Information Sciences with Applications, Vol. 3, 2024
https://doi.org/10.61356/j.iswa.2024.3232

Adaptive Analysis of Machine Learning Algorithms on Medical Datasets for Disease Prediction

Salwa Elsayed [1,*], Mahmoud M. Ismail [1], Amal F. Abdel-Gawad [1], and Israa Mohamed [2]

[1] Department of Decision Support, Faculty of Computers and Informatics, Zagazig University, Zagazig, 44519, Egypt.
Emails: salwa.elsayed@fci.zu.edu.eg, mmsabe@zu.edu.eg, amgawad2001@yahoo.com, afbadaawi@zu.edu.eg.
[2] Faculty of Engineering and Computer Sciences, King Salman International University, South Sinai, Egypt;
israasalem@zu.edu.eg.

* Correspondence: salwa.elsayed@fci.zu.edu.eg.

Abstract: This paper aims to analyze and evaluate the performance of different classification algorithms on medical datasets and gives some tips and guidelines for choosing the right algorithm for the dataset. For this purpose, twelve medical datasets were selected from the UCI repository, including breast cancer data, chronic kidney disease, cryotherapy, hepatitis, immunotherapy, Indian Liver Patient Dataset (ILPD), liver disorders, pima diabetes, risk factors cervical cancer, statlog (heart) dataset, dermatology and lung cancer. The selected classifiers for performance evaluation are Bagging, IBK, J48, JRip, Multilayer Perceptron (MP), Naïve Bayes (NB), Support Vector Machine (SVM), ZeroR and Voting Frequency Intervals (VFI). The algorithms were implemented on datasets using the WEKA tool and 10-fold cross-validation. Experimental results and analysis for all measures were introduced to guideline researchers who seek to achieve and how to balance different measures. Additionally, challenges and future directions were presented at the end of this paper.

Keywords: Medical Datasets; Machine Learning Algorithms; Support Vector Machine; Bagging; Voting Frequency Intervals; Naïve Bayes.

1. Introduction

Machine learning algorithms are very powerful tools for analyzing, diagnosing, and predicting diseases. These algorithms proved their ability to solve many problems and predict diseases by obtaining higher accuracy than other methods, so, many researchers have paid attention to apply them. The healthcare industry collects a huge amount of data that is not correctly mined with high quality. The electronic health record (EHR) is one way to keep the full history of the patient records to analyze the data for the future. The EHR includes significantly large data in the form of X-rays and key observations by doctors about fitness and vital signs. Modern big data EHR systems contain different channels of data sources from pharmacies, nursing, radiology units, and hospitals through a network. There are a number of models that must be filled through administrative work to register in a traditional environment of the clinic. However, EHR can automate a large part of this management business, including patient admission and exit models, invoices, check-in and out-of-patient bills, and patient demographics. The laboratory systems, pharmacy systems, computerized physician application input systems, and coding systems for regulating healthcare data are integrated into certain categories of effective analysis and radiology systems in central EHR systems. A number of machine learning algorithms can be applied to various virtual profiles with the help of coding systems to diagnose diseases such as atrial fibrillation, dementia, clopidogrel, and type 2 diabetes. By analyzing these data and extracting hidden patterns, more useful information will be obtained that will help medicines make the correct decision in real-time. Despite the diverse work in machine learning algorithms, no work has been done to determine exactly which algorithm will be more suitable for the dataset on hand [1-7]. In healthcare, sensors constantly collect big data from the subject to identify patterns in dataset parts and to further process these data through machine learning algorithms. Decision trees play an important role in healthcare. Decision trees are heavily utilized in the diagnosis of diseases in healthcare. In some cases, diagnosis requires constant monitoring of autonomic neuropathy. The identification of cardiovascular autonomic neuropathy through sensor data is the key to understanding the vital signs of diabetes. Analysis of these data can be carried out through decision trees and group methods. This analysis helps provide advanced diet and treatment plans for the subject.

Figure 1 shows different types of machine learning [8]. As seen in Figure 1, machine learning can be divided into four types: supervised learning, unsupervised learning, semi-supervised learning, and reinforcement learning. Supervised learning includes employing a model to develop a mapping among input samples and the target variable. There are two essential types of supervised learning: classification, which includes predicting a class label, and regression, which includes predicting a numerical value. There are many supervised algorithms, such as naïve Bayes, decision trees, linear regression, support vector machines, and many others. Unsupervised learning is a data-driven model that forecasts results using unlabeled data as input. The two essential types of unsupervised learning are clustering and association. Semi-supervised learning is supervised learning in which the training data involve too little labeled data and a large number of unlabeled data. Unlike supervised learning, semi-supervised learning aims to make efficient use of all data, not simply labeled data. Semi-supervised learning can be used in classification and clustering. Reinforcement learning is when an agent works in an environment and should learn to work using feedback. Reinforcement learning can be used in classification and control. An implementation of a reinforcement problem is playing a game where the agent's objective is to score well and make moves while receiving feedback in the form of penalties or rewards.

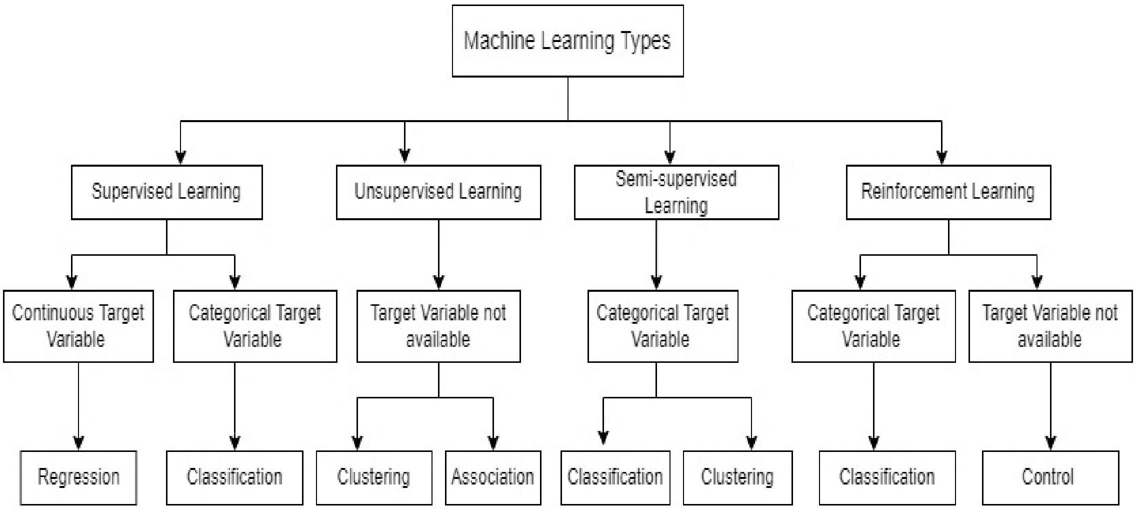

Figure 1. Types of machine learning algorithms [8].

2. Literature Review

The selected classifiers in this study are bagging, IBK, J48, JRip, multilayer perceptron (MP), naïve Bayes (NB), support vector machine (SVM), ZeroR, and voting frequency intervals (VFI). These algorithms were briefly reviewed. Bagging, also known as bootstrap aggregation, is an ensemble learning method that is typically used to minimize contrast within a noisy dataset. In bagging, a random sample of data is selected in a training group with replacement, meaning that individual data points can be selected more than once. After creating many data samples, these models are trained independently, depending on the type of task regression or classification [9]. The IBk algorithm makes just-in-time predictions for a test instance. IBk utilizes a distance measure to locate k “close” instances in the training data for every test instance and employs those selected instances to create a prediction. IBK has two main advantages: it is fast and can give higher accuracy than other algorithms [10]. The J48 algorithm is one of the best machine learning algorithms for categorically and continuously examining data. The algorithm involves a technique called imputation, which addresses missing values and helps in solving the problem of missing values. The algorithm also provides a process of tree pruning, which is conducive to creating small trees and preventing overfitting of the data. Additionally, the algorithm offers the subtree replacement process, which reduces the classification error after overwriting the subtree with a leaf [11].

JRip uses a propositional rule learner named (Repeated Incremental Pruning to Produce Error Reduction (RIPPER)) that guarantees errors to be minimal. The algorithm has four main stages: growing a rule, pruning, optimization, and selection. It takes the instances in the dataset and evaluates them in increasing order. Another merit is that the algorithm becomes shorter and simpler by trimming any useless parts of the rule. This makes the algorithm easier to understand and interpret than other algorithms [12]. Multilayer perceptrons (MPs) are a class of feed-forward artificial neural networks (ANNs). MP models are the most basic deep neural networks and consist of three types of layers: the input layer, output layer, and hidden layer. Prediction and classification are performed by the output layer. MP can be used to deal with the high computing power requirements of modern deep learning architectures [13]. The naïve Bayes (NB) algorithm is a supervised learning algorithm based on the Bayes theorem. It predicts membership probabilities for each class such as the probability that a given data point pertains to a particular class. The class with the highest probability is considered the most likely class. NB has many advantages, such as it is fast and usually ready to learn, it requires less training data, it is highly scalable, it can deal with both continuous and discrete data and it is easy to update on the arrival of new data [14].

Support vector machine (SVM) is a supervised machine learning algorithm that can be used for both classification and regression. However, it is mainly used in classification problems. In the SVM algorithm, we draw each data element as a point in n-dimensional space (where n is the number of features you have) with the value of each parameter being a certain coordinate value. After that, we make the classification by finding the superior level that distinguishes the two categories [15]. ZeroR is the simplest target-based rating method and ignores all forecasters. ZeroR simply predicts the majority class. Although there is no predictive power in ZeroR, it is useful for determining baseline performance as a standard for other classification methods [16]. The voting frequency intervals (VFI) algorithm is a classification algorithm that depends on the principle of voting frequency intervals. In VFI, each training instance is represented as a vector of features. Then, feature intervals are created for every feature. A set of values for a given feature where the same subset of class values is observed and represented by an interval. Therefore, two neighboring intervals constitute different classes. This algorithm has two main phases: the training phase and the classification phase [17].

In this paper, a comprehensive review of the classification of 12 medical datasets taken from the UCI machine learning repository [18] is introduced. The selected datasets are breast cancer data, chronic kidney disease, cryotherapy, hepatitis, immunotherapy, Indian Liver Patient Dataset (ILPD), liver disorders, pima diabetes, risk factor cervical cancer, statlog (heart) dataset, dermatology, and lung cancer. We will present an overview of the selected dataset and the most recent research on it.

