Arabian Journal for Science and Engineering (2019) 44:3781–3800
https://doi.org/10.1007/s13369-018-3670-8

RESEARCH ARTICLE - COMPUTER ENGINEERING AND COMPUTER SCIENCE

A Fibrosis Diagnosis Clinical Decision Support System Using Fuzzy Knowledge

Sara Sweidan[1] · Shaker El-Sappagh[2,3] · Hazem El-Bakry[1] · Sahar Sabbeh[2,4] · Farid A. Badria[5] · Kyung-Sup Kwak[3]

Received: 10 July 2018 / Accepted: 28 November 2018 / Published online: 13 December 2018

Abstract

Liver cirrhosis, the end stage of chronic liver disease, is one of the major risk factors for the development of liver cancer, and may result in premature death. This research proposes a fuzzy fibrosis decision support (F2DS) system. It is a fuzzy knowledge-based expert system for liver fibrosis stage prediction. F2DS is carefully based on a set of knowledge acquisition and machine learning techniques. In addition, the system depends on domain expert knowledge for designing the membership functions and validating the fuzzy knowledge base. It depends on a suitable list of 17 symptoms, and laboratory test features that can accurately and significantly describe fibrosis patients. The experimental results of the expert system were obtained using a real dataset from the Liver Institute, Mansoura University, Egypt, of 119 patients infected by chronic viral hepatitis C. The performance of the system was evaluated with many metrics, achieving a testing accuracy of 95.7%. The evaluation of proposed fuzzy expert system shows its capability of diagnosing the stages of liver fibrosis with a high degree of accuracy, and it can be embedded as a component in a healthcare system to assist physicians in their daily practice. In addition, students training in medicine can benefit from this system.

Keywords Clinical decision support · Fuzzy decision tree · Fuzzy logic · Machine learning · Liver fibrosis

1 Introduction

The hepatitis C virus (HCV) is the main reason of liver diseases including hepatitis C and liver cancer. HCV currently infects approximately 2% of the world's population [1]. In most countries, the prevalence of HCV is between 1 and 3%. However, the situation is more critical in Egypt, because it has the highest prevalence of HCV in the world. HCV reached 14.7% in 2008 [1,2] and 10.0% in 2015 [3]. In addition, in Egypt about 80% of the infected individuals become chronic carriers of the virus (i.e., 11 million anti-HCV-positive persons). Moreover, approximately 10% to 20% of patients are likely to develop significant liver fibrosis and cirrhosis within 20–30 years [4]. The liver biopsy remains the current standard for assessment and diagnosis of fibrosis and cirrhosis [2]. This is a surgical procedure accompanied by sampling errors up to 30% of all biopsies [5]. There are viral logical features of HCV used as indicators to predict the stages of histological liver damage such as serum alanine aminotransferase/aspartate aminotransferase (ALT/AST) levels, direct and total serum bilirubin, albumin, platelet count, and international normalized ratio (INR) of blood clotting. On the other hand, some noninvasive methods (AST/ALT, APRI,

✉ Sara Sweidan
sweera20@yahoo.com

Shaker El-Sappagh
shaker_elsapagh@yahoo.com

Hazem El-Bakry
helbakry5@yahoo.com

Sahar Sabbeh
m-sabbeh@yahoo.com; sfsabbeh@uj.edu.sa

Farid A. Badria
faridbadria@yahoo.com

Kyung-Sup Kwak
kskwak@inha.ac.kr

1 Information Systems Department, Faculty of Computers and Information Sciences, Mansoura University, Mansoura, Egypt

2 Faculty of Computers & Informatics, Benha University, Benha, Egypt

3 Department of Information and Communication Engineering, Inha University, Incheon, South Korea

4 Faculty of Computing and Information Technology, University of Jeddah, Jeddah, Saudi Arabia

5 Faculty of Pharmacy, Mansoura University, Mansoura, Egypt

FIB-4, etc.) utilize serum markers to detect the five stages of fibrosis: f_0, negative fibrosis; f_1, mild fibrosis; f_2, significant fibrosis; f_3, cirrhosis; and f_4, significant cirrhosis [6].

Patient diagnosis is difficult, especially in the information-overload environment of electronic health record systems where the number of patients is increasing rapidly [7]. Moreover, chronic disease requires medical experts to do continuous monitoring. In developing countries like Egypt, full availability of medical experts to cover a large number of patients cannot be achieved. Automated systems, such as clinical decision support systems (CDSSs), can help to overcome this issue [6]. The intent of this assignment is to build a CDSS to help physicians to estimate liver fibrosis in HCV patients. The resulting system has three advantages:

1. It could reduce costs and waiting times for healthcare centers because physicians do not need to conduct liver biopsy tests [5].
2. It permits early detection of the stages of liver fibrosis, especially for patients with a high probability of developing the disease [7].

The number of healthcare applications for a CDSS has increased in the last few years [7–15]. Accurate diagnosis is one of the most important problems in medicine. The relationship between diagnosis and clinical treatment protocols is evident in healthcare, as we explain in our survey [16]. Some studies tried to solve the medical diagnosis problem by using classical methods [2,11,17,18]. Hashem et al. [2] proposed a methodology to reduce the need for liver biopsies in cirrhosis patients. They used various data mining techniques including chi-square for feature selection, decision tree for rule induction, and artificial neural network (ANN) for final classification step. Their methodology used advanced fibrotic biomarkers for predicting the fibrosis stage. Orczyk and Porwik [11] developed pruned C4.5 classifiers to diagnose the liver fibrosis and compared their performance statistically with other different machine learning techniques, which were tested on the same dataset. They concluded that the C4.5 had better performance than other techniques. Gorunescu et al. [14] proposed a new model for the diagnosis of liver fibrosis. They developed multilayer perceptron network (MLP) where a combination of medical attributes is given as input to the MLP incorporated with genetic algorithms for the optimization process. Raoufy et al. [17] developed backpropagation ANN with three layers to diagnose cirrhosis in HBV chronic infected patients. Sartakhti et al. [19] employed support vector machine (SVM) classifiers to diagnose the hepatitis disease, and they used simulated annealing (SA) to optimize the performance.

Fuzzy logic was first applied in the medical applications [6,7,16] after Zadeh [20] introduced fuzzy logic. In reality, most medical data are vague; both physicians and patients describe conditions using uncertain and imprecise terms. The application of fuzzy expert systems in the medical domain started in 1985, and the exponential growth in published papers asserts its effectiveness [12]. Adlassnig [21] used fuzzy set theory in medical diagnosis. Adlassning stated that physicians, like others, can make mistakes, and could not be correct or certain about the diagnosis they made [21]. For this reason, Adlassnig extended his work by using fuzzy sets in medical diagnosis. Fuzzy set theory has been used to diagnose different diseases, including diabetes mellitus [7,22,23], kidney disease [12], and heart diseases [24]. Farokhzad and Ebrahimi [13] have proposed a model based on the association rule and adaptive neuro-fuzzy inference system (ANFIS) to diagnose liver sickness. El-Sappagh et al. [25] compared the performance of ANFIS with fuzzy analytical hierarchy process to diagnose fibrosis. Gadaras and Mikhailov [15] have proposed a fuzzy classification model based on automatically extract weighted fuzzy rules and combinations of membership functions for all fuzzy features. They evaluated their method by using different datasets in different medical domains (breast cancer, diabetes, and BUPA liver) and obtained accuracies of 96.5%, 92.9%, and 90.3%, respectively. Tsipouras et al. [8] proposed a model for coronary artery disease diagnosis by extracting a set of fuzzy rules from an initial dataset. They used C4.5 decision tree for rule induction, and then transformed the extracted crisp rules into fuzzy rules. Anooj [24] developed a fuzzy classification model for heart diseases. This model applied decision tree to generate rules and used the Mamdani FIS to classify patients [26]. To diagnose chronic liver disease, Kumar and Sahoo [27] developed a novel model based on SVM, rule induction, ANN, naïve Bayes, and decision tree techniques to classify liver disease. They evaluated the proposed model with *K*-fold cross-validation technique where decision tree model had the best performance. Because medical data are always uncertain and vague in nature, fuzzy rule-based systems (FRBSs) are the most intuitive technique for handling the nature of medical data [7]. The following are some advantages of FRBSs.

1. FRBSs use a domain knowledge representation formalism that can be easily understood by physicians.
2. FRBSs handle the uncertainty and imprecision inherited in medical data.
3. FRBSs give the degree of investigation of the classification output. These values are of interest for physicians.
4. As asserted in many previous studies [8], FRBSs usually provide higher classification accuracies than other techniques such as classical decision trees.

It is familiar that a medical data have to deal with imprecision and uncertainty, because diagnosis is not a matter of exact numerical values [28]. Physicians usually work with linguistic assertions based on ranges of values. Usually, the used indicators have some based intervals, and these intervals were built in fuzzy corresponding to normal-versus-bad states, because they depend on other features related to each patient. As a result, our study follows the FRBS technique to propose this CDSS system. The knowledge base consists of a set of fuzzy rules that estimate whether a new patient has a risk liver fibrosis or not.

In diagnosis problems, physicians are interested not only in the output but also in how the system get the required output. In addition, FRBSs have the ability to explain their results. They are called white box systems, because a domain expert can intuitively understand the systems' results. In contrast, neural network (NN) and linear programming (LP) models usually attain high classification accuracy, but making the decision is a black box system process [14]. Our methodology uses a combination of techniques, including a set of machine learning techniques to prepare the medical dataset, a fuzzy decision tree (FDT) technique for fuzzy rule induction that will be discussed later in the next section, and a Mamdani fuzzy inference system (FIS) [26]. This study is based on a complete list of medical features to accurately diagnose the fibrosis stage, including *demographics* (e.g., residence, occupation, gender, and age), *laboratory tests* (e.g., platelet count, white blood cell, hemoglobin, PCR, SGPT, SGOT, serum bilirubin, serum albumin, and serum ferritin), *signs* (e.g., ascites, spleen, lesions, and portal vein), and *symptoms* (e.g., appetite, dyspnea, vomiting, diarrhea, indigestion, vision problem, fatigue, jaundice, and skin pigmentation).

In this study, we build an F2DS CDSS to predict the liver fibrosis stages in HCV patients. The study depends on a real dataset collected from the electronic health records of Liver Institute, Mansoura University, Egypt. The dataset has been utilized to generate the list of system's fuzzy rules. The used fuzzy linguistic variables and fuzzy sets are modeled based on domain experts' knowledge and the most recent clinical practice guidelines. The fuzzy knowledge base is generated by training a fuzzy decision tree algorithm on the preprocessed and high-quality dataset. The proposed F2DS is based on the Mamdani fuzzy inference system in order to diagnose new cases. The remainder of this paper is organized as follows: Sect. 2 discusses some preliminaries. Section 3 presents the problem description. Section 4 presents the proposed framework, Sect. 5 discusses the results and evaluation, and Sect. 6 provides a discussion of the system, a conclusion, and suggests future work.

2 Preliminaries

To make the paper self-contained, next section discusses some basic concepts about the topics covered.

2.1 Fuzzy System

The fuzzy set theory [20] was introduced by Zadeh in 1965 to address vague, blurred, or imprecise concepts [29]. Fuzzy logic is a form of multi-valued logic derived from fuzzy set theory, and is used to deal with reasoning that is approximate, nonlinear, vague, and ill-defined. While in classical set theory, elements either belong to a set or not, in fuzzy set theory, elements can belong to a certain degree. More formally, classical sets are defined by characteristic functions.

Definition 1 *Fuzzy set* A over a universe of discourse X is defined by membership function (MF) μ_A (or simply$A$), which maps each element x to a value between [0, 1], as shown in Eq. 1:

$$\mu_A(x) : X \rightarrow [0, 1] \tag{1}$$

where A is the fuzzy set, μ_A is the degree of membership $\in [0, 1]$,$x \in X$, and $\mu_A(x)$ is a MF. Fuzzy set A can be defined as a set of ordered pairs:$A = \{x/\mu_A(x) | \forall x \in X, \mu_A(x) \in [0, 1]\}$ [29]. The crisp set operations, such as complement, intersection, and union, are extended to fuzzy sets, and these operations are performed by negation function, t-norm function, and t-conorm function, respectively. Fuzzy sets are used to define fuzzy linguistic variables in fuzzy expert systems.

