



Distance measure for Pythagorean fuzzy sets with varied applications

Juthika Mahanta¹ · Subhasis Panda²

Received: 23 December 2020 / Accepted: 5 July 2021 / Published online: 5 August 2021
© The Author(s), under exclusive licence to Springer-Verlag London Ltd., part of Springer Nature 2021

Abstract

Distance measure is one of the research hotspot in Pythagorean fuzzy environment due to its quantitative ability of distinguishing Pythagorean fuzzy sets (PFSs). Various distance functions for PFSs are introduced in the literature and have their own pros and cons. The common thread of incompetency for these existing distance functions is their inability to distinguish highly uncertain PFSs distinctly. To tackle this point, we introduce a novel distance measure for PFSs. An added advantage of the measure is its simple mathematical form. Moreover, superiority and reasonability of the prescribed definition are demonstrated through proper numerical examples. Boundedness and nonlinear behaviour of the distance measure is established and verified via suitable illustrations. In the current scenario, selecting an antivirus face-mask as a preventive measure in the COVID-19 pandemic and choosing the best school in private sector for children are some of the burning issues of a modern society. These issues are addressed here as multi-attribute decision-making problems and feasible solutions are obtained using the introduced definition. Applicability of the distance measure is further extended in the areas of pattern recognition and medical diagnosis.

Keywords Pythagorean fuzzy sets (PFSs) · Distance measure and similarity · Medical diagnosis · Pattern recognition · Multi-attribute decision making

1 Introduction

The fuzzy sets [1] allow an element under consideration to have a partial membership degree of belongingness to a set instead of the classical cases of “belongs to” or “not belongs to”. This revolutionary idea of fuzzy sets which can accommodate uncertainty naturally has been studied and applied extensively by researchers across different communities. Later, it was realized that, like the partial grade of membership for belongingness, the non-membership grade is equally important. Thereby, intuitionistic fuzzy sets (IFSs) were introduced and studied by Atanossov [2–5] which is a generalization of fuzzy set. IFSs consider degrees of membership (μ) as well as non-membership (ν), with their sum restricted to a value less than

equals to one, for an element in the universe of discourse. Obviously, for fuzzy sets the non-membership grade is always fuzzy complement of the membership grade. Considering μ as abscissa and ν as ordinate, the area under the line $\mu + \nu = 1$ in the first quadrant geometrically represents the admissible region for IFSs. There are instances where the sum of membership and non-membership grades exceeds one, viz. $\mu = \nu = \frac{1}{\sqrt{2}}$, and these can't be represented by IFSs. It necessitates a further generalization of IFSs to encompass such cases. Subsequently, Pythagorean fuzzy sets (PFSs) were introduced by Yager [6, 7] which allow a larger admissible area with sum of the squares of μ and ν is less than equals to one. It is evident that, every fuzzy set is an IFS, each IFS is a PFS, but not conversely. PFSs allow more accessible area than IFSs, so it can efficiently and accurately handle uncertainty. Because of that, PFSs have drawn the attention of researchers and are being applied in myriad fields like pattern recognition [8–12], decision making [13–19], medical diagnosis [20–24] and others [25–31], resulting in improved outcomes, since inception.

✉ Juthika Mahanta
juthika@math.nits.ac.in

¹ Department of Mathematics, NIT Silchar, Silchar, Assam 788010, India

² Department of Physics, NIT Silchar, Silchar, Assam 788010, India

To estimate the similarity or difference between PFSs, various distance and similarity measures have been introduced and studied in the literature. These measures are successfully employed to different regime of applications; few among them are delineated here. On the basis of the Minkowski distance measure, Chen [32] introduced a distance for PFSs and applied it to problems of Internet stock and R&D project investment along with some other real-world problems. In [33], the authors proposed several distance measures for PFSs and Pythagorean fuzzy numbers and demonstrated their usefulness via examples. Wei & Wei [23] presented ten different types of similarity measures for PFSs based on cosine function and illustrated their serviceability in medical diagnosis & pattern recognition problems. Ejegwa [34] extended the distance measures for IFSs, viz, Hamming, Euclidean, normalized Hamming, and normalized Euclidean distances, and similarities to PFSs and applied them to multi-criteria decision-making problems and multi-attribute decision-making problems. Taking advantage of the Jensen–Shannon divergence, a new divergence measure for PFSs was introduced in [21]. This measure was further established as a superior tool than the existing ones and was used to solve medical diagnosis problems. Generalized Pythagorean fuzzy normalized Hamming distance, the generalized Pythagorean fuzzy normalized Hausdorff distance and the generalized hybrid Pythagorean fuzzy normalized distance were introduced in [27]. Peng [35] came up with new Pythagorean distance and similarity measures [36] to overcome the deficiencies of existing such measures. In [9], the author modified the distance function introduced by Zhang & Xu in [37] by normalizing it and validated the axiomatic definition of a metric for modified version which was missing in [37].

Summarizing the existing literature, we observe that most of the distance functions for PFSs are introduced as a generalization of IFS counterparts. Few among them are not even normalized and thereby were reintroduced with suitable normalization. On the other hand, the complex mathematical form of some distance measures hinders their applicability and popularity as a handy mathematical tool. In addition to that, none among the existing measures could distinctly distinguish highly uncertain PFSs, i.e. PFSs with very small values of membership and non-membership grades. One encounters such PFSs when minimal knowledge or information is available about a system. So, there is a need for a suitable distance measure to deal with such PFSs. To fill the above-stated lacunae, we introduce a new distance function with simple mathematical form for PFSs which can effectively handle highly uncertain PFSs and is equally efficient in other cases also. Therefore, the main contributions of this work are

- introduction of a new distance measure for PFSs.
- construction of simple mathematical form of it to avoid calculation complicity.
- ability of handling effectively PFSs with higher degree of uncertainty and equal efficiency in other cases also.
- applicability of the introduced definition in the fields of multi-attribute decision making (MADM), pattern recognition & medical diagnosis.

The paper is organized as follows. Section 2 recalls the mathematical background of related concepts. The new definition of distance measure for PFSs is proposed and some properties are studied in Sect. 3. Different characteristics of the proposed measure are verified via various numerical examples in Sect. 4. Applicability of the prescribed distance function in various fields is elaborated in Sect. 5 followed by the conclusions in the last section.

2 Mathematical background

In this section, we quickly recall some of the basic definitions and results required for our further studies.

Definition 2.1 [2] Let X be a finite universe of discourse. An IFS K in X is defined as $K = \{\langle \mu_K(x), \nu_K(x) \rangle | x \in X\}$, $\mu_K : X \rightarrow [0, 1]$ and $\nu_K : X \rightarrow [0, 1]$ are, respectively, the membership and non-membership functions, such that $0 \leq \mu_K(x) + \nu_K(x) \leq 1, \forall x \in X$. The degree of hesitancy is given by $\pi_K(x) = 1 - \mu_K(x) - \nu_K(x)$.

Definition 2.2 [6, 13] Let X be a finite universe of discourse. A Pythagorean fuzzy set (PFS) K in X is defined as $K = \{\langle K_Y(x), K_N(x) \rangle | x \in X\}$, $K_Y : X \rightarrow [0, 1]$ and $K_N : X \rightarrow [0, 1]$, respectively, indicate the guaranteed membership and non-membership functions, such that $0 \leq K_Y^2(x) + K_N^2(x) \leq 1, \forall x \in X$. The degree of hesitancy is given by $K_H(x) = \text{Hes}(x) = \sqrt{1 - K_Y^2(x) - K_N^2(x)}$.

