Journal of Artificial Intelligence and Data Mining (JAIDM), x(x): xxx-xxx, xxxx

Shahrood University of Technology

Journal of Artificial Intelligence and Data Mining (JAIDM)
Journal homepage: http://jad.shahroodut.ac.ir

Research paper

An Intelligent Fuzzy System for Diabetes Disease Detection using Harris Hawks Optimization

Zahra Asghari Varzaneh, Soodeh Hosseini*

Department of Computer Science, Faculty of Mathematics and Computer, Shahid Bahonar University of Kerman, Kerman, Iran.

Article Info

Article History:
Received 31 October 2022
Revised 09 December 2022
Accepted 26 March 2023

DOI:10.22044/JADM.2023.12383.2383

Keywords:
Fuzzy Expert System, Harris Hawks Optimization, Membership Functions, Disease Diabetes.

Corresponding author: so_hosseini@uk.ac.ir(S. Hosseini).

Abstract

This paper proposes a fuzzy expert system for diagnosing diabetes. In the proposed method, at first, the fuzzy rules are generated based on the Pima Indians Diabetes Database (PIDD), and then the fuzzy membership functions are tuned using the Harris Hawks Optimization (HHO). In the experimental dataset, PIDD with the age group from 25-30 is initially processed and the crisp values are converted into fuzzy values in the stage of fuzzification. The improved fuzzy expert system increases the classification accuracy, which outperforms several famous methods for diabetes disease diagnosis. The HHO algorithm is applied to tune fuzzy membership functions to determine the best range for fuzzy membership functions and increase the accuracy of fuzzy rule classification. The experimental results in terms of accuracy, sensitivity, and specificity prove that the proposed expert system has a higher ability than other data mining models in diagnosing diabetes.

1. Introduction

Diabetes is known as a common disease affecting millions of people around the world [1]. According to the latest statistics released by the World Health Organization (WHO) and the International Diabetes Federation, one in four people over the age of 60 have diabetes [2]. Diabetes is a serious disease that affects almost every organ in the body like the heart, eyes, kidney, skin, nerves, blood vessels, and foot [3]. Early detection and prevention of this disease reduces mortality, and prevents and reduces its complications. Today, physicians face a large amount of medical data. Data mining tools can be widely used in medical sciences. Data mining technique is a way to automatically analyze data, and identify hidden patterns among them [4]. Various data mining methods have been widely used in medical research works that can help physicians diagnose diseases and reduce errors. Expert systems in patient health care are used to improve the quality and efficiency of health data in patient care [5, 6].

A Fuzzy Expert System (FES) is an intelligent system. This system applies fuzzy logic for reasoning and calculations. In fact, FES is a type of expert system that uses fuzzy logic instead of dual value logic [7]. The FES consists of a set of membership functions and rules that are collectively used for reasoning. Fuzzy systems are based on perception or rules, and the core of each fuzzy system is a knowledge base consisting of fuzzy rules. When there is an increase in the possible number of rules of the FES, experts find it difficult to define the complete rule set. Also the performance of the system can be increased by tuning of membership function using optimization algorithms [8, 9]. Harris Hawks Optimization (HHO) is a metaheuristic algorithm inspired by the behavior of intelligent birds named Hawk

Harris [10]. This species has a mechanism that allows them to hunt even when escaping. In this paper, an FES is used to diagnose diabetes disease. This system uses if-then rules, and due to its membership functions with precise degrees of affiliation, they have a high ability to determine fuzzy rules and can help diagnose diabetes with good accuracy. The HHO algorithm is applied to tune fuzzy membership functions to determine the best range for fuzzy membership functions, thus increasing the accuracy of fuzzy rule classification. The experimental results prove that the proposed expert system has a higher ability than other data mining models in diagnosing diabetes. In the rest of the paper, the related works about the diagnosis of diabetes methods are briefly explained in Section 2. Section 3 presents the theoretical background regarding FES and HHO, respectively. Section 4 describes the details of the proposed method. In Section 5, the experimental results are presented. Finally, conclusions and future work are presented in Section 6.

