International Journal of E-Health and Medical Communications
Volume 12 • Issue 2 • March-April 2021

Artificial Bee Colony and Deep Neural Network-Based Diagnostic Model for Improving the Prediction Accuracy of Diabetes

Anand Kumar Srivastava, ABES Engineering College, Ghaziabad, India & Jaypee University of Information Technology, Waknaghat, India

Yugal Kumar, Jaypee University of Information Technology, Waknaghat, India

https://orcid.org/0000-0003-3451-4897

Pradeep Kumar Singh, Jaypee University of Information Technology, Waknaghat, India

https://orcid.org/0000-0002-7676-9014

ABSTRACT

A large number of machine learning approaches are implemented in healthcare field for effective diagnosis and prediction of different diseases. The aim of these machine learning approaches is to build automated diagnostic tool for helping the physician as well as monitor the health status of patients. These diagnostic tools are widely adopted in intensive care unit for life expectancy of patients. In this study, an effort is made to design an automated diagnostic model for the diagnosis and prediction of diabetes patients. The proposed diagnostic model is designed using artificial bee colony (ABC) algorithm and deep neural network (DNN) technique, called ABC-DNN-based diagnostic model. The ABC algorithm is applied to determine the relevant features for diabetes prediction and diagnosis while DNN technique is adopted for the prediction and diagnosis of diabetes affected patients. The performance of proposed diagnostic model is tested over Pima Indian Diabetes dataset and evaluated using accuracy, sensitivity, specificity, F-measure, Kappa, and area under curve (AUC) parameters. Further, 10-fold and 50-50% training-testing method are considered to assess the performance of proposed diagnostic model. The experimental results of proposed ABC-DNN model is compared with DNN technique and several existing diabetes studies. It is observed that proposed ABC-DNN model achieves 94.74% accuracy rate using 10-fold method.

KEYWORDS

Artificial Bee Colony, Deep Neural Network, Diagnosis, Diagnostic Model, Feature Selection

DOI: 10.4018/IJEHMC.2021030102

1. INTRODUCTION

Due to advancement in health care system, life expectancy of human is increased tremendously in last two decades. But, several challenges are associated with healthcare systems. These challenges are lack of medical information, diagnostic errors, inadequate data, irrelevant features, data threaten etc. In literature, several expert systems, decision diagnostic system, electronic health record system, prediction system are presented (Doi, 2007; Meystre et al., 2008; Rangayyan et al., 2007). These systems considerably help the physicians for disease diagnosis. The term disease diagnosis refers to determine the disease through symptoms and it can be formulated as classification of medical data for decision making. Such decision making systems requires higher computing power to process the large medical dataset. The decision capabilities of these systems depend on amount of training data. The aim of these systems is to minimize the physician errors. Such systems can enhance the decision capabilities of physician as well as improve the diagnostic accuracy. It is noticed that many computer-aided diagnostic tools are presented in literature to help the physicians. Large number of machine learning techniques is incorporated in diagnostic tools to improve the prediction rate (Kavakiotis et al., 2017; Nahar et al., 2013; Nilashi et al., 2017; Shickel et al., 2017). It is seen that all features are not equally important for decision making process. The irrelevant features can also affect the prediction accuracy of the algorithm. Hence, several researchers also focus on the identification of relevant features for disease diagnosis and prediction. Several features selection algorithms are also presented in literature (Akay, 2009; Inbarani et al., 2014; Lin & Hsieh, 2015; Subanya & Rajalaxmi, 2014). Features selection can be defined as identification of most relevant set of features from the medical datasets. The aim of feature selection techniques is to reduce computational cost and also improve the accuracy rate of diagnostic process. These techniques are also integrated with machine learning methods and can be either supervised or unsupervised. The aim of these feature selection techniques is to compute a weight function for each feature of the medical dataset. The weight function is used to determine the relevant features form datasets. Some of feature selection techniques to improve predictive accuracy are listed as. A particle swarm optimization (PSO) and ABC based feature selection algorithm to determine the optimal set of features is reported in (Lin & Hsieh, 2015). To determine the optimal set of features for support vector machine (SVM), Subanya et al. (2014) applied ABC based feature reduction algorithm on heart dataset. Akay applied F-score technique to determine optimal set of features for breast cancer dataset (Akay, 2009). Inbarani et al. (2014) adopted rough set based approach to remove redundant features from medical datasets. Further, this approach is integrated with PSO for improving the classification accuracy. Hence, feature selection is an important issue for building an efficient and effective diagnostic model. Further, it is also limit the number of input features in the diagnostic model in order to produce good predictive results (Akay, 2009). Hence, in this work, an attempt is made to select the optimum features for deep neural network model to improve the diagnostic accuracy. To address the same, ABC based feature selection method is designed to determine the optimum set of features. Further, this feature selection method is integrated with DNN technique. Finally, a diagnostic model is developed using ABC based feature selection method and DNN technique, called ABC-DNN diagnostic model. The performance of proposed ABC-DNN diagnostic model is evaluated using Pima Indian Diabetes disease dataset. It is noticed that proposed ABC-DNN diagnostic model provides better results than DNN technique. The main contributions of this study are summarized as follows:

1. To develop an efficient and effective diagnostic model for improving the diagnostic accuracy of diabetes disease.
2. ABC based feature selection method is designed to select the relevant feature for the diagnosis of diabetes disease.
3. DNN technique is adopted for the prediction of diabetes disease using reduced feature set.

4. This paper explore the impact of DNN technique on the accuracy issue of diabetes disease using feature selection and not using the feature selection technique. This paper also investigates the impact of feature selection technique on a diagnostic model either the performance of diagnostic model is improved or not improved. Hence, this paper also described the role of feature selection technique in the field of healthcare.

Rest of paper is organized as follows. Section 2 discusses works reported on disease diagnosis and prediction using machine learning approaches. Section 3 presents the proposed ABC-DNN based Diagnostic system for diabetes disease. Experimental results of proposed ABC-DNN system are presented in section 4. Section 5 concludes the entire work.

2. RELATED WORKS

This section describes the recent works reported on disease diagnosis and prediction using machine learning approaches.

Lin et al. (2019) developed a clinical diagnostic score system for the prediction of coronary artery spasm in patients. In this system, multi variable analysis is performed to identify the patients with coronary disease. The performance of the proposed system is evaluate using nine hundred seventy-six patients. It is noticed that proposed diagnostic score system achieves ninety-six percent accuracy rate.

An ensemble learning framework is designed to diagnosis the thromboembolism (Sabra et al., 2020). The proposed framework consists of venous thromboembolism (VTE) ontology, semantic extraction, sentiment assessment and an ensemble classifier. The ensemble classifier is the combination of the multi-layer perceptron neural network (MLPNN) and SVM. The performance of proposed framework is investigated on clinical data of two hundred fifty patients. It is seen that proposed ensemble framework obtains higher F-measure rate as compared to other classifiers.

Ahmed and Acharjya adopted cuckoo search and rough set theory to design a hybrid approach for the accurate prediction of heart disease (Acharjya, 2020). The cuckoo search technique is used to determine the main features of heart disease. Whereas, rough theory is applied to develop the inference rule for heart disease prediction. The performance of proposed hybrid approach is tested on six hundred three patients. It is observed that proposed approach develops fifty-three rule for the prediction of heart disease and provides state of art prediction results.

Abdar et al. (2019) presented a hybrid machine learning technique i.e. N2Genetic optimizer for accurate prediction of coronary artery disease. It is seen that N2Genetic optimiser is used as training technique for SVM technique, called N2Genetic-nuSVM. The well-known Z-Alizadeh Sani dataset is considered to evaluate the performance of the aforesaid technique. Authors claimed that proposed N2Genetic-nuSVM technique achieves 93.08% accuracy rate.

