JOURNAL OF INFORMATION AND COMMUNICATION TECHNOLOGY
http://e-journal.uum.edu.my/index.php/jict

How to cite this article:
AL-Behadili, H. N. K., & Ku-Mahamud, K. R. (2021). Fuzzy unordered rule using greedy hill climbing feature selection method: An application to diabetes classification. *Journal of Information and Communication Technology, 20*(3), 391-422. https://doi.org/10.32890/jict2021.20.3.5

Fuzzy Unordered Rule Using Greedy Hill Climbing Feature Selection Method: An Application to Diabetes Classification

[1]Hayder Naser Khraibet AL-Behadili & [2]Ku Ruhana Ku-Mahamud
[1]Computer Science Department, Shatt Al-Arab University Collage, Iraq
[2]School of Computing, Universiti Utara Malaysia, Malaysia

hayderkhraibet@sa-uc.edu.iq
ruhana@uum.edu.my

Received: 25/10/2020 Revised: 27/1/2021 Accepted: 1/2/2021 Published: 11/6/2021

ABSTRACT

Diabetes classification is one of the most crucial applications of healthcare diagnosis. Even though various studies have been conducted in this application, the classification problem remains challenging. Fuzzy logic techniques have recently obtained impressive achievements in different application domains, especially medical diagnosis. Fuzzy logic technique is unable to deal with data of a large number of input variables in constructing a classification model. In this research, a fuzzy logic technique using greedy hill climbing feature selection methods was proposed for the classification

of diabetes. A dataset of 520 patients from the Hospital of Sylhet in Bangladesh was used to train and evaluate the proposed classifier. Six classification criteria were considered to authenticate the results of the proposed classifier. Comparative analysis proved the effectiveness of the proposed classifier against Naive Bayes, support vector machine, K-nearest neighbor, decision tree, and multilayer perceptron neural network classifiers. Results of the proposed classifier demonstrated the potential of fuzzy logic in analyzing diabetes patterns in all classification criteria.

Keywords: Data mining, diabetes, feature selection, fuzzy logic, machine learning.

INTRODUCTION

Medical diagnostic is the operation of classifying which condition explains a person's status into a distinct and separate disease (Sisodia & Sisodia, 2018). It is frequently related to medical context being implicit. With technological development, different devices could be used for monitoring and collecting medical data about a specific disease (Vitabile et al., 2019). These data could be used in the future for determining and making medical decisions regarding prognosis and treatment to improve accuracy, reliability, and diagnostic speed.

Diabetes could affect one out of four people over the age of 65 years (Morgan, 2018). According to the World Health Organization, this disease has affected over 246 million people worldwide, and this number is expected to increase to 380 million by 2025 (Durgadevi & Kalpana, 2018). It is common knowledge that some people are unaware of having this disease. Thus, diabetes has been categorized as the fifth deadliest disease in the world, with no coming treatment in sight. Diabetes could be controlled when detected early. By contrast, late diagnosis could lead to potential complications and, over time, may lead to kidney disease, stroke, heart disease, foot problems, eye problems, and nerve damage (Centers for Disease Control and Prevention, 2017). With the rise of artificial intelligence and its continued advent into the healthcare and medical diagnostic sectors, the cases of diabetes and their symptoms are well controlled and detected.

Artificial intelligence or machine intelligence in the medical field refers to the simulation of the intelligence of medical experts in machines or computers programmed to act like experts and imitate their experience for medical diagnosis (David, 2016). Artificial intelligence is widely used to derive interesting patterns from medical data (Sharif et al., 2019; Ngan et al., 1999). Various artificial intelligence methods are used by experts in analyzing diabetes. Different data mining algorithms have been recently used to detect diabetes using the principles of artificial intelligence, machine learning, and statistical methods, such as K-nearest neighbor (KNN) (Saxena et al., 2014), ant colony optimization (ACO) (Ganji & Abadeh, 2011), support vector machine (SVM) (Barakat et al., 2010), artificial neural network (ANN) (Smith et al., 1988), and decision tree (Kaur & Chhabra, 2014). Different studies show the success of these classification algorithms; nevertheless, the classification models created by those classifiers are of a complex mathematical model, which are considered incomprehensible and opaque to humans. This weakness prevents the usage of these classifiers in various real-life domains, whereby both comprehensibility and classification accuracy are needed. According to the authors' knowledge, no comprehensive research has been conducted to combine fuzzy unordered rule algorithm (FURIA) (Hühn & Hüllermeier, 2009) together with greedy hill climbing feature selection method (Venkatesh & Anuradha, 2019) to detect early-stage diabetes. Therefore, the transparency and clarity of FURIA (i.e., fuzzy rules induction) and the feature reduction ability of greedy hill climbing method have improved the performance of classification and detection accuracy of diabetes.

In this paper, a comparative study between popular classification algorithms and the proposed technique is presented. This research aims to provide high classification accuracy and a comprehensive comparative result of different classification techniques in diabetes. The rest of this paper is structured as follows. The related works with different classification techniques in diabetes are described in the Related Works section. Then, the proposed hybrid fuzzy unordered rule using the greedy hill climbing feature selection method is provided. The evaluation of the proposed method used in this paper is explained, followed by the analysis of the results in the Results and Discussion section. The final section concludes the research and provides some possible future research directions.

Related Works

Diabetes has been studied in the literature by various researchers from countless aspects. As a life-threatening disease, early diagnosis of diabetes could save many lives. Artificial intelligence has been widely applied in areas of medical diagnosis and diabetes prediction to ensure an accurate and meaningful result. This involves either pattern recognition or classification system. Fuzzy logic has also been used to handle noisy, irrelative, and ambiguous data to improve the classification performance of many feature selection algorithms.

Fuzzy system constructs large knowledge for diabetes detection. Vieira et al. (2012) proposed a fuzzy extension criterion as a searching strategy to allow more flexible data into fuzzy space, which enables a variety of features to be considered during the optimization process. The extension criteria are used to solve the problem of multi-objective optimization. This problem occurs when a minimum number of features is obtained, which reduces the classification accuracy. Therefore, the research provides new objective functions (i.e., fuzzy) with wide flexibility in solving the problem of multi-objective feature selection. The UCI datasets with a diverse number of features (ranging from 9 to 279) and sample size (ranging between 178 and 699) have been used in the evaluation of the proposed fuzzy objective functions. The proposed fuzzy approach exhibits high classification performance for the majority of the datasets.

Cai et al. (2016) applied a fuzzy criterion in multi-objective unsupervised learning by using hybridized filter-wrapper approach. This method allowed for an active approach to select features from the data and to avert misunderstanding of overlapping features in an unsupervised multi-objective clustered problem. An experiment was carried out using benchmark datasets that showed a superior performance of the proposed method in both number of features and accuracy. Jalali, Nasiri, and Minaei (2009) presented a wrapper-based feature selection method based on consistency measure function and fuzzy logic. The work projected the full dataset into a fuzzy space, and then the consistency measures selected the best feature subset. Evaluation of the proposed method was demonstrated by testing nine datasets from a real-world problem. It showed that all numbers of features were reduced with higher classification performance in five

out of nine datasets. In the remaining datasets, equal classification performance was obtained.

Nosrati and Eftekhari (2014) employed fuzzy similarity measures integrated with multi-objective genetic algorithm (GA) for feature selection and classification problem to find the optimal set of features (subset). The performance of this method was evaluated by using UCI datasets, and the results showed superior performance as compared to correlation-based feature selection methods.

