

Medical diagnosis with the aid of using fuzzy logic and intuitionistic fuzzy logic

Satyajit Das¹ · Debashree Guha¹ · Bapi Dutta¹

Published online: 21 May 2016 © Springer Science+Business Media New York 2016

Abstract The objective of the present study is to develop/establish a web-based medical diagnostic support system (MDSS) by which health care support can be provided for people living in rural areas of a country. In this respect, this research provides a novel approach for medical diagnosis driven by integrating fuzzy and intuitionistic fuzzy (IF) frameworks. Subsequently, based on the proposed approach a web-based MDSS is developed. The proposed MDSS comprises of a knowledge base (KB) and intuitionistic fuzzy inference system (IFIS). Based on the observation that medical data cannot be described with both precision and certainty, a medical KB is constructed in the form of a set of if-then decision rules by employing both fuzzy and IF logics. After constructing the medical KB, a new set of patients is considered for diagnosing the diseases. For each patient, linguistic values of the patients' symptoms are considered as inputs of the proposed IFIS and modeled by using the generalized triangular membership functions. Subsequently, integrated fuzzy and IF rule-based inference system is used to find a valid conclusion for the new set of patients. In a nutshell, in this paper fuzzy rule-based and IFS based inference systems are combined for better and more realistic representation of uncertainty of the medical diagnosis problem and for more accurate diagnostic result. The method is composed of following four steps: (1) the modeling of antecedent part of the rules, which consist of linguistic assessments of the patients' symptoms provided by the doctors/medical experts with their corresponding confidence levels, by using generalized fuzzy numbers; (2) the modeling of consequent part, which reveals the degree of association and the degree of non-association of diseases into the patient, by using IFSs; (3) the use of IF aggregation operator in inference process; (4) the application of relative closeness function to find the final crisp output for a given diagnosis. Finally, the applicability of the proposed approach is illustrated with a suitable case study. This article has also justified the proposed approach by using similarity measurement.

Keywords Intuitionistic fuzzy set · Linguistic variable · Medical knowledge base · Fuzzy reasoning scheme · Medical diagnosis

1 Introduction

Health care access, affordability and quality are problems for developing countries of the world and a large number of individuals do not receive the quality care that they need [1]. Generally, rural public health systems across the country face difficulty in attracting, retaining, and ensuring the regular presence of highly trained medical professionals. It is well known that many doctors are not willing to work in the rural areas due to lack of facilities, even if they are paid high salaries. Nevertheless, many health problems that rural people suffer from are preventable and easily treatable [2]. In this respect, a simple computer-aided symptom-based diagnostic application is needed to improve the delivery of health services in rural areas, to guide the paramedic/nurse in handling common ailments directly by administering simple remedies; and only refer to secondary care for the more complex problems. In recent times, these are increasingly possible due to changes in information and

Debashree Guha deb1711@yahoo.co.in

¹ Department of Mathematics, IIT Patna, Patna 800013, India

medical technologies. Over the last decades, there have been numerous implementations of computational intelligence in medicine [3–5] for diagnosing a disease. Computer technology can be used to organize, store and retrieve medical knowledge base (KB) as needed and, thus, can improve the accessibility of curative, preventive and promotive health services [6] to poor and unreached groups.

In spite of the increasing scientific evolution of both information technology and medical science, the inherent imprecision of the later makes the amalgamation of these technologies a rather difficult task. The main sources of this natural uncertainty are insufficient understanding of biological systems and their interconnections, the ambiguity of medical results and measurements [7]. As complex decision making, disease diagnosis involves a number of symptom analysis, and in this process, we have to face uncertainty which essentially lies in human expressions. For example, suppose at a consultation session [8] to evaluate a set of symptoms for disease diagnosing purpose, a doctor asks a patient about his(er) condition. When a patient describes his(er) condition, there may be inexactness of information about source of disease, period of suffering or the symptoms of disease. The doctor has to deal with this kind of incomplete and imperfect information. Therefore, in medical diagnosis problems, the primary task may be regarded as the proper management of incomplete information. Basically, the natural evolution of various diseases and the uncertain nature of medical data require a consistent framework that can model the uncertainty or imprecision, intrinsically connected to the medical problem, by using variables and multiple class memberships and facilitate approximate reasoning. This evidently makes the fuzzy logic (FL) a valuable tool for representing medical concepts by treating them as fuzzy sets [9]. Researchers have made several advancements regarding the applications of fuzzy approach in the field of medical diagnosis problems [10–14]. Fuzzy set assigns to each member of the universe of discourse a degree of membership between zero and one. By considering the non-membership degree to the fuzzy set, Atanassov [15] introduced intuitionistic fuzzy set (IFS) which implicates the fact that the non-membership degree is not always compliment to the membership degree. There may arise some hesitations. Thus, there are some cases where IFS theory is more suitable to deal with incomplete or inexact information present in real-world applications [16-18]. In view of this, IFS theory has been employed to cope with uncertainty in medical diagnosis problems successfully [16, 19–21].

Based upon the observation that medical data are often imprecise, incomplete and vague so that using standalone IFS may not improve the accuracy of diagnosis, the aim of this present study is to propose a new algorithmic approach for solving medical diagnosis problem by utilizing both fuzzy and intuitionistic fuzzy frameworks. We will provide a brief overview of the relevant works for the medical diagnosis problems in Section 2, including the motivation and ideas of the proposed approach.

The paper is planned as follows: Section 3 describes some useful results that are used in our proposed medical diagnostic support system (MDSS). The proposed MDSS is presented in Section 4. In Section 5, a numerical example is presented. In Section 6, a discussion about significance and usefulness of the proposed approach is described. Further, a similarity measurement is applied to validate the diagnostic results of our approach. A sensitivity analysis is also conducted in this Section to examine the impact of small perturbations in the values of confidence level of the doctor regarding evaluation of patients' symptoms, on the diagnostic results. A web-based tool for MDSS is introduced in Section 7 while Section 8 concludes the discussion.

2 Related works and motivation

In this section, a brief review of the related works is carried out followed by a discussion about the motivation behind this work.

2.1 Related works

Nowadays, fuzzy rule-based model has been widely and successfully used in medical reasoning [14, 22–24]. As per our knowledge, several rule-based medical diagnosis systems have been published. Some of those works are given below.

By utilizing fuzzy set theory and fuzzy production rules, a weighted fuzzy reasoning algorithm for handling medical diagnostic problems was introduced in [22]. In 2001, Yao and Yao [23] proposed a fuzzy decision making method for medical diagnosis on the basis of fuzzy number and compositional rule of inference. Moreover, by utilizing fuzzy classification trees, Chiang et al. [25] proposed a medical decision support system for colon polyp screening. For the management of tropical diseases, Obot and Uzoka [24] used fuzzy rule base framework. A neuro-fuzzy rule base network methodology was applied in medical diagnosis in [26]. In 2009, Gadaras and Mikhailov described a methodology of medical diagnosis by using fuzzy rule-based classification technique [7]. Maio et al. [27] presented a medical diagnosis system based on an ontology model, which represents a relation between medical disease and it's symptomatology in a qualitative manner by using fuzzy labels. Fenza et al. [28] developed an architecture for context-aware service discovery in the health care domain that exploits synergy among intelligent agent technology, semantic web models and computational intelligence techniques. Aribarg et al. [29] introduced a classification method by using simulated annealing for efficient medical diagnosis. An effective heart disease diagnosis was proposed in [30] with the help of framing fuzzy rules by using support sets. By using fuzzy inference system, Singh et al. [31] diagnosed the arthritis disease. A web-based decision support system by using fuzzy rule base for the diagnosis of typhoid fever was developed by Samuel et al. [14]. Caponetti et al. [32] proposed an approach based on fuzzy mathematical morphology to delineate the segment images of human oocyte from the entire image. Furthermore, fuzzy rule base system was used to solve lung disease [33], breast cancer [34], abdomen pain [35], disease in human brain [36], malaria [37], vision defects [38], etc.

On the other hand, as mentioned in the introduction, among all the higher order fuzzy sets, IFS has been used effectively in medical diagnosis problems, due to its capability in modeling hesitation of human mind. For example, in 2001, De et al. [16] formulated a medical diagnosis system by using composition of intuitionistic fuzzy relation (IFR). Kharal [19] presented a homeopathic drug selection by using the same composition rule. Szmidt and Kacprzyk [21] proposed a medical diagnosis process by distance measures of IFSs. Further, they used similarity measure [39] of IFSs for supporting medical diagnostic reasoning. Again in [40–42] several medical diagnostic approaches were developed based on similarity measures under intuitionistic fuzzy environment. Applying an improved intuitionistic fuzzy cross-entropy approach, Hung [43] solved a medical diagnosis problem. In 2012, Zhang et al. [44] presented a new medical diagnosis process by utilizing score functions on IFSs with double parameter. Hung and Tuan [20] pointed out some shortcomings of [16] by providing appropriate examples. Recently, few research articles are published in this domain, where disease diagnosis is performed based on intuitionistic fuzzy soft sets [45, 46] and type-2 fuzzy sets [47, 48] based operations.

