IEEE Access
Multidisciplinary : Rapid Review : Open Access Journal

Received 2 September 2022, accepted 12 September 2022, date of publication 15 September 2022,
date of current version 22 September 2022.

Digital Object Identifier 10.1109/ACCESS.2022.3206790

RESEARCH ARTICLE

Using Epidemic Modeling, Machine Learning and Control Feedback Strategy for Policy Management of COVID-19

KARTIK NARAYAN[1], (Student Member, IEEE), HEENA RATHORE[2], (Senior Member, IEEE), AND FAYCAL ZNIDI[3], (Senior Member, IEEE)
[1]Department of Computer Science and Engineering, Indian Institute of Technology Jodhpur, Jodhpur 342037, India
[2]Department of Computer Science, Texas State University, San Marcos, TX 78666, USA
[3]Department of Electrical Engineering, Texas A&M University Texarkana, Texarkana, TX 75503, USA

Corresponding author: Heena Rathore (heena.rathore@ieee.org)

ABSTRACT Coronavirus disease (COVID-19) is one of the world's most challenging pandemics, affecting people around the world to a great extent. Previous studies investigating the COVID-19 pandemic forecast have either lacked generalization and scalability or lacked surveillance data. City administrators have also often relied heavily on open-loop, belief-based decision-making, preventing them from identifying and enforcing timely policies. In this paper, we conduct mathematical and numerical analyses based on closed-loop decisions for COVID-19. Combining epidemiological theories with machine learning models gives this study a more accurate prediction of COVID-19's growth, and suggests policies to regulate it. The Susceptible, Infectious, and Recovered (SIR) model was analyzed using a machine learning model to estimate the optimal constant parameters, which are the recovery and infection rates of the coupled nonlinear differential equations that govern the epidemic model. To modulate the optimized parameters that regulate pandemic suppression and mitigation, a systematically designed feedback-based strategy was implemented. We also used pulse width modulation to modify on-off signals in order to regulate policy enforcement according to established metrics, such as infection recovery ratios. It was possible to determine what type of policy should be implemented in the country, as well as how long it should be implemented. Using datasets from John Hopkins University for six countries, India, Iran, Italy, Germany, Japan, and the United States, we show that our 30-day prediction errors are almost less than 3%. Our model proposes a threshold mechanism for policy control that divides the policy implementation into seven states, for example, if Infection Recovery Ratio (IRR) >80, we suggest a complete lockdown, vs if 10<IRR<20, we suggest encouraging people to stay at home and organizations to work at 50% capacity. All countries which implemented a policy control strategy at an early stage were accurately predicted by our model. Furthermore, it was determined that the implementation of closed-loop strategies during a pandemic at different times effectively controlled the pandemic.

INDEX TERMS Covid-19, pandemic, policy management, SIR, machine learning, regression, control theory, transfer function.

I. INTRODUCTION

Human populations have been affected by communicable diseases since ancient times [1]. Two of the most ancient and deadly diseases of humanity, tuberculosis and malaria, ravaged Ancient Egypt for more than 5,000 years and are still a major health problem today [2]. In 2009, a new A/H1N1 influenza virus emerged, causing the first global pandemic in 40 years [3]. Within the first two months of the outbreak, the disease had spread to more than 70 countries

The associate editor coordinating the review of this manuscript and approving it for publication was Yiqi Liu.

This work is licensed under a Creative Commons Attribution 4.0 License. For more information, see https://creativecommons.org/licenses/by/4.0/

with more than 30,000 confirmed cases [4] thereby adversely impacting human lives. Despite two years of pandemic prevalence, people are still adapting to its after effects. As of May 28, 2022, there had been 8.5 million confirmed cases and more than 1 million deaths in the United States (US) alone [5]. COVID-19 is one of the many communicable diseases that threaten us to this day, despite all the technological advances. Also, the findings in [6] suggest that the resurgence of the virus, resulting in the second wave, may be attributable to an incongruous behavior on the part of Italian residents in following recommended health measures, despite government-run preventive programs. It is obvious that there is a pressing need to take action against global pandemic, which requires non-pharmaceutical interventions such as social distancing, testing, and contact tracing, as well as pharmaceutical research, which includes designing vaccines and studying the protein that can contain the virus [7].

In order to study the epidemics, a variety of mathematical models have been proposed, which includes studies on Susceptible, Infected, Recovered (SIR), Susceptible, Exposed, Infected, Recovered (SEIR), and others [8]. These study model to stabilize and diminish the rate of propagation of a communicable virus in the event of a pandemic outbreak. However, these models are simple, which makes it easy to calculate, it likely oversimplifies complex disease processes such as Covid-19 due to its simplicity [9]. Many data scientists and health specialists have faced the challenge of modeling the COVID-19 outbreak and related projections for the number of infected cases, deaths, and recovered cases [10]. Consequently, experts in epidemiological modeling have difficulty predicting COVID-19's future [11]. Furthermore, it is shown epidemiological modeling is not usually viewed from an engineering perspective, which is why we think public officials might greatly benefit from one and better manage this deadly disease. On the other hand, the use of Artificial Intelligence and Deep Learning methods [12] for modeling numerous types of epidemic communicable diseases in different application scenarios has increased in recent years, making it very difficult for specialists to determine which of these models to use for prediction [13].

In the light of the recent pandemic, techniques have been proposed for managing public health strategies [14], [15], thereby focusing on implementing control measures to keep the reproduction rate (R_0) [16] below 1. Mitigation, which focuses on slowing the pandemic spread, and suppression, which focuses on reversing growth, are the two basic approaches to control the spread of a pandemic [17]. This generally relies on keeping the pandemic (R_0) below a specific value and is typically done by enforcing policies such as social distancing. One technique for implementing social-distancing measures is an on-off approach, where city officials relax some restrictions when the number of new cases requiring intensive care is below a threshold, tightening them otherwise. However, these officials generally resort to an open loop, principle-based decision-making process [18], thereby struggling to identify the optimal policy at the right time. This can have the unintended consequences of causing economic activities to plummet, thereby putting millions of jobs at risk. The aim of this work is to assist city officials in making data-informed decisions to keep people safe, while sustaining economic activities during any pandemic outbreak.

This work addresses how effectively we can modulate and manage public health policies by developing hybrid models which are not just purely data-driven, but are also based on control feedback and time based monitoring of the country. Furthermore, how can we measure the pandemic contagiousness using novel threshold mechanism to control the pandemic. In our research, we explored how feedback can help stabilize and slow the progression of this deadly viral infection. Using engineering principles, we have come up with a practical approach to provide policymakers with concrete guidance, one that takes both medical and socioeconomic factors into account. We relied on feedback-based strategies to control the outbreak and manage the longer-term caseload effectively. The goal of this paper is to develop a fundamental understanding of interventions' impacts on the pandemic spread and their ability to forecast its progression. We present a novel data-based modeling approach that is equally efficient in controlling and suppressing the spread of viruses. In this study, we address the following challenges: (1) how to evaluate public health policies using models that are not only data-driven, but also based on control feedback strategy, and (2) how can we calculate the Infection Recovery Ratio (IRR) to assess the contagiousness of the pathogens as explained in [19]. (3) how can we analyze the best-fit model for predicting the COVID-19 outbreak using different analytical techniques. This work addresses these issues and provides the following contribution:

- In this study, a SIR model is employed to investigate COVID-19's behavior. The machine learning model is used to optimize both infection rates and recovery rates.
- The proposed IRR is used to guide city officials on which policies to implement at what stage.
- Additionally, we use a modulated control signal and a linear time-invariant transfer function to optimize the duty cycle and the width of the on-off signals for policy control. This model presents an enumeration of IRR values that is based on heuristics. Thus, the enumeration of values is categorized into various states corresponding to various policies. It is based on the pandemic control signal that the model suggests which policy and time-frame the government should follow.
- We also used polynomial regression analysis to forecast the values of the susceptible, infected and recovered data. The study helped in determining at which time period the control policy should be regulated in order to control the pandemic.
- The model was tested on six countries, namely India, Iran, Italy, Germany, Japan and the US. The results show that the model was accurate in predicting the number of actual cases for all countries except US and Japan. Furthermore, based on IRR threshold, policy control

signals can be applied to the countries for controlling the pandemic.

The paper is organized as follows: Section II describes the related work on epidemic modelling and pandemic control. Section III describes the building blocks of the proposed work. Section IV presents the performance analysis of the proposed work, and Section V summarizes the critical appraisal of the proposed model against the existing state of the art. This is followed by the limitations of the study and the conclusions described in Section VI and VII respectively.

II. RELATED WORK

There are numerous approaches to modeling a pandemic [20]; in general, models for epidemic forecasting can be classified into two broad categories: some approaches focus purely on mathematical modeling (communicable disease models) [21], while others use data analytics (data driven models) [22]. Moreover, highly referenced data analytics work relies exclusively on data obtained from centralized health monitoring agencies to provide high-level, summarized data analysis [23], [24].

The communicable disease models use classical differential equations to study the spread of disease. In these models, the population is divided into several categories and mathematical rules are applied to determine how individuals move between three main compartments, namely, they are susceptible to the virus, they are infected by the virus, or they have recovered from the virus. Epidemiological modeling has become commonplace in scientific, administrative, and social networks, and is forming the basis for everything from public policies (such as shutting down schools and small businesses) to personal decisions (such as where to travel). The SIR model [20] is a commonly cited mathematical model which calculates the rate of increase in the number of infections by multiplying the current infection rate by the current susceptible rate. Furthermore, the SIR model incorporates the effect of recovery, i.e., when an infected population becomes immune after a period of time. Researchers have also studied its effects in terms of preventing the spread of a communicable virus during a pandemic outbreak [20], [25]. Data-driven models, such as neural networks, broad sets of inputs, such as mobility, demographics, and medical capacity, are used to predict outputs, such as deaths and hospitalizations. In light of the recent pandemic, several strategies have been suggested for managing public health issues [14], [26], such as implementing control measures to keep the R_0 below 1 [15], [27]. The use of machine learning in pandemic situations is to improve the accuracy of prediction for both infectious and non-infectious disease screening [28]. Some of the recent machine learning approaches for pandemic detection and control utilizes supervised and unsupervised learning algorithms [29], [30], [31]. The main limitation is that most machine learning techniques utilize only supervised learning algorithms that use pre-acquired, labeled datasets with limited emphasis on the pandemic control and policy management.

