ISSN 2350-1022

International Journal of Recent Research in Mathematics Computer Science and Information Technology
Vol. 9, Issue 2, pp: (33-41), Month: October 2022 – March 2023, Available at: www.paperpublications.org

Machine Learning and its use in Covid-19 Detection: A Review

Satyendra Kumar Sagar

DOI: https://doi.org/10.5281/zenodo.7457985

Published Date: 19-December-2022

Abstract: **Coronavirus disease (COVID-19), an infectious disease strike the whole world and people of all age groups. It can cause common symptoms as fever, dry cough, tastelessness, loss of smell and can even lead to breath shortening, and even death. This can be symptomatic as well as asymptomatic in some patients and thus lead to increased communicability. Machine Learning is a latest trend currently useful in almost all research areas. In this study we reviewed papers that implements Machine learning techniques in COVID-19 diagnosis. Articles from all the popular digital libraries are included with higher citation both from researchers and industries. In this article we provide the state of the art of the covid diagnosis with advanced and non contact techniques and this can guide both clinicians and technologists. These techniques are more powerful since it reduce the dependency on the covid test kits and also the human made errors. Moreover, disease can be diagnosed with high accuracy from covid data, in less time and reduced cost using the biomedical images or clinically collected records.**

Keywords: **COVID-19, Corona Virus, Machine Learning, Convolutional Neural Network, Deep Neural Network, Computed Tomography, X-ray images.**

I. INTRODUCTION

In 2019 The outbreak of disease commonly known as Covid-19 i.e., coronavirus disease first strike peoples in China and then across the world. This disease is highly inter-person contagiousness and a it is a form of induced pneumonia. This disease have symptoms like cough, cold, fever, but some severe symptoms like shortening in breath can even lead to death with major organ failures [1]. However the vast connectivity of humans with technology and the world is even reached a level it was never before. It infected more than 100 million people around the world and lead to 2.35 million deaths. These numbers are increasing continuously.

This pandemic caused huge loss of manpower, economy due to lockdown and loss of productivity. This outbreak required huge efforts to develop the test kits such as reverse transcription polymerase chain reaction (RT-PCR) kits [2], antigen test or rapid test kits. However these tests take time, and requires huge resources and this is where Artificial intelligence or machine learning (ML) plays an essential role in covid case classification. These ML models can be used to predict infectious cases and recovery rates using chest x-ray, CT scans [3] or blood samples. In the first wave patients with severe symptoms were taken to hospital or intensive care unit. However the actual asymptomatic people (those who did not had symptoms) could not seek any medical assistance and remained undetected and uncounted. But they infected a huge population again. During the second wave many infectives along with the asymptomatic infectives were tested and got registered. The variants of covid are changing with time and their symptoms are also varying making it hard to detect.

Different ML models exists such as regression models, decision tree, support vector machine (SVM), clustering methods, and Convolutional neural network and many more [4]. In addition, access to the patient's private data can violate the privacy since Ml models requires huge data for training. But since in this pandemic the dataset is released by many organizations for public support, hence using this technique was possible. Some authors have used concept of federated learning, transfer

ISSN 2350-1022

International Journal of Recent Research in Mathematics Computer Science and Information Technology
Vol. 9, Issue 2, pp: (33-41), Month: October 2022 – March 2023, Available at: www.paperpublications.org

learning in the standard models. Federated learning is an effective solution for data privacy, centralized computation, and high computation power.

This pandemic requires the need of wearing masks. Wearing mask raised the scope of crimes as people can not be identified with masks. ML techniques can also be used to detect faces with masks and identify the person.

Using ML models can help in pre screening of patients before test, non contact methods for diagnosis and low cost solution. Moreover, successful screening of contaminated patients, is subject to doctors error (human error) and also the current test have higher negative false reports around 15-20% [5]. Moreover, RT-PCR test have low sensitivity.

The purpose of this study is to detect the role of ML techniques and algorithms in investigating and various purposes that deals with COVID-19. This review has included articles from the year 2019 to 2021 and all the related articles in different digital libraries such as Science Direct, IEEE, Springer, and MDPI etc. We investigated and suggested different future directions in this aspect and summarize our key findings. ML can lead to better investigations, diagnosis, prediction, and discrimination.

The overview of the review paper is as follows, section 2 gives brief description f ML types and techniques. Section 3 gives the summary of different works using the ML techniques and section 4 gives a brief overview of datasets. In section 5 conclusion and future scope of ML in covid detection is provided.

