

# Prediction analysis on the pre and post COVID outbreak assessment using machine learning and deep learning

Ishaan Walecha, Divya Jain

Department of Computer Science, NorthCap University, Haryana, India

## Article Info

### Article history:

Received Dec 24, 2021

Revised May 30, 2022

Accepted Jun 11, 2022

### Keywords:

Corona virus

Deep learning

Machine learning

Post COVID

Pre COVID

Predictive analysis

## ABSTRACT

In this time of a global urgency where people are losing lives each day in a large number, people are trying to develop ways/technology to solve the challenges of COVID-19. Machine learning (ML) and artificial intelligence (AI) tools have been employed previously as well to the times of pandemic where they have proven their worth by providing reliable results in varied fields this is why ML tools are being used extensively to fight this pandemic as well. This review describes the applications of ML in the post and pre COVID-19 conditions for contact tracing, vaccine development, prediction and diagnosis, risk management, and outbreak predictions to help the healthcare system to work efficiently. This review discusses the ongoing research on the pandemic virus where various ML models have been employed to a certain data set to produce outputs that can be used for risk or outbreak prediction of virus in the population, vaccine development, and contact tracing. Thus, the significance and the contribution of ML against COVID-19 are self-explanatory but what should not be compromised is the quality and accuracy based on which solutions/methods/policies adopted or produced from this analysis which will be implied in the real world to real people.

This is an open access article under the [CC BY-SA](https://creativecommons.org/licenses/by-sa/4.0/) license.



## Corresponding Author:

Ishaan Walecha

Department of Computer Science, NorthCap University

Sector 23A, Gurgaon, Haryana, India

Email: walechai019@gmail.com

## 1. INTRODUCTION

In the last few years, we have been seeing the rise of new viruses that can challenge the human quality of life by spreading globally. The recent of this being the novel coronavirus. In 2020 on March 11, World Health Organization (WHO) declared a global pandemic emergency situation with 118,326 cases, 4,292 deaths being caused by SARS-CoV2 infection which led to the public health crisis of a scale unseen before [1]. This growing pandemic began in Wuhan, China in 2019 which has not only pushed the world to adapt to their “new normal” also it has created immense pressure on the healthcare system and economy all around the world [2]–[6]. On June 22, 2021, WHO had released the data of 178,503,429 cases and 3,872,457 deaths being caused by COVID-19 worldwide [7].

Since the outbreak of this new SARS CoV2, scientists and the medical industry around the world everywhere have started to face this infectious disease. Infection of virus affects both our upper and lower respiratory systems that lead to chronic obstructive pulmonary disease COPD and lung disorders. Major symptoms arising from the COVID-19 infection include body ache, losing the ability to taste and smell, nausea, fever, congestion, and sore throat. Various other important factors including ethnic, cultural, and demographic behaviors, social distance, and quarantine measures, have a significant impact in reducing the risk of infection with the illness.

In an age of digitalization, through the use of data acquisition, machine learning (ML), and computing infrastructure, artificial intelligence (AI) applications are spreading into the fields that were once considered to be the sole province of human expertise. ML is a collection of technologies that analyzes data by detecting patterns in it. ML technologies, in contrast to traditional techniques of pattern recognition, rely on AI to identify patterns, become self-improving, and become even more efficient when more data is accessible. Hence, ML serves to be an important, adaptable, innovative means through which this pandemic can be brought under control [8].

In this review, we are discussing the contribution performed by recent technologies to tackle this pandemic at every stage. We witness the role of these modern technologies in the prediction of the outbreak, risk, and mortality. Globally extreme funds have been poured in by various countries for the development of the vaccine, treatment of the disease where we observe ML plays a big role in identifying the most probable targets [9]–[11]. AI is the future, as they say, satisfying this ML provides us the means to prepare a healthcare management system that is well equipped to fight the battle in the future. Figure 1 depicts the pre-COVID-19 and post-COVID-19 analyses.

