International Journal of Research in Engineering and Science (IJRES)
ISSN (Online): 2320-9364, ISSN (Print): 2320-9356
www.ijres.org Volume 11 Issue 4 ǁ April 2023 ǁ PP. 111-116

Prediction of Covid-19 Confirmed, Death and Cured Cases in India

NarmadhaVarshini N ,Priyadhatshini D ,YamunaDevi V,Yasotha S, P.Yogashree

*Department of Computer Science and Engineering, Vivekanandha College of Engineering for Women, Namakkal, Tamilnadu, India. Email: priyadharshini8866@gmail.com *Corresponding author*

Abstract
The global proliferation of COVID-19 has placed mankind in danger. Because of the disease's high infectivity and transmissibility, the resources of some of the world's most powerful economies are being taxed. The capacity of machine learning algorithms can anticipate the number of forthcoming patients impacted by COVID-19, which is now regarded as a possible threat to humanity. In this work, four conventional forecasting models, least absolute shrinkage and selection operator LR were utilized to anticipate the COVID-19 hazardous elements. Each model makes three sorts of predictions: the number of newly infected cases, the number of fatalities, and the number of recoveries. However, in the cannot predict an exact outcome for the patients. To address the problem, the suggested technique employing LR predicts the number of COVID-19 cases in the following 30 days as well as the influence of preventative measures such as social isolation and lockout on the spread of COVID-19.

Keywords: *Coronavirus; COVID-19; respiratory tract; multi-class classification; random forest*

Date of Submission: 20-03-2023 Date of acceptance: 04-04-2023

I. INTRODUCTION

1.1 COVID-19

COVID-19, the global pandemic, has exposed human society's vulnerability to severe infectious illnesses as well as the challenge of treating this problem in a globally integrated complex system. COVID-19 impacted almost 100 nations in a few of weeks. As a result, the entire human species should not only work together to defeat the pandemic, but also make reasonable plans to return to work and production based on the real circumstances in each location, as well as conduct geographical risk assessments. Many attempts have been made to develop an appropriate and rapid method of detecting infected people at an early stage. Guan et colleagues discovered bilateral pulmonary parenchymal ground-glass and consolidative pulmonary opacities, occasionally with a rounded shape and a peripheral lung distribution, after performing chest CT scans on 21 COVID19 infected individuals in China. As a result, COVID-19 diagnosis may be portrayed as an image segmentation issue to extract the disease's major characteristics. Coronavirus Disease 2019 (COVID-19), caused by a new coronavirus, is rapidly spreading over the world. As of April 9, 2020, it had infected over 1,436,000 persons in over 200 nations and territories.

1.2 FUTURE FORECASTING

Forecasting is the practice of creating future predictions based on past and present data, most typically using trend analysis. A common example would be the estimate of some variable of interest at some future date. Prediction is a similar but broader phrase. Both terms can relate to formal statistical approaches that use time series, cross-sectional, or longitudinal data, as well as less formal judgmental procedures. In hydrology, for example, the phrases "forecast" and "forecasting" are frequently reserved for estimates of values at specific future periods, but the term "prediction" is used for more general predictions, such as the number of times floods will occur over a lengthy period. Forecasting and prediction rely heavily on risk and uncertainty; it is usually regarded good practice to highlight the degree of uncertainty attached to projections. In any event, the data must be current if the forecast is to be as precise as feasible. In certain circumstances, the data needed to forecast the variable of interest is predicted itself.

1.3 SUPERVISED MACHINE LEARNING

The machine learning job of learning a function that translates an input to an output based on example input-output pairs is known as supervised learning. It derives a function from labeled training data, which

consists of a collection of training samples. Each example in supervised learning is a pair consisting of an input item (usually a vector) and a desired output value (also called the supervisory signal). A supervised learning algorithm examines the training data and generates an inferred function that can be used to map fresh samples. In an ideal case, the algorithm will be able to accurately predict the class labels for unknown examples. This necessitates that the learning algorithm generalize from the training data to previously encountered scenarios in a "reasonable" manner. Concept learning is a term used to describe the parallel task in human and animal psychology.

