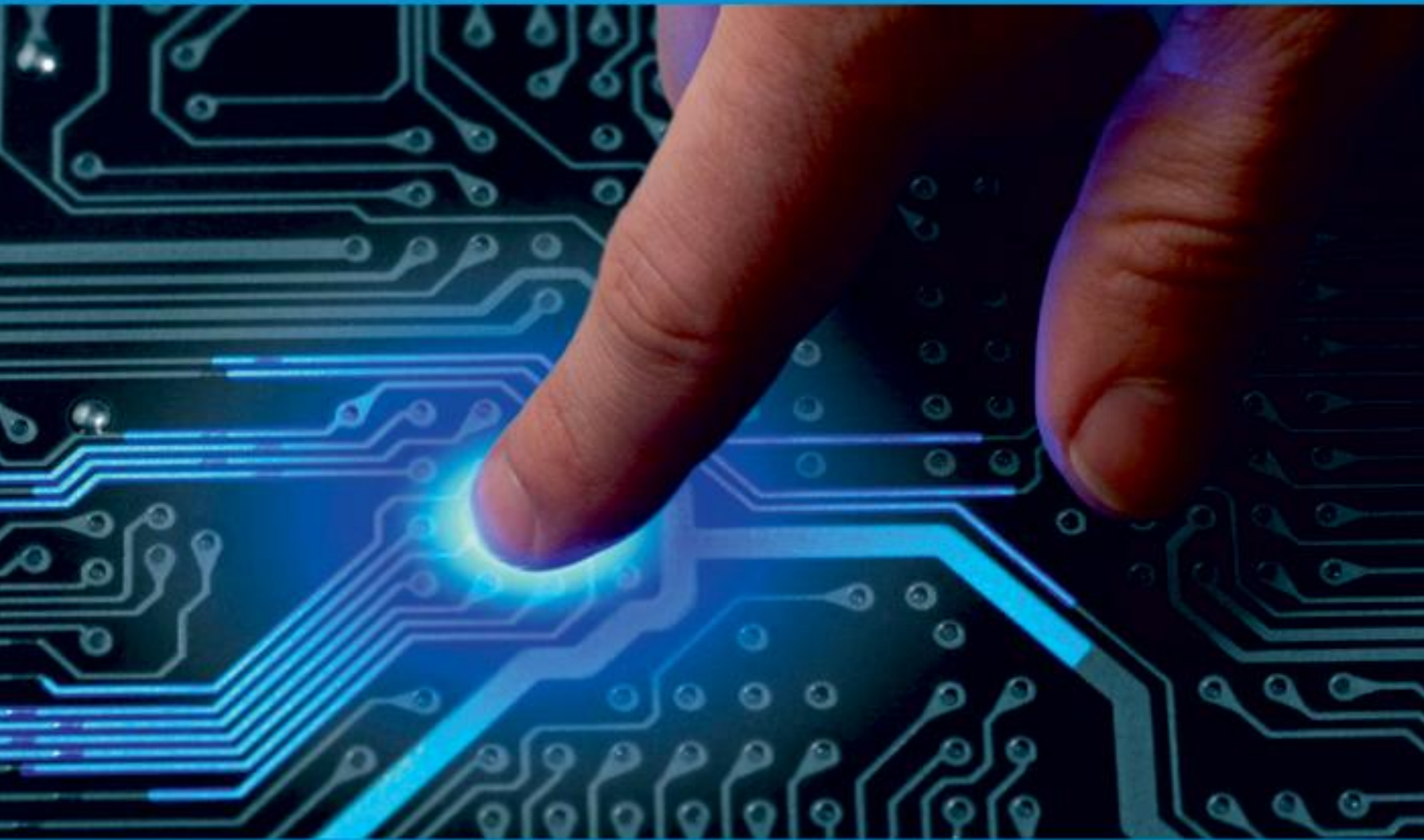
IJIRCCE

e-ISSN: 2320-9801 | p-ISSN: 2320-9798



INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN COMPUTER & COMMUNICATION ENGINEERING

Volume 11, Issue 5, May 2023

INTERNATIONAL STANDARD SERIAL NUMBER INDIA

Impact Factor: 8.379

International Journal of Innovative Research in Computer and Communication Engineering

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.379 |

|| Volume 11, Issue 5, May 2023 ||

| DOI: 10.15680/IJIRCCE.2023.1105173|

Covid-19 Patient Count Prediction using deep Learning models

Maithali Sambhaji Patil, Prof. Bhagyashree Dhakulkar

Department of Computer Engineering, Dr. D Y Patil School of Engineering and Technology, Lohegaon, Pune, India

ABSTRACT–Among the world's most serious problems was the COVID-19 pandemic, known colloquially as the coronavirus epidemic. The covid sickness is a novel, highly contagious lung disease. Both global health activities and data integration are essential for determining and mitigating potential health risks due to this virus. These efforts prioritize preventative measures over reactive ones, especially in the face of outbreaks of communicable diseases. For prevention measures to be successful, it is necessary to have an accurate assessment of the disease's present position, its rate of spread, its likely progression, and its estimated prevalence. Thus, especially for communicable diseases, precise quantitative advanced data analysis is an essential part of governmental healthcare efforts. Consequently, this study paper presents a beneficial strategy for attaining covid rate of survival forecasting using deep learning approaches. For this purpose, this paper describes a precise and effective technique for predicting the prevalence of infection in covid. To forecast the survival rate of covid, we employ a method based on a combination of the Convolutional Neural Network-Long Short-Term Memory and the Decision Tree. Numerous experiments have confirmed the method's efficacy and shown the superiority of its application.