II. MACHINE LEARNING

Machine learning (ML) is a method of data analysis to automate a data driven model. It is a subset of where machines learn from data, identify patterns and make decisions with minimal human intervention.These algorithms are more reliable than humans but they specifically depends on the input data provided to learn. The data provided is divided into 3 parts one for training, one for validation and other for testing. Based on the accuracy on the testing data optimum parameters and model are chosen. Always higher accuracy model are preferred because they have less error and are more close to actual predictions.

Types of Machine Learning Algorithms: Machine learning algorithms can be broadly classified into three sub-types as shown in Figure 1. These are supervised learning methods, unsupervised learning methods and reinforcement learning methods. There is two types of data, one is labeled and the other is unlabeled data.

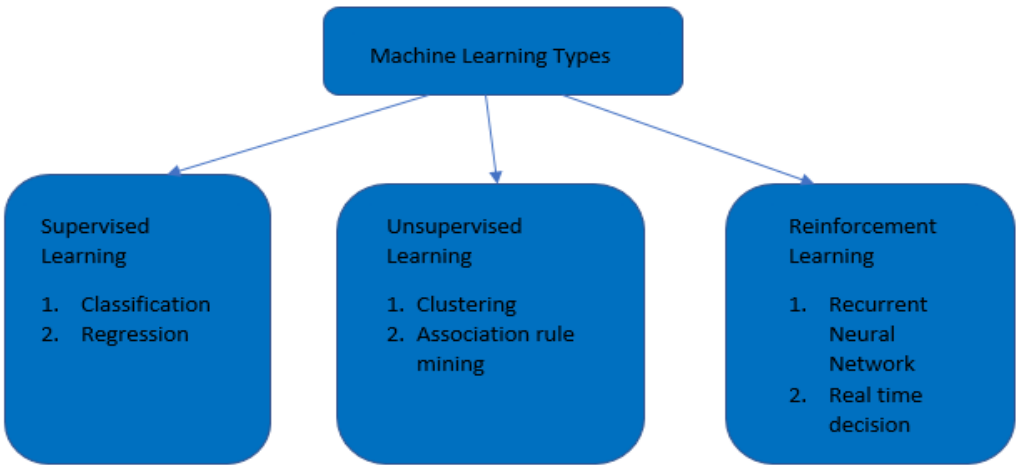

Figure 1: Types of Machine learning algorithms.

Supervised learning methods: These machine learning algorithms have labeled data i.e. input is given with the output value. Since the output is given within an training example it is known as supervised learning algorithms. Many regression and classification problems falls under this category. Labeled data has both the input and output parameters in a completely machine-readable pattern, but requires a lot of human labor to label the data, to begin with. It is one of the most basic types of ML algorithms. Even though the data needs to be labeled accurately for this method to work, supervised learning is extremely powerful when used in the right circumstances.

ISSN 2350-1022

International Journal of Recent Research in Mathematics Computer Science and Information Technology
Vol. 9, Issue 2, pp: (33-41), Month: October 2022 – March 2023, Available at: www.paperpublications.org

This is used to find relationship between the features or parameters to establish a cause and effect relation in the dataset. We can also say that it finds relationship between input and output features. The model trained is then tested on the dataset and testing accuracy is examined. Once the accuracy is significantly high it can be used further on real dataset. In general there are different types of supervised machine learning algorithms such as regression, support vector machine, convolutional neural network, and many more. Some of them are explained in detail below:

- Linear Regression Techniques: The relation between the input and output is linear in this learning and approach and the parameters determined are slope and intercept. Training data is used to learn automatically the value of slope and intercept such that almost all the samples passes through this line or the error is minimum between the predicted value and actual value. Figure 2 shows an example of linear regression. The relation is given as follows for linear regression:

$Y= ax+b$

Where y is output, x is input, a and b are slope and intercept respectively.

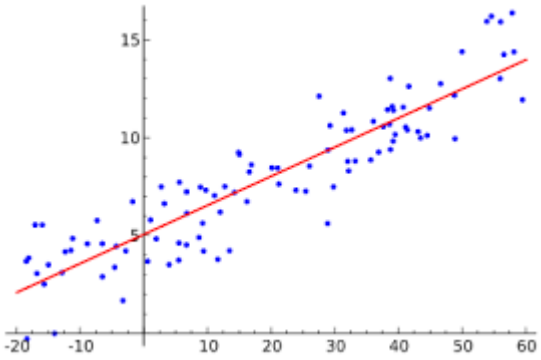

Figure 2: Example of linear regression[6].