2. Related Works

There have been a lot of studies reported in the literature to predict diabetes. More specifically, the use of approaches like Early Neural Network (ENN) [11], C4.5 rules [12], and k-NN classification [13] have been previously reported. Lee [14] developed an FES-based decision support program to predict diabetes. This system is based on fuzzy logic, and is created with a five-layer fuzzy ontology. Importantly, the proposed fuzzy expert system can work effectively for diabetes decision support applications. Senthilkumar [15] developed an FES that aimed to diagnose diabetes using a fuzzy verdict mechanism using the concept of fuzzification with a triangular membership function. The proposed method works more effectively for diabetes applications than the previously developed ones. Jain and Raheja [16] used a fuzzy verdict mechanism in the Fuzzy Logic-based Diabetes Diagnosis System (FLDDS). It considers the information collected from patients as an input dataset. This system generates rules based on physicians' knowledge, and uses them to predict diabetes. However, the evaluation results have not been investigated on large datasets. Shirali *et al.* [17] improved the diagnosis of diabetes using fuzzy systems. The system used in their research work was Sugeno fuzzy inference systems and intelligent algorithms firefly. The proposed method enables the use of a few simple fuzzy rules to detect diabetes with a good accuracy. The experimental results show that this method is more accurate than the existing algorithms in this field on the PID standard dataset. In [18], the authors presented a method for the diagnosis of diabetes using an artificial neural network and a neuro-fuzzy approach. According to the results, higher precision was obtained following the use of the neuro-fuzzy approach. Abedian *et al.* [19] proposed a data mining method for the diagnosis of diabetes based on native data from one of the specialized diabetes centers in Mashhad. Its database contains 254 independent records based on 13 characteristics. In their study, among different protocols used for pattern recognition, the SVM method displayed a significantly better performance. In 2020, a data mining-based model was introduced for the diagnosis of diabetes using data mining and a heuristic method combining neural networks and particle swarm intelligence [20]. According to the results, the sensitivity and accuracy of the proposed method are better and higher than previous similar methods. In [21], the authors proposed a hybrid model using the fuzzy decision tree and the bee colony algorithm called ABC-FDT. In ABC-FDT, the optimal number of fuzzy sets is considered for each feature to provide the best feature segmentation. The aim is to increase the accuracy of the classification and reduce the complexity of calculations.

Aamir *et al.* [22] used fuzzy logic to develop an interpretable model for early diagnosis of diabetes. Fuzzy logic was combined with the cosine domain method and made two fuzzy classifications. After that, fuzzy rules were designed based on these classifiers. The results showed that the proposed model worked well in the accurate diagnosis of diabetes. Gundluru *et al.* [23] designed a deep learning model with principal component analysis (PCA) for dimensionality reduction, and to extract the most important features, the Harris hawks optimization algorithm was used further to optimize the classification and feature extraction process. In [24], a predictive diabetes diagnosis model based on intelligent fuzzy inference rules (IFIR_PDDM) was proposed. In the proposed IFIR_PDDM model, first, a fuzzy membership function of medical recommendations was created. Medical experts then verified the mining-based rules using the decision tree rule induction technique. The proposed model predicted the risk of diabetes using fuzzy inference, and diagnosed diabetes in less time.

3. Preliminaries

3.1. Fuzzy computing

Fuzzy computing was proposed by Zadeh in 1965 AD [25]. Fuzzy computing provides a way to compute uncertain and ambiguous data and information, while providing an inference and reasoning mechanism based on a set of "if-then" rules. Fuzzy systems combine the concepts of fuzzy set theory and fuzzy logic with each other, and provide a framework for presenting linguistic knowledge with uncertainty, and have two main characteristics that have increased their popularity; one is that they are suitable for approximate reasoning, especially for systems for which mathematical modeling is difficult, and the other is that fuzzy logic allows decisions to be made using uncertain information with the help of linguistic variables [26].

3.2. Fuzzy expert system

The fuzzy expert system consists of five parts. The first part is the user interface that receives information about the system input variables from a database. The second section is the base of fuzzy rules, the third section is the fuzzy generator unit, the fourth section is the fuzzy inference motor, and the fifth part is the Defuzzifier unit. In the fuzzy phase, the input values are converted to fuzzy numbers through the fuzzy generator unit. Inference motors can infer outputs using fuzzy rules and operators. The Defuzzifier produces an output with a definite value of the fuzzy sets that are the output of the inference motor [27]. Figure 1 shows the structure of the system in general.

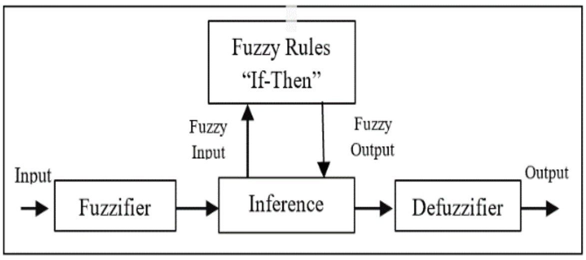

Figure 1. General architecture of the fuzzy expert system.