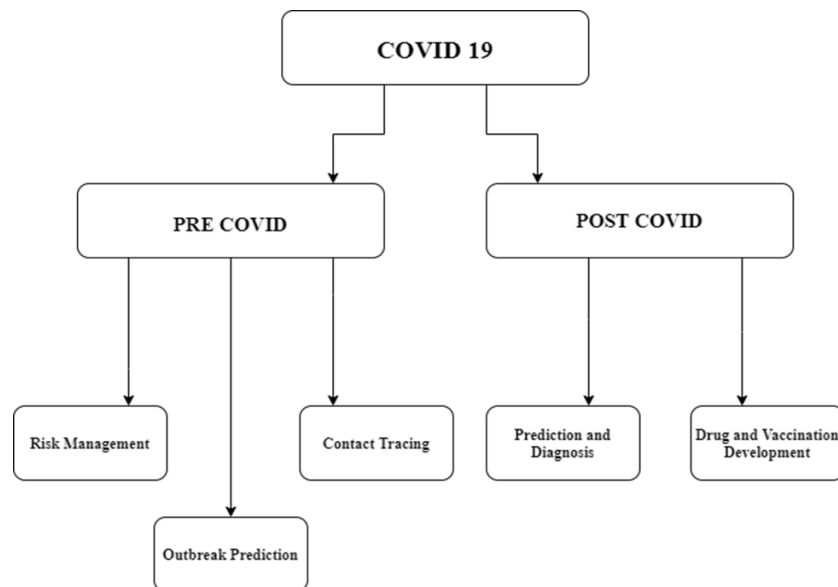


Figure 1. Pre COVID-19 and post COVID-19 assessment

## 2. PRE-COVID ASSESSMENT

### 2.1. Risk management

It is very important to classify the patients based on how serious is the case to give them proper care which is need by them to recover from the coronavirus disease. Various research work has been performed implying ML tools/algorithms [12] to forecast APACHE-II risk-prediction score namely neural network, Random Forest [13], and classification and regression decision tree (CRT). A database of 6,995 patients was extracted from hospital records of which 162 were found to be positive. They discovered that 25 (15.4%) of the 162 patients had critical COVID-19. In predicting coronavirus, ML models outpaced every other criterion, including the ratings of APACHE II, with 88.0% sensitivity, 92.7% precision, and 92.0% accuracy [14].

Immediate ICU requirement is a must for critical COVID-19 patients to save their lives. Only 20-30% of COVID-19 patients require hospitalization and from them, only 5-10% require critical care in the ICU so to identify those 5-10% of the patient a tool has been created by the scientists with the help of ML which predict the risk of the patient and if they need ICU or not. To train the random forest model [15], data such as vitals, analysis of the nurse, data of the lab, and electrocardiograms were utilized. Cheng *et al*, [16] classified COVID-19 virus sequences with 100% precision, and discovered an important relationship in a very short period among 5,000 genomes of the virus, with the help of only raw DNA sequence data and no advanced medical information, training, or gene or genome annotations.

In another study similar work was performed to detect COVID-19 patients which are under risk where data were examined from a hospital in London of 879 positive SARS-CoV2 patients who were admitted from January to May 2020. From electronic health records, the students collected anonymous population results, physiological therapeutics, and laboratory samples. They used the data available in the

initial phase of the patient so that they can evaluate if the patient needs high-priority invasive care that could save the patient life. Multivariate logistic regression [17] random forest and extreme gradient boosted trees were employed [18] on the data of the patients. They classified the type of care required by the patient in three clinical methods, patients who required severe care, patients who need the ventilators, and in-hospital mortality, and the results generated depicted that patients that needed severe care were 15%, 7% required ventilators, and 31% were mortality cases of the hospital which was highly accurate [19]. Similar studies were performed that gave us clear evidence of utilization of ML algorithm for risk management in the healthcare sector for COVID-19 patients.

## 2.2. Outbreak prediction

Even before the corona pandemic ML has been widely used to predict the outbreak of other diseases as well like Ebola virus disease (EVD). This virus began to spread in early 2014 to west African countries and had infected a large population causing deaths one such study was conducted in which they used the IDEA model (incidence decay with exponential adjustment) to investigate the dynamics of epidemics. The initial development patterns of the Ebola virus epidemic that occurred in 2014 in the region of West African were close to the previous outbreak situation caused by Ebola, according to these findings. The loss of control apparent, especially in Liberia, is concerning, with disease processes increasing in an essentially unchecked exponential fashion. Similar types of research were done in [20]–[22] the past for EVD and SARS [23]. These types of models can also be used for corona.

The dataset of 10 densely populated countries, in particular India, Pakistan, Germany, Ethiopia, China, Democratic Republic of Congo, Philippines, Nigeria, Bangladesh, and Indonesia was sourced and 9 different models were used for the prediction of the outbreak. Auto-regressive moving average (ARMA), auto-regressive integrated moving average (ARIMA), linear regression (LR), linear regressor polynomial (LRP), Bayesian ridge polynomial regressor (BRR), support vector regressor (SVR), random forest regressor (RFR) XG boost regressor (XGB) and holt-winters (HW) exponential smoothing. Results produced from the study proved implied that a ML algorithm can forecast the highs or lows in the cases for each nation while the accuracy produced may vary [24].