- Decision Tree: Decision Tree is a flow chart like tree-shaped structure that represents sets of decisions and is able to generate rules for the classification of a data set. Decision tree has three types of nodes namely are Root Node, Internal Node and Leaf or Terminal Node. A record enters the tree in the root node [7]. The root node is applied to the test to determine which internal node will find the underlying record. There are several algorithms for the initial test, but the goal is always the same: choose the test that best discriminates between the target classes. This process is repeated until the record reaches a leaf node. All the records that end up at a given leaf of the tree are classified the same way, and each leaf node is assigned a class label. In the Figure 3 a decision tree is shown with root nodes and leaf nodes.

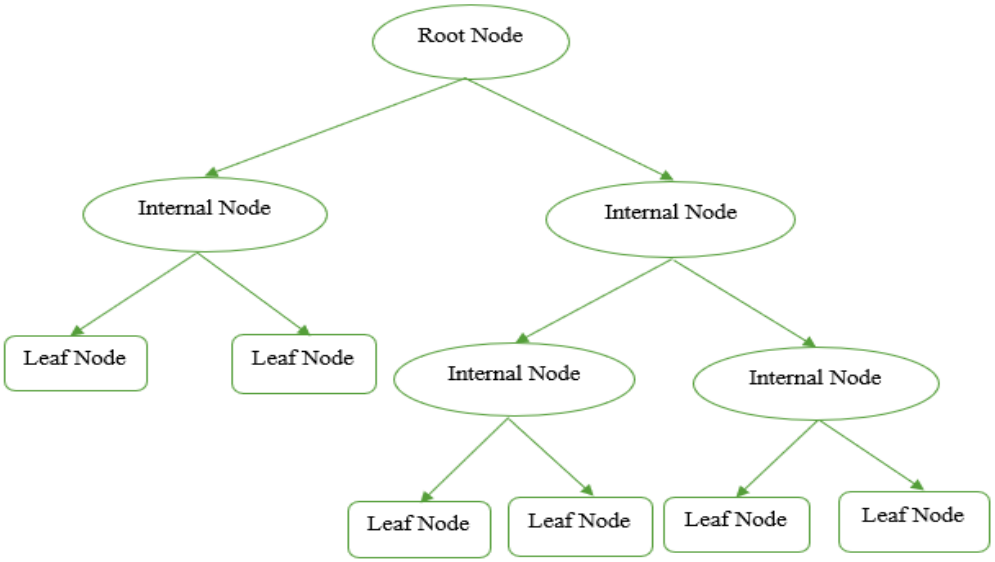

Figure 3: Decision tree.

ISSN 2350-1022

International Journal of Recent Research in Mathematics Computer Science and Information Technology
Vol. 9, Issue 2, pp: (33-41), Month: October 2022 – March 2023, Available at: www.paperpublications.org

- Artificial Neural Network: An artificial neural network (ANN) is comes from the concept and working of human brain and neurons. It simulate the functioning of a human brain. An artificial neural network often just called a neural network. ANN is an efficient and most used approach in the field of the computing system. The idea behind the ANN is based on Biological Neurons in the human brain [8]. There are three layers an input layer, one or more hidden layers and an output layer. There is web of neurons processing units that make up an ANN, which in turn consist of inputs and outputs. For the provided inputs ANN learns to produce the desired output. The training process is performed using backpropagation from which all the weights are learned. Usually ANN are used for modelling non-linear problems. They find applications in many areas like health, classification, industries, etc. Figure 4 shows the image of ANN.

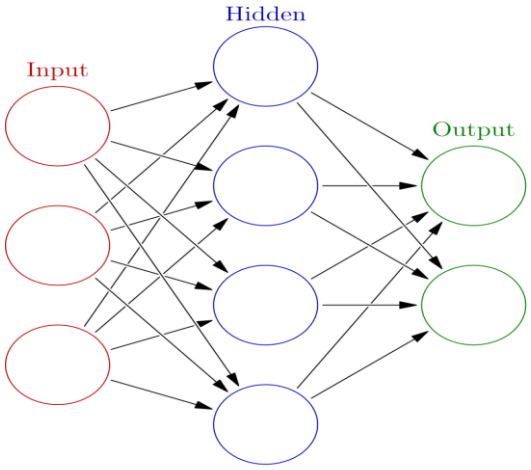

Figure 4: Artificial neural network [8].