Hedjazi et al. (2010) applied fuzzy logic to solve efficiency and operation safety of sensors in the industrial plant domain. Fuzzy logic was used in the learning algorithm for sensor situation identification. It was employed for feature selection in choosing the optimal number of sensors that were either operational or faulty. The work used centered binomial as a membership function for attribute selection. The results showed that the method had a highly accurate performance.

El-Alfy and Al-Obeidat (2014) developed a multi-criteria fuzzy classifier hybrid with greedy attribute selection method for network anomaly detection. This hybrid technique had a significant impact on the performance of intrusion-detection systems. The proposed hybrid system enhanced the detection rates for different kinds of intrusions and reduced the number of selected attributes to about 74 percent. A swarm intelligence classifier called FCS-ANTMINER for diabetes diagnosis was developed by Ganji and Abadeh (2011). The classifier was based on a combination of fuzzy logic and ACO in order to extract a predictive model. The classification performance was compared with state-of-the-art classifiers used in the literature, and the result for classification was 84.24 percent.

Another classification system combined hybrid ACO and fuzzy logic for diabetes classification. The result showed that the proposed method outperformed the baseline classifiers in terms of classification accuracy. This research also provided a detection expert system for diabetes (Tnv & Gundabathina, 2016). Beloufa and Chikh (2013) presented an artificial swarm intelligence algorithm called modified artificial bee colony (ABC) algorithm for diabetes classification. They used this modified algorithm to create an optimal fuzzy classifier by searching for optimal fuzzy rules and membership functions

simultaneously on the basis of classification accuracy and high readability. The experimental result showed that the proposed fuzzy classifier based on modified ABC algorithm could be a helpful tool for diabetes diagnosis (Beloufa & Chikh, 2013). Jain and Raheja (2015) combined fuzzy verdict mechanism in the fuzzy logic system for diabetes diagnosis. The proposed mechanism was introduced to enhance the result's accuracy by providing a decision on whether the patients were suffering from diabetes or not and to produce a comprehensive result. This research considered urine as an important parameter for diabetes disease. The obtained classification result was promising at 87.02 percent.

Other artificial intelligence methods have been applied in areas of medical diagnosis and diabetes prediction to ensure accurate and meaningful results. Smith et al. (1988) proposed an ANN model for diabetes prediction. ANN is considered one of the best algorithms for data classification problems. The research was conducted on Pima Indians, who were considered a high-risk population of having diabetes. The result of the classification accuracy was 76 percent. The research by Temurtas et al. (2009), which used multilayer ANN to solve the classification problem, produced a 79.62 percent accuracy. Multilayer perceptron (MLP) has been known to contain more than one layer in its structure, which helps in producing a good performance. Kahramanli and Allahverdi (2008) developed a hybrid system consisting of ANN and fuzzy neural network. In the research, two medical datasets were used (Cleveland heart disease and Pima Indian diabetes) to evaluate the performance of this hybridization. The proposed hybrid system enhanced the classification accuracy with 84.24 percent for diabetes. Jaganathan et al. (2007) proposed an improvement on quick reduct as data pre-processing (i.e., reduction methods). Then, they applied swarm ACO algorithms for diabetes prediction. The classification improved to 76.58 percent when compared with the original ACO algorithm. This pre-processing stage improved the quality of data by selecting the most related attributes to be used for classification.

The research by Mei et al. (2017) demonstrated a personalized hypoglycemic medication classification for diabetes. The dataset was from the Electronic Health Records (EHR) repository of China consisting of 21,796 patients with diabetes. This research used hierarchical recurrent neural network and compared the performance

with that of a logistic regression classification algorithm. The result showed that the proposed technique outperformed the logistic regression. In a study by Saxena et al. (2014), KNN classification algorithm was used on a dataset from Stanford University repository, which comprised 11 patient attributes. Different K values were used in the experiment, which showed that the best result of K = 5, with 75 percent prediction accuracy. Three classification algorithms were proposed by Sisodia and Sisodia (2018) for diabetes diagnosis. These algorithms were Naive Bayes, SVM, and decision tree. The highest classification accuracy obtained was 76.30 percent using the Naive Bayes classifier. Perveen et al. (2016) used the same traditional J48 decision tree with AdaBoost ensemble and bagging methods. The database was collected by Canadian Primary Care Sentinel Surveillance Network. The result showed that AdaBoost ensemble with decision tree was better than bagging with decision tree. Another research was conducted on Egyptian patients with diabetes (Karim et al., 2016). This research added a new factor, which was age. This study aimed to classify people who would have the disease or not as an early warning before reaching the critical phase. The decision tree classification algorithm was used with 84 percent accuracy. However, the result was not compared with other classification algorithms. Real data from 100 patients were collected for the prediction of diabetes types in a research conducted by Rahman et al. (2014). A healthcare system was developed, which consisted of AdaBoost method with random committee classifier to supply a service for patients with diabetes. The system produced a result with 81.0 percent accuracy, and the future direction was to add a feedback method on the system to increase the user satisfaction level.

Aishwarya and Anto (2014) proposed a hybridization of GA and extreme learning machine for a medical expert system. GA was used in this system as a feature selection method. The accuracy of the proposed medical system was promising at 89.54 percent as compared to that of the other existing results in previous studies. In another research conducted by Kaur and Kumari (2019), five different classification models for diabetes diagnosis were developed. These models were multifactor dimensionality reduction, KNN, ANN, linear kernel for SVM (SVM-linear), and radial basis function kernel for SVM. The research was conducted using R data tools on the Pima Indian dataset. The best classification result was found in SVM-linear

and KNN. For future research directions, the researchers proposed Boruta wrapper algorithm as the feature selection technique before building the diabetes prediction model. A recent study proposed by Faniqul et al. (2019) considered new factors for diabetes detection at the early stage. The dataset was collected from patients with diabetes at the Hospital of Sylhet in Bangladesh. The data consisted of information regarding 520 patients and included 16 main factors. The result of the classification accuracy was investigated using three classifiers, namely random forest, logistic regression, and Naive Bayes. The result obtained showed that random forest had the best classification accuracy on this dataset. Mishra et al. (2020) proposed a hybrid classification technique that used the variation of GA based on multilayer perceptron called enhanced and adaptive-GA-multilayer perceptron (EAGA-MLP). This work introduced a new fitness function called chromosome swapping and a variation of mutation called Restrics mutate. The proposed method outperformed different variants of MLP, which are: GA-MLP, E-GA-MLP, A-GA-MLP, and EAGA-MLP.

In summary, the existing classification techniques used in the literature have two types, as summarized in Table 1. The first type is considered as a black box (e.g., SVM, ANN, KNN, random forest, logistic regression, Naive Bayes, and random committee). These techniques produce high classification accuracy, but the classification models are not understandable. The other type (i.e., decision tree and fuzzy logic) produces an understandable model with low classification accuracy. This weakness prevents the usage of these classifiers in diabetes, whereby both comprehensibility and classification accuracy are needed. Thus, an extension to decision tree classifier called FURIA was introduced, and it outperformed the other classifiers of the second type (Hühn & Hüllermeier, 2009). In addition, the greedy hill climbing feature selection method has an ability to find the main factors and explain the relationships between these factors. Therefore, this research has introduced a new FURIA with greedy hill climbing feature selection method to diagnose diabetes. Testing was performed with the most related diabetes symptoms dataset. These data were collected from the Hospital of Sylhet in Bangladesh by Faniqul et al. (2019). The dataset was directly collected from patients who are recently diagnosed with diabetes or those who still have not suffered from diabetes but with some symptoms.