Narayan and Sathiyamoorthy developed a recommender system for prediction and identifying the heart disease (Narayan & Sathiyamoorthy, 2019). The proposed recommender system is the combination of the Fourier transformation and machine learning technique. In this work, three machine learning technique is used for prediction of heart disease. These techniques are artificial neural network (ANN), naïve bayes (NB) and SVM. The performance of the proposed recommender system is evaluated using real time dataset. The experimental results showed that proposed recommender system is more capable to predict and identify the heart disease as compared to existing ones.

Bucholc et al. (2019) designed a computerized decision support system for predicting the accuracy of the Alzheimer's disease. The proposed decision support system consists of six machine learning techniques. These machine learning techniques are Kernel Ridge Regression (KRR), Support Vector Regression, and k-Nearest Neighbor for regression and SVM, Random Forest (RF), and k-Nearest Neighbor (k-NN). It is observed that KKR and SVM provide better classification rate.

An ensemble neuro fuzzy technique is employed for diagnosis of hepatitis disease (Nilashi, Ahmadi, Shahmoradi et al, 2019). The proposed ensemble method works in three steps. In first step, Non-linear Iterative Partial Least Squares method is applied for dimension reduction. In second step, self-organize map method is used for clustering task. In third step, neurro fuzzy inference system is used for predicting the hepatitis disease. Authors claimed that proposed ensemble neuro fuzzy technique provides superior results than Neural Network, adaptive neural fuzzy inference system (ANFIS), k-NN and SVM techniques.

Zhao et al. (2019) developed an integrated model based on the logistic regression (LR) and SVM for the detection of colorectal cancer. In the proposed model, LR method is applied to capture the more relevant features of colorectal cancer. Whereas, SVM with different kernel functions such as linear, sigmoid, radial basis and polynomial are adopted for the detection of colorectal cancer. It is observed that SVM with radial basis function obtains higher accuracy rate as compared to rest of these.

Nilashi et al. (2019) designed a medical decision support system for monitoring and diagnosis of real time Parkinson's disease patients. The proposed medical decision support system is the combination of dimension reduction technique and ensemble learning method. In the proposed system, singular value decomposition technique is implemented as dimension reduction technique. Whereas, adaptive neuro-fuzzy inference system is adopted for the diagnosis of Parkinson's disease patients. The experimental results showed that proposed system is effectively used to diagnosis and monitoring of Parkinson's disease patients.

Ali et al (2019) developed an automated diagnostic system based on the statistical method and deep neural network for prediction of heart disease. The statistical method is used to eliminate the irrelevant features of heart disease dataset whereas, deep neural network is configured for the prediction of heart disease. The performance of automated diagnostic system is compared with ANN and DNN models. It is stated that proposed diagnostic model obtains superior results than ANN and DNN models.

An intelligent risk prediction system based on fuzzy temporal rules is presented for detection of breast cancer (Kanimozhi et al., 2019). The proposed prediction system employs the fuzzy rule based classification for detection of breast cancer. Some temporal constraints are also imposed in this system. It is observed that proposed risk prediction system computes more accurate risk for breast cancer.

Gokulnath and Shantharajah presented a hybrid technique based on genetic algorithm (GA) and SVM for the prediction of heart disease, called GA-SVM (Gokulnath & Shantharajah, 2019). In this technique, genetic algorithm is used to select the more relevant features of heart disease. Whereas, SVM is implemented for prediction of heart disease. Cleveland heart disease dataset is used to evaluate the performance of the GA-SVM technique. It is observed that proposed technique provides more accurate results as compared to rest of techniques.

Kirk et al. (2017) developed a decision making system called DEAR (Detect, Evaluate, Assess and Recommend action). The basic aim of this proposed system is to detect the change in the environment, to provide the risk assessment and gives real-time recommendations to control the outbreak of mosquito diseases.

Devarajan et al. (2019) implemented a healthcare system for the treatment of the Parkinson's disease. The proposed model analyzed the voice samples of the patients to recommend the proper treatment. Fog computing is used as middle layer between the cloud server and end user in the proposed system. Further the Fuzzy K-nearest Neighbor classifier (FKNC) and Case-based Reasoning classifier (CBRC) is utilized to classify the non- Parkinson patients and Parkinson patients. The proposed system also generated an immediate alert in the case of abnormality. The proposed system is tested on Parkinson dataset using Accuracy, F-measure, Sensitivity, Specificity and Precision parameter. The experimental results showed that the performance of proposed system is better than existing healthcare systems.

Peker presented a decision support system for improving the accuracy of medical diagnosis (Peker, 2016). In this work, a new attribute selection method based of k- medoids clustering is applied for attribute weighting. In classification phase, SVM method is applied for diagnosis of medical data. The effectiveness of proposed DSS is tested on three medical datasets i.e. Patkinson's Disease, Heart

Disease and Liver Disorder. It is stated that combination of k-medoids and SVM approach outperforms than reported methods in literature.

Huda et al. (2016) developed a hybrid ensemble method for imbalance medical data especially for brain tumor diagnosis. The hybrid ensemble method is the combination of feature selection technique and ensemble classification approach. ANN with input gain measurement approximation is applied for selecting better attributes to improve the accuracy of ensemble classification method. In classification stage, decision tree based bootstrap method is applied to classify the imbalance data. Authors claimed that proposed hybrid ensemble method achieves higher accuracy rate in comparison to other popular machine learning algorithms.

Alickovic and Subasi designed a hybrid classification algorithm based on genetic algorithm and rotation forest for breast cancer diagnosis (Aličković & Subasi, 2017). This algorithm works in two stages. In first stage, genetic algorithm is applied to determine relevant set of features. In second stage, rotation forest algorithm is adopted for diagnosis of breast cancer. Two well-known breast cancer datasets are considered to evaluate the effectiveness of proposed algorithm. It is noticed that proposed algorithm obtains 99% accuracy rate.

A rule base classification model is presented to predict the Parkinson's disease (Kumar & Sahoo, 2013). Several rules are designed for accurate prediction of Parkinson's disease. Authors claimed that the proposed rile based classification model achieves more than ninety-eight percent accuracy rate as compared to NB, SVM and LR techniques.

A personal health record based system is developed for monitoring of diabetes disease using mobile environment (Kumar et al., 2019). The objective of the proposed system is to improve the life style of patients and also overcome the probability of chronic diabetes disease. In this work, a graphical user interface (GUI) based application is also developed to track the status of diabetic patients.

For early detection of the dengue disease, Gambhir et al. (2018) evaluated the performance of three well-known machine learning classifiers. These approaches are ANN, decision tree (DT) and NB techniques. It is noticed that ANN techniques obtains better results than DT and NB in terms of accuracy, sensitivity and specificity. In continuation of their work, PSO-ANN based diagnostic model is also developed for early detection of dengue disease (Gambhir et al., 2017). In the proposed diagnostic model, PSO technique is applied to optimize the weight function of ANN method.

A home health monitoring system is designed for early detection of diabetes disease (Chatrati et al., 2020). The proposed monitoring system is combination of the decision making technique and machine learning approaches. Authors claimed that proposed monitoring system is capable to segregate the diabetes patients effectively and also send real time alert messages to diabetes patients.

To predict the diabetes mellitus, devi et al. (2020) applied farthest first clustering algorithm and sequential minimal optimization classifier. In this work, farthest first clustering algorithm is applied to group the data in one cluster. While, sequential minimal optimization is adopted to diagnose the diabetes mellitus. It is reported that proposed integrated approach provides better results as compared to existing approaches.

A comparative study of different machine learning techniques adopted for heart disease prediction is reported in (Srivastava et al., 2018). Authors discuss pros and cons of each machine learning technique in brief. Srivastava et al. (2020) designed a rule based monitoring system for accurate prediction of diabetes disease. The proposed system consists of various rules for the identification of diabetes disease. These rules are derived using clinical and non-clinical symptoms of diabetes disease and a panel of doctors certifies these rules. The proposed monitoring system achieves 90.4% accuracy rate.