2.2 The motivation and ideas

Now although, as evident from the above literature review, research articles of medical diagnosis problems are available on purely fuzzy rule-based methods and purely IFS based models. The amalgamation of these two existing approaches is an issue that remains looked into. The limitations of IFSs based medical diagnosis problems are that the information of previous diagnosis about patients are not utilized in most of the works. On the other hand, intrinsic uncertainty of medical problems cannot be tackled solely by fuzzy rule-based models. Thus, the next step of evolution in solving medical diagnosis problem, requires a paradigm that combines both fuzzy rule-based and IFS based approaches. However, the problem of integrating these two approaches in medical diagnosis problem setting is not yet attended to so well and this is the aspect that has motivated the present work. In other words, our idea in this article is proposing a hybrid method between fuzzy rule-based system and IFS based approach, to solve medical diagnosis problem, which can improve the effectiveness of medical treatment as well as improve the accuracy of the diagnosis.

For this purpose, first a medical KB is required to utilize the previous diagnoses of patients. In the process of making the medical KB, it is realized that in the evaluation of a set of symptoms for disease diagnosis purpose, medical expert may feel comfortable to express his(er) opinion in natural language. The confidence level [49] of the doctor in his(er) own opinion is also an important aspect. For example, in a hospital two doctors are consulting to assess a patient's current physical condition. One expert says "I have 'full confidence' that (s)he will be recovered albeit slowly". Another says that "I am 'not so confident' how his(er) recovery will be". It is to be noted here that, the diagnosis of a patient may be dependent on the confidence level of a doctor. In such scenario, generalized fuzzy number (GFN) [50] is a more appropriate tool to represent linguistic opinion of the doctor as it considers the degree of confidence of doctor's opinion. With this viewpoint, in this paper, a medical KB is constructed in the form of a set of fuzzy decision rules. To form the KB, five symptoms, such as, Temperature, Headache, Stomach pain, Cough and Chest pain and five diseases, such as, Viral Fever, Malaria, Typhoid, Stomach problem and Chest problem are considered. The antecedent part of the rules consist of linguistic evaluation of patients' symptoms with different degrees of confidences of the doctors and the consequent part represents the degree of association and non-association of diseases into the patients. To model this type of situation, we propose a KB in which the variables in the antecedent part are linguistic and they are quantified by GFNs. This means that the values of the patients' symptoms are expressed in natural language term which easily represent the knowledge of the medical experts. The variables in the consequent part are quantified by IFSs, which consist of membership and non-membership degrees. The membership degree captures the degree of association of the disease to the patient and non-membership degree represents the degree of nonassociation of the disease to the patient. Thus, both FL and intuitionistic fuzzy logic (IFL) provide a rigorous framework for verbal representation of medical concepts that can be then embedded in meaningful fuzzy rules. Such rules can be easily absorbed, verified, adapted and possibly expanded by medical experts, and used for the development of MDSS, which could be valuable in the process of medical diagnosis. Then the new input data sets from patients are collected as their observed symptoms. To determine a proper diagnosis for each patient with the given values of tested symptoms, an appropriate intuitionistic fuzzy inference system (IFIS) based on the proposed medical KB is used.

Finally, the research proposes a web-based MDSS by using our proposed algorithm for the primary diagnosis of diseases, such as, *Malaria, Viral fever, Typhoid, Stomach problem and Chest problem* of patients having *Temperature, Headache, Stomach pain, Cough and Chest pain* as symptoms. The system is developed with the aim of improving the accessibility of the primary health care facilities, particularly to improve accessibility for the poor and unreached groups of the rural areas in developing countries of the world, where there are shortages of doctors and lack of proper diagnostic infrastructure. The proposed system will assist medical-personnel, especially in rural areas, to ensure in providing quality at all levels of health and medical care services.

3 Basic concepts

Due to inherent vagueness in medical diagnosis problems, medical experts always feel comfortable to express their opinions in natural languages. So, the first task is to quantify these linguistic expressions. Since, in this present study, during development of medical KB the other aim is to capture the degree of confidences of experts' opinions, naturally the most suitable approach to quantify the opinions is with the help of GFNs. Further, in this work IFSs are used to model the possibility and impossibility of diseases into the patient (a brief justification and deeper discussion for using GFNs and IFSs are provided in Section 2.2). Thus, a primary discussion of various concepts, including fuzzy number, GFN, generalized triangular fuzzy number (GTFN) and IFS are mathematically introduced in Appendix A so as to facilitate further study.

The following section establishes a convenient way of constructing generalized triangular membership function [51] for linguistic values of decision variables involved in the antecedent part of the rule base.

3.1 Membership functions for the linguistic value of decision variable

In this section, we revise the concepts of linguistic variable and its representation. In this article, the medical KB is designed in terms of if-then rules, where the decision variables in antecedent part of the rules expressed by linguistic evaluation of patients' symptoms and these variables can be treated as linguistic variables [52]. A linguistic variable may be regarded either as a variable whose value is a fuzzy number or as a variable whose values are defined in linguistic terms. Let *u* be the name of the linguistic variable and T(u) is the term set of *u*, i.e., the set of names of linguistic values of *u* with each value being a fuzzy number defined on *U*. In this work, it is assumed that the values of each of the linguistic variables, namely, $u_1, u_2, ..., u_n$ are defined in the interval $[a, b] \subset \mathbb{R}$. Let U = [a, b] and the term set T(u) consists of $K + 1(K \ge 2)$ terms as follows:

$$T(u) = \{low_1, around(a + \beta), around(a + 2\beta),, around(a + (K - 1)\beta), high_K\}$$

where $\beta = (b - a)/K$ and each term may be represented [51] with the help of generalized triangular membership functions { $\mu_{\tilde{A}_1}(u), \mu_{\tilde{A}_2}(u), ..., \mu_{\tilde{A}_{K+1}}(u)$ } (Fig. 1) of the following form:

$$\mu_{\widetilde{A}_1}(u) = \mu_{low_1}(u) = \begin{cases} \frac{(b-u)w_1}{(b-a)}, \text{ for } a \le u \le b, \\ 0, & \text{otherwise} \end{cases}$$
(1)

The GTFN \widetilde{A}_1 can be expressed as $\widetilde{A}_1 = [(a, a, b); w_1].$

$$\mu_{\widetilde{A}_k}(u) = \mu_{around(a+k\beta)}(u)$$

$$= \begin{cases} \frac{(u-a)w_k}{k\beta}, & \text{for } a \le u \le a+k\beta, \\ w_k, & \text{for } u = a+k\beta, \\ \frac{(b-u)w_k}{(b-a-k\beta)}, & \text{for } a+k\beta \le u \le b, \\ 0, & \text{otherwise} \end{cases}$$
(2)

For $1 \le k \le K - 1$ and each of the corresponding GTFN \widetilde{A}_k is denoted as $\widetilde{A}_k = [(a, a + k\beta, b); w_k]$ and

$$\mu_{\widetilde{A}_{k+1}}(u) = \mu_{high_k}(u) = \begin{cases} \frac{(u-a)w_{k+1}}{(b-a)}, & \text{for } a \le u \le b, \\ 0, & \text{otherwise} \end{cases}$$
(3)

The GTFN \widetilde{A}_{k+1} is denoted as $\widetilde{A}_{k+1} = [(a, b, b); w_{k+1}]$. Where $w_1, w_2, ..., w_{k+1} (\in [0, 1])$ denote the corresponding confidence level of the expert.

The novelty of this term set arrangement is that while very little knowledge is available about the boundaries



Fig. 1 Membership function for the linguistic variable with $\left(k+1\right)$ terms when w=1

 Table 1
 The term set

Term set	Fuzzy numbers
Very Low(VL)	(0,0,1)
Low(L)	(0,0.25,1)
Medium(M)	(0,0.5,1)
High(H)	(0,0.75,1)
Very High(VH)	(0,1,1)

of individual term, each term is stretched over the whole domain, though the mid values of each of the terms are situated at a fixed distance apart.

For example, we consider a term set as $T(u) = \{\text{Very low}, \text{Low, Medium, High, Very high}\}$ where each term in T(u) is defined by a fuzzy number in the universe of discourse [0,1]. By using (1)–(3), each term of T(u) can be transformed to associate fuzzy number as given in Table 1 (for the sake of simplicity, in Table 1 and Fig. 1, the confidence level of the expert is assumed as one).

The next section describes the aggregation operators involved in the proposed IFIS.

3.2 Logical connective operators

In fuzzy reasoning scheme to find the matching degree among the new fuzzy input data and each given rule, conjunctions and disjunctions are two basic operations. The realizations of conjunction and disjunction are facilitated via triangular norm (t-norm) and triangular conorm (tconorm), respectively, as they offer a very wide class of binary functions of conjunction and disjunction in fuzzy context [53]. Moreover, the selection of suitable t-norm and t-conorm operators depend on the user behavior towards the aggregation and specific requirement of the decision scenario [54]. The definitions of t-norm and t-conorm are given as

Definition 1 [51, 55] A t-norm T is a function from $[0, 1] \times [0, 1] \rightarrow [0, 1]$ such that it is symmetric, associative, non decreasing in each argument and $T(a, 1) = a \forall a \in [0, 1]$.

Definition 2 [55] A t-conorm S is a function from $[0, 1] \times [0, 1] \rightarrow [0, 1]$ such that it is symmetric, associative, non decreasing in each argument and $S(a, 0) = a \forall a \in [0, 1]$.

However, in the computation process of our proposed IFIS, we utilize minimum and maximum operators as these two operators are basic t-norm and t-conorm respectively.