Table 1

Related Studies on the Classification of Diabetes with Different Techniques

References	Dataset	Method	Accuracy
Vieira et al. (2012)	Eight UCI datasets	Fuzzy logic	-
Cai et al. (2016)	Six UCI datasets	Fuzzy criterion in multi-objective unsupervised learning	-
Jalali et al. (2009)	Nine real-world problems	Fuzzy logic + Consistency measures	-
Nosrati Nahook and Eftekhari (2014)	Six UCI datasets	Fuzzy similarity measures + Multi-objective genetic algorithm	-
Hedjazi et al. (2010)	Complex industrial plants	Fuzzy logic + Fuzzy based feature selection	-
El-Alfy and Al-Obeidat (2014)	Intrusion detection	Fuzzy logic + Greedy attribute selection	-
Ganji and Abadeh (2011)	Pima Indians	Fuzzy logic + ACO	84.24%
Tnv and Gundabathina (2016)	Private hospital in India	Fuzzy logic + ACO	-
Beloufa and Chikh (2013)	Pima Indians	Fuzzy logic + Modified ABC	84.21%
Jain and Raheja (2015)	Pima Indians	Fuzzy verdict mechanism + Fuzzy logic	87.02%
Smith et al. (1988)	Pima Indians	ANN	76%
Temurtas et al. (2009)	Pima Indians	MLANN	79.62%
Kahramanli and Allahverdi (2008)	Pima Indians	ANN+FNN	84.24%
Jaganathan et al. (2007)	Pima Indians	ACO + Improved quick reduct	76.58%

(continued)

Mei et al. (2017)	EHR repository of China	HRNN	-
Saxena et al. (2014)	Stanford University repository	KNN	75%
Sisodia and Sisodia (2018)	Pima Indians	Naive Bayes SVM Decision tree	76.30%
Perveen et al. (2016)	Canadian Primary Care Sentinel Surveillance Network (CPCSSN)	AdaBoost + Decision tree Bagging + Decision tree	-
Karim et al. (2016)	Egyptian diabetes patients	Decision tree	84%
Rahman et al. (2014)	Dataset collected from a local diabetes hospital in Pakistan	AdaBoost + random committee	81.0%
Aishwarya and Anto (2014)	Pima Indians	GA + ELM	89.54%
Kaur and Kumari (2019)	Pima Indians	1- Multifactor dimensionality reduction 2-KNN 3-ANN 4-SVM-linear 5-SVM-RBF	83% 88% 86% 89% 84%
Faniqul et al. (2019)	Hospital of Sylhet	1-Random forest 2-Logistic regression 3-Naive Bayes	97.4% 92.4% 87.4%
Mishra et al. (2020)	Pima Indians	EAGA-MLP	94.7%

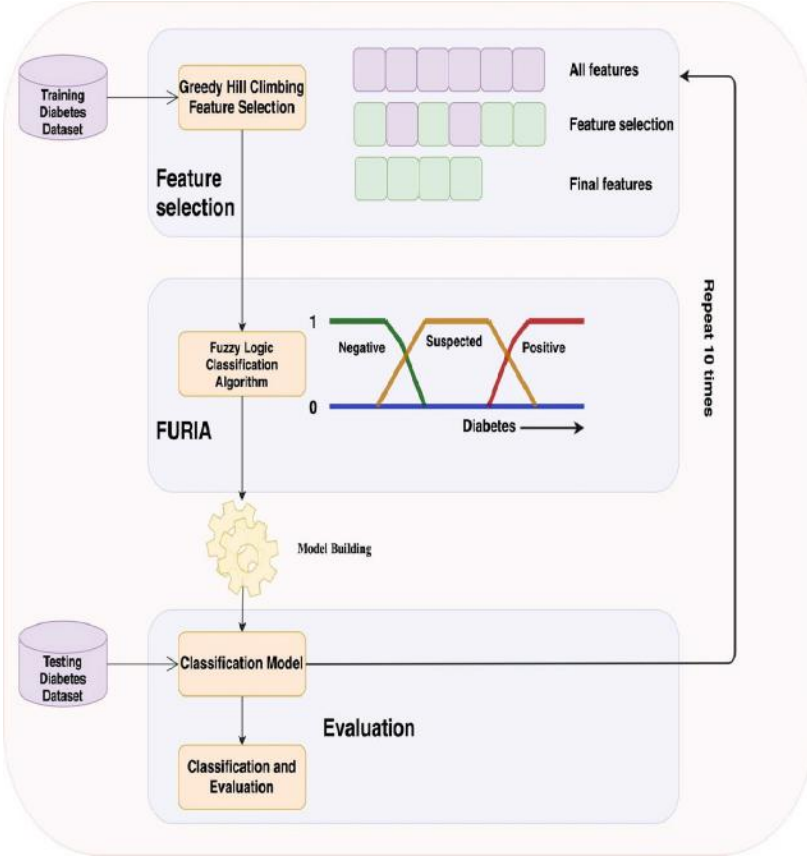
Proposed Hybrid Fuzzy Unordered Rule WITH Greedy Hillclimbing Feature Selection Method

A general framework of the proposed diabetes classification model is presented in Figure 1. The hybrid fuzzy unordered rule and greedy hill climbing feature selection method were used to collect the most

related features from the data for diabetes diagnosis to produce a simple classification model. The two main stages of diabetes diagnosis are: i) feature selection method, in which the greedy hill climbing used a correlation-based filter to eliminate the irrelative features; and ii) classification model, which was responsible for extracting patterns.

Figure 1

General Framework for the Proposed Diabetes Classification Model



The outline of the proposed classifier algorithm, which consists of 24 steps, is shown in Algorithm 1.

Algorithm 1: Greedy Fuzzy Unordered Rule algorithm for classification

```
Input     Original arff dataset
Output    The selected features of arff dataset
Step 1:   While (folds>10)
              // Greedy Hill climbing for Feature Selection starts here
Step 2:       Create set of Attributes {};
Step 3:       Measure Attribute amount of Unpredictability ();
Step 4:       Repeat
Step 5:           Pick out worst attribute ();
Step 6:           Delete worst attribute ();
Step 7:           AttributesSet {}* = AttributesSet {}-Worst attribute;
Step 8:           If fitness (AttributesSet {}* ) > fitness (AttributesSet {});
Step 9:               AttributesSet {} = AttributesSet {}*;
Step 10:          End If
Step 11:      Until Stopping Criteria is met
                  // Greedy Hill climbing for Feature Selection ends here
                  // Fuzzy Unordered Rule starts here
Step 12:      Class Selection ();
Step 13:      Learn classification rules (fuzzification) for selected class;
Step 14:      While (StoppingCriteria == false)
Step 15:          RuleGrowing Method();
Step 16:          if (StoppingCriteria == true)
Step 17:              Delete the newly created rule;
Step 18:          End If
Step 19:      End While
Step 20:      Calculate attribute purity;
Step 21:      Calculate rules confidence degrees;
Step 22:      Evaluate & Apply rule stretching;
Step 23:      Perform rule pruning;
Step 24:  End While
                  // Fuzzy Unordered Rule ends here
```

The first step began by dividing the dataset into ten equal-size parts (folds). In the training stage, nine parts were used, and the remaining

part was used in the testing stage. Two lists were developed in the second step to understand the feature selection method. A list of features is denoted by *Attributes (A) = { a1, a2, a3,..., an}*, which represented various feature values, and another list of correlation values for each features value is denoted as *CorrelationValue (CV) = { υ1, υ2, υ3 ,..., υn}*, which determined the probability value of each attribute. The best attribute subset consisted of attributes highly correlated to the classification class (target class) and uncorrelated with each other. The third step involved measuring the uncertainty and unpredictability in the classification model, and was conducted using entropy. The entropy of *A* attributes is given by Equations 1 and 2 (Venkatesh & Anuradha, 2019):

$$\text{Entropy (A|X)} = -\sum_{x \in X} P(x) \sum_{a \in A} P(a|x)\, log2\,(P(a|x)), \quad (1)$$

$$\text{Entropy (A)} = -\sum_{a \in A} P(a)\, log2\,(P(a)) \quad (2)$$

where the entropy of *A* after observing another variable *X*.