3.3. Harris hawks optimization

The HHO algorithm is a new meta-heuristic algorithm inspired by hawk behavior in nature, and has the two main steps of exploration and exploitation.

3.3.1. Exploration step

At this stage, the algorithm seeks to discover the search space, and update the position of the hawks based on a random operator or based on the position of other hawks, which is formulated as follows:

$$x(t+1) = \begin{cases} x_{rand}(t) - r_1 \left| x_{rand}(t) - 2r_2 x(t) \right|, & q \geq 0.5 \\ \left(x_{rabbit}(t) - x_m(t)\right) - r_3\left(LB + r_4(UB - LB)\right), & q < 0.5 \end{cases} \quad (1)$$

where q and $r_1 - r_4$ are random numbers in the interval [0, 1]. x_{rabbit} is the location of the rabbit (prey). $x(t+1)$ is the position of i^th hawk in (t + 1)^th iteration. The average location of the Hawks is . Moreover, UB and LB indicate the upper and lower boundaries of the variables.

Table 1. Attributes of PIDD.

Abbreviation	Full name
DBP	Diastolic blood pressure
Glucose	Plasma glucose concentration In 2-hours OGTT
Pregnant	Number Of times pregnant
TSFT	Triceps skin fold thickness
Age	Age
DPF	Diabetes pedigree function
BMI	Body mass index
INS	2-hour serum insulin
DM	Diabetes mellitus

The escaping energy of each prey is indicated as E, which is calculated as follows:

$$E = 2E_0\left(1 - \frac{t}{t_{max}}\right) \quad (2)$$

where E_0 is a random number in [-1, 1]. If $|E| \geq 1$, exploration is done, and if $|E| < 1$, exploitation is done.

3.3.2. Exploitation step

In this step, the Harris hawks perform the surprise pounce to attack prey. The number (r) indicates the chance of the prey to successfully escape ($r < 0.5$) or not successfully escape ($r \geq 0.5$) before the surprise pounce.

When $r \geq 0.5$ and $|E| \geq 0.5$, the prey has enough energy, and tries to escape but only gets tired . At the same time, Harris hawk surrounds the prey, and then attacks it with a sudden blow by the following rules:

$$x_i^{t+1} = \Delta x_i^t - E\left|Jx_{rabbit} - x_i^t\right| \quad (3)$$

$$\Delta x_i^t = x_{rabbit} - x_i^t, \quad (4)$$

The difference between the location vector of the prey and the current individual is indicated by Δx_i^t.

$J = 2(1 - r_5)$ is the jump strength of the prey throughout the escaping phase.
When r ≥ 0.5 and |E| < 0.5, the prey is exhausted, and has no escaping strength. For this case, the current location is updated using:

$$x_i^{t+1} = x_{rabbit} - E\left|\Delta x_i^t\right|. \quad (5)$$

Considering r < 0.5 and |E| ≥ 0.5, the prey has enough energy to escape before the attack comes as a surprise. This is formulated by:

$$Y = x_{rabbit}(t) - E\left|Jx_{rabbit}(t) - x(t)\right|. \quad (6)$$

The Levy flight is used to simulate the rapid, irregular, and abrupt movement of Harris hawks when chasing the prey as follows:

$$Z = Y + S \times Levy(D), \quad (7)$$

where Levy is the Levy flight function [28], S is a random vector with the size 1 × D, and D is the dimensions of the problem.
The final step of the Harris hawks is to update the positions of the hawks as follows:

$$x(t+1) = \begin{cases} Y, & if\ F(Y) < F(x(t)) \\ Z, & if\ F(z) < F(x(t)) \end{cases} \quad (8)$$

where:

$$Y = x_{rabbit}(t) - E\left|Jx_{rabbit}(t) - x_m(t)\right| \quad (9)$$

$$Z = Y + S \times Levy(D) \quad (10)$$

4. Proposed Method

This section describes the FES designed for the diagnosis of diabetes disease.

4.1. UCI Pima Indians diabetes database

This paper uses data from the National Institute of Diabetes and Digestive and Kidney Diseases (NIDDK) of Pima Indians to implement the proposed method [29]. The database contains 768 instances of female patients of Pima Indian heritage. This dataset contains eight attributes with integer values. All experiments were performed on these eight attributes; therefore, this dataset includes eight disease symptoms and a target variable (class) that indicates the presence or absence of disease and is indicated by values of 1 and 0, respectively. The dataset information includes the details given in Table 1. The data with the age group from 25-30 is taken to test the fuzzy determination mechanism.

