International Journal of Early Childhood Special Education (INT-JECSE)
DOI:10.9756/INTJECSE/V14I5.809 ISSN: 1308-5581 Vol 14, Issue 05 2022

Performance Investigation of Artificial Intelligence and Machine Learning Approaches in Breast Cancer Detection Due to the Coronavirus (COVID-19) Pandemic

Devanand
M.Tech.Scholar
Computer Science & Engineering Department, Veer Bahadur Singh Purvanchal University, Jaunpur
Dileep Kumar Yadav
Assistant Professor
Computer Science & Engineering Department ,Veer Bahadur Singh Purvanchal University, Jaunpur
Sanjeev Gangwar
Assistant Professor
Computer Application Department ,Veer Bahadur Singh Purvanchal University, Jaunpur
Saurabh Pal
Professor
Computer Application Department ,Veer Bahadur Singh Purvanchal University, Jaunpur

Abstract
The novel coronavirus disease (COVID-19) has spread as a pandemic across 219 countries, wreaking havoc on health care, socioeconomic conditions, and international connections. The study's main goal is to give current technological features of artificial intelligence (AI) and other important technologies, as well as their implications for dealing with COVID-19 and preventing the pandemic's disastrous repercussions. This article introduces AI approaches that have made significant contributions to health care, then highlights and categorises their applications in combating COVID-19, such as detection and diagnosis, data analysis and treatment procedures, research and drug development, social control and services, and outbreak prediction. The research looks at the relationship between technology and epidemiology, as well as the possible effects of technology in health care with the advent of machine learning and natural language processing techniques.

Keywords
artificialintelligence(AI),coronavirusdisease(COVID-19),deeplearning(DL),healthcare,machinelearning(ML)technology

Introduction: The year 2020 began with the advancement of several digital technologies that will benefit health care. These technologies1, which include the internet of things (IoT) with fifth generation (5G) networks,2,3, bigdata,4, artificial intelligence (AI), including machine learning (ML) and deep learning (DL),5,6, and blockchain technology7, are being used to address problems in traditional health care systems and the pandemic. 8 The globe is currently facing a global health disaster caused by the coronavirus illness (COVID-19). 9 COVID-19, which was caused by a novel coronavirus (severe acute respiratory syndrome coronavirus-2 [SARS-CoV-2]), was identified using the World identification, isolation, rapid management, spread prediction, and contact tracking systems. 13 The primary challenges, however, include delays in viral tests, treatments, or medicines, as well as providing services to key zones. The primary goal is now to detect and diagnose the virus as early as possible, to monitor and nurse contacts continuously, to analyse epidemiological and medical reports from patients, and to track the progress of treatment procedures and drugs. Industry 4.0 technologies, such as AI, 5G-based IoT devices, and other digital technologies, are critical for health, social, and economic performance in the fight against the coronavirus. These technologies are capable of offering enhanced digital solutions for addressing difficulties throughout the disaster14-17 and alleviating the global health crisis caused by this disease. 1,7,18,19AI is one of the promising health-care technologies for better understanding and tackling the COVID-19 situation.

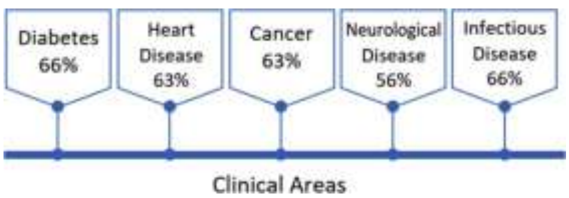

Figure 1 depicts the major therapeutic domains where AI and machine learning show the most potential.

According to the WHO, there will be 132 million people on the planet on April 4, 2021. 11,12,30 According to the Worldometer, which was updated on April 4, 2021, the 2019-nCoV is presently fast growing and upsetting 219 countries and territories, as well as 2 international conveyances (Figure 1). 11 On February 11, 2020, the WHO formally dubbed the disease COVID-19 and the virus that causes it SARS-CoV-2, and declared a worldwide health emergency. The 2019-nCoV virus has a significant genetic similarity to the severe acute respiratory syndrome (SARS) virus, which caused a pandemic in 2002. 31

