MACHINE LEARNING AND DEEP LEARNING TECHNIQUES IN ANALYZING AND PREDICTING COVID-19: A SURVEY

P.Punitha
Assistant Professor
Vignana Bharathi Institute of Technology
punithaarabati@gmail.com

B. Kranthi Kiran
Associate Professor
JNTUH, Jagitial
kranthikiran9@gmail.com

Abstract:

The coronavirus disease (COVID-19), popularly called SARS-CoV-2 (severe acute respiratory syndrome) was first reported in December, 2019 in Wuhan, China. As it has spread globally, millions of people have been seriously affected. This coronavirus has mutated into many variants such as Aplha, Beta, Gamma, Delta, and Omicron over a period of two years. Having made an assessment of COVID-19 scenario, the World Health Organization has declared it as a pandemic. Many researchers use machine learning and deep learning techniques to diagnose COVID-19 disease and other associated disorders. Many reputed medical research and development organizations responded actively to this extraordinary challenge and developed vaccines that protect people from SARS-CoV-2. The government of India started undertaking the uphill task of the administration of coronavirus vaccination on 16 January 2021. The first, second, and booster doses of governmentally approved vaccination have been given to approximately 1.9 billion people of India. It is to be noted that at least one shot of vaccination has been given to 93% of the population above the age of 12 years. And, 83% of the population above the age of 12 years has been fully vaccinated. This study looks at how Machine Learning and Deep Learning philosophies might be utilized to predict various diseases. It mainly focuses on the literature available in the prediction and risk analysis of COVID-19.

***Keywords**: SARS-CoV-2, COVID-19, Machine Learning Techniques, Deep Learning techniques, vaccine*

1.Introduction

Coronavirus disease (2019) is a highly infectious disease caused by the SARS-CoV-2. The first coronavirus case was registered in December 2019 in Wuhan, China. This contagious disease has spread like wild fire across the globe and turned out to be COVID-19 pandemic. The coronavirus symptoms are cough, fever, fatigue, headache, difficulty in breathing, and loss of taste and smell. These symptoms may begin one to fourteen days after getting exposed to the coronavirus. However, one-third of infected people do not develop any conspicuous symptoms. Of those who develop these symptoms are identified as patients and 81% of patients have developed mild to moderate symptoms like mild pneumonia, whereas 14% of patients have developed severe symptoms and 5% have developed critical symptoms such as shock, respiratory failure, etc. It has been ascertained that the older people infected with coronavirus have been found to be at a higher risk of developing serious symptoms. Some people have continued to experience severe post-COVID-19 complications for months after their recovery. Serious research is being carried out to probe into the long-standing effects of the infectious disease.

COVID-19 is a serious problem faced all the people throughout the world. Its impact is more on elders and on people with chronic diseases. Vaccination saved many people from getting infected. There are many side effects i.e., post COVID complications in many people. As of now WHO not declared COVID-19 as endemic.

Fig 1: COVID-19 vaccination process in India (Courtesy: Asia Briefing Ltd.)

The rest of this paper is coordinated as follows: The second section depicts a recent study of ML and DL techniques for COVID-19. Section 3 gives a review about various DL approaches used for COVID-19 prediction and analysis. Segment 4 makes the inference and future work.

2.Recent study of Machine Learning , Deep learning techniques for COVID-19

ML and DL models are used in health care to predict diseases and help decision-makers to handle diseases. The use of Machine learning techniques in health care research has implemented various tasks, such as predicting survival rate of heart diseases, cancer, kidney diseases, etc. Many ML and DL researchers are working on COVID-19 prediction and analysis.

Some research papers focused on predicting disease based on symptoms of individual health condition and some papers focused on prediction of number COVID-19 cases and deaths based on previous data.

Acheme, I.D. and Vincent, O.R.[8] proposed research work on data science techniques using ML classification algorithms to predict the survivability of COVID-19 patients. This study utilized the dataset of Nigeria COVID-19 cases.This study shown that patients survivability depends on health condition and age.Zaid Abdi Alkareem Alyaasseri et al.[14] presented an intensive study on machine learning and deep learning techniques used to diagnose and to identify COVID-19 outbreak. The work considered various highly reputed databases such as IEEE, Springer, MPDI and Elsevier.