For breast cancer, breast cancer is the abnormal growth of cells coating breast lobes or channels. These cells increase uncontrollably and can extend to other parts of the body. Both men and women can get hurt with breast cancer, but it is not common in men. In this paper, [19], the authors showed how DNA methylation can play a vital role in organizing gene expression and may aid as a biomarker for treatment response and survival from breast cancer. In this research, an experiment was performed on 83 women with locally advanced breast cancer, and the results showed that DNA methylation analysis could be used as a predictive tool to forecast survival outcomes in breast cancer patients. In this paper, [20], a comprehensive study was carried out on 63,424 American women using surveillance and epidemiology. The goal of this paper was to compare mortality after a second primary cancer (SPC) among breast cancer survivors to mortality among women after a first primary cancer (FPC). The results showed that mortality after certain SPCs would be higher than mortality after FPCs of the same type. In this review paper [21], the authors showed that aging parameters that result from a geriatric assessment can forecast survival, toxicity, and patient-reported outcomes in older patients with breast cancer. This can help in clinical practice every day to determine patients at risk of primary mortality, high risk of treatment toxicity, or weak functional outcome after treatment. Many papers have attempted to analyze breast cancer disease to find corresponding treatments [22-24].

Chronic kidney disease (CKD): Chronic kidney disease is a kind of kidney disease in which there is continuing loss of renal function over a period of time. Early on, there are typically no symptoms. If someone has CKD, he will be at increased risk of heart attack, particularly if he smokes or is weighty. The authors of this paper [25] showed that there is growing evidence that the path and repercussions of chronic kidney disease are different in men and women, and even the points of separation to determine chronic kidney disease may vary. However, the same methods have been used to diagnose CKD in men and women until now. CKD was more prevalent in men than in women, and the majority of people with CKD were in the 45-64 age group. According to this study [26], the latest assessment of the number of people who had chronic kidney disease dates back to 2010, when 14% of Spanish adults (6.7 million) had CKD. The Global Burden of Disease (GBD) predicted that chronic kidney disease will be the fifth-largest global cause of death by 2040 [27]. The fast growth in the number of people with CKD leads to growth in the number of lives lost, which is a large risk to the world [28, 29].

For cryotherapy, cryotherapy is the use of severe cold to freeze and eliminate abnormal tissues. Doctors use it to treat several skin diseases (such as warts and skin marks) and some cancers, including prostate, cervical, and liver cancers. This treatment is also called cryoablation. The authors of this paper [30] aimed to examine whether cryotherapy was chosen as a mock procedure to minimize symptoms of chronic rhinitis. The results showed that cryotherapy seems to be effective in patients with high symptoms who were resistant to other medical and surgical treatments, regardless of their rhinitis causes. Additionally, in this article [31], the authors ensured that the use of cryotherapy during the aggravation period of the experimental gout model reduced joint edema, enhanced the strength of the grip of the limbs, and reduced the total number of white blood cells. There are many articles in which authors applied cryotherapy for different diseases with more positive results [32, 33].

For hepatitis: hepatitis refers to inflammation of the liver. The liver is an important organ that handles nutrients, filters the blood, and struggles with infections. When the liver is inflamed, its function is reduced. There are many situations that can lead to hepatitis, such as excessive alcohol, toxins, certain medications and some medical conditions. However, hepatitis usually occurs due to a virus. In the United States, the most public kinds of viral hepatitis are hepatitis A, hepatitis B, and hepatitis C. In this research [34], the authors presented a study that compared cases of children with idiopathic acute hepatitis with basic case numbers in the previous 5 years. Additional epidemiological, immunological and clinical studies are needed using metagenomic sequencing and other techniques, such as immunomodulation, RNA expression characterization and toxicological analysis, to identify pathogens, risk factors and the evolution of this continuing epidemic. Additionally, this review [35] presented a broad outline for autoimmune hepatitis (AIH) and displayed AIH type 1 and AIH type 2. The results of this review concluded that most patients respond well to standard immunosuppressive treatment. You should take a look at these articles [36-38] to know more about hepatitis disease and its different types.

For immunotherapy: immunotherapy or biological therapy is disease treatment that amplifies or suppresses the immune system. Immunotherapies used to provoke the immune response are classified as stimulating immunotherapy, while immunotherapies that suppress the immune response are classified as repressive immunotherapy [39]. The Figure explains how immunotherapy can be used for cancer treatment. This review [40] presented an overview of the options for immunotherapy in patients with COVID-19. The authors of this paper aimed to support clinical decision-making and to aid in examining immunotherapy in other acute infections by introducing an algorithm that helps clinicians in using immunotherapy strategies. Immunotherapy is used in the treatment of different diseases, such as hepatocellular carcinoma, tumors, and cancer [41-43].

For the Indian liver patient dataset, liver disease is the tenth most common cause of death in India. According to the World Health Organization (WHO), liver disease may hurt every Indian in 5 Indians, posing a great danger to the lives of people in India [44]. ILPD is a commonly used open-source dataset that involves biochemical markers in a group of patients. A number of these patients had liver diseases, and the others did not. Many papers have introduced ILPD to predict liver disease. In this paper, [45], the authors presented an overview of classification models based on this dataset and concluded that logistic regression (LR) and random forest (RF) models can give the best results with accuracy greater than 70%. Additionally, in this article [46], the authors tried random forest, support vector machine, and AdaBoost algorithms, with random forest providing the greatest accuracy of 93%. According to this recent research [47], algorithms that predict liver diseases require scrutinizing possible biases. Therefore, the authors discussed sex disparities that exist in ILPD and concluded that biases should be evaluated in the early stages of machine learning to give a vision of inequalities in existing clinical practice. ILPD has received a great deal of attention as an attempt to achieve the highest accuracy in predicting liver disease [48-50].

For liver disorders, the liver is the body's largest internal organ. The liver should be safeguarded from disease since it is frequently exposed to dangerous substances. Hepatitis, cirrhosis, liver tumors, liver abscesses, and other conditions are examples of different forms of liver disorders. In this research [51], the Japanese Society of Gastroenterology (JSGE) and the Japan Society of Hepatology (JSH) introduced a new version of the clinical practice guidelines for liver cirrhosis as a joint recommendation. This condensed version discussed the problems with treating liver cirrhosis and its various complications in clinical settings. For liver tumors, in this review paper [52], the primary pathogenic mechanisms of benign liver tumors were explained, along with their therapeutic implications. This study concluded that a thorough evaluation by radiologists, hepatologists, physicians, and pathologists is necessary for the proper functioning of benign liver tumors. There are various mechanisms used for the treatment of liver tumors [53, 54].

For Pima Indian diabetes: diabetes is a long-term disease in which the body becomes resistant to the hormone insulin, which turns food into glucose. The objective is to predict the start of diabetes. Diabetes, which is typically classified into Type 1 and Type 2 diabetes, affects a large number of individuals globally [55]. The authors of this paper [56] presented the Internet of Medical Things (IoMT) environment as a platform for an electronic diagnosis system based on machine learning (ML) algorithms, specifically for the diagnosis of diabetes mellitus (type 2 diabetes). Using the Pima Indians diabetes dataset, naïve Bayes, random forest, and J48 decision tree models were trained and tested. The results concluded that the random forest classifier can beat both the naïve Bayes and J48 decision trees with accuracy (79.57%). Additionally, the authors of this article [57] proposed an effective prediction approach for diabetes mellitus prediction by utilizing a deep 1D convolutional neural network (DCNN) on an unbalanced dataset with missing values. The suggested algorithm was trained and tested on the Pima Indian Diabetes dataset, and the results concluded that DCNN can outperform the other selected machine learning algorithms. The authors of this study [58] implemented binary logistic regression, multilayer perceptron, and k-nearest neighbor for diabetes datasets. The results of this study showed that the k-nearest neighbor algorithm is the best algorithm compared to other algorithms. Different methods have been proposed to predict Pima Indian diabetes and obtain the best results [59, 60].

Cervical cancer, a cancer that starts in the cervix is called cervical cancer. It is caused by cells that can invade or disseminate to different places of the body growing abnormally. A risk factor is something that raises a person's likelihood of having a specific illness, such as cancer. Some risk factors can be changed, such as environmental or lifestyle risk factors, while others, such as hereditary and family history risk factors, cannot. Figure 2 shows the common risk factors for cervical cancer [61]. This paper [62] sought to assess cervical cancer's most recent global prevalence, risk factors, and temporal trends for various nations and age categories. The authors concluded that cervical cancer is more prevalent in places with low and middle human development index (HDI). Cervical cancer has become more common and more lethal in some populations, although its overall burden has been reduced. Additionally, the main goal of this research [63] is to reveal potential predictors of defined risk factors such as age, cancer stage, smoking status, number of partners, usage of contraceptives, and yearly gynecological exams. The research explored and evaluated the levels of anxiety and depression in patients with cervical cancer. The results showed that patients with cervical cancer are at a higher risk of developing psychological disorders than the general population; hence, screening programs are necessary for them. These papers [64, 65] present more details about women's knowledge of cervical cancer risk factors and perioperative venous thromboembolism in cervical cancer patients.

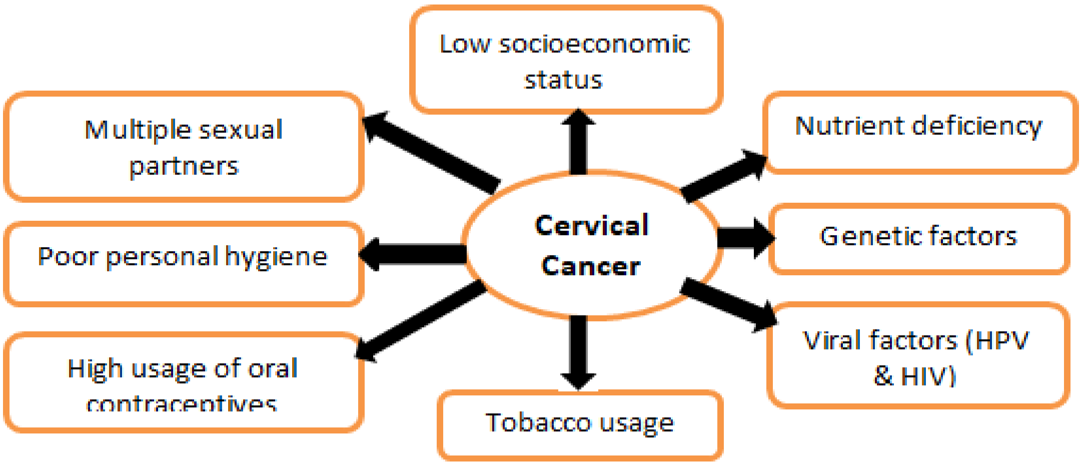

Figure 2. Common risk factors of cervical cancer [61].

