

# Explainable Models for Epidemiological Forecasting

Shayar Shah

sshah630@gatech.edu

Georgia Institute of Technology

Atlanta, Georgia, USA

Aryan Pariani

apariani3@gatech.edu

Georgia Institute of Technology

Atlanta, Georgia, USA

Aravind Vengarai

avs9@gatech.edu

Georgia Institute of Technology

Atlanta, Georgia, USA

Conference'17, July 2017, Washington, DC, USA

© 2023 Association for Computing Machinery.

ACM ISBN 978-x-xxxx-xxxx-x/YY/MM...\$15.00

<https://doi.org/10.1145/nnnnnnnn.nnnnnnn>

## ABSTRACT

The increased dependence on complex deep learning and statistical models for epidemiological forecasting, particularly for COVID-19, has led to the development of sophisticated algorithms. However, the opaque nature of these models often obscures the understanding of the factors influencing their predictions. Our project aims to apply Explainable AI techniques to interpret time-series forecasts of COVID-19 case counts by statistical and deep learning models, focusing on identifying the most predictive features from a comprehensive multivariate dataset. This dataset, spanning January 2020 to December 2023 and collected by *Our World in Data*, will be analyzed for several countries. Our objective is to understand how key influencing features vary across different epidemic phases, including abrupt outbreaks, and across various countries. We will employ correlation studies and other metrics to examine feature contributions during diverse epidemic stages and in the most severely affected countries.

## KEYWORDS

Explainable AI, Predictive Feature Extraction, Time-series forecasting, LSTM

### ACM Reference Format:

Shayar Shah, Aryan Pariani, and Aravind Vengarai. 2023. Explainable Models for Epidemiological Forecasting. In *Proceedings of ACM Conference (Conference'17)*. ACM, New York, NY, USA, 12 pages. <https://doi.org/10.1145/nnnnnnnn.nnnnnnn>

## 1 INTRODUCTION

The novel coronavirus's rapid proliferation during the COVID-19 pandemic has significantly influenced statistical and machine learning research, enhancing epidemiological forecasting, such as predicting future case counts in large-scale epidemics. Researchers have applied numerous statistical and deep learning models to forecast accurately and support timely interventions. Yet, the complex nature of these models often hides the rationale behind their predictions, challenging the extraction of key features for early signal detection and targeted policy-making. The interpretability of these models is crucial as they inform governmental and health organization decisions, particularly for preemptive actions during a pandemic's critical phases.

In forecasting COVID-19's spread, factors like government policies, vaccination rates, and healthcare data are pivotal, though their combined effects may vary across nations. Recognizing the

potential of a nuanced approach, we investigate if aggregated features, such as the interplay between vaccination trends and policy strictness, enhance predictive accuracy. This study is driven by the necessity for in-depth feature analysis alongside advanced time-series forecasting. By analyzing a comprehensive dataset from *Our World in Data*, we aim to predict pandemic trends with greater insight through meticulous feature extraction.

## 2 RESPONSE TO MILESTONE COMMENTS

To improve readability, we have increased the font size to ensure all text within the figure is easily legible. In response to the suggestion for a more informative caption, we have revised the caption for Figure 1. The new caption explicitly outlines the key observations from the data. Addressing the suggestion of focus on differences in regions, we have conducted thorough analyses of the different regions and the varying feature significance across each region. These results are discussed in section 6.2, Global Explainability Analysis and section 6.2, Local Explainability Analysis. We have also analyzed the varying feature significance across 3 distinct critical phases of COVID-19 in the USA in section 6.4, Cross-Phase Analysis. Additionally, we have placed a strong emphasis on interpretability as suggested through clear and concise descriptions of our methodologies and findings, utilization of visual aids like charts and graphs to present our results in a clear manner and detailed explanations of the significance of our results in the context of global epidemiological trends

## 3 PROBLEM STATEMENT / CONTRIBUTIONS

Our work seeks to answer: "What are the primary feature configurations driving model predictions in epidemiological forecasting, and to what extent and direction?" Understanding how these critical factors vary across different epidemic phases and countries is vital. The project's main objectives are:

### (1) Multivariate Time-Series Forecasting of daily new COVID-19 case counts:

Train and run different statistical and deep learning models, specifically VAR and LSTM, for *multivariate time-series forecasting of COVID-19 case counts* individually for every one of *four countries - Japan, USA, France, and Italy*. The forecasting models will be based on an extensive multivariate dataset with features like vaccinations, number of ICU patients, and measure of government response stringency, spanning three full years - from *January 3, 2020 until December 31, 2022*.

### (2) Global Explainability Analysis:

Identify and rank the features that are *overall* most predictive of COVID-19 case count forecasts by our models for every country *for the entire 3-year period of the pandemic ("global")*. This analysis

- will be based on the ***Pearson correlation coefficient*** statistical metric, that can shed light on the relative degree and direction of contribution of different features.
- (3) **Local / Phased Explainability Analysis:** Identify and rank the features that are predictive of COVID-19 case count forecasts by our models for every country ***during notable country-specific, phases of the pandemic, spanning short 21-day periods ("local")***. This analysis will be based on ***SHAP (SHapley Additive exPlanations) values***. We pay special attention to the predictive features during the onset of outbreaks in every country, especially the Omicron strain.
- (4) **Cross-Regional + Cross-Phase Comparative Analysis:** Finally, we will compare all above explainability analysis results among the four countries and try to make sense of the cross-country differences. We will also explore and understand the temporal dynamics of the contributions of features and their varying importance during multiple different phases of the pandemic for one country - USA.

By answering these questions, we aim to not only make predictions more understandable but also contribute to enhancing trust and transparency in AI-driven epidemiological forecasts, to facilitate effective and timely government responses for future epidemics.

## 4 LITERATURE REVIEW

To contextualize our work on "Explainable Models for Epidemiological Forecasting," we will review relevant research papers that contribute to the field of epidemiological prediction and provide insights into the challenges and opportunities in this domain.

Solayman et al. [10] present a significant advancement in COVID-19 prediction by deploying machine learning models, including a hybrid CNN-LSTM approach and explainable AI techniques like the LIME framework. This work complements our original research as it addresses the critical need for transparent and interpretable models in understanding and predicting the dynamics of the COVID-19 pandemic. By demonstrating the effectiveness of the CNN-LSTM model in capturing the temporal aspects of symptom progression and applying explainable AI, this study provides valuable insights into feature interpretability—a crucial aspect of epidemiological forecasting.

The work by Devaraj et al. [1] underscores the potential of deep learning models, notably the Stacked Long Short-Term Memory (SLSTM) algorithm, in achieving highly accurate predictions of COVID-19 cases, with an error rate of under 2%. We intend to leverage these findings to enhance the explainability of our epidemiological forecasting models. Moreover, considering various variables, including temperature, rainfall, demonstrates the models' ability to capture complex epidemiological dependencies, and we aim to draw inspiration from these insights by incorporating similar factors and relationships into our models to ensure they provide interpretable, data-driven forecasts for epidemiological scenarios.

The work by Mishra et al. [5] delves into advanced mathematical and deep learning models for short-term COVID-19 incidence and mortality forecasting. Mishra et al. expand the traditional SEIR model by including hospitalization, vaccination, and quarantine

data, a concept pertinent to our goal of enhancing model explainability by considering a broader range of factors in epidemiological scenarios.

Mathonsi and van Zyl [4] introduce the Multivariate Exponential Smoothing Long Short-Term Memory (MES-LSTM) hybrid model, an extension of the ES-RNN model for multivariate time series forecasting. Their main objective is to enhance forecast accuracy and prediction interval construction across various applications, including COVID-19 morbidity datasets. While we emphasize improving model interpretability, Mathonsi and van Zyl's work aligns with the broader trend of enhancing forecasting accuracy, a crucial aspect of epidemiological modeling. The work by Nikparvar et al. [7] employs a deep LSTM neural network for county-level COVID-19 forecasting, addressing challenges in accurate prediction during a dynamic pandemic with spatial variations. In contrast, our focus is on developing explainable epidemiological forecasting models, emphasizing the interpretation of predictions from various models, including deep learning. Nikparvar's insights into deep LSTM applications, mobility data, and multivariate time series can enrich our understanding of feature selection and model interpretability in deep learning for epidemiological forecasting.

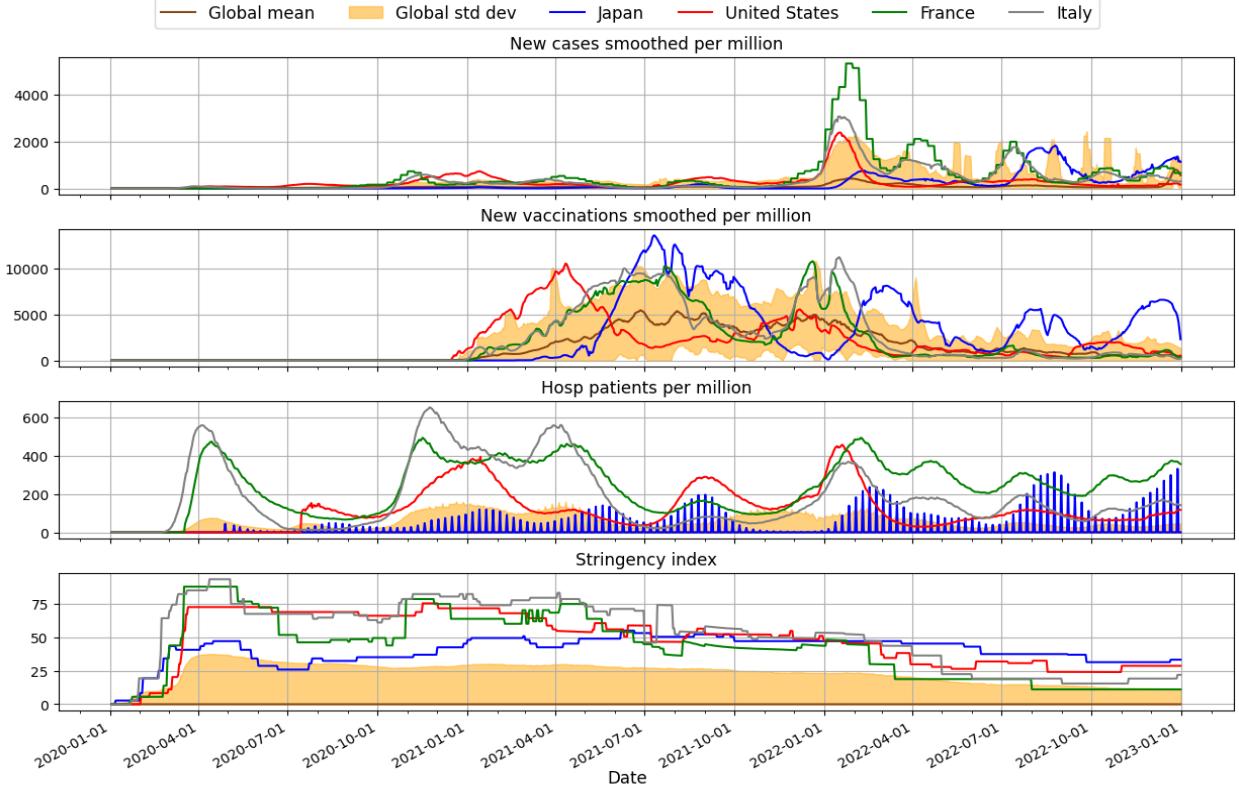
The work by Rodriguez et al. [8] presents a comprehensive framework for real-time COVID-19 forecasting using deep learning techniques. This offers avenues to enhance our models' predictive capabilities, potentially leading to more precise disease spread forecasts. DeepCOVID's ability to handle noisy and heterogeneous data aligns with a core challenge in epidemiological forecasting. We can apply similar data preprocessing and cleaning strategies to improve our models' robustness with complex real-world data. Addressing data quality issues can provide more reliable insights into disease transmission dynamics. Nguyen et al. [6] introduce the innovative BeCaked model in their study, aiming to address the critical question of which factors primarily influence epidemiological forecasts. BeCaked employs various statistical and deep learning models, including VAR and LSTM, to conduct time-series forecasting for COVID-19 cases. We seek to enhance the interpretability of our predictions and gain a deeper insight into the factors guiding epidemiological forecasting by using their findings.

Sherratt et al. [9] study on multi-model ensemble forecasts of COVID-19 cases in European nations introduces a comprehensive forecasting framework that includes the establishment of a centralized forecasting hub, standardized forecasts, and evaluation metrics like the Weighted Interval Score (WIS), which will serve as a foundational guide for our epidemiological forecasting efforts. By adopting their methodologies, we aim to enhance the transparency and reliability of our predictions. Furthermore, we will leverage their insights into ensemble methods, particularly favoring the use of median ensembles, to improve accuracy of our predictive models.