It is clear from the definitions that the admissible area under Pythagorean membership grades is greater than that of intuitionistic membership grades.

Unless mentioned otherwise, from now onward any PFS, say $K = \{\langle K_Y(x), K_N(x) \rangle | x \in X\}$ is denoted simply as K for brevity. The guaranteed membership K_Y , non-membership K_N and the degree of hesitancy K_H are always associated with the PFS K in this notation.

2.1 Operations

Let $\mathcal{PFS}(X)$ denote the class of all Pythagorean fuzzy sets defined on the finite universe of discourse X . The following operations [6, 13] are valid for all members $K, L \in \mathcal{PFS}(X)$,

$$K = L \Leftrightarrow K_Y(x) = L_Y(x) \& K_N(x) = L_N(x) \forall x \in X, \tag{1}$$

$$K \subseteq L \Leftrightarrow K_Y(x) \leq L_Y(x) \& K_N(x) \geq L_N(x) \forall x \in X, \tag{2}$$

wesay $K \subset L \Leftrightarrow K \subseteq L, K \neq L,$

the complement of K is defined as

$$K^c = \{ \langle K_N(x), K_Y(x) \rangle \mid x \in X \}, \tag{3}$$

$$K \cup L = \{ \langle \max[K_Y(x), L_Y(x)], \min[K_N(x), L_N(x)] \rangle \mid x \in X \}, \tag{4}$$

$$D_{CN}(K, L) = \left[\frac{1}{2n} \sum_{i=1}^n \left(|K_Y^2(x) - L_Y^2(x)|^\beta + |K_N^2(x) - L_N^2(x)|^\beta + |K_H^2(x) - L_H^2(x)|^\beta \right) \right]^{\frac{1}{\beta}}, \tag{8}$$

where $\beta \geq 1$ is a distance parameter. For $\beta = 1$ we get the Hamming distance and $\beta = 2$ gives the Euclidean distance.

4. The normalized PFSJS distance

$$D_{FW}(K, L) = \frac{1}{n} \sqrt{\frac{1}{2} \left[\sum_p K_p^2(x_i) \log \left(\frac{2K_p^2(x_i)}{K_p^2(x_i) + L_p^2(x_i)} \right) + \sum_p L_p^2(x_i) \log \left(\frac{2L_p^2(x_i)}{K_p^2(x_i) + L_p^2(x_i)} \right) \right]}, \tag{9}$$

$$K \cap L = \{ \langle \min[K_Y(x), L_Y(x)], \max[K_N(x), L_N(x)] \rangle \mid x \in X \}. \tag{5}$$

where $K_H^2(x) = 1 - K_Y^2(x) - K_N^2(x)$ and $p \in \{Y, N, H\}$.

5. Normalized Hausdorff distance [38]

$$D_{HN}(K, L) = \frac{1}{n} \sum_{i=1}^n \max[|K_Y^2(x_i) - L_Y^2(x_i)|, |K_N^2(x_i) - L_N^2(x_i)|]; \tag{10}$$

2.2 Distance measures

Below we list some of the popularly used definitions of distance measure for Pythagorean fuzzy sets. The following normalized versions of distance functions between two PFSs K, L were introduced in [21].

1. Normalized Hamming distance

$$D_{HmN}(K, L) = \frac{1}{2n} \sum_{i=1}^n (|K_Y^2(x_i) - L_Y^2(x_i)| + |K_N^2(x_i) - L_N^2(x_i)| + |K_H^2(x_i) - L_H^2(x_i)|); \tag{6}$$

6. Zhang and Xu distance [37]

$$D_{ZX}(K, L) = \frac{1}{2} \sum_{i=1}^n (|K_Y^2(x_i) - L_Y^2(x_i)| + |K_N^2(x_i) - L_N^2(x_i)| + |K_H^2(x_i) - L_H^2(x_i)|); \tag{11}$$

7. Modified Zhang and Xu distance [9]

$$D_{MZX}(K, L) = \frac{1}{n} D_{ZX}(K, L), \tag{12}$$

where $D_{ZX}(K, L)$ denotes Zhang and Xu distance as mentioned in Eq. (11).

2. The Euclidean distance

$$D_{EN}(K, L) = \sqrt{\frac{1}{2n} \sum_{i=1}^n \left((K_Y^2(x_i) - L_Y^2(x_i))^2 + (K_N^2(x_i) - L_N^2(x_i))^2 + (K_H^2(x_i) - L_H^2(x_i))^2 \right)}; \tag{7}$$

3. The normalized Chen’s distance

This list is by no means exhaustive and few of the popular distance measures are mentioned above. We have compared our results with those obtained through some of the aforesaid distances.

3 Proposed definition with mathematical properties

Here we introduce a novel distance measure for PFSs and study some of its mathematical properties.

Definition 3.1 Given a finite universe $X = \{x_i \mid i = 1, 2, \dots, n\}$, distance between any two PFSs K & L is defined as

$$D(K, L) = \frac{1}{n} \sum_{i=1}^n \frac{|K_Y^2(x_i) - L_Y^2(x_i)| + |K_N^2(x_i) - L_N^2(x_i)|}{K_Y^2(x_i) + L_Y^2(x_i) + K_N^2(x_i) + L_N^2(x_i)}, \tag{13}$$

where $K = \{\langle K_Y(x_i), K_N(x_i) \rangle \mid x_i \in X\}$ and $L = \{\langle L_Y(x_i), L_N(x_i) \rangle \mid x_i \in X\}$.

Then we have the following theorem.

Theorem 3.2 Let K_1, K_2 and K_3 be three PFS defined in the universe of discourse X . The introduced distance measure D satisfies the following properties:

- (i) Boundedness: $0 \leq D(K_1, K_2) \leq 1, \forall K_i$;
- (ii) Reflexivity: $D(K_1, K_1) = 0 \forall K_i$;
- (iii) Symmetry: $D(K_1, K_2) = D(K_2, K_1)$;
- (iv) Separability: $D(K_1, K_2) = 0$ iff $K_1 = K_2$;

A distance function satisfying the conditions (ii), (iii) & (iv) of Theorem 3.2 is known as semi-metric in the literature [39].

Proof

(i) It is evident that,

$$\max \left[\frac{|K_Y^2(x_i) - L_Y^2(x_i)| + |K_N^2(x_i) - L_N^2(x_i)|}{K_Y^2(x_i) + L_Y^2(x_i) + K_N^2(x_i) + L_N^2(x_i)} \right] = 1$$

for each x_i

$$\Rightarrow \frac{1}{n} \sum_{i=1}^n \frac{|K_Y^2(x_i) - L_Y^2(x_i)| + |K_N^2(x_i) - L_N^2(x_i)|}{K_Y^2(x_i) + L_Y^2(x_i) + K_N^2(x_i) + L_N^2(x_i)} \leq 1.$$

We have $K_Y(x_i), L_Y(x_i) \in [0, 1] \forall x_i \in X$. So $D(K, L)$ is a nonnegative quantity and hence we conclude $0 \leq D(K, L) \leq 1$.

- (ii) Obvious implication from equality of two PFSs.
- (iii) Follows from the definition of distance function.
- (iv) Using reflexivity property, we are left to show $D(K_1, K_2) = 0 \Rightarrow K_1 = K_2$.