In the fourth step, the algorithm began the iteration, which was terminated using the stopping criteria of either the classifier arriving at the limit number of iterations or no improvement achieved for the predefined number of iterations. The fifth and sixth steps focused on searching for the worst attribute in the neighborhood and removing it to create a different subset that would be based on the current subset. In the seventh step, a new list of features is denoted as *AttributesSet {}**, which represented the set of features after deleting the selected attribute. The eighth step involved checking if the deletion of the attribute increased the quality of the current subset (i.e., attribute fitness). In the ninth step, the new subset was saved as the current subset if and only if the quality increased, and the stage proceeded to the tenth step.

The fuzzy unordered rule classification stage initiated on the 12th step. Fuzzy unordered rule algorithm adopted the separate-and-conquer approach and was a modified version of the well-known repeated

incremental pruning to produce an error reduction (RIPPER) algorithm (Cohen, 1995). The RIPPER algorithm, used to learn patterns for rule construction, is an extension of the incremental reduced error pruning (IREP) rule induction classifier (Mohamed et al., 2012). RIPPER enhanced IREP in many aspects, such as it is able to deal with multinomial classification problems (Hülhn & Hüllermeier, 2010). In the FURIA classifier, the order of the classes was not applicable, indicating that the default rule used in the majority classifiers was irrelevant. In the 12th step, FURIA selected a specific class to create a fuzzy rule to it in accordance with a list of covered training instances. In the 13th step, fuzzification was conducted to create fuzzy rule by using the fuzzy logic concept. FURIA calculated the best fuzzification of membership in terms of purity. In addition, FURIA selected terms from the data to be added to the rule in accordance with maximizing the information gain (*IG*) criterion, which measured the improvement of the rule performance for the specific class. *IG* was measured according to the following Equation 3 (Hülhn & Hüllermeier, 2010):

$$IG = P_i \times \left(\log_2\left(\frac{P_i}{P_i+N_i}\right) - \log_2\left(\frac{P}{P+N}\right)\right) \quad (3)$$

where and represent the number of negative and positive instances, respectively, covered by the construction rule. In the same way, *P* and *N* represent the number of negative and positive instances, respectively, covered by the default rule.

The 14th–19th steps involved the main FURIA algorithm loop (while) with stopping criteria of uncovered instances by specific fuzzy rules. These steps included checking if the rule error was more than or equal to 0.5, then the stopping criteria = true. If so, the algorithm would delete the newly created rule. The 20th step also involved fuzzification for realizing the largest purity antecedent. Fuzzification was repeated until all antecedents were fuzzified. Fuzzification played a crucial role in classifying new instances even though the purity on the training data did not vary in the fuzzification step. The relevant instance of the antecedent may change in this step. Thus, recalculation was conducted in each iteration. In the 21st step, all rules were evaluated by computing the confidence degrees on the basis of a certainty factor.

The rule stretching method was performed in the 22nd step to exploit the antecedents of the rule. This method consisted of two factors. For the first factor, each rule was treated as a list of antecedents <*α1,α2,....,αm*> instead of a set of antecedents {*α1,α2,....,αm*}. This method aimed to reflect the most important antecedents in the rule. In general, the list of antecedents is denoted as <*α1,α2,...,αk*>, where $k \leq m$, k is represented as $j - 1$, and αj is the antecedent not satisfied by the query instance. The evaluation was measured using Equation 4 (Hühn & Hüllermeier, 2009):

$$\frac{Tp + 1}{Tp + N + 2} \times \frac{K + 1}{m + 2}, \tag{4}$$

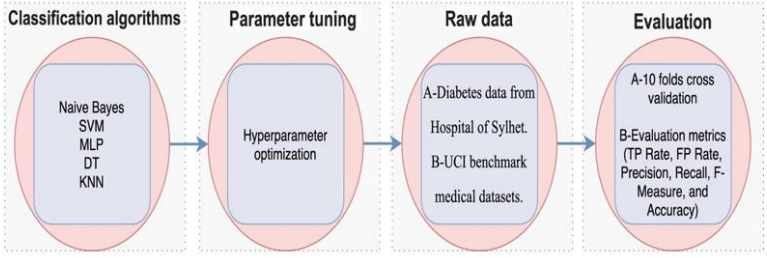
where *Tp* is the number of true positive instances, and *N* the number of true negative instances covered by the fuzzy rule. The second factor was rule stretching, which considered the degree of generalization: too short rules were deducted, as removal of the antecedents decreased the rule's relevance. Laplace Correction was responsible for determining the quality of the rule on the basis of the antecedents' number, and preference was given to longer rules. In the 23rd step, the pruning method was applied to remove all unwanted antecedents from a rule to create the replacement, except in the situation wherein the pruning method would delete all antecedents from the rule, thereby creating a default rule to cover the set of instances.

EVALUATION OF THE PROPOSED METHOD

This section describes the research methodology of this study. The state-of-the-art diabetes classification algorithms are explained, followed by description of the parameters and their values used in the algorithms. Lastly, the datasets and their characteristic are described. Figure 2 shows the main steps to develop a classification model.

Figure 2

Essential Steps in the Research Methodology to Develop a Classification Model



State-of-the-art Diabetes Classification Algorithm

Pattern data classification was utilized to group each case of data into one of the predefined sets of classes. Classification is a data mining task that accurately classifies each item of a target class. The most famous classification algorithms used for diabetes predication will be explained in this study.

The Naive Bayes classification algorithm, a statistical classification method based on the Bayes theorem, has the ability to classify the probability of membership of the target class (Faniqul et al., 2019; Irfan et al., 2018). Support vector machine (SVM) is a supervised classification algorithm that analyzes a set of input cases and generates non-probabilistic binary linear models. SVM uses input cases to recognize patterns that could predict a target class (Aydin et al., 2011; Barakat et al., 2010). Multilayer perceptron neural network (MLPNN) is a feedforward type of ANN. It is a statistical nonlinear data modeling method. This method has the ability to find patterns in the data by dealing with complicated relationships between input and output (Temurtas et al., 2009; Tkáč & Verner, 2015).

Decision tree is a popular data mining method that classifies the target class on the basis of multiple input variables. The classifier builds a decision tree model in accordance with the internal node that is equivalent to input variable, with each possible input variable value

at the edge of the decision tree nodes. Decision tree consists of a leaf representing the value of target variables, and the decision tree pattern is based on the path from the root to leaf (Hssina et al., 2014; Perveen et al., 2016). KNN classifies the data cases based on similar cases that were previously classified. It decides the target class of new given data by examining the KNN of the most equal neighbors and assigns the same class (Kaur & Kumari, 2019; Liao & Vemuri, 2002).