## 5 METHODOLOGY

### 5.1 Dataset: Collection, Processing and Analysis

**5.1.1 Source:** Our main dataset comes from "Our World in Data" (OWID). OWID is a popular source that's been collecting data about the COVID-19 pandemic since its start. Their dataset has a lot of information, not just about the number of cases, but also about how countries responded, how many people got tested or vaccinated, and other important factors. From January 2020 to September 2023,



**Figure 1: Comparative analysis of COVID-19 time-series data of daily new case count, daily new vaccination count, number of hospital patients, and government response stringency across the four chosen countries - "JPN", "USA", "FRA", "ITA"**

the dataset gives us a clear picture of how the pandemic changed over time. This dataset is publicly available through OWID's official GitHub repository [3] [2].

**5.1.2 Features & Description:** The dataset from OWID is comprehensive, with categories ranging from epidemiological data to policy and stringency metrics. Specific features of note include:

- **Epidemiological Data:** Captures the pandemic's progression with metrics like confirmed cases, ICU admissions, and reproduction rate.
- **Testing Data:** Captures the testing landscape with total tests, new tests, and positivity rates.
- **Vaccination Metrics:** Captures the global vaccination effort, including doses administered and number of individuals vaccinated.
- **Policy Responses:** Captures governmental actions, such as school closures and travel restrictions.
- **Demographic Data:** Captures metrics like population, median age, and GDP per capita.

**5.1.3 Data preprocessing & cleaning:** Firstly, although we had access to this dataset with daily incremental updates as of now (late October 2023), we decided to keep only the data recorded for all 3 years of 2020-2022, starting from the beginning of COVID-19 testing (January 3<sup>rd</sup>, 2020) until December 31<sup>st</sup>, 2022, to maintain disease

relevance. Given the vastness of our dataset and inherent challenges of real-world data collection with 67 feature columns, some of which were relevant only temporarily (especially vaccination-related features) throughout the pandemic thus far, our dataset was guaranteed to have missing values. Upon conducting a missing value analysis, there was an expectedly significant fraction of missing values for some feature columns. The top 5 feature columns with the most missing (null) values were found to be: "weekly\_icu\_admissions", "weekly\_icu\_admissions\_per\_million", "excess\_mortality\_cumulative\_absolute", "excess\_mortality\_cumulative", and "excess\_mortality", each with around 96%-97% missing entries.

Since the country identifier and all the features in our dataset relevant to forecasting new case counts are all quantitative, i.e. of numerical data type, we imputed all missing / null values in these columns with value 0. To verify that this was a sound decision, we found zero 1.5 IQR outliers (valued beyond 1.5 times interquartile range) before and after imputation in every column, confirming that missing value imputation with 0's did not significantly skew the distribution.

**5.1.4 Data Analysis (post processing):** We first analyzed the country case distribution in our cleaned dataset, and select four countries located on vastly different longitudes, with low fractions of imputed data and with some of the highest case counts in the world - "Japan", "USA", "France" and "Italy". Thus, we filter the dataset to include

only time-series data for these countries for our study, as mentioned before. We further analyze the time-series trends for the entire 2020-2022 period. Figure 1 shows time-series plots, for every one of the 4 countries, of counts of daily new cases per million of the country population as well as features that we suspect to be correlated to this case count, including daily new vaccinations and new hospital patients per million, and the government response stringency index. We can visually infer that there is some degree of prominent correlation during some time periods in rising case counts trends with trends of other features. So, we would expect to see, as an initial hypothesis, that some of these features might prove to be the most crucial for the case count forecasts.

Furthermore, we can identify time periods with unusually high spikes in new case counts that could have most desirably been prevented, especially Q1 of 2022 in USA, FRA, ITA and Q3 of 2022 in JPN, during the onset spread of the highly contagious Omicron variant in each country. Thus, based on these plots, we select 3 important time periods for evaluating our comparative cross-phase local forecasting analysis for USA: the Omicron onset phase starting around January 2022, a peak phase before vaccines began to be distributed (before December 2020), and a more recent, declining phase in 2022. It will be interesting to see the differences in the most predictive feature set for every forecast across the three phases with different feature trends.

## 5.2 Forecasting Approaches and Models

Time-series data has a temporal sequence, and to preserve this sequence, we cannot split the data randomly. Instead, we use a method called "rolling-window" or "walk-forward validation" for model validation. In this approach, we start with a training set of a certain size and validate the model on the next time step. We then drop the datapoint at the least recent time-step and expand the training set to include the datapoint for the new time step and repeat the process. This method ensures that we always predict the future based on the past, without any data leakage.

Specifically, for our project, we focus on forecasting the variable "**new\_cases\_smoothed\_per\_million**" for a given day, which represents new confirmed cases of COVID-19 (7-day smoothed) per 1,000,000 people. Additionally, for our multivariate forecasting models, we establish a fixed  $n$ -day window immediately preceding the forecast day. This window includes both the case count data and seven carefully selected time-series features. These features are chosen for their potential significance in reflecting pandemic dynamics and response measures, which may influence the spread of COVID-19 and its case counts. The features include:

- **icu\_patients\_per\_million:** The number of COVID-19 patients in ICUs per million people, reflecting the severity and impact of the pandemic on healthcare systems.
- **reproduction\_rate:** Real-time estimate of effective reproduction rate of COVID-19, indicating how quickly the virus is spreading.
- **hosp\_patients\_per\_million:** The number of COVID-19 patients in hospital per million, showing the current burden on healthcare facilities due to COVID-19.
- **new\_people\_vaccinated\_smoothed\_per\_hundred:** Daily number of people receiving their first vaccine dose (7-day

smoothed) per 100 people in the population , indicative of vaccination progress.

- **new\_vaccinations\_smoothed\_per\_million:** New COVID-19 vaccination doses administered (7-day smoothed) per, providing insight into the pace of vaccination efforts.
- **stringency\_index:** A composite measure based on nine response indicators including school closures and travel bans, reflecting the level of government response.
- **tests\_per\_case:** Tests conducted per new confirmed case of COVID-19 (7-day smoothed), an indicator of testing intensity and case detection.

Given the nature of the problem at hand, which deals with time-series forecasting for COVID-19 cases, we leverage a statistical model and a deep learning model that have shown promise in time-series forecasting on epidemiological data. We select VAR and LSTM as our forecasting models. VAR, being a multivariate extension of the Autoregression model, captures linear dependencies between multiple time series, making it an ideal choice for our dataset with interconnected variables. On the other hand, LSTM, a type of recurrent neural network, excels in learning patterns and dependencies from long sequences, making it apt for forecasting longer horizons with potential non-linear interactions. Below is an elaboration on VAR and LSTM and how we plan to use them for our forecasting analysis:

- **Vector Autoregression (VAR):** VAR is a statistical model used to capture the linear interdependencies among multiple time series [12]. It can be expressed as:

$$Y_t = c + A_1 Y_{t-1} + \dots + A_p Y_{t-p} + e_t$$

where  $Y_t$  is a vector of endogenous variables,  $c$  is a vector of constants,  $A_1, A_2, \dots, A_p$  are matrices of coefficients, and  $e_t$  is a vector of error terms. Leveraging VAR for forecasting involves:

- (1) Testing all time series to ensure stationarity (using **Augmented Dickey-Fuller** test). If any series is non-stationary, we difference it until it becomes stationary.
- (2) Selecting the lag order ' $p$ ' for the VAR model.
- (3) Estimating the VAR model and validating its assumptions, like no autocorrelation in the residuals.
- (4) Using the model for forecasting the next ' $h$ ' steps.

- **Long Short-Term Memory (LSTM):** LSTM networks are a type of recurrent neural network (RNN) architecture that are adept at handling long-term dependencies in sequential data [11]. They utilize memory cells to store and recall information, making them particularly suitable for time-series forecasting. LSTMs, being a deep learning model, require the data to be normalized or standardized. Leveraging LSTMs for forecasting includes:

- (1) Normalizing the time series; transforming the time series data into a supervised learning problem, where we use the last ' $n$ ' time steps to predict the next time step.
- (2) Designing and Training LSTM on transformed data.
- (3) Using the trained LSTM to forecast next ' $h$ ' steps and then inverse-transforming predictions to get original scale.

With these methodologies in place, we are equipped to understand the dynamics of COVID-19 case counts over different phases

and the impact of various factors on its spread. For hyperparameter tuning, we propose the use of **Bayesian Optimization**. To gauge the performance and predictive accuracy of our models, the dataset will be split temporally into training and validation sets, ensuring that the sequences are preserved. This will allow us to not only train our models but also assess their performance in different pandemic phases. The performance metrics we utilize for evaluation are:

- **Mean Absolute Error (MAE):** Defined as:

$$MAE = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i|$$

where  $y_i$  are the actual values,  $\hat{y}_i$  are the predicted values, and  $n$  is the number of samples.

- **Root Mean Square Error (RMSE):** Defined as:

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

### 5.3 Predictive Feature Explainability Analysis

We conduct two different types of thorough analyses for purposes of analyzing explainability, specifically the most predictive features and how much and in what direction they contribute to forecast predictions, based on different explainability metrics:

- **Global Analysis :** In our global analysis, we utilize **Pearson Correlation Coefficients** to evaluate the degree of overall positive or negative correlation of every feature with forecast predictions *by both VAR and LSTM models for the entire three-year analysis period for every country*.
- **Local / Phased Analysis :** In our local or phased analysis, we utilize **SHAP values** to evaluate the degree of positive or negative contribution of every feature to forecast predictions *by only the LSTM model for the country-specific 21-day period of fastest growth* in target "new\_cases\_smoothed\_per\_million" count for every country, which coincides with the Omicron outbreak onset periods. Furthermore, for our cross-phase analysis for USA, we evaluate two additional significant phases: a pre-vaccination peak in 2020, and a more recent phase of fastest decline in 2022.

Now, we delve into the two mentioned explainability metrics that help us gain insight into the importance of specific features in the prediction and signals of epidemic-characteristic case counts:

- **Pearson Correlation Coefficient:** Correlation, a statistical measure, captures the strength and direction of the linear relationship between two variables. The most common metric, Pearson's correlation coefficient ( $r$ ), ranges from -1 to 1, with values close to 1 indicating a strong positive linear relationship and values close to -1 indicating a strong negative linear relationship. Given a feature,  $x_i$ , and our target variable,  $y_i$  (case counts), the Pearson's correlation coefficient is calculated as:

$$r = \frac{\sum (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum (x_i - \bar{x})^2 \sum (y_i - \bar{y})^2}}$$

where  $\bar{x}$ ,  $\bar{y}$  are the means of the feature and the target variable, respectively. A high absolute value of  $r$  for a feature

suggests that changes in the feature could be strongly associated with changes in our target.

- **SHAP (SHapley Additive exPlanations) Values:** SHAP values provide a unified measure of feature importance across different types of machine learning and deep learning models - so it is only suitable for our LSTM model in our study. They decompose a prediction's deviation from the mean prediction into contributions from each feature. For prediction  $p$  and a feature set  $S$ , SHAP is defined as:

$$\phi_j = \sum_{S \subseteq N \setminus \{j\}} \frac{|S|!(|N| - |S| - 1)!}{|N|!} [f(S \cup \{j\}) - f(S)]$$

where  $\phi_j$  is the SHAP value for feature  $j$ ,  $N$  is the set of all features,  $f(S)$  is the model's output for the feature set  $S$ . In simpler terms, for each feature, the SHAP value is the average contribution of that feature to every possible prediction, weighted by the likelihood of that set of features occurring. Features with higher absolute SHAP values are considered more influential for the model's predictions. By ranking features based on their SHAP values, we can determine which ones are the most predictive. Furthermore, SHAP values can be visualized using SHAP force plots, offering insights into the effects of feature values on predictions. **Positive SHAP values** push predictions higher than the mean prediction, while **negative SHAP values** push predictions lower.

## 6 RESULTS & DISCUSSION

### 6.1 Multivariate Time-Series Forecasting

This section delves into the results of our multivariate time-series forecasting, where we employed both statistical and deep learning models, specifically Vector Autoregression (VAR) and Long Short-Term Memory (LSTM) networks. We focused on forecasting the variable *new\_cases\_smoothed\_per\_million* for a given day, representing smoothed new confirmed cases of COVID-19 per million people.

#### 6.1.1 Data Preparation and Model Configuration.

- **Data Split:** We adhered to an 80%-20% train-test split, maintaining the temporal order of the dataset without shuffling, to prevent look-ahead bias.
- **Scaling:** We utilized MinMax scaling for data normalization, fitting the scaler on the training set and transforming both training and test sets. This preprocessing step was crucial for effective modeling, particularly in VAR and LSTM models, where scaled data significantly impacts model performance.