Definition 2 Let R be the rule base made of a set of *OR*-connected multi-input single-output (MISO) rules,R^1, R^2, $\ldots$, R^n, i.e.,

$$R = \left\{R^1, R^2, \ldots, R^n\right\} \tag{2}$$

where the single fuzzy rule is denoted by $R^n \in R$. The standard form of rule R^n is IF $<$*Fuzzy-Antecedent*$>$ THEN $<$*Fuzzy-Consequent*$>$, which is written as follows:

$$R^n : x_1^n is A_1^n \wedge \cdots \wedge x_M^n \text{ is } A_M^n \text{ THEN } y^n \text{ is } B^n : \beta_i \tag{3}$$

where $x_k^n is A_k^n$, which is $k's$ fuzzy expression or condition, $\wedge$ is the *AND* fuzzy operator, $x_1^n, x_2^n, \ldots, x_M^n$are the M input fuzzy variables, y is the output fuzzy variable, A_k^n and $B^n$are the linguistic labels of x_k^n and y, respectively, $k = (1, 2, \ldots, M)$, and β_i is the degree of reliability of the rule. With the selection of$M$, we determine the input variables required for each rule.

Definition 3 Rules are fired using an inference algorithm such as Rete [29]. The computed consequences related to the same variable are aggregated (typically, using the maximum). Then, the output variables can be *defuzzified*. A defuzzified number is usually represented by a defuzzification function, such as the center of gravity (COG), which can be determined using the moment of area method, defined as:

$$\mathrm{COG} = \left(\frac{\int_X x \mu_B(x)\,\mathrm{d}x}{\int_X \mu_B(x)\,\mathrm{d}x} \right) \tag{4}$$

where $\mu_B(x)$ is the aggregated value of the fuzzy variable B over the universe of discourse X.

2.2 Fuzzy Decision Tree

A decision tree [30] is a formalism for expressing mapping from attribute values to classes, and consists of tests or attribute nodes linked to two or more sub-trees and leaves or decision nodes labeled with a class, which indicate the decision. The main advantage of the decision tree approach is that it is easy to follow any path through the tree. The tree can be expressed as a set of rules [31]. Iterative Dichotomiser (ID3) [31] and classification and regression trees (CART) [32] are most typical kinds of decision tree induction algorithms, which work through recursive partitioning. Their methodology is to partition the sample space in a data-driven manner, and represent the partitions as trees. ID3 is designed to deal with discrete domain, and requires prior partitioning; CART is designed to deal with a continuous numeric domain, but does not require prior partitioning because the conditions in the tree depend on a threshold [33].

The fuzzy decision tree construction method is included in fuzzy modeling. It depends on a classical decision tree in building effective branch attributes and their splitting interval to determine fuzzy partitions for input variables, thus allowing rule extraction. There are many techniques for building fuzzy decision trees [31–36]. One of the most suitable methods is to build a classical decision tree and then extend it to a fuzzy one, as will be discussed in the following subsections.

2.2.1 Building a Classical Decision Tree

Decision tree development requires attribute value space partitioning into A classes where $A \geq 2$ in each node N. The algorithm by Quinlan [30] for building the decision tree has the ability to divide attribute value space into the required number of nominal partitions, calculate for each partition the information gain, and next iterate recursive partitioning with the highest information gain. The decision tree root contains all data samples, D_i, with no restrictions. Each child node N_i is recursively split by partitioning its samples. A node become a leaf when its samples come from a unique class, or after a specific condition occurs, such as the node having a specific number of samples. Information gain is the most popular mechanism for attribute selection [30]. Algorithm 1 illustrates the classical decision tree creation process.

Algorithm 1: Classical decision tree construction

Input: Dataset D_i, decision node N

Begin

Step 1: For node N, compute the expected information from class decision partitioning

$I^N = -\sum_i^{|d|}(p_i . \log_2 p_i)$ // where d is the decisions set, p_i is the probability that training set samples in the represented node of class i //

Step 2: For each attribute A_i, $A_i \in \{A_1, A_2, \ldots, A_d\}$,

- Split the domain interval into initial partitions (classes) a_{ij} , $j \geq 2$
- Compute the information gain, $I^{N|a_{ij}}$, for each attribute with all possible partitioning combinations

Step 3: Select the attribute A_i with maximized information gain

$info(A_i) = I^N - \sum_j^{|D_i|}(w_j . I^{N|a_{ij}})$ //where w_j is relative weights of samples at child node a_{ij} to all samples in N, and D_i is a symbolic domain of the attribute//

Step 4: Split node N according to the selected attribute

Step 5: Return the partitioning of attributes at each child node with the largest information gain

Step 6: For each resulting node

If the stopping condition was not reached, then

Go to Step 1

End

2.2.2 Fuzzification of the Resulting Decision Tree

A fuzzy decision tree differs from a classical decision tree because it uses splitting criteria based on fuzzy restrictions for dataset D_i, represented by a set of attributes, A_i, for an object. Some attributes having discrete nominal domains are

characterized by crisp values, and others having a continuous numeric domain are characterized by crisp values, interval values, and fuzzy numbers [34].

Definition 4 Let D_i be a set of data samples at node N [34]:

$$D_i = \left\{ \left(x, \mu_{D_i}(x) \right) | x \in U \right\} \tag{5}$$

where x denotes data, and $\mu_{D_i}(x)$ denotes the membership degree of x to D_i.

Definition 5 Let $D_i^{v|A}$ be a fuzzy subset of D_i such that

$$\begin{aligned} D_i^{v|A} &= \left\{ \left(x, \mu_{D_i^{v|A}}(x) \right) \middle| \mu_{D_i^{v|A}}(x) \right. \\ &\left. = f\left(\mu_{D_i}(x), M(x, A, v) \right), x \in \sup(D_i) \right\} \end{aligned} \tag{6}$$

where A is a branching attribute, v is a trapezoidal fuzzy number for numeric attribute A, $M(x, A, v)$ denotes the matching degree of the attribute A value of x to branching v, and $f\left(\mu_{D_i}(x), M(x, A, v)\right)$ is an aggregation function by T-norm operators [34].

Boundary points P_i and trapezoidal fuzzy sets $Trap(a, b, c, d)$ are responsible to locate fuzzy partitioning intervals between two boundary points belonging to different classes to achieve overlapping [34]. Algorithm 2 shows partitioning for a numeric attribute.

2.2.3 Rule Extraction

Each rule is extracted from the decision tree as a path from the root node traversing down to the leaf node. Assume the rules $r_1, r_2, \ldots, r_i$ are extracted from a decision tree; the antecedent part of rule r_i is a list of fuzzy conditions of the form cond_1 AND cond_2 AND ... AND cond_n for n conditions. Each fuzzy condition, $cond_j$, has the form$(attISval)$, where att is a fuzzy attribute, and val is a fuzzy set for attribute att. The consequent part is the target classification output [35]. The logic operation $\bigvee$ used to sum all fuzzy concepts representing the antecedents of the rules with the consequent, as illustrated in Definition 2.

The proposed framework constructs a classical decision tree with an ID3 algorithm; then, it extends it to a fuzzy decision tree after defining fuzzy sets for each attribute. We implement the FDT in our system because it is capable to adapt to different kinds of output (continuous numerical output or discrete output), and it represents classes by using suitable splitting criteria based on fuzzy restrictions.

3 Problem Statement

This section provides the descriptions of the problem from the medical expert's point of view and of the medical dataset used.

Algorithm 2: Numerical attribute partitioning

Require: Input attribute partition $A_i = \{A_i | i = 1, \ldots, n\}$
1: Sort boundary point $P_j = \{P_j | j = 1, \ldots, p\}$
2: $iter = 1$; $\forall i\ n_i$=1 // i = dimension of fuzzy set, $i = \{1, \ldots, p_j\}$//
3: **if** $n_i^{max} \leq M_i$ **then** // M_i, initial size of fuzzy set //
generate $D_i^{v|A}$ fuzzy partitions of size n_i // remove inner boundary points in the same fuzzy set//
4: **end if**
5: **for** $D_i^{v|A}$ **do** calculate $I^{N|D_i^{v|A}}$
6: **if** $I^{N|D_i^{v|A}} < I^{N|a_{ij}}$ **then** remove $D_i^{v|A}$
7: **else** store fuzzy sets as a base system
8: **end if**
9: **end for**
10: **while** $iter \leq iter_{max}$ **do**
11: $iter = iter + 1$
12: **for** $1 \leq i \leq p$ **do**
13: $n_i = n_i + 1$
14: repeat 3 to 9
15: **end for**
16: keep $fuzzy\ system_{iter}$
17: **end while**

3.1 The Problem Description

Fibrosis is the scarring of connective tissue, changes the architecture of certain organs, and disrupts normal function. In the liver, it results from chronic liver diseases, such as viral hepatitis [36]. The degree of fibrosis is divided into five stages: f_0, negative fibrosis; f_1, mild fibrosis; f_2, significant fibrosis; f_3, cirrhosis; and f_4, significant cirrhosis [6]. Liver fibrosis is positively related to clinical implications in the management of chronic liver disease. As a result, it is critical to analyze all demographic, laboratory, and clinical data to make the diagnosis, save money and lives, and make decisions on treatment. The liver biopsy is a procedure where pieces of liver tissue are taken to be sent for examination, considering that adult biopsy sample just 1/50,000 of the liver tissues [5]. This procedure costs from US$1000 to US$2200 [5]. In addition, complications include internal bleeding and pain, and delayed results are often expected. Noninvasive serum biomarkers are patient friendly due to their immediate results, but its limitation is lower performance in diagnoses of significant fibrosis [4]. A poor diagnosis leads to significant liver fibrosis, cirrhosis, and end-stage liver failure. In this study, we develop a knowledge-based fuzzy expert system to assist physicians in liver fibrosis stage prediction. The outcomes of this study help intervene in the progress of HCV infections and help prevent complications, which may impose financial and health burdens on patients. More complicated and costly options, like liver transplantation or chemotherapy, will be the only solutions in the absence of such an informative study. Moreover, early and right decisions save money and shed more light on conditions for a better understanding of the infection's progress. The CDSS helps prevent aggravation of a viral infection into liver cancer, which subsequently leads to premature death.

3.2 The Dataset Description

The study dataset was obtained from the Liver Institute, Mansoura University, Egypt. Data were extracted from electronic health records of 119 patients, all infected with chronic HCV. Data concerned the analysis of patient demographics, laboratory tests, and symptoms. The studied dataset comprises 27 features that are relevant to HCV disease. They were selected according to domain expert knowledge and the most recent standard clinical practice guidelines [37]. Almost all participating patients were in the age range of 16–65 years, distributed into 80 males (67.5%) and 39 females (32.5%). The used dataset includes the most related clinical features to the clinical decision process and when predicting the status for each patient case. A detailed description of the used dataset with its features is presented in Table 1.

It has nine quantitative variables and 18 qualitative variables. Broadly, the missing data represent 13% of the whole dataset, and only nine patients had 100% complete information in all of the fields (7%). This is normal in the medical field. To enhance its quality, the dataset was preprocessed using machine learning techniques before being used.

4 The Proposed F2DS Framework

This section discusses the proposed F2DS system in detail. It generates its knowledge base using an FDT technique, and it uses the Mamdani fuzzy inference mechanism to predict patient states. In Fig. 1, the components of the proposed system with their relationships were illustrated. It presents input and output of the F2DS as the raw patient description data and the fibrosis stage prediction, respectively.

4.1 Data Preprocessing

The dataset is represented as data matrix, the rows of the matrix correspond to patient cases, and the columns to predictor features. Data quality has significant implications for the quality of diagnosis results. Preprocessing of the dataset is necessary to remove problems associated with medical data, like redundancy, noise, and missing data.

