Forecasting confirmed COVID-19 cases in Bangladesh using SARIMA, LSTM and SIR model and their comparison

Submitted By

Name: Nazmus Sakib

ID: 2017-2-54-011

and

Name: Akib Naved

ID: 2016-2-50-033

This Thesis Paper is Submitted in Partial Fulfillment of the Requirements of the Degree of Bachelor of Science in " Informations & Communications Engineering".

Department of Electronics & Communication Engineering

East West University

Approval

The thesis paper titled “Forecasting confirmed COVID-19 cases in Bangladesh using SARIMA, LSTM and SIR model and their comparison” submitted by Nazmus Sakib (ID: 2017-2-54-011) and Akib Naved (ID: 2016-2-50-033) to the Department of Electronics and Communications Engineering, East West University, Dhaka, Bangladesh has been accepted as satisfactory for the partial fulfillment of the requirements for the degree of Bachelor of Science in Information and Communications Engineering and approved as to its style and contents.

Approved By

____

Supervisor
S. M. Raiyan Chowdhury
Lecturer
Department of Electronics Communications Engineering
East West University
Dhaka, Bangladesh

Declaration

We declare that our work has not been previously submitted and approved for the award of a degree by this or any other University. As per of my knowledge and belief, this paper contains no material previously published or written by another person except where due reference is made in the paper itself. We hereby, declare that the work presented in this thesis paper is the outcome of the investigation performed by us under the supervision of S. M. Raiyan Chowdhury, Lecturer, Department of Electronics Communications Engineering, East West University, Dhaka, Bangladesh.

Countersigned

____

Supervisor name

S. M. Raiyan Chowdhury

Signature

____

Name: Nazmus Sakib

ID: 2017-2-54-011

____

Name: Akib Naved

ID: 2016-2-50-033

Acknowledgement

We would like to express our deepest gratitude to Our thesis supervisor, S. M. Raiyan Chowdhury for guiding us with valuable advice during our thesis investigation. His guidance, support and encouragement helped us in completing this thesis paper. Thanks to the East West University for numerous courses that helped us complete this project. And finally, thanks to parents, and numerous friends who endured this long process with us, always offering support and love.

Abstract

Every year there has been a new variant of Covid-19 that has been disrupting social lives, taking the lives of human and damaging economy of many countries. So it has been a crucial role for policymakers to make policies based on how the Covid-19 will impact the future. By understanding the forecasting of epidemic developments, policymakers can make policies that can reduce the severity of the impact the policies do to the economy of the country. There has been many researches regarding this topic. Among many forecasting methods, three methods have been mentioned the most by the scholars. They are SARIMA model, LSTM model and one of the Compartment Model of Epidemiology known as SIR model. Here we will compare these three models and decide which is the best for forecasting of Covid cases.

Keywords: Epidemiology, Neural Networks, Forecasting, LSTM, SIR, Compartmental Model, ARIMA model, SARIMA model, Comparison, Bangladesh

Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Background

Coronavirus is from the family of viruses of MERS (Middle East Respiratory Syndrome) and SARS (Severe Acute Respiratory Syndrome)[1]. The sizes of these viruses are 80 to 120 nm in diameter which is said to be one of the largest among RNA viruses[2]. This is an animal virus that can be found mostly in bats. Coronavirus 2019 which is also known as COVID 19 is a new strain of coronavirus which started to spread during late 2019. First case of Covid-19 was found in Wuhan, Hubei Province of China[3]. After that this pandemic began to spread out exponentially and became one of the most dangerous and fatal pandemics in human history[4].

1.2 Problem statement

As the number of Covid cases are increasing rapidly, it has become necessary to take measures to slow down the spreading of the virus and prepare for the worst case scenario just in case it doesn't slow down. Various methods such as deep learning models, time forecasting methods, mathematical models are being used to forecast the confirmed cases and help people understand

and manage this situation. But due to virus mutating constantly, it has become harder to forecast confirmed cases. Many scholars are now focused on innovation that is needed for analysis and prediction of epidemic trend of Covid-19. Some machine learning RNN models such as LSTM and GRU are the most commonly used models in prediction of the cases because the dataset is sequential data.

1.3 Motivation

There is a record of large scale pandemic which happened in 1920s, known as spanish flu[5]. But since it was during the era when data related to the pandemic was mainly recorded in paper, the accessibility to the data was low and it was difficult to do research on those data as well. Almost 100 years later, we have another large pandemic now, which is known as Covid-19 . But this is the digital era where everything is recorded and is stored in virtual storage. Due to this, the demand of the research related to Covid has also skyrocketed. Many research has already been published related to Covid-19. One of such problems that has been researched upon is the forecasting of the confirmed cases. With the accurate forecasting of Covid-19 trends, many problems can be reduced such as preparing more hospital beds if the pandemic becomes worse in the future, the severity of different measures taken by government that also impacts economy of the country can be reduced if pandemic becomes better etc. This is why we decided to choose this as a topic we should be researching upon.

1.4 Thesis organization

This thesis consists of 9 chapters. Short summary of each chapter is given below:

- In chapter 1, we introduced our topic, which problem we are focusing on and what is our motivation to do this research.

- In chapter 2, we gave a short summary of similar papers that also did a research on the related topic.
- In chapter 3, we discussed shortly about deep learning and RNN.
- In chapter 4, we discussed the history of SIR model in short and the equations of SIR.
- In chapter 5, we discussed about the seasonal time forecasting method known as SARIMA.
- In chapter 6, we discussed about how we conducted the experiment and the reliability of the dataset.
- In chapter 7, we changed some parameters of each of the forecasting method and decided which fits the dataset best.
- In chapter 8, we varied the training and testing dataset for each model and then showed and compared the predictions.
- In chapter 9, we focused on conclusion, our research limitations and the future scope of our model.

Chapter 2

Related Works

Yi-Cheng Chen, Ping-En Lu[6] proposed and implemented time dependent SIR model for different countries such as USA, UK, France, Iran, Spain, Italy, Germany, Republic of Korea. They were able to show that their model is better than traditional SIR model and also more robust.

Lin Feng [7] implemented hybrid of SEIR and LSTM/GRU model to show prediction of transmission rate, recovery rate and reproduction number. The results showed that the variance of the parameters predicted by the hybrid model is smaller than theoretical value and doesn't fluctuates much.

Zhifang Liao, Peng Lan [8] Implemeneted another variant of SIR model which is known as SIRVD and Combined it with different variants of deep learning methods such as Vanilla LSTM, Stacked LSTM, Bidirectional LSTM and GRU to predict confirmed cases which showed promising results for short term and medium term predictions.

Yi Kang [9] proposed a SEIR model to predict infected cases that was adopted with substantial public health interventions which performed better than SEIR model without substantial public health interventions and LSTM RNN model. The researcher emphasized on the importance of public health interventions to reduce the spreading of covid in early stages to reduce the number of infections.

Philip Nadler, Rossella Arcucci [10] proposed a neural SIR model and compared it with dynamic

SIR, univariate LSTM and multivariate LSTM models. The Neural Sir Model performed better in most of the comparisons. This research paper showed that how accurate the forecasting can be if the forecasting accuracy of LSTM networks is combined with epidemiological model dynamics of the SIR Model.

Shuo Feng, Zebang Feng [11] implemented a hybrid model of SEIR model and vanilla RNN model and a simple DNN model to predict the trends. Due to the virus variation, the research had some limitations. Even with limitations SIR predicted well for short term predictions in shorter area with smaller populations and with the AI model predicted well for short term predictions in larger area with larger population.

Sarbhan Singh, Bala Murali Sundram [12] forecasted daily confirmed cases using ARIMA models. They choose a suitable ARIMA model by using three stages of verification - model identification, parameters estimation and model verification. This research showed the effectiveness of ARIMA models during early stages of pandemic. Despite having limited data points, ARIMA models proved to be effective and since it's the easiest method than other models, it is easier to implement.

Vasilis Papastefanopoulos, Pantelis Linardatos [13] compared different time series method for forecasting such as ARIMA, Prophet, HWAAS, NBEATS, Gluonts and TBAT. According to the study TBAT performed significantly better than other models such as Prophet, Gluonts and N-BEATS. But since not every country has the same condition as others, the performance for each country was not good as expected. This research also proposed that population related statistics and measurement of social distancing needs to be implemented for a better forecasting.

Anjir Ahmed Chowdhury, Khandaker Tabin Hasan [14] used LSTM and ANFIS network to predict and analyse covid in Bangladesh. ANFIS is a modified variant of Artificial Neural Network(ANN). It is one of the high performing neural networks in both computing and techniques that are used for learning non-linearity. According to the result LSTM performed better than ANFIS, But due to the lack of validation regarding infected cases in Bangladesh, they couldn't ensure the validity of results.

Babacar Mbaye Ndiaye, Lena Tendeng [15] implemented a hybrid model of SIR and a machine learning technique known as prophet to forecast the Covid cases. it is a procedure for forecasting time series data where non-linearity trends are usually fit with yearly, weekly or daily seasonality. The forecasting results are satisfactory but still needs to be improved with the better grasp on the input behaviour.

Farah Shahid, Aneela Zameer [16] used different deep learning model for predictions such as LSTM, Bidirectional LSTM and GRU including other time forecasting methods such as SVR and ARIMA and then forecasted cases for different countries. The results show deep learning models performed better than SVR and ARIMA models and Bi-LSTM outperforms every other models including deep learning models.

Sharif Noor Zisad , Mohammad Shahadat Hossain [17] proposed a variation of the SEIR model combined with LSTM. In this paper, they added another differential equation for quarantined compartment to generalized SEIR model. The equation for the basic reproduction number of this model is also different than generalized equation. With results, they were able to prove that their proposed model works better than traditional SEIR model.

Chapter 3

Deep Learning

3.1 Introduction

As we progress more into the modern era, Deep Learning is becoming more and more important. Amidst the ongoing crisis of COVID Pandemic, Deep learning also has been used for many problems related to COVID. Such as chatbots, covid detection using lung ultra-sound imagery, forecasting covid cases and sentimental analysis on covid. There are many others but these four uses are the most popular use of deep learning. Below we will give a summary for each of them:

1. **Making chatbots:** chatbots are gaining popularity now-a-days in different fields such as healthcare, education, marketing, entertainments etc. It is to manage the huge number of user requests during pandemics. It has lessened the workload of doctors by a lot while helping the patients indiscriminately. One of these chatbots is made by some researchers using the BERT Model which is already pre-trained and developed by Google researchers[18].

2. **Detecting Covid:** This is more important than others because of how fast the pandemic is spreading. to detect Covid, radiological images of the chest and RT-PCR test can be used also. So this falls under image classification problem and CNN is specially known to solve such problems with precision. One of these neural model is developed by researchers using

ResNet50 models[19].

3. **Sentimental Analysis:** People all over the world have different sentiments regarding Covid because it affected many people's mind by isolating them. And also people's sentiment regarding Covid vaccination is changing . One of the research was done on Indian people to analyse their sentiment regarding Covid[20].

4. **Forecasting Covid cases:** This is also an important use of deep learning because Government, Medical Organization and hospitals have to prepare in advance against many waves of Covid pandemic. But it has becoming more and more difficult to predict with precision because of the virus mutation. New variants of Covid is being detected every year. LSTM which is known for forecasting different time series data can also be used for forecasting Covid cases. But before we discuss about LSTM, we will discuss in details about ANN, RNN and the problems of RNN and how LSTM overcomes those problems.