Further, to obtain the inferred outcome of each rule and overall output of the system, we need to facilitate the operations of multiplication of IFS value by a scalar and aggregation of several IFS values. In this work, we utilize continuous Archimedean t-norm and t-conorm based addition and multiplication rules, proposed by Beliakov et al. [56], as they produce more consistent result to aggregate IFS values. In this regard, a brief description of the concepts of IFS operations and aggregation operators for IFSs based on continuous Archimedean t-norm and t-conorm, which are extensively used in the proposed IFIS, are briefly reviewed in Appendix B.

4 Proposed medical diagnosis approach

In this section, we present the proposed approach for medical diagnosis. First, we shall describe how we interpret medical diagnosis problem mathematically.

Given two sets $X = \{x_1, x_2..., x_n\}$ and $Y = \{y_1, y_2..., y_m\}$ where X is a set of symptoms and Y is a set of diseases, respectively. The values n and $m \in \mathbb{N}$ are numbers of symptoms and diseases, respectively. The interrelations among the symptoms and diseases are characterized by using a set of decision rules that constitute the medical KB. Subsequently, consider a new set, namely $P = \{P_1, P_2, ..., P_t\}$, of t number of patients with symptoms $\{x_1, x_2..., x_n\}$. The medical diagnosis problem aims to provide an accurate and timely diagnosis of the diseases $\{y_1, y_2..., y_m\}$ to the new set of patients.

4.1 Medical knowledge base

In our proposed approach, a medical KB is required, which can be expressed as a set of inference rules in the form 'if antecedent and then consequent' specifying a relationship between the inputs (symptoms) and outputs (diseases). Consequently, it can provide useful information to physicians/medical experts for better diagnosis of diseases while a new set of symptoms is observed. The medical KB can be carefully formulated after a detailed discussion about the various aspects of Viral Fever, Malaria, Typhoid etc., with the concerned medical experts.

During the discussion with the medical experts, it is realized that to evaluate a set of symptoms, when doctor asks a patient regarding his(er) condition, s(he) (i.e., patient) is not to be so confident to describe the conditions. The confidence level of the patient is inherently involved in the linguistic expressions which are used for expressing their conditions. As a result, when medical expert provides the judgments based on the patient's symptoms, the confidence level of the patient is intrinsically connected to the expert's advice. Such linguistic expressions can be then suitably captured by GFN as it considers the degree of confidence of expert's opinions. It is also worth to mention that sometimes an expert is also in hesitation to provide judgments regarding the possibility of disease into a patient. These observations form the background of the present study where the medical KB is established by using a set of decision rules based on FL and IFL and thus, the expertise of the doctor is captured when diagnosing the patient.

In the decision rules that constitute the medical KB, symptoms are considered in antecedent part and diseases in consequent part. The antecedent part of the rules involve linguistic evaluation of patient's symptoms associated with the degree of confidence of the doctor and the consequent part reveals the state of the patient in terms of diagnosis. For example, consider the following system of p fuzzy rules

$$R_i : If x_1 is \widetilde{A}_{i1} and x_2 is \widetilde{A}_{i2}, ..., x_n is \widetilde{A}_{in}$$

then possibility of y_1 is C_{1i}, y_2 is $C_{2i}, ..., y_m$ is C_{mi} . (4)

where i = 1, 2, ..., p (number of rules), GTFNs \tilde{A}_{ij} (j = 1, 2, ..., n), representing the values of the linguistic variables { $x_1, x_2, ..., x_n$ }, are defined in the universe [0, 1]. The computation procedure of constructing membership functions of GTFNs is described in Section 3.1. The set { $y_1, y_2, ..., y_m$ } represents the set of diseases and IFSs $C_{ki} = (\mu_{y_{ki}}, \nu_{y_{ki}})(k = 1, 2, ..., m)$ represent the degrees of possibility and impossibility of disease y_k , described in i^{th} rule. More specifically, $\mu_{y_{ki}}$ represents the degree of association of disease y_k into the patient and $\nu_{y_{ki}}$ represents the degree of non-association of disease y_k into the patient.

4.2 Intuitionistic fuzzy inference system

This section presents intuitionistic fuzzy inference engine (IFIE) which is the core of the proposed IFIS. In this part, the discussed medical KB is applied to diagnose the disease for a new patient and the process is described below.

Inputs: It is assumed that the medical KB, consists of p fuzzy rules (as given in (4)), receives a new set of linguistic values of symptoms, as fuzzy inputs denoted as $\{\widetilde{U}_1, \widetilde{U}_2, ..., \widetilde{U}_n\}$.

Matching degree: The degree to which the fuzzy input data \tilde{U}_j (j = 1, 2, ..., n) matches the i^{th} rule R_i is computed by using a conjunction operator (t-norm T) and a disjunction operator (t-conorm S) as follows [57]:

$$\alpha_{i} = T(S_{x \in \widetilde{U}}(T_{j=1,2,...,n}(\mu_{\widetilde{A}_{ij}(x)}, \mu_{\widetilde{U}_{j}(x)})), i = 1, 2, ..., p$$
(5)

Output of each rule: Subsequently, the inferred outcome of each rule is defined as an intuitionistic fuzzy value based on the concept of Larsen's [58] product implication operator and computed by using Definition B.1 (Appendix B) as follows:

$$\alpha_i(\mu_{y_{ki}},\nu_{y_{ki}}) = \left(h^{-1}\left(\alpha_i h(\mu_{y_{ki}})\right), g^{-1}\left(\alpha_i g(\nu_{y_{ki}})\right)\right).$$

Larsen's [57, 58] product implication operator is used for finding the association and non-association of diseases into the patient. The reason for using such kind of implication is that, it is the most commonly used implication operator and is simpler in nature.

4.3 Aggregation of the rule outcomes

Eliciting the corresponding output obtained from each rule R_i for the fuzzy inputs $\{\widetilde{U}_1, \widetilde{U}_2, ..., \widetilde{U}_n\}$, the overall system output is computed as follows:

We apply the intuitionistic fuzzy arithmetic mean (IAM) (defined in (13)) that combines the outcome (calculated in the previous step) for each rule in the rule base. Thus, the aggregated outcome $C_k(k = 1, 2, ..., m)$ is computed as

$$C_{k} = IAM(\alpha_{1}(\mu_{y_{k1}}, \nu_{y_{k1}}), \alpha_{2}(\mu_{y_{k2}}, \nu_{y_{k2}}), ..., \alpha_{p}(\mu_{y_{kp}}, \nu_{y_{kp}}))$$

$$= \left(1 - \frac{1}{p} \sum_{i=1}^{p} \alpha_{i}(1 - \mu_{y_{ki}}), \frac{1}{p} \sum_{i=1}^{p} \alpha_{i}\nu_{y_{ki}}\right)$$

$$= (\mu_{y_{k}}, \nu_{y_{k}})(k = 1, 2, ..., m)$$
(6)

where,

$$\mu_{y_k} = 1 - \frac{1}{p} \sum_{i=1}^{p} \alpha_i (1 - \mu_{y_{ki}})$$
⁽⁷⁾

and

$$\nu_{y_k} = \frac{1}{p} \sum_{i=1}^{p} \alpha_i \nu_{y_{ki}}$$
(8)

Here, $p = \text{total number of rules in the rule base, } C_k$ is the aggregated outcome for the *k*th disease $y_k (k = 1, 2, ..., m)$.

More specifically, μ_{y_k} and ν_{y_k} are averages of strength and non-strength of kth (k = 1, 2, ..., m) disease into a patient, and computed by (7) and (8), respectively. Hence, μ_{y_k} reveals the degree of association of disease y_k into a patient and ν_{y_k} reveals the degree of non-association of disease y_k into a patient.

Thus, the association and non-association of the diseases, namely $y_1, ..., y_m$ into the patient are obtained. Then final diagnosis of the patient needs to be evaluated. For this purpose, a process is required which translates the output of the IFIE into crisp values, and this crisp diagnostic result is mostly expected by medical experts for proper analysis and explanation. For this purpose, we introduce the 'most possible disease' and we apply a relative closeness function Fover the aggregated outcome with respect to each disease. This function will determine the relative closeness of each disease with respect to the 'most possible disease'.

4.4 Final crisp output for finding the accurate diagnostic result

First, we shall describe how we interpret 'most possible disease' mathematically. If the possibility of a disease is represented by IFS $I^+ = (1, 0)$, then it implies that the possibility of attacking that disease is cent percent, i.e., the patient is suffering from that particular disease. Similarly, if the possibility of a disease is represented by IFS $I^- = (0, 1)$, then it implies that the non-possibility of attacking that disease is cent percent, i.e., the patient disease is cent percent, i.e., the possibility of attacking that the non-possibility of attacking that disease is cent percent, i.e., the patient is not suffering from that particular disease.

In order to find out the final diagnosis, the relative closeness coefficient [59] for each disease with respect to 'most possible diseases' can be computed as follows:

$$F(C_k) = \frac{d(C_k, I^-)}{d(C_k, I^-) + d(C_k, I^+)} (k = 1, 2, ..., m)$$
(9)

where, $d(C_k, I^+) = \sqrt{(\mu_{y_k} - 1)^2 + (\nu_{y_k} - 0)^2}$, $d(C_k, I^-) = \sqrt{(\mu_{y_k} - 0)^2 + (\nu_{y_k} - 1)^2}$ are computed by using Euclidean distances of C_k from $I^+ = (1, 0)$ and $I^- = (0, 1)$, respectively.

