



Healthcare Diseases Classification Based on Machine Learning Algorithms: A Review

Ahmed Jameel Mohammed*, Adnan M. Abdulazeez**

* IT Dept., College of Akre, Akre University for Applied Sciences, Akre, Iraq

** IT Dept., Duhok Technical College, Duhok Polytechnic University, Duhok, Iraq

*Corresponding Email: ahmed.mohammed@dpu.edu.krd

ABSTRACTS

Researchers have increasingly focused on applying machine learning algorithms to enhance healthcare operations in the past few years. Machine learning has become increasingly popular and has shown to be a viable strategy for raising the standard of healthcare, preventing disease transmission and early disease detection, reducing hospital operational expenses, aiding government healthcare programs, and enhancing healthcare efficiency. This review offers a succinct and well-structured summary of machine learning research that has been done in the field of healthcare. Specifically, the emphasis is placed on the examination of non-communicable illnesses, which pose a significant risk to public health and rank among the primary contributors to global mortality. Moreover, the COVID-19 pandemic, which is among the world's deadliest illnesses and has recently been formally declared a public health emergency, is included. This study aims to assist health sector researchers in choosing appropriate algorithms. After conducting a comprehensive investigation, it was shown that the Decision Tree (DT), Gaussian Naive Bayes (GNB), and Random Forest (RF), algorithms had the highest performance in healthcare classification, achieving a remarkable accuracy rate of 100%. In most tests, the

ARTICLE INFO

Article History:

Received 09 May 2024

Revised 28 Jun 2024

Accepted 02 Jul 2024

Available online 24 Jul 2024

Publication date 01 Dec 2024

Keywords: Communicable Diseases, Non-Communicable Diseases, Healthcare, Machine Learning Algorithms

Random Forest (RF) and Support Vector Machine (SVM) demonstrated consistently better performance.

© 2024 Tim Konferensi UNIKOM

1. INTRODUCTION

The health sector has seen considerable changes in recent years due to the advancement of technology. Technological progress has improved existing health technologies and introduced new health technologies to the sector. Machine learning algorithms, which work with data sets and offer a realistic approach to events, have started to get a lot of interest from scholars (Ali et al., 2020).

The increasing demand for satisfactory and high-quality healthcare services must be met by healthcare providers. Process improvements in health institutions reduce the costs of the institutions, while patient satisfaction and service quality can be increased. In this direction, there are many international studies in the literature focusing on the improvement of health processes (Ray et al., 2022). Studies such as classifying, predicting, and diagnosing diseases have been the focus of researchers in the health sector. The most crucial components of managing an illness are early detection of warning signs, early diagnosis, and disease classification; meanwhile, examining disease findings and sending patients for more testing lowers the risk of death. Also, depending on the examinations, Classifying the disease and diagnosing it quickly improves the patient's quality of life (Hossain, et al., 2020; Haji et al., 2021).

Decision-making in health care is a complex process as it brings together various aspects and involves different stakeholders. Machine learning algorithms, one of the most widely used methods in disease prediction studies, provide very quick results (Jayatilake and Ganegoda, 2021). However, machine learning, which has many algorithms, also shows the performance percentage of the algorithm at the end of the output obtained. On the other hand, while researchers are confused about which algorithm to apply to the data, it is sometimes seen that they lose time in the algorithm they choose (Dietterich, 2000).

Despite the fact that non-communicable diseases play a big role in machine learning research, it is observed that researchers experience some complexities when choosing machine learning algorithms (Ferdousi et al., 2021). For this purpose, this study aims to analyze the threats to public health in general and non-communicable diseases, which rank highest among the reasons why people die, and summarize the best-performing algorithms with high accuracy. In this context, many machine learning studies, such as early diagnosis, disease classification, drug classification, and disease prediction studies, have been classified in detail. In addition, COVID-19 disease, the disease, it is also included and has just been formally declared a public health emergency due to its high

transmissibility. The study analysed machine learning research conducted in the realm of health, focusing on specific diseases.

This literature's primary goal is to assist health sector researchers in selecting the right algorithms. In this study, machine learning algorithm studies applied in healthcare services are classified in order to prevent the complexity and loss of time in algorithm selection. In addition, it is a compilation of studies that contain a lot of information, such as algorithm performance rates as a result of the applied algorithm. Furthermore, separate tables were created for diseases that threaten public health and rank first in the list of reasons of death in the world, and information about the performance results of the algorithm was also included. The study includes a total of 9 diseases and many different algorithms.

2. REVIEW OF NON-COMMUNICABLE DISEASES

Non-communicable diseases have significant impacts on the health and social welfare systems, leading to premature death and long-term illness or disability. The primary cause of death in this study is non-communicable diseases, which pose a concern to public health. Additionally, it intends to summarize the most effective algorithms that demonstrate high accuracy. In this context, many machine learning studies, such as early diagnosis, disease classification, and disease prediction studies, have been classified in detail.

2.1. Heart Disease

A systematic review of research has been showed to analyse machine learning

studies in healthcare within the scope of the review study. Heart conditions top the list of fatal illnesses. Globally, heart disease is a leading cause of death, accounting for 17.9 million deaths annually, according to data from the World Health Organization ([Anon n.d., 2023](#)).

For this reason, studies such as the classification of heart disease, timely diagnosis, and risk of heart disease are important. As shown in Table 1, it summarizes machine learning algorithms for heart disease that were analyzed within the scope of the review, including the algorithm with performed the best accuracy rate, the problem addressed, the algorithms compared, and the results.

In ([Ahmed et al., 2020](#)), the authors used a dataset of Cleveland heart patients for the year 2016, consisting of many data such as age, gender, blood sugar values, cholesterol values, and maximum heartbeat, and applied four algorithms to the dataset. Distinguish words with the title "hrtdis" from tweet streams and store them in Kafka. As a result, it aimed to predict whether a post contained symptoms of heart disease or not. The RF classifier outperformed better performance than other models, obtaining the highest accuracy of 94.9%.

In ([Mung and Phyu, 2020](#)), the authors applied four algorithms separately to data sets consisting of lung cancer, heart disease, diabetes, and cervical cancer. To evaluate the dataset, compare it with Ensemble Learning (EL) NB, DT, and K-NN, and algorithms. The heart disease dataset contains 370 patients and 14 attributes. The Ensemble learning algorithm gave the highest accuracy results in heart disease.

In (Li et al., 2020), the authors used a dataset with 297 patients and 13 features to diagnose heart disease patients. Many methods were used for heart disease prediction. As machine learning classifiers, the LR, ANN, K-NN, DT, NB, and SVM algorithms were employed. The best techniques for model evaluation and hyperparameter adjustment have been learned through the application of the cross-validation method. The classifiers' efficacy was evaluated using the features that feature selection algorithms had chosen. Four common feature selection techniques were used by the researchers on the dataset: FCMIM, MRMR, LASSO, and LLBFS. Features Selected by the Proposed algorithm with SVM algorithm provide a high accuracy rate.

In (Sridhar and Kirubakaran, 2021), the authors studied the Cleveland heart patient dataset from the University of California database to find the best classifier to predict the diagnosis of heart disease patients, analyzed 252 out of 297 patients. Applied RNN and CNN algorithms to a heart disease dataset. The accuracy result of the RDNN algorithm was 97.78%.

In (Jothi et al., 2021) the authors used DT and K-NN algorithms to predict the level of risk for heart disease. The risk of heart disease was determined, and individuals who might have heart disease were predicted. The heart disease dataset includes 13 medical parameters such as age, gender, fast blood glucose, chest pain, etc. The results were produced using Python programming. As a result of the two algorithms, it was observed that the algorithm with the highest percentage of accuracy was the Decision Tree (DT) model, the accuracy rate of this algorithm was 81%.

In (Waris and Koteeswaran, 2021), the authors enhanced the Python K-NN classifier with the goal of early identification and prediction of heart disease. Heart patients' dataset includes many pieces of information, such as smoking, eating habits, diabetes, and blood pressure. Pre-processing is the first phase in their five-step study format, the second step is the calculation of the data, the third step is the application of the K-NN algorithm, the fourth step is to complete the class prediction phase, and the last step is to establish the algorithm's accuracy. The Novel K-NN algorithm gave better result, and the accuracy of algorithm was 93%.

In (Hassan et al., 2021), to estimate the risk of heart disease, the scientists employed the SVM, NB, and DT algorithms. Finding the best machine learning model to forecast heart disease accurately was the goal of their research. The analysis demonstrates that the DT algorithm outperforms NB and SVM in terms of accuracy while also requiring minimum training time. The DT algorithm achieves an impressive accuracy of 98.3%.

In (Taylan et al., 2023), the authors wanted to make heart disease prediction more convenient and reliable. The following methods tested: Gradient Boosting, K-NN, RF, DT, NB, and LR. Following extensive testing, the Logistic Regression algorithm emerged as the sole dominant algorithm, consistently obtaining accuracies of 91.6% and 90.8% in the majority of tests. However, in the last dataset, the RF algorithm outperformed all others, achieving the greatest accuracy of 98.6% in all tests.

In (Basha et al., 2023), in their investigation, the authors employed a

variety of machine learning algorithms, including K-NN, NB, DT, SVM, and LR. Their study's primary goal is to create a model that can most accurately forecast heart disease. The algorithm with the best

performance was the SVM algorithm, with an accuracy of 80%. The comparison algorithms and results are summarized in Table 1.

Table 1. Machine Learning Algorithm Used for Heart Disease.