3.2 Artificial Neural Network:

Artificial Neural Network(ANN) is one of the basic concept of deep learning and this is based on the idea of human brain. Just like how human brain makes decision based on what they see and feel, ANN also tries to simulate such operation in a virtual environment. The process of ANN maybe similar to a biological neuron but they are not completely similar. Most of ANN have three common layers to do these operation. These are: a input layer, hidden layers and a output layer. Input layer takes information, hidden layers process those information and output layer makes prediction. There can be one or multiple neurons in these layers. Each of these layers can be fully connected to each other or partially connected to each other. The connections are assigned a weight which represents the strength of the connection of the nodes that are connected. A basic model of ANN is given below:

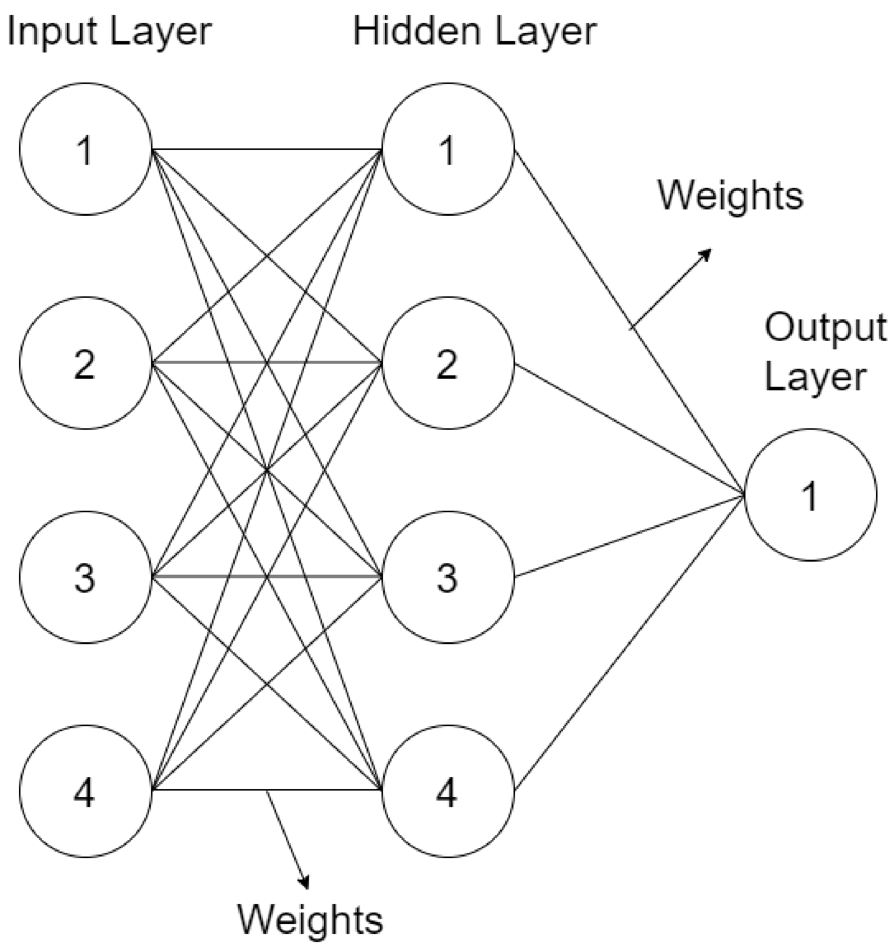

Figure 3.1: A basic three layered ANN model

Since most of our decisions in our life represent a non-linear information, ANN also has to adapt to such non-linearities. To adapt to such non-linearities, activation functions are used in ANN. Two of the activation functions that are used in this research are sigmoid function and hyperbolic tangent function.

Sigmoid Function: This is one of the basic activation function of ANN. After taking input from the node this function maps it to between 0 and 1. This output can be the importance of the information of input or the probability of the input belonging to a certain class.

$$f(x) = \frac{1}{1 + e^{-x}} \tag{3.1}$$

Hyperbolic tangent function(tanh): The tanh activation function is slightly different from sigmoid function. This activation function maps the outputs between -1 and 1. This is better than sigmoid since the output provided by it is zero centered.

$$f(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}} \tag{3.2}$$

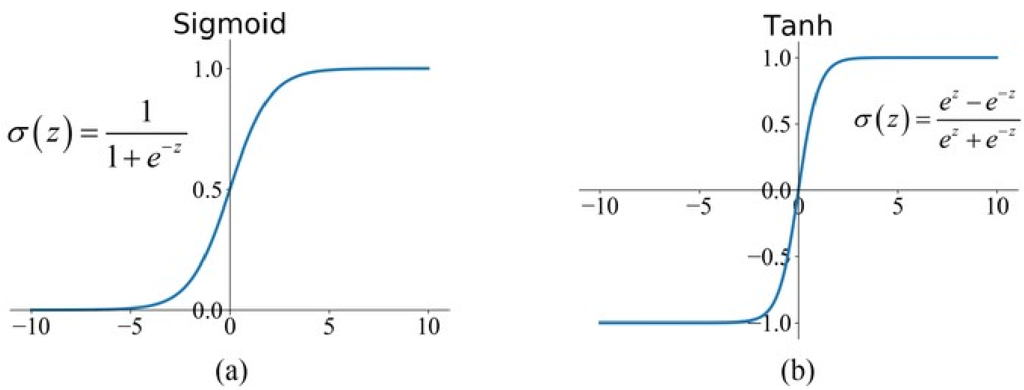

Figure 3.2: Image of (a) sigmoid and (b) tanh activation function[21].

3.3 Recurrent Neural Network:

Recurrent Neural Network(RNN) is a variation of ANN which has memory. RNN model can save relevant data from past information and predicts output based on those past data. This is why this model is called Recurrent neural network. In conventional ANN time or sequence doesn't matter so it simply gives output for the input that has been taken at that time. Inputs taken at different time has no connection between them. Hence for problems that depend on time and sequence, ANN is not practical. This is where RNN becomes useful. RNN basically takes inputs in different time or sequence and gives output based on them while forwarding the data to the next node in the layer. RNN has been used for many sequence related problem such as real-time learning control scheme[22], forecasting of solar radiation[23], sentimental analysis, chat-bots, machine translation etc.

Figure 3.3 represents the single neuron of a RNN layer which repeats after a time. here x_t is the present input and $W_{x(t)}$ is the corresponding weight matrix. The present output depends on both present input and memory of the previous input. After producing the present output, the memory of the present input is passed onto next node for future predictions.

RNN model is used specially for problems related to sequential data[24][25][26]. But RNN has problems known as vanishing gradient problem and exploding gradient problem [27]. Due to this, RNN can't process long term memory information and since our data is a long sequence of Covid cases, the vanilla RNN won't be that much effective [28]. Thats why we can use another variant of the model known as LSTM which solves the problem of RNN effectively[29].

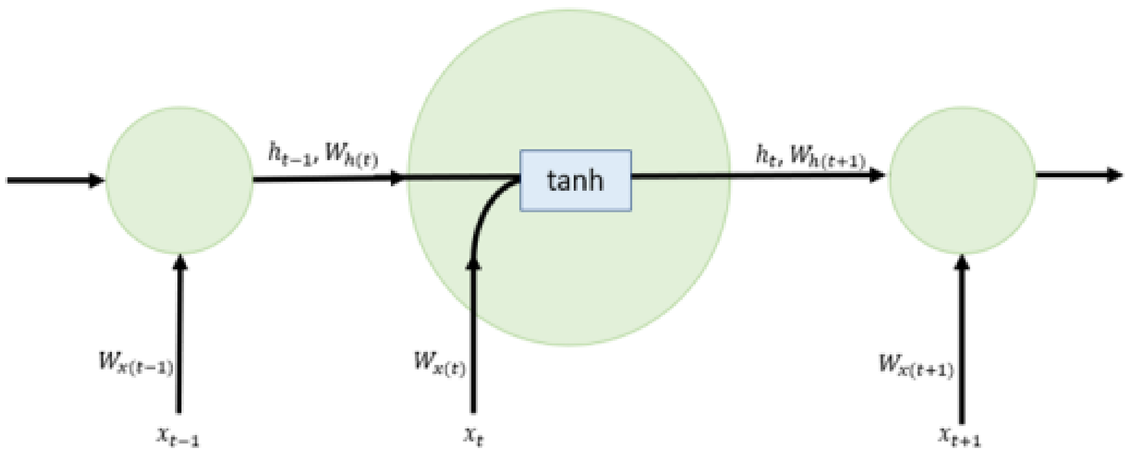

Figure 3.3: The repeating module in a RNN network [7]

3.4 Long-Short-Term-Memory:

Long-Short-Term-Memory neural network model is another variant of RNN, This model solves the problem of short term dependency of RNN and is specially good for long sequence of data. There are four gates which defines the LSTM network. Cell state keeps only relevant data from previous states and forwards these data to future states and Gates decide which data to forget and which data to keep in the memory. The information that are being stored or being forgotten depends on relevance of the data to the newest output. There are mainly four gates in a single LSTM neuron. These are forget, input, cell and output gate.

In the figure 3.4 we can see the representation of LSTM's single neuron. Input here is denoted by $x(t)$, the hidden state from the previous timestep is denoted by $h_{(t-1)}$ and the cell state from previous timestep which is also the memory of LSTM is denoted by by $C_{(t-1)}$. The forget gate, input gate, cell gate and output gate is represented by f_t, i_t, C_t, o_t. The hidden state of the present timestep h_t and cell state of the present timestep C_t forwards the relevant data to the future timesteps. Below we will talk about how each gate works in short:

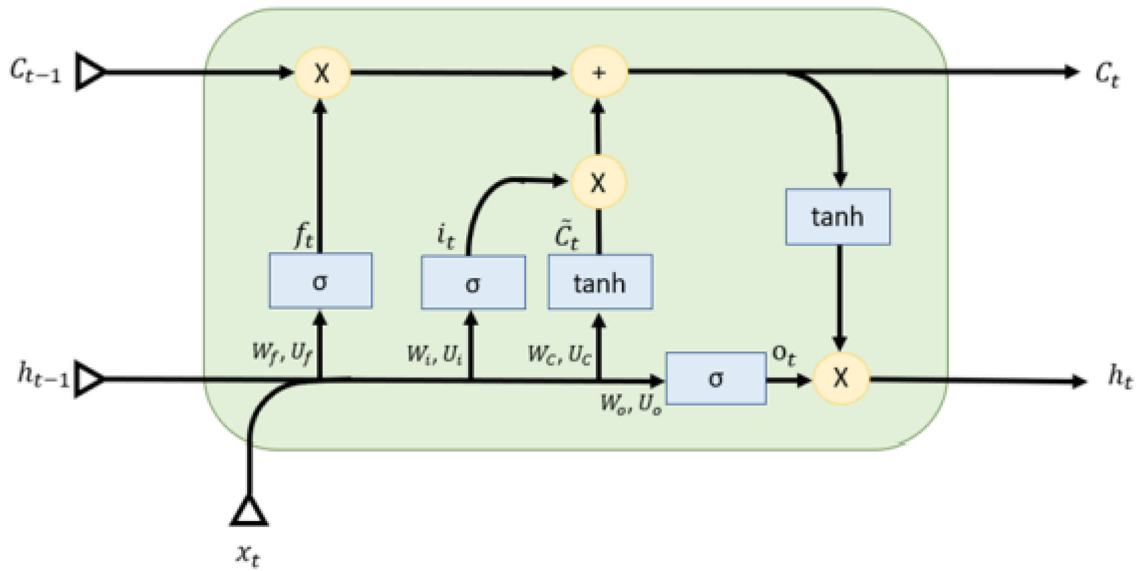

Figure 3.4: The repeating module in a LSTM Network [7]

1. **Forget gate:** This gate decides which data to keep in memory and which data to forget. That's why sigmoid function is used here, The data that are being forgotten is multiplied by 0 and the data or the portion of the data that is being stored is multiplied by non-zero values.

$$f_t = \sigma(W_f[h(t-1), x_t] + b_f) \tag{3.3}$$

2. **Input gate:** This gate decides which data needs to be updated. Here first the data goes through both sigmoid and tanh functions and then both of them are multiplied and forwarded to the cell state.

$$i_t = \sigma(W_i[h(t-1), x_t] + b_i) \tag{3.4}$$

$$\hat{C}_t = \tanh(W_c[h(t-1), x_t] + b_c) \tag{3.5}$$

3. **Cell gate:** This gate updates the cell state from previous timestep and forwards the new cell state to the future timestep.

$$C_t = f_t * C(t-1) + i_t * \hat{C}_t \tag{3.6}$$

4. **Output gate:** This gate predicts and gives output of the current timestep based on present cell state and present input. The output multiplied with filtered cell state is forwarded to the next timestep as hidden state.

$$O_t = \sigma(W_o[h(t-1), x_t] + b_o) \tag{3.7}$$

$$h_t = O_t * \tanh(C_t) \tag{3.8}$$

Chapter 4

Epidemiology and SIR model

4.1 Introduction

SIR model was developed first by Kermack and McKendrik in 1927 [30]. This model used the data from Indian plague epidemic in the early 1900s. When the pandemic started British Empire organized an Indian Plague Commission for studying the plague. When many researchers were trying to build a model for Epidemic, Kermack and McKendrick released a paper in 1927. They built this model using some of the mathematical derivations. Single variable calculus was used to solve the equations. This paper also included a section where they added a figure that they got after using the data from Indian plague pandemic[31]. After that students saw the potential of the model and they were also researching the model to improve it. Before covid this model was used for researching Ebola in West Africa. After this model became well-known, Many variations of this model were developed. For example SIS Model, SIRD Model, SIRV Model, MSIR model, SEIR model, SEIS model etc [32]. This model is now used to simulate different scenerios of covid such as spread of the disease, case prediction, determining parameters such as transmission rates and recovery rates etc.