2.1 Technical Gap

Through the extensive literature survey, it is noticed that large number of machine learning techniques have been reported for disease diagnosis and prediction. It also noted that ANN, SVM, NB and DT are widely adopted for diagnosis task. But, it is stated that optimization of ANN and SVM is not an easy task. Various optimization techniques are studied to optimize the weight and bias functions of ANN and SVM techniques. The optimization of weight and bias functions improves the diagnostic rate. Moreover, it is also noticed that performance of classifier also depends on the features of a dataset. Various studies showed that all features of a dataset are not equally important. The selection of relevant features has significant impact on the performance of classifiers. Irrelevant features can degrade the performance of classifiers. Hence in this work, ABC based feature selection method is applied to determine optimal set of features. Further, the concept of deep neural network (DNN) is implemented diagnostic task. DNN is significantly different from ANN and gives better diagnostic rate as compared ANN.

3. PROPOSED ABC-DNN BASED DIAGNOSTIC MODEL

This section describes the ABC-DNN based diagnostic model for diagnosis and prediction of diabetes disease. The proposed diagnostic model is the combination of the ABC based feature selection method and DNN technique. The subsection 3.1 describes the details of ABC based feature selection algorithm which is used to compute the more relevant features of diabetes disease. Subsection 3.2 describes the DNN technique which is used as classifier in the proposed diagnostic system. Figure 1 illustrates the block diagram of proposed ABC-DNN based diagnostic model.

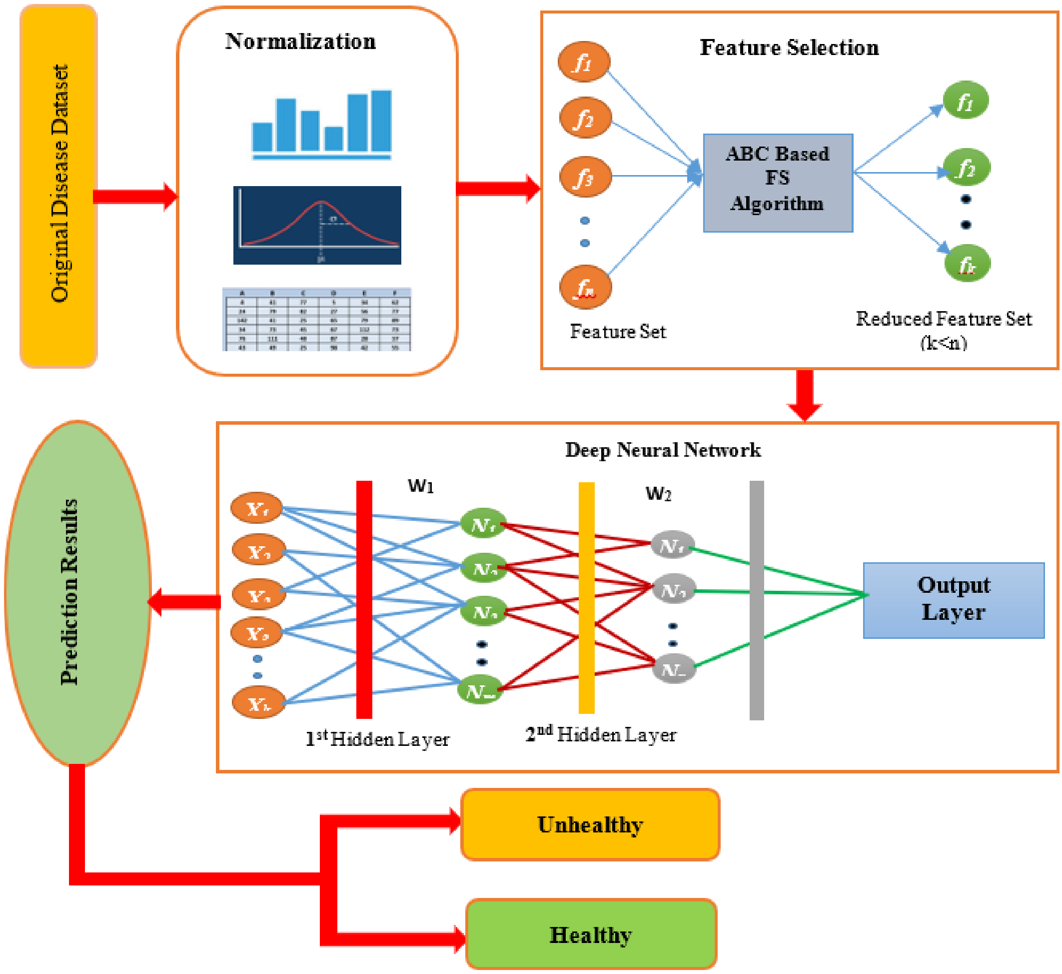
3.1 ABC Based Feature Selection (FS) Algorithm

This subsection describes a feature selection algorithm based on artificial bee colony (ABC). The aim of the feature selection algorithm is to determine the more relevant features of diabetes disease. The feature selection algorithm can be either supervised or unsupervised in nature. This work considers the unsupervised ABC feature section algorithm for determine the relevant features. ABC is a stochastic optimization method inspired through honey bees which is used to optimize numerical function problems (Karaboga, 2005). This algorithm is widely adopted for solving variety of optimization problem such as clustering, function optimization, constrained and unconstrained optimization problems (Cao et al., 2019; Dedeturk & Akay, 2020; Li et al., 2019; Sahoo, 2017). ABC algorithm consists of three types of bees such as employed bees, onlooker bees and scout bee. In ABC, employed bees is equal to number of onlooker bees and there is a single scout bee in bee colony. The works of these bees are described as employed bees are responsible to determine the location of food sources, and further, collects the information regarding food sources and handover this information to onlooker bees. The main concern of the onlooker bees is to evaluate the quality of food sources discovered through employed bees. These bees are also explored the new food source location, if previously discovered food source quality is not good. Scout bee generates new food source location, if previously generated food source is abandoned. In ABC algorithm, employed and onlooker bees perform the exploitation process whereas the exploration is performed by the scout bee. 3The proposed ABC based feature selection algorithm consists of four phases- initialization phase, employed bees phase, onlooker bees phase, abandoned food source or scout bee phase. The detailed steps of ABC feature selection algorithm are given in Algorithm 1. The flowchart of the ABC based feature selection algorithm is presented in Figure 2.

3.2 Deep Neural Network (DNN)

DNN can be characterized as a neural network with finite level of complexity. It consists of complicate mathematical model for processing of the information and contains multiple layer (Ali et al., 2019). DNN integrates both of feature selection and classification task itself to build good decision making. DNN obtains good attention from the research community in recent time. DNN can be described as an input layer (I) for input, hidden layer (L) for processing and an output layer (O) for final outcome. DNN is developed in python programming language using tf.contrib.learn.DNNClassifier library from Google. DNN having advantage over conventional neural network because it is designed using

Figure 1. Block Diagram of the Proposed ABC-DNN based Diagnostic Model



the wide set of trials and every trial modifies number of hidden layers, activation function, learning steps and number of neurons of DNN configuration. The accuracy of DNN is validated using the test set in every manual configuration. In this work, rectified linear unit activation function is used to generate all neuron layers for DNN classifier. The DNN classifiers is described through eq. 7 and further, this equation is also defined the activation function:

$$F\left(x\right) = x' = \max\left(0, x\right) \tag{7}$$

In abovementioned equation, x denotes an input to a neuron, also called ramp function. Moreover, the activation function is associated with smooth approximation which is given in eq. 8:

$$F\left(x\right) = \ln\left[1 + \exp\left(x\right)\right] \tag{8}$$

Algorithm 1. ABC based Feature Selection Algorithm

/* Initialization Phase */	
Step 1:	Initialize user defined parameters of the proposed algorithm such food sources, limit, maximum cycle number, colony size and upper bounce. For i= 1: K * K is the total number of food sources */ Initialize the food source within the boundary of given dataset in random order. Send the employed bees to the food source. End For Iteration = 0
Step 2:	While (Stopping condition is not met)
/* Start Employed Bees Phase*/	
Step 3:	For i= 1: K /* i represents the ith employed bee and K represents total number of food sources*/ Discover the position of new food source in the neighbouring of old food source using eq. 1. $V_{i,j} = X_{i,j} + \varnothing_{i,j} \left(X_{i,j} - X_{k,j} \right)$ (1) Compute the fitness of the new generated food source using eq. 2. $fit_i = \frac{1}{1 + f_i}$ (2) A greedy selection is made between the old and new food source positions and keep the best one End For
Step 4:	Calculate the probability values for each food source. For i=1: K Compute the probability value associated with each food source using eq. 3. $P_i = \frac{fit_i}{\sum_{i=1}^{SN} fit_i}$ (3) End For
/* Start the Onlooker Bees Phase*/	