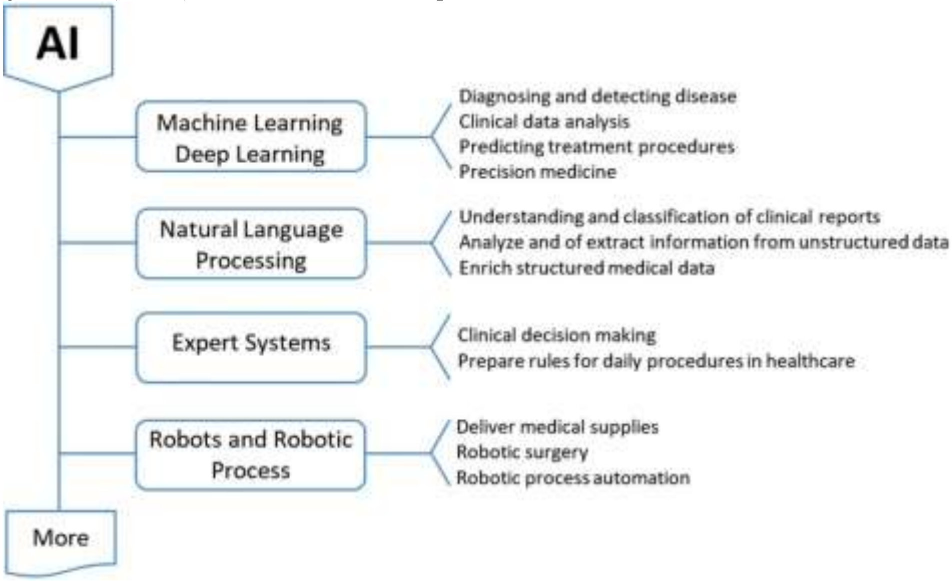

Figure 2 depicts some components of artificial intelligence (AI) technology in health care.

MachineLearningandDeepLearning
The most important type of AI is machine learning (ML), which has applications in a variety of situations in health care, including diagnosing patient characteristics, predicting treatment methods, and precision medicine. 35 DL or neural network models used in clinical data analysis and disease diagnosis are examples of composite forms of ML. 36,37

Processing of Natural Language
NLP applications include clinical document design, building, identification, interpretation, and classification, as well as research publications. The NLP may examine unstructured medical data about patients, organise radiological investigation reports, transcribe patient interactions, and extract information from unstructured data to improve structured medical data. 38,39 It is used in conjunction with the ML-based knowledge graph to discover a link between the coronavirus and prospective therapeutic candidates and to generate outputs. 40,41

Systems of Expertise
Expert systems, which include human experts and knowledge engineers, are frequently used for clinical decision making37, and providers of electronic medical records develop a set of rules with daily operations in health care.UnderstandingandDetecting*To prevent the spread of infectious diseases, viruses shouldbe tracked earlier. Quick detection of virus helps for patientisolation, contact tracing, treatment, and the delivery of emer-gency alerts to others. AI-based systems can learn to detectvirusoutbreaksby analyzingnewsreports,socialmediaposts,andotheroutbreak-relateddocuments.AIsystemBlueDot,[42]usingMLalgorithms,wasusedtotrackanddetectinfectiousdiseaseoutbreaksin Wuhan,China.Another AI predictive system, ClosedLoop C-19 index,[43]using expert knowledge in health care can identify thosepeople who are at the highest risk of critical complicationsfromCOVID-19.AI-basedsurveillancesystems,usingfacialrecognitiontoolsandtemperaturesensors,canidentifypeoplewhohavefeversa ndwhohavecoronavirus.Severalsearch engines, social networking sites, and media sites usingML approaches can be used to track flu-related disease out-breaks in real time.[44] But the main challenge here is datasafetyandsecurity.Itisrequiredtoestablishastandardpro-tocoltopermitcommunicationbetweendevicesandsystemswithoutcompromisingdataintegrity.*

Treatment, Diagnosis, and Genome Analysis
Rapid and accurate detection of COVID-19 infections can save lives and prevent future outbreaks. AI technology has a greater impact on disease detection and diagnosis. 20 It can gather data from numerous clinical reports, evaluate DNA sequences, and provide therapy recommendations. 45 To identify COVID-19 instances, a DL method46 and a supervised neural network model (known as COVNet)47,48 were created to assess patients' chest x-ray and computed tomography (CT) scan pictures. 49 The model can identify between COVID-19(+)