4.2 SIR Model(Kermack-McKendrick Model):

This is one of the most basic of SIR models and it is also known as Kermack-McKendrick Model.

Figure 4.1: SIR (Susceptible, Infected, Recovered) model

In this model, there are three dependent variables $S(t)$, $I(t)$ and $R(t)$ which are controlled by three differential equations below:

$$\frac{dS(t)}{dt} = -\frac{\beta \cdot S(t) \cdot I(t)}{n} \tag{4.1}$$

$$\frac{dI(t)}{dt} = \frac{\beta \cdot S(t) \cdot I(t)}{n} - \gamma \cdot I(t) \tag{4.2}$$

$$\frac{dR(t)}{dt} = \gamma \cdot I(t) \tag{4.3}$$

- Here, S(t) represents the percentage of population that is exposed to the disease,
- I(t) represents the percentage of population that is already infected with disease,
- R(t) represents the percentage of population that has recovered from the disease and has an immunity against the disease
- β represents the disease transmission rate
- γ represents the recovery rate
- n is the total population

We can also note that

$$S(t) + I(t) + R(t) = n \tag{4.4}$$

Another important parameter, called basic reproduction number, is given by:

$$R_0 = \frac{\beta \cdot S(t)}{\gamma} \tag{4.5}$$

It is a well known value in epidemiology. This reproducing number is also used to determine if the disease begins spreading in a population or not. If $R_0 > 1$ then infection will start spreading and if $R_0 < 1$, then infection will not start spreading[33]. This model neglects the time varying property of β and γ. Due to this, it is difficult to predict using the traditional SIR model. This is why we will use a time-dependent model proposed by Yi-Cheng Chen, Ping-En Lu [6].

4.3 Discrete Time-dependent SIR Model:

In this model, transmission rate $\beta(t)$ and recovery rate $\gamma(t)$ are assumed to be also functions of time. So if we replace β with $\beta(t)$ and γ with $\gamma(t)$ in the differential equations 4.1, 4.2 and 4.3, we get the equations below:

$$\frac{dS(t)}{dt} = -\frac{\beta(t) \cdot S(t) \cdot I(t)}{n} \tag{4.6}$$

$$\frac{dI(t)}{dt} = \frac{\beta(t) \cdot S(t) \cdot I(t)}{n} - \gamma(t) \cdot I(t) \tag{4.7}$$

$$\frac{dR(t)}{dt} = \gamma(t) \cdot I(t) \tag{4.8}$$

the above equations will still satisfy the equation 4.4 which means the sum of S, I and R will still be equal to total population n. Now we can assume time to be a discrete variable because we are predicting for cases on daily basis so time variable isn't continuous anymore. So if we use the equations above for discrete time variable, then the equations 4.6, 4.7 and 4.8 become

$$S(t+1) - S(t) = -\frac{\beta(t) \cdot S(t) \cdot I(t)}{n} \tag{4.9}$$

$$I(t+1) - I(t) = \frac{\beta(t) \cdot S(t) \cdot I(t)}{n} - \gamma(t) \cdot I(t) \tag{4.10}$$

$$R(t+1) - R(t) = \gamma(t) \cdot I(t) \tag{4.11}$$

Here since the infection spread is very low at start, we can assume $S(t)$ = n, so equation 4.10 becomes

$$I(t+1) - I(t) = \beta(t) \cdot I(t) - \gamma(t) \cdot I(t) \tag{4.12}$$

From equation 4.11 we get the derivation of $\gamma(t)$ and from equation 4.11 and 4.12 we get the derivation of $\beta(t)$:

$$\gamma(t) = \frac{R(t+1) - R(t)}{I(t)} \tag{4.13}$$

$$\beta(t) = \frac{[I(t+1) - I(t)] + [R(t+1) - R(t)]}{I(t)} \tag{4.14}$$

We track transmission rate $\beta(t)$ and recovery rate $\gamma(t)$ using ridge regression and then train the ridge regression. Then we predict $\beta(t)$ and $\gamma(t)$ using ridge regression and with these predicted $\beta(t)$ and $\gamma(t)$ values, we can predict confirmed cases of next day using below equations which can again be derived from 4.13 and 4.14:

$$I(t+1) = (1 + \beta(t) - \gamma(t)) \cdot I(t) \tag{4.15}$$

$$R(t+1) = R(t) + \gamma(t) \cdot I(t) \tag{4.16}$$

4.4 Ridge Regression:

Ridge regression is used to estimate the coefficients of multiple regression models[34]. This method is specially useful for models that have multiple highly correlated independent variable. it was first introduced by Hoerl and Kennard [[35], [36]] and is specially popular for solving problems that have multi-collinear variables. In these problems the ordinary least squares(OLS) unbiased estimators can become a huge problem if they become unstable due to large variance[37]. Ridge regression includes a penalty or tuning parameter which is known as alpha term. By tuning this alpha term we can achieve accurate predictions. In Intra Model Analysis chapter we used different values of alpha to achieve precise estimation. Below with figures we will show how it affects the prediction(the alpha parameters for beta and gamma prediction in each SIR model is given in 7.4 table):

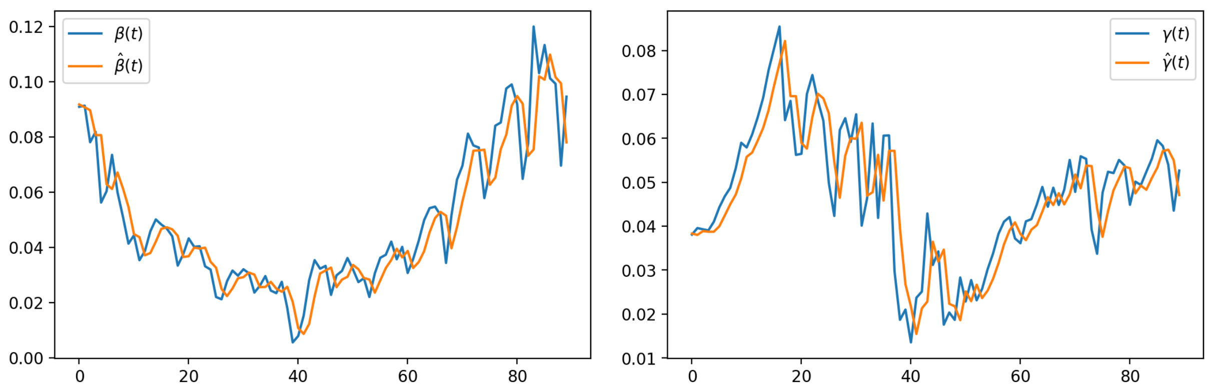

Figure 4.2: Comparison of predicted beta and gamma using SIR1

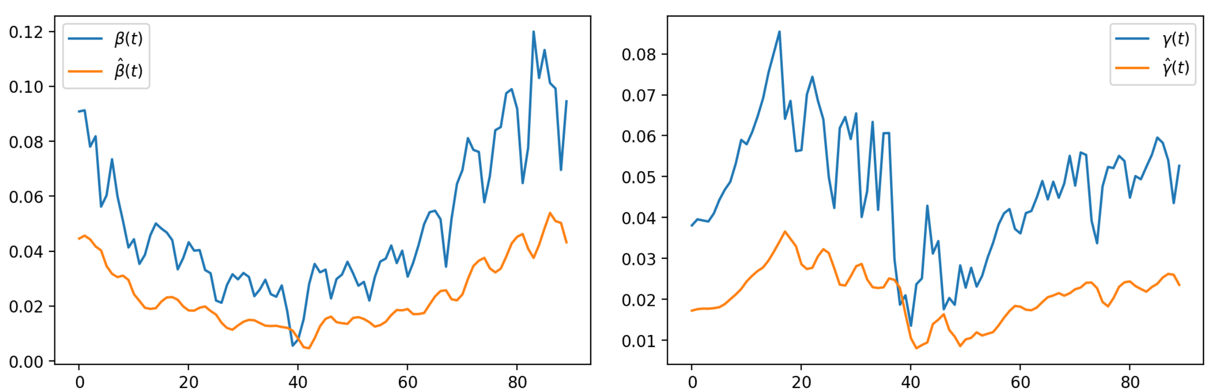

Figure 4.3: Comparison of predicted beta and gamma using SIR2

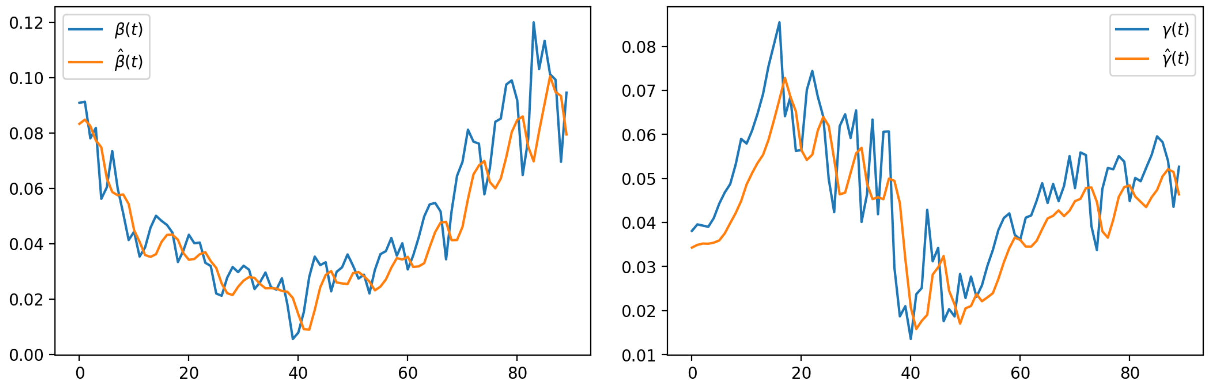

Figure 4.4: Comparison of predicted beta and gamma using SIR3

In above figures blue lines are the actual beta and gamma and orange lines are predicted beta and gamma values.

Chapter 5

SARIMA Model

5.1 Introduction

Analysis of time series data has become important recently because of the large amount of data we have accessed to now-a-days. It has been used in fields such as medicines, weather, economics, astronomy etc[38]. Due to the importance of accurate time forecasting , many forecasting method was introduced. ARIMA is one of the most well known time series analysis tool that has been used in many forecasting such as rainfall forecasting[39], aircraft failure forecasting [40] etc. It was first introduced first by Box and Jenkins in 1970[41]. It is also known as Box-Jenkins model. During that time this time analysis tools has been described as revolutionary in the studies related to time forecasting. If some requirements are met, this model can give optimal outputs. SARIMA is a modification of ARIMA, When a time series dataset shows a seasonal pattern, SARIMA is used. otherwise ARIMA is used. And in our dataset we see a seasonal pattern. At the beginning of every year, there is a sharp rise in confirmed cases and at the end of the year the rise decays. Since our dataset shows a seasonal pattern, we will use SARIMA model. But first we need to learn how SARIMA works so we will introduce ARIMA first, then will talk about SARIMA.

5.2 ARIMA model

Here in ARIMA,

- AR stands for Auto Regression, it can be symbolized by p, this p depends on past values. Let's assume Y_t is the current value and the past values which it depends on are Y_{t-1} , Y_{t-2}, Y_{t-3} etc. So we can express Y_t as $Y_t = f(Y_{t-1}, Y_{t-2}, Y_{t-3}, .., \epsilon_t)$ where ϵ_t is the error term. A common representation of AR(p) model can be represented below :

$$Y_t = \beta_0 + \beta_1 * Y_{t-1} + \beta_2 * Y_{t-2} + \beta_3 * Y_{t-3} + ... + \epsilon_t \tag{5.1}$$

- MA stands for moving average, it can be symbolized by q, this q depends only on random error terms. Let's assume Y_t is the current value and the error terms which it depends on are ϵ_{t-1} , ϵ_{t-2}, ϵ_{t-3} etc. So we can express Y_t as $Y_t = f(\epsilon_{t-1}, \epsilon_{t-2}, \epsilon_{t-3}, ...)$. A common representation of MA(q) model can be represented below :

$$Y_t = \beta_0 + \epsilon_t + \phi_1 * \epsilon_{t-1} + \phi_2 * \epsilon_{t-2} + \phi_3 * \epsilon_{t-3} + ... + \phi_q * \epsilon_{t-q} \tag{5.2}$$

- I stands for Integrated, it can be symbolized by d, this d determines how much degree of difference is needed for making the dataset stationary.