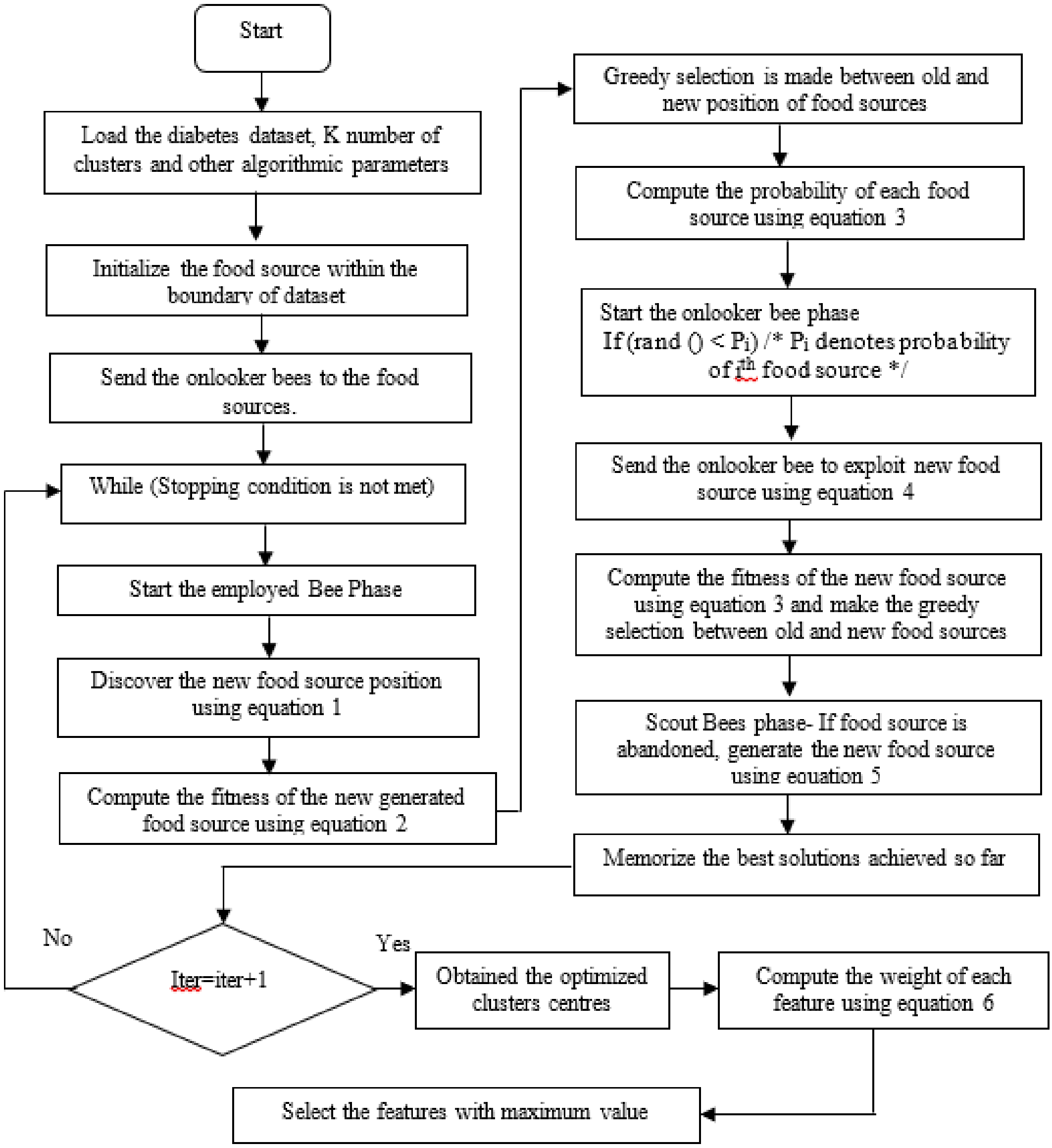
Step 5:	For i = 1: K /* i represents the ith onlooker bee and K represents total number of food sources */
	If (rand () < P_i) /* P_i denotes probability of ith food source */
	Send the onlooker bee to exploit new food source in the neighbouring of old position using eq. 4.
	$V_{i,j} = X_{i,j} + \varnothing_{i,j} \left(X_{i,j} - X_{k,j} \right) + \varphi_{i,j} \left(Y_j - X_{k,j} \right)$ (4)
	Compute the fitness of the new food source using equation 3.
	A greedy selection is made between the old and new food source and keep the best one.
	Else
	i=i+1
	End If
	End For

/* Scout Bee Phase*/	

Step 6:	If (the food source is not upgraded further using limit)
	Send a scout bee to generate the new position of food source using eq. 5.
	$X_{new} = X_{best} + rand\left[0,1\right]\left(X_{best} - X_{curr} \right)$ (5)
	End If

Step 7:	Memorize the best solution achieved so far. Iteration = iteration + 1 End While Obtain the final optimized cluster centres.
Step 8:	Compute the weight of each feature using equation 6. $f_i = \left(\sum_{i=1}^{d} \sum_{j=1}^{h} \sum_{k=1}^{k} \frac{X_{ih}}{C_{ik}} \right) \times \frac{1}{d}$ (6)
Step 9:	Select the features with maximum value.

Figure 2. Flow chart of the ABC based feature selection algorithm



In prediction task, the new representation of input layer is illustrated through hidden layers and can be defined using eq. 9:

$$x_{t+1} = m\left[w_t x_t + b_t\right] \tag{9}$$

x_{t+1} denotes the (t+1)th hidden layer, w_t denotes the weight of ith hidden layer, b_t denotes the bias of ith hidden layer and m denotes activation function.

4. EXPERIMENTAL RESULTS AND DISCUSSION

This section demonstrates the simulation results of proposed ABC-DNN based diagnostic model. The performance of proposed model is evaluated using Pima indian diabetes dataset. Furthermore, 10 cross fold and 50-50 percent training and testing method is used to validate the performance of diagnostic model. The simulation results of proposed system are described in terms of average of the ten independent run. The parameters tuning of the DNN method are taken same as reported in (Ali et al., 2019). A total seven hundred sixty-eight patient's data are collected for the accurate diagnosis of diabetes disease. The prediction accuracy of the proposed model is described through accuracy, sensitivity, specificity, kappa, AUC and f-measure parameters. These parameters are computed using equations 10-13:

$$Accuracy = \frac{TP + TN}{TP + FP + TN + FN} \times 100 \tag{10}$$

$$Sensitivity = \frac{TP}{TP + FN} \times 100 \tag{11}$$

$$Specificity = \frac{TN}{FP + TN} \times 100 \tag{12}$$

$$F - Measure = \frac{2 \times Precision \times Recall}{Precision + Recall} \tag{13}$$

Table 1 illustrates the simulation results of proposed ABC-DNN diagnostic model and DNN technique. The simulation results are evaluated using 10-cross fold and 50-50% training and testing techniques. It is observed that proposed diagnostic model obtains 94.74% accuracy rate using 10-fold method. Where, the accuracy rate using all features i.e. DNN technique is 82.67%. Hence, it is stated that proposed ABC based feature selection method improves the accuracy rate of DNN technique. On the analysis of 50-50% training-testing technique, it is also seen that proposed ABC-DNN diagnostic model achieves higher accuracy rate as compared to DNN technique. It is observed that ABC based feature selection method improves the accuracy rate of DNN upto twelve percent. It is also stated that proposed diagnostic model achieves higher sensitivity, specificity, f-measure, kappa and AUC rate as compared to DNN technique. Furthermore, it is noted that kappa statistics is significantly improved using ABC-DNN technique. It is also revealed that there is significant difference between the performances of 10 fold and 50-50% techniques in terms of ABC-DNN model.

