

RESEARCH

Open Access



Improve data classification performance in diagnosing diabetes using the Binary Exchange Market Algorithm

Faranak Rezaei, Maryam Abbasitabar*, Shirin Mirzaei, Zahra Kamari Direh, Sahar Ahmadi, Zahra Azizi and Darya Danialy

*Correspondence:
Maryam.
Abbasitabar2020@yahoo.com
Department of Electrical
Engineering and Medical
Equipment, Imam Reza
Center of Applied Science
and Technology (IRCAST),
Eslamabad-e Gharb,
Kermanshah, Iran

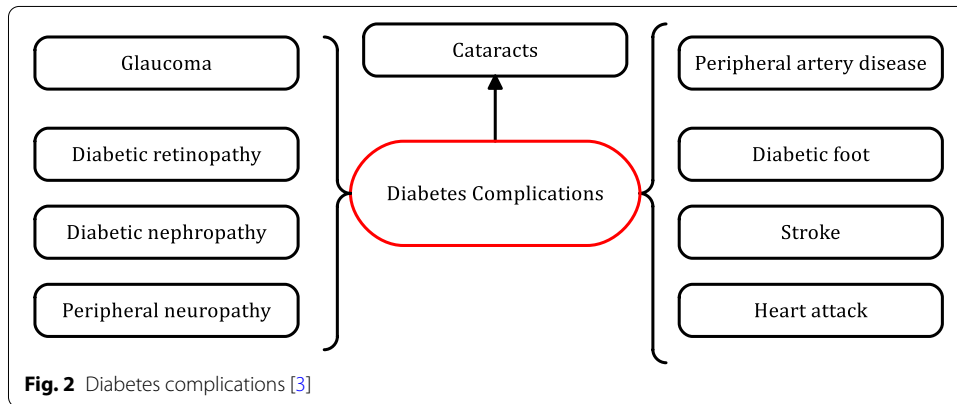
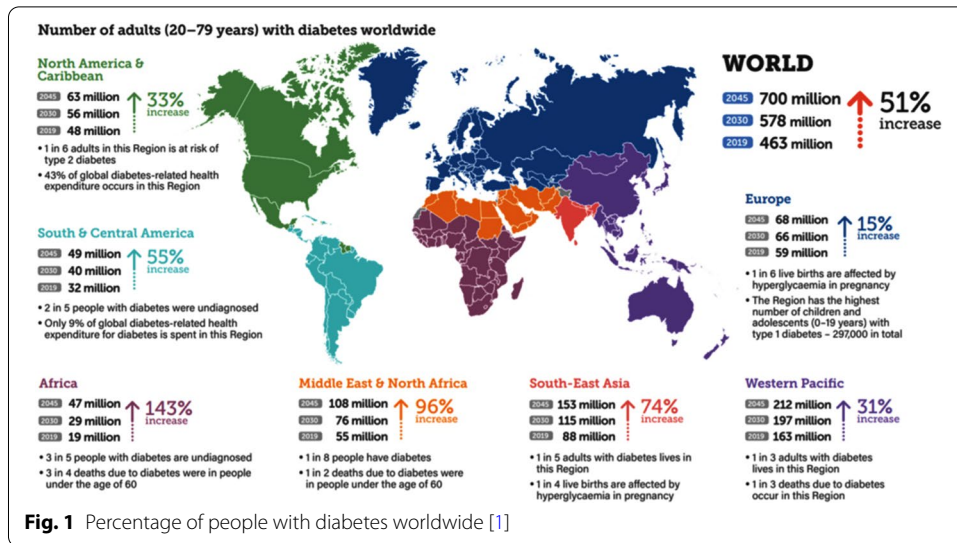
Abstract

Today's lifestyle has led to a significant increase in referrals to medical centers to diagnose various diseases. To this end, over the past few years, researchers have turned to new diagnostic methods, including data mining and artificial intelligence, intending to facilitate the detection process and increase reliability. The high volume of data available in medical centers can be considered one of the main problems in using these methods. The optimal selection of essential and influential features reduces the maximum dimension for better diagnosis with more reliability of results. In this paper, a new approach uses a Binary Exchange Market Algorithm (BEMA) to identify essential and practical features in the diabetes dataset and determine the best algorithm binary function (type of sigmoid function) to improve the performance of the EMA algorithm is presented. For validation and efficiency of the proposed BEMA algorithm, several SVM, KNN, and NB classification models have been used to train and test the final model. The results obtained from the evaluations show that the proposed BEMA-SVM combined method has a better performance than the previous methods to improve accuracy and offer an effect equivalent to 98.502%. Also, to provide better results and more reliability than the proposed method, researchers can use a combination of several classes with the proposed method, which is outside the scope of this study.

