IJARSCT

ISSN (Online) 2581-9429

International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)

Impact Factor: 6.252

Volume 2, Issue 2, March 2022

Covid-19 Detection Using Machine Learning and Deep Learning

Ms. S. S. Wankhede[1], Suvidhnya Dondalkar[2], Ravindra Farkade[3], Ashutosh Joshi[4], Ayush Barapatre[5], Ashwin Meshram[6]

Assistant Professor, Department of Computer Science & Engineering[1]

UG Students, Department of Computer Science & Engineering[2,3,4,5,6]

Priyadarshini Bhagwati College of Engineering, Nagpur, Maharashtra, India

Abstract: *The current COVID-19 pandemic threatens human life, health, and productivity. AI plays an essential role in COVID-19 case classification as we can apply machine learning models on COVID-19 case data to predict infectious cases and recovery rates using chest x-ray. Accessing patient's private data violates patient privacy and traditional machine learning model requires accessing or transferring whole data to train the model. In recent years, there has been increasing interest in federated machine learning, as it provides an effective solution for data privacy, centralized computation, and high computation power. In this paper, we studied the efficacy of federated learning versus traditional learning by developing two machine learning models (a federated learning model and a traditional machine learning model) using Keras and TensorFlow federated, we used a descriptive dataset and chest x-ray (CXR) images from COVID19 patients. During the model training stage, we will try to identify which factors affect model prediction accuracy and loss like activation function, model optimizer, learning rate, number of rounds, and data Size, we kept recording and plotting the model loss and prediction accuracy per each training round, to identify which factors affect the model performance, and we found that softmax activation function and SGD optimizer give better prediction accuracy and loss, changing the number of rounds and learning rate has slightly effect on model prediction accuracy and prediction loss but increasing the data size did not have any effect on model prediction accuracy and prediction loss. finally, we build a comparison between the proposed models' loss, accuracy, and performance speed, the results demonstrate that the federated machine learning model has a better prediction accuracy and loss but higher performance time than the traditional machine learning model.*

Keywords: Chest X-ray, Covid19, Convolutional Neural Network, etc.

I. Introduction

The Coronavirus Disease 2019 (COVID-19) pandemic continues to have a devastating effect on the health and well-being of the global population, caused by the infection of individuals by the severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2). A critical step in the fight against COVID-19 is effective screening of infected patients, such that those infected can receive immediate treatment and care, as well as be isolated to mitigate the spread of the virus. The main screening method used for detecting COVID19 cases is reverse transcriptase-polymerase chain reaction (RT-PCR) testing, which can detect SARSCoV-2 ribonucleic acid (RNA) from respiratory specimens (collected through a variety of means such as nasopharyngeal or oropharyngeal swabs). While RTPCR testing is the gold standard as it is highly specific, it is a very time-consuming, laborious, and complicated manual process that is in short supply. The world has not yet fully recovered from this pandemic and the vaccine that can effectively treat Covid-19 is yet to be discovered. However, to reduce the impact of the pandemic on the country's economy, several governments have allowed a limited number of economic activities to be resumed once the number of new cases of Covid has dropped below a certain level. As these countries cautiously restarting their economic activities, concerns have emerged regarding workplace safety in the new post-Covid-19 environment. To reduce the possibility of infection, it is advised that people should avoid any person-to-person contact such as shaking hands and they should maintain a distance of at least 1 meter from each other. Furthermore, the sensitivity of RT-PCR testing is highly variable and have not been reported in a clear and consistent manner to date2, and initial findings in China showing relatively poor sensitivity3. Furthermore, subsequent findings showed highly variable positive rate depending on how the specimen was collected as well as decreasing positive rate with time after symptom onset.

II. Literature Survey

2.1. A literature Review on COVID-19 Disease Diagnosis from Respiratory Sound Data

Dawei Wang et al. in this research has described the epidemiological, demographic, clinical, laboratory, radio-logical and treatment data from Zhongnan Hospital, Wuhan China. The data was analysed and documented to be used to track the infections. The author gives better insights about the radio-logical and treatment data that could be used for our prediction of COVID-19 in our model.