For heart disease, a variety of heart-related disorders are referred to as heart diseases. The following are some examples of heart diseases: blood vessel disease, such as coronary artery disease, arrhythmias, congenital heart defects, heart valve disease, heart muscle disease, and heart infection. A heart attack is an injury to the heart muscle brought on by a decrease in blood flow brought on by artery blockages. The authors of this review paper [66] tried to cover the epidemiology and pathology of the most common cardiovascular diseases (CVDs) worldwide. Additionally, they show how well-known risk factors affect the emergence of CVDs and the efficacy of preventive measures. Additionally, methods for preventing and treating CVDs are reviewed. According to this research [67], there are significant gender variations in conventional and psychological risk factors as well as age of onset. Some risk factors, such as diabetes, hypertension, hypercholesterolemia, and obesity, raise the risk of CVD in women more than they do in males, and socioeconomic and psychological factors also appear to have a stronger effect on CVD in women. Additionally, the authors of this paper [68] proposed a machine-learning framework to predict heart disease using different algorithms. The framework is executed using five algorithms: RF, NB, SVM, Hoeffding Decision Tree, and Logistic Model Tree (LMT). The results showed that RF can give the highest accuracy for predicting heart disease. Currently, many papers can advise humans to avoid heart disease or at least reduce the number of people with the disease, and these articles also provide more mechanisms for the treatment of heart diseases [69, 70].

For dermatology: the area of medicine that deals with the skin is called dermatology. Dermatology is the diagnosis and treatment of conditions affecting the skin, hair, or nails. The authors of this research [71] want to provide a summary of the ethical considerations that are currently being debated concerning dermatological machine learning healthcare applications (ML-HCAs). They discovered that ML-HCAs provide stakeholder-specific risks that must be taken into account individually for patients, medical professionals, and society. Additionally, this paper [72] presented a review of dermatology and cosmetology approaches to acne scar reduction techniques. The results showed that recent dermatology and cosmetic surgeries employ procedures with few side effects so that patients can quickly resume their normal daily activities. Atrophic scars can be reduced with the help of appropriate dermatological care, good skincare practices, and prompt application of cosmetological procedures. Skin diseases are many and varied, as we have noticed from the figure, and these diseases have a great lack of research between diagnosis, prevention methods, and treatment methods [73, 74].

For lung cancer, a cancer that starts in the cells of the lung is called lung cancer. It results in chest pain, hoarseness, headaches, difficulty breathing, bloody coughing, and weight loss. To date, there is no clear explanation for the high mortality rates from lung cancer, and the incidence of people with multiple primary lung cancers (MPLCs) is increasing [75]. Additionally, it is very important to distinguish between MPLC and intrapulmonary lung cancer metastasis to ensure the correct diagnosis. In 2020, an important study was presented for inherited lung cancer, which showed that the essential risk factor related to this disease is tobacco smoke, but approximately 10%–25% of lung cancer cases involve people who have never smoked [76]. Therefore, in this study, the authors discussed previous analyses of lung cancer in non-smokers and concluded that many affected patients exhibit an obvious inherited propensity, and multiple genes have been linked to an elevated chance of developing lung cancer. According to [77], a predictive length of stay (LOS) framework for lung cancer patients using machine learning techniques was introduced. The framework proposed to deal with imbalanced datasets for classification-based approaches using electronic healthcare records. Many authors have shown the risk of smoking, which is related to lung cancer, and others have presented some recommendations for the treatment of lung cancer [78, 79].

3. Methodology

First, a comprehensive review of machine learning algorithms was performed. After that, twelve medical datasets were chosen from the UCI repository, and nine classifiers were selected for performance evaluation. The experiment was performed using the Waikato Environment for Knowledge Analysis (WEKA) tool. After loading data on Weka, data were preprocessed to convert various formats into the required ones. Then, the selected classifiers were implemented on datasets with 10-fold cross-validation for testing. After that, different accuracy measures were obtained from the results, such as accuracy, precision, recall, F-measure, mean absolute error, ROC area, and PRC area, and reported for comparison. Finally, the best algorithm was determined for each dataset based on accuracy measures.

4. Results and Analysis

Tables (1-7) and Figures (3-9) show the results of different accuracy measures obtained from implementing the selected classifiers on the selected datasets. In machine learning, accuracy is defined as the ratio of correct predictions to all predictions made. This seems to be enough as a measure of the performance of the machine learning model, which, however, turns out to be incomplete upon closer examination. Accuracy can be calculated as follows:

$$\text{Accuracy} = \frac{TP+TN}{TP+TN+FP+FN} \tag{1}$$

Where TP = true positives, TN = true negatives, FP = false positives, and FN = false negatives. As illustrated in Table 1 and Figure 3, the most suitable algorithm for breast cancer data is SVM with an accuracy of 96.57%, which is higher than the accuracies of the other algorithms. For chronic kidney disease, VFI can beat all other algorithms and obtain an accuracy of 100%. Bagging is the most suitable algorithm for cryotherapy, with an accuracy of 93.33%. According to hepatitis, SVM and VFI give the same accuracy of 85.16%, which is the highest among all other algorithms. For immunotherapy, bagging is better than the other algorithms, with an accuracy of 84.44%. SVM and ZeroR are more suitable for ILPD with the same highest accuracy of 71.36%. For liver disorders, MP can achieve higher performance than other algorithms, with an accuracy of 71.59%. According to Pima Diabetes, NB has the highest accuracy of 76.3%. For the risk factors for cervical cancer, bagging, and JRip yielded the highest accuracy of 96.15%. SVM can give an accuracy of 84.07%, which is the highest among the algorithms for the statlog dataset. For dermatology, NB can give the highest accuracy of 97.27%. Finally, bagging and VFI give the same highest accuracy of 81.25% for lung cancer.

Table 1. Accuracy of the selected algorithms on medical datasets.

Dataset/Algorithm	Bagging	IBK	J48	JRip	MP	NB	SVM	ZeroR	VFI
1-Breast Cancer Data	95.85	95.14	94.56	96.28	95.85	95.99	**96.57**	65.52	96.43
2-Chronic Kidney Disease	98.75	95.75	99	97.75	99.75	95	62.50	62.50	**100**
3-Cryotherapy	88.89	90	**93.33**	87.78	87.78	83.33	55.56	53.33	91.11
4-Hepatitis	64.52	80.64	83.87	63.23	62.58	84.52	**85.16**	79.35	**85.16**
5-Immunotherapy	**84.44**	70	82.22	82.22	80	76.67	78.89	78.89	75.56
6-Indian Liver Patient Dataset	69.3	64.49	68.78	66.38	68.95	55.75	**71.36**	**71.36**	65.87
7-Liver Disorders	69.57	62.9	68.7	66.67	**71.59**	55.36	58.26	57.97	60.29
8-Pima diabetes	75.78	70.18	73.83	76.04	75.39	**76.3**	65.10	65.10	64.45
9-Risk factors cervical cancer	**96.15**	94.41	95.1	**96.15**	94.76	88.69	93.59	93.59	91.72
10-Statlog (Heart) Dataset	80	75.19	76.67	80.74	77.41	83.7	**84.07**	55.55	80
11-Dermatology	36.89	94.53	93.99	88.79	96.45	**97.27**	95.35	30.60	95.90
12-Lung-cancer	**81.25**	68.75	78.12	78.13	65.63	78.12	65.62	71.87	**81.25**

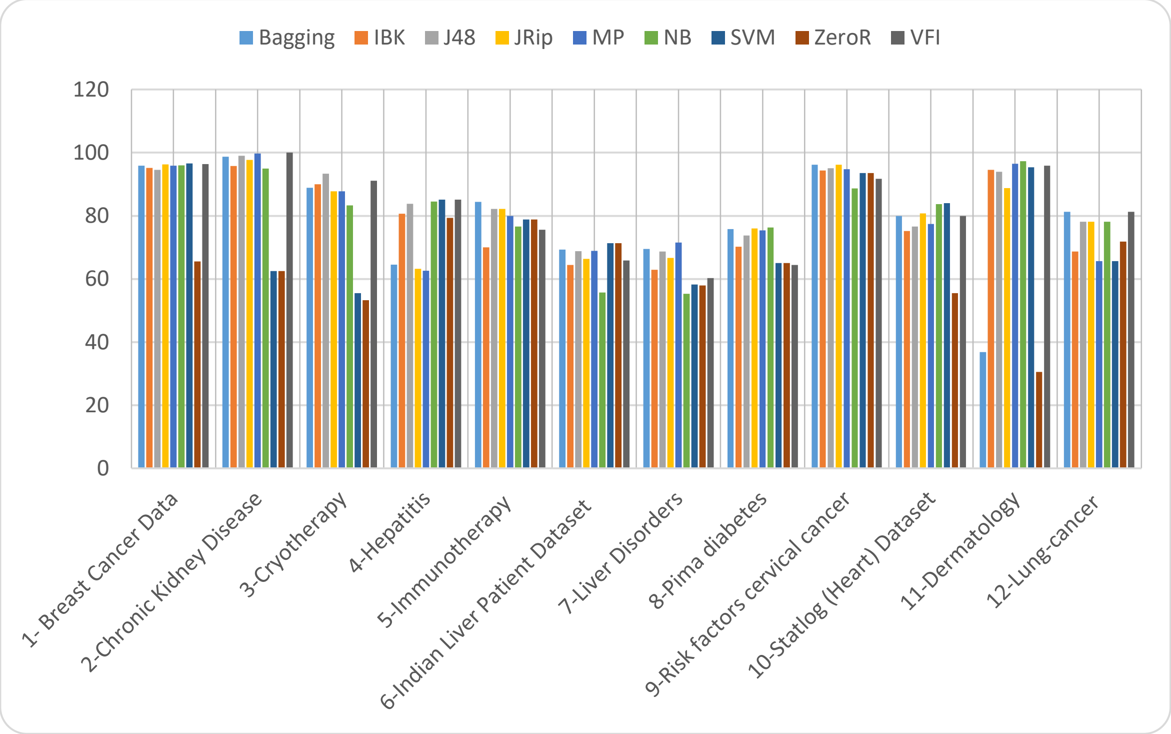

Figure 3. Accuracy of the selected algorithms on medical datasets.