The experimental results of proposed ABC-DNN based diagnostic model is also compared with existing diabetes prediction model and several machine learning approaches. Table 2 presents the experimental results of proposed diagnostic model and other existing diabetes models using accuracy parameter. It is revealed that ABC-DNN model obtains higher accuracy rate as compared to rest of existing studies. Hence, it is stated that proposed ABC-DNN based diagnostic model is one of the efficient and effective method for computer-aided diagnostic system. This model also improves the

Table 1. Simulation results of DNN and ABC-DNN on diabetes dataset

Features	Parameters	10 Cross Fold	50-50% Training and Test
Using All Features (DNN)	ACC	82.67 ± 4.53	81.06 ± 4.87
	Sensitivity	88.93± 5.14	87.41 ± 4.74
	Specificity	78.58 ± 6.13	74.59 ± 5.81
	f-measure	78.12	75.67
	Kappa	71.23	68.43
	AUC	81.42	79.08
Using Attribute Weighting (ABC-DNN)	ACC	94.74 ± 2.56	91.61 ± 3.46
	Sensitivity	95.52 ± 4.18	93.16 ± 3.94
	Specificity	92.06 ± 2.35	90.04 ± 4.23
	f-measure	89.31	85.94
	Kappa	91.28	87.56
	AUC	92.36	88.27

accuracy rate of diabetes disease. Further, it is noticed that proposed diagnostic model obtains higher accuracy arte using 10-fold technique rather than 50-50% training-testing technique.

Figure 3(a) illustrates the data instances of diabetes dataset using age, pedigree function and BMI features. While, figure 3(b) denotes the prediction of diabetes dataset in two classes i.e. Diabetes and Non-Diabetes. The prediction of diabetes and non-diabetes patients is done through the proposed ABC-DNN based diagnostic model. Figures 4-5 demonstrate the experimental results of the ABC-DNN based diagnostic model and DNN technique using 10-fold method. Figure 4 shows the experimental results of ABC-DNN model and DNN technique using accuracy, sensitivity and specificity parameters. Whereas, Figure 5 depicts the simulation results using kappa, AUC and F-measure parameters. It is seen that proposed ABC-DNN based model achieves better quality results than DNN technique. Figures 6-7 show the experimental result of ABC-DNN model and DNN technique using 50-50% training-testing method. It is noticed that integration of ABC based feature selection method improves the performance of DNN significantly. Hence, it is stated that ABC-DNN based diagnostic model outperforms than DNN technique.

5. CONCLUSION

This paper presents an ABC-DNN based diagnostic model for the diagnosis and prediction of diabetes disease. The proposed diagnostic model is designed using ABC based feature selection method and deep neural network technique. In proposed model, ABC based feature selection method is used to determine the relevant features of diabetes disease. Further, the DNN technique is adopted for diagnosis and prediction of diabetes disease. The performance of proposed diagnostic model is evaluated using Pima Indian Diabetes dataset. The different performance measures like accuracy, sensitivity, specificity, AUC, F-measure and Kappa are considered to assess the performance of ABC-DNN based diagnostic model. Furthermore, experimental results are evaluated using 10-fold and 50-50% training-testing techniques. It is observed that ABC-DNN based diagnostic model provides better results than DNN method. The experimental results of ABC-DNN model is also compared with thirty-one existing diabetes studies. It also revealed that proposed diagnostic model achieves higher accuracy rate than existing studies. Furthermore, it is noticed that 10-fold method is more suitable than 50-50% training-testing method. Hence, it is concluded that proposed ABC-DNN

Table 2. Performance comparison of proposed KhmAW-SVM method and previous studies reported in literature on PIMA indian diabetes dataset

Sr. No.	Study	Algorithm	Accuracy (%)
1	Deng and Kasabov (2003)	ESOM	78.4
2	Polat et al. (2008)	LS-SVM,	78.21
3	Temurtas et al. (2009)	MLNN with LM	79.62
4	Kayaer and Yıldırım (2003)	GRNN	80.21
5	Carpenter and Markuzon (1998)	ARTMAP-IC	81
6	Temurtas et al. (2009)	MLNN with LM	82.37
7	Dogantekin et al. (2010)	LDA-ANFIS	84.61
8	Bozkurt et al. (2014)	DTDN	76
9	Yilmaz et al. (2014)	Modified K-Means Clustering	93.71
10	Zhu et al. (2015)	Multiple Factors Weighted Combination	91
11	Dogantekin et al. (2010)	Various methods	59.5 and 77.7
12	Ramezani et al. (2018)	LANFIS	88.05
13	Orkcu and Bal (2011)	Real-coded Genetic Algorithm	77.6
14	Luukka (2007)	Similarity Classifier + Feature Extraction	75.97
15	Isa et al. (2011)	Clustered-Hybrid MLP	80.59
16	Ozcift and Gulten (2013)	Rotation Forest Ensemble Classifier	74.47
17	Aslam et al. (2014)	Genetic Programming+K-Nearest Neighbour	80.5
18	Seera and Lim (2014)	Fuzzy Max-Min NN-CART Random Forest	78.39
19	Belle et al. (2015)	Radial Basis Function Classifier	76.7
20	Wang et al. (2015)	Improved Electro magnetism like Mechanism	77.21
21	Seera et al. (2015)	Hybrid Fuzzy ARTMAP-CART model	87.64
22	Zhu et al. (2015)	Multiple Factors Weighted Combination	93
23	Ding et al. (2015)	Extreme Learning Machine	77.63
24	Mohapatra et al. (2015)	Improved Cuckoo Search based ELM	78.5
25	Feng et al. (2015)	Variable Coded Hierarchical Fuzzy Classification	79.17
26	Luukka (2011)	Similarity Classifier using PCA and Entropy	75.82
27	Polat and Gunes (2007a)	Fuzzy-Artificial immune recognition system	84.42
28	Polat and Gunes (2007b)	PCA + ANFIS	89.47
29	Ghazavi and Liao (2008)	Fuzzy Modelling with Selected Features	77.65
30	Polat et al. (2008)	Generalized discriminant analysis Least square SVM	82.05
31	Kahramanli and Allahverdi (2006)	ANN + FNN	84.24
32	**Our Study**	**ABC-DNN (50-50 training and test set)**	**91.61**
33	**Our Study**	**ABC-DNN (10 cross fold)**	**94.74**

Figure 3. (a) Illustrate the distribution of Pima Indian Diabetes Disease data, while (b) Shows diabetes prediction using ABC-DNN method

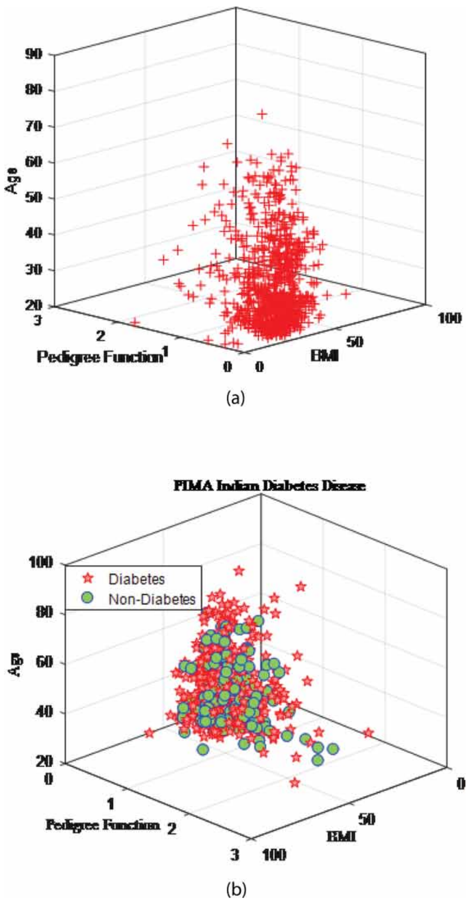


Figure 4. Simulation results of DNN and proposed ABC-DNN based diagnostic model using accuracy, sensitivity and specificity parameters (10-fold method)