4.1.1 Anonymization, *UoM*, and Normalization

Anonymization is the process of removing any data that can identify the patient. All personal data, such as name, address, and ID numbers, are removed for privacy. As detailed in Table 1, we selected a unified unit of measurement (UoM) for each laboratory test. All feature values are converted to the selected UoM. We handle that difference with union units of measurement for each laboratory test. Data integration helps reduce and avoid redundancies and inconsistencies in the resulting dataset. We followed a manual process for schema integration, i.e., an attribute ID for each patient.

4.1.2 Handling Missing Data

In the developed systems, missing or unknown values are unavoidable in the real-life medical data. For example, the results of some laboratory tests may not be available because a healthcare institution cannot conduct the test, or the physician may have decided it was useless, or the test might have been skipped due to high cost [38]. The model learning process must deal with these missing data. In our study, there is no patient case with more than 50% of the values missing. Furthermore, features in the dataset with 25% or more of the values missing were dropped. As detailed in Table 1, from a set of 27 features, we dropped five: residence, occupation, indigestion, vision problems, and skin pigmentation. This reduced the overall missing data from 13 to 3.6% of the

Table 1 Patient feature descriptions, (data type: N = Numerical, C = Categorical, O = Ordinal)

Feature type	Feature name	Data type	UoM	Normal range	Min-Mean-Max	F. No.
Demographics	Residence	Qualitative/C	–	{Rural, urban}	–	1
	Occupation	Qualitative/C	–	{doctor, officer}	–	2
	Gender	Qualitative/C	Male, female	1, 2	–	3
	Age	Qualitative/N	–	19–82	19, 49.2, 82	4
Laboratory tests	Platelet count	Quantitative/N	10^3/cm	150–450	14, 184.2, 416	5
	White blood cell	Quantitative/N	U/L	4–11	1.1, 5.9, 10.3	6
	Hemoglobin	Quantitative/N	g/dL	12–16	9.6, 13.72, 18.1	7
	PCR	Quantitative/N	IU/mL	100,000–1000.000	50, 1.074.487, 20.040.062	8
	SGPT (ALT)	Quantitative/N	U/L	0–49	10, 63.27, 226	9
	SGOT (AST)	Quantitative/N	U/L	0–34	5, 58.182, 211	10
	Serum bilirubin (SB)	Quantitative/N	mg/dL	0.0–1.2	0.3, 1.009, 2.2	11
	Serum albumin (SA)	Quantitative/N	g/dL	3.2–4.8	0.7, 3.729, 5	12
	Serum ferritin	Quantitative/N	ng/mL	22–300	19.2, 221.95, 670	13
Signs	Ascites	Qualitative/C	No, Mild	1, 2	–	14
	Spleen	Qualitative/C	Normal, Enlarge	1, 2	–	15
	Lesions	Qualitative/O	Yes, No	1, 2	–	16
	Portal vein	Qualitative/C	Yes, No	1, 2	–	17
Symptoms	Appetite	Qualitative/O	Absent, rare, bad	1, 2, 3	–	18
	Dyspnea	Qualitative/O	Absent, rare, bad	1, 2, 3	–	19
	Vomiting	Qualitative/O	Absent, rare, bad	1, 2, 3	–	20
	Diarrhea	Qualitative/O	Absent, rare, bad	1, 2, 3	–	21
	Indigestion	Qualitative/O	Absent, rare, bad	1, 2, 3	–	22
	Vision problems	Qualitative/C	Yes, No	1, 2	–	23
	Fatigue	Qualitative/C	Absent, rare, bad	1, 2, 3	–	24
	Jaundice	Qualitative/O	Absent, rare, bad	1, 2, 3	–	25
	Skin pigmentation	Qualitative/O	Absent, rare, bad	1, 2, 3	–	26
Diagnosis	Fibrosis stage	Qualitative/O	F_0, F_1, F_2, F_3	0, 1, 2, 3	–	27

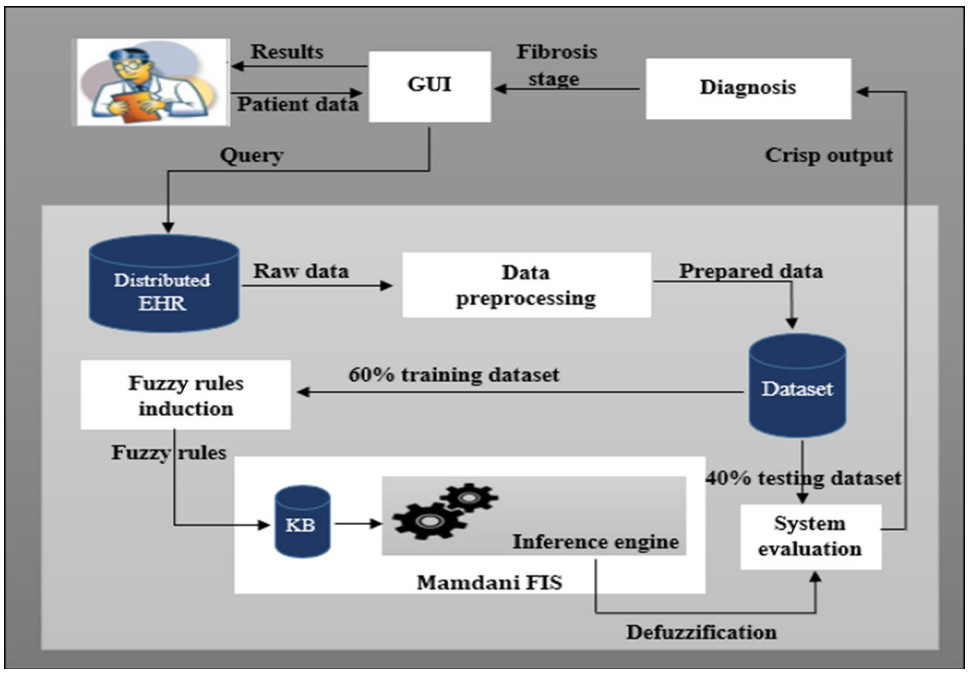


Fig. 1 The proposed F2DS framework

Table 2 Matching patient cases example

C#	F6	F7	F8	F9	F10	F11	F12	F13	F14	F15	F16	F17	F18	F19	F20	F21
$x77$	5.3	14.5	110	85	120	1.5	3	260	1	1	1	1	3	1	1	1
$x103$	5.3	14	250	–	46	2.4	3.7	297	1	1	1	–	3	1	–	1
$x116$	–	13.6	59	120	92	0.9	4.8	–	2	1	3	2	2	2	3	1

whole dataset. Complete cases rose from nine to 58 complete cases, or 48% of the dataset.

Hot deck imputation is a method for solving the problem of missing data (recipients) by replacing each missing value with an observed value from a similar unit (donor) in the same dataset [39]. This method of imputed values uses actually occurring values, not constructed values. The imputation method depends on collecting complete cases in a pool; then, we vote for a case with a high similarity in order to provide the missing data. Consider a patient case, P_{xi}, represented by an m-dimensional input vector,$x_i = [x_1, x_2, \ldots, x_m]^p$, given the patient cases represented by P_{x116}, P_{x103}, and P_{x77}, as shown in Table 2. Patient case P_{x103} has missing data in three features (f_9, f_{17}, and f_{20}) and is more similar to P_{x77} than P_{x116}.

The used similarity criterion is the heterogeneous Euclidean-overlap metric (HEOM) [40]. The HEOM distance between them is calculated by:

$$d\left(x_a, x_b\right) = \sqrt{\sum_{j=1}^{n} d_j\left(x_{aj}, x_{bj}\right)^2} \tag{7}$$

where $d_j\left(x_{aj}, x_{bj}\right)$ is the distance between x_a and x_b on the jth attribute. For a discrete attribute, the distance function d_j assigns a value of 0 for same values; otherwise, the value is 1, (d_j {0, 1}). For a quantitative attribute, we use $max\left(x_j\right)$ and $min\left(x_j\right)$ as the maximum and minimum values, respectively. The differences are almost between 0 and 1 (d_j [0, 1]). Now, so we have a dataset with 119 complete patient cases.

4.1.3 Feature Selection

Feature selection is frequently used as a data preprocessing step. Its goal is to improve prediction accuracy with selection of a subset of attributes from data that are most relevant for construction of the model system without decreasing the final result and without creating a new combination of attributes [41]. There are two general approaches for selection algorithms: the filter approach and the wrapper approach. The filter approach is independent of learning algorithms, and it uses general characteristics of the dataset to evaluate and rank features. The wrapper approach evaluates the goodness of features using the predictive accuracy of a predetermined learning algorithm [42]. In our study, the first phase is feature clustering to divide the original set into clusters (i.e., demographic data, laboratory test data, sign data, and symptom data) to make inner clusters correlate with each other. The target output is liver fibrosis diagnosis stage, Y_f, divided into four categories: $Y_f = \{f_0, f_1, f_2, f_3\}$. We need to select the relevant features significant to the target class from each cluster, and mute out features that are not useful in order to construct the final feature subset. Filter approach methods are implemented in our study. These methods can be classified into three groups: instance-based (e.g., relief and relief F), statistically based (e.g., chi-square), and entropy-based (e.g., information gain and gain ratio) [42]. Among the chosen techniques from each group, the relief gain algorithm [43] assess the quality of features due to how their value distinguishes between instances that are near each other; the chi-square [41] statistics method measures the lack of independence between term t and category Y; and gain ratio [44] is an enhancement of information gain (its function is to simultaneously maximize the feature's information gain and minimize its entropy)—then, it ranks the most relevant features repeated in the feature selection algorithms. Table 3 shows description data and statistical values of all observed attributes in our dataset.

As shown in Table 3, we use different feature selection algorithms to compare j clusters by separately ranking each algorithm on the whole dataset. Waikato Environment for Knowledge Analysis (Weka) software [45] is used in the implementation of various Ftechniques over N datasets to compare the results of them and optimize the accuracy of prediction. As described in Table 3, we used the Friedman test to compare the selected techniques separately [42,46]. The best performance technique got the first rank. Then, for each technique, average rank R_i was calculated across all datasets, with test statistics of the Friedman test computed as:

$$\gamma_F^2 = \frac{12N}{k\left(k+1\right)}\left[\sum_{i=1}^{k} R_i^2 - \frac{k\left(k+1\right)^2}{4}\right] \tag{8}$$

where $R_i = \frac{1}{N}\sum_i r_i^j$, with r_i^j as the rank of the ith algorithm on dataset x_j. According to Eq. 8, the best performing algorithm is the chi-square. So, we dropped some features from one cluster (polymerase chain reaction (PCR), serum albumin (SA), serum ferritin, hemoglobin) resulting from the feature selection technique with a 20% selected features ratio.

Table 3 Description data and statistical values

Feature type	Feature name	Instance-based group Relief FN = 119	Statistical theory group χ^2 (x^2) $N = 119$	Entropy-based group gain ratio $N = 119$
Demographic	Gender	− 0.00244	7.0342	0.074
	Age	0.00875	23.404	0.21
Laboratory tests	Platelet count	0.05788	55.208	0.433
	White blood cell	0.00008	13.602	0.323
	Hemoglobin	− 0.00656	0	0
	PCR	0.00374	0	0
	SGPT (ALT)	0.01964	25.526	0.206
	SGOT (AST)	0.0470	61.162	0.429
	Serum bilirubin (SB)	0.19818	73.345	0.33
	Serum albumin (SA)	0.02490	0	0
	Serum ferritin	− 0.00637	0	0
Sign	Ascites	− 0.0113	0.9914	0.008
	Spleen	− 0.0047	3.8895	0.033
	Lesions	0.02345	12.176	0.182
	Portal vein	0.02841	20.846	0.287
Symptom	Appetite	− 0.0277	1.98	0.009
	Dyspnea	− 0.0196	3.946	0.027
	Vomiting	0.0289	16.335	0.108
	Diarrhea	0.00002	13.691	0.095
	Fatigue	0.0028	5.155	0.034
	Jaundice	0.2770	60.341	0.257

4.2 Fuzzy Rules Induction

Fuzzy knowledge is defined in fuzzy rules form (IF condition THEN conclusion) to express the relationships between fuzzy parameters. There are many fuzzy rule-learning methods [47]. The most natural method is to extract rules from a domain expert [13]. However, domain experts are often not available. In addition, it is difficult to collect a complete and non-contradicting set of rules from experts. The alternative method is to utilize a machine learning algorithm to induce a rule set based on data. We have 119 cases in a dataset with improved and high-quality data. After finishing the feature selection process, we preprocessed the dataset with a subset of 17 features. In addition, the dataset was divided randomly into two subsets of training (60% of the samples) and testing (40% of the samples) sets.