Keywords: Diabetes, Data Mining, Binary Exchange Market Algorithm (BEMA), Support Vector Machine (SVM), Feature Selection (FS)

Introduction

Diabetes can be considered a metabolic disease in which a person's body cannot produce enough insulin or respond appropriately to the insulin produced. Disorders such as diabetes, hypertension, and hyperlipidemia are called diseases of the new civilization. Urbanization, the restriction of human activities, and the increase in the consumption of high-calorie substances have led to new diseases. Diabetes is vital in terms of the number of patients and its type. Available statistics and forecasts show that the number of people with diabetes worldwide is between 20 and 79 years old and will increase



from 463 million in 2019 to 700 million in 2045. The percentage of diabetes worldwide is shown in Fig. 1 [1].

According to the above description, diabetes can be divided into two categories [2]:

- Type 1: This type includes 10% of people, who are insulin-dependent.
- Type 2: This type affects 90% of people with diabetes. This type of body produces some insulin for the body’s activities, but this amount is not enough for all the body’s needs. In general, type 2 diabetes is the most common model and is usually asymptomatic or has mild symptoms. This can cause a person with diabetes to be unaware of their condition for years and irreparable damage to various parts of the patient’s body. A summary of these injuries can be seen in Fig. 2 [3].

Due to the rapid spread of this disease, accuracy in diagnosing diabetes is of particular importance. There are many different ways to diagnose diabetes, but they impose many costs on patients. In this regard, researchers have proposed methods that are inexpensive and have early detection. One of these computer simulations is Data Mining (DM) techniques. DM is formed with the concept of extracting information

or patterns and determining specific relationships in a large amount of data in one or more databases. This method can be used effectively in predicting and diagnosing patients quickly and cheaply [4].

These new technologies highlight challenges for both patients and health care providers, as they must be able to analyze diabetes data to make it usable. There are several critical factors in the accurate prediction and analysis of diabetes, even in a computer simulation. Combining these factors and their properties increases the chances of precise prediction. These key factors that pose challenging issues in diagnosing and predicting diabetes can be categorized as follows [5]:

- **Data:** This challenge includes the availability of accurate and quality data, data collection and sharing, data privacy and security, data integration from heterogeneous sources, data access and storage
- **Clearing and preprocessing data:** This challenge includes selecting the correct data, clearing the data, selecting and extracting features, reducing dimensions, removing data noise, converting data, and integrating data.
- **Diagnosis and prediction techniques:** This challenge also includes general and global methods, clinical and public usability, evaluation of existing approaches concerning recent new data sets, robust software tools, development of online extraction tools and real-time forecasting, model selection Suitable are the integration of models from different fields and efficiency and accuracy.

Related works

Given the importance of this issue in recent years, researchers have developed data mining methods to solve one or more simultaneous challenges to diagnose diabetes early. We will review some of the most important of these researches.

In reference [2], the diagnosis of diabetes has been made by analyzing data in patterns and classifying using a Decision Tree (DT). In this study, the DT algorithm is responsible for analyzing the patterns in the data, and the results of the evaluations show an accuracy of 79.86%.

In the diagnosis of diabetes using algorithms such as Naive Bayes (NB), Logistic Regression (LR), DT, and Random Forest (RF) were done in reference [6], and the results were compared with each other. These comparisons showed that the LR algorithm obtained the best diagnostic performance with 80.43% accuracy and a mean squared error of 0.3748 among the compared algorithms.

KNN, NB, SVM, and DT algorithms have been used to diagnose diabetes in reference [7]. The simulation results in diagnosing diabetes in NB algorithms, KNN, SVM, and DT are 69.68, 31.30, 70.86, and 67.17, respectively.

Reference [8] used SVM classifiers, Artificial Neural Network (ANN), DT, KNN, and NB to diagnose type 2 diabetes. The simulations showed that the ANN, SVM, DT, KNN, and NB achieved 97.44, 81.19, 95.03, 90.85, and 91.60% accuracy in diagnosing diabetes. The results show that ANN is more efficient in diagnosing diabetes than other algorithms.

Reference [9] can be considered a new method for diagnosing diabetes. SVM and EA select the best parts in the proposed approach to reduce the data dimension. The proposed method succeeded in diagnosing diabetes with 95.81% accuracy.

The possible diagnosis of diabetes in patients using SVM, DT, and NB algorithms is discussed in the reference [10]. The Indian Diabetes Database is their reference in these tests. The use of an NB classification algorithm shows an accuracy of 76.30%.

In the reference [11], using the predicted significant features, various tools have been provided to determine and select the part for clustering in the prediction and according to the law of the association. Feature selection in this study was made through the Principal Component Analysis method. ANN algorithms, RF, and Clustering Techniques (CT) have been used to classify diabetes here. Among the mentioned methods, the RF method had 74.7%, the ANN had an accuracy of 75.7%, and the K-means Clustering algorithm had an accuracy of 73.6%.

Prediction algorithms have used Machine Learning (ML) algorithms to find the optimal classification [12]. Among the methods used, DT and RF algorithms have the highest accuracy with 98.2% and 98%, respectively.