2.2. AI Enabled Preliminary Diagnosis for COVID-19 from Cough Samples via an App

According to the results obtained from the Systematic Literature Review (SLR), RQ1 could not be answered thoroughly. In many works, a clear comparison between various machine learning algorithms has been conducted deliberately but the conclusion couldn't be achieved.

2.3. The Continuing 2019-nCoV Epidemic Threat of Novel Coronaviruses to Global Health

The latest 2019 novel coronavirus outbreak in Wuhan, China A systematic literature through the guidelines of Claes Wohlin and Barbara Kitchenham, has been conducted to analyze and answer RQ1. This literature review focuses on the understanding of several machine learning algorithms and also identifying appropriate machine learning algorithms that can be used for prediction.

2.4 Severe Acute Respiratory Syndrome-Related Coronavirus: The Species and its Viruses

A statement of the Coronavirus Study Group The present outbreak of lower respiratory tract infections, including respiratory distress syndrome, is the third spillover, in only two decades, of an animal coronavirus to humans resulting in a major epidemic.

3.5 Clinical Features of Patients Infected with 2019 Novel Coronavirus in Wuhan

China Recent cluster of pneumonia cases in Wuhan, China, was caused by a novel beta coronavirus, the 2019 novel coronavirus (2019- nCoV). We report the epidemiological, clinical, laboratory, and radiological characteristics and treatment and clinical outcomes of these patients.

III. Methodology

In this section, we present the proposed methodology for classifying an X-ray as being of a healthy patient or a patient affected by COVID-19. First, we describe the datasets of images used in this study. Then, we explain the process of feature extraction, which is based on the transfer learning theory. After that, we present the classification techniques applied and the steps of their training process. Lastly, we define the metrics we use to evaluate the results and to compare it to other approaches. Fig. 1 presents the infographics of the proposed approach; each step is explained in the next subsections.

3.1 Preprocessing

After collecting the database from the user we need to preprocess that image. For removing low-frequency background noise, normalizing the intensity of the individual particle images we use the preprocessing technique, by using MATLAB Firstly, we convert RGB images into grey scale images (rgb to gray converter). This will convert RGB images to high intensity Grey scale images. In this step we can perform noise removal and segmentation operation. The main aim of preprocessing is to reduce unwanted distortion and an improvement in input data (sign language images). The Image preprocessing technique uses the considerable redundancy in images. Neighboring pixel corresponding to one object in

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
Impact Factor: 6.252
Volume 2, Issue 2, March 2022

real image has adjusted some or similar brightness value. For preprocessing median filter is used for reduce "salt and pepper" noise in images. By using median filter smoothing, sharpening and edge enhancement operation are performed. The main use of median filter is that it can run through signal entry by entry or it can replace each entry with median of neighboring entries. Each output pixel contains the median value in the 3-by3 neighborhood around the corresponding pixel in the input image. g(x , y)={ $\boldsymbol{1iff(x, y)} \geq \boldsymbol{T0Otherwise}$ Here g (x, y) represents threshold image pixel at (x, y) and f (x, y) represents grayscale image pixel at (x, y).

3.2 Segmentation

Training on skin segmentation dataset. We used the pores and skin segmentation dataset from UCI containing approximately 2,00000 points for training the usage of gaining knowledge of algorithms like SVM. The trained models are then used to phase out the non-skin labeled pixels.

3.3 Feature Extraction

Feature extraction is a set of methods that map input features to new output features. It is a process of dimensionality reduction by which an initial set of raw data is reduced to more manageable groups for processing. A characteristic of these large data sets is a large number of variables that require a lot of computing resources to process. Feature selection is for filtering irrelevant or redundant features from your dataset. Feature extraction describes the relevant shape information contained in a pattern so that the task projected of classifying the pattern is made easy by a formal procedure. In pattern recognition and in image processing, feature extraction is a special form of dimensionality reduction.