Authors	Year	The Problem Handled	Comparison Algorithms	Results of Comparison	Accuracy (%)
Ahmed et al.	2020	Finding the optimal algorithm that provides better accuracy for heart disease	NB, DT, K-NN, Ensemble Learning	Ensemble Learning algorithm was the best algorithm for all dataset.	99.93
Mung & Phyu	2020	Finding the best classification to predict diagnosis of heart patients	DT, RF, SVM, LR	Random Forest (RF) algorithm is the best predictor.	94.9
Li et al.	2020	Identifying heart disease in a timely and efficient manner	ANN, SVM, NB, DT, LR, K-NN,	SVM algorithm is the best predictor.	92.37
Sridhar & Kirubakaran	2021	Finding the best classification to predict diagnosis of heart patients	RDNN, HNN	It was decided that the Novel RDNN algorithm was the best algorithm	97.78
Jothi et al.	2021	Estimating the risk level of heart disease	K-NN, DT	The aim is to establish a system that calculates the patient's probability of getting heart disease and the DT is the best.	81
Waris & Koteeswar	2021	To predict diagnosis of heart patients	Novel K-NN, K-NN	It was decided that the Novel K-NN algorithm was the best algorithm.	93
Hassan et al.	2021	Finding the best classification to predict diagnosis of heart patients	DT, NB, SVM	DT algorithm is the best predictor.	98.3

Authors	Year	The Problem Handled	Comparison Algorithms	Results of Comparison	Accuracy (%)
Taylan et al.	2023	Finding the best classification to predict diagnosis of heart patients	KNN, LR,GB, RF, NB,DT, LR	Random Forest (RF) algorithm is the best predictor.	98.6
Basha et al.	2023	To predict diagnosis of heart patients	SVM, KNN, LR, NB, DT	SVM algorithm is the best predictor.	80

2.2. Diabetes Disease

In 2015, there were 229 million people living with diabetes worldwide. By 2050, that figure is projected to rise to around 1.3 billion people (Ong et al., 2023). Diabetes treatment is not only inadequate but also very costly. Therefore, it is very important to find the causes of diabetes, diagnose the disease early, and take precautions. Different algorithms for diabetes disease came together and are summarized in Table 2, with the methods with the maximum accuracy rate, the problem addressed, and the algorithms compared. As a result of the studies, Linear Discriminant Analysis (LDA), RF, NB, K-NN, SVM, DT, LR algorithms were obtained.

In (Lukmanto et al., 2019), the authors used feature selection method and fuzzy SVM were employed for the purpose of classifying and detecting diabetes. The Pima dataset was utilized for analysis. The dataset includes many data points, such as blood pressure, glucose values, BMI, and insulin values. As a result, the fuzzy SVM algorithm's accuracy was 89.02%.

In (Mujumdar and Vaidehi, 2019), the authors used the Pima Indian dataset used to analyze glucose, BMI, age, insulin, etc. The aimed of study to design

a diabetes prediction model to classify diabetes in a dataset. The researchers observed that LR gave the greatest accuracy of 96%, and in addition to the pipeline results, the AdaBoost classifier was better model with an accuracy of 98.8%. The Extra Trees Classifier, RF, Gradient Boosting, LDA, and AdaBoost algorithms were used.

In (Nguyen et al., 2019), the goal of the study the authors carried out was to forecast the development of type 2 diabetes. Its goal was to use public hospital registry data for the US population to estimate the development of diabetes in patients. The dataset consisted of data from 9948 patients with type 2 diabetes between 2009 and 2011. 80% of the dataset samples were used as the training, and 20% of the dataset was used as the testing. The dataset exhibits class imbalance, which was addressed by applying the SMOTE oversampling technique to the minority samples. Using an ensemble model without the SMOTE oversampling method, it achieved 82.12% accuracy. While using the SMOTE oversampling method, it did not improve accuracy but increased sensitivity from 31.17% to 49.40%.

In (Tigga and Garg, 2020), the authors used an online system to collect data on the risk of diabetes in 952 samples with an 18-question survey in India. Numerous characteristics are included in the dataset, including age, gender, BMI, smoking and alcohol consumption trends, and sleep habits. The SVM, LR, NB, RF, DT, and K-NN algorithms were used to discover the algorithm with the maximum accuracy rate. The RF algorithm is the best predictor with an accuracy of 94.10%.

In (Viloria et al., 2020), the authors used a Colombian patient data set. The researchers tried to predict the probability of patients having diabetes by using the SVM algorithm. The acquired findings encompassed outcomes such as diabetes, a tendency to diabetes, and the absence of diabetes. The SVM algorithm attained a good accuracy percentage of 99.2%.

In (Reddy et al., 2020), the authors used many classification algorithms in their study, such as NB, K-NN, SVM, and LR. The Pima dataset used, planned a diabetes prediction model for diabetes

classification in the dataset. The main aim of their study is to design a model that predicts diabetes with maximum accuracy. With an accuracy of 98.48%, the RF algorithm demonstrated the best performance.

In (Febrian et al., 2023), the authors' study conducted a comparison of two NB and K-NN method in order to predict diabetes. This prediction was using supervised machine learning and a number of the health parameters in the dataset. Depending on the result, the Naive Bayes algorithm performed higher than KNN, at 76.07%.

In (Uddin et al., 2023), the authors' developed a model was designed to assist in the estimate of diabetes. The model consists of various machine learning techniques, including DT, LR, RF, NB, K-NN, and SVM. The random forest method achieved excellent results, attaining a 97% accuracy rate on the diabetes dataset of 2019 and an 80% accuracy rate on the Pima Indian dataset. The comparison algorithms are summarized in Table 2.

Table 2. Machine Learning Algorithm Used for Diabetes.

Authors	Year	The Problem Handled	Comparison Algorithms	Results of Comparison	Accuracy (%)
Lukmanto et al.	2019	Classifying diabetes	-	It shows 89.02% accuracy in predicting diabetic patients. The SVM used.	89.02
Mujumdar & Vaidehi	2019	To predict a model for diabetes classification	AdaBoost, Gradient Boost, RF, LR, DT, LDA	The algorithm that gives the best solution is the AdaBoost algorithm.	98.8

Authors	Year	The Problem Handled	Comparison Algorithms	Results of Comparison	Accuracy (%)
Nguyen et al.	2019	Predicting patients with diabetes onset	-	Using an ensemble model achieves good results.	82.12
Tigga and Garg	2020	To predict diabetes risk	LR, RF, DT, K-NN, SVM, NB	The RF was the best estimator.	94.1
Viloria et al.	2020	Predicting the likelihood of patients having diabetes	-	The SVM showed results that were positive.	99.2
Reddy et al.	2020	To design a model that predicts diabetes with high accuracy	SVM, K-NN, LR, NB, RF	The RF showed the highest performance compared to other classifiers.	98.48
Febrian et al.	2023	To predict diagnosis of diabetes	KNN, NB	NB algorithm is the best predictor.	76.07
Uddin et al.	2023	To design a model that predicts diabetes with high accuracy	LR, SVM, DT, NB, Lnr, KNN, RF	The algorithm that gives the best solution is the RF algorithm.	97

2.3. Liver Disease

The liver, one of the most important organs in the body, is located in the upper left corner of the abdominal cavity. Numerous conditions affect the liver, including fatty liver, liver cancer, hepatitis, and typical liver disease issues. The increasing prevalence of chronic diseases, prolongation of life expectancy, and developments in science and engineering accelerate innovations in health technology (Behera et al., 2023). The review study looked for machine learning algorithms that compared different algorithms for liver disease and found the most accurate one. The problem that the studies tried to solved, the results of models were compared as shown in Table 3.

In (Gatos et al., 2017), the authors employed an ultrasound shear wave electrography (SWE) imaging dataset in conjunction with a classification algorithm to classify chronic liver disease. The dataset consists of liver pictures obtained from a total of 126 individuals. Various techniques were employed on a dataset consisting of 126 patients. These techniques included pre-processing, the calculation of five stiffness values cluster segmentation, feature selection, and feature extraction. With the SVM model, the best value was obtained with the highest accuracy (93.5%) in the classification of chronic liver patients according to subject discrimination.

In (Gogi and Vijayalakshmi, 2018), the authors used many classification algorithms for the purpose of classifying or predicting liver patients. There has been discovered an approximately one-to-one area under the curve (AUC). A confusion matrix shows the performance of the classification model on dataset. The dataset contains 572 patients, it has 394 samples were predicted as "yes" and 178 samples as "no.". The researchers employed the DT, LR, SVM, and Linear Discriminator algorithms. Using Matlab programming, the best algorithm is logistic regression, with ROC of 0.93 and accuracy of 95.8%.

In (Thaiparnit et al., 2018), the authors performed classification for seven types of liver disease and used 359 patient data points to classify liver disease. To evaluate the classification, set up five-fold cross-validation using the test hypothesis. Data rules section, OneR Rule, DT Stump technique, and Tree REPTree RF techniques were used. The RF algorithm gave better performance than other classification models.

In (Wu et al., 2019), the authors created a model for estimate fatty liver disease (FLD) and early detection of liver disease using ANN, LR, NB, and RF. The RF model performed better performance than other classification models with an accuracy of 87.47%. It has been stated that the application of the RF model in the clinical setting can help doctors stratify fatty liver patients for early treatment, surveillance, primary prevention, and management.

In (Srivenkatesh, 2019), the authors classified liver infections in order to forecast liver infections using various machine learning algorithms and to

determine which algorithm was the most effective. The dataset utilized in this study comprised information from 583 patients diagnosed with liver disease, encompassing ten distinct variables. The data was obtained from the Kaggle platform. In this study, five types of classifiers were applied to examine liver disease: K-NN, NB, RF, LR, and SVM. As a result of the study, the LR model showed higher performance than other classification algorithms.

In (Shi et al., 2021), the authors conducted a retrospective evaluation of liver patient data spanning from January 2013 to May 2017. The study used SVM, LR, NB and SVM algorithms to classify this dataset of 672 consecutive patients who underwent arcus aortic surgery. As a result of this study, the LR model showed better performance than other classification models.

In (Md et al., 2023), the authors presented a new architecture that utilizes ensemble learning and improved pre-processing techniques to forecast liver illness using the ILPD dataset. The pre-processed data has been trained using various ensemble learning techniques, including extra tree, bagging, RF, gradient boosting, XGBoost, and stacking. The suggested model, utilizing the additional tree classifier and random forest, demonstrated superior performance compared to other methods, achieving maximum accuracy of 91.82% and 86.06%, respectively.