Combining the concepts of AR,I and MA, we get the equation below as the representation of ARIMA(p,d,q) model:

$$Y_t = \beta_0+\beta_1*Y_{t-1}+\beta_2*Y_{t-2}+\beta_3*Y_{t-3}+....+\beta_p*Y_{t-p}*\epsilon_t+\phi_1*\epsilon_{t-1}+\phi_2*\epsilon_{t-2}+\phi_3*\epsilon_{t-3}+....+\phi_q*\epsilon_{t-q} \tag{5.3}$$

There is also a term known as white noise, it is a time series dataset with mean value = 0, constant standard deviation and correlation lags = 0. it is purely random in nature and ARIMA model can't be used forecasting for this dataset[42].

5.3 SARIMA model

SARIMA(p,d,q)(P,D,Q)[S] which is also known as Seasonal ARIMA is a ARIMA model where both seasonal and non seasonal factors are considered. It includes both seasonal and non-seasonal components in a multiplicative method. Here (p,d,q) are consecutively non-seasonal AR order, non-seasonal differencing and non-seasonal MA order and (P,D,Q) are consecutively seasonal AR order, seasonal differencing and seasonal MA order and S is the number of periods per season. For our dataset we used S=12 because our forecasting shows a monthly seasonality.

Chapter 6

Methods and Materials

6.1 Dataset:

After the covid-19 pandemic has started governments have been constantly releasing statistical reports regarding covid cases and all of these statistical reports are public so that researchers can do analysis on these datas to predict how this virus will act in future and give a warning if there is a possibility that the situation will become worse. Worldometer collects most of these data such as daily cases and total cases etc. from all over the country[43]. Center for Systems Science and Engineering (CSSE) at JHU(John Hopkins University) also has data repository for such datasets and they are also recorded in a time series format which is beneficial for us as we are forecasting the confirmed cases of covid in Bangladesh [44]. The dataset ranges from March 8, 2020 to July 4, 2021. There are 530 daily records of deaths, recovered and confirmed cases in Bangladesh. For training and testing different models, dataset will be split and varied. The dataset will have four columns containing dates, confirmed, recovered and deaths. the following table is the preview of the dataset:

	A	B	C	D
1	Date	Confirmed	Recovered	Deaths
2	6/1/2021	802305	742151	12660
3	6/2/2021	804293	744065	12694
4	6/3/2021	805980	746035	12724
5	6/4/2021	807867	747758	12758
6	6/5/2021	809314	749425	12801
7	6/6/2021	810990	751322	12839
8	6/7/2021	812960	753240	12869
9	6/8/2021	815282	755302	12913
10	6/9/2021	817819	757569	12949
11	########	820395	759630	12989
12	########	822849	761916	13032
13	########	824486	764024	13071
14	########	826922	766266	13118
15	########	829972	768830	13172
16	########	833291	771073	13222
17	########	837247	773752	13282
18	########	841087	776466	13345
19	########	844970	778421	13399
20	########	848027	780146	13466
21	########	851668	782655	13548
22	########	856304	785482	13626
23	########	861150	788385	13702
24	########	866877	791553	13787
25	########	872935	794783	13868
26	########	878804	797559	13976
27	########	883138	800854	14053
28	########	888406	804103	14172

Figure 6.1: Selected dataset

6.2 Data Preprocessing:

This pre-processing is only for the LSTM model to ensure that we train our model properly because unscaled data with large range of values can slow down learning rate and also the convergence of our neural network can result in bad outcome. There is a built in function in python library just for this purpose. With this we will scale each of the numerical value for the input to range [0-1], The equation minmaxscaler function in python uses to scale the data is given below:

$$x_{scaled} = \frac{x - x_{min}}{x_{max} - x_{min}} \tag{6.1}$$

Here,

1. x_{scaled} is our new value

2. x is the original cell value

3. x_{min} is the minimum value of the column

4. x_{max} is the maximum value of the column

6.3 Applying algorithms for different models:

Now moving onto the most vital part of our research methodology which is to apply the three models we have selected for comparison. These models are LSTM, Time dependent SIR model and Seasonal ARIMA. SARIMA is a statistical analysis model, LSTM is RNN (recurrent neural network based) architecture and SIR model is a mathematical model for infectious disease. All of them can be used to predict confirmed cases but their theoretical basis for predicting is different from each other. After the results are acquired from the different models, we will discuss it in the result and analysis chapter and compare them to find the most suitable model that can predict the confirmed cases more accurately. flowchart below shows how we will do the research:

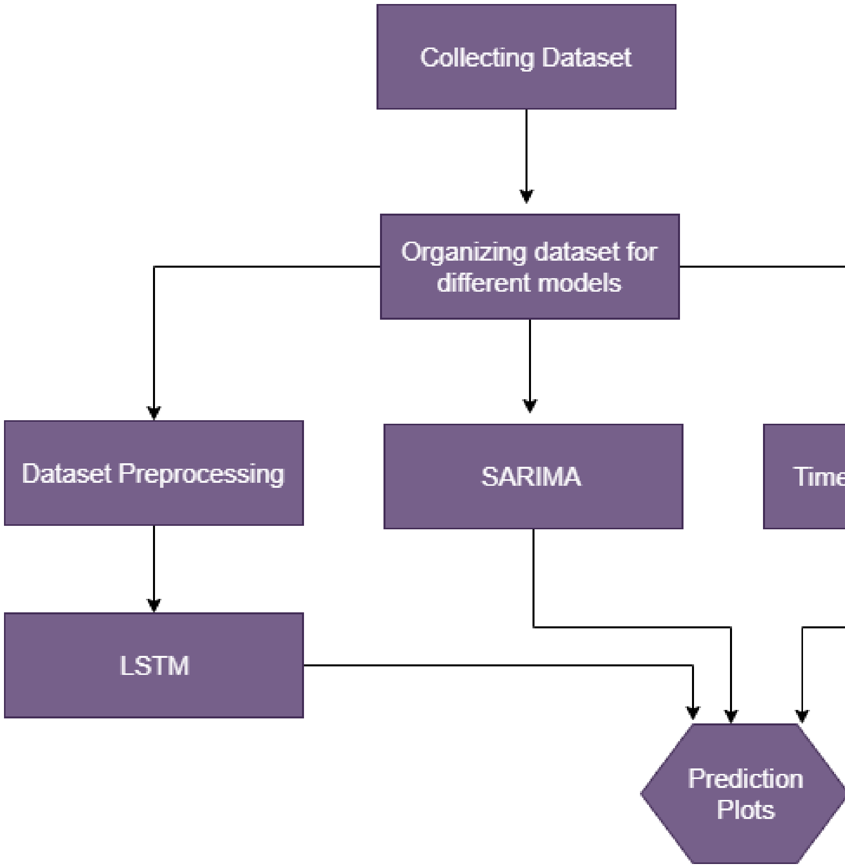

Figure 6.2: Research Methodology

6.3.1 Selected models from LSTM, SIR and SARIMA:

We have selected LSTM3, SIR1 and SARIMA1 based on their performance. Details of these models have been discussed in the next chapter.

6.3.2 Codes:

Codes for these models are publicly available on github: SIR Model [45], LSTM Model[46], SARIMA Model[47]

Chapter 7

Intra Model Analysis

7.1 Introduction

To compare each model, first we have to find best of each model. Without properly evaluating the parameters of each model, we won't be able to compare them properly. That's why we will take some models with different parameters. For example, for LSTM we can increase, decrease number of layers and neuron units to check what works the best and using that model we will compare with other models later. Here each model is unique. ARIMA is one of the time series analysis tool in statistics study. Seasonal ARIMA is the improved version of ARIMA that forecasts based on seasonal elements of the dataset. Here we will vary the values of (p,d,q) and (P,D,Q,m) for different models to compare them and select the best one out of them. SIR is a compartmental model in epidemiology and it is mainly a mathematical model that tries to predict things as how the disease spreads or the total number of infected people or the duration of the epidemic disease [32]. Here we used ridge regression to estimate beta(transmission rate) and gamma(recovering rate) and predict the confirmed cases based on that. Here we can vary the the parameters of ridge regression such as alpha to differentiate the models and compare them. Below we are giving details of comparisons of the intra model analysis.

7.2 Simple LSTM Model Comparison:

Here, For optimizer, we selected adam as it is one of the most used and most reliable among the optimizers. Adam combines the benefit of both AdaGrad(Adaptipe Gradient) and RMSProp (Root Mean Square Propagation) to converge quickly as possible and get the desired output in short amount of time [48]. Since we are trying to solve a regression problem, we used MSE(Mean Square Error) as our loss function. For LSTM and dense layers we used tanh activation functions. The reason for using tanh is that research shows that tanh still gives the best output compared to others[49]. There will be one output layer with one neuron to predict the results. We are using 50 epochs and 5 steps per epoch for the LSTM models. We will take 3 models and we will be changing number of layers and the types of layers(Dense or LSTM). Below we are giving summary of each model:

LSTM Model	Layers	Units in each layer	Total Parameters
LSTM1	LSTM,LSTM,Dense,Dense	150,64,64,1	150465
LSTM2	LSTM,LSTM,LSTM,Dense	150,64,64,1	179329
LSTM3	LSTM,LSTM,Dense	150,64,1	146305

