Determining COVID-19 dynamics using Physics Informed Neural Networks

Simanga Gwebu [1] Joseph Malinzi [1,2] Sandile Motsa [1]

[1] Faculty of Science and Engineering, Department of Mathematics, University of Eswatini, Private bag 4, Kwaluseni, Swaziland

[2]Institute of Science, Durban University of Technology, Durban 4000, South Africa

Abstract

The Physics Informed Neural Networks framework is applied to the understanding of the dynamics of Coronavirus of 2019. To provide the governing system of equations used by the framework, the Susceptible-Infected-Recovered-Death mathematical model is used. The study focused on finding the patterns of the dynamics of the disease which involves predicting the infection rate, recovery rate and death rate; thus predicting the active infections, total recovered, susceptible and deceased at any required time. The study used data that was collected on the dynamics of COVID-19 from the Kingdom of Eswatini between March 2020 and September 2021. The obtained results showed less errors thus making highly accurate predictions.

1 Introduction and Background

1.1 Introduction

Coronavirus disease of 2019 (COVID-19) is a diseases caused by the Severe Acute Respiratory Syndrome Coronavirus 2 (SARS-CoV-2) virus [4]. It has spread and caused havoc at a global scale and was officially declared a pandemic by the World Health Organization (WHO) in 11 March 2020 [5, 6]. This virus is a member of the Beta coronavirus family which makes it highly

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 30 November 2021 doi:10.20944/preprints202111.0555.v1

likely to cause severe symptoms and fatal [8]. Due to the massive impact of the viral disease there is a agent need to understand its dynamics [3, 5, 12–14].

The spread of diseases like COVID-19 can be modelled using systems of ordinary differential equations (ODEs). Data analysis methods have also been used to understand the spreading patterns of COVID-19 including the machine learning (ML) approach [3, 15]. The viability of using artificial intelligence (AI) methods to solve ODEs has been disputed since these equations are governed by scientific laws which are never instilled during the training process [1]. A new framework called the Physics Informed Neural Networks (PINNS) has been developed and aims to eliminate this problem. It also achieves high prediction accuracy from small sized datasets [2].

The Physics Informed Neural Network framework has been applied to multiple studies of COVID-19. One study applied PINNS to estimate the spread of COVID-19, which also considers the aid of quarantine controls. The study employed the Susceptible - Exposed - Infected - Removed as the governing systems of equations. It mainly tries to understand the benefits of the implementations of COVID-19 restrictions [12]. Another study was conducted where the parameters were time-varying, it employed the Susceptible-Infected-Recovered-Deceased (SIRD) model as the governing systems of equations. This model however was developed using a recurrent neural network [3].

The PINNS is a Artificial Neural Network (ANN) framework which during the training process exposes the developed neural network to data sets and governing laws. The governing laws are provided in the form of ODEs or PDEs to the model [1, 2, 35]. Understanding the dynamics of COVID-19 is crucial in reducing the effects of the virus. AI models have been developed, but most of these models tend to require a lot of training data to achieve high accuracy. This is however not possible for cases like the newly discovered COVID-19 since it has a small data set. Other models only fit the given data rendering them less accurate in making future predictions. This brings about the need for development of models that can use small size data to make accurate predictions on the dynamics of the diseases spread.

The aim of the study is to use the Physics Informed Neural Networks framework to determine the dynamics of COVID-19. The PINNS mathematical model used is the study is the SIRD model. The study focuses on determining the average rates of the virus' contraction, recovery and deaths. It

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 30 November 2021 doi:10.20944/preprints202111.0555.v1

uses it to determine the active infections, total recovered, susceptible and deceased at any required time. The study uses data from the Kingdom of Eswatini obtained between March 2020 and September 2021 to train the neural network.

Outline of the study The remainder of the document has four sections, the first is the literature review section. This section contains a review of some studies pertaining the use of the PINNS. The following section contains the methodology section which contains the analysis of the mathematical and physics informed neural network framework. The section that follows is the results and simulations section, which contains the results and errors obtained and the analysis. There is also the conclusion with recommendations of future studies.

1.2 Background

Artificial Neural Networks are the basis of deep learning, a branch of AI and ML [17]. They are computational models created as an attempt to harness the capabilities of both the human brains and computers [2, 21]. A common structure of ANNS consists of nodes arranged in layer format joined by connectors [22]. The first layer is called the input layer, it receives data in vector format and passes a dot product of the connector weight and received data to adjacent node [23]. The dot product in the nodes is multiplied by the activation function [24]. The activation function is a mathematical function which changes the input values to a non-linear format. This process is called the feedforward process [23]. The process which takes the error and adjusts the weights during the training process is called backpropagation.

Definition 1. *A feedforward neural network with a total of N neurons arranged in a single layer is a function $y : \mathbb{R}^d \to \mathbb{R}$ of the form*

$$y(t) = \sum_{i=1}^{N} \alpha_i \sigma(w_i^T + b_i),$$

where $t \in \mathbb{R}^d, \alpha_i, b_i \in \mathbb{R}$. σ is the activation function, w_i are weights for each neuron multiplied to input value t. α_i are neural network weights and are applied to the output of each neuron in the layer and b_i is the bias of each neuron.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 30 November 2021 doi:10.20944/preprints202111.0555.v1

There are numerous activation functions used in neural networks. The study employees the tangent hyperbolic function (tanh).

Definition 2. *A tangent hyperbolic activation function is a function* $\sigma : \mathbb{R} \to \mathbb{R}$ *such that*

$$\sigma(t) \to \begin{cases} 1 & as \quad t \to \infty, \\ -1 & as \quad t \to -\infty. \end{cases}$$

A key feature of neural networks is that during training they adjust the internal values such that they can solve any given problem to some degree of accuracy. By definition neural networks are discriminatory functions [22]. This property thus make the neural network have suitable properties to be a universal approximator.

Theorem 1.1. *If the* σ *in the neural network definition is a continuous, then the set of all neural networks is dense in a space of continuous discriminatory functions function with domain* C *on* I_n *denoted by* $C(I_n)$*. Where* I_n *is an n-dimensional unit cube.*

Proof. Let $\mathcal{N} \subset C(I_n)$ be the set of neural networks. Where $\mathcal{N}$ is a linear subspace of $C(I_n)$. To show that $\mathcal{N}$ is dense in $C(I_n)$, we show that its closure is $C(I_n)$. By contradiction, suppose $\overline{\mathcal{N}} \neq C(I_n)$. Then $\overline{\mathcal{N}}$ is a closed proper subspace of $C(I_n)$. □

Two main approaches used to make predictions are the mathematical modelling approach and the data based approach. These two models have different advantages and disadvantages. Mathematical models used in the estimation of any process are mainly derived from underlying processes [1]. These models thus conform to governing laws, this gives them a guided output which, if given the correct initial values always provides accurate answers. The key disadvantages though is that mathematical models do not account for any unforeseen changes, a weakness in real time process analysis [33].

Data models including ML algorithms determine patterns from input data and provide an output. To understand these patterns efficiently, they require larger data sets [2]. This means given a small data set other closely relevant data may be used increasing the margin of error. The need for large data and intensive training process also brings about a need for massive processing power which is expensive [34]. Data can also be reduced to fit the available processing power which can

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 30 November 2021 doi:10.20944/preprints202111.0555.v1

compromise results.

2 Review of Studies on Physics Informed Neural Networks

The newly developed Physics Informed Neural Networks framework has been used to analyses and simulate multiple models. This section reviews some of the studies of Physics Informed Neural Networks. Some of the studies also cover COVID-19 related studies.

2.1 Physics Informed Deep Learning for Traffic State Estimation

The Physics Informed Neural Networks framework has been used in the analysis of real time traffic states. Traffic state estimation is the process of estimating traffic variables using partial data. These traffic variables include f which is the traffic flow rate, v the vehicle's average speed rate and ρ which is the vehicle density. The aim of traffic states analysis is to achieve better road planning and understanding. This includes early detection of vehicle congestion blockades and high transportation demand. An example is a detection of a sudden drop in the average speed v would indicate severe congestion or an accident [1].

The approaches that are used to conduct these traffic estimations are mainly mathematical and data driven approaches; the study uses the data driven approach. However, since the use of data driven approaches like machine learning require a lot of data. This data has a drawback as it requires a lot of sensors and other equipment which are very expensive. This forces transportation planners to collect data only in cost effective areas thus leading to the collection of noisy data. To mitigate these challenges the study employees a physics informed neural network approach.

In the development of the mathematical models, they set the variables based on the data collected, q is the indicated number of vehicles crossing a location at time. The average speed v is obtained by finding the mean speed of vehicles and the vehicle density ρ is obtained as the number of vehicles in a particular road distance. $N(x,t)$ is the cumulative traffic flow which is the total number of vehicles that pass a particular point x by time t. $q(x,t)$ is a partial differential equation of cumulative flow which represent the flow with respect to time (t). Density $\rho(x,t)$ is a partial

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 30 November 2021 doi:10.20944/preprints202111.0555.v1

differential equation of cumulative flow with respect to x. The mathematical representations of the densities is

$$q(x,t) = \frac{\partial N(x,t)}{\partial t}, \tag{1}$$

$$\rho(x,t) = \frac{-\partial N(x,t)}{\partial x}. \tag{2}$$

The conservation law states

$$\frac{\partial qN(x,t)}{\partial x} + \frac{\partial \rho N(x,t)}{\partial t} = 0 \tag{3}$$

The relationship between the stated variables is:

$$v(\rho) = v_f \left(1 - \frac{\rho}{\rho_m}\right), \tag{4}$$

$$q(\rho) = \rho v_f \left(1 - \frac{\rho}{\rho_m}\right). \tag{5}$$

where v_f = traffic free flow and ρ_m = maximum traffic flow

The cost function $(J_D L)$ is used to improve the accuracy of the neural network; it is calculated by obtaining the mean square error (MSE) of N number of outputs at point x at time t. $\rho^*(x,t)$ is the neural network's prediction and $\rho(x,t)$ is the genuine value. The implementation of the physics informed neural network then additionally finds the MSE J_{PHY} which is with regard to the conservation of the stated conservation laws.

$$J_{DL} = \frac{1}{N}\sum_{i=1}^{N} |\rho(x,t) - \rho^*(x,t)|^2 \tag{6}$$

$$J_{PHY} = \frac{1}{N}\sum_{i=1}^{N} |v_f(1 - \frac{2\rho(x,t)}{\rho_m}) + \frac{\partial \rho(x,t)}{\partial x}) + \frac{\partial \rho(x,t)}{\partial t})|^2 \tag{7}$$

The physics informed neural network is then used to optimize the neural network and a parameter μ is added to give the neural network an adjustment weight. They then added the PINNS

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 30 November 2021 doi:10.20944/preprints202111.0555.v1

implementation equation.

$$J = \mu J_{DL} + (1 - \mu) J_{PHY} \tag{8}$$

The accuracy is then calculated using the Frobenius norm to measure the accuracy of the neural network. The model was tested using varying data sizes and collection locations and there were positive results [1].