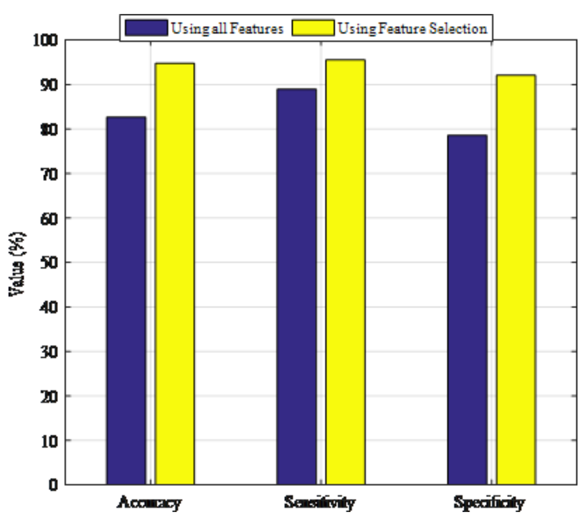


Figure 5. Simulation results of DNN and proposed ABC-DNN based diagnostic model using F-measure, Kappa and AUC (10-fold technique)

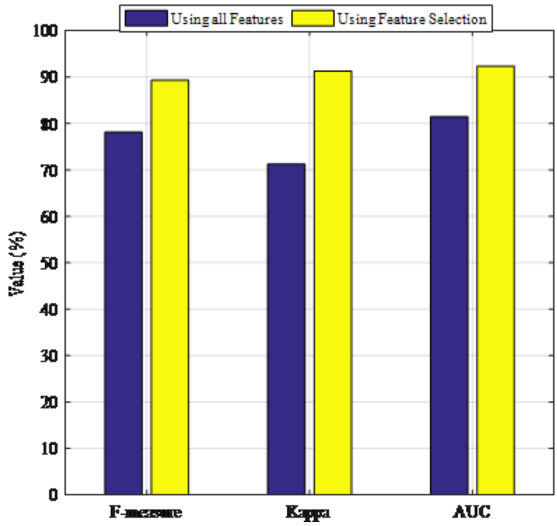
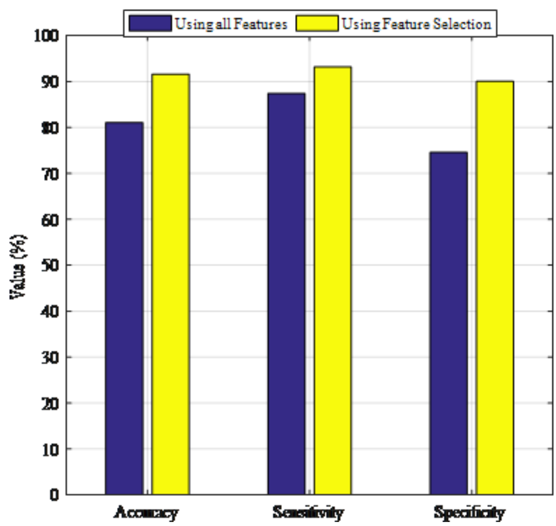
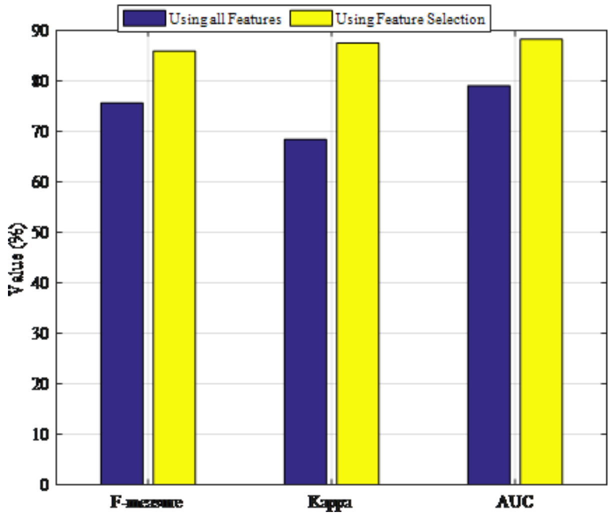


Figure 6. Simulation results of DNN and proposed ABC-DNN based diagnostic model using accuracy, sensitivity and specificity parameters (50-50% training-testing method)



based diagnostic model is an effective and efficient diagnostic model for diagnosis and prediction of diabetes disease. In future, DNN method will be applied for the accurate diagnosis of different diseases like Parkinsons, liver disease, heart disease etc. Further, the weight function of DNN method will optimize using meta-heuristic techniques. The different feature selection technique will be adopted for the selection of relevant features.

Figure 7. Simulation results of DNN and proposed ABC-DNN based diagnostic model using F-measure, Kappa and AUC (50-50% training-testing method)



REFERENCES

Abdar, M., Książek, W., Acharya, U. R., Tan, R. S., Makarenkov, V., & Pławiak, P. (2019). A new machine learning technique for an accurate diagnosis of coronary artery disease. *Computer Methods and Programs in Biomedicine*, *179*, 104992. doi:10.1016/j.cmpb.2019.104992 PMID:31443858

Acharjya, D. P. (2020). A Hybrid Scheme for Heart Disease Diagnosis Using Rough Set and Cuckoo Search Technique. *Journal of Medical Systems*, *44*(1), 27. doi:10.1007/s10916-019-1497-9 PMID:31828437

Akay, M. F. (2009). Support vector machines combined with feature selection for breast cancer diagnosis. *Expert Systems with Applications*, *36*(2), 3240–3247. doi:10.1016/j.eswa.2008.01.009

Ali, L., Rahman, A., Khan, A., Zhou, M., Javeed, A., & Khan, J. A. (2019). An Automated Diagnostic System for Heart Disease Prediction Based on ${\chi^{2}} $ Statistical Model and Optimally Configured Deep Neural Network. *IEEE Access: Practical Innovations, Open Solutions*, *7*, 34938–34945. doi:10.1109/ACCESS.2019.2904800

Aličković, E., & Subasi, A. (2017). Breast cancer diagnosis using GA feature selection and Rotation Forest. *Neural Computing & Applications*, *28*(4), 753–763. doi:10.1007/s00521-015-2103-9

Aslam, M. W., Zhu, Z., & Nandi, A. K. (2013). Feature generation using genetic programming with comparative partner selection for diabetes classification. *Expert Systems with Applications*, *40*(13), 5402–5412. doi:10.1016/j.eswa.2013.04.003

Bozkurt, M. R., Yurtay, N., Yilmaz, Z., & Sertkaya, C. (2014). Comparison of different methods for determining diabetes. *Turkish Journal of Electrical Engineering and Computer Sciences*, *22*(4), 1044–1055. doi:10.3906/elk-1209-82

Bucholc, M., Ding, X., Wang, H., Glass, D. H., Wang, H., Prasad, G., & Finn, D. P. et al. (2019). A practical computerized decision support system for predicting the severity of Alzheimer's disease of an individual. *Expert Systems with Applications*, *130*, 157–171. doi:10.1016/j.eswa.2019.04.022 PMID:31402810

Cao, Y., Lu, Y., Pan, X., & Sun, N. (2019). An improved global best guided artificial bee colony algorithm for continuous optimization problems. *Cluster Computing*, *22*(2), 3011–3019. doi:10.1007/s10586-018-1817-8

Carpenter, G. A., & Markuzon, N. (1998). ARTMAP-IC and medical diagnosis: Instance counting and inconsistent cases. *Neural Networks*, *11*(2), 323–336. doi:10.1016/S0893-6080(97)00067-1 PMID:12662841

Chatrati, S. P., Hossain, G., Goyal, A., Bhan, A., Bhattacharya, S., Gaurav, D., & Tiwari, S. M. (2020). Smart home health monitoring system for predicting type 2 diabetes and hypertension. *Journal of King Saud University-Computer and Information Sciences*.