4.2. Fuzzy inference system

The conversion from crisp to fuzzy input is known as fuzzification [30]. Each crisp input is converted to its fuzzy equivalent using a family of membership functions. For the PIDD, each case has eight input attributes and one output attribute, listed in Table 1, and each attribute can be constructed as a fuzzy variable with some fuzzy numbers. The relationship in the PIDD concerning age is given in Table 2 [31].

Table 2. Description of fuzzy relation.

Fuzzy relation	Description
R ≥ FZ (Age 0_25)	Very very Young
R ≥ FZ (Age 25_30)	Very young
R ≥ FZ (Age 30_35)	More or less young
R ≥ FZ (Age 35_40)	Slightly young
R ≥ FZ (Age 40_45)	Slightly old
R ≥ FZ (Age 45_50)	More or less old
R ≥ FZ (Age 50_55)	Very old
R ≥ FZ (Age 55_60)	Very very old

The fuzzy process used in this paper is the Mamdani fuzzy inference system; the crisp set of rules is transformed into a fuzzy model using a triangular membership function. It is given as follows:

$$f(x) = \begin{cases} 0, & x \le a \\ (x-a)/(b-a), & a < x \le b \\ (c-x)/(c-b), & b < x < c \\ 0, & x > c \end{cases} \quad (11)$$

where a and b determine the lower and upper bounds, respectively, while c is the peak of the triangle. Thus the rules are represented by fuzzy triangular membership functions. By fuzzification of the crisp input values, membership values and degrees are obtained. These obtained fuzzy values are processed in a fuzzy verdict mechanism. The output values are sent to the defuzzification phase unit using the center method according to (12), and from this unit, the final values will be as output [32].

$$DM_i = (\Sigma_{(i=1)}^n Z_i \mu(Z_i)) / (\Sigma_{(i=1)}^n \mu(Z_i)) \quad (12)$$

where Z_i is the center gravity of output membership functions, and $\mu(Z_i)$ is the number of fuzzy numbers of the output fuzzy variable DM.

In this paper, fuzzy variables glucose, INS, BMI, PD, and age have been used as input fuzzy variables, and DM variable as a fuzzy variable has been used to define fuzzy membership functions as well as various fuzzy rules to define a fuzzy expert system. Figure 2 and Figure 3 show examples of the membership functions of the fuzzy input variable "DPF" and the output variable "DM", respectively.

In the fuzzy inference system defined in this study, the DPF variable is defined as one of the input variables. Fuzzy membership functions are a set of input variables. In this study, low, medium, and high membership functions are defined for the input variables. However, the DM variable is an output variable whose values are independent from the input variables [31].

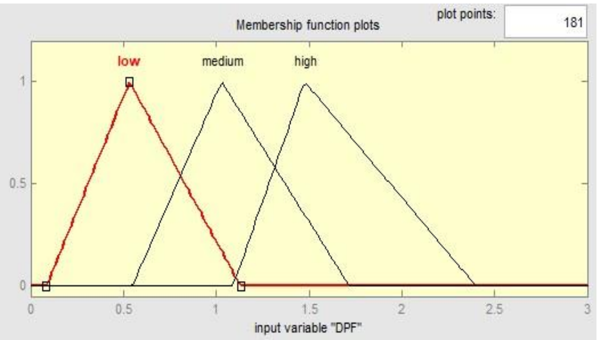

Figure 2. Membership function of input variable "DPF".

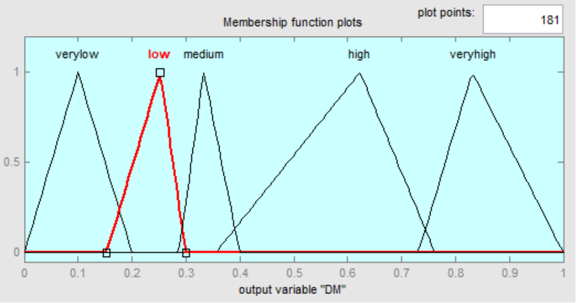

Figure 3. Membership function of output variable "DM".