In (Suárez et al., 2023), the authors' aim of the study was to create a method for the detection of liver fibrosis risk. An extremely effective prediction model was developed for this particular goal using the extreme gradient boosting (XGB)

technique. The proposed method, XGB, outperformed existing machine learning classification algorithms in terms of accuracy. The XGB technique achieved

the highest balanced accuracy, with a score of 93.16%. The comparison algorithms are summarized in Table 3.

Table 3. Machine Learning Algorithm Studies Used for Disease.

Authors	Year	The Problem Handled	Comparison Algorithms	Results of Comparison	Accuracy (%)
Gatos et al.	2017	To classifying chronic liver disease.	-	The best algorithm is SVM.	93.5
Gogi & Vijayalakshmi	2018	To classifying liver patients	SVM, LR, LDA,DT	The best algorithm is LR.	95.8
Thaiparnit et al.	2018	To classifying liver patients	-	The RF algorithm that gives the best results.	75.76
Wu et al.	2019	Predicting fatty liver disease	RF, NB, ANN, LR	The best algorithm is RF.	87.47
Srivenkatesh	2019	Liver infection prediction classify	RF,K-NN, NB, LR, SVM	The best algorithm is LR.	76.27
Shi et al.	2021	Examining the dataset of liver patients who underwent aortic arch surgery	LR, SVM, RF, NB	Logistic Regression has performed better.	77
Md et al.	2023	To detecting Liver Disease	GB, XGBoost, bagging, TC,RF, ET	The algorithm that gives the best solution is the TC algorithm.	91.82
Suarez et al.	2023	To detecting Liver Disease	SVM, BLDA, LR, KNN, DT, XGB	The algorithm that gives the best solution is the XGB algorithm.	93.16

2.4. Thyroid Disease

The thyroid is a butterfly-shaped organ located in the front part of the neck. The hormones of the thyroid gland enter the bloodstream and regulate metabolism, controlling its speed. Thyroid disease poses a big risk to human life (Abbad et al., 2021). Therefore, studies conducted in this field are very important. In this review, machine learning algorithms in which different algorithms for thyroid disease are combined are analyzed. The most accurate algorithm is selected as the one with the highest accuracy rate, the problem addressed, and the algorithm compared with the results. It is summarized in Table 4.

In (Tyagi et al., 2018), to identify thyroid problems early on, the authors used machine learning techniques. The dataset providing information about thyroid patients was utilized. The dataset includes information such as age, gender, and thyroid hormone values. Many machine learning algorithms were used, such as DT, SVM, and K-NN, to predict those at risk of developing thyroid disease. The algorithm that gives the best results is SVM, which has an accuracy of 99.63%.

In (Olatunji et al., 2021), the authors planned to build a system to detect thyroid disease at very early stages (pre-symptomatic stage). The dataset from King Fahad Hospital in Saudi Arabia includes thyroid patients and various blood diseases. There are two classes in

the dataset: normal patients and thyroid patients. In the dataset containing 14 attributes, there are 109 thyroid patients and 109 normal patients. The ANN, SVM, RF, and NB algorithms were used in the study. The algorithm that gives the best results is RF, which has an accuracy of 90.91%.

In (Chaganti et al., 2022), the authors' goal was to forecast thyroid based on feature selection by using a variety of machine learning methods. The datasets come from the University of California, Irvine's (UCI) machine learning library. The dataset contains 7200 samples and 21 attributes. As result of comparison among algorithms showed that, the algorithm that gave the best results was RF, which has an accuracy of 99%.

In (Alshayeji, 2023), the authors suggested approach seeks to overcome current constraints in the field, including insufficiently detailed feature analysis, limited visualization capabilities, inadequate prediction accuracy, and unreliable results to predict early thyroid risk. Furthermore, the issue of overfitting was resolved with the use of 5-fold cross-validation and data balancing utilizing the SMOTE. The proposed model demonstrated exceptional performance, achieving an accuracy of 99.5%, a sensitivity of 99.39%, and a specificity of 99.59% when used with the boosting method.

In (Devi et al., 2023), the research's investigation uses a dataset from the University of California Irvine (UCI) machine-learning repository, consisting of 3163 patient records and 23 independent features. The BCCM predicted probabilities of classifiers to demonstrate a remarkable accuracy of 99.7%.

Table 4. Machine Learning Algorithm Used for Thyroid Disease.

Authors	Year	The Problem Handled	Comparison Algorithms	Results of Comparison	Accuracy (%)
Tyagi et al.	2018	Predicting thyroid disease	SVM, K-NN, DT	SVM was the algorithm that gave the best results.	99.63
Olatunji et al.	2021	Predicting the possibility of thyroid disease	RF, ANN, SVM, NB	The best performing algorithm is RF.	90.91
Chaganti et al.	2022	To predict thyroid disease	RF, LR, GBM, ADA, SVM	The best performing algorithm is RF.	99
Alshayeji	2023	To predict early thyroid risk	LR, NN, DT, SVM, KNN, Bagging, Boosting	The algorithm that gives the best solution is the proposed model with Boosting algorithm.	99.5
Devi et al.	2023	Predicting the possibility of thyroid disease	LR, KNN, SVM, GNB, DT, RF, GB, RC, SGB, BCCM, BCRM	The algorithm that gives the best solution is the proposed model with BCCM.	99.7

2.5. Breast Cancer

Breast cancer, one of the leading gynaecological diseases, is one of the most fatal diseases; early diagnosis reduces the risk of death. Therefore, studies carried out in this field are very important. The review study provides a summary of machine learning algorithms focused on breast cancer, encompassing

several algorithms (Tewari et al., 2022; Khorshid & Abdulazeez, 2021). The algorithm with the highest accuracy rate and the particular problem it targets are shown in Table 5, the algorithms that were compared, and the corresponding outcomes. NB, RF, SVM, DT, LR, K-NN, LDA, ANN, Radial Basis Function Kernel SVM algorithms were applied in the studies.

In (Karthikeyan et al., 2020), the authors' study, employed three primary algorithms in conjunction with the WBC dataset with the objective of determining the optimal machine learning algorithm. The algorithms used were K-NN, RF, and NB. In order to find the best classification accuracy, compared the efficiency and effectiveness of these algorithms in terms of accuracy, precision, sensitivity, and specificity, and as a result, the K-NN algorithm was found to be the best performing algorithm compared to the other algorithms, with an accuracy of 95.90%.

In (Vaka et al., 2020), the authors aimed to present a new method to detect breast cancer using machine learning techniques. The proposed method produced very accurate and efficient results compared to existing methods. The DNNS method: the proposed method is based on the support value in a deep neural network. For comparison, machine learning algorithms such as NB and SVM were used. The experimental findings demonstrate that the suggested

DNNS method outperforms the current approaches by a substantial margin.

In (Sarkar et al., 2021), the authors performed miRNA analysis to classify breast cancer. In breast cancer, tumour suppressor microRNAs and downstream signalling pathways are responsible for the tumour. In order to obtain this list of microRNAs, machine learning algorithms were used. The study consists of two stages. In the first stage, by looking at the accuracy of classifying the dataset, Using ANN, SVM, DT, K-NN, RF, NB, and DISCR algorithms, the best accuracy was selected. In the second stage, the classification of breast cancer subtypes was performed. Compared to the other six methods, the Random Forest (RF) algorithm yielded comparatively higher accuracy (76.5761%).

In (Gopal et al., 2021), Using Internet of Things devices, the authors employed machine learning algorithms for the purpose of classifying and selecting breast cancer features. The researchers employed the Wisconsin Breast Cancer Dataset (WBCD) for their analysis. There are 32 features and 569 examples in the dataset. LR, MLP, and RF classifiers were used, and the MLP gave better accuracy rate than others.

In (Wu and Hicks, 2021), the authors used two main machine learning algorithms on their Wisconsin breast cancer dataset. It aimed to classify the disease of breast cancer. It was suggested to compare the NB and K-NN algorithms, and cross-validation was used to assess

each algorithm's correctness. After making an accurate comparison between the algorithms, the K-NN algorithm achieved a higher efficiency of 97.51%.

In (Al-Azzam and Shatnawi, 2021), the authors aimed to evaluate and compare the performance and accuracy of classification algorithms for breast cancer prediction. Nine machine learning classification algorithms were employed, encompassing both supervised and semi-supervised learning approaches. These are: GNB, RF, LR, LSVM, DT, RBFSVM, Xgboost, K-NN, and Gradient Boosting, were used on the Wisconsin Diagnostic Cancer dataset. K-fold cross-validation was used, and the hyperparameters were tuned, to guarantee the model's dependability. The best accuracy of 98% was attained by the semi-supervised LR method and the supervised KNN model technique.

In (Khorshid et al., 2021), The authors' goal was to assess and contrast the algorithms for performance classification. The following classifiers' performances are compared: Weighted

K-NN, SVM, LR, K-NN, and GNB. A source of the dataset was the UCI Machine Learning Repository. This project's main goal is to classify women with breast cancer using machine learning algorithms, with an emphasis on accuracy. In comparison to the other classifiers, the results show that weighted K-NN (96.7%) attained the highest level of accuracy.

In (Ebrahim et al., 2023), the National Cancer Institute (NIH), in the United States, provided the authors with a dataset comprised of 1.7 million datasets. The accuracy evaluation took into account both deep learning and traditional learning techniques. The study included ensemble techniques (ET), logistic regression (LR), support vector machines (SVM), classical decision trees (DT), linear discriminants (LD), and logistic regression (LR). For comparison, the study used recurrent neural network (RNN), deep neural network (DNN), and neural network algorithms. With a score of 98.7%, the DT algorithm had the highest accuracy.

Table 5. Machine Learning Algorithm Used for Breast Cancer.