Table 7.1: LSTM1, LSTM2 and LSTM3 details

7.2.1 Prediction and error comparison of the LSTM Models:

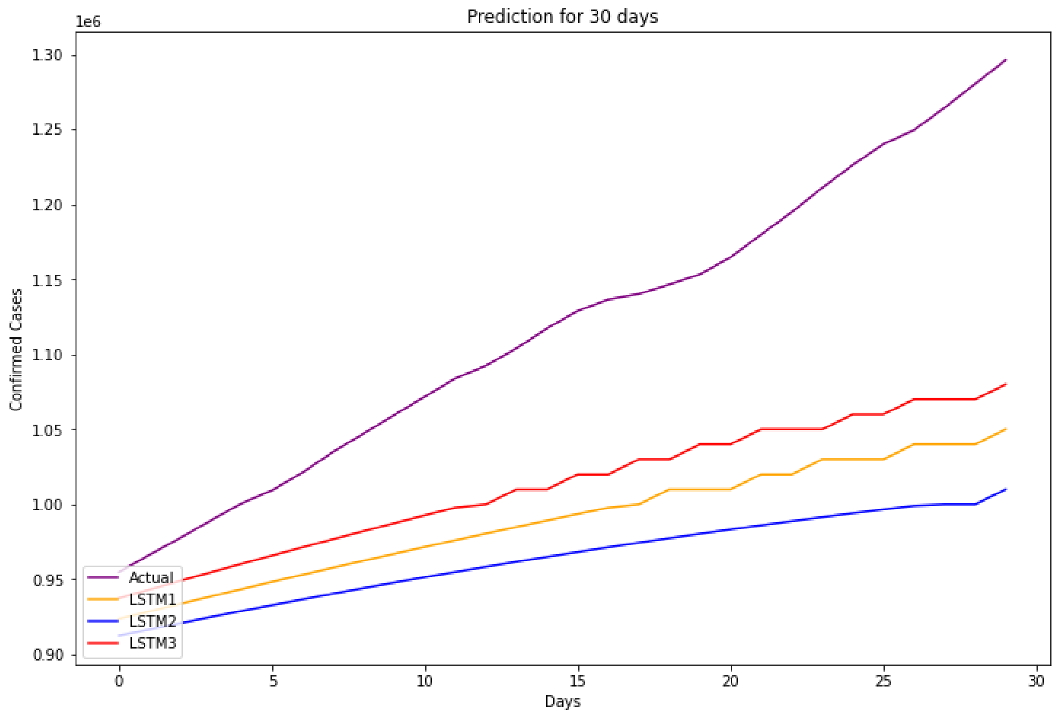

Figure 7.1: Comparison of predicted confirmed cases using LSTM1, LSTM2 and LSTM3

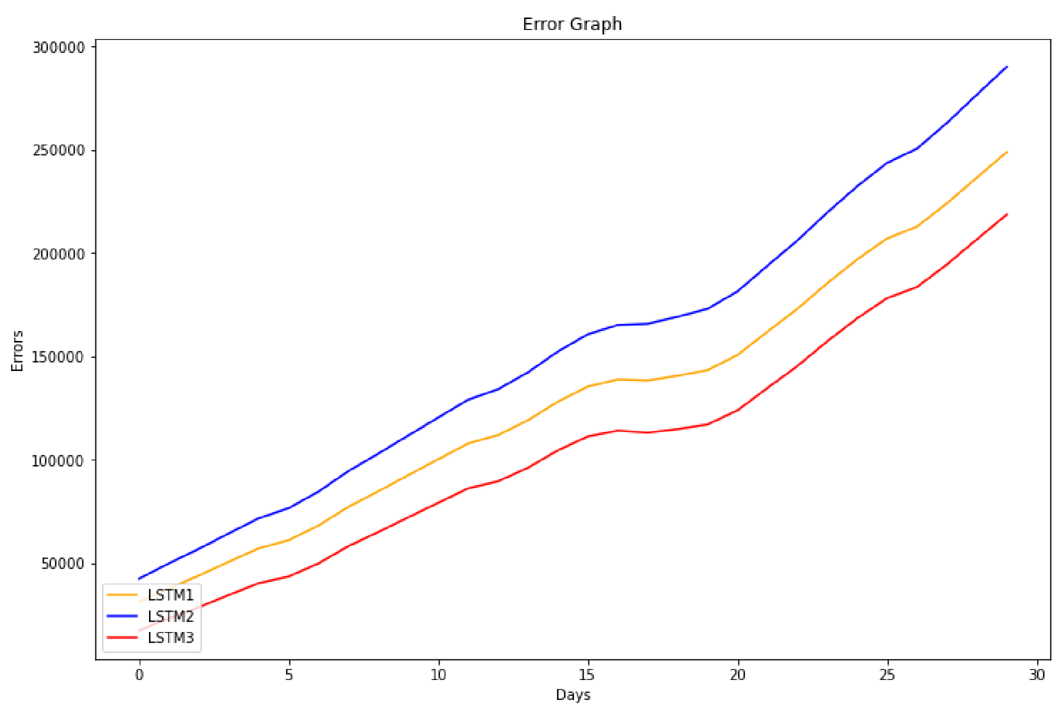

Figure 7.2: Comparison of errors of LSTM1, LSTM2 and LSTM3

Model	RMSE
LSTM1	142574.8027
LSTM2	169319.1603
LSTM3	**119683.1874**

Table 7.2: RMSE Comparison of LSTM1, LSTM2, LSTM3

7.2.2 Comparison analysis of the LSTM Models:

We see that out of three models we selected the LSTM3 shows the lowest error per day and also has the lowest RMSE of the three models selected so we will use the LSTM3 model as the only LSTM model to compare with other non-LSTM models.

7.3 Time-Dependent SIR Model Comparison:

Here, we just have to vary alpha of the ridge regression. Alpha is the regularization strength that helps to improve the conditioning of the problem and reduces the variance of the estimate[50].We are predicting the beta and gamma first and then calculate the predicted confirmed cases with the predicted beta and gamma using the equations of time-dependent SIR [6]. Below we have given the details of the models of SIR :

SIR Model	alpha for predicting beta	alpha for predicting gamma
SIR1	0.003765	0.001675
SIR2	0.87354	0.76457
SIR3	0.07854	0.06478

Table 7.3: SIR1, SIR2 and SIR3 details

7.3.1 Prediction and error comparison of the SIR Models:

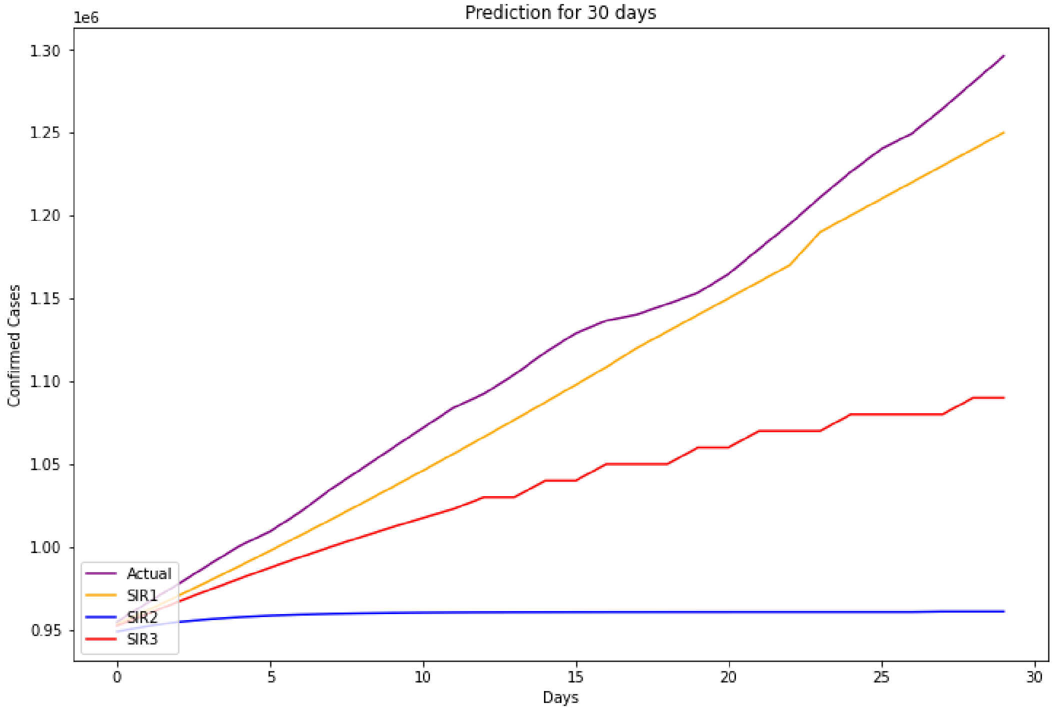

Figure 7.3: Comparison of predicted confirmed cases using SIR1, SIR2 and SIR3

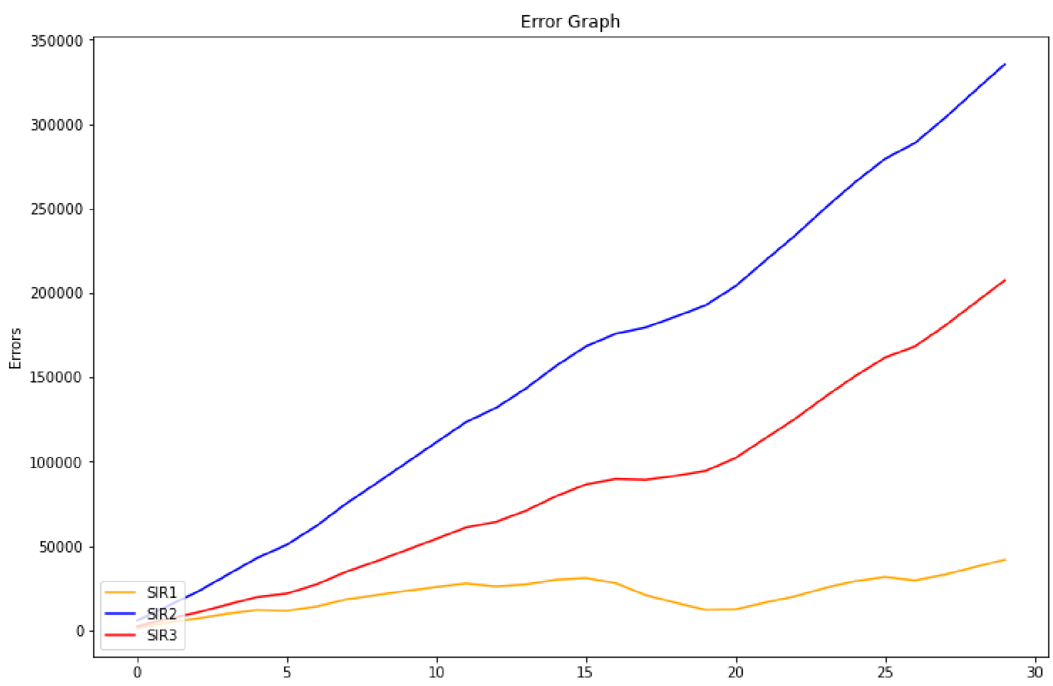

Figure 7.4: Comparison of errors of SIR1, SIR2 and SIR3

Model	RMSE
SIR1	**23685.96475**
SIR2	185502.4321
SIR3	103094.0045

Table 7.4: RMSE Comparison of SIR1, SIR2, SIR3

7.3.2 Comparison analysis of the SIR Models:

We see that out of three models we selected the SIR1 shows the lowest error per day and also has the lowest RMSE among the three models selected so we will use the SIR1 model as the only SIR model to compare with other non-SIR models.

7.4 Seasonal Arima Comparison:

For seasonal arima model we have to select a proper order for trend elements (p,d,q) and seasonal elements (P,D,Q,m) where

- p = auto regression order for the trend elements
- d = difference order for the trend elements
- q = moving average order for the trend elements
- P = auto regression order for the seasonal elements
- D = difference order for the seasonal elements
- Q = moving average order for the seasonal elements
- m = number of steps for a single seasonal period

Here, the prediction depends on the (p,d,q) and (P,D,Q,m) orders. So we will vary the orders and see what performs better. Below we have given the details of the models of SARIMA:

SARIMA Model	(p,d,q)	(P,D,Q,m)
SARIMA1	(5,1,5)	(4,1,1,12)
SARIMA2	(3,1,0)	(1,1,0,12)
SARIMA3	(0,1,0)	(0,1,0,12)

Table 7.5: SARIMA1, SARIMA2 and SARIMA3 details

7.4.1 Prediction and error comparison of the SARIMA Models:

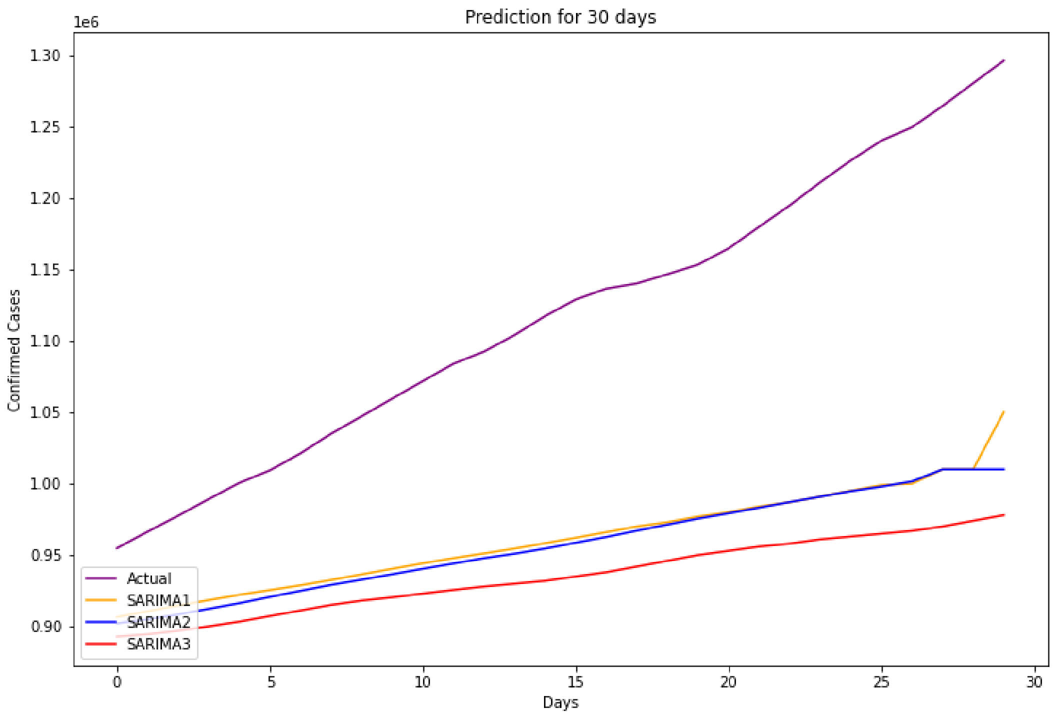

Figure 7.5: Comparison of predicted confirmed cases using SARIMA1, SARIMA2 and SARIMA3