Now $D(K_1, K_2) = 0$

$$\Rightarrow \frac{1}{n} \sum_{i=1}^n \frac{|K_{1Y}^2(x_i) - K_{2Y}^2(x_i)| + |K_{1N}^2(x_i) - K_{2N}^2(x_i)|}{K_{1Y}^2(x_i) + K_{2Y}^2(x_i) + K_{1N}^2(x_i) + K_{2N}^2(x_i)} = 0$$

$$\Rightarrow |K_{1Y}^2(x_i) - K_{2Y}^2(x_i)| + |K_{1N}^2(x_i) - K_{2N}^2(x_i)| = 0$$

$$\Rightarrow K_{1Y}^2(x_i) - K_{2Y}^2(x_i) = 0 \ \& \ K_{1N}^2(x_i) - K_{2N}^2(x_i) = 0$$

$$\Rightarrow K_{1Y}(x_i) = K_{2Y}(x_i) \ \& \ K_{1N}(x_i) = K_{2N}(x_i) \quad \forall x_i$$

$$\Rightarrow K_1 = K_2.$$

4 Numerical examples and characteristics

4.1 Boundedness and nonlinearity

- ◇ Consider two PFSs K and L defined on $X = \{x\}$, where $K = \langle a, b \rangle, L = \langle b, a \rangle$. As K and L are PFSs, we must have $a^2 + b^2 \leq 1$. Distance between such K and L using Eq. (13), is portrayed in Fig. 1 (left) and it is verified that, the distance is 0 iff $a = b$, i.e. when $A = B$. We adhere to the fact that, the PFS L is the complement of K and a mirror reflection about the 45° line will take PFS L to PFS K . This geometrical interpretation infers that, the PFSs lying on $a = b$ line are self-complement and thereby result in null distances as evident from Fig. 1 (left). It further verifies that the proposed distance function is bounded between 0 and 1. Additionally, for a selected value of b , say 0.2, Fig. 1 (right) portrays the nonlinear behaviour of the distance function. In a similar manner, for a given value of a , the proposed distance yields similar characteristics against b .
- ◇ Let us take a PFS, $K = \langle a, b \rangle$ defined on $X = \{x\}$. Being the membership & non-membership values of a PFS, $a, b \in [0, 1]$ and $a^2 + b^2 \leq 1$. Considering a^2, b^2 as the membership & non-membership grades, respectively, we construct another PFS, say $L = \langle a^2, b^2 \rangle$. Now, we calculate the distance between PFSs K & L using Eq. (13) and plot its variation in Fig. 2, for all allowed values of a and b . Boundedness and nonlinearity of the proposed distance are apparent here.

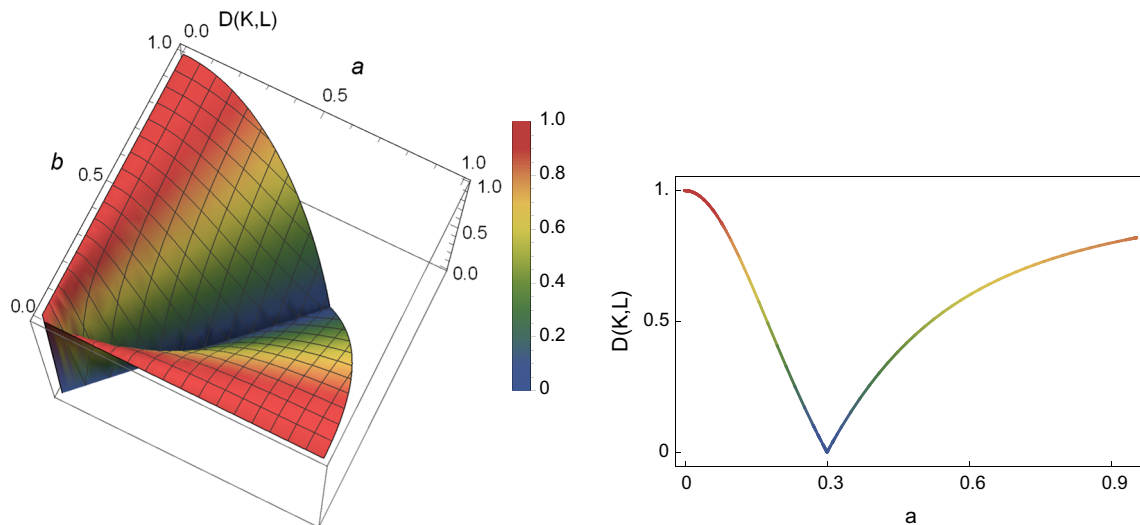


Fig. 1 Variation of distance function $D(K, L)$ with respect to (left) a and b . (right) a for a fixed $b = 0.2$

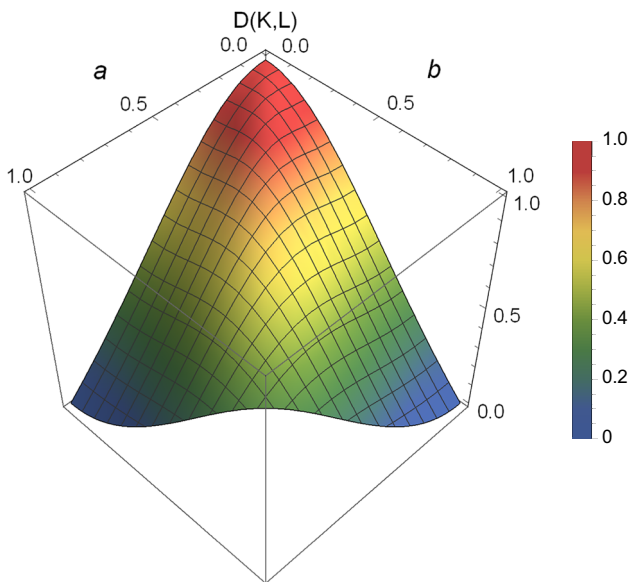


Fig. 2 Variation of $D(K, L)$ against all feasible values of a & b

◇ Next, consider the PFSs, $K = \langle a, b \rangle$ and $L = \langle \beta, 1 - \beta \rangle$ defined on $X = \{x\}$. For such a choice of L , the distance between them now varies against a , b and β . Without loss of generality, following four choices of K are considered

- (i) $K = \langle 1, 0 \rangle$, (ii) $K = \langle 0, 1 \rangle$,
- (iii) $K = \langle 0.4, 0.8 \rangle$ and (iv) $K = \left\langle \frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}} \right\rangle$.

and variations of $D(K, L)$ with respect to β are portrayed in Fig. 3. It is evident from the figure that the distance function D is bounded for all allowed values of β and its variation shows nonlinearity in all the cases.

4.2 Superiority

To demonstrate the superiority of the introduced distance function, a comparison among results obtained by our method and few popular distance functions is executed. For this purpose, let us take three PFSs, K, L, M defined on $X = \{x\}$, where $K = \langle .01, .02 \rangle, L = \langle .02, .02 \rangle, M = \langle .1, .2 \rangle$. As mentioned earlier, the proposed function is sensitive to highly uncertain PFSs, thereby, without loss of generality, aforesaid three PFSs are chosen as exemplary. Then, we calculate the distances $D(M, K)$ and $D(M, L)$ using Eq. (13) and few other distance functions (using Eqs. (6)–(10)). The results are displayed in Table 1. It is evident that the distance measures D_{HMN}, D_{EN} & D_{HN} couldn't differentiate K & L from M . However, D_{CN} and D_{FW} could distinguish, but they fall short when the results are rounded off to second decimal places only. On the other hand, the proposed function could clearly identify them as distinct sets even after rounding off to second decimal place. This affirms the superiority of our proposed definition.