2.2 Physics Informed Neural Networks for Power Systems

This study implemented the physics informed neural network in the analysis of power generation. The power generation process involves the use of generators which are driven by various energy forms including wind and water. The analysis and understanding of a real time power generation is vital in the understanding of the amount of power the generators are producing [33]. The use of data models to analyses the power production and use of mathematical models to make the estimations in not new. However they all have drawbacks, the use of data and machine learning models requires massive data. To eliminate the noisy data there is also a need to have the data processed by experts before use. This analysis model also requires the development of complex neural network designs [33].

The study thus introduces the use of physics informed neural network to implement a training process dependent on both data and physics laws. The study uses a single machine infinite bus (SMIB) system, which is a generator with only one generator. The parameters and variables in the equation include the inertia constant m_1, d_1 damping coefficient, B_{12} is the entry of the bus sustenance. P_1 is power generated by the generator, V_1 and V_2 are voltage magnitudes of buses 1 and 2, σ_1, σ_2 represent the voltage angle behind reactance $\sigma^{\cdot}$ is the angular frequency of generators. Thus the resulting function is

$$f_\sigma(t,P_1) = m_1\sigma^{\cdot\cdot} + d_1\sigma^{\cdot} + B_{12}V_1V_2\sin(\sigma) - P_1. \tag{9}$$

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 30 November 2021 doi:10.20944/preprints202111.0555.v1

Equation 9 is used as the governing equation in the implementation of the physics informed neural network. The model adjusts σ, $\sigma^{\cdot}$ and P_1 between $[P_{min},P_{max}]$ during the learning process. The model was simulated using data sets that were created using computer models and showed very positive results [33].

2.3 Neural Network aided quarantine control model estimation of global COVID-19 spread

The study provides two deep learning models to make approximations of the parameters of the spread of COVID-19.These models are used to make forecasts using data from the USA, China (Wuhan), Italy and South Korea. The deep learning models used in the models are both of physics informed neural network format. The PINNS is employed to eliminate the problems associated with machine learning and normal neural network models. These problems include the over fitting of data, the need for high processing power and the need for extra data from other pandemics or diseases spread such as MERS and SARS in this case. Artificial neural networks are also complex making it hard to comprehend how the final approximation is attained. However, PINNS make the process easier and thus ease the understanding and analysis of COVID-19's spread. The study mainly tries to understand the benefits of the implementations of COVID-19 restrictions [12].

The system of ODE's used in the first model is that of a SEIR model. In the model S represents the number of of people who are Susceptible in the population, E represents the number of Exposed people in the population, I represent the number of people who are Infected or active cases and R represents number of people who have been removed. The first employed model does not account for the impacts that can be imposed by the policies that have been implemented to reduce the spread of COVID-19.

The common mathematical models however do not account for these in the predictions they make; making it hard to account for variables or elements such as over crowding, social distancing and other policies which may have been implemented by the different countries. The main policies highlighted by the authors include the use of police to enforce proper social distancing in traffic crossings, shops and other places. It also focuses on the shutdown of public transport, trains and

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 30 November 2021 doi:10.20944/preprints202111.0555.v1

airports. Thus to account for these multiple policies and have a better prediction the study uses real data. This study is conducted using data and estimations. The study also estimates the effective reproduction rate. The first model:

$$\frac{dS(t)}{dt} = -\frac{\beta S(t)I(t)}{N}, \tag{10}$$
$$\frac{dE(t)}{dt} = \frac{\beta S(t)I(t)}{N} - \sigma E(t), \tag{11}$$
$$\frac{dI(t)}{dt} = \sigma E(t) - \gamma I(t), \tag{12}$$
$$\frac{dR(t)}{dt} = \gamma I(t). \tag{13}$$

Subjected to the initial conditions, $S = S_0$, $I = I_0$, $R = R_0$ and $E = E_0$.

The second model used in the study accounts for quarantine control. The model thus introduces a time dependent variable $T(t) = Q(t) \times I(t)$. This also changes the effective reproduction rate to $R_t = \frac{\beta}{\gamma + Q(t)}$ .The parameter $Q(t)$ is also determined using a separate neural network which takes in the data of Time, Susceptible, Exposed, Infected and Recovered as input data. The model processes the data in a 2-layer network with 10 nodes per layer and uses a ReLu activation function ($NN(W,U)$). The determined $Q(t)$ is then put in the Physics Informed Neural Network which uses the model below to make the approximations of the model.

$$\frac{dS(t)}{dt} = -\frac{\beta S(t)I(t)}{N}, \tag{14}$$
$$\frac{dI(t)}{dt} = \frac{\beta S(t)I(t)}{N} - (\gamma - Q(t))I(t)$$
$$\frac{dI(t)}{dt} = \frac{\beta S(t)I(t)}{N} - (\gamma - NN(W,U))I(t), \tag{15}$$
$$\frac{dR(t)}{dt} = \gamma I(t), \tag{16}$$
$$\frac{dT(t)}{dt} = Q(t)I(t) = NN(W,U)I(t). \tag{17}$$

Subjected to the initial conditions, $S = S_0$, $I = I_0$, $R = R_0$ and $T = T_0$.

The results that were attained by the study showed that the first model which does not account

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 30 November 2021 doi:10.20944/preprints202111.0555.v1

for imposed restrictions had approximations which were bigger than the real values. This means it approximated that the virus would be more catastrophic. The second model achieved a better fit showing that the imposed restrictions have had a positive impact in the spread of COVID-19. The model was also comprehensible providing parameters which can be used to make future predictions [12].

2.4 Identification and prediction of time-varying parameters of COVID-19 model: a data-driven deep learning approach

This study focused on finding the parameters of an SIRD model which are time based rather than to the average parameter [3]. This study also used a deep learning model and specifically a Physics Informed Neural Network. The virus spreading model employed is that study is an SIRD where S represents the number of people who are Susceptible, I represents the Infected people or active cases, R represents the number of Recovered people and D represents the number of Deaths. Where β is the spreading rate, γ is the recovery rate and δ is the death rate [49].

$$\frac{dS(t)}{dt} = -\frac{\beta S(t) I(t)}{N}, \tag{18}$$

$$\frac{dI(t)}{dt} = \frac{\beta S(t) I(t)}{N} - \gamma I(t) - \delta I(t), \tag{19}$$

$$\frac{dR(t)}{dt} = \gamma I(t), \tag{20}$$

$$\frac{dD(t)}{dt} = \quad \delta I(t). \tag{21}$$

The model is subjected to the initial values, $S = S_0$, $I = I_0$, $R = R_0$ and $D = D_0$.

This model does however employ a different orientation of neural networks called a recurrent neural network; a class of artificial neural network. These types of networks have a connection between the nodes forming a directed graph with a long and temporal sequence. This allows the network to exert a temporal dynamic behaviour which allows them to have a kind of Long Shot Temporal Memory (LSTM). The model takes in real life data as input data and uses a Physics Informed neural Network to approximate the parameters.The output values are taken as the input

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 30 November 2021 doi:10.20944/preprints202111.0555.v1

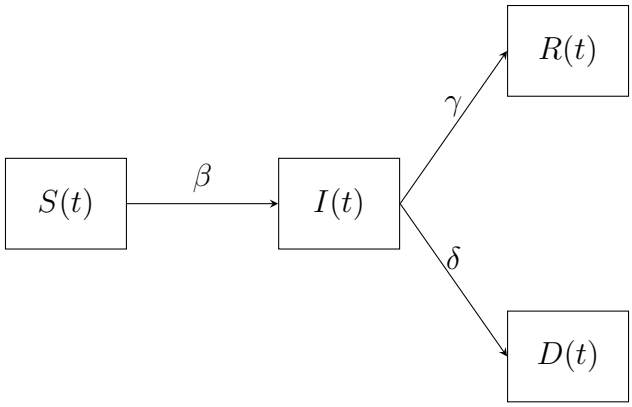

Figure 1: A schematic flow diagram representing an Susceptible - Infected - Recovered - Dead (SIRD) COVID-19 transmission.

of the LSTM recurrent neural network and are used to make predictions of future parameters. The output values which have been obtained are then substituted in the SIRD model. This is a new and advanced model of neural network and provides a better approximation [3].

3 Methodology

The governing laws of the Physics Informed Neural Networks framework are provided as mathematical equations. This section covers the development and evaluation of the key mathematical models of focus. The mathematical model serve as the assumed physics laws the model should adhere to.

The Susceptible - Infected - Recovered - Deceased (SIRD) model used assumes that the population can assume four states, Susceptible (S), Infected (I), Recovered (R) and Deceased (D). The susceptible population is the group which can contract the virus, this contraction occurs at the rate β. The infected population is the population group that has contacted the virus and it's still active. The infected group can be removed to either assume a recovered population at the rate γ or deceased population at the rate δ. This means δ is the death rate, β is the infection rate and γ is the recovery rate. Figure 1 shows the resulting COVID-19 transmission SIRD flow diagram.

From the flow diagram in Figure 1 we obtain the system.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 30 November 2021 doi:10.20944/preprints202111.0555.v1

$$\frac{dS(t)}{dt} = -\frac{\beta S(t)I(t)}{N}, \tag{22}$$

$$\frac{dI(t)}{dt} = \frac{\beta S(t)I(t)}{N} - \gamma I(t) - \delta I(t), \tag{23}$$

$$\frac{dR(t)}{dt} = \gamma I(t), \tag{24}$$

$$\frac{dD(t)}{dt} = \delta I(t). \tag{25}$$

The model is subjected to the initial values, $S = S_0$, $I = I_0$, $R = R_0 = 0$ and $D = D_0 = 0$.

As an accuracy and optimization aid, studies and implementations of neural networks have show that the use of numbers less than 1 is better. Hence we need to rescale the given data to assume values between 0 and 1 through the non-dimentionalization process.