DOI:10.9756/INTJECSE/V14I5.809 ISSN: 1308-5581 Vol 14, Issue 05 2022

and COVID-19() patients with high accuracy. 50 Several AI startups have introduced AI-powered diagnosis solutions to aid clinicians in viral identification and disease outbreak monitoring. In the United States, XR Health51 has announced a virtual reality tele-health support group to help people who have been isolated due to the virus, so that individuals with comparable ailments can get assistance from each other and from doctors. In addition, DL methods that can process massive and complicated genomic information are used in SARS-CoV-2 genomic sequencing. 52–54 For example, Google DeepMind's AI algorithms can aid in the understanding of virus protein structure and the development of new illness treatment processes. 55 Because of dataset volatility, new demographics, and the negative repercussions of new algorithms on health outcomes, AI systems for diagnosis, analysis, and therapy encounter logistical challenges in implementation.

PredictionandForecastingthe Outbreak

AI has been employed for keeping track of and predicting howCOVID-19spreadsovertimeandspace.[21]Akhtaretal[56]pre-sented a dynamic artificial neural network prototype to forecastthe span of the COVID-19 pandemic. This approach wasapplied for the prediction of the 2015 Zika virus pandemic.HealthMap[57] and BlueDot[42] have been developed using MLalgorithms that can efficiently predict the outbreak of thevirus.Wangetal[58]proposedaprognosismodelforinfluenzadiagnosis, depending on analysis of the real-time Twitter infor-mationthatcanhelptopreventfutureoutbreaks.[59]TheXGBoost model[60] is another ML-based prediction modelthat was used to diagnose a patient disease caused by the coro-navirus. The model is convenient for analyzing whether aperson should be isolated due to the COVID-19. Severalstudieshavebeenperformedtogathertrainingdatafromthepresent pandemic and to develop a precise prediction modelforthepandemic.[61]

SpeedingupResearchandDrugDiscovery

AmazonWebServices(AWS)developedanewsearchengine,CORD-19,[64]usingMLandNLPthatcanhelpscien-tistsandresearcherssearchhugevolumesofresearcharticlesanddocumentsquickly.InCalifornia,anAI-basedcompany,C3.ai,[65] recently founded a research consortium aimed attackling COVID-19 by utilizing AI and related strategies infinding the virus spread, forecasting its evolution, developingnew drugs, and fighting future outbreaks.[66] AI and ML canidentify and predict viral proteins of 2019-nCoV for drugdevelopment, support that can speed up the discovery ofmedications and vaccines to treat COVID-19.[54,55,67] An AIcompany in the United Kingdom, BenevolentAI,[68] aimed atdiscovering drugs and immediate medicines for COVID-19by employing AI, DL, and other cutting-edge technologies.EvenbeforetheCOVID-19outbreak,AIhadpotentialcontri-butionstohealthcarefornewdrugdiscovery.[69,70]Generally,AI is still in its initial stages in this area, and the predictionaccuracy of ML methodologies, together with more real-worldrelevance,willcontinuetodevelop.

PublicHealthManagementandServices

SeveralAI-basedactivitieshavethepotentialtocontributetomanagement of the pandemic, such as helping doctors tomonitorhealthcrisesandhandlemultiplepatientsinhospitals,scanninginfectedpeopleinpublicspacesusingthermalim aging, measuring social distancing and lockdown proce-dures,andprovidinghealthconsultationservicestoraiseaware-ness around the world. Also, AI-based systems and virtualreality technology provide social services for COVID-19patients and help to educate others. Such systems, therefore,illustrate how society and public administrations can benefitfromtheuseofAItechnology.AI-baseddronesareusedforsafe and fast delivery of medical supplies in critical zones[71];AI-enabledrobotsareusedforcleaning,sterilizing,anddeliver-ingfoodandmedicinetoavoidhuman-to-humancontact;andML-based UVD robots using ultraviolet light can disinfectandautonomouslykillbacteriaandvirusestolimitthespreadof coronavirus. The major threats against the safety of thepeople and AI devices as well as their interference with airtrafficareneededtobeaddressed.