Unsupervised Learning: Data used here is unlabeled data and this has either only one or none of the parameters in a machine-readable form. This negates the need for human labor but requires more complex solutions. Using these algorithms without humans interference one can learn the relation and pattern among the data. These learning results in finding the hidden structures and relationships among data. One of the common type of unsupervised learning method is k-means clustering. This algorithm uses unlabeled data (i.e., data without defined categories or groups). The goal of this algorithm is to find groups in the data, with the number of groups represented by the variable ***K***. The algorithm works iteratively to assign each data point to one of ***K*** groups based on the features that are provided. Data points are clustered based on feature similarity. Figure 5 shows an K-means clustering an unsupervised learning algorithm.

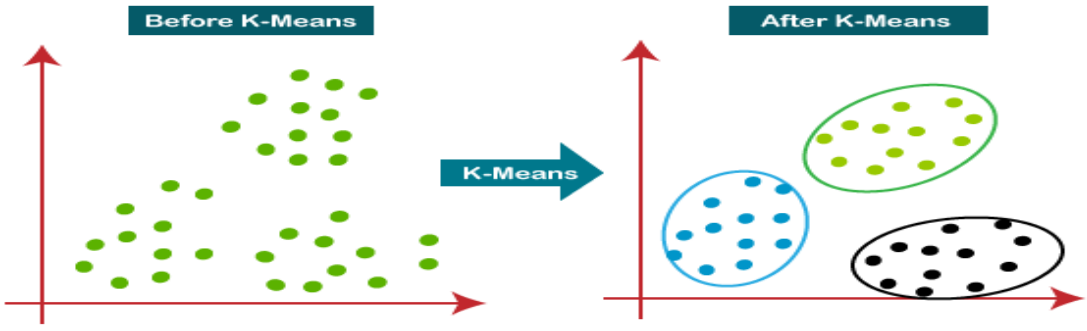

Figure 5: K-means clustering an unsupervised learning algorithm [9].

Reinforcement learning: In the previous methods a training data is available to train a machine but for scenarios where data is not available, reinforcement learning comes into picture. This requires no prior dataset and directly takes inspiration from how human beings learn from data in their lives. It features an algorithm that improves upon itself and learns from new situations using a trial-and-error method. Favorable outputs are encouraged or 'reinforced', and non-favorable outputs are discouraged or 'punished'. We can say their is reward for right answers and punishment for wrong answers.

ISSN 2350-1022

International Journal of Recent Research in Mathematics Computer Science and Information Technology
Vol. 9, Issue 2, pp: (33-41), Month: October 2022 – March 2023, Available at: www.paperpublications.org

III. LITERATURE REVIEW

In this section we present the current state of the art of techniques of artificial intelligence (machine learning) used in Covid-19. This disease is causing a global health crisis. Figure 6 shows the different works done with application of ML in COVID-19.

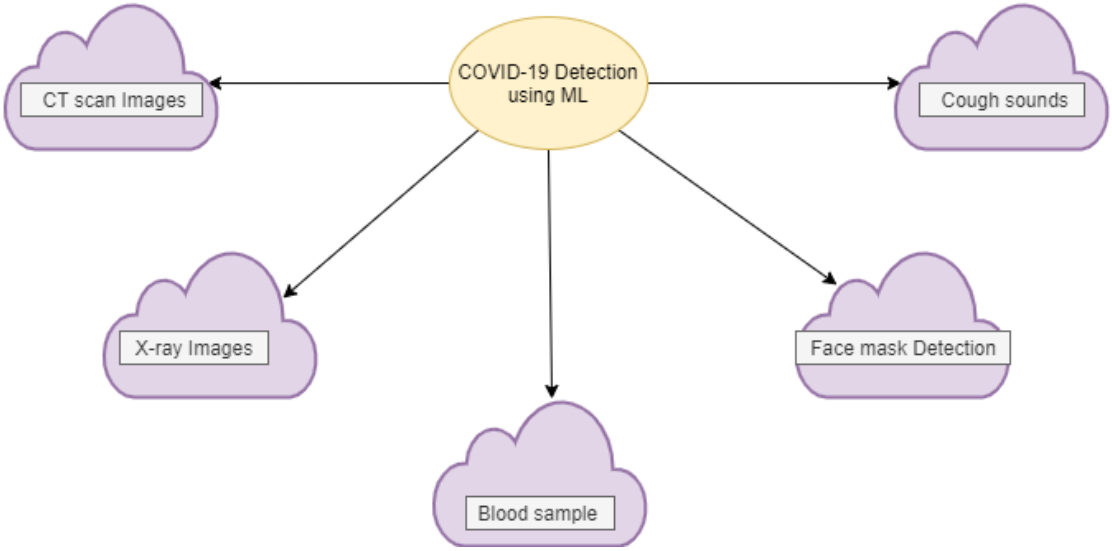

Figure 6: Works done using the ML in COVID-19.