Then the patient's diseases can be determined according to greater relative closeness coefficient value defined in (9). However, for any two diseases y_1 and y_2 , if the corresponding relative closeness coefficient values are equal and maximum, i.e., $F(C_1) = F(C_2)$, then it implies that the patient may possess both the diseases y_1 and y_2 .

A brief flow-chart of the proposed medical diagnosis process is shown in Fig. 2.

5 Case study

Suppose in a given pathology, there are four patients, namely Arka, Bumba, Chandra and Dip whose symptoms are recorded by a routine case taking practice. We use the proposed medical diagnostic approach for diagnosing some diseases, such as, *Malaria, Viral fever, Typhoid, Stomach problem and Chest problem.* For these four patients, we also consider *Temperature, Headache, Stomach pain, Cough and Chest pain* as symptoms for the aforementioned diseases.

Our next aim is to construct a medical KB in the form of rule base. In literature, there are many ways of framing the rules. One of the popular ways to formulate the rules is using expert's knowledge [14, 35, 37] and this process is widely used in many applications. In this context, to construct the rule base, the required data set is collected via survey from a local government hospital at Patna, India, after a consultation with the concerned medical experts about the various aspects of the aforementioned diseases. We consider this data as a primary level of our survey. However, in future



Fig. 2 Flowchart of the inference system

our aim is to modify this data via survey in sub-centres, rural family welfare-centre, block primary health-centre and other health-centres in the state.

Based on the collected data, a medical KB is constructed by using ten fuzzy decision rules $(\mathbf{R_1} - \mathbf{R_{10}})$ which are described below.

- R1 If Temp is H with confidence level (CL) 0.7 and Hdc is M with CL 0.8 and Stp is VL with CL 0.8 and Cough is VH with CL 0.5 and Chp is L with CL 0.9 then possibility of Vfev is (0.9,0.0) and Mal is (0.5,0.3) and Typ is (0.4,0.6) and Stpr is (0.3,0.7) and Chpr is (0.1,0.8)
- R₂ If Temp is H with CL 0.6 and Hdc is L with CL 1.0 and Stp is VL with CL 0.8 and Cough is M with CL 0.8 and Chp is VH with CL 0.3 then possibility of Vfev is (0.8,0.0) and Mal is (0.3,0.7) and Typ is (0.8,0.2) and Stpr is (0.1,0.8) and Chpr is (0.1,0.9)
- R₃ If Temp is VH with CL 0.9 and Hdc is M with CL 0.4 and Stp is VL with CL 1.0 and Cough is H with CL 1.0 and Chp is L with CL 0.7 then possibility of Vfev is (0.6,0.3) and Mal is (1.0,0.0) and Typ is (0.2,0.6) and Stpr is (0.1,0.7) and Chpr is (0.1,0.9)



Fig. 3 Matching degree calculation from Rule-1

- R₄ If Temp is VH with CL 0.7 and Hdc is L with CL 0.8 and Stp is VL with CL 1.0 and Cough is M with CL 1.0 and Chp is H with CL 0.7 then possibility of Vfev is (0.6,0.3) and Mal is (0.8,0.0) and Typ is (0.2,0.5) and Stpr is (0.1,0.8) and Chpr is (0.2,0.8)
- R₅ If Temp is M with CL 0.9 and Hdc is VH with CL 0.7 and Stp is H with CL 0.8 and Cough is L with CL 0.9 and Chp is VL with CL 0.9 then possibility of Vfev is (0.5,0.4) and Mal is (0.5,0.3) and Typ is (0.9,0.0) and Stpr is (0.1,0.7) and Chpr is (0.0,0.9)
- R₆ If Temp is M with CL 0.7 and Hdc is VH with CL 0.5 and Stp is L with CL 0.8 and Cough is H with CL 0.4 and Chp is VL with CL 0.8 then possibility of Vfev is (0.4,0.4) and Mal is (0.3,0.5) and Typ is (0.8,0.1) and Stpr is (0.3,0.7) and Chpr is (0.1,0.9)
- R₇ If Temp is VL with CL 0.8 and Hdc is L with CL 0.9 and Stp is VH with CL 0.9 and Cough is H with CL 0.3 and Chp is M with CL 0.5 then possibility of Vfev is (0.1,0.7) and Mal is (0.3,0.5) and Typ is (0.4,0.4) and Stpr is (1.0,0.0) and Chpr is (0.3,0.5)
- R₈ If Temp is VL with CL 0.7 and Hdc is M with CL 0.4 and Stp is VH with CL 0.8 and Cough is L with CL 0.5 and Chp is H with CL 0.4 then possibility of Vfev is (0.2,0.7) and Mal is (0.1,0.7) and Typ is (0.2,0.7) and Stpr is (0.8,0.1) and Chpr is (0.5,0.4)
- **R9** If Temp is L with CL 0.7 and Hdc is VL with CL 1.0 and Stp is M with CL 0.5 and Cough is H with CL 0.4 and Chp is VH with CL 0.9 then possibility of Vfev is (0.1,0.6) and Mal is (0.2,0.6) and Typ is (0.3,0.6) and Stpr is (0.2,0.8) and Chpr is (1.0,0.0)
- R₁₀ If Temp is L with CL 0.5 and Hdc is VL with CL 0.9 and Stp is H with CL 0.6 and Cough is M with CL 0.7

and Chp is VH with CL 0.7 then possibility of Vfev is (0.2,0.7) and Mal is (0.1,0.6) and Typ is (0.2,0.6) and Stpr is (0.4,0.5) and Chpr is (0.8,0.1)

where, VL=Very Low, L=Low, M=Medium, H=High, VH=Very High; Temp=Temperature, Hdc=Headache, Stp=Stomach Pain, Chp=Chest Pain; Vfev=Viral fever, Mal=Malaria, Typ=Typhoid, Stpr=Stomach problem and Chpr=Chest problem and CL=Confidence level.

As mentioned earlier, in the present study the rules that constitute the rule base are carefully formulated with the assistance of medical experts from a local government hospital at Patna, namely, Mahavir Vaatsalya Aspatal, Patna, India. The fuzzy rules can be further refined by embodying doctor/medical experts' domain knowledge, who may decide to modify or delete some rules, or even add new one [60].

In our rule base, to express antecedent part of the rules, a linguistic term set having five polar terms, such as, Very Low, Low, Medium, High and Very High is utilized. The graphical representations of antecedent parts of rule R_1 are shown in Fig. 3, where, the blue line represents the clauses of antecedents; the black dotted line represents the input, whereas the green for matching degree.

Fuzzy input data (i.e., linguistic evaluations of new symptoms) for four patients are given in Table 2.

To compute the matching degree, we consider t-norm T and t-conorm S as the minimum operator and maximum operator, respectively. The calculated matching degree for each rule corresponding to the aforementioned patients' symptoms is shown in Table 3.

The average of each rule's output for every patient is computed by using (6) and the results are given in Table 4.

Patient Temperature Headache Stomach pain Cough Che	
	est-pain
Arka (VH, 0.9) (H, 1) (M, 0.4) (L, 0.8) (VL	L, 1)
Bumba (VL, 1) (M, 0.9) (H, 0.8) (VL, 0.9) (L,	1)
Chandra (VH, 0.7) (H, 0.8) (L, 0.9) (H, 0.8) (VI	L, 0.9)
Dip (H, 0.8) (M, 1) (M, 0.7) (VH, 0.6) (L,	1)

 Table 2
 Input data

Table 3Matching degree ofthe rules

	Arka	Bumba	Chandra	Dip
α_1	0.3404	0.32	0.4324	0.5
α_2	0.2307	0.245	0.225	0.24489
α3	0.5	0.4	0.4571	0.4
α_4	0.4827	0.307	0.4613	0.4
α_5	0.2857	0.3	0.3	0.3111
α_6	0.3555	0.334	0.4	0.4
α_7	0.2909	0.276	0.3	0.4
α ₈	0.3478	0.381	0.35	0.31578
α9	0.3555	0.3348	0.399	0.4363
α_{10}	0.3414	0.459	0.34	0.4102

The relative closeness coefficient value with respect to each disease for each patient is calculated by using (9) and given in Table 5.

The normalized relative closeness coefficient value of each disease for each patient is shown in Table 6.

Decision From Tables 5 and 6, it is clear that, Arka suffers from malaria whereas Bumba, Chandra and Dip suffer from viral fever.

6 Further analysis

6.1 Some important observations towards the proposed approach

Some important analyzes of the novelty and significance to the proposed method are summarized below.

- This study encompasses the idea of both fuzzy rulebased and IFSs based medical diagnosis models to enhance the accuracy of the final diagnostic result. Furthermore, the proposed method could handle all the issues of the standalone fuzzy rule and IFS based methods.
- From the mentioned literature survey (Section 2.1), it is observed that in the medical KB, none has considered the idea of capturing confidence level of the medical experts in their opinions. The present work uses the idea for the first time, where the medical experts' confidence

levels are taken into consideration in the formulation of medical KB. Further, the proposed method makes the diagnosis more flexible by modeling the decision part (i.e., consequent part) by using IFS. An IFS can model the chance of occurring and non-occurring of a disease into a patient and this is very close to the real-life situations. Thus, the proposed method can also be regarded as a consistent and meaningful framework to address medical diagnosis problem.

