Recent Advances in Computer-Aided Medical Diagnosis Using Machine Learning Algorithms With Optimization Techniques

5 authors, including:

Taki Hasan Rafi
Hallym University

15 PUBLICATIONS   **17** CITATIONS

Raed Shubair
New York University Abu Dhabi

428 PUBLICATIONS   **4,947** CITATIONS

Faisal Farhan
Ahsanullah University of Science & Tech

5 PUBLICATIONS   **2** CITATIONS

Md.Ziaul Hoque
Ahsanullah University of Science & Tech

3 PUBLICATIONS   **2** CITATIONS

Some of the authors of this publication are also working on these related projects:

Project Antennas and RF Sensors for Wireless & Medical Applications View project

Project IEEE-UAE Distinguished Seminar Series View project

Recent Advances in Computer-Aided Medical Diagnosis Using Machine Learning Algorithms With Optimization Techniques

TAKI HASAN RAFI[ID][1], **RAED M. SHUBAIR**[ID][2], **(Senior Member, IEEE), FAISAL FARHAN**[1], **MD. ZIAUL HOQUE**[1], **AND FARHAN MOHD QUAYYUM**[1]

[1]Department of Electrical and Electronic Engineering, Ahsanullah University of Science and Technology, Tejgaon, Dhaka 1208, Bangladesh
[2]Department of Electrical and Computer Engineering, New York University (NYU), Abu Dhabi, United Arab Emirates

Corresponding author: Raed M. Shubair (raed.shubair@nyu.edu)

ABSTRACT Artificial intelligence is a spectacular part of computer engineering that has earned a compelling diversion in the field of medical data classification due to its state-of-art algorithmic strength and learning capabilities. Machine Learning is a major sub-domain of artificial intelligence, where it has become one of the most promising fields in computer science. In recent years, there is a large spectrum of healthcare and biomedical data that has been growing intensely. Due to the huge labeled or unlabeled data, it is important to have a compact and robust machine learning solution for classification. Several optimizers have been deployed to improve the inclusive performance of machine learning models. The classification of machine learning models depends on several factors. This comprehensive review paper aims to insight into the current stage of optimized machine learning success on medical data classification. An increasing number of unstructured medical data has been utilizing in machine learning algorithms to predict intuitions. But it is difficult to inherent immense intuition from those data. So machine learning researchers have utilized state-of-art optimizers and novel feature selection techniques to overcome and emend the performance accuracy. We have highlighted some recent literature, which exhibits the robust impact of optimizers and feature selection on machine learning techniques on medical data characterization. On the other hand, a clean-cut introduction on machine learning and theoretical outlook of widely utilized optimization techniques like genetic algorithm, gray wolf optimization, and particle swarm optimization are discussed for initial understanding of the optimization techniques.

INDEX TERMS Breast cancer, diabetes, parkinson disease, heart disease, machine learning, optimizers.

I. INTRODUCTION

Machine learning (ML) focuses on the receiving and self-learning policy by manipulating the algorithms. Machine learning is also very impressive on managing the large spectrum of data, whether it's labeled or unlabeled. It has very immersive applications in medical data classification in large-scale optimization [1]. It is responsible of taking information and intuition from the training dataset [2]. These applications have been utilized with expanding accomplishment to foresee persistent forecasts in numerous different zones of medication, for example, hospital re-admission prediction, breast cancer prediction, diabetic difficulties, cardiovascular mortality, and so on.

The associate editor coordinating the review of this manuscript and approving it for publication was Juan Liu[ID].

In machine learning algorithms, optimizations are needed to reduce the computation and extracting important features from the dataset. It can impact the overall classification performance of machine learning techniques. However, optimization techniques can play a vital role in constructing machine learning algorithms for large-amount of data. Because of their broad application and appealing theoretical features, optimization methods have risen in popularity in machine learning. Because of the rising complexity of machine learning models, current assumptions must be analyzed. That's why optimization in machine learning is extremely useful.

Medical diagnosis is a challenging task in machine learning domain. A large amount of medical data is accessible to utilize for machine learning models. Medical data classification is based on the human medical expert perspective.

But there are some errors, unwanted biasness can be occurred while classifying the data. So robust intelligent solution can decrease address this problem into a more compact way [3]. In medical research field, a scoring prediction system is utilized to predict the disease risk of the patients [4]–[10].

In this review article, we have considered four diseases to review their medical data classification utilizing optimizers prior to machine learning models. The four diseases are breast cancer, heart disease, diabetes and parkinson disease. The main reason behind taking these four diseases because machine learning techniques are widely and mostly used in diagnosis of these four disease comparing to other diseases like kidney, lung and others. A brief introductions of these diseases are presented.

Breast cancer, which is the 2nd leading death reasons among all cancers in women. Multifarious intellect reasons are involved behind the causes of appearing breast cancer [11]. Breast cancer is mainly seen in 40 years old ladies [12]. A timely robust screening process in needed to diminish the mortality rate of breast cancer [13]. Precautionary treatment can easily increase the chances of getting rid of breast cancer [14]. However, breast cancer can be diagnosed by mammogram and structural symptoms. Different intelligent medical diagnosis systems are developed to avoid the human errors in the screening process [15]. In recent years, machine learning and deep learning have stimulated an enormous effort in medical classification systems.

Heart disease considered as the most deadliest diseases among all countries. It is drastically increasing in a undefined order globally [23]. Heart disease is also named as Cardiovascular disease. Which caused 17 million deaths globally, according to WHO. Multi health attributes are considered such as blood pressure, cholesterol, auscultation to determine cardiovascular condition. Regular health monitoring is required to prevent any unwanted circumstances [24]. Most common initial symptoms are deliberated such as fatigue and chest pain for an early assumption to cardiovascular disorder [25]. Lifestyle elements can dominate the risk of a patient. According to the investigation process of suspicious cardiovascular disorder, physicians have required some medical tests such as blood pressure, previous heart disorder history in the family or the patient itself, chest X-ray, ECG report etc [26]. Data mining techniques have been stimulating a tremendous success on heart disease prediction in a proficient manner. In recent years, due to accessible medical heart disease data, several clinical decision systems are developed to analyze individual medical condition for the patients. It helps profoundly to the patients as well as the physicians in healthcare domain [27].

Diabetes, a common disease globally. It increases the blood sugar level from the expected norm [35]. Sugar is considered as the main source of energy in human body, when it comes to uncontrolled sugar production in the body if creates diabetes [36]. Rising number of diabetes patients in lower-middle income countries are very concerning [37]. Several severe medical conditions can occur due to uncontrolled diabetes such as blood pressure, organ failure, heart disease, kidney failure and so many other conditions. Which leads to high risk of death. Type - 2 diabetes leads to high risk and effects the insulin reproduction in the human body [38]. Diabetes can regulate by the certain lifestyle changes, desired drugs and In next few years, diabetes will cause 25% expanded death rate globally. Utilizing the scoring method can segregate and differentiate the low-high risk patients [39]. However, frequent diagnosis is essential to determine patient's health condition about diabetes. Data mining techniques are widely utilized to thrive the clinical support system for the patients.

Parkinson is one of the incurable diseases in the world and many people suffer from it, resulting in memory loss and other symptoms. It is a progressive neuro-degenerative motor system disorder. For this reason, early detection is much needed so that precautions can be taken at the early stage. Around 90% of parkinson patients have examined with dysphonia, recent studies said. Vocal valuation is needed to determine parkinson disease [46]. It causes for crumbling of movement such as, unsteadiness, discourse, for example, trouble articulating sounds, and decreased volume and pitch range, and on psychological perception of human [47]. A tremendous success can be achieved if the patient gets proper meditation and treatment [48]. In past decades, machine learning researchers are trying to build a user friendly and intelligent PD detection support system. Several state-of-art techniques are applied to solve this problem [51].

Data mining techniques have attained a massive success on medical data classification. But the desired datasets are not primarily congruent for the data mining techniques to build a robust solution. Due to an inconsistent and small dataset, it is not easy to gain quality features from the dataset. Regardless to this, we need adaptable machine learning techniques with suitable and quality features to build decision support systems. Different optimizers are enforced to get key features. Machine learning models with optimizers are performing very commendable comparing the results without optimizers [52].

In this review paper, we have considered to review some recent literature of medical data classification using optimizers. This review will give intuitions of current scenario of the impact of optimizers in machine learning performance in medical data classification. Key contributions of this review paper are –

- A holistic overview of recent works on optimization techniques used in machine learning for medical data classification is demonstrated.
- Recent development of machine learning on medical data classification is provided.
- Most extensive available datasets for breast cancer, heart disease, diabetes and parkinson disease are discussed while reviewing literature.
- Limitations and future scopes of optimization techniques are discussed in this review paper.
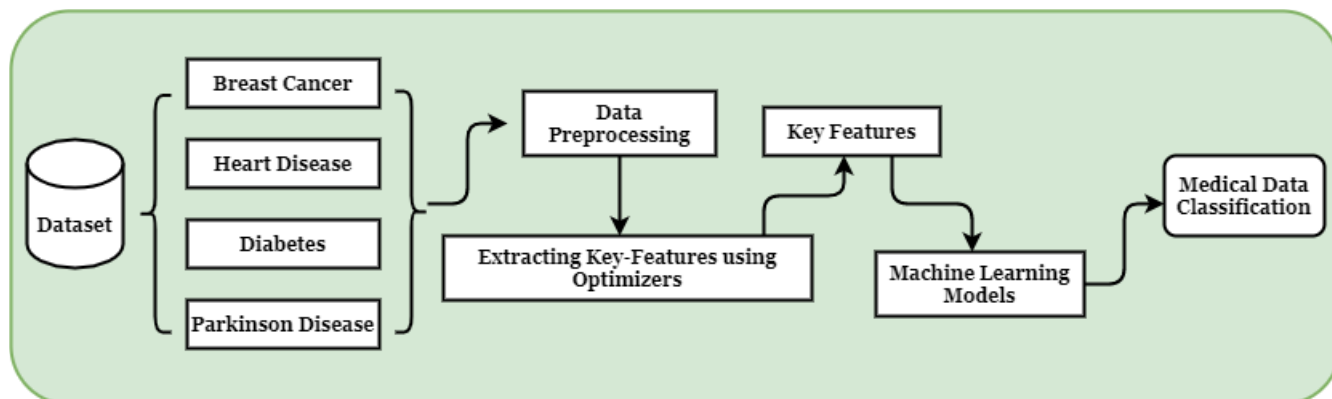
The following paper is oriented as, section II describes the overall background on machine learning and optimization techniques, section III has the taxonomy of this article, section IV summarizes the recent literature on disease prediction utilizing machine learning and optimizer techniques. And lastly section V and VI represent the future direction, discussion and conclusion.

II. BACKGROUND ON MACHINE LEARNING

Machine learning has the capability to learn from the data. We train a large set of data to machine learning models. Where it finds the patterns from the data [45]. In recent years, there are a large spectrum of applications of machine learning techniques. In this section a brief introduction of machine learning classification is presented. Figure 2 shows an overview of machine learning classification.

A. SUPERVISED LEARNING

Supervised learning is a dominant learning process, where the output builds upon the labeled input [53]. In ideal cases, the supervised algorithms find a specific pattern from the labeled input data. This learning process has computed as to minimize the error by altering the error for a productive result. It mostly concedes the best outcome while the predicted features are known. Labeled breast cancer, heart disease, diabetes prediction are the ideal example of supervised learning, where the prediction class is known [54].

B. UNSUPERVISED LEARNING

Unsupervised learning relied on the unlabeled data as the input. It urges to predict the output utilizing the unlabeled input data. It is also considered as adaptive learning by its extreme learning process [55]. Clustering technique is considered as the recurrent method of unsupervised learning. Where it barely relied on the probabilistic clustering technique. Although, unsupervised learning is mostly used in unlabeled segmentation [56].

C. REINFORCEMENT LEARNING

Reinforcement learning is a learning process where an agent take multidimensional decisions for picking the right

action [57]. It is marginally varied from the supervised learning. Reward and penalty are the major stage to take action for an agent in an environment. There are four stages of a reinforcement environment. These stages are action-rule, performance component, reinforcement component and discovery component. Q-learning is the widely applicable reinforcement learning variant [58].

D. RECOMMENDER SYSTEMS

Recommender systems can be characterized as a learning procedures by righteousness of which online client can alter their locales to meet client's preference [59]. Apparently, it is an intelligent computer based system that learns from the client's searching and preference. Youtube video recommendation, facebook friend suggestions are the ideal example of recommender systems. Currently, it has a large scale applications in online marketing, social sites etc [60].

III. BACKGROUND ON OPTIMIZATION TECHNIQUES
A. GENETIC ALGORITHM

Genetic Algorithm (GA) is one of the state-of-the-art algorithms that utilize both combinatorial and numerical optimizations and was developed by John Holland in the 1970's. They have numerous applications in the research fields such as management, engineering, identifying several medical diseases and etc. The prime three operators of GA are selection, crossover and mutation [61].

The two operators crossover and mutation utilizes the search space to figure out the solutions and each solution is displayed as standalone chromosome with numerous alleles embedded with material of genetics that evaluates the object function's value of fitness. In case of crossover, it generates two kinds of chromosomes that is known as offspring and comes from two parents. To build heterogeneousness in the search space of the population pool a known probability is given on the offspring [62]. Also, it can be seen that the population, residing in individual generation, is calculated utilizing the function known as fitness function. In the next step the selection process is carried out where it can be seen that the low fitness chromosomes are gotten rid of by the high fitness chromosomes. Some of the well-known processes for
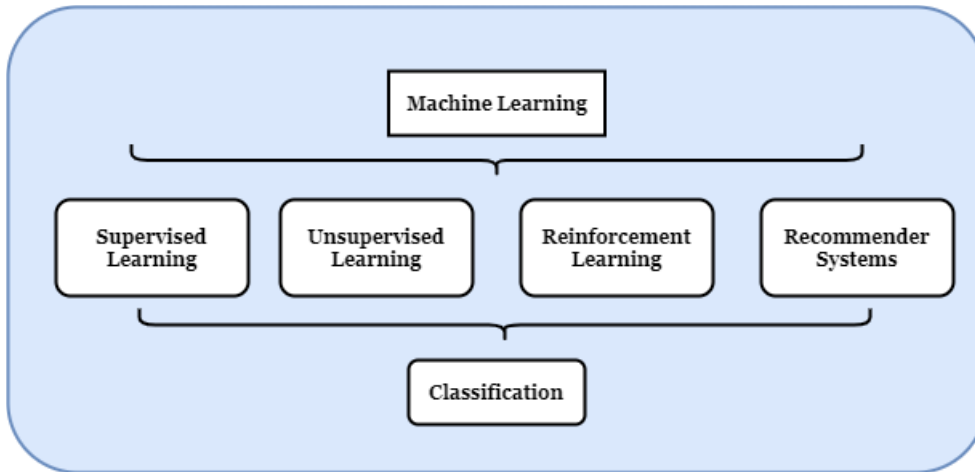
FIGURE 2. An overview of machine learning classification techniques.

selection or generation are Roulette-wheel selection, Boltzmann selection, Tournament selection, Rank selection and Steady-state selection. Although, it can be seen that selection alone cannot generate new diversifications in the population pool so to eliminate this crossover and mutation methods are introduced. In crossover there are numerous types such as are one point crossover, two-point crossover, uniform crossover, multi-point crossover and average crossover. Mutation can be defined as the unique change of the point assigned to the certain gene. Furthermore, mutation can be utilized to eliminate premature convergence in the local optima. Mutation is established in several processes which are creep mutations, random gene mutations, heuristic mutation and etc. The whole process is repeated until the desired termination criterion is met. In this way the latest generation is created from the previous generation [63].