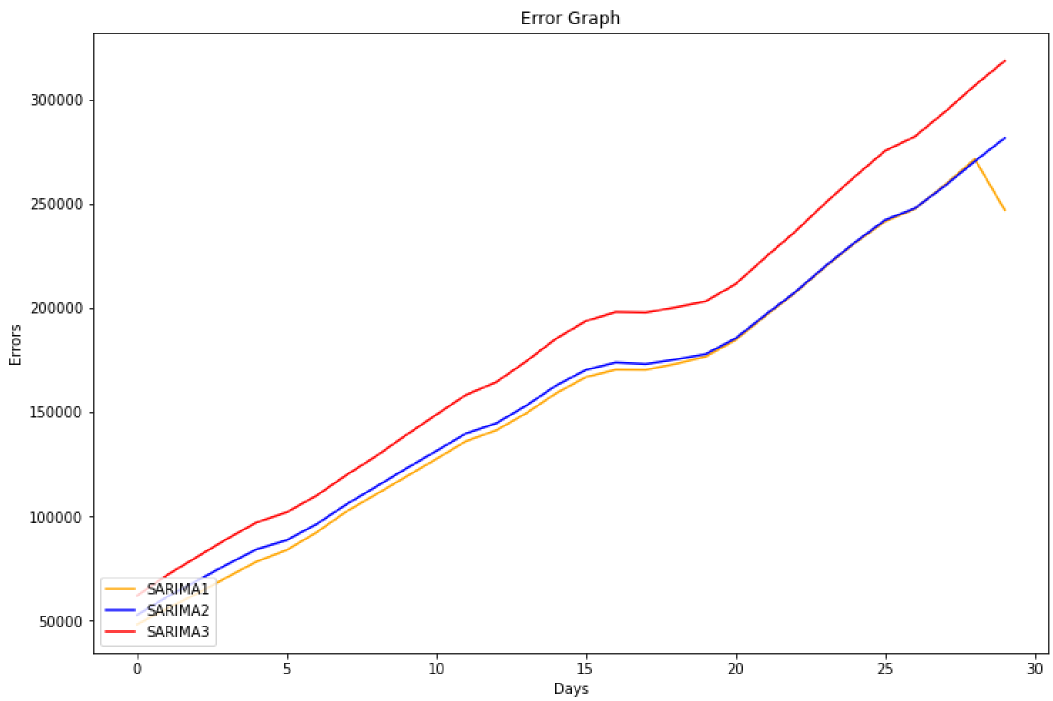

Figure 7.6: Comparison of errors of SARIMA1, SARIMA2 and SARIMA3

Model	RMSE
SARIMA1	**169407.7557**
SARIMA2	172956.2956
SARIMA3	196831.212

Table 7.6: RMSE Comparison of SARIMA1, SARIMA2, SARIMA3

7.4.2 Comparison analysis of the SARIMA Models:

We see that its pretty difficult to determine which is better because all of them gives Error graph and prediction graph similar to each other. But SARIMA1 has comparatively lower RMSE than other models even if its not by lot. So we will use the SARIMA1 model as the only SARIMA model to compare with other non-SARIMA models.

Chapter 8

Experimental Results and Analysis

8.1 Result:

For this project, we have 530 total data in dataset of confirmed cases in Bangladesh. The dataset starts from January 22, 2020 and ends at July 4, 2021. Our prediction will start from July 5, 2021. and will end at August 3,2021. So we will predict for 30 days. Table below shows the actual number of confirmed cases, deaths and recovered from July 5,2021:

Date	Confirmed	Recovered	Deaths
7/5/2021	954881	839082	15229
7/6/2021	966406	844515	15392
7/7/2021	977568	850502	15593
7/8/2021	989219	856346	15792
7/9/2021	1000543	862384	16004
7/10/2021	1009315	868139	16189
7/11/2021	1021189	874501	16419

Table 8.1: Actual number of Confirmed cases, recovered and deaths from 7/5/2021 to 7/11/2021 in Bangladesh

8.1.1 Result Analysis Method:

After we have processed our dataset, we trained those datasets using our 3 models such as LSTM, SIR and SARIMA. For training we split the data into two parts. One for training and one for testing. We tested the data for last 15,30,90 data and we used the rest of the data for training. Then we plotted all the prediction for each model along with the actual confirmed cases in one plot for comparison and also plotted error graphs to observe how many errors we are getting per day for each model. We also calculated RMSE for all models and compared them. We observed that SIR works best among these 3 models and overall had lower RMSE.

8.1.2 Predicting for confirmed cases using SARIMA, LSTM and SIR:

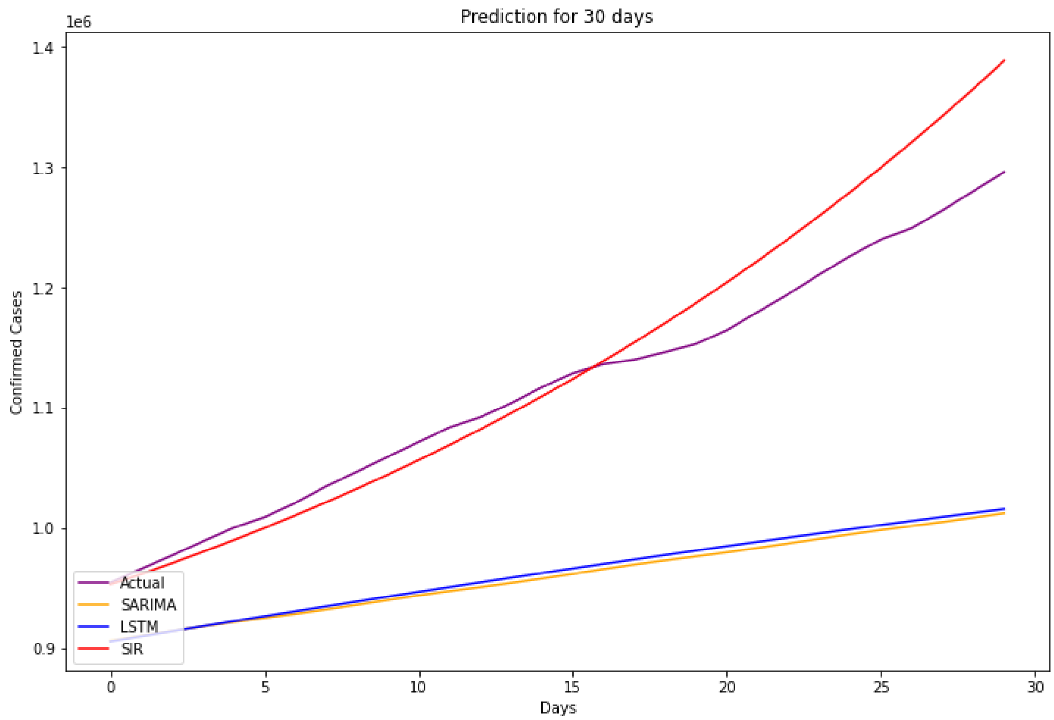

Figure 8.1: Comparison of predicted confirmed cases using SARIMA, LSTM and SIR with last 15 data

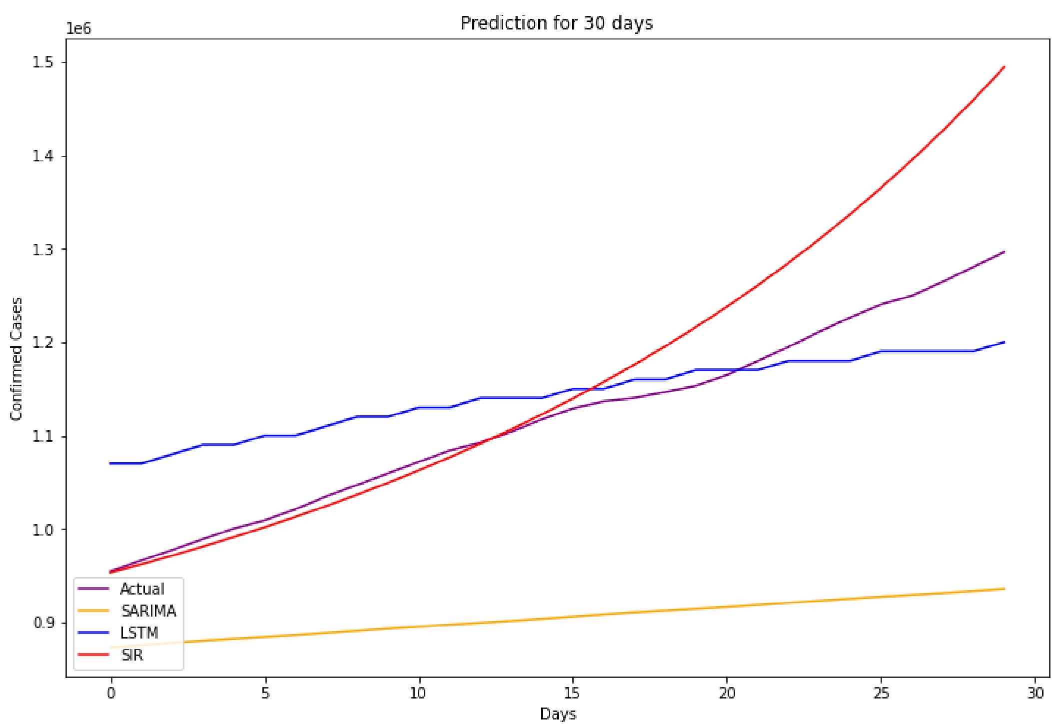

Figure 8.2: Comparison of predicted confirmed cases using SARIMA, LSTM and SIR with last 30 data

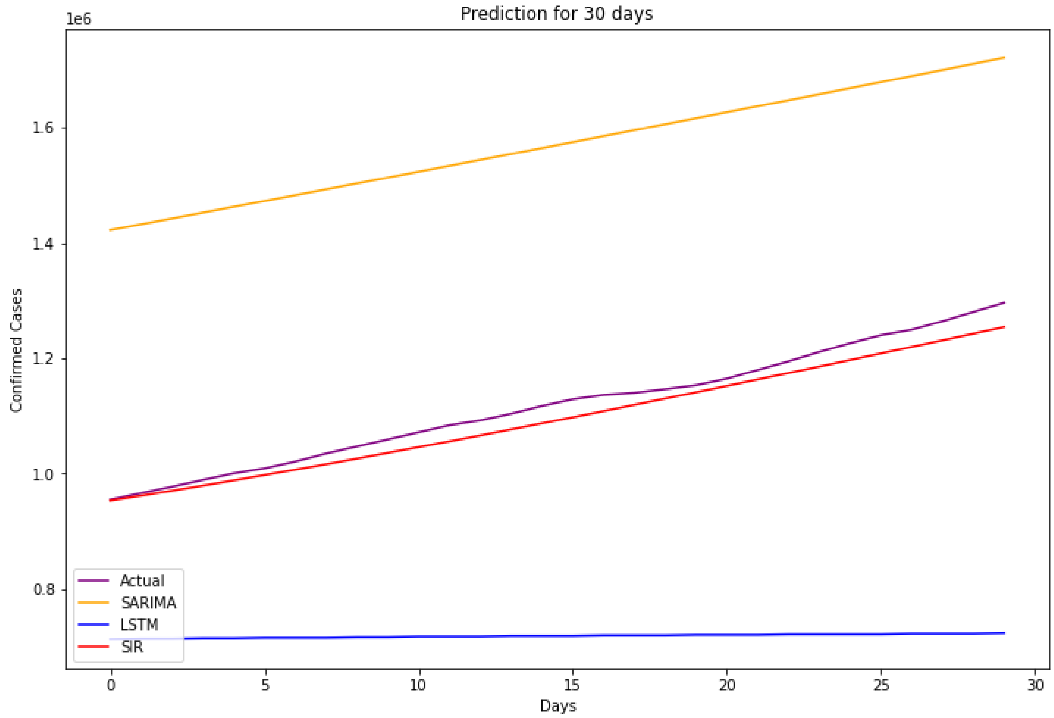

Figure 8.3: Comparison of predicted confirmed cases using SARIMA, LSTM and SIR with last 90 data