It has been observed that most of the studies, concentrated mainly on either classical epidemiological models or machine learning models for COVID-19 pandemic predictions, both of which have limitations in generality and scalability, as well as a paucity of monitoring data [8]. Furthermore, it is particularly difficult to calculate mortality rates among reported cases (case fatality ratio) in the early stages of an epidemic. Therefore, these inaccuracies and biases can be carried over to the estimates of the impact of the public health measures that are being taken to contain COVID-19 in the community [32]. Researchers have also applied feedback-based control theory to control the number of infections by monitoring (R_0) in conjunction with the number of fatalities [18]. Although R_0 is a standard measure that can be utilized to measure the disease spread, it does not indicate the severity of infectious disease, nor does it indicate the rapidity of a pathogen's spread [33]. This paper uses a novel metric to measure the disease's contagiousness.

It is shown in [34], [35], and [36] that simple social distancing control actions are analyzed for controlling the impact of pandemic (more specifically the universal single interval social distancing is utilized). These previous studies presented different techniques for determining the optimal single interval control action, in order to minimize the infected peak prevalence rate which is defined as maximum proportion of infected individuals. Almost all of these proposals assume a full lockdown, a scenario that is somewhat unrealistic.

The aim of this study is to develop a hybrid communicable disease, data-driven, and control feedback theory strategy, a standard tool in control engineering, to address the limitations outlined above. Even though the aforementioned models are useful for predicting epidemic spread, they lack the granularity necessary to analyze individual behaviors during epidemics and analyze the relationship between individual decisions and epidemic spread. Therefore, such a high-level analysis is of limited use to city officials in adjusting public health policy guidelines [37].

III. PROPOSED WORK

In this section, we provide details of the proposed work with respect to the SIR Model, machine learning and the control feedback strategy.

A. EVALUATION OF THE SIR MODEL

Multiple mathematical models have been developed to examine the spread of infection. One of these models is the SIR model, which is composed of three coupled nonlinear ordinary differential equations. The model assumes that population size is constant and is divided into three parts: Susceptible (s), Infected (i) and Recovered (r). The SIR differential equations can be expressed as follows [38]:

$$\frac{ds}{dt} = -k_1 s(t) i(t) \tag{1}$$

$$\frac{di}{dt} = -k_2 i(t) + k_1 s(t) i(t) \tag{2}$$

TABLE 1. Metrics used in and obtained from SIR curves for different countries.

Country	Start Date	Initial S, I, R Values	Initial k_1, k_2 values	IRR	k_1	k_2	Time coordinate (intersection point of infected and recovered curve)	PW
Germany	03/22/20	200000, 1, 1	0.001, 0.001	116.64	0.1720	0.0015	-	50
India	03/25/20	3000000, 1, 1	0.001, 0.001	13.87	0.1031	0.0075	-	30
Iran	03/14/20	75000, 1, 1	0.001, 0.001	3.52	0.2533	0.0720	67.70724716	21
Italy	03/09/20	200000, 1, 1	0.001, 0.001	10.95	0.2100	0.0191	95.84883635	49
Japan	03/22/20	18000, 1, 1	0.001, 0.001	5.61	0.1452	0.0259	109.40949619	39
US	03/19/20	2000000, 1, 1	0.001, 0.001	147.58	0.1580	0.0011	-	50

$$\frac{dr}{dt} = k_2 i(t) \tag{3}$$

where $s(t)$ = susceptible population at time t, $i(t)$ = infected population at time t and $r(t)$ = recovered population at time t, k_1 = infection rate, and k_2 = recovery rate.

In these differential equations, the ratio of constant coefficients, k_1 (the infection rate of the pathogen) and k_2 (its recovery rate), will determine IRR which is k_1/k_2 as described in [19]. This paper employs a machine learning (ML) model as discussed in section III-B to predict the optimal values for k_1 and k_2, which are the constants of the SIR coupled nonlinear ordinary differential equation, in order to further analyze the SIR curves. By using these optimal values and the IRR value, we were able to control the Pulse Width (PW) parameter, outlined in section III-C1. This was accomplished by examining the intersection point between infected and recovered cases, followed by applying a transfer function to determine a cutoff-off frequency and then using that as an on-off rule signal. Here, policies are enforced when predicted infected patients are greater than predicted recovered patients, and lifted when vice versa, as described in section III-C1.

B. ML MODELS

We have used ML to predict SIR model and to optimize k_1 and k_2 values.

1) ML APPROACH FOR SIR

ML is applied to the SIR model as described aforementioned by fitting the SIR curves of a specific time period into the ML model. The usage information serves as training data and predicts the curve's behavior in the future. Using the predicted curve characteristics, the government can develop a policy that balances the pandemic with the nation's economic activities. The ML model uses a custom Loss Function (LF) based on the weighted sum of root mean squared errors of both active and recovered cases. The LF was developed using the active cases, which include both confirmed cases and deaths, along with recovered cases. This approach improves curve fitting optimization since it takes into account both active and recovered case errors. Here, the LF is defined as follows:

$$E_1 = \sqrt{\frac{(a'(t) - a(t))^2}{T}} \tag{4}$$

$$E_2 = \sqrt{\frac{(r'(t) - r(t))^2}{T}} \tag{5}$$

$$LF = \alpha E_1 + (1 - \alpha) E_2 \tag{6}$$

where E_1 = RMSE error of active cases over a time period T,
E_2 = RMSE error of recovered cases over a time period T,
$a(t)$ = number of actual active cases at time t,
$a'(t)$ = number of predicted active cases at time t,
$r(t)$ = number of actual recovered cases at time t,
$r'(t)$ = number of predicted recovered cases at time t,
T = total number of days in consideration,
α = constant which determines the weights of the total LF.
Specifically, we calculate the LF for training data spanning over T days, also referred to as time period T. In addition, t refers to the day of that particular time period.

ML is trained to minimize the LF to find optimal values of k_1 and k_2. To minimize the LF, the limited memory Broyden–Fletcher–Goldfarb–Shanno with bound-constrained optimization algorithm (L-BFGS-B) algorithm [39] is used with the minimize optimization function in the scikit-learn library [40]. Through the training of the ML model, it is possible to determine the optimal values for k_1 and k_2, which are the constants in the SIR model.Using SIR differential equations in conjunction with ML strategies, the model was trained to predict SIR curves with k_1 and k_2 being optimized after training.

SIR graphs include curves representing actual data, as well as curves representing S, I, and R curves, which are fitted curves obtained from training the model, along with a prediction for a near future date. In this study, the initial values of I and R are taken from real-world data, and the initial value of the susceptible population is calculated by using the appropriate ratio when compared with the number of cases in mainland China.

Table 1 shows that the default values for k_1 and k_2 which are set as 0.001. Observations have shown that when the initial values of the S, I, and R populations are changed, the curves change and take on a new shape and behavior. This behavior will be described in the results section. By using ML, the k_1 and k_2 parameters of the SIR model are optimized to find the point of intersection for infected and recovered cases. A simple, but elegant method for controlling the multiple input signals related to the various trigger points is to look at the intersection point between the patients who are infected and those who have recovered. These trigger points are denoted by phases, with phase (P_1) policies being enforced when predicted infected patients exceed predicted recovered patients and phase (P_2) policies being unenforced when predicted infected patients fall below predicted recovered patients. Feedback control can be used to modulate signals between phases, as discussed in section III-C.

2) REGRESSION ANALYSIS

The polynomial regression model [41] was evaluated as a tool for forecasting and predicting the pandemic confirmed, recovered, and infected data as represented by function of (x). The output of the polynomial regression function is as follows:

$$f(x) = c_0 + c_1x + c_2x^2 \ldots . c_nx^n \tag{7}$$

where n is the degree of the polynomial and c is a set of coefficients. We divided the confirmed, recovered, and infected compartments into training and testing phases, during which each compartment was trained independently with no correlation between the curves. Each of the polynomial curves can have different polynomial degrees, which are selected according to fitting of the training data, but each polynomial curve of a country may not necessarily have the same polynomial degree. In this study, we optimized each curve separately for the regression model, based on the LF, which is the root mean square error. The polynomial regression model was then used to predict the future time period of 10-20 days, with the regression graph featuring actual data curves in addition to predicted confirmed, infected, and recovered curves.

C. CONTROL FEEDBACK STRATEGY

The objective is to identify a model that determines what policies and time-frames the government should adopt. In order to attain this objective, we introduce the concept of switching strategy to effectively manage the transition between these phases utilizing control theory techniques [34]. A key part of our argument consists of three key points: specifically, we propose a modulating strategy for on-off policy, present an elegant method for gathering feedback based on the SIR model's predicted parameters, and offer a systematic method for controlling disease phase transitions.

