COVID-19 PREDICTION MODEL: A LITERATURE SURVEY

*Deepti Malhotra, **Gurinder Kaur Sodhi

*Department of Electronics & Communication Engineering, Desh Bhagat University (Punjab) India 147301
**Department of Electronics & Communication Engineering, Desh Bhagat University (Punjab) India 147301
Corresponding Author EmailID: deeptimalhotra1981@gmail.com

Abstract
Believed to have been originated Chinese province Wuhan in December 2019, the coronavirus havesaid to cause 95 million cases with overall death rate of 2%of overall cases(as per Jan 2022).
This fast spreading pandemic virus poses a challenge at world level and proposes serious danger to people's health as well as the economy. With time and regions this virus has undergone several mutations resulting in rise of various other viruses, OMICRON being the latest.The most common and widely faced threat in this disease was in the case of asymptomatic patients, the ones who showed no symptoms and yet were carriers of corona virus. In recent times,many researchers have started exploring various methods to predict the disease on the basis of medical parameters. Few of the commonly tools used are machine learning and artificial intelligence. The present paper aims to compile the various models used by researchers in last year in predicting COVID.
Keywords: covid19, prediction techniques, artificial intelligence, machine learning

1.Introduction
1.1 Covid 2019
The coronavirus originated in Chinese province, Wuhan in December 2019. The coronavirus disease 2019 is a virus of the coronavirus family. Studies suggest that COVID-19 encompasses medicaltra its similar to SARS-CoV [1]. Major symptoms are fever and cough, while gastrointestinal signs are rare. The absence of fever in novel coronavirus- infected people is more common than in patients diseased with the same virus, that is, MERS corona virus (2%) and SARS corona virus (1%). Thus, non-febrile patients are likely to be exempted from a monitoring device with the main attention on diagnosing fever.

1.2 Machinelearning
The healthcare sector is a huge sector requiring medical data whose collection and processing is required in real-time.
The emergence of AI has proved to be a technological breakthrough of the 21th century. It includes many applications in areas ranging from weather forecasting, astrophysical investigation to self-directed systems. Machine learning based prediction has gained hugeadmiration in the last few years. The course of learning is basically learning from past work experience or observations, such as instances, or instructions, to see patterns in data. This process makes use of examples to improve the decision-making ability of the system[2].
1.3 Covid-19 Disease Prediction using Machine-Learning

The epidemiological data containing the number of people who suffered with this disease previously and the total population is employed in the predictive models of corona virus 2019. Different parameters namely latency period and probability of healing are utilized for predicting the infection trends. But, these models are inefficient to reflect diverse socio-static and economic factors due to which the course of the virus is affected. All stages of this predictive framework are as;Data collection, Relevant dataset, Pre-processing, Feature Selection and Machine learning
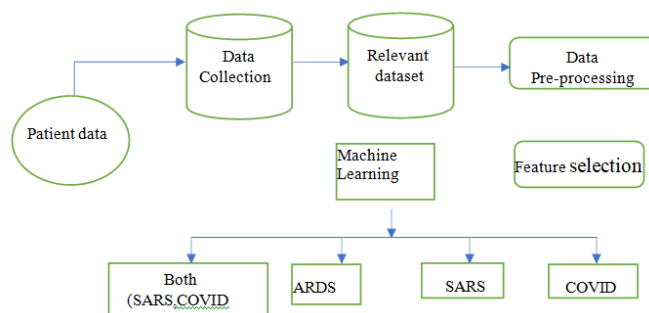

Fig. 1.1.General Composition of COVID-19 prediction model
Following are the some commonly used Covid-19 outbreak prediction models.
K-Nearest neighbor algorithm
Logistic Growth Model
Long short term memory

2.LITERATURE REVIEW
Pratima Kumari, et.al (2020) intended ML (machine learning) method called ANN in order to predict the

Coronavirus Disease 2019 outbreak in India for the first time [8]. In addition, a mathematical curve fitting system was adopted for determining the performance of the intended method. The development of the epidemic at diverse transmission rates for analyzing the impact of preventive measures such as lockdown and social distancing. The intended method was proved appropriate to predict the development rate of metrics of corona disease at lower MAPE values and cumulative deceasedcases.

Yasin Khan, et.al (2020) formulated a CNN algorithm to carry out a multi-class classification of the chest X-ray images of Coronavirus Disease 2019 as infected patients,patients of pneumonia and healthy persons [9]. Moreover, the Monte-Carlo simulation was utilized on the original data distributions for improving the data so that the disease was predicted. Moreover, COVID-19 pandemic was forecasted using LR of the components of GMM. X-ray pneumonia images. The accuracy obtained from the formulated algorithm was computed 100% in training phase and 96.66% in the testingphase.