We implemented a fuzzy decision tree as a rule-induction algorithm to extract the most suitable rule base from training dataset. We implemented the FDT with a gain ratio technique to split nodes in order to generate the prediction rules. The algorithm itself extracts a group of fuzzy rules on the base linguistic variable, $a_i^j ?A$, which leads to classifying patients P into four output subsets, c_i. A fuzzy rule can be defined as:

$$\text{if } A^1 \text{ is } a_i^1 \text{ and } A^2 \text{ is } a_i^2 \text{ and } \ldots \text{ and } A^n \text{ is } a_i^n \text{ then } C \text{ is } c_i \quad (9)$$

The system generated a set of 74 fuzzy rules from the training dataset. These rules were stored in the F2DS knowledge base (KB). Table 4 presents a small sample of this KB. Each fuzzy set is presented with appropriate notations, as we discuss later.

Starting from root node (i.e., splitting-point) N, fibrosis stage includes all training dataset items, S_k. The FDT uses a recursive procedure to split each node into C_i child nodes. $S_k^{v|A}$ represents fuzzy subsets in S_k, which were selected for the split according to the information gain (IG) algorithm. Equation 10 calculates the IG of each attribute. The entropy [35] has been used to select an attribute as branching attribute. A large entropy value means a more impure dataset. The best chosen attribute generates less impurity in a child node. Finally, the attribute can be qualified by IG measure, which is derived from impurity measure:

$$\text{Gain}\,(S_k, A) = \text{Entropy}\,(S_k) - \sum_{v \in \text{itemA}} \frac{\left|C_{S_k}^{v|A}\right|}{\left|C_{S_k}^{A}\right|} \text{Entropy}\left(S_k^{v|A}\right) \quad (10)$$

Table 4 Sample of generated fuzzy rules

Rule #	Age	Gender	PLT	WBC	ALT	AST	SB	Ascites	Spleen	Portal vein	Lesion	Appetite	Dyspnea	Diarrhea	Fatigue	Vomiting	Jaundice	Fibrosis stage
	IF (age is A and gender is M and SB is H and ……..and vomiting is A then fibro stage is F_1																	
R_1	O	-	L	L	VH	-	L	-	-	NO	-	-	B	B	-	A	A	$\boldsymbol{F_1}$
R_2	-	M	-	L	-	VH	H	M	N	-	NO	B	-	-	-	-	-	$\boldsymbol{F_2}$
R_3	O	M	VL	L	VH	VH	H	NO	N	-	-	-	-	-	-	-	-	$\boldsymbol{F_3}$
R_4	A	M	-	L	-	VH	H	NO	N	NO	NO	-	-	-	-	B	-	$\boldsymbol{F_3}$
R_5	A	M	-	L	-	VH	H	NO	N	NO	NO	-	-	R	-	A	-	$\boldsymbol{F_2}$
R_6	A	M	-	L	-	VH	H	NO	N	NO	NO	-	R	A	-	A	R	$\boldsymbol{F_3}$
R_7	O	-	-	-	-	-	L	-	-	NO	-	-	-	-	-	A	A	$\boldsymbol{F_0}$
R_8	A	M	-	L	-	-	N	-	-	-	-	B	A	A	-	R	R	$\boldsymbol{F_1}$
R_9	Y	M	-	L	-	VH	H	NO	N	-	-	-	-	-	-	-	-	$\boldsymbol{F_2}$
R_{10}	-	M	-	L	-	VH	H	M	N	-	NO	B	-	-	-	-	-	$\boldsymbol{F_3}$
R_{11}	-	F	-	-	-	VH	H	-	-	-	-	-	-	-	-	-	-	$\boldsymbol{F_3}$
R_{12}	-	M	-	L	-	VH	H	-	E	-	-	-	-	-	-	-	-	$\boldsymbol{F_2}$
R_{13}	-	M	-	L	-	-	N	-	-	-	-	-	A	-	B	A	-	$\boldsymbol{F_0}$
R_{14}	A	M	-	-	-	H	L	-	-	NO	-	-	A	A	-	A	A	$\boldsymbol{F_1}$
R_{15}	-	M	-	L	-	VH	H	-	E	-	-	-	-	-	-	-	-	$\boldsymbol{F_0}$
R_{16}	-	M	-	L	-	VH	H	M	N	-	-	-	-	-	-	-	B	$\boldsymbol{F_3}$

where

$$\text{Entropy}\,(S_k) = -\sum_i p_i^s . \log_2 p_i^{S_k}$$

$$p_i^{S_k} = \frac{C_{S_k}^i}{C_{S_k}}, \quad \text{and} \quad C_{S_k} = \sum_i C_{S_k}^i$$

where A is the branching attribute, and v denotes the branching level for numeric attribute A. The entropy for $S_k^{v|A}$ is calculated as $\text{Entropy}\left(S_k^{v|A}\right) = -\sum_i p_i^{S_k^{v|A}} . \log_2 p_i^{S_k^{v|A}}$. As information gain increases, the impurity of the child node decreases. The results after the training dataset help select the SB attribute to be the root node with the highest gain ratio; thus, its fuzzy attributes are selected as branching level attribute v = (lowSb, normalSb, highSb). For illustration, we use only trapezoidal fuzzy sets: trap (a, b, c, d). currently, the implemented tree is not pruned. Each internal node has a branching level, v, excluding nodes that its data samples not satisfying the fuzzy restrictions.

Fuzzy membership function is generated by decision tree induction method, so there is no need to specify it before construction the decision tree. After finishing fuzzy decision tree construction, the predictive score can be calculated when each atom ends with a decision node. The predictive score is the probability of the target indicator in the decision node of the trained model. We used a method to measure the degree of confidence with which data x belong to target class C_i as follows: $cf\,(x, c) = \frac{k(x,c)}{\max_i k(x,i)}$ [34].

The value of credibility (confidence) of the fuzzy rule represents how often it is likely to be true. Only the fuzzy rules whose credibility is more than or equal to the threshold $\varepsilon \geq 0.6$ are likely to be true. Figure 2 represents a part of the proposed FDT; leaf nodes contain more than one class label and are not always 1. Data may reach multiple leaf nodes from the root. For instance, the leftmost leaf suggests that the patient has a probability equals to 0.2 for fibrosis being absent, and 1.0 for mild fibrosis.

4.3 Fuzzy Inference Engine

In the fuzzy inference process, parallel if then rules are fired concurrently according to the input values to induce the diagnosed variable. The new input case is classified by Mamdani inference procedure into one of four output classes: f_0 (fibrosis absent), f_1 (mild fibrosis), f_2 (significant fibrosis), and f_3 (cirrhosis). The inference process is achieved by using fuzzy linguistic variables and the generated fuzzy rules, as explained in the following section. The inference process is illustrated in Fig. 3. The physician uses a graphical user interface (GUI)-based medical application to collect the patient's current state and collect the patient history from the medical record. These crisp data are entered into the Mamdani fuzzy inference mechanism, which has three main phases: fuzzification, inference engine, and defuzzification. The purpose of fuzzification phase is the transforming of crisp input into fuzzy. Inference engine phase performs approximate reason-

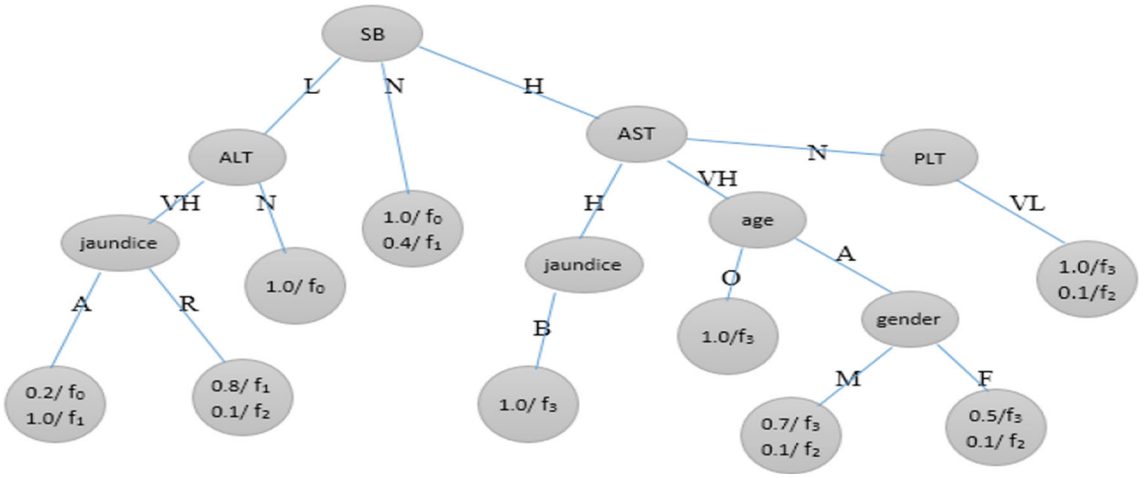


Fig. 2 Part of the induction FDT

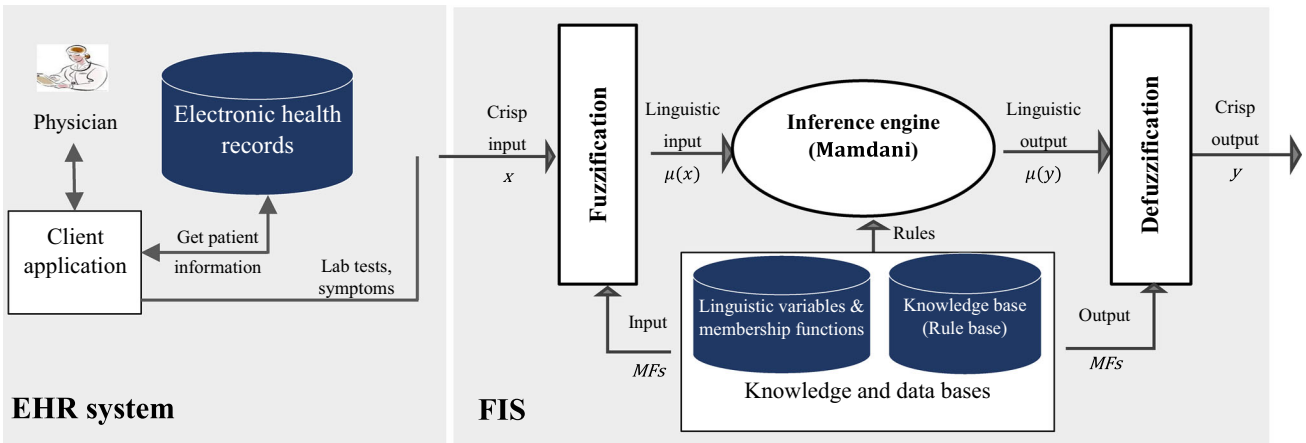


Fig. 3 Mamdani fuzzy inference process

ing in order to achieve a desired diagnosis for a human. It uses the defined membership function for all parameters to interprets the rule base with the implication operators, and realizes mapping from input fuzzy sets to output fuzzy sets. Defuzzification phase consists of getting non-fuzzy output from fuzzy output, also using membership functions. The output of the system is again crisp output. The KB was previously created. Next, we discuss the formulation of linguistic variables and values.