The control feedback strategy acts as a policy control signal, whereby policies are initiated (enforced) when predicted infected patients exceed predicted recovered patients, and subsequently unenforced when the reverse is true. Since the IRR is governed by the ratios k_1, the infection rate, and k_2, the recovery rate, these values are analyzed to adjust the PW parameters of the on-off signal, outlined in section III-C1. In addition, we use linear time-invariant transfer function to optimize the duty cycle and the width of the on-off signals for policy control. This model presents an enumeration of IRR values that is based on heuristics. Thus, the enumeration of values is categorized into various states corresponding to various policies. It is based on the pandemic control signal that the model suggests which policy and time-frame the government should follow. In the following sections, we will discuss the concept of modulated control signal and the linear time-invariant transfer function.

1) CONTROL SIGNAL

The Pulse Width Modulation (PWM) technique is used to reproduce the amplitude of an analog input signal by generating pulses of variable width. With PWM phase control, the phases are driven by a series of "ON-OFF" pulses, while varying the duty cycle (a percentage of time when the output voltage is "ON" as compared to "OFF") while maintaining the frequency. A PWM signal's ON time can be adjusted (or modulated) as desired, since it is a digital, unipolar square wave signal. Generally, duty cycle (D) is defined as the ratio of ON time to signal period, that is, $D = M/T$, where M is the duty cycle time and T is the duration time. There is a range between 0 and 1 and D can also be expressed in percentage terms, i.e. from 0% to 100%. Control is achieved by varying the duty cycle of the control signal. Figure 1 shows the PWM signal with two basic time periods. Here, frequency (F) is the reciprocal of the duration time, F = 1/T with the standard unit of frequency represented as Hertz (Hz).

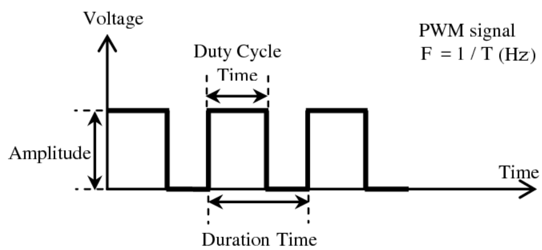


FIGURE 1. PWM signal with two basic time periods.

In the SIR model, the constant value determines the PW of the control signal, enabling the government to decide which policy to implement. In addition, the constants k_1 and k_2 of those differential equations are also used to calculate the IRR, which is defined as k_1/k_2. In this study, a square wave is used as the control signal, while a PWM signal controls when the signal is turned on or off. Digital signals can be either 0 or 1, so we used a signal value of 1 to represent complete lockdown, in other words, strict restriction policies must be implemented, and 0 to represent no restriction. Actions are defined as enforcement of a policy within a specific phase. Figure 2 illustrates the phase transitions.

Specifically, the PW is calculated as the difference in x coordinates between the first date in the training data and the intersection point of the SIR curves for infected and recovered patients. The start date is the first date in the training data that is considered in training the model. Thus, pandemic control can only be achieved if real cases follow the path of the predicted curves after this tipping point. It is often the case that the curves do not intersect in the near future due to the IRR being high or an epidemic ravaging the country. Therefore, instead of using a large PW in such cases, we use thresholds of IRR values to define the PW as follows:

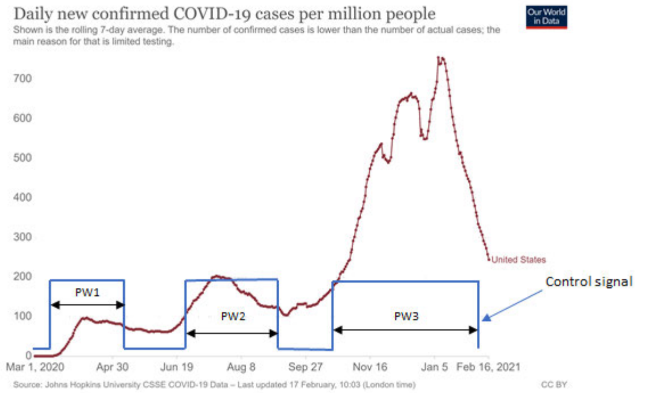


FIGURE 2. Definition of the phase repetition interval [Reproduced from [43]].

- A default PW of 50 days is applied when the infected and recovered cases do not intersect in the near future and IRR > 100.
- The default value is set at 40 days when $60 < \text{IRR} < 100$.
- The default value is 30 days when IRR < 60.

Following the control signal, the government then applies its policies in phases to the population. In the case of a simple square wave, it is suggested that a lockdown is implemented during the duty cycle width, along with strict policies for the on-off periods. The lockdown is abruptly imposed on the public without warning, just as the off policy is abruptly retracted since all restrictions are revoked at once [34]. Nevertheless, the goal is to implement restrictions gradually so that they have the least economic impact on the population and maintain a balance between public health and economic needs. In the next section, we will introduce the transfer function, along with the control signal, to present an enumeration of IRR values based on heuristics. Therefore, the enumeration of values is divided into a variety of states corresponding to various policies. It uses a pandemic control signal that passes through a transfer function controlled by IRR values to suggest what policies and timescales the government should follow.

2) TRANSFER FUNCTION

The control signal's output is modeled by a linear time invariant transfer function [42]. The input $x(t)$ is used to create a general linear time invariant system, and the output $y(t)$ is used to obtain the bilateral Laplace transform $x(t)$ and $y(t)$ [42]:

$$X(s) = \mathcal{L}(x(t)) = \int_{-\infty}^{\infty} x(t)e^{-st}dt \tag{8}$$

$$Y(s) = \mathcal{L}(y(t)) = \int_{-\infty}^{\infty} y(t)e^{-st}dt \tag{9}$$

where $X(s)$ and $Y(s)$ are the Laplace transform of $x(t)$ and $y(t)$, respectively, and s is a complex number that represents the frequency parameter. The transfer function $H(s)$ represents the relationship between the output signal $Y(s)$ of a control system and the input signal $X(s)$, for all possible input values, and is defined as follows [42]:

$$H(s) = \frac{Y(s)}{X(s)} \tag{10}$$

We will assume the system is a simple single-pole filter in this article, with the following characteristics [42]:

$$H(s) = \frac{\omega_0}{s + \omega_0} \tag{11}$$

With the above equations where ω_0 is proposed as:

$$\omega_0 = \frac{k_1 ln(IRR)}{2} \tag{12}$$

Modulated control signals can be passed through a transfer function, where the PW of the input signal and the pole of the transfer function are set by parameters of the SIR model. If only square waves are used, it is suggested that the public be locked down together with strict on-off policies, but instead, the lockdown is abruptly imposed without warning and the off policy is suddenly retracted since all restrictions are revoked at once. Instead, a control signal through a transfer function is used to allow a gradual implementation of policies.

Following the filtering phase, the control signal enters the state of imposition of the most restrictive measures that have the least economic impact on the population and maintain a balance between public health and economic needs. Restrictions are implemented progressively, depending on IRR and k_1. For instance, in a country with a high IRR imposing complete lockdown occurs more quickly, with fewer intermediate state of pandemic control. As the y-value changes within the PW from 0 to 1, indicating a specific policy should be implemented by the government. The gradual implementation of policies ensures that economic activities are not disrupted and stricter policies are only imposed when necessary. In Figure 3, phase transitions and state transitions are illustrated by using a modulated signal through a transfer function while considering the parameters of the SIR model.

Table 1 provides the data for feeding into the ML model for each of the six countries included in the study, as well as the IRR, k_1, and k_2 outputs of the ML model. The optimal values are obtained using the L-BFGS-B algorithm after a model

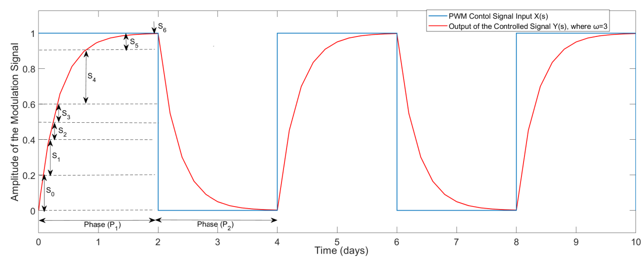


FIGURE 3. Phase transitions and state transition illustration where states (S_0-S_6) are defined in Section IV-C.

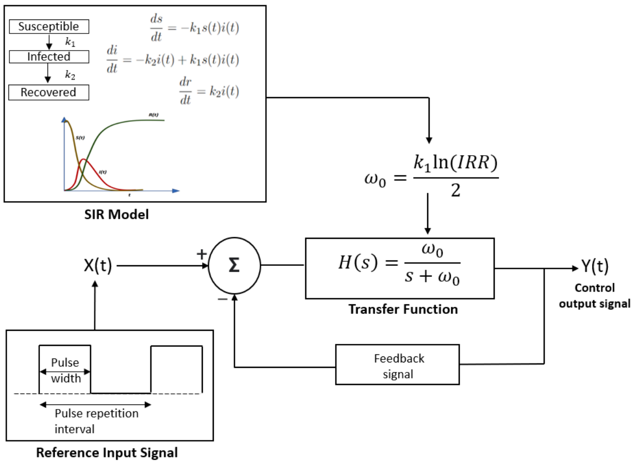


FIGURE 4. Closed loop control system for policy enforcement.

has been fitted using SIR differential equations. Additionally, it displays the x-coordinate of the point of intersection of each country's infected and recovered curve, as well as the PW of the control signal.