Next example demonstrates that the applicability of the introduced definition is not limited to highly uncertain PFSs only.

4.3 Counter-intuitive case

We consider pairs of PFSs $A_1 = \{\langle .10, .10 \rangle, \langle .10, .10 \rangle\} = A_2; A_3 = \{\langle .714, .640 \rangle, \langle .640, .557 \rangle\} = A_4$.

- $B_1 = \{\langle .02, .02 \rangle, \langle .02, .01 \rangle\};$
- $B_2 = \{\langle .01, .01 \rangle, \langle .02, .02 \rangle\};$

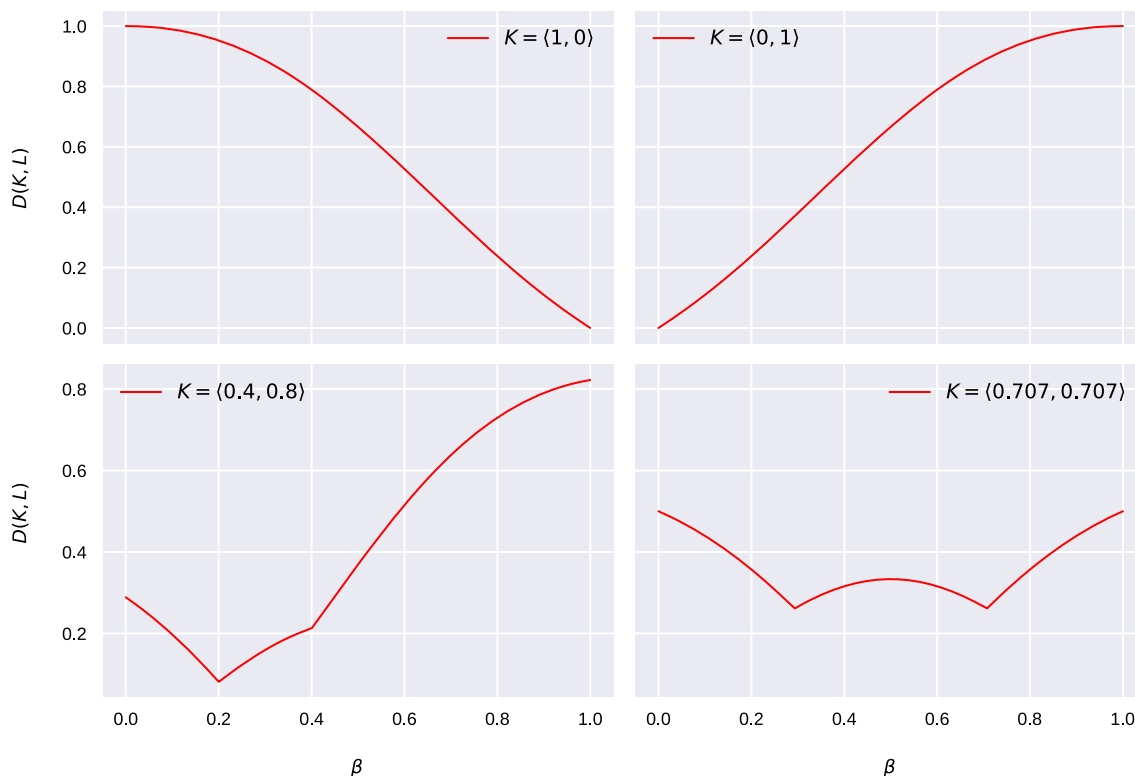


Fig. 3 Variation of $D(K, L)$ with respect to β for some selected K

Table 1 Comparison of different distance measures

Distance	D	D_{HmN}	D_{EN}	$D_{CN}(\beta = 5)$	D_{FW}	D_{HN}
$D(M, K)$.980	.049	.045	.046	.128	.040
$D(M, L)$.969	.049	.045	.045	.126	.040

The results for proposed distance functions are marked in bold face

Table 2 Comparison of distance measures for different pairs of PFSs A & B

A_i	A_1	A_2	A_3	A_4
B_i	B_1	B_2	B_3	B_4
D	.94	.95	.090	.135
D_{HmN}	.02	.02	.150	.150
D_{EN}	.02	.02	.141	.141
D_{HN}	.01	.01	.100	.100
$D_{CN}(\beta = 5)$.02	.02	.155	.155
D_{FW}	.08	.08	.150	.111

$B_3 = \{ \langle .640, .640 \rangle, \langle .714, .640 \rangle \};$
 $B_4 = \{ \langle .781, .557 \rangle, \langle .557, .458 \rangle \},$ such that $A_1 \neq A_3$ and all B_i 's are distinct. Distances between such pairs of PFSs are evaluated using the proposed definition (Eq. (13)) & some

popular distance measures (using Eqs. (6)–(10)) with a comparison given in Table 2.

It is evident from Table 2 that, our distance function can clearly distinguish such pairs by comparing the values in columns 1 & 2 and 3 & 4, respectively, while others fail to do so. Results obtained via the proposed distance are marked in bold faces in the table.

5 Applications of the proposed definition

5.1 Multi-attribute decision-making (MADM) problems

In an MADM problem, one tries to make the best possible decision out of finite alternatives, by taking into account a collection of clashing attributes with preferred weights. Here all the alternatives against each attribute are represented by PFSs and the concept of similarity measure is used to obtain a decision. Summarizing the scores of evaluators for each attribute concerning every alternative the representation is obtained. Weights for the attributes have to be determined a priori. A decision is achieved by implementing the following algorithm.

- Step 1: Write the PFS representation of the alternatives, say, $M_i = \{ \langle M_Y^{(i)}(x_j), M_N^{(i)}(x_j) \rangle | x_j \in X \},$

Table 3 PFS representation of different mask types

Mask type	A_1	A_2	A_3	A_4
M_1	$\langle 0.73, 0.29 \rangle$	$\langle 0.8, 0.12 \rangle$	$\langle 0.28, 0.44 \rangle$	$\langle 0.28, 0.68 \rangle$
M_2	$\langle 0.29, 0.61 \rangle$	$\langle 0.54, 0.63 \rangle$	$\langle 0.27, 0.68 \rangle$	$\langle 0.30, 0.63 \rangle$
M_3	$\langle 0.54, 0.49 \rangle$	$\langle 0.44, 0.56 \rangle$	$\langle 0.61, 0.54 \rangle$	$\langle 0.73, 0.42 \rangle$
M_4	$\langle 0.39, 0.64 \rangle$	$\langle 0.34, 0.43 \rangle$	$\langle 0.45, 0.31 \rangle$	$\langle 0.23, 0.61 \rangle$
M_5	$\langle 0.53, 0.29 \rangle$	$\langle 0.45, 0.66 \rangle$	$\langle 0.73, 0.44 \rangle$	$\langle 0.60, 0.63 \rangle$
M_6	$\langle 0.10, 0.25 \rangle$	$\langle 0.32, 0.27 \rangle$	$\langle 0.43, 0.37 \rangle$	$\langle 0.60, 0.60 \rangle$