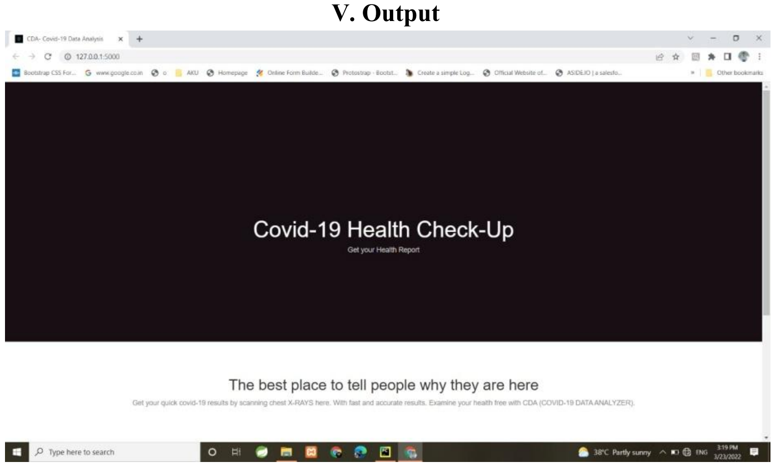
3.4 Classification

It refers to a predictive modeling problem where a class label is predicted for a given example of input data. Classification is a task that requires the use of machine learning algorithms that learn how to assign a class label to examples from the problem domain.

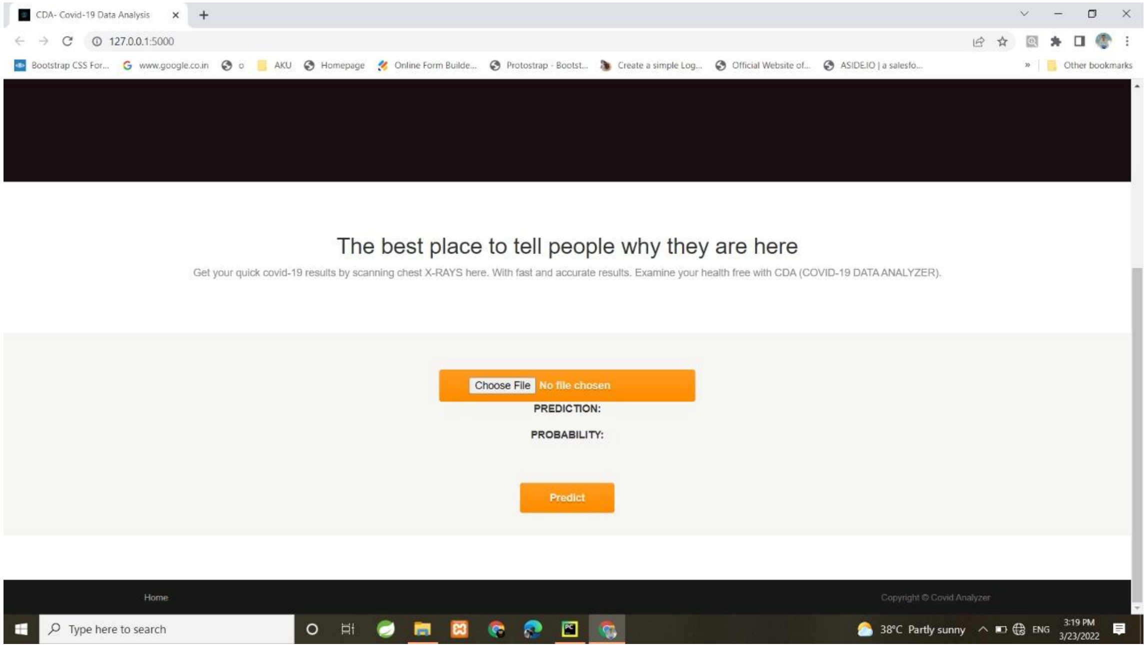
IV. Software Requirements

- Operating System: Window 7 Or Later
- Language: Python 3.9.7
- Open-cv (Library)
- Python IDE
- PyCharm 2020.3 Or Above Version