Nooshin Ayoobi et al.[20] analyzed three deep learning and their respective bidirectional extensions to predict the new cases of COVID-19 and death rate using time series dataset Abdelhafid Zeroual et al.[22] made an extensive study on five deep learning models namely Recurrent Neural Network (RNN), Long short term memory (LSTM), Bidirectional LSTM (BiLSTM), Variational Auto Encoder (VAE) and Gated recurrent units (GRUs) .These models are used in forecasting number of confirmed cases and recovered cases with the forecasting horizon of 17 days. Johns Hopkins dataset of five countries Spain, Italy, France, China, the USA and Australia. The performance of each model is verified in terms of evaluation metrics such as MAE, RMSE, MAPE, RMSLE, and EV. The results have shown that VAE (Variational Auto Encoder) has performed better.

3. Deep Learning techniques used for COVID-19 prediction and analysis

Many scholars has presented ML and DL techniques for COVID-19 prediction and analysis. Deep learning techniques are widely used nowadays for COVID-19. This section focuses on different deep learning techniques.

The following figure shows the deep learning model.

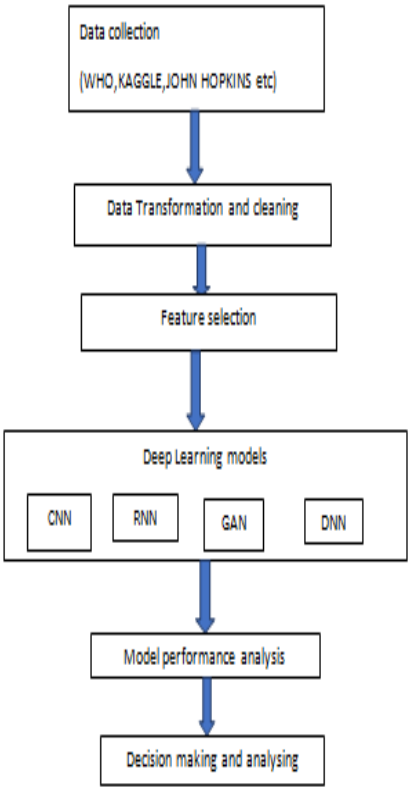

Fig 1.2: Deep learning model

Ketu and Mishra [21] introduced the CNN-LSTM model to predict the current status of COVID-19. This work utilized CNN for extracting the useful information and LSTM for extracting the dependencies of the short and long term. The main purpose of this survey paper was to investigate the COVID-19 across India. To analyze the outcomes, the data was collected from MoHFW, GoI and Arogya Setu Application from the period of 30th January 2020 to 10th June 2020. The experimental analysis was carried out for the various states of India of actual and predicted values and the statistical measures like MSE, RMSE, R^2 and MAPE performances were considered. Finally, it was proved that the hybrid CNN-LSTM obtained better results than CNN and LSTM respectively.

Zeroual et al. [22] provided a comparative analysis of DL models for forecasting the number of new COVID-19 , recovered cases. The methods like RNN, LSTM, Bi-LSTM, GRU and VAE (Variational autoencoder) were applied to forecast COVID-19 on the basis of small volume of data.These models captured related patterns and time variant characteristics of the prior and future tendancy of COVID-19. The analysis of confirmed and recovered cases were obtained from the countries like USA, Italy, Spain, Australia China, and France. The performance like MAE, RMSLE, MAPE and RMSE were compared and the DL model VAE obtained better performance.

Satpathy et al. [23] predicted the mortality rate and risks accompanied in COVID-19 infected people. The data was collected from WHO and pre-processed; then the forecasting approaches like average model, auto ARIMA, seasonal ARIMA, Facebook prophet, Holt's linear and Holt's winter were considered. The metrics like MAE, MSE, R^2 and RMSE were compared and proved that auto ARIMA was effectively and accurately forecasted the COVID-19. The RMSE and R^2 values obtained by auto ARIMA was 3.15 and 0.87 respectively.

Table 1: ML and DL methods in prediction and risk analysis

Authors Name	Methods	Process	Findings
Acheme, I.D. and Vincent, O.R.	Decision tree,Random forest model and Gradient boosting	This study utilized the dataset of Nigeria	This results shown that patients survivability depends on health condition

		COVID-19 cases	and age.
Abdelhafid Zeroual et al[22]	RNN, LSTM, BiLSTM, VAE, GRU	These models are used in forecasting number of confirmed cases and recovered cases with the forecasting horizon of 17 days. Johns Hopkins dataset of five countries Spain, Italy, France, China, the USA and Australia	The performance of each model is verified in terms of RMSE,MAE,MAPE,EV and RMSLE. The results shown that VAE (Variational Auto Encoder) performed better.
Rahil Mazlo umi et al.[2]	KNN, Decision tree (DT), logistic regression (LR), SVM,RF, SGD,bagging classifier (BC), adaptive boosting (AdaBoost)	Dataset of blood samples from 306 infected patients is taken from Tangji hospital of China. Patients' survival or death is analyzed based on clinical parameters using eight machine learning algorithms.	The hypothesis tests showed that RF,SVM, DT,KNN and AdaBoost produced more accurate results.
Mohammadreza Nemati et al.[9]	SVM,GB	This work analysed statistical models and ML	The Gradient Boosting survival model performed better.