#### 6.1.2 VAR Model Insights.

- **Lag Order Selection:** The optimal number of lags, or "lag order," for the VAR model was determined using the Akaike Information Criterion (AIC), striking a balance between model complexity and goodness of fit.

#### 6.1.3 LSTM Model Insights.

- **Historical Window:** The LSTM model was formatted to use a 20-day historical window, allowing each prediction for a given day to be based on data from the previous 20 days.

This configuration aimed to capture short-term temporal dependencies vital for time-series forecasting.

- **Model Architecture:** The LSTM model comprised a sequence of layers, including an initial LSTM layer with 128 nodes, followed by two fully connected layers with 64 and 32 nodes, respectively, and concluding with a final output layer. This architecture facilitated the learning of complex patterns in the data.
- **Training Parameters:** The model was trained for 100 epochs with a batch size of 8, using the mean absolute error (MAE) as the loss function and the Adam optimizer with a learning rate of 0.007.

#### 6.1.4 Performance Metrics.

- **Evaluation:** The performance metrics used for both VAR and LSTM models were Mean Absolute Error (MAE) and Root Mean Squared Error (RMSE). The results, detailed in Table 1 of the report, demonstrated robust and satisfactory performance for both models across all analyzed countries.

Model/Country	JPN	USA	FRA	ITA
<b>Mean Absolute Error (MAE)</b>				
VAR	448.487	72.264	172.839	88.437
LSTM	596.257	163.931	170.202	103.345
<b>Root Mean Square Error (RMSE)</b>				
VAR	633.451	95.006	208.542	114.720
LSTM	691.426	165.187	200.414	126.039

**Table 1: Testing Error Metric Results - MAE and RMSE - for VAR and LSTM Models for all countries**

## 6.2 Global Explainability Analysis

In our global analysis, we leveraged Pearson Correlation Coefficients (PCC) to assess the correlation of various features with COVID-19 case count forecasts made by VAR and LSTM models across four countries: Japan, USA, France, and Italy, over the entire three-year period of the pandemic analyzed for our study. The results of the top 4 predictive features (out of the total 7) and their corresponding **positive** / **negative** PCCs, sorted according to the absolute value of their correlation with daily case count forecasts, as shown in Table 2, provide intriguing insights into how different features correlate with the predictions. A more detailed summary of all PCC values of all seven features with forecasts by both models is shown in the Appendix section in Table 5.

### Country-Wise Analysis:

- **Japan (JPN):** The positive correlation of new vaccinations with case counts in Japan, seen in both VAR and LSTM models, is somewhat counterintuitive. This anomaly could be attributed to the late biggest surge of COVID-19 cases in Japan, due to its extreme eastern longitudinal location compared to other countries, particularly during the late Omicron variant outbreak in Japan in Q3 of 2022. The aggressive vaccination campaigns due to global precedent set by the

Omicron outbreak in other countries of the world in during this period might have coincided with the rise in cases when the Omicron outbreak had only just started in Japan. This observation suggests that standard vaccines were less effective against newer strains like Omicron, underscoring the necessity for booster shots and targeted vaccines as the virus evolved. This hypothesis is further backed by evidence in the local analysis for the Omicron onset period in Japan.

- **United States (USA):** In the USA, hospital patients and ICU patients showed strong positive correlations, indicating the burden on healthcare systems was a significant indicator of case counts. Interestingly, tests\_per\_case showed a negative correlation, serving as a reliable inverse indicator of case count trends.
- **France (FRA) and Italy (ITA):** Similar to the USA, hosp\_patients showed a strong correlation with case counts for both these countries. However, more importantly, new\_vaccinations shows very consistent and exclusive strong negative correlations with case counts for both these countries and for both VAR and LSTM models. This negative correlation could indicate that as vaccination rates increased, the number of new COVID-19 cases decreased. This trend is expected and aligns with the general understanding of vaccine efficacy. The negative correlation of tests\_per\_case reaffirms its role as an inverse indicator.

**Common Predictive Features Across Countries:** Across all countries, **hospital patients** and **ICU admissions** consistently emerged as most significant predictors of COVID-19 case counts, reflecting the direct impact of the pandemic on healthcare systems. This confirms our initial hypothesis that we conjured in our time-series data analysis, based on our earlier visualization of similar trends in hospital patients and new daily cases presented in Figure 1. The negative correlation of tests\_per\_case across all countries confirms its inverse relationship with case counts, acting as a sanity check and reinforcing the reliability of our results.

The stringency\_index showed varying degrees of correlation across different countries and models. In Italy (ITA) and the United States (USA), as observed in the VAR model, and in Japan (JPN) as observed in the LSTM model, the stringency\_index exhibited a negative correlation with COVID-19 case counts. This might indicate that stricter government measures potentially led to a decrease in the number of cases, which aligns with the expected impact of stringent policies.

**Cross-Country Differences and Outliers:** The positive correlation of new vaccinations in Japan presents a unique case, likely influenced by the specific timing and nature of the pandemic's progression in the country. This highlights the importance of considering geographical and temporal factors in epidemiological modeling, especially considering Japan's vastly different longitudinal location compared to the other countries.

The strong negative correlation of new\_vaccinations exclusively seen in Italy (ITA) and France (FRA) compared to the lack thereof in USA and JPN can perhaps hint that in countries like France and Italy, vaccination campaigns were rolled out more effectively and at scale than in Japan (JPN) or USA. Increased vaccination coverage could have led to a significant reduction in case counts. Additionally,

this could also highlight the difference in public opinion about or cautiousness against vaccine efficacy in France and Italy compared to USA and Japan.

The negative correlation of reproduction\_rate observed in LSTM model predictions, particularly in Italy, might be due to the model's ability to capture complex, non-linear interactions in the data. This could indicate a lag or indirect effect where initial high reproduction rates might lead to stricter measures and behavioral changes, eventually reducing case counts.

Overall, these findings underline the significance of localized factors in shaping the pandemic's trajectory and the response effectiveness. They also illustrate the nuanced nature of correlations in epidemiological data, where direct causation is not always evident, and multiple factors can interplay in complex ways.

### 6.3 Local / Phased Explainability Analysis

In our phased local analysis, we applied SHAP (SHapley Additive exPlanations) values to analyze the most predictive features during the 21-day span of the sharpest COVID-19 case rises in each country. Using SHAP values from LSTM model forecasts, we assessed feature influence on case count predictions at the end of each nation-specific period. The LSTM model, predicting the following day's cases from the past 20 days' data, essentially evaluates features over these 21-day intervals. These intervals often corresponded with the initial phase of the Omicron variant's spread, as shown in Figure 1, with the timing varying across countries. Our local SHAP analysis during each country's initial Omicron phase is summarized in Table 3, with comprehensive SHAP force plots for Japan, Italy, the USA, and France provided in the Appendix (Figures 3, 2, 4, 5).

#### Country-specific insights:

- **Japan (JPN):** Japan's positive SHAP values for hospital and ICU patients align with global trends, where these indicators reflect the escalating impact of the pandemic. However, the positive correlation with new vaccinations during the Omicron outbreak suggests a more complex scenario. It may indicate that increased vaccination efforts coincided with a surge in cases, possibly reflecting a delay in the vaccines' effectiveness or a response to the late arrival of the Omicron wave in Japan. Notably, the reproduction rate, which quantifies how quickly the virus is spreading, also showed a positive SHAP value, indicating its significant role during the surge, but it was not among the top predictors.
- **United States (USA):** The United States exhibits the most considerable SHAP values for ICU patients, emphasizing the extreme pressure on healthcare systems. Hospital admissions are a close second, reinforcing their predictive strength. Although the reproduction rate doesn't feature as a top predictor, its role in the spread of the virus is implicitly understood through its impact on hospital and ICU admissions.
- **France (FRA):** In France, the pattern holds, with high SHAP values for hospital and ICU patients, signifying their predictive reliability. Tests per case show a positive yet low SHAP value, hinting at its marginal role in predicting case counts compared to the more substantial healthcare indicators. The reproduction rate, again, is not highlighted, possibly overshadowed by the direct healthcare indicators.

- **Italy (ITA):** Italy presents an intriguing case where decreased vaccinations correlate with a significant increase in case counts, as indicated by the negative SHAP value. This suggests that decreased vaccinations and perhaps a less effective vaccination campaign played a role in the outbreak.

**Common Predictive Features Across Countries:** Across all countries, **hospital patients** and **ICU admissions** are consistently significant predictors of COVID-19 case counts, with high positive SHAP values indicating the direct impact of the pandemic on healthcare systems. These metrics are central to understanding and forecasting the trends of COVID-19.

**Cross-Country Differences and Outliers:** Japan's positive SHAP value for new vaccinations during the Omicron outbreak contrasts with Italy's negative value, suggesting variations in the timing and effectiveness of vaccination campaigns. The lower magnitude of SHAP values for tests per case in the USA and France, compared to their healthcare system indicators, suggests that while testing is important, it may not be as immediate a predictor of case trends as the direct impact on healthcare facilities. The increasing reproduction rate, while a significant factor in Japan's COVID-19 surge, showed limited predictive value in other countries.

In summary, the SHAP value-based analysis during the Omicron outbreak reveals that while the burden on healthcare systems is a universal predictor of COVID-19 case trends, the role of vaccinations and testing varies significantly between countries, offering vital insights for pandemic response strategies.

### 6.4 Cross-Phase Explainability Analysis (USA only)

For our cross-phase explainability analysis, a specialized extension of the prior local analysis, we investigate possible variance in predictive features deduced from local SHAP value analysis of LSTM forecasts, focusing on temporal dynamics in different phases of the pandemic in the USA. Specifically, our local SHAP-based explainability analysis is limited to three distinct 21-day periods or phases of the pandemic progression in the USA: a peak phase in 2020 (first year of the pandemic) before vaccines were distributed, the same peak phase (Q1 of 2022) during the Omicron outbreak onset period analyzed in the previous section, and finally, the phase of the fastest declining case count after the first Omicron onset peak. The results of the features deduced (based on SHAP values) to be predictive of case count trends in each of these three phases in the USA are summarized in Table 4.

#### Pre-Vaccine Peak Phase (2020):

- Key features: ICU Patients, Hospital Patients, Tests per Case.
- Notable trend: Relatively higher SHAP values for ICU and Hospital Patients, reflecting their critical role in predicting case counts under strained healthcare systems.

#### Peak Phase During Omicron Outbreak (Q1 2022):

- Key features: ICU Patients, Hospital Patients, Tests per Case.
- Notable trend: Continued importance of healthcare system indicators, albeit with a slightly different magnitude, suggesting ongoing challenges despite advancements in treatments and vaccination.

#### Phase of Fastest Decline Post-Omicron Peak (2022):

Country	Model	Top Feature 1	Top Feature 2	Top Feature 3	Top Feature 4
JPN	VAR	new_vaccinations (0.245)	tests_per_case (-0.226)	hosp_patients (0.181)	icu_patients (0.108)
	LSTM	new_vaccinations (0.250)	tests_per_case (-0.246)	stringency_index (-0.225)	hosp_patients (0.205)
USA	VAR	hosp_patients (0.818)	icu_patients (0.664)	tests_per_case (-0.250)	stringency_index (-0.114)
	LSTM	hosp_patients (0.603)	icu_patients (0.506)	reproduction_rate (-0.282)	tests_per_case (-0.116)
FRA	VAR	hosp_patients (0.432)	tests_per_case (-0.316)	new_vaccinations (-0.214)	icu_patients (0.209)
	LSTM	hosp_patients (0.447)	tests_per_case (-0.307)	reproduction_rate (-0.274)	new_vaccinations (-0.164)
ITA	VAR	tests_per_case (-0.394)	hosp_patients (0.247)	new_vaccinations (-0.114)	stringency_index (-0.094)
	LSTM	tests_per_case (-0.386)	new_vaccinations (-0.285)	reproduction_rate (-0.267)	hosp_patients (0.179)

Table 2: Global Analysis: Top 4 most predictive features based on Pearson Correlation Coefficient (PCC) of features with LSTM and VAR forecasts for every country (sorted in decreasing order of absolute PCC - red = positive PCC, blue = negative PCC)

Country	Pred. Date	Predictive Features (Deduced from & sorted by absolute SHAP value for LSTM forecast)
JPN	July 29, 2022	1. hosp_patients (0.640), 2. icu_patients (0.446), 3. new_vaccinations (0.396), 4. reproduction_rate (0.321)
USA	Jan 3, 2022	1. icu_patients (0.679), 2. hosp_patients (0.665), 3. tests_per_case (0.063)
FRA	Jan 10, 2022	1. hosp_patients (0.657), 2. icu_patients (0.548), 3. tests_per_case (0.039)
ITA	Jan 10, 2022	1. new_vaccinations (-0.850), 2. hosp_patients (0.418), 3. icu_patients (0.369), 4. tests_per_case (0.022)

Table 3: Local / Phased Analysis: Top most predictive features based on SHAP values of contributions of features corresponding to LSTM model forecasts for given end date of country-specific 21-day period of fastest growth in daily COVID-19 case count for every country (sorted in decreasing order of absolute SHAP value - red = positive SHAP, blue = negative SHAP). All these country-specific periods expectedly coincide with the onset of the highly contagious Omicron variant outbreak in each country.

