Machine Learning Approaches For Disease Prediction:– A Review

Conference Paper · June 2022

DOI: 10.1109/AIC55036.2022.9848838

4 authors, including:

Sudha Yadav
Amity University
7 PUBLICATIONS **46** CITATIONS
SEE PROFILE

Yudhvir Singh
Maharshi Dayanand University
104 PUBLICATIONS **746** CITATIONS
SEE PROFILE

Vivek Jaglan
74 PUBLICATIONS **779** CITATIONS
SEE PROFILE

2022 IEEE World Conference on Applied Intelligence and Computing (AIC) | 978-1-6654-7988-2/22/$31.00 ©2022 IEEE | DOI: 10.1109/AIC55036.2022.9848838

Machine Learning Approaches For Disease Prediction:- A Review

Sudha
Computer Science & Engineering
University Institute of Engineering and Technology, Maharshi Dayanand University, Rohtak, India
sudhayadav.91@gmail.com

Harkesh Sehrawat
Computer Science & Engineering
University Institute of Engineering and Technology, Maharshi Dayanand University, Rohtak, India
sehrawat_harkesh@mdu.ac.in

Yudhvir Singh
Computer Science & Engineering
University Institute of Engineering and Technology, Maharshi Dayanand University, Rohtak, India
dr.yudhvirs@gmail.com

Vivek Jaglan
Computer Science & Engineering
DPG Insititute of Technology and Managment
Gurgaon, India
jaglanvivek@gmail.com

Abstract: **Over recent years, disease prediction catches the attention of researcher's awareness to cover a large range in medical as well as computer science field. Therefore, several models have been constructed for many different-different diseases diagnose and their forecasting. These models utilise an assortment of patient features to assess the likelihood of results over a definite interval of time and have capability to make better decision making. Patients' health database contain large amount of information regarding particular disease and several laboratory test results. It has become essential to discover hidden patterns from those longitudinal health-related databases, and machine learning algorithms are playing a vital role to achieve this task. These algorithms assure the superior accuracy of observation and identification of disease. This paper highlighting various diseases, whose diagnose and prediction have been done through machine learning algorithms. It conveys concentration in the direction of machine learning algorithms and attributes that are used for the prediction of diseases and decision-making process accordingly.**

Keywords— machine learning (ML), disease prediction, prediction models, SVM (support vector machine), NB (Naïve bayes), DT (Decision tree), RF (regression, random forest), ANN (artificial neural network.

I. INTRODUCTION

Artificial Intelligence, as its name indicates human intelligence which can be implemented by machine. In the field of computer science, it means the capability of machine to imitate clever nature by itself, using machine learning [1, 2]. AI makes computer able to think and greatly more intelligent. Machine learning lies within AI as it is a subpart of AI field. [3]. Machine learning (ML) is an important technology which is scattering in numerous computer science engineering and application areas, as well as has been frequently used in image processing, pattern recognition, natural language processing (NLP), cyber security, medical science and several other areas. In medical science, healthcare system is one of the most imperative research applications of ML [4]. Patients' health-related database contain large amount of information. Healthcare firms maintain these longitudinal databases and researchers can use these databases to extract valuable information as per need [5]. Sometimes patients not able to describe their medical issues precisely and on the basis of that result of laboratory tests can direct to some extent of error. Physicians can't identify the disease accurately as per patient description and they may not have knowledge of all areas. To deal with such issues, it is essential to build up a disease diagnose or prediction system that unites medical information with computer science technology to construct the major outcomes and can help society [6]. To get better quality of medical data, save in medical costs, and decrease fluctuations in patient rates, ML models are useful to achieve all these things. Therefore, to investigate diagnostic analysis, these models are frequently used as compared with other conventional methods.

II. MACHINE LEARNING AND ITS TYPES

ML is a means to convert information into valuable knowledge. Patient health records consist of large amount of data, which have some hidden patterns inside it. ML methods are useful to evaluate and interpret those numerous hidden patterns to predict future outcomes and can be helpful for better decision making. Categories of machine learning algorithms which are useful for diseases prediction are representing through figure 1.