D. CONTROL FEEDBACK STRATEGY FOR POLICY MANAGEMENT

By coupling PWM and linear time-invariant transfer function, a control feedback strategy for policy management can be formed whereby specific policies can be applied gradually and turned on and off according to policy enforcement needs. The framework of the proposed work is shown in Figure 4. As seen from figure, the model has four components namely SIR model, input signal, feedback signal and transfer function. SIR model is used to predict the number of susceptible, infected, and recovered cases. Machine learning is then used to predict the optimal values of k_1 and k_2 from the SIR model. Predicted values of k_1 and k_2 values are later used to determine IRR values. IRR values determine which policies should be implemented and how aggressively they should be implemented. The poles of the linear time-invariant transfer function (ω_0) are then determined by the IRR values. The transfer function allows us to experiment with different policy control strategies based on predicted values of the SIR model at different orders of differentiation. PWM acts like "control knobs" phases, which turn on (enforce) policies when predicted infected patients exceed predicted recovered patients and turn off (will not enforce) policies when predicted recovered patients surpass predicted infected patients. PWM controls phase transitions based on a series of "ON-OFF" pulses while varying the duty cycle according to the intersection point between infected and recovered patients. To make the transition between these phases as smooth as possible, a control feedback strategy has been designed to manage the transition between these phases. By coupling the transfer function and PWM, restrictions can be implemented gradually to maintain the balance between public health and economic interests. Initial control signals are square waves, then PWM signals regulate when the signal is ON-OFF and consequently transfer function is used to implement the policy at different states depending on optimized IRR values. In this way, policy enforcement can be implemented gradually and turned on/off as necessary, rather than abruptly imposed on the public without warning during the phase.

IV. RESULTS

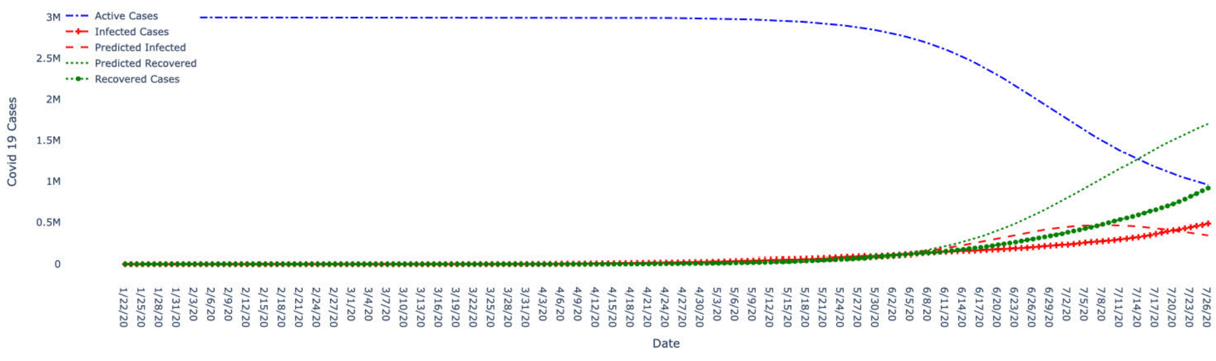
A. DATASET

SIR curves were plotted using dataset from John Hopkins University (JHU), which includes a time series of COVID-19 cases [43]. The dataset contains country-specific distributions of confirmed cases, deaths, and recovered cases, as well as day-specific counts for each category. The dataset is a time-series dataset which covers covid information for a duration of 510 days from 1/22/20 - 6/14/21. The instances of data or number of days used for training varies for each country and we match the training data duration in such a way that data before the onset of covid-spread in a country is used for training and the SIR curves are predicted for the next 30 days (which is the size of testing data). The instances used in training in days can be calculated by subtracting the start date with 22/1/20. Start date can be obtained from Table 1. For instance- training data size for Germany is 60 days = 3/22/20 - 1/22/20 and the number of days for which the curves are predicted (test data size) = 30 days.

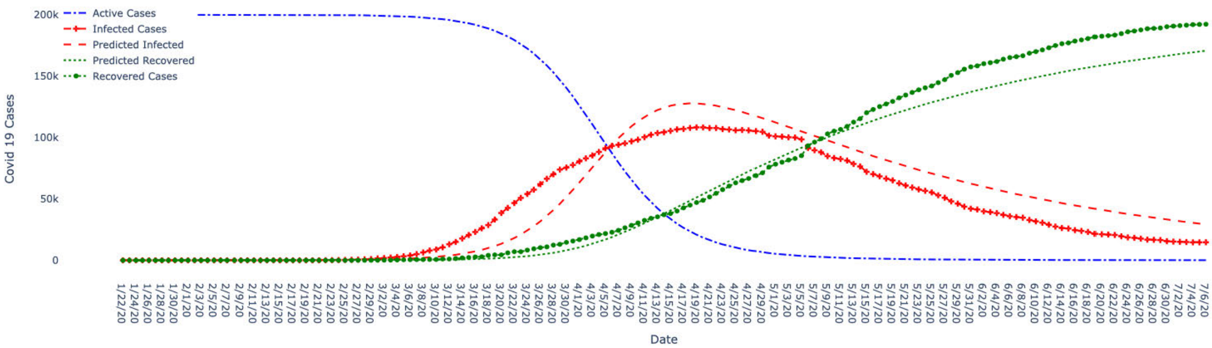
B. SIR CURVES ON ACTUAL DATA

To demonstrate the predictability of SIR curves, six countries were selected from around the world, including the US, India, Iran, Italy, Germany, and Japan. The COVID-19 pandemic curve is drawn from January 22, 2020 forward, with each country having its own demographics, social habits, and policies regarding the pandemic.

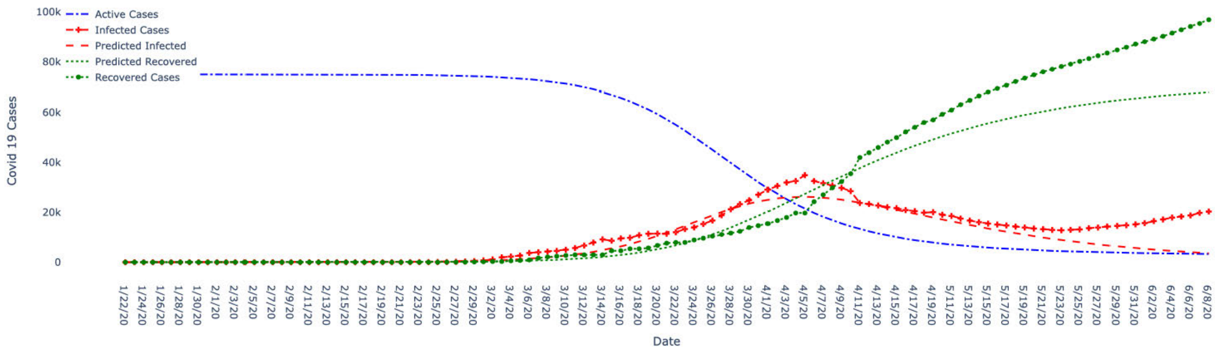
Figures 5(a)-(c) and Figures 6(a)-(c), show the SIR graphs for India, Italy, Iran, Germany, the US and Japan, respectively. Here, the blue line represents the number of active cases in the country. The red dashed line and red dashed rectangle represents the predicted behavior of the infected curve and the actual behavior of infected cases across the country respectively. The green dotted line and the green dotted circle represents the predicted behavior of recovered cases and actual recovered cases in the country respectively. Plotting SIR graphs of India, Italy, Iran, and Germany show the least variation between actual and predicted values. By contrast, Japan's and the U.S.'s 'actual' values do not match with the regression predictions. It is because these countries have implemented some restrictive measures to combat the spread of the disease. In the year 2020, the spread in the US was growing at a faster rate, and the IRR reached 150, which was alarming. Our hypothesis suggests that a nationwide lockdown could have prevented widespread disruptions caused by the COVID-19 pandemic. The plot shows that, in the US, imposing a lockdown would have decreased the



(a) SIR Curves for India



(b) SIR Cuves for Italy



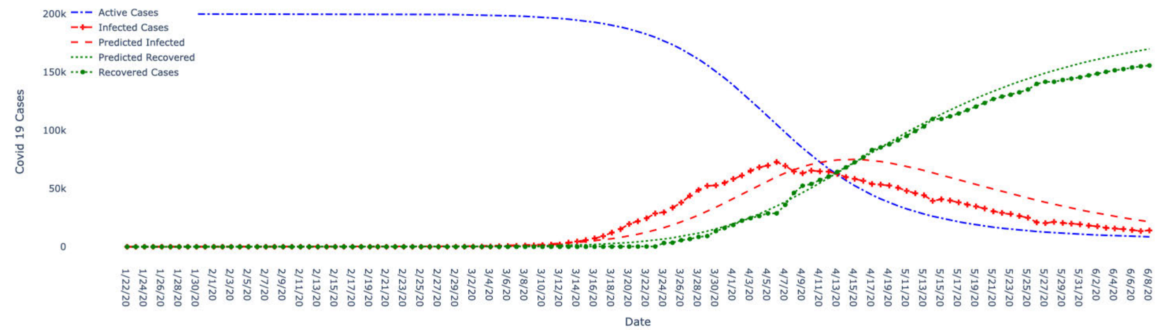
(c) SIR Curves for Iran

FIGURE 5. SIR curves and its prediction compared to actual data for (a) India (b) Italy, and (c) Iran. The blue line represents the number of active cases in the country. The red dashed line represents the predicted behavior of the infected curve. The red dashed rectangle represents the actual behavior of infected cases across the country. The green dotted line represents the predicted behavior of recovered cases. The green dotted circle represents actual recovered cases in the country.