In modeling this fuzzy expert system, it is necessary to write fuzzy inference rules. These rules show how the defined fuzzy sets relate to each other, and how they affect the final result. The knowledge related to the determination of inputs, outputs, and inference rules has been obtained through the knowledge of experts in the field of diabetes diagnosis. The rules are used in this paper given as follows [31]:

1- If (Glucose is Glucoselow) or (INS is INSlow) or (BMI is BMIlow) or (DPF is DPFlow) or (Age is Ageyoung), then (DM is DMverylow).

2- If (Glucose is Glucoselow) or (INS is INSlow) or (BMI is BMIhigh) or (DPF is DPFlow) or (Age is Ageyoung), then (DM is DMlow).

3- If (Glucose is Glucosemedium) or (INS is INShigh) or (BMI is BMIhigh) or (DPF is DPFmedium) or (Age is Ageyoung), then (DM is DMmedium).

4- If (Glucose is Glucosehigh) or (INS is INSmedium) or (BMI is BMIhigh) or (DPF is DPFhigh) or (Age is Ageyoung), then (DM is DMhigh).

5- If (Glucose is Glucoselow) or (INS is INSlow) or (BMI is BMImedium) or (DPF is DPFlow) or (Age is Ageyoung), then (DM is DMverylow).

The rule viewer of the generated FIS is shown in Figure 4.

As shown in Figure 4, by changing the input values, different outputs are obtained that can help physicians to make decisions.

Figure 4. FES generated by fuzzy rules.

4.3. Fuzzy membership optimization

The rule base is the main part of the fuzzy inference system, and the quality of results in a fuzzy system depends on the fuzzy rules. The rules are generated by experts easily but the MFs are difficult to obtain. Tuning of MFs is a time-consuming process. The fuzzy systems can be formulated as a space search problem, where each point in the space corresponds to a rule set and MFs. This makes evolutionary algorithms such as Genetic Algorithms (GAs) [33] and particle swarm optimization (PSO) [34], better choices for searching these spaces.

This paper proposes an HHO-based fuzzy expert system. The fuzzy determination mechanism executes rules with a fuzzy operator to make a decision on the possibility of individuals suffering from diabetes and to present the knowledge with descriptions. The fuzzy membership functions are tuned by HHO.

For each fuzzy membership function, there are three parameters, as shown in Figure 5: C (center), L (left), and R (right) correspond to the original membership function, where C′, L′, and R′ refer to the center, left, and right of the adjusted membership function. For the adjustment of membership functions, the following equations are defined:

$$\begin{aligned} C' &= (C + k_i) - w_i \\ L' &= (L + k_i) - w_i \\ R' &= (R + k_i) - w_i \end{aligned} \tag{13}$$

where k_i and w_i adjustment coefficients. This optimization influence on shrink or stretch of triangular membership functions.

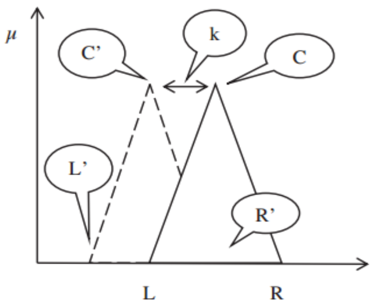

Figure 5. Fuzzy membership parameters.

5. Evaluation of system performance

In this paper, the HHO algorithm was used to tune fuzzy membership functions to diagnose diabetes. The initial population size was 30 members, and the number of iterations of the algorithm to find the best parameters was 100 times. To evaluate the efficiency of the proposed fuzzy classifier, the dataset is divided into two groups consisting of training and testing datasets. In this study, 70% of the dataset was selected as the training set, and 30% of it was selected as the test data.

The designed system was developed in Matlab's Simulink to the measuring scale for performance. The efficiency of the proposed expert system is calculated based on criteria such as accuracy, sensitivity, and specificity using Equations (14–16) [35, 36].

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \tag{14}$$

$$Sensitivity = \frac{TP}{TP + FN} \tag{15}$$

$$Specificity = \frac{TN}{TP + FN} \tag{16}$$

where the True Positive (TP) and the True Negative (TN) denote the correct classification. False Positive (FP) is the outcome when the predicted class is yes (or positive) and the actual class is no (or negative). Still, a False Negative (FN) is the outcome when the predicted class is no (or negative) and the actual class is yes (or positive). Table 3 lists the various outcomes of a two-class prediction.

Table 3. Classification accuracy obtained with different methods.