A. Supervised Learning:

As its name indicates some supervision is needed to predict future events. Supervised learning use labeled data (input and output variables). Dataset of input parameters is partitioned into two portions: training and testing. Algorithms learn through training process which involves mapping of input and output variable. After completion of training phase, testing dataset applied to algorithms to predict future outcomes and to assess the performance of ML approaches. Decision tree (DT), Regression, Random forest (RF), Support vector machine (SVM), Artificial neural network (ANN), Bayesian are examples of supervised machine learning [3,7].

978-1-6654-7988-2/22/$31.00 ©2022 IEEE
DOI: 10.1109/aic.2022.114

B. Unsupervised Learning:

This learning type contains unlabeled data (made up of only input variables without any output variables), and the algorithm tries to examine the intrinsic constitution of data from input parameters. Main goal of algorithms is to learn the underlying distribution of data. The unsupervised learning methods mainly work for clustering of data and dimensionality reduction problems. Examples are K-means, Hierarchical clustering, Fuzzy C means (FCM), Principle component analysis (PCA) [3, 7].

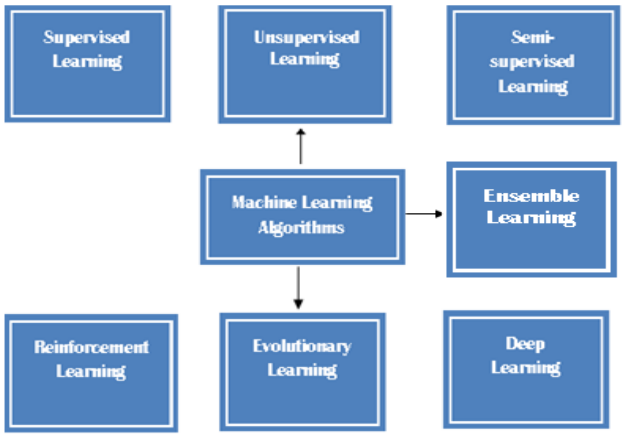

Figure 1: Different Modes of Machine learning

C. Semi-supervised Learning:

This learning uses both (labeled and unlabeled) types of data. For training purpose, partial labeled data (typically a fewer quantity of labeled-data with an enormous quantity of unlabeled-data) is used. Generative Model, Self-Training, Transductive SVM are some examples of semi-supervised learning [7].

D. Reinforcement Learning:

This learning approach contains a problem and a solution, where problem refers to an object and solution refers to an algorithm. Algorithm will update if outcome is invalid, but does not notify that how to correct it. It will discover and assess diverse prospective until it discovers the correct outcomes. It will revise its information on the basis of previous state and response attained from the problem. Q-learning method is an example of reinforcement learning [3].

E. Evolutionary Learning:

This learning type stimulated by biological evolution and environment. The common process begins with an unsystematic set of solutions. These solutions will go through a no. of repetitions for improvement and finally new solutions will generates. The key role is to select the best options in ever repetition based on fitness function. When optimal solution is achieved then this recursive process will stop. Genetic algorithm (GA) is an example of evolutionary learning [3].

F. Deep Learning:

This division of ML is based on the collection of algorithms. These learning algorithms correspond to multiple processing layers and have advanced abstraction. Deep graph with several processing layers, constructed with many linear and nonlinear transformations, is used in it. This learning method determines complicated arrangement in longitudinal data sets with the help of back-propagation algorithm to specify that how a machine should transform its inside parameters that are used to compute the representation in each layer from the representation in the previous layer [8].

G. Ensemble Learning:

Ensemble learning is also known as hybrid approach of learning. In this learning type, a no. of different-different algorithms is combined to form a single learning method. These individual algorithms may be Decision tree, Naïve bayes, Neural network, Support vector machine, K-means etc. Bagging and Boosting are two popular ensemble learning techniques [7].

III. DISEASES PREDICTED USING MACHINE LEARNING ALGORITHMS

A lot of diseases have been predicted using machine learning techniques to improve healthcare system and to help physician for identification and diagnosis of diseases. A variety of models have been developed by researchers for detection and risk prediction of diseases using algorithms of different modes of machine learning. Table 1 presents various diseases which have been predicted through machine learning algorithms.