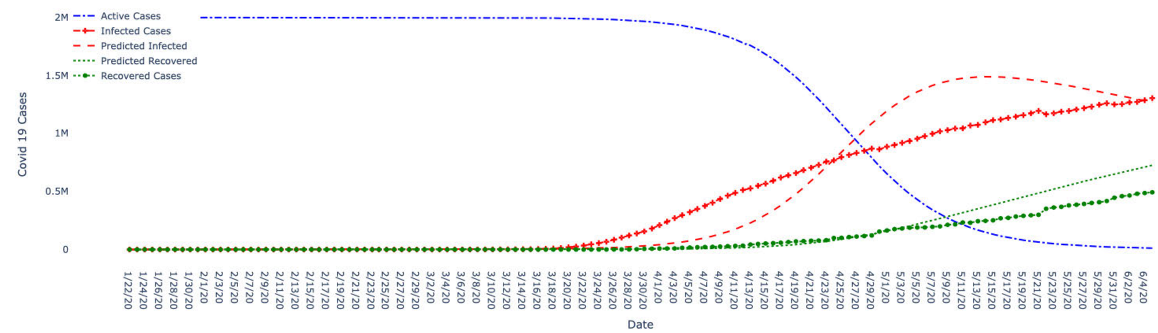
rate of infection, which was consistently increasing. Despite the regression model's prediction, the Japanese authorities were able to prevent worse conditions than those predicted. A major reason for Japan's success is the government's insistence on mask usage. Japan's excellent medical infrastructure, together with those measures, enabled it to control the pandemic early on, as shown in Figure 6(c).

C. POLICY SIGNAL GRAPHS

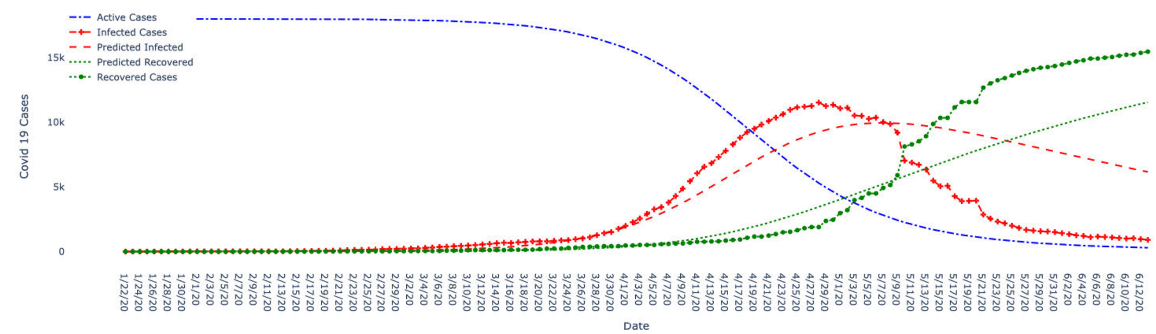
The policy signal graphs for the US, Japan, Italy, Iran, Germany, and India are shown in Figures 7(a)-(f) respectively. The black line represents the proposed policy control signal to be implemented by the country. The black policy control signal is the output signal of the transfer function when the red square wave is fed through as input. Policy signal graphs illustrate which policies should be implemented based on indicators like infection rate and IRR, as well as how decisions should change as the policy signal changes. Dashed lines on these graphs represent square wave signals with PWs as in Table 1, while solid lines represent the control signal passing through a single-pole filter with ω_0 as the pole. As the final control signal is not a square wave, it suggests gradual



(a) SIR Curves for Germany



(b) SIR Curves for US



(c) SIR Curves for Japan

FIGURE 6. SIR curves and its prediction compared to actual data in (a) Germany, (b) US, and (c) Japan. The blue line represents the number of active cases in the country. The red dashed line represents the predicted behavior of the infected curve. The red dashed rectangle represents the actual behavior of infected cases across the country. The green dotted line represents the predicted behavior of recovered cases. The green dotted circle represents actual recovered cases in the country.

implementation and lifting of all policy measures. It also suggests implementing restrictions in a phased or stepped manner.

Depending on IRR thresholds, policy implementation is divided into seven states represented by S. The policies and thresholds are outlined below.

1) S_0, if $0 <$IRR≤ 2: No restrictive policies should be implemented
2) S_1, if $2 <$IRR≤ 5: Wear mask and sanitize your hand regularly
3) S_2, if $5 <$IRR≤ 10: Practice social distancing at public places
4) S_3, if $10 <$IRR≤ 20: Encourage people to stay at home. Companies to work with 50% capacity
5) S_4, if $20 <$IRR≤ 50: Weekend lockdown policy should be implemented.
6) S_5, if $50 <$IRR≤ 80: Weekend lockdown policy should be implemented. Also, only essential services should be running in the country and for fixed time.
7) S_6, if IRR> 80: Complete lockdown.

The control signal's y-coordinates are used to determine which policy to adopt. If a country has a high IRR factor, such as the US, there is less time spent in S_0 to S_5 states and, therefore, reaching S_6 state is relatively faster than with

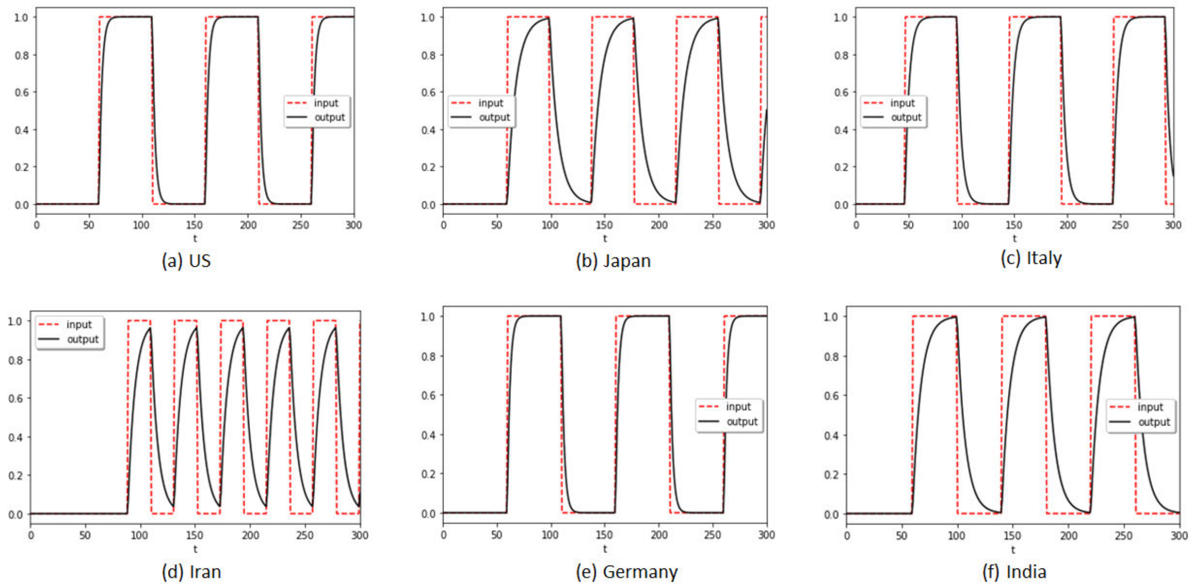


FIGURE 7. The predicted policy control signal to be implemented in (a) US, (b) Japan, (c) Italy, (d) Iran, (e) Germany, and (f) India. The dotted red signal is a square with PW as shown in Table 1. The black line represents the proposed policy control signal to be implemented by the country. The black policy control signal is the output signal of the transfer function when the red square wave is fed through as input.

TABLE 2. State and control signal relationship.

y-coordinate of the control signal	state of the country
0 to 0.2	S_0
0.2 to 0.4	S_1
0.4 to 0.5	S_2
0.5 to 0.6	S_3
0.6 to 0.9	S_4
0.9 to 1	S_5
1	S_6

respect to Japan, which has a low IRR factor. As seen from the PW of Japan, the majority part of the PW indicates the country falls into the state S_3 to S_5 state. In contrast to Iran, the control signal does not indicate a total lockdown within the next 30 days.

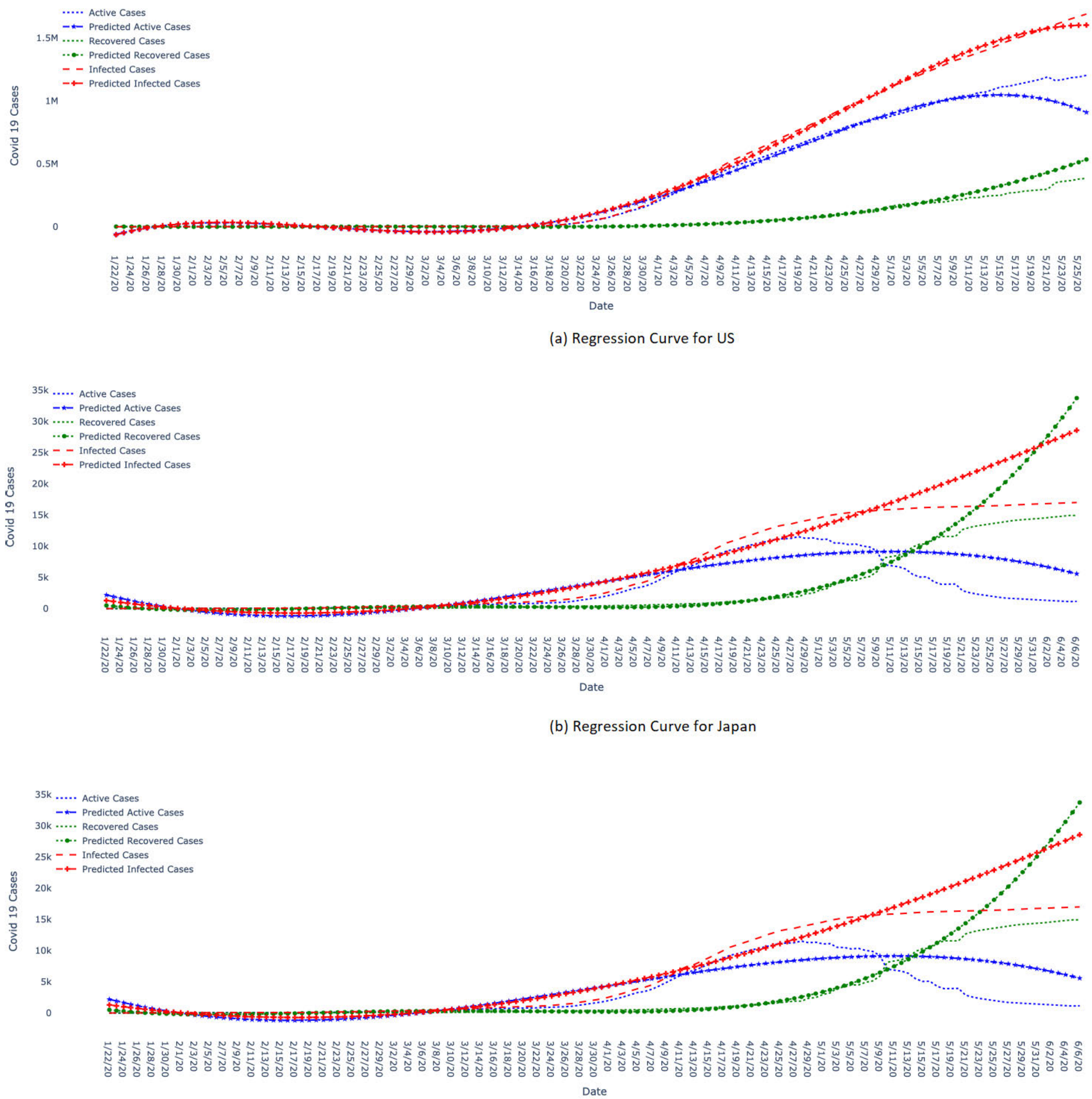
In Table 2, we show the y-coordinate of the control signal and the policy that a country needs to follow to attain a given state. Depending on the pandemic control signal's x-coordinate interval, the country determines how long it will take to implement a particular policy.