II. LITERATURE REVIEW

2.1 USING EXISTING CELLULAR WIRELESS NETWORK FUNCTIONALITIES TO DETECT REGIONS AT RISK FOR COVID-19 SPREAD.

In this research, Alaa A. R. Alsaeedy and Edwin K. P. Chong et al. suggest a novel technique for identifying places with high population density and mobility that are at risk of spreading COVID-19. Crowded areas with actively moving individuals (referred to as at-risk areas) are more likely to transmit the disease, especially if they contain asymptomatic infected persons alongside healthy ones. Methods: Our system detects at-risk areas by leveraging current cellular network features such as handover and cell (re)selection, which are utilized to provide smooth coverage for mobile end-user equipment (UE). We leverage pre-existing cellular-network functions to control end-user movement and assure smooth coverage. Because almost everyone owns a cellular mobile device (known as user equipment (UE)), these function as always-on human trackers. More particular, as the number and mobility of UEs increases, so does the number and mobility of individuals. According to a recent study, SARS-CoV-2 can survive in the air for up to three hours (as aerosols) after being exhaled by infected patients while speaking, coughing, or even breathing, whether symptomatic or not. We are especially concerned about the situation in which infectious persons are present in regions with a high density of continually moving people.

2.2 USING MACHINE LEARNING TO QUANTIFY COVID-19 CONTENT IN THE ONLINE HEALTH OPINION WAR

A large quantity of possibly deadly COVID-19 disinformation is surfacing online, according to Richard f. Sear,

Nicolas Velásquez et al. We apply machine learning to assess COVID-19 material among online opponents of conventional health advice, namely vaccines ("anti-vax"). We discovered that the anti-vax group is having a less concentrated argument about COVID-19 than the pro-vaccination ("pro-vax") community. The anti-vax community, on the other hand, demonstrates a greater range of "flavours" of COVID-19 issues, and so can appeal to a broader cross-section of those seeking COVID-19 counsel online, such as those skeptical of a mandated fast-tracked COVID-19 vaccination or those seeking alternative cures. As a result, the anti-vax movement appears to be in a better position to gain new supporters in the future than the pro-vax community. This is important because if a COVID-19 vaccine is not widely adopted, the globe will fall short of providing herd immunity, leaving countries vulnerable to future COVID-19 outbreaks. We present a mechanistic model that explains these findings and may aid in determining the potential success of intervention techniques. Our technique is scalable, addressing the pressing issue that social media platforms have in analyzing massive amounts of online health misinformation and deception.

2.3 DEEP LEARNING WITH WEAK SUPERVISION FOR COVID-19 INFECTION DETECTION AND CLASSIFICATION FROM CT IMAGES

In this study, Shaopinghu, Yuangao, and colleagues offer Since late December 2019, Wuhan, China has been experiencing an outbreak of a new coronavirus illness (COVID-19), which has since spread globally. Although COVID-19 is an urgently treated condition, it can be lethal, with a mortality rate of 4.03% in China and 13.04% in Algeria and 12.67% in Italy (as of 8th April 2020). As a result of significant alveolar injury and gradual respiratory failure, the beginning of severe sickness may result in death. Although laboratory testing, such as reverse transcription polymerase chain reaction (RT-PCR), is the gold standard for clinical diagnosis, false negatives can occur. Furthermore, in a pandemic situation, a lack of RT-PCR testing resources may cause the ensuing clinical decision and treatment to be delayed. In such cases, chest CT imaging has shown to be an invaluable tool for both diagnosis and prognosis of COVID-19 patients. We propose a poorly supervised deep learning technique for identifying and categorizing COVID-19 infection from CT scans in this work. The suggested technique can reduce the need for manual labeling of CT images while still detecting infections and distinguishing COVID-19 patients from non-COVID-19 cases.

2.4 COVID-19 RISK ASSESSMENT BASED ON MULTISOURCE DATA FROM A GEOGRAPHICAL PERSPECTIVE

In this research, Yan Zhang, Yingbing L et al. argued that Corona Virus Disease 2019 (COVID-19) cases in Wuhan were eradicated, and the epidemic situation was mostly under control. Such public health infectious illness has a significant impact on the national economy. At the moment, several nations and regions throughout the world are still in epidemic mode, and it is critical to assess the infection situation and travel danger in the region. To assess the surrounding circumstances on a reasonably fine scale, and then reasonable zoning decisions to support the restoration of production and activity. In this study, indications for assessing the COVID-19 outbreak were created utilizing data from many sources. The GeoDetector model and the decision tree model were used to compute an evaluation of 736 fine-grained grids. The study discovered that the risk level was substantially greater in older areas than in modern communities; population density was the most important factor of infection.