A. Breast cancer:

Sharma et al. [9] constructed a prediction model for breast cancer disease. In this paper, three ML algorithms RF, K-nearest neighbours (K-NN), Naïve bayes (NB) were used to detect disease through mammography of patients. Among all three approaches K-NN gave highest detection accuracy 95.90% when used with 10 fold cross validation method. Ganggayah et al. [10] focused on factors prediction for breast cancer patients using six classifiers and compared their performance also. Gupta et al. [11] presented prediction model using K-NN, SVM, Logistic Regression (LR), DT, RF and Deep learning. Authors used 30 features of patients details and mammography as database. Deep learning technique using adam gradient descent learning gave best accuracy 98.24% result.

B. Asthma:

Finkelstein et al. [12] has executed a work for early prediction of asthma attack. Three machine learning classifiers Adaptive bayesian network (ABN), SVM, NB

used for this work. In terms of performance measures, sensitivity of NB, ABN and SVM was 0.80, 1.00, and 0.84, and specificity 0.77, 1.00, and 0.80 respectively. Wang et al. [13] utilized the deep learning technique ANN to predict asthma-related emergency visits. Yang et al. [14] presented a novel deep learning time-sensitive attentive neural network (TSANN) to make better prediction for asthma exacerbation risk. A large size electronic health records of the patient (age 18-80 yr). As compared to other existing models, this proposed work achieved AUC value 0.7003.

C. Brain cancer:

Aaswad et al. [15] has presented a work of brain cancer detection using machine learning approach (tenserflow). A database of MRI images of patients was used to detect brain cancer. Vijayakumar, et al. [16] utilized a novel machine learning technique, capsule neural network for classification of brain cancer using database of MRI images. Training accuracy 98.4% and testing accuracy 95.3% achieved through this proposed work. Brunese et al. [16] used ensemble learning method for detection of brain cancer from a database of MRI images. This method attained 99% accuracy.

D. Liver:

Mafazalyaqeen, et al. [17] proposed a work for liver disease prediction using genetic algorithm with boosted C5.0 classification technique. This work attained 92.93% accuracy. Vasan et al. [18] has predicted liver disease using machine learning approaches. NB and SVM two classifiers used for prediction purpose.

E. Heart Disease:

Miao et al. [20] has presented work for coronary heart disease diagnosis using ensemble learning with adaptive boosting technique. Higher testing accuracy attained 96.72% with SUH dataset. Haq et al.[21] used six different classifiers LR, SVM, ANN, DT, NB, K-NN and RF for construction of a hybrid intelligent system for heart disease prediction.

F. Infection:

Parreco et al. [23] worked for prediction of bloodstream infections and mortality using supervised ML methods. Three classifiers deep learning , LR, and gradient boosted trees used for prediction along with a number of attributes. LR gave highest AUC of 0.722 ±0.048. Taylor et al. [24] has predicted infections (urinary tract) in the emergency department with machine learning using six classifiers extreme gradient boosting (XGboosting), RF, adaptive boosting (ADP), SVM, NN, and LR along with a good selection of several clinical attributes.

G. Diabetes:

Perveen et al. [25] authors implemented machine learning techniques for prognostic modelling and prevention of diabetes. Three machine learning methods Naive Bayes, SVM and DT used for this purpose. Analysis work presented that Naive Bayes gave better performance with the maximum accuracy (76.30%). Suyash et al. [26] executed ANN technique which gave 92% accuracy.

H. Kidney Disease:

Anusorn et al. [27] utilized four supervised ML algorithms for predictive analysis of chronic kidney disease. 24 clinical attributes were considered for prediction work. On the basis of performance analysis, highest accuracy 98.3% of proposed work attained through SVM. Alasker et al. [28] focused on detection of chronic kidney disease using five different ML methods i.e. KNN, Back propagation neural network, DT, NB, one class classifier along with large number of clinical attributes. Almasoud et al. [29] executed four machine learning approaches LR, SVM, RF, Gradient boosting for prediction of kidney disease and compare their performance. 99.1% accuracy achieved through gradient boosting.