D. REGRESSION CURVES ON ACTUAL DATA

The polynomial regression model is utilized to plot and forecast active cases, infected cases, and recovered cases for each of the six countries. Figures 8(a)-(c) and 9(a)-(c) illustrate the regression curves for the US, Japan, Italy, Iran, Germany, and India. Here, starred dashed blue line and dotted blue line represents the predicted behaviour of active cases and the actual behaviour of active cases in the country respectively. Crossed dashed red line and dashed red line represents the predicted behaviour of infected curve and the actual behaviour of infected cases in the country respectively. Circled dashed green line and dashed green line represents the predicted behaviour of recovered curve and the actual behaviour of recovered cases in the country respectively. From the figures it can be seen that fitting the polynomial regression model to training data is more accurate in countries where the pandemic is just beginning or growing rapidly than in countries that do not exhibit those behaviors. When compared to India and the US, other countries where the number of infected cases is decreasing have more error while fitting the polynomial regression model and gives relatively poor results for the duration of the pandemic. In those countries applying the policy when rate of change of infection increases the recovery rate would be much beneficial rather than applying at a later stage. For instance, in Japan and Italy the policy could have started at a early month of March rather than waiting till the end of the month. We can also conclude that if the regression model is applied at the beginning of an outbreak it can give much better and accurate results and applying policies at those stages could effectively control the pandemic.

E. COMPARATIVE ANALYSIS BETWEEN SIR AND REGRESSION

Based on the SIR model and the regression model, we have modeled the COVID-19 pandemic. Based on our findings, SIR model is better at predicting curves over a long period of time than regression models, whereas regression models are better at predicting curves in the early phases of viral outbreaks. In the SIR model, the curves corresponding to susceptible, infected, and recovered are related



(a) Regression Curve for US

(b) Regression Curve for Japan

(c) Regression Curve for Italy

FIGURE 8. Regression curves and its prediction compared to actual data in (a) US, (b) Japan, and (c) Italy. Starred dashed blue line represents the predicted behaviour of active cases. Dotted blue line represents the actual behaviour of active cases in the country. Crossed dashed red line represents the predicted behaviour of infected curve. Dashed red line represents the actual behaviour of infected cases in the country. Circled dashed green line represents the predicted behaviour of recovered curve. Dashed green line represents the actual behaviour of recovered cases in the country.

since they are modeled using differential equations that are related. Thus, the constants of SIR differential equations determine their curve behavior. Since SIR differential equations and constants govern those equations, all curves in the SIR model are dependent on each other. On the other hand, the Active, Recovered, and Infected curves plotted using the regression method are independent of each other since they are not related by any mathematical condition. Consequently, since each curve has its own training set, the behavior of one doesn't affect the behavior of another.

In order to compare the models and their curve behavior, two countries were selected:

- *Case 1*: Italy: the selected period includes all the phases of the pandemic, and thus reveals the rise and fall of the infected curves (See Figure 8 (c)).
- *Case 2*: India: the selected time interval only includes the beginning of the pandemic, so the infected curve is only shown for the rising phase of the epidemic (See Figure 9 (c)).

In the second case, India's curves only correspond to the period around the outbreak of the pandemic. As seen in

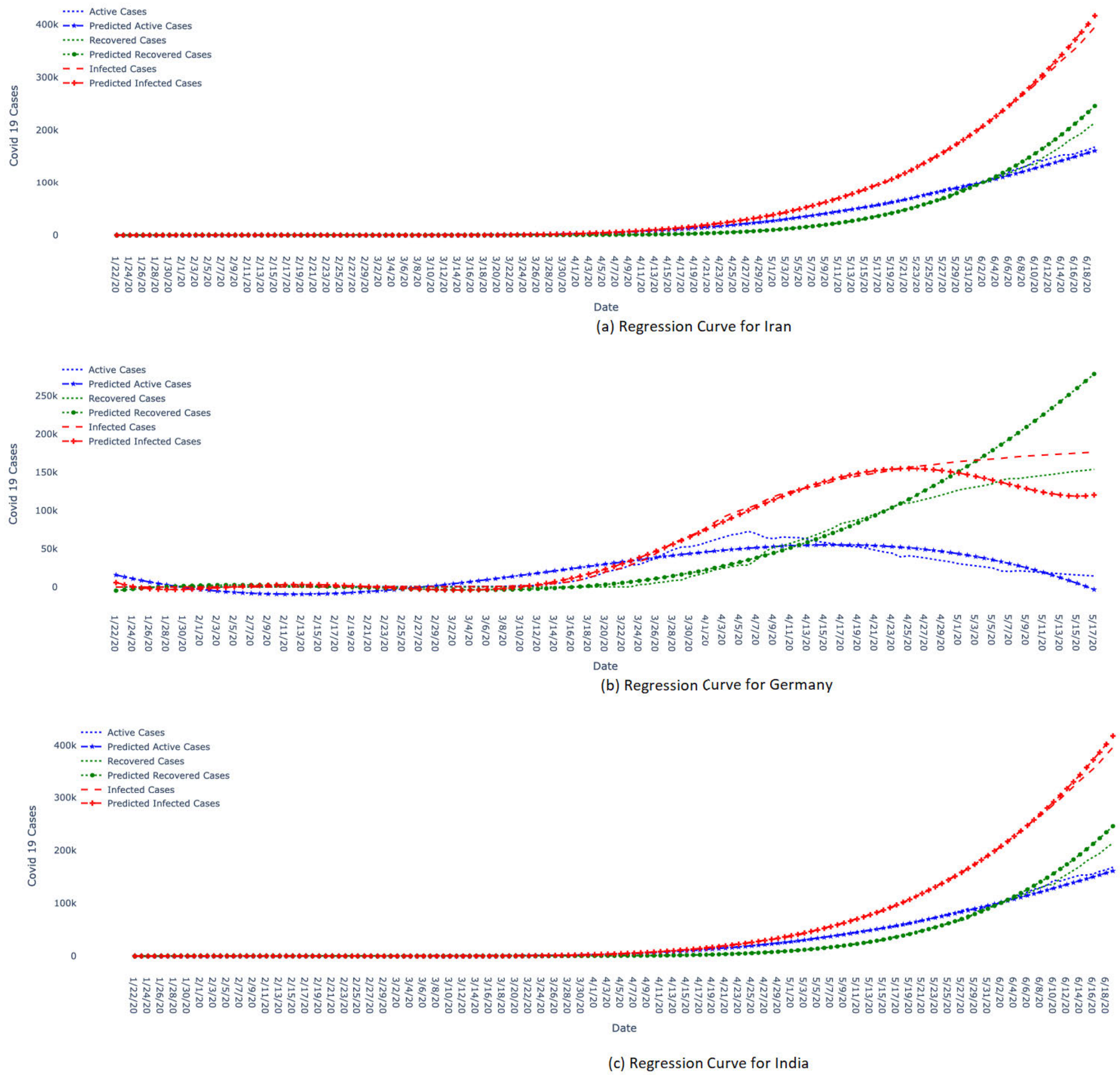


FIGURE 9. Regression curves and its prediction compared to actual data in (a) Iran, (b) Germany, and (c) India. Starred dashed blue line represents the predicted behaviour of active cases. Dotted blue line represents the actual behaviour of active cases in the country. Crossed dashed red line represents the predicted behaviour of infected curve. Dashed red line represents the actual behaviour of infected cases in the country. Circled dashed green line represents the predicted behaviour of recovered curve. Dashed green line represents the actual behaviour of recovered cases in the country.

Figure 9(c), the regression model performs better than the SIR model for India. The SIR model, in contrast, performs significantly better at predicting the downfall of recovered cases in Italy than does a regression model. From the simulation results, we can conclude that the SIR model performs well when a longer period of time is taken into consideration or when the time interval encompasses more phases of the pandemic.