- Key features: Tests per Case, ICU Patients, Hospital Patients (all with high absolute SHAP values but negative, indicating a decrease in these metrics).
- Notable trend: The negative SHAP values for these features reflect their decreasing trends in this phase, signaling a decline in pandemic severity. The high absolute values of SHAP for Tests per Case highlight its significant role in indicating the effective control of the virus spread during this phase.

**Special Observation for the Fastest Decline Phase:** In this phase, further analysis based on the SHAP force plots reveals the isolated impact of changes in less predictive features like stringency\_index on the forecasted downward trend. The stringency\_index increased two-fold from 0.2 to 0.4 in the USA, indicating aggressively ramped-up government measures in response to the Omicron outbreak. This increase in stringency correlated with the decreased case count forecast, evidenced in the SHAP force plot for stringency\_index in this phase - Figure 6 in Appendix section.

## 7 CONCLUSION

In conclusion, our extensive research underscores that hospital and ICU admissions are paramount in predicting COVID-19 case trends, as consistently evidenced across global, local, and cross-phase analyses. These healthcare system metrics, crucially highlighted by their significant SHAP values, should be central to epidemic monitoring and forecasting.

Additionally, new vaccinations emerge as the second most predictive feature, emphasizing the vital role of vaccination campaigns in influencing the course of the pandemic. The stringency index and reproduction rate, while varying in their significance, contribute

Phase	Pred. Date	Predictive Features
Pre-Vaccine	Dec 2, 2020	1. icu_patients (0.792), 2. hosp_patients (0.626), 3. tests_per_case (0.150)
Fastest Increase (Omicron onset peak)	Jan 3, 2022	1. icu_patients (0.679), 2. hosp_patients (0.665), 3. tests_per_case (0.063)
Fastest Decline (After Omicron peak)	Mar 4, 2022	1. tests_per_case (-0.515), 2. icu_patients (-0.225), 3. hosp_patients (-0.207)

Table 4: Cross-Phase Analysis: Top predictive features based on SHAP values of contributions of features corresponding to LSTM model forecasts for different phases or periods of the pandemic for USA (sorted in decreasing order of absolute SHAP value - red = positive SHAP, blue = negative SHAP).

valuable insights into government response efficacy and the virus's transmission dynamics.

These findings advocate for a multifaceted approach in epidemic management. Governments and health organizations should prioritize real-time tracking of hospital admissions and vaccination data for effective epidemic forecasting and intervention planning. By integrating these key predictive features into their strategies, policymakers can enhance their preparedness and responsiveness, leading to more effective control and management of public health crises. This study reaffirms the importance of a comprehensive, data-driven approach in understanding and tackling epidemics.

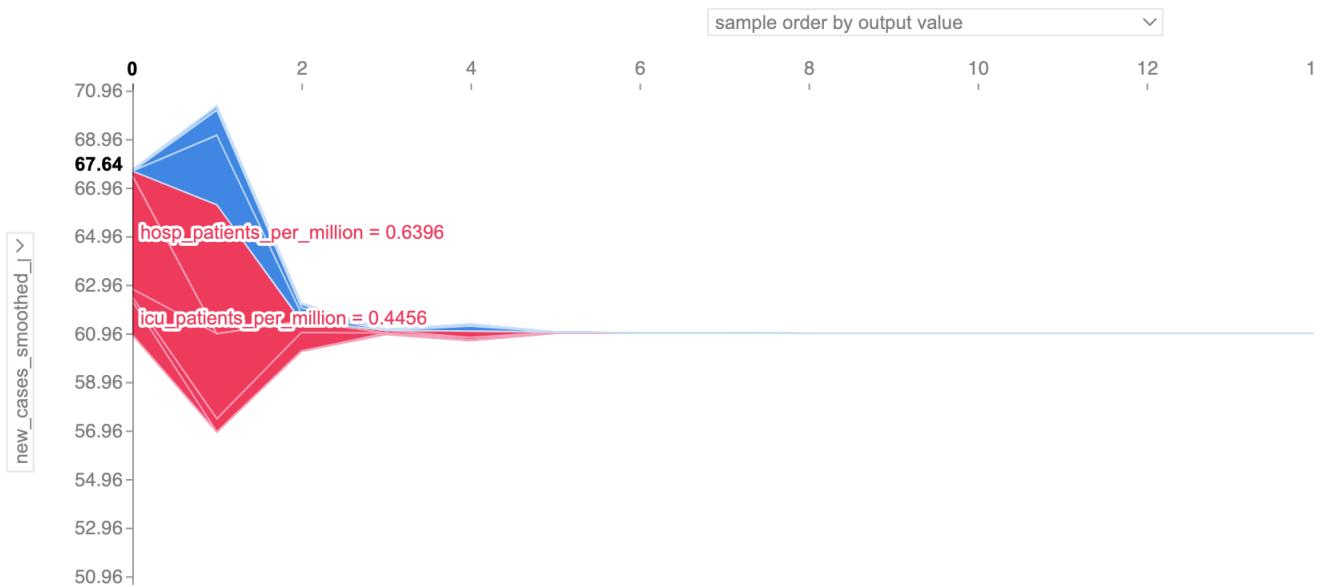
## REFERENCES

- [1] Jayanthi Devaraj, Rajvikram Madurai Elavarasan, Rishi Pugazhendhi, G.M. Shafiqullah, Sumathi Ganesan, Ajay Kaarthic Jeysree, Irfan Ahmad Khan, and Eklas Hossain. 2021. Forecasting of COVID-19 cases using deep learning models: Is it reliable and practically significant? *Results in Physics* 21 (2021), 103817. <https://doi.org/10.1016/j.rinp.2021.103817>
- [2] Joe Hasell, Edouard Mathieu, Diana Beltekian, et al. 2020. A cross-country database of COVID-19 testing. *Scientific Data* 7 (2020), 345. <https://doi.org/10.1038/s41597-020-00688-8>
- [3] Edouard Mathieu, Hannah Ritchie, Esteban Ortiz-Ospina, et al. 2021. A global database of COVID-19 vaccinations. *Nature Human Behaviour* 5 (2021), 947–953. <https://doi.org/10.1038/s41562-021-01122-8>
- [4] Thabang Mathonsi and Terence L. van Zyl. 2022. A Statistics and Deep Learning Hybrid Method for Multivariate Time Series Forecasting and Mortality Modeling. *Forecasting* 4, 1 (2022), 1–25. <https://doi.org/10.3390/forecast4010001>
- [5] Suryanshi Mishra, Tinku Singh, Manish Kumar, and Satakshi. 2023. Multivariate time series short term forecasting using cumulative data of coronavirus. *Evolving Systems* (2023), 1–18.
- [6] Duc Q. Nguyen, Nghia Q. Vo, Thinh T. Nguyen, et al. 2022. BeCaked: An Explainable Artificial Intelligence Model for COVID-19 Forecasting. *Sci Rep* 12 (2022), 7969. <https://doi.org/10.1038/s41598-022-11693-9>
- [7] B. Nikparvar, M.M. Rahman, F. Hatami, et al. 2021. Spatio-temporal prediction of the COVID-19 pandemic in US counties: modeling with a deep LSTM neural network. *Sci Rep* 11 (2021), 21715. <https://doi.org/10.1038/s41598-021-01119-3>
- [8] A. Rodríguez, A. Tabassum, J. Cui, J. Xie, J. Ho, P. Agarwal, B. Adhikari, and B. A. Prakash. 2021. DeepCOVID: An Operational Deep Learning-driven Framework for Explainable Real-time COVID-19 Forecasting. In *Proceedings of the AAAI Conference on Artificial Intelligence*, Vol. 35. 15393–15400. <https://doi.org/10.1609/aaai.v35i17.17808>
- [9] Katharine Sherratt, Hugo Gruson, Rok Grah, et al. 2023. Predictive performance of multi-model ensemble forecasts of COVID-19 across European nations. *eLife* 12 (2023), e81916. <https://doi.org/10.7554/eLife.81916>
- [10] Sanzida Solayman, Sk. Azmiara Aumi, Chand Sultana Mery, Muktadir Mubassir, and Riasat Khan. 2023. Automatic COVID-19 Prediction using Explainable Machine Learning Techniques. *International Journal of Cognitive Computing in Engineering* 4 (2023), 36–46. <https://doi.org/10.1016/j.ijcce.2023.01.003>
- [11] Ralf C. Staudemeyer and Eric Rothstein Morris. 2019. Understanding LSTM – a tutorial into Long Short-Term Memory Recurrent Neural Networks. (2019). arXiv:1909.09586 [cs.NE]
- [12] James H. Stock and Mark W. Watson. 2001. Vector Autoregressions. *Journal of Economic Perspectives* 15, 4 (December 2001), 101–115. <https://doi.org/10.1257/jep.15.4.101>

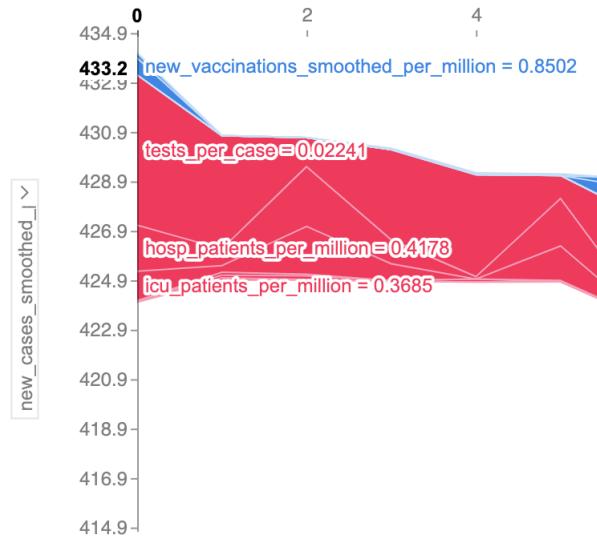
## 8 APPENDIX (NEXT PAGE ONWARDS)

Country	Model	icu_patients	reproduction_rate	hosp_patients	new_people_vaccinated	new_vaccinations	stringency_index	tests_per_case
JPN	VAR	0.107729	0.007737	0.180703	0.001807	0.245047	-0.028810	-0.225630
	LSTM	0.095168	-0.154007	0.204755	-0.003464	0.250026	-0.225034	-0.246298
USA	VAR	0.663982	0.009238	0.818069	0.016836	-0.053772	-0.114198	-0.250081
	LSTM	0.506415	-0.282252	0.602685	0.016063	-0.020632	-0.058847	-0.116270
FRA	VAR	0.208859	-0.010934	0.432336	-0.028362	-0.214295	-0.060527	-0.316253
	LSTM	0.113928	-0.273564	0.446505	-0.019303	-0.163998	-0.061879	-0.307374
ITA	VAR	0.076757	0.017491	0.247071	0.002477	-0.113772	-0.094205	-0.394261
	LSTM	-0.007949	-0.266659	0.179372	-0.027531	-0.285053	-0.044861	-0.385535

**Table 5: (APPENDIX) Pearson Correlation Coefficients (PCCs) (rounded to 6 decimal places) of all 7 features with multivariate LSTM and VAR model forecasts - red = positive PCC, blue = negative PCC)**



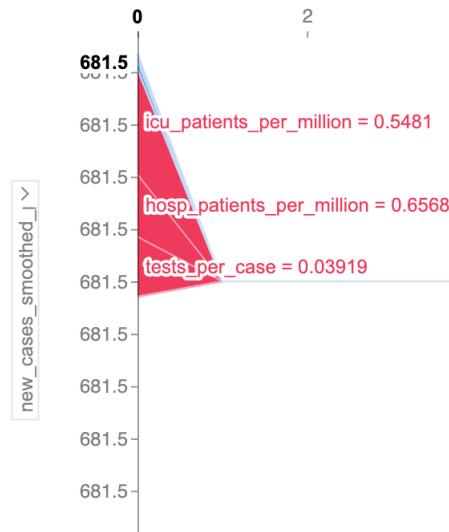
**Figure 2: (APPENDIX) SHAP force plot for Japan (JPN) highlighting predictive features during the Omicron outbreak onset period.**