Table 4 Distances of M_i from M^+/M^- and relative similarity measures

	M_1	M_2	M_3	M_4	M_5	M_6
$D(M^+, M_i)$	0.306	0.632	0.322	0.577	0.323	0.540
$D(M^-, M_i)$	0.497	0.104	0.424	0.288	0.429	0.514
S_r	0.513	0.464	0.507	0.480	0.508	0.498

Table 5 PFS representation of different schools

Schools	A_1	A_2	A_3	A_4	A_5
S_1	$\langle 0.7, 0.6 \rangle$	$\langle 0.9, 0.4 \rangle$	$\langle 0.8, 0.4 \rangle$	$\langle 0.6, 0.5 \rangle$	$\langle 0.5, 0.4 \rangle$
S_2	$\langle 0.6, 0.5 \rangle$	$\langle 0.8, 0.3 \rangle$	$\langle 0.9, 0.1 \rangle$	$\langle 0.7, 0.6 \rangle$	$\langle 0.5, 0.6 \rangle$
S_3	$\langle 0.8, 0.4 \rangle$	$\langle 0.7, 0.5 \rangle$	$\langle 0.6, 0.2 \rangle$	$\langle 0.4, 0.3 \rangle$	$\langle 0.7, 0.2 \rangle$
S_4	$\langle 0.6, 0.4 \rangle$	$\langle 0.9, 0.2 \rangle$	$\langle 0.8, 0.5 \rangle$	$\langle 0.7, 0.5 \rangle$	$\langle 0.8, 0.3 \rangle$
S_5	$\langle 0.8, 0.5 \rangle$	$\langle 0.7, 0.4 \rangle$	$\langle 0.9, 0.3 \rangle$	$\langle 0.4, 0.5 \rangle$	$\langle 0.6, 0.5 \rangle$

Table 6 Relative similarity measures corresponding to w_1 & w_2

$S_r(S_i)$	$S(D)$	$S_r(S_j)$				
		$S_r(S_1)$	$S_r(S_2)$	$S_r(S_3)$	$S_r(S_4)$	$S_r(S_5)$
$1 - D$	w_1	0.5007	0.5016	0.4975	0.5093	0.5010
	w_2	0.4986	0.4987	0.5013	0.5096	0.5012
$1 - D$	w_1	0.5013	0.5031	0.4953	0.5180	0.5018
$1 + D$	w_2	0.4974	0.4975	0.5024	0.5183	0.5023
$\frac{e^{-D} - e^{-1}}{1 - e^{-1}}$	w_1	0.5011	0.5025	0.4962	0.5145	0.5015
	w_2	0.4979	0.4980	0.5019	0.5148	0.5018

popularly known as decision matrix. Here, $M_Y^{(i)}$ and $M_N^{(i)}$ represent the guaranteed membership and non-membership grades, respectively, for the i^{th} alternative, M_i .

Step 2: Construct positive ideal PFS using Eq. (4) from the alternatives as

$$M^+ = \bigcup_i M_i. \tag{14}$$

Step 3: Similarly construct negative ideal PFS using Eq. (5) as

$$M^- = \bigcap_i M_i. \tag{15}$$

Step 4: Calculate the distances $D(M_i, M^+)$ and $D(M_i, M^-)$. For a given weight vector $w = \{w_i, 1 \leq i \leq n\}$ the distance measure should be interpreted as the weighted distance given by

$$D(K, L) = \frac{1}{n} \sum_{i=1}^n w_i \frac{|K_Y^2(x_i) - L_Y^2(x_i)| + |K_N^2(x_i) - L_N^2(x_i)|}{K_Y^2(x_i) + L_Y^2(x_i) + K_N^2(x_i) + L_N^2(x_i)}. \tag{16}$$

Step 5: Measure the similarity S , using the distance calculated in the step 4. Here various functional forms of similarity measure can be explored.

Step 6: Evaluate relative similarity S_r for M_i through the formula

$$S_r(M_i) = \frac{S(D_+^{(i)})}{S(D_+^{(i)}) + S(D_-^{(i)})}, \tag{17}$$

where $D_+^{(i)} = D(M_i, M^+)$ and $D_-^{(i)} = D(M_i, M^-)$ are estimated using Eq. (16).

Step 7: Rank the alternatives M_i in ascending order of relative similarity values.

5.1.1 Face mask selection

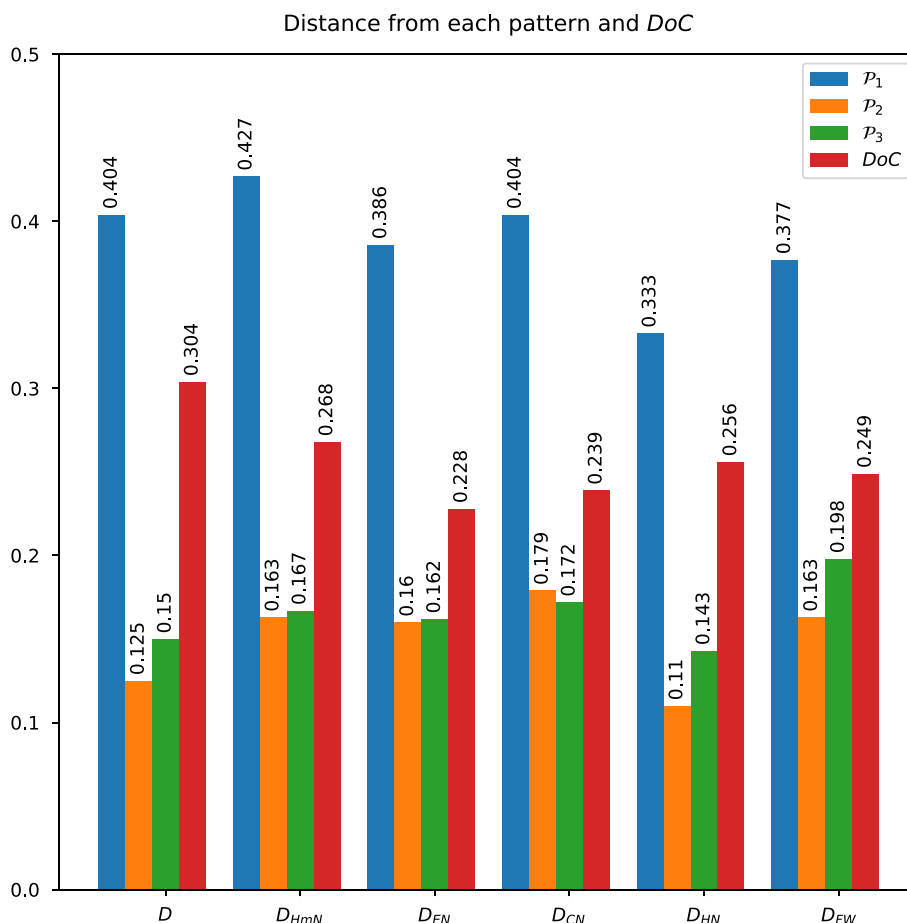
In the COVID-19 pandemic, wearing a face-mask has become an integral part of the life. There are plenty of face-masks available in the market, few among them are disposable medical masks (M_1), medical surgical masks (M_2), particulate respirators (N95) (M_3), ordinary non-medical masks (M_4), medical protective masks (M_5) and gas masks (M_6). The set of clashing attributes for each mask type are leakage rate (A_1), reutilizability (A_2), quality of raw material (A_3) and filtration capability (A_4). To buy a face-mask a person will put his/her preferred weight on the four attributes which leads to an MADM problem. The PFS representation of each attribute for every mask type is given in Table 3 (for more details please refer to [40]).