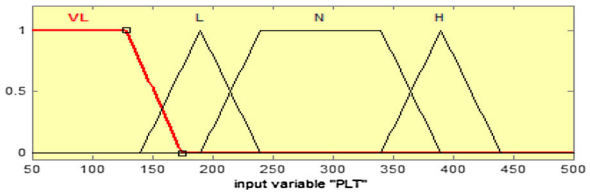
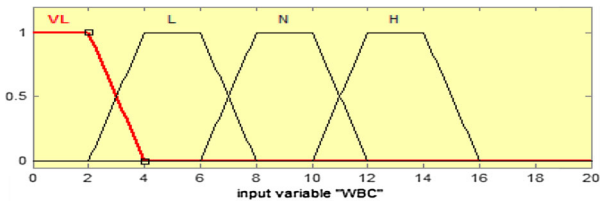
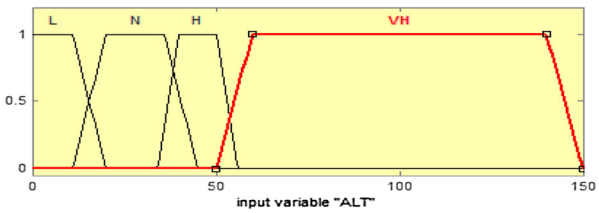
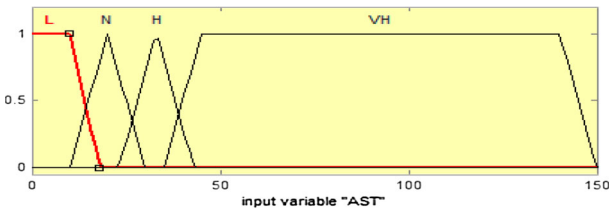
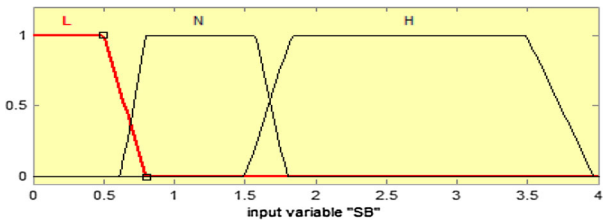
4.3.1 Fuzzy Variables

Our system is built on a complete set of 17 medical features that completely and accurately describe a fibrosis patient. Each of these medical features represents a fuzzy variable. They are age, ALT, AST, platelet count (PLT), white blood cells (WBC), serum bilirubin (SB), gender, portal vein, ascites, liver lesions, spleen, appetite, jaundice, fatigue, dyspnea, vomiting, and diarrhea. Each variable has a universe of discourse, and a type of either numerical, categorical, or ordinal. Tables 5 and 6 illustrate each fuzzy input variable with its data interval range and the associated fuzzy terms.

4.3.2 Fuzzy Sets

This step is responsible for determining the degree to which the input/output variables belong to each of the appropriate fuzzy sets. Each input variable, $A_i^j \varepsilon A^n \varepsilon A$, has membership interval degrees $C_i\left(A_i^j\right)$, and for output variable $c_i ?C$ according to the medical domain knowledge expert. In the present work, all laboratory tests have been divided into five categories (namely very low, low, normal, high, and very high) as detailed in Table 5. Symptoms are divided into three categories (absent, rare, and bad); signs are divided into two categories (yes and no), as detailed in Table 6. Output variables are divided into four categories (absent, mild, significant, and cirrhosis), as detailed in Table 7. For

Table 5 Sample of fuzzy sets description for quantitative input variables

Variable	Fuzzy set	Notation	Fit vector	$\mu(x)$
PLT range: (50–500) ID: x_4	V-low	VL	(50, 128, 175)	
	Low	L	(140, 190, 240)	
	Normal	N	(190, 240, 340, 390)	
	High	H	(340, 390, 440)	
WBC range: (0–20) ID: x_5	V-low	VL	(0, 2, 4)	
	Low	L	(2,4, 6, 8)	
	Normal	N	(6, 8, 10, 12)	
	High	H	(10, 12, 14, 16)	
ALT range: (0–150) ID: x_1	Low	L	(0, 11, 20)	
	Normal	N	(11, 20, 36, 45)	
	High	H	(34, 40, 50, 56)	
	V-high	VH	(50, 60, 140, 150)	
AST range: (0–150) ID: x_2	Low	L	(0, 10, 18)	
	Normal	N	(10, 20, 30)	
	High	H	(23, 33, 43)	
	V-high	VH	(35, 45, 140, 150)	
SB range: (0–4) ID: x_3	Low	L	(0, 0.5, 0.8)	
	Normal	N	(0.6, 0.8, 1.5, 1.8)	
	High	H	(1.5, 1.8, 3.5, 3.9)	

every fuzzy variable, a collection of fuzzy sets represent the terms of the associated linguistic variables. Therefore, the universe of discourse, UOD^i, of each selected variable X^i is divided into a certain number of s_i fuzzy sets described by membership functions$\mu^i_{s_i}$. The partition of an input variable domain into sharp intervals can be defined by many methods, including using intervals obtained from an expert's knowledge [12], the domain partition into a number of equal intervals, by decision tree algorithms, or by supervised or unsupervised clustering methods [48]. In our study, the fuzzy terms of linguistic variables are defined according to medical domain expert knowledge. Then, the partitions are optimized by k-means [49] and decision tree [34], with the k-means clustering method being the simplest, using a centroid-based approach to minimize intra-cluster variation [49]. It classifies total objects x into k clusters, with similarity based on Euclidean distance. We applied k-means and decision tree methods to generate the fuzzy partitions, and then compared them with the expert-based designed membership functions. In our study, we developed different membership function shapes MFs, according to the data type of each variable. The variables with discrete or categorical values were modeled using singleton membership functions with the same number of values, including gender, sign, and symptom. Numerical

Table 6 Fuzzy set description for discrete input variables

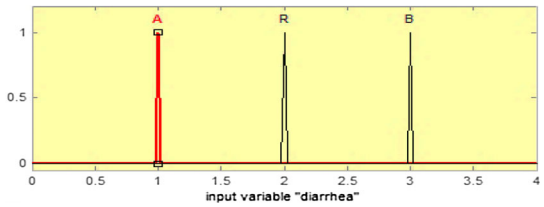
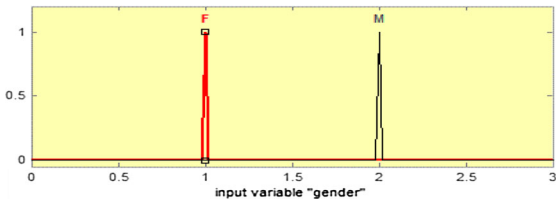
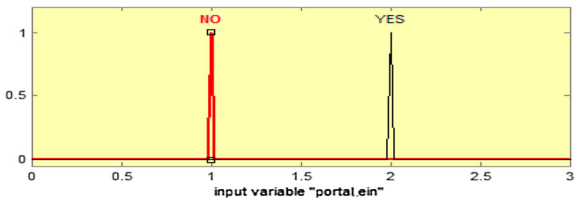
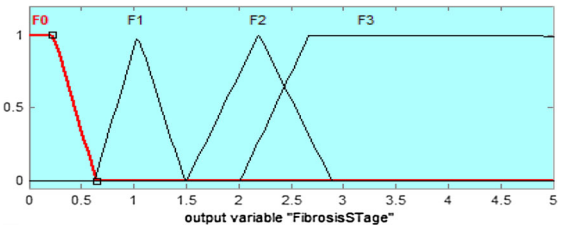
Variable	Fuzzy set	Notation	Fit vector	$\mu(x)$
Symptom range: (0–4) ID: $x_8, x_9, x_{10}\ x_{12}, x_{13}, x_{16}$	Absent	A	(1, 1, 1)	
	Rare	R	(2, 2, 2)	
	Bad	B	(3, 3, 3)	
Gender range: (0–3) ID: x_{15}	Male	M	(1, 1, 1)	
	Female	F	(2, 2, 2)	
Sign range: (0–3) ID: x_6, x_7, x_{11}, x_{17}	No	No	(1, 1, 1)	
	Yes	Yes	(2, 2, 2)	

Table 7 Fuzzy set description for the target variable

Variable	Fuzzy set	Notation	Fit vector	$\mu(x)$
Fibrosis stage range:(0–5) code:x_{18}	Absent	F_0	(0, 0.23, 0.64)	
	Mild	F_1	(0.62, 1, 1.5)	
	Significant	F_2	(1.5, 2.1, 2.7)	
	Cirrhosis	F_3	(2, 2.75, 5)	

features with a continuous data domain, including SB, AST, ALT, PLT, WBC, and age, are represented by trapezoidal MFs and triangular MFs. Triangular MFs use three parameters, which reflect the difference between a number and an interval as $x\ \langle x_a, x\rangle\ x_b$, as illustrated in Eq. 11. Left-shoulder MFs only decrease MFs like $x > x_a$; right-shoulder ones are increasing MFs like $x < x_a$ [50].

Trapezoidal MFs are provided in order to overcome the limitation of linear models between input and output variables, and obtain MFs with increasing/decreasing segments. Trapezoidal MFs can be represented by four parameters, as seen in Eq. 12 [50], where $a \le b \le c \le d$ are real number parameters of fuzzy set A, and the parameters b and c of each membership function are obtained from the minimum and maximum values of each fuzzy set. Parameter a of the first fuzzy set is the same value as$b$. Parameter d of the last fuzzy set in some features has the same value as parameter c.

$$\mu(x; a, b, c) \begin{cases} 0 & x < a, \\ \frac{x-a}{b-a} & a \le x < b, \\ \frac{c-x}{c-b} & b \le x < c, \\ 0 & x \ge c. \end{cases} \tag{11}$$

$$\mu(x; a, b, c, d) \begin{cases} 0 & x\ \langle a, x\rangle\ d, \\ \frac{x-a}{b-a} & a \le x < b, \\ 1 & b \le x < c, \\ \frac{d-x}{d-c} & c \le x \le d. \end{cases} \tag{12}$$

Tables 5, 6 and 7 present the input and output variables with their shapes. Table 5 presents a sample of quantitative input X_n variables, where each attribute $x_i \, \varepsilon X_n$ is fuzzified into linguistic terms, $A_k = \left\{a_{k,1}, \ldots, a_{k,i}, \ldots, a_{K,Ik}\right\}$, and $\mu_a\,(x_i)$ indicates the degree of compatibility for value x_i with the concept modeled by fuzzy set A.

The qualitative input variables, X_v, can be represented using binary MFs. The fuzzy set with a continuous membership function is discretized. The idea of discretizing the continuous MFs of fuzzy sets and the use of a discrete representation of fuzzy sets is presented in Table 6. We chose triangular MFs to represent fuzzy sets where all the parameters are equal. The categorical attribute X_v is categorized into a set of linguistic (i.e., categorical) values, $x_i = \{b_1, \ldots, b_i, \ldots., b_k\}$, for k fuzzy sets. For example, the range of symptom fuzzy sets (absent, rare, and bad) is shown in Table 6 as singleton functions, $\mu_{\text{Absent}}\,(x) = 1$, $\mu_{\text{rare}}\,(x) = 2$, $\mu_{\text{bad}}\,(x) = 3$, where $k = 3$, and $x_i = \{\text{absent, rare, bad}\}$.