In [13], The ML technique using DM with the help of Law Classifiers is presented. To evaluate the accuracy, the LR method was used in the results to predict the classification. This evaluation shows that age, gender, and body mass are the main causes of diabetes.

ANN, KNN, and SVM algorithms have been studied in [14] to classify diabetic patients. The deposit method has been used to determine the best characteristics in reducing the volume of data. This study shows the SVM algorithm performs better in data classification than different algorithms. DM methods SVM, NB, J48, RF, and KNN in the reference [15] have been used to increase diagnosing diabetes. KNN algorithm with 98.07% accuracy in diagnosis has a higher efficiency than the others.

Research highlights

In previous research, each of them has sought to provide a way to improve it along with various classification methods. One of the most critical challenges is selecting the optimal feature to reduce the maximum data dimension while maintaining essential and practical data. This issue is still one of the most important research topics pursued by various researchers. Accordingly, in this paper, a new method using the Binary Exchange Market Algorithm (BEMA) to identify essential and practical features (a discrete problem) of the diabetes data set to reduce the data dimension along with determining the best binary function of the algorithm (type Sigmoid function) is presented. The proposed BEMA is performed in the best possible way in generating and organizing random numbers due to having two search operators and two absorber operators. Optimization algorithms have various limitations, but compared to other algorithms, this algorithm has fewer limitations and difficulties in terms of being confined and not being able to find the optimal point, respectively. These features are due to the rapid convergence of local optimizations (exploratory problem) and converging to different responses (extraction problem) each time the program is run. To evaluate the performance of the proposed BEMA, several types of SVM, KNN and NB have been used to train and test the final model.

Materials and methods

In this section, first, the Exchange Market Algorithm (EMA) and binarizing of the algorithm used in this paper are explained briefly. Then, the proposed method based on the algorithms is described. Evaluation parameters are provided at the end of this section as well.

Exchange Market Algorithm (EMA)

The EMA is a new optimization algorithm introduced in 2014 [16]. An evolutionary optimization algorithm has been developed by studying the stock exchange performance by market elites. In this algorithm, shareholders buy and sell different stocks in different market conditions. It is assumed that shareholders compete to become the most successful sellers in the ranking list. It is considered here that people with low ratings usually face reasonable risks to make more profit, and it is assumed that intelligent shareholders act like successful stock market people.

In general, there is a certain number of stocks (optimization problem variable). At the same time, each person tries to buy a cost-effective part of these stocks (initialization of problem variables). These variables are determined in each iteration. Then, the algorithm is calculated according to the stocks purchased (the variable indicates the optimization problem by the algorithm), which is to maximize the possible profit in the market. The operation steps of the algorithm are as follows:

- 1) Determining the initial values and attributing them to the shareholders (in this step, the number of shares, stock values, the number of initial shareholders, and the required number of iterations are determined according to the type of optimization problem, and the shareholders take a number of any random stocks.
- 2) Calculating the cost of shareholders and their ranking (at this stage, members are evaluated and, according to the total validity of their shares, divided into three separate groups to identify different groups of shareholders. It should be noted that the group refers to members with high, middle, and lower ranks of shareholders and is part of the total population. The name of this group is only due to some notable changes in them.
- 3) Applying changes in the shares of the second group in equilibrium market conditions (at this stage, the members of the first group or members with high rank and top members remain unchanged, according to Eq. (1) Change stocks).

$$pop_j^{group(2)} = r \times pop_{1,i}^{group(1)} + (1 - r) \times pop_{2,i}^{group(1)} \tag{1}$$

$$i = 1, 2, 3, \dots, n_i \text{ and } j = 1, 2, 3, \dots, n_j$$

- 4) Apply changes in the stocks of third group members in equilibrium market conditions (The members of this group have a low ranking in terms of performance, and their stock values change according to Eq. (3)).

$$S_K = 2 \times r_1 \times (pop_{i,1}^{group(1)} - pop_k^{group(3)}) + 2 \times r_1 \times (pop_{i,2}^{group(1)} - pop_k^{group(3)}) \tag{2}$$

$$pop_k^{group(3),new} = pop_k^{group(3)} + 0.8 \times s_k \tag{3}$$

- 5) Determining shareholder costs and ranking them (until now, the goal was to find the optimal point, and there was no market fluctuation. At this stage, the main population of shareholders is evaluated in terms of performance and is sorted based on the value of the member. Next, the share of middle and low-ranking members is divided into separate groups based on the changes)
- 6) Trading of group members' shares under market fluctuation conditions using Eq. (4) (at this stage, the first group members remain unchanged as the top members. The middle members or the second group trades their shares according to Eq. (4) and therefore changes their shares)

$$\Delta_{nt1} = n_{t1} - \delta + (2 \times r \times \mu \times \eta_1) \tag{4}$$