From Table 3, positive and negative ideal PFSs are calculated using Eqs. (14) & (15), respectively, as

$$M^+ = \{ \langle 0.73, 0.25 \rangle, \langle 0.80, 0.12 \rangle, \langle 0.73, 0.31 \rangle, \langle 0.73, 0.42 \rangle \};$$

$$M^- = \{ \langle 0.10, 0.64 \rangle, \langle 0.32, 0.66 \rangle, \langle 0.27, 0.68 \rangle, \langle 0.23, 0.68 \rangle \}.$$

Fig. 4 PFS distances of a test sample from patterns



By putting equal weight on each attribute, i.e. $w = \{.25, .25, .25, .25\}$, we calculate the distances between M_i and M^+/M^- which are listed in Table 4. Using the distance calculated, measure the similarity between M_i and M^+/M^- via the operator $S(D) = 1 - D$. Further the relative similarity $S_r(M_i)$ is obtained using Eq. (17) and results are displayed in Table 4.

Finally, using Step 7, the ranking of face-mask is obtained as follows

$$M_1 \succ M_5 \succ M_3 \succ M_6 \succ M_4 \succ M_2.$$

Therefore, the best choice for the face-mask turned out to be M_1 , i.e. disposable medical masks which is consistent with the existing result [40]. A buyer may put more stress on a particular attribute compared to rest, say the attribute - filtration capability (A_4) to select a face-mask. So, the weight vector modifies to, $w = \{0.2, 0.1, 0.1, 0.6\}$ for the attributes, $A = \{A_1, A_2, A_3, A_4\}$. Following the steps of the algorithm, the best choice of face-mask is M_3 i.e. particulate respirators (N95) mask. Above discussion demonstrates the effect of individual preference on attributes in the outcome of decision-making process.

5.1.2 School selection [38]

Nowadays in spite of having good public sector schools, the trend among parents is to send their children to private sector. Of late private sector schools just spring up like mushrooms. So selecting a suitable private school has become a daunting task for parents. Let us consider a collection of five private sector schools, designated as S_i where $i = 1, 2, \dots, 5$, in a locality. Considering education- alists and experts opinion the following attributes of the schools are found to be the most influencing and significant [38]: personalized attention (A_1), academic accolades (A_2), skilful and up-to-date teachers (A_3), parents and community engagement (A_4) and technical facilities (A_5). The PFS representation of each school against the attributes is detailed in Table 5.

Following the algorithm, we obtain

$$S^+ = \{\langle 0.8, 0.4 \rangle, \langle 0.9, 0.2 \rangle, \langle 0.9, 0.1 \rangle, \langle 0.7, 0.3 \rangle, \langle 0.8, 0.2 \rangle\}$$

$$S^- = \{\langle 0.6, 0.6 \rangle, \langle 0.7, 0.5 \rangle, \langle 0.6, 0.5 \rangle, \langle 0.4, 0.6 \rangle, \langle 0.5, 0.6 \rangle\}$$

Choices of weightage to the attributes are subjective. Here, we consider two distinct weight vectors, say $w_1 = \{0.18, 0.32, 0.23, 0.13, 0.14\}$, a case of unequal

Table 7 PFS representation of symptoms

Name of patient	S_1	S_2	S_3	S_4	S_5
Ali	$\langle 0.8, 0.1 \rangle$	$\langle 0.6, 0.1 \rangle$	$\langle 0.2, 0.8 \rangle$	$\langle 0.6, 0.1 \rangle$	$\langle 0.1, 0.6 \rangle$
Baruah	$\langle 0.0, 0.8 \rangle$	$\langle 0.4, 0.4 \rangle$	$\langle 0.6, 0.1 \rangle$	$\langle 0.1, 0.7 \rangle$	$\langle 0.1, 0.8 \rangle$
Chatterjee	$\langle 0.8, 0.1 \rangle$	$\langle 0.8, 0.1 \rangle$	$\langle 0.0, 0.6 \rangle$	$\langle 0.2, 0.7 \rangle$	$\langle 0.0, 0.5 \rangle$
Deka	$\langle 0.6, 0.1 \rangle$	$\langle 0.5, 0.4 \rangle$	$\langle 0.3, 0.4 \rangle$	$\langle 0.7, 0.2 \rangle$	$\langle 0.3, 0.4 \rangle$

Table 8 Symptom characteristics for diagnosis

Disease	S_1	S_2	S_3	S_4	S_5
D_1	$\langle 0.4, 0.0 \rangle$	$\langle 0.3, 0.5 \rangle$	$\langle 0.1, 0.7 \rangle$	$\langle 0.4, 0.3 \rangle$	$\langle 0.1, 0.7 \rangle$
D_2	$\langle 0.7, 0.0 \rangle$	$\langle 0.2, 0.6 \rangle$	$\langle 0.0, 0.9 \rangle$	$\langle 0.7, 0.0 \rangle$	$\langle 0.1, 0.8 \rangle$
D_3	$\langle 0.3, 0.3 \rangle$	$\langle 0.6, 0.1 \rangle$	$\langle 0.2, 0.7 \rangle$	$\langle 0.2, 0.6 \rangle$	$\langle 0.1, 0.9 \rangle$
D_4	$\langle 0.1, 0.7 \rangle$	$\langle 0.2, 0.4 \rangle$	$\langle 0.8, 0.0 \rangle$	$\langle 0.2, 0.7 \rangle$	$\langle 0.2, 0.7 \rangle$
D_5	$\langle 0.1, 0.8 \rangle$	$\langle 0.0, 0.8 \rangle$	$\langle 0.2, 0.8 \rangle$	$\langle 0.2, 0.8 \rangle$	$\langle 0.8, 0.1 \rangle$

weightage and $w_2 = \{0.2, 0.2, 0.2, 0.2, 0.2\}$, an example of equal emphasis to each attribute. Using w_1 & w_2 , we evaluate the weighted distance between S_i and S^+/S^- using