I. Lung Cancer:

Monsi et al. [30] has implemented Deep Convolutional neural network (CNN) as ML technique for estimation of lung disease. A database of patients chest X-ray was used for prediction work. Delzell et al. [31] used different classifiers of linear, nonlinear and ensemble types of learning for predicting and extracting relevant features of lung cancer from a database of CT scan of patients chest.

J. Alzheimer:

Lodha et al. [34] has constructed model for diagnosis of disease. Five ML classifiers SVM, Gradient boost, KNN, RF, NN used for this purpose and compare their performance also. Imaging database was used for extracting relevant features for diagnosis of Alzheimer disease. 98.36% was highest accuracy achieved through NN. Neelaveni et al. [35] executed only two classifiers SVM, DT for prediction of alzheimer disease. Based on comparison, 85% accuracy attained through SVM.

K. Anemia:

Jaiswal et. al [37] worked with three different machine learning classifiers NB, RF and DT for anemia disease prediction and compared their performance also to find out a best classifier for this work. Complete blood count results were as features for anemia prediction; based on performance comparison, NB gave best accuracy results 96.09%. Khan et al. [38] used 5 ML algorithms linear discriminant analysis (LDA), CART, k-NN, SVM, and LR for prediction of childhood anemia.

L. Thyroid:

Tyagi et al. [39] has focused on thyroid disease prediction using machine learning methods. Four ML methods ANN,

KNN, SVM, DT have used in this work in which SVM attained highest accuracy 99.63% as compared to other methods. Yadav et al. [40] has used DT, RF, classification and regression tree (CART), bagging ensemble method.

M. Parkinson:

Mathur et al. [41] has constructed a forecasting model for Parkinson disease using various ML algorithms (SVM, ANN, KNN, MLP) that helped to enhance the performance of datasets and played a critical function to make the early estimation of disease at correct time. Performance analysis demonstrated that combination of KNN with MLP gave better accuracy 91.38%. Senturk et al. [42] presented a diagnosis model for Parkinson disease. CART, SVM and ANN are three machine learning algorithms used for classification.

N. Alopecia:

Kapoor et al. [43] implemented computerized categorization method for early examination of alopecia using ML. ANN method was used for this purpose. Accuracy achieved 91%. Solam et al. [44] performed cluster analysis for development of a prediction model with the grading system.

O. Gastric Cancer:

Junichi et al. [45] executed XGBoost for calculation of upcoming risk associated with gastric cancer. Xgboost gave highest AUC value (0.899). Peng et al. [46] has executed ANN, RF, Adaboost, Gradient Boost Machine (GBM), DT and GBM framework—Light GBM ML methods for gastric cancer risk prediction.

P. COVID-19:

Gomathi et al. [47] implemented AutoML (Auto- machine learning) approach for prediction of COVID-19 pandemic in India. 98.6% accuracy achieved in this work. Alakus et al. [48] has presented his work for prediction COVID-19 infection. Deep Learning method were executed for this.