Output attribute X_O is fuzzified into linguistic variables $C_k = \left\{c_{k,1}, \ldots, c_{k,i}, \ldots, c_{K,Ik}\right\}$. As shown in Table 7, the target output fuzzy sets with parameters are identified by medical domain expert according to the AST to Platelet Ratio Index (APRI) scale [6]. An example of fuzzification for attribute $\text{SB}(x_3)$ is shown in Table 5, where membership function is represented by trapezoidal shape of Eq. 12. Consider the fuzzy sets $\text{SB}_{\text{low}}(p)$, $\text{SB}_{\text{normal}}(p)$, $\text{SB}_{\text{high}}(p)$: $R \rightarrow [0, 1]$, membership degree $a_{3,1}\,(p) = \text{low}_{\text{SB}}\,(p)$ equals 1 for $\text{SB}(p) < 0.6$, $\frac{0.8-\text{SB}(p)}{0.8-0.6}$ for $0.6 \leq \text{SB}\,(p) \leq 0.8$, and equals 0 for $\text{SB}(p) > 0.8$. Membership degree $a_{3,2}\,(p) = \text{normal}_{\text{SB}}\,(p)$ equals 0 for $\text{SB}(p) < 0.6$ or $\text{SB}(p) > 1.8$, equals $\frac{\text{SB}(p)-0.6}{0.8-0.6}$ for $0.6 \leq \text{SB}\,(p) < 0.8$, equals 1 for $0.8 \leq \text{SB}\,(p) < 1.5$, and equals $\frac{1.8-\text{SB}(p)}{1.8-1.5}$ for $1.5 \leq \text{SB}\,(p) \leq 1.8$. Membership degree $a_{3,3}\,(p) = \text{high}_{\text{SB}}\,(p)$ equals 0 for $\text{SB}(p) < 1.5$ or $\text{SB}(p) > 3.97$, equals $\frac{\text{SB}(p)-1.5}{1.8-1.5}$ for $1.5 \leq \text{SB}\,(p) < 1.8$, and equals 1 for $1.8 \leq \text{SB}\,(p) < 3.97$.

Finally, we have a knowledge base composed of fuzzy rules and linguistic variables. In FIS, the fuzzy rule is inferred by input vector x using a fuzzy inference engine.

5 Results and Evaluation

The aim of this study is to solve the prediction problem with liver fibrosis stages for HCV patients using a set of input parameters from an HCV patient dataset. To achieve this target, we implemented an F2DS. This is a new medical fuzzy inference system for liver fibrosis diagnosis. We constructed a prediction model by discovering the fuzzy rules using a fuzzy rule reasoning method from experimental datasets, and generalized the relationship functions for both input and output parameters. The study was applied to a real dataset to show how it can be utilized for real medical diagnosis. Input parameters are x: {ALT, AST, SB, PLT, WBC, ascites, spleen, portal vein, lesions, appetite, dyspnea, fatigue, jaundice, diarrhea, vomiting, age, gender}, and output parameter y is the class of target disease, as shown in Table 7. This paper develops the F2DS using a set of steps as follows: input fuzzification, generating membership function MFs, extracting fuzzy rules, and output defuzzification. In the fuzzification step, trapezoidal triangular MFs are used to determine the degree that input/output variables belong to each of the appropriate fuzzy sets. In the defuzzification step, we used the center of gravity Eq. 4, which calculates the center of the area under the curve. In order to evaluate the proposed expert system, we firstly studied the impact of the input variables on the output decisions. Secondly, we used the test dataset to check each case's diagnosis against the diagnosis of expert physician and APRI test's diagnosis. The results were calculated by comparison with the expert diagnoses to estimate the simulation between system's decision and pathology examinations. Thirdly, the proposed system was evaluated by using the tenfold cross-validation technique. Finally, we compared the proposed system's results with the results of previous studies. The results of the evaluation process for a patient infected by chronic HCV who has cirrhosis are shown in Fig. 4. The chart of rules for the Mamdani FIS is shown with the parameters used. We implement the Mamdani fuzzy inference method through a fuzzy logic toolbox provided in Matlab $R2012a$ software, based on fuzzy rules generated by the FDT. The Mamdani FIS is usually used in particular for decision support applications because of its expressiveness and interpretability of the output [51].

We developed the fuzzy rule-based system for predicting the diagnosis class, and considered the appropriate MFs for input fuzzification and output defuzzification in the FIS. We have four output classes, and 74 fuzzy rules are used for the prediction model. As shown in Fig. 1, we used 40% of the dataset in the evaluation process. The test dataset was made up of 47 data samples, demonstrating that patients had all stages of fibrosis. Each test case is described by 17 patient features. Each feature is considered a dimensional system with its membership functions. The manner of data selection was random. The rule base was inferred from crisp input to get the final crisp diagnosis.

5.1 Performance

Figure 5 illustrates the relationships between some of the most important input variables and the output decision variable according to the constructed knowledge base. The choice of only two variables derives from the need to show the behavior of 3D surfaces representing the functions of the posterior probabilities of the output class. The surface emphasizes the separation of the variable space into regions

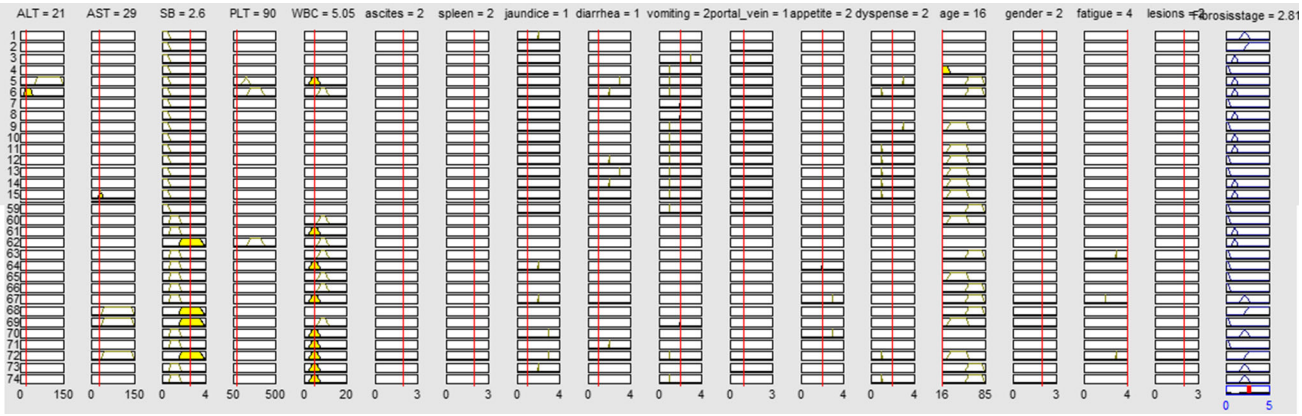


Fig. 4 The inference process in the Matlab toolbox

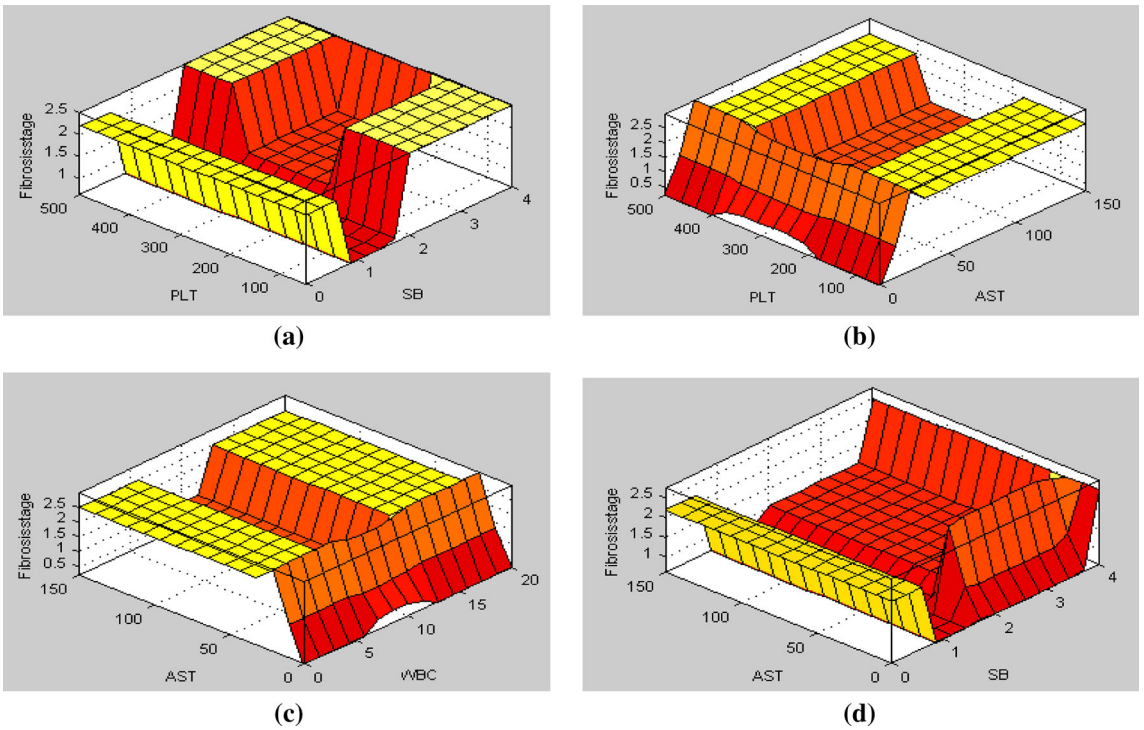


Fig. 5 Response surface plots by using different variables **a** PLT, SB, **b** AST, PLT, **c** WBC, AST, and **d** AST, SB

associated with different output classes, f_i: i.e., f_0 is red, f_1 is orange, f_2 is gold, and f_3 is yellow. Figure 5 is a 3D graph showing how the variations in variables affect the liver fibrosis stages.

Figure 5a represents the effect of serum bilirubin and platelet count in the diagnosis process due to their importance in the medical diagnosis. The intersection point x between the normal SB value and any PLT values belongs to *fibrosis absent*, even with a low PLT; the region of the intersection between normal PLT values and any SB belongs to *mild fibrosis*, even with a high SB. The region of the intersection between a high SB and a low PLT is for a *cirrhosis* diagnosis. In Fig. 5b, we study the relationship between platelet count (PLT) and aspartate aminotransferase (AST). The system represents the importance of PLT in the diagnosis process; the top view of the surface emphasizes each output class's probabilities, and the *cirrhosis* diagnosis is when PLT is low and AST is high, but mild fibrosis, f_1, is represented in the normal AST range. In Fig. 5c, we studied the relationship between white blood cells (WBC) and AST. The surface shows that when WBC decreases and AST increases, the probability of *cirrhosis* increases. This is medically intuitive, which asserts the accuracy of our system. In Fig. 5d, the variables AST and SB emphasized the separation regions between different output classes. The intersection regions between high SB and high AST increase the probability of *cirrhosis*. The intersec-

Table 8 Evaluation of the proposed F2DS

Term	Cirrhosis (f_3)		Absent (f_0)		Mild (f_1)		Significant (f_2)	
	F2DS	APRI	F2DS	APRI	F2DS	APRI	F2DS	APRI
TP	10	6	15	13	9	10	9	5
TN	35	36	32	30	36	25	34	36
FP	2	0	0	2	0	11	2	0
FN	0	4	0	2	2	1	2	6
CE	0.04	0.166	0	0.08	0.04	0.25	0.08	0.12

tion between normal SB and normal AST is a high probability of *fibrosis absent* and *mild fibrosis*. These variables are considered standard measures, and were met in the process of diagnosing liver fibrosis stages.

After a global analysis of all the possible scenarios for test response surface plots, and with the combination of 17 variables, it was determined that the variables with the highest impact in liver fibrosis diagnosis are PLT, SB, and AST. If these variables are compared with the rest of the input variables, probabilities can be up to 2.5, which is the highest degree of fibrosis, as can be observed. This is a key indicator that allows us to assume that these variables have a greater impact on the output.