Table 1. Analysis of different models based on classification accuracy

Country	Best model	Accuracy (%)
Bangladesh	LRP	86.45
India	ARMA	99.26
China	XGB	82
Pakistan	BRR	87.91
Germany	ARIMA	85.39
Nigeria	ARMA	98.06
Ethiopia	ARMA	99.93
Democratic Republic of Congo	LRP	91.96
Philippines	SVR	50.54
Indonesia	ARIMA	97.72

Due to the absence of the data standard models are less accurate for a long-time estimate. To overcome this Ardabili *et al.* [25] compared ML and soft computing models for forecasting this epidemic. Various types of ML models were implemented on the data sourced from China, the USA, Iran, Germany, and Italy [26] which was extracted from the world meters website for total cases over 30 days and this study resulted that multi-layered perceptron (MLP) and adaptive network-based fuzzy inference system (ANFIS) were the best models which gave the most accurate results [25].

## 2.3. Contact tracing

The most peculiar step to prevent the transmission and spread of the COVID-19 virus is tracing the contacts [27] COVID-19 as we know is a disease that spread from droplets of saliva, or discharges from the nose through contact transmission as supported by WHO reports [28]. Various applications were developed in smartphones which are easily accessible to people to contacts digitally. these apps used different means of technology such as mobile monitoring data, global positing system, contact details, Bluetooth, network-based API, and card purchase data, all such efforts created a contract tracing mechanism that was completely digitalized that became extremely useful over other non-digital methods as this could operate itself in the current scenario and that too quicker than the non-digital methods. All these tools are programmed to tell if someone is vulnerable to the virus with the help of ML and AI, by using the data of an individual and their recent touch chain [29].

Table 2. Countries and their applications

S no.	Country	Contact tracing app
1	India	Aarogya Setu
2	UAE	Trace COVID
3	Australia	COVID Safe
4	Italy	Immuni
5	Singapore	Trace Together
6	UK	NHS COVID-19 App
7	China	Conjunction with Alipay

AI-powered sensors are being used by several governments and healthcare networks throughout the world to help with enhanced triage. Baidu, a Chinese technology company, has created a no-contact infrared sensor gadget that can detect persons with fever even in crowds in this manner. Similarly, in conjunction with Care.ai, Tampa General Hospital in Florida has placed an AI system at its gates to identify patients who may be suffering from COVID-19 symptoms. By installing cameras at gates, the technology escorts a face thermal scan and identifies other indicators, such as perspiration and discoloration, to thwart travelers with flu [30]. COVID-19 voice detector is an AI-based app that detects infection in human speech. Patients may interact with medical professionals in a simulated world thanks to augmented reality and virtual reality technologies [31].

### 3. POST-COVID ASSESSMENT

#### 3.1. Prediction and diagnosis

Various research projects have evidently depicted the use of ML tools for the prediction of coronavirus using the images produced by X-rays of the chest in the population, example – In a study where researchers had used 2 classifiers logistics regression (LR) and convolutional neural networks (CNN) on the dataset available online [32], [33] and integrated generative adversarial network (GAN) to have 500 Xray in total. In addition, a dimensionality reduction method focused on principal component analysis was also applied.

The outcome of the study shows that the suggested CNN and LR models obtained an accuracy of 97.6% and 95.2%, respectively, without feature extraction and with only 233ms of preparation. An accuracy of 100% was obtained by the CNN system using a dataset with 0.99 variances. This approach is extremely useful to identify the corona patients in an economically viable manner for the developing countries that can't afford testing kits to identify the disease at a mass scale. The key advantage of the deep learning (DL) system was to eliminate the painstaking and labor-intensive do-it-yourself (DIY) characteristics thereby enhancing the accuracy of classification based on a data-driven approach [34].

COVIDX-Net is another new, deep-learning platform used by radiologists to diagnose COVID-19 patients with their X-rays. It has been tested on fifty X-rays of the chest dataset with 50% positive and 50% negative COVID-19 cases which were provided by Dr. Cohen and Dr. Rosebrock. The COVIDX-Net tool is comprised of seven models varying in their deep coevolutionary network architectures VGG19 [35], DenseNet121 [36], InceptionV3 [37], ResNetV2 [38], [39], Inception-ResNet-V2 [40], Xception [41], and MobileNetV2 [42]. As a result, they suggested that the VGG19 and DenseNet201 models be used to detect patients using X-rays of patients [43].