V. COMPARATIVE ANALYSIS

In this section we provide the qualitative analysis of our work with the existing state of the art.

Epidemiological models such as the SIR model lack heterogeneity and can easily be programmed and analyzed [44]. There is also a shared opinion that SIR models are not completely appropriate due to their inflexibility [45]. Unlike existing SIR and SEIR models [46], our model provides the ability to predict the critical parameters, such as infection rate and recovery rate, so that it can adapt to changing policy requirements. We observe, for example, that if city-wide lockdown is applied at a particular instant, i.e., IRR thresholds, it can lower the transmission rate substantially in comparison to other epidemiological models. Our method incorporates the SIR model along with the feedback-based control signal and transfer function for pandemic control. For instance, in Japan and Italy the policy could have started early during the month of March rather than waiting till the end of the month. In mid-March, for example, the US has an IRR of $147.58 > 80$ (i.e., State S_6). In this case, our policy suggested complete lockdown to reduce the spread of COVID-19. However, in reality, the government started shutting down the public schools in New York city and Ohio started closing restaurants and bars [47]. The complete lockdown took a while to be imposed in those states of the US. Similarly for Germany, we computed IRR as $116.64 > 80$ (i.e., State S_6) which is a complete lockdown phase. However, Germany imposed a partial lockdown in late March [48].

In contrast, data-driven models require enormous amounts of data to model behavior and improve learning and are computationally expensive. ML models used for COVID-19 forecasting, as discussed in [49] and [50], do not efficiently exploit the fact that the data is time-series data. It does not include continuous feedback and does not take that into account when forecasting. Pandemic forecasts are subject to uncertainty caused by various unknown factors at the time. Our model which is based on feedback learning has the advantage of reducing the impact of uncertainty in models; as a result, if it is designed carefully, a strategy can still be effective despite inaccurate models.

A commonly cited approach [17], authors used on-off mechanisms to ease social-distancing measures, by keeping the pandemic R_0 below a certain threshold, where some restrictions are lifted as the number of new intensive care cases drops and are reinstated when it exceeds the threshold. Such an approach is more robust to uncertainties concerning both the R_0 and the severity of the virus, as well as more flexible and resilient than solutions that are based on fixed duration. The on-off mechanism is an example of a feedback system, in which the feedback variable is the number of intensive care unit patients. By contrast, the on-off approach implies that the lockdown is abruptly imposed on the public without warning, as the off policy is abruptly retracted when all restrictions are withdrawn at once. The major drawback to using this type of on-off control mechanism is that it can create oscillations, if overly too aggressively, overwhelm the healthcare system's ability to effectively treat serious conditions. On the other hand, our study uses a feedback mechanism coupled with a transfer function by enumerating IRR values using heuristics to gradually introduce restrictions so as to ensure that they have the least economic impact on the population and maintain a balance between health and prosperity. For example, in Italy, the IRR on March 9 was 10.95 (i.e., State S_3), where the policy suggested that the city officials encourage people to stay at home and businesses to operate at 50% capacity. However, in reality, due to the COVID-19 pandemic in Italy, the Italian government imposed a national lockdown or quarantine on 9 March 2020 that restricted all movements of the population except for necessity, work, and health reasons. As the lockdown is abruptly imposed on the public without warning, the healthcare system was overwhelmed, and this can lead to its inability to treat COVID-19. Nonetheless, as the policy suggested in [17] is only based on a on-off policy and does not take gradual implementation into account, it would recommend a full lockdown within the same period instead of the above measure. This is because the infection rate is higher than the recovery rate. As a result of gradual implementation, the health care system does not overload and is capable of effectively treating pandemic.

VI. LIMITATIONS OF THE STUDY

In the paper, data analysis has been conducted using Covid-19, 2020 datasets. In addition, the study examines threshold mechanisms for implementing state policies during pandemic outbreaks. The threshold values for these policies were determined with a limited dataset. We plan to use ML in the future to calculate and predict the threshold for implementing pandemic policies during outbreaks of pandemics.

VII. CONCLUSION

This paper utilizes SIR, feedback strategy based on control theory and regression models to study the behavior of COVID-19 pandemic. It was deduced that the SIR model is best suited for analyzing long-term trends in the spread of diseases, whereas the regression model provides better results during the outbreak phase than the SIR model. The results can be quantified using mean squared error between the predicted values and the actual values. At an early stage of the pandemic, it is evident from the results shown in the paper that there is a large difference between the predicted and the actual number of infected people curves. Based on the graphical result and the mean squared error, it shows that the SIR model cannot provide a useful early prediction of the epidemic in this case. However, we use the SIR parameters along with feedback based control theory in order to implement policy guidelines in the form of phases and states. Control feedback strategy helped to determine not only the type of policy that should be implemented in the country but also the length of the time it should be implemented for. The model was comparatively evaluated with regression analysis in understanding when and where the pandemic strategy should be employed. As a result of this study, the country officials would be able to control the pandemic and not impact the economy negatively. In countries such as India and the US, the SIR curves do not converge because the infection rate increases much faster than the recovery rate. In addition, it was found that the SIR model along with machine learning accurately computed the actual and predicted cases for all countries with the exception of Japan and US since these countries did not implement a policy control strategy at an early stage, whereas India, Iran, Italy, and Germany did. Furthermore, based on SIR parameters and feedback strategy based on control theory the states that we proposed can be effectively used by the public officials. It was found that if these state policies were implemented at the right time, the pandemic would have been controlled without affecting the economy of all the aforementioned countries.

For future work, we plan to use ML for predicting IRR values. Furthermore, the idea for the control feedback strategy for policy management can be tried in different use-cases for pandemic control, two of which are presented here. On account of pandemic spread, it is basic to know the spatio temporal feedback to investigate the effect of individuals relocating between different countries or clusters. Such movement presents vulnerabilities in the model which can be concentrated on utilizing the idea proposed in this paper. Besides, there are various COVID-19 episodes in which an infected individual probably infected 80% or a greater amount of individuals in room in only a couple hours [51], though in different cases it was considerably less infectious. Such an over-dispersive and super-spreading conduct of this infection brings imbalanced dissemination of cases [52]. This sort of situation, shifting back and forth between being infectious and noninfectious, caught by over-dispersion factor, can be concentrated on utilizing the idea proposed in this paper.

REFERENCES

[1] *Infectious Disease Emergence: Past, Present, and Future*, Institute of Medicine (U.S.) Forum on Microbial Threats. Microbial Evolution and Co-Adaptation: A Tribute to the Life and Scientific Legacies of Joshua Lederberg: Workshop Summary, National Academies Press (U.S.), Washington, DC, USA, 2009, p. 5.

[2] A. Lalremruata, "Molecular identification of falciparum malaria and human tuberculosis co-infections in mummies from the fayum depression (lower egypt)," *PLoS ONE*, vol. 8, no. 4, 2013, Art. no. e60307.

[3] S. Al Hajjar and K. McIntosh, "The first influenza pandemic of the 21st century," *Annals Saudi Med.*, vol. 30, no. 1, pp. 1–10, 2010.

[4] Y. Yang, "The transmissibility and control of pandemic influenza A (H1N1) virus," *Science*, vol. 326, no. 5953, pp. 729–733, 2009.

[5] *COVID-19 CoronaVirus Pandemic*, Worldometer, Worldometers, Geneva, Switzerland, Oct. 2020.

[6] L. Casini and M. Roccetti, "A cross-regional analysis of the COVID-19 spread during the 2020 Italian vacation period: Results from three computational models are compared," *Sensors*, vol. 20, no. 24, p. 7319, 2020.

[7] J. Viana, "Controlling the pandemic during the SARS-CoV-2 vaccination rollout," *Nature Commun.*, vol. 12, no. 1, pp. 1–15, 2021.

[8] H. Wang, "Predicting the epidemics trend of COVID-19 using epidemiological-based generative adversarial networks," *IEEE J. Sel. Topics Signal Process.*, vol. 16, no. 2, pp. 276–288, Feb. 2022.

[9] J. Tolles and T. Luong, "Modeling epidemics with compartmental models," *J. Amer. Med. Assoc.*, vol. 323, no. 24, pp. 2515–2516, 2020.

[10] P. Khrapov and A. Loginova, "Mathematical modelling of the dynamics of the Coronavirus COVID-19 epidemic development in China," *Int. J. Open Inf. Technol.*, vol. 8, no. 4, pp. 13–16, 2020.

[11] J. E. Gnanvi, K. V. Salako, G. B. Kotanmi, and R. G. Kakaï, "On the reliability of predictions on COVID-19 dynamics: A systematic and critical review of modelling techniques," *Infectious Disease Model.*, vol. 6, pp. 258–272, Jan. 2021.

[12] F. Shahid, A. Zameer, and M. Muneeb, "Predictions for COVID-19 with deep learning models of LSTM, GRU and Bi-LSTM," *Chaos, Solitons Fractals*, vol. 140, pp. 110–212, Nov. 2020.

[13] R. Gupta, G. Pandey, and S. K. Pal, "Comparative analysis of epidemiological models for COVID-19 pandemic predictions," *Biostatist. Epidemiol.*, vol. 5, no. 1, pp. 1–23, 2021.

[14] C. R. MacIntyre, "Case isolation, contact tracing, and physical distancing are pillars of COVID-19 pandemic control, not optional choices," *Lancet Infectious Diseases*, vol. 20, no. 10, p. 1105, 2020.

[15] S. C. Kamerlin and P. M. Kasson, "Managing coronavirus disease 2019 spread with voluntary public health measures: Sweden as a case study for pandemic control," *Clin. Infectious Diseases*, vol. 71, no. 12, pp. 3174–3181, 2020.