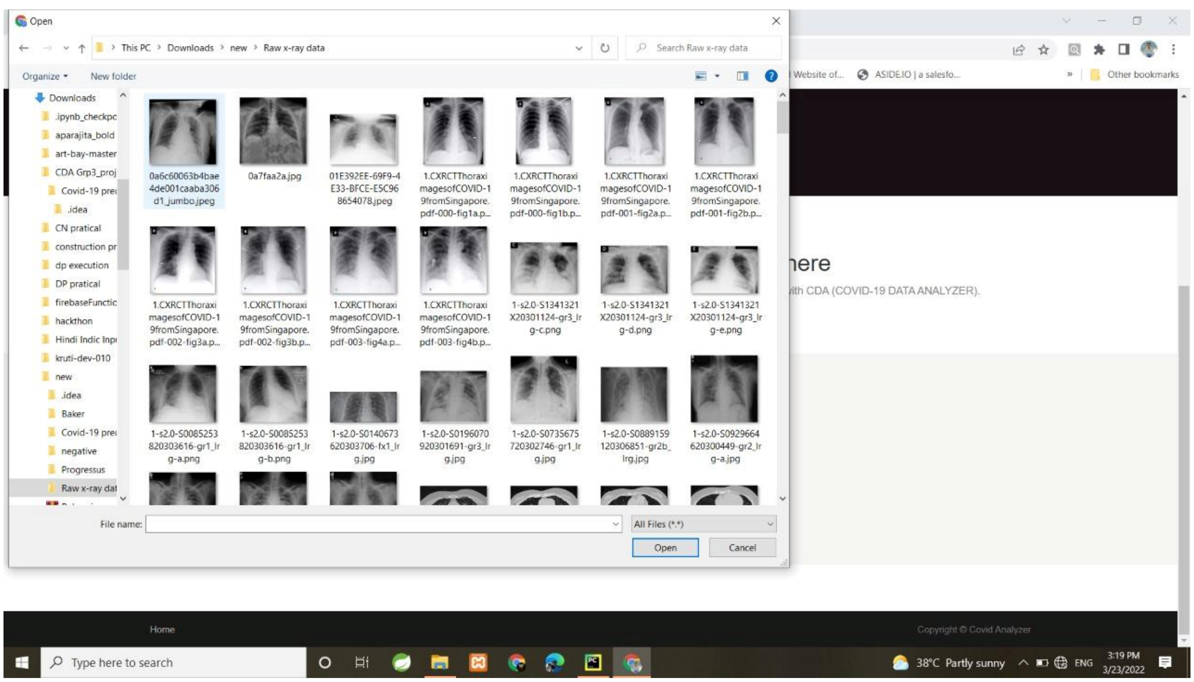
V. Output



Output 1: Home Page 1

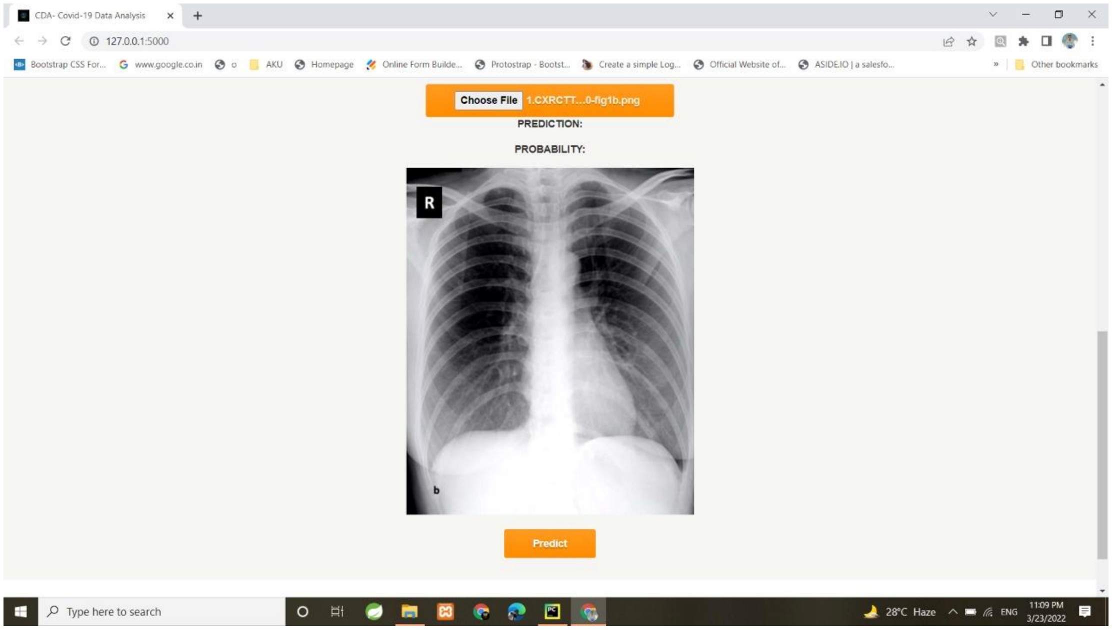


Output 2: Home Page 2

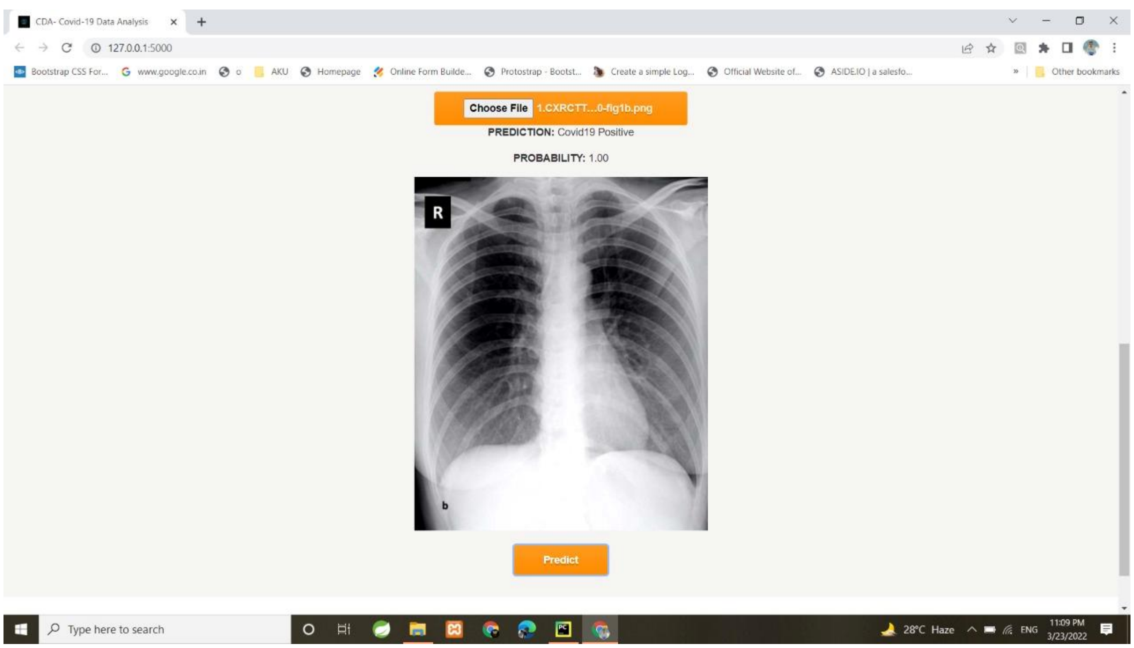


Output 3: Fetching Image for Processing

DOI: 10.48175/IJARSCT-2951

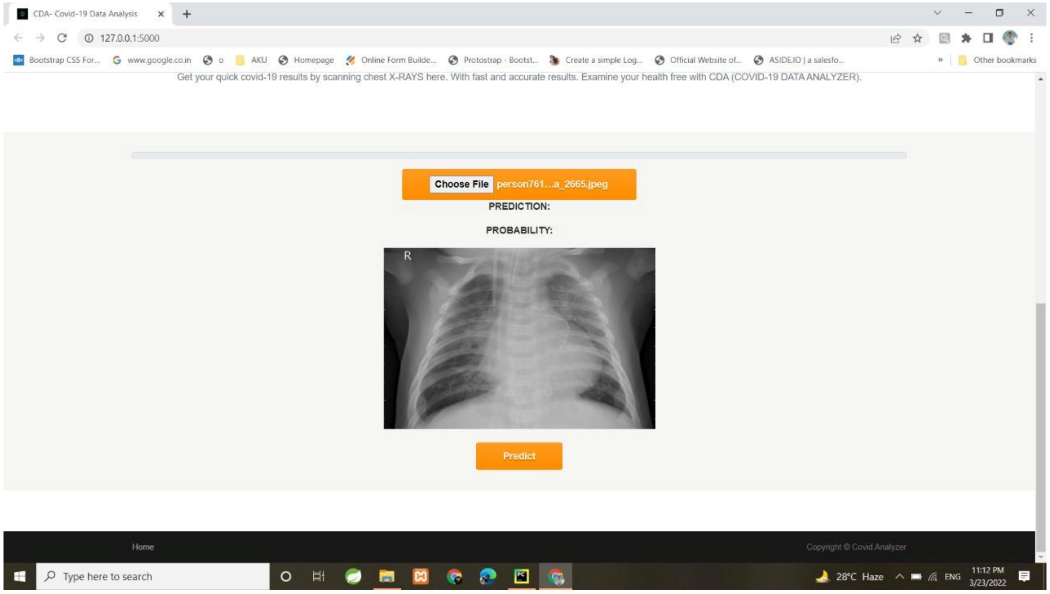


Output 4: Fetched Image for Processing 1

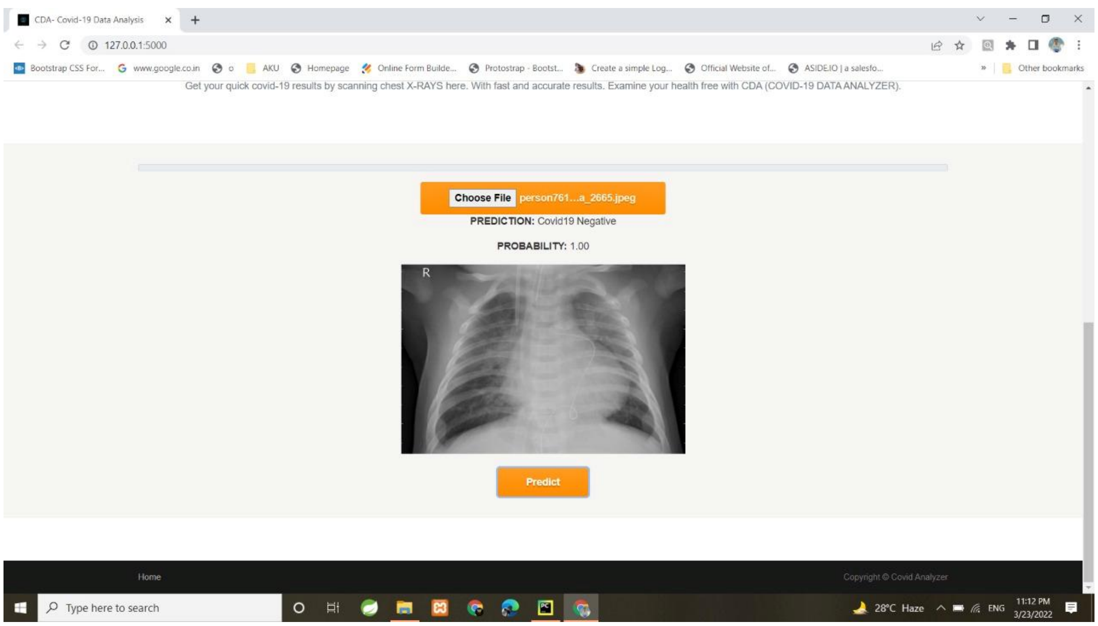


Output 5: Output Image (Positive)

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
Volume 2, Issue 2, March 2022
Impact Factor: 6.252



Output 6: Fetched Image For Processing 2



Output 7: Output Image (Negative)

VI. CONCLUSION

The results showed that the convolutional neural network with minimized convolutional and fully connected layers is capable of detecting COVID-19 imag within the two AUC scores of 96.51 and 96.33%, respectively. In addition, the