2.5 A HYBRID COVID-19 DETECTION MODEL BASED ON AN IMPROVED MARINE PREDATORS ALGORITHM AND A DIVERSITY REDUCTION STRATEGY BASED ON RANKING

Many nations are challenged by the medical resources necessary for COVID-19 detection, as proposed by

Mohamed Abdel-Basset, Reda Mohamed, and colleagues in this study, necessitating the creation of a low-cost, quick method to detect and diagnose the virus efficiently for a large number of tests. Although a chest X-Ray scan is a valuable candidate tool, the pictures produced by the scans must be precisely and rapidly examined if a significant number of tests are to be processed. COVID-19 induces bilateral pulmonary parenchymal ground-glass and consolidative pulmonary opacities, with a rounded shape and a peripheral lung distribution in certain cases. The goal of this study is to quickly extract from chest X-Ray pictures comparable tiny areas that may have COVID-19 diagnostic traits. As a result, this research offers a hybrid COVID-19 detection model for X-Ray picture segmentation based on an enhanced marine predators algorithm (IMPA). The ranking-based diversity reduction (RDR) technique is utilized to improve the IMPA's performance and achieve better results in less iteration.

III. EXISTING SYSTEM

COVID 19 is now seen as a possible hazard to mankind. In this work, four conventional prediction models were utilized to predict COVID-19 hazardous elements, including linear regression (left to right), at least complete summary, and choose operator, Support Vector Machine (SVM). Predictions are produced based on each of the models, such as the number of new infections, fatalities, and recurrences during the following 10 days. The study's findings provide a promising mechanism for using these approaches in the current setting of COVID 19 infection. Predictions are produced based on each of the models, such as the number of new infections, fatalities, and recurrences during the following 10 days. The study's findings provide a promising mechanism for using these approaches in the current setting of COVID 19 infection. COVID-19 does not appear to have a serious impact on children; several paediatric wards have been focused on the urgency of COVID-19-related concerns. As a result, many other acute and chronic disorders, particularly those that are infrequent, may receive insufficient care. This lack of interest can lead to serious complications or even death, especially in children.

IV. PROPOSED SYSTEM

Machine learning approaches were effective for prediction because they automatically extracted important characteristics from training samples, used activation from the previous time step as input for the current time step, and used network self-connections. According to the results of the model analysis, we believe that the early stage emergency intervention actions, such as blocking, restricting the flow of people, and enhancing assistance, had a critical restraining influence on the original expansion of the epidemic. Continued investment in diverse medical resources to guarantee that suspected patients may be identified and treated in a timely way is a very successful preventative and treatment technique. The AI calculation LR every one of these calculation are utilized and the best calculation are arranged in the r-squared blunder and the changed r-squared mistake.

4.1 DATA

The data contains the total number of confirmed cases, total number of fatalities, total number of newly confirmed cases, and total number of healed patients per province. We also utilized data on recent diagnoses in South Korea, Iran, and Italy, which contains data from official announcements from various nations. All data is from the daily case report, and the data is updated once every day.

4.2 ESTIMATION PROCESS

The Basic reproduction number varies substantially between control stages and has a direct impact on control intensity. Furthermore, the virus's incubation time has a direct impact on the virus's transmission pace. These

two parameters must be calculated. According to current research, unrestrained Basic reproduction. As a result, we selected the valuation range in the appropriate range. The value range for the regulated Basic reproduction number was chosen to be [0, 1.5].

4.3 DATA-DRIVEN METHODS TO PREDICT COVID-19

The data was used (when the first incidence of COVID-19 was reported in India), with 80% of the data used for training and the remaining 20% for predicting and validation. The resultant figure depicts the total number of confirmed cases, the observed data is the data used for training, the official data (green line) is the official data available, and the predicted data represents the total number of confirmed cases forecast. According to this graph, the predicted number of total verified positive cases roughly matches the existing official data.

4.4 DATA PRE PROCESSING

Data preprocessing is a technique for converting raw data into a clean data collection. The dataset is frequently inadequate, inconsistent, and/or deficient in specific behaviors or patterns, and it is rife with mistakes. Data preparation is a tried-and-true way for dealing with such problems.

4.5 PREDICTION OF ACCURACY

This approach is appropriate for use with predicting neural networks or characteristic data such as infection event or non-event binomial effects. The forecast accuracy of different metrics may be utilized for a variety of applications. They include the rate at which normal (unpredicted forecast accurately predicts sensitivity (non-infectious disease), accuracy (expected percentage of predicted trend), positive predictive value, negative predictive value (correctly predicted infection rate is), and the ratio. Expected forecasts are a measure of the chance that the growth in the whole process surpasses the individual's accuracy).