[16] J. H. Jones, "Notes on R0," *California: Dept. Anthropolog. Sci.*, vol. 323, pp. 1–19, May 2007.

[17] N. M. Ferguson, "Impact of non-pharmaceutical interventions (NPIs) to reduce COVID-19 mortality and healthcare demand," Imperical College London, London, U.K., Tech. Rep. 9, 2020.

[18] G. Stewart, K. Heusden, and G. A. Dumont, "How control theory can help us control COVID-19," *IEEE Spectr.*, vol. 57, no. 6, pp. 22–29, Jun. 2020.

[19] H. Rathore, S. Ranwa, and A. Samant, "Modular network effects on communicable disease models," in *Proc. 6th Asia Modeling Symp.*, 2012, pp. 126–131.

[20] R. Goel and R. Sharma, "Mobility based SIR model for pandemics-with case study of COVID-19," in *Proc. IEEE/ACM Int. Conf. Adv. Social Netw. Anal. Mining (ASONAM)*, Dec. 2020, pp. 110–117.

[21] L. James, "The use and misuse of mathematical modeling for infectious disease policymaking: Lessons for the COVID-19 pandemic," *Med. Decis. Making*, vol. 41, no. 4, pp. 379–385, 2021.

[22] E. Kuhl, "Data-driven modeling of COVID-19-Lessons learned," *Extreme Mech. Lett.*, vol. 40, Oct. 2020, Art. no. 100921.

[23] E. Dong, H. Du, and L. Gardner, "An interactive web-based dashboard to track COVID-19 in real time," *Lancet Infect Dis*, vol. 20, no. 5, pp. 533–534, 2020.

[24] S. Shilo, H. Rossman, and E. Segal, "Axes of a revolution: Challenges and promises of big data in healthcare," *Nature Med.*, vol. 26, no. 1, pp. 29–38, 2020.

[25] D. Acemoglu, V. Chernozhukov, I. Werning, and M. D. Whinston, *A Multi-Risk SIR Model With Optimally Targeted Lockdown (No. w27102)*. Cambridge, MA: National Bureau of Economic Research, 2020.

[26] C. Liu, "A microscopic epidemic model and pandemic prediction using multi-agent reinforcement learning," 2020, *arXiv:2004.12959*.

[27] R. Goel and R. Sharma, "Mobility based SIR model for pandemics-with case study of COVID-19," in *Proc. IEEE/ACM Int. Conf. Adv. Social Netw. Anal. Mining (ASONAM)*, Dec. 2020, pp. 110–117.

[28] S. Tiwari, P. Chanak, and S. K. Singh, "A review of the machine learning algorithms for COVID-19 case analysis," *IEEE Trans. Artif. Intell.*, early access, Jan. 11, 2022, doi: 10.1109/TAI.2022.3142241.

[29] A. Dairi, F. Harrou, and Y. Sun, "Deep generative learning-based 1-SVM detectors for unsupervised COVID-19 infection detection using blood tests," in *IEEE Trans. Instrum. Meas.*, vol. 71, pp. 1–11, 2022.

[30] H. Wang, G. Tao, J. Ma, S. Jia, L. Chi, H. Yang, Z. Zhao, and J. Tao, "Predicting the epidemics trend of COVID-19 using epidemiological-based generative adversarial networks," *IEEE J. Sel. Topics Signal Process.*, vol. 16, no. 2, pp. 276–288, Feb. 2022.

[31] O. Semenov, E. Agu, K. Pahlavan, and Z. Su, "COVID-19 social distance proximity estimation using machine learning analyses of smartphone sensor data," *IEEE Sensors J.*, vol. 22, no. 10, pp. 9568–9579, May 2022.

[32] A. Vespignani, H. Tian, C. Dye, J. O. Lloyd-Smith, R. M. Eggo, M. Shrestha, S. V. Scarpino, B. Gutierrez, M. U. Kraemer, J. Wu, and K. Leung, "Modelling COVID-19," *Nature Rev. Phys.*, vol. 2, no. 6, pp. 279–281, 2020.

[33] P. L. Delamater, E. J. Street, T. F. Leslie, Y. T. Yang, and K. H. Jacobsen, "Complexity of the basic reproduction number (R0)," *Emerg. Infect. Dis.*, vol. 25, no. 1, pp. 1–4, Jan. 2019.

[34] M. Sadeghi, J. M. Greene, and E. Sontag, "Universal features of epidemic models under social distancing guidelines," *Annu. Rev. control*, vol. 51, pp. 426–440, Jan. 2021.

[35] S. Federic and G. Ferrari, "Taming the spread of an epidemic by lockdown policies," *J. Math. Econ.*, vol. 93, Mar. 2021, Art. no. 102453.

[36] D. H. Morris, "Optimal, near-optimal, and robust epidemic control," *Commun. Phys.*, vol. 4, no. 1, pp. 1–8, 2021.

[37] K. Leung, "First-wave COVID-19 transmissibility and severity in China outside Hubei after control measures, and second-wave scenario planning: A modelling impact assessment," *Lancet*, vol. 395, no. 10233, pp. 1382–1393, 2020.

[38] C. Ji and D. Jiang, "Threshold behaviour of a stochastic SIR model," *Appl. Math. Model.*, vol. 38, nos. 21–22, pp. 5067–5079, 2014.

[39] J. L. Morales and J. Nocedal, "Remark on algorithm 778: L-BFGS-B: Fortran subroutines for large-scale bound constrained optimization," *ACM Trans. Math. Softw.*, vol. 38, no. 1, pp. 1–4, 2011.

[40] F. Pedregosa, "Scikit-learn: Machine learning in Python," *J. Mach. Learn. Res.*, vol. 12, pp. 2825–2830, Oct. 2011.

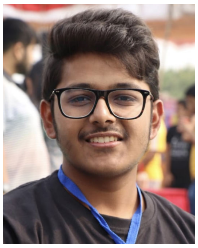
[41] R. M. Heiberger and E. Neuwirth, "Polynomial regression," in *R Through Excel*. New York, NY, USA: Springer, 2009, pp. 269–284.

[42] A. V. Oppenheim, J. R. Buck, and R. W. Schafer, *Discrete-Time Signal Processing*. Upper Saddle River, NJ, USA: Prentice-Hall, 2001.

[43] (2021). *John Hopkins Data*. [Online]. Available: https://coronavirus.jhu.edu/map.html

[44] A. Adiga, "Mathematical models for COVID-19 pandemic: A comparative analysis," *J. Indian Inst. Sci.*, vol. 100, no. 4, pp. 793–807, 2020.

[45] C. Gourieroux and Y. Lu, "SIR model with stochastic transmission," 2020, *arXiv:2011.07816*.
[46] A. Kaddar, A. Abta, and H. T. Alaoui, "A comparison of delayed SIR and SEIR epidemic models," *Nonlinear Anal., Model. Control*, vol. 16, no. 2, pp. 181–190, 2011.
[47] *CDC Museum COVID-19 Timeline*, Centers for Disease Control and Prevention, (CDC), Atlanta, GA, USA, 2021.
[48] R. Bosen and J. Thurau, "COVID: How Germany battles the pandemic—A chronology," DW Made for Minds, 2021. [Online]. Available: https://p.dw.com/p/3vTRN
[49] Y. Alali, F. Harrou, and Y. Sun, "A proficient approach to forecast COVID-19 spread via optimized dynamic machine learning models," *Sci. Rep.*, vol. 12, pp. 1–20, Feb. 2022.
[50] R. K. Mojjada, A. Yadav, A. V. Prabhu, and Y. Natarajan, "Machine learning models for COVID-19 future forecasting," *Mater. Today, Proc.*, 2020, doi: 10.1016/j.matpr.2020.10.962.
[51] M. Hutson, "The mess behind the models: Too many of the COVID-19 models led policymakers astray. Here's how tomorrow's models will get it right," *IEEE Spectr.*, vol. 57, no. 10, pp. 30–35, Sep. 2020.
[52] Z. Tufekci, "This overlooked variable is the key to the pandemic," The Atlantic, vol. 30, 2020.



KARTIK NARAYAN (Student Member, IEEE) is currently pursuing the bachelor's degree with the Department of Computer Science and Engineering, Indian Institute of Technology Jodhpur, Jodhpur. His research interests include machine learning, artificial intelligence, computer vision, and deep learning.

HEENA RATHORE (Senior Member, IEEE) received the Ph.D. degree in computer science and engineering from the Indian Institute of Technology Jodhpur, India. She is currently an Assistant Professor at Texas State University. Academically, she has worked as an Assistant Professor in instruction at The University of Texas at San Antonio, and a Visiting Assistant Professor at Texas A&M University Texarkana, Texarkana. She also worked as a Data Scientist and the Program Manager at Hiller Measurements, Austin, TX, USA. Prior to that, she worked as a Postdoctoral Researcher for the US-Qatar Joint Collaborative Project between Temple University, University of Idaho, and Qatar University. She did her Ph.D. studies on a Tata Consultancy Services Research Scholarship. Her research interests include cognitive AI, cybersecurity of cyber physical systems, and biologically inspired systems.

FAYCAL ZNIDI (Senior Member, IEEE) received the B.S. and M.S. degrees in electrical engineering from The University of Tennessee, USA, in 2009 and 2010, respectively, and the Ph.D. degree in electrical engineering from the University of Arkansas, USA, in 2019. He is currently an Assistant Professor at Texas A&M University Texarkana, USA. He has also contributed to several industry projects during his professional career at Tunisian Company of Electricity and Gas (STEG). His research interests include operations research, optimization in power networks, power system dynamics and controls, power system protection, microgrid operation, and machine learning in power systems.