KEYWORDS—Pearson correlation, K Nearest Neighbor, Convolutional Neural Networks-Long Short Term Memory, Decision Tree.

I. INTRODUCTION

It has been several years since the reemergence of an old pandemic was discovered. Initially, this was spotted in China. The World Health Organization dubbed the illness induced by the epidemic viral infection, Severe Acute Respiratory Syndrome Coronavirus 2, covid. There has been a steady progression of this illness into a worldwide catastrophe. With a surge of covid sufferers, several nations' healthcare infrastructures buckled under the strain of an unexpectedly large number of patients. Due to a lack of resources, a sickness of this magnitude poses a serious danger to the lives of those affected.

Furthermore, if governments can use connection monitoring to identify possible covid patients in advance, they may take precautions to separate these high-risk individuals. In the long run, this minimizes the spread of the illness and the quantity of individuals who need treatment. Manually interaction monitoring of positive covid instances is the standard practice for governments. The time and resources required by several people, particularly medical professionals, make this strategy impractical and ineffective. It is a lengthy process that takes approximately 3 days for every new observation. Whenever the illness is propagating fast, it is pointless to invest 3 days identifying a new patient. Numerous methods already present to simplify the process of tracking for lost contacts. Unfortunately, we do find a few problems with the current methods.

Forecasting the acute risk of just about any illness at a preliminary phase is an essential undertaking with numerous impacts, including lowering the risk of dying, minimizing the use of healthcare procedures, and aiding in the decision-making of clinicians. This increase may be attributed to a few factors: the delayed detection of cases, the high incidence of covid, and the inability of several institutions to deal with the epidemic. Thus, anticipating the number of covid patients is an important task with several favorable aspects, including such ensuring that every patient receives the appropriate level of service for his magnitude, attempting to make efficient use of medical resources by giving prime importance to high-risk patient populations, and helping doctors make choices that would enhance the quality of medical care every patient receives.

In specific, an accurate risk forecasting model could aid personalized medicine initiatives by allowing for patient-specific therapeutic care, hence boosting the likelihood of a full recovery. Additionally it would help emergency rooms better manage the flow of patients, minimizing long waits. Thus, a great deal of study has been devoted to determining whether or not medical, analytical, and radiographic variables may be used to forecast patient prognosis. A recent assessment of covid patient count prediction systems found that the majority of these are skewed against one element or

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.379 |

|| Volume 11, Issue 5, May 2023 ||

| DOI: 10.15680/IJIRCCE.2023.1105173|

another, as will be detailed below, despite the fact that encouraging findings have indeed been published by multiple authors.

The society has seen the consequences of novel covid variations, according to research by Vivek Kumar Prasad et al [1]. As nothing more than a result of the devastation wreaked upon community hospital infrastructure by the Covid latest wave, it is imperative that all nations establish and implement appropriate care regulations, as well as ensure the appropriate transmission of vital strategic planning. The authors offer a method, titled "Availability of Beds for covid," that uses ensembles prediction to estimate the prospective need for resources like beds & ventilation systems. Methodologies of prediction were examined by the authors. The hospital bed as well as ventilator needs throughout every area are accessed using an artificial neural network algorithm.

In light of the present epidemic, Ertugrul Karaçuha et al. [2] offer concrete recommendations for predicting incidence counts in order to effectively prepare medical resources for victims. This simulation could be applied to the current covid situation as well as to future possible regional or global epidemics. This research uses a mathematical model to forecast the total estimated incidence of documented cases, fatalities, and recoveries from the covid epidemic. Gaussian analysis was used to predict the maximum number of patients and forecast the course of the epidemic based on daily fluctuations in epidemic statistics. Unfortunately, although the suggested Gaussian forecasting approach does well for maximum forecasting, it tends to under predict the long-term outcomes. Using the various probabilities may help to make more accurate predictions.