4.6 CLASSIFICATION

For each data set point, the classification approach predicts the target class. A risk factor can be connected with patients using the classification strategy by studying their illness patterns.

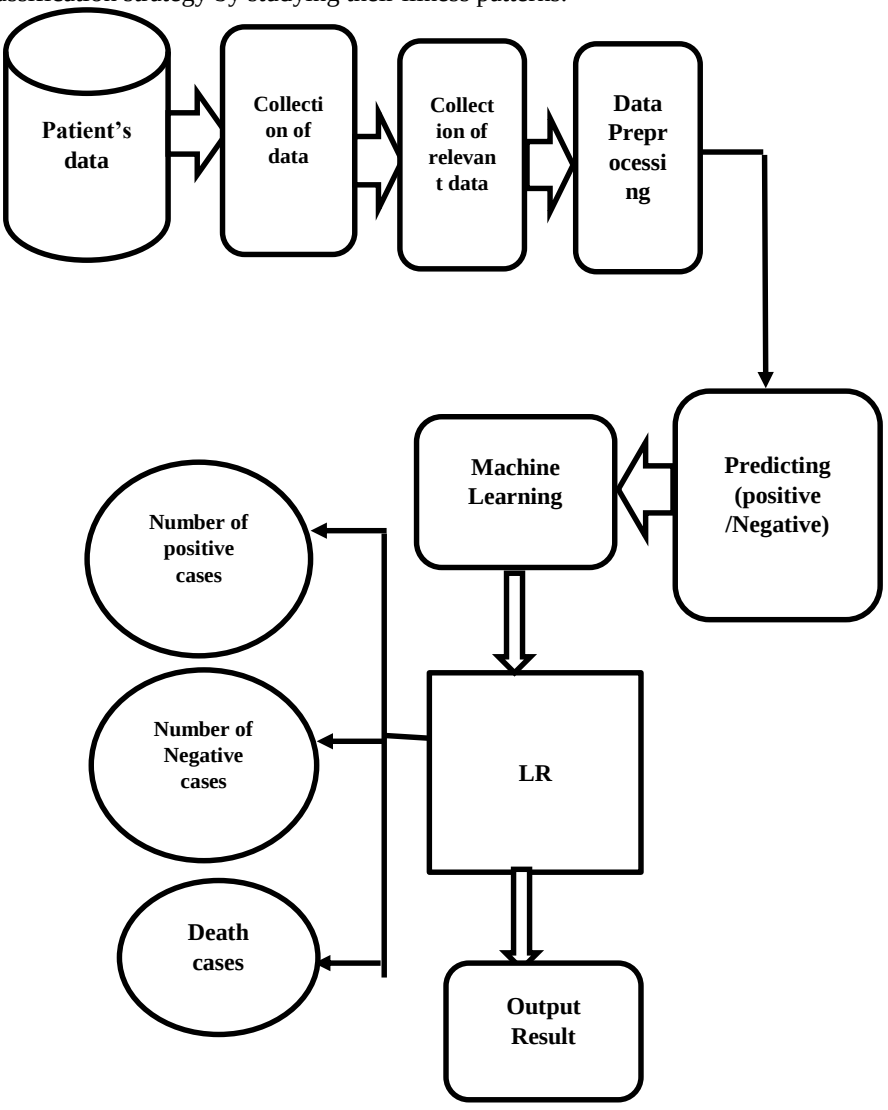

Figure 1 : Classification

V. RESULT ANALYSIS

To develop a structure for the future deciding of the amount of cases impacted by COVID-19 using AI procedures. The dataset used for the assessment contains information about the step by step reports of the amount of as of late tainted cases, the amount of recoveries, and the amount of passings in light of COVID-19 all throughout the planet. As the death rate and avowed cases are extending bit by bit which is an upsetting situation for the world. The amount of people who can be affected by the COVID-19 pandemic in different countries of the world isn't prominent. This assessment is an undertaking to calculate the amount of people that can be affected similarly as new sullied cases and passing's including the amount of expected recoveries for the impending 10 days. Four AI models LR, LASSO, SVM, ES and LSTM have been used to predict the amount of as of late debased cases, the amount of passings, and the amount of recuperations.