$w = \frac{S}{N}, \quad x = \frac{I}{N}, \quad y = \frac{R}{N}, \quad z = \frac{D}{N}, \quad t = q$

Thus

$S = wN, \quad I = xN, \quad R = yN, \quad D = zN$

Subsisting in the SIRD model we obtain

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 30 November 2021 doi:10.20944/preprints202111.0555.v1

$$\frac{d(wN)}{dq} = -\frac{\beta(wN)(xN)}{N}$$
$$\frac{d(xN)}{dq} = \frac{\beta(wN)(xN)}{N} - \gamma xN - \delta xN$$
$$\frac{d(yN)}{dq} = \gamma xN$$
$$\frac{d(zN)}{dq} = \delta xN$$

Hence the resulting system is.

$$\frac{dw}{dq} = -\beta wx$$
$$\frac{dx}{dq} = \beta wx - \gamma x - \delta x$$
$$\frac{dy}{dq} = \gamma x$$
$$\frac{dz}{dq} = \delta x$$

3.1 The neural network

The resulting neural network that we develop takes a single input value of time t. The input is passed through the layers with weights $W_{i,j}$ where i is the position of the start node and j is the position of the ending node. These weights form a matrix and at every node the product of the weight and time is subjected to an activation function $\tanh$ denoted by σ. The output nodes of the model are $S(t)$, $I(t)$,$R(t)$ and $D(t)$ which form the output layer. x is the sum of the products of W and t.

$$\sigma(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}} \tag{26}$$

The representation of a neural network matrix with m layers and n nodes per layer.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 30 November 2021 doi:10.20944/preprints202111.0555.v1

3.1.1 Residual of model's equations

The residual error of an ODE is the difference between the right hand side and the left hand side of an ODE. In the development of PINNS the residual error is used to calculate the loss function of the neural network. From the SIRD model we obtain four residual error functions. From Equation 22 we obtain Res_S, the residual error of the susceptible population. From Equation 23 we obtain Res_I which is the residual error of the infected population. Equation 24 gives the residual error of the recovered population Res_R. From Equation 25 we obtain Res_D the residual error of the deceased population.

$$Res_S = \frac{dS(t)}{dt} + \frac{\beta S(t)I(t)}{N}, \tag{27}$$

$$Res_I = \frac{dI(t)}{dt} - \frac{\beta S(t)I(t)}{N} + \gamma I(t) + \delta I(t), \tag{28}$$

$$Res_R = \frac{dR(t)}{dt} - \gamma I(t), \tag{29}$$

$$Res_D = \frac{dD(t)}{dt} - \delta I(t). \tag{30}$$

3.1.2 The loss function

To optimize a neural network through back propagation, a loss function has to be first obtained. For the PINNS we developed, we find the loss function $loss_T$ by obtaining the sum of two loss functions $loss_1$ and $loss_2$. $loss_1$ is the sum of the mean square errors of susceptible population $MSE_{Soutput}$, the mean square errors of the infected population $MSE_{Ioutput}$, the mean square errors of the recovered population $MSE_{Routput}$ and the mean square errors of the deceased population $MSE_{Doutput}$. $MSE_{Soutput}$ is the mean square error of the difference of the predicted susceptible $S^*(t_i)$ and the actual data value S_i. $MSE_{Ioutput}$ is the mean square error of the difference of the predicted infected $I^*(t_i)$ and the actual data value I_i. $MSE_{Routput}$ is the mean square error of the difference of the predicted recovered $R^*(t_i)$ and the actual data value R_i. $MSE_{Doutput}$ is the mean square error of the difference of the predicted deceased $D^*(t_i)$ and the actual data value D_i.

$loss_2$ is the sum of the mean square errors of susceptible population residual error MSE_{Sres} given by the mean square errors of the infected population residual error MSE_{Ires}, the mean square

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 30 November 2021 doi:10.20944/preprints202111.0555.v1

errors of the recovered population residual error MSE_{Rres} and the mean square errors of the deceased population residual error MSE_{Dres}.

$$loss_1 = MSE_{Soutput} + MSE_{Ioutput} + MSE_{Routput} + MSE_{Doutput} \tag{31}$$

$$loss_2 = MSE_{Sres} + MSE_{Ires} + MSE_{Rres} + MSE_{Dres} \tag{32}$$

$$loss_T = loss_1 + loss_2 \tag{33}$$

$$MSE_{Sres} = \frac{1}{M}\sum_{i=1}^{M} |Res_S|^2 \tag{34}$$

$$MSE_{Ires} = \frac{1}{M}\sum_{i=1}^{M} |Res_I|^2 \tag{35}$$

$$MSE_{Rres} = \frac{1}{M}\sum_{i=1}^{M} |Res_R|^2 \tag{36}$$

$$MSE_{Dres} = \frac{1}{M}\sum_{i=1}^{M} |Res_D|^2 \tag{37}$$

$$MSE_{Soutput} = \frac{1}{M}\sum_{i=1}^{M} |S^*(t_i) - S_i|^2 \tag{38}$$

$$MSE_{Ioutput} = \frac{1}{M}\sum_{i=1}^{M} |I^*(t_i) - I_i|^2 \tag{39}$$

$$MSE_{Routput} = \frac{1}{M}\sum_{i=1}^{M} |R^*(t_i) - R_i|^2 \tag{40}$$

$$MSE_{Doutput} = \frac{1}{M}\sum_{i=1}^{M} |D^*(t_i) - D_i|^2. \tag{41}$$

From the input of time, to the matrix of layers, the output layer and the residual functions we obtain a neural network which resembles Figure 2

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 30 November 2021 doi:10.20944/preprints202111.0555.v1

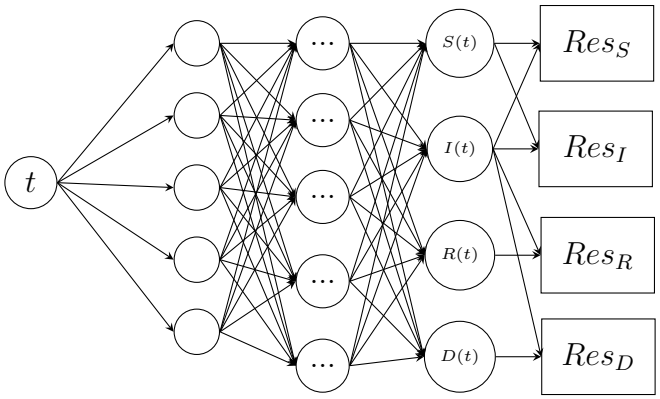

Figure 2: A schematic representation of the Physics informed neural network which takes an input of time (t) and outputs Susceptible(S), Infected (I), Recovered (R) and Deceased D. The output is subjected to PINN.

3.2 Basic model properties

The following section provides the analysis of the mathematical model. It gives a detailed evaluation on the properties of the model and its expected behaviour such as finding the reproduction number which is the minimum number of transmissions expected for a pandemic to occur and the system of ODE's sensitivity analysis.

3.2.1 Reproduction Number R_0

To understand COVID-19 which has become a pandemic we need to determine the minimum rate at which secondary infections should occur for a pandemic to occur. The reproduction number R_0 is also the rate which any spread below would stop the spread.The following is its derivation.

$$0 < \beta S_0 I_0 - (\gamma + \delta) I_0 \tag{42}$$

$$0 < \beta S_0 - (\gamma + \delta) \tag{43}$$

$$\beta S_0 < (\gamma + \delta) \tag{44}$$

$$R_0 = \frac{\beta S_0}{\gamma + \delta} \tag{45}$$

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 30 November 2021 doi:10.20944/preprints202111.0555.v1

3.2.2 SIRD Model Analysis

The sensitivity analysis of the mathematical model also provides some of the Key properties of the model such as the projected maximum number of infections I_{max}. From Equation 22,Equation 23,Equation 24 and Equation 25 we find the maximum number of infected individuals that can occur at a particular time. First we divide Equation 22 and Equation 23.

$$\frac{dI(t)}{dS(t)} = -1 + \frac{\gamma + \delta}{\beta S}$$

Integrating we obtain

$$I + S - \frac{\gamma + \delta}{\beta}(\ln S) = I_0 + S_0 - \frac{\gamma + \delta}{\beta}(\ln S_0)$$

To obtain the maximum value we find a point where the Equation 3.2.2 is equal to zero, and we determine it occurs when $S = \frac{\beta}{\gamma+\delta}$.

$$\begin{aligned} I + S - \frac{\gamma + \delta}{\beta}(\ln S) &= I_0 + S_0 - \frac{\gamma + \delta}{\beta}(\ln S_0) \\ I_{max} + S - \frac{\gamma + \delta}{\beta}(\ln S) &= I_0 + S_0 - \frac{\gamma + \delta}{\beta}(\ln S_0) \\ I_{max} + \frac{\beta}{\gamma + \delta} - \frac{\gamma + \delta}{\beta}(\ln(\frac{\beta}{\gamma + \delta})) &= I_0 + S_0 - \frac{\gamma + \delta}{\beta}(\ln S_0) \end{aligned}$$

$$\begin{aligned} I_{max} &= I_0 + S_0 - \frac{\gamma + \delta}{\beta}[1 + (\ln(\frac{\beta}{\gamma + \delta})S_0)] \\ I_{max} &= I_0 + S_0 - \frac{\gamma + \delta}{\beta}[1 + (\ln(R_0))]. \end{aligned}$$

Now we obtain the amount of people that we expect to eventually get infected. The fate of infected individuals is that they either recover RR_{end} or die D_{end} hence we find the expected number of people to either recover or die.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 30 November 2021 doi:10.20944/preprints202111.0555.v1

$$\overline{R}_{end} = RR_{end} + D_{end},$$
$$\overline{R}_{end} = S_0 + I_0 - S_{end}.$$

where

$$S_{end} = \frac{\gamma + \delta}{\beta} ln(S_{end}),$$
$$S_{end} = I_0 + S_0 - \frac{\gamma + \delta}{\beta} ln(S_0)$$

Where S represents susceptibles, I represents Infected, R represents Recovered and D represents Deceased. Then $N = S + I + R + D$ the total number of the population. β is the infection rate, γ is the recovery rate and δ is the death rate. Since in the analysis or in the model Recovered and Deceased individuals have the same effects on the model, we group them as R representing removed with the removal rate of $(\gamma + \delta)$. Such that we have,

$$\frac{dS(t)}{dt} = -\frac{\beta S(t) I(t)}{N},$$
$$\frac{dI(t)}{dt} = \frac{\beta S(t) I(t)}{N} - (\gamma + \delta) I(t),$$
$$\frac{dR(t)}{dt} = (\gamma + \delta) I(t).$$

Which can be rewritten as,

$$\frac{dS(t)}{dt} = -\frac{\beta S(t) I(t)}{N}$$
$$\frac{dI(t)}{dt} = \frac{\beta S(t) I(t)}{N} - (\gamma + \delta) I(t)$$
$$R = N - I - S$$