According to precision, precision is defined as the ratio of correct positive predictions of all cases classified as positive and is calculated as follows:

$$\text{Precision} = \frac{TP}{TP+FP} \quad (2)$$

As seen from Table 2 and Figure 4, bagging can give the best precision for lung cancer, with a precision of 0.851, while JRip can give better results for cryotherapy and Pima diabetes than other algorithms, with precision values of 1.000 and 0.684, respectively. NB gives the best precision for dermatology with 0.974, while SVM gives better precision for immunotherapy, Indian Liver Patient Dataset, liver disorders, risk factors cervical cancer, and statlog (heart) dataset with 0.789, 0.714, 0.757, 0.936 and 0.841, respectively. ZeroR gives the best precision for immunotherapy, the Indian Liver Patient Dataset, and risk factors for cervical cancer, with values of 0.789, 0.714, and 0.936, respectively. VFI can give better precision for breast cancer data, chronic kidney disease, and hepatitis with 0.967, 1.000 and 0.861, respectively, but IBK, J48, and MP failed to obtain any better results for the datasets. Precision determines the proportion of correct positive predictions, so it is very important because high precision permits doctors to determine treatments based on the correct understanding of the patient's disease.

Table 2. Precision of the selected algorithms on medical datasets.

Dataset/Algorithm	Bagging	IBK	J48	JRip	MP	NB	SVM	ZeroR	VFI
1- Breast Cancer Data									
2-Chronic Kidney Disease	0.927	0.937	0.918	0.928	0.931	0.914	0.966	0.655	0.967
3-Cryotherapy	0.974	0.898	0.993	0.979	0.993	0.882	0.625	0.625	1.000
4-Hepatitis	0.952	0.915	0.977	1.000	0.911	0.811	0.758	0.533	0.911
5-Immunotherapy	0.632	0.794	0.825	0.618	0.581	0.853	0.847	0.630	0.861
6-Indian Liver Patient	0.692	0.250	0.60	0.60	0.539	0.40	0.789	0.789	0.765
Dataset	0.444	0.40	0.441	0.337	0.435	0.389	0.714	0.714	0.692
7-Liver Disorders	0.715	0.682	0.702	0.699	0.726	0.702	0.757	0.336	0.596
8-Pima diabetes	0.677	0.579	0.632	0.684	0.659	0.677	0.651	0.651	0.652
9-Risk factors for cervical	0.662	0.578	0.607	0.649	0.60	0.331	0.936	0.936	0.922
cancer	0.820	0.782	0.788	0.810	0.807	0.839	0.841	0.309	0.807
10-Statlog (Heart) Dataset	0.352	0.947	0.94	0.888	0.965	0.974	0.954	0.940	0.959
11-Dermatology	0.851	0.667	0.768	0.768	0.645	0.775	0.645	0.517	0.813
12-Lung-cancer									

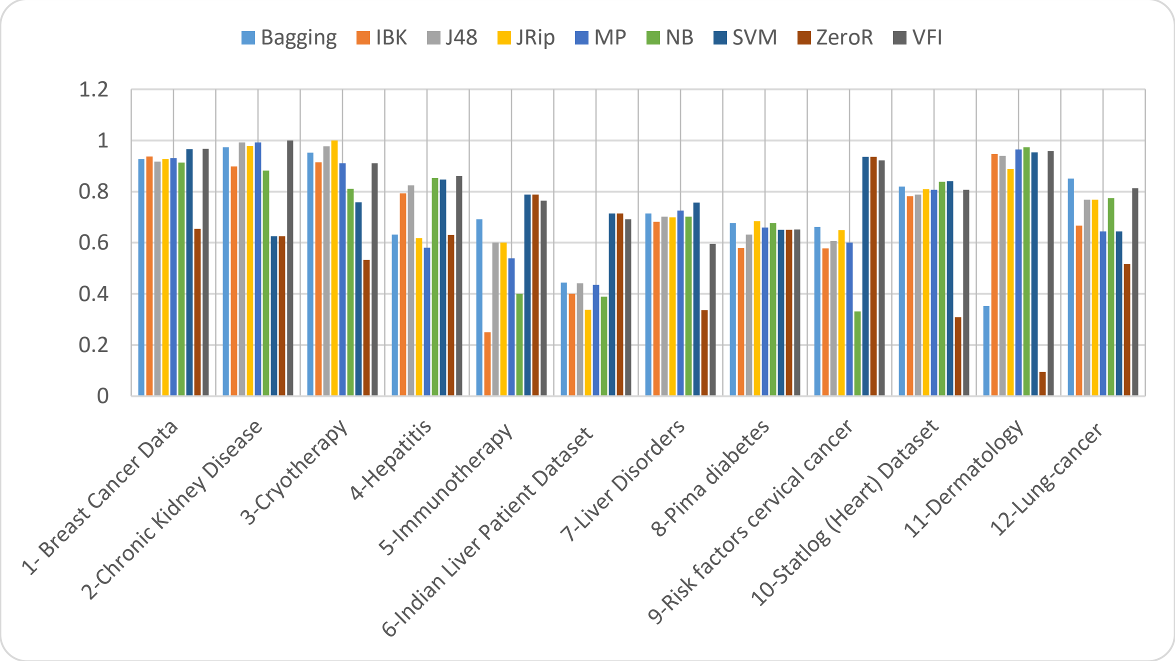

Figure 4. Precision of the selected algorithms on medical datasets.

Rendering to recall, recall measures the ratio of actual positive labels that the model has correctly selected. Recall is calculated as follows:

$$\text{Recall} = \frac{TP}{TP+FN} \tag{3}$$

As shown in Table 3 and Figure 5, bagging gave the best recall for lung cancer, with a recall of 0.813, while J48 gave better results for the Indian Liver Patient Dataset, liver disorders, and risk factors for cervical cancer than the other algorithms, with values of 0.714, 0.687 and 0.962, respectively. MP obtained the best recall for cryotherapy, immunotherapy and Pima diabetes with 0.978, 0.800 and 0.684, respectively, while NB gave the best recall for breast cancer data and dermatology with 0.974 and 0.973, respectively. SVM gives better recall for hepatitis, Indian Liver Patient Dataset and statlog (heart) dataset than other algorithms with 0.852, 0.714 and 0.841, respectively while ZeroR gives better recall for only Indian Liver Patient Dataset with 0.714. VFI can obtain better recall for chronic kidney disease, hepatitis and lung cancer with 1.000, 0.852 and 0.813, respectively, but IBK and JRip failed to obtain any better results for the datasets. Recall and Precision frequently clash. In other words, raising one score may result in a drop in the other. Therefore, the choice of whether to increase recall or precision is situational and greatly influenced by the kind of problem being addressed.

Table 3. Recall of the selected algorithms on medical datasets.

Dataset/Algorithm	Bagging	IBK	J48	JRip	MP	NB	SVM	ZeroR	VFI
1- Breast Cancer Data	0.665	0.953	0.927	0.933	0.966	0.974	0.966	0.655	0.964
2-Chronic Kidney Disease	0.998	0.998	0.995	0.998	0.995	0.988	0.625	0.625	1.000
3-Cryotherapy	0.867	0.900	0.889	0.856	0.978	0.911	0.556	0.533	0.911
4-Hepatitis	0.800	0.806	0.839	0.813	0.655	0.845	0.852	0.794	0.852
5-Immunotherapy	0.756	0.778	0.789	0.789	0.800	0.756	0.789	0.789	0.756
6-Indian Liver Patient Dataset	0.671	0.667	0.714	0.703	0.545	0.626	0.714	0.714	0.659
7-Liver Disorders	0.581	0.629	0.687	0.552	0.471	0.554	0.583	0.58	0.603
8-Pima diabetes	0.595	0.654	0.651	0.667	0.684	0.678	0.651	0.651	0.645
9-Risk factors cervical cancer	0.957	0.949	0.962	0.960	0.952	0.922	0.936	0.936	0.917
10-Statlog (Heart) Dataset	0.581	0.752	0.767	0.756	0.804	0.837	0.841	0.556	0.8
11-Dermatology	0.369	0.945	0.94	0.888	0.964	0.973	0.954	0.306	0.959
12-Lung-cancer	0.813	0.688	0.781	0.781	0.656	0.781	0.656	0.719	0.813

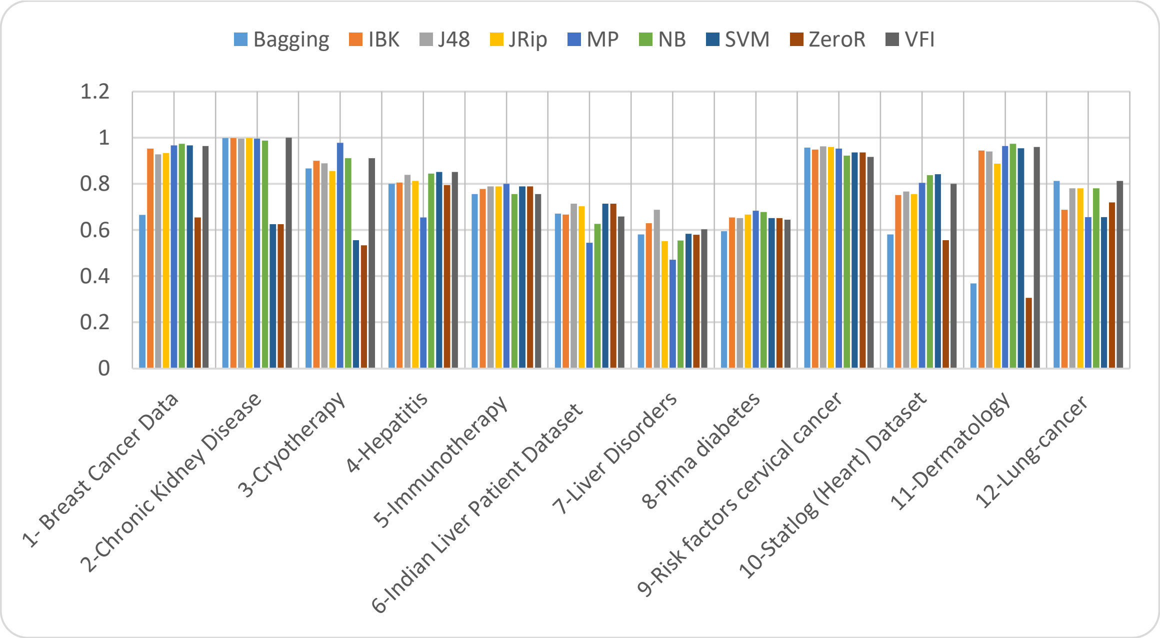

Figure 5. Recall the selected algorithms on medical datasets.