Table 3. Analysis of different models based on classification accuracy

S no.	Model	Accuracy (%)
1	VGG19	90
2	DenseNet121	90
3	InceptionV3	50
4	ResNetV2	70
5	Inception-ResNet-V2	80
6	Xception	80
7	MobileNetV2	60

Another study used deep learning techniques to develop a model that could distinguish COVID 19 patients from IAVP and stable patients based on pulmonary computed tomography images. Many CNN [44], [45] models were used in this analysis to characterize computed tomography datasets and measure the COVID-19 risk. Total 618 computed tomography images were gathered together containing 219 images from 110 COVID-19 patients, 224 images from the viral pneumonia Influenza-A patients, and 175 images from healthy individuals while 528 computed tomography images were employed for confirmation and training purposes that includes 189 samples from COVID-19 infected/diseased people, 194 images from IAVP

patients and 145 images of healthy individuals. Test range consisted of the remaining 90 CT sets which included 30 corona patients, 30 IAVP cases, and 30 stable patients. Various pathogens were differentiated morphologically by utilizing an image classification model. In this analysis, two different classification models were used and compared in this study. The first network was focused on linking the mechanism of the location to attention in the full-connection layer, and the second model was based on relative traditional residual network (Res-Net) [38]. Models having the location and attention mechanism were compared with ones that lacked this linkage which revealed that the location-attention mechanism provides a better alternative to differentiate coronavirus patients from others. An average precision of 86.7% was obtained from CT cases, therefore, proving that ML can be a valuable companion/screening instrument for clinical frontline physicians dealing with the pandemic [46].

### 3.2. Drug and vaccine development

The development of an efficient curative plan is an urgent need to treat rapidly growing pandemics. Since we are still struggling to find the cure for the disease it is very essential to create an effective strategy to make medically accepted medicines to cure SARS-CoV-2 which can be made by using ML, either to make a new medicine or use it for medical trials of the present drug which is used to cure of SARS-CoV-2 [47]. Various research projects have been taken upon to achieve this goal using ML. Atazanavir is one such drug that is being widely used on corona patients came up in one such recent study conducted by Beck *et al.* [48] where they worked on molecule transformer-drug target interaction (MT-DTI) model that evaluated the interactions between our drug of interest that is commercially available with the potential targets resulting in values that determined their binding affinities to each other. These values were generated based on the sequence of their amino acids (FASTA) and target protein's chemical sequences (SMILES) to list the drugs that are not only food drug agency (FDA) approved but also have antiviral properties to destroy the SARS-COV2 INFECTION by inhibiting its functioning. Atazanavir, which is the antiretroviral drug for the treatment of HIV, proved to be the most potent drug at the end of this study by generating a Kd value of 94.94 nM against COVID-19. After that remdesivir has Kd of 113.13 nM, followed by efavirenz Kd of 199.17 nM, after that ritonavir Kd of 204.05 nM and dolutegravir Kd of 336.91 nM were some of the chemical substances which can be used against SARS-CoV2. Some of the other medicament, such as Kaletra (lopinavir/ritonavir), were also discovered to be effective [48].

For preventive preparation, prevention, and treatment, large virus outbursts necessitate the clarification of their place in taxonomy and the sequence of their genetic material. Therefore, researchers discovered the coronavirus genome sequence and combined this along with ML to categorize entire COVID-19 virus genomes. 5,000 unique viral genomic sequences were examined for which data was extracted from various sources like NCBI, Virus-host DB, GISAID. For genome studies, they used machine learning with digital signal processing (MLDSP) [49] and MLDSP GUI [50], which included an approach based on a decision tree along with this Spearman's rank correlation coefficient which was used for the confirmation of results. This method achieved 100 % accuracy for classifying the COVID-19 sequences and found the best relationship from 5,000 viral genomes using DNA sequence data without biological knowledge. This study suggests that it can be used for critical periods during the outbreak of the virus [51].

Another application of ML would be the prediction of practicable synthetic antibodies that could neutralize the virus which can be evidently concluded from the work of Magar, R., Yadav, P., & Farimani, A. B who extracted their dataset from the viruses such as Hepatitis, HIV, Dengue, SARS, Ebola, of antibody-antigen sequences and they have merged it with patient clinical/biochemical IC50 data. They utilized the Virus Net dataset, which carries 1933 samples of 15 different viruses taken from the Compile, Analyze, and Tally NAb panels (CATNAP) database from the Los Alamos National Laboratory (LANL) [52], [53]. Various means that were employed to predict whether the antibody will kill the virus were XGBoost, random forest, multilayer perceptron, support vector machine (SVM), and logistic regression. Considering the sequence of antibodies that neutralized virus 2589 mutant strains of anticorps sequences were developed which could be served as a possible candidate for antibodies. It was predicted that eighteen antibodies were beneficial for the neutralization of the virus. To assess the permanence and viability of the suggested antibody structures, molecular dynamics (MD) simulations of every structure were performed, which helped us screen 9 possible antibodies that could neutralize the virus. The performance in terms of their stability was observed as XGBoost with 90.57% then RF 89.18% followed by LR 81.17% then MLP 78.23% and at last SVM 75.49 further IN VITRO experiments can be performed to validate and assess the efficiency of these predicted antibodies for their action against SARS-CoV-2 virus [54].