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 30 November 2021 doi:10.20944/preprints202111.0555.v1

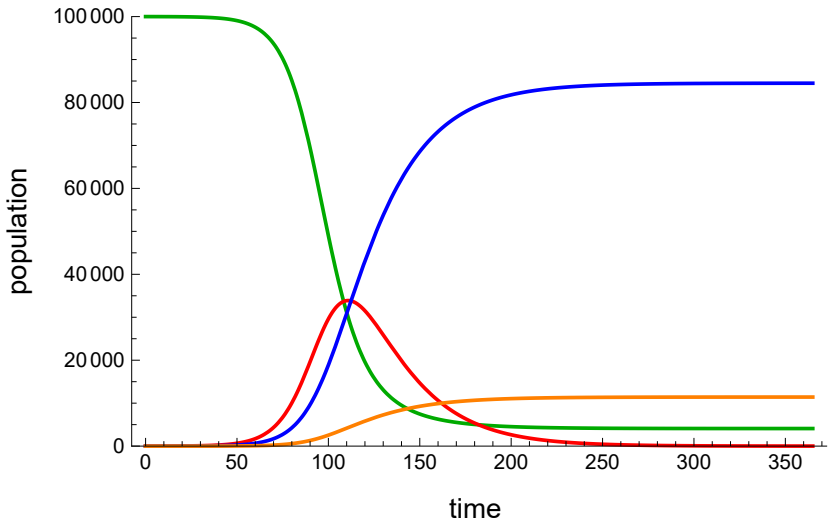

Figure 3: A Mathematica generated graph simulation of an example SIRD model. The green represents Susceptible population, blue represents the recoveries, red is the active infected population and orange is the deceased population.

4 Findings and Discussions

This section contains the results obtained from the simulations of the Physics Informed Neural Networks framework. It also provides a detailed analysis of the results and changes in accuracy with varying parameters such as data size. The data is from national daily updates and a Google studio analysis page by University of Eswatini and Wits Ithemba Labs [52].

4.1 Simulation Using Mathematica Generated Data

To test the model and validate the PINNS we first generated artificial data of an SIRD model using Mathematica. The advantage of data of this kind was that its less noisy. We initialized the model with a susceptible population of 100,000, 0 recoveries and deaths and 5 infections. The average infection rate is 0.14, average recovery rate was 0.037 and the average death rate was 0.005 obtaining the results in Figure 3.

4.1.1 PINNS model of Mathematica results

The results of this model discussed above from a PINNS of 3 layers with 30 nodes per layer.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 30 November 2021 doi:10.20944/preprints202111.0555.v1

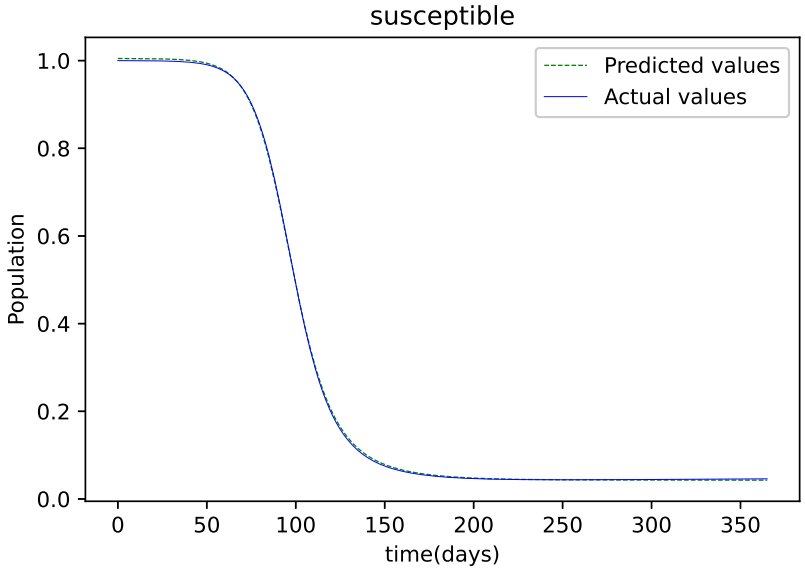

Figure 4: The resulting graph of the predicted values of the Susceptible and the actual values of the susceptible from the Mathematica generated data.

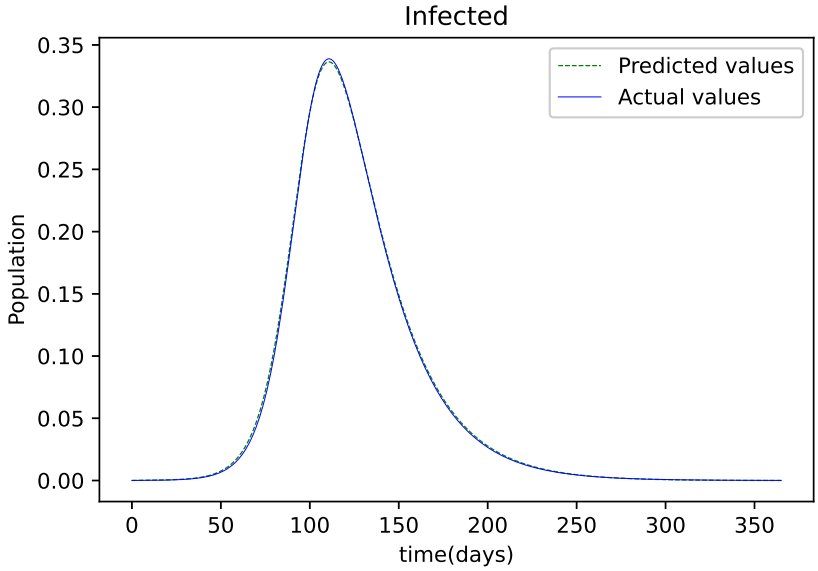

Figure 5: The resulting graph of the predicted values of the Infected population and the actual values of the infected population from the Mathematica generated data

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 30 November 2021 doi:10.20944/preprints202111.0555.v1

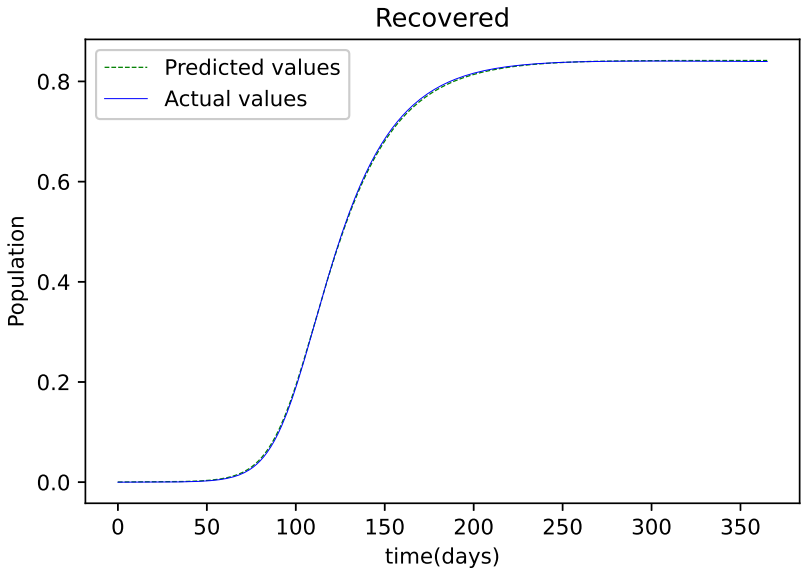

Figure 6: The graphs shows the results of the predicted values of the Recovered and the actual values of the recovered from the Mathematica generated data.

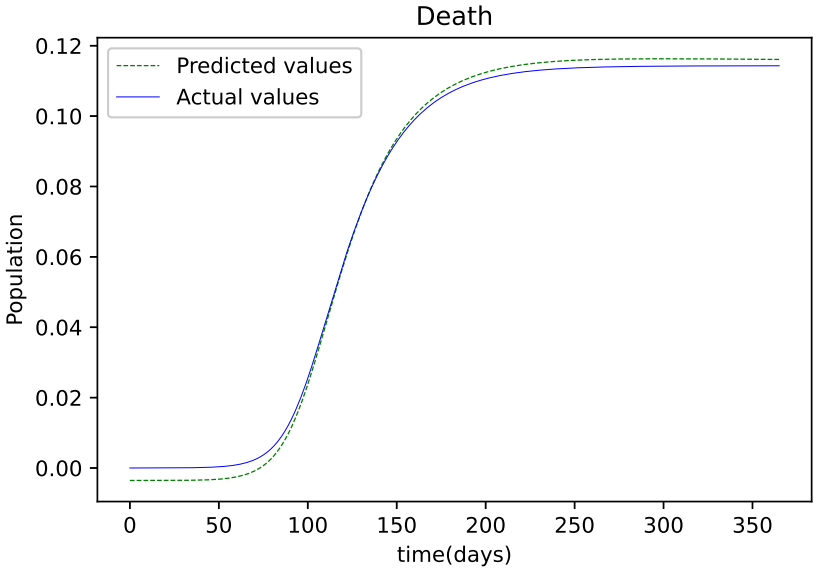

Figure 7: The resulting graph of the predicted values of the Deceased population and the actual values of the deceased population from the Mathematica generated data.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 30 November 2021 doi:10.20944/preprints202111.0555.v1

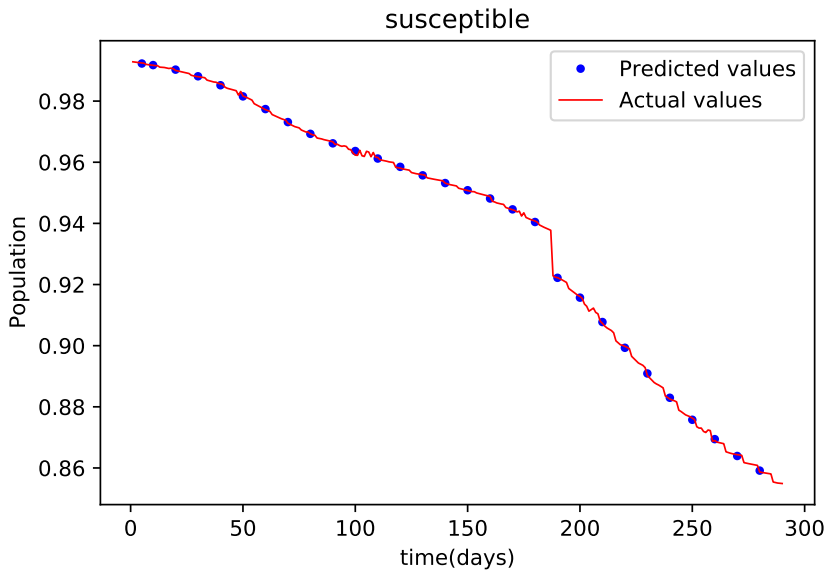

Figure 8: This graph shows a comparison of the predicted values of susceptible population and the actual data of susceptible population for the State of Alabama.