Affording to the F-measure, the F-measure is an evaluation metric used to state the performance of a machine learning model. It gives aggregated information about the precision and recall of the model. The F-measure is calculated as follows:

$$\text{F} - \text{measure} = \frac{2*Recall*Precision}{Recall+Precision} \tag{4}$$

As shown in Table 4 and Figure 6, J48 gave better results for immunotherapy, the Indian Liver Patient Dataset, liver disorders and Pima diabetes than the other algorithms, with F-measures of 0.882, 0.833, 0.68 and 0.789, respectively, while MP obtained the best F-measure for cryotherapy, with an F-measure of 0.978. NB gives the best F-measure for breast cancer data and dermatology with 0.974 and 0.973, respectively, but SVM gives a better F-measure for immunotherapy, Indian Liver Patient Dataset, Pima diabetes, risk factors cervical cancer and statlog (heart) dataset than other algorithms with 0.882, 0.833, 0.789, 0.967, and 0.84, respectively. ZeroR gives a better F-measure for immunotherapy, Indian Liver Patient Dataset, Pima diabetes, and risk factors for cervical cancer, with values of 0.882, 0.833, 0.789 and 0.967, respectively, but VFI can obtain a better F-measure for chronic kidney disease, hepatitis and lung cancer, with values of 1.000, 0.855 and 0.813, respectively. Bagging, IBK and JRip failed to obtain any better results for the datasets. The F-measure is very important when comparing algorithms because a high F-measure indicates a high value for both recall and precision.

Table 4. F-measure of the selected algorithms on medical datasets.

Dataset/Algorithm	Bagging	IBK	J48	JRip	MP	NB	SVM	ZeroR	VFI
1- Breast Cancer Data	0.553	0.953	0.927	0.933	0.966	0.974	0.966	0.792	0.965
2-Chronic Kidney Disease	0.998	0.998	0.995	0.997	0.995	0.988	0.769	0.769	1.000
3-Cryotherapy	0.867	0.900	0.889	0.855	0.978	0.910	0.419	0.696	0.911
4-Hepatitis	0.737	0.799	0.825	0.804	0.792	0.848	0.849	0.702	0.855
5-Immunotherapy	0.679	0.761	0.882	0.743	0.762	0.764	0.882	0.882	0.760
6-Indian Liver Patient Dataset	0.643	0.654	0.833	0.662	0.564	0.636	0.833	0.833	0.670
7-Liver Disorders	0.573	0.629	0.68	0.455	0.415	0.544	0.432	0.425	0.544
8-Pima diabetes	0.547	0.621	0.789	0.626	0.721	0.678	0.789	0.789	0.648
9-Risk factors cervical cancer	0.957	0.940	0.964	0.963	0.950	0.926	0.967	0.967	0.920
10-Statlog (Heart) Dataset	0.566	0.752	0.767	0.753	0.804	0.837	0.84	0.397	0.801
11-Dermatology	0.358	0.945	0.94	0.888	0.965	0.973	0.954	0.143	0.959
12-Lung-cancer	0.776	0.674	0.766	0.766	0.650	0.777	0.65	0.601	0.813

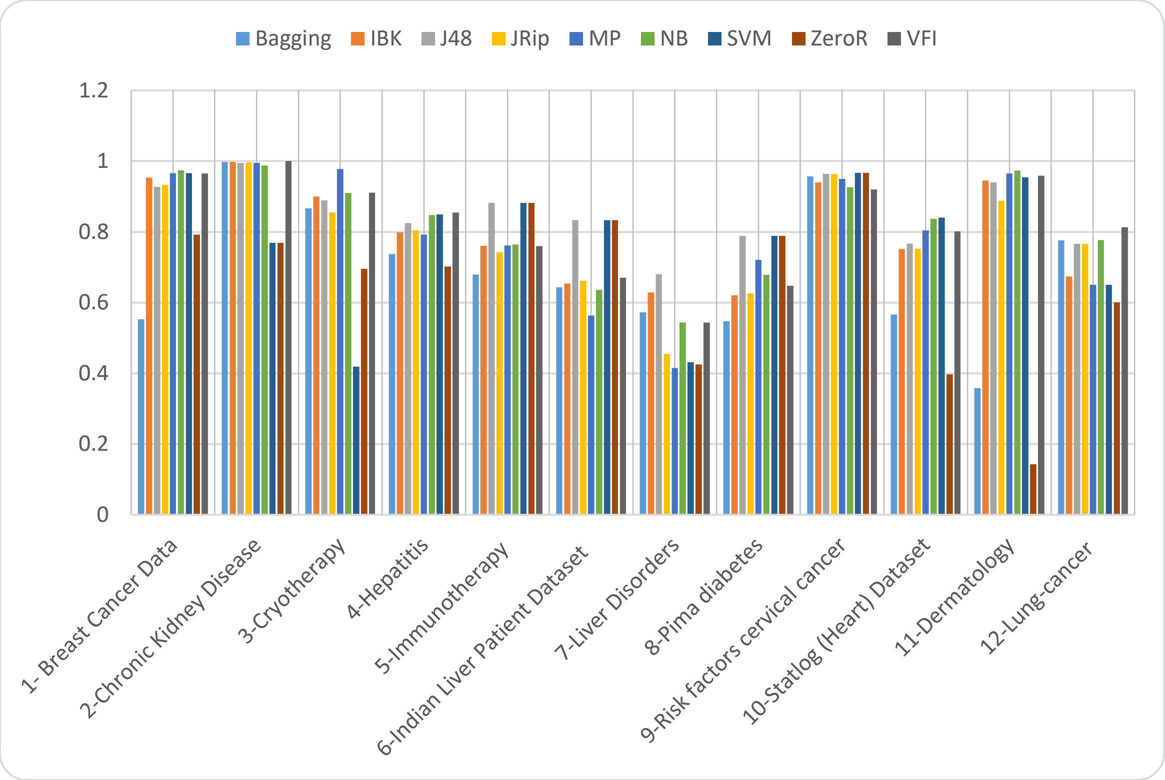

Figure 6. F-measure of the selected algorithms on medical datasets.

According to the mean absolute error (MAE), MAE is the average absolute error between the actual observation and the predicted observation. The lower the value of MAE of an algorithm, the better the algorithm among all others. The MAE is calculated as follows:

$$\frac{\sum|y_{i}- x_{i}|}{N} \qquad (5)$$

Where ∑ is a Greek symbol for summation, yi is the predicted value for the ith observation, xi is the real value for the ith observation and N is the total number of observations. Calculating MAE for machine learning models is very important to determine if the model is trained well or not. If the model was trained well, MAE would be low and result in an efficient model. As seen from Table 5 and Figure 7, J48 can beat all other algorithms and obtain 0.3673 for liver disorders. MP can give better results for cryotherapy and risk factors for cervical cancer compared to the other algorithms with 0.0622 and 0.0489, respectively, while NB obtains the best MAE for breast cancer data, Pima diabetes and dermatology with 0.0276, 0.3456 and 0.0109, respectively. SVM gives the best MAE for hepatitis, immunotherapy, Indian Liver Patient Dataset and statlog (heart) dataset with 0.1484, 0.2111, 0.2864 and 0.1593, respectively but, VFI gives better MAE for chronic kidney disease and lung cancer with 0.0016 and 0.2117, respectively. Bagging, IBK, JRip and ZeroR failed to obtain any better results for the datasets.

Table 5. MAE of the selected algorithms on medical datasets.

Dataset/Algorithm	Bagging	IBK	J48	JRip	MP	NB	SVM	ZeroR	VFI
1- Breast Cancer Data	0.4243	0.0478	0.0841	0.0849	0.0341	0.0276	0.0343	0.452	0.1299
2-Chronic Kidney Disease	0.0062	0.0052	0.005	0.0025	0.0069	0.0153	0.375	0.4689	0.0016
3-Cryotherapy	0.225	0.1092	0.0818	0.1952	0.0622	0.1538	0.4444	0.4983	0.1169
4-Hepatitis	0.3164	0.1979	0.2029	0.2577	0.452	0.1661	0.1484	0.3302	0.3793
5-Immunotherapy	0.328	0.2455	0.3333	0.2914	0.2264	0.3201	0.2111	0.3373	0.2325
6-Indian Liver Patient Dataset	0.3905	0.3591	0.4088	0.3937	0.4545	0.3919	0.2864	0.4091	0.3289
7-Liver Disorders	0.451	0.3718	0.3673	0.4876	0.5291	0.4597	0.4174	0.4874	0.4889
8-Pima diabetes	0.4478	0.3941	0.4544	0.4377	0.3651	0.3456	0.349	0.4545	0.3638
9-Risk factors cervical cancer	0.0525	0.0562	0.0558	0.0549	0.0489	0.0815	0.0641	0.121	0.2618
10-Statlog (Heart) Dataset	0.4393	0.2502	0.274	0.3142	0.1951	0.1835	0.1593	0.4939	0.4814
11-Dermatology	0.239	0.0228	0.0264	0.0456	0.0144	0.0109	0.2233	0.2665	0.1033
12-Lung-cancer	0.3074	0.3516	0.2552	0.2926	0.3275	0.2376	0.3438	0.4125	0.2117

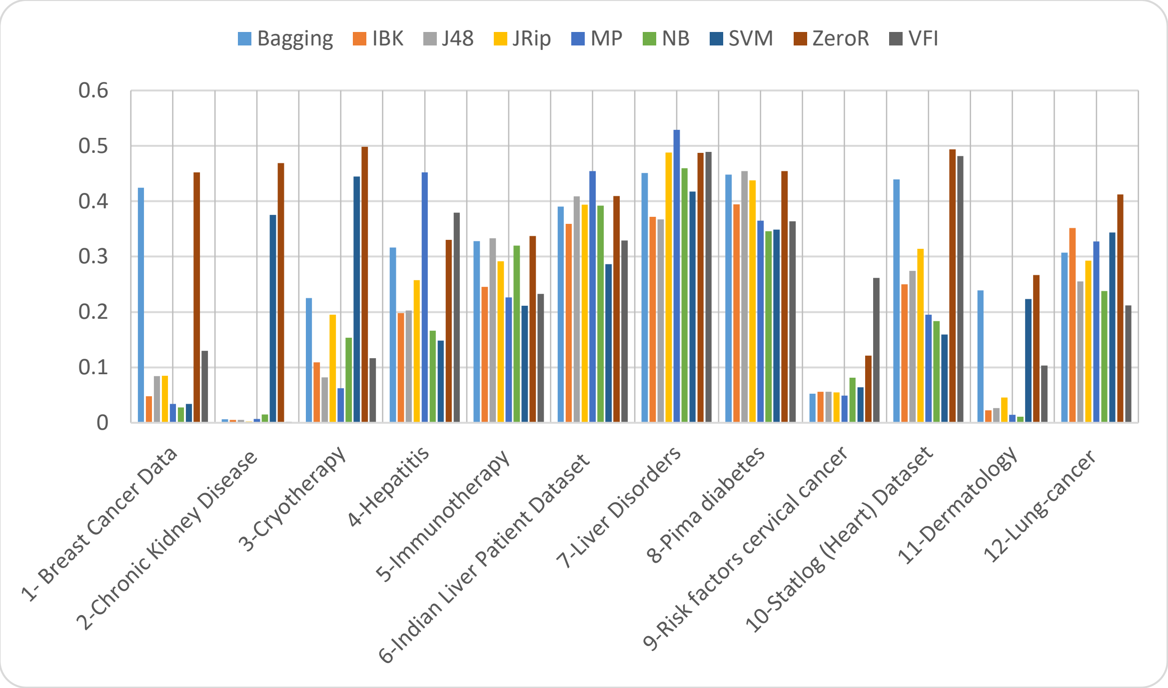

Figure 7. MAE of the selected algorithms on medical datasets.