$$\mu = \frac{t_{pop}}{n_{pop}} \tag{5}$$

$$n_{t1} = \sum_{y=1}^n s_{ty}, y = 1, 2, 3, \dots, n \tag{6}$$

$$\eta_1 = n_{t1} \times g_1 \tag{7}$$

$$g_1^k = g_{1,max} - \frac{g_{1,max} - g_{1,min}}{iter_{max}} \times k \tag{8}$$

$$\Delta_{nt2} = nt_2 - \delta \tag{9}$$

- 7) Trading stocks of members of the third group using Eq. (10) under market fluctuation conditions (at this stage, unlike what happened in the previous stage, shareholders trade some stocks according to Eq. (10) without considering their total stocks)

$$\Delta_{nt3} = (4 \times r_s \times \mu \times \eta_2) \tag{10}$$

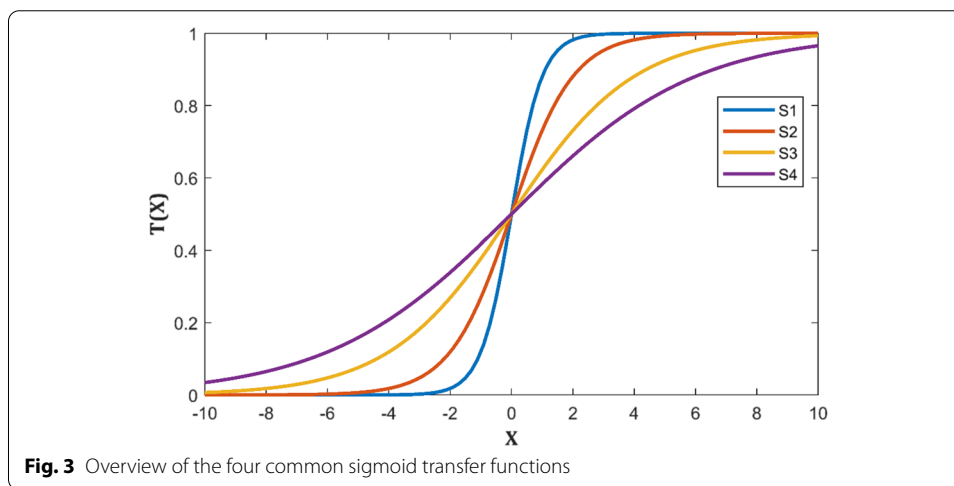
$$r_3 = (0.5 - rand) \tag{11}$$

$$\eta_2 = n_{t1} \times g_2 \tag{12}$$

- 8) Here, the end conditions are checked, and in case of dissatisfaction with the end conditions of the program, it is referred to as stage 2 (at this stage, the market volatility conditions end, and the shareholder evaluation program is performed. If the conditions are not satisfied, it is returned to stage 2. Otherwise, if the final requirements are met, the program operation will end).

Table 1 Mathematical formulation of four common sigmoid transfer functions

Function	Transfer function	Transfer function in EMA
S1	$T(x) = \frac{1}{1+e^{-2x}}$	$T(BEMA_i^d(t)) = \frac{1}{1+e^{-2BEMA_i^d(t)}}$
S2	$T(x) = \frac{1}{1+e^{-x}}$	$T(BEMA_i^d(t)) = \frac{1}{1+e^{-BEMA_i^d(t)}}$
S3	$T(x) = \frac{1}{1+e^{(-\frac{x}{2})}}$	$T(BEMA_i^d(t)) = \frac{1}{1+e^{(-\frac{BEMA_i^d(t)}{2})}}$
S4	$T(x) = \frac{1}{1+e^{(-\frac{x}{3})}}$	$T(BEMA_i^d(t)) = \frac{1}{1+e^{(-\frac{BEMA_i^d(t)}{3})}}$



Binarizing the algorithm

The feature selection problem becomes discrete because the algorithms generate continuous numbers and improve them. In order to binary the algorithm, sigmoid functions are used, as seen in Table 1, the output diagram of which is shown in Fig. 3. Choosing the optimal type in solving any feature selection problem is also a challenge. In this step, number one is used to select an attribute from a data set, and the number zero is used to display it if the attribute is not selected. In Table 1, $BEMA_i^d$ shows the continuous value of solution i in the population of the algorithm in dimension d in iteration t .

Figure 3. shows that the data received from the four sigmoid functions are continuously between 0 and 1. The output of the feature selection problem must be discrete. In most discretization papers, a threshold is taken after the output of the four sigmoid functions, which is usually the best case in meta-heuristic algorithms using the rand function. Therefore, according to Eq. (13), after applying four sigmoid functions to the continuous solutions of the EMA, a threshold for the final conversion of the algorithm to a discrete binary state occurs.

$$BEMA_i^d(t + 1) = \begin{cases} 1 & \text{if } rand(0, 1) < sg(BEMA_i^d(t)) \\ 0 & \text{if } rand(0, 1) \geq sg(BEMA_i^d(t)) \end{cases} \tag{13}$$

In Eq. (13), the rand function (0,1) produces a number between zero and one uniform distribution type. All the EMA solutions will be converted to a discrete binary classification based on this equation.