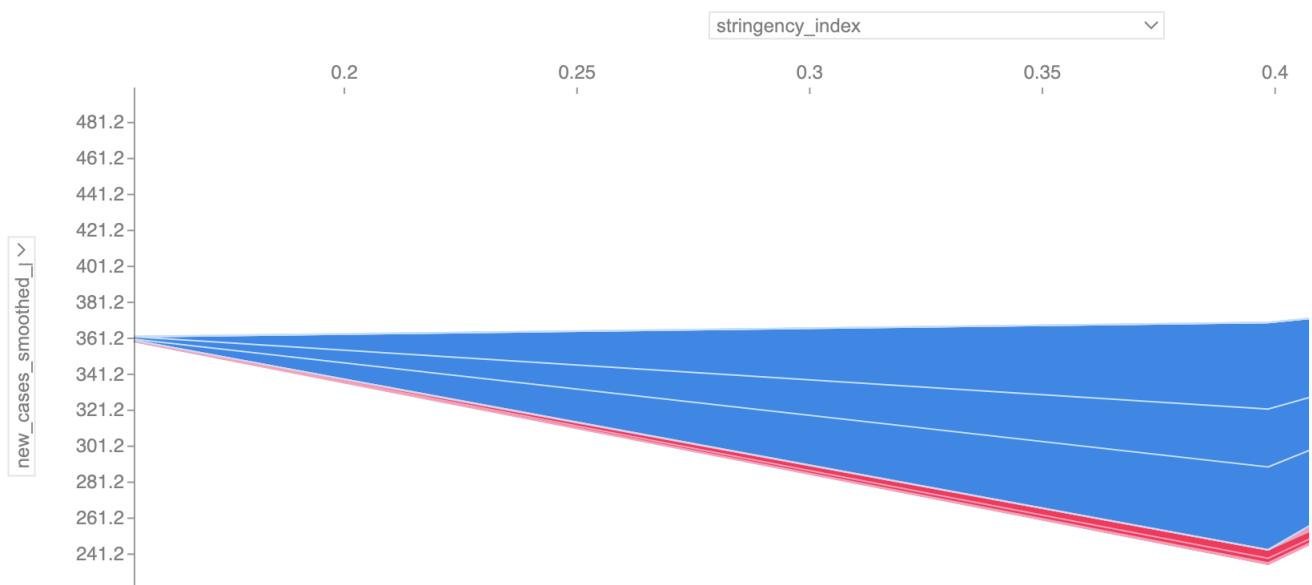
**Figure 3: (APPENDIX) SHAP force plot for Italy (ITA) showing key features influencing COVID-19 case count predictions during the Omicron outbreak.**



**Figure 4: (APPENDIX) Force plot of SHAP values for the USA, indicating features with the most impact in forecasting case counts during the Omicron phase.**



**Figure 5: (APPENDIX) Visualization of SHAP values for France (FRA) depicting influential features during the rapid increase of cases in the Omicron outbreak.**



**Figure 6: (APPENDIX) SHAP force plot illustrating the impact of the stringency index during the phase of fastest decline in COVID-19 cases (after first Omicron onset peak) in the USA. The plot shows the increase in stringency index correlating with the decrease in case counts.**

# Explainable Models for Epidemiological Forecasting

Shayar Shah  
sshah630@gatech.edu  
Georgia Institute of Technology  
Atlanta, Georgia, USA

Aryan Pariani  
apariani3@gatech.edu  
Georgia Institute of Technology  
Atlanta, Georgia, USA

Aravind Vengarai  
avs9@gatech.edu  
Georgia Institute of Technology  
Atlanta, Georgia, USA

Conference'17, July 2017, Washington, DC, USA  
© 2023 Association for Computing Machinery.  
ACM ISBN 978-x-xxxx-xxxx-x/YY/MM...\$15.00  
<https://doi.org/10.1145/nnnnnnnn.nnnnnnn>

## ABSTRACT

The increasing reliance on intricate deep learning and statistical or machine learning models for epidemiological forecasting has prompted the development of accurate, albeit complex, algorithms for understanding disease spread, particularly COVID-19. However, the black-box nature of these models poses challenges in deciphering the precise factors leading to their predictions. The objective of our project is to employ Explainable AI techniques to interpret the time-series case count forecasts by statistical and deep learning models, and determine the most predictive features from a multivariate time-series dataset for the COVID-19 pandemic collected by *Our World in Data* from January 2020 to September 2023. We aim to explore this using models like vector autoregression and LSTM. Furthermore, this study aims to comprehend how these influential features vary across different epidemic phases and during abrupt outbreaks. The investigation will employ both correlation studies and other metrics to analyze the features' contributions during diverse outbreaks and pandemic phases.

## KEYWORDS

Explainable AI, Predictive Feature Extraction, Time-series forecasting, LSTM

### ACM Reference Format:

Shayar Shah, Aryan Pariani, and Aravind Vengarai. 2023. Explainable Models for Epidemiological Forecasting. In *Proceedings of ACM Conference (Conference'17)*. ACM, New York, NY, USA, 6 pages. <https://doi.org/10.1145/nnnnnnnn.nnnnnnn>

## 1 INTRODUCTION

The novel coronavirus's rapid proliferation during the COVID-19 pandemic has significantly influenced statistical and machine learning research, enhancing epidemiological forecasting, such as predicting future case counts in large-scale epidemics. Researchers have applied numerous statistical and deep learning models to forecast accurately and support timely interventions. Yet, the complex nature of these models often hides the rationale behind their predictions, challenging the extraction of key features for early signal detection and targeted policy-making. The interpretability of these models is crucial as they inform governmental and health organization decisions, particularly for preemptive actions during a pandemic's critical phases.

In forecasting COVID-19's spread, factors like government policies, vaccination rates, and healthcare data are pivotal, though their combined effects may vary across nations. Recognizing the

potential of a nuanced approach, we investigate if aggregated features, such as the interplay between vaccination trends and policy strictness, enhance predictive accuracy. This study is driven by the necessity for in-depth feature analysis alongside advanced time-series forecasting. By analyzing a comprehensive dataset from *Our World in Data*, we aim to predict pandemic trends with greater insight through meticulous feature extraction.

## 2 RESPONSE TO PROPOSAL COMMENTS

We have focused directly on the comments made in the graded proposal, and the changes made to address those can be easily observed. As recommended, we have added more relevant references in our related works and expanded on others to ensure that there is sufficient background presented to understand each aspect of our work thoroughly.

We addressed the first suggestion by adding more references and expanding on our literature review section by discussing 2 more papers (one by Nguyen et al. and another by Sherratt et al.).

We address the second suggestion in Section 4.1.3, and we elaborate on the pre-and post-cleaning statistics of our dataset, which initially had significant missing values in key columns like "weeklyIcuAdmissions" and "excessMortality," amounting to 96%-97% missing data. After imputing zeroes for missing numerical data, our outlier analysis confirmed the integrity of the dataset's distribution.

## 3 PROBLEM STATEMENT FORMULATION

We aim to answer the pressing question in this work: "Which configurations of factors or features primarily drive models in their predictions for epidemiological forecasting?" Delving deeper, understanding how these critical factors change across various phases of an epidemic, especially before and during unexpected outbreaks like the delta and omicron variants of COVID-19, becomes essential. Overall, this project aims to:

- (1) Train and run different statistical and deep learning models like VAR and LSTM for time-series forecasting of COVID-19 cases, based on an extensive multivariate dataset with features like vaccinations, number of ICU patients, and measure of government response stringency, spanning from January 2020 until September 2023, as of the present.
- (2) Identify and rank the features most predictive of COVID-19 case count predictions by models, while assessing different statistical metrics like correlation and SHAP values that can shed light on the relative degrees of contribution of different features.
- (3) Understand the temporal dynamics of these features and how their importance varies during different phases or outbreaks of the pandemic.

By answering these questions, we aim to not only make predictions more understandable but also contribute to enhancing trust and transparency in AI-driven epidemiological forecasts.

## 4 LITERATURE REVIEW

To contextualize our work on "Explainable Models for Epidemiological Forecasting," we will review relevant research papers that contribute to the field of epidemiological prediction and provide insights into the challenges and opportunities in this domain.

Solayman et al. [10] present a significant advancement in COVID-19 prediction by deploying machine learning models, including a hybrid CNN-LSTM approach and explainable AI techniques like the LIME framework. This work complements our original research as it addresses the critical need for transparent and interpretable models in understanding and predicting the dynamics of the COVID-19 pandemic. By demonstrating the effectiveness of the CNN-LSTM model in capturing the temporal aspects of symptom progression and applying explainable AI, this study provides valuable insights into feature interpretability—a crucial aspect of epidemiological forecasting.

The work by Devaraj et al. [1] underscores the potential of deep learning models, notably the Stacked Long Short-Term Memory (SLSTM) algorithm, in achieving highly accurate predictions of COVID-19 cases, with an error rate of under 2%. We intend to leverage these findings to enhance the explainability of our epidemiological forecasting models. Moreover, considering various variables, including temperature, rainfall, demonstrates the models' ability to capture complex epidemiological dependencies, and we aim to draw inspiration from these insights by incorporating similar factors and relationships into our models to ensure they provide interpretable, data-driven forecasts for epidemiological scenarios.

The work by Mishra et al. [5] delves into advanced mathematical and deep learning models for short-term COVID-19 incidence and mortality forecasting. Mishra et al. expand the traditional SEIR model by including hospitalization, vaccination, and quarantine data, a concept pertinent to our goal of enhancing model explainability by considering a broader range of factors in epidemiological scenarios.

Mathonsi and van Zyl [4] introduce the Multivariate Exponential Smoothing Long Short-Term Memory (MES-LSTM) hybrid model, an extension of the ES-RNN model for multivariate time series forecasting. Their main objective is to enhance forecast accuracy and prediction interval construction across various applications, including COVID-19 morbidity datasets. While we emphasize improving model interpretability, Mathonsi and van Zyl's work aligns with the broader trend of enhancing forecasting accuracy, a crucial aspect of epidemiological modeling.

The work by Nikparvar et al. [7] employs a deep LSTM neural network for county-level COVID-19 forecasting, addressing challenges in accurate prediction during a dynamic pandemic with spatial variations. In contrast, our focus is on developing explainable epidemiological forecasting models, emphasizing the interpretation of predictions from various models, including deep learning. Nikparvar's insights into deep LSTM applications, mobility data, and multivariate time series can enrich our understanding of feature

selection and model interpretability in deep learning for epidemiological forecasting.

The work by Rodríguez et al. [8] presents a comprehensive framework for real-time COVID-19 forecasting using deep learning techniques. This offers avenues to enhance our models' predictive capabilities, potentially leading to more precise disease spread forecasts. DeepCOVID's ability to handle noisy and heterogeneous data aligns with a core challenge in epidemiological forecasting. We can apply similar data preprocessing and cleaning strategies to improve our models' robustness with complex real-world data. Addressing data quality issues can provide more reliable insights into disease transmission dynamics.

Nguyen et al. [6] introduce the innovative BeCaked model in their study, aiming to address the critical question of which factors primarily influence epidemiological forecasts. BeCaked employs various statistical and deep learning models, including VAR and LSTM, to conduct time-series forecasting for COVID-19 cases. We seek to enhance the interpretability of our predictions and gain a deeper insight into the factors guiding epidemiological forecasting by using their findings.

Sherratt et al. [9] study on multi-model ensemble forecasts of COVID-19 cases in European nations introduces a comprehensive forecasting framework that includes the establishment of a centralized forecasting hub, standardized forecasts, and evaluation metrics like the Weighted Interval Score (WIS), which will serve as a foundational guide for our epidemiological forecasting efforts. By adopting their methodologies, we aim to enhance the transparency and reliability of our predictions. Furthermore, we will leverage their insights into ensemble methods, particularly favoring the use of median ensembles, to improve the accuracy of our predictive models.