- In this work, to capture the confidence levels of medical experts, the values of each of the linguistic variables in the antecedent are represented by using GFNs. It is worth noticing that, each GFN [51] in a term set is stretched over the whole domain [0,1]. Due to the same base of each of the GFNs, every rule provides a nonzero matching degree (for the non-zero fuzzy input) and it ensures that each and every rule could be fired and consequently, it provides better efficiency to the output. This research intends to address this issue, in an effort to provide a more accurate diagnosis.
- In this paper, we have used t-norm, t-conorm and intuitionistic fuzzy aggregation operator in the IFIS, which means that our method has a solid theoretical foundation. The method is simple in terms of computation compared to the other existing methods.
- The computational complexity of the proposed method is evaluated as follows: if there are 'n' number of symptoms, 'p' number of rules and 'm' number of diseases, then our algorithm requires ((3n - 1)p + (5p + 1)m +10m' arithmetic operations for producing diagnostic

Table 4	Average	output
---------	---------	--------

Patient	<i>C</i> ₁	<i>C</i> ₂	<i>C</i> ₃	C_4	<i>C</i> ₅
Arka	(0.8037,0.1451)	(0.8072,0.1334)	(0.7889,0.1605)	(0.7605,0.2098)	(0.7598,0.2212)
Bumba	(0.8064, 0.1448)	(0.8006,0.1405)	(0.8025, 0.1521)	(0.7796,0.1925)	(0.7810,0.1989)
Chandra	(0.8067,0.1418)	(0.8000,0.1416)	(0.8026, 0.1541)	(0.7735,0.1999)	(0.7702,0.2096)
Dip	(0.7827,0.1593)	(0.7768, 0.1544)	(0.7725,0.1749)	(0.7436,0.2278)	(0.7410,0.2357)

Patient\disease	Vfev	Mal	Тур	Stpr	Chpr
Arka	0.8278	0.8347	0.8129	0.7750	0.7692
Bumba	0.8294	0.8180	0.8240	0.7932	0.7909
Chandra	0.8301	0.8272	0.8232	0.7865	0.7801
Dip	0.8100	0.8080	0.7975	0.7576	0.7525

Table 5 Relative closeness coefficient value

result. In addition, for a fixed number of rules, with the increase in number of symptoms and diseases, the computational burden to generate diagnostic results grows in linear order. So, there will be no big change in the computational complexity in the case of more diseases and symptoms as the complexity of our algorithm is of polynomial order.

6.2 Validation of results using similarity measure

The proposed method makes the diagnosis flexible by accommodating the idea of confidence level and incorporating the concept of IFS during the diagnostic process. As per our knowledge, none has considered these ideas, therefore, direct comparison to any other method is not possible. However, in an attempt to assess the accuracy of the results of our proposed method, we will apply similarity measurements in this section.

Researchers have introduced various similarity measures [50, 61–63] for GTFNs. One of such measures, employed in this work to validate the results of our case study, is Hwang and Yang's [62] similarity measure.

In our medical KB, variables in the antecedent part of each rule are represented as GTFNs. The new inputs (symptoms of new patients) are also modeled as GTFNs. First, the similarity between antecedent of each rule and input data is computed by utilizing similarity measure. For instance, suppose we have a rule

$$R_1 : If X_1 is A_{11}, X_2 is A_{12}, ..., X_5 is A_{15} then Y_1 is C_{11}, ..., Y_5 is C_{15}$$

and one new input with five clauses $P_{11}, P_{12}, ..., P_{15}$. Now find the similarity values, such as, $S(\widetilde{A}_{11}, P_{11}), S(\widetilde{A}_{12}, P_{12}), ..., S(\widetilde{A}_{15}, P_{15})$ by [62]. Then the overall similarity between the set of new inputs and

 Table 6
 Normalized relative closeness coefficient value

antecedent of rule R_1 is computed as

$$S(R_1) = \frac{S(\widetilde{A}_{11}, P_{11}) + S(\widetilde{A}_{12}, P_{12}) + \dots + S(\widetilde{A}_{15}, P_{15})}{5}$$
(10)

In this way, we can compute the overall similarity between the set of new inputs and antecedent of each rule R_i (i = 1, 2, ..., p). Then, by observing the similarity values, one particular rule is selected which has maximum overall similarity. Then the consequent part of the selected rule is considered for diagnosis of the disease for the new set of inputs. For this purpose, based on the consequent part of the selected rule, the relative closeness coefficient value for each disease is computed in a way similar to the method described in Section 4.4.

After applying, the similarity measurement method to our case study (mentioned in Section 5), the obtained diagnostic results are described in Table 7.

From the Table 7, it is observed that, Arka suffers from malaria, Bumba suffers from stomach problem, and Chandra and Dip suffer from viral fever. This result is almost same as described in our proposed approach.

Note Similarity measurement method allows to validate the results of the medical diagnosis problem given in the case study. However, by using this method strength of activation of the if-part of each rule for new input data can not be computed. Thus, the standalone similarity measurement method may not improve the accuracy of final diagnosis. With this observation, similarity measurement method can be viewed as a validation approach, wherein our aim is to get an idea for estimating the possibility of diseases intuitively by observing the maximum similarity value between the new input data and the given rules.

Patient\disease	Vfev	Mal	Тур	Stpr	Chpr
Arka	0.2059	0.2077	0.2022	0.1928	0.1914
Bumba	0.2045	0.2017	0.2032	0.1957	0.1949
Chandra	0.2051	0.2044	0.2034	0.1943	0.1928
Dip	0.2063	0.2058	0.2032	0.1930	0.1917

 Table 7 Results by similarity measure

Patient	Selected	Closeness co	Closeness coefficient				
	rule	Vfev	Mal	Тур	Stpr	Chpr	
Arka	Rule-4	0.6484	0.8649	0.2533	0.1566	0.1566	Malaria
Bumba	Rule-8	0.2533	0.2171	0.2533	0.8434	0.5495	Stomach problem
Chandra	Rule-1	1.0000	0.7167	0.6484	0.2171	0.1566	Viral fever
Dip	Rule-1	1.0000	0.7167	0.6484	0.2171	0.1566	Viral fever

6.3 Sensitivity analysis

Sensitivity analysis addresses to the answer of the question 'How sensitive is the overall decision with respect to small changes in the input values?' In other words, sensitivity analysis is the systematic investigation regarding the impact of potential changes in the input values on the final decision. Here we conduct a sensitivity analysis to explore the effect of variation of confidence level of medical experts about the symptoms of patients, on the final diagnostic result. Without loss of generality, we assume that linguistic values (strengths of the symptoms are expressed by linguistic values) are quantified by GTFNs and they are of the form $S_i = [(a_i, b_i, c_i); w_i]$, where w_i represents confidence level of doctors/medical experts with respect to corresponding symptoms. A slight variation in the original confidence level provided by medical experts is computed as follows:

 $S_i = [(a_i, b_i, c_i); w_i - \Delta_i]$, where Δ_i is variation value and its range is $[0, \min w_i]$ with $0 \le w_i - \Delta_i \le 1$.

Sensitivity analysis is performed to determine the influence of doctor's confidence level in evaluating the symptoms of patients, on the final diagnostic result. The result of the sensitivity analysis, i.e., the variation of strength of diseases into the patients Arka, Bumba, Chandra and Dip are depicted in Figs. 4, 5, 6 and 7, respectively.

From Fig. 4, we observe that, when the confidence level of doctors in evaluating five symptoms of the patient Arka, vary from 0 to 0.02, the strengths of the two diseases, such as, Viral fever and Malaria, are changed. But over the range [0,0.4], strengths of other three diseases, such as, Typhoid, Stomach problem and Chest problem remain constant. It demonstrates that, diseases Viral fever and Malaria are more sensitive than other three diseases for the patient Arka.

From Fig. 5, we observe the following. During evaluation of symptoms of the patient Bumba, for different values of confidence levels of doctors in the range [0,0.05], strengths of Stomach problem and Chest problem vary whereas in the range [0.05,0.2] strengths of Malaria and Typhoid diseases vary. But throughout the range [0,0.8], strength of Viral fever remain constant. Hence, for the patient Bumba, Malaria, Typhoid, Stomach problem and Chest problem are more sensitive than the Viral fever.

From Fig. 6, it is observed that, if the confidence levels of five symptoms of the patient Chandra vary from 0.08 to 0.12, then strengths of Malaria and Typhoid are changed.



Fig. 4 Diagnostic result of Arka sensitive to confidence level of symptoms provided by Doctors





Variation value of confidence level of symptoms of Dip

Fig. 6 Diagnostic result of Chandra sensitive to confidence level of symptoms provided by Doctors

Fig. 7 Diagnostic result of Dip sensitive to confidence level of symptoms provided by Doctors

In the range [0.1,0.16], strengths of diseases Viral fever, malaria and Typhoid vary, whereas throughout the range [0,0.7], strength of Chest problem remain fixed.

From Fig. 7, it is clear that, if the confidence levels regarding five symptoms of the patient Dip vary from 0 to 0.06, then strengths of Viral fever and Malaria are changed. Again, other three diseases Typhoid, Stomach problem and Chest problem remain unchanged over the range. It states that, Viral fever and Malaria are more sensitive than other three diseases for the patient Dip.

Finally, in order to explore the applicability of the proposed method, a web-based MDSS for diagnosing diseases is developed in the following section.

