness of countries and more sophisticated tools to guide the efforts of public health officials and governments. A wealth of data collected during the COVID-19 pandemic and advanced AI methodologies are allowing researchers to expand our collective knowledge and guide these efforts. This paper builds on the research that highlighted the importance of the non-pandemic predictors in the assessment of pandemic risk [35]. It expands into pandemic predictors and utilizes COVID-19 actual data to refine forecasting tools for future outbreaks. This research was conducted to identify the vaccination inflection points and the time needed to reach the critical cumulative vaccination rate thresholds to observe continuous decrease of the case fatality rates. It was conducted both at the aggregate and country levels, signaling the turnaround point in the pandemic.

The analysis of the actual COVID-19 historical data shows that all of the countries, in aggregate, had the highest fatality rates during the first year of the pandemic. Implementation of the pandemic measures, such as masks, social distancing, school and workplace lockdowns, and testing, had a significant impact on the initial lowering of the case fatality rates. With the introduction of first vaccines in December of 2020, the case fatality rates decreased even further, often reflected as steep downward slopes in graphs. Several types of vaccines were available at the time of the initial vaccination: genetically engineered messenger RNA, viral vector vaccines, and protein subunit vaccine [63]. The initial vaccinations were delivered as single dose or 2-dose vaccines, followed by single dose booster vaccines to improve already established immunity. The first booster dose was approved for use in the third quarter of 2021, followed by two in 2022, and one in 2023, for a total of four booster doses, in developed countries [16]. As of today, there are approximately 40 different vaccines that were approved by regulatory agencies for full emergency use authorization across different countries [36].

In the dataset used for this research, 65% of countries started their vaccination efforts in December 2020, and 35% started in January 2021. On average, looking at the mean values, the time to reach the primary vaccination inflection point, the first reduction in the case fatality rate post vaccination, was on day 83.27 at the vaccination rate of 31%. The secondary vaccination inflection point, representing the most significant and continuous CFR drop post vaccination, was reached at day 339.31 at the average vaccination rate of 67.8%. All four parameters had a very large range, signalizing the presence of outliers. Median values indicate a shorter time to reach the PVIP (57.5 days), lower vaccination rate at PVIP (6.05%), a longer time to reach the SVIP (355.5 days) and a higher overall vaccination rate at SVIP (71.25%). Countries with the mid-level GDP per capita implemented their vaccination campaigns in the most successful way, securing the shortest times to reach both vaccination inflection points looking at both mean and median values. Regarding the individual countries, Finland was the first country to reach the PVIP in only 15 days with the vaccination rate of 1.1%, while Romania had the longest wait, reaching the PVIP in 367 days with a vaccination rate of 27.8%. The UK observed the most significant CFR reduction (SVIP) in the shortest amount of time in 161 days (vaccination rate 63.8%), while Romania reached the same point in 560 days (vaccination rate 41.6%). The highest vaccination rate at SVIP was achieved in Portugal (89%) on November 16, 2021. The SVIP score was developed to categorize countries based on their actual time to achieving secondary vaccination inflection point. The majority of countries reached the SVIP between 300-370 days, while the countries with the lowest score (shorter time to SVIP) had the highest median vaccination rates. This tool can help with the interpretation of changes in the dynamics of the pandemic.

Looking specifically at the US (Supplement Table S1), PVIP was achieved after 94 days post-vaccination, at a vaccination rate of 24.8% and CFR of 1.8%. The most significant reduction in CFR (SVIP) was achieved at 363 days after the vaccination start date or 269 days after the PVIP date. The vaccination threshold at the SVIP was 70.3% with the CFR of 1.59%. The CFR reduced from 1.82 at the time of start of vaccination to 1.59 at the time when it reached the SVIP. In most countries, including the US, priority for COVID-19 vaccination was given to health care workers, residents and personnel of long-term care facilities, elderly patients, and patients with certain comorbidities. The US was grouped with Switzerland, Luxembourg, and Ireland in the high GDP per capita countries (>$50,000). Within this group, the US was the first country to start the vaccination campaign; however, it needed a longer time than other countries to reach both vaccination inflection points. These results were most likely influenced by the impact of widespread anti-vaccine campaigns, scientific misinformation, and overall lack of readiness of certain parts of the population to support government efforts [21, 22, 23].