The treatment for the coronavirus in a human cell has been explored by a profound learning model and ML methodology. This study was conducted to identify human cells inevitably depending on patients and treatment doses. The RxRx19a Dataset research contained over 300,000 reported trials in the cell of

humans having SARS CoV2 virus and more than 1,660 authorized food and drug administration (FDA) medicines [55].

It was a three-step process that began with the conversion of 1024 cell features to distinct numerical to form a digital image, then moved on to the training phase, which included ML algorithms such as SVM decision trees (DT), and ensemble methods for numerical features and deep CNN that were employed for converting the features of image, and finally, model's accuracy was tested and evaluated for concentration level prediction and treatment classification was done in the end phase using traditional ML algorithms that produced results that the Area under the curve for oseltamivir-carboxylate therapy was 73% when using DT, 84% when using SVM [56], and 86% when using ensemble, but Deep learning significantly enhanced the research accuracy and precision measurement, as evidenced by their scores of 98.05% and 96.52%, respectively. The deep convolutional neural network (DCNN) surpassed these standard approaches, scoring 98.2% in forecasting care concentration levels, compared to 96.4 and 97.3 for DT and SVM, respectively. However, they found that the ensemble approach surpassed the DCNN in terms of research precision, scoring 98.5% [57].

#### 4. CONCLUSION AND FUTURE SCOPE

After the occurrence of the COVID-19 pandemic, the whole system has been put to test. Healthcare infrastructure and researchers have been burdened extremely especially in developing areas to sustain their systems and work effectively to save lives. People all around the globe, ubiquitously are fighting against the pandemic in their ways. medical industries and our scientists are working vigorously day and night to eradicate the issue by breaking the chain of transmission, attend patients in large numbers, develop a vaccine, and test kits, so that the human race could survive, this is where ML and AI come into the picture for our rescue employment of ML tools has not only significantly contributed to the fulfillment of the above tasks but also provides us the facility to analyze the data from multiple angles which gives us a perception, clear solutions to this real-world problem. This paper discusses the various applications of modern technologies algorithms/models that have been implied on the data/samples collected from various sources to develop the vaccine, methods for outbreak prediction, contact tracing, risk management, and diagnosis.

Since we are in a pandemic, that affects the mass population hence data is being generated in large amounts but this also is not a hindrance to understanding the disease process while ML tools are present as they still allow data analysis and provide rapid identification of pattern using AI which would not have been possible while implying traditional tools or methods of mathematics and statistics as they would definitely prove to be time-consuming. But handling a large amount of data should not affect the quality and accuracy of results to help the system work efficiently. Immature data should not be fed in the database that creates ambiguous results or creates noise in the data. So from the studies conducted so far, This can be concluded that ML tools provide us relief in this pandemic by providing us the means of treatment and also help us prevent the spread of disease. In near future, ML has great potential to sustain and support our healthcare infrastructure for its efficient working and thereby reducing human intervention. In the future, the research can be extended by analyzing more deep learning models for COVID-19 outbreak prediction.