Method	Accuracy	Sensitivity	Specificity
Kalpana and Senthilkumar [31]	0.76	0.853	0.69
ENN	0.66	0.772	0.624
C4.5 rules	0.67	0.784	0.627
CART	0.708	0.811	0.674
k-NN	0.712	0.835	0.65
ID3	0.697	0.789	0.723
Our method by PSO	0.708	0.794	0.682
Our method by COA	0.753	0.846	0.728
Our method by HOA	0.73	0.840	0.673
Our method by HHO	**0.787**	**0.886**	**0.736**

As shown in Table 3, the use of a fuzzy inference system increases the accuracy of classification. The experimental results on the PIDD dataset show that the proposed fuzzy inference is better than other classification algorithms including ENN, C4.5, CART, k-NN, and ID3 in terms of accuracy, sensitivity, and specificity criteria. The numerical results obtained to compare the efficiency of algorithms show that fuzzy expert systems have higher accuracy and efficiency than other classification models. Moreover, the proposed method for tuning fuzzy membership functions uses four well-known metaheuristic algorithms that HHO algorithm has a higher efficiency than other algorithms. In addition, the results obtained for the criteria of sensitivity and specificity show that the proposed system based on the hybrid of HHO with the fuzzy classifier system can improve efficiency. Figure 6 shows the convergence curve of the proposed algorithm. In each iteration, the HHO algorithm seeks to find the best values to find the *k* and *w* parameters in (13) to find the best membership functions for the fuzzy variables that categorize accurately.

In the initial iterations of the algorithm, the fitness value (accuracy) of the fuzzy classifier is low, and after several steps, this value is improved and increased. The convergence curve stops after several consecutive iterations and converges to an optimal value. This convergence makes it impossible to achieve a better fitness value in successive iterations. The number of repetitions for the HHO algorithm is considered equal to 100, and Figure 6 shows 20 iterations of the algorithm.

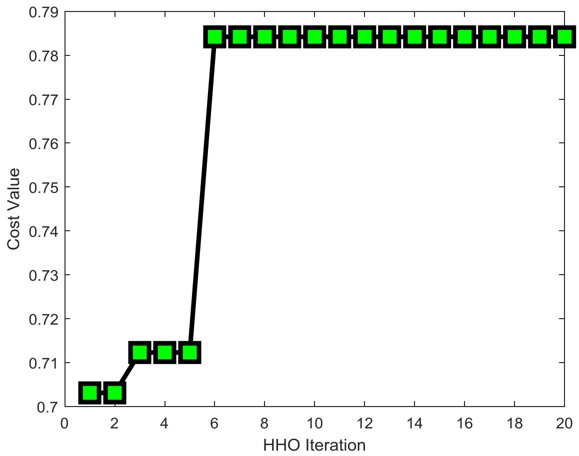

Figure 6. Proposed method by the HHO algorithm.

6. Conclusion

This paper used an HHO to tune the fuzzy membership functions (MFs) for diabetes disease diagnosis. In the experimental dataset, PIDD with the age group from 25-30 was initially processed and the crisp values were converted into fuzzy values in the stage of fuzzification. The proposed system based on a hybrid of Harris hawks optimization (HHO) with the fuzzy expert system shows its performance compared to other methods in terms of accuracy, sensitivity, and specificity. This system enables us to employ optimized fuzzy rules in the development process of a diabetes disease detection expert system effectively.

References

[1] M. Nazarzadeh, Z. Bidel, and A. Sanjari Moghaddam, “Meta-analysis of diabetes mellitus and risk of hip fractures small study effect,” *Osteoporos Int*, vol. 27, pp. 229, 2016.

[2] Global Burden of Disease Collaborative Network. Global Burden of Disease Study 2019. Results. Institute for Health Metrics and Evaluation. 2020. (https://vizhub.healthdata.org/gbd-results/).

[3] H. Temurtas, N. Yumusak, and F. Temurtas, “A comparative study on diabetes disease diagnosis using neural networks,” *Expert Systems with Applications*, vol. 36, pp. 8610–8615, 2009.

[4] A. Iyer, S. Jeyalatha, and R. Sumbaly, "Diagnosis of diabetes using classification mining techniques," in press, 2015.

[5] J. Haddadnia, J. Vahidi, A. Gharakhani, and A. Fouzi, “Fuzzy Diagnosis of Diabetes Mellitus Based on the Comprehension Rules and Characteristics Based on Combination of Data Mining Systems and Intelligent Intelligent Algorithms,” *International Conference on Non-linear Modeling and Optimization*, 2011.

[6] S. Elsappagh, M. Elmogy, and AM. Riad AM, “A fuzzy ontology oriented case-based reasoning framework for semantic diabetes diagnosis,” *Artif Intell Med*, vol.14, pp. 92-5, 2015.