Fig. 5 PFS distances of each symptom from all possible diagnosis

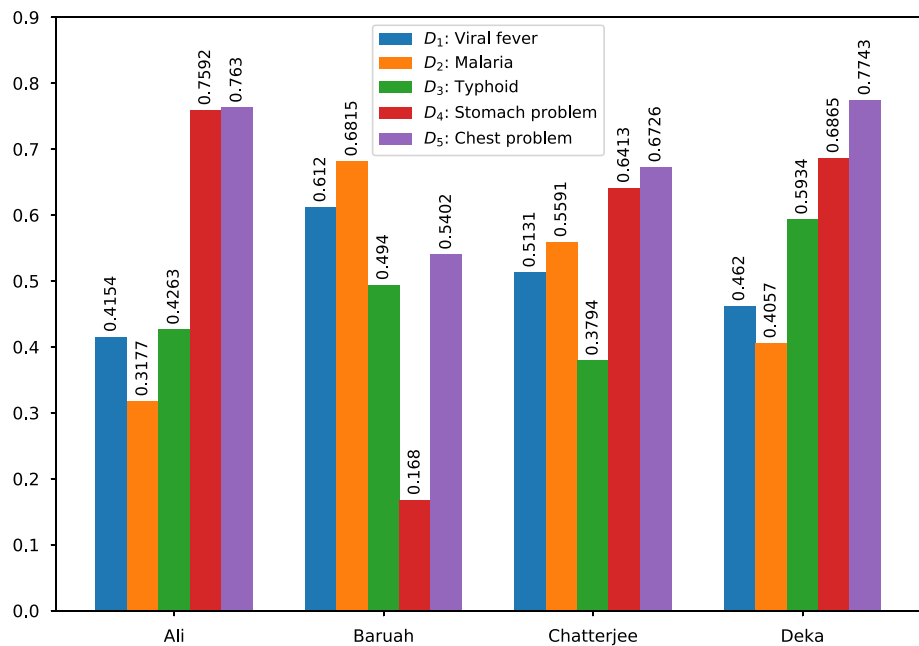


Table 9 Comparison of prediction

Patient	Our prediction	Others' prediction
Ali	Malaria	Malaria [21, 24, 42–48, 50], Viral fever [49]
Baruah	Stomach problem	Stomach problem [21, 24, 42–50]
Chatterjee	Typhoid	Typhoid [21, 42, 43, 45–50], Malaria [24, 44]
Deka	Malaria	Viral fever [21, 42, 45, 47, 48, 50], Malaria [24, 43, 44, 46, 49]

Eq. (16). These weighted distances are further used to obtain similarity measures and subsequently relative similarity measures. We have employed three distinct functional forms of similarity measure ($S(D)$) as listed in Table 6.

Higher the value of relative similarity, better is the rank. On that note,

- ◇ ranking corresponding to w_1 is $S_4 \succ S_2 \succ S_5 \succ S_1 \succ S_3$. The best choice of the private sector school turns out to be S_4 which is matching with the existing literature [38].
- ◇ Ranking corresponding to w_2 is $S_4 \succ S_3 \succ S_5 \succ S_2 \succ S_1$. Interestingly, here also the best option is found to be S_4 ; however positions of the rest have altered.

It is to be observed that, different functional forms of similarity measures have provided same ranking irrespective of the weight vector.

5.2 Pattern recognition

A pattern recognition problem is addressed using the proposed distance measure. The problem is consisting of a test sample $\mathcal{S} = \{\langle 0.7, 0.7 \rangle, \langle 0.3, 0.8 \rangle, \langle 0.7, 0.2 \rangle\}$, which is to be classified as one among three given patterns $\mathcal{P}_1 = \{\langle 0.5, 0.6 \rangle, \langle 0.7, 0.3 \rangle, \langle 0.8, 0.5 \rangle\}$, $\mathcal{P}_2 = \{\langle 0.6, 0.6 \rangle, \langle 0.4, 0.8 \rangle, \langle 0.6, 0.1 \rangle\}$ and $\mathcal{P}_3 = \{\langle 0.6, 0.7 \rangle, \langle 0.4, 0.9 \rangle, \langle 0.6, 0.4 \rangle\}$. The steps to be followed in the process are:

- ◇ Step 1: Obtain the PFS representations of the patterns as well as the sample.
- ◇ Step 2: Calculate the distance of \mathcal{S} from each of $\mathcal{P}_1, \mathcal{P}_2$ and \mathcal{P}_3 via Eq. (13).
- ◇ Step 3: Classify the test sample as the pattern from which minimum value is obtained in step 2.
- ◇ Step 4: Calculate degree of confidence (*DoC*) [41] using the formula :

$$DoC = \sum_{j=1, j \neq j_0}^n |D(\mathcal{P}_j, \mathcal{S}) - D(\mathcal{P}_{j_0}, \mathcal{S})|, \quad (18)$$

where \mathcal{P}_{j_0} is the classified pattern for the test sample \mathcal{S} . Greater value of *DoC* guarantees the higher confidence level of the prediction. It is to be noted that, the process can be readily applied for a problem with n numbers of patterns.

Figure 4 includes the distances of \mathcal{S} from the three given patterns using our proposed definition and *DoC* against this prediction. From Fig. 4, it is evident that the test sample \mathcal{S} has minimum distance from pattern \mathcal{P}_2 and hence is classified as \mathcal{P}_2 . For comparison step 2 is repeated using $D_{HmN}, D_{EN}, D_{CN}, D_{HN}$ & D_{FW} and the results are displayed in Fig. 4. Except D_{CN} , others have predicted the test sample as pattern \mathcal{P}_2 , whereas the distance formula D_{CN} with $\beta = 4$ has classified it as \mathcal{P}_3 . Our result coincides with the conclusions drawn through majority of the distance measures. Besides that, the proposed distance has the highest value of *DoC* which assures the confidence level of the classification.

5.3 Medical diagnosis [24, 42–50]

Suppose there are four patients $P = \{\text{Ali, Baruah, Chatterjee, Deka}\}$ having the symptoms $S = \{S_1: \text{Temperature}, S_2: \text{Headache}, S_3: \text{Stomach pain}, S_4: \text{Cough}, S_5: \text{Chest pain}\}$. The symptoms of the patients represented by PFS are listed in Table 7. The set of

possible diagnosis $D = \{D_1: \text{Viral fever}, D_2: \text{Malaria}, D_3: \text{Typhoid}, D_4: \text{Stomach problem}, D_5: \text{Chest problem}\}$ is represented in Table 8.

Now we calculate the distance between S and D for each of the patients using the proposed definition (Eq. (13)) and have obtained the results as displayed in Fig. 5. In Table 9 we provide a comparison of our prediction with that of few existing results obtained via either IFS- or PFS-based distance measures and thereby confirming our predictions with the majority of existing ones.

6 Conclusions

Here, we have introduced a distance measure for PFSs whose simplicity in the mathematical form is an extra boon to it. The salient feature of the distance measure is its efficacy in distinguishing PFSs with high hesitancy. Supremacy of the distance function is established via suitable numerical examples and reasonability is tested by beating the counter-intuitive cases. Nonlinearity & bounded characteristics are proved and verified through proper examples. Applicability of the distance measure is established in various fields like pattern recognition, medical diagnosis and multi-attribute decision-making problems. We believe, the proposed distance function will find its serviceability in new avenues of application.

Acknowledgements The authors would like to acknowledge anonymous reviewers for their valuable suggestions which enrich the manuscript a lot.

Declarations

Conflict of interest The authors have no conflicts of interest to declare that are relevant to the content of this article.

References

1. Zadeh LA (1965) Fuzzy sets. *Inf Control* 8(3):338–353
2. Atanassov KT (1986) Intuitionistic fuzzy sets. *Fuzzy Sets Syst* 20(1):87–96
3. Atanassov KT (1994) New operations defined over the intuitionistic fuzzy sets. *Fuzzy Sets Syst* 61(2):137–142
4. Atanassov KT (1999) Intuitionistic fuzzy sets. In *Intuitionistic fuzzy sets*. Springer, pp 1–137
5. Atanassov KT (1999) Interval valued intuitionistic fuzzy sets. In *Intuitionistic fuzzy sets*. Springer, pp 139–177
6. Yager RR (2013) Pythagorean fuzzy subsets. In 2013 joint IFSA world congress and NAFIPS annual meeting (IFSA/NAFIPS). IEEE, pp 57–61
7. Yager RR (2016) Properties and applications of pythagorean fuzzy sets. In *Imprecision and uncertainty in information representation and processing*. Springer, pp 119–136