Dedeturk, B. K., & Akay, B. (2020). Spam filtering using a logistic regression model trained by an artificial bee colony algorithm. *Applied Soft Computing*, *91*, 106229. doi:10.1016/j.asoc.2020.106229

Deng, D., & Kasabov, N. (2003). On-line pattern analysis by evolving self-organizing maps. *Neurocomputing*, *51*, 87–103. doi:10.1016/S0925-2312(02)00599-4

Devarajan, M., & Ravi, L. (2019). Intelligent cyber-physical system for an efficient detection of Parkinson disease using fog computing. *Multimedia Tools and Applications, 78*(23), 1-25.

Devi, R. D. H., Bai, A., & Nagarajan, N. (2020). A novel hybrid approach for diagnosing diabetes mellitus using farthest first and support vector machine algorithms. *Obesity Medicine*, *17*, 100152. doi:10.1016/j.obmed.2019.100152

Ding, S., Zhao, H., Zhang, Y., Xu, X., & Nie, R. (2015). Extreme learning machine: Algorithm, theory and applications. *Artificial Intelligence Review*, *44*(1), 103–115. doi:10.1007/s10462-013-9405-z

Dogantekin, E., Dogantekin, A., Avci, D., & Avci, L. (2010). An intelligent diagnosis system for diabetes on linear discriminant analysis and adaptive network based fuzzy inference system: LDA-ANFIS. *Digital Signal Processing*, *20*(4), 1248–1255. doi:10.1016/j.dsp.2009.10.021

Doi, K. (2007). Computer-aided diagnosis in medical imaging: Historical review, current status and future potential. *Computerized Medical Imaging and Graphics*, *31*(4-5), 198–211. doi:10.1016/j.compmedimag.2007.02.002 PMID:17349778

Feng, T. C., Li, T. H. S., & Kuo, P. H. (2015). Variable coded hierarchical fuzzy classification model using DNA coding and evolutionary programming. *Applied Mathematical Modelling*, *39*(23-24), 7401–7419. doi:10.1016/j.apm.2015.03.004

Gambhir, S., Malik, S. K., & Kumar, Y. (2017). PSO-ANN based diagnostic model for the early detection of dengue disease. *New Horizons in Translational Medicine*, *4*(1-4), 1–8. doi:10.1016/j.nhtm.2017.10.001

Gambhir, S., Malik, S. K., & Kumar, Y. (2018). The Diagnosis of Dengue Disease: An Evaluation of Three Machine Learning Approaches. *International Journal of Healthcare Information Systems and Informatics*, *13*(3), 1–19. doi:10.4018/IJHISI.2018070101

Ghazavi, S. N., & Liao, T. W. (2008). Medical data mining by fuzzy modeling with selected features. *Artificial Intelligence in Medicine*, *43*(3), 195–206. doi:10.1016/j.artmed.2008.04.004 PMID:18534831

Gokulnath, C. B., & Shantharajah, S. P. (2019). An optimized feature selection based on genetic approach and support vector machine for heart disease. *Cluster Computing*, *22*(6), 14777–14787. doi:10.1007/s10586-018-2416-4

Huda, S., Yearwood, J., Jelinek, H. F., Hassan, M. M., Fortino, G., & Buckland, M. (2016). A hybrid feature selection with ensemble classification for imbalanced healthcare data: A case study for brain tumor diagnosis. *IEEE Access: Practical Innovations, Open Solutions*, *4*, 9145–9154. doi:10.1109/ACCESS.2016.2647238

Inbarani, H. H., Azar, A. T., & Jothi, G. (2014). Supervised hybrid feature selection based on PSO and rough sets for medical diagnosis. *Computer Methods and Programs in Biomedicine*, *113*(1), 175–185. doi:10.1016/j.cmpb.2013.10.007 PMID:24210167

Isa, N. A. M., & Mamat, W. M. F. W. (2011). Clustered-hybrid multilayer perceptron network for pattern recognition application. *Applied Soft Computing*, *11*(1), 1457–1466. doi:10.1016/j.asoc.2010.04.017

Kanimozhi, U., Ganapathy, S., Manjula, D., & Kannan, A. (2019). An intelligent risk prediction system for breast cancer using fuzzy temporal rules. *National Academy Science Letters*, *42*(3), 227–232. doi:10.1007/s40009-018-0732-0

Karaboga, D. (2005). An idea based on honey bee swarm for numerical optimization (Vol. 200). Technical report-tr06, Erciyes University, Engineering Faculty, Computer Engineering Department.

Kavakiotis, I., Tsave, O., Salifoglou, A., Maglaveras, N., Vlahavas, I., & Chouvarda, I. (2017). Machine learning and data mining methods in diabetes research. *Computational and Structural Biotechnology Journal*, *15*, 104–116. doi:10.1016/j.csbj.2016.12.005 PMID:28138367

Kayaer, K., & Yıldırım, T. (2003, June). Medical diagnosis on Pima Indian diabetes using general regression neural networks. In *Proceedings of the international conference on artificial neural networks and neural information processing (ICANN/ICONIP)* (*Vol. 181*, p. 184). Academic Press.

Kirk, R. A., & Kirk, D. A. (2017, July). Introducing a Decision Making Framework to Help Users Detect, Evaluate, Assess, and Recommend (DEAR). In *Action Within Complex Sociotechnical Environments. In International Conference on Human Interface and the Management of Information* (pp. 223-239). Springer.

Kumar, Y., & Sahoo, G. (2013). Prediction of different types of liver diseases using rule based classification model. *Technology and Health Care*, *21*(5), 417–432. doi:10.3233/THC-130742 PMID:23963359

Kumar, Y., Yadav, G., Singh, P. K., & Arora, P. (2019). A PHR-Based System for Monitoring Diabetes in Mobile Environment. In *Mobile Solutions and Their Usefulness in Everyday Life* (pp. 129–144). Springer. doi:10.1007/978-3-319-93491-4_7

Li, J. Q., Song, M. X., Wang, L., Duan, P. Y., Han, Y. Y., Sang, H. Y., & Pan, Q. K. (2019). Hybrid Artificial Bee Colony Algorithm for a Parallel Batching Distributed Flow-Shop Problem With Deteriorating Jobs. *IEEE Transactions on Cybernetics*. Advance online publication. doi:10.1109/TCYB.2019.2943606 PMID:31603832

Lin, K. C., & Hsieh, Y. H. (2015). Classification of medical datasets using SVMs with hybrid evolutionary algorithms based on endocrine-based particle swarm optimization and artificial bee colony algorithms. *Journal of Medical Systems*, *39*(10), 119. doi:10.1007/s10916-015-0306-3 PMID:26289628

Lin, Y., Qin, H., Chen, R., Liu, Q., Liu, H., & Dong, S. (2019). A comprehensive clinical diagnostic score system for prediction of coronary artery spasm in patients with acute chest pain. *IJC Heart & Vasculature*, *22*, 205–209. doi:10.1016/j.ijcha.2019.02.001 PMID:30963096

Luukka, P. (2007). Similarity classifier using similarity measure derived from Yu's norms in classification of medical data sets. *Computers in Biology and Medicine*, *37*(8), 1133–1140. doi:10.1016/j.compbiomed.2006.10.005 PMID:17129582

Luukka, P. (2011). Feature selection using fuzzy entropy measures with similarity classifier. *Expert Systems with Applications*, *38*(4), 4600–4607. doi:10.1016/j.eswa.2010.09.133