TABLE I. SEVERAL DISEASE PREDICTIONS DONE BY MACHINE LEARNING TECHNIQUES

Predicted disease	Authors and Published year	Machine Learning Methods	Attributes used	Accuracy
Breast Cancer	Sharma et al. [9] 2018	RF, K-NN, NB	Mammography	95.90% by K-NN
	Ganggayah et al. [10] 2019	DT, RF, LR, NN, SVM, Extreme gradient boost	23 attributes like marital status, stage of cancer, diagnose method etc.	82.7% by RF
	Gupta et al. [11] 2020	K-NN, SVM, LR,DT, RF and Deep learning	Mammography (30 features)	98.24% by deep learning using
Asthma	Finkelstein et al. [12] 2016	ABN, SVM, NB	Cough, Wheeze, Chest tightness, Physical activity, Sputum production, Use of inhaler, Shortness of breath, Cold, Estimation of asthma/day, Medicine, Sleep disturbance	Highest sensitivity and specificity achieved through ABN was 1.00 and 1.00
	Wang et al. [13] (2019)	Deep learning method ANN	Demographic data and clinical visits	AUC of ANN was 0.845
	Yang et al. [14] (2020)	A novel time sensitive alternative neural network (TSANN)	Demographic data and clinical visits	AUC value of proposed model was 0.7003
Brain Cancer	Aaswad, et al. [15] 2018	CNN with 5 layers	MRI images	Training accuracy 99% and validation accuracy 98%
	Vijayakumar, et al. [16] 2019	Capsule neural network	MRI images	Training accuracy 98.4% and testing accuracy 95.3%
	Brunese et al. [17] 2020	Ensemble learning	MRI images	Attained 99% accuracy
Liver	Mafazalyaqeen, et al. [18] 2017	Genetic algorithm	Age, Sex, TB, Alkphos, Sgpt Alamine, DB, TP, Sgot Aspartate, A/G Ratio, ALB	Accuracy achieved 92.93%
	Vasan et al. [19] 2019	NB and SVM	TB, DB, TP, A/G proportion, Albumin, SGPT, Alkaline Phosphatase and SGOT	Not given
Heart Disease	Miao et. al. [20] 2016	Ensemble learning using adaptive boosting technique	Age, Sex , CP, Htn, Chol, Restecg, Ekgmo, Ekgday etc	Higher testing accuracy was 96.72% with SUH dataset
	Haq et al. [21] 2018	LR, K-NN, ANN, SVM, DT, NB, and RF	Age, Sex, Cardiography, Chest pain, BP, Cholesterol, FBS, Heart rate etc	LR gave highest accuracy 89%
	Latha et al. [22] 2019	Ensemble learning with NB, PART, C4.5, MLP	Age, Sex , CP, Cholesterol, Oldpeak, Slope etc	Proposed model accuracy was 85.48%

Infection	Parreco et. al. [23] 2018	LR, Gradient boosted trees, and Deep learning	Paralysis , COPD, Diabetes, Obesity, Weight loss, Liver disease, Renal failure, AIDS, Electrolyte disorder etc	LR gave highest AUC of 0.722 ±0.048
	Taylor et al. [24] 2018	RF, XGboosting, ADP, SVM, NN, and LR	Demographic information, Vitals, Past medical history, Laboratory results, Structured historical and Physical exam results, Medications, Chief complaint	87.5% by XGboost
Diabetes	Perveen et al. [25] 2019	DT, SVM and NB	Age, Number of times pregnant, Class, Diastolic blood pressure, Plasma glucose concentration, Skin fold thickness , BMI, Diabetes pedigree function, 2-Hour serum insulin	NB outperforms with the highest accuracy of 76.30%
	Suyash et al. [26] 2019	ANN	Same as Perveen et al.[25]	Accuracy achieved 92%
Kidney Disease	Anusorn et al. [27] 2016	K-NN, SVM, LR, DT	Clinical data (24 attributes)	accuracy (98.3%) by SVM
	Alasker et al. [28] 2017	Back propagation neural network, K-NN, DT, NB, one class classifier	Demographic and clinical attributes	NB produces 99.36%
	Almasoud et al. [29] 2019	LR, SVM, RF, Gradient boosting	Clinical data	99.1% accuracy using gradient boosting
Lung Disease	Monsi et al. [30] 2019	Deep CNN	Chest X-ray images	Not given
	Delzell et al. [31] 2019	Different classifiers of linear, nonlinear and ensemble types	CT scan	AUC (0.747) of elastic net was higher
Stress	Disha et al. [32] 2020	NB, LR, MLP, Bayes Net, J48 and RF	Age, Sex, Weight, Family issues, Emotions etc	Bayes Net gave 88.59% accuracy
	Prince et al. [33] 2020	NB, BN, K-star, LNN, MLP, RBFN, RF and J48	Basic attributes collected through questionnaires	RBFN gave highest accuracy 96.17%
Alzheimer	Lodha et al. [34] 2018	SVM, Gradient boost, K-NN, RF, NN	Imaging data	98.36% by NN
	Neelaveni et al. [35] 2020	SVM, DT	Age, No. of visits, Education, MMSE (Mini Mental State Examination) etc	85% by SVM
Sickle cell disease	DhafarHamed et al. [36] 2017	MLP, SVM, RF, and LR	Demographics data and clinical data	sensitivity 96% and specificity 98%, using MLP
Anemia	Jaiswal et al. [37] 2019	NB, RF, DT	Complete blood count(CBC)	NB gave best accuracy result 96.09%
	Khan et al. [38] 2019	Linear discriminant analysis (LDA), CART, K-NN, SVM, and LR	Demographics details and clinical data	Accuracy 68.53% attained by RF
Thyroid	Tyagi et al. [39] 2018	ANN, KNN, SVM, DT	Age , TSH, Sex, TBG, T3, T4U, FTI	SVM attained highest accuracy 99.63%
	Yadav et al. [40] 2020	DT, RF, classification and regression tree (CART), Bagging ensemble method	Age, Sex and many others clinical variables	98%, 99%, and 93% and 100% using proposed bagging ensemble method
Parkinson	Mathur et al. [41] 2019	KNN with MLP	11 features selected from clinical dataset	Accuracy achieved 91.28%
	Senturk et al. [42] 2020	CART, ANN and SVM	22 phonetic features used	93.84% using SVM
Alopecia	Kapoor et al. [43] 2018	Neural network (NN)	Length of hair, Nail brittleness, Damage, Hair Follicle	Accuracy achieved 91%
	Solam et al. [44] 2019	Cluster analysis	Disease duration, Basic details, Family history of AA, History of atopic diseases, Extrascalp hair loss, Primary treatment, and Progress during follow-up	Not given