#### REFERENCES




- [1] C. Sohrabi *et al.*, "World Health Organization declares global emergency: A review of the 2019 novel coronavirus (COVID-19)," *International Journal of Surgery*, vol. 76, pp. 71–76, Apr. 2020, doi: 10.1016/j.ijsu.2020.02.034.
- [2] WHO, "Coronavirus disease (COVID-2019) press briefings," 2020. .
- [3] N. Zhu *et al.*, "A novel coronavirus from patients with pneumonia in China, 2019," *New England Journal of Medicine*, vol. 382, no. 8, pp. 727–733, Feb. 2020, doi: 10.1056/NEJMoa2001017.
- [4] N. Zhu *et al.*, "Early transmission dynamics in Wuhan, China, of novel coronavirus-infected pneumonia," *New England Journal of Medicine*, vol. 382, no. 8, pp. 727–733, 2020.
- [5] J. Cohen and D. Normile, "New SARS-like virus in China triggers alarm," *Science*, vol. 367, no. 6475, pp. 234–235, Jan. 2020, doi: 10.1126/science.367.6475.234.
- [6] T. Lupia, S. Scabini, S. Mormese Pinna, G. Di Perri, F. G. De Rosa, and S. Corcione, "2019 novel coronavirus (2019-nCoV) outbreak: A new challenge," *Journal of Global Antimicrobial Resistance*, vol. 21, pp. 22–27, Jun. 2020, doi: 10.1016/j.jgar.2020.02.021.
- [7] WHO, "Coronavirus disease (COVID-19)," 2020. .
- [8] A. Bansal, R. P. Padappayil, C. Garg, A. Singal, M. Gupta, and A. Klein, "Utility of artificial intelligence amidst the COVID 19 pandemic: a review," *Journal of Medical Systems*, vol. 44, no. 9, p. 156, Sep. 2020, doi: 10.1007/s10916-020-01617-3.
- [9] D. Jain and V. Singh, "Diagnosis of breast cancer and diabetes using hybrid feature selection method," in *2018 Fifth International Conference on Parallel, Distributed and Grid Computing (PDGC)*, Dec. 2018, pp. 64–69, doi: 10.1109/PDGC.2018.8745830.
- [10] D. Jain and V. Singh, "A novel hybrid approach for chronic disease classification," *International Journal of Healthcare Information Systems and Informatics*, vol. 15, no. 1, pp. 1–19, Jan. 2020, doi: 10.4018/IJHISI.2020010101.
- [11] D. Jain and V. Singh, "Feature selection and classification systems for chronic disease prediction: A review," *Egyptian Informatics Journal*, vol. 19, no. 3, pp. 179–189, Nov. 2018, doi: 10.1016/j.eij.2018.03.002.
- [12] M. T. Hagan, H. B. Demuth, and M. H. Beale, *Neural networks in a softcomputing framework*, vol. 2. London: Springer-Verlag, 2006.