A powerful model using improved Feature extraction and Machine Learning techniques was created by Jayashree Piri et al [3]. Using our FS method, even very simple classifiers like random forest and K closest neighbors proved able to reliably predict the condition of COVID-19 individuals. Throughout this piece, researchers provide a novel and improved composite multi-objective optimization for solving Features extraction issues. The proposed method incorporates a wrapper and filters modeling into a centralized platform in the hopes of optimizing the benefits of each kind, building on the most recent chimp optimization algorithm strategy merged with binary harris hawk optimization. On top of it as well, researchers have shown that the existence of two repositories and the management of a parameter may make improved binary multi-objective hybrid filter wrapper chimp optimization run more slowly than expected. Therefore, researchers want to investigate other fitness responsibilities in the near future to keep overall performance maintained without extending the total amount of time spent jogging.

This research article's literature review is found in the second part. Section 3 describes the suggested strategy, while Section 4 thoroughly evaluates the findings obtained. This study article is finalized in the section 5 including the extent of the future improvements.

II. Literature Survey

The Susceptible Infected Recovered methodology used for estimating COVID-19 casualties has been examined by Onder Tutsoy et al. [4]. As a result, the innovative, all-encompassing Susceptible Infected Dead paradigm has already been presented, dissected, and defended. Characteristics in the suggested Susceptible Infected Dead model are not stated directly; rather, they are determined intuitively using optimizations. The fatalities in Turkey are used to ensure that the model accurately represents the real-world system. Global death tolls come in a variety of shapes and sizes, but they always have common features including peaks, advances, and declines. As a result, the suggested approach may be readily adapted for use in various jurisdictions around the globe.

An excellent method for differentiating between steady state condition chest, infectious pneumonia, and covid radiography is proposed by Karen Panetta et al. [5], which uses a machine learning-based methodology and a unique feature extraction classifier, Contour Fibonacci-p sequences. The primary benefit of this description is that it eliminates noise in images automatically while simultaneously encrypting textural arrangements with varying interruptions, alignments, and forms. Results from simulated analyses of the whole radiograph dataset demonstrate that the suggested technique outperforms both Deep Learning and the traditional Fibonacci signifier in terms of recall. Improved results were obtained using the suggested feature descriptor contrasted to many of the Deep Learning techniques, with results being on par with those obtained using more conventional techniques. The suggested method is a machine learning model; as such, it does not need any specialist hardware, takes minimal time to train, and yields a stable prototype with respectable sensing techniques from very modest data sources.

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.379 |

|| Volume 11, Issue 5, May 2023 ||

| DOI: 10.15680/IJIRCCE.2023.1105173|

According to Furqan Rustam et al. [6], the COVID-19 contagion is very vulnerable and might spark a major international disaster. Numerous scientists and public sectors all across the globe are worried that a sizable section of the global population might be impacted by the epidemic. In this research, we present an ML-based approach for forecasting the likelihood of a worldwide COVID-19 epidemic. The work investigated a database comprising real historical information on a daily basis and generates forecasts for the next weeks employing machine learning strategies. The report's findings show that, considering the existing state of the prediction sector and the type and amount of the dataset, ES is the most effective. The mortality rate may be predicted and confirmed in certain circumstances using linear regression as well as LASSO's prediction abilities. Those different models agree that mortality ratios will escalate in the days that followed, and that the pace of recovery would stall.

In order to distinguish between deceased and released covid patient populations, Lin Wang et al. [7] extracted potent characteristics from standard statistical physiological parameters and established a flow functionality categorisation Lasso Linear regression framework using the significant medical data sources of covid patient populations provided by the data storage facility. Using a time-lapse characteristic of physiological parameters and the idea of blood oxygen saturation, the authors of the proposed Lasso model of logistic regression retrieved improved characteristics about the patterns alterations. Blood oxygen levels fluctuate throughout the week, which inspired the name "Bad Days." Findings from this study suggest that blood oxygen trends may be used to predict which individuals will need to be admitted to critical care and when particular medicines should be started.

The following are areas where Meditya Wasesa et al. [8] make contributions. To begin with place, this study suggested deep learning-based algorithms to forecast the everyday electricity usage of a small grid-based structure throughout covid. Furthermore, the study showed how digital Google Mobility as well as Google Trends plus covid information may be used to better estimate a building's overall expenditure, in this case a genuine instructive micro grid-based facility. The results confirmed that using information retrieval statistics like Google Trends may enhance the precision of energy consumption projections for both residential and learning institutions. Additionally, this research verified that the incorporation of covid information not only significantly influence the reliability of the power demand forecast at the national scale while also being appropriate for building settings.

According to Lace Padilla et al [9], infographics may either assist individuals to comprehend the situation and make informed decisions throughout outbreaks or sow confusion and skepticism. In this, researchers show how to use several prediction modeling principles to boost confidence without sacrificing effectiveness in movement comprehension. Public-facing infographics that include unpredictability boost comprehension and endurance for unpredictability in future occurrences, in while also providing crucial health - related information. Numerous inconsistent quantitative findings have also been reported by the researchers. In particularly, researchers selected to provide triggers that varied specific predictions, and, at occasionally, projections would deliver outcomes that were neither in line with the pattern. Since the authors did not produce the information or algorithms, there are a lot of possible explanations for the variation. To evaluate the repeatability of these results, additional work is required to systematically manipulate information and modify the features of projections to uncover origins of this inconsistency.