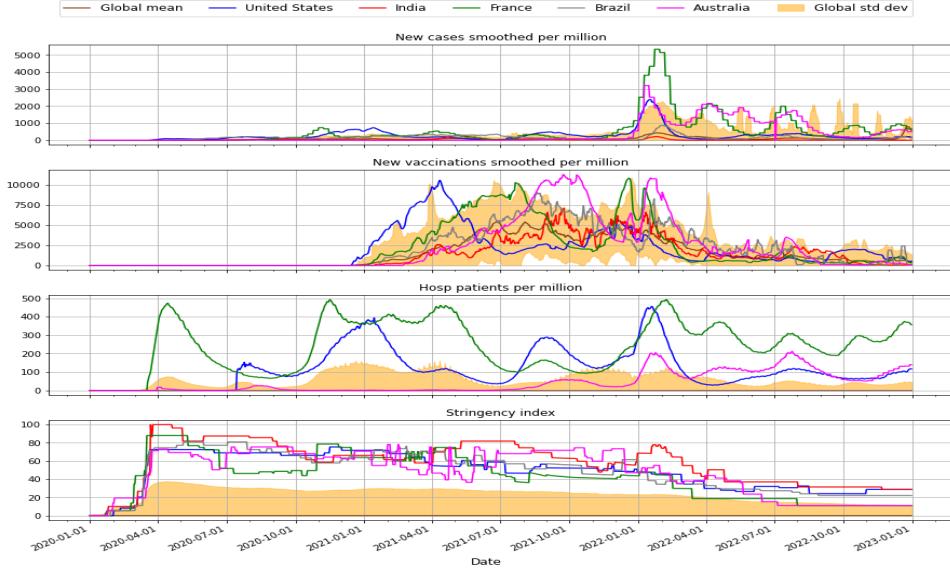
## 5 METHODOLOGY

### 5.1 Dataset: Collection, Processing and Analysis

**5.1.1 Source:** Our main dataset comes from "Our World in Data" (OWID). OWID is a popular source that's been collecting data about the COVID-19 pandemic since its start. Their dataset has a lot of information, not just about the number of cases, but also about how countries responded, how many people got tested or vaccinated, and other important factors. From January 2020 to September 2023, the dataset gives us a clear picture of how the pandemic changed over time. This dataset is publicly available through OWID's official GitHub repository [3] [2].

**5.1.2 Features & Description:** The dataset from OWID is comprehensive, with categories ranging from epidemiological data to policy and stringency metrics. Specific features of note include:

- **Epidemiological Data:** Captures the pandemic's progression with metrics like confirmed cases, ICU admissions, and reproduction rate.
- **Testing Data:** Captures the testing landscape with total tests, new tests, and positivity rates.
- **Vaccination Metrics:** Captures the global vaccination effort, including doses administered and number of individuals vaccinated.
- **Policy Responses:** Captures governmental actions, such as school closures and travel restrictions.



**Figure 1: Comparative Analysis of COVID-19 Impact and Response Measures in Different Countries**

- **Demographic Data:** Captures metrics like population, median age, and GDP per capita.

**5.1.3 Data preprocessing & cleaning:** Firstly, although we had access to this dataset with daily incremental updates as of now (late October 2023), we decided to keep only the data recorded for all 3 years of 2020-2022, starting from the beginning of COVID-19 testing (January 3<sup>rd</sup>, 2020) until December 31<sup>st</sup>, 2022, to maintain disease relevance. Given the vastness of our dataset and inherent challenges of real-world data collection with 67 feature columns, some of which were relevant only temporarily (especially vaccination-related features) throughout the pandemic thus far, our dataset was guaranteed to have missing values. Upon conducting a missing value analysis, there was an expectedly significant fraction of missing values for some feature columns. The top 5 feature columns with the most missing (null) values were found to be: "weekly\_icu\_admissions", "weekly\_icu\_admissions\_per\_million", "excess\_mortality\_cumulative\_absolute", "excess\_mortality\_cumulative", and "excess\_mortality", each with around 96%-97% missing entries.

Since the country identifier and all the features in our dataset relevant to forecasting new case counts are all quantitative, i.e. of numerical data type, we imputed all missing / null values in these columns with value 0. To verify that this was a sound decision, we found zero 1.5 IQR outliers (valued beyond 1.5 times interquartile range) before and after imputation in every column, confirming that missing value imputation with 0's did not significantly skew the distribution.

**5.1.4 Data Analysis (post processing):** We first analyzed the country case distribution in our cleaned dataset, and find the country with the highest total case count in every one of the 5 major continents: United States (North America), Brazil (South America), India (Asia), France (Europe) and Australia. Thus, we filter the dataset to include only time-series data for these countries for our study. We further analyze the time-series trends for the entire 2020-2022

period. Figure 1 shows time-series plots, for every one of the 5 countries, of counts of daily new cases per million of the country population as well as features that we suspect to be correlated to this case count, including daily new vaccinations and new hospital patients per million, and the government response stringency index. We can visually infer that there is some degree of prominent correlation during some time periods in rising case counts trends with trends of other features. So, we would expect to see, as an initial hypothesis, that some of these features might prove to be the most crucial for the case count forecasts.

Furthermore, we can identify time periods with unusually high spikes in new case counts that could have most desirably been prevented, especially Q1 of 2022 during the onset spread of the highly contagious Omicron variant. Thus, based on these plots, we select 3 important time periods for evaluating our phased forecasting analysis for every country: the Omicron onset phase starting around January 2022, a peak phase before vaccines began to be distributed (before December 2020), and a more recent peak phase in the second half of 2022. It will be interesting to see the differences in the most predictive feature set for every forecast across the three phases with different feature trends.

## 5.2 Forecasting Methods and Models

Time-series data has a temporal sequence, and to preserve this sequence, we cannot split the data randomly. Instead, we use a method called "rolling-window" or "walk-forward validation" for model validation. In this approach, we start with a training set of a certain size and validate the model on the next 'h' time steps, where 'h' is the forecasting horizon. We then expand the training set to include those 'h' time steps and repeat the process. This method ensures that we always predict the future based on the past, without any data leakage.

**Three-Phased Forecasting Analysis:** For our analysis, we divide the time period into three significant phases: the pre-vaccination

peak, the Omicron onset, and a recent peak in the second half of 2022. In each phase, we train our models on the data leading up to the peak, validate on the peak period to understand the model's performance, and finally identify the most predictive features for that phase.

Given the nature of the problem at hand, which deals with time-series forecasting for COVID-19 cases, we leverage a statistical model and a deep learning model that have shown promise in time-series forecasting on epidemiological data. We select VAR and LSTM as our forecasting models instead of Prophet, Transformers and Random Forests. VAR, being a multivariate extension of the Autoregression model, captures linear dependencies between multiple time series, making it an ideal choice for our dataset with interconnected variables. On the other hand, LSTM, a type of recurrent neural network, excels in learning patterns and dependencies from long sequences, making it apt for forecasting longer horizons with potential non-linear interactions. In comparison, while Transformers are powerful for sequence data, they might be over-complex for our specific forecasting needs and can be computationally expensive. Prophet, although robust, is tailored more for univariate time series with strong seasonal effects. Finally, Random Forests, being a tree-based method, might not inherently capture temporal dependencies as effectively as VAR or LSTM in the context of our project's requirements. Below is an elaboration on VAR and LSTM and how we plan to use them for our forecasting analysis:

- **Vector Autoregression (VAR):** VAR is a statistical model used to capture the linear interdependencies among multiple time series [12]. It can be expressed as:

$$Y_t = c + A_1 Y_{t-1} + \dots + A_p Y_{t-p} + e_t$$

where  $Y_t$  is a vector of endogenous variables,  $c$  is a vector of constants,  $A_1, A_2, \dots, A_p$  are matrices of coefficients, and  $e_t$  is a vector of error terms. Leveraging VAR for forecasting involves:

- (1) Testing all time series to ensure stationarity. If any series is non-stationary, we difference it until it becomes stationary.
- (2) Selecting the lag order 'p' for the VAR model.
- (3) Estimating the VAR model and validating its assumptions, like no autocorrelation in the residuals.
- (4) Using the model for forecasting the next 'h' steps.

- **LSTM (Long Short-Term Memory):** LSTM networks are a type of recurrent neural network (RNN) architecture that are adept at handling long-term dependencies in sequential data [11]. They utilize memory cells to store and recall information, making them particularly suitable for time-series forecasting. LSTMs, being a deep learning model, require the data to be normalized or standardized. Leveraging LSTMs for forecasting includes:

- (1) Normalizing the time series; transforming the time series data into a supervised learning problem, where we use the last 'n' time steps to predict the next time step.
- (2) Designing and Training LSTM on transformed data.
- (3) Using the trained LSTM to forecast next 'h' steps and then inverse-transforming predictions to get original scale.

With these methodologies in place, we are equipped to understand the dynamics of COVID-19 case counts over different phases

and the impact of various factors on its spread. For hyperparameter tuning, we propose the use of **Bayesian Optimization**. To gauge the performance and predictive accuracy of our models, the dataset will be split temporally into training and validation sets, ensuring that the sequences are preserved. This will allow us to not only train our models but also assess their performance in different pandemic phases. The performance metrics are:

- **Mean Absolute Error (MAE):** Defined as:

$$MAE = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i|$$

where  $y_i$  are the actual values,  $\hat{y}_i$  are the predicted values, and  $n$  is the number of samples.

- **Root Mean Square Error (RMSE):** Defined as:

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

- **Mean Absolute Percentage Error (MAPE):** Defined as:

$$MAPE = \frac{100\%}{n} \sum_{i=1}^n \left| \frac{y_i - \hat{y}_i}{y_i} \right|$$

### 5.3 Predictive Feature Identification and Evaluation

Here, we delve into the techniques that help us gain insight into the importance of specific features in the prediction and signals of epidemic-characteristic case counts:

- **Correlation:** Correlation, a statistical measure, captures the strength and direction of the linear relationship between two variables. The most common metric, Pearson's correlation coefficient ( $r$ ), ranges from -1 to 1, with values close to 1 indicating a strong positive linear relationship and values close to -1 indicating a strong negative linear relationship. Given a feature,  $x_i$ , and our target variable,  $y_i$  (case counts), the Pearson's correlation coefficient is calculated as:

$$r = \frac{\sum (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum (x_i - \bar{x})^2 \sum (y_i - \bar{y})^2}}$$

where  $\bar{x}, \bar{y}$  are the means of the feature and the target variable, respectively. A high absolute value of  $r$  for a feature suggests that changes in the feature could be strongly associated with changes in our target.

- **SHAP (SHapley Additive exPlanations) Values:** SHAP values provide a unified measure of feature importance across different types of models. They decompose a prediction's deviation from the mean prediction into contributions from each feature. For prediction  $p$  and a feature set  $S$ , SHAP is defined as:

$$\phi_j = \sum_{S \subseteq N \setminus \{j\}} \frac{|S|!(|N| - |S| - 1)!}{|N|!} [f(S \cup \{j\}) - f(S)]$$

where  $\phi_j$  is the SHAP value for feature  $j$ ,  $N$  is the set of all features,  $f(S)$  is the model's output for the feature set  $S$ . In simpler terms, for each feature, the SHAP value is the

average contribution of that feature to every possible prediction, weighted by the likelihood of that set of features occurring. Features with higher absolute SHAP values are considered more influential for the model's predictions. By ranking features based on their SHAP values, we can determine which ones are the most predictive. Furthermore, SHAP values can be visualized using SHAP summary plots, offering insights into the effects of feature values on predictions. Positive SHAP values push predictions higher than the mean prediction, while negative values push predictions lower.

In conclusion, using both correlation and SHAP values, we will comprehensively evaluate the features' importance for predicting COVID-19 cases, shedding light on those that play pivotal roles in the models' forecasts.

## 6 INCOMPLETE OUTCOMES

Our primary focus in the next sprint will be implementing our forecasting model using VAR and LSTMs. Once our forecasting model is in place, we will proceed to the crucial phases of training and testing to obtain the model's forecasts of number of cases.

Once we have the forecasted number of cases data, we will delve into predictive feature identification and analysis part of the project. This step involves constructing a model to pinpoint the specific features or factors that seem to exert a substantial influence on the pandemic's evolution over time and across various regions. Upon gathering all our results, we plan to conduct an informative analysis. We will conduct a comparative study to ascertain how the importance of different features evolves as the pandemic progresses through its distinct phases. Furthermore, we will explore geographical disparities by comparing results among the selected countries.

By the end of the semester, we aim to complete the following with respect to our "Explainable models for Epidemiological Forecasting" project:

- Model Development: Implement LSTM, VAR for time-series forecasting. Compare their performance in terms of accuracy and forecast reliability.
- Explainability Integration: Incorporate Explainable AI (XAI) techniques into our models to ensure that the forecasts are not just accurate but also interpretable. This will assist stakeholders in understanding why certain predictions are made and build trust in the model's outputs.
- Phased Analysis: Analyze the importance of the identified features during different phases of the epidemic (beginning, peak, end/dying out, outbreaks of specific variants). We expect to have preliminary findings, at the very least, on how features contribute differently during these phases.
- Correlational Study: Conduct a comprehensive correlational analysis between identified significant features and the progression of the pandemic during the given time phases and across the given regions, thus improving how we understand how certain factors directly influence outbreak dynamics.
- Extension to Outliers: If time permits, we'll explore the impact of sudden outbreaks and shifts in data distributions on feature importance and prediction accuracy.