It is important to recognize that the impact of vaccination is dependent on many factors, such as speed of implementation of the campaign, availability of vaccines, acceptance of the vaccine by the targeted population, and others. The direct impact of vaccination at the population level will often lag and the data may show some initial misalignment that can be explained. For example, the most significant reduction in CFR in the US was observed in December 2021, signaling the turnaround of the pandemic in the US, with the steady decline of the ratio of total infections and death cases. However, in the next few months in 2022 there was a significant increase in new infections and deaths [20]. While the vaccination rate in the US at that time was reaching 70%, it can be assumed that the increase in new cases was caused by several factors, such as the delay in immunity development post vaccination, breakthrough infections, lack of booster vaccination, higher vulnerability of the unvaccinated population, relaxation of pandemic measures and, most significantly, the emergence of new variants with limited immunity coverage from existing vaccines (e.g., omicron variant BA.2.86 that emerged in Nov of 2021).

To understand the findings and applications of this research, it is important to examine the potential variables that may have influenced the results. This research indicates that there are four most important factors that influence the vaccination, and the case fatality rates in any country. Two are non-pandemic variables (not immediately influenced by the pandemic): percentage of people in the population who are 65 years of age or older, and the life expectancy of the population. An additional two variables are pandemic variables: percentage of people who had a confirmed COVID-19 infection with testing, and the level and scope of the pandemic measures that were implemented. The final ranking order of importance of the four variables was: *stringency_index*, *aged_65_older*, *life_expectancy*, and *positive_rate*. It would be expected that the same factors would be the most important in a potential new pandemic as well, due to the increased vulnerability of the elderly and sick patient populations to any infectious disease, and the importance of the infection transmission rates and the speed of implementation of response measures. Simply put, if the vaccines were available at the outbreak of COVID-19 and campaigns were implemented fast in targeted populations, the world would not have a pandemic. This is critical learning highlighting the need to take care of the most vulnerable parts of the populations and implementing appropriate procedures for testing, vaccination, and other public health measures.

This research may have been influenced by the inherited challenges of the vaccination process. Published literature highlights the challenges introduced by the disparity in the distribution of COVID-19 vaccines, where the majority of the vaccines were initially delivered to high – and upper middle-income countries vs lower-income countries [64]. This was evident by the differences in times to first vaccination inflection point, demonstrating that lower-income countries had a higher case fatality rate and needed a longer time to observe the CFR reduction as a result of vaccinations, than the higher-income countries. Lack of availability of sufficient doses of vaccines, less organized execution of vaccine campaigns, including the order of vaccination (elderly and immunocompromised population) may have also influenced the results across countries. In addition, factors affecting vaccination acceptance, confidence in safety and efficacy and the risk of side effects, preference for natural immunity, scientifically sounding misinformation, as well as different cultures and political systems, also played a role in the observed vaccination patterns, spread of infection, and mortality of COVID-19 [65, 66, 67, 68, 69, 70].

All foundational forecasting methodologies utilized in this research (ARIMA, Prophet, and LSTM) showed good accuracy and precision, with only small numerical differences in results, relative to the actual values. They performed well and continue to be a true foundational platform for time series forecasting. Utilization of enhancement features to improve limitations of foundational models is already an established approach, and customizing enhancement based on specificities of data allows for more robust analyses. Combining models into hybrids of foundational or foundational with enhancement models is a newer approach requiring validation. The two hybrid forecasting models (double hybrid: SARIMA-Bidirectional LSTM, and triple hybrid: SARIMA-Prophet- Bidirectional LSTM) utilized in this research are both novel models and their validation was conducted by comparing them to foundational models alone, to each other, and to the actual historical data. They both performed well with high accuracy and precision, and better than the foundational models.

However, the performance and accuracy of the triple hybrid SARIMA- Prophet-Bidirectional LSTM model was superior to other models. In addition, the anomaly and volatility detection analyses, conducted using Isolation Forest and GARCH models, validated performance of all forecasting methodologies, reporting all indicators within the typical and acceptable ranges. In summary, all foundational and hybrid models used for forecasting showed comparable results at the primary and secondary vaccination inflection timepoints and performed with high accuracy relative to the actual data. The best performance was observed with the novel triple hybrid SARIMA-Prophet-Bidirectional LSTM, indicating that hybrid models, combining models with enhanced capabilities, can result in higher accuracy and greater sophistication in analysis. Ability to predict the vaccination inflection point and measure its immediate, as well as the most pronounced impacts, allows for a deeper understanding of the dynamics between the vaccination and case fatality rates. In addition, it is important to remember that the data for this research was trained based on the specificities of the COVID-19 pandemic. For the use of these forecasting models for future pandemics, they may need to be re-trained with the data specific to the new pandemic.