Proposed method

Flowchart The proposed method with the aim of optimal feature selection using the BEMA for the diagnosis of diabetes is presented in Fig. 4.

According to the flowchart of Fig. 4, the method is classified into the following phases:

- First phase: calling, normalizing, and segmenting data

In this section, the data are called first, and then the data is normalized using the minimum–maximum method and Eq. (14). Placing data in a specific domain is done by normalization. Ignoring this step will reduce the correct diagnosis in the objective function. Therefore, the application of this method causes the various dimensions to be examined relatively by the algorithm, and the effect of one is not more than the others.

$$Z = \frac{X - \min(x)}{\max(x) - \min(x)} \tag{14}$$

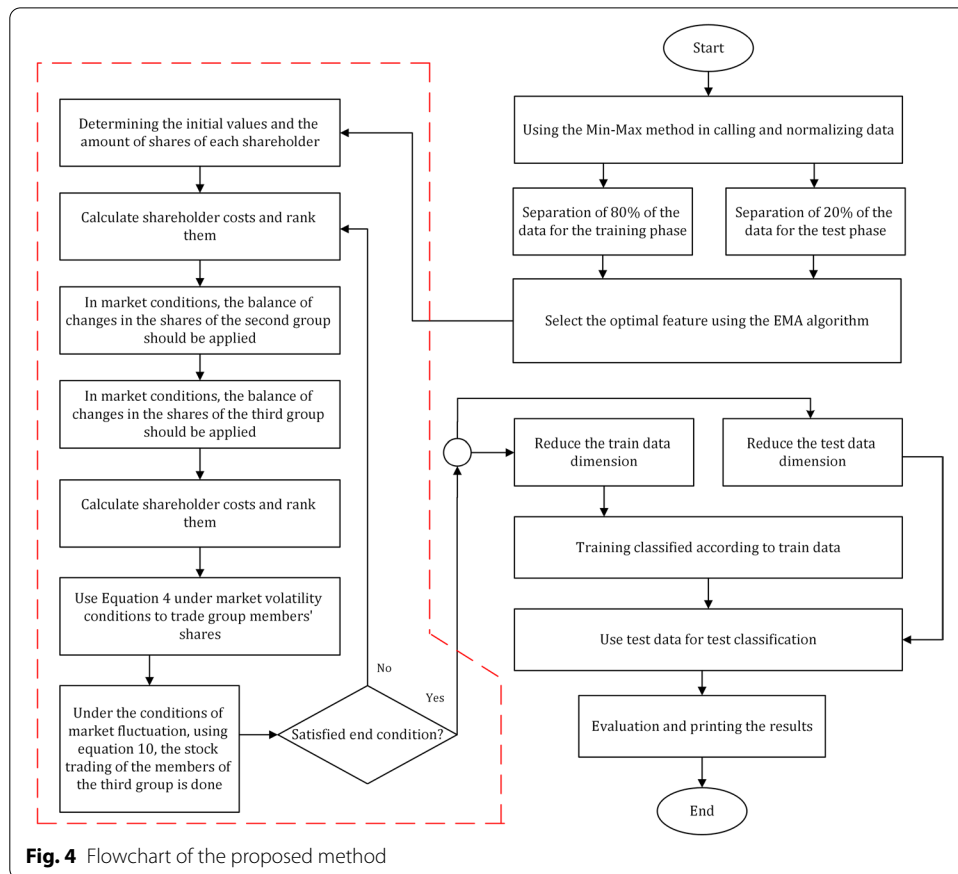


Fig. 4 Flowchart of the proposed method

After normalizing the data, the two sets of training and testing divide the data by 80% to 20%, respectively.

- Phase 2: Optimal selection of sigmoid function type, optimal feature selection, and reduction of data dimension

In this phase, the most optimal type of sigmoid function is selected to binary the algorithm to achieve the best type of feature selection (reducing data volume and increasing detection accuracy) in the diabetes dataset. In this step, first, the algorithm generates a set of initial solutions. According to the algorithm process described in the second part (Using Eqs. (1) to (12)), optimization will be performed to select the optimal type of sigmoid function and the optimal selection of features. Finally, several important and practical features remain to reduce the maximum dimension of the data while the critical information is not deleted.

- **Phase 3: Training, testing, and classification**

After reducing the volume of data by the optimally selected attributes, the desired classifications, several types of SVM, KNN and NB have been used to train the data and tests it in the next step use of the criteria presented in relations (15) to (18) to evaluate the performance of the algorithm in the optimal attribute selection step.