The key symptom in this disease is cough and thus in [10] authors recorded coughs through smartphone of the healthy (COVID-19 negative) and COVID-19 positive persons. They found that covid positive coughs samples are 15%–20% shorter than non-covid coughs. They classified and discriminated the cough using machine learning method and this is a non-contact screening, easy to apply and reduce the load of covid tests. They could provide early self-isolation for suspected cough suggestive of COVID-19. Using six different ML techniques with leave-*p*-out cross-validation scheme to train and evaluate, best results were obtained using the Resnet50 classifier.

Computer Tomography (CT) [11] image are used to detect of level of Corona Virus Disease. Commonly known as CT scan, the manual judgement of these images are subject to doctor's accuracy and it can be improved by using a deep classification based on convolution and deconvolution local enhancement [12]. These operation, enhances the contrast between local lesion region and abdominal cavity of COVID-19 and obtain middle level features. By this, it can effectively determine whether the feature vector in each feature channel contains the image features of COVID-19.

Some authors [13] used super resolution CT scans images into very deep, super-residual neural networks to enhance lung CT scan efficiency. They used existing pre-trained models to decrease the training time. Using the pre-processing step improved the overall performance of the parameters.

The hand crafted feature extraction from the CT scans can also be used such as Grey Level Co-occurrence Matrix (GLCM), Local Directional Pattern (LDP), Grey Level Run Length Matrix (GLRLM), Grey-Level Size Zone Matrix (GLSZM), and Discrete Wavelet Transform (DWT). They formed the dataset by making patches (16x16, 32x32, 48x48 and 64x64) of complete images [14] and extracted feature from patches, increasing the overall accuracy using SVM. GLSZM method gave the best accuracy of 99.68%.

CT scans can also be also be used for suspected RT-PCR tests and using scattergram images can help to achieve this. In [15] authors proposed Scat-NET which is 25 layered CNN model integrated with scattergram images . These images are frequently used to reveal the numbers of neutrophils, eosinophils, basophils, lymphocytes and monocytes. These are measurements to evaluate disease symptoms, and the relationship. They revealed that patients with negative RT-PCR but positive CT test resulted to be positive. They proposed that the Scat-NET model can be arranged at the CT scan test.

However automatic detection of disease (COVID-19) from CT scans are subject to large datasets, ambiguity in the characteristics and the model accuracy or recall. Hence authors in [16] propose a method to diagnose CT scans with high

recall and accuracy. Many a times recall is better measure than precision thus they explore a trade-off among them. They used the concept of pre-trained ML models i.e. transfer learning approach. The proposed stacked ensemble use four CNN models: VGG-19, ResNet-101, DenseNet-169 and WideResNet-50-2.

Another work that used digital chest x-ray radiographs to automatically detect COVID-19 pneumonia patients using Deep CNN [17]. The Deep CNN model is Inception V3 with transfer learning gave an overall accuracy of 98% and above.

Federated learning is another approach of distributed learning and thus in [18], authors integrated it for covid detection. They identified the factors affecting model accuracy and loss like activation function, model optimizer, learning rate, number of rounds, and data size during the model training stage. They found that softmax activation function and SGD optimizer give better prediction accuracy and loss. Change in the number of rounds and learning rate has slight effect on accuracy and loss whereas increase in the data size have no effect. They specifically demonstrated that the federated ML model has a better prediction accuracy and loss but higher performance time than the traditional ML model.

X-ray images are another diagnosis method for determining the covid in the lungs. It is difficult to classify between chest X-ray images of common Pneumonia, Covid positive, and healthy lungs and thus [19] used a classifier ensemble technique. They utilized Choquet fuzzy integral to increase accuracy of individual classifier. They used transfer learning approach to train the base CNN classifiers (with two dense layers and one softmax layer) using InceptionV3, DenseNet121, and VGG19.