7 A web-based tool for supporting medical diagnosis

One of the key challenges in today's world is to provide primary healthcare for the people living in rural areas. As mentioned in the introduction, delivering health services in such areas is quite difficult due to lack of adequate medical infrastructure and scarcity of trained doctors. The advancement of web technologies and wide access of the internet, open up a new avenue for delivering primary medical services via web-based applications. The major advantages of a web-based system are that it can be accessed from anywhere via internet and can perform specific complex decision making task to assist users. In view of this, we develop a web-based tool to diagnose five common diseases based on the proposed medical diagnosis approach. Basically, the system is designed to support primary health workers working in rural health service centers, to diagnose the disease at initial level. Thus, the system can be regarded as a way of complementing the inadequate number of medical experts in developing countries. The architecture of the web-based MDSS is depicted in Fig. 8.

Fig. 8 Architecture of the web based system

Welcome to Diagnosis Section

tiont Ilinter	u and Destan sharrations	
illent mistr	y and Doctor obsevations	

Symptoms	Linguistic Label	Confidence
Body Temperature	H 🗸	0.8 🖵
Headeche	M 🔶	1.0 👻
Stomach Pain	М 👻	0.7 👻
Cough	VH 🗸	0.6 🖵
Chest Pain	L 🗸	1.0 👻
Diagnose	Patient Re	set

Fig. 9 Interactive web page for diagnosis

The system consists of three components: user interface, decision making module and data management module. User interface helps the primary health workers to interact with the system via interactive web pages. First, the health workers complete examining the patients and collect data related to the signs and symptoms of diseases. Then, they provide their assessments for each of the symptoms (Fig. 9). These assessments are basically GFNs and used as new inputs of the system for diagnosis the diseases of the new patients.

The decision making module is responsible for processing each set of inputs by using the proposed medical KB and IFIS. When a health worker feeds the system with values representing signs and symptoms of diseases based on his(er) observations regarding a patient and directs the system to diagnosis the disease by clicking the button 'Diag-





Fig. 10 Diagnosis result of the system

nosis patient' (Fig. 9), consequently this decision making module is started running on a web server on the background of the system and finally, produces the diagnosis result (Fig. 10). Figure 10 shows that the patient with ID '01' has viral fever with 81 % possibility. The data management module is responsible for storing all the data generated during this process.

The system is developed in Netbeans integrated development environment (IDE) by using HyperText Markup Language (HTML), Java (J2SE), Java server pages (JSP), JavaScript programming languages with My Structured Query Language (MYSQL) as the database management systems. To demonstrate the method that drives the proposed system, we evaluate it by using the collected medical data of fifteen patients as presented in Table 8. For each of the patients, the diagnostic result is shown in the last column of Table 8.

The following observations on the proposed web-based MDSS, we would like to share with our readers.

In most of the rural areas of developing countries, an insufficient number of doctors has enhanced the mortality of patients who are suffering from various diseases [5]. As an illustration, in rural areas of India sometimes specialist doctors may come to the health-centre only once in a week. The waiting time for treatments in health-centre of rural areas sometimes takes a few days or weeks. By that time, the condition of patients gets worse as the disease may have already spread out. As many of the diseases could have been cured at the early stage, the patients may have to suffer for the rest of their life.

With this observation in the background, a web-based MDSS is developed by emulating human intelligence, which could be used to assist the health workers in predicting the diseases, without consulting the specialists directly. Health workers will be trained on how to use such a system for diagnosing patients as a way of complementing the insufficient number of medical experts/specialists in rural areas [14]. The software will not replace the specialist or doctor [5], as it is developed to assist medical practitioner and health workers in diagnosing and predicting the patient's condition from certain rules without spending several hours. The proposed system can assist/helps the medical workers of health-centre in predicting the disease

Table 8	Web	based	diagnostic	result
---------	-----	-------	------------	--------

Patient ID	Temperature	Headache	Stomach pain	Cough	Chest pain	Diagnostic result
01			()(() 7)			N7:1 6
01	(H, 0.8)	(M, 1)	(M, 0.7)	(VH, 0.6)	(L, 1)	viral fever
02	(L, 0.3)	(VL, 0.8)	(M, 0.8)	(H, 0.5)	(H, 0.4)	Typhoid
03	(VL, 0.9)	(L, 0.5)	(M, 0.8)	(M, 0.7)	(H, 0.8)	Viral fever
04	(VL, 0.3)	(L, 0.4)	(M, 0.8)	(H, 0.6)	(VH, 0.9)	Typhoid
05	(VH, 0.7)	(H, 0.8)	(L, 0.9)	(H, 0.8)	(VL, 0.9)	Viral fever
06	(L, 0.2)	(VL, 0.8)	(VH, 0.8)	(M, 0.3)	(H, 0.7)	Stomach problem
07	(H, 0.7)	(VH, 0.8)	(M, 0.7)	(L, 0.8)	(M, 0.4)	Malaria
08	(L, 0.1)	(VL, 0.7)	(M, 0.7)	(H, 0.5)	(H, 0.6)	Typhoid
09	(H, 0.8)	(L, 0.3)	(M, 0.6)	(M, 0.7)	(H, 0.4)	Viral fever
10	(VH, 0.9)	(H, 1)	(M, 0.4)	(L, 0.8)	(VL, 1)	Malaria
11	(L, 0.5)	(VL, 1)	(H, 0.7)	(M, 0.6)	(VH, 0.8)	Viral fever
12	(VL, 0.8)	(L, 0.9)	(VH, 0.9)	(H, 0.3)	(M, 0.4)	Viral fever
13	(VH, 0.9)	(M, 0.4)	(VL, 1)	(H, 1)	(L, 0.7)	Malaria
14	(VL, 0.6)	(H, 0.9)	(L, 0.6)	(H, 0.8)	(VH, 0.9)	Viral fever
15	(H, 0.6)	(M, 0.5)	(L, 0.8)	(VL, 0.8)	(H, 0.7)	Malaria

which can provide early diagnosis before starting a proper medical treatment. The system can reduce the chances of developing serious illness as well as avoid unnecessary tests. If the patients' diseases are not diagnosed in early stage then they would not have the appropriate medical treatment, they have to buy several drugs, may go for several medical tests, etc., which may imply an over cost for the patients [10].

In this respect, the proposed web-based MDSS can save time, avoid unnecessary drugs and tests, reduce extra cost and help patients.

8 Conclusion

This study has presented a novel algorithm to solve medical diagnosis problem by integrating fuzzy and intuitionistic fuzzy frameworks. The research contribution has two phase: first phase executes a new diagnostic approach and validates the final outcome. Second phase develops a web-based MDSS for the diagnosis of diseases by adapting the proposed approach. This method has considered a set of five related diseases with a set of common symptoms. The advantages of this approach can be pointed out as follows: (i) the proposed methodology presents an attempt to develop a medical diagnostic approach considering medical experts' confidence levels, (ii) the proposed medical KB has a higher utility due to its ability to include possibility and nonpossibility of diseases during the diagnostic process, (iii) the method is simple in terms of computation, (iv) the membership construction method makes an effort so that each rule could be fired for providing better efficiency to the output, (v) the proposed algorithm is used for developing web-based MDSS.

In general, it may be concluded that combination of GFN and IFS in medical knowledge base makes the medical diagnostic model developed in the present work more realistic. This is because the use of GFN and IFS imply that both confidence level of patients in expressing their conditions and hesitation of medical experts in providing the degree of possibility of diseases-are being taken into consideration simultaneously and thus, enhance the capability of handling uncertainty in the medical diagnosis problems. It is important to remark here that, this is a reasonable starting point, since, we could not find any ready-made benchmark data which can fit in the proposed MDSS. This is because the rule base of our medical diagnostic model is not only framed by using fuzzy numbers but also IFSs. Thus, in the present study, we have fitted the approach by a short amount of collected data (with the assistantship of doctors) and validated the result of our case study by similarity measurement method.

In future, we would like to extend the proposed MDSS with a large set of samples and symptoms collected via

survey in sub-centre, rural family welfare-centre, block primary health-centre and other health-centres in our state and analyze the results in comparison with the conventional approaches.

Acknowledgments We are very grateful to the Editors and the anonymous Reviewers for their insightful and constructive comments and suggestions for the improvement of the manuscript. The authors would like to thank Dr. Rajesh Ranjan, Mahavir Vaatsalya Aspatal, Patna, India for his precious assistance in data analysis of our research. We especially appreciate Dr. Tapomay Dutta Majumder, for his utmost co-operation in rule base construction from collected data. The first author gratefully acknowledges the financial support provided by Ministry of Human Resource Development, Government of India and third author gratefully acknowledges the financial support provided by the Council of Scientific and Industrial Research, New Delhi, India, under Award 09/1023(007)/2011-EMR-I.

Appendix A

Here, we review definitions and concepts of fuzzy number, GFN and IFS that are required for developments of the present study. We start by recalling the definition of a fuzzy number.

Definition A.1 [64] A fuzzy set defined on real line \mathbb{R} is called a fuzzy number if it is convex, normal and having a piecewise continuous membership function with bounded support.