[7] J. Buckley, W. Siler, D. J. F. s. Tucker, and systems, “A fuzzy expert system,” vol. 20, no. 1, pp. 1-16, 1986.

[8] Garcia, M. A. *et al.*, “ESDIABETES (AN EXPERT SYSTEM IN DIABETES),” *CCSC: South Central Conference, by the Consortium for Computing in Small Colleges*, pp. 166–175, 2001.

[9] Ganesh Kumar, P. *et al.*, “Design of a fuzzy expert system for microarray data classification using a novel genetic swarm algorithm,” *Expert Systems with Applications*. No. 39, pp. 1811–1821, 2012.

[10] A. A. Heidari, S. Mirjalili, H. Faris, I. Aljarah, M. Mafarja, and H. Chen, "Harris hawks optimization: Algorithm and applications," *Future Generation Computer Systems*, vol. 97, pp. 849-872, 2019.

[11] J. W. Smith, J. E. Everhart, W. C. Dickson, W. C. Knowler, and R. Johannes, “ Using the ADAP learning algorithm to forecast the onset of diabetes mellitus,” *In Proceedings of the 12th symposium on computer applications and medical care, Los Angeles, CA, USA*, pp. 261–265, 1998.

[12] K. Polat, S. Günes, and A. Arslan, “A cascade learning system for classification of diabetes disease: Generalized discriminate analysis and least square support vector machine,” *Expert Systems with Applications*, vol. 34, pp. 482–487, 2008.

[13] K. Saxena, Z. Khan, and S. Singh, "Diagnosis of diabetes mellitus using k nearest neighbor algorithm," vol. 2, pp. 36-43, 2014.

[14] C. S. Lee, “A Fuzzy Expert System for Diabetes Decision Support Application,” *IEEE transactions on systems, man, and cybernetics*, vol. 41, no. 1, 2011.

[15] M. Kalpana and A.V. Senthilkumar, “Fuzzy Expert System for Diabetes using Fuzzy Verdict Mechanism,” *Internationl Journal of Advanced Networking and Applications*, vol. 3, pp. 1128–1134, 2011.

[16] V. Jain and S. Raheja, “Improving the Prediction Rate of Diabetes using Fuzzy Expert System,” *I.J. Information Technology and Computer Science*, vol. 10, pp. 84-91, 2015.

[17] M. Shirali *et al.*, "Improvement diagnosis of diabetes using a combination of sugeno fuzzy inference systems and firefly algorithms," *Iranian Journal of Diabetes and Metabolism*, vol. 15, no. 3, pp. 172-176, 2016.

[18] I. Zabbah, A. Eskandari, Z. Sardari, and A. Noghandi, “ Diagnosis of Diabetes using Artificial

Neural Network and Neuro-Fuzzy approach," *Journal of Torbat Heydariyeh University of Medical Sciences*, vol. 6, no. 2, 2018.

[19] I. Abedian, A. Ayoobi, H. Ghaffary, and I. Zabbah, " Diagnosis of diabetes by using a data mining method based on native data," *Journal of Torbat Heydariyeh University of Medical Sciences*, vol. 7, no. 1, pp. 1-14, 2019.

[20] A. Panah and S. Fallahpour, "Predicting Type2 Diabetes Using Data Mining Algorithms," *J Mazandaran Univ Med Sci*, vol. 30, no. 191, pp. 22-30, 2020.

[21] M. Mojarad, E. Hajizadegan, and M. Gurkani, "Diagnosis of Diabetes Using Bee Colony Algorithm and Fuzzy Decision Tree," *Int. J. Sci. Res. in Network*, vol. 9, pp. 3, 2021.

[22] K. M. Aamir, L. Sarfraz, M. Ramzan, M. Bilal, J. Shafi, and M. Attique, "A Fuzzy Rule-Based System for Classification of Diabetes," *Sensors*, vol. 21, no. 23, p. 8095, Dec. 2021.

[23] N. Gundluru et al., "Enhancement of Detection of Diabetic Retinopathy Using Harris Hawks Optimization with Deep Learning Model," *Computational Intelligence and Neuroscience*, vol. 2022, 2022.

[24] P. Nagaraj and P. Deepalakshmi, "An intelligent fuzzy inference rule‐based expert recommendation system for predictive diabetes diagnosis," *International Journal of Imaging Systems and Technology*, 2022.