In [64] the authors have suggested that GA deals with variables that are string-coded rather than normal ones because in string-coded variables the coding for the search space is distinctive and advantageous. In [65] the authors wanted to propose a method where the controlling parameters are generated into the transformation in such a way where the proven fitter stand alone do not untimely dominate the given population. A numerical approach to elaborate the accomplishment of the several techniques. Here is an example of the proposed method:

$$F(X) = \frac{X_1 X_2 - X_3 X_4}{X_5^2} - \frac{X_6 X_7 (X_8 + X_9)}{X_{10}} \quad (1)$$

Side constraints for the design variables are $0.1 \leq X_i \leq 5, i = [1, 10]$. From here it can be seen that the following equation surge a ten-variable function.

B. GREY WOLF OPTIMIZATION
Grey Wolf Optimization (GWO) is a type of meta-heuristic algorithm used for classification and prediction of data. Recently most medical applications involve GWO optimizer for prediction of results with maximized accuracy [66]. By observing how grey wolves conduct their hunting and

how their leadership chain is followed authors [67] proposed a mathematical model that imitates the behavior of grey wolves. Based on the hunting behavior and leadership hierarchy there are four types of wolves namely alpha wolves (α), beta wolves (β), delta wolves (δ) and omega wolves (ω). The mathematical model of the optimizer is also created following this hierarchy.

Alpha wolves are on the first level leading the hierarchy guides the hunting process and the other pack of wolves follow their guidance. The main task of this level of wolves is decision making [37].

Beta wolves which acts as advisors of the first level wolves help them in decision making. Alpha control is transmitted by beta wolves to the whole packet and the responses are forwarded back to the alpha [68].

Delta wolves also known as scouts monitors and protects the pack. The wolves in this pack don't belong to any other category. They look after the territory and alerts the pack if there is any danger [69]. Omega wolves are at the lowest category and they follow the superior wolves.

The mathematical model in GWO illustrates the behavior of the grey wolves. Following mathematical equations representing the hunting behavior are

$$\vec{D} = |\vec{C} \cdot \vec{X}_p(t) - \vec{X}(t) \quad (2)$$
$$\vec{X}(t+1) = \vec{X}_-(t) - \vec{A} \cdot \vec{D} \quad (3)$$
$$\vec{A} = 2\vec{a} \cdot \vec{r}_1 - \vec{a} \quad (4)$$
$$\vec{C} = 2 \cdot \vec{r}_2 \quad (5)$$

Here, t refers to current run. The coefficient vectors are represented by A and C. X_p indicates the position vector of prey and X indicates position of the grey wolf. r_1, r_2 represents random vectors and the component a linearly decreases during the course of iteration and a denotes as distance controlling parameter and D denotes as $|CX_p(t)X(t)|$.

C. PARTICLE SWARM OPTIMIZATION
Particle swarm optimization (PSO) deals with the complex and complicated problems in a way that establish a

connection between the environment and respective simple agents [70]. At the year of 1995 two great men were involved in the invention if PSO. Among these two men one was Russel Eberhant, electrical engineer and the other one is a socio-psychologist, James Kennedy. These two man established a metaheuristic optimization which were lead by the particle swarm. The whole process is based on movement of each particle and the iterations of them. The nearest one to the optimum let know the others about its place, so that trajectory can be updated [71].

There has been many research done on particle swarm optimization. Framework of combinational optimization can be called the most effective one among them [72]. Different objective and particles made by research space need to be defined to apply PSO. The main purpose of this technique is to find out the optimum by moving those particles, Each particles contain their own position coordinates. While iterations occurs, there comes a speed which allows to move the particle. So the particle change their positions and it gets updated to the nearest one which is considered as its best position. So the evolving process helps it to find the optimum one [73].

When a set of particles deals with the best criterion particle, each of them gets to know their position which were also best visited. They are also aware of the must used value of criterion and coordinates. After optimal scheduling, the value was given to the function objective because each of them needed to be compared based on the optical and criterion values of the particle.

Here goes the objective function for PSO which has been used in different kind of research methods also [74].

$$r = \frac{\sum_{i=1}^{n}(x_i - \overline{x})(y_i - \overline{x})}{\sqrt{\sum_{i=1}^{n}(x_i - \overline{x})^2}\sqrt{\sum_{i=1}^{n}(y_i - \overline{y})^2}} \quad (6)$$

Here, n represents the amount of samples. The sample mean for x and y was represented by

$$\overline{x} = \frac{1}{n}\sum_{i=1}^{n} x_i, y_i \quad (7)$$

Each one of a population in the PSO algorithm is called particle. While considering a standard PSO every particle modify its position and speed in every iteration depending on their experience. Which is denoted as *pbest*. For all particles best experience it is noted by *gbest*. They are written in equations below. The cost function which has been defined previously evaluates the performance result of all particles.

$$v^i[t+1] = w \cdot v^i[t] + c_1 r_1(p^{ibest}[t] - p^i[t]$$
$$+ c_2 r_2(p^{gbest}[t] - p^i[t] \quad (8)$$
$$p^i[t+1] = p^i[t] + v^i[t+1] \quad (9)$$

Here, i = 1,2,.... N. By N we get the amount that how many population are there. The speed vector is represented by $v^i[t]$ at $[t]th$ iterations. $p^i[t]$ of the *ith* particle is the present position. For the best position previously of the ith particle it

is represented as $p^i best[t]$. Again, for the whole particle best position it is represented with $p^g best[t]$. 'w' was utilized to control the global and local search pressure. Here, c_1, c_2 are called the cognitive parameters and social parameters. And r_1, r_2 are used as any random numbers between 0 and 1.

IV. TAXONOMY
In last few years, researchers have tried to address the healthcare problem by using machine learning models. Several machine learning models are extensively utilized for detection some common medical conditions. But due to explicit large amount of medical data, machine learning models are not performing well in some cases. But researchers are trying to improve the overall performance of the machine learning models by utilizing optimizers. In this review paper, we have tried to address some recent advances of machine learning optimizers. Taxonomy of this paper is as following.

- Formal introduction of the diseases.
- Recent works on machine learning with optimizers.
- Performance analysis and future direction of the review.

V. LITERATURE ON MEDICAL DATA CLASSIFICATION
In this comprehensive review article, we have contemplated four extensive diseases for conducting the evaluation and insight to the machine learning algorithms with optimization techniques. In this section, some recent literatures are presented.

A. BREAST CANCER
Tabrizchi *et al.* [11] proposed in their paper that for early stage detection of breast cancer they categorized the cancer into two groups of benign and malignant and suggesting a novel collective learning technique utilizing Multi-Verse Optimizer (MVO) and Gradient Boosting Decision Tree (GBDT). To enhance the feature selection they suggested a new effective method based on MVO and GBDT parameters. The two data sets that were utilized in their research are of Wisconsin Diagnostic Breast Cancer and Wisconsin Breast Cancer. Gradient Boosting Decision Tree is a strong collective method that was first invented by Friedman and it's a state of the art technique because it converts the weak classifier into robust ones by mixing them together. They found out that GBDT is quite different from other known techniques since it utilizes function space for the diligence of optimization. Moreover, this technique is more elastic, scalable and strong against the complications of non-linear problems when compared to other linear models for example linear regression. An optimizer was first suggested having a metaheuristic algorithm and is utilized in many researches nowadays to evaluate different complications in various applications. It can be seen that compared to other algorithms, MVO possessed a robust capability for the optimization with lesser control variables. Finally, they have concluded that with the help of MVO and GBDT they have developed an efficient and effective strong classifier for the optimization of the breast cancer datasets with high precision.

Punitha *et al.* [12] proposed an enhanced optimization framework for deep learning algorithms to characterize breast cancer in an efficient way. They named it as IABC-EMBOT. Bee colony and monarchy butterfly optimization algorithms have been combined and specifically designed for breast cancer recognition. They have executed IABC-EMBOT utilizing the MATLAB with the addition of neural network toolbox and required techniques of back propagation that were worked out during the analysis. In their approach for optimization of datasets for the detection of breast cancer cells they have utilized MLP network and consequently for the improved collection the MLP network is trained with IABC-EMBOT approach. The MLP network is divided into three layers each contributing to achieve higher precision and sensitivity of datasets. They have used winner-take-all approach, to enhance and proficient the datasets of cancer cells, during the method of execution of IABC-EMBOT. The datasets were obtained from Wisconsin Breast Cancer Database (WBCD) which consists of two classes. To achieve extraordinary precision of optimization, they have utilized ABC and IMBO algorithms. Finally, they have concluded that with an integrated ABC- BFA (Bacterial Foraging Algorithm) the technique of intelligent breast cancer detection can be achieved with higher accuracy of sensitivity and speed.

Wuniri *et al.* [13] proposed that feature selection is an important part to build a breast cancer classifier for the preventive diagnosis. They have suggested a method known as wrapper method that is an incorporated framework in which the feature selection datasets regarding breast cancer are implanted with the Bayesian classifiers. For handling the discrete and continuous features they have proposed two approach and they are as following (i) A naïve approach regarding the discrete features (ii) For the continuous features it is kernel probability density estimation. These in turn helps to direct to feature-type-aware hybrid in Bayesian classifiers. In their research they have utilized genetic algorithm (GA) to acquire a near optimal subset where it can be seen that Area Under the Curve (AUC) metrics and it's equivalent classifiers produced a good result. The experiments were conducted with the continuous Wisconsin diagnostic breast cancer dataset and the real breast cancer dataset for Chinese women. Also, the algorithm's convergence is improved by the one class-F-score. One class-F-score is another commonly used metric that calculates the features upon their merits and is responsible for crossover operation, population initialization and mutation operation guidance in GA. They have embraced one class-F-score over the original F-score as it can perform well with the breast cancer datasets since they are immensely lopsided. In the experiment with the WDBC datasets, it was found out that with the Big F algorithm the results were better than compared to Bayesian classifier - based Genetic algorithm (BG). Moreover, Big-F algorithm accuracy and efficiency were quite high for the breast cancer diagnosis resulting the researchers to choose it over BG algorithm. Finally, they concluded that Big-F algorithm outperforms, in feature selection and add rapidly, the old-fashioned

BG algorithm and with the supervision of one class-F-score the precision is quite high and they have plan to expand its reach to other cancer datasets.

Memon *et al.* [14] suggested machine learning-based diagnostic system which commendably identifies, the menacing and compassionate people in the environment of IoT, for the early stage identification of the breast cancer diagnosis. They have utilized recursive feature selection algorithm to improve the further precision of obtaining the datasets and building a strong cataloguing for the classification system. To achieve the outstanding prognostic model for the arrangement of the datasets they have utilized testing/training split method. Furthermore, Matthews's correlation coefficient, F1-score, and execution time, which were automatically computed, have been tested for the performance of the classifier and they have used WBCD for acquiring the cancer datasets. To scheme the machine learning predictive model system, they have worked out SVM and Recursive feature elimination (REF) FS algorithm for the accurate detection of the marked classification of the cancer cells in people. They have finally concluded that their proposed approach is quite accurate in detection of the early stage of breast cancer with huge precision sensitivity and specificity.

Ronoud and Asadi [15] have suggested a diagnosis system utilizing the Deep Belief Network to extract the breast cancer cell data's. They have explained that if DBN is embedded with extreme learning machine (ELM) classifier, it can overcome the challenges with the two well-known methods E (T)-DBN-BP-ELM and E (T)-DBN-ELM-BP. To tackle another challenge of the number of neurons and hidden layers it can be seen that architecture optimization is done by the Genetic algorithm (GA). They have proposed another method, in which the challenges are evaluated, E (TW)-DBN utilizing GA. DBN is developed with different layers of RBM and RBM is an artificial neural network consisting of two layers: (i) Single visible layer (ii) Single hidden layer. The ELM is utilized in this particular research as it does not have to tune the hidden layer which in turn helps to build a connection between input layer and hidden layer easily and the computation done in ELM is based upon efficient least squares method. In the first fine tuning test, ELM is utilized by the model E (T)-DBN-ELM-BP for the selection of weights which in turn produces a more suitable initial point in BP algorithm. In E (T)-DBN-BP-ELM, the second method, the two algorithms BP and ELM labor for the tuning of fine steps in first and second. For the last method E (TW)-DBN, the DBN training is done by GA and it can be seen that the network architecture is enhanced by GA for all the three methods. For the assessment of algorithm's performance is done by the two most famous and accessible data sets which are Breast Cancer Wisconsin—Original (WBCO) and Breast Cancer Wisconsin—Diagnostic (WDBC).Finally, they concluded that the suggested three models performed well in the diagnosis but the first two models, E (T)-DBN-BP-ELM and E (T)-DBN-ELM-BP respectively, accomplished quite

outstanding result compared to E (TW)-DBN in the detection with high accuracy of cancer cells.