Currently, Reverse transcription polymerase chain reaction (rRT-PCR) is standard test for covid-19. But it requires 3-4 hours to generate results and higher false-negative rates (15-20%). Moreover these test require certified laboratories, costly equipment and trained persons to test the patients. Thus the authors used two ML models to classify hematochemical values from routine blood exams [20]. They discriminated between covid positive or not based on the clinical interpretations of blood tests samples.

Another work [21] that used blood samples first extracted eleven important features/indices from blood using random forest algorithm. This method achieved significant accuracy that is similar to CT scan images.

People wearing masks in the covid pandemic cannot be identified easily thus using ML models peoples with face mask can be identified. Authors in [22] used a hybrid ML model for face mask detection with two components. They first extracted features from Resnet50 model and then classified face masks using decision trees, Support Vector Machine (SVM), and ensemble algorithm.

In some cases the textual clinical reports can be classified into four classes. This can be achieved using ML algorithms with feature engineering techniques like Term frequency/inverse document frequency (TF/IDF), Bag of words (BOW) and report length [23]. ML classifiers are implemented on these features like Logistic regression and Multinomial Naive Bayes classifiers.

In another work [24], the identification of the severity of covid can be done to facilitate the risk estimation. 32 highly associated features to detect covid-19 severeness. However further, inter-feature redundancies among the 32 features was identified and finally selected 28 features were used to train the model. They achieved an overall accuracy of 81.48%.

Some authors use a software tool comprised of unsupervised Latent Dirichlet Allocation (LDA) and other ML methods to analyze the Twitter data in Arabic. They aim to detect government pandemic measures and public concerns during the COVID-19 pandemic [25]. They described the tool in detail with architecture, software components, and algorithms. They collected 14 million tweets from the Kingdom of Saudi Arabia in 4 months 1 Feb to 1 June 2020. They detected timewise progression of events from the public.

Feature extraction and ML can be used to predict the affinity of protein sequences of ten viruses to three categories of RNA sequences [26]. RNAs can be involved in the protein-RNA complexes stored in the RCSB database, the human miRNAs deposited at the mirBase database, and the lncRNA deposited in the LNCipedia database. They found that evolution tries to conserve key viral proteins involved in the replication and transcription ans prunes their interaction capability. Clustering techniques can be used to probe the conformational plasticity of the RBD using long molecular dynamics (MD) simulation [27]. Table 1 gives the summary of works in the above section.

ISSN 2350-1022

International Journal of Recent Research in Mathematics Computer Science and Information Technology
Vol. 9, Issue 2, pp: (33-41), Month: October 2022 – March 2023, Available at: www.paperpublications.org

Table 1: Summary of works in the literature review.

Paper	Focus	Dataset	ML	Performance parameter
Pahar et al. [10}	Detection from Cough sounds	Coswara dataset (92 covid positive and 1079 Healthy) + self collected data in South Africa (18 covid positive + 26 normal)	Regression, k-nearest neighbour, SVM, MLP, CNN, long short-term memory (LSTM) and Resnet50.	Accuracy 98%
Fang et al. [12]	CT scan images	COVID-CT dataset 143 patients 1460 images (Petuum researchers, University of California).	Feature extraction of region of interest (ROI) of in images	Sensitivity=98%, Specificity=96%, Positive & negative predictive value=98%, Precision=97%.
Arora et al. [13]	Lung CT-SCAN images	SARS-COV-2 CT-Scan and Covid-CT Scan	Pre-trained models used such as XceptionNet, MobileNet, InceptionV3, DenseNet, ResNet50, and VGG16	Precision with MobileNet 94.12% and 100% .
Barstugan et al. [14]	Abdominal Computed Tomography (CT) images	150 CT images	Feature extraction from handcrafted methods, Support Vector Machines (SVM)	accuracy was obtained as 99.68%
Tuncer et al. [15]	CT scan and Scattergram images	335 patients	CNN (depth 25)	Overall accuracy 92.4%.
Jangam et al. [16]	CT scan images		Transferred learning using VGG-19, ResNet-101, DenseNet-169 and WideResNet-50-2	
Asif et al. [17]	digital chest x-ray images	The dataset consists of 864 COVID-19, 1345 viral pneumonia and 1341 normal chest x-ray images	Inception V3 with transfer learning	overall accuracy of 98%
Salam et al. [18]	Factors affecting the classification accuracy	descriptive dataset and chest xray (CXR) images	Compared federated learning and traditional learning	NA
Dey et al. [19]	Classifies chest X-ray images	IEEE and Kaggle datasets CMSC-678-ML-Project GitHub dataset	Choquet fuzzy integral and CNN InceptionV3, DenseNet121, and VGG19	Recall=99.00%, precision=99.00%, F-score=99.00%, Accuracy=99.02%.
Davide et al. [20]	Blood tests samples	279 patients who, after being admitted to the San Raffaele Hospital (Milan, Italy) 177 resulted positive, whereas 102 received a negative response.	Decision Tree model	Accuracy=82% to 86%, Sensitivity=92% to 95%.
Wu et al. [21]	Blood samples	49 clinical available blood test data	Random forest	Accuracy=98% Sensitivity=95%, specificity=96.97%
Loey et al. [22]	face mask detection	Real-World Masked Face Dataset (RMFD), the Simulated Masked Face Dataset (SMFD), and the Labeled Faces in the Wild (LFW)	Resnet50, decision trees, SVM, and ensemble algorithm	Accuracy by SVM classifier is 99.64% in RMFD, 99.49% in SMFD and 100% in LFW.
Khanday et al. [23]	Classification from the textual clinical reports		Feature extraction using Term frequency/inverse document frequency (TF/IDF), Bag of words (BOW) and report length. Logistic regression and Multinomial Naive Bayes classifiers	Accuracy=96.2%