- [13] L. Breiman, "Bagging predictors," *Machine Learning*, vol. 24, no. 2, pp. 123–140, 1996, doi: 10.1023/A:1018054314350.
- [14] D. Assaf *et al.*, "Utilization of machine-learning models to accurately predict the risk for critical COVID-19," *Internal and Emergency Medicine*, vol. 15, no. 8, pp. 1435–1443, Nov. 2020, doi: 10.1007/s11739-020-02475-0.
- [15] S. Basu, K. Kumbier, J. B. Brown, and B. Yu, "Iterative random forests to discover predictive and stable high-order interactions," *Proceedings of the National Academy of Sciences*, vol. 115, no. 8, pp. 1943–1948, Feb. 2018, doi: 10.1073/pnas.1711236115.
- [16] C. F.-Y. *et al.*, "Using machine learning to predict ICU transfer in hospitalized COVID-19 patients," *Journal of Clinical Medicine*, 2020.
- [17] F. Zhou *et al.*, "Clinical course and risk factors for mortality of adult inpatients with COVID-19 in Wuhan, China: a retrospective cohort study," *The Lancet*, vol. 395, no. 10229, pp. 1054–1062, Mar. 2020, doi: 10.1016/S0140-6736(20)30566-3.
- [18] H. Zou and T. Hastie, "Regularization and variable selection via the elastic net," *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, vol. 67, no. 2, pp. 301–320, Apr. 2005, doi: 10.1111/j.1467-9868.2005.00503.x.
- [19] H. FS *et al.*, "Early risk assessment for COVID-19 patients from emergency department data using machine learning," 2021.
- [20] D. N. Fisman, T. S. Hauck, A. R. Tuite, and A. L. Greer, "An IDEA for short term outbreak projection: nearcasting using the basic reproduction number," *PLoS ONE*, vol. 8, no. 12, p. e83622, Dec. 2013, doi: 10.1371/journal.pone.0083622.
- [21] K. Roosa *et al.*, "Real-time forecasts of the COVID-19 epidemic in China from February 5th to February 24th, 2020," *Infectious Disease Modelling*, vol. 5, pp. 256–263, 2020, doi: 10.1016/j.idm.2020.02.002.
- [22] J. D. Kelly *et al.*, "Real-time predictions of the 2018–2019 Ebola virus disease outbreak in the Democratic Republic of the Congo using Hawkes point process models," *Epidemics*, vol. 28, p. 100354, Sep. 2019, doi: 10.1016/j.epidem.2019.100354.
- [23] D. Fisman, E. Khoo, and A. Tuite, "Early epidemic dynamics of the West African 2014 Ebola outbreak: estimates derived with a simple two-parameter model," *PLoS Currents*, 2014, doi: 10.1371/currents.outbreaks.89c0d3783f36958d96ebbae97348d571.
- [24] A. Khakharia *et al.*, "Outbreak prediction of COVID-19 for dense and populated countries using machine learning," *Annals of Data Science*, vol. 8, no. 1, pp. 1–19, Mar. 2021, doi: 10.1007/s40745-020-00314-9.
- [25] S. F. Ardabili *et al.*, "COVID-19 outbreak prediction with machine learning," *Algorithms*, vol. 13, no. 10, p. 249, Oct. 2020, doi: 10.3390/a13100249.
- [26] Worldometer, "Reported cases and deaths by country, territory, or conveyance," 2020.
- [27] WHO, "Contact tracing in the context of COVID-19," 2020. .
- [28] WHO, "Health topic, coronavirus disease overview," 2020. .
- [29] S. Lalmuanawma, J. Hussain, and L. Chhakhuak, "Applications of machine learning and artificial intelligence for COVID-19 (SARS-CoV-2) pandemic: A review," *Chaos, Solitons & Fractals*, vol. 139, p. 110059, Oct. 2020, doi: 10.1016/j.chaos.2020.110059.
- [30] Vivek Kumar, "Healthcare Centers Are Turning To Ai To Combat Covid-19," *AnalyticsInsight*, 2020. .
- [31] S. Kushwaha *et al.*, "Significant applications of machine learning for COVID-19 pandemic," *Journal of Industrial Integration and Management*, vol. 05, no. 04, pp. 453–479, Dec. 2020, doi: 10.1142/S2424862220500268.
- [32] Y. Ren, L. Li, and Y. Jia, "New method to reduce COVID-19 transmission-the need for medical air disinfection is now," *Journal of Medical Systems*, vol. 44, no. 7, p. 119, Jul. 2020, doi: 10.1007/s10916-020-01585-8.
- [33] D. Fisher and D. Heymann, "Q&A: The novel coronavirus outbreak causing COVID-19," *BMC Medicine*, vol. 18, no. 1, p. 57, Dec. 2020, doi: 10.1186/s12916-020-01533-w.
- [34] J. Rasheed, A. A. Hameed, C. Djeddi, A. Jamil, and F. Al-Turjman, "A machine learning-based framework for diagnosis of COVID-19 from chest X-ray images," *Interdisciplinary Sciences: Computational Life Sciences*, vol. 13, no. 1, pp. 103–117, Mar. 2021, doi: 10.1007/s12539-020-00403-6.
- [35] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," *3rd International Conference on Learning Representations, ICLR 2015 - Conference Track Proceedings*, Sep. 2014.
- [36] G. Huang, Z. Liu, L. Van Der Maaten, and K. Q. Weinberger, "Densely connected convolutional networks," in *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Jul. 2017, vol. 2017-Janua, pp. 2261–2269, doi: 10.1109/CVPR.2017.243.
- [37] C. Szegedy, V. Vanhoucke, S. Ioffe, J. Shlens, and Z. Wojna, "Rethinking the inception architecture for computer vision," in *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Jun. 2016, vol. 2016-Decem, pp. 2818–2826, doi: 10.1109/CVPR.2016.308.
- [38] H. Kaiming, Z. Xiangyu, R. Shaoqing, and S. Jian, "Deep residual learning for image recognition," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2006, pp. 770–778.
- [39] E. C. Too, L. Yujian, S. Njuki, and L. Yingchun, "A comparative study of fine-tuning deep learning models for plant disease identification," *Computers and Electronics in Agriculture*, vol. 161, pp. 272–279, Jun. 2019, doi: 10.1016/j.compag.2018.03.032.
- [40] C. Szegedy, S. Ioffe, V. Vanhoucke, and A. A. Alemi, "Inception-v4, inception-ResNet and the impact of residual connections on learning," in *31st AAAI Conference on Artificial Intelligence, AAAI 2017*, 2017, pp. 4278–4284.
- [41] F. Chollet, "Xception: Deep learning with depthwise separable convolutions," in *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Jul. 2017, vol. 2017-Janua, pp. 1800–1807, doi: 10.1109/CVPR.2017.195.
- [42] M. Sandler, A. Howard, M. Zhu, A. Zhmoginov, and L.-C. Chen, "Mobilenetv2: Inverted residuals and linear bottlenecks," in *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*, Jun. 2018, pp. 4510–4520, doi: 10.1109/CVPR.2018.00474.
- [43] E. E.-D. Hemdan, M. A. Shouman, and M. E. Karar, "COVIDX-Net: A framework of deep learning classifiers to diagnose COVID-19 in X-Ray images," 2020.
- [44] M. H. Hesamian, W. Jia, X. He, and P. Kennedy, "Deep learning techniques for medical image segmentation: achievements and challenges," *Journal of Digital Imaging*, vol. 32, no. 4, pp. 582–596, Aug. 2019, doi: 10.1007/s10278-019-00227-x.
- [45] M. Akagi *et al.*, "Deep learning reconstruction improves image quality of abdominal ultra-high-resolution CT," *European Radiology*, vol. 29, no. 11, pp. 6163–6171, Nov. 2019, doi: 10.1007/s00330-019-06170-3.
- [46] X. Xu *et al.*, "A deep learning system to screen novel coronavirus disease 2019 pneumonia," *Engineering*, vol. 6, no. 10, pp. 1122–1129, Oct. 2020, doi: 10.1016/j.eng.2020.04.010.
- [47] Y. Ge; *et al.*, "A data-driven drug repositioning framework discovered a potential therapeutic agent targeting COVID-19," *bioRxiv*.
- [48] B. R. Beck, B. Shin, Y. Choi, S. Park, and K. Kang, "Predicting commercially available antiviral drugs that may act on the novel coronavirus (SARS-CoV-2) through a drug-target interaction deep learning model," *Computational and Structural Biotechnology Journal*, vol. 18, pp. 784–790, 2020, doi: 10.1016/j.csbj.2020.03.025.
- [49] G. S. Randhawa, K. A. Hill, and L. Kari, "ML-DSP: Machine learning with digital signal processing for ultrafast, accurate, and scalable genome classification at all taxonomic levels," *BMC Genomics*, vol. 20, no. 1, p. 267, Dec. 2019, doi: 10.1186/s12864-019-5571-y.
- [50] G. S. Randhawa, K. A. Hill, and L. Kari, "MLDSP-GUI: An alignment-free standalone tool with an interactive graphical user interface for DNA sequence comparison and analysis," *Bioinformatics*, vol. 36, no. 7, pp. 2258–2259, Apr. 2020, doi: 10.1093/bioinformatics/btz918.
- [51] G. S. Randhawa, M. P. M. Soltysiak, H. El Roz, C. P. E. de Souza, K. A. Hill, and L. Kari, "Machine learning using intrinsic genomic signatures for rapid classification of novel pathogens: COVID-19 case study," *PLoS ONE*, vol. 15, no. 4, p. e0232391, Apr. 2020, doi: 10.1371/journal.pone.0232391.