The results of this research can guide countries in the assessment of the pandemic risk and inform public health policy makers in creating measures to minimize the impact of any potential infectious disease pandemic on the people, environment, and socio-economic systems. It is determined that countries can achieve a maximum vaccination rate of 70% with milder measures, and that 90% can be reached only with strict mandates imposed by governments [71]. This highlights the need to plan, organize and execute efficient vaccination campaigns, and improve surveillance and monitoring to substantially reduce morbidity and mortality and avoidance of breakdown of health care systems in countries to control potential new pandemics [72, 73, 74, 75, 76, 77, 78, 79, 80].

V. CONCLUSION AND FUTURE RESEARCH

In conclusion, the research conducted in this paper will add to the knowledge base in the areas of machine and deep learning, and public health. It demonstrates that the novel hybrid time series forecasting models, combining foundational models with enhanced features, provides better performance and higher accuracy over traditional foundational models. The performance and accuracy of the triple hybrid SARIMA-Prophet-Bidirectional LSTM model was superior to other models and was successfully validated with anomaly and volatility detection analyses. In addition, it shows that 42% of countries had seen an immediate effect of vaccination in <50 days, and 23.1% of countries reached the most pronounced impact in <300 days, suggesting the need for improvements. Applying advanced AI methodologies to forecast time to country specific vaccination inflection points, and assessing the vaccination rates relative to the case fatality rates, can provide another useful tool to guide countries in their pandemic risk preparedness.

This paper has several limitations that can be utilized to guide further research, such as: (1) inherited limitations and variabilities of the vaccination campaigns in different countries (supply, distribution, new variants reducing the effectiveness of current vaccines; (2) differences in the health system infrastructures, speed and scope of implementation of other pandemic measures across countries; (3) limitations of the Our Word In Data dataset (e.g., size, completeness, and accuracy, due to the voluntary data reporting and possible underreporting of infection and death cases; and (4) selection of machine and deep learning methodologies and enhancements.

Conflict of Interest

Marco M. Vlajnic and Steven J. Simske declare that there is no conflict of interest.

REFERENCES

1. Muttappallymyalil et al (2022) Vaccination Rate and Incidence of COVID-19 and Case Fatality Rate (CFR): A Correlational Study Using Data From 2019 to 2021. 22 Aug 20. doi: 10.7759/cureus.28210.
2. World Health Organization. (2023). *WHO coronavirus (COVID-19) dashboard*. World Health Organization. Retrieved April 27, 2023, from https://covid19.who.int/.
3. Hospital for Special Surgery. (2021, November 5). *HSS News*. HSS Study Identifies Risk Factor for "Long-Haul" COVID-19 in People with Rheumatic Diseases. Retrieved April 27, 2023, from https://news.hss.edu/hss-study-identifies-risk-factors-for-long-haul-covid-19-in-people-with rheumatic-diseases/.
4. American Hospital Association. (2021). *American Hospital Association homepage: AHA*. American Hospital Association. Retrieved April 27, 2023, from https://www.aha.org/.
5. “Score Data Collection Tool - Who.” *World Health Organization*, World Health Organization, from www.who.int/data/data-collection-tools/score# Accessed 8 Oct. 2023.
6. Magazzino C, Mele M and Schneider N (2020) The relationship between air pollution and COVID-19-related deaths: an application to three French cities. Applied Energy 279, 115835.
7. Huang J et al. (2021) The oscillation-outbreaks characteristic of the COVID-19 pandemic. National Science Review 8. https://doi.org/10.1093/nsr/nwa b100.
8. Coccia M. (2021). Effects of the spread of COVID-19 on public health of polluted cities: results of the first wave for explaining the