8.1.3 Comparison Analysis:

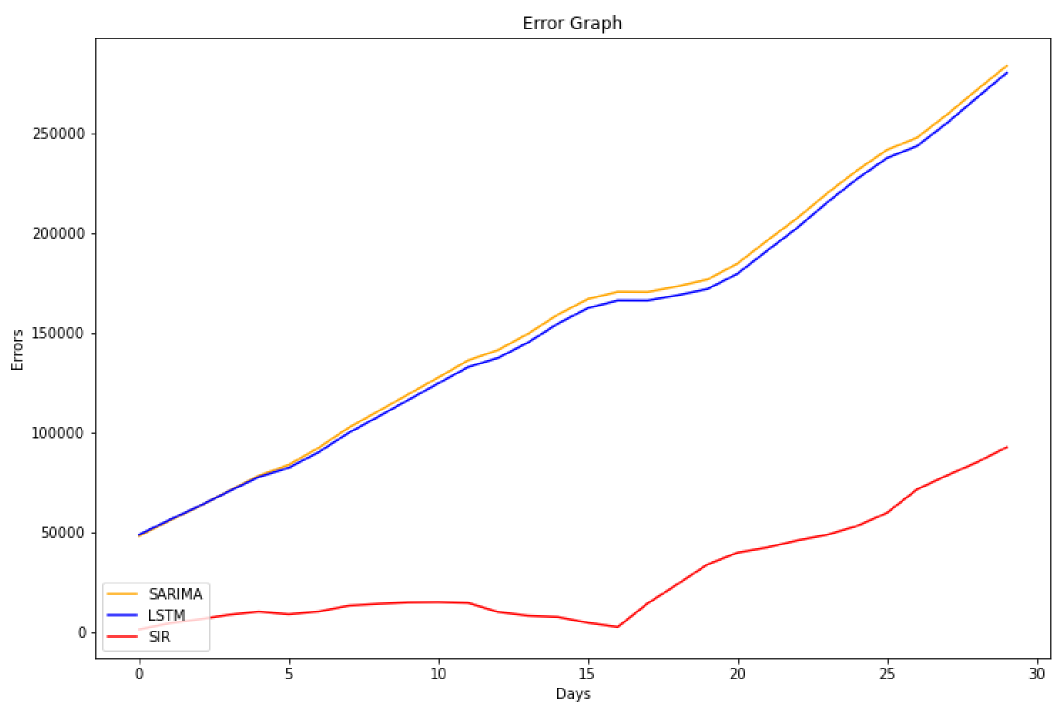

Figure 8.4: Comparison of errors of SARIMA, LSTM AND SIR for last 15 data as test dataset

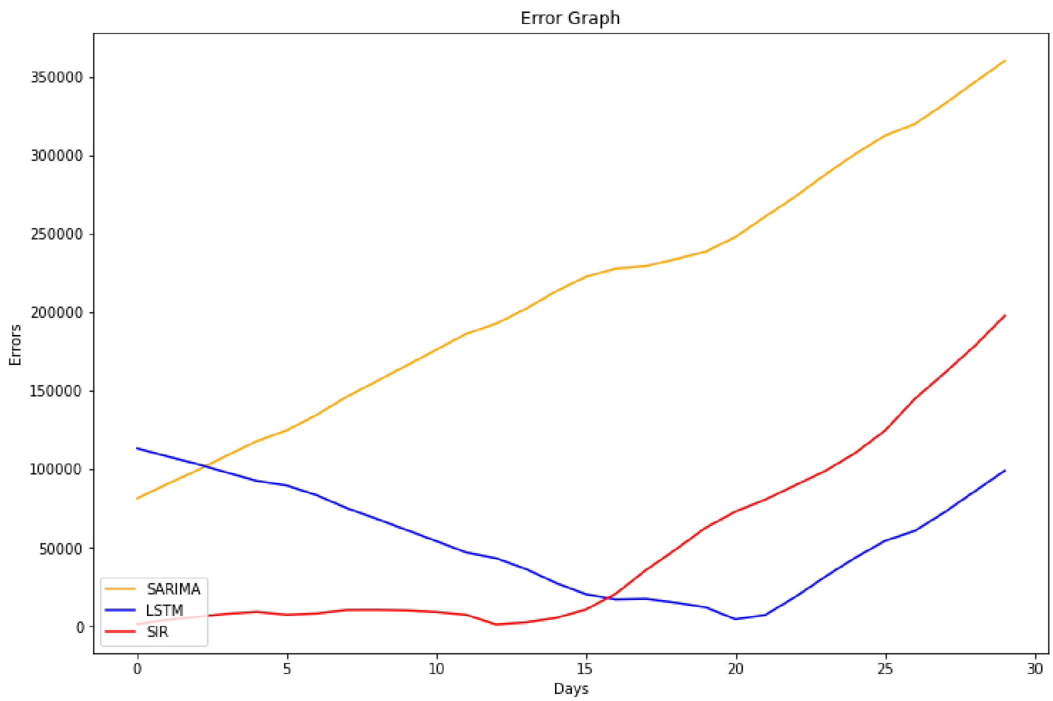

Figure 8.5: Comparison of errors of SARIMA, LSTM AND SIR for last 30 data as test dataset

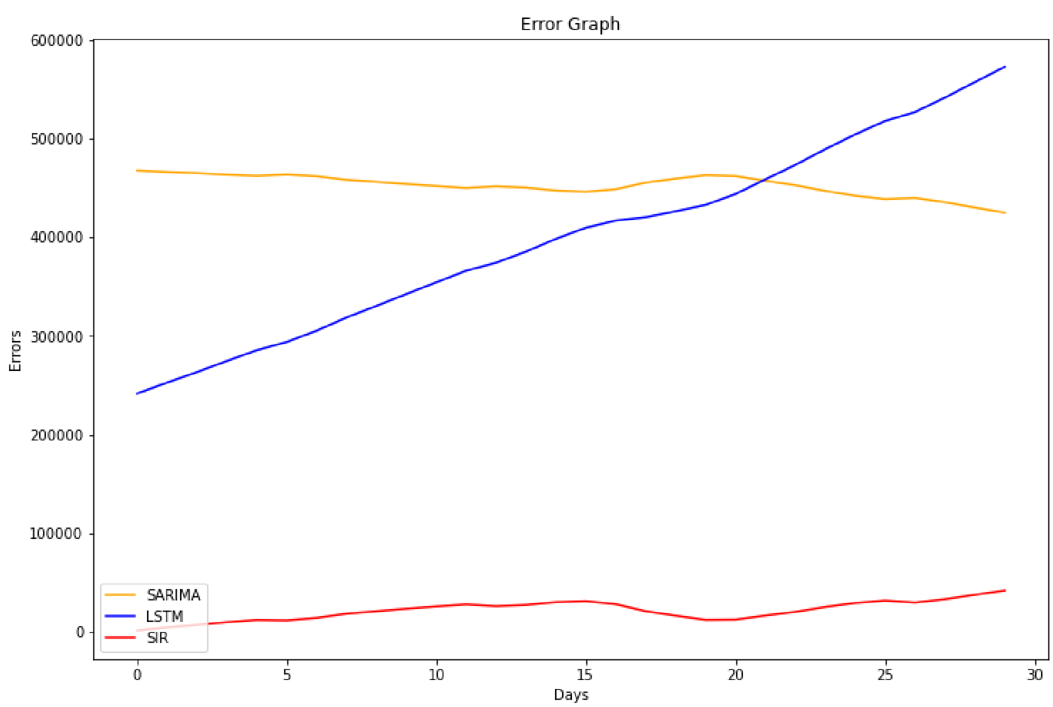

Figure 8.6: Comparison of errors of SARIMA, LSTM AND SIR for last 90 data as test dataset

From the above figures, we see SIR is the most stable model out of the three, LSTM performs slightly better when it is predicting with 30 tested data, But Seasonal Arima doesn't perform good at all. From the above figures we can confirm that SARIMA and LSTM are not stable because recovered cases and death cases aren't factors here. Due to that, These predictions can quickly stray away from predicting close to the accurate result. This is why vanilla LSTM or simple Seasonal ARIMA isn't enough for these predictions. Thus there needs to be a modified version of these models that takes recovered cases and death cases into consideration and makes prediction based on those factors. table below shows RMSE for each models with different train and test data:

Predicted with	SARIMA	LSTM	SIR
last 15 data	**171303.8644**	167813.8499	38730.29948
last 30 data	227376.0861	**64676.81792**	78416.16899
last 90 data	452367.0258	410239.32	**23685.96475**

Table 8.2: RMSE Comparison of LSTM, SARIMA, SIR

As we see from the table above, RMSE for Seasonal ARIMA is the highest among all other models and it is very unreliable. The lowest RMSE is obtained by using SIR model with the tested data length of 90. LSTM did perform better when it used tested data length of 30 and it is third lowest RMSE among the models. Here RMSE of SIR ranges from aprroximately 24000 to 79000, RMSE of LSTM model ranges from approximately 65000 to 411000 and RMSE of SARIMA Model ranges from approximately 172000 to 453000. So we can say SIR model works better for prediction of confirmed cases of covid overall. Since SARIMA's lowest RMSE is 171303.8644, it can be considered the worst model for predicting confirmed cases.

Chapter 9

Conclusion

9.1 Research Challenges

We faced many challenges due to models with different theoretical basis. Since the core concept of each model is different, it was difficult to understand the models properly to predict confirmed cases. Another challenge was adjusting dataset for each model. Even though dataset organizing isn't difficult at all, it is time consuming. Since the dataset wasn't that large, with the increase of tested data, length of trained data shrank, which didn't help some models in providing accurate output. For example, LSTM had better result after we increased tested data length to 30 but when we increased the tested data length to 90, the output became even less accurate due to the decreased length of the training dataset.

9.2 Limitations

We used three different models to predict the confirmed cases of Bangladesh and compared them. But these models are the basic model from the same branch. For example, LSTM model we used was the univariate LSTM model. There are other variations of Deep Learning Based models which can be used. Such as Bidirectional LSTM,GRU etc. Similarly we used SIR model which is the simplest compartmental model. There are many variations of epidemic models such as SEIR, SIRVD, MSIR, MSEIR, MSEIRS [32]. Due to the time limitation, we weren't able to study more about the variations of these models to implement them in our thesis.

And because we used simple models to predict confirmed cases, The RMSE was much larger than our expected RMSE. Only Model that had the lowest RMSE was SIR model which was around 24000. This is still a big error and if we want to prepare for future waves of covid(such as the third wave in 2022 caused by new variant omicron), then we have to be more accurate while predicting. And most important limitation is that due to small dataset, we can't make predictions for a long time gap.

9.3 Future Scope

We are proposing a hybrid model by combining LSTM and SIR model. In SIR model, In SIR model, we used ridge regression to predict beta and gamma for future days but what if we used LSTM model to predict beta and gamma and then predict the confirmed cases using the predicted beta and gamma. This hybrid model can be used for more precise prediction.

In this thesis, we used the dataset of Bangladesh but the results may differ for different countries. So to make more accurate prediction as possible the model shown in the flowchart can be used. Because in LSTM we only used confirmed cases to predict future cases and since deaths and recovered cases are unknown to LSTM model, it can't produce predictions as accurate as we expected. If we make a hybrid model of LSTM and SIR where we combine the forecasting accuracy of LSTM with the epidemiological model dynamics of SIR, we can predict more accurately.

While trying to implement this hybrid model, there was a problem we faced. We have to consider that beta and gamma suffers from multicollinearity so while predicting the beta and gamma, we have to implement a method that considers their high correlation.