Kadam *et al.* [16] suggested feature ensemble learning based on Sparse Auto encoders and Softmax Regression (FE-SSAE-SM model) for early identification of breast cancer datasets and the dataset that was used is Wisconsin Diagnostic Breast Cancer (WDBC).The experimental analysis was evaluated on both (FE-SSAE-SM model) and Stacked Sparse Autoencoders and Softmax Regression based (SSAE-SM) model. It was seen that (FE-SSAE-SM model) performed well in comparison to (SSAE-SM) model and other well-known algorithms. In (SSAE-SM) model the end level of stacked autoencoders is supplied to softmax level for categorization. But it's seen that various representations pops when stacked sparse auto encoder based DNN are taught. The efficiency of (FE-SSAE-SM model) is quite high because of fine tuning was done on the input data layer, feature representation layer, and final Softmax layer. The experiment of the suggested models were carried out in MATLAB and the outcome is (FE-SSAE-SM model) performed well with high sensitivity, precision and specificity.

Supriya and Deepa [17] suggested a Breast Cancer (BC) diagnosis system. They have utilized an optimized artificial neural network. For handling the missing attributes and duplicate data, they used RMA and Hadoop MapReduce techniques respectively. Feature selection is utilized to measure performance deviation in the proposed and other machine learning techniques. Features are considered by exploiting Modified Dragonfly Algorithm (MDA). They gathered the dataset from Wisconsin Breast Cancer Database (WBCD), which consists of two individual datasets. Each of the dataset has 34 and 32 attributes by contrast. They cultivated their proposed ANN model by using GWO technique. The GWO is based on bio-inspired algorithm. The algorithm mimics the activities of the grey wolves to survive as a group (pack). Also, the GWO algorithm is sectored as 3 parts: (i) Encircling Prey (ii) Hunting (iii) Attacking the Prey. After feature selection, they obtained higher accuracy. Finally, their experimental finding are validated using IWDT (Improved Weighted-Decision Tree) algorithm.

Wang *et al.* [18] build a technique known as Improved Random Forest (RF)-based rule extraction (IRFRE) which can eliminate the drawbacks of the certain methods in describing the causes of the diagnosis. With the help of the decision tree collected for the cancer (breast) diagnosis, a precise and explicable classification rules can be derived for IRFRE. For the experimental analysis they have devised three steps to develop the method: (i) utilizing the Random Classifier for different decision tree models for the development of production containing plentiful decision rules present (ii) to separate decision rules from the trained trees a rule extraction step is developed (iii) at last an enhanced multi-objective evolutionary algorithm (MOEA) is given the task to find out the best deal among the explanation and precision for an optimal ruled for the constituent rule. The constructed method is analyzed on three datasets of the breast cancer

which are Wisconsin Diagnostic Breast Cancer (WDBC) dataset, Wisconsin Original Breast Cancer (WOBC) dataset, and Surveillance, Epidemiology and End Results (SEER). In IRFRE histories of past histories is combined with new findings so that the analysis takes place faster and all data are collected for computing from patients and doctors of various facilities. With the help of MOEA, it's seen that a precise and understandable rule predictor can be developed based on the extracted rule of Random Forest technique. For an unique sets of solutions in the optimization process to be created they suggested that an improved tool for elimination of identical individuals and recreating various parents. Overall, the experimental outcomes suggested that the proposed method can clarify the black-box methods and the performance is quite high compared to various famous methods. Finally, they concluded that even with high precision the method can still deduce to Random Forest on three different datasets with enhancement.

Liu *et al.* [19] suggested in their paper to eliminate the huge cost linked with detecting a false negative compared to false positive, a different breast cancer intelligent diagnosis method was suggested by them. In their proposed algorithm feature selection is utilized as information gain directed simulated annealing genetic algorithm wrapper (IGSAGAW). Also, cost sensitive support vector machine (CSSVM) learning method is used as for the withdrawal of the top m optimal and for the order of the features is achieved accordingly to IG method. Even though their method of feature selection can acquire the remarkable precision as well as the least misclassification cost and decreasing the complications related with SAGASW technique. The datasets that were used in their paper are Wisconsin Original Breast Cancer (WBC) and Wisconsin Diagnostic Breast Cancer (WDBC). In this paper their main goals were of their suggested classification are to examine the enactment of IGSAGAW on selection of feature and also to find out the effect of CSSVM method. From their experimental analysis it can be seen that the precision of IGSAGAW is quite high on feature selection also the misclassification cost of IGSAGAW performed the best. Finally, they concluded that even though their model (IGSAGAW+CSSVM) were not as much precise when compared to IGSAGAW+BP but the computational time required for their suggested method is quite low thus becoming optimum for the diagnosis for breast cancer detection with low cost.

Sangaiah and Vincent Antony Kumar [20] have suggested an algorithm which is a hybrid-type detection system for the breast cancer diagnosis. For the early detection of breast cancer, the suggested algorithm utilized a ReliefF quality reduction with entropy based genetic algorithm. In their experiment, the datasets were used from Wisconsin Breast Cancer Dataset (WBCD) which have different characteristics organized accordingly. They proposed since original relief method can only work on nominal and numerical properties, also they were unable to covenant with data that were not complete and restricted to two –class. To solve this, they have

utilized ReliefF which is a broad version of the original Relief method and discard all the shortcomings of the original. ReliefF is not restricted by data types, rapid and can work quite well in noisy surroundings. In classification system, for the detection of the most applicable features, genetic algorithm is being utilized. Entropy is defined as the similarity of datasets and supplies additional perception about the gain of information. In this suggested technique by them, entropy is being utilized as the detection of the applicable features. This helps to build a strong correlation between the input and output and discarded the extra components in the input features. Their proposed method for breast cancer diagnosis which is dependent on ReliefF and Entropy GA technique is found to be quite highly accurate for small features as well as the time required to build the model is also quite low.

Haung [21] suggested a system of high precision for the detection of breast cancer and for this they developed an improved machine learning framework. Mainly, in this framework fruit fly optimization algorithm (FOA) is being utilized with Levy flight (LF) method (LFPA) for the purpose of optimization for the purpose of two vital features of support vector machine (SVM) and construct LFOA-based SVM (LFOA-SVM) for breast cancer diagnosis. The fruit fly optimization is defined as a meta-heuristic method that is enthused by the imitating characteristics of a fruit fly. In FOA the solution of candidate is unanimously developed in the solution space and consequently each fruit fly will apprise their formation with accordance to the fruit fly mode. LFOA-SVM model optimizes the inner bound which in turn vigorously alters the parameters of the SVM with the help of LFOA. Consequently, this optimized parameters are provided into the SVM prediction model for the outer loop to carry out the classification task of breast cancer diagnosis. All the experiments of the models were done on MATLAB and the dataset that was used is Wenzhou people's Hospital (2004-2015). They found out that the center point uniqueness of LFOA-SVM that improves the eminence for FOA lies in collaboration of FOA with LF technique. Finally, they concluded that LFOA-SVM technique performs the best among the other similar techniques and can cause early diagnosis of the cancer with high precision that can help the doctors.

Abdar [22] suggested a nested collective model that utilized the Stacking and Vote (Voting) as the classifiers mixture of models in their ensemble technique for the identification of breast cancer tumors. ''Classifiers'' and ''Meta Classifiers'' are the two contents of their each ensemble classifier. In their paper, they have built a two layer nested classifiers where the Meta Classifiers contains at least two various algorithms. They piloted their experiment on the Wisconsin Diagnostic Breast Cancer (WDBC) and it can be seen that their proposed model worked very well with high precision and efficiency compared to other related models. Also, among their proposed methods SV-BayesNet-3-MetaClassifier and SV-Naïve Bayes-3 MetaClassfier the later takes a reduced amount of time to construct which help to enhance the efficiency even though both of them have the same precision. This less time

taken to construct and high precision of the proposed model helps to identify the cancer tumors at a very early stage.

B. HEART DISEASE
Tama *et al.* [23] has introduced a new method of Coronary Heart Disease (CHD) prediction with machine learning techniques. They have built a two- tier ensemble where some of them were used as base to others. They have used Gradient Boost Machine, Random Forest Classifier and Extreme Gradient Boosting. The data sets that had been used during this research were Cleveland, Statlog, Hungarain. An optimized feature was selected for the significant sets of data. Later on, there has been a statistical test too. The experiment showed it outperformed any other models in terms of F1, accuracy and AUC.

Nilashi [24] has established a detective model to detect the heart disease based on machine learning. This model was established by supervised and unsupervised machine learning techniques. Fuzzy SVM and Principal Component Analysis (PCA) was used along with two imputation processes. The two imputation techniques were basically used for the missing value imputation. Moreover, they have applied the increased FSVM and PCA for increased learning of the data. It was done to lessen the computation time. This was related to disease prediction. This experiment was done on two real data sets of data. The whole experiment showed that increased FSVM helps to gain better accuracy for detecting a heart disease. Also, it stated that it helps to lessen the time of computation.

Sharma *et al.* [25] has deal with the modified artificial plant optimization (MAPO) technique. Plant optimization (MAPO) technique has been used as an optimal feature selector. It was combined with other machine learning techniques. By this process it can find out the rate of heart by utilizing the data set of fingertip video. This results in the prediction of coronary heart disease at that moment. The dataset was made from a long time ago and all noises were filtered. Use of Standard Error Estimate and Pearson Correlation was seen here. They give the estimation of 2.418 and 0.9541 respectively. The prediction was utilized on two other different dataset and MAPO has been applied again to detect coronary heart disease. The experiment result showed a significant outcome from MAPO which lessened the dimensionality. Compared to other dimensionality, there was a reduction of 81.25%. So basically it outperformed the other optimizer with far better accuracy.

Reddy *et al.* [26] has established a hybrid genetic algorithm and a fuzzy logic classifier for heart disease detection. By this, the heart disease can be detected in the early stage. This hybrid technique was comprised of rough set data and the fuzzy rule classification model. The fuzzy classification were optimized by the algorithm called genetic algorithm. In the rough set, the important data were selected, which will cause heart disease. Then an adaptive genetic algorithm with fuzzy logic (AGAFL) hybrid classifier predicts the disease. The whole test was done on the data set of the UCI heart

TABLE 1. A summary of most recent works for breast cancer identification using machine learning models with optimizers.

Authors	Year	Dataset	Data Description	Data Splitting	Optimization Techniques for Diagnosis	Evaluation Metrics
[11]	2020	WBC, WDBC	WBC, 10 Features and 699 Instances, WDBC, 32 Features and 569 Instances	K-fold Cross-Validation	MVO with GBDT	WBC: Accuracy = 97.13%, Specificity = 97.16%, Sensitivity = 97.09%, WBDC: Accuracy = 98.76%, Specificity = 97.64%, Sensitivity = 99.43%
[12]	2019	WBCD	Attributes = 32, 34, Classes = 2	-	MLP network trained with ABC and IMBO	Accuracy = 97.53%, Sensitivity = 96.75%, Specificity = 97.04%
[13]	2019	WDBC, RBCD	WDBC, Attributes = 10, Features = 30. RBCD, Features = 4	RBCD, Training = 70%, Testing = 30%	One class-F-classifier with Big-F algorithm	Accuracy=97.10%, Sensitivity = 92.40%, Specificity = 100%, AUC = 99.4
[14]	2019	WBCD	Instances = 569, Features = 32	70% Training, 30% Testing	REF with SVM classifier	Accuracy = 99%, Specificity = 99%, Sensitivity = 98%
[15]	2019	WBCO, WDBC	WBCO, 683 Instances, Attributes = 9, Classes = 2, WDBC = 569 Instances, Attributes = 30, Classes = 2	80% Training. 20% Testing	GA with E(T)-DBN-BP-ELM and E(T)-DBN-ELM-BP & , E(TW)-DBN	E(T)-DBN-BP-ELM: Accuracy = 99.45%, Sensitivity = 100%, Specificity = 99.17% E(T)-DBN-ELM-BP: Accuracy = 99.75%, Sensitivity = 100%, Specificity = 98.80% E(TW)-DBN: Accuracy = 98.05% , Sensitivity = 99.46%, Specificity = 97.295%
[16]	2019	WDBC	Features = 32, Class Distribution (B/M) = 357/212	10-fold-cross-validation	FE-SSAE-SM model	Accuracy = 98.60%, Sensitivity = 97.16%, Specificity=99.44%
[17]	2019	WBCD	Attributes = 32, 34, Classes = 2	k-fold-5 cross validation	Grey Wolf Optimization	Accuracy = 99.20%, Precision = 99.55%, Recall = 99.55% (with Feature Selection)

disease dataset. The results had better output and outran all the other techniques.

Khourdifi and Bahaj [27] have worked with Fast Correlation-Based Feature (FCBF) to outrun the repeating features so that they can improve the quality of the classification of heart disease. They have also performed different classification such as Random Forest, SVM, Multilayer Perceptron, Naïve Bayes and K-nearest neighbor. Moreover the utilization of artificial neural networks was there along with particle Swarm Optimization by combining the Ant Colony Optimization (ACO) process. This hybrid mixture approach was applied to the set of data of heart disease. They have

TABLE 1. *(Continued.)* A summary of most recent works for breast cancer identification using machine learning models with optimizers.

[18]	2019	WDBC, WOBC, SEER	WDBC, 569 Instances, Features = 32, WOBC, 699 Instances	10 K - fold Cross-Validation	IRFRE with MOEA	WDBC: Accuracy = 95.09%, Sensitivity = 93.4%, Specificity = 96.09%, WOBC: Accuracy = 96.44%, Sensitivity = 96.3%, Specificity = 96.73%, SEER: Accuracy = 80.45%, Sensitivity = 78.29%, Specificity = 82.49%
[19]	2019	WBC, WDBC	WBC: Attributes = 10, Class Distribution (B/M) = 458/241, WDBC: Attributes = 32, Class Distribution (B/M) = 357/212	10-fold-cross-validation	IGSAGAW - CSSVM	Accuracy = 95.7%
[20]	2018	WBC	Features = 75	-	Entropy based Genetic Algorithm	Proposed RF-RGA: Accuracy = 82.05%, Precision = 83%, Sensitivity = 73%, Specificity = 88%
[21]	2018	Wenzhou people's Hospital	Features = 14, Class Distribution (B/M) = 232/238	10-fold-cross-validation	LFOA-based SVM	Accuracy = 93.83%, Sensitivity = 91.22%, Specificity = 96.53%
[22]	2018	WDBC	Features = 32, Class Distribution (B/M) = 357/21	K-fold cross - validation	SV-BayesNet-3-MetaClassifier and SV-Naïve Bayes-3 MetaClassfier	SV-BayesNet-3-MetaClassifier: Accuracy = 98.07%, Time required = 2.89 sec, SV-Naïve Bayes-3 MetaClassfier: Accuracy = 98.07%, Time required = 2.56 sec

also studied different machine learning algorithms and compared their results. Thus they got the accuracy of 99.65% for the hybrid model consisting of FCBF, ACO and PSO. The results were satisfying compared to the process of machine learning.