In attempt to anticipate the likelihood of catastrophic outcomes for patients with COVID19, Elena Casiraghi et al [10] set outside to construct a statistical method capable of processing medical, diagnostic, and analytical data associated with these individuals. The laboratory and clinical data were obtained at the moment of every patient's admission to the Emergency, whereas the 4 radiographic readings were subsequently appraised from the individuals CXR, perhaps by aggregating radiological specialists' judgments. The results show that by integrating the Boruta technique with possible combination feature extraction integrated in RFs, reliable functionality may be attained. The impact of misclassification is mitigated and superior outcomes are attained when the chosen extracted features is fed to RFs limited to operate on equitable bootstrapped examples, as compared to Association rule trees or generalized linear algorithms.

An updated prediction paradigm for the sensitivity and increase in risk of covid patient populations is proposed by Safynaz Abdel Fattah Sayed et al [11] to aid clinicians, healthcare facilities, as well as other medical institutions in prioritizing whichever individuals require care and consideration first. A accessible collection of X-ray images from individuals diagnosed with covid illness served as the inspiration for the suggested model. The database is separated into three sensitivity classifications: strong, medium and moderate sensitivity categories. Patients in the highest severity category are at risk of death; those in the moderate category will likely necessitate intensive care; and those in the lowest category will likely recover without hospitalization.

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.379 |

|| Volume 11, Issue 5, May 2023 ||

| DOI: 10.15680/IJIRCCE.2023.1105173|

For the purpose of evaluating verified, fatal, and healed cases of covid throughout the several regions of India, Vishan Kumar Gupta et al. [12] investigate 5 models using machine learning containing 3 key characteristics. Documented cases, deaths, and recoveries are used as systematic evaluation. In this case, machine learning techniques do not incorporate any supplementary data from other models or template architectures. Precision is measured across all variants. As a result of the extensive tests, researchers decided to choose the random forest approach as the highest predictive method for predicting the outcomes of our different scenarios since it surpasses other approaches to machine learning. For every one of these predictions, the random forest algorithm performed almost as well as a generalized linear, and this reliability can be measured using K-fold stepwise regression.

III. PROPOSED SYSTEM

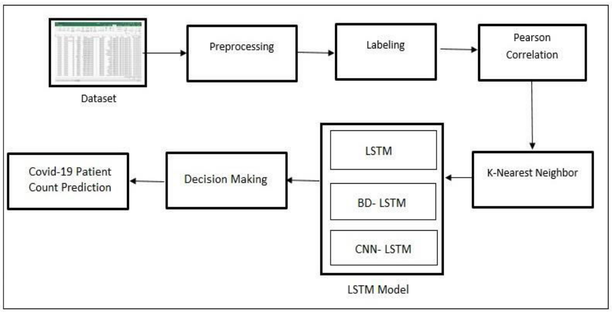


Figure 1: System Overview

The research framework for the predicting of Covid-19 instances has indeed been recognized using deep learning strategies in the combination of Convolutional Neural Networks as well as Long Short Term Memory. The assumption has been validated by the execution of a series of processes, which may be seen in steps given below.

Step 1: Dataset Collection and Data preprocessing –The dataset is provided as an input to this phase of the technique, which marks the beginning of the process that will be followed by the suggested strategy. At the beginning of the process of implementing the suggested model, a dataset comprising India's individual districts and saved in.csv format is loaded into the system as an input. This information was gathered with the help of the website located at the following URL: https://api.covid19india.org/csv/latest/districts.csv. This dataset contains the fields Date, State, District, Recovered, Deceased, Other, and tested in addition to other fields. The parameter in the interface receives the location to the dataset as a quantity that is hard coded into the program. In order to read the dataset as a two-dimensional list, the columns and rows first need to be split, which is done by pursuing the path.

The preparation of the dataset is carried out so that the data may be conditioned, and so that the implementation effectiveness of the suggested forecasting model can be enhanced. The first stage of the preprocessing has indeed been completed, which consisted of determining the input parameters included within the dataset. Following the identification of these parameter categories, the next phase of the strategy involves using the dataset to determine the degree to which the qualities are correlated with one another by using the Pearson Correlation methodology.