Choujun Zhan, et.al (2021) projected a ML (machine learning) mechanism in order to forecast the COVID-19 on the basis of BLS (Broad Learning System) [10]. The significant attributes were extracted using RF (Random Forest). Therefore, RF-Bagging-BLS (Random- forest-Bagging Broad Learning System) was constructed by integrating the bagging strategy with BLS with the objective of predicting the COVID-19 pandemic. The projected mechanism performed more effectively with regard to RMSE (relative mean square error), MAD and MAPE in comparison with other models. Thus, the projected mechanism outperformed the traditionalmodels.

Leonardo et. al(2021) introduced an ANN with the objective of predicting the number of covid confirmed cases and deaths and also the future seven days for the time series of Brazil, Portugal and the United States [12]. The simulation results depicted that the introduced method was capable to predict the confirmed cases and deaths from Coronavirus Disease 2019. The ANN utilized with

a specific test set provided 50% higher MSE as compared to the Artificial Neural Network with a random test set. The least MSE was acquired for all scenarios when the sigmoidal and linear activation functions were integrated with the LMfunction.

Durga Prasad Kavadi, et.al (2020) recommended a PDR-NML technique to predict Coronavirus Disease 2019 which was spread all over the world [13]. The optimal metrics were investigated efficiently using

PPDLR model. Subsequently, the NML was adopted for the normalized attributes so as the accurate predictions were achieved. The results indicated the superiority of the recommended technique over existing schemes in the Indian population and it was presented as a promising tool to make predictions for other countries.

Ankan Ghosh Dastider, et.al (2021) investigated a CNN (convolutional neural network) algorithm in which the fusion of an autoencoder network and separable convolutional branches was done with a modified DenseNet-201 network [14]. Hence, a vigorous, noise- free classification system was constructed. The accuracy of investigated algorithm was evaluated using a five-fold cross-validation system. The in-depth result analysis proved that when LSTM layers were utilized, the classification performance was enhanced. This algorithm provided performed more effectively as compared to conventional techniques. The extensive analysis confirmed the applicability of the investigated algorithm to detect the severity scores of Coronavirus Disease 2019 from LUS images.

Abhishek Dixit, et.al (2021) established a 3-step approach that employed K-means clustering and extracted the attributes in the data pre-processing stage [15]. The subsequent stage focused on optimizing the selected attributes. For this, a novel feature optimization technique was deployed on the basis of hybrid DEA (differential evolution algorithm) and PSO. Thereafter, the SVM made the utilization of optimized attributes. The empirical outcomes demonstrated that the established approach yielded the accuracy around 99.34% and their robustness and adaptability of this approach was also proved in detecting the patients suffered from COVID-19.

Elena Casiraghi, et.al (2020) constructed a computerized mechanism whose major intend was to extract the most relevant radiological and laboratory variables for improving the process of predicting the patient risk [16]. After that, an adequate ML (machine learning) model was put forward. A robust and consistent variable was selected by integrating Boruta with RF (Random Forest) in a 10-fold cross-validation system for achieving a variable importance estimate that was impartial when the surrogates were available. The selection of significant variables was done for training the RF algorithm whose robustness and efficacy was presented. The implementation of the developed system was easy and assisted the emergency departments in predicting the rapid and accurate risk in COVID-19 patients.

Anwar Jarndal, et.al (2020) discussed that a model was formulated for estimating the number of deaths in the affected cases on the basis of documented number of older diabetic and smoking cases [17]. The GPR algorithm was adopted for building a model. A comparative analysis was done on this algorithm against ANN (Artificial Neural Network) model. A trustworthy data gathered from WHO for distinct countries was utilized in this algorithm. The adopted algorithm was efficient to estimate the number of deaths because of any arbitrary number of inputs. The effective measures were planned for mitigating the number of deaths using this algorithm.

Huan Zhao, et.al (2021) designed the BP (back propagation neural network) algorithm for analyzing and predicting the growth rate of Coronavirus Disease 2019 [18]. Initially, the gathering of epidemic data of Italy was done from WHO. Subsequently, R and MATALB were employed to analyze a number of indicators. A computer language was executed to put together the results obtained from multiple indicators for predicting the COVID-19 in Italy. The predictive outcomes indicated that the trend of Italy epidemic situation in the study period was found good. But, the measures were required to prevent and control the corona disease. Eventually, a comparison of fitting value of the designed algorithm was done with real value and the fitting degree was obtained upto0.99.