Parthiban and Subramanian [28] have introduced a new technique called coactive neuro-fuzzy inference system (CANFIS) for the detection of heart disease. The model was collaborated with the machine learning techniques of neural network capabilities and the fuzzy logic. This full mixed qualitative technique was integrated with the help of genetic processes. It helps to detect the presence of that disease. The results of this model were evaluated based on classification, perfectness, and the performance of training. At last the result has shown a better potential accuracy and efficiency.

Gokulnath and Shantharajah [29] have introduced an optimizer based on the Support Vector Machine (SVM) classifier function. Genetic algorithm (GA) is one of the uses of this objective function. By this way many more significant characteristics can be chosen for predicting heart disease. There have been some good experimental outcomes when they compared the GA-SVM with CFS, Relief, Filtered subset, consistency subset, gain ratio etc feature selectivity techniques. The analysis for the operation of receiver features was performed to determine how the SVM classifier works. It came out pretty good. Moreover the dataset collected from Cleveland was processed in MATLAB along with the desired framework.

Nalluri *et al.* [30] have proposed a hybrid system which utilizes optimization of classifiers. SVM and Multilayer perceptron technique. Other than that 3 recent algorithms was deployed for the optimization of the parameters. Moreover,

it leads to six alternative hybrid disease diagnosis systems. It can also be called a hybrid intelligent system (HISs). Sensitivity, accuracy, namely, objectives and specificity has been compared with the other techniques. The desired model was evaluated on the benchmark of 11 sets of data. The experimental result showed that it performs better in terms of our criteria than other techniques. Furthermore, a statistical test was deduced to substantiate the efficiency of the results.

Liu [31] has focused on the study to help the process of diagnosis of heart problems by utilizing a classification system which is hybrid. Rough set and Relief f methods were the key of this whole experiment. The desired system contained two sub models. One is a system of classification with ensemble classifiers and other is the selectivity system of characteristics of RFRS. Three stages were introduced to the RFRS system. These are data discretization, feature extraction using the Relief F, feature reduction using the heuristic rough set of reduction. The second system was based on the C4.5 classifier. For the jackknife cross validation process 92.59% accuracy was obtained. It showed the technique gives better performance compared to other classification techniques.

Bashir *et al.* [32] has proposed a novel classifier ensemble framework that has been based on increased approach of bagging along with the scheme of voting which is weighted multi objectively for analysis and prediction the disease of heart. By using this model, many limitations were overcome as well as resolved some performance issues. It was compared to the five other quadratic discriminant analysis, Naïve Bayes, SVM, linear regression, Instance based learner. The authors have utilized five different types of sets of datas for the experiment, validation and evaluation. Those were public data. Upon the comparison of different classifiers, the proposed model has been validated by ANOVA statistics and ten-fold cross validation. Also, the results showed an accuracy of 84.16%. 93.29% sensitivity, specificity of 96.70% were obtained with the proposed model. F-critical was lower than f-ratio and p-valued came out smaller than 0.05 as 95% results proved significant for all the datasets.

Long *et al.* [33] have briefly discussed and dealt with one of the heart disease diagnosis systems which was based on rough sets attribute reduction and a logic system called internal type-2 fuzzy (IT2FLS). The attribute reduction which has been based on rough sets had to undergo the integration process. For IT2FLS it was the same, didn't change. The goal was to deal with the uncertainties and challenges of big datasets. The IT2FLS used a learning process which is hybrid by consisting of the algorithm called fuzzy C mean clustering. When it comes to work with the dataset which are high dimensional, the process of learning is expensive. To find out the optical reduction, firefly algorithm was used. It was seen that, it increased the working performance of IT2FLS. The result of the experiment has a very improved and dominant significance compared to other machine learning techniques such as SVM, Naïve Bayes, artificial neural network. Thus it was told to be very useful for heart disease diagnosis.

C. DIABETES

Mishra *et al.* [34] has proposed a new hybrid attribute optimization algorithm known as Enhanced and Adaptive Genetic Algorithm (EAGA) for diabetes diagnosis. Authors used the optimization algorithm on Pima Indian diabetes diagnosis. Furthermore the EAGA model was used with Multilayer Perceptron (MLP) and this hybrid classification approach was implemented for diabetes diagnosis. This EAGA-MLP model was also compared with several existing studies and the performance was evaluated based on some important performance metrics. Experimental result showed that this model has the maximum accuracy rate and least computational time among all the previous classification algorithms.

Massaro *et al.* [35] has focused on applying Long Short-Term Memory (LSTM) network for predicting health status of diabetes patients. The proposed LSTM network in this study is also suitable for DSS platforms. Authors used the Pima Indians Diabetes Dataset that contains several medical predictor variables and a single target variable. Traditional LSTM neural networks and a new approach LSTM artificial records-AR were integrated into the information system to collect patient information and data. The LSTM artificial data was adopted for improving the training dataset in order to obtain better accuracy. Then a prediction model was created based on co-occurrent analysis of several attributes. Results showed that by using LSTM-AR- approach a better accuracy can be found compared to traditional LSTM approach and Multi-Layer Perceptron (MLP) applied on the datasets.

Choubey *et al.* [36] evaluated the performance of several classification methods with Principal Component Analysis (PCA) and Particle swarm optimization (PSO) for diabetes. PCA which is basically a feature reduction method can identify which features are relevant in the feature set of linearly dependent features. PSO method that is also a popular stochastic optimization technique also used for feature reduction. Authors used the Pima Indian Diabetes Dataset and Localized Diabetes Dataset to analyze their research work. In the first approach of their research they utilized several classification methods namely Logistic Regression, K-Nearest Neighbor (KNN), ID3 DT, C4.5 DT, Naive Bayes, KStar on the datasets. In another approach authors used PCA and PSO before classification of dataset for feature reduction. After that classification of the dataset was done using the same classification methods. The comparative analysis between the two approaches showed that the later approach that used PCA and PSO algorithms resulted in higher accuracy and less computational time. Authors also suggested that this proposed approach can also be applied efficiently for early diagnosis of other medical diseases.

Shankar Babu *et al.* [37] presented a diabetes prediction model by using the theory of fuzzy rule. The optimization that the authors used for this model is grey wolf optimization. This optimizer gives better and efficient accuracy in prediction. In this optimizer the first step is to initialize the grey wolf optimizer population and initialization of alpha wolves and beta wolves. The fitness value is then

TABLE 2. A summary of most recent works for heart disease identification using machine learning models with optimizers.

Authors	Year	Dataset	Data Description	Data Splitting	Optimization Techniques for Diagnosis	Evaluation Metrics
[23]	2020	Z-Alizadeh Sani, Statlog, Cleveland, and Hungarian	Features = 54, 13, 13, 13 Instances = 303, 261, 303, 294	10-fold cross validation	Particle Swarm optimization	Accuracy = 83.90%, 76.82%
[24]	2020	Cleveland and Statlog from UCI	Attributes = 13, Classes = 2	80% Training. 20% Testing	GA with E(T)-DBN-BP-ELM and E(T)-DBN-ELM-BP & , E(TW)-DBN	Accuracy = 87%, 88%, Specificity = 81.31%, 82.35%, Sensitivity = 86.93%, 82.35%
[25]	2019	Fingertip's Video Dataset, Kaggle.	Attributes = 14	-	Artificial Plant Optimization Algorithm,Modified Artificial Plant Optimization (MAPO)	Accuracy = Dimensionality was reduced to 81.2%
[26]	2019	Switzerland, Hungarian and Cleveland dataset from UCI.	Attributes = 14	-	AGAFL	Accuracy = 90%, 91%, 89%, Specificity = 90%. 88%, 75%, Sensitivity = 91%, 92% ,97%
[27]	2018	Heart Disease dataset from UCI machine learning repository.	Attributes = 14	Clustering algorithms, 10-fold cross validation.	Particle swarm optimization (PSO), FCBF, Ant colony optimization (ACO)	Accuracy = 99.65% with KNN, 99.6% with RF.
[28]	2018	Cleveland heart-disease database.	Attributes = 14	-	CANFIS and Genetic Algorithm	-
[29]	2018	Cleveland heart disease database.	Attributes = 11	Ten fold cross validation and ANOVA statistics.	Genetic algorithm (GA)	Accuracy = 83.70%
[30]	2017	Cleveland and Statlog from UCI	Attributes = 11	Ensemble classification techniques	Support vector machine (SVM) and multilayer perceptron (MLP) technique.	Accuracy = 85.8%, 85.9%, 85%, 82.4%, 81.3%, 96.99%, 87.1%, 97.7%, 93.8%, 77.60%, 73%
[31]	2017	Statlog (Heart) dataset, obtained from the UCI database.	Attributes = 13	Data discretization, feature extraction using the ReliefF algorithm	ReliefF and Rough Set (RFRS)	Accuracy = 92.59% using the ensemble classifiers when k = 100
[32]	2015	SPECT, SPECTF, Heart Disease, Statlog, Eric from UCI	Attributes = 23, 45, 14, 14, 8	Ten fold cross validation and ANOVA statistics.	Bootstrap aggregation with multi-objective optimized voting	Accuracy = 84.16%, Sensitivity = 93.29%, Specificity = 96.70%
[33]	2015	Balanced dataset and unbalanced dataset from UCI Repository Machine Learning Database.	Attributes = 13, 44, Classes = 2	k-fold-5 cross validation	Chaos firefly algorithm, Binary particle Swarm optimization and Rough sets based attribute reduction.	Accuracy = 87.2%

calculated. After maximum iteration the values are updated and the fitness value is calculated again. The classification result is finally obtained by using the updated fuzzy rules. Authors also used ant colony optimization for predicting the model. After performance analysis of both optimizers, authors found that GWO gives much better results

in terms of accuracy, precision and recall than Ant Colony Optimization (ACO).

Bani-Hani *et al.* [38] implemented a Recursive General Regression Neural Network(R-GRNN) oracle in their research for prediction and diagnosis of diabetes. GNN which is an enhanced oracle of GRNN takes into account the

predictions of individually trained classifiers as well as the first GRNN oracle. It then gives one superior prediction as output based on the error rate for each classifier that was taken from a set of observations. Dataset used in this research was taken from Pima Indians Diabetes Dataset. It contains data of about 768 observed patients among which 268 have diabetes. For feature selection and hyperparameter optimization authors utilized Genetic Algorithm (GA). They used several classifiers along with R-GNN Oracle namely Support Vector Machine (SVM), Random Forest(RF), Probabilistic Neural Network (PNN), Multilayer Perceptron (MLP), K-Nearest Neighbor (KNN), Gaussian Naive Bayes (GNB) and GRNN Oracle. After performance evaluation it was observed that the proposed model (R-GNN) resulted in the highest accuracy, AUC and sensitivity than the other models. However in terms of specificity optimized MLP resulted in higher value than R-GNN. For future study authors suggested applying hyperparameter optimization and feature selection instead of utilizing feature selection based on optimized hyperparameters from all the features.

Alirezaei *et al.* [39] proposed a hybrid optimization algorithm for reducing noise and data dimension in diagnosis. In their research authors analyzed PIMA Indian Type-2 dataset and applied several methods to figure out which gives the best classification accuracy. After using KNN algorithm for imputing missing data, K-means algorithm was utilized for detecting and deleting outliers. For data pre-processing they used four meta-heuristic algorithms namely Non-dominated sorting genetic algorithm (NSGA-II), multi-objective particle swarm optimization (MOPSO), multi-objective firefly (MOFA), and multi-objective imperialist competitive algorithm (MOICA) that helped to obtain effective and significant features. Finally a SVM was implemented for classification of data. Authors used several performance metrics for comparing the algorithms. After performance comparison analyses it was seen that MOFA proved to be the best algorithm among the four optimizers in terms of the performance metrics.

Cui *et al.* [40] utilized an improved support vector machine algorithm to predict readmission of diabetic patients. Authors investigated readmission rate of diabetic patients in 130 United States hospitals. Due to imbalance of data, they utilized an efficient SMOTE-based class imbalance processing method. The support vector machine classifier (SVM) was used to classify the data. To optimize the SVM classifier GA was used. GA is implemented to tune three kernel functions linear, polynomial and sigmoid kernel to provide better accuracy. It also helps in searching optimal parameters for RBF-based support vector machines. Authors also compared their proposed method with other methods like LACE score, Naive Bayes, Decision Tree, Logistic Regression and Back Propagation Neural Network (BPNN). Experimental results showed that their proposed diabetic readmission prediction method proved to be superior to the other compared methods.

Paul and Choubey [41] have used GA-RBF NN classification system on Pima Indian Diabetes Dataset (PIDD) for

the diagnosis of diabetes. In the first stage of their work, authors used GA for feature selection. It helped to remove insignificant features. Moreover, it provides less computational time with increased accuracy. In the next stage Radial Basis Function Neural Network (RBF - NN) was utilized for classification of data. This classification method was compared with some earlier methods and it was observed that the later requires less cost and improved training and classification. Authors also suggested to use this classification system for several other medical diseases.

Maniruzzaman [42] applied four classification techniques to diagnose and classify diabetes dataset. Authors proposed Gaussian process classification (GPC) technique in their study compared with some existing classification techniques. GPC technique is introduced in the study as it is capable of handling several problems in regression and classification like complex data types, linearity of classical methods, curse of dimension etc. Moreover uncertainty in unknown functions can also be easily handled using the Gaussian process. Authors used the classifiers in the Pima Indian diabetes dataset. Apart from GPC classification technique the other classification techniques that were used are LDA, QDA and NB. Then performance evaluation was done based on the parameters: classification accuracy (ACC), sensitivity (SE), specificity (SP), positive predictive value (PPV), negative predictive value (NPV). Performance analysis showed that GPC classification has the best outcome results among all the four classifiers.