Common Parameters

In all classifiers successfully used to classify diabetes, the values of the input parameters of the classification algorithms are shown in Table 2.

Table 2

Parameters of Classifiers

Classifiers	Parameter Value							
Naive Bayes	batchSize	numDecimalPlaces		Test mode				
	100	2		10-folds				
SVM	batchSize	epsilon	Kernel type	toleranceparameter		numDecimalPlaces		Test mode
	100	1.0	Polynomial Kernel	0.001		2		10-folds
MLP	batchSize	numDecimalPlaces		learningRate		NumOfEpochs	hidden layer	Test mode
	100	2		0.3		500	1	10-folds
Decision tree	batchSize	numDecimalPlaces		ConfidenceFactor		Minimum number of instances per leaf		Test mode
	100	2		0.25		2		10-folds
KNN	batchSize	numDecimalPlaces		K	SearchAlgorithm	Distance Function		Test mode
	100	2		3	LinearNNsearch	Euclidean distance		10-folds

Dataset Details

The diabetes dataset was obtained from patient information at the Hospital of Sylhet in Bangladesh. This dataset has been used to investigate patients with early-stage diabetes in accordance with the World Health Organization criteria. The dataset consisted of 520 cases with 16 attributes (i.e., features) and was divided into positive and negative classes. However, the target for this research was to identify the main attributes (factors) expected to be highly related with the occult development of diabetes. The details of the dataset are shown in Table 3, which included the name of feature, number of values for each attribute, number of values in each class, and the type of attributes.

Table 3

Characteristics of Diabetes Data

No.	Name of Feature	Attribute Value	Type of Attributes
1	Age	6	Discrete
2	Gender	2	Discrete
3	Genital thrush	2	Discrete
4	Polydipsia	2	Discrete
5	Obesity	2	Discrete
6	Partial paresis	2	Discrete
7	Polyphagia	2	Discrete
8	Polyuria	2	Discrete
9	Visual blurring	2	Discrete
10	Delayed healing	2	Discrete
11	Irritability	2	Discrete
12	Itching	2	Discrete
13	Weakness	2	Discrete
14	Muscle stiffness	2	Discrete
15	Alopecia	2	Discrete
16	Sudden weight loss	2	Discrete
17	Class	2	Discrete

Other experiments were performed using nine UCI small and medium size medical benchmark datasets in ARFF Weka's format to test the

performance of the proposed classifier. These datasets are popular medical datasets in the literature, and they have different attribute numbers, which lie between 9 and 19. The datasets exhibited binary and multi-label classes. They also have different instance numbers within the range of 106–768. The main characteristics of the medical benchmark datasets are listed in Table 4.

Table 4

Description of Medical Datasets

Name of dataset	Description		
	Number of Attributes	Number of Instances	Number of Classes
BreastCancer	9	286	2
BreastTissue	10	106	6
ClevelandHeartDisease	13	303	5
HeartStatlog	13	270	2
Hepatitis	19	155	2
Lymphography	18	148	4
WisconsinBreastCancer	9	699	2
Diabetes	16	520	2
Pima Indians	9	768	2

Results and Discussion

This section evaluated the performance of the proposed classification algorithm. In the first step, it presented the experimental methods, the evaluation criteria used in the experiments, and the majority of diabetes classification algorithms. In the second step, experiments were conducted to obtain connections between the values of the input parameters for the proposed technique and achieve good classification results. Finally, the results were compared with those of baseline classification algorithms in diabetes diagnosis.

The well-known k-fold cross-validation test was used to partition the dataset into ten parts (Al-behadili et al., 2020; Gupta et al., 2016; Hairuddin et al., 2020) In each test, one part was used for the testing set, while the rest was used for the training set. The test was repeated ten times, with different datasets for testing each time. The average classification performance of the ten runs indicated the performance of the classifier. For more robust analysis results, the evaluation was measured by the combination of facts and values of six evaluation criteria. These criteria were true positive (TP) rate, false positive (FP) rate, precision, recall, F-measure, and accuracy. In addition, the input parameters of the algorithm are as follows:

- Fold's, responsible for instance numbers used for pruning the fuzzy rule.
- MinNo, the minimum total weight of the instances covered by each discovered rule.
- Optimizations, the number of optimizations run.
- Uncovered instance methods, the methods that deal with the uncovered instances.

The fold's parameter was responsible for determining the instance number used for pruning the fuzzy rule. In accordance with the experimental result, the fold's parameter could influence the classification accuracy at a maximum percentage of 1 percent. Figure 3 represents the values of the folds. The best value was obtained when the number of folds was medium.

Figure 4 presents the results obtained from the classification accuracy's point-of-view if the minimum total weight of the instances (i.e., min No) in a rule were 1, 2, 3, 4, 5, 6, 7, 8, 9, or 10 and the fold's parameter as kept at fixed value = 5. The best classification result was selected as minNo = 2.

Figure 3

Instance Numbers Used for Pruning the Fuzzy Rule

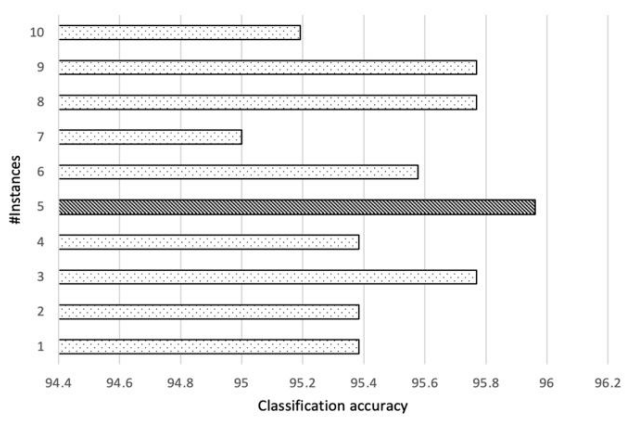


Figure 4

Minimum Total Weight of the Instances in a Rule

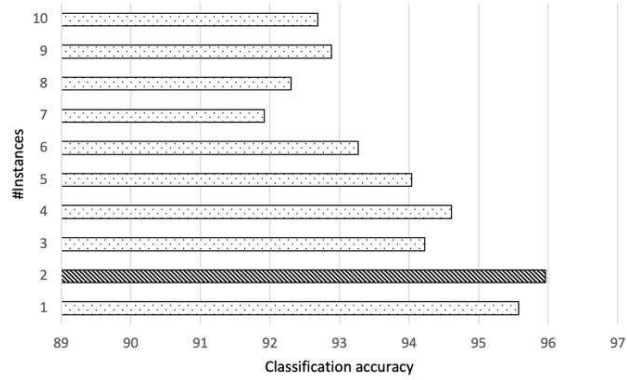


Figure 5 depicts the number of optimizations run on this parameter, which was mostly responsible for the learning process and determined the growth of the execution times in the learning process. In this experiment, the fold's parameter was kept at fixed value = 5 and minNo = 2. The best classification result was obtained with the

number of optimizations = 4.