Authors	Year	The Problem Handled	Comparison Algorithms	Results of Comparison	Accuracy (%)
Karthikeyan et al	2020	Classifying breast cancer	NB, K-NN, RF	The K-NN algorithm gave a better result.	95.90
Vaka et al.	2020	Early detection of breast cancer	SVM, NB, DNNS	The DNNS algorithm gave the best result.	97.21

Authors	Year	The Problem Handled	Comparison Algorithms	Results of Comparison	Accuracy (%)
		provide diagnostics			
Sarkar et al.	2021	Predicting breast cancer subtype	KN-NN, DT, RF, SVM, ANN, NB, DISCR	RF provided accuracy that was relatively better than the other six techniques.	76.58
Gopal et al.	2021	Ensuring early predict of breast cancer	RF, MLP, LR	The MLP was found to be the best-performing algorithm.	98
Wu and Hicks	2021	Classifying breast cancer	NB, K-NN, RF	The K-NN algorithm gave the best result.	97.51
Al-Azzam & Shatnawi	2021	Early detection of breast cancer provide diagnostics	LR, GNB, SVM, RBF SVM, DT, RF, K-NN Xgboost, Gradient Boosting,	The KNN model for a and LR algorithms gave the highest accuracy.	98
Khorshid et al.	2021	Classifying breast cancer breast	SVM, KNN, LR, Weighted K-NN, GNB,	The algorithm that gives the best solution is the Weighted K-NN algorithm.	96.7
Ebrahim et al.	2023	Classifying breast cancer	DT, LD, LR, SVM, ET, DNN, RNN, PNN	The algorithm that gives the best solution is the DT algorithm.	98.7

2.6. Cardiac Arrest

Cardiac arrest is when the heart no longer performs its pumping function. Cardiac arrest is different from a heart attack. A heart attack is usually a reaction

to fat accumulating in the heart vessels. Cardiac arrest occurs suddenly and unexpectedly (Majumder et al., 2019). Therefore, studies conducted in this field are very important. Table 6 summarizes the machine learning algorithms for cardiac arrest that were analyzed within the literature review, comprising the problem solved, the methods compared, the algorithm with the highest accuracy rate, and the outcomes. Table 6 shows the following algorithms: NB, RF, SVM, DT, LR, K-NN, ANN, and GNB.

In (Chauhan et al., 2019), the authors aimed to examine different machine learning algorithms to predict the possibility of cardiac arrest. After applying SVM, RF, DT, LR, and ANN algorithms to the dataset to predict the occurrence of cardiac arrest in patients, it was determined that the best-performing algorithm was ANN, and its accuracy was 85%.

In (Chang et al., 2019), the authors, created two tasks to detect cardiac arrest. In the first task, The AUROC curve was used, and in the second, The region beneath the precision-recall curve was computed using machine learning techniques. Static data includes values such as age, gender, height, weight, temperature, etc., while dynamic data includes vital signs such as mean arterial pressure, systolic blood pressure, pulse rate, respiratory rate, body temperature, etc. The NB, SVM, and RF algorithms were used. The RF gave better performance than others.

In (Kwon et al., 2019), the authors conducted a survival analysis of patients who experienced cardiac arrest after discharge. Its goal was to create a DCAPS that could forecast illness recovery and discharge survival. In validation data, the accuracy of DCAPS's area under the receiver operating characteristic curve (AUROC) for predicting neurological recovery was 95.3%, Logistic Regression was 94.7% accurate, Random Forest (RF) was 94.3% accurate, Support Vector Machine (SVM) was 93.0% accurate, and conventional methods from a previous study were 81.7% accurate. DCAPS surpassed the conventional approach and other machine learning techniques in its ability to predict neurological improvement and survival for the correct discharge of OHCA patients.

In (Javan et al., 2019), the authors aimed to predict cardiac arrests due to sepsis in adult patients. It also looked into how the vital sign time-series dynamics affected the cardiac arrest prediction. Classifiers like XGBoos, MLP, K-NN, SVM, DT, LR, GNB, and RF are among the machine learning methods that are employed. The best-performing algorithm was Random Forest (RF).

In (Hirano et al., 2021), the authors aimed to develop a patient's cardiac arrest prediction model. Japan gathered data on 43,350 people who suffered from cardiac arrest outside of a hospital, and patients younger than 18 years of age or patients whose cardiac arrest occurred due to a factor were not included. The results of LR, RF, SVM, and MLP

algorithms were compared. The results are within the 95% confidence interval and are evaluated with a 5% margin of error. In validation analyses, 86.6% accuracy was determined for SVM, 87.7% for RF, and 88.8% for MLP classifiers. It was determined that the best estimator was the Multilayer Perceptron classifier.

In (Safa and Pandian, 2021), the authors aimed to calculate the stress analysis of cardiac arrest patients. The study consists of temperature, blood pressure, pulse, and stress-related cardiac patients' data. The prediction model for stress analysis trained DT, SVM and K-NN algorithms. A dataset related to heart disease was used to evaluate the algorithms. The outcome demonstrates how much more effective the K-NN classification method was than the SVM and DT algorithms.

In (Shashikant and Chetankumar, 2023), the authors aimed to develop a patient's cardiac arrest prediction model

based on smokers by analysing heart rate variability (HRV) factors. They aimed to assess and compare the predictive accuracy of DT, LR, and RF models in identifying the occurrence of cardiac arrest among individuals who smoke. The random forest model demonstrated the highest accuracy in classification, followed by the decision tree, while logistic regression exhibited the lowest accuracy in classification. The accuracy performance of the random forest was 93.61%.

In (Javeed et al., 2023), the authors planned to build a system to predict death in cardiac patients to tackle this issue. The newly suggested method utilized a χ^2 statistical model to prioritize the characteristics within the dataset. There are 368 samples altogether and 55 attributes in the dataset that the suggested model is based on. Comparing the suggested model, χ^2 _RF, to other techniques, it obtained the highest accuracy of 94.59%.

Table 6. Machine Learning Algorithm Used for Cardiac Arrest.

Authors	Year	The Problem Handled	Comparison Algorithms	Results of Comparison	Accuracy (%)
Chauhan et al.	2019	Predicting the possibility of cardiac arrest	SVM, RF, DT, LR, ANN	The best performing algorithm is Artificial Neural Networks (ANN).	85
Chang et al.	2019	Predicting the possibility of cardiac arrest	NB, SVM, RF	The best performing algorithm is RF	88

Authors	Year	The Problem Handled	Comparison Algorithms	Results of Comparison	Accuracy (%)
Kwon et al.	2019	To analyze the survival of patients with cardiac arrest after discharge.	LR, RF, SVM, DCAPS	The best performing algorithm is DCAPS.	97.6
Javan et al.	2019	Predicting cardiac arrest due to infection	K-NN, GNB, SVM, DT, LR, Gradient boosting, XGBoost, RF	The best performing algorithm is RF.	95
Hirano et al.	2021	Predicting the possibility of cardiac arrest	LR, SVM, RF, MLP	The best performing algorithm is MLP.	88
Safa & Pandian	2021	Analysing the stress of cardiac arrest patients	K-NN, SVM, DT	The best performing algorithm is K-NN.	85.98
Shashikanth & Chetankumar	2023	Predicting the possibility of cardiac arrest	LR, DT, RF	The algorithm that gives the best solution is the RF algorithm.	93.61
Javeed et al.	2023	to build a system to predict in cardiac arrest	NB, LR, DT, RF, KNN, SVM., Proposed Model	The algorithm that gives the best solution is the Proposed Model algorithm.	94.59

2.7. Chronic Kidney Disease

When the kidney loss of function over time is called chronic kidney failure or chronic kidney disease. Loss of kidney function over time puts human life at risk. Studies carried out in this field are of great importance ([Dritsas and Trigka,](#)

[2022](#)). Table 7 shows the results, the problem is solved, and the algorithms that were compared. It also shows which algorithm had the highest accuracy rate. Table 9 shows the following algorithms:

RF, GNB, SVM, K-NN, DT, ANN, MLP, LDA, LR algorithms.

In (Rabby et al., 2019), the authors used classification algorithms to diagnose chronic kidney disease. In one part, the dataset was divided into training and testing. Many classification algorithms have been used. The algorithms with high accuracy rates were the RF, DT, and GNB algorithms. These algorithms, which show 100% accuracy, can classify chronic kidney patients. The researchers also created a mobile application that utilizes the most effective output outcomes classifier technique to forecast kidney disease based on patient reports.

In (Almansour et al., 2019), the authors aim of the study is to diagnose chronic kidney patients early. Comparison by applying classification algorithms to a dataset of 400 patients and 24 features relevant to the diagnosis of chronic kidney disease. The analysis was completed with ANN and SVM algorithms. According to the empirical results obtained from the applications, the accuracy of the ANN algorithm was 99.75% and the accuracy of the SVM algorithm was 97.75%. Although the results of both algorithms were good, the algorithm gave the highest accuracy rate was the ANN algorithm.

In (Chaithra et al., 2023), the authors investigated the early diagnosis of chronic kidney disease. A system has been created to predict patients who are at risk of developing chronic kidney disease. The dataset contains 445 chronic

kidney patients. The study utilizes an online data set obtained from the UCI Machine Learning Repository as well as a real-time data set collected from Khulna City Medical College. The RF had an accuracy of 97.12%, and AN had an accuracy of 94.5%. As a result of the study, the algorithm with a high probability of accuracy was the Random Forest (RF) algorithm.

In (He et al., 2021), the authors aimed to kidney injury frequently occurs as a complication following liver transplantation and serves as a sign of unfavorable prognosis. The dataset used was from a total of 493 patients who underwent brain death post-liver transplantation (DCDLT). Compared patients with and without acute kidney injury. LR, DT, RF, and SVM algorithms were used for prediction. The algorithm with the highest accuracy rate was determined to be RF.