IV. DATASETS

Different authors, hospitals, clinics all are collecting data of the covid patients at their end. But some general datasets for Covid-19 can be downloaded from the following sources. Kaggle [28] provides the Open Research Dataset (CORD-19) from the White House and a coalition of leading research groups. It covers over 500,000 scholarly articles. It is good enough to apply recent advances in natural language processing and other AI techniques to in support of the fight against COVID.

An open-access data and computational resources are being provided by federal agencies, including NIH, public consortia, and private entities. One such dataset is provided by Data Science Strategy [29]. These resources are being aggregated and posted for scientific and public health interests. Google also provides an open access dataset for researchers.

Google has produced a hosted repository of public datasets, like covid-19 open dataset, the Global Health Data from the World Bank, and OpenStreetMap data, free to access and query through our Covid-19 public datatset program to aid researchers, data scientists, and analysts in the effort to combat COVID-19 [30].

V. CONCLUSION AND FUTURE SCOPE

With the sudden outbreak in the world due to COVID-19, it has rapid global health concern. However the current state of testing lack in fast inference or test reports and it is higher in cost. ML plays a vital role in early diagnosis and timely treatment of covid. It is almost no cost testing method and also require no special training of the personals. This technology advancement of AI and ML has influenced every field of life even the medical field and shown promising results in health care. The decision making and diagnosis of covid can be easily achieved using CT scan images, X-ray images and cough sounds by analysing these data. These tools can carry out preliminary assessment of suspected patients and help them to get timely treatment and quarantine suggestion. Supervised learning show better outcomes than than Unsupervised learning methods.

However since the covid phase is changing with the varients, different symptoms are arriving and thus recurrent supervised learning can be better used to achieve superior accuracy. We have encountered works with transfer learning, federated learning and unsupervised learning but reinforcement learning is a less explored area.

Also incremental learning in this domain can play vital role to generate upto the mark models. utilized for superior accuracy. However ML is subject to the availability of the epidemiological data strongly, since it affects the reliability of several mathematical models.

Moreover adding better acceleration methods can be deployed to run these models on low resource devices.

REFERENCES

[1] Sudre, Carole H., et al. "Attributes and predictors of Long-COVID: analysis of COVID cases and their symptoms collected by the Covid Symptoms Study App." *Medrxiv* (2020).

[2] Banaganapalli, Babajan, et al. "Multilevel Systems Biology Analysis of Lung Transcriptomics Data Identifies key miRNAs and Potential miRNA Target Genes for SARS-CoV-2 Infection." *Computers in Biology and Medicine* (2021): 104570.

[3] Roberts, Michael, et al. "Machine learning for COVID-19 detection and prognostication using chest radiographs and CT scans: a systematic methodological review." *arXiv preprint arXiv:2008.06388* (2020)

[4] Nabavi, Shahabedin, et al. "Medical Imaging and Computational Image Analysis in COVID-19 Diagnosis: A Review." *Computers in Biology and Medicine* (2021): 104605.

[5] Kwekha-Rashid, Ameer Sardar, Heamn N. Abduljabbar, and Bilal Alhayani. "Coronavirus disease (COVID-19) cases analysis using machine-learning applications." *Applied Nanoscience* (2021): 1-13.