Figure 6 portrays the classification result based on three methods performed for uncovered instances. These methods were rule stretching, vote for the most frequent class, and reject the decision and abstain. In this experiment, the fold, minNo, and optimization parameters were kept at fixed values of 5, 2, and 4, respectively. The best result was obtained with voting for the most frequent class. These methods could influence the classification accuracy at a maximum percentage of 1.73 percent.

Figure 5

Number of Optimizations Run

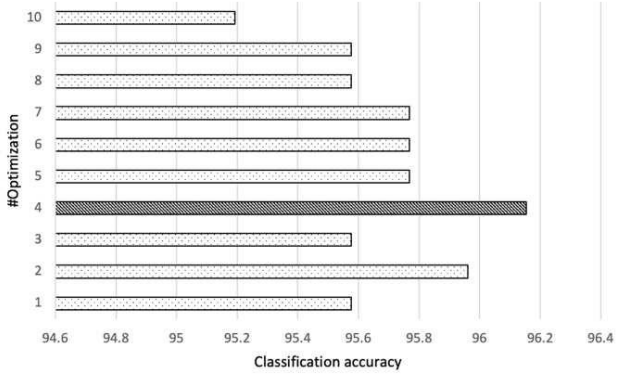
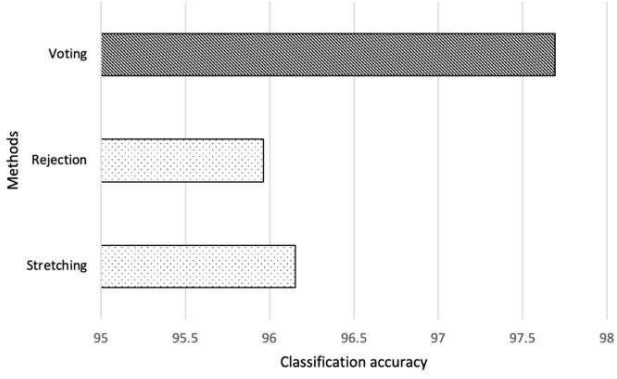


Figure 6

Methods Performed for Uncovered Instances



This section compares the results of the hybrid classifier with those of the state-of-the-art classification algorithms. These classifiers included Naive Bayes, SVM, MLPNN, KNN, and decision tree. An experiment on early-warning diabetes at the Hospital of Sylhet in Bangladesh was conducted for all classification algorithms. In the first evaluation stage, Table 5 indicates the experimental results of the average prediction performance in 10-fold cross-validation. For each criterion, the best result was written in bold. Table 5 displays that the proposed method was the best among all other classifiers in all evaluation criteria. Best classification accuracies are as highlighted.

Table 5

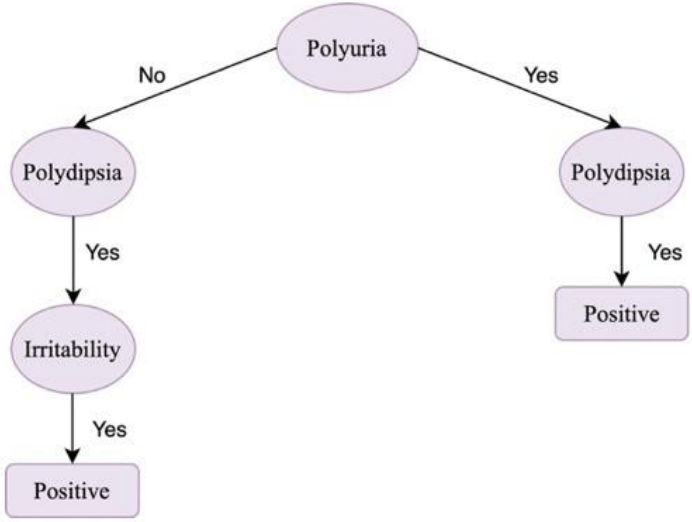
Classifier Performance for Diabetes Classification

Classifiers	Class	TP Rate	FP Rate	Precision	Recall	F-Measure	Accuracy
Proposed Technique	Positive	0.988	0.040	0.975	0.988	0.981	
	Negative	0.960	0.013	0.980	0.960	0.970	**97.692**
	Average	**0.977**	**0.028**	**0.977**	**0.977**	**0.977**	
Naive Bayes	Positive	0.863	0.100	0.932	0.863	0.896	
	Negative	0.900	0.138	0.804	0.900	0.849	87.692
	Average	0.877	0.114	0.883	0.877	0.878	
SVM	Positive	0.938	0.100	0.938	0.938	0.938	
	Negative	0.900	0.063	0.900	0.900	0.900	92.307
	Average	0.923	0.086	0.923	0.923	0.923	
MLPNN	Positive	0.975	0.030	0.981	0.975	0.978	
	Negative	0.970	0.025	0.960	0.970	0.965	97.307
	Average	0.973	**0.028**	0.973	0.973	0.973	
KNN	Positive	0.922	0.040	0.974	0.922	0.947	
	Negative	0.960	0.078	0.885	0.960	0.921	93.653
	Average	0.937	0.055	0.939	0.937	0.937	
Decision tree	Positive	0.922	0.040	0.974	0.922	0.947	
	Negative	0.960	0.078	0.885	0.960	0.921	95.384
	Average	0.954	0.044	0.955	0.954	0.954	

Furthermore, the proposed method found 11 significant factors out of 16 factors that were available in the dataset. These factors were age, gender, polyuria, polydipsia, sudden weight loss, polyphagia, itching, irritability, delayed healing, muscle stiffness, and alopecia. The best results and discovered factors were achieved by using greedy hill climbing feature selection method. It used the power of information theory (i.e., entropy) to find the most effective factors for diabetes. In addition, the proposed classification model provided the highest TP rate, precision, recall, F-measure, and classification accuracy for all evaluation measurements. It also produced the lowest FP rate. Thus, the proposed technique was more suitable for diabetes diagnosis than the other classifiers. In addition, Figure 7 shows the fuzzy rules that were obtained by using this experiment for the purpose of diabetes detection.

Figure 7

Example of Fuzzy Classification Model Constructed by the Proposed Method



The performance of the commonly used classification techniques on other medical benchmarking datasets was checked to compare the results of the proposed hybrid classifier. Table 6 shows that the proposed hybrid classifier was better than MLPNN, KNN, and

decision tree in all datasets. The proposed hybrid classifier was better than SVM in seven datasets. Furthermore, the proposed classifier outperformed Naive Bayes in six out of nine medical datasets. In comparison with all classifiers, the proposed hybrid classifier achieved the highest results in five datasets. It also acquired the second-best performance in four datasets. Meanwhile, the second-best classifier was Naive Bayes, with three datasets. SVM achieved the best results in two datasets. MLPNN, KNN, and decision tree classifiers obtained the lowest results in all datasets.

Looking at the overall average ranking of classification accuracy in Table 6, the proposed classifier performed the best, as expected from the usage of the greedy hill climbing feature selection method together with the fuzzy logic that chose the most related features, which increased the classification accuracy.