Evaluation Parameters

In each of the requirements given in Eqs. (15) to (18), there are four parameters, the definition of each of which represents the values of the confusion matrix parameters as follows:

- TN: Indicates the number of records whose actual category is negative, and the classification algorithm has correctly identified the variety as unfavorable.
- TP: Indicates the number of records whose actual category is positive, and the classification algorithm correctly identifies the variety as favorable.
- FN: Indicates the number of records whose actual batch is positive, and the batch classification algorithm has erroneously detected as unfavorable.
- FP: Indicates the number of records whose actual batch is negative, and the batch classification algorithm has erroneously identified as favorable.
- Accuracy: This criterion is the ratio of the number of correctly predicted classes and shows the degree of accuracy of prediction. The following equation shows the calculation of this criterion.

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \tag{15}$$

- Sensitivity: This criterion indicates the accuracy of the prediction model and means classes that are ready for prediction error. Therefore, this criterion indicates the ability of the algorithm to detect positive categories. The following equation shows the calculation of this criterion.

$$Sensitivity = \frac{TP}{TP + FN} \tag{16}$$

- **Specificity:** This criterion indicates the efficiency of classification incorrect prediction and thus no involvement with the disease. The following equation shows the calculation of this criterion.

$$Specificity = \frac{TN}{FP + TN} \tag{17}$$

- **F-Measure:** This criterion is an appropriate parameter for evaluating classification quality. This criterion describes the weighted average between the two quantities of sensitivity and specificity. The following equation also shows the calculation of this criterion.

$$F - Measure = 2 \times \frac{Sensitivity \times Specificity}{Sensitivity + Specificity} \tag{18}$$

Results

In this section, first, the dataset employed in this research is explained. Then, classification results achieved using the proposed method are presented.

Dataset

In this study, the Indian PIMA diabetes dataset was used to evaluate the proposed method for diagnosing type 2 diabetes [17]. This is a non-linear dataset of Indian women 21 years of age and older and is contained in the Kaggle dataset in reference [18]. This dataset includes the test data of 768 individuals with eight attributes with real integers, according to Table 2. Also, two diagnostic classes (class 1 for healthy people and class 2 for people with diabetes) have been used to determine whether people are healthy or sick.

Results of the proposed algorithms and comparison

In order to simulate the proposed method, MATLAB software is used as a well-known fourth-generation software environment for performing numerical calculations.

In this paper, the BEMA is used in the proposed method to select the type of sigmoid function optimally. The parameters set in this algorithm are such that after

Table 2 Features related to the diabetes database [18]

Feature description	Number of times pregnant
Number of times pregnant	[0–17]
Plasma glucose concentration a 2 h in an oral glucose tolerance test (mg/dl)	[0–199]
Diastolic blood pressure (mm Hg)	[0–122]
Triceps skin fold thickness (mm)	[0–99]
2-h serum insulin (mu U ml)	[0–864]
Body mass index (weight in kg/(height in m) ²)	[2.42–18.2]
Diabetes pedigree function	[0.078–1]
Age (years)	[21–72]
Class	[1–2]

Table 3 Selected features by the proposed method [19]

Feature name	Feature description	Selective feature
Pregnancies	Number of times pregnant	–
Glucose	Plasma glucose concentration a 2 h in an oral glucose tolerance test (mg/dl)	✓
BloodPressure	Diastolic blood pressure (mm Hg)	–
SkinThickness	Triceps skin fold thickness (mm)	✓
Insulin	2h serum insulin (mu U.ml)	–
BMI	Body mass index (weight in kg/(height in m)^2)	✓
DiabetesPedigreeFunction	Diabetes pedigree function	–
Age	Age (years)	–

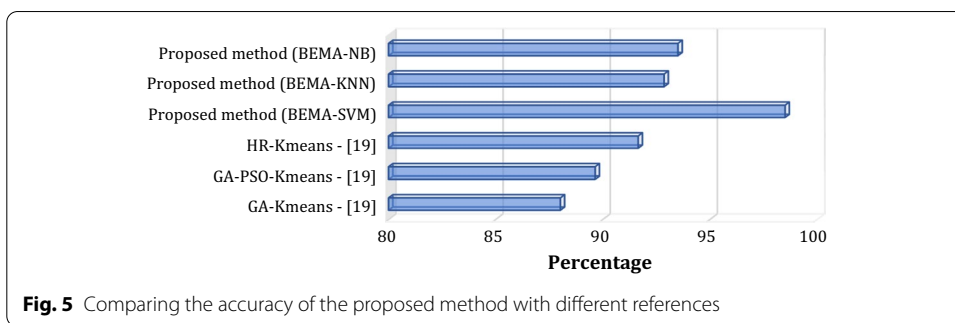
Table 4 Compare the results of the proposed method