Karegowda *et al.* [43] integrated Genetic Algorithm and Black Propagation network (BPN) to implement a hybrid model. In the model they used Genetic Algorithm for optimizing and initializing the connection weights of Black Propagation network. GA is utilized for improving the performance of BPN in many ways. It is an efficient search method that is able to find out large search spaces that can be used with BPN for determining the number of hidden nodes and hidden layers. It can also select relevant feature subsets, determine the learning rate, momentum and optimize the network connection weights of Black Propagation network. For finding out the relevant features, authors used Gain ratio and Correlation based feature selection. The dataset that was used for this model was taken from Pima Indians Diabetes Database (PIDD). The classification of this dataset was done using the hybrid GA-BPN model. This model showed significant improvement in classification accuracy.

Ganji and Abadeh [44] presented FCS-ANTMINER which is a new classification algorithm obtained by combining Ant Colony Optimization (ACO) and Fuzzy logic. They used this classification algorithm for diabetes disease diagnosis. In their research they used the Pima Indian Diabetes dataset (PID) and implemented FCS-ANTMINER on this dataset. The FCS-ANTMINER consists of a training stage and a testing stage. During the training stage, an ACO algorithm is applied where fuzzy rules are generated through training patterns. In the testing stage, for classifying the test patterns, a fuzzy inference engine is utilized

by the test classifier. After using the proposed classifier (FCS-ANTMINER) the result showed that it has the highest classification accuracy compared to several recognized classification algorithms.

D. PARKINSON DISEASE

Tomar *et al.* [45] has utilized Least Squares Twin Support Vector Machine (LSTSVM) for Parkinson disease diagnosis. This classifier was used on Parkinson disease dataset. LSTSVM classifier is very much efficient as it takes less time for computation and provides better generalization ability. For feature selection Particle Swarm Optimization (PSO) was used. The parameters were also optimized using PSO. Authors evaluated the classification model based on some parameter metrics namely accuracy, sensitivity and specificity. It was observed that LSTSVM+PSO classifier model gave the best performance result when used with Gaussian Kernel function. Datasets were taken from UCI. The experiment was resulted in accuracy, specificity and sensitivity.

Sharma *et al.* [46] proposed a modified version of grey wolf optimization for diagnosing Parkinson's disease. Modified Grey Wolf Optimization (MGWO) which is a special type of Grey Wolf Optimization (GWO) is utilized in the research for feature selection. MGWO takes a group of features as input and gives reduced and relevant features as output. This optimizer helps to get better accuracy result for recognizing Parkinson's disease. Authors implemented their proposed model on several dataset like voice, handwriting and speech. The proposed algorithm was compared with Optimized Cuttlefish Algorithm (OCFA) and it was observed that MGWO algorithm maximized the accuracy result and using minimum features for PD diagnosis. These data contain different voice, speech and handwriting. The proposed system helped to detect the illness with a 94.83% accuracy containing 98.28% rate of detection along with 16.03% alarm rate. Later on these were also compared to the result of Optimized Cuttlefish algorithm (OCFA). But the proposed model showed better results.

Shahsavari [47], classification of Parkinson's Disease (PD) patient dataset was analyzed using the Extreme Learning Machine (ELM) model. This model acts as a type of feed-forward neural network having a single hidden layer for classifying the patients. ELM provides faster computation and better generalization. Moreover, it has the ability of universal approximation. Authors used Hybrid Particle Swarm Optimization (PSO) for feature selection. This optimizer helps to extract relevant features and the ELM-based model is also improved through it. Dataset used for this research was taken from a PD patient's medical information dataset. Then the proposed model was implemented on the dataset. Experimental results showed that the proposed model leads in Accuracy, Recall, Precision and F-score than the other compared models.

Cai *et al.* [48] proposed an optimal SVM based on bacterial foraging optimization (BFO) for the identification of Parkinson Disease. The fruitfulness of BFO+SVM was proved to be precise on the datasets of PD established on vocal computation. Comparison of the suggested model (SVM+BFO) was done with other optimization techniques such as SVM based on the grid search method and an SVM based on PSO. BFO is categorized as three main procedures which are (i) Chemotaxis (ii) Reproduction (iii) Elimination-dispersal. There are mainly main kernel functions such as linear kernel, polynomial kernel, sigmoid kernel and radial based kernel. The RF-BFO-SVM framework was constructed to distinguish between the PD patients from the healthy controls by mixing the feature selection with the proposed model (BFO+SVM). All the experiments were executed on MATLAB and the evaluation metrics of the proposed model was precise and satisfactory.

Sankara Babu *et al.* [37] suggested in their paper that the vital goal of their work was to identify the various diseases at an early stage. They proposed utilizing GWO and auto encoder based Recurrent Neural Network (GWO+RNN) for identification purposes. GWO is being utilized for feature selection purpose and RNN technique is being utilized for the identification of diseases. GWO is chosen over other state-of-the-art because it eliminates unnecessary and inessential characteristics ominously. This is done after the features are sent out to the RNN classifier. The datasets that were used in their paper are Hungarian, Cleveland, PID, mammographic masses, Switzerland. From their experiment, they observed that GWO+RNN technique's effect on identification of diseases accomplished enhanced result compared to similar techniques like Group Search Optimizer and Fuzzy Min-Max Neural Network (GFMMNN) approach. The simulation was done on PyCharm and the evaluation metrics were evaluated, suggesting that the proposed model GWO+RNN method is an accurate model.

Shen [49] proposed a Fruit Fly Optimization based SVM (FFO-SVM) was implemented for medical diagnosis classification. Authors used four datasets namely Parkinson Dataset, Pima Indians Diabetes Dataset (PIDD), Wisconsin breast cancer dataset and Thyroid dataset. The fruit fly optimizer follows a few steps to work. After parameter and population initialization it evaluates the population. Then after replacing the smell concentration value with Fitness function it finds the maximal smell concentration. After keeping the maximal concentration and going through several iterative optimization it gives the best optimized value. In this study, FOA optimizer was used for optimizing the SVM parameters and SVM model performs classification task based on the optimal parameters gained by FOA technique. Authors also used GA-SVM, BFO-SVM and Grid-SVM on the same datasets to compare with the proposed model. Performance analysis showed that the proposed method (FOA-SVM) gave the maximum classification accuracy with lease computation time. Authors also suggested to use this model for further medical diagnosis.

Qiang Li *et al.* [50] suggested in their paper a unified system known can be developed, grey wolf optimization (IGWO) and kernel extreme learning machine (KELM) termed as

TABLE 3. A summary of most recent works for diabetes identification using machine learning models with optimizers.

Authors	Year	Dataset	Data Description	Data Splitting	Optimization Techniques for Diagnosis	Evaluation Metrics
[34]	2020	Pima Indians Diabetes Database (PIDD)	Total Instances: 768, Total attributes: 9	-	Enhanced and Adaptive Genetic Algorithm (EAGA)	Accuracy: 94.7%, Precision: 91%, Recall: 89.8% F1-score: 90.4%
[35]	2019	Pima Indians Diabetes Database (PIDD)	Total instances: 768	Training Dataset= 80%, Testing Dataset= 20%	LSTM-AR Model	AUC: 89%, Accuracy: 84%, Loss: 50%
[36]	2019	Pima Indians Diabetes Database (PIDD), Localized Diabetes Dataset (LDD)	Total cases for PIDD: 768, Positive instances: 268, Negative instances: 500No.of features: 8, No. of classes: 2, Total cases for LDD: 1058, Positive instances: 753 Negative instances: 305, No.of features: 12, No. of classes: 2	Training = 70%, Testing = 30%	Principal Component Analysis (PCA) and Particle Swarm Optimization (PCO)	Highest Accuracy, PCA-C4.5 DT accuracy: 95.58 % PCA-C4.5 DT ROC: 0.981, PSO-C4.5 DT accuracy: 94.00 %, PSO-C4.5 DT ROC: 0.938
[37]	2019	Pima Indians Diabetes Database (PIDD)	Number of features: 8, Number of classes: 2	-	Grey Wolf Optimization	Accuracy: 81.158, Precision: 76.24, Recall: 80.48
[38]	2019	Pima Indians Diabetes Database (PIDD)	Observed Patients: 768, Diabetic Patient: 268, Independent features: 8, Dependent feature: 1	-	Recursive General Regression Neural Network (R-GRNN)	Accuracy: 81.14, AUC: 0.86, Sensitivity: 63.80, Specificity: 89.14
[39]	2019	Pima Indians Diabetes Database (PIDD)	Total Observation: 768, Features: 9	-	Non-dominated sorting genetic algorithm(NSGA-II), multi-objective particle swarm optimization (MOPSO), multi-objective firefly (MOFA), multi-objective imperialist competitive algorithm (MOICA)	Accuracy for NSGA-II: 98%, Accuracy for MOFA: 100%, Accuracy for MOPSO: 94.64% Accuracy for MOICA: 100%, Sensitivity for NSGA-II: 100%
[40]	2018	Health Facts Database	Total medical records: 8756, Total features: 50	-	GA-SVM with RBF	Accuracy: 0.81, Sensitivity: 0.82, Specificity: 0.79, F1-: 0.819, G-mean: 0.81
[41]	2017	Pima Indians Diabetes Database (PIDD)	Number of instances: 768, Number of attributes: 8	Training Dataset= 70%, Testing Dataset= 30%	GA-RBF NN	Accuracy: 76.0, ROC: 0.813
[42]	2017	Pima Indians Diabetes Database (PIDD)	Total Number of patients: 768, Diabetic patients: 268, Number of attributes: 9	-	Gaussian process classification (GPC)	ACC: 81.97%, SE: 91.79%, SP: 63.33%, PPV: 84.91%, NPV: 62.50%

IGWO-KELM, for the detection of medical diseases. The selection of feature is utilized for the objective of uncovering the optimal feature subset of the records of the medical. They at first embraced the genetic algorithm (GA) to produce mixed initial positions and Grey wolf optimization is utilized for the refresh of current positions of habitants in the secluded space. This results of achieving the best feature subset for the improved classification based on KELM. After experimenting the suggested model on the two common diagnosis pattern and evaluating different evaluation metrics, it can be observed that the suggested model IGWO-KELM outperformed the other two equivalent models which are GWO and

TABLE 3. *(Continued.)* A summary of most recent works for diabetes identification using machine learning models with optimizers.

[43]	2011	Pima Indians Diabetes Database (PIDD)	Total cases: 768, No missing, value cases: 392, Positive cases: 130, Negative Cases: 262	Training Dataset: 60 %, Testing, Dataset: 40%	Genetic Algorithm and Back Propagation Network (BPN)	Classification accuracy with all attributes: 77.707 %, Classification accuracy with Decision Tree: 84.076 %, Classification accuracy with GA-CFS: 84.713 %
[44]	2011	Pima Indians Diabetes Database (PIDD)	Total Instances: 768, Features: 8, Classes: 2	-	FCS-ANTMINER	Classification accuracy: 84.24%

GA. Finally, they have concluded that since their suggested method has high precision and efficiency on the two datasets so it can be utilized for early detection and in the future for similar related practical medical diseases.

Wu [51] has focused on Deep Brain Stimulation (DBS). They have approached the Radial Basis Function Neural Network (RBFNN) along with Particle Swarn Optimizer (PSO). Also the use of principal component analysis (PCA) and Local Field Potential (LFP) was seen in this research paper. The whole system was trained by real life Parkinson patient. Thus they found out the detection of accuracy can be rise up to 89%. Moreover, there have been a comparison among RBFNN based PSO & RBFNN. But RBFNN based PSO was far behind in terms of performance with a fair reduction in computation.

VI. DISCUSSIONS, CHALLENGES AND FUTURE DIRECTIONS

In this review paper, we have discovered some recent literature on medical data classification utilizing optimized machine learning algorithms. Several optimization techniques are involved nowadays in machine learning algorithms to achieve better results. Presently, there are a huge span of data globally. But most of the datasets are unstructured and not well organized for machine learning algorithms to reuse. Even some datasets have lower quality to extract key features to achieve an outstanding accuracy. In this regard, various optimization techniques are incorporating to extract key features from the datasets. This review insights some recent works and demonstrates the impact of optimizers in machine learning systems. The whole review article introduces the overall optimization impact, data pre-processing methods, performance evaluation techniques in different datasets. There are several considerate optimization technique are utilized extensively in recent works. These widely used optimization techniques are genetic algorithms, particle swarm and grey wolf optimization techniques. However, these are base techniques. But most of the have used ensemble optimization techniques using these base optimization techniques.

For the review convenience, Table 1 provides an overview of recent advances of machine learning with optimizars in breast cancer prediction. From the table, we can see [11]–[22] have utilized several optimization methods. In most literature, they have utilized WBC, WBCD and WBCO datasets. K-fold validation based dataset splitting process is used. Where, only [11] is utilized 5 k-fold cross validation used. But in [14] and [15] have utilized only manual data splitting method. However, in [18]–[22] have utilized 10 k-fold validation. In the [13], they have achieved the highest accuracy of 99.20%, which is higher than all other recent works. They used WBCD dataset, and 5 k-fold based data splitting. By contrast, they have enforced grey wolf optimization technique to achieve this remarkable performance. Where the most works have the accuracy ranging from 82% to 99%. On the other hand, genetic algorithm based E(T)-DBN-BPELM have performed well in [15] with an accuracy of 99.45% in WBCO dataset. Moreover, In [79]–[86], the authors have discussed about and deployed traditional machine learning and deep learning models to detect breast cancer without optimizers. Although, they have also achieved a decent outcome. But with optimization solutions are more robust for detection breast cancer.

Table 2 demonstrates heart disease classification using machine learning with optimizations. This is an comprehensive outlook of recent machine learning based classification task for heart disease prediction. However, from the table we can see [23]–[33] have utilized several optimization methods. Where most common datasets that are utilized such as UCI Heart Disease dataset, Cleveland dataset and Starlog heart disease datasets. In [24], [29], [30] and [33], they have utilized 10 k fold-validation for dataset splitting and validation. Only [23] have utilized 5 fold-cross validation in their experiment. In [24] have achieved the highest accuracy of 99.65% by using the particle swarm, ant colony and FCBF optimization techniques. However, they have validated two machine learning algorithms such as K-NN and Random Forest. Where K-NN achieved the highest accuracy. On the other hand, most of the accuracy ranging from 81% to 99%. In [87]–[93], the authors have discussed about and deployed traditional machine learning and deep learning models to detect heart disease without optimizers. Although, they have also achieved a decent outcome. But with optimization solutions are more robust for detecting heart disease.

TABLE 4. A summary of most recent works for parkinson disease identification using machine learning models with optimizers.