8. Peng X, Garg H (2019) Multiparametric similarity measures on pythagorean fuzzy sets with applications to pattern recognition. *Appl Intell* 49(12):4058–4096
9. Ejegwa PA (2020) Modified Zhang and Xu's distance measure for pythagorean fuzzy sets and its application to pattern recognition problems. *Neural Comput Appl* 32(14):10199–10208
10. Ejegwa PA, Awolola JA (2019) Novel distance measures for pythagorean fuzzy sets with applications to pattern recognition problems. *Granul Comput*, pp 1–9
11. Ullah K, Mahmood T, Ali Z, Jan N (2020) On some distance measures of complex pythagorean fuzzy sets and their applications in pattern recognition. *Complex Intell Syst* 6(1):15–27
12. Singh S, Ganie AH (2020) On some correlation coefficients in pythagorean fuzzy environment with applications. *Int J Intell Syst* 35(4):682–717
13. Yager RR, Abbasov AM (2013) Pythagorean membership grades, complex numbers, and decision making. *Int J Intell Syst* 28(5):436–452
14. Yager RR (2013) Pythagorean membership grades in multicriteria decision making. *IEEE Trans Fuzzy Syst* 22(4):958–965
15. Liu Y, Liu J, Qin Y (2020) Pythagorean fuzzy linguistic muirhead mean operators and their applications to multiattribute decision-making. *Int J Intell Syst* 35(2):300–332
16. Verma R, Merigó JM (2019) On generalized similarity measures for pythagorean fuzzy sets and their applications to multiple attribute decision-making. *Int J Intell Syst* 34(10):2556–2583
17. Garg H (2016) A new generalized pythagorean fuzzy information aggregation using Einstein operations and its application to decision making. *Int J Intell Syst* 31(9):886–920
18. Ren P, Zeshui X, Gou X (2016) Pythagorean fuzzy TODIM approach to multi-criteria decision making. *Appl Soft Comput* 42:246–259
19. Zhang X (2016) Multicriteria pythagorean fuzzy decision analysis: a hierarchical qualiflex approach with the closeness index-based ranking methods. *Inf Sci* 330:104–124
20. Garg H (2016) A novel correlation coefficients between pythagorean fuzzy sets and its applications to decision-making processes. *Int J Intell Syst* 31(12):1234–1252
21. Xiao F, Ding W (2019) Divergence measure of pythagorean fuzzy sets and its application in medical diagnosis. *Appl Soft Comput* 79:254–267
22. Thao NX (2019) A new correlation coefficient of the pythagorean fuzzy sets and its applications. *Soft Comput*, pp 1–12
23. Wei G, Wei Yu (2018) Similarity measures of pythagorean fuzzy sets based on the cosine function and their applications. *Int J Intell Syst* 33(3):634–652
24. Ejegwa PA (2020) Improved composite relation for pythagorean fuzzy sets and its application to medical diagnosis. *Granul Comput* 5(2):277–286
25. Reformat MZ, Yager RR (2014) Suggesting recommendations using pythagorean fuzzy sets illustrated using netflix movie data. In *International conference on information processing and management of uncertainty in knowledge-based systems*. Springer, pp 546–556
26. Ejegwa PA (2019) Pythagorean fuzzy set and its application in career placements based on academic performance using max-min-max composition. *Complex Intell Syst* 5(2):165–175
27. Li Z, Mao L (2019) Some novel similarity and distance measures of pythagorean fuzzy sets and their applications. *J Intell Fuzzy Syst* 37(2):1781–1799
28. Gülçin B, Fethullah G (2019) A novel approach integrating ahp and copras under pythagorean fuzzy sets for digital supply chain partner selection. *IEEE Trans Eng Manag*
29. Kumar R, Edalatpanah SA, Jha S, Singh R (2019) A pythagorean fuzzy approach to the transportation problem. *Complex Intell Syst* 5(2):255–263
30. Li Z, Wei G (2019) Pythagorean fuzzy heronian mean operators in multiple attribute decision making and their application to supplier selection. *Int J Knowl Based Intell Eng Syst* 23(2):77–91
31. Khalifa AEL et al (2020) Characterizing solution for stock portfolio problem via pythagorean fuzzy approach. *Environ Energy Econ Res* 4(2):127–134
32. Chen T-Y (2018) Remoteness index-based pythagorean fuzzy vikor methods with a generalized distance measure for multiple criteria decision analysis. *Inf Fus* 41:129–150
33. Li D, Zeng W (2018) Distance measure of pythagorean fuzzy sets. *Int J Intell Syst* 33(2):348–361
34. Ejegwa PA (2018) Distance and similarity measures for pythagorean fuzzy sets. *Granul Comput*, pp 1–14
35. Peng X (2019) New similarity measure and distance measure for pythagorean fuzzy set. *Complex Intell Syst* 5(2):101–111
36. Adabitabar FM, Agheli B, Baloui JE (2020) A new similarity measure for pythagorean fuzzy sets. *Complex Intell Syst* 6(1):67–74
37. Zhang X, Zeshui X (2014) Extension of topsis to multiple criteria decision making with pythagorean fuzzy sets. *Int J Intell Syst* 29(12):1061–1078
38. Hussian Z, Yang M-S (2019) Distance and similarity measures of pythagorean fuzzy sets based on the hausdorff metric with application to fuzzy topsis. *Int J Intell Syst* 34(10):2633–2654
39. Wilson WA (1931) On quasi-metric spaces. *Am J Math* 53(3):675–684
40. Yang Z, Li X, Garg H, Qi M (2020) Decision support algorithm for selecting an antivirus mask over covid-19 pandemic under spherical normal fuzzy environment. *Int J Environ Res Public Health* 17(10):3407
41. Hatzimichailidis AG, Papakostas GA, Kaburlasos VG (2012) A novel distance measure of intuitionistic fuzzy sets and its application to pattern recognition problems. *Int J Intell Syst* 27(4):396–409
42. Cheng C, Xiao F, Cao Z (2019) A new distance for intuitionistic fuzzy sets based on similarity matrix. *IEEE Access* 7:70436–70446
43. Own C-M (2009) Switching between type-2 fuzzy sets and intuitionistic fuzzy sets: an application in medical diagnosis. *Appl Intell* 31(3):283
44. De SK, Biswas R, Roy AR (2001) An application of intuitionistic fuzzy sets in medical diagnosis. *Fuzzy Sets Syst* 117(2):209–213
45. Wei C-P, Wang P, Zhang Y-Z (2011) Entropy, similarity measure of interval-valued intuitionistic fuzzy sets and their applications. *Inf Sci* 181(19):4273–4286
46. Luo M, Zhao R (2018) A distance measure between intuitionistic fuzzy sets and its application in medical diagnosis. *Artif Intell Med* 89:34–39
47. Szmjdt E, Kacprzyk J (2001) Intuitionistic fuzzy sets in intelligent data analysis for medical diagnosis. In *International conference on computational science*. Springer, pp 263–271
48. Mondal K, Pramanik S (2015) Intuitionistic fuzzy similarity measure based on tangent function and its application to multi-attribute decision making. *Glob J Adv Res* 2(2):464–471
49. Szmjdt E, Kacprzyk J (2004) A similarity measure for intuitionistic fuzzy sets and its application in supporting medical diagnostic reasoning. In *International conference on artificial intelligence and soft computing*. Springer, pp 388–393
50. Peng X, Yuan H, Yang Y (2017) Pythagorean fuzzy information measures and their applications. *Int J Intell Syst* 32(10):991–1029