RelatedDatasets

- Data is critical in assisting the public, governmental, and health-care sectors in combating the COVID-19 pandemic.
- • GISAID data72 from all influenza viruses and the coronavirus, including virus genomic sequences and linked medical and epidemiological data, was utilised to predict protein structures and RNA sequencing.
- 73
- • A Coronavirus genome sequence dataset in the Kaggle repository74 was utilised for RNA sequencing of a sample of bronchoalveolar lavage fluid from a patient identified by 2019-nCoV, a novel RNA virus species from the Coronaviridae family.
- 73
- The CORD-19 dataset, comprised of over 29 000 research papers on COVID-19, SARS, Middle East respiratory syndrome (MERS), and allied viruses in the Kaggle repository,75 was used to extract the most relevant disease information.
- • The daily COVID-19 case count files are provided by the GitHub coronavirus repository76, and all data

DOI:10.9756/INTJECSE/V14I5.809 ISSN: 1308-5581 Vol 14, Issue 05 2022

operations are vectorized, allowing users to generate new CSV, JSON, or Pickle files.

- • COVID-19 Korean dataset77 from Korea, utilised for data visualisation with features such as displaying infected patient paths and regional patient count.
- • COVID-19 image dataset78, a public dataset in the GitHub repository, contains chest x-rays and CT scans of patients with COVID-19 or other viral and bacterial pneumonias like MERS, SARS, and acute respiratory distress syndrome. Data is collected from various open sources, clinics, and physicians and used to create AI-based applications for understanding and detecting contamination.
- • Coronavirus Tweets dataset in Kaggle repository79 contains Twitter data from the hashtags #coronavirus, #coronavirusoutbreak, #coronavirusPandemic, #covid19, #epitwitter, and #ihavecorona and is used to anticipate coronavirus outbreaks.
- • COVID-19 coronavirus dataset in European Union Open Data Portal80 contains public information about the COVID-19 pandemic, including daily updates and global cases. The European Center for Disease Prevention and Control released the dataset.
- • COVID-19 pandemic data in the Hum Data repository81 demonstrates global COVID-19 data with confirmed cases and deaths in locations with humanitarian response plans; data is provided from the WHO..

ArtificialIntelligenceApproachesinHealthCare

AImethodscanbecategorizedinto2mainclassesinhealthcareapplications:ML,includingDL,andNLPapproaches.

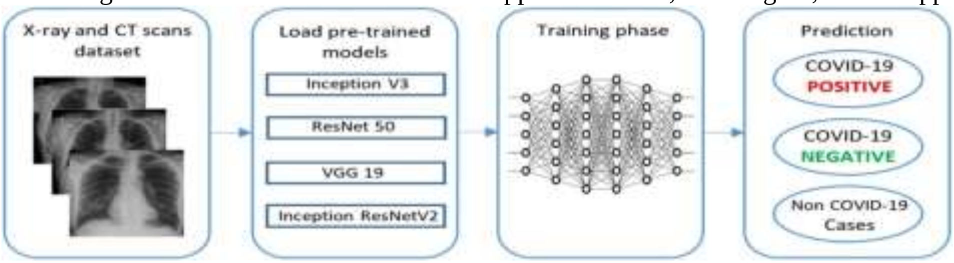

Figure3.DL-basedframeworksforCOVID-19detectionanddiagnosis.Abbreviations:DL,deeplearning;COVID-19,coronavirusdisease.

MachineLearningModelsforDetectionandDiagnosis

Image analysis for medical data feature selection, classification, segmentation, and lesion identification. 97–100 DL models are used to assess medical imaging and signal data such as x-ray, CT, and magnetic resonance imaging. These findings have a significant impact on the detection, identification, and diagnosis of a variety of disorders, including cancer, brain tumours, diabetes, and pneumonia. 50,101–105 Recently, reverse transcription-polymerase chain reaction has become a common test method for detecting and classifying respiratory viruses. 106 Although this standard method of diagnosis detects COVID-19, computer-aided ML approaches aid in speedier identification and diagnosis. As a result, speedy and cost-effective COVID-19 detection solutions rely on AI and ML-based algorithms. 107,108 for coronavirus clinical image processing 109–113 Most studies employ x-ray and CT scan pictures as input to ML and DL prototypes for early diagnosis of virus-infected cases. 49,50,114,115

Several pretrained DL models have been published for detecting COVID-19-infected patients using chest x-ray and CT scan pictures, including ResNet50, InceptionV3, VGG-19, and Inception ResNetV2.