[25] L. A. ZADEH, "Fuzzy sets," *Information and Control*, vol. 8, pp. 338-353, 1965.

[26] W. Siler, J. J. Buckley, "Fuzzy expert systems and fuzzy reasoning," *John Wiley & Sons*, 2005.

[27] J. A. Chirinos, A. Veerani, J. P. Zambrano, A. Schob, G. Perez, A. J., Mendez, and S. Chakko, "Evaluation of comorbidity scores to predict all-cause mortality in patients with established coronary artery disease," *International journal of cardiology*, vol. 117, no.1, pp. 97-102, 2007.

[28] X.S. Yang, "Nature-inspired metaheuristic algorithms," Luniver press, 2010.

[29] C. S. Lee and M. H. Wang, "Ontology-based intelligent healthcare agent and its application to respiratory waveform recognition," *Expert System Applications*, vol. 33, no. 3, pp. 606–619, 2007.

[30] M. Fasanghari and G. A. Montazer, "Design and implementation of fuzzy expert system for Tehran Stock Exchange portfolio recommendation," *Expert Systems with Applications*, vol. 37, pp. 6138-6147, 2010.

[31] M. Kalpana, A.V. Senthilkumar, "Fuzzy Expert System for Diagnosis of Diabetes Using Fuzzy Determination Mechanism," *International Journal of Computer Science & Emerging Technologies*, vol., no. 6, 354-361, 2011.

[32] I. Saritas, A. Ozkan, N. Allahverdi, and M. Argindogan, "Determination of the drug dose by fuzzy expert system in treatment of chronic intestine inflammation," *Springer Science+ Business Media J Intell Manuf*, vol. 20, pp. 169-176, 2009.

[33] D.E. Goldberg, "Genetic Algorithms in Search, Optimization, and Machine Learning", *Addison-Wesley, Reading, MA*, 1989.

[34] J. Kennedy, R.C. Eberhart, "Particle swarm optimization," *InProc. 1995 IEEE international conference on neural networks*, pp. 1942–1948, Australia, 1995.

[35] M. Hossin and M. N. Sulaiman, "A review on evaluation metrics for data classification evaluations," *International journal of data mining & knowledge management process*, vol. 5, no. 2, p. 1, 2015.

[36] O. Oladimeji, "Detecting Breast Cancer through Blood Analysis Data using Classification Algorithms," *Journal of Artificial Intelligence & Data Mining (JAIDM)*, vol. 9, no. 3, pp. 351-359, 2021.

یک سیستم فازی هوشمند برای تشخیص بیماری دیابت با استفاده از الگوریتم شاهین هریس

زهرا اصغری ورزنه و سوده حسینی*

گروه علوم کامپیوتر، دانشکده ریاضی و کامپیوتر، دانشگاه شهید باهنر کرمان، کرمان، ایران.

ارسال ۲۰۲۲/۱۰/۳۱؛ بازنگری ۲۰۲۲/۱۲/۰۹؛ پذیرش ۲۰۲۳/۰۳/۲۶

چکیده:

این مقاله یک سیستم خبره فازی برای تشخیص دیابت پیشنهاد می‌کند. در روش پیشنهادی، ابتدا قوانین فازی بر اساس پایگاه داده دیابت هندی‌های پیما (PIDD) تولید می‌شوند و سپس توابع عضویت فازی با استفاده از بهینه‌سازی شاهین هریس (HHO) تنظیم می‌شوند. در مجموعه داده تجربی، ابتدا PIDD با گروه سنی ۳۰-۲۵ سال پردازش شده و مقادیر قطعی در مرحله فازی سازی به مقادیر فازی تبدیل می شوند. سیستم خبره فازی بهبود یافته، دقت طبقه بندی را افزایش می دهد که از چندین روش معروف برای تشخیص بیماری دیابت بهتر عمل می‌کند. الگوریتم HHO برای تنظیم توابع عضویت فازی به منظور تعیین بهترین محدوده برای توابع عضویت فازی و افزایش دقت طبقه‌بندی قوانین فازی اعمال می‌شود. نتایج تجربی از نظر معیارهای دقت (accuracy)، حساسیت (sensitivity) و ویژگی (specificity) ثابت می‌کند که سیستم خبره پیشنهادی توانایی بالاتری نسبت به سایر مدل‌های داده کاوی در تشخیص بیماری دیابت دارد.

کلمات کلیدی: سیستم خبره فازی، الگوریتم شاهین هریس، توابع عضویت، بیماری دیابت.