Table 6

Classifier Performance for Medical Datasets

Dataset		Proposed Technique	Naive Bayes	SVM	MLPNN	KNN	Decision tree
Breast-Cancer	Accuracy	**75.174**	73.426	70.629	68.181	73.076	73.426
	Rank	**1**	2.5	5	6	4	2.5
BreastTissue	Accuracy	**97.169**	94.339	64.15	94.339	91.509	95.283
	Rank	**1**	3.5	6	3.5	5	2
Cleveland-HeartDisease	Accuracy	**83.828**	83.498	82.508	80.858	79.538	78.217
	Rank	**1**	2	3	4	5	6
HeartStatlog	Accuracy	82.963	**83.333**	82.222	81.111	81.481	81.851
	Rank	2	**1**	3	6	5	4
Hepatitis	Accuracy	85.161	84.516	**87.096**	84.516	82.58	81.935
	Rank	2	3.5	**1**	3.5	5	6

(continued)

Lymphography	Accuracy	83.108	**83.783**	**83.783**	81.081	82.432	78.378
	Rank	3	**1.5**	**1.5**	5	4	6
WisconsinBreastCancer	Accuracy	96.423	**96.995**	96.28	95.708	96.137	95.851
	Rank	2	**1**	3	6	4	5
Diabetes	Accuracy	**97.692**	87.692	92.307	97.307	93.653	95.384
	Rank	**1**	6	5	2	4	3
Pima Indians	Accuracy	**89.22**	75.26	67.18	74.21	70.31	74.34
	Rank	**1**	2	6	4	5	3
Average Rank		**1.5**	2.9	3.8	4.2	4.4	4.2

CONCLUSION

Diabetes detection is a serious medical problem in the real world. Therefore, early-stage detection of diabetes plays an important role in treatment. In this research, a new hybrid classifier was proposed for diabetes classification on the basis of the feature selection method. The proposed classifier comprised two main steps: a) greedy hill climbing feature selection method; and b) generation of a fuzzy unordered rule list for diabetes pattern classification. The simulation results of the proposed hybrid classifier could deal with diabetes diagnosis problems at an early stage. The proposed classifier also showed the ability to select good features that improved the classification accuracy. In addition, it demonstrated the importance of feature selection in diabetes detection and showed that the performance of the classification became better after taking more consideration for this method. For validation purposes, the proposed classifier outperformed the other state-of-the-art classifiers, such as Naive Bayes, SVM, MLPNN, KNN, and decision tree. For future research, the usage of stochastic local search algorithms (i.e., variable neighborhood search, guided local search, and iterated local search) together with swarm intelligence algorithms (i.e., particle swarm optimization, ACO, and ABC) could be implemented to improve the feature selection method, increase the classification rate, and reduce the error rate. Another research direction will be to focus on collecting extra information to discover new potential elements to be integrated.

ACKNOWLEDGMENT

The researchers would like to thank the Malaysian Ministry of Higher Education for financially supporting this research under Grant Scheme (TRGS /1/2018/UUM/02/3/3 (S/O code14163).

REFERENCES

Aishwarya, S. S., & Anto, S. (2014). A medical expert system based on genetic algorithm and extreme learning machine for diabetes disease diagnosis. *International Journal of Science, Engineering and Technology Research (IJSETR)*, *3*(5), 1375–1380.

Al-behadili, H. N. K., Ku-Mahamud, K. R., & Sagban, R. (2020). Hybrid ant colony optimization and iterated local search for rules-based classification. *Journal of Theoretical and Applied Information Technology*, *98*(04), 657–671.

Al-Behadili, H. N. K., Sagban, R., & Ku-Mahamud, K. R. (2020). Adaptive parameter control strategy for ant-miner classification algorithm. *Indonesian Journal of Electrical Engineering and Informatics (IJEEI)*, *8*(1), 149–162. https://doi.org/10.11591/ijeei.v8i1.1423

Aydin, I., Karakose, M., & Akin, E. (2011). A multi-objective artificial immune algorithm for parameter optimization in support vector machine. *Applied Soft Computing Journal*, *11*(1), 120–129. https://doi.org/10.1016/j.asoc.2009.11.003

Barakat, N., Bradley, A. P., & Barakat, M. N. H. (2010). Intelligible support vector machines for diagnosis of diabetes mellitus. *IEEE Transactions on Information Technology in Biomedicine*, *14*(4), 1114–1120. https://doi.org/10.1109/TITB.2009.2039485

Beloufa, F., & Chikh. (2013). Design of fuzzy classifier for diabetes disease using modified artificial bee colony algorithm. *Computer Methods and Programs in Biomedicine*, *112*(1), 92–103. https://doi.org/10.1016/j.cmpb.2013.07.009

Cai, F., Wang, H., Tang, X., Emmerich, M., & Verbeek, F. J. (2016). Fuzzy criteria in multi-objective feature selection for unsupervised learning. *Procedia Computer Science*, *102*(August), 51–58. https://doi.org/10.1016/j.procs.2016.09.369

Centers for Disease Control and Prevention, U. D. of H. and H. S. (2017). National diabetes statistics report, 2017. Estimates of diabetes and its burden in the United States background. *Division of Diabetes Translation*. https://doi.org/10.2196/jmir.9515

Cohen, W. (1995). Fast effective rule induction. In *Machine Learning Proceedings*, *2435* (pp. 115–123).

David D. Luxton. (2016). *An introduction to artificial intelligence in behavioral and mental health care.* Academic Press.

Durgadevi, M., & Kalpana, R. (2018). Performance analysis of classification approaches for the prediction of type II diabetes. In *2017 9th International Conference on Advanced Computing, ICoAC 2017* (pp. 339–344). https://doi.org/10.1109/ICoAC.2017.8441197

El-Alfy, E. S., & Al-Obeidat, F. (2014). A multicriterion fuzzy classification method with greedy attribute selection for anomaly-based intrusion detection. *Procedia Computer Science*, *34*, 55–62. https://doi.org/10.1016/j.procs.2014.07.037

Faniqul, I., Rahatara, F., Sadikur, R., & Humayra, B. (2019). Likelihood prediction of diabetes at early stage using data mining techniques. In *International Symposium, ISCMM 2019* (p. 154). Springer. https://doi.org/10.1007/978-981-13-8798-2_12

Ganji, M. F., & Abadeh, M. S. (2011). A fuzzy classification system based on ant colony optimization for diabetes disease diagnosis. *Expert Systems with Applications*, *38*(12), 14650–14659. https://doi.org/10.1016/j.eswa.2011.05.018

Gupta, A., Mohammad, A., Syed, A., & N., M. (2016). A comparative study of classification algorithms using data mining: Crime and Accidents in Denver City the USA. *International Journal of Advanced Computer Science and Applications*, *7*(7), 374–381. https://doi.org/10.14569/ijacsa.2016.070753

Hairuddin, N., Yusuf, L., & Othman, M. (2020). Gender classification on skeletal remains: Efficiency of metaheuristic algorithm method and optimized back propagation neural network. *Journal of Information and Communication Technology*, *19*(2), 251–277. https://doi.org/10.32890/jict2020.19.2.5

Hedjazi, L., Kempowsky-Hamon, T., Despènes, L., Le Lann, M. V., Elgue, S., & Aguilar-Martin, J. (2010). Sensor placement and fault detection using an efficient fuzzy feature selection

approach. In *Proceedings of the IEEE Conference on Decision and Control* (pp. 6827–6832). https://doi.org/10.1109/CDC.2010.5717254

Hssina, B., Merbouha, A., Ezzikouri, H., & Erritali, M. (2014). A comparative study of decision tree ID3 and C4.5. *International Journal of Advanced Computer Science and Applications*, *2*, 13–19.