Rendering the ROC area, the ROC area, also called the AUC (area under the ROC curve), is an evaluation metric that aids in verifying the ability of a model to distinguish the classes (positive and negative classes). The ROC curve is plotted between two parameters: the true positive rate and the false positive rate. The AUC ranges from 0 to 1, and the best algorithm has the highest AUC. As seen from Table 6 and Figure 8, bagging can beat all other algorithms and obtain 0.947 and 0.807 for the risk factors of cervical cancer and lung cancer, respectively. MP can give better results for breast cancer data, chronic kidney disease, and cryotherapy compared to other algorithms with 0.993, 1.000, and 0.996, respectively, while NB obtains better results for the ROC area for breast cancer data, chronic kidney disease, hepatitis, liver disorders, Pima diabetes, the statlog (heart) dataset and dermatology with 0.993, 1.000, 0.854, 0.664, 0.737, 0.891 and 0.999, respectively. VFI gives acceptable results for chronic kidney disease, immunotherapy, and the Indian Liver Patient Dataset, with 1.000, 0.732, and 0.670, respectively. IBK, J48, JRip, SVM, and ZeroR failed to obtain any better results for the datasets.

Table 6. ROC area of the selected algorithms on medical datasets.

Dataset/Algorithm	Bagging	IBK	J48	JRip	MP	NB	SVM	ZeroR	VFI
1- Breast Cancer Data	0.581	0.983	0.960	0.949	0.993	0.993	0.967	0.496	0.984
2-Chronic Kidney Disease	0.997	0.998	0.996	0.997	1.000	1.000	0.500	0.500	1.000
3-Cryotherapy	0.926	0.970	0.955	0.869	0.996	0.993	0.524	0.464	0.970
4-Hepatitis	0.516	0.732	0.701	0.670	0.496	0.854	0.500	0.467	0.727
5-Immunotherapy	0.589	0.704	0.470	0.564	0.707	0.721	0.500	0.470	0.732
6-Indian Liver Patient Dataset	0.592	0.626	0.491	0.559	0.567	0.638	0.500	0.491	0.670
7-Liver Disorders	0.602	0.652	0.492	0.475	0.540	0.664	0.500	0.492	0.609
8-Pima diabetes	0.520	0.604	0.497	0.575	0.712	0.737	0.500	0.497	0.676
9-Risk factors cervical cancer	0.947	0.817	0.879	0.912	0.913	0.920	0.500	0.476	0.616
10-Statlog (Heart) Dataset	0.609	0.831	0.809	0.759	0.869	0.891	0.833	0.500	0.672
11-Dermatology	0.618	0.990	0.976	0.966	0.998	0.999	0.980	0.487	0.982
12-Lung-cancer	0.807	0.442	0.708	0.531	0.696	0.773	0.500	0.401	0.729

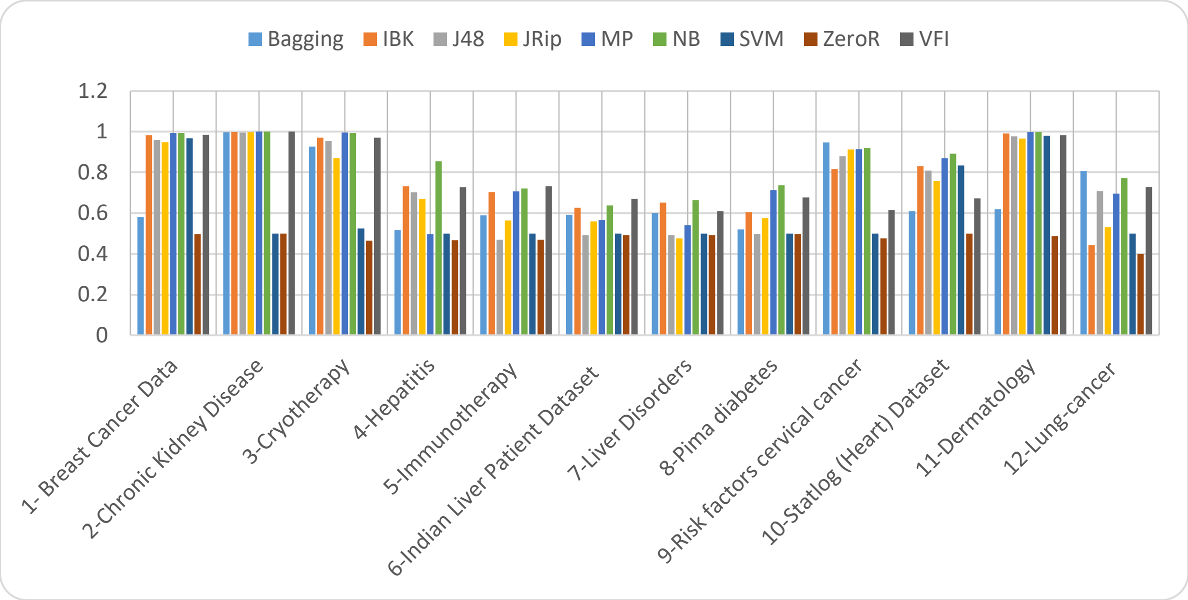

Figure 8. ROC area of the selected algorithms on medical datasets.

According to the PRC area, the PRC area is the area under the precision-recall curve (PR curve), where the PR curve is a graph with precision values on the y-axis and recall values on the x-axis. The best algorithm has a higher area under the PR curve. As seen from Table 7 and Figure 9, bagging can beat all other algorithms and obtain 0.972 and 0.824 for the risk factors of cervical cancer and lung cancer, respectively. MP can give better results for breast cancer data, chronic kidney disease, and cryotherapy compared to other algorithms with 0.993, 1.000, and 0.996, respectively, while NB obtains better results for breast cancer data, chronic kidney disease, hepatitis, liver disorders, Pima diabetes, the statlog (heart) dataset and dermatology with 0.993, 1.000, 0.895, 0.674, 0.747, 0.888 and 0.996, respectively. VFI gives better results for chronic kidney disease, immunotherapy and the Indian Liver Patient Dataset, with 1.000, 0.804, and 0.721, respectively. IBK, J48, JRip, SVM and ZeroR failed to obtain any better results for the datasets.

Table 7. PRC area of the selected algorithms on medical datasets.

Dataset/Algorithm	Bagging	IBK	J48	JRip	MP	NB	SVM	ZeroR	VFI
1- Breast Cancer Data									
2-Chronic Kidney Disease	0.624	0.980	0.949	0.939	0.993	0.993	0.953	0.546	0.982
3-Cryotherapy	0.994	0.997	0.993	0.996	1.000	1.000	0.531	0.531	1.000
4-Hepatitis	0.927	0.961	0.942	0.841	0.996	0.993	0.521	0.486	0.972
5-Immunotherapy	0.689	0.786	0.777	0.750	0.546	0.895	0.672	0.660	0.798
6-Indian Liver Patient	0.710	0.775	0.658	0.704	0.782	0.793	0.667	0.658	0.804
Dataset	0.644	0.669	0.587	0.623	0.652	0.696	0.591	0.587	0.721
7-Liver Disorders	0.599	0.640	0.509	0.498	0.565	0.674	0.513	0.509	0.622
8-Pima diabetes	0.566	0.613	0.544	0.595	0.723	0.747	0.546	0.544	0.690
9-Risk factors for cervical	0.972	0.941	0.952	0.956	0.962	0.959	0.880	0.877	0.911
cancer	0.616	0.811	0.769	0.722	0.856	0.888	0.788	0.506	0.661
10-Statlog (Heart) Dataset	0.316	0.966	0.930	0.879	0.990	0.996	0.944	0.197	0.922
11-Dermatology	0.824	0.598	0.732	0.641	0.767	0.812	0.596	0.561	0.760
12-Lung-cancer									

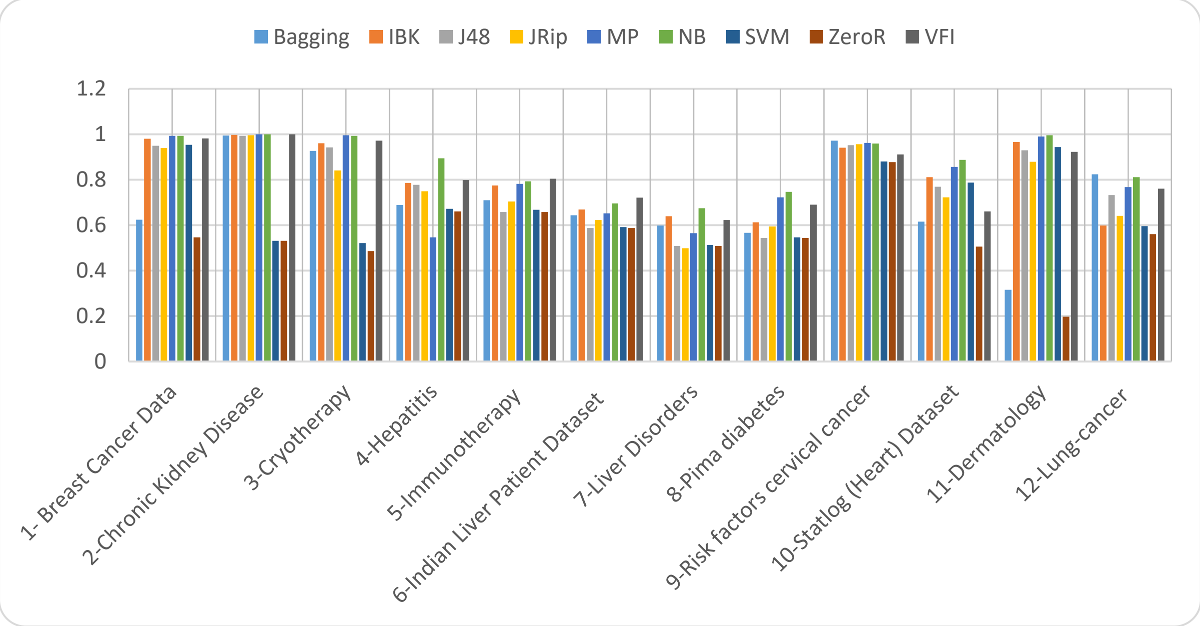

Figure 9. PRC area of the selected algorithms on medical datasets.