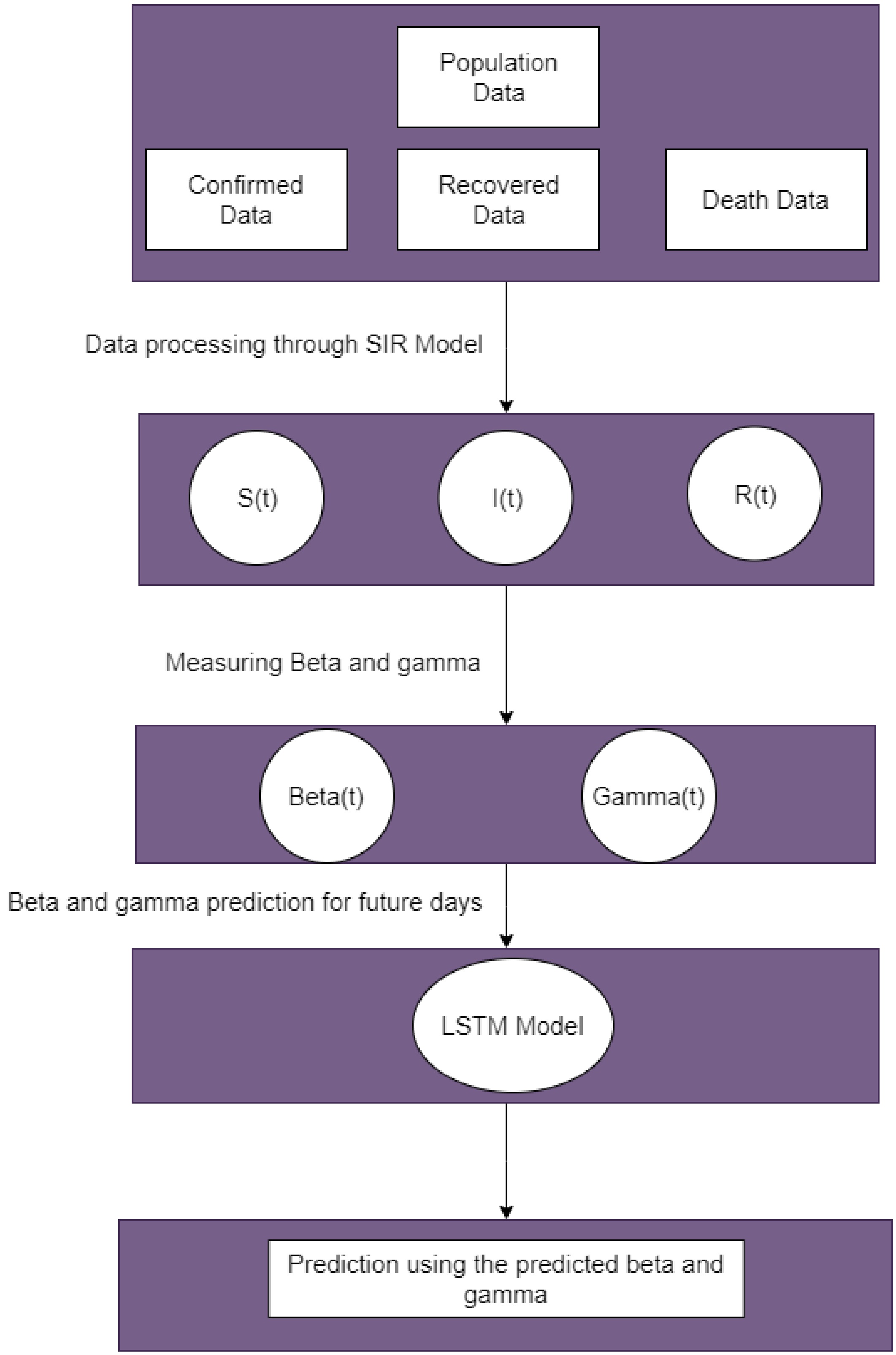

Figure 9.1: Hybrid Model based on SIR and LSTM

9.4 Conclusion

The aim of this thesis was to compare the models for forecasting confirmed cases in Bangladesh. With more accurate prediction, country can prepare better against the upcoming third wave caused by new variant of covid, also known as omricorn. Extra medical camps for covid patients can be built just in case the rate of transmission exceeds the amount of patients hospitals in Bangladesh can hold. This is the first time pandemic of this scale has happened during the era of IOT. Every data gathered from this pandemic is invaluable and can be used to control the outcome of future pandemics. To achieve such goal, we decided to research on the dataset related to Covid and compare the models to determine which can predict confirmed cases in Bangladesh accurately and we determined SIR model is the best for forecasting confirmed cases. Finally we developed a plan to build a model that can predict the cases with more precision. Our future research will expand upon what we learnt in this thesis and ensure that we can achieve the accuracy we expected.

Bibliography

[1] P. Wadhwa, Aishwarya, A. Tripathi, P. Singh, M. Diwakar, and N. Kumar, "Predicting the time period of extension of lockdown due to increase in rate of covid-19 cases in india using machine learning," *Materials Today: Proceedings*, vol. 37, 08 2020.

[2] "Coronavirus," https://en.wikipedia.org/wiki/Coronavirus.

[3] B. News, "Coronavirus disease named covid-19," https://www.bbc.com/news/world-asia-china-51466362, 2020.

[4] S. Roychoudhury, A. Das, P. Sengupta, S. Dutta, S. Roychoudhury, A. P. Choudhury, A. B. F. Ahmed, S. Bhattacharjee, and P. Slama, "Viral pandemics of the last four decades: Pathophysiology, health impacts and perspectives," *International Journal of Environmental Research and Public Health*, vol. 17, no. 24, 2020. [Online]. Available: https://www.mdpi.com/1660-4601/17/24/9411

[5] Wikipedia, "List of epidemics," https://en.wikipedia.org/wiki/List_of_epidemics.

[6] Y.-C. Chen, P.-E. Lu, C.-S. Chang, and T.-H. Liu, "A time-dependent sir model for covid-19 with undetectable infected persons," *IEEE Transactions on Network Science and Engineering*, vol. 7, no. 4, pp. 3279–3294, 2020.

[7] L. Feng, "Seir model combined with lstm and gru for the trend analysis of covid-19," *Middle Tennessee State University*, 2021.

[8] Z. Liao, P. Lan, X. Fan, B. Kelly, A. Innes, and Z. Liao, "Sirvd-dl: A covid-19 deep learning prediction model based on time-dependent sirvd," *Computers in Biology and Medicine*, vol. 138, p. 104868, 09 2021.

[9] Y. Kang, "Prediction models of covid-19 based on seir and lstm rnn," *AIP Conference Proceedings*, vol. 2350, no. 1, p. 020037, 2021. [Online]. Available: https://aip.scitation.org/doi/abs/10.1063/5.0048409

[10] P. Nadler, R. Arcucci, and Y. Guo, "A neural sir model for global forecasting," vol. 136, pp. 254–266, 11 Dec 2020. [Online]. Available: https://proceedings.mlr.press/v136/nadler20a.html

[11] S. Feng, Z. Feng, C. Ling, C. Chang, and Z. Feng, "Prediction of the covid-19 epidemic trends based on seir and ai models," *PLOS ONE*, vol. 16, no. 1, pp. 1–15, 01 2021. [Online]. Available: https://doi.org/10.1371/journal.pone.0245101

[12] S. Singh, B. Sundram, K. Rajendran, K. Law, T. Aris, H. Ibrahim, C. Dass, and B. Gill, "Coronavirus pandemic forecasting daily confirmed covid-19 cases in malaysia using arima models," *The Journal of Infection in Developing Countries*, vol. 14, pp. 971–976, 10 2020.

[13] V. Papastefanopoulos, P. Linardatos, and S. Kotsiantis, "Covid-19: A comparison of time series methods to forecast percentage of active cases per population," *Applied Sciences*, vol. 10, p. 3880, 06 2020.

[14] A. Chowdhury, T. Hasan, and K. Hoque, "Analysis and prediction of covid-19 pandemic in bangladesh by using anfis and lstm network," *Cognitive Computation*, vol. 13, 05 2021.

[15] B. Ndiaye, L. Tendeng, and D. Seck, "Analysis of the covid-19 pandemic by sir model and machine learning technics for forecasting," 04 2020.

[16] 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, p. 110212, 08 2020.

[17] S. N. Zisad, M. S. Hossain, M. S. Hossain, and K. Andersson, "An integrated neural network and seir model to predict covid-19," *Algorithms*, vol. 14, no. 3, 2021. [Online]. Available: https://www.mdpi.com/1999-4893/14/3/94

[18] E. Amer, A. Hazem, O. Farouk, A. Louca, Y. Mohamed, and M. Ashraf, "A proposed chatbot framework for covid-19," pp. 263–268, 2021.

[19] Z. KARHAN and F. AKAL, "Covid-19 classification using deep learning in chest x-ray images," pp. 1–4, 2020.

[20] P. Sv, D. R. Ittamalla, and G. Deepak, "Analyzing the attitude of indian citizens towards covid-19 vaccine – a text analytics study," *Diabetes Metabolic Syndrome: Clinical Research Reviews*, vol. 15, 02 2021.

[21] J. Feng, X. He, Q. Teng, C. Ren, H. Chen, and Y. Li, "Reconstruction of porous media from extremely limited information using conditional generative adversarial networks," *Physical Review E*, vol. 100, 09 2019.

[22] T. W. S. Chow and Y. Fang, "A recurrent neural-network-based real-time learning control strategy applying to nonlinear systems with unknown dynamics," *IEEE Transactions on Industrial Electronics*, vol. 45, no. 1, pp. 151–161, Feb 1998.

[23] A. Mellit and Alessandro Massi Pavan, "A 24-h forecast of solar irradiance using artificial neural network: Application for performance prediction of a grid-connected PV plant at Trieste, Italy," *Solar Energy*, vol. 84, pp. 807–821, December 2010.

[24] A. Hassan, I. Shahin, and M. B. Alsabek, "Covid-19 detection system using recurrent neural networks," pp. 1–5, 2020.

[25] G. Petneházi, "Recurrent neural networks for time series forecasting," 12 2018.

[26] H. Hewamalage, C. Bergmeir, and K. Bandara, "Recurrent neural networks for time series forecasting: Current status and future directions," 09 2019.

[27] S. Hochreiter, "The vanishing gradient problem during learning Recurrent neural nets and problem solutions," *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems*, vol. 06, pp. 107–116, 1998.

[28] Y. Bengio, P. Simard, and P. Frasconi, "Learning long-term dependencies with gradient descent is difficult," *IEEE Transactions on Neural Networks*, vol. 5, no. 2, pp. 157–166, March 1994.

[29] S. Hochreiter and J. Schmidhuber, "Long short-term memory," *Neural computation*, vol. 9, pp. 1735–80, 12 1997.

[30] K. W. Ogilvy and M. A. G., "A contribution to the mathematical theory of epidemics," pp. 700–721, 1927.

[31] C. Cano, "The sir models, their applications, and approximations of their rates," *Mathematics Senior Capstone Papers*, 2020. [Online]. Available: https://digitalcommons.latech.edu/mathematics-senior-capstone-papers/18

[32] "Compartmental models in epidemiology," https://en.wikipedia.org/wiki/Compartmental_models_in_epidemiology.

[33] "Basic reproduction number," https://en.wikipedia.org/wiki/Basic_reproduction_number.

[34] "Ridge regression," https://en.wikipedia.org/wiki/Ridge_regression.

[35] "Ridge regression: Biased estimation for nonorthogonal problems," *Technometrics*, vol. 42, no. 1, pp. 80–86, 2000. [Online]. Available: http://www.jstor.org/stable/1271436

[36] "Ridge regression: Applications to nonorthogonal problems," *Technometrics*, vol. 12, no. 1, pp. 69–82, 1970. [Online]. Available: http://www.jstor.org/stable/1267352

[37] S. Mardikyan, E. Çetin, S. Mardikyan, and E. Çetin, "Efficient choice of biasing constant for ridge regression," *International Journal of Contemporary Mathematical Sciences*, vol. 3, 01 2008.

[38] O'Reilly, "Chapter 1. time series: An overview and a quick history," https://www.oreilly.com/library/view/practical-time-series/9781492041641/ch01.html.

[39] M. D. Moloy, M. Chowdhury, M. Binyamin, and S. Mondal, "Using sarima approach to modeling and forecasting monthly rainfall in bangladesh," 01 2018.

[40] Y. Yang, H. Zheng, and R. Zhang, "Prediction and analysis of aircraft failure rate based on sarima model," pp. 567–571, 2017.

[41] E. Stellwagen and L. Tashman, "Arima: The models of box and jenkins," *Foresight: Int. J. Appl. Forecast.*, pp. 28–33, 01 2013.

[42] M. Mahan, C. Chorn, and A. Georgopoulos, "White noise test: detecting autocorrelation and nonstationarities in long time series after arima modeling," 06 2015.

[43] "Worldometer. bangladesh coronavirus: 1,581,083 cases and 28,048 deaths," https://www.worldometers.info/coronavirus/country/bangladesh/, 2021.

[44] "Covid-19 data repository by the center for systems science and engineering (csse) at johns hopkins university," -https://github.com/CSSEGISandData/COVID-19, 2021.

[45] "A time-dependent sir model for covid-19 with undetectable infected persons," https://github.com/PingEnLu/Time-dependent_SIR_COVID-19/.

[46] "Covid forecasting using lstm," https://github.com/bnsreenu/python_for_microscopists/blob/master/166b_COVID_forecasting_using_LSTM.py.

[47] "Intro to time series forecasting using arima," https://github.com/bnsreenu/python_for_microscopists/blob/master/163-Intro_to_time_series_Forecasting_using_ARIMA.py.

[48] "Gentle introduction to the adam optimization algorithm for deep learning," https://machinelearningmastery.com/adam-optimization-algorithm-for-deep-learning/.

[49] A. Akandeh and F. Salem, "Simplified long short-term memory recurrent neural networks: part i," 07 2017.

[50] "sklearn-1.0.2 documentation," https://scikit-learn.org/stable/modules/generated/sklearn.linear_model.Ridge.html.