- [52] H. Yoon *et al.*, "CATNAP: A tool to compile, analyze and tally neutralizing antibody panels," *Nucleic Acids Research*, vol. 43, no. W1, pp. W213–W219, Jul. 2015, doi: 10.1093/nar/gkv404.
- [53] CATNAP, "HIV sequence database." 2022.
- [54] R. Magar, P. Yadav, and A. Barati Farimani, "Potential neutralizing antibodies discovered for novel corona virus using machine learning," *Scientific Reports*, vol. 11, no. 1, p. 5261, Dec. 2021, doi: 10.1038/s41598-021-84637-4.
- [55] K. Heiser<sup>1</sup> *et al.*, "Identification of potential treatments for COVID-19 through artificial intelligence-enabled phenomic analysis of human cells infected with SARS-CoV-2," *bioRxiv*, p. 2020.04.21.054387, 2020.
- [56] D. Jain and V. Singh, "A two-phase hybrid approach using feature selection and Adaptive SVM for chronic disease classification," *International Journal of Computers and Applications*, vol. 43, no. 6, pp. 524–536, Jul. 2021, doi: 10.1080/1206212X.2019.1577534.
- [57] N. E. M. Khalifa, M. H. N. Taha, G. Manogaran, and M. Loey, "RETRACTED ARTICLE: A deep learning model and machine learning methods for the classification of potential coronavirus treatments on a single human cell," *Journal of Nanoparticle Research*, vol. 22, no. 11, p. 313, Nov. 2020, doi: 10.1007/s11051-020-05041-z.

## BIOGRAPHIES OF AUTHORS



**Ishaan Walecha**    is pursuing his Bachelor of Technology in Computer Science from The NorthCap University. He can be contacted at email: walechai019@gmail.com.



**Dr. Divya Jain**    is an active and dedicated resource in the Computer Science Department. She is currently working as a University Fellow in the Computer Science Department from 2015. She did her graduation in B. Tech Computer Science and Engineering with distinction from Maharishi Dayanand, University, Rohtak in 2012 and M. Tech in Computer science and Engineering in 2014 from ITM university, Gurugram. She did Diploma in .NET Technologies from NIIT Ltd. She is currently pursuing Ph.D. from NCU, Gurugram in the field of data mining and machine learning. She has published one book and many research papers in reputed International Journals and Conferences. Her core subjects are Data Mining and Machine Learning. She can be contacted at email: divyajain1890@gmail.com.