Gastric cancer	Junichi et al. [45]	XGBoost	Demographic and clinical attributes	Xgboost gave highest AUC value (0.899)
	Peng et al. [46] 2020	Adaboost, Gradient Boost Machine (GBM), DT, ANN, RF, and GBM framework--LightGBM	Patient ID, State-County Age, Race, Sex, Grade Radiation Sequence with Surgery Radiation Recode, Site_Labeled , Summ-Surg, Reginal nodes positive(1988+), Chemotherapy Survival	85% by custom ensemble
COVID-19	Gomathi et. al. [47] 2020	AutoML	age, gender, clinical manifestation, travel and contact with COVID patients	98.6% accuracy
	Alakus et al.[48] 2020	Deep Learning	Laboratory data	86.66% of proposed model

IV. CONCLUSION

The current study examines the several different machine learning algorithms used to design predictive models for 18 human diseases. Analysis present that supervised machine learning algorithms is most widely used with the addition of easy and simple predictive modelling. Along with supervised learning, ensemble learning unsupervised learning, deep learning also implemented by the researchers to improve accuracy of prediction models. In medical science, the execution of these models definitely can facilitate to give enhanced health services and improve physicians decision-making. It's elementary to validate the diverse algorithms on the basis of a specific problem, and re-evaluation process could support to analyze the performance.

REFERENCES

[1] Hamet, Pavel, and Johanne Tremblay. "Artificial intelligence in medicine." *Metabolism* 69 (2017): S36-S40.

[2] Battineni, Gopi, et al. "Applications of machine learning predictive models in the chronic disease diagnosis." *Journal of personalized medicine* 10.2 (2020): 21.

[3] Fatima, Meherwar, and Maruf Pasha. "Survey of machine learning algorithms for disease diagnostic." *Journal of Intelligent Learning Systems and Applications* 9.01 (2017): 1.

[4] Jiang, Min, et al. "A study of machine-learning-based approaches to extract clinical entities and their assertions from discharge summaries." *Journal of the American Medical Informatics Association* 18.5 (2011): 601-606.

[5] Saranya, G., and A. Pravin. "A comprehensive study on disease risk predictions in machine learning." *International Journal of Electrical and Computer Engineering* 10.4 (2020): 4217.

[6] M. Jiang, Y. Chen, M. Liu, S. T. Rosenbloom, S. Mani, J. C. Denny, and H. Xu, "A study of machine-learning-based approaches to extract clinical entities and their assertions from discharge summaries," *J. Am Med Inform Assoc*, vol. 18, no. 5, pp. 601–606, 2011.