Authors	Year	Dataset	Data Description	Data Splitting	Optimization Techniques for Diagnosis	Evaluation Metrics
[45]	2020	Parkinson Dataset	Total attributes: 23	-	Particle Swarm Optimization	Accuracy: 97.95%
[46]	2018	Voice dataset, Handwriting dataset, Speech dataset	Hand PD Dataset: Instances count: 632, Features count: 13, Missing values: N/A, Speech PD Dataset: Instances count: 195, Features count: 23, Missing values: N/A, Voice PD Dataset:, Instances count: 1040, Features count: 26, Missing values: N/A	Data Training: 70%, Data Testing: 30%	Grey Wolf Optimization	Accuracy: 94.83%, Detection rate: 98.28%, False Alarm: 16.03%.
[47]	2016	Parkinson Disease patient's medical information dataset	No. of selected features: 12	-	ELM + Hybrid Particle Swarm Optimization	Accuracy: 88.72 %, Recall: 94.33 % Precision: 90.48%, F1-score: 92.36 %
[48]	2017	Parkinson Dataset	Features = 22, Classes = 2	5-fold-Cross validation	BFO+SVM	Accuracy= 96.89, Sensitivity=98.75, Specificity=90.83
[49]	2016	Parkinson Dataset	Attributes = 9, 8, 22, 5, Classes = 2, 2, 2, 3	-	Fruit Fly Optimization	ACC: 0.9627, AUC: 0.9626, Sensitivity: 0.9624, Specificity: 0.9659
[50]	2017	Parkinson and Wisconsin diagnostic breast cancer (WDBC)	Parkinson: 195 samples,Features=22, WDBC: Attributes=32	10-fold-Cross-Validation	IGWO-KELM	Parkinson: Accuracy=97.45, Sensitivity=98.08, Specificity=96.67, WDBC: Accuracy=95.7, Sensitivity=94.88, Specificity=96.75
[51]	2010	John Radcliffe Hospital, Oxford, UK.	-	Two individual stage based data splitting	RBFNN + PCA + Particle Swarm Optimization	PCA + PSO + RBFNN - 88.92%

Table 3 provides an outlook of recent advances of machine learning algorithms with optimizers in diabetes prediction. From the table we can see [34]–[44] have utilized several optimization techniques to enhance the overall prediction accuracy. Around all recent works have utilized Pima Indians Diabetes Database (PIDD). This dataset is extensively used in research community. In [40], they have utilized multi-objective particle swarm optimization (MOPSO) for optimization purpose. They have achieved the highest accuracy of near 100% in the Pima Indians Diabetes Database (PIDD) which is higher than any other recent works. Principle component analysis ensemble with particle swarm optimization technique has been adapted by [37], where they have achieved 95.58% accuracy using C4.5 Decision Tree model. As per the literatures, the dataset splitting is done

manually. On the other hand, in [44], Enhanced and Adaptive Genetic Algorithm (EAGA) based optimization also achieved a higher accuracy of 94.7%. Most of the accuracy ranging from 76% to 99%. In [94]–[100], the authors have discussed about and deployed traditional machine learning and deep learning models to detect diabetes without optimizers. Although, they have also achieved a decent outcome. But with optimization solutions are more robust for detecting diabetes.

Table 4 demonstrates the overall scenario of machine learning based parkinson disease prediction with optimizers. From the table we can see [45]–[52] have utilized several optimization techniques with machine learning models to enhance the overall performance of the model. However, there are numerous datasets are used. Internally developed datasets are also used in their experiment. Parkinson Dataset is the most

common, utilized by [46], [49], [50]. However, in most cases particle swarm optimization based optimization techniques are used in the most literature. In parkinson dataset [46], Butter Fly Optimization based SVM has achieved accuracy of 96.89%. However, particle swarm optimization is also performed better with an accuracy of 97.95%. Which is higher than all recent works. In [101]–[109], the authors have discussed about and deployed traditional machine learning and deep learning models to detect parkinson disease without optimizers. Although, they have also achieved a decent outcome. But with optimization solutions are more robust for detecting parkinson disease.

Machine learning has become very fascinating in healthcare domain. Due a large volume of healthcare data, it is difficult for the medical doctors to identify a specific disease efficiently. However, to tackle this issue, dedicated AI researchers are applying various statistical techniques in healthcare data to classify them accordingly. Therefore, in future more robust datasets should be prepared to achieve a sophisticated result in computer-aided medical diagnosis. However, authors believe that there is a very high possibility to achieve desired results by producing optimized AI models in medical data classification area. Moreover, it will also help the physicians to make diagnosis decisions more accurately and improve the overall healthcare policies. Thus, optimization techniques in machine learning can play a very significant role to perceive better outcome rather than replacing the physicians.

Optimization techniques are mathematically standard. It needs good mathematical background to come up with a new optimization technique. However, all existing optimization cannot deliver desired output in the machine learning models. So it is important to figure out which optimization technique can be the best choice for a specific model and dataset. These are the key challenges on optimization techniques in machine learning domain. These challenges can be addressed for future remarks.

VII. LIMITATIONS ON THIS REVIEW

This paper is a comprehensive review of machine learning optimization techniques in medical data classification. Recent advances of optimization techniques used in literature are highlighted. On the other hand, a contemporary introduction of machine learning and most extensively used optimization techniques are highlighted in this review paper for better intuition to the readers.

However, there are still some limitations in this review paper. Due to less number of recent works on optimization techniques for machine learning, this paper is still lack of robust outlook for making further decisions on determining better optimization technique for the medical data classification and best feature selection. Hopefully, in future more research will work on optimization techniques so that a robust review paper with more recent works will be highlighted. But in short, this review paper is still a far-reaching outlook of

current advances. So researchers can use this as an initial reference.

VIII. CONCLUSION

This review paper involves the current aspect of machine learning model with optimization techniques to classify the medical data. We tried to give a holistic outlook of some recent works on predicting breast cancer, heart disease, diabetes and parkinson disease. The main intend of this review paper is to feature recent studies of optimization techniques involved in machine learning for a better prediction system. This paper will help the researchers and industrial engineers to analyze the better optimization technique for medical data classification. This review paper has been structured into several parts. Firstly, we started with a formal introduction of this review and diseases. Then an introductory background has been discussed in the next section with demonstrates the classification of machine learning, deep learning and optimization techniques. After that, some literature are presented of each disease. Which allows the prospect of current scenario of optimization techniques. In near future, researchers can utilize the optimization techniques to build compact solution for intelligent decision system making. In our observation, some common datasets have been used for these experiments. More elaborate and large scale datasets are needed to validate the overall performance of the optimization techniques for the machine learning models.

REFERENCES

[1] A. X. Du, S. Emam, and R. Gniadecki, "Review of machine learning in predicting dermatological outcomes," *Frontiers Med.*, vol. 7, p. 266, Jun. 2020, doi: 10.3389/fmed.2020.00266.

[2] G. Winter, "Machine learning in healthcare," *Brit. J. Healthcare Manage.*, vol. 25, no. 2, pp. 100–101, Feb. 2019, doi: 10.12968/bjhc.2019.25.2.100.

[3] M. M. Islam, M. R. Haque, H. Iqbal, M. M. Hasan, M. Hasan, and M. N. Kabir, "Breast cancer prediction: A comparative study using machine learning techniques," *Social Netw. Comput. Sci.*, vol. 1, no. 5, pp. 1–14, Sep. 2020, doi: 10.1007/s42979-020-00305-w.

[4] S. A. Balamurugan and M. Salomi, "A predictive risk level classification of diabetic patients using deep learning modified neural network," *J. Ambient Intell. Hum. Comput.*, vol. 12, no. 7, pp. 7703–7717, 2020, doi: 10.1007/s12652-020-02490-1.

[5] S. Liu, R. Zhang, X. Shang, and W. Li, "Analysis for warning factors of type 2 diabetes mellitus complications with Markov blanket based on a Bayesian network model," *Comput. Methods Programs Biomed.*, vol. 188, May 2020, Art. no. 105302, doi: 10.1016/j.cmpb.2019.105302.

[6] H. Elayan, R. M. Shubair, and A. Kiourti, "Wireless sensors for medical applications: Current status and future challenges," in *Proc. 11th Eur. Conf. Antennas Propag. (EUCAP)*, Mar. 2017, pp. 2478–2482, doi: 10.23919/EuCAP.2017.7928405.

[7] M. Y. ElSalamouny and R. M. Shubair, "Novel design of compact low-profile multi-band microstrip antennas for medical applications," in *Proc. Loughborough Antennas Propag. Conf. (LAPC)*, Nov. 2015, pp. 1–4, doi: 10.1109/LAPC.2015.7366015.

[8] D. Bazazeh and R. Shubair, "Comparative study of machine learning algorithms for breast cancer detection and diagnosis," in *Proc. 5th Int. Conf. Electron. Devices, Syst. Appl. (ICEDSA)*, Dec. 2016, pp. 1–4, doi: 10.1109/ICEDSA.2016.7818560.

[9] M. I. AlHajri, N. T. Ali, and R. M. Shubair, "Classification of indoor environments for IoT applications: A machine learning approach," *IEEE Antennas Wireless Propag. Lett.*, vol. 17, no. 12, pp. 2164–2168, Dec. 2018, doi: 10.1109/LAWP.2018.2869548.

[10] M. AlHajri, N. Ali, and R. Shubair, "Indoor localization for IoT using adaptive feature selection: A cascaded machine learning approach," *IEEE Antennas Wireless Propag. Lett.*, vol. 18, pp. 2306–2310, 2019.

[11] H. Tabrizchi, M. Tabrizchi, and H. Tabrizchi, "Breast cancer diagnosis using a multi-verse optimizer-based gradient boosting decision tree," *Social Netw. Appl. Sci.*, vol. 2, no. 4, pp. 1–19, Apr. 2020, doi: 10.1007/s42452-020-2575-9.

[12] S. Punitha, A. Amuthan, and K. S. Joseph, "Enhanced monarchy butterfly optimization technique for effective breast cancer diagnosis," *J. Med. Syst.*, vol. 43, no. 7, pp. 1–14, Jul. 2019, doi: 10.1007/s10916-019-1348-8.

[13] Q. Wuniri, W. Huangfu, Y. Liu, X. Lin, L. Liu, and Z. Yu, "A generic-driven wrapper embedded with feature-type-aware hybrid Bayesian classifier for breast cancer classification," *IEEE Access*, vol. 7, pp. 119931–119942, 2019, doi: 10.1109/access.2019.2932505.

[14] M. H. Memon, J. P. Li, A. U. Haq, M. H. Memon, and W. Zhou, "Breast cancer detection in the IoT health environment using modified recursive feature selection," *Wireless Commun. Mobile Comput.*, vol. 2019, Nov. 2019, Art. no. 5176705, doi: 10.1155/2019/5176705.

[15] S. Ronoud and S. Asadi, "An evolutionary deep belief network extreme learning-based for breast cancer diagnosis," *Soft Comput.*, vol. 23, no. 24, pp. 13139–13159, Dec. 2019, doi: 10.1007/s00500-019-03856-0.

[16] V. J. Kadam, S. M. Jadhav, and K. Vijayakumar, "Breast cancer diagnosis using feature ensemble learning based on stacked sparse autoencoders and softmax regression," *J. Med. Syst.*, vol. 43, no. 8, pp. 1–11, Aug. 2019, doi: 10.1007/s10916-019-1397-z.

[17] M. Supriya and A. J. Deepa, "A novel approach for breast cancer prediction using optimized ANN classifier based on big data environment," *Health Care Manage. Sci.*, vol. 23, no. 3, pp. 414–426, Sep. 2020, doi: 10.1007/s10729-019-09498-w.

[18] S. Wang, Y. Wang, D. Wang, Y. Yin, Y. Wang, and Y. Jin, "An improved random forest-based rule extraction method for breast cancer diagnosis," *Appl. Soft Comput.*, vol. 86, Jan. 2020, Art. no. 105941, doi: 10.1016/j.asoc.2019.105941.

[19] N. Liu, E.-S. Qi, M. Xu, B. Gao, and G.-Q. Liu, "A novel intelligent classification model for breast cancer diagnosis," *Inf. Process. Manage.*, vol. 56, no. 3, pp. 609–623, May 2019, doi: 10.1016/j.ipm.2018.10.014.

[20] I. Sangaiah and A. V. A. Kumar, "Improving medical diagnosis performance using hybrid feature selection via relieff and entropy based genetic search (RF-EGA) approach: Application to breast cancer prediction," *Cluster Comput.*, vol. 22, no. S3, pp. 6899–6906, May 2019, doi: 10.1007/s10586-018-1702-5.

[21] H. Huang, X. Feng, S. Zhou, J. Jiang, H. Chen, Y. Li, and C. Li, "A new fruit fly optimization algorithm enhanced support vector machine for diagnosis of breast cancer based on high-level features," *BMC Bioinf.*, vol. 20, no. S8, pp. 1–14, Jun. 2019, doi: 10.1186/s12859-019-2771-z.

[22] M. Abdar, M. Zomorodi-Moghadam, X. Zhou, R. Gururajan, X. Tao, P. D. Barua, and R. Gururajan, "A new nested ensemble technique for automated diagnosis of breast cancer," *Pattern Recognit. Lett.*, vol. 132, pp. 123–131, Apr. 2020, doi: 10.1016/j.patrec.2018.11.004.

[23] B. A. Tama, S. Im, and S. Lee, "Improving an intelligent detection system for coronary heart disease using a two-tier classifier ensemble," *BioMed Res. Int.*, vol. 2020, Apr. 2020, Art. no. 9816142, doi: 10.1155/2020/9816142.

[24] M. Nilashi, H. Ahmadi, A. A. Manaf, T. A. Rashid, S. Samad, L. Shahmoradi, N. Aljojo, and E. Akbari, "Coronary heart disease diagnosis through self-organizing map and fuzzy support vector machine with incremental updates," *Int. J. Fuzzy Syst.*, vol. 22, no. 4, pp. 1376–1388, Jun. 2020, doi: 10.1007/s40815-020-00828-7.

[25] P. Sharma, K. Choudhary, K. Gupta, R. Chawla, D. Gupta, and A. Sharma, "Artificial plant optimization algorithm to detect heart rate & presence of heart disease using machine learning," *Artif. Intell. Med.*, vol. 102, Jan. 2020, Art. no. 101752, doi: 10.1016/j.artmed.2019.101752.