## 7 GENERAL DIFFICULTIES

Our data exploration and cleaning have progressed smoothly, with the dataset proving to be comprehensive and of high quality. No significant outliers were present, and the preprocessing was limited mainly to addressing missing values. However, we encountered a challenge in dataset reduction, which involved selecting the most pertinent data from numerous countries. Our strategy focused on the top five countries, each representing a different continent. This approach aimed to provide a targeted and varied dataset that would yield clear insights, minimize extraneous information, and facilitate the identification of trends pertinent to the pandemic's impact across these regions.

In preparation for forecasting, we must now segment the data appropriately. The segmentation, informed by our Data Analysis, will delineate time phases for focused forecasting, enhancing the robustness and fault tolerance of our model by enabling it to predict various global scenarios. Identifying these meaningful segments was initially tricky; however, it became more straightforward when considering vaccination rates, case numbers, the timing of vaccine rollouts, and the emergence of new COVID variants.

Looking ahead to the forecasting phase, we anticipate the complexity of managing an extensive feature set, which may complicate the interpretation of our results. To address this, we will prioritize the analysis and exclusion of features that appear to be extraneous, thereby streamlining our final forecasting task.

## 8 CONTRIBUTION

The completion of this project is a collective effort, with each member leading a specific aspect of the project. All members actively contributed to each other's sections.

Shayar leads data procurement, cleaning, and preprocessing, ensuring data quality and model development. Aryan is responsible for developing deep learning models, and training regimen. Aravind will determine evaluation metrics, led the inference process, and conducted a correlational study for extended analysis.

## 9 EXPECTED TIMELINE OF REMAINING WORK

Dates	Activities
Nov 5 - Nov 12	Initial Forecasting Model Implementation and Testing. Incorporate Milestone report feedback
Nov 12 - Nov 18	Predictive Feature Model and Explainability. Correlational study of predictive features.
Nov 19 - Nov 25	Explore sudden outbreaks and shifts in data, synthesize findings from correlational study, and begin drafting the end-of-semester report.
Nov 26 - Dec 7	Finalize report and prepare a presentation.

Table 1: Expected Timeline

## REFERENCES

- [1] Jayanthi Devaraj, Rajvikram Madurai Elavarasan, Rishi Pugazhendhi, G.M. Shafiu-llah, Sumathi Ganesan, Ajay Kaarthic Jeysree, Irfan Ahmad Khan, and Eklas Hossain. 2021. Forecasting of COVID-19 cases using deep learning models: Is it reliable and practically significant? *Results in Physics* 21 (2021), 103817. <https://doi.org/10.1016/j.rinp.2021.103817>
- [2] Joe Hasell, Edouard Mathieu, Diana Beltekian, et al. 2020. A cross-country database of COVID-19 testing. *Scientific Data* 7 (2020), 345. <https://doi.org/10.1038/s41597-020-00688-8>
- [3] Edouard Mathieu, Hannah Ritchie, Esteban Ortiz-Ospina, et al. 2021. A global database of COVID-19 vaccinations. *Nature Human Behaviour* 5 (2021), 947–953. <https://doi.org/10.1038/s41562-021-01122-8>
- [4] Thabang Mathonsi and Terence L. van Zyl. 2022. A Statistics and Deep Learning Hybrid Method for Multivariate Time Series Forecasting and Mortality Modeling. *Forecasting* 4, 1 (2022), 1–25. <https://doi.org/10.3390/forecast4010001>
- [5] Suryanshi Mishra, Tinku Singh, Manish Kumar, and Satakshi. 2023. Multivariate time series short term forecasting using cumulative data of coronavirus. *Evolving Systems* (2023), 1–18.
- [6] Duc Q. Nguyen, Nghia Q. Vo, Thinh T. Nguyen, et al. 2022. BeCaked: An Explainable Artificial Intelligence Model for COVID-19 Forecasting. *Sci Rep* 12 (2022), 7969. <https://doi.org/10.1038/s41598-022-11693-9>
- [7] B. Nikparvar, M.M. Rahman, F. Hatami, et al. 2021. Spatio-temporal prediction of the COVID-19 pandemic in US counties: modeling with a deep LSTM neural network. *Sci Rep* 11 (2021), 21715. <https://doi.org/10.1038/s41598-021-01119-3>
- [8] A. Rodríguez, A. Tabassum, J. Cui, J. Xie, J. Ho, P. Agarwal, B. Adhikari, and B. A. Prakash. 2021. DeepCOVID: An Operational Deep Learning-driven Framework for Explainable Real-time COVID-19 Forecasting. In *Proceedings of the AAAI Conference on Artificial Intelligence*, Vol. 35. 15393–15400. <https://doi.org/10.1609/aaai.v35i17.17808>
- [9] Katharine Sherratt, Hugo Gruson, Rok Grah, et al. 2023. Predictive performance of multi-model ensemble forecasts of COVID-19 across European nations. *eLife* 12 (2023), e81916. <https://doi.org/10.7554/eLife.81916>
- [10] Sanzida Solayman, Sk. Azmiara Aumi, Chand Sultana Mery, Muktadir Mubassir, and Riasat Khan. 2023. Automatic COVID-19 Prediction using Explainable Machine Learning Techniques. *International Journal of Cognitive Computing in Engineering* 4 (2023), 36–46. <https://doi.org/10.1016/j.ijcce.2023.01.003>
- [11] Ralf C. Staudemeyer and Eric Rothstein Morris. 2019. Understanding LSTM – a tutorial into Long Short-Term Memory Recurrent Neural Networks. (2019). arXiv:1909.09586 [cs.NE]
- [12] James H. Stock and Mark W. Watson. 2001. Vector Autoregressions. *Journal of Economic Perspectives* 15, 4 (December 2001), 101–115. <https://doi.org/10.1257/jep.15.4.101>

# Explainable Models for Epidemiological Forecasting

Shayar Shah

sshah630@gatech.edu

Georgia Institute of Technology

Atlanta, Georgia, USA

Aryan Pariani

apariani3@gatech.edu

Georgia Institute of Technology

Atlanta, Georgia, USA

Aravind Vengarai

avs9@gatech.edu

Georgia Institute of Technology

Atlanta, Georgia, USA

Conference'17, July 2017, Washington, DC, USA

© 2023 Association for Computing Machinery.

ACM ISBN 978-x-xxxx-xxxx-x/YY/MM...\$15.00

<https://doi.org/10.1145/nnnnnnnn.nnnnnnnn>

## ABSTRACT

The increasing reliance on intricate deep learning and statistical or machine learning models for epidemiological forecasting has prompted the development of accurate, albeit complex, algorithms for understanding disease spread, particularly COVID-19. However, the black-box nature of these models poses challenges in deciphering the precise factors leading to their predictions. The objective of our project is to employ Explainable AI techniques to interpret the time-series predictions of several proposed statistical and deep learning models, and determine the most predictive features from a multivariate time-series dataset for the COVID-19 pandemic collected by *Our World in Data* from January 2020 to September 2023. We aim to explore this using models like vector autoregression, LSTM, transformers, and more. Furthermore, this study aims to comprehend how these influential features vary across different epidemic phases and during abrupt outbreaks. The investigation will employ both correlation studies and other metrics to analyze the features' contributions during diverse outbreaks and pandemic phases.

## KEYWORDS

Explainable AI, Predictive Feature Extraction, Time-series forecasting, LSTM

### ACM Reference Format:

Shayar Shah, Aryan Pariani, and Aravind Vengarai. 2023. Explainable Models for Epidemiological Forecasting. In *Proceedings of ACM Conference (Conference'17)*. ACM, New York, NY, USA, 5 pages. <https://doi.org/10.1145/nnnnnnnn.nnnnnnnn>

## 1 INTRODUCTION

The rapid spread of the novel coronavirus in the deadly global COVID-19 pandemic has been a focal point for statistical and machine learning research, pushing the boundaries of epidemiological forecasting, i.e. predicting rates of future case and death counts that are characteristic of deadly, large-scale epidemics. Various statistical and deep learning models have been employed throughout the pandemic, striving for accurate predictions to aid in timely interventions. However, the intricate, black-box nature of these models often obscures their decision-making processes, making it difficult to identify specific predictive features or signals. This understanding can otherwise facilitate identifying early signals for future epidemics and can guide policy for targeted intervention, by closely monitoring such signals within the data so as to be able to take timely steps to prevent future epidemics. As governments

and health organizations make critical decisions based on these forecasts, the interpretability of the models becomes of paramount importance for intervention in different phases of any pandemic, especially before and during the onsets of peak outbreak phases.

In the case of epidemiological forecasting for COVID-19, variables like testing rates, vaccination pace, and population mobility might directly influence disease spread, but the configurations of combinations of these variables can provide a more comprehensive view. For instance, a feature representing the rate of change of vaccinations combined with testing positivity rates might offer a more predictive power than considering them individually. We derive motivation for this work from recognizing this need. While leveraging state-of-the-art time-series modeling and prediction is essential, diving deeper into the data to extract the most impactful features is equally crucial. Utilizing the comprehensive dataset provided by *Our World in Data*, our goal is not just to predict future spread of pandemics but to do so with a deeper understanding by harnessing feature extraction.

## 2 PROBLEM STATEMENT FORMULATION

We aim to answer the pressing question in this work: "Which configurations of factors or features primarily drive models in their predictions for epidemiological forecasting?" Delving deeper, understanding how these critical factors change across various phases of an epidemic, especially before and during unexpected outbreaks like the delta and omicron variants of COVID-19, becomes essential. Overall, this project aims to:

- (1) Train and run different statistical and deep learning models like ARIMA and LSTM's for time-series forecasting of COVID-19 cases and deaths, based on an extensive multivariate dataset with features like vaccinations, number of ICU patients, and measure of government response stringency, spanning from January 2020 until September 2023, as of the present.
- (2) Identify and rank the features most predictive of COVID-19 case and death count predictions by models, while assessing different statistical metrics like correlation and SHAP values that can shed light on the relative degrees of contribution of different features.
- (3) Understand the temporal dynamics of these features and how their importance varies during different phases or outbreaks of the pandemic.

By answering these questions, we aim to not only make predictions more understandable but also contribute to enhancing trust and transparency in AI-driven epidemiological forecasts.

### 3 LITERATURE REVIEW

To contextualize our work on "Explainable Models for Epidemiological Forecasting," we will review relevant research papers that contribute to the field of epidemiological prediction and provide insights into the challenges and opportunities in this domain.

Solayman et al. [6] present a significant advancement in COVID-19 prediction by deploying machine learning models, including a hybrid CNN-LSTM approach and explainable AI techniques like the LIME framework. This work complements our original research as it addresses the critical need for transparent and interpretable models in understanding and predicting the dynamics of the COVID-19 pandemic. By demonstrating the effectiveness of the CNN-LSTM model in capturing the temporal aspects of symptom progression and applying explainable AI, this study provides valuable insights into feature interpretability—a crucial aspect of epidemiological forecasting.

The work by Devaraj et al. [1] underscores the potential of deep learning models, notably the Stacked Long Short-Term Memory (SLSTM) algorithm, in achieving highly accurate predictions of COVID-19 cases, with an error rate of under 2%. We intend to leverage these findings to enhance the explainability of our epidemiological forecasting models. Moreover, considering various variables, including temperature rainfall, demonstrates the models' ability to capture complex epidemiological dependencies, and we aim to draw inspiration from these insights by incorporating similar factors and relationships into our models to ensure they provide interpretable, data-driven forecasts for epidemiological scenarios.

The work by Mishra et al. [3] delves into advanced mathematical and deep learning models for short-term COVID-19 incidence and mortality forecasting. Mishra et al. expand the traditional SEIR model by including hospitalization, vaccination, and quarantine data, a concept pertinent to our goal of enhancing model explainability by considering a broader range of factors in epidemiological scenarios.

Mathonsi and van Zyl [2] introduce the Multivariate Exponential Smoothing Long Short-Term Memory (MES-LSTM) hybrid model, an extension of the ES-RNN model for multivariate time series forecasting. Their main objective is to enhance forecast accuracy and prediction interval construction across various applications, including COVID-19 morbidity datasets. While we emphasize improving model interpretability, Mathonsi and van Zyl's work aligns with the broader trend of enhancing forecasting accuracy, a crucial aspect of epidemiological modeling.

The work by Nikparvar et al. [4] employs a deep LSTM neural network for county-level COVID-19 forecasting, addressing challenges in accurate prediction during a dynamic pandemic with spatial variations. In contrast, our focus is on developing explainable epidemiological forecasting models, emphasizing the interpretation of predictions from various models, including deep learning. Nikparvar's insights into deep LSTM applications, mobility data, and multivariate time series can enrich our understanding of feature selection and model interpretability in deep learning for epidemiological forecasting.