[7] Ayon Dey, "Machine Learning Algorithms: A Review" *International Journal of Computer Science and Information Technologies*, Vol. 7 (3) , 2016, 1174-1179.

[8] LeCun, Yann, Yoshua Bengio, and Geoffrey Hinton. "Deep learning." *nature* 521.7553 (2015): 436-444.

[9] Sharma, Shubham, Archit Aggarwal, and Tanupriya Choudhury. "Breast cancer detection using machine learning algorithms." *2018 International Conference on Computational Techniques, Electronics and Mechanical Systems (CTEMS)*. IEEE, 2018.

[10] Ganggayah, Mogana Darshini, et al. "Predicting factors for survival of breast cancer patients using machine learning techniques." *BMC medical informatics and decision making*19.1 (2019): 1-17.

[11] Gupta, Puja, and Shruti Garg. "Breast cancer prediction using varying parameters of machine learning models." *Procedia Computer Science* 171 (2020): 593-601.

[12] Finkelstein, Joseph, and In cheol Jeong. "Machine learning approaches to personalize early prediction of asthma exacerbations." *Annals of the New York Academy of Sciences*1387.1 (2017): 153.

[13] Wang, Xiao, et al. "Deep Learning Models to Predict Pediatric Asthma Emergency Department Visits." *arXiv preprint arXiv:1907.11195* (2019).

[14] Xiang, Yang, et al. "Asthma Exacerbation Prediction and Risk Factor Analysis Based on a Time-Sensitive, Attentive Neural Network: Retrospective Cohort Study." *Journal of medical Internet research* 22.7 (2020): e16981.

[15] Sawant, Aaswad, et al. "Brain cancer detection from mri: A machine learning approach (tensorflow)." *Brain* 5.04 (2018).

[16] Vijayakumar, T. "Classification of brain cancer type using machine learning." *Journal of Artificial Intelligence* 1.02 (2019): 105-113.

[17] Brunese, Luca, et al. "An ensemble learning approach for brain cancer detection exploiting radiomic features." *Computer methods and programs in biomedicine* 185 (2020): 105134.]

[18] Hassoon, Mafazalyaqeen, et al. "Rule optimization of boosted c5. 0 classification using genetic algorithm for liver disease prediction." *2017 international conference on computer and applications (icca)*. IEEE, 2017.

[19] Durai, Vasan, Suyan Ramesh, and Dinesh Kalthireddy. "Liver disease prediction using machine learning." *Int. J. Adv. Res. Ideas Innov. Technol* 5.2 (2019): 1584-1588.

[20] Miao, Kathleen H., Julia H. Miao, and George J. Miao. "Diagnosing coronary heart disease using ensemble machine learning." *Int J Adv Comput Sci Appl (IJACSA)* (2016).

[21] Haq, Amin Ul, et al. "A hybrid intelligent system framework for the prediction of heart disease using machine learning algorithms." *Mobile Information Systems* 2018 (2018).

[22] Latha, C. Beulah Christalin, and S. Carolin Jeeva. "Improving the accuracy of prediction of heart disease risk based on ensemble classification techniques." *Informatics in Medicine Unlocked* 16 (2019): 100203.

[23] Parreco, Joshua P., et al. "Predicting central line-associated bloodstream infections and mortality using supervised machine learning." *Journal of critical care* 45 (2018): 156-162.

[24] Taylor, R. Andrew, et al. "Predicting urinary tract infections in the emergency department with machine learning." *PloS one*13.3 (2018): e0194085.

[25] Perveen, Sajida, et al. "Prognostic modeling and prevention of diabetes using machine learning technique." *Scientific reports*9.1 (2019): 1-9.

[26] Srivastava, Suyash, et al. "Prediction of diabetes using artificial neural network approach." *Engineering Vibration, Communication and Information Processing*. Springer, Singapore, 2019. 679-687.

[27] Charleonnan, Anusorn, et al. "Predictive analytics for chronic kidney disease using machine learning techniques." *2016 management and innovation technology international conference (MITicon)*. IEEE, 2016.