In (Swain et al., 2023) the authors' objective of the study is to create a machine-learning model capable of utilizing publicly accessible data to predict the likelihood of chronic kidney disease. Data pre-processing techniques were applied to this dataset to create a universal model. Among the several applied learning techniques, SVM and RF demonstrated the most favorable outcomes in terms of false-negative rates and test accuracy, achieving rates of 99.33% and 98.67%, respectively. Nevertheless, the SVM outperformed RF

in terms of performance when assessed using 10-fold cross-validation.

In (Kumar et al., 2023), the authors introduced an innovative deep-learning model that integrates a fuzzy deep neural network for detecting and predicting kidney illness. The dataset originates from the Changhua Christian Hospital in

Taichung, Taiwan. After removing patients' identifying information, a dataset consisting of 5617 records from January 1, 2000, to July 27, 2017, was acquired. The results indicate that the suggested model achieves a remarkable accuracy of 99.23%, surpassing the performance of existing techniques.

Table 7. Machine Learning Algorithm Used for Chronic Kidney Disease.

Authors	Year	The Problem Handled	Comparison Algorithms	Results of Comparison	Accuracy (%)
Rabby et al.	2019	To diagnose chronic kidney disease early	Gradient boosting, K-NN, ANN, LR DT, GNB, RF, SVM, RF, AdaBoost, LD	The best-performing algorithms were the DT, RF, and GNB algorithms.	100
Almansour et al.	2019	To diagnose chronic kidney disease early	ANN, SVM	The best-performing algorithm was the ANN algorithm.	99
Chaithra et al.	2020	To diagnose chronic kidney disease early	RF, ANN	The best-performing algorithm was the RF algorithm.	97.12
He et al.	2021	To create a system for predict chronic kidney disease.	LR, RF, SVM, DT	The best-performing algorithm was the RF algorithm.	79
Swain et al.	2023	To diagnose chronic kidney disease early	SVM, RF	The best-performing algorithm was the SVM algorithm.	99.3
Kumar et al.	2023	To Detecting kidney disease	-	The best performing algorithm is fuzzy deep neural network	99.23

2.8. Alzheimer's Disease

Alzheimer's disease (AD) is a geriatric condition. Some parts of the brain begin to become increasingly damaged over time, and as a result, it is a disease that disrupts all daily activities and behaviors, especially memory (Franciotti et al., 2023). A compilation study looked at machine learning algorithms that combine different algorithms for Alzheimer's disease. Table 8 shows the results, the problem is solved, and the algorithms that were compared. It also shows which algorithm had the highest accuracy rate. Table 8 shows the following algorithms: NB, RF, SVM, DT, K-NN, ANN, MLP, Extreme Learning Machine (ELM), Extreme Learning Machine (GPR), Partial Least Squares (PLS), and SVM-DA algorithms.

In (Lodha et al., 2018), the authors aimed to analyze the brain images of Alzheimer's patients. Neuropsychological and objective assessments were investigated to see if there was a relationship between them. To detect the disease early, machine learning classification algorithms have been used. SVM, ANN and RF algorithms were used. The algorithm with the better accuracy rate was the ANN algorithm, with an accuracy of 98.36%.

In (Zhang et al., 2019), the authors aimed to identify early detection of patients with Alzheimer's disease. In addition to clinical information, voxel-based morphometry (VBM) parameters and tissue parameters were used in the

AD data in the neuroimaging initiative. Used 10-fold cross-validation to estimate the performance of models. The proposed methods were applied to data from 58 patients with Alzheimer's disease and 94 normal controls and achieved a classification accuracy of up to 96% with the ELM model. The results for the other three models are 82% PLS, 79% GPR, and 75% SVM. The results were optimal in distinguishing Alzheimer's patients from normal control patients, and therefore this study may be useful for the diagnosis of Alzheimer's.

In (Neelaveni and Devasana, 2020), the authors used machine learning algorithms to diagnose Alzheimer's disease early by using psychological parameters such as age, number of visits, MMSE, and education of AD patients. The MMSE score, age, number of visits, and education of the patients are important parameters that greatly assistance in predicting the AD disease. The SVM and DT algorithms were tested with 70% training and 30% test datasets. The best performing algorithm was the SVM algorithm, with an accuracy of 85%.

In (Khan and Zubair, 2022), the authors used data from the open access imaging studies series OASIS database of MRI brain images for analysis. The dataset consists of 343 samples of MRI sessions involving 150 subjects. To classify longitudinal brain MRI data for Alzheimer's diagnosis into two classes: with and without dementia, MMSE,

CDR, and ASF were used in the analysis. The planned machine learning pipeline created a classifier system with data transformation and feature selection techniques embedded in the experimental and data analysis designs. The best-performing algorithm was the RF algorithm, with an accuracy of 86.84%.

In (Rangaswamy et al., 2020), the authors aimed to find the best classification algorithm to diagnose Alzheimer's disease. It has compiled a dataset of 20,401 harmful and 37,452 normal genetic variations from the GWAS and GTEx portals, respectively. The RFECV used to select important features was followed by forward feature selection to distinguish between harmful and neutral variants, and the RF classifier was used and it has an accuracy of 81.21%.

In (Uysal and Ozturk, 2020), the authors proposed a system that can be used to decide between patients with AD, mild cognitive impairment (MCI), and cognitive normal (CN). For the established model, the dataset contains 482 patients, divided into 160 (33.2%) for testing and 322 (66.7%) for training. A dataset was generated using gender, age, diagnosis, and measurements of right and left hippocampus volume. Prediction performances were evaluated by classifying them among all diagnostic groups. Since the left hippocampus volume value gave more successful

prediction results than the right hippocampus in every model, it has examined a new area by adding the age and gender parameters that did not include the right hippocampus value, and this proved that the success rate increased. KNN performed with 80% accuracy, and GNB performed with 82% accuracy when the right hippocampus volume value was not included. The gender and age features had positive effects for each algorithm.

In (Ryzhikova et al., 2021), the authors developed a new method to diagnose AD based on cerebrospinal fluid. Raman Spectra of Cerebrospinal Fluid (CSF) samples were obtained from 21 individuals diagnosed with AD and 16 healthy control (HC) subjects. The ANN and SVM-DA were used for differentiation purposes, and the most successful results observed were the differentiation of Alzheimer's disease and healthy control subjects with 84% accuracy.

In (Salehi et al., 2023), the authors used LSTM to MRI data in order to address the limitations of traditional methods for detecting AD. The main goal of the study was to design a model for detecting Alzheimer's disease early. They obtained a batch of MRI data from the Kaggle source, which was used to train the LSTM. The performance of the model they developed was exceptional, with an AUC of 0.97 and an accuracy of 98.62%.

Table 8. Machine Learning Algorithm Used for Alzheimer's Disease.

Authors	Year	The Problem Handled	Comparison Algorithms	Results of Comparison	Accuracy (%)
Lodha et al.	2018	Analyzing brain images of Alzheimer's patients	SVM, RF, ANN	ANN was the algorithm that gave the best results.	98.36
Zhang et al.	2019	Detecting Alzheimer's disease early	SVM, GPR, PLS, ELM	The best performing algorithm is ELM	96
Neelaveni & Devasana	2020	Detecting Alzheimer's disease early	DT, SVM	The best performing algorithm is SVM	85
Khan & Zubair	2022	Categorizing brain MRI data for Alzheimer's diagnosis and dichotomizing into dementia and non-dementia.	ExtraTree, DT, NuSVC, AdaBoost, GradientBoosting, GNB, Ridge, K-NN, RF	RF was the algorithm that gave the best results.	86.84
Rangaswamy et al.	2020	To diagnose Alzheimer's disease.	XgBoost, AdaBoost, ANN, RF	The best performing algorithm is RF	81.21
Uysal & Ozturk 2020	2020	To separate between patients with AD, MCI, and CV	K-NN, LR, SVM, DT, RF, GNB	The best performing algorithms are GNB and KNN	80 / 82
Ryzhikova et al.	2021	Developing a new method to diagnose	ANN, SVM-DA	The best performing algorithms are ANN and SVM-DA	84

Authors	Year	The Problem Handled	Comparison Algorithms	Results of Comparison	Accuracy (%)
		Alzheimer's Disease			
Salehi et al.	2023	To detect Alzheimer's disease early	-	The algorithm that gives the best solution is the proposed LSTM model.	98.62

3. REVIEW OF COMMUNICABLE DISEASES

Small organisms like bacteria, viruses, fungi, or parasites that enter the body can cause illnesses known as communicable diseases. There are many different types of infectious diseases and transmission methods. In this review study, the COVID-19 disease, which is highly transmissible and has been officially designated as a public health crisis in recent times, is also encompassed.

3.1. COVID-19 Disease

The World Health Organization named the infectious disease caused by the SARS-CoV-2 virus COVID-19 in February 2020. The World Health Organization declared this pandemic a public health emergency of international concern on January 30, 2020. At first, it was not considered a serious case, but the rapid spread of the infection and its high mortality rate soon caused many problems in healthcare systems all over the world (Sun et al., 2020).

In this study, different algorithms for COVID-19 disease, which were screened within the scope of this study, came together. machine learning algorithm with the highest accuracy, and the

problem addressed, the comparison algorithms and results are summarized in Table 9. As a result of the search, SVM, RF, DT, ANN, LR, K-NN, Polynomial Regression (PR) algorithms were used in the studies.

In (Gambhir et al., 2021), the authors analyzed the transmission trend of COVID-19 disease in India. The dataset from the Indian Ministry of Health and Family Welfare and data from January 22, 2020, to June 24, 2020, a total of 154 days, were used. Compared SVM and PR. The increase in cases was predicted for July and August and showed approximately 93% accuracy.

In (Bayat et al., 2021), the authors aimed to develop a model, when the number of tests is not sufficient to diagnose SARS-CoV-2 disease, other methods are tried to identify the disease. They used a machine learning classification models to analyze the relationship between test results and the test date of SARS-CoV-2 with 20 routine laboratory tests collected over a 2-day period. The dataset was used for SARS-CoV-2, it contains 75,991 patients receiving inpatient and outpatient treatment. It has been compared with positive and negative results. As a result of the study, the accuracy was 86.4%.