The graph in Figure 3 is the resulting graph of the data used for the artificial data for the early training process. Figure 4 is a data fitting graph of the susceptible population and there is a good fit which means there is a substantially small sized error. Figure 5 is the obtained results and the graphs have a good fit which means they have a small error. The graph in Figure 6 is that of the results of the recovered and it also has a good fit and less errors. Figure 7 is the resulting graph of deceased, this graph has a good fit, but is less accurate compared to the other graphs.

4.2 PINNS Simulations of Alabama State Data

For further validation of the model we tested the SIRD model using a data set from the American State of Alabama. The data set covered about 300 days and the simulation was on a 3 layer neural network with each layer having 30 nodes each and 1 million iterations were conducted.

Figure 8 is a resulting graph obtained using data fitting for the susceptible population and there is a good fit which means there is minimal sized error. Figure 9 is the resulting graph of the data fitting, it has a good fit which means they have a small error. The graph in Figure 10 is that of the results of the recovered and it also has a good fit and less errors. Figure 11 is the resulting graph

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 30 November 2021 doi:10.20944/preprints202111.0555.v1

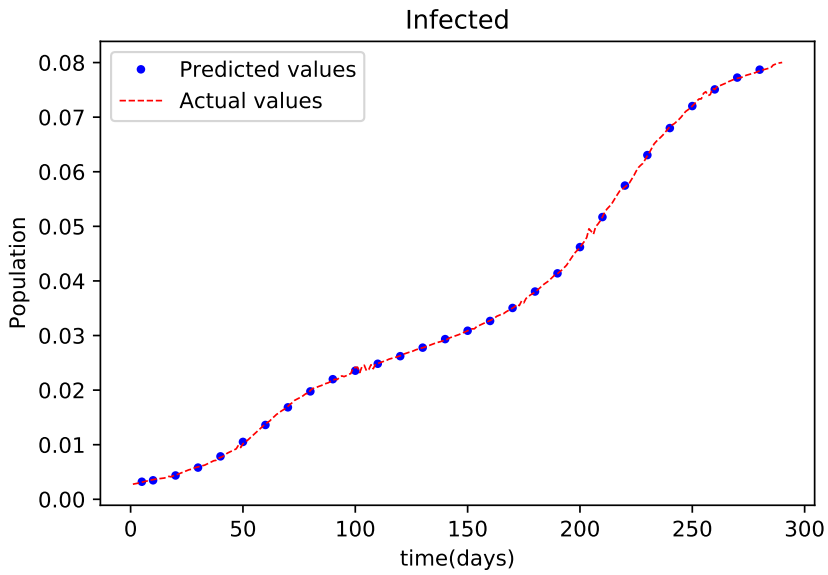

Figure 9: The graph shows a comparison of the predicted infected population values and the actual data infected population for the State of Alabama.

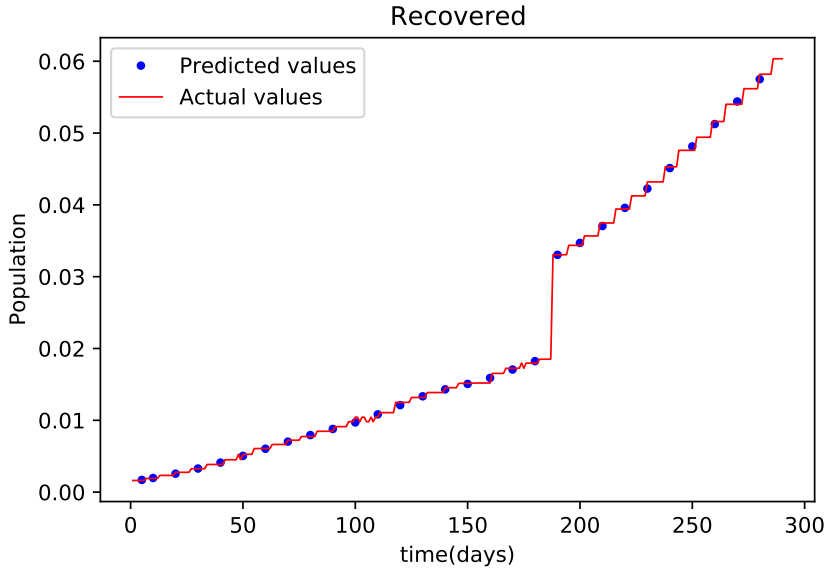

Figure 10: This graph shows a comparison of the predicted values of recovered population and the actual data of recovered population for the State of Alabama.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 30 November 2021 doi:10.20944/preprints202111.0555.v1

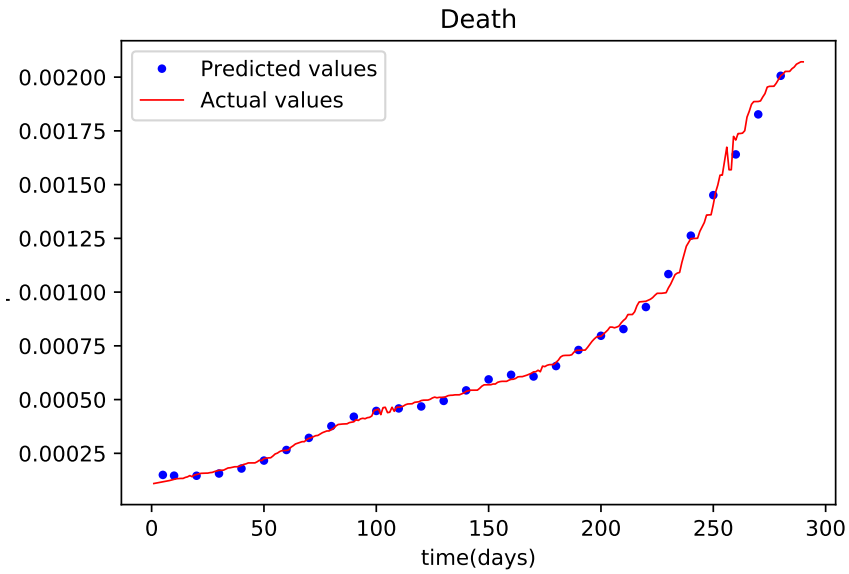

Figure 11: The graph shows a comparison of the predicted deceased population values and the actual data deceased population for the State of Alabama.

of deceased, this graph has a good fit, however it has a bigger error compared to the other graphs.

4.3 PINNS simulation of A model using 170 data points

After using a small sized data set to train the model of 30% of the available data results were obtained. Figure 12 is a resulting graph obtained only for data fitting purposes for the susceptible population and there is a good fit which, hence a small sized error. Figure 13 is the obtained graph of the data fitting, it has a good fit which means they have a small error. The graph in Figure 14 is that of the results of the recovered and it also has a good fit and less errors. Figure 15 is the resulting graph of deceased, this graph has a good fit, however it has a bigger error compared to the other graphs. The overall outcome shows that as much as the fitting has less errors, they are more larger compared to the cases where bigger size data was used.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 30 November 2021 doi:10.20944/preprints202111.0555.v1

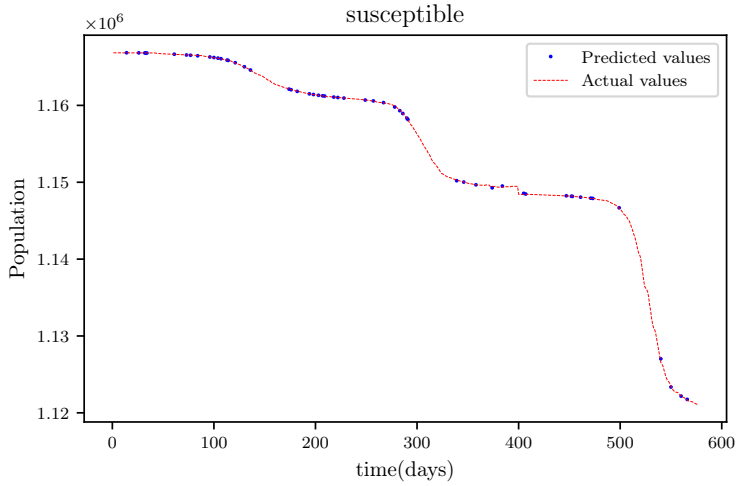

Figure 12: This graph shows a comparison of the predicted values of susceptible population and the actual data of susceptible population for a 130 data points.

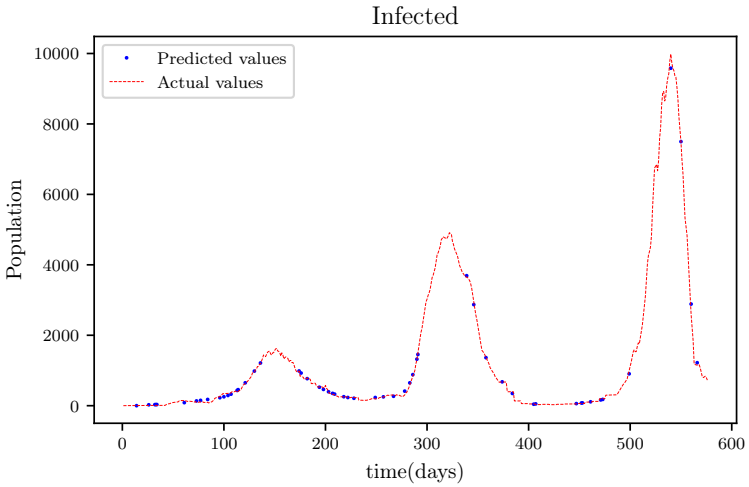

Figure 13: The graph shows a comparison of the predicted infected population values and the actual data infected population for a 130 data points.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 30 November 2021 doi:10.20944/preprints202111.0555.v1

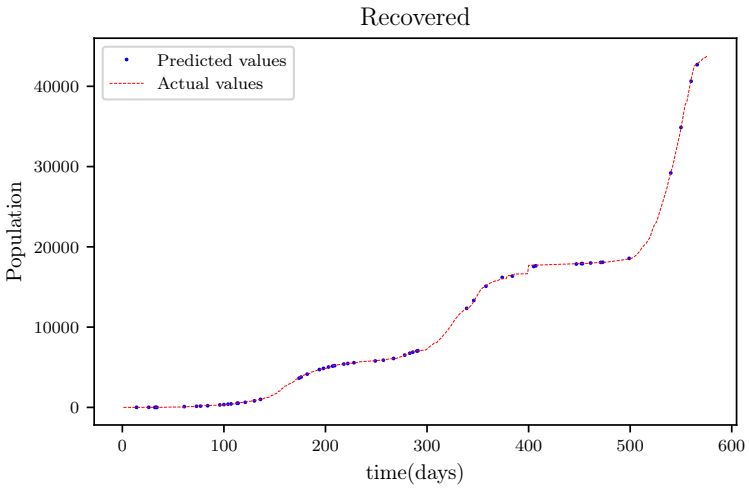

Figure 14: This graph shows a comparison of the predicted values of recovered population and the actual data of recovered population for a 130 data points.