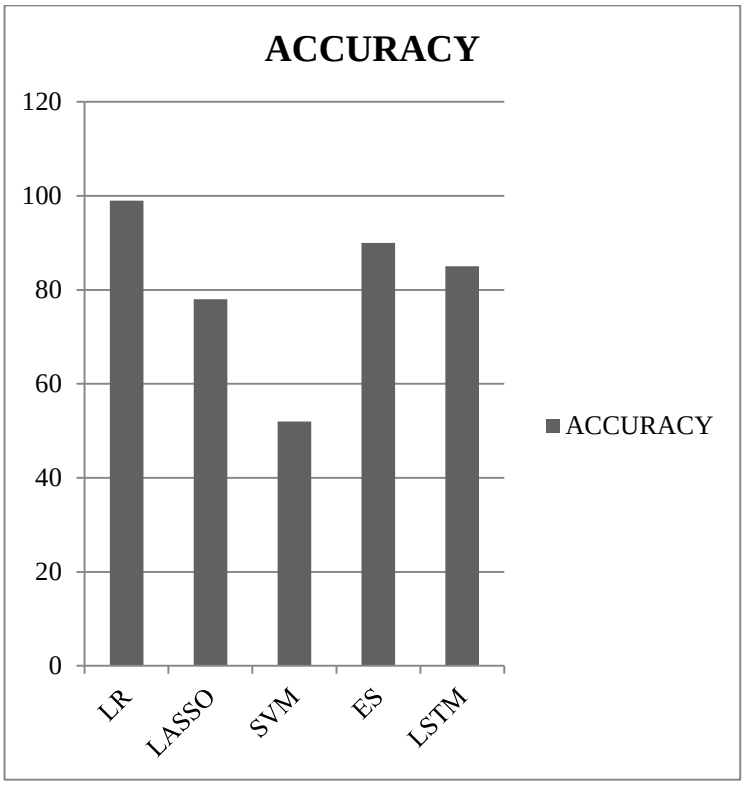

Figure 2 : Accuracy

The plots of asserted cases, passing's, and recoveries on the underlying four sheets followed by the plot of certifiable situation amassed from the genuine data reports of the analyzing season of the assessment in the fifth sheet. The results in the charts show that the ML models used in this assessment befit the assessing task making the course towards the comfort of the examination and future investigation of the near nature.

ALGORITHM	ACCURACY
LR	99
LASSO	78
SVM	52
ES	90
LSTM	85

Figure 3 : Accuracy Table

VI. CONCLUSION

A data-driven forecasting/estimation strategy was utilized to anticipate the likely number of COVID-19 positive cases in India during the following 30 days. Using and curve fitting, the number of recovered cases LR LASSO, SVM, ES, LSTM daily positive cases, and died cases was also estimated. The effect of preventative measures such as social isolation and lockdown has also been demonstrated, indicating that these preventive

measures can considerably minimize the transmission of the infection. Although this method frequently requires sufficient data to support it, in the early stages of epidemic transmission, this method can still be used to more accurately predict the indicators of epidemic transmission in the short term, so that intervention control can be provided at all levels of departments and policy implementation can provide short-term emergency prevention programmes. For various factors and in different places, the forecast results of three distinct mathematical models differ. In general, the Logistic model may have the best fitting effect among the three models.

REFERENCES

[1]. Alsaeedy, A. A. R., and E. Chong (2020). Using Existing Cellular Wireless Network Functionalities, Detecting Regions at Risk of COVID-19 Spread. 1-1 in the IEEE Open Journal of Engineering in Medicine and Biology.

[2]. R. F. Sear, N. Velasquez, R. Leahy, N. J. Restrepo, S. El Oud, N. Gabriel,... Johnson, N. F. (2020). Using machine learning to quantify COVID-19 material in the online health opinion war. 1-1, IEEE Access.

[3]. Hu S., Gao Y., Niu Z., Jiang Y., Li L., Xiao X.,... Yang G. (2020). Deep Learning with Weak Supervision for COVID-19 Infection Detection and Classification from CT Images. 1-1, IEEE Access.

[4]. Y. Zhang, Y. Li, B. Yang, X. Zheng, and M. Chen (2020). COVID-19 Risk Assessment Using Multisource Data from a Geographical Perspective. 1-1, IEEE Access.

[5]. M. Abdel-Basset, R. Mohamed, M. Elhoseny, R. K. Chakrabortty, and M. Ryan (2020). A hybrid COVID-19 detection model that employs an upgraded marine predator algorithm as well as a ranking-based diversity reduction method. 1-1, IEEE Access.