dejà vu in the second wave of COVID-19 pandemic and epidemics of future vital agents. *Environmental science and pollution research international*, *28*(15), 19147–19154. https://doi.org/10.1007/s11356-020-11662-7.
9. Coccia, M. Pandemic Prevention: Lessons from COVID-19. *Encyclopedia* 2021, *1*, 433-444. https://doi.org/10.3390/encyclopedia1020036.
10. Coccia M. (2021). The impact of first and second wave of the COVID-19 pandemic in society: comparative analysis to support control measures to cope with negative effects of future infectious diseases. *Environmental research*, *197*, 111099. https://doi.org/10.1016/j.envres.2021.111099.
11. Coccia M. (2022). COVID-19 pandemic over 2020 (withlockdowns) and 2021 (with vaccinations): similar effects for seasonality and environmental factors. *Environmental research*, *208*, 112711. https://doi.org/10.1016/j.envres.2022.112711.
12. Coccia M. (2022). Optimal levels of vaccination to reduce COVID-19 infected individuals and deaths: A global analysis. *Environmental research*, *204*(Pt C), 112314. https://doi.org/10.1016/j.envres.2021.112314.
13. Coccia M. (2022). Preparedness of countries to face COVID-19 pandemic crisis: Strategic positioning and factors supporting effective strategies of prevention of pandemic threats. *Environmental research*, *203*, 1116 78. https://doi.org/10.1016/j.envres.2021.111678.
14. Aldila D et al. (2021) Impact of early detection and vaccinat strategy in COVID-19 eradication program in Jakarta, Indonesia. BMC Research Notes 14, 132.
15. Fontanet and Cauchemez (2020) COVID-19 herd immunity: where are we? Nat Rev Immunol 2020 Oct;20(10):583-584.doi: 10.1038/s41577-020-00451-5.
16. Centers for Disease Control and Prevention. (2023, December 7). *Use of updated COVID-19 vaccines 2023–2024 formula for persons aged ≥6 months: Recommendations of the Advisory Committee on Immunization Practices - United States, September 2023*. Centers for Disease Control and Prevention. https://www.cdc.gov/mmwr/volumes/72/wr/mm7242e1.htm.
17. Magazzino, C., Mele, M., & Coccia, M. (2022). A machine learning algorithm to analyse the effects of vaccination on COVID-19 mortality. *Epidemiology and infection*, *150*, e168. https://doi.org/10.1017/S0950268822001418.
18. R. Jeffrey Melton, Robert C. Sinclair (2021) COVID-19 Infection Rates Are Related to Population Rates of Vaccination: A Response to Subramanian and Kumar. https://www.researchgate.net/publication/355929758.
19. Chen Y. T. (2023). Effect of vaccination patterns and vaccination rates on the spread and mortality of the COVID-19 pandemic. *Health policy and technology*, *12*(1), 100699. https://doi.org/10.1016/j.hlpt.2022.100699.
20. E. Mathieu, H. Ritchie, L. Rodés-Guirao, C. Appel, C. Giattino, J. Hasell, B. Macdonald, S. Dattani, D. Beltekian, E. Ortiz-Ospina, and M. Roser. (2020). *Coronavirus Pandemic (COVID-19)*. [Online]. Available: https://ourworldindata.org/coronavirus.
21. Kwok, S. W. H., Vadde, S. K., & Wang, G. (2021). Tweet Topics and Sentiments Relating to COVID-19 Vaccination Among Australian Twitter Users: Machine Learning Analysis. *Journal of medical Internet research*, *23*(5), e26953. https://doi.org/10.2196/26953.
22. Lincoln, T. M., Schlier, B., Strakeljahn, F., Gaudiano, B. A., So, S. H., Kingston, J., Morris, E. M. J., & Ellett, L. (2022). Taking a machine learning approach to optimize prediction of vaccine hesitancy in high income countries. *Scientific reports*, *12*(1), 2055. https://doi.org/10.1038/s41598-022-05915-3.
23. Liew, T. M., & Lee, C. S. (2021). Examining the Utility of Social Media in COVID-19 Vaccination: Unsupervised Learning of 672,133 Twitter Posts. *JMIR public health and surveillance*, *7*(11), e29789. https://doi.org/10.2196/29789.
24. Kir et al (2022) Augmented Artificial Neural Network Model for the COVID-19 Mortality Prediction: Preliminary Analysis of