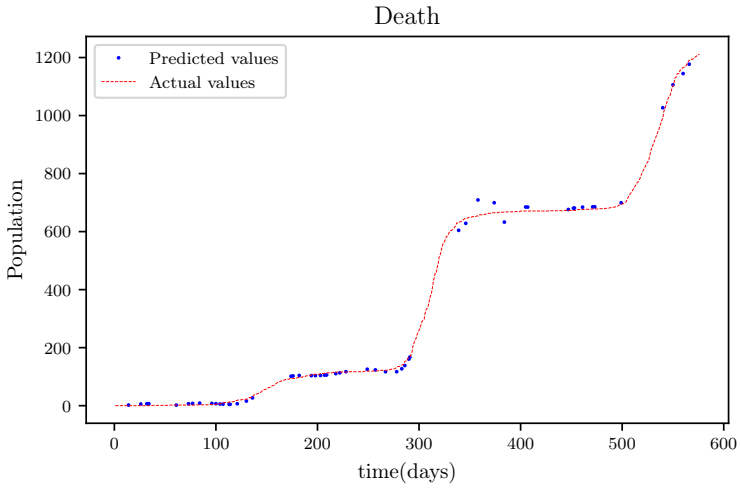

Figure 15: The graph shows a comparison of the predicted deceased population values and the actual data deceased population for a 130 data points.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 30 November 2021 doi:10.20944/preprints202111.0555.v1

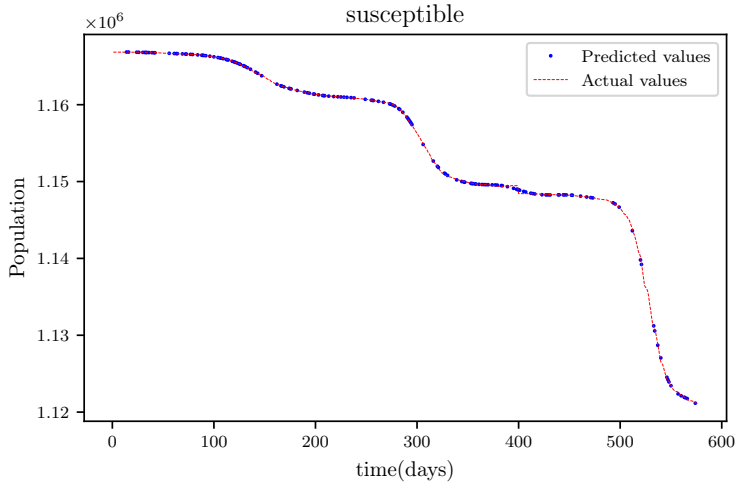

Figure 16: This graph shows a comparison of the predicted values of susceptible population and the actual data of susceptible population for a 530 data points.

4.4 PINNS simulation of A model using all available data points at the time (576 data points)

The simulation were conducted with 5,00,000 iterations and 4 layers and each layer having 30 nodes. The size of the data set used was 576 which was the maximum days data available at the time.

Figure 16 is the resulting graph for data fitting purposes for the susceptible population and there is a small sized error and a good fitting. Figure 17 is the obtained graph of the data fitting, it has a good fit which means they have a small error. The graph in Figure 18 is that of the results of the recovered and it also has a good fit and less errors. Figure 19 is the resulting graph of deceased, this graph has a good fit, however it has a bigger error compared to the other graphs. The overall outcome shows that as much as the fitting has less errors, they are more larger compared to the cases where bigger size data was used.

4.5 PINNS simulation forecasting 30 days

Simulations using the 3 layers of 30 nodes per layer was conducted and there were 5,000,000 iterations made during the training.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 30 November 2021 doi:10.20944/preprints202111.0555.v1

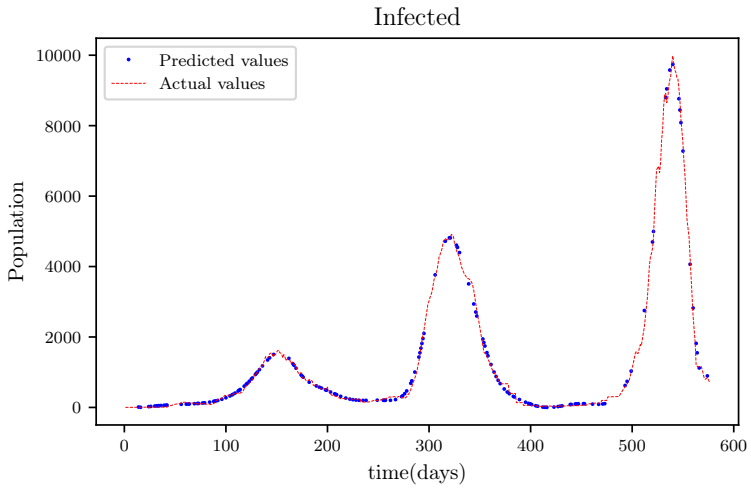

Figure 17: The graph shows a comparison of the predicted infected population values and the actual data infected population for a 530 data points.

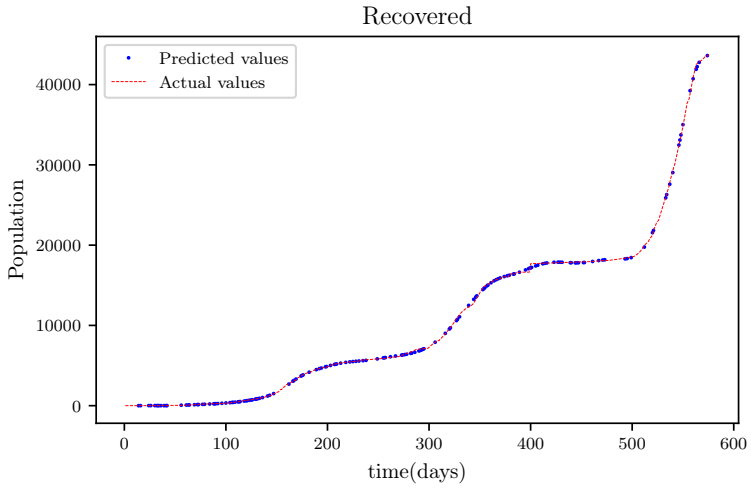

Figure 18: This graph shows a comparison of the predicted values of recovered population and the actual data of recovered population for a 530 data points.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 30 November 2021 doi:10.20944/preprints202111.0555.v1

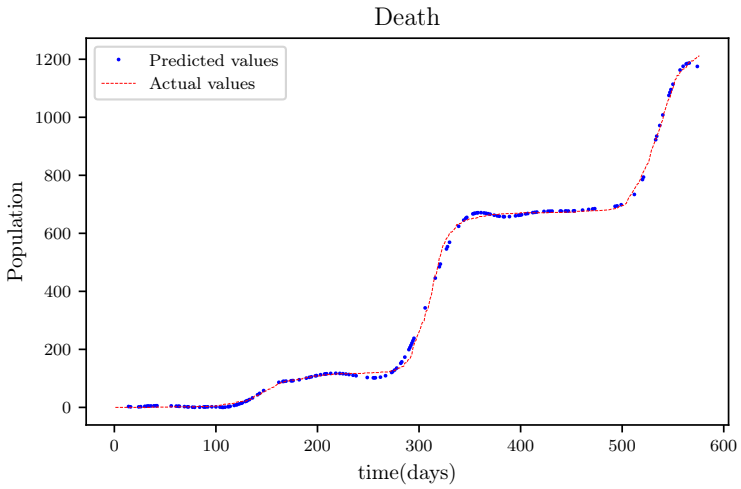

Figure 19: The graph shows a comparison of the predicted deceased population values and the actual data deceased population for a 530 data points.

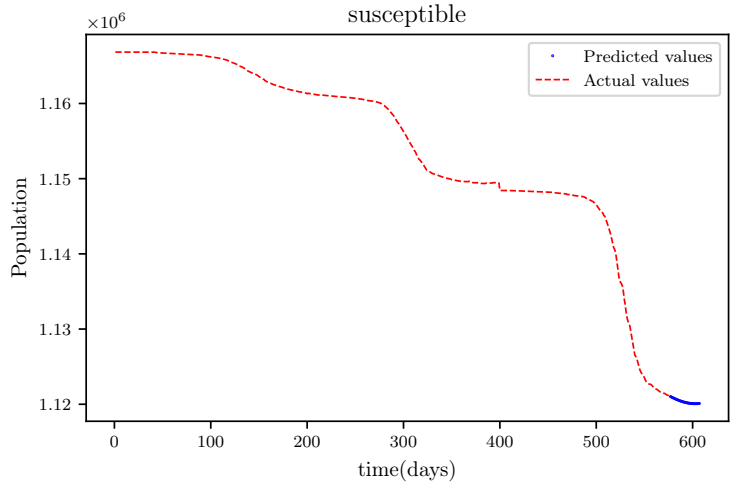

Figure 20: This graph shows a comparison of the predicted values of susceptible population and the actual data of susceptible population for a SIRD model with future predictions.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 30 November 2021 doi:10.20944/preprints202111.0555.v1

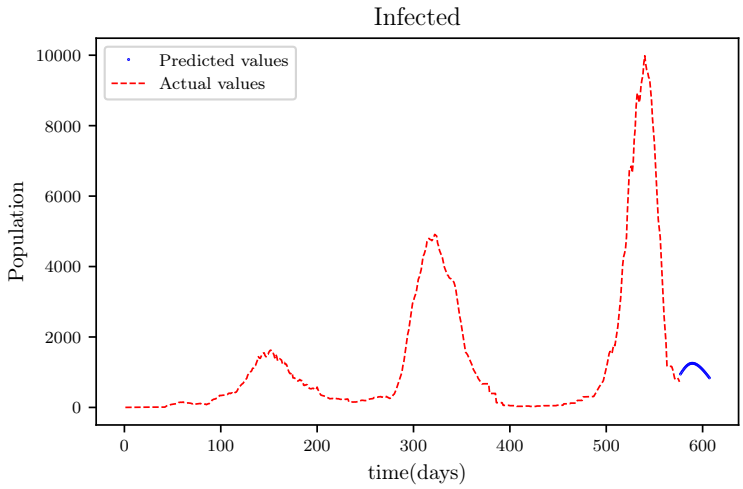

Figure 21: The graph shows a comparison of the predicted infected population values and the actual data infected population for a SIRD model with future predictions.