second proposed architecture, which had the second-lightest architecture, is capable of detecting COVID-19 in three-class, COVID-19/Pneumonia/Normal images, with a macro-averagedF1 score of 94.10%. Therefore, the use of AI-based automated high-accuracy technologies may provide valuable assistance to doctors in diagnosing COVID-19.

REFERENCES

[1]. Ali Al-Hazmi. Challenges presented by MERS corona virus, and SARS corona virus to global health. Saudi journal of biological sciences, 23(4):507–511, 2016. Publisher: Elsevier.

[2]. Sina F Ardabili, Amir Mosavi, PedramGhamisi, Filip Ferdinand, Annamaria R Varkonyi-Koczy, Uwe Reuter, Timon Rabczuk, and Peter M Atkinson. Covid-19 outbreak prediction with machine learning. Available at SSRN 3580188, 2020.

[3]. Hiba Asri, Hajar Mousannif, Hassan Al Moatassime, and Thomas Noel. Using machine learning algorithms for breast cancer risk prediction and diagnosis. Procedia Computer Science, 83:1064–1069, 2016.

[4]. Taiwo Oladipupo Ayodele. Types of machine learning algorithms. New advances in machine learning, pages 19–48, 2010.

[5]. Taiwo Oladipupo Ayodele. Types of machine learning algorithms. New advances in machine learning, pages 19–48, 2010. Publisher: InTech . [6] Liu, Boyi, et al. Experiments of federated learning for covid-19 chest x-ray images." arXiv preprint arXiv:2007.05592 (2020).

[6]. Fiorino Gionata, et al. Inflammatory bowel disease care in the COVID-19 pandemic era: the Humanitas, Milan, experience." Journal of Crohn's and Colitis 14.9 (2020): 1330–1333. pmid:32211765.

[7]. WieczorekMichał, Jakub Siłka, and Marcin Woźniak. Neural network powered COVID-19 spread forecasting model." Chaos, Solitons Fractals 140 (2020): 110203.

[8]. Zhang Weishan, et al. Dynamic fusion-based federated learning for COVID-19 detection. IEEE Internet of Things Journal (2021).

[9]. Lian, Xiangru, et al. Can decentralized algorithms outperform centralized algorithms? a case study for decentralized parallel stochastic gradient descent. arXiv preprint arXiv:1705.09056 (2017).

[10]. Yang Qiang, et al. Federated machine learning: Concept and applications. ACM Transactions on Intelligent Systems and Technology (TIST) 10.2 (2019): 1–19.

[11]. Pang, Junjie, et al. Collaborative City Digital Twin For Covid-19 Pandemic: A Federated Learning Solution. arXiv preprint arXiv:2011.02883 (2020).

[12]. Afshar Parnian, et al. Covid-caps: A capsule networkbased framework for identification of covid-19 cases from xray images. Pattern Recognition Letters 138 (2020): 638– 643. pmid:32958971