In (Saha et al., 2021), the authors aimed to diagnose COVID-19 by examining patients' chest X-ray images. It has proposed an automatic detection scheme called EMCNet and developed the model using a CNN. The dataset was separated into two parts: training and testing. The dataset consists of 4600 patients in total. After data preprocessing and data splitting completed the CNN model. As obtained outputs, it compared with other algorithms. The CNN algorithm reached an accuracy rate of 96.52%. The RF, SVM, DT, and AdaBoost machine learning algorithms were used. These algorithms outputs have provided good results. The classification algorithm that performed the best performance in this classification was SVM with an accuracy of 96.96%.

In (Pourhomayoun and Shakibi, 2021), the authors used a dataset consisting of 307,382 patient samples and more than 2,670,000 laboratory-confirmed COVID-19 patients from 146 countries around the world. The study aimed to identify who should be taken care of first and who has a higher priority for hospitalization in hospitals and medical facilities. In addition, they also proposed an artificial intelligence model that would help prevent overcrowding in the system and eliminate delays in providing necessary care. The results show an overall accuracy of 89.98% in predicting mortality. Many classification algorithms were used, including ANN, SVM, RF, DT, LR) and KNN to predict death rate in COVID-19 patients. The best-performing algorithm was Artificial Neural Networks (ANN), which had an accuracy rate of 89.98%.

In (Alves et al., 2021), the authors aimed to develop a model for predicting

the diagnosis of COVID-19 through machine learning algorithms based on routine blood test results using a publicly available anonymous dataset. The raw dataset contains 5644 samples and 111 features. The DT and RF algorithms were used to decide the model using the criteria graph. As a result of the study, the algorithm with the highest accuracy rate was the RF algorithm, with an accuracy of 88%.

In (Liu et al., 2023), the authors aimed to develop a classification model for predicting the diagnosis of COVID-19 through machine learning algorithms. The objective is to utilize omics data from a substantial cohort to initially predict the COVID-19 status (positive or negative) of patients and subsequently assess the severity of the disease. In the COVID-19 diagnosis, the multilayer perceptron model had the greatest AUC of 0.99, while the LR model achieved the highest F1-score of 0.95. Regarding the job of predicting severity, they attained the best level of accuracy, specifically 0.76, utilizing a logistic regression model. The study discovered that machine learning models exhibited superior performance when applied to integrated multi-omics data as opposed to single-omics data.

In (Arowolo et al., 2023), the authors machine learning techniques, specifically ABC and SVM, to forecast COVID-19 for an IoT data system. The system underwent evaluation using a confusion matrix, yielding the following results: 95% accuracy, 97% precision, and 96% F1 score for ABC-LSVM. Additionally, the system achieved 97% accuracy, 100% precision, and 97% F1 score, for ABC-Q-SVM. Fetching pertinent data from IoT systems prior to classification has been found to be advantageous.

Table 9. Machine Learning Algorithm Used for COVID-19 Disease.

Authors	Year	The Problem Handled	Comparison Algorithms	Results	Accuracy (%)
Gambhir et al.	2021	Identifying a COVID-19 diagnosis by examining chest X-ray images	RF, SVM, DT, AdaBoost	SVM is the best predictor.	96.96
Bayat et al.	2020	To predict the risk of death from COVID-19 disease	SVM, ANN RF, DT LR, K-NN	The ANN algorithm for predicting death rates is the best estimator.	89.98
Saha et al.	2021	The transmission trend of COVID-19 disease in India was analyzed	PR, RF,SVM	The increase in cases was predicted for July and August. SVM is the best predictor.	93
Pourhoma youn & Shakibi	2021	Diagnosing SARS-CoV-2 patients	-	It may provide a complementary method (XGBoost) for determining COVID-19 infection status.	-
Alves et al.	2021	Diagnosing COVID-19 disease by looking at Brazilian hospital data	DT, RF	Random Forest (RF) algorithm is the best predictor.	88
Liu et al.	2023	To detect the risk of death from COVID-19 disease	SVM, RF,LR, KNN, RUS, MLP	LR is the best predictor.	76
Arowolo et al.	2023	To detect possible cases of COVID-19	ABC-LSVM, ABC-Q-SVM	ABC-Q-SVM is the best predictor.	97

4. CONCLUSION

The objective of this study is to compare machine learning classification algorithms used in health studies to discover the best-performing algorithm. In this context, machine learning classification studies carried out in the field of healthcare in the last 6 years are summarized and classified. The detailed literature study focused on machine learning studies applied to diseases that threaten public health and are among the top causes of death in the world.

Additionally, a COVID-19 disease, which is on the list of the world's deadliest diseases and has been declared a public health emergency in recent years, is also included. It is clear that studies on diseases such as early diagnosis, disease classification, drug classification, and epidemic prevention make great contributions to the health sector. Although these studies reduce the risk of

death for patients, improves the quality of life.

However, due to the scope and architecture, it is difficult to choose a forecasting and demand decision model suitable for the complexity of the system. It is thought that the study will contribute significantly to the relevant literature by shedding light on researchers working in this field.

The best performing classification algorithms in healthcare are DT, RF, and GNB, with an accuracy of 100%. However, in the majority of cases, RF and SVM exhibited the highest degree of performance. As a outcome of the study, it is seen that machine learning algorithms are mostly used in diagnosing and classifying diseases.

REFERENCES

- Abbad Ur Rehman, H., Lin, C. Y., Mushtaq, Z., & Su, S. F. (2021). Performance analysis of machine learning algorithms for thyroid disease. *Arabian Journal for Science and Engineering*, 1-13. doi: 10.1007/s13369-020-05206-x.
- Ahmed, H., Younis, E. M., Hendawi, A., & Ali, A. A. (2020). Heart disease identification from patients' social posts, machine learning solution on Spark. *Future Generation Computer Systems*, 111, 714-722. doi: 10.1016/j.future.2019.09.056.
- Al-Azzam, N., & Shatnawi, I. (2021). Comparing supervised and semi-supervised machine learning models on diagnosing breast cancer. *Annals of Medicine and Surgery*, 62, 53-64. doi: 10.1016/j.amsu.2020.12.043.
- Ali, F., El-Sappagh, S., Islam, S. R., Kwak, D., Ali, A., Imran, M., & Kwak, K. S. (2020). A smart healthcare monitoring system for heart disease prediction based on ensemble deep learning and feature fusion. *Information Fusion*, 63, 208-222. doi: 10.1016/j.inffus.2020.06.008.

- Almansour, N. A., Syed, H. F., Khayat, N. R., Altheeb, R. K., Juri, R. E., Alhiyafi, J., ... & Olatunji, S. O. (2019). Neural network and support vector machine for the prediction of chronic kidney disease: A comparative study. *Computers in biology and medicine*, 109, 101-111. doi: 10.1016/j.compbimed.2019.04.017.
- Alshayegi, M. H. (2023). Early Thyroid Risk Prediction by Data Mining and Ensemble Classifiers. *Machine Learning and Knowledge Extraction*, 5(3), 1195-1213.
- Alves, M. A., Castro, G. Z., Oliveira, B. A. S., Ferreira, L. A., Ramírez, J. A., Silva, R., & Guimarães, F. G. (2021). Explaining machine learning based diagnosis of COVID-19 from routine blood tests with decision trees and criteria graphs. *Computers in Biology and Medicine*, 132, 104335. doi: 10.1016/j.compbimed.2021.104335.
- Anon. n.d. "Cardiovascular Diseases." https://www.who.int/health-topics/Cardiovascular-Diseases#tab=tab_1. ACCESSES.2023
- Arowolo, M. O., Ogundokun, R. O., Misra, S., Agboola, B. D., & Gupta, B. (2023). Machine learning-based IoT system for COVID-19 epidemics. *Computing*, 105(4), 831-847. doi: 10.1007/s00607-022-01057-6.
- Basha, Y. H., Nassif, A. B., & AlShabi, M. (2022, June). Predicting heart failure disease using machine learning. In *Smart Biomedical and Physiological Sensor Technology XIX* (Vol. 12123, pp. 75-84). SPIE. doi: 10.1117/12.2632634.
- Bayat, V., Phelps, S., Ryono, R., Lee, C., Parekh, H., Mewton, J., ... & Holodniy, M. (2021). A severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) prediction model from standard laboratory tests. *Clinical Infectious Diseases*, 73(9), e2901-e2907. doi: 10.1093/cid/ciaa1175.
- Behera, M. P., Sarangi, A., Mishra, D., & Sarangi, S. K. (2023). A Hybrid Machine Learning algorithm for Heart and Liver Disease Prediction Using Modified Particle Swarm Optimization with Support Vector Machine. *Procedia Computer Science*, 218, 818-827. doi: 10.1016/j.procs.2023.01.062.
- Chaganti, R., Rustam, F., De La Torre Díez, I., Mazón, J. L. V., Rodríguez, C. L., & Ashraf, I. (2022). Thyroid disease prediction using selective features and machine learning techniques. *Cancers*, 14(16), 3914. doi: 10.3390/cancers14163914.
- Chaganti, R., Rustam, F., De La Torre Díez, I., Mazón, J. L. V., Rodríguez, C. L., & Ashraf, I. (2022). Thyroid disease prediction using selective features and machine learning techniques. *Cancers*, 14(16), 3914. doi: 10.3390/cancers14163914.