Step 2: Pearson Correlation –The Pearson Correlation is used to assess the degree of connection that exists between two or more qualities, and its results are then put to use to discover which characteristics have the least amount of correlation. This produces a correlation matrix, which may be helpful in determining the mix of traits to be chosen for the next step. After this has been established, the characteristics that have a reduced correlation are removed, and new correlation estimates are computed. The following equation 1, which is used to calculate the Pearson correlation, may be found underneath.

$$r = \frac{\sum(x_i-\bar{y})(y_i-\bar{y})}{\sqrt{\sum(x_i-\bar{x})^2 \sum(y_i-\bar{y})^2}} \text{ ------- (1)}$$

Where,
x_i=values of x variable
y_i= values of y variable

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.379 |

|| Volume 11, Issue 5, May 2023 ||

| DOI: 10.15680/IJIRCCE.2023.1105173|

$\bar{x}$= mean of x variable values
$\bar{y}$= mean of y variable values

Only after dataset is collected into the double dimension list, many columns that are deemed to be useless are eliminated from the list. These columns include State, other, and tested. During the process of data segmentation, which is going to be addressed in the next phase, date, district, confirmed, recovered, and deceased are going to be selected as preprocessed list items.

Step 3: K nearest Neighbors – The data that was acquired in the past is now being utilized in order to get the most useful clusters possible with the information. The dataset is used in the K nearest neighbor classification clustering method in order to locate the data points which are the most closely related to each other. In order to accomplish this goal, the magnitude of K is first determined, and then the distances between points are computed, followed by the placement of suitable centroids. As a consequence of this, a list of distances will be produced that is, in essence, ordered in the increasing order of the distances. Following the completion of the previous phase, the resulting list is cut in half, and the highest list is then sent to the subsequent step for the purpose of predicting the infection rate using the CNN-LSTM method. Through using equation 2 that is provided below, the Euclidean distance has indeed been determined.

$$ED = \sqrt{\sum(AT_i - AT_j)^2} \_\_\_\_\_\_\_\_ (2)$$

Where,
ED=Euclidian Distance
A_{Ti}=Attribute at index i
A_{Tj}= Attribute at index j

Step 4: Data segregation –At this point, the preprocessed list is transformed into a segmented list that is particular to a region. A data segmented list is used in order to transform the period into numerical quantities. Additionally, this process turns the string form of information that corresponds to the lengths into a Date type of information. This phase makes it possible for the algorithm to simply arrange or choose rows for the obtained list based on the time periods that were supplied.

Step 3: Training and testing data Creation –At first, the strategy that has been presented makes predictions about the 10 leading regions for the specific type of scenario using the preprocessed list that has been used to create the training and testing data list. This is achieved by employing the sort values function, which requires the input of two variables: the sorting integer, as well as the type. After the 10 leading regions have been chosen, the lists of those who will be trained and tested are compiled.

The information is selected first from preprocessed list in a list somewhere in the center of the two date ranges that were provided. After then, two separate lists are compiled, one for the date, and the second for the identity of the particular region. After that, the method train test split is used in order to build the training and testing sets (). This technique requires the additional characteristics in order to be carried out: X, Y, a testing size of 0.15, and a reorganizing Boolean variable of false. This is due to the fact that the algorithm does not rearrange the information based on the date. After that, the random data for the train as well as the test is created and distributed among 4 distinct lists: train x, train y, test x, and test y.

They are able to easily ingest various types of data because to the fact that both the train and test collections are converted to ordinal values. In order to strike a healthy balance between the test and training lists, the MinMaxScaler() method is called upon. This process standardizes the data by assigning upper and lower boundaries to the range of values. Following the completion of the normalization process, the lists are molded into a single dimension by means of the use of the reshape function ( ). The reshaped data are input into several CNN-LSTM models in order to provide predictions about infection rates for the respective date periods during the COVID-19 pandemic.

Step 4: Convolutional Neural Networks and Long Short Term Memory –The following is the neural network, and its model parameters are as follows: train x, train y, test x, test y, numerical object of normalization, test date, and case category. The capitalize() method is used to properly capitalize the case of words. The conv1D is supplemented with a 32-bit filter that has a kernel capacity of one and uses the ReLU activate function. Additional layer with the same characteristics and activation function that tanh is applied, this time to the one-dimensional representation. A max

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.379 |

|| Volume 11, Issue 5, May 2023 ||

| DOI: 10.15680/IJIRCCE.2023.1105173|

pooling one-dimensional layer that has a pool size of 1 is put into operation here. In contrast to the extra variables, a dropout of 30% is applied across the last hidden layer and the output layer. This takes place between the two layers using the symbol (or 0.3).

The model is built up in many layers, each of which is generated with a unique combination of characteristics. The succeeding second layer employs a filter with a size of 64 and a kernel value of 1, along with comparable padding using tanh as the activation function. The maximum amount of pooling may also be accomplished with a pool size of one and a dropout layer of thirty-five percent, written as 0.35. In addition to the maximum pooling layer, a last one-dimensional convolution layer is implemented with the value 128 serving as the filter. The remainder of the variables are kept the same as they were for the layers that came before it. A new dropout layer is created with a rate of forty percent, which is written as 0.4.