[6] https://en.wikipedia.org/wiki/Linear_regression

[7] Safavian, S. Rasoul, and David Landgrebe. "A survey of decision tree classifier methodology." *IEEE transactions on systems, man, and cybernetics* 21.3 (1991): 660-674.

[8] https://en.wikipedia.org/wiki/Artificial_neural_network

[9] https://www.analyticsvidhya.com/blog/2021/04/k-means-clustering-simplified-in-python/

[10] Pahar, Madhurananda & Klopper, Marisa & Warren, Robin & Niesler, Thomas. (2021). COVID-19 Cough Classification using Machine Learning and Global Smartphone Recordings. Computers in Biology and Medicine. 135. 104572. 10.1016/j.compbiomed.2021.104572.

[11] Buzug, Thorsten M. "Computed tomography." *Springer handbook of medical technology*. Springer, Berlin, Heidelberg, 2011. 311-342.

[12] Fang, Lingling, and Xin Wang. "COVID-19 deep classification network based on convolution and deconvolution local enhancement." *Computers in Biology and Medicine* (2021): 104588.

[13] Arora, Vinay, et al. "Transfer learning-based approach for detecting COVID-19 ailment in lung CT scan." *Computers in biology and medicine* (2021): 104575.

[14] Barstugan, Mucahid, Umut Ozkaya, and Saban Ozturk. "Coronavirus (covid-19) classification using ct images by machine learning methods." *arXiv preprint arXiv:2003.09424* (2020).

[15] Tuncer, Seda Arslan, et al. "Scat-NET: COVID-19 Diagnosis with a CNN Model using Scattergram Images." *Computers in Biology and Medicine* (2021): 104579.

[16] Jangam, Ebenezer, and Chandra Sekhara Rao Annavarapu. "A stacked ensemble for the detection of COVID-19 with high recall and accuracy." *Computers in Biology and Medicine* 135 (2021): 104608.

[17] Asif, Sohaib, et al. "Classification of COVID-19 from Chest X-ray images using Deep Convolutional Neural Network." *2020 IEEE 6th International Conference on Computer and Communications (ICCC)*. IEEE, 2020.

[18] Abdul Salam, Mustafa, Sanaa Taha, and Mohamed Ramadan. "COVID-19 detection using federated machine learning." *Plos one* 16.6 (2021): e0252573.

[19] Dey, Subhrajit, et al. "Choquet Fuzzy Integral-based Classifier Ensemble Technique for COVID-19 Detection." *Computers in Biology and Medicine* (2021): 104585.

[20] Brinati, Davide, et al. "Detection of COVID-19 infection from routine blood exams with machine learning: a feasibility study." *Journal of medical systems* 44.8 (2020): 1-12.

[21] Wu, Jiangpeng, et al. "Rapid and accurate identification of COVID-19 infection through machine learning based on clinical available blood test results." *MedRxiv* (2020).

[22] Loey, Mohamed, et al. "A hybrid deep transfer learning model with machine learning methods for face mask detection in the era of the COVID-19 pandemic." *Measurement* 167 (2021): 108288.

[23] Khanday, Akib Mohi Ud Din, et al. "Machine learning based approaches for detecting COVID-19 using clinical text data." *International Journal of Information Technology* 12.3 (2020): 731-739.

[24] Yao, Haochen, et al. "Severity detection for the coronavirus disease 2019 (COVID-19) patients using a machine learning model based on the blood and urine tests." *Frontiers in cell and developmental biology* 8 (2020): 683.

[25] Alomari, Ebtesam, et al. "COVID-19: Detecting government pandemic measures and public concerns from Twitter arabic data using distributed machine learning." *International Journal of Environmental Research and Public Health* 18.1 (2021): 282.

[26] Lanjanian, Hossein, et al. "High-throughput analysis of the interactions between viral proteins and host cell RNAs." *Computers in biology and medicine* 135 (2021): 104611.

[27] Kumar, Vikash, Haiguang Liu, and Chun Wu. "Drug repurposing against SARS-CoV-2 receptor binding domain using ensemble-based virtual screening and molecular dynamics simulations." *Computers in biology and medicine* 135 (2021): 104634.

[28] https://www.kaggle.com/allen-institute-for-ai/CORD-19-research-challenge

[29] https://datascience.nih.gov/covid-19-open-access-resources

[30] https://cloud.google.com/blog/products/data-analytics/free-public-datasets-for-covid19