- Chaithra, A. S., Chandana, D. K., Chetana, S. M., & Greeshma, N. (2022, December). Risk Prediction of Chronic Kidney Disease Using Machine Learning Algorithms. In *International Conference on Data Science, Machine Learning and Applications* (pp. 333-338). Singapore: Springer Nature Singapore. doi: 10.1007/978-981-99-2058-7_30.
- Chang, H. K., Wu, C. T., Liu, J. H., Lim, W. S., Wang, H. C., Chiu, S. I., & Jang, J. S. R. (2019, June). Early detecting in-hospital cardiac arrest based on machine learning on imbalanced data. In *2019 IEEE International Conference on Healthcare Informatics (ICHI)* (pp. 1-10). doi: 10.1109/ICHI.2019.8904504.
- Chauhan, U., Kumar, V., Chauhan, V., Tiwary, S., & Kumar, A. (2019, July). Cardiac arrest prediction using machine learning algorithms. In *2019 2nd international conference on intelligent computing, instrumentation and control technologies (ICICICT)* (Vol. 1, pp. 886-890). doi: 10.1109/ICICICT46008.2019.8993296.
- Devi, M. S., Kumar, V. D., Brezulianu, A., Geman, O., & Arif, M. (2023). A Novel Blunge Calibration Intelligent Feature Classification Model for the Prediction of Hypothyroid Disease. *Sensors*, 23(3), 1128. doi: 10.3390/s23031128.
- Dietterich, T. G. (2000, June). Ensemble methods in machine learning. In *International workshop on multiple classifier systems* (pp. 1-15). Berlin, Heidelberg: Springer Berlin Heidelberg.
- Dritsas, E., & Trigka, M. (2022). Machine learning techniques for chronic kidney disease risk prediction. *Big Data and Cognitive Computing*, 6(3), 98. doi: 10.3390/bdcc6030098.
- Ebrahim, M., Sedky, A. A. H., & Mesbah, S. (2023). Accuracy Assessment of Machine Learning Algorithms Used to Predict Breast Cancer. *Data*, 8(2), 35. doi: 10.3390/data8020035.
- Febrian, M. E., Ferdinan, F. X., Sendani, G. P., Suryanigrum, K. M., & Yunanda, R. (2023). Diabetes prediction using supervised machine learning. *Procedia Computer Science*, 216, 21-30.
- Ferdousi, R., Hossain, M. A., & El Saddik, A. (2021). Early-stage risk prediction of non-communicable disease using machine learning in health CPS. *IEEE Access*, 9, 96823-96837. doi: 10.1109/ACCESS.2021.3094063.
- Franciotti, R., Nardini, D., Russo, M., Onofrj, M., Sensi, S. L., & Alzheimer's Disease Neuroimaging Initiative. (2023). Comparison of Machine Learning-based approaches to predict the conversion to Alzheimer's disease from Mild Cognitive Impairment. *Neuroscience*, 514, 143-152. doi: 10.1016/j.neuroscience.2023.01.029.

- Gambhir, E., Jain, R., Gupta, A., & Tomer, U. (2020, September). Regression analysis of COVID-19 using machine learning algorithms. In *2020 International conference on smart electronics and communication (ICOSEC)* (pp. 65-71). doi: 10.4171/automata-1/11.
- Gatos, I., Tsantis, S., Spiliopoulos, S., Karnabatidis, D., Theotokas, I., Zoumpoulis, P., ... & Kagadis, G. C. (2017). A machine-learning algorithm toward color analysis for chronic liver disease classification, employing ultrasound shear wave elastography. *Ultrasound in medicine & biology*, 43(9), 1797-1810. doi: 10.1016/j.ultrasmedbio.2017.05.002.
- Gogi, V. J., & Vijayalakshmi, M. N. (2018, July). Prognosis of liver disease: Using Machine Learning algorithms. In *2018 International Conference on Recent Innovations in Electrical, Electronics & Communication Engineering (ICRIEECE)* (pp. 875-879). doi: 10.1109/ICRIEECE44171.2018.9008482.
- Gopal, V. N., Al-Turjman, F., Kumar, R., Anand, L., & Rajesh, M. (2021). Feature selection and classification in breast cancer prediction using IoT and machine learning. *Measurement*, 178, 109442. doi: 10.1016/j.measurement.2021.109442.
- Haji, S. H., Abdulazeez, A. M., Zeebaree, D. Q., Ahmed, F. Y., & Zebari, D. A. (2021, July). The impact of different data mining classification techniques in different datasets. In *2021 IEEE Symposium on Industrial Electronics & Applications (ISIEA)* (pp. 1-6). doi: 10.1109/ISIEA51897.2021.9510006.
- Hassan, N. S., Abdulazeez, A. M., Saeed, J. N., Zeebaree, D. Q., Al-Zebari, A., & Ahmed, F. Y. (2021, July). A compassion of three data mining algorithms for heart disease prediction. In *2021 IEEE Symposium on Industrial Electronics & Applications (ISIEA)* (pp. 1-6). doi: 10.1109/ISIEA51897.2021.9509985.
- He, Z. L., Zhou, J. B., Liu, Z. K., Dong, S. Y., Zhang, Y. T., Shen, T., ... & Xu, X. (2021). Application of machine learning models for predicting acute kidney injury following donation after cardiac death liver transplantation. *Hepatobiliary & Pancreatic Diseases International*, 20(3), 222-231. doi: 10.1016/j.hbpd.2021.02.001.
- Hirano, Y., Kondo, Y., Sueyoshi, K., Okamoto, K., & Tanaka, H. (2021). Early outcome prediction for out-of-hospital cardiac arrest with initial shockable rhythm using machine learning models. *Resuscitation*, 158, 49-56. doi: 10.1016/j.resuscitation.2020.11.020.
- Hossain, M. A., Ferdousi, R., & Alhamid, M. F. (2020). Knowledge-driven machine learning based framework for early-stage disease risk prediction in edge environment. *Journal of Parallel and Distributed Computing*, 146, 25-34. doi: 10.1016/j.jpdc.2020.07.003.

- Javan, S. L., Sepehri, M. M., Javan, M. L., & Khatibi, T. (2019). An intelligent warning model for early prediction of cardiac arrest in sepsis patients. *Computer methods and programs in biomedicine*, 178, 47-58. doi: 10.1016/j.cmpb.2019.06.010.
- Javeed, A., Saleem, M. A., Dallora, A. L., Ali, L., Berglund, J. S., & Anderberg, P. (2023). Decision Support System for Predicting Mortality in Cardiac Patients Based on Machine Learning. *Applied Sciences*, 13(8), 5188. doi: 10.3390/app13085188.
- Jayatilake, S. M. D. A. C., & Ganegoda, G. U. (2021). Involvement of machine learning tools in healthcare decision making. *Journal of healthcare engineering*, 2021. doi: 10.1155/2021/6679512.
- Jothi, K. A., Subburam, S., Umadevi, V., & Hemavathy, K. (2021). WITHDRAWN: Heart disease prediction system using machine learning. doi: 10.1016/j.matpr.2020.12.901.
- Karthikeyan, B., Gollamudi, S., Singamsetty, H. V., Gade, P. K., & Mekala, S. Y. (2020). Breast cancer detection using machine learning. *Int. J. Adv. Trends Comput. Sci. Eng*, 9(2), 981-984.
- Khan, A., & Zubair, S. (2022). An improved multi-modal based machine learning approach for the prognosis of Alzheimer's disease. *Journal of King Saud University-Computer and Information Sciences*, 34(6), 2688-2706. doi: 10.1016/j.jksuci.2020.04.004.
- Khorshid, S. F., & Abdulazeez, A. M. (2021). Breast cancer diagnosis based on k-nearest neighbors: a review. *PalArch's Journal of Archaeology of Egypt/Egyptology*, 18(4), 1927-1951.
- Khorshid, S. F., Abdulazeez, A. M., & Sallow, A. B. (2021). A comparative analysis and predicting for breast cancer detection based on data mining models. *Asian Journal of Research in Computer Science*, 8(4), 45-59. doi: 10.9734/ajrcos/2021/v8i430209.
- Kumar, K., Pradeepa, M., Mahdal, M., Verma, S., RajaRao, M. V. L. N., & Ramesh, J. V. N. (2023). A Deep Learning Approach for Kidney Disease Recognition and Prediction through Image Processing. *Applied Sciences*, 13(6), 3621. doi: 10.3390/app13063621.
- Kwon, J. M., Jeon, K. H., Kim, H. M., Kim, M. J., Lim, S., Kim, K. H., ... & Oh, B. H. (2019). Deep-learning-based out-of-hospital cardiac arrest prognostic system to predict clinical outcomes. *Resuscitation*, 139, 84-91. doi: 10.1016/j.resuscitation.2019.04.007.

- Li, J. P., Haq, A. U., Din, S. U., Khan, J., Khan, A., & Saboor, A. (2020). Heart disease identification method using machine learning classification in e-healthcare. *IEEE access*, 8, 107562-107582. doi: 10.1109/ACCESS.2020.3001149.
- Liu, X., Hasan, M. R., Ahmed, K. A., & Hossain, M. Z. (2023). Machine learning to analyse omic-data for COVID-19 diagnosis and prognosis. *BMC bioinformatics*, 24(1), 7. doi: 10.1186/s12859-022-05127-6.
- Lodha, P., Talele, A., & Degaonkar, K. (2018, August). Diagnosis of alzheimer's disease using machine learning. In *2018 fourth international conference on computing communication control and automation (ICCUBEA)* (pp. 1-4). IEEE. doi: 10.1109/ICCUBEA.2018.8697386.
- Lukmanto, R. B., Nugroho, A., & Akbar, H. (2019). Early detection of diabetes mellitus using feature selection and fuzzy support vector machine. *Procedia Computer Science*, 157, 46-54. doi: 10.1016/j.procs.2019.08.140.
- Majumder, A. K. M., ElSaadany, Y. A., Young, R., & Ucci, D. R. (2019). An energy efficient wearable smart IoT system to predict cardiac arrest. *Advances in Human-Computer Interaction*, 2019. doi: 10.1155/2019/1507465.
- Md, A. Q., Kulkarni, S., Joshua, C. J., Vaichole, T., Mohan, S., & Iwendi, C. (2023). Enhanced Preprocessing Approach Using Ensemble Machine Learning Algorithms for Detecting Liver Disease. *Biomedicines*, 11(2), 581. doi: 10.3390/biomedicines11020581.
- Mujumdar, A., & Vaidehi, V. (2019). Diabetes prediction using machine learning algorithms. *Procedia Computer Science*, 165, 292-299. doi: 10.1016/j.procs.2020.01.047.
- Mung, P. S., & Phyu, S. (2020, February). Effective analytics on healthcare big data using ensemble learning. In *2020 IEEE Conference on Computer Applications (ICCA)* (pp. 1-4). IEEE. doi: 10.1109/ICCA49400.2020.9022853.
- Neelaveni, J., & Devasana, M. G. (2020, March). Alzheimer disease prediction using machine learning algorithms. In *2020 6th international conference on advanced computing and communication systems (ICACCS)* (pp. 101-104). IEEE. doi: 10.1109/ICACCS48705.2020.9074248.
- Nguyen, B. P., Pham, H. N., Tran, H., Nghiem, N., Nguyen, Q. H., Do, T. T., ... & Simpson, C. R. (2019). Predicting the onset of type 2 diabetes using wide and deep learning with electronic health records. *Computer methods and programs in biomedicine*, 182, 105055. doi: 10.1016/j.cmpb.2019.105055.
- Olatunji, S. O., Alotaibi, S., Almutairi, E., Alrabae, Z., Almajid, Y., Altabee, R., ... & Alhiyafi, J. (2021). Early diagnosis of thyroid cancer diseases using