[26] G. T. Reddy, M. P. K. Reddy, K. Lakshmanna, D. S. Rajput, R. Kaluri, and G. Srivastava, "Hybrid genetic algorithm and a fuzzy logic classifier for heart disease diagnosis," *Evol. Intell.*, vol. 13, no. 2, pp. 185–196, Jun. 2020, doi: 10.1007/s12065-019-00327-1.

[27] Y. Khourdifi and M. Bahaj, "Heart disease prediction and classification using machine learning algorithms optimized by particle swarm optimization and ant colony optimization," *Int. J. Intell. Eng. Syst.*, vol. 12, no. 1, pp. 242–252, Feb. 2019, doi: 10.22266/ijies2019.0228.24.

[28] L. Parthiban and R. Subramanian, "Intelligent heart disease prediction system using CANFIS and genetic algorithm," *Int. J. Biol. Med. Sci.*, vol. 3, no. 3, pp. 157–160, 2008.

[29] C. B. Gokulnath and S. P. Shantharajah, "An optimized feature selection based on genetic approach and support vector machine for heart disease," *Cluster Comput.*, vol. 22, no. 6, pp. 14777–14787, Nov. 2019, doi: 10.1007/s10586-018-2416-4.

[30] M. R. Nalluri and D. S. Roy, "Hybrid disease diagnosis using multiobjective optimization with evolutionary parameter optimization," *J. Healthcare Eng.*, vol. 2017, Jul. 2017, Art. no. 5907264, doi: 10.1155/2017/5907264.

[31] X. Liu, X. Wang, Q. Su, M. Zhang, Y. Zhu, Q. Wang, and Q. Wang, "A hybrid classification system for heart disease diagnosis based on the RFRS method," *Comput. Math. Methods Med.*, vol. 2017, Jan. 2017, Art. no. 8272091, doi: 10.1155/2017/8272091.

[32] S. Bashir, U. Qamar, and F. H. Khan, "BagMOOV: A novel ensemble for heart disease prediction bootstrap aggregation with multi-objective optimized voting," *Australas. Phys. Eng. Sci. Med.*, vol. 38, no. 2, pp. 305–323, Jun. 2015, doi: 10.1007/s13246-015-0337-6.

[33] N. C. Long, P. Meesad, and H. Unger, "A highly accurate firefly based algorithm for heart disease prediction," *Expert Syst. Appl.*, vol. 42, no. 21, pp. 8221–8231, 2015, doi: 10.1016/j.eswa.2015.06.024.

[34] S. Mishra, H. K. Tripathy, P. K. Mallick, A. K. Bhoi, and P. Barsocchi, "EAGA-MLP—An enhanced and adaptive hybrid classification model for diabetes diagnosiss," *Sensors*, vol. 20, no. 14, pp. 1–34, 2020, doi: 10.3390/s20144036.

[35] A. Massaro, V. Maritati, D. Giannone, D. Convertini, and A. Galiano, "LSTM DSS automatism and dataset optimization for diabetes prediction," *Appl. Sci.*, vol. 9, no. 17, pp. 10–13, 2019, doi: 10.3390/app9173532.

[36] D. K. Choubey, P. Kumar, S. Tripathi, and S. Kumar, "Performance evaluation of classification methods with PCA and PSO for diabetes," *Netw. Model. Anal. Health Informat. Bioinf.*, vol. 9, no. 1, Dec. 2020, Art. no. 5, doi: 10.1007/s13721-019-0210-8.

[37] B. S. Babu, A. Suneetha, G. C. Babu, Y. J. N. Kumar, and G. Karuna, "Medical disease prediction using grey wolf optimization and auto encoder based recurrent neural network," *Period. Eng. Nat. Sci.*, vol. 6, no. 1, pp. 229–240, 2018, doi: 10.21533/pen.v6i1.286.

[38] D. Bani-Hani, P. Patel, and T. Alshaikh, "An optimized recursive general regression neural network Oracle for the prediction and diagnosis of diabetes," *Global J. Comput. Sci. Technol.*, vol. 19, no. 2, pp. 1–11, May 2019, doi: 10.34257/gjcstdvol19is2pg1.

[39] M. Alirezaei, S. T. A. Niaki, and S. A. A. Niaki, "A bi-objective hybrid optimization algorithm to reduce noise and data dimension in diabetes diagnosis using support vector machines," *Expert Syst. Appl.*, vol. 127, pp. 47–57, Aug. 2019, doi: 10.1016/j.eswa.2019.02.037.

[40] S. Cui, D. Wang, Y. Wang, P.-W. Yu, and Y. Jin, "An improved support vector machine-based diabetic readmission prediction," *Comput. Methods Programs Biomed.*, vol. 166, pp. 123–135, Nov. 2018, doi: 10.1016/j.cmpb.2018.10.012.

[41] S. Paul and D. K. Choubey, "GA_RBF NN: A classification system for diabetes," *Int. J. Biomed. Eng. Technol.*, vol. 23, no. 1, p. 71, 2017, doi: 10.1504/ijbet.2017.10003045.

[42] M. Maniruzzaman, N. Kumar, S. Islam, H. S. Suri, A. S. El-Baz, J. S. Suri, and M. M. Abedin, "Comparative approaches for classification of diabetes mellitus data: Machine learning paradigm," *Comput. Methods Programs Biomed.*, vol. 152, pp. 23–34, Dec. 2017, doi: 10.1016/j.cmpb.2017.09.004.

[43] A. G. Karegowda, A. S. Manjunath, and M. A. Jayaram, "Application of genetic algorithm optimized neural network connection weights for medical," *Soft Comput.*, vol. 2, no. 2, pp. 15–23, 2011.

[44] M. F. Ganji and M. S. Abadeh, "A fuzzy classification system based on ant colony optimization for diabetes disease diagnosis," *Expert Syst. Appl.*, vol. 38, no. 12, pp. 14650–14659, 2011, doi: 10.1016/j.eswa.2011.05.018.

[45] D. Tomar, B. R. Prasad, and S. Agarwal, "An efficient Parkinson disease diagnosis system based on least squares twin support vector machine and particle swarm optimization," in *Proc. 9th Int. Conf. Ind. Inf. Syst. (ICIIS)*, Dec. 2014, pp. 1–6, doi: 10.1109/ICIINFS.2014.7036603.

[46] P. Sharma, S. Sundaram, M. Sharma, A. Sharma, and D. Gupta, "Diagnosis of Parkinson's disease using modified grey wolf optimization," *Cognit. Syst. Res.*, vol. 54, pp. 100–115, May 2019, doi: 10.1016/j.cogsys.2018.12.002.

[47] M. K. Shahsavari, "Using extreme learning machine and hybrid," in *Proc. 4th Int. Conf. Control. Instrum., Autom.*, Jan. 2016, pp. 27–28.

[48] Z. Cai, J. Gu, and H.-L. Chen, "A new hybrid intelligent framework for predicting Parkinson's disease," *IEEE Access*, vol. 5, pp. 17188–17200, 2017, doi: 10.1109/ACCESS.2017.2741521.

[49] L. Shen, H. Chen, Z. Yu, W. Kang, B. Zhang, and H. Li, "Evolving support vector machines using fruit fly optimization for medical data classification," *Knowl.-Based Syst.*, vol. 96, pp. 61–75, Mar. 2016, doi: 10.1016/j.knosys.2016.01.002.

[50] X. Qiang, T. Li, H. Chen, H. Huang, X. Zhao, Z. Cai, C. Tong, and W. Liu, "An enhanced grey wolf optimization based feature selection wrapped kernel extreme learning machine for medical diagnosis," *Comput. Math. Methods Med.*, vol. 2017, Jan. 2017, Art. no. 9512741.

[51] D. Wu, "Prediction of parkinson's disease tremor onset using a radial basis function neural network based on particle swarm optimization," *Int. J. Neural Syst.*, vol. 20, no. 2, pp. 109–116, 2010, doi: 10.1142/S0129065710002292.

[52] S. Das, A. Dey, A. Pal, and N. Roy, "Applications of artificial intelligence in machine learning: Review and prospect," *Int. J. Comput. Appl.*, vol. 115, no. 9, pp. 31–41, Apr. 2015.

[53] V. Nasteski, "An overview of the supervised machine learning methods," *Horizons B*, vol. 4, pp. 51–62, Dec. 2017, doi: 10.20544/HORIZONS.B.04.1.17.P05.

[54] N. Kolla and M. Giridhar Kumar, "Supervised learning algorithms of machine learning: Prediction of brand loyalty," *Int. J. Innov. Technol. Explor. Eng.*, vol. 8, no. 11, pp. 3886–3889, 2019, doi: 10.35940/ijitee.J9498.0981119.

[55] M. Usama, J. Qadir, A. Raza, H. Arif, K.-L.-A. Yau, Y. Elkhatib, A. Hussain, and A. Al-Fuqaha, "Unsupervised machine learning for networking: Techniques, applications and research challenges," *IEEE Access*, vol. 7, pp. 65579–65615, 2019, doi: 10.1109/ACCESS.2019.2916648.

[56] D. Dhall, R. Kaur, and M. Juneja, "Machine learning: A review of the algorithms and its applications," in *Proceedings of ICRIC* (Lecture Notes in Electrical Engineering), vol. 597, P. Singh, A. Kar, Y. Singh, M. Kolekar, and S. Tanwar, Eds. Cham, Switzerland: Springer, 2019, pp. 47–63, doi: 10.1007/978-3-030-29407-6_5.

[57] L. P. Kaelbling, M. L. Littman, and A. W. Moore, "Reinforcement learning: A survey," *J. Artif. Int. Res.*, vol. 4, no. 1, pp. 237–285, Jan. 1996.

[58] W. Qiang and Z. Zhongli, "Reinforcement learning model, algorithms and its application," in *Proc. Int. Conf. Mech. Sci., Electric Eng. Comput. (MEC)*, Aug. 2011, pp. 1143–1146, doi: 10.1109/MEC.2011.6025669.

[59] S. P. Sahu, A. Nautiyal, and M. Prasad, "Machine learing algorithms for recommender system—A comparative analysis," *Int. J. Comput. Appl. Technol. Res.*, vol. 6, no. 2, pp. 97–100, Feb. 2017, doi: 10.7753/ijcatr0602.1005.

[60] I. Portugal, P. Alencar, and D. Cowan, "The use of machine learning algorithms in recommender systems: A systematic review," *Expert Syst. Appl.*, vol. 97, pp. 205–227, May 2015, doi: 10.1016/j.eswa.2017.12.020.

[61] A. Thengade and R. Dondal, "Genetic algorithm–survey paper" in *Proc. MPGI Nat. Multi Conf. Int. J. Comput. Appl.*, Jan. 2012, pp. 887–975.

[62] H. Lingaraj, "A study on genetic algorithm and its applications," *Int. J. Comput. Sci. Eng.*, vol. 4, pp. 139–143, Oct. 2016.

[63] K. F. Man, K. S. Tang, and S. Kwong, "Genetic algorithms: Concepts and applications," *IEEE Trans. Ind. Electron.*, vol. 43, no. 5, pp. 519–534, 1996, doi: 10.1109/41.538609.

[64] S. Katoch, S. S. Chauhan, and V. Kumar, "A review on genetic algorithm: Past, present, and future," *Multimedia Tools Appl.*, vol. 80, no. 5, pp. 8091–8126, 2020.

[65] T. R. Schneider, "A genetic algorithm for the identification of ormationally invariant regions in protein molecules," *Acta Crystallogr. Sect. D Biol. Crystallogr.*, vol. 58, no. 2, pp. 195–208, 2002, doi: 10.1107/S0907444901019291.

[66] E. Emary, H. M. Zawbaa, and C. Grosan, "Experienced gray wolf optimization through reinforcement learning and neural networks," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 29, no. 3, pp. 681–694, Mar. 2018, doi: 10.1109/TNNLS.2016.2634548.

[67] Q. Al-Tashi, H. Md Rais, S. J. Abdulkadir, S. Mirjalili, and H. Alhussian, "A review of grey wolf optimizer-based feature selection methods for classification," in *Proc. Conf. Evol. Mach. Learn. Techn.*, Jan. 2020, pp. 273–286, doi: 10.1007/978-981-32-9990-0_13.

[68] T. A. Rashid, D. K. Abbas, and Y. K. Turel, "A multi hidden recurrent neural network with a modified grey wolf optimizer," *PLoS ONE*, vol. 14, no. 3, Mar. 2019, Art. no. e0213237.

[69] S. K. Mosavi, E. Jalalian, F. Soleimenian, and U. Branch, "A comprehensive survey of grey wolf optimizer algorithm and its application," *Int. J. Adv. Robot. Expert Syst.*, vol. 1, no. 6, pp. 23–45.

[70] L. Zajmi, F. Y. H. Ahmed, and A. A. Jaharadak, "Concepts, methods, and performances of particle swarm optimization, backpropagation, and neural networks," *Appl. Comput. Intell. Soft Comput.*, vol. 2018, pp. 1–7, Sep. 2018, doi: 10.1155/2018/9547212.

[71] S. J. Berlin and M. John, "Particle swarm optimization with deep learning for human action recognition," *Multimedia Tools Appl.*, vol. 79, nos. 25–26, pp. 17349–17371, Jul. 2020, doi: 10.1007/s11042-020-08704-0.

[72] A. R. Syulistyo, D. M. Jati Purnomo, M. F. Rachmadi, and A. Wibowo, "Particle swarm optimization (PSO) for training optimization on convolutional neural network (CNN)," *J. Comput. Inf.*, vol. 9, no. 1, p. 52, Feb. 2016, doi: 10.21609/jiki.v9i1.366.

[73] H. N. Abdull Hamed, S. M. Shamsuddin, and N. Salim, "Particle swarm optimization for neural network learning enhancement," *J. Teknol.*, vol. 49, no. 1, pp. 13–26, Dec. 2008, doi: 10.11113/jt.v49.194.

[74] F. E. Fernandes Junior and G. G. Yen, "Particle swarm optimization of deep neural networks architectures for image classification," *Swarm Evol. Comput.*, vol. 49, pp. 62–74, Sep. 2019, doi: 10.1016/j.swevo.2019.05.010.

[75] M. A. S. A. Husaini, M. H. Habaebi, S. A. Hameed, M. R. Islam, and T. S. Gunawan, "A systematic review of breast cancer detection using thermography and neural networks," *IEEE Access*, vol. 8, pp. 208922–208937, 2020, doi: 10.1109/ACCESS.2020.3038817.