Hühn, J., & Hüllermeier, E. (2009). FURIA: An algorithm for unordered fuzzy rule induction. *Data Mining and Knowledge Discovery*, *19*(3), 293–319. https://doi.org/10.1007/s10618-009-0131-8

Hülhn, C., & Hüllermeier, E. (2010). An analysis of the FURIA algorithm for fuzzy rule induction. *Studies in Computational Intelligence*, *262*, 321–344. https://doi.org/10.1007/978-3-642-05177-7_16

Irfan, M., Uriawan, W., Kurahman, O. T., Ramdhani, M. A., & Dahlia, I. A. (2018). Comparison of Naive Bayes and K-Nearest Neighbor methods to predict divorce issues. In *IOP Conference Series: Materials Science and Engineering* (Vol. 434, No. 1, p. 012047). https://doi.org/10.1088/1757-899X/434/1/012047

Jaganathan, P., Thangavel, K., Pethalakshmi, A., & Karnan, M. (2007). Classification rule discovery with ant colony optimization and improved quick reduct algorithm. *IAENG International Journal of Computer Science*, *February*, 286–291. http://www.scopus.com/inward/record.url?eid=2-s2.0-84888273942&partnerID=tZOtx3y1

Jain, V., & Raheja, S. (2015). Improving the prediction rate of diabetes using fuzzy expert system. *International Journal of Information Technology and Computer Science*, *7*(10), 84–91. https://doi.org/10.5815/ijitcs.2015.10.10

Jalali, L., Nasiri, M., & Minaei, B. (2009). A hybrid feature selection method based on fuzzy feature selection and consistency measures. In *Proceedings - 2009 IEEE International Conference on Intelligent Computing and Intelligent Systems, ICIS 2009* (Vol. 1, pp. 718–722). https://doi.org/10.1109/ICICISYS.2009.5358395

Kahramanli, H., & Allahverdi, N. (2008). Design of a hybrid system for the diabetes and heart diseases. *Expert Systems with Applications*, *35*(1–2), 82–89. https://doi.org/10.1016/j.eswa.2007.06.004

Karim, O., Yasser, K., & Thanaa, R. (2016). Early predictive system for diabetes mellitus disease. In *Industrial Conference on Data Mining* (pp. 420–427). https://doi.org/10.1007/978-3-319-41561-1

Kaur, G., & Chhabra, A. (2014). Improved J48 classification algorithm for the prediction of diabetes. *International Journal of Computer Applications*, *98*(22), 13–17. https://doi.org/10.5120/17314-7433

Kaur, H., & Kumari, V. (2019). Predictive modelling and analytics for diabetes using a machine learning approach. *Applied Computing and Informatics*, *February 2019*. https://doi.org/10.1016/j.aci.2018.12.004

Liao, Y., & Vemuri, V. R. (2002). Use of k-nearest neighbor classifier for intrusion detection. *Computers and Security*, *21*(5), 439–448. https://doi.org/10.1016/S0167-4048(02)00514-X

Mei, J., Zhao, S., Jin, F., Zhang, L., Liu, H., Li, X., Xie, G., Li, X., & Xu, M. (2017). Deep diabetologist: Learning to prescribe hypoglycemic medications with recurrent neural networks. *Studies in Health Technology and Informatics*, *245*, 1277. https://doi.org/10.3233/978-1-61499-830-3-1277

Mishra, S., Tripathy, H. K., Mallick, P. K., Akash, K. B., & Paolo, B. (2020). EAGA-MLP – An enhanced and adaptive hybrid classification model for diabetes diagnosis. *Sensors*, *20*(14), 1–31. https://doi.org/10.3390/s20144036

Mohamed, W. N. H. W., Salleh, M. N. M., & Omar, A. H. (2012). A comparative study of reduced error pruning method in decision tree algorithms. In *Proceedings - 2012 IEEE International Conference on Control System, Computing and Engineering, ICCSCE 2012* (pp. 392–397). https://doi.org/10.1109/ICCSCE.2012.6487177

Morgan, N. (2018). Diabetic Neuropathy. In *University of North Dakota*. https://doi.org/10.1177/004947550203200403

Ngan, P. S., Wong, M. L., Lam, W., Leung, K. S., & Cheng, J. C. Y. (1999). Medical data mining using evolutionary computation. *Artificial Intelligence in Medicine*, *16*(1), 73–96. https://doi.org/10.1016/S0933-3657(98)00065-7

Nosrati Nahook, H., & Eftekhari, M. (2014). A new method for feature selection based on fuzzy similarity measures using multi objective genetic algorithm. *Journal of Fuzzy Set Valued Analysis*, *2014*, 1–12. https://doi.org/10.5899/2014/jfsva-00162

Perveen, S., Shahbaz, M., Guergachi, A., & Keshavjee, K. (2016). Performance analysis of data mining classification techniques to predict diabetes. *Procedia Computer Science*, *82*, 115–121. https://doi.org/10.1016/j.procs.2016.04.016

Rahman, A., Muhammad, S., Muhammad, I., & Byeong, K. (2014). Prediction of diabetes mellitus based on boosting ensemble modeling. In *International Conference on Ubiquitous Computing and Ambient Intelligence* (pp. 1–8). https://doi.org/10.1007/978-3-319-13102-3

Saxena, K., Khan, Z., & Singh, S. (2014). Diagnosis of diabetes mellitus using K Nearest Neighbor algorithm. *International Journal of Computer Science Trends and Technology (IJCST)*, *2*(4), 36–43.

Sharif, N., Ahmad, N., Ahmad, N., Desa Mat, W., Helmy, K., Ang, W., & Abidin, I. (2019). A fuzzy rule-based expert system for asthma severity identification in emergency department. *Journal of Information and Communication Technology*, *18*(4), 415–438. https://doi.org/10.32890/jict2019.18.4.2

Sisodia, D., & Sisodia, D. S. (2018). Prediction of diabetes using classification algorithms. *Procedia Computer Science*, *132*(Iccids), 1578–1585. https://doi.org/10.1016/j.procs.2018.05.122

Smith, J. W., Everhart, J. E., Dickson, W. C., Knowler, W. C., & Johannes, R. S. (1988). Using the ADAP learning algorithm to forecast the onset of diabetes mellitus. In *Proceedings - Annual Symposium on Computer Applications in Medical Care* (pp. 261–265).

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. https://doi.org/10.1016/j.eswa.2008.10.032

Tkáč, M., & Verner, R. (2015). Artificial neural networks in business: Two decades of research. In *Applied Soft Computing*, *38*, 788–804. Elsevier B.V. https://doi.org/10.1016/j.asoc.2015.09.040

Tnv, M., & Gundabathina, J. (2016). Fuzzy classification rules generation with ant colony optimization for diabetes diagnosis. *International Journal of Emerging Trends & Technology in Computer Science (IJETTCS)*, *5*(5), 39–44.

Venkatesh, B., & Anuradha, J. (2019). A review of feature selection and its methods. *Cybernetics and Information Technologies*, *19*(1), 3–26. https://doi.org/10.2478/CAIT-2019-0001

Vieira, S., Sousa, J., & Kaymak, U. (2012). Fuzzy criteria for feature selection. *Fuzzy Sets and Systems*, *189*(1), 1–18. https://doi.org/10.1016/j.fss.2011.09.009

Vitabile, S., Marks, M., Stojanovic, D., Pllana, S., Molina, J. M., Krzyszton, M., ... & Ilic, A. S. (2019). Medical data processing and analysis for remote health and activities monitoring. In *High-Performance Modelling and Simulation for Big Data Applications* (pp. 186–220). https://doi.org/10.1007/978-3-030-16272-6