- computational intelligence techniques: A case study of a Saudi Arabian dataset. *Computers in Biology and Medicine*, 131, 104267. doi: 10.1016/j.compbiomed.2021.104267.
- Pourhomayoun, M., & Shakibi, M. (2021). Predicting mortality risk in patients with COVID-19 using machine learning to help medical decision-making. *Smart health*, 20, 100178. doi: 10.1016/j.smhl.2020.100178.
- Rabby, A. S. A., Mamata, R., Laboni, M. A., & Abujar, S. (2019, July). Machine learning applied to kidney disease prediction: Comparison study. In *2019 10th international conference on computing, communication and networking technologies (ICCCNT)* (pp. 1-7). IEEE. doi: 10.1109/ICCCNT45670.2019.8944799.
- Rangaswamy, U., Dharshini, S. A. P., Yesudhas, D., & Gromiha, M. M. (2020). VEPAD- Predicting the effect of variants associated with Alzheimer's disease using machine learning. *Computers in Biology and Medicine*, 124, 103933. doi: 10.1016/j.compbiomed.2020.103933.
- Ray, S., Lall, S., Mukhopadhyay, A., Bandyopadhyay, S., & Schönhuth, A. (2022). Deep variational graph autoencoders for novel host-directed therapy options against COVID-19. *Artificial Intelligence in Medicine*, 134, 102418. doi: 10.1016/j.artmed.2022.102418.
- Reddy, D. J., Mounika, B., Sindhu, S., Reddy, T. P., Reddy, N. S., Sri, G. J., ... & Kora, P. (2020). WITHDRAWN: Predictive machine learning model for early detection and analysis of diabetes. doi: 10.1016/j.matpr.2020.09.522.
- Ryzhikova, E., Ralbovsky, N. M., Sikirzhytski, V., Kazakov, O., Halamkova, L., Quinn, J., ... & Lednev, I. K. (2021). Raman spectroscopy and machine learning for biomedical applications: Alzheimer's disease diagnosis based on the analysis of cerebrospinal fluid. *Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy*, 248, 119188. doi: 10.1016/j.saa.2020.119188.
- Safa, M., & Pandian, A. (2021). Applying machine learning algorithm to sensor coupled IoT devices in prediction of cardiac stress-An integrated approach. *Materials Today: Proceedings*. doi: 10.1016/j.matpr.2021.02.698.
- Saha, P., Sadi, M. S., & Islam, M. M. (2021). EMCNet: Automated COVID-19 diagnosis from X-ray images using convolutional neural network and ensemble of machine learning classifiers. *Informatics in medicine unlocked*, 22, 100505. doi: 10.1016/j.imu.2020.100505.
- Salehi, W., Baglat, P., Gupta, G., Khan, S. B., Almusharraf, A., Alqahtani, A., & Kumar, A. (2023). An approach to binary classification of alzheimer's disease using LSTM. *Bioengineering*, 10(8), 950. doi: 10.3390/bioengineering10080950.

- Sarkar, J. P., Saha, I., Sarkar, A., & Maulik, U. (2021). Machine learning integrated ensemble of feature selection methods followed by survival analysis for predicting breast cancer subtype specific miRNA biomarkers. *Computers in Biology and Medicine*, 131, 104244. doi: 10.1016/j.combiomed.2021.104244.
- Shashikant, R., & Chetankumar, P. (2023). Predictive model of cardiac arrest in smokers using machine learning technique based on Heart Rate Variability parameter. *Applied Computing and Informatics*, 19(3/4), 174-185. doi: 10.1016/j.aci.2019.06.002.
- Shi, S., Lei, G., Yang, L., Zhang, C., Fang, Z., Li, J., & Wang, G. (2021). Using machine learning to predict postoperative liver dysfunction after aortic arch surgery. *Journal of Cardiothoracic and Vascular Anesthesia*, 35(8), 2330-2335.
- Sridhar, G. M., & Kirubakaran, A. P. (2021). Heart disease and optimal prediction of attacks using hybrid machine learning algorithm: A survey. *Materials Today: Proceedings*. doi: 10.1016/j.matpr.2020.12.865.
- Srivenkatesh, M. (2019). Performance evolution of different machine learning algorithms for prediction of liver disease. *Int. J. Innov. Technol. Explor. Eng*, 9(2), 1115-1122. doi: 10.35940/ijitee.l3619.129219.
- Suárez, M., Martínez, R., Torres, A. M., Ramón, A., Blasco, P., & Mateo, J. (2023). A Machine Learning-Based Method for Detecting Liver Fibrosis. *Diagnostics*, 13(18), 2952. doi: 10.3390/diagnostics13182952.
- Sun, P., Lu, X., Xu, C., Sun, W., & Pan, B. (2020). Understanding of COVID-19 based on current evidence. *Journal of medical virology*, 92(6), 548-551. doi: 10.1002/jmv.25722.
- Swain, D., Mehta, U., Bhatt, A., Patel, H., Patel, K., Mehta, D., ... & Manika, S. (2023). A Robust Chronic Kidney Disease Classifier Using Machine Learning. *Electronics*, 12(1), 212.
- Taylan, O., Alkabaa, A. S., Alqabbaa, H. S., Pamukçu, E., & Leiva, V. (2023). Early prediction in classification of cardiovascular diseases with machine learning, neuro-fuzzy and statistical methods. *Biology*, 12(1), 117. doi: 10.3390/biology12010117.
- Tewari, Y., Ujjwal, E., & Kumar, L. (2022, April). Breast Cancer Classification Using Machine Learning. In *2022 2nd International Conference on Advance Computing and Innovative Technologies in Engineering (ICACITE)* (pp. 01-04). IEEE.
- Thaiparnit, S., Chumuang, N., & Ketcham, M. (2018, November). A comparative study of clasification liver dysfunction with machine learning. In *2018 International*

Joint Symposium on Artificial Intelligence and Natural Language Processing (iSAI-NLP) (pp. 1-4). IEEE.

- Tigga, N. P., & Garg, S. (2020). Prediction of type 2 diabetes using machine learning classification methods. *Procedia Computer Science*, 167, 706-71. doi: 10.1016/j.procs.2020.03.336.
- Tyagi, A., Mehra, R., & Saxena, A. (2018, December). Interactive thyroid disease prediction system using machine learning technique. In *2018 Fifth international conference on parallel, distributed and grid computing (PDGC)* (pp. 689-693). IEEE. doi: 10.1109/PDGC.2018.8745910.
- Uddin, M. A., Islam, M. M., Talukder, M. A., Hossain, M. A. A., Akhter, A., Aryal, S., & Muntaha, M. (2023). Machine Learning Based Diabetes Detection Model for False Negative Reduction. *Biomedical Materials & Devices*, 1-17. doi: 10.1007/s44174-023-00104-w.
- Uysal, G., & Ozturk, M. (2020). Hippocampal atrophy based Alzheimer's disease diagnosis via machine learning methods. *Journal of Neuroscience Methods*, 337, 108669. doi: 10.1016/j.jneumeth.2020.108669.
- Vaka, A. R., Soni, B., & Reddy, S. (2020). Breast cancer detection by leveraging Machine Learning. *Ict Express*, 6(4), 320-324. doi: 10.1016/j.ict.2020.04.009.
- Viloria, A., Herazo-Beltran, Y., Cabrera, D., & Pineda, O. B. (2020). Diabetes diagnostic prediction using vector support machines. *Procedia Computer Science*, 170, 376-381. doi: 10.1016/j.procs.2020.03.065.
- Waris, S. F., & Koteeswaran, S. (2021). WITHDRAWN: Heart disease early prediction using a novel machine learning method called improved K-means neighbor classifier in python. doi: 10.1016/j.matpr.2021.01.570.
- Wu, C. C., Yeh, W. C., Hsu, W. D., Islam, M. M., Nguyen, P. A. A., Poly, T. N., ... & Li, Y. C. J. (2019). Prediction of fatty liver disease using machine learning algorithms. *Computer methods and programs in biomedicine*, 170, 23-29. doi: 10.1016/j.cmpb.2018.12.032.
- Wu, J., & Hicks, C. (2021). Breast cancer type classification using machine learning. *Journal of personalized medicine*, 11(2), 61. doi: 10.3390/jpm11020061.
- Zhang, F., Tian, S., Chen, S., Ma, Y., Li, X., & Guo, X. (2019). Voxel-based morphometry: improving the diagnosis of Alzheimer's disease based on an extreme learning machine method from the ADNI cohort. *Neuroscience*, 414, 273-279. doi: 10.1016/j.neuroscience.2019.05.014.