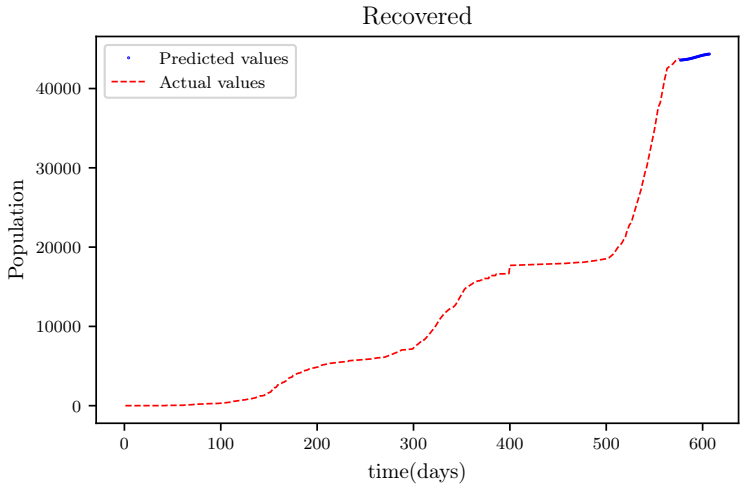

Figure 22: This graph shows a comparison of the predicted values of recovered population and the actual data of recovered population for a SIRD model with future predictions.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 30 November 2021 doi:10.20944/preprints202111.0555.v1

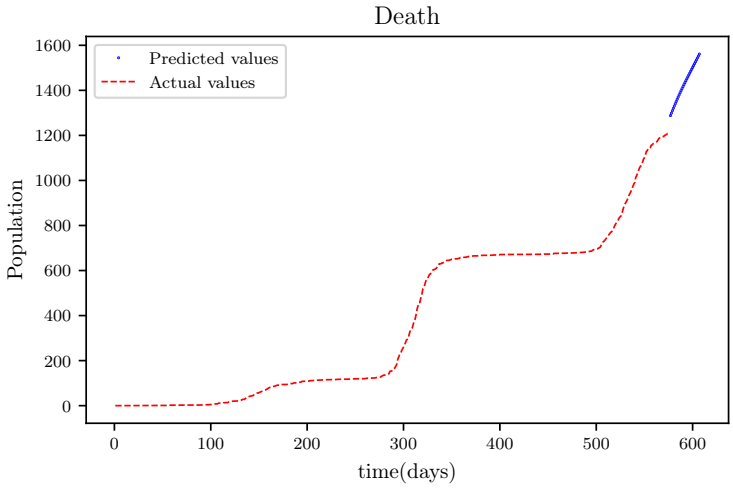

Figure 23: The graph shows a comparison of the predicted deceased population values and the actual data deceased population for a SIRD model with future predictions.

Figure 20 is a resulting graph on the forecasting of 30 days of the susceptible population, the numbers seem to continue dropping which is the expected results. Figure 21 is the obtained graph showing the forecasted data of the active infections which shows a curved format. The graph in Figure 22 is that of the results of the forecasted recoveries. Figure 23 is the resulting graph of the forecasted deceased population. The forecast also determined that the maximum expected infections I_{max}, susceptibles population expected at the end of the spread of the disease $S_{end}$and the total expected recoveries R_{end}.

$I_{max} = 72{,}121$

$S_{end} = 1{,}094{,}719$

$R_{end} = 70{,}274$

4.6 Deep Learning Sensitivity Analysis

The Physics Informed neural Network framework developed in the study is mainly affected by four factors. These are the number of iterations performed during training, the size of the training data used, the total number of layers in the model and the number of nodes in each layer. To conduct the sensitivity analysis first the default model is setup, number of iterations = 200 000, size of training

 doi:10.20944/preprints202111.0555.v1

316 data $= 400$, number of layers $= 3$ and number of nodes $= 30$.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 30 November 2021 doi:10.20944/preprints202111.0555.v1

Table 1: Results of the mean square error analyzing of varying number of iterations and number of layers in the simulations

.3Iterations	Number of layers		
	2	4	8
100 000	3.397×10^{-6}	1.996×10^{-7}	5.461×10^{-8}
200 000	2.434×10^{-6}	1.871×10^{-7}	3.866×10^{-8}
400 000	2.098×10^{-7}	2.340×10^{-8}	3.584×10^{-9}
800 000	1.454×10^{-7}	1.621×10^{-8}	3.055×10^{-9}

Table 2: Results of the mean square error analysing of varying number of nodes in a layer and number of layers in the simulations

.3 Nodes	Number of layers		
	2	4	8
10	1.870×10^{-7}	2.098×10^{-8}	4.783×10^{-9}
20	2.494×10^{-7}	2.243×10^{-8}	4.131×10^{-9}
40	2.144×10^{-7}	2.830×10^{-8}	4.723×10^{-9}
80	2.789×10^{-7}	2.941×10^{-8}	8.794×10^{-9}

Table 3: Results of the mean square error analysing of varying sizes of data points and number of layers in the simulations

.3Data size	Number of layers		
	2	4	8
100	2.269×10^{-7}	2.070×10^{-8}	3.438×10^{-9}
150	2.121×10^{-7}	3.473×10^{-8}	4.123×10^{-9}
200	3.364×10^{-7}	2.904×10^{-8}	1.015×10^{-9}
350	3.214×10^{-7}	4.470×10^{-8}	3.440×10^{-9}

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 30 November 2021 doi:10.20944/preprints202111.0555.v1

Table 4: Results of the mean square error analysing of varying number of iterations and number of nodes per layer in the simulations

.3Nodes	Number of Iterations		
	100 000	400 000	800 000
10	8.991×10^{-7}	3.655×10^{-8}	5.824×10^{-8}
20	2.361×10^{-6}	2.323×10^{-8}	2.390×10^{-8}
40	5.144×10^{-7}	7.846×10^{-8}	2.632×10^{-8}
80	7.642×10^{-7}	5.921×10^{-8}	3.327×10^{-8}

Table 5: Results of the mean square error analyzing of varying number of iterations and data size per layer in the simulations

.3Data size	Number of Iterations		
	100 000	400 000	800 000
100	1.292×10^{-6}	1.188×10^{-7}	2.752×10^{-8}
150	2.536×10^{-6}	3.273×10^{-8}	2.843×10^{-8}
200	1.110×10^{-6}	1.491×10^{-8}	2.064×10^{-8}
350	1.063×10^{-6}	3.399×10^{-8}	1.661×10^{-8}

Table 6: Results of the mean square error analyzing of varying sizes of data points iterations and number of nodes in layers in the simulations

.3Data size	Number of Nodes		
	10	40	80
100	6.102×10^{-8}	4.882×10^{-8}	1.533×10^{-7}
150	4.574×10^{-8}	3.127×10^{-8}	1.823×10^{-7}
200	4.053×10^{-8}	3.386×10^{-8}	5.401×10^{-8}
350	1.277×10^{-7}	3.318×10^{-8}	1.231×10^{-8}

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 30 November 2021 doi:10.20944/preprints202111.0555.v1

Then multiple simulations were conducted setting all the model variables as default, only vary two parameters and take a record of all the mean square errors. Only a single simulation of the model setup is conducted. The initial parameters are randomly set for each scenario which means there are some uncontrolled margins of errors. For such cases a contingency only allows one extra simulation trial.

From Table 1 the results of the study show that as all other parameters remain the same an increase in the number of iterations reduces the margin of error in the same layer. The results also vividly shows that an increase in the number of layers also reduces the error. That means that when the number of layer and iterations are increased much better accuracy is achieved. Table 2 results show that the increase in number of layers reduces the error. The size of the error in each layer also varies just like the changes with the number of layers. From the results it's observed that the change in the number of nodes has limited effects. Table 3 shows that an increase in the number of hidden layers reduces the margin error. It also shows that an increase in the size of data also reduces the margin of error.This shows that an increase in both the data size and number of iterations will reduce the margin of error.

Table 4 results shows that as the number of nodes per layer is increased the margin of error is reduced. The results also show that the as the number of iterations are increased the error also reduces. Table 5 shows that as the number of iterations increases the error is reduced. It also shows that as the error is random smaller iterations but as the iterations increase the larger data set achieve less errors. Table 6 shows that as the data increase per number of nodes the error is reduced, however as the number of nodes are increased the size of the error increases.

4.7 Discussion

The results obtained showed high accuracy especially in data fitting compared to mathematical model approach. One other advantage of this model is that the as it makes predictions it also predicts the wave behaviour of the active infected population. In comparison to other PINNS approaches the model achieved an relatively similar results on being surpassed slightly by a convolutional neural network time varying model [12, 49]. The biggest limit that the method has is

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 30 November 2021 doi:10.20944/preprints202111.0555.v1

that it uses historical data. This makes it loss efficiency as it gets used to forecast data especially if the forecasted period is long since other factors previously unforeseen get introduced such as new species.

5 Conclusion

The aim of the study was to analyse the dynamics of COVID-19 using the Physics Informed Neural Networks framework; A framework which exposes a neural network to both data and governing equations during the training process. The mathematical model employed as the governing equations of the PINNS training model was the Susceptible- Infected - Recovered - Deceased. The data used for conducting the simulations was from the kingdom of Eswatini. Simulations of PINNS were conducted results were obtained and viewed in the form of tables and graphs.These simulations included the use of artificial data to validate the model.

The first simulation of this model included the use of only 170 data points which gave accurate result. These results were less accurate compared to results obtained from simulations of larger datasets but, were reasonably accurate proving that the framework can achieve accurate results even from small size datasets [49]. A larger dataset containing all the data available at the time was also used. The results were obtained with highly accurate predictions, compared to conventional deep learning [12, 49]. The model was also more accurate than mathematical models used in various other studies [37]. Other simulations involved forecasting prospective SIRD values.