Nanning Zheng, et.al (2020) suggested a hybrid AI (artificial-intelligence) technique in order to forecast the Coronavirus Disease 2019 [19]. At first, an ISI (improved susceptible-infected) model was presented for estimating several infection rates so that the transmission laws and the growth trend were analyzed. Later, the effects of measures taken to prevent and control the disease were taken in account for embedding the NLP (natural language processing) and LSTM network into the presented model with the objective of developing the suggested hybrid technique to predict coronavirus disease. The outcomes obtained in experimentation depicted that the suggested technique was capable of mitigating the errors of the predictive results.

Safa Bahri, et.al (2020) intended a LSTM (Long Short-Term Memory)-based predictive framework to predict the future cumulative number of recovered cases of CoronaVirus Disease 2019 seven days earlier, in India, USA and Italy [20]. This framework was suitable for the dynamic of evolution of recovered cases of the observed Coronavirus Disease 2019 and predicting the disease accurately with a total error within 1.46 to 2.65%. The experimental results revealed the effectiveness of intended framework which provided a minimal error below 3%. This model was proved as an effective tool for practitioners for monitoring the occurrence of corona disease all over theworld.

Teddy Mantoro, et.al (2020) introduced two techniques namely SVR and SIR system for predicting COVID-19 [21]. The disease was predicted in the best-case as well as worst-case scenario. The initial scenario included the current daily case as a maximum case and another country's maximum case were contained in the latter scenario utilized. The SVR algorithm provided diverse end of epidemic. The SIR model was utilized to validate the similar end of epidemic on January 2021 in both cases. However, a rapid increase of infectious people out of 450,000 people was found in initial scenario to 5,500,000 people in latter scenario.

The prediction obtained using the introduced approach offered an insight for the policy maker to deal with the coronavirus diseasepandemic.

B. Prabha, et.al (2021) constructed a novel ML (machine learning) method known as HDDP (Hybrid Disease Detection Principle) which was developed using two traditional ML model namely CNN (Convolutional Neural Network) and the AdaBoost Classifier [22]. The ML factors were considered for the CT image of lung that was pre-processed, its attributes were extracted and the classification was done. The constructed method had potential for handling the CT images on the basis of these effective image processing strategies. Hence, the COVID-19 disease was predicted and a report was sent to the respective user with properaccuracy ratio. This approach focused on predicting the disease effectively and assisted the doctors and patients successfully.

Wencheng Zou, et.al (2020) devised the ML (machine learning) techniques namely LR (linear regression) and NN (neuron networks) to predict the number of positive cases of Coronavirus Disease 2019 [23]. The LR algorithm was deployed for gathering the data of a state so that the corona disease was predicted. The number of infections was predicted using the data taken from Georgia and Massachusetts. The dataset was split into3 consecutive time periods and diverse models were trained to be suitable for each corresponding data for comparing MSE (mean square error) values. Moreover, the general viewpoint despite of three time periods depicted the superiority of single variable linear regression (LR) over accurately in comparison with the traditional technique with regard toaccuracy.

References

[1] Yifan Yang, Wenwu Yu, Duxin Chen, "Prediction of COVID-19 spread via LSTM and the deterministic SEIR model", 2020, 39th Chinese Control Conference(CCC)

[2] Zainab Abbas AbdulhusseinAlwaeli, Abdullahi Abdu Ibrahim, "Predicting Covid-19 Trajectory Using Machine Learning", 2020, 4th International Symposium on Multidisciplinary Studies and Innovative Technologies(ISMSIT)

[3] Suraj Bodapati, Harika Bandarupally, M Trupthi, "COVID-19 Time Series Forecasting of Daily Cases, Deaths Caused and Recovered Cases using Long Short Term Memory Networks", 2020, IEEE 5th International Conference on Computing Communication and Automation(ICCCA)

[4] Hanqing Chao, Xi Fang, Pingkun Yan, "Integrative analysis for COVID-19 patient outcome prediction", 2020, Medical ImageAnalysis

[5] Danish Rafiq, Suhail Ahmad Suhail, Mohammad Abid Bazaz, "Evaluation and prediction of COVID-19 in India: A case study of worst hit states", 2020, Chaos, Solitons &Fractals

[6] Farah Shahid, AneelaZameer, Muhammad Muneeb, "Predictions for COVID-19 with deep learning models of LSTM, GRU and Bi-LSTM", 2020, Chaos, Solitons &Fractals