COVID-19 patients' imaging data were gathered for the experiment from the GitHub78 and Kaggle87 public sources. The residual neural network (ResNet) architecture is a modified version of the CNN model116, and ResNet50 is a 50-layer neural network-based model trained on the ImageNet dataset that employs a technique of skip links between levels known as residual learning. 89 InceptionV3,117, a CNN model with a fully connected neural network, enhances the usage of managed resources within the network. Google's Xception118 CNN model was an upgraded version of the Inception model. VGG-19,119, which consists of 19 layers of deep neural network architectures with low convolution filters, was designed to achieve high accuracy in health care applications with large amounts of visual data. 120 Another pre-trained architecture includes a deep neural network using the Inception ResNetV2 model, which was trained on the patients' image dataset and supplied the output as a list of predicted disease class probabilities. 121 Figure 4 depicts a schematic depiction of standard CNN architectures for COVID-19 detection and diagnosis, including pre-trained ResNet50, InceptionV3, VGG-19, and Inception ResNetV2 architectures. It offers three forms of detection, including COVID-19(+), COVID-19(), and others. This model assists health professionals in determining which instances should be studied using normal methods and which management measures should be used to differentiate COVID-19 and non-COVID-19 patients..

DOI:10.9756/INTJECSE/V14I5.809 ISSN: 1308-5581 Vol 14, Issue 05 2022

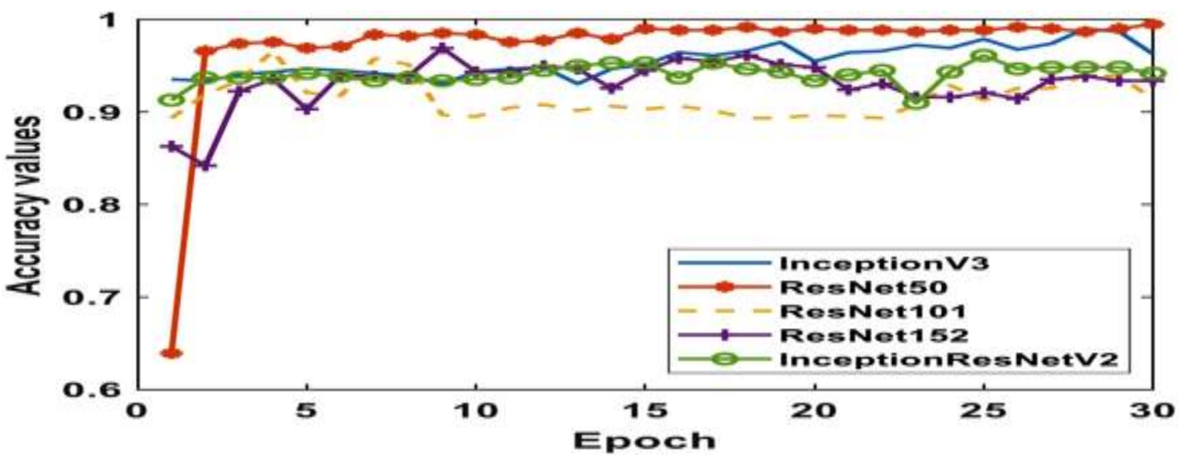

Figure4Thecomparisonplotsoftestingaccuracyofthe5pre-trainedconvolutionalneuralnetwork(CNN)-basedmodels.112 patients using the chest x-ray images, and this approach achieved an accuracy rate of 97.56%.The pretrained ResNet50 proposed by Song et al[123] achieved a success rateof86%usingCTimages.AnotherstudyproposedbyXuetal[124]achievedanaccuracyof86.7%inidentifyingCOVID-19 cases utilizing the ResNet model. Wang et al[49]usedCTimagesinthemodifiedinceptionmodelandachieveda sorting precision of 82.9%. In the study,[125] a deep neuralnetworkmodelwasimplementedthatofferedtheprecisediag-nosticsformulticlustering,suchasCOVID-19(+),pneumonia,andnone,usingchestx-rayimages.Thisapproachappliedthe combined ResNet50–support vector machine (SVM) models andperformedtheclassificationwith95.38%accuracy.