In addition, the system performance was measured with two of the most widely used metrics: squared classification error index (SCE) in Eq. 13, and classification error index (CE) in Eq. 14 [48]:

$$\mathrm{SCE} = \frac{1}{N} * \frac{1}{K} * \sum_{i=1}^{N} \sum_{k=1}^{K} \left(\alpha_i^k - \delta_i^k\right)^2 \tag{13}$$

$$\mathrm{CE} = \frac{1}{N} * \sum_{i=1}^{N} y_i \tag{14}$$

where N is the number of data samples, y_i is 1 if the ith case is correctly classified (0, otherwise), K is the number of output classes, α_i^k is the activation of the kth class for the ith sample, and δ_i^k is 1 for the correct class (0, otherwise). Our domain expert guided the evaluation process and interpretation of the results. The system's SCE value is measured with Eq. 13. The system has an SCE of 0.125, while CE is simply the percentage of wrongly classified data samples. The system has an error ratio of 0.025. The system's classification accuracy is calculated with acc $= 1 -$ CE. The system achieved accuracy of 95.7%.

5.2 Evaluation by Measured Terms

Next, the experimental results of the F2DS for risk prediction were evaluated using measured terms [24], and these terms have the following semantics.

- *True Positive (TP)* is the number of cases correctly identified as cirrhosis.
- *True Negative (TN)* is the number of cases correctly identified as non-cirrhosis.
- *False Negative (FN)* is the number of cases incorrectly identified as non-cirrhosis.
- *False Positive (FP)* is the number of cases incorrectly identified as cirrhosis.

These measured terms were used in measuring performance according to all four output diagnosis classes. The *APRI* is a simple medical test to predict liver fibrosis with acceptable diagnosis accuracy. It employs two laboratory tests in one formula, Eq. 15 [6]. Nevertheless, it has limitations, such as its inability to identify high levels of fibrosis, and because laboratories establish different values for the upper limit of normal (ULN) [52]:

$$\mathrm{APRI} = \left(\frac{\mathrm{AST} * \left(\frac{1}{\mathrm{ULN}}\right)}{\mathrm{PLT}}\right) * 100 \tag{15}$$

where ULN $= 35$ U/L is the upper normal range. Table 8 presents an evaluation comparison between F2DS and APRI using measured terms and mean square error for each output class.

These results were used to measure the performance of the proposed system compared with the APRI results. The comparison is presented in Fig. 6. The results are reported in different terms considering accuracy ACC $= \frac{\mathrm{TP+TN}}{\mathrm{TP+TN+FN+FP}}$, sensitivity TPR $= \frac{\mathrm{TP}}{\mathrm{TP+FN}}$, and specificity SPC $= \frac{\mathrm{TN}}{\mathrm{TN+FP}}$, false positive rate FPR $= \frac{\mathrm{FP}}{\mathrm{FP+TN}}$, and false negative rate FNR $= \frac{\mathrm{FN}}{\mathrm{TP+FN}}$, F_1Score $= \frac{2\mathrm{TP}}{2\mathrm{TP+FP+FN}}$, and Matthews correlation coefficient

$$\mathrm{MCC} = \frac{\mathrm{TP} * \mathrm{TN\text{-}FP} * \mathrm{FN}}{\sqrt{(\mathrm{TP}+\mathrm{FP})\,(\mathrm{TP}+\mathrm{FN})\,(\mathrm{TN}+\mathrm{FP})\,(\mathrm{TN}+\mathrm{FN})}}$$

[24]. The used test dataset has no missing values, and without 10-fold cross-validation [53].

Figure 6 presents the comparison results between F2DS and the APRI medical test, showing accuracies of 95.7%

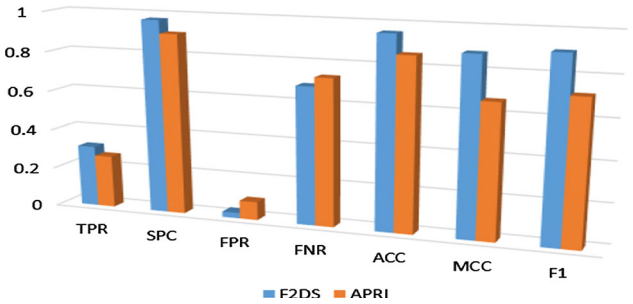


Fig. 6 Comparison results between F2DS and APRI

and 86%, respectively. The sensitivities are 31% and 26.4%, respectively. The specificities are 97.2% and 90.7%, respectively. False positive rates are 2.7% and 9.2%, false negative rates are 68% and 73%, f_1 scores are 90% and 71%, and Matthews correlation coefficients are 88% and 66%, respectively. As shown in Fig. 6, these results emphasize the medical significance of our proposed system. It achieves more accurate results than APRI.

5.3 Evaluation by Tenfold Cross-Validation

In this section, we evaluate the system with the whole dataset against the clinical diagnosis by physicians. The 10-fold cross-validation technique was employed to examine the model's performance due to its advantage of random dataset division. Tenfold cross-validation is a statistical technique useful in determining the robustness of a model [54]. The dataset was divided into 10 subsets, and the holdout method is repeated 10 times. Each time, one of the 10 subsets is used as the test set, and other nine subsets are combined together to form the training subset. First, we trained the algorithm using the training set in each fold and measured *TPR*, *SPC*, *FNR*, and *FPR* using the test dataset. Then, we calculated the average performance of the 10 measured values. The disadvantages with this method are that it takes a lot of time and requires a lot of computation to make the evaluation.

We aim to predict fibrosis with high levels {significant, cirrhosis}. To measure the performance in a more accurate manner, we minimized the number of classes in the decision variables by dividing the values into two groups for positive class ($f_2 + f_3$) and negative class ($f_0 + f_1$). The dataset was divided into $n_k = n/K$ cases, where n number of all data samples, and $K = 10$ is the number of folds; n_k=12 is the size of the kth data sample. Equation 16 is used to get cv_k cross-validation:

$$\mathrm{cv}_k = \frac{1}{K} * \sum_{k=1}^{K} \mathrm{MSE}_k \tag{16}$$

where mean square error MSE_k is the error rate for each test fold. We then added them together to get the final results. The system achieved an average accuracy of 93% when predicting the four output classes without integration. The accuracy result after integration was 94.1%.

6 Discussion

This study described the applicability of a fuzzy expert system in diagnosing liver fibrosis with its different stages, even at the first level of cirrhosis. The model can handle uncertainty at both input and output. The evaluation process was based on a dataset of 119 real HCV patients. The patients were clinically and pathologically diagnosed by medical experts. We used a fuzzy decision tree algorithm to construct the fuzzy rule knowledge base. This knowledge base has been refined by removing the conflicting rules. Conflicting rules had the same IF parts but different THEN parts. Rules having the higher weights were appended to the knowledge base, and the other rules were excluded. The final rule set had a complete coverage to our diagnosis domain, and they were highly interoperable by domain experts. In addition, the generated rules presented a good match between the decision from the system and the pathology examinations. Prediction results of our F2DS system were compared to other classifiers, which are mentioned in the literature. Generally, classical classification algorithms often require large datasets in the training phase. Researchers in [17] stated that backpropagation ANN with three layers achieved accuracy of 91.3%. However, ANN has many limitations in medical domain. It works as a black box and cannot provide a clear and direct interpretation of its decisions. Researchers in [14] proposed an evolutionary system for the classification of chronically infected HCV patients. When their system adjusted weights of input patterns before network training, the problem of permutation phenomenon appeared [18]. They mentioned that their model achieved 46% accuracy, and 49% after applying the feature selection techniques. The framework of [11] combined a filter feature selection method with decision tree algorithm for classification. A dataset with 290 features has been used; as a result, the pruned decision tree did not work as required. Kumar and Sahoo [27] presented a classification model based on SVM, rule induction, decision tree, naïve Bayes, and ANN techniques. Their average accuracies were 71.36%, 82.68%,71.12%, 55.29%, and 70.67%, respectively. These systems had accuracies of 82.33%, 94.34%, 98.46%, 82.16%, and 79.07%, respectively, after applying rule-based system. In addition, Hashem et al. [2] used dataset of 335 infected patients to build their classification system based on six crisp rules generated using classical decision tree. The resulting system is error prone if we change the weights of input features, and the generated crisp rules are not applicable for the nature of medical data.

Table 9 An empirical analysis of the literature review in liver fibrosis diagnosis

References	Dataset/feature	No. of rules	Technique	Year	Accuracy (%)
Gadaras and Mikhailov [15]	345/8	8	FIS	2009	89.4
Hashem et al. [2]	355/9	6	DT + ANN	2010	91.7
Raoufy et al. [17]	144/12	N/A	DT	2011	91.3
Gorunescu et al. [14]	722/25	N/A	GA + DT	2012	49
Kumar and Sahoo [27]	583/12	20	DT	2013	98.4
Farokhzad and Ebrahimi [13]	583/6	243	ANFIS	2016	70
Orczyk and Porwik [11]	290/7	N/A	DT	2016	71
Proposed F2DS	119/17	74	FDT + FIS	2018	94.1

On the other hand, fuzzy expert systems achieved high accuracy, even when the machine learning techniques used to generate fuzzy rules are based on small datasets [25], see Table 9. Models of [2,22] applied crisp rules to generate the decision support systems. In references [13,15], the fuzzy rules were collected only from medical experts. The accuracy and completeness of the rules depend only on the experience of physician, and these experiences are different from one expert to another. The completeness and accuracy of the resulting knowledge bases affect the decision support system accuracy. Since fuzzy mapping rule provides a functional mapping between fuzzy input and fuzzy output, the crisp relationship between input and output is very complicated when they are developed in fuzzy system. Fuzzy mapping rules work similar to human initiatives in mapping each relation. Fuzzy rules can be generated from their equivalent crisp rules [8]. However, the generated rules are non-overlapping. Tsipouras et al., [8] have transformed the crisp rules into fuzzy rules. They have achieved 80% accuracy using crisp rules and 85% accuracy using fuzzy rules implementation.

Anooj [24] followed another idea to generate the fuzzy system knowledge base. Their decision tree algorithm was applied on the input data, where were in a fuzzy numbers format. In this research, fuzzy number is generated by decision tree induction method. Malmir et al. [12] have developed fuzzy rule-based system based on Mamdani FIS for kidney diseases. The system achieved 87.5% performance accuracy for 40 samples and 22 input features; the implemented rule base was defined by only medical experts.

Having said that, our proposed system is based on a set of features related significantly to the liver fibrosis disease, i.e., non-expensive routine blood laboratory tests, patient physical examination, and demographic data. The *ALT*, *AST*, and *PLT* are the most used metrics in the medical literatures, especially for liver diseases [2,11,13–15,18,22], where they have been used to measure the damages in liver. *Age* is a critical feature in the proposed system because it has also a critical medical weight. For example, old males are 30% of infected patients diagnosed with high levels of fibrosis. The *SB* is a critical feature in the rule set; it participated in all induced rules. White blood cells feature is not participated in most generated rules, but it is also proven to be an important indicator for cirrhosis prediction, as illustrated in Sect. 5.1. The liver fibrosis symptom features are related to the diagnosis process especially jaundice and appetite. After the experiment, the jaundice level is linked with SB laboratory test. Jaundice bad level appeared when patient had a high SB degree. Signs had vital roles in the cirrhosis diagnosis process, where the patient has at least one of them especially portal vein. The proposed F2DS system achieved an overall accuracy of 94.1% for liver fibrosis stages domain.

The experimental results of the system assert the applicability of the approach to diagnoses of other diseases. Furthermore, due to the simplicity of the system, it can be implemented in a mobile application for remote patient monitoring. Having said that, our system has some drawbacks that will be handled in our advanced work. First, there are some other medical factors that were not used in the proposed system, including the alkaline phosphatase level test (ALP). In addition, some diseases that have direct or indirect effects on liver status, such as thyroid disease, obesity, and diabetes, were not considered. Second, for integration and interoperability between the CDSS and electronic health records [28], a semantic ontology will be considered in future improvements to the system.