References	Algorithm	Accuracy %	Sensitivity %	Specificity %	F-Measure %
[19]	HR-Kmeans	91.65	91.11	50	64.54
	GA-PSO-Kmeans	89.64	86.65	75.33	80.59
	GA-Kmeans	88.02	83.73	50	62.95
Without Feature selection	BEMA-SVM	79.221	100	0	88.406
	BEMA-KNN	81.169	100	0	89.606
	BEMA-NB	83.766	100	0	91.166
With Feature selection	BEMA-SVM	98.502	100	90.625	98.785
	BEMA-KNN	92.857	96.8	75.862	95.652
	BEMA-NB	93.506	93.023	96	96

various implementations and determining the critical parameters of the BEMA in receiving the best response from the objective function to increase the accuracy of diagnosis by considering the training set and test with a ratio of 80% to 20% in the data classification, the adjustment values of the parameters are as follows: the number of iterations is equal to 20, and the number of population is equal to 100. Number of the population associated with group 1 with non-oscillating mode ($0.2 \times$ number of population), number of the population associated with group 2 with non-oscillating mode ($0.3 \times$ number of population), number of the population related to group 3 with non-oscillating mode ($0.5 \times$ number of population), number of the population related to group 1 with oscillating mode ($0.2 \times$ number of population), number of the population related to group 2 with oscillating mode ($0.3 \times$ number of population), number of the population related to group 3 with oscillating mode ($0.5 \times$ number of population), and $g_1 = g_2 = [1 \times 10^{-1}, 5 \times 10^{-2}]$.

Specific features and optimal selection in Table 3 are the results of implementing the proposed method.

The results show that three features out of eight main features have been selected as optimal features, equivalent to a 62.5% reduction in the number of features. After



selecting and reducing the data dimension by the desired features, the three classification algorithms, SVM, KNN, and NB, have been trained and tested. Table 4 shows the results of this test and its evaluation with reference [19], the results of which are compared with eight other methods and show high-performance results.

To better display, the results obtained by comparing the accuracy of the results of the proposed method in both cases without selecting the feature and by selecting the feature with references [19] are shown in Fig. 5.

The results of Table 4 and Fig. 5 show that in addition to determining the appropriate algorithm for the optimal selection of features, selecting the type of classification can also improve the results. This improvement is visible due to examining the three categories, SVM, KNN, and NB, alongside the proposed BEMA algorithm.

According to the obtained results, it was found that the proposed BEMA algorithm, by selecting 3 optimal features, was able to detect the accuracy more accurately than the reference results [19], which in addition to the three methods presented in it (GA-Kmeans, GA-PSO-Kmeans and HR-Kmeans) has been compared with 5 references [17, 20–23]. However, in the reference [19], by selecting three features, the proposed method has an accuracy of 91.65%. The proposed method has an accuracy of 98.502 percent, has been able to improve by 6.95 percent in accuracy. The only common feature between the proposed and reference method [19] is glucose, which indicates the high importance of this feature in diagnosing diabetes. These cases show the efficiency of the proposed method in diagnosing diabetes due to the optimal use of the BEMA algorithm with two purposes of exploration and extraction. Among the evolutionary search algorithms, the optimal use of two factors of exploration and extraction is significant. It is also essential to strike a balance between these two factors. Here, exploration means that the search is universal, meaning that the algorithm requires more investigation.

Extraction means getting better answers around an answer; exploration seeks a global search. But extraction requires small and significant changes that change the current response. In general, methods’ computational complexity and performance depend on their application. Therefore, according to the application of the proposed BEMA-SVM algorithm for the diagnosis of diabetes based on the Indian Diabetes Database PIMA, it was able to perform better than the HR-Kmeans-KNN method presented in the reference [19] and be competitive with it.

Conclusion

This paper examines a variety of data mining algorithms to diagnose diabetes and its effects. It was shown here that the proposed method algorithm in determining the optimal type of sigmoid function and optimal properties using the SVM algorithm has the highest accuracy and application in predicting diabetes. This result indicates the ability to use the proposed BEMA in the optimal use of two factors of exploration and extraction in determining the optimal results. The results show that the factors of blood plasma glucose concentration, triceps skin thickness, and body mass index had the most significant impact on the diagnosis of diabetes. Compared with other previous methods, the proposed method is at an acceptable level with an accuracy of 98.502% detection. It may help assist medical professionals in making treatment decisions.

Reference [19] is used to compare the results of the method proposed in this paper. The type of algorithms for selecting the optimal properties by keeping the classification type, i.e., KNN, the constant, has been determined. But in the proposed method of this article, the opposite of this issue has been examined, and this issue has been one of the reasons for the superiority of the proposed method. Because combining a super-innovative algorithm with its proper classification can provide more acceptable results. Therefore, researchers can use a combination of techniques presented in this paper and reference [19] as well as simultaneous adjustment of classification parameters instead of their default settings in the future.

Proper generation and organization of random numbers due to the existence of two attracting operators and two effective and efficient search operators, the ability to select the search area and thus the ability to optimize different types of problems, converge to precisely the same answers each time the program runs, not staying in local optimized and therefore very high ability to derive the global optimal point from the advantages of the proposed BEMA. Thus, observing these results, the use of this algorithm as an efficient and reliable tool to solve engineering and applied problems is presented.