Although ML-based imaging approaches show a significantrole in the analysis of COVID-19 cases, the irregularities indatasets remain the prime challenge in COVID-19 diagnosisfromimagedata.

Machine Learning and Natural LanguageProcessingModelforClinicalDecisionMakingandManagement

Datacollectedfromvarioussourcesarenotalwaysstructuredandrequire NLPproceduresto extractdatafeaturesfromdisor-ganized data sources to enrich structured medical data.[38]ML-basedapproachesarethenappliedtoanalyzestructuredclinicaldata,providetheclusteringofpatients'reports,andsub-sequentlyassesstheprobabilityofdiseaseoutcomes.[39]Thus,NLPproceduresgeneratemachine-readablestructuredandclus-tereddatafromtexts,andthentheclustereddatacanbeana-lyzed by ML techniques to provide outputs, as shown inFigure7.[126]Thefigureillustratestheprocessflowofstructureddata generation from clinical documents, including images,genetic data, electronic medical records, and electrophysiologi-caldata,throughNLPandML,tosuggestclinicalactivities.

AI-based models can be applied to solve clinical problemsandtoassistmedicalpracticesinprimarydetection,identifi-cation,diagnosis,treatment,andmanagementofcriticalCOVID-19 cases.[127–129] Another AI-based model is showninFigure8.DL(ordeepneuralnetwork)approachescouldbe useful tools for the differential diagnosis of COVID-19,andthegenomicvariantsfromnormalorsevereCOVID-19 patients can be classified by the ML analyzer topredict potential COVID-19-infected patients.[53] This modelcouldalsobeusedtomonitorcriticallyillpatientsandassist physicians in decision making about further therapy.Although ML models provide significant improvements inclinicaldecisionmaking,alargevolumeofstudiesisneededtoensuretheseoutcomesandtoincreasetheiraccep-tanceinhealthcaresectors.

DOI:10.9756/INTJECSE/V14I5.809 ISSN: 1308-5581 Vol 14, Issue 05 2022

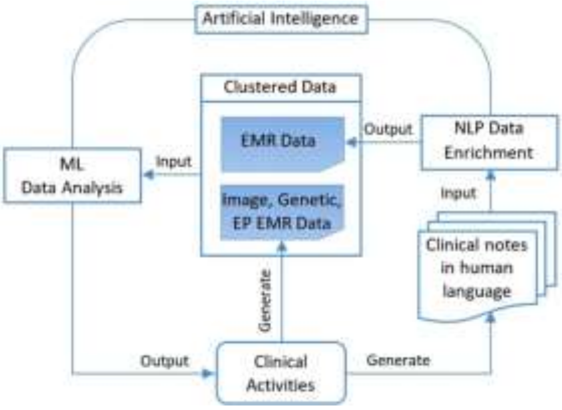

Figure5.AI-basedclinicalmodels,includingNLP-baseddataenrichmentandML-baseddataanalysis.[122]
Abbreviations:AI,artificialintelligence;NLP,naturallanguagepro-cessing;ML,machinelearning.
MachineLearningforAcceleratingDrugDevelopment
Because there are no specific COVID-19 treatments available, it is critical to find quick and efficient solutions to reduce COVID-19 infection and virus incidence. Magar et al133 presented an AI-based approach that integrates big data and medical expertise with ML to identify anti-body sequences that can prevent coronavirus progression. 106 Figure 10,133 depicts a visual diagram of the proposed model, which consists of four primary modules: training dataset, feature extraction, learning, and selecting hypothetical candidates. This ML model was trained using a dataset of 1831 antigen and antibody sequences from various viruses (including H1N1, Dengue, SARS, Ebola, and HIV) from the CATNAP tool91 and 102 samples from the RCSP protein data bank92, for a total of 1933 samples.