The obtained results conclude that the model is well suited to make predictions of values within the training period. This thus means that the model is well suited in data fitting, where there are some days where data was not collected,was wrongly inserted or was lost. The other benefit of the framework is that it returns the spreading rate, death rate and recovery rate which were setup to serve as the adjusting variables for the PINNS part of the neural network. One limit however is that increasing the number of forecasted days leads to diminishing accuracy. This also occurs since the predictions are linked to the spread rate, death rate and recovery rate patterns which are determined using the old available data. The model does also predict the wave format displayed by the active infections. This concludes that the model is highly suitable for making short-term

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 30 November 2021 doi:10.20944/preprints202111.0555.v1

predictions, but for long-term purposes predictions can be made with a sustainable margin of error.

One big limitation faced during the research was the shortage of data and process power. Thus for further research we recommend the development of a model which will group each of the SIRD populations using age, since it has been shown that different age groups are affected by the disease differently. Using the well segmented data each of the parameters or rates can then be defined as per age group. We also recommend that a model similar to the one used in the study be tested only at a larger scale using higher processing power.

References

[1] Huang, J., & Agarwal, S. (2020, September). Physics informed deep learning for traffic state estimation. In 2020 IEEE 23rd International Conference on Intelligent Transportation Systems (ITSC) (pp. 1-6). IEEE.

[2] Raissi, Maziar, Paris Perdikaris, and George E. Karniadakis. "Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations." Journal of Computational Physics 378 (2019): 686-707.

[3] Long, Jie, A. Q. M. Khaliq, and K. M. Furati. "Identification and prediction of time-varying parameters of COVID-19 model: a data-driven deep learning approach." International Journal of Computer Mathematics just-accepted (2021): 1-19.

[4] Moriarty, Leah F., et al. "Public health responses to COVID-19 outbreaks on cruise ships?worldwide, February?March 2020." Morbidity and Mortality Weekly Report 69.12 (2020): 347-352

[5] Chen, Yi-Cheng, et al. "A time-dependent SIR model for COVID-19 with undetectable infected persons." IEEE Transactions on Network Science and Engineering 7.4 (2020): 3279-3294.

[6] Lu, Hongzhou, Charles W. Stratton, and Yi?Wei Tang. "Outbreak of pneumonia of unknown etiology in Wuhan, China: the mystery and the miracle." Journal of medical virology 92.4 (2020): 401-402.

[7] Akpan, I.J., Aguolu, O.G. and Ezeume, I.C., 2021. Overcoming the Challenge of Communicating the Concept and Science of SARS-CoV-2 and COVID-19 to Non-Experts. Challenge, 64(2), pp.117-131.

[8] Guo, Yan-Rong, et al. "The origin, transmission and clinical therapies on coronavirus disease 2019 (COVID-19) outbreak?an update on the status." Military Medical Research 7.1 (2020): 1-10.

[9] Warren, George W., Ragnar Lofstedt, and Jamie K. Wardman. "COVID-19: the winter lockdown strategy in five European nations." Journal of Risk Research 24.3-4 (2021): 267-293.

[10] Oduro, Bismark, and Vusi Mpendulo Magagula. "COVID-19 intervention models: An initial aggressive treatment strategy for controlling the infection." Infectious Disease Modelling 6 (2021): 351-361.

[11] Huang, Chaolin, et al. "Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China." The lancet 395.10223 (2020): 497-506.

[12] Dandekar, R., & Barbastathis, G. (2020). Neural Network aided quarantine control model estimation of global Covid-19 spread. arXiv preprint arXiv:2004.02752.

[13] Bloomgarden, Zachary T. "Diabetes and COVID?19." Journal of Diabetes 12.4 (2020): 347-348.

[14] Lauer, Stephen A., et al. "The incubation period of coronavirus disease 2019 (COVID-19) from publicly reported confirmed cases: estimation and application." Annals of internal medicine 172.9 (2020): 577-582.

[15] Ivorra, Benjamin, et al. "Mathematical modeling of the spread of the coronavirus disease 2019 (COVID-19) taking into account the undetected infections. The case of China." Communications in nonlinear science and numerical simulation 88 (2020): 105303.

[16] Thabet, Sabri TM, et al. "Study of transmission dynamics of COVID-19 mathematical model under ABC fractional order derivative." Results in Physics 19 (2020): 103507.

[17] Goodfellow, Ian, et al. Deep learning. Vol. 1. No. 2. Cambridge: MIT press, 2016.

[18] Jordan, Michael I., and Tom M. Mitchell. "Machine learning: Trends, perspectives, and prospects." Science 349.6245 (2015): 255-260.

[19] Kotsiantis, Sotiris B., I. Zaharakis, and P. Pintelas. "Supervised machine learning: A review of classification techniques." Emerging artificial intelligence applications in computer engineering 160.1 (2007): 3-24.

[20] Mijwel, Maad M. "Artificial neural networks advantages and disadvantages." Retrieved from LinkedIn https//www. linkedin. com/pulse/artificial-neuralnet Work (2018).

[21] Barto, Andrew G., and Thomas G. Dietterich. "Reinforcement learning and its relationship to supervised learning." Handbook of learning and approximate dynamic programming 10 (2004): 9780470544785.

[22] Li, S., Ma, B., Chang, H., Shan, S., & Chen, X. (2018, July). Continuity-discrimination convolutional neural network for visual object tracking. In 2018 IEEE International Conference on Multimedia and Expo (ICME) (pp. 1-6). IEEE.

[23] Dey, Ayon. "Machine learning algorithms: a review." International Journal of Computer Science and Information Technologies 7.3 (2016): 1174-1179.

[24] Sra, Suvrit, Sebastian Nowozin, and Stephen J. Wright, eds. Optimization for machine learning. Mit Press, 2012.

[25] Raschka, Sebastian. Python machine learning. Packt publishing ltd, 2015.

[26] Shen, D., Wu, G., & Suk, H. I. (2017). Deep learning in medical image analysis. Annual review of biomedical engineering, 19, 221-248.

[27] James, Gareth, et al. "Unsupervised learning." An introduction to statistical learning. Springer, New York, NY, 2021. 497-552.

[28] Kaelbling, Leslie Pack, Michael L. Littman, and Andrew W. Moore. "Reinforcement learning: A survey." Journal of artificial intelligence research 4 (1996): 237-285.

[29] Likhachev, Maxim, and Dave Ferguson. "Planning long dynamically feasible maneuvers for autonomous vehicles." The International Journal of Robotics Research 28.8 (2009): 933-945.

[30] Fagnant, Daniel J., and Kara Kockelman. "Preparing a nation for autonomous vehicles: opportunities, barriers and policy recommendations." Transportation Research Part A: Policy and Practice 77 (2015): 167-181.

[31] Mishra, R., & Tripathi, S. P. (2021). Deep learning based search engine for biomedical images using convolutional neural networks. Multimedia Tools and Applications, 80(10), 15057-15065.

[32] Preuer, K., Lewis, R. P., Hochreiter, S., Bender, A., Bulusu, K. C., & Klambauer, G. (2018). DeepSynergy: predicting anti-cancer drug synergy with Deep Learning. Bioinformatics, 34(9), 1538-1546.

[33] Misyris, G. S., Venzke, A., & Chatzivasileiadis, S. (2020, August). Physics-informed neural networks for power systems. In 2020 IEEE Power & Energy Society General Meeting (PESGM) (pp. 1-5). IEEE.

[34] Mao, Zhiping, Ameya D. Jagtap, and George Em Karniadakis. "Physics-informed neural networks for high-speed flows." Computer Methods in Applied Mechanics and Engineering 360 (2020): 112789.

[35] Jagtap, Ameya D., and George Em Karniadakis. "Extended Physics-Informed Neural Networks (XPINNs): A Generalized Space-Time Domain Decomposition Based Deep Learning Framework for Nonlinear Partial Differential Equations." Communications in Computational Physics 28.5 (2020): 2002-2041.

[36] Kazory, Amir, Claudio Ronco, and Peter A. McCullough. "SARS-CoV-2 (COVID-19) and intravascular volume management strategies in the critically ill." Baylor University Medical Center Proceedings. Vol. 33. No. 3. Taylor & Francis, 2020.

[37] Inui, Shohei, et al. "Chest CT findings in cases from the cruise ship diamond princess with Coronavirus disease (COVID-19)." Radiology: Cardiothoracic Imaging 2.2 (2020): e200110.

[38] Dlamini, Wisdom M., et al. "Spatial risk assessment of an emerging pandemic under data scarcity: A case of COVID-19 in Eswatini." Applied Geography 125 (2020): 102358.

[39] Hiransha, M., Gopalakrishnan, E. A., Menon, V. K., & Soman, K. P. (2018). NSE stock market prediction using deep-learning models. Procedia computer science, 132, 1351-1362.

[40] Goshima, K., & Takahashi, H. (2016). Quantifying news tone to analyze the Tokyo Stock Exchange with deep learning. Security Analysis Journal, 54(3), 76-86.

[41] Nabipour, M., Nayyeri, P., Jabani, H., Mosavi, A., & Salwana, E. (2020). Deep learning for stock market prediction. Entropy, 22(8), 840.

[42] Wang, Jialan, et al. "Bankruptcy and the COVID-19 Crisis." Available at SSRN 3690398 (2020).

[43] Donthu, Naveen, and Anders Gustafsson. "Effects of COVID-19 on business and research." Journal of business research 117 (2020): 284.

[44] Sharma, Abhinav, and Juan Luis Nicolau. "An open market valuation of the effects of COVID-19 on the travel and tourism industry." Annals of tourism research (2020).

[45] Ozili, Peterson K. "Covid-19 pandemic and economic crisis: The Nigerian experience and structural causes." Journal of Economic and Administrative Sciences (2020).

[46] Bendavid, Eran, et al. "Assessing mandatory stay?at?home and business closure effects on the spread of COVID?19." European journal of clinical investigation 51.4 (2021): e13484.

[47] Trogen, Brit, and Arthur Caplan. "Risk Compensation and COVID-19 Vaccines." (2021).

[48] Kissler, Stephen M., et al. "Social distancing strategies for curbing the COVID-19 epidemic." medRxiv (2020).

[49] Long, J., Khaliq, A. Q. M., & Furati, K. M. (2021). Identification and prediction of time-varying parameters of COVID-19 model: a data-driven deep learning approach. International Journal of Computer Mathematics, (just-accepted), 1-19.

[50] Yegnanarayana, Bayya. Artificial neural networks. PHI Learning Pvt. Ltd., 2009.

[51] Medsker, Larry R., and L. C. Jain. "Recurrent neural networks." Design and Applications 5 (2001).

[52] Univercity of Eswatini, https://datastudio.google.com/reporting /b847a713-0793-40ce-8196-e37d1cc9d720 /page/2a0LB