Vaccination in Turkey. DOI: 10.35377/saucis.05.01.999373.
25. Guha, S., Kodipalli, A. (2023). Deep Learning Sequence Models for Forecasting COVID-19 Spread and Vaccinations. In: Tistarelli, M., Dubey, S.R., Singh, S.K., Jiang, X. (eds) Computer Vision and Machine Intelligence. Lecture Notes in Networks and Systems, vol 586. Springer, Singapore. https://doi.org/10.1007/978-981-19-7867-8_29.
26. Noroozi-Ghaleini, E., & Shaibani, M. J. (2023). Investigating the effect of vaccinated population on the COVID-19 prediction using FA and ABC-based feed-forward neural networks. *Heliyon*, *9*(2), e13672. https://doi.org/10.1016/j.heliyon.2023.e13672.
27. Rashed, E. A., Kodera, S., & Hirata, A. (2022). COVID-19 forecasting using new viral variants and vaccination effectiveness models. *Computers in biology and medicine*, *149*, 105986. https://doi.org/10.1016/j.compbiomed.2022.105986.
28. Xue, J., Chen, J., Hu, R., Chen, C., Zheng, C., Su, Y., & Zhu, T. (2020). Twitter Discussions and Emotions About the COVID-19 Pandemic: Machine Learning Approach. *Journal of medical Internet research*, *22*(11), e20550. https://doi.org/10.2196/20550.
29. Cresswell, K., Tahir, A., Sheikh, Z., Hussain, Z., Domínguez Hernández, A., Harrison, E., Williams, R., Sheikh, A., & Hussain, A. (2021). Understanding Public Perceptions of COVID-19 Contact Tracing Apps: Artificial Intelligence-Enabled Social Media Analysis. *Journal of medical Internet research*, *23*(5), e26618. https://doi.org/ 10.2196/26618.
30. Hussain, A., Tahir, A., Hussain, Z., Sheikh, Z., Gogate, M., Dashtipour, K., Ali, A., & Sheikh, A. (2021). Artificial Intelligence-Enabled Analysis of Public Attitudes on Facebook and Twitter Toward COVID-19 Vaccines in the United Kingdom and the United States: Observational Study. *Journal of medical Internet research*, *23*(4), e26627. https://doi.org/10.2196/26627
31. Cheng, C., Jiang, W. M., Fan, B., Cheng, Y. C., Hsu, Y. T., Wu, H. Y., Chang, H. H., & Tsou, H. H. (2023). Real-time forecasting of COVID-19 spread according to protective behavior and vaccination: autoregressive integrated moving average models. *BMC public health*, *23*(1), 1500. https://doi.org/10.1186/s12889-023-16419-8.
32. Dhamodharavadhani, S., & Rathipriya, R. (2023). Vaccine rate forecast for COVID-19 in Africa using hybrid forecasting models. *African health sciences*, *23*(1), 93–103. https://doi.org/10.4314/ahs.v23i1.11.
33. Kumar, R., Gupta, M., Agarwal, A., Mukherjee, A., & Islam, S. M. N. (2023). Epidemic efficacy of Covid-19 vaccination against Omicron: An innovative approach using enhanced residual recurrent neural network. *PloS one*, *18*(3), e0280026. https://doi.org/10.1371/journal.pone.0280026.
34. Liu, Longjian. (2018). Biostatistical Basis of Inference in Heart Failure Study. 10.1016/B978-0-323-48558-6.00004-9.
35. M. M. Vlajnic and S. J. Simske, "Accuracy and Performance of Machine Learning Methodologies: Novel Assessments of Country Pandemic Vulnerability Based on Non-Pandemic Predictors," in IEEE Access, vol. 11, pp. 90575-90594, 2023, doi: 10.1109/ACCESS.2023.3307495.
36. World Health Organization. (2023, August 8). *Status of covid-19vaccines within who EUL/PQ evaluation process*. Status of COVID-19 Vaccines within WHO EUL/PQ evaluation process. https://extranet.who.int/prequal/sites/default/files/document_files/Status_COVID_VAX_08August2023.
37. Thomas, S. J., Moreira, E. D., Jr, Kitchin, N., Absalon, J., Gurtman, A., Lockhart, S., Perez, J. L., Pérez Marc, G., Polack, F. P., Zerbini, C., Bailey, R., Swanson, K. A., Xu, X., Roychoudhury, S., Koury, K., Bouguermouh, S., Kalina, W. V., Cooper, D., Frenck, R. W., Jr, Hammitt, L. L., ... C4591001 Clinical Trial Group (2021). Safety and Efficacy of the BNT162b2 mRNA Covid-19 Vaccine through 6 Months. *The New England journal of medicine*, *385*(19), 1761–1773. https://doi.org/10.1056/NEJMoa2110345.
38. Katella, K. (2023, October 5). *Comparing the COVID-19 vaccines: How are they different?*. Yale Medicine. https://www.

yalemedicine.org/news/covid-19-vaccine-comparison.