		based approaches based patient time duration in hospital.	
Chandra et al.[16]	LSTM, Bi-LSTM, and encoder-decoder (ED)-LSTM	The data was captured from first and second waves and this work predicted that the chance of another wave was less.	The experimental analysis proved that ED-LSTM obtained better RMSE and MSE values for the cases for univariate and multivariate when compared to LSTM and Bi-LSTM models.
Marzouk et al.[17]	CNN, LSTM and MLP (multilayer perceptron)	DL techniques to predict COVID-19 outbreaks in the country of	Among the three Dl models, the LSTM model obtained better results and this LSTM with optimized
		Egypt. Initially, the data was standardized and the neural networks conducted time series analysis by dividing data in every subsequence.	parameter value was provided for forecasting the spread of epidemic.
Abbasimehr and Paki [19]	LSTM, multi-head attention and CNN with Bayesian optimization (BO)	The experimentation was conducted using the dataset obtained from m Johns Hopkins University's Coronavirus Resour	Among the three models, multi-head attention obtained better SMAPE value of 0.25 and 2.59 for the short and long term predictions.

		ce Center.	

4.Conclusion and future work

The COVID-19 pandemic has changed everyone's life. Many frontline workers and Governments are working towards the control of virus and to speed up vaccination process. In this paper a comprehensive study of machine learning and deep learning techniques were presented. All the studies include datasets of x-ray images, CT scan images, and time series data. This paper mainly included best feature selection, ML, DL techniques. In future work, the following points is to be considered, i) Analysing the number of people affected by COVID-19 age and chronic disease wise. ii) There is a strong need to develop ML, DL techniques for post COVID-19 complications. iii) There is a need to develop DL models to predict and analyse new COVID-19 variants and waves.

References

[1] Quiroz-Juárez, M.A., Torres-Gómez, A., Hoyo-Ulloa, I., León-Montiel, R.D.J. and U'Ren, A.B., 2021. Identification of high-risk COVID-19 patients using machine learning. Plos one, 16(9), p.e0257234.

[2] Mazloumi, R., Abazari, S.R., Nafarieh, F., Aghsami, A. and Jolai, F., 2022. Statistical analysis of blood characteristics of COVID-19 patients and their survival or death prediction using machine learning algorithms. Neural Computing and Applications, pp.1-15.

[3] Sinha, A. and Rathi, M., 2021. COVID-19 prediction using AI analytics for South Korea. Applied Intelligence, 51(12), pp.8579-8597.

[4] Gallo Marin, B., Aghagoli, G., Lavine, K., Yang, L., Siff, E.J., Chiang, S.S., Salazar-Mather, T.P., Dumenco, L., Savaria, M.C., Aung, S.N. and Flanigan, T., 2021. Predictors of COVID-19 severity: a literature review. Reviews in medical virology, 31(1), pp.1-10.

[5] Shwet, K. and Mishra, P.K., 2021. A hybrid deep learning model for COVID-19 prediction and current status of clinical trials worldwide. Computers, Materials, & Continua, pp.1896-1919.

[6] Arroio, A., 2020. The value of education in the context of COVID-19 pandemic. Problems of Education in the 21st Century, 78(3), pp.309-313.

[7] Pinter, G., Felde, I., Mosavi, A., Ghamisi, P. and Gloaguen, R., 2020. COVID-19 pandemic prediction for Hungary; a hybrid machine learning approach. Mathematics, 8(6), p.890.

[8] Acheme, I.D. and Vincent, O.R., 2021. Machine-learning models for predicting survivability in COVID-19 patients. In Data Science for COVID-19 (pp. 317-336). Academic Press.

[9] Nemati, M., Ansary, J. and Nemati, N., 2020. Machine-learning approaches in COVID-19 survival analysis and discharge-time likelihood prediction using clinical data. Patterns, 1(5), p.100074.

[10] Alimadadi, A., Aryal, S., Manandhar, I., Munroe, P.B., Joe, B. and Cheng, X., 2020. Artificial intelligence and machine learning to fight COVID-19. Physiological genomics, 52(4), pp.200-202.