The work by Rodríguez et al. [5] presents a comprehensive framework for real-time COVID-19 forecasting using deep learning techniques. This offers avenues to enhance our models' predictive capabilities, potentially leading to more precise disease spread forecasts. DeepCOVID's ability to handle noisy and heterogeneous data aligns with a core challenge in epidemiological forecasting. We can apply similar data preprocessing and cleaning strategies to improve our models' robustness with complex real-world data. Addressing data quality issues can provide more reliable insights into disease transmission dynamics.

### 4 PROPOSED APPROACH

#### 4.1 Dataset

**4.1.1 Source:** Our main dataset comes from "Our World in Data" (OWID). OWID is a popular source that's been collecting data about the COVID-19 pandemic since its start. Their dataset has a lot of information, not just about the number of cases or deaths, but also about how countries responded, how many people got tested or vaccinated, and other important factors. From January 2020 to September 2023, the dataset gives us a clear picture of how the pandemic changed over time. This dataset is publicly available through OWID's official GitHub repository.

**4.1.2 Features & Description:** The dataset from OWID is comprehensive, with categories ranging from epidemiological data to policy and stringency metrics. Specific features of note include:

- Epidemiological Data: Capturing the pandemic's progression with metrics like confirmed cases, deaths, ICU admissions, and reproduction rate.
- Testing Data: Revealing the testing landscape with total tests, new tests, and positivity rates.
- Vaccination Metrics: Chronicling the global vaccination effort, including doses administered and number of individuals vaccinated.
- Policy Responses: A reflection of governmental actions, such as school closures and travel restrictions.
- Demographic Data: Providing contextual understanding with metrics like population, median age, and GDP per capita.

**4.1.3 Data preprocessing & cleaning:** Given the vastness of our dataset and inherent challenges of real-world data collection, preprocessing will be very useful. Datasets were pruned to ensure consistent representation across regions, with certain territories excluded due to lack of continuous data. This section will serve to identify days with missing data, followed by suitable imputation strategies.

#### 4.2 Predictive Methods and Models

Given the nature of the problem at hand, which deals with time-series forecasting for COVID-19 cases and deaths, we leverage a suite of models that have shown promise in the domain of sequential data. Below is an elaboration on each:

- **Random Forests:** Random forests are an ensemble learning method that constructs multiple decision trees during training and outputs the mode of the classes or mean prediction

of individual trees for classification or regression, respectively. They are known to handle missing values and achieve higher accuracy by reducing overfitting.

- **Vector Autoregression (VAR):** VAR is a statistical model used to capture the linear interdependencies among multiple time series. It can be expressed as:

$$Y_t = c + A_1 Y_{t-1} + \dots + A_p Y_{t-p} + e_t$$

where  $Y_t$  is a vector of endogenous variables,  $c$  is a vector of constants,  $A_1, A_2, \dots, A_p$  are matrices of coefficients, and  $e_t$  is a vector of error terms.

- **LSTM (Long Short-Term Memory):** LSTM networks are a type of recurrent neural network (RNN) architecture that are adept at handling long-term dependencies in sequential data. They utilize memory cells to store and recall information, making them particularly suitable for time-series forecasting.
- **Prophet:** Developed by Facebook, Prophet is an open-source forecasting tool tailored for producing high-quality forecasts for time series with strong multiple seasonality patterns. It captures trends, seasonality, and holidays, providing intuitive parameters that can be easily tuned.
- **Transformers:** Transformers are a deep learning architecture introduced for natural language processing tasks but have shown proficiency in various domains, including time series. They rely on self-attention mechanisms to weigh the importance of different parts of the input data.

For hyperparameter tuning, we propose the use of **Bayesian Optimization**: Bayesian optimization is a model-based approach for finding the minimum of a function that is expensive to evaluate. By building a probabilistic model of the objective function and using it to select the most promising hyperparameters to evaluate in the true objective function, it efficiently reaches better results with fewer function evaluations. The Gaussian Process is often used as a probabilistic model in this technique.

To gauge the performance and predictive accuracy of our models, the dataset will be split temporally into training and validation sets, ensuring that the sequences are preserved. This will allow us to not only train our models but also assess their performance in different pandemic phases. The performance metrics are:

- **Mean Absolute Error (MAE):** Defined as:

$$MAE = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i|$$

where  $y_i$  are the actual values,  $\hat{y}_i$  are the predicted values, and  $n$  is the number of samples.

- **Root Mean Square Error (RMSE):** Defined as:

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

- **Mean Absolute Percentage Error (MAPE):** Defined as:

$$MAPE = \frac{100\%}{n} \sum_{i=1}^n \left| \frac{y_i - \hat{y}_i}{y_i} \right|$$

- **Mean Percentage Error (MPE):** Defined as:

$$MPE = \frac{100\%}{n} \sum_{i=1}^n \left( \frac{y_i - \hat{y}_i}{y_i} \right)$$

### 4.3 Predictive Feature Identification and Evaluation

Here, we delve into the techniques that help us gain insight into the importance of specific features in the prediction and signals of epidemic-characteristic case and death counts:

- **Correlation:** Correlation, a statistical measure, captures the strength and direction of the linear relationship between two variables. The most common metric, Pearson's correlation coefficient ( $r$ ), ranges from -1 to 1, with values close to 1 indicating a strong positive linear relationship and values close to -1 indicating a strong negative linear relationship. Given a feature,  $x_i$ , and our target variable,  $y_i$  (either case or death counts), the Pearson's correlation coefficient is calculated as:

$$r = \frac{\sum (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum (x_i - \bar{x})^2 \sum (y_i - \bar{y})^2}}$$

where  $\bar{x}, \bar{y}$  are the means of the feature and the target variable, respectively.

A high absolute value of  $r$  for a feature suggests that changes in the feature could be strongly associated with changes in our target.

- **SHAP (SHapley Additive exPlanations) Values:** SHAP values provide a unified measure of feature importance across different types of models. They decompose a prediction's deviation from the mean prediction into contributions from each feature.

For a given prediction  $p$  and a feature set  $S$ , the SHAP value is defined as:

$$\phi_j = \sum_{S \subseteq N \setminus \{j\}} \frac{|S|!(|N| - |S| - 1)!}{|N|!} [f(S \cup \{j\}) - f(S)]$$

where  $\phi_j$  is the SHAP value for feature  $j$ ,  $N$  is the set of all features,  $f(S)$  is the model's output for the feature set  $S$ . In simpler terms, for each feature, the SHAP value is the average contribution of that feature to every possible prediction, weighted by the likelihood of that set of features occurring. Features with higher absolute SHAP values are considered more influential for the model's predictions. By ranking features based on their SHAP values, we can determine which ones are the most predictive. Furthermore, SHAP values can be visualized using SHAP summary plots, offering insights into the effects of feature values on predictions. Positive SHAP values push predictions higher than the mean prediction, while negative values push predictions lower.

In conclusion, using both correlation and SHAP values, we will comprehensively evaluate the features' importance for predicting COVID-19 cases and deaths, shedding light on those that play pivotal roles in the models' forecasts.

## 5 EXPECTED OUTCOMES

By the end of the semester, we aim to achieve the following with respect to our "Explainable models for Epidemiological Forecasting" project:

- Feature Importance Identification: Establish a clear understanding of which features in the COVID-19 dataset are most predictive for forecasting cases and deaths.
- Model Development: Implement statistical and machine learning models like LSTM, VAR, transformers, ARIMA, SLSTM, and extended SEIR models. Compare their performance in terms of accuracy and forecast reliability.
- Explainability Integration: Incorporate Explainable AI (XAI) techniques into our models to ensure that the forecasts are not just accurate but also interpretable. This will assist stakeholders in understanding why certain predictions are made and build trust in the model's outputs.
- Phased Analysis: Analyze the importance of the identified features during different phases of the epidemic (beginning, peak, end/dying out, outbreaks of specific variants). We expect to have preliminary findings, at the very least, on how features contribute differently during these phases.
- Correlational Study: Conduct a comprehensive correlational analysis between identified significant features and the progression of the pandemic, ensuring we understand how certain factors directly influence outbreak dynamics.
- Extension to Outliers: If time permits, we'll explore the impact of sudden outbreaks and shifts in data distributions on feature importance and prediction accuracy.
- Documentation & Preliminary Report: Compile all findings, methodologies, and insights into a comprehensive report, detailing the project's progression, challenges faced, solutions implemented, and results obtained. This report will serve as a foundation for any further research or refinement in the future.

By achieving these objectives, we hope to not only enhance the accuracy of epidemic forecasts but also make them more interpretable and trustworthy, bridging the gap between complex data and actionable insights.

## 6 CONTRIBUTION

The completion of this project is a collective effort, with each member leading a specific aspect of the project. While roles are delineated for clarity, every member is integral to each phase, providing input, expertise, and support regardless of the primary lead.

### (1) Shayar:

- *Primary Responsibility:* Data Procurement, Cleaning, and Preprocessing.
  - Responsible for retrieving the required datasets, ensuring that the data is cleaned, applying any necessary preprocessing techniques (handling missing values, potential outliers, and data normalization or standardization) and aiding in the feature selection process.

### (2) Aryan:

- *Primary Responsibility:* Model Development and Training.
  - Aryan will take the lead in developing the deep learning models, initializing the parameters, and setting up

the training regimen. He will also oversee the feature selection process.

### (3) Aravind:

- *Primary Responsibility:* Evaluation Metrics Selection, Model Inference, and Correlational Study.
  - Aravind will determine evaluation metrics, lead the inference process, and undertake the correlational study for the extended analysis.

## 7 EXPECTED TIMELINE

The project milestone report is due on Nov 4th, 2023 and the final project report is due on Dec 2, 2023. We have created a rough expected timeline while keeping those important deadlines in mind.

### October 1st - October 14th:

- *Data Collection and Preliminary Analysis*
  - Download the complete dataset from the OWID GitHub repository.
  - Initial data cleaning, handling missing values, and basic statistical analysis.
  - Decide on the initial set of models.

### October 15th - October 28th:

- *Initial Model Implementation and Testing*

### October 29th - November 4th:

- *Model Refinement and Milestone Report Preparation*

### November 5th - November 18th:

- *Feedback Incorporation and Extension*
  - Incorporate Milestone report feedback.
  - Explore sudden outbreaks and shifts in data.
  - Synthesize findings from the correlational study.
  - Begin drafting the end-of-semester report.

### November 19th - November 25th:

- *Final Model Testing and Evaluation*

### November 26th - December 1st:

- *Report Refinement and Presentation Preparation*
  - Finalize the report.
  - Prepare a presentation.

### December 2nd - December 7th:

- *Final Project Submission and Presentation*

## REFERENCES

- [1] Jayanthi Devaraj, Rajvikram Madurai Elavarasan, Rishi Pugazhendhi, G.M. Shafiqullah, Sumathi Ganesan, Ajay Kaarthic Jeysree, Irfan Ahmad Khan, and Eklaas Hossain. 2021. Forecasting of COVID-19 cases using deep learning models: Is it reliable and practically significant? *Results in Physics* 21 (2021), 103817. <https://doi.org/10.1016/j.rinp.2021.103817>
- [2] Thabang Mathonsi and Terence L. van Zyl. 2022. A Statistics and Deep Learning Hybrid Method for Multivariate Time Series Forecasting and Mortality Modeling. *Forecasting* 4, 1 (2022), 1–25. <https://doi.org/10.3390/forecast4010001>
- [3] Suryanshi Mishra, Tinku Singh, Manish Kumar, and Satakshi. 2023. Multivariate time series short term forecasting using cumulative data of coronavirus. *Evolving Systems* (2023), 1–18.
- [4] B. Nikparvar, M.M. Rahman, F. Hatami, et al. 2021. Spatio-temporal prediction of the COVID-19 pandemic in US counties: modeling with a deep LSTM neural network. *Sci Rep* 11 (2021), 21715. <https://doi.org/10.1038/s41598-021-01119-3>
- [5] A. Rodriguez, A. Tabassum, J. Cui, J. Xie, J. Ho, P. Agarwal, B. Adhikari, and B. A. Prakash. 2021. DeepCOVID: An Operational Deep Learning-driven Framework for Explainable Real-time COVID-19 Forecasting. In *Proceedings of the AAAI Conference on Artificial Intelligence*, Vol. 35. 15393–15400. <https://doi.org/10.1609/aaai.v35i17.17808>

[6] Sanzida Solayman, Sk. Azmiara Aumi, Chand Sultana Mery, Muktadir Mubassir, and Riasat Khan. 2023. Automatic COVID-19 Prediction using Explainable Machine

Learning Techniques. *International Journal of Cognitive Computing in Engineering* 4 (2023), 36–46. <https://doi.org/10.1016/j.ijcce.2023.01.003>