The LSTM architecture is implemented by making use of one hundred blocks, each of which has its activation function set to tanh and its return sequencing set to true. Following that, the dropout layer of 25 percent, or 0.25 percent, was introduced. The dropout frequency of 20%, or 0.2, has been implemented, and the secondary layer of the LSTM uses fifty blocks. The other

characteristics, however, have not been altered, and they remain the same. Beyond this step, the network is flattened, and a dense layer and additional dropout layer with parameters of 20% and 0.2 are introduced. The kernel dimension of the dense layer is set to 6, as well as the activation function is set to softmax. A last dense layer is given a

CNN LSTM	
Layer	**Activation**
CONV 1D 32 Samples,Kernel=1	relu
CONV 1D 32 Samples,Kernel=1	relu
Max Pooling 1D	
Dropout 30%	
CONV 1D 64 Samples,Kernel=1	relu
Max Pooling 1D	
Dropout 35%	
CONV 1D 128 Samples,Kernel=1	Tanh
Max Pooling 1D	
Dropout 40%	
LSTM 50	
Dropput 25%	
Flatten	
Dense 6	Tanh
Dropout 20%	
Dense 1	None
Adam Optimizer	
Batch size 10	
Epochs 250	

kernel capacity of 1 and is not given an activation function when it is first started. Figure 2 has design schematics that present the whole architecture, which may be seen there.

Figure 2: Architecture for CNN LSTM

The used tanh and ReLU activation function are depicted in equation 1 and 2.

$$tanh = \frac{(e^{x}-e^{-x})}{(e^{x}+e^{-x})} \_\_\_\_\_\_ (3)$$

$$Relu=max\ (0,\ x) \_\_\_\_\_ (4)$$

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.379 |

|| Volume 11, Issue 5, May 2023 ||

| DOI: 10.15680/IJIRCCE.2023.1105173|

Where x is the input attributes values

After being built using the adam optimizer, the network would then be trained using a total of 1000 epochs as well as a batch size of 10. The resulting file will have the extension.h5 and will contain the trained algorithm. After that, this file is employed in order to carry out the testing that is covered in the subsequent phase of the method.

Step 5: Decision Making – The testing is carried out for the forecasts made for a certain set of test data. In addition to the testing data for a certain Indian district during a given period of time, the model file is used as an input in the analysis process. The algorithm that has been trained is now being employed in order to achieve the goal of effectively realizing the forecast. After the CNN-LSTM neural network model has been established, the predict() function of the network object is employed to determine the predictions for a certain test list. A graph that was generated by the matplotlib API object displays the findings that were projected for the experiment.

IV. RESULTS AND DISCUSSIONS

The prediction methodology for Covid-19 infection rate forecasting utilizing three distinct kinds of LSTM framework is implemented through using python programming language. For such objective of the execution, a Windows computer equipped with 8 GB of main memory and 1 TB of secondary storage is utilized to run Spyder IDE. The suggested model has been put into practice using the dataset that was obtained from the following URL: https://api.covid19india.org/csv/latest/districts.csv. The dataset includes district-specific information for confirmed cases, dead cases, and recovered cases from the COVID 19 investigation. For the objective of conducting experiments, the seven regions with the highest number of confirmed cases have been chosen, as indicated in figure 3 below.

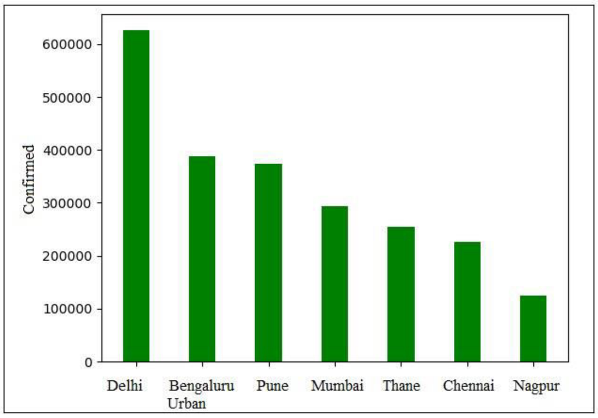


Figure 3: Top 7 Districts for Confirmed cases

Within those areas, the LSTM paradigm, the Bi-directional LSTM framework, and the Convolution LSTM framework are all actively used as LSTM applications. Data from the previous days are now being utilized for the training and testing purposes, and forecasts are now being generated for the next seven days. The resulting estimates are reported in the following table 1, and figures 4, 5, and 6 accordingly represent the related graphs.