[76] K. B, "Breast cancer detection using machine learning," *Int. J. Adv. Trends Comput. Sci. Eng.*, vol. 9, no. 2, pp. 981–984, Apr. 2020, doi: 10.30534/ijatcse/2020/12922020.

[77] L. Hussain, W. Aziz, S. Saeed, S. Rathore, and M. Rafique, "Automated breast cancer detection using machine learning techniques by extracting different feature extracting strategies," in *Proc. 17th IEEE Int. Conf. Trust, Secur. Privacy Comput. Commun.*, Aug. 2018, pp. 327–331, doi: 10.1109/TrustCom/BigDataSE.2018.00057.

[78] H. Dhahri, E. A. Maghayreh, A. Mahmood, W. Elkilani, and M. F. Nagi, "Automated breast cancer diagnosis based on machine learning algorithms," *J. Healthcare Eng.*, vol. 2019, Nov. 2019, Art. no. 4253641, doi: 10.1155/2019/4253641.

[79] A. F. M. Agarap, "On breast cancer detection: An application of machine learning algorithms on the Wisconsin diagnostic dataset," in *Proc. 2nd Int. Conf. Mach. Learn. Soft Comput.*, 2018, pp. 5–9, doi: 10.1145/3184066.3184080.

[80] M. R. Al-Hadidi, A. Alarabeyyat, and M. Alhanahnah, "Breast cancer detection using K-nearest neighbor machine learning algorithm," in *Proc. 9th Int. Conf. Develop. Syst. Eng. (DeSE)*, Aug. 2016, pp. 35–39, doi: 10.1109/DeSE.2016.8.

[81] M. Tahmooresi, A. Afshar, B. Bashari Rad, K. B. Nowshath, and M. A. Bamiah, "Early detection of breast cancer using machine learning techniques," *J. Telecommun. Electron. Comput. Eng.*, vol. 10, nos. 2–3, pp. 21–27, 2018.

[82] T. H. Rafi, "A comparative analysis of hyper-parameter tuned supervised machine learning algorithms on breast cancer prediction," *Ann. Eng.*, vol. 1, no. 1, p. 3, Aug. 2020.

[83] S. Mohan, C. Thirumalai, and G. Srivastava, "Effective heart disease prediction using hybrid machine learning techniques," *IEEE Access*, vol. 7, pp. 81542–81554, 2019, doi: 10.1109/ACCESS.2019.2923707.

[84] S. Pouriyeh, S. Vahid, G. Sannino, G. De Pietro, H. Arabnia, and J. Gutierrez, "A comprehensive investigation and comparison of machine learning techniques in the domain of heart disease," *Proc. IEEE Symp. Comput. Commun.*, Feb. 2017, pp. 204–207, doi: 10.1109/ISCC.2017.8024530.

[85] A. U. Haq, J. P. Li, M. H. Memon, S. Nazir, and R. Sun, "A hybrid intelligent system framework for the prediction of heart disease using machine learning algorithms," *Mobile Inf. Syst.*, vol. 2018, Dec. 2018, Art. no. 3860146, doi: 10.1155/2018/3860146.

[86] N. M. Lutimath, C. Chethan, and B. S. Pol, "Prediction of heart disease using machine learning," *Int. J. Recent Technol. Eng.*, vol. 8, no. 2, pp. 474–477, 2019, doi: 10.35940/ijrte.B1081.0982S1019.

[87] J. P. Li, A. U. Haq, S. U. Din, J. Khan, A. Khan, and A. Saboor, "Heart disease identification method using machine learning classification in E-Healthcare," *IEEE Access*, vol. 8, pp. 107562–107582, 2020, doi: 10.1109/ACCESS.2020.3001149.

[88] S. Safdar, S. Zafar, N. Zafar, and N. F. Khan, "Machine learning based decision support systems (DSS) for heart disease diagnosis: A review," *Artif. Intell. Rev.*, vol. 50, no. 4, pp. 597–623, Dec. 2018, doi: 10.1007/s10462-017-9552-8.

[89] J.-J. Beunza, E. Puertas, E. García-Ovejero, G. Villalba, E. Condes, G. Koleva, C. Hurtado, and M. F. Landecho, "Comparison of machine learning algorithms for clinical event prediction (risk of coronary heart disease)," *J. Biomed. Informat.*, vol. 97, Sep. 2019, Art. no. 103257, doi: 10.1016/j.jbi.2019.103257.

[90] D. Dutta, D. Paul, and P. Ghosh, "Analysing feature importances for diabetes prediction using machine learning," in *Proc. IEEE 9th Annu. Inf. Technol., Electron. Mobile Commun. Conf. (IEMCON)*, Nov. 2018, pp. 924–928, doi: 10.1109/IEMCON.2018.8614871.

[91] I. Kavakiotis, O. Tsave, and A. Salifoglou, "Machine learning and data mining methods in diabetes research," *Comput. Struct. Biotechnol. J.*, vol. 15, no. 9, pp. 104–116, 2017, doi: 10.1016/j.csbj.2016.12.005.

[92] A. Yahyaoui, A. Jamil, J. Rasheed, and M. Yesiltepe, "A decision support system for diabetes prediction using machine learning and deep learning techniques," in *Proc. 1st Int. Informat. Softw. Eng. Conf. (UBMYK)*, Nov. 2019, pp. 1–4, doi: 10.1109/UBMYK48245.2019.8965556.

[93] M. K. Hasan, M. A. Alam, D. Das, E. Hossain, and M. Hasan, "Diabetes prediction using ensembling of different machine learning classifiers," *IEEE Access*, vol. 8, pp. 76516–76531, 2020, doi: 10.1109/ACCESS.2020.2989857.

[94] M. Maniruzzaman, M. J. Rahman, B. Ahammed, and M. M. Abedin, "Classification and prediction of diabetes disease using machine learning paradigm," *Health Inf. Sci. Syst.*, vol. 8, no. 1, pp. 1–14, Dec. 2020, doi: 10.1007/s13755-019-0095-z.

[95] L. Kopitar, P. Kocbek, L. Cilar, A. Sheikh, and G. Stiglic, "Early detection of type 2 diabetes mellitus using machine learning-based prediction models," *Sci. Rep.*, vol. 10, no. 1, Dec. 2020, Art. no. 11981, doi: 10.1038/s41598-020-68771-z.

[96] S. Perveen, M. Shahbaz, K. Keshavjee, and A. Guergachi, "Prognostic modeling and prevention of diabetes using machine learning technique," *Sci. Rep.*, vol. 9, no. 1, 2019, Art. no. 13805, doi: 10.1038/s41598-019-49563-6.

[97] A. Ul Haq, J. Li, M. H. Memon, J. Khan, S. U. Din, I. Ahad, R. Sun, and Z. Lai, "Comparative analysis of the classification performance of machine learning classifiers and deep neural network classifier for prediction of Parkinson disease," in *Proc. 15th Int. Comput. Conf. Wavelet Act. Media Technol. Inf. Process. (ICCWAMTIP)*, Dec. 2018, pp. 101–106, doi: 10.1109/ICCWAMTIP.2018.8632613.

[98] C. Gao, H. Sun, T. Wang, M. Tang, N. I. Bohnen, M. L. T. M. Müller, and T. Herman, "Model-based and model-free machine learning techniques for diagnostic prediction and classification of clinical outcomes in Parkinson's disease," *Sci. Rep.*, vol. 8, no. 1, 2018, Art. no. 7129, doi: 10.1038/s41598-018-24783-4.

[99] R. Z. U. Rehman, S. Del Din, Y. Guan, A. J. Yarnall, J. Q. Shi, and L. Rochester, "Selecting clinically relevant gait characteristics for classification of early Parkinson's disease: A comprehensive machine learning approach," *Sci. Rep.*, vol. 9, no. 1, 2019, Art. no. 17269, doi: 10.1038/s41598-019-53656-7.

[100] R. Alkhatib, M. O. DIab, C. Corbier, and M. El Badaoui, "Machine learning algorithm for gait analysis and classification on early detection of Parkinson," *IEEE Sensors Lett.*, vol. 4, no. 6, Jun. 2020, Art. no. 6000604, doi: 10.1109/LSENS.2020.2994938.

[101] S. Shetty and Y. S. Rao, "SVM based machine learning approach to identify Parkinson's disease using gait analysis," in *Proc. Int. Conf. Inventive Comput. Technol. (ICICT)*, Coimbatore, India, Aug. 2016, pp. 1–5, doi: 10.1109/INVENTIVE.2016.7824836.

[102] Z. Karapinar Senturk, "Early diagnosis of Parkinson's disease using machine learning algorithms," *Med. Hypotheses*, vol. 138, Jan. 2020, Art. no. 109603, doi: 10.1016/j.mehy.2020.109603.

[103] S. A. Mostafa, A. Mustapha, M. A. Mohammed, R. I. Hamed, and N. Arunkumar, "Examining multiple feature evaluation and classification methods for improving the diagnosis of Parkinson's disease," *Cogn. Syst. Res.*, vol. 54, pp. 90–99, May 2019, doi: 10.1016/j.cogsys.2018.12.004.

[104] D. Bazazeh, R. M. Shubair, and W. Q. Malik, "Biomarker discovery and validation for Parkinson's disease: A machine learning approach," in *Proc. Int. Conf. Bio-Eng. Smart Technol. (BioSMART)*, Dec. 2016, pp. 1–6.

TAKI HASAN RAFI received the bachelor's degree in electrical engineering from Ahsanullah University of Science and Technology, Dhaka, Bangladesh. He will be starting the Ph.D. degree in computer science in this year. He has published around ten research articles in various peer-reviewed journals and IEEE flagship conferences. Prior to his research involvement, he spends couple of years in a start-up and an internet service provider company in Bangladesh. His research interests include deep learning in medical applications, generative adversarial networks, spiking neural networks, and human–computer interaction. He also works as a reviewer in various SCI journals and IEEE conferences.

RAED M. SHUBAIR (Senior Member, IEEE) received the B.Sc. degree (Hons.) in electrical engineering from Kuwait University, Kuwait, in June 1989, and the Ph.D. degree (Hons.) in electrical engineering from the University of Waterloo, Canada, in February 1993, for which he received the University of Waterloo Distinguished Doctorate Dissertation Award.

He is currently a Full Professor of electrical engineering affiliated with New York University (NYU), Abu Dhabi. His current and past academic and research appointments also include Massachusetts Institute of Technology (MIT), Harvard University, and the University of Waterloo. He was a Full Professor of electrical engineering with Khalifa University (formerly, Etisalat University College), United Arab Emirates, from 1993 to 2017, during which he received several times the Excellence in Teaching Award and Distinguished Service Award. He has over 380 publications in the form of articles in peer-reviewed journals, papers in refered conference proceedings, book chapters, and U.S. patents. His publication span several research areas, including 6G and terahertz communications, modern antennas and applied electromagnetics, signal and array processing, machine learning, the IoT and sensor localization, medical sensing, and nano-biomedicine. He is a fellow of MIT Electromagnetics Academy and a Founding Member of MIT Scholars of the Emirates. He is a standing member of the editorial boards of several international journals and serves regularly for the steering, organizing, and technical committees of IEEE flagship conferences in Antennas, Communications, and Signal Processing, including several editions of IEEE AP-S/URSI, EuCAP, IEEE GloablSIP, IEEE WCNC, and IEEE ICASSP. He is also a Board Member of European School of Antennas and the Regional Director of the IEEE Signal Processing Society in IEEE Region 8 Middle East. He is a Founding Member of five IEEE society chapters in United Arab Emirates, which are the IEEE Communication Society Chapter, the IEEE Signal Processing Society Chapter, the IEEE Antennas and Propagation Society Chapter, the IEEE Microwave Theory and Techniques Society Chapter, and the IEEE Engineering in Medicine and Biology Society Chapter. He was a recipient of several international awards, including the Distinguished Service Award from ACES Society, USA, and MIT Electromagnetics Academy, USA. He organized and chaired numerous technical special sessions and tutorials in IEEE flagship conferences. He delivered more than 60 invited speaker seminars and technical talks in world-class universities and flagship conferences. He has served as the TPC Chair for IEEE MMS2016 and IEEE GlobalSIP 2018 Symposium on 5G satellite networks. He served as the Founding Chair for the IEEE Antennas and Propagation Society Educational Initiatives Program. He is the Founder and the Chair of IEEE at New York University at Abu Dhabi. He is an Officer of IEEE ComSoc Emerging Technical Initiative (ETI) on Machine Learning for Communications. He is the Founding Director of the IEEE UAE Distinguished Seminar Series Program for which he was selected to receive, along with Mohamed AlHajri of MIT, the 2020 IEEE UAE Award of the Year. He served as an Invited Speaker for U.S. National Academies of Sciences, Engineering, and Medicine Frontiers Symposium. He holds several leading roles in the international professional engineering community. He is also an Editor of IEEE JOURNAL OF ELECTROMAGNETICS, RF AND MICROWAVES IN MEDICINE AND BIOLOGY, and IEEE OPEN JOURNAL OF ANTENNAS AND PROPAGATION.

FAISAL FARHAN received the bachelor's degree in electrical and electronic engineering from Ahsanullah University of Science and Technology, Dhaka, Bangladesh, in January 2021. He is currently pursuing the Ph.D. degree in machine learning and artificial intelligence. He is also a full-time Lecturer with the Department of Electrical and Electronic Engineering, Ahsanullah University of Science and Technology. He has worked on several academic and independent research projects. His volunteering affiliations include IEEE AUST Student Branch, IEEE AUST WIE Section, and AUST EEE SOCIETY. His research interests include signal processing, speech processing, brain–computer interface, machine learning, and applied artificial intelligence. He was a recipient of the 2021 Dean's List of Honour Award at Ahsanullah University of Science Technology for his academic excellence.

FARHAN MOHD QUAYYUM received the bachelor's degree in electrical and electronic engineering from Ahsanullah University of Science and Technology, Dhaka, Bangladesh, where he is currently pursuing the M.Sc. degree in electrical engineering. His research interests include machine learning, deep learning, and speech recognition.

● ● ●

MD. ZIAUL HOQUE received the bachelor's degree in electrical and electronics engineering from Ahsanullah University of Science and Technology, Dhaka, Bangladesh, where he is currently pursuing the M.Sc. degree in electrical engineering with fully-funded scholarships. His research interests include machine learning, signal processing, and pattern recognition. Thus these have been always the core of his research.