However, one of the challenges of the proposed method in using the BEMA is the optimal selection of the parameters of this algorithm, which in this paper has been done by trial and error. In future research, methods such as Taguchi can solve this problem. Using the Taguchi method, various response functions can be estimated based on the specified factors. Estimated results help us obtain the factors that lead to the best results for the intended test. Another challenge of the BEMA is its high execution time compared to other meta-heuristic algorithms, which is twice the execution time of different algorithms in optimization.

Acknowledgements

The authors would like to thank Mansour Moradi for his generous support and encouragement during the preparation this manuscript.

Author contributions

All the corresponding authors contributed equally to the conduct of the present study. All authors read and approved the final manuscript.

Funding

Not applicable.

Availability of data and materials

<https://www.kaggle.com/uciml/pima-indians-diabetes-database>

Declarations

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

Received: 27 September 2021 Accepted: 28 March 2022

Published online: 25 April 2022

References

1. <https://diabetesatlas.org/data/en/world/>
2. Iyer A, Jeyalatha S, Sumbaly R. Diagnosis of diabetes using classification mining techniques. arXiv preprint arXiv:1502.03774. 2015.
3. <https://www.shreyashhospital.org/latest-update/complications-of-dia/7>
4. Endocrine and Metabolism Research Institute. Education of Type1 Diabetes. 2010. <http://emri.tums.ac.ir/upfiles/60782275.pdf>
5. Khan FA, et al. Detection and Prediction of Diabetes using Data Mining: A Comprehensive Review. IEEE Access. 2021.
6. Nabi M, Wahid A, Kumar P. Performance analysis of classification algorithms in predicting diabetes. Int J Adv Res Computer Sci. 2017;8:3.
7. Patel PPS, Priya B, Himanshu DP. Analyze Data Mining Algorithms for Prediction of Diabetes. Int J Eng Develop Res. 2017;5(3):466–73.
8. Heydari M, Timuri M, Mohammad Z, et al. Comparison of various classification algorithms in the diagnosis of type2 diabetes in iran-the international standard for diabetes mellitus. Int J Diabetes Develop Ctries. 2016;36(2):167–73.
9. Baghi Yazdel R, Jaali E, Khodaei E, Habibi M. Methods of dealing with missing data: advantages, disadvantages, theoretical approaches and application of software. Higher Educ Lett. 2016;9(33):11–37.
10. Sisodia D, Sisodia DS. Prediction of diabetes using classification algorithms. Procedia Computer Sci. 2018;132:1578–85.
11. Alam TM, Iqbal MA, Ali Y, Wahab A, Ijaz S, Baig TI, Abbas Z. A model for early prediction of diabetes. Inform Med Unlocked. 2019;16:100204.
12. Sneha N, Gangil T. Analysis of diabetes mellitus for early prediction using optimal features selection. J Big Data. 2019;6(1):13.
13. Sohail MN, Jiadong R, Muhammad MU, Chauhdary ST, Arshad J, Verghese AJ. An accurate clinical implication assessment for diabetes mellitus prevalence based on a study from Nigeria. Processes. 2019;7(5):289.
14. Kaur H, Kumari V. Predictive modelling and analytics for diabetes using a machine learning approach. Appl Comput Inform. 2020;23:2.
15. Alpan K, Ilgi GS. Classification of Diabetes Dataset with Data Mining Techniques by Using WEKA Approach. In: 2020 4th International Symposium on Multidisciplinary Studies and Innovative Technologies (ISMSIT) (pp. 1–7). IEEE; 2020.
16. Ghorbani N, Babaei E. Exchange market algorithm. Appl Soft Comput. 2014;19:177–87.
17. Beloufa F, Chikh MA. Design of fuzzy classifier for diabetes disease using Modified Artificial Bee Colony algorithm. Comput Methods Programs Biomed. 2013;112(1):92–103.
18. <https://www.kaggle.com/uciml/pima-indians-diabetes-database>
19. Li X, Jusheng Z, Fatemeh S. Improving the accuracy of diabetes diagnosis applications through a hybrid feature selection algorithm. Neural Process Lett. 2021;45:1–17.
20. Choubey DK, Kumar P, Tripathi S, Kumar S. Performance evaluation of classification methods with PCA and PSO for diabetes. Netw Model Anal Heal Inf Bioinf. 2020;9(1):5.
21. Hasan S, Shamsuddin SM. Multi-strategy learning and deep harmony memory improvisation for self-organizing neurons. Soft Comput. 2019;23(1):285–303.
22. Choubey DK, Kumar M, Shukla V, Tripathi S, Dhandhanika VK. Comparative analysis of classification methods with PCA and LDA for diabetes. Curr Diabetes Rev. 2020;16(8):833–50.
23. Mohamadi M, Tab FA, Soltanian K (2019) Evolutionary feature selection based on semi-local search. In: 2019 4th international conference on pattern recognition and image analysis (IPRIA), pp 228–233

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.