Meystre, S. M., Savova, G. K., Kipper-Schuler, K. C., & Hurdle, J. F. (2008). Extracting information from textual documents in the electronic health record: A review of recent research. *Yearbook of Medical Informatics*, *17*(01), 128–144. doi:10.1055/s-0038-1638592 PMID:18660887

Mohapatra, P., Chakravarty, S., & Dash, P. K. (2015). An improved cuckoo search based extreme learning machine for medical data classification. *Swarm and Evolutionary Computation*, *24*, 25–49. doi:10.1016/j.swevo.2015.05.003

Nahar, J., Imam, T., Tickle, K. S., & Chen, Y. P. P. (2013). Computational intelligence for heart disease diagnosis: A medical knowledge driven approach. *Expert Systems with Applications*, *40*(1), 96–104. doi:10.1016/j.eswa.2012.07.032

Narayan, S., & Sathiyamoorthy, E. (2019). A novel recommender system based on FFT with machine learning for predicting and identifying heart diseases. *Neural Computing & Applications*, *31*(1), 93–102. doi:10.1007/s00521-018-3662-3

Nilashi, M., Ahmadi, H., Shahmoradi, L., Ibrahim, O., & Akbari, E. (2019). A predictive method for hepatitis disease diagnosis using ensembles of neuro-fuzzy technique. *Journal of Infection and Public Health*, *12*(1), 13–20. doi:10.1016/j.jiph.2018.09.009 PMID:30293875

Nilashi, M., Ibrahim, O., Ahmadi, H., & Shahmoradi, L. (2017). An analytical method for diseases prediction using machine learning techniques. *Computers & Chemical Engineering*, *106*, 212–223. doi:10.1016/j.compchemeng.2017.06.011

Nilashi, M., Ibrahim, O., Samad, S., Ahmadi, H., Shahmoradi, L., & Akbari, E. (2019). An analytical method for measuring the Parkinson's disease progression: A case on a Parkinson's telemonitoring dataset. *Measurement*, *136*, 545–557. doi:10.1016/j.measurement.2019.01.014

Orkcu, H. H., & Bal, H. (2011). Comparing performances of backpropagation and genetic algorithms in the data classification. *Expert Systems with Applications*, *38*(4), 3703–3709. doi:10.1016/j.eswa.2010.09.028

Peker, M. (2016). A decision support system to improve medical diagnosis using a combination of k-medoids clustering based attribute weighting and SVM. *Journal of Medical Systems*, *40*(5), 116. doi:10.1007/s10916-016-0477-6 PMID:27000777

Polat, K., & Güneş, S. (2007a). An improved approach to medical data sets classification: Artificial immune recognition system with fuzzy resource allocation mechanism. *Expert Systems: International Journal of Knowledge Engineering and Neural Networks*, *24*(4), 252–270. doi:10.1111/j.1468-0394.2007.00432.x

Polat, K., & Güneş, S. (2007b). An expert system approach based on principal component analysis and adaptive neuro-fuzzy inference system to diagnosis of diabetes disease. *Digital Signal Processing*, *17*(4), 702–710. doi:10.1016/j.dsp.2006.09.005

Polat, K., Güneş, S., & Arslan, A. (2008). A cascade learning system for classification of diabetes disease: Generalized discriminant analysis and least square support vector machine. *Expert Systems with Applications*, *34*(1), 482–487. doi:10.1016/j.eswa.2006.09.012

Ramezani, R., Maadi, M., & Khatami, S. M. (2018). A novel hybrid intelligent system with missing value imputation for diabetes diagnosis. *Alexandria Engineering Journal*, *57*(3), 1883–1891. doi:10.1016/j.aej.2017.03.043

Rangayyan, R. M., Ayres, F. J., & Desautels, J. L. (2007). A review of computer-aided diagnosis of breast cancer: Toward the detection of subtle signs. *Journal of the Franklin Institute*, *344*(3-4), 312–348. doi:10.1016/j.jfranklin.2006.09.003

Sabra, S., Malik, K. M., Afzal, M., Sabeeh, V., & Charaf Eddine, A. (2020). A hybrid knowledge and ensemble classification approach for prediction of venous thromboembolism. *Expert Systems: International Journal of Knowledge Engineering and Neural Networks*, *37*(1), e12388. doi:10.1111/exsy.12388

Sahoo, G. (2017). A two-step artificial bee colony algorithm for clustering. *Neural Computing & Applications*, *28*(3), 537–551. doi:10.1007/s00521-015-2095-5

Seera, M., & Lim, C. P. (2014). A hybrid intelligent system for medical data classification. *Expert Systems with Applications*, *41*(5), 2239–2249. doi:10.1016/j.eswa.2013.09.022

Seera, M., Lim, C. P., Tan, S. C., & Loo, C. K. (2015). A hybrid FAM–CART model and its application to medical data classification. *Neural Computing & Applications*, *26*(8), 1799–1811. doi:10.1007/s00521-015-1852-9

Shickel, B., Tighe, P. J., Bihorac, A., & Rashidi, P. (2017). Deep EHR: A survey of recent advances in deep learning techniques for electronic health record (EHR) analysis. *IEEE Journal of Biomedical and Health Informatics*, *22*(5), 1589–1604. doi:10.1109/JBHI.2017.2767063 PMID:29989977

Srivastava, A. K., Kumar, Y., & Singh, P. K. (2020). A Rule-Based Monitoring System for Accurate Prediction of Diabetes: Monitoring System for Diabetes. *International Journal of E-Health and Medical Communications*, *11*(3), 32–53. doi:10.4018/IJEHMC.2020070103

Srivastava, A. K., Singh, P. K., & Kumar, Y. (2018, November). A Taxonomy on Machine Learning Based Techniques to Identify the Heart Disease. In *International Conference on Next Generation Computing Technologies* (pp. 13-25). Springer.

Subanya, B., & Rajalaxmi, R. R. (2014, February). Feature selection using Artificial Bee Colony for cardiovascular disease classification. In *2014 International Conference on Electronics and Communication Systems (ICECS)* (pp. 1-6). IEEE. doi:10.1109/ECS.2014.6892729

Temurtas, H., Yumusak, N., & Temurtas, F. (2009). A comparative study on diabetes disease diagnosis using neural networks. *Expert Systems with Applications*, *36*(4), 8610–8615. doi:10.1016/j.eswa.2008.10.032

Van Belle, V., & Lisboa, P. (2014). White box radial basis function classifiers with component selection for clinical prediction models. *Artificial Intelligence in Medicine*, *60*(1), 53–64. doi:10.1016/j.artmed.2013.10.001 PMID:24262978

Wang, K. J., Adrian, A. M., Chen, K. H., & Wang, K. M. (2015). An improved electromagnetism-like mechanism algorithm and its application to the prediction of diabetes mellitus. *Journal of Biomedical Informatics*, *54*, 220–229. doi:10.1016/j.jbi.2015.02.001 PMID:25677947

Yan, G., Ma, G., Lv, J., & Song, B. (2006, January). Combining independent component analysis with support vector machines. In *2006 1st International Symposium on Systems and Control in Aerospace and Astronautics* (pp. 4-pp). IEEE.

Yilmaz, N., Inan, O., & Uzer, M. S. (2014). A new data preparation method based on clustering algorithms for diagnosis systems of heart and diabetes diseases. *Journal of Medical Systems*, *38*(5), 48–59. doi:10.1007/s10916-014-0048-7 PMID:24737307

Zhao, D., Liu, H., Zheng, Y., He, Y., Lu, D., & Lyu, C. (2019). A reliable method for colorectal cancer prediction based on feature selection and support vector machine. *Medical & Biological Engineering & Computing*, *57*(4), 901–912. doi:10.1007/s11517-018-1930-0 PMID:30478811

Zhu, J., Xie, Q., & Zheng, K. (2015). An improved early detection method of type-2 diabetes mellitus using multiple classifier system. *Information Sciences*, *292*, 1–14. doi:10.1016/j.ins.2014.08.056