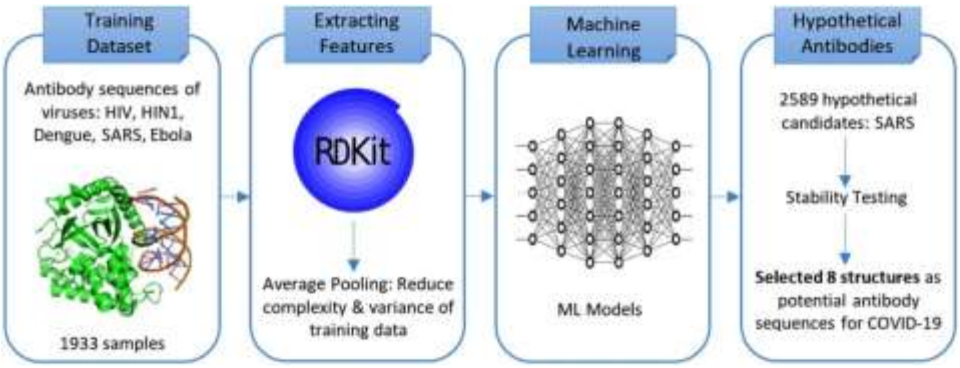

Figure6.Adata-drivenframeworkfordiscoveringantibodysequencestotreatcoronavirusdisease(COVID-19).
And sequence features against SARS, and among these candi-dates, eventually, 8structures were taken as prospective anti-body sequences forneutralizing COVID-19.
Because COVID-19 is caused by a new coronavirus, a thorough understanding of viral architecture aids in therapeutic development. Developing a vaccine for a new disease is costly and time-consuming; ML techniques can help to expedite the entire process. 136 As previously noted, ML combined with NLP approaches enables rapid processing of large amounts of medical data and pulls crucial information from the data, allowing for novel structural predictions and the repurposing of existing medications. The study and its findings could aid in the development of effective COVID-19 medications and vaccines. AI aids in the acceleration of the entire pharmaceutical process, from disease prediction to therapeutic discovery (Figure 11). 55 ML and DL methods have the ability to help in drug development and biological research, but some limitations should be noted, such as expertise about selecting optimal models, the requirement for huge training datasets, and high training computational costs.

DOI:10.9756/INTJECSE/V14I5.809 ISSN: 1308-5581 Vol 14, Issue 05 2022

Name of the Algorithms	Averages	Precision	Recall	F1- Score	Support
K-NN	Macro. Average.	0.74	0.65	0.64	114
	Weighted Average	0.73	0.70	0.67	114
Logistic Regression	Macro Average	0.29	0.50	0.37	114
	Weighted Average	0.35	0.59	0.44	114
Decision Tree	Macro Average	0.94	0.95	0.95	114
	Weighted Average	0.95	0.95	0.95	114
Random Forest	Macro Average	0.96	0.96	0.96	114
	Weighted Average	0.96	0.96	0.96	114
Support Vector Machine	Macro Average	0.29	0.50	0.37	114
	Weighted Average	0.35	0.59	0.44	114
Naïve Bayes	Macro Average	0.63	0.51	0.41	114
	Weighted Average	0.62	0.60	0.47	114

Table1. Results of Prediction, Recall, F1-Score While Using Various algorithms with Breast cancer Dataset

Measure	**Formula**
Accuracy, Recognition Rate	TP+TN ……… P+N
Error, Misclassification Rate	FP+FN. ………. P+N
Sensitivity, True Positive Rate Recall	TP…………. P
Specificity, True Negative Rate	TN…………. N
Precision	TP………. TP+FP
F, F1, F-Score, Harmonic Mean of Precision and Recall	2*Precision*recall…… Precision + Recall

Table2. Problem Solving With Help of the Measurement and Formula of the Breast Cancer Prediction