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.379 |

|| Volume 11, Issue 5, May 2023 ||

| DOI: 10.15680/IJIRCCE.2023.1105173|

			LSTM			
Delhi	Bangaluru U	Pune	Mumbai	Thane	Chennai	Nagpur
638258	404242	400487	321408	277620	234653	145112
638207	404119	402665	321438	277455	234671	146162
638013	403954	401283	320076	276505	234530	145658
638081	404041	401689	320496	276814	234554	146059
638133	404091	401904	320789	276980	234614	146267
638115	404086	401950	320563	276799	234625	146264
638060	404053	401904	320376	276658	234604	146232
			BiDirectional LSTM			
Delhi	Bangaluru U	Pune	Mumbai	Thane	Chennai	Nagpur
637987	404178	402610	319367	276492	234541	146653
638272	404222	402160	321997	277448	234709	146545
638068	404001	401355	320126	276605	234536	145831
638079	404042	401746	320271	276748	234571	146201
638095	404090	402353	320512	276906	234600	146271
638094	404092	401976	320581	276953	234598	146112
638087	404067	401447	320392	276872	234595	145960
			Conv LSTM			
Delhi	Bangaluru U	Pune	Mumbai	Thane	Chennai	Nagpur
638045	403966	400988	320182	276540	234679	145845
638214	404193	404055	321735	277032	234606	147403
638119	404009	401342	320401	276699	234520	145985
638120	404027	401504	320444	276714	234555	146029
638126	404054	401667	320517	276738	234581	146059
638132	404067	401740	320559	276767	234589	146088
638132	404065	401775	320570	276797	234584	146106

Table 1: Different LSTM model prediction outcomes

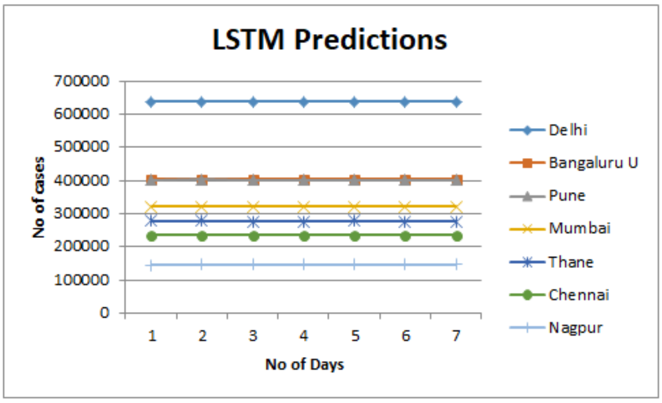


Figure 4: Prediction outcomes for LSTM

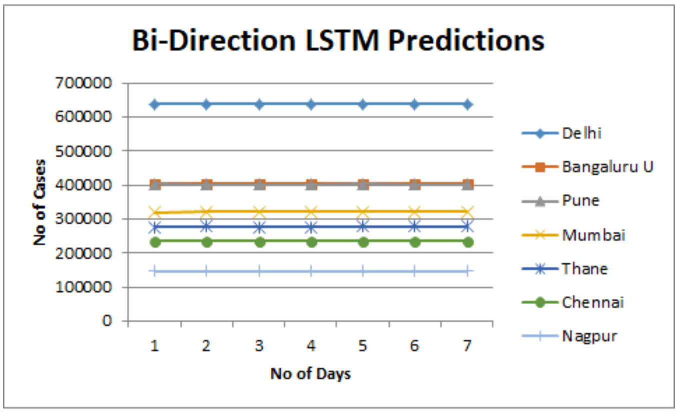


Figure 5: Prediction outcomes for Bidirectional LSTM

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.379 |

|| Volume 11, Issue 5, May 2023 ||

| DOI: 10.15680/IJIRCCE.2023.1105173|

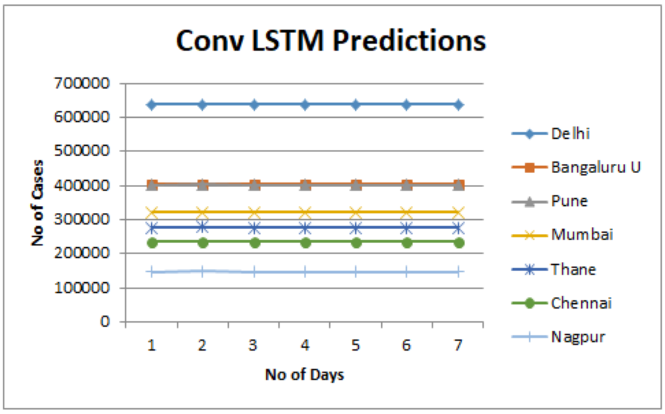


Figure 6: Prediction outcomes for Convolutional LSTM

The performance reliability assessment of the three LSTMs that were built tends to measure their performance by acquiring their average absolute discrepancy in between produced forecasts and real forecasts for the specified dates. This is done in order to measure how well the models are doing.

V. CONCLUSION AND FUTURE SCOPE

The proposed methodology for the goal of obtaining precise and reliable forecasts for covid-19 rate of infection has indeed been incorporated with the aid of Long Short Term Memory. The input dataset is used as an input and then preprocessed to eliminate any data redundancy before being sent on to the subsequent module for the purpose of partition. After the data have been preprocessed, they are separated according to districts, and then they are sent to the LSTM component. The LSTM component is the one that is the recipient of the segmented data, which is afterwards normalized and mapped before being divided into sections for the objectives of testing and training. During training, the LSTM component is exposed to three distinct categories of cases: confirmed, recovered, and dead. The Deep Learning framework emerges as a solution in this framework that allows for an efficacious predictive model of the Covid-19 rates of infection to be accomplished through the deployment of three distinct kinds of LSTM models. These models are the basic LSTM, the Conditional LSTM, as well as the Bi-Directional LSTM. In this circumstance, the Deep Learning framework comes into play and allows for an impactful forecasting of the Covid-19 rates of infection. The accuracy of the Covid-19 transmission rate forecasts has been shown to be proven by the results of the experiments.