7 Conclusion

In this paper, we proposed and implement a framework for fuzzy rule-based system in the medical domain. The framework is based on an interpretable knowledge base, which considers both expert knowledge and knowledge extracted from data. We proposed a new knowledge-based system for prediction of liver fibrosis stages using a fuzzy reasoning technique. Euclidean-overlap metric and chi-square, respectively, were used for handling missing data and feature selection in the dataset. We used entropy measure for

fuzzy rules induction. The developed diagnosis module consists of an expert system and a fuzzy inference system to perform the diagnostic task. A set of 74 rules was defined using the patients' dataset, as well as expert knowledge in the disease domain. We evaluated the knowledge-based system on a real dataset of 119 cases; the dataset was taken from HCV patients' clinical records at the Liver Institute of Mansoura University. The developed expert system uses rules to diagnose stages of liver fibrosis in patients based on their laboratory tests, symptoms, signs, and demographic data. The combination of an expert system and fuzzy logic forms an expert fuzzy system, and achieved good prediction accuracy of 95.7% for liver fibrosis stages. The developed method in this study can be developed for other pathologies as a decision support system. There is still plenty of work to do to enhance liver fibrosis diagnosis. In future studies, we will solve problems by adding a semantic dimension based on Web Ontology Language (OWL). This will be achieved by using a fuzzy ontology technique.

References

1. Amer, F.; Gohar, M.; Yousef, M.: Epidemiology of hepatitis C virus infection in Egypt. Int. J. Trop. Dis. Health **7**(3), 119–131 (2015)
2. Hashem, A.; Rasmy, E.; Wahba, K,; Shaker, O.: Prediction of the degree of liver fibrosis using different pattern recognition techniques. In: 5th Cairo International Biomedical Engineering Conference, pp. 210–214. IEEE (2010)
3. Ministry of Health and Population [Egypt]: El-Zanaty and Associates [Egypt] and ICF International, Egypt Health Issues Survey, p. 2015. Cairo: Ministry of Health and Population and ICF International, Egypt and Rockvilla, Maryland, USA (2015)
4. Sebastiani, G.; Gkouatsos, K.; Pantopoulos, K.: Chronic hepatitis C, and liver fibrosis. World J. Gastroenterol. **20**(32), 11033–11053 (2014)
5. Sanai, F.; Keeffe, E.: Liver biopsy for histological assessment-the case against. Saudi J. Gastroenterol. **16**(2), 124–132 (2010)
6. Badria, F.; Gabr, S.: Prediction of liver fibrosis and cirrhosis among egyptians using noninvasive index. J. Pure Appl. Microbiol. **1**(1), 45–50 (2007)
7. Saleh, E.; Valls, A.; Moreno, A.; Romero, P.: Diabetes retinopathy risk estimation using fuzzy rules on electronic health record data. In: Modeling Decision for Artificial Intelligence MDAI Lecture Notes in Computer Science vol. **2016**, no. (9880), pp. 263–274 (2016)
8. Tsipouras, M.; Fotiadis, D.; Naka, K.; Michalis, L.: Automated diagnosis of coronary artery disease based on data mining and fuzzy modeling. IEEE Trans. Inf. Technol. Biomed. **12**(2), 447–458 (2008)
9. Satarkar, S.; Ali, M.: Fuzzy expert system for the diagnosis of common liver disease. Int. Eng. J. Res. Dev. **1**(1), 2–7 (2014)
10. Bohacik, J.; Zabovsky, M.: Dissimilarity measure for comparison of fuzzified instance and its application in a fuzzy rule-based system for heart failure domain. In: 14th International Symposium on Applied Machine Learning and Informatics, pp. 339–344. IEEE (2016)
11. Orczyk, T.; Porwik, P.: Liver fibrosis diagnosis support system using machine learning methods. Adv. Comput. Syst. Secur. **395**, 111–121 (2015)
12. Malmir, B.; Amini, M.; Chang, S.: A medical decision support system for disease diagnosis under uncertainty. Expert Syst. Appl. **88**, 95–108 (2017)
13. Farokhzad, M.; Ebrahimi, L.: A novel adaptive neuro-fuzzy inference system for the diagnosis of liver disease. Int. J. Acad. Res. Comput. Eng. IJARCE **1**(1), 61–66 (2016)
14. Gorunescu, F.; Belciug, S.; Gorunescu, M.; Badea, R.: Intelligent decision making for liver fibrosis stadialization based on tandem feature selection and evolutionary driven neural network. Expert Syst. Appl. **39**, 12824–12832 (2012)
15. Gadaras, I.; Mikhailov, L.: An interpretable fuzzy rule-based classification methodology for medical diagnosis. Artif. Intell. Med. **47**(1), 25–41 (2009)
16. Sweidan, S.; Elbakry, H.; Elsappagh, S.; Sabah, S.; Mastorakis, N.: viral hepatitis diagnosis: a survey of artificial intelligent techniques. Int. J. Biol. Biomed. **1**, 106–116 (2016)
17. Raoufy, M.; Vahdani, P.; Alavian, S.; Fekri, S.; Eftekhari, P.; Gharibzadeh, S.: A novel method for diagnosing cirrhosis in patients with chronic hepatitis B: artificial neural network approach. J. Med. Syst. **35**, 121–126 (2011)
18. Gong, T.; Tuson, A.: Enhanced forma analysis of permutation problems. In: Proceedings of the 9th Annual Conference on Genetic and Evolutionary Computation- GECCO'07, pp. 923–930 (2007)
19. Sartakhti, J.; Zangooei, M.; Mozafari, K.: hepatitis disease diagnosis using a novel hybrid method based on support vector machine and simulated annealing (SVM-SA). Comput. Method Program Biomed. **I08**, 570–579 (2012)
20. Zadeh, L.: Fuzzy sets. Inf. Control **8**, 338–353 (1965)
21. Adlassnig, K.: Fuzzy set theory in medical diagnosis. IEEE Trans. Syst. Man Cybern. **16**(2), 260–265 (1986)
22. El-Sappagh, S.; Elmogy, M.; Riad, A.M.: A fuzzy-ontology oriented case-based reasoning framework for semantic diabetes diagnosis. Artif. Intell. Med. **65**(3), 179–208 (2015)
23. El-Sappagh, S.; Alonso, J.; Ali, F.; Ali, A.; Jang, J.; Kwak, K.: An ontology-based interpretable fuzzy decision support system for diabetes diagnosis. IEEE Access **6**, 37371–94 (2018)
24. Anooj, P.: Implementing decision tree fuzzy rules in clinical decision support system after comparing with fuzzy based and neural network based systems. In: IEEE, IT Convergence and security ICITCS, 2013 International Conference (2013). https://doi.org/10.1109/ICITCS.2013.6717845
25. El-Sappagh, S.; Ali, F.; Ali, A.; Hendawi, A.; Badria, F.; Suh, D.: Clinical decision support system for liver fibrosis prediction in hepatitis patients: A case comparison of two soft computing techniques. IEEE Access **6**, 52911–52929 (2018)
26. Mamdani, E.; Assillian, S.: An experiment in linguistic synthesis with a fuzzy logic controller. Int. J. Man Mach. Stud. **7**(1), 1–13 (1975)
27. Kumar, Y.; Sahoo, G.: Prediction of different types of liver diseases using rule based classification model. Technol. Health Care **21**, 417–432 (2013)
28. El-Sappagh, S.; Elmogy, M.: an encoding methodology for medical knowledge using SNOMED CT ontology. J. King Saud Univ. Comput. Inf. Sci. **28**(3), 311–329 (2016)
29. Forgy, C.: A fast algorithm for the many pattern/many object pattern match problems. Artif. Intell. **19**(1), 17–37 (1982)
30. Quinlan, J.: C4.5: Programs for machine learning. In: Machine Learning, Vol. 16, Issue 3, pp. 235–240. Morgan Kaufmann Publishers Inc (1993)
31. Quinlan, J.: Induction of decision trees. Mach. Learn. **1**, 81–106 (1986)
32. Jang, J.: Structure determine in fuzzy modeling: a fuzzy CART approach. In: Proceedings of the IEEE Conference on Fuzzy Systems, pp. 480–485 (1994)
33. Janikow, C.: Fuzzy decision trees: issues and methods. IEEE Trans. Syst. Man Cybern B **28**(1), 1–14 (1998)

34. Lee, K.; Lee, J.; Kwang, H.: A fuzzy decision tree induction method for fuzzy data. In: IEEE International Fuzzy System Conference Proceedings, pp. 16–21 (1999)
35. Wang, T.; Dalee, H.: Construction a fuzzy decision tree by integrating fuzzy sets and entropy. In: 5th WSEAS International Conference on Applied Computer Science, pp. 306–311 (2006)
36. Suk, K.; Kim, D.: staging of liver fibrosis or cirrhosis: the role of hepatic venous pressure gradient Measurement. World J. Hepatol. **7**(3), 607–615 (2015)
37. Kanwal, F.: Quality of care assessment in chronic liver disease. Clin. Liver Dis. **4**(6), 149–152 (2014)
38. Almuhaideb, S.; Menai, M.: Impact of preprocessing on medical data classification. Front. Comput. Sci. **10**(6), 1082–1102 (2016)
39. Andridge, R.; Little, R.: A review of hot-deck imputation for survey non-response. Int. Stat. Rev. **78**(1), 40–64 (2011)
40. Pereira, E.; Estevez, D.: Automatic classification of respiratory patterns involving missing data imputation Techniques. Biosyst. Eng. **138**, 65–76 (2015)
41. Mohsin, M.; Hamdan, A.; Bakar, A.: An evaluation of feature selection technique for dendrite cell algorithm. In: IEEE, IT Convergence and Security ICITCS, pp. 1–5 (2015). https://doi.org/10.1109/ICITCS.2014.7021732
42. Liu, S.; Chen, X.; Liu, W.; Chen, J.; Gu, Q.; Chen, D.: FECAR: A feature selection framework for software defect prediction. IEEE 38th Annual International Computers, Software and Application Conference, pp. 426–435 (2014)
43. Kira, K.; Rendell, L.: A practical approach to feature selection. In: 9th International Workshop On Machine Learning, pp. 249–256 (1992)
44. Rahman, S.; Mutalib, S.; Khanafi, N.; Ali, A.: Exploring feature selection and support vector machine in Text Categorization. In: IEEE, 16th International Conference on Computational Science and Engineering, pp. 1101–1104 (2013)
45. Markov, Z.; Russell, I.: An introduction to the WEKA data mining system. ACM SIGCSE Bull. **38**(3), 367–368 (2006)
46. Lessmann, S.; Baesens, B.; Mues, C.; Pietsch, S.: Benchmarking classification models for software defect prediction: a proposed framework and novel findings. IEEE Trans. Softw. Eng. **34**(4), 485–496 (2008)
47. Santana, E.; Giraldo, G.: A knowledge-based expert system for scheduling in services systems. Appl. Comput. Sci. Eng. **657**, 212–224 (2016)
48. Pota, M.; Esposito, M.; Pietro, G.: Designing rule-based fuzzy systems for classification in medicine. Knowl. Based Syst. **124**(c), 105–132 (2017)
49. MacQueen J.: Some methods for classification and analysis of multivariate observations. In: 5th Berkeley Symposium on Mathematical Statistics and Probability, vol. I., pp. 281–297 (1967)
50. Pota, M.; Esposito, M.: degrees of freedom and advantages of different rule-based fuzzy systems. Adv. Appl. Pure Math. **29**, 107–114 (2014)
51. Kaur, A.; Kaur, A.: comparison of Mamdani-type and Sugeno-type fuzzy inference systems for air conditioning system. Int. J. Soft Comput. Eng. IJSCE **2**(2), 323–325 (2012)
52. Khan, D.; Zuhra, F.; Khan, F.; Mubarak, A.: Evaluation of diagnostic accuracy of APRI for prediction of fibrosis in hepatitis C patients. J. Ayub Med. Coll. Abbottabad **20**, 122–126 (2008)
53. Burman, P.: Estimation of optimal transformations using v-fold cross-validation and repeated learning-testing methods. Indian J. Stat. **52**(3), 314–345 (1990)
54. Kohavi, R.: A study of cross-validation and bootstrap for accuracy estimation and model selection. In: Proceeding of the Fourteenth International Joint Conference on Artificial Intelligence, vol. **2**, no. (12), pp. 1137–1143 (1995)