39. F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, and E. Duchesnay, "Scikit-learn: Machine Learning in Python," *J. Mach. Learn. Res.*, vol. 12, pp. 2825–2830, Nov. 2011.
40. Nicholson, C., Beattie, L., Beattie, M., Razzaghi, T., & Chen, S. (2022). A machine learning and clustering-based approach for county-level COVID-19 analysis. *PloS one*, *17*(4), e0267558. https://doi.org/10.1371/journal.pone.0267558.
41. Rguibi, M. A., Moussa, N., Madani, A., Aaroud, A., & Zine-Dine, K. (2022). Forecasting Covid-19 Transmission with ARIMA and LSTM Techniques in Morocco. *SN computer science*, *3*(2), 133. https://doi.org/10.1007/s42979-022-01019-x
42. Facebook Open Source. (2023). *Prophet: Forecasting at scale*. Prophet. https://facebook.github.io/prophet/.
43. Pyo, J., Pachepsky, Y., Kim, S., Abbas, A., Kim, M., Kwon, Y. S., Ligaray, M., & Cho, K. H. (2023). Long short-term memory models of water quality in inland water environments. *Water research X*, *21*, 100207. https://doi.org/10.1016/j.wroa.2023.100207.
44. G. Li, N. Yang, A Hybrid SARIMA-LSTM Model for Air Temperature Forecasting. *Adv. Theory Simul.* 2023, 6, 2200502. https://doi.org/10.1002/adts.202200502.
45. Jin, Y.; Wang, R.; Zhuang, X.; Wang, K.; Wang, H.; Wang, C.; Wang, X. Prediction of COVID-19 Data Using an ARIMA-LSTM Hybrid Forecast Model. *Mathematics* 2022, *10*, 4001. https://doi.org/10.3390/math10214001.
46. Y. -C. Jin, Q. Cao, K. -N. Wang, Y. Zhou, Y. -P. Cao and X. -Y. Wang, "Prediction of COVID-19 Data Using Improved ARIMA-LSTM Hybrid Forecast Models," in *IEEE Access*, vol. 11, pp. 67956-67967, 2023, doi: 10.1109/ACCESS.2023.3291999.
47. Hema Priya, N., Adithya Harish, S.M., Ravi Subramanian, N., Surendiran, B. (2022). Covid-19: Comparison of Time Series Forecasting Models and Hybrid ARIMA-ANN. In: Rathore, V.S., Sharma, S.C., Tavares, J.M.R., Moreira, C., Surendiran, B. (eds) Rising Threats in Expert Applications and Solutions. Lecture Notes in Networks and Systems, vol 434. Springer, Singapore. https://doi.org/10.1007/978-981-19-1122-4_59.
48. de Araújo Morais, L. R., & da Silva Gomes, G. S. (2022). Forecasting daily Covid-19 cases in the world with a hybrid ARIMA and neural network model. *Applied soft computing*, *126*, 109315. https://doi.org/10.1016/j.asoc.2022.109315.
49. Wan Mohamad Nawi, W. I. A., K Abdul Hamid, A. A., Lola, M. S., Zakaria, S., Aruchunan, E., Gobithaasan, R. U., Zainuddin, N. H., Mustafa, W. A., Abdullah, M. L., Mokhtar, N. A., & Abdullah, M. T. (2023). Developing forecasting model for future pandemic applications based on COVID-19 data 2020-2022. *PloS one*, *18*(5), e0285407. https://doi.org/10.1371/journal.pone.0285407.
50. Borges, D., & Nascimento, M. C. V. (2022). COVID-19 ICU demand forecasting: A two-stage Prophet-LSTM approach. *Applied soft computing*, *125*, 109181. https://doi.org/10.1016/j.asoc.2022.109181.
51. Long, B.; Tan, F.; Newman, M. Forecasting the Monkeypox Outbreak Using ARIMA, Prophet, NeuralProphet, and LSTM Models in the United States. *Forecasting* 2023, *5*, 127-137. https://doi.org/10.3390/forecast5010005.
52. Shastri, S., Singh, K., Kumar, S., Kour, P., & Mansotra, V. (2020). Time series forecasting of Covid-19 using deep learning models: India-USA comparative case study. *Chaos, solitons, and fractals*, *140*, 110227. https://doi.org/10.1016/j.chaos.2020.110227.
53. Devaraj, J., Madurai Elavarasan, R., Pugazhendhi, R., Shafiullah, G. M., Ganesan, S., Jeysree, A. K., Khan, I. A., & Hossain, E. (2021). Forecasting of COVID-19 cases using deep learning models: Is it reliable and practically significant?. *Results in physics*, *21*, 103817. https://doi.org/10.1016/j.rinp.2021.103817.