[7] Anwar Jarndal, Saddam Husain, Omar Zaatar, Talal Al Gumaei, Amar Hamadeh, "GPR and ANN based Prediction Models for COVID-19 Death Cases", 2020, International Conference on Communications, Computing, Cybersecurity, and Informatics(CCCI)

[8] Pratima Kumari, Durga Toshniwal, "Real-time estimation of COVID-19 cases using machine learning and mathematical models - The case of India", 2020, IEEE 15th International Conference on Industrial and Information Systems(ICIIS)

[9] Yasin Khan, Pritam Khan, Sudhir Kumar, Jawar Singh, Rajesh M. Hegde, "Detection and Spread Prediction of COVID-19 from Chest X-ray Images using Convolutional Neural Network-Gaussian Mixture Model", 2020, IEEE 17th India Council International Conference (INDICON)

[10] Choujun Zhan, Yufan Zheng, Haijun Zhang, Quansi Wen, "Random-Forest-Bagging Broad Learning System with Applications for COVID-19 Pandemic", 2021, IEEE Internet of Things Journal

[11] Yifan Yang, Wenwu Yu, Duxin Chen, "Prediction of COVID-19 spread via LSTM and the deterministic SEIR model", 2020, 39th Chinese Control Conference(CCC)

[12] Leonardo Sestrem de Oliveira, Sarah Beatriz Gruetzmacher, João Paulo Teixeira, "COVID-19 Time Series Prediction", 2021, Procedia ComputerScience

[13] Durga Prasad Kavadi, Rizwan Patan, Amir H. Gandomi, "Partial derivative Nonlinear Global Pandemic Machine Learning prediction of COVID 19", 2020, Chaos, Solitons & Fractals

[14] Ankan Ghosh Dastider, Farhan Sadik, Shaikh Anowarul Fattah, "An integrated autoencoder-based hybrid CNN-LSTM model for COVID-19 severity prediction from lung ultrasound", 2021, Computers in Biology andMedicine

[15] Abhishek Dixit, Ashish Mani, Rohit Bansal, "CoV2-Detect-Net: Design of COVID-19 predictionmodelbasedonhybridDE-PSOwithSVMusingChestX-rayimages",2021, Information Sciences

[16] Elena Casiraghi, Dario Malchiodi, Gabriella Trucco, Marco Frasca, Luca Cappelletti, Tommaso Fontana,"Explainable Machine Learning for Early Assessment of COVID-19 Risk Prediction in Emergency Departments", 2020, IEEEAccess

[17] Anwar Jarndal, Saddam Husain, Omar Zaatar, Talal Al Gumaei, Amar Hamadeh, "GPR and ANN based Prediction Models for COVID-19 Death Cases", 2020, International Conference on Communications, Computing, Cybersecurity, and Informatics(CCCI)

[18] Huan Zhao, Yichi Li, Shenglan Chu, Suling Zhao, Chenglin Liu, "A COVID-19 Prediction Optimization Algorithm Based on Real-time Neural Network Training—Taking Italy as an Example", 2021, IEEE Asia-Pacific Conference on Image Processing, Electronics and Computers(IPEC)

[19] Nanning Zheng, Shaoyi Du, Jianji Wang, He Zhang, Wenting Cui, Zijian Kang, Tao Yang, Bin Lou, Yuting Chi, Hong Long, Mei Ma, Qi Yuan, Shupei Zhang, Dong Zhang, FengYe,JingminXin,"PredictingCOVID-19inChinaUsingHybridAIModel",2020,IEEE Transactions on Cybernetics

[20] Safa Bahri, MoetezKdayem, NesrineZoghlami, "Deep Learning for COVID-19 prediction", 2020 4th International Conference on Advanced Systems and Emergent Technologies(IC_ASET)

[21] Teddy Mantoro, RahmadyaTriasHandayanto, Media Anugerah Ayu, Jelita Asian, "Prediction of COVID-19 Spreading Using Support Vector Regression and Susceptible Infectious Recovered Model", 2020, 6th International Conference on Computing Engineering and Design(ICCED)

[22] B. Prabha, Sandeep Kaur, HarikumarPallathadka, "Intelligent predictions of Covid disease-basedonlungCTimagesusingmachinelearningstrategy",2021,MaterialsToday:
Proceedings

[23] Wencheng Zou, "The COVID-19 Pandemic Prediction in the US Based on Machine Learning", 2020, International Conference on Public Health and Data Science(ICPHDS)