DOI:10.9756/INTJECSE/V14I5.809 ISSN: 1308-5581 Vol 14, Issue 05 2022

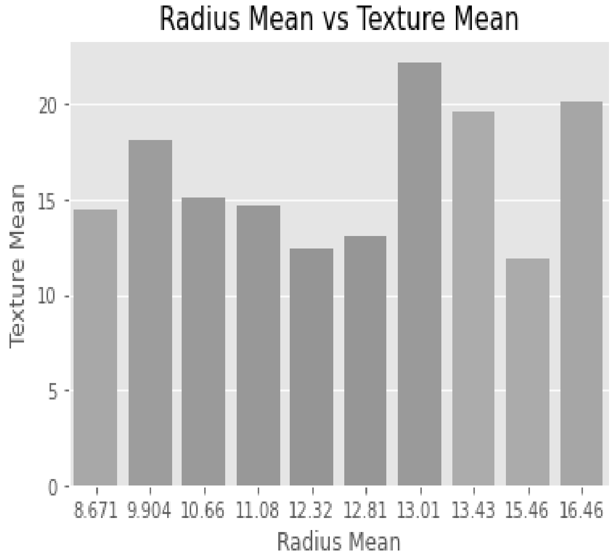

Fig.3 Radius Mean And Texture Mean help BCDS.

EXPERIMENTAL RESULT

Fig. 3 Impact of the test data on the diagnostic result to observe and correlation between attributes

The goal of this research project on Machine Learning Algorithms on Breast Cancer Data is to better understand the impact of test data on diagnostic outcomes as well as the relationship between attributes. Set both the accuracy and training score to 100%. who have everlasting breast cancer and have not activated the algorithm based on the computation located among the KNN, SVM, N.B., Random Forest, Logistic Regression, and Decision Tree approach used in this proceeding? Random Forest has the highest accuracy of 96.61 percent, while Decision Tree has the lowest accuracy of 94.73 percent. Random Forest has a 100% Training Score, while Decision Tree has a 99.34 percent Training Score[10].

Name of Algorithms	Training Score in %	Accuracy in %
K-Nearest Neighbor	80.21%	70.17%
Support Vector Machine	63.73%	58.77%
Naïve Bays	63.29%	56.29%
Random Forest	100.00%	96.49%
Logistic Regression	63.73%	58.77%
Decision Tree	99.34%	94.73%

Table.4 Find out of the Training Score and Accuracy Help of Machine Learning Algorithm.

CONCLUSION

The Information Cultivate system is used to examine the region. The study employed machine learning techniques such as Random Forest and Decision Tree, Support Vector Machine, K-Nearest Neighbor, Logistic Regression, and Naive Bayes. To predict which people with recurrent breast cancer regret infection and which do not, a Random Forest and Decision Tree technique is applied. According to the most recent statistics, the random forest classifier outperforms other models in terms of precision and execution time in predicting the best show outcome.

REFERENCES

1. Nikita Rane, Jean Sunny, Rucha Kanade, Prof. Sulochana Devi, Breast Cancer Classification And Prediction Using Machine Learning vol.09 Issue 02, February-2020.
2. Md. Milon Islam, Hasib Iqbal, Md. Rezwanul Haque and Md. Kamrul Hasan, Prediction Of Breast Cancer Using Support Vector Machine And K-Nearest Neighbors, 2017 IEEE Region 10 Humanitarian Technology Conference R-10-HTC.
3. Dr. Stephanie J Seidler, Dr. DE. Huber, Overview of Diagnosis and Treatment of Breast Cancer in Young Women. Gynecology 2020.
4. MeriemAmrane, SalihaOukid, Ikram Gagaoua, TolgaEnsari, Breast Cancer Classification Using Machine Learning78-1-5386-5135-3/18 IEEE.
5. Chuck Easttom, Sudip Thapa, Justin Lawson, A Comparative Study Of Machine Learning Algorithms For Use In Breast Cancer Studies, 978-1-7281-3783-4/20 IEEE.
6. Dr. MukteviSrivenkatesh, Prediction of Breast Cancer Disease Using Machine Learning, International Journal of Innovative Technology and Exploring Engineering, Volume-9 Issue-4, February 2020.
7. KNN With Different Distance Measures And Classification Rules Using Machine Learning, European Journal of Molecular and Clinical Medicine volume07, Issue 03, 2020.
8. “Breast Cancer Wisconsin (Diagnostic) Data Set”, Sep 2019, [online] Available:
9. UCI, “Breast Cancer Wisconsin (Diagnostic) Data Set,” Kaggle, 25-Sep 2016. [online] Available: https://www.kaggle.com/uciml/breast -cancer Wisconsin - data.
10. “U.S Breast Cancer Statics”, breastcancer.org, Sep 2019,[online] Available: https://www.breastcancer.org/symptoms/understand_bc/statistics/.