To represent expert's opinion, in a more flexible manner, GFN is better to use as it captures confidence level of expert. The definition of GFN is given as follows:

Definition A.2 [50] A GFN $\widetilde{A} = [(a, b, c, d); w]$, where a,b,c,d are real numbers and $0 < w \le 1$, is a fuzzy subset of the real numbers \mathbb{R} with a membership function $\mu_A(u)$ which satisfies the following conditions:

- i) µ_Ã(u) is a continuous function from ℝ to the closed interval [0, w].
- ii) $\mu_{\widetilde{A}}(u) = 0$, where $-\infty < u \le a$.
- iii) $\mu_{\widetilde{A}}(u)$ is monotonic increasing function in [a,b].
- iv) $\mu_{\widetilde{A}}(u) = w$, where $b \le u \le c$.
- v) $\mu_{\widetilde{A}}(u)$ is monotonic decreasing function in [c,d].
- vi) $\mu_{\widetilde{A}}(u) = 0$, where $d \le u < \infty$.

If $\mu_{\widetilde{A}}(u)$ is a linear function of u and also b = c then \widetilde{A} is called a GTFN. Without any loss of generality, throughout the work to represent the linguistic assessments of patients' symptoms, provided by the medical experts, in a flexible manner, we consider GTFNs. When a GTFN \widetilde{A} is used to represent the expert's opinion, the value w indicates the level of confidence of the medical expert.

Definition A.3 [15] An IFS *A* in the universe \mathcal{U} can be expressed as a set of ordered triple, $A = \{(u, \mu_A(u), \nu_A(u)) : u \in \mathcal{U}\}$, where $\mu_A(u) : \mathcal{U} \to [0, 1]$ is the degree of belongingness and $\nu_A(u) : \mathcal{U} \to [0, 1]$ is the degree of non-belongingness of *u* in *A*. They satisfy the relation $0 \le \mu_A(u) + \nu_A(u) \le 1 \forall u \in U$. The quantity $\pi_A(u) = 1 - \mu_A(u) - \nu_A(u)$ is called the degree of hesitation (indeterminacy) of *u* in *A*.

A brief description of the concepts of IFS operations and aggregation operators for IFSs based on continuous Archimedean t-norm and t-conorm, are briefly reviewed in Appendix B.

Appendix B

It is well known that, a strict Archimedean t-norm has a representation by means of its additive generator [56] $g: [0,1] \rightarrow [0,\infty]$ such that g(1) = 0 as $T(x, y) = g^{-1}(g(x) + g(y))$

Similarly, a strict Archimedean t-conorm is expressed through its additive generator [56] $h : [0, 1] \rightarrow [0, \infty]$ as $S(x, y) = h^{-1}(h(x) + h(y))$

With the above analysis, pointwise operations for IFSs based on a given strict Archimedean t-norm T (generated by an additive generator g) and strict Archimedean t-conorm S (generated by an additive generator h) are introduced as follows:

Definition B.1 [54, 56] Let $a_1 = (\mu_{a_1}, \nu_{a_1})$ and $a_2 = (\mu_{\tilde{a}_2}, \nu_{\tilde{a}_2})$ be two IFSs and $\alpha \ge 0$ be a scalar, and let a strict t-norm *T* and strict t-conorm *S* be generated by additive generators g and h, respectively. Then

(i)
$$a_1 \oplus a_2 = (S(\mu_{a_1}, \mu_{a_2}), T(\nu_{a_1}, \nu_{a_2}))$$

 $= (h^{-1}(h(\mu_{a_1}) + h(\mu_{a_2})), g^{-1}(g(\nu_{a_1}) + g(\nu_{a_2})))$
(ii) $a_1 \otimes a_2 = (T(\mu_{a_1}, \mu_{a_2}), S(\nu_{a_1}, \nu_{a_2}))$

$$= \left(g^{-1}\left(g(\mu_{a_1}) + g(\mu_{a_2})\right), h^{-1}\left(h(\nu_{a_1}) + h(\nu_{a_2})\right)\right)$$

(iii)
$$\alpha a_1 = \left(h^{-1}\left(\alpha h(\mu_{a_1})\right), g^{-1}\left(\alpha g(\nu_{a_1})\right)\right)$$

(iv)
$$a_1^{\alpha} = (g^{-1}(\alpha g(\mu_{a_1})), h^{-1}(\alpha h(\nu_{a_1})))$$

To combine several IFSs into a single IFS, an aggregation function [53] is required which can be defined as: a function $f : [0, 1]^n \rightarrow [0, 1]$ is called aggregation function of *n* arguments if it satisfies the properties (i) f(0, 0, ..., 0) = 0 and f(1, 1, ..., 1) = 1; (ii) if $x \le y$ then $f(x) \le f(y) \forall x, y \in [0, 1]^n$. In this work, we use intuitionistic fuzzy weighted arithmetic mean (IWAM) operator to aggregate IFSs. Based on arithmetic operations of IFSs defined in Definition B.1, IWAM operator [56] is defined as **Definition B.2** [56] Let $a_i = (\mu_{a_i}, \nu_{a_i})(i = 1, 2, ..., n)$ be a collection of IFSs. An IWAM operator is a mapping IWAM : $(\mathbb{R}^+)^n \to \mathbb{R}^+$ which is defined by

$$IWAM(a_1, a_2, ..., a_n) = \omega_1 a_1 + \omega_2 a_2 + ... + \omega_n a_n$$

= $\left(h^{-1}\left(\sum_{i=1}^n \omega_i h(\mu_{a_i})\right), g^{-1}\left(\sum_{i=1}^n \omega_i g(\nu_{a_i})\right)\right)$ (11)

where $\omega = (\omega_1, \omega_2, ..., \omega_n)^T$ be the weight vector of $a_i (i = 1, ..., n), \omega_i > 0$, and $\sum_{i=1}^n \omega_i = 1$.

Definition B.3 [56] If in particular, $h^{-1} = h = 1 - Id$ and $g^{-1} = g = Id$ on [0, 1], and that $\sum_{i=1}^{n} \omega_i = 1$ ensures the argument of $h^{-1} \in [0, 1]$, then IWAM and IAM operator reduce to the following forms:

$$IWAM(a_1, a_2, ..., a_n) = \left(1 - \left(\sum_{i=1}^n \omega_i (1 - \mu_{a_i})\right), \sum_{i=1}^n \omega_i v_{a_i}\right)$$
(12)

If the weight vector $\omega = (\frac{1}{n}, \frac{1}{n}, ..., \frac{1}{n})^T$, then IWAM operator reduces to IAM operator and consequently (12) becomes

$$IAM(a_1, a_2, ..., a_n) = \left(1 - \left(\frac{1}{n}\sum_{i=1}^n (1 - \mu_{a_i})\right), \frac{1}{n}\sum_{i=1}^n \nu_{a_i}\right)$$
(13)

References

- West DM (2013) Improving health care through mobile medical devices and sensors. Brookings Institution Policy Report
- 2. Bhandari L, Dutta S (2007) Health infrastructure in rural India. India Infrastructure Report
- Szolovits P, Patil RS, Schwartz WB (1988) Artificial intelligence in medical diagnosis. Ann Intern Med 108(1):80–87
- Jarvis-Selinger S, Bates J, Araki Y, Lear SA (2011) Internet-based support for cardiovascular disease management. Int J Telemed Appl 2011:1–9
- Ishak WHW, Siraj F (2002) Artificial intelligence in medical application: an exploration. Health Inform Eur J 16:1–9
- 6. Kumar G (2009) Health sector reforms in India. Manohar Publication
- Gadaras I, Mikhailov L (2009) An interpretable fuzzy rule-based classification methodology for medical diagnosis. Artif Intell Med 47(1):25–41
- Dwivedi A, Mishra D, Kalra PK (2006) Handling uncertaintiesusing probability theory to possibility theory, http://www.iitk.ac. in/directions/feb2006/PRINTMISHRAK.pdf
- 9. Zadeh LA (1965) Fuzzy sets. Inf Control 8(3):338-353
- Sanz J, Galar M, Jurio A, Brugos A, Pagola M, Bustince H (2014) Medical diagnosis of cardiovascular diseases using an intervalvalued fuzzy rule-based classification system. Appl Soft Comput 20:103–111