Using a cloud infrastructure and a Generative adversarial Neural Network, it will be possible in the future to apply the algorithm to enormous datasets that include hundreds of characteristics for each area of the globe. These datasets will cover the entire world.

REFERENCES

[1] V. K. Prasad et al., "ABV-CoViD: An Ensemble Forecasting Model to Predict Availability of Beds and Ventilators for COVID-19 Like Pandemics," in IEEE Access, vol. 10, pp. 74131-74151, 2022, doi: 10.1109/ACCESS.2022.3190497.
[2] E. Karaçuha et al., "Modeling and Prediction of the Covid-19 Cases With Deep Assessment Methodology and Fractional Calculus," in IEEE Access, vol. 8, pp. 164012-164034, 2020, doi: 10.1109/ACCESS.2020.3021952.
[3] J. Piri, P. Mohapatra, H. K. R. Singh, B. Acharya and T. K. Patra, "An Enhanced Binary Multiobjective Hybrid Filter-Wrapper Chimp Optimization Based Feature Selection Method for COVID-19 Patient Health Prediction," in IEEE Access, vol. 10, pp. 100376-100396, 2022, doi: 10.1109/ACCESS.2022.3203400.
[4] O. Tutsoy, Ş. Çolak, A. Polat and K. Balikci, "A Novel Parametric Model for the Prediction and Analysis of the COVID-19 Casualties," in IEEE Access, vol. 8, pp. 193898-193906, 2020, doi: 10.1109/ACCESS.2020.3033146.

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.379 |
|| Volume 11, Issue 5, May 2023 ||
| DOI: 10.15680/IJIRCCE.2023.1105173|

[5] K. Panetta, F. Sanghavi, S. Agaian and N. Madan, "Automated Detection of COVID-19 Cases on Radiographs using Shape-Dependent Fibonacci-p Patterns," in IEEE Journal of Biomedical and Health Informatics, vol. 25, no. 6, pp. 1852-1863, June 2021, doi: 10.1109/JBHI.2021.3069798.
[6] F. Rustam et al., "COVID-19 Future Forecasting Using Supervised Machine Learning Models," in IEEE Access, vol. 8, pp. 101489-101499, 2020, doi: 10.1109/ACCESS.2020.2997311.
[7] L. Wang et al., "A Time-Series Feature-Based Recursive Classification Model to Optimize Treatment Strategies for Improving Outcomes and Resource Allocations of COVID-19 Patients," in IEEE Journal of Biomedical and Health Informatics, vol. 26, no. 7, pp. 3323-3329, July 2022, doi: 10.1109/JBHI.2021.3139773.
[8] M. Wasesa et al., "Predicting Electricity Consumption in Microgrid-Based Educational Building Using Google Trends, Google Mobility, and COVID-19 Data in the Context of COVID-19 Pandemic," in IEEE Access, vol. 10, pp. 32255-32270, 2022, doi: 10.1109/ACCESS.2022.3161654.
[9] L. Padilla, R. Fygenson, S. C. Castro and E. Bertini, "Multiple Forecast Visualizations (MFVs): Trade-offs in Trust and Performance in Multiple COVID-19 Forecast Visualizations," in IEEE Transactions on Visualization and Computer Graphics, vol. 29, no. 1, pp. 12-22, Jan. 2023, doi: 10.1109/TVCG.2022.3209457.
[10] E. Casiraghi et al., "Explainable Machine Learning for Early Assessment of COVID-19 Risk Prediction in Emergency Departments," in IEEE Access, vol. 8, pp. 196299-196325, 2020, doi: 10.1109/ACCESS.2020.3034032.
[11] S. A. -F. Sayed, A. M. Elkorany and S. Sayed Mohammad, "Applying Different Machine Learning Techniques for Prediction of COVID-19 Severity," in IEEE Access, vol. 9, pp. 135697-135707, 2021, doi: 10.1109/ACCESS.2021.3116067.
[12] V. K. Gupta, A. Gupta, D. Kumar and A. Sardana, "Prediction of COVID-19 confirmed, death, and cured cases in India using random forest model," in Big Data Mining and Analytics, vol. 4, no. 2, pp. 116-123, June 2021, doi: 10.26599/BDMA.2020.9020016.

INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN COMPUTER & COMMUNICATION ENGINEERING

9940 572 462 6381 907 438 ijircce@gmail.com