As seen from all results, it is not easy to determine the best algorithm for a dataset because it is difficult to balance all measures. In the context of diagnostics and medicine, accuracy is important to determine the ability of an algorithm to predict diseases but not enough alone. In some diseases, such as cancer, improving recall is crucial since it is preferable to classify someone as being positive for cancer when they do not actually have cancer than the opposite. In other problems, it is more important to balance the recall and precision (F-measure) to obtain more accurate predictions. Additionally, MAE is important to determine the error ratio between actual and predicted values. Additionally, the ROC and PRC areas are important for determining the ability of a model to distinguish the classes. Authors guideline researchers to take care when choosing an algorithm for a dataset because it is about people's lives.

5. Challenges and Future Directions

The major challenge when using machine learning algorithms is understanding machine learning models. It is very important to know which parameters (variables) cause model predictions, especially in the field of healthcare. If we know these important variables, this leads to training the model with a small number of variables and does not waste time. Another challenge is progressive advantage, which describes the value that the machine learning model adds in comparison to the state of the art. Additionally, machine learning models must have a higher degree of security against attacks, especially when they are applied on mobile devices.

According to future trends, there are many algorithms aiming to aid us in understanding how machine learning models make predictions. However, many researchers have applied explainable artificial intelligence and showed its importance in providing effective explanations for predictions. Additionally, different techniques for feature selection can be used to develop models, such as wrapper and filter methods. Researchers should evaluate machine learning models with different methods, not just accuracy, especially in the field of healthcare.

6. Conclusions

In this study, the power of machine learning algorithms was explained first. Then, an overview of the selected classifiers and datasets was presented. After that, twelve medical datasets were selected from the UCI repository for testing, and nine classification machine learning algorithms were selected for performance evaluation. An overview of the selected classifiers and datasets is presented. Classification algorithms were implemented on medical datasets by using the WEKA tool and 10-fold cross validation. The results were reported in terms of accuracy, precision, recall, F-measure, mean absolute error, ROC area and PRC area. The results conclude that the most suitable algorithm for each dataset is as follows: SVM for breast cancer, VFI for chronic kidney disease, J48 for cryotherapy, SVM and VFI for hepatitis, Bagging for immunotherapy, SVM and ZeroR for ILPD, MP for liver disorders, NB for Pima diabetes, Bagging and JRip for risk factors for cervical cancer, SVM for statlog, NB for dermatology, Bagging and VFI for lung cancer. This study also concludes that determining the appropriate algorithm for each dataset is a challenging task. Experimental results and analysis for all measures were introduced to guideline researchers who seek to achieve and how to balance different measures. Additionally, challenges and future directions for machine learning are presented in the last section of this paper.

Author Contributions

All authors contributed equally to this work.

Funding

This research was conducted without external funding support.

Ethical approval

This article does not contain any studies with human participants or animals performed by any of the authors.

Conflicts of Interest

The authors declare that there is no conflict of interest in the research

.Institutional Review Board Statement

Not applicable.

Informed Consent Statement

Not applicable.

Data Availability Statement

Not applicable

References

[1] G. Shilimkar and S. Pisal, "Disease Prediction Using Machine Learning," 2021.

[2] R. Sala, M. G. Zambetti, F. Pirola, and R. Pinto, "How to select a suitable machine learning algorithm: a feature-based, scope-oriented selection framework," in 23rd Summer School" Francesco Turco"-Industrial Systems Engineering 2018, 2018, vol. 2018, pp. 87-93: AIDI-Italian Association of Industrial Operations Professors.

[3] A. S. Kwekha-Rashid, H. N. Abduljabbar, and B. Alhayani, "Coronavirus disease (COVID-19) cases analysis using machine-learning applications," Applied Nanoscience, pp. 1-13, 2021.

[4] M. T. Huyut and H. Üstündağ, "Prediction of diagnosis and prognosis of COVID-19 disease by blood gas parameters using decision trees machine learning model: a retrospective observational study," Medical Gas Research, vol. 12, no. 2, p. 60, 2022.

[5] P. Ghosh et al., "Efficient prediction of cardiovascular disease using machine learning algorithms with relief and LASSO feature selection techniques," IEEE Access, vol. 9, pp. 19304-19326, 2021.

[6] M. Diwakar, A. Tripathi, K. Joshi, M. Memoria, and P. Singh, "Latest trends on heart disease prediction using machine learning and image fusion," Materials Today: Proceedings, vol. 37, pp. 3213-3218, 2021.

[7] G. Chaubey, D. Bisen, S. Arjaria, and V. Yadav, "Thyroid disease prediction using machine learning approaches," National Academy Science Letters, vol. 44, no. 3, pp. 233-238, 2021.

[8] R. Sengupta, D. Sengupta, A. K. Kamra, and D. J. A. I. Pandey, "ARTIFICIAL INTELLIGENCE AND QUANTUM COMPUTING FOR A SMARTER WIRELESS NETWORK," vol. 7, no. 19, p. 2020, 2020.

[9] M. M. Megdad, S. S. Abu-Naser, and B. S. J. I. J. o. A. I. S. R. Abu-Nasser, "Fraudulent Financial Transactions Detection Using Machine Learning," vol. 6, no. 3, 2022.

[10] E. Ropelewska, K. Sabanci, M. F. J. E. F. R. Aslan, and Technology, "Preservation effects evaluated using innovative models developed by machine learning on cucumber flesh," pp. 1-9, 2022.

[11] D. Kshirsagar, S. J. J. o. A. I. Kumar, and H. Computing, "A feature reduction based reflected and exploited DDoS attacks detection system," vol. 13, no. 1, pp. 393-405, 2022.

[12] K. Karataş, I. Arpaci, Y. J. E. Yildirim, and U. Society, "Predicting the Culturally Responsive Teacher Roles With Cultural Intelligence and Self-Efficacy Using Machine Learning Classification Algorithms," p. 00131245221087999, 2022.

[13] A. A. Alnuaim et al., "Human-Computer Interaction for Recognizing Speech Emotions Using Multilayer Perceptron Classifier," vol. 2022, 2022.

[14] N. A. Mansour, A. I. Saleh, M. Badawy, H. A. J. J. o. a. i. Ali, and h. computing, "Accurate detection of Covid-19 patients based on Feature Correlated Naïve Bayes (FCNB) classification strategy," vol. 13, no. 1, pp. 41-73, 2022.

[15] M. Tanveer, T. Rajani, R. Rastogi, Y.-H. Shao, and M. J. A. o. O. R. Ganaie, "Comprehensive review on twin support vector machines," pp. 1-46, 2022.

[16] J. Kalezhi, M. Chibuluma, C. Chembe, V. Chama, F. Lungo, and D. J. R. i. E. Kunda, "Modelling Covid-19 infections in Zambia using data mining techniques," vol. 13, p. 100363, 2022.

[17] S. Jindal, M. Sachdeva, and A. K. J. K. J. o. S. Kushwaha, "Performance evaluation of machine learning based voting classifier system for human activity recognition: 10.48129/kjs. splml. 19189," 2022.

[18] C. Blake, "UCI repository of machine learning databases," http://www. ics. uci. edu/~ mlearn/MLRepository. html, 1998.

[19] C. A. Pedersen et al., "DNA methylation changes in response to neoadjuvant chemotherapy are associated with breast cancer survival," vol. 24, no. 1, pp. 1-14, 2022.

[20] Z. Deng, M. R. Jones, M.-C. Wang, and K. J. n. B. C. Visvanathan, "Mortality after second malignancy in breast cancer survivors compared to a first primary cancer: a nationwide longitudinal cohort study," vol. 8, no. 1, pp. 1-10, 2022.

[21] W. G. van der Plas-Krijgsman et al., "Predicting disease-related and patient-reported outcomes in older patients with breast cancer-a systematic review," vol. 12, no. 5, pp. 696-704, 2021.

[22] I. Meattini et al., "Integrating radiation therapy with targeted treatments for breast cancer: From bench to bedside," p. 102417, 2022.

[23] Mahmoud Ibrahim, Sentiment Analysis for Fake News Detection in Online Media Networks: A survey, fusion techniques and quality metrics, Journal of Neutrosophic and Information Fusion, Vol. 1 , No. 2 , (2023) : 44 68 (Doi : https://doi.org/10.54216/NIF.010205)

[24] Rose Aljanada,Ghadeer W. Abukhalil,Aseel M. Alfaisal,Raghad M. Alfaisal, Adoption of Google Glass technology: PLS-SEM and machine learning analysis, Journal of International Journal of Advances in Applied Computational Intelligence, Vol. 1 , No. 1 , (2022) : 08-22 (Doi : https://doi.org/10.54216/IJAACI.010101)

[25] B. Fernandez-Fernandez et al., "Gender, albuminuria and chronic kidney disease progression in treated diabetic kidney disease," vol. 9, no. 6, p. 1611, 2020.

[26] M. Gorostidi et al., "Chronic kidney disease in Spain: Prevalence and impact of accumulation of cardiovascular risk factors," vol. 38, no. 6, pp. 606-615, 2018.

[27] M. E. Rodríguez-Ortiz, C. Pontillo, M. Rodríguez, P. Zürbig, H. Mischak, and A. J. S. r. Ortiz, "Novel urinary biomarkers for improved prediction of progressive eGFR loss in early chronic kidney disease stages and in high risk individuals without chronic kidney disease," vol. 8, no. 1, pp. 1-11, 2018.

[28] Aseel M. Alfaisal,Aisha Zare,Afrah Alshaafi,Rose Aljanada,Raghad M. Alfaisal,Ghadeer W. Abukhalil, Predicting the actual use of social media sites among university communicators: using PLS-SEM and ML approaches, Journal of International Journal of Advances in Applied Computational Intelligence, Vol. 1 , No. 1 , (2022) : 23-33 (Doi : https://doi.org/10.54216/IJAACI.010102)

[29] B. Bikbov et al., "Global, regional, and national burden of chronic kidney disease, 1990–2017: a systematic analysis for the Global Burden of Disease Study 2017," vol. 395, no. 10225, pp. 709-733, 2020.

[30] A. G. Del Signore, J. B. Greene, J. L. Russell, D. M. Yen, E. M. O'Malley, and R. J. Schlosser, "Cryotherapy for treatment of chronic rhinitis: 3-month outcomes of a randomized, sham-controlled trial," in International Forum of Allergy & Rhinology, 2022, vol. 12, no. 1, pp. 51-61: Wiley Online Library.