54. Z. Li, Y. Wang, Y. Wang, Y. Zheng and H. Su, "Covid-19 Epidemic Trend Prediction Based on CNN-StackBiLSTM," *2022 IEEE 11th Data Driven Control and Learning Systems Conference (DDCLS)*, Chengdu, China, 2022, pp. 970-975, doi: 10.1109/DDCLS55054.2022.9858588.
55. R. R. Maaliw, Z. P. Mabunga and F. T. Villa, "Time-Series Forecasting of COVID-19 Cases Using Stacked Long Short-Term Memory Networks," *2021 International Conference on Innovation and Intelligence for Informatics, Computing, and Technologies (3ICT)*, Zallaq, Bahrain, 2021, pp. 435-441, doi: 10.1109/3ICT53449.2021.9581688.
56. Ali, F., Ullah, F., Khan, J. I., Khan, J., Sardar, A. W., & Lee, S. (2023). COVID-19 spread control policies based early dynamics forecasting using deep learning algorithm. *Chaos, solitons, and fractals*, *167*, 112984. https://doi.org/10.1016/j.chaos.2022.112984.
57. Sah, Sweeti & B, Surendiran & Dhanalakshmi, R. & Mohanty, Sachi & Alenezi, Fayadh & Polat, Kemal. (2022). Forecasting COVID-19 Pandemic Using Prophet, ARIMA, and Hybrid Stacked LSTM-GRU Models in India. Computational and Mathematical Methods in Medicine. 2022. 1-19. 10.1155/2022/1556025.
58. Liu, S., Wan, Y., Yang, W., Tan, A., Jian, J., & Lei, X. (2022). A Hybrid Model for Coronavirus Disease 2019 Forecasting Based on Ensemble Empirical Mode Decomposition and Deep Learning. *International journal of environmental research and public health*, *20*(1), 617. https://doi.org/10.3390/ijerph20010617.
59. Da Silva, T. T., Francisquini, R., & Nascimento, M. C. V. (2021). Meteorological and human mobility data on predicting COVID-19 cases by a novel hybrid decomposition method with anomaly detection analysis: A case study in the capitals of Brazil. *Expert systems with applications*, *182*, 115190. https://doi.org/10.1016/j.eswa.2021.115190.
60. Lesouple, Julien & Baudoin, Cédric & Spigai, M. & Tourneret, Jean-Yves. (2021). Generalized Isolation Forest for Anomaly Detection. Pattern Recognition Letters. 149. 10.1016/j.patrec.2021.05.022.
61. H. Xu, G. Pang, Y. Wang and Y. Wang, "Deep Isolation Forest for Anomaly Detection," in IEEE Transactions on Knowledge and Data Engineering, vol. 35, no. 12, pp. 12591-12604, 1 Dec. 2023, doi: 10.1109/TKDE.2023.3270293.
62. Cai G, Wu Z, Peng L. Forecasting volatility with outliers in Realized GARCH models. *Journal of Forecasting*. 2021; 40: 667– 685. https://doi.org/10.1002/for.2736
63. Assistant Secretary for Public Affairs (ASPA). (2023, December 15). *Covid-19 vaccines*. HHS.gov. https://www.hhs.gov/coronavirus/covid-19-vaccines/index.html.
64. OCHA. (2021, April 20). *Unequal vaccine distribution self-defeating, World Health Organization chief tells Economic and Social Council's Special ministerial meeting - world*. ReliefWeb. https://reliefweb.int/report/world/unequal-vaccine-distribution-self-defeating-world-health-organization-chief-tells.
65. Lane, S., MacDonald, N. E., Marti, M., & Dumolard, L. (2018). Vaccine hesitancy around the globe: Analysis of three years of WHO/UNICEF Joint Reporting Form data-2015-2017. *Vaccine*, *36*(26), 3861–3867. https://doi.org/10.1016/j.vaccine.2018.03.063.
66. Larson H. J. (2018). The state of vaccine confidence. *Lancet (London, England)*, *392* (10161), 2244–2246. https://doi.org/10.1016/S0140- 6736(18)32608-4.
67. Logan, J., Nederhoff, D., Koch, B., Griffith, B., Wolfson, J., Awan, F. A., & Basta, N. E. (2018). 'What have you HEARD about the HERD?' Does education about local influenza vaccination coverage and herd immunity affect willingness to vaccinate?. *Vaccine*, *36*(28), 4118–4125. https://doi.org/10.1016/j.vaccine.2018.05.037.
68. Kim, D., Keskinocak, P., Pekgün, P., & Yildirim, İ. (2022). The balancing role of distribution speed against varying efficacy levels of COVID-19 vaccines under variants. *Scientific reports*, *12*(1), 7493. https://doi.org/10.1038/s41598-022-11060-8.