[28] Alasker, Haya, et al. "Detection of kidney disease using various intelligent classifiers." *2017 3rd international conference on science in information technology (ICSITech)*. IEEE, 2017.
[29] Almasoud, Marwa, and Tomas E. Ward. "Detection of chronic kidney disease using machine learning algorithms with least number of predictors." *International Journal of Soft Computing and Its Applications* 10.8 (2019).
[30] Monsi, Justin, et al. "XRAY AI: Lung Disease Prediction Using Machine Learning." *International Journal of Information* 8.2 (2019).
[31] Delzell, Darcie AP, et al. "Machine learning and feature selection methods for disease classification with application to lung cancer screening image data." *Frontiers in oncology* 9 (2019): 1393.
[32] Disha Sharma et al. "Stress prediction of students using machine learning" International Journal of Mechanical and Production Engineering Research and Development (2020) 5609-5620.
[33] Prince Kumar, "Assessment of anxiety, diperession and stress using machine learning".
[34] Lodha, Priyanka, Ajay Talele, and Kishori Degaonkar. "Diagnosis of alzheimer's disease using machine learning." *2018 Fourth International Conference on Computing Communication Control and Automation (ICCUBEA)*. IEEE, 2018.
[35] Neelaveni, J., and MS Geetha Devasana. "Alzheimer disease prediction using machine learning algorithms." *2020 6th International Conference on Advanced Computing and Communication Systems (ICACCS)*. IEEE, 2020.
[36] Abd, DhafarHamed, et al. "The utilisation of machine learning approaches for medical data classification and personal care system management for sickle cell disease." *2017 Annual Conference on New Trends in Information & Communications Technology Applications (NTICT)*. IEEE, 2017.
[37] Jaiswal, Manish, Anima Srivastava, and Tanveer J. Siddiqui. "Machine learning algorithms for anemia disease prediction." *Recent Trends in Communication, Computing, and Electronics. Springer, Singapore*, 2019. 463-469.
[38] Khan, Jahidur Rahman, et al. "Machine learning algorithms to predict the childhood anemia in Bangladesh." *Journal of Data Science* 17.1 (2019): 195-218.
[39] Ankita Tyagi, Ritika Mehra "Interactive Thyroid Disease Prediction System Using Machine Learning Technique"
[40] Yadav, Dhyan Chandra, and Saurabh Pal. "Prediction of thyroid disease using decision tree ensemble method." *Human-Intelligent Systems Integration* 2.1 (2020): 89-95.
[41] Mathur, Richa, Vibhakar Pathak, and Devesh Bandil. "Parkinson disease prediction using machine learning algorithm." *Emerging Trends in Expert Applications and Security*. Springer, Singapore, 2019. 357-363.
[42] Senturk, Zehra Karapinar. "Early diagnosis of Parkinson's disease using machine learning algorithms." *Medical hypotheses* 138 (2020): 109603.
[43] Kapoor, Ishita , Mishra , Anju, "Automated classification method for early diagnosis of alopecia using machine learning." *International Conference on Computational Intelligence and Data Science* (ICCIDS 2018) 437–443.
[44] Lee, Solam, et al. "Topographic phenotypes of alopecia areata and development of a prognostic prediction model and grading system: a cluster analysis." *JAMA dermatology* 155.5 (2019): 564-571.
[45] Taninaga, Junichi, et al. "Prediction of future gastric cancer risk using a machine learning algorithm and comprehensive medical check-up data: A case-control study." *Scientific reports* 9.1 (2019): 1-9.
[46] Liu, Peng, et al. "Two staged prediction of gastric cancer patient's survival via machine learning techniques." *Proc. 7th Int. Conf. Artif. Intell. Appl.*. 2020.
[47] Gomathi, S., et al. "Pattern analysis: predicting COVID-19 pandemic in India using AutoML." *World Journal of Engineering* (2020).
[48] Alakus, Talha Burak, and Ibrahim Turkoglu. "Comparison of deep learning approaches to predict COVID-19 infection." *Chaos, Solitons & Fractals* 140 (2020): 110120.