- 11. Zadeh HG, Pakdelazar O, Haddadnia J, Rezai-Rad G, Mohammad-Zadeh M (2012) Diagnosing breast cancer with the aid of fuzzy logic based on data mining of a genetic algorithm in infrared images. Middle East J Cancer 3(4):119–129
- Tsakonas A, Dounias G, Von Keyserlingk DG, Axer H (2001) Hybrid computational intelligence for handling diagnosis of aphasia. In: European symposium on intelligent technologies, hybrid systems and their implementation on smart adaptive systems. Eunite. Citeseer, pp 13–14
- Sanchez E (1996) Truth-qualification and fuzzy relations in natural languages, application to medical diagnosis. Fuzzy Sets Syst 84(2):155–167
- Samuel O, Omisore M, Ojokoh B (2013) A web based decision support system driven by fuzzy logic for the diagnosis of typhoid fever. Expert Syst Appl 40(10):4164–4171
- Atanassov KT (1986) Intuitionistic fuzzy sets. Fuzzy Sets Syst 20(1):87–96
- De SK, Biswas R, Roy AR (2001) An application of intuitionistic fuzzy sets in medical diagnosis. Fuzzy Sets Syst 117(2):209–213
- Das S, Dutta B, Guha D (2015) Weight computation of criteria in a decision-making problem by knowledge measure with intuitionistic fuzzy set and interval-valued intuitionistic fuzzy set. Soft Comput 1–22. doi:10.1007/s00500-015-1813-3
- Papakostas GA, Hatzimichailidis AG, Kaburlasos VG (2013) Distance and similarity measures between intuitionistic fuzzy sets: a comparative analysis from a pattern recognition point of view. Pattern Recogn Lett 34(14):1609–1622
- Kharal A (2009) Homeopathic drug selection using intuitionistic fuzzy sets. Homeopathy 98(1):35–39
- Hung KC, Tuan HW (2013) Medical diagnosis based on intuitionistic fuzzy sets revisited. J Interdiscip Math 16(6):385–395
- Szmidt E, Kacprzyk J (2001) Intuitionistic fuzzy sets in some medical applications. In: Computational intelligence theory and applications. Springer, pp 148–151
- Chen SM (1994) A weighted fuzzy reasoning algorithm for medical diagnosis. Decis Support Syst 11(1):37–43
- Yao JF, Yao JS (2001) Fuzzy decision making for medical diagnosis based on fuzzy number and compositional rule of inference. Fuzzy Sets Syst 120(2):351–366
- Obot OU, Uzoka F-ME (2008) Fuzzy rule-based framework for the management of tropical diseases. Int J Med Eng Inf 1:7–17
- Chiang IJ, Shieh MJ, Hsu JY-J, Wong J-M (2005) Building a medical decision support system for colon polyp screening by using fuzzy classification trees. Appl Intell 22(1):61–75
- Obot OU, Uzoka F-ME (2009) A framework for application of neuro-case-rule base hybridization in medical diagnosis. Appl Soft Comput 9(1):245–253
- Maio CD, Loia V, Fenza G, Gallo M, Linciano R, Morrone A (2011) Fuzzy knowledge approach to automatic disease diagnosis. In: Proceedings of IEEE international conference on fuzzy systems (FUZZ), pp 2088–2095
- Fenza G, Furno D, Loia V (2012) Hybrid approach for contextaware service discovery in healthcare domain. J Comput Syst Sci 78(4):1232–1247
- Aribarg T, Supratid S, Lursinsap C (2012) Optimizing the modified fuzzy ant-miner for efficient medical diagnosis. Appl Intell 37(3):357–376
- Sundarapandian V, Ephzibah E (2012) Framing fuzzy rules using support sets for effective heart disease diagnosis. Int J Fuzzy Logic Syst 2:11–16
- Singh S, Kumar A, Panneerselvam K., Vennila JJ (2012) Diagnosis of arthritis through fuzzy inference system. J Med Syst 36(3):1459–1468
- 32. Caponetti L, Castellano G, Basile MT, Corsini V (2014) Fuzzy mathematical morphology for biological image segmentation. Appl Intell 41(1):117–127

- Phuong NH, Kreinovich V (2001) Fuzzy logic and its applications in medicine. Int J Med Inform 62(2):165–173
- 34. Zhou SM, John RI, Wang X-Y, Garibaldi JM, Ellis IO (2008) Compact fuzzy rules induction and feature extraction using svm with particle swarms for breast cancer treatments. In: IEEE congress on evolutionary computation (CEC). IEEE, pp 1469– 1475
- Sajja PS, Shah DM (2010) Knowledge based diagnosis of abdomen pain using fuzzy prolog rules. J Emerg Trends Comput Inf Sci 1(2):55–60
- Baig F, Khan MS, Noor Y, Imran M (2011) Design model of fuzzy logic medical diagnosis control system. Int J Comput Sci Eng 3(5):2093–2108
- Djam X, Wajiga G, Kimbi Y, Blamah N (2011) A fuzzy expert system for the management of malaria. Int J Pure Appl Sci Technol 5(2):84–108
- Dahiya V, Trivedi J, Sajja PS (2012) Designing of fuzzy rule base for vision defects. Int J Comput Technol Appl 3(2):734–737
- Szmidt E, Kacprzyk J (2004) A similarity measure for intuitionistic fuzzy sets and its application in supporting medical diagnostic reasoning. In: Artificial intelligence and soft computing-ICAISC. Springer, pp 388–393
- 40. Ye J (2011) Cosine similarity measures for intuitionistic fuzzy sets and their applications. Math Comput Model 53(1):91–97
- Hung KC, Wang PK (2014) An integrated intuitionistic fuzzy similarity measures for medical problems. Int J Comput Intell Syst 7(2):327–343
- Yanni W, Yaping D, Fancheng M (2013) Similarity measure of intuitionistic trapezoidal fuzzy numbers and its application for medical diagnosis. In: 32nd Chinese Control Conference (CCC). IEEE, pp 8567–8571
- Hung KC (2012) Medical pattern recognition: applying an improved intuitionistic fuzzy cross-entropy approach. Adv Fuzzy Syst 2012:1–6
- 44. Zhang Z, Yang J, Ye Y, Hu Y, Zhang Q (2012) A type of score function on intuitionistic fuzzy sets with double parameters and its application to pattern recognition and medical diagnosis. Procedia Eng 29:4336–4342
- Das S, Kar S (2014) Group decision making in medical system: an intuitionistic fuzzy soft set approach. Appl Soft Comput 24:196– 211
- 46. Zhang H, Shu L, Liao S (2014) Generalized trapezoidal fuzzy soft set and its application in medical diagnosis. J Appl Math 2014:1– 12
- Own CM (2009) Switching between type-2 fuzzy sets and intuitionistic fuzzy sets: an application in medical diagnosis. Appl Intell 31(3):283–291
- 48. Lee H, Chen YP (2014) Skin cancer extraction with optimum fuzzy thresholding technique. Appl Intell 40(3):415–426
- Guha D, Chakraborty D (2010) A new approach to fuzzy distance measure and similarity measure between two generalized fuzzy numbers. Appl Soft Comput 10(1):90–99
- Chen SJ, Chen SM (2003) Fuzzy risk analysis based on similarity measures of generalized fuzzy numbers. IEEE Trans Fuzzy Syst 11(1):45–56
- Chakraborty D, Guha D, Dutta B Multi-objective optimization problem under fuzzy rule constraints using particle swarm optimization. Soft Comput. doi:10.1007/s00500-015-1639-z
- 52. Zadeh LA (1975) The concept of a linguistic variable and its application to approximate reasoning-II. Inf Sci 8(4):301-357
- 53. Beliakov G, Pradera A, Calvo T (2007) Aggregation functions: a guide for practitioners. Springer, Berlin
- Xia M, Xu Z, Zhu B (2012) Some issues on intuitionistic fuzzy aggregation operators based on Archimedean t-conorm and tnorm. Knowl-Based Syst 31:78–88

867

- 55. Klement EP, Mesiar R, Pap E (2000) Triangular norms, vol 8. Springer Science and Business Media
- Beliakov G, Bustince H, Goswami DP, Mukherjee UK, Pal NR (2011) On averaging operators for Atanassov's intuitionistic fuzzy sets. Inf Sci 181(6):1116–1124
- 57. Lee KH (2006) First course on fuzzy theory and applications, vol 27. Springer Science and Business Media
- Larsen PM (1980) Industrial applications of fuzzy logic control. Int J Man Mach Stud 12(1):3–10
- Boran FE, Genc S, Kurt M, Akay D (2009) A multi-criteria intuitionistic fuzzy group decision making for supplier selection with topsis method. Expert Syst Appl 36(8):11363–11368
- 60. Castellano G, Fanelli AM, Mencar C (2003) Discovering human understandable fuzzy diagnostic rules from medical data. In: Proceedings of European symposium on intelligence technologies, hybrid systems and their implementation on smart adaptive systems, pp 227–233
- Yong D, Wenkang S, Feng D, Qi L (2004) A new similarity measure of generalized fuzzy numbers and its application to pattern recognition. Pattern Recogn Lett 25(8):875–883
- 62. Hwang CM, Yang MS (2014) New similarity measures between generalized trapezoidal fuzzy numbers using the Jaccard Index. Int J Uncertain Fuzziness Knowl Based Syst 22(6):831–844
- Vicente E, Mateos A, Jiménez A (2013) A new similarity function for generalized trapezoidal fuzzy numbers. In: Proceedings of artificial intelligence and soft computing. Springer, pp 400–411
- Cheng CH (1998) A new approach for ranking fuzzy numbers by distance method. Fuzzy Sets Syst 95(3):307–317

Satyajit Das received his M.Sc. degree in Mathematics with Computer Applications from National Institute of Technology, Durgapur, India in 2012. Presently, he is Research Fellow in the Department of Mathematics in Indian Institute of Technology, Patna, India. His current research area includes: Fuzzy optimization, Multi-criteria decision making and Intuitionistic fuzzy set theory.

Debashree Guha received her B.Sc. and M.Sc. degree in Mathematics from Jadavpur University, Calcutta, India in 2003 and 2005, respectively. She received her Ph.D degree in Mathematics from Indian Institute of Technology, Kharagpur, India in 2011.

Presently, she is the Assistant Professor of Department of Mathematics in Indian Institute of Technology, Patna, India. Her current research interest includes: Multi-attribute decision making, Fuzzy mathematical programming, Aggregation operators and Fuzzy logic. Her research results have been published in the Computer and Industrial Engineering, Applied Soft Computing, IEEE Transactions on Fuzzy Systems, Soft Computing, among others.

Bapi Dutta received his M.Sc. degree in industrial mathematics and informatics from Indian Institute of Technology, Roorkee, India in 2010. He is currently Ph.D candidate in the department of mathematics in Indian Institute of Technology, Patna, India. His current research interests includes: aggregation operators, multi-attribute decision making.