69. Paul, E., Steptoe, A., & Fancourt, D. (2021). Attitudes towards vaccines and intention to vaccinate against COVID-19: Implications for public health communications. *The Lancet regional health. Europe*, *1*, 100012. https://doi.org/10.1016/j.lanepe.2020.100012.
70. Brumfiel, G. (2021, April 7). *Vaccine refusal may put herd immunity at risk, researchers warn*. NPR. https://www.npr.org/sections/health-shots/2021/04/07/984697573/vaccine-refusal-may-put-herd-immunity-at-risk-researchers-warn.
71. Coccia M. (2022). Improving preparedness for next pandemics: Max level of COVID-19 vaccinations without social impositions to design effective health policy and avoid flawed democracies. *Environmental research*, *213*, 113566. https://doi.org/10.1016/j.envres.2022.113566.
72. Akamatsu, T., Nagae, T., Osawa, M., Satsukawa, K., Sakai, T., & Mizutani, D. (2021). Model-based analysis on social acceptability and feasibility of a focused protection strategy against the COVID-19 pandemic. *Scientific reports*, *11*(1), 2003. https://doi.org/10.1038/s41598-021-81630-9
73. Lalmuanawma, S., Hussain, J., & Chhakchhuak, L. (2020). Applications of machine learning and artificial intelligence for Covid-19 (SARS-CoV-2) pandemic: A review. *Chaos, solitons, and fractals*, *139*, 110059. https://doi.org/10.1016/j.chaos.2020.110059.
74. Peiffer-Smadja, N., Rawson, T. M., Ahmad, R., Buchard, A., Georgiou, P., Lescure, F. X., Birgand, G., & Holmes, A. H. (2020). Machine learning for clinical decision support in infectious diseases: a narrative review of current applications. *Clinical microbiology and infection : the official publication of the European Society of Clinical Microbiology and Infectious Diseases*, *26*(5), 584–595. https://doi.org/10.1016/j.cmi.2019.09.009
75. Putrino, A., Raso, M., Magazzino, C., & Galluccio, G. (2020). Coronavirus (COVID-19) in Italy: knowledge, management of patients and clinical experience of Italian dentists during the spread of contagion. *BMC oral health*, *20*(1), 200. https://doi.org/10.1186/s12903-020-01187-3.
76. Vaishya, R., Javaid, M., Khan, I. H., & Haleem, A. (2020). Artificial Intelligence (AI) applications for COVID-19 pandemic. *Diabetes & metabolic syndrome*, *14*(4), 337–339. https://doi.org/10.1016/j.dsx.2020.04.012.
77. Xylogiannopoulos, K. F., Karampelas, P., & Alhajj, R. (2021). COVID-19 pandemic spread against countries' non-pharmaceutical interventions responses: a data-mining driven comparative study. *BMC public health*, *21*(1), 1607. https://doi.org/10.1186/s12889-021-11251-4.
78. Magazzino, C., Mele, M., & Schneider, N. (2021). Assessing a fossil fuels externality with a new neural networks and image optimization algorithm: the case of atmospheric pollutants as confounders to COVID-19 lethality. *Epidemiology and infection*, *150*, e1. https://doi.org/10.1017/S095026882100248X.
79. Romeo, L., & Frontoni, E. (2022). A Unified Hierarchical XGBoost model for classifying priorities for COVID-19 vaccination campaign. *Pattern recognition*, *121*, 108197. https://doi.org/10.1016/j.patcog.2021.108197.
80. Wu, H., Banerjee, R., Venkatachalam, I., Chougale, P. (2022). Impact of Interventional Policies Including Vaccine on COVID-19 Propagation and Socio-economic Factors: Predictive Model Enabling Simulations Using Machine Learning and Big Data. In: Arai, K. (eds) Intelligent Systems and Applications. IntelliSys 2021. Lecture Notes in Networks and Systems, vol 296. Springer, Cham. https://doi.org/10.1007/978-3-030-82199-9_60.
81. *US coronavirus vaccine tracker*. USAFacts. (2024, January 8). https://usafacts.org/visualizations/covid-vaccine-tracker-states/.
82. Office for National Statistics. (2023, March 26). *Coronavirus (COVID-19) latest insights: Vaccines*. Coronavirus (COVID-19) latest insights- Office for National Statistics. https://www.ons.gov.uk /peoplepopulationandcommunity/healthandsocialcare/condition

sanddiseases/articles/coronaviruscovid19latestinsights/vaccines.
83. Hochreiter, S. & Schmidhuber, J"urgen, 1997. Long short-term memory. Neural computation, 9(8), pp.1735–1780. https://doi.org/10.1162/neco.1997.9.8.1735.