[11] Sardar, R., Sharma, A. and Gupta, D., 2021. Machine learning assisted prediction of prognostic biomarkers associated with COVID-19, using clinical and proteomics data. Frontiers in genetics, 12.

[12] Chahar, S. and Roy, P.K., 2021. COVID-19: A Comprehensive Review of Learning Models. Archives of Computational Methods in Engineering, pp.1-26.

[13] Rasheed, J., Jamil, A., Hameed, A.A., Al-Turjman, F. and Rasheed, A., 2021. COVID-19 in the age of artificial intelligence: a comprehensive review. Interdisciplinary Sciences: Computational Life Sciences, 13(2), pp.153-175.

[14] Alyasseri, Z.A.A., Al-Betar, M.A., Doush, I.A., Awadallah, M.A., Abasi, A.K., Makhadmeh, S.N., Alomari, O.A., Abdulkareem, K.H., Adam, A., Damasevicius, R. and Mohammed, M.A., 2022. Review on COVID-19 diagnosis models based on machine learning and deep learning approaches. Expert systems, 39(3), p.e12759.

[15] Elghamrawy, S.M., Hassanien, A.E. and Vasilakos, A.V., 2021. Genetic-based adaptive momentum estimation for predicting mortality risk factors for COVID-19 patients using deep learning. International Journal of Imaging Systems and Technology.

[16] Chandra, R., Jain, A. and Singh Chauhan, D., 2022. Deep learning via LSTM models for COVID-19 infection forecasting in India. PloS one, 17(1), p.e0262708.

[17] Marzouk, M., Elshaboury, N., Abdel-Latif, A. and Azab, S., 2021. Deep learning model for forecasting COVID-19 outbreak in Egypt. Process Safety and Environmental Protection, 153, pp.363-375.

[18] Alakus, T.B. and Turkoglu, I., 2020. Comparison of deep learning approaches to predict COVID-19 infection. Chaos, Solitons & Fractals, 140, p.110120.

[19] Abbasimehr, H. and Paki, R., 2021. Prediction of COVID-19 confirmed cases combining deep learning methods and Bayesian optimization. Chaos, Solitons & Fractals, 142, p.110511.

[20] Ayoobi, N., Sharifrazi, D., Alizadehsani, R., Shoeibi, A., Gorriz, J.M., Moosaei, H., Khosravi, A., Nahavandi, S., Chofreh, A.G., Goni, F.A. and Klemeš, J.J., 2021. Time series forecasting of new cases and new deaths rate for COVID-19 using deep learning methods. Results in Physics, 27, p.104495.

[21] Ketu, S. and Mishra, P.K., 2022. India perspective: CNN-LSTM hybrid deep learning model-based COVID-19 prediction and current status of medical resource availability. Soft Computing, 26(2), pp.645-664.

[22] Zeroual, A., Harrou, F., Dairi, A. and Sun, Y., 2020. Deep learning methods for forecasting COVID-19 time-Series data: A Comparative study. Chaos, Solitons & Fractals, 140, p.110121.

[23] Satpathy, S., Mangla, M., Sharma, N., Deshmukh, H. and Mohanty, S., 2021. Predicting mortality rate and associated risks in COVID-19 patients. Spatial Information Research, 29(4), pp.455-464.

[24] P.Punitha, B.Kranthi Kiran, DATA MINING IN HEALTH CARE: A SURVEY FOCUSED ON COVID-19, Aut Aut Research Journal, ISSN NO: 0005-0601, Volume XII, Issue XII, December/2021.

[25]. Majeed A, Hwang SO. Data-Driven Analytics Leveraging Artificial Intelligence in the Era of COVID-19: An Insightful Review of Recent Developments. Symmetry. 2022; 14(1):16. https://doi.org/10.3390/sym14010016

[26]. Shorten C, Khoshgoftaar TM, Furht B. Deep Learning applications for COVID-19. J Big Data. 2021;8(1):18. doi: 10.1186/s40537-020-00392-9. Epub 2021 Jan 11. PMID: 33457181; PMCID: PMC7797891.

[27]. Alafif T, Tehame AM, Bajaba S, Barnawi A, Zia S. Machine and Deep Learning towards COVID-19 Diagnosis and Treatment: Survey, Challenges, and Future Directions. Int J Environ Res Public Health. 2021 Jan 27;18(3):1117. doi: 10.3390/ijerph18031117. PMID: 33513984; PMCID: PMC7908539.

[28]. Nayak, J, Naik, B, Dinesh, P. et al. Significance of deep learning for COVID-19: state-of-the-art review. Res. Biomed. Eng. 38, 243–266 (2022). https://doi.org/10.1007/s42600-021-00135-6.