[31] Ismail Eyad Samara, Intelligent systems and AI techniques: Recent advances and Future directions, Journal of International Journal of Advances in Applied Computational Intelligence, Vol. 1 , No. 2 , (2022) : 30-45 (Doi : https://doi.org/10.54216/IJAACI.010202)

[32] C. Peyronnel et al., "Effects of local cryotherapy on systemic endothelial activation, dysfunction, and vascular inflammation in adjuvant-induced arthritis (AIA) rats," vol. 24, no. 1, pp. 1-16, 2022.

[33] S. Farhadi-Tooli et al., "Droplet-vitrification cryotherapy and thermotherapy as efficient tools for the eradication of apple chlorotic leaf spot virus and apple stem grooving virus from virus-infected quince in vitro cultures," vol. 162, no. 1, pp. 31-43, 2022.

[34] Ahmed Sleem,Ibrahim Elhenawy, An Attentive Convolutional Recurrent Network for Fake News Detection, Journal of International Journal of Advances in Applied Computational Intelligence, Vol. 2 , No. 1 , (2022) : 08-14 (Doi : https://doi.org/10.54216/IJAACI.020101)

[35] B. Terziroli Beretta-Piccoli, G. Mieli-Vergani, D. J. C. Vergani, and m. immunology, "Autoimmmune hepatitis," vol. 19, no. 2, pp. 158-176, 2022.

[36] M. M. Mücke and S. J. J. o. H. Zeuzem, "The recent outbreak of acute severe hepatitis in children of unknown origin–what is known so far," 2022.

[37] M. Iannacone and L. G. J. N. R. I. Guidotti, "Immunobiology and pathogenesis of hepatitis B virus infection," vol. 22, no. 1, pp. 19-32, 2022.

[38] J. M. Baker et al., "Acute hepatitis and adenovirus infection among children—Alabama, October 2021–February 2022," vol. 71, no. 18, p. 638, 2022.

[39] R. Bai et al., "Mechanisms of cancer resistance to immunotherapy," vol. 10, p. 1290, 2020.

[40] F. L. van de Veerdonk et al., "A guide to immunotherapy for COVID-19," vol. 28, no. 1, pp. 39-50, 2022.

[41] B. Scheiner et al., "Prognosis of patients with hepatocellular carcinoma treated with immunotherapy–development and validation of the CRAFITY score," vol. 76, no. 2, pp. 353-363, 2022.
[42] J. Pan et al., "Self‐Blockade of PD‐L1 with Bacteria‐Derived Outer‐Membrane Vesicle for Enhanced Cancer Immunotherapy," vol. 34, no. 7, p. 2106307, 2022.
[43] J. Kim et al., "Thermosensitive hydrogel releasing nitric oxide donor and anti-CTLA-4 micelles for anti-tumor immunotherapy," vol. 13, no. 1, pp. 1-13, 2022.
[44] D. Mondal, K. Das, and A. J. C. L. D. Chowdhury, "epidemiology of liver Diseases in india," vol. 19, no. 3, p. 114, 2022.
[45] M. B. Priya, P. L. Juliet, P. J. I. R. J. o. E. Tamilselvi, and Technology, "Performance analysis of liver disease prediction using machine learning algorithms," vol. 5, no. 1, pp. 206-211, 2018.
[46] S. Pahari and D. K. Choubey, "Analysis of Liver Disorder by Machine Learning Techniques," in Soft Computing: Theories and Applications: Springer, 2022, pp. 587-601.
[47] I. Straw, H. J. B. h. Wu, and c. informatics, "Investigating for bias in healthcare algorithms: a sex-stratified analysis of supervised machine learning models in liver disease prediction," vol. 29, no. 1, 2022.
[48] E. N. Mutlu, A. Devim, A. A. Hameed, and A. Jamil, "Deep Learning for Liver Disease Prediction," in Mediterranean Conference on Pattern Recognition and Artificial Intelligence, 2022, pp. 95-107: Springer.
[49] K. Gupta, N. Jiwani, N. Afreen, and D. Divyarani, "Liver Disease Prediction using Machine learning Classification Techniques," in 2022 IEEE 11th International Conference on Communication Systems and Network Technologies (CSNT), 2022, pp. 221-226: IEEE.
[50] L. Ganu and B. Arun, "Hybrid Combination of Machine Learning Techniques for Diagnosis of Liver Impairment Disease in Clinical Decision Support System," in Advanced Machine Intelligence and Signal Processing: Springer, 2022, pp. 479-492.
[51] H. Yoshiji et al., "Evidence-based clinical practice guidelines for liver cirrhosis 2020," vol. 56, no. 7, pp. 593-619, 2021.
[52] J.-C. Nault, V. Paradis, M. Ronot, J. J. N. R. G. Zucman-Rossi, and Hepatology, "Benign liver tumours: understanding molecular physiology to adapt clinical management," pp. 1-14, 2022.
[53] J.-C. Nault et al., "Non-invasive diagnosis and follow-up of benign liver tumours," vol. 46, no. 1, p. 101765, 2022.
[54] C. Desdouets, A. J. N. R. G. Lujambio, and Hepatology, "Decoding therapy resistance in liver tumours: a giant leap," vol. 19, no. 2, pp. 83-84, 2022.
[55] I. A. Bashir, A. A. Mariod, R. Banu, T. B. J. C. R. i. N. Elyas, and F. Science, "Significance of Health Related Predictors of Diabetes in Pima Indians Women," vol. 7, no. 2, p. 350, 2019.
[56] V. Chang, J. Bailey, Q. A. Xu, Z. J. N. C. Sun, and Applications, "Pima Indians diabetes mellitus classification based on machine learning (ML) algorithms," pp. 1-17, 2022.
[57] S. A. Alex, J. Nayahi, H. Shine, V. J. N. C. Gopirekha, and Applications, "Deep convolutional neural network for diabetes mellitus prediction," vol. 34, no. 2, pp. 1319-1327, 2022.
[58] S. Selvakumar, K. S. Kannan, and S. GothaiNachiyar, "Prediction of diabetes diagnosis using classification based data mining techniques," International Journal of Statistics and Systems, vol. 12, no. 2, pp. 183-188, 2017.
[59] M. E. Pavkov, R. L. Hanson, W. C. Knowler, P. H. Bennett, J. Krakoff, and R. G. J. D. c. Nelson, "Changing patterns of type 2 diabetes incidence among Pima Indians," vol. 30, no. 7, pp. 1758-1763, 2007.
[60] S. C. B. Jaganathan et al., "Machine Learning for Smartphone-Based Early Detection of Diabetic Disease in Pima Indians Diabetes Database," vol. 13, no. 1, pp. 780-796, 2022.
[61] R. A. Husain and V. J. A. o. g. h. Ramakrishnan, "Global variation of human papillomavirus genotypes and selected genes involved in cervical malignancies," vol. 81, no. 5, pp. 675-683, 2015.
[62] J. Huang et al., "Global distribution, risk factors, and recent trends for cervical cancer: a worldwide country-level analysis," vol. 164, no. 1, pp. 85-92, 2022.
[63] S. Tosic Golubovic et al., "Risk Factors and Predictive Value of Depression and Anxiety in Cervical Cancer Patients," vol. 58, no. 4, p. 507, 2022.
[64] H. Zhao, Y. Peng, M. Lv, Y. Shi, S. J. M. Zhang, and C. Oncology, "Incidence and risk factors of perioperative venous thromboembolism in patients with cervical cancer," vol. 16, no. 6, pp. 1-6, 2022.

[65] P. R. Garg et al., "Women's Knowledge on Cervical Cancer Risk Factors and Symptoms: A Cross Sectional Study from Urban India," vol. 23, no. 3, pp. 1083-1090, 2022.

[66] G. D. Flora and M. K. J. C. p. d. Nayak, "A brief review of cardiovascular diseases, associated risk factors and current treatment regimes," vol. 25, no. 38, pp. 4063-4084, 2019.

[67] A. M. J. D. i. c. n. Möller-Leimkühler, "Gender differences in cardiovascular disease and comorbid depression," 2022.

[68] P. Motarwar, A. Duraphe, G. Suganya, and M. Premalatha, "Cognitive approach for heart disease prediction using machine learning," in 2020 International Conference on Emerging Trends in Information Technology and Engineering (ic-ETITE), 2020, pp. 1-5: IEEE.

[69] K. J. T. J. o. C. I. Musunuru, "Moving toward genome-editing therapies for cardiovascular diseases," vol. 132, no. 1, 2022.

[70] J. M. Guirguis-Blake, C. V. Evans, L. A. Perdue, S. I. Bean, and C. A. J. J. Senger, "Aspirin use to prevent cardiovascular disease and colorectal cancer: updated evidence report and systematic review for the US Preventive Services Task Force," vol. 327, no. 16, pp. 1585-1597, 2022.

[71] T. Willem et al., "Risks and benefits of dermatological machine learning healthcare applications–an overview and ethical analysis," 2022.

[72] K. Chilicka, M. Rusztowicz, R. Szyguła, and D. J. J. o. C. M. Nowicka, "Methods for the Improvement of Acne Scars Used in Dermatology and Cosmetology: A Review," vol. 11, no. 10, p. 2744, 2022.

[73] G. Fredman, L. Skov, M. Mann, and B. J. D. Dyring-Andersen, "Towards precision dermatology: emerging role of proteomic analysis of the skin," vol. 238, no. 2, pp. 185-194, 2022.

[74] D. Aickara, A. M. Bashyam, R. O. Pichardo, and S. R. J. J. o. D. T. Feldman, "Topical methotrexate in dermatology: A review of the literature," vol. 33, no. 1, pp. 512-517, 2022.

[75] A. M. Romaszko and A. J. A. C. E. M. Doboszyńska, "Multiple primary lung cancer: a literature review," vol. 27, no. 5, pp. 725-730, 2018.

[76] V. T. L. de Alencar, M. N. Formiga, and V. C. C. J. E. de Lima, "Inherited lung cancer: a review," vol. 14, 2020.

[77] B. Alsinglawi et al., "An explainable machine learning framework for lung cancer hospital length of stay prediction," Scientific reports, vol. 12, no. 1, pp. 1-10, 2022.

[78] P. Zhang et al., "Association of smoking and polygenic risk with the incidence of lung cancer: a prospective cohort study," vol. 126, no. 11, pp. 1637-1646, 2022.

[79] E. Heitzer et al., "Recommendations for a practical implementation of circulating tumor DNA mutation testing in metastatic non-small-cell lung cancer," vol. 7, no. 2, p. 100399, 2022

Received: 07 Dec 2023, **Revised:** 19 Mar 2024,

Accepted: 18 Apr 2024, **Available online:** 22 Apr 2024.