Modified Szmidt and Kacprzyk’s Intuitionistic Fuzzy Distances and their Applications in Decision-making


Abstract

Intuitionistic fuzzy models are significant in resolving decision-making. Distance measures under intuitionistic fuzzy environment are reliable techniques deployed to express the application of IFSs. Some approaches of estimating distances between IFSs have been explored by Szmidt and Kacprzyk, where the complete parameters of IFSs are considered. Albeit, the distance operators lack reliability because of certain setbacks. In this paper, we modified Szmidt and Kacprzyk’s distance operators between IFSs to enhance reliability in terms of applications. Some theorems are given to substantiate the validity of the modified intuitionistic fuzzy distance operators. Furthermore, decision-making cases of pattern recognition and disease identification are discussed using the Szmidt and Kacprzyk’s distances and their improved versions where information are represented in intuitionistic fuzzy pairs. From the study, it is observed that the modified Szmidt and Kacprzyk’s distance operators between IFSs yield better results compare to the Szmidt and Kacprzyk’s distance operators between IFSs.

Keywords: Decision-making, Distance measure, Intuitionistic fuzzy set, Pattern recognition, Medical diagnosis

Article History:
Received: 12 December 2021
Received in revised form: 21 April 2022
Accepted for publication: 24 April 2022
Published: 29 May 2022

©2022 Journal of the Nigerian Society of Physical Sciences. All rights reserved.
Communicated by: T. Latunde

1. Introduction

Decision-making is a critical task enmeshed with vagueness. With the introduction of fuzzy sets by Zadeh [1], the problem of vagueness has been considerably tackled to enhance the solution of many decision-making problems including medical diagnosis, career determination, pattern recognition among others. Fuzzy set theory although relevant has a setback in the sense that it considers only the membership degree \( \mu \) (MD) of the case under consideration. Because of this setback, Atanassov [2] proposed a generalized fuzzy set called intuitionistic fuzzy set (IFS). IFS is described by membership degree \( \mu \), non-membership degree \( \nu \) and intuitionistic fuzzy index \( \varpi \) with the property that their sum is unity.

IFSs have been applied in sundry cases [3, 4, 5]. Chen et al. [6] discussed fuzzy queries process based on intuitionistic fuzzy social networks, De et al. [7] applied IFSs to medical decision via composite relation, Ejegwa and Onasanya [8] improved the intuitionistic fuzzy composite relation in [7] with application to medical diagnosis, and Liu and Chen [9] presented a group decision-making based on Heronian aggregation operators of IFSs. Several information measures have been studied to
enhance the application of IFSs in real-life problems. Ejegwa [10] proposed a new correlation coefficient of IFSs and applied the measure to solve multi-criteria decision-making (MCDM) problems. Many correlation coefficients of IFSs have been proposed and applied to several decision-making problems [11, 12, 13, 14, 15]. Medical diagnostic problems were solved based on intuitionistic fuzzy correlation coefficient as seen in [16, 17, 18]. Garg [19] presented a correlation coefficient under intuitionistic multiplicative environment with application in decision-making process, and a new correlation coefficient of IFSs based on the connection number of set pair analysis was studied with application [20]. A robust technique of computing the correlation coefficients of complex IFSs and their applications in decision-making were discussed in [21]. TOPSIS method based on correlation coefficient under intuitionistic fuzzy soft sets was discussed and with application [22]. The idea of aggregation operators has been applied in many cases of decision-making [23, 24, 25, 26], and fuzzy soft set-valued maps has been introduced and applied to homotopy [27].

Similarly, the concepts of similarity and distance measures under intuitionistic fuzzy context have been discussed as reliable information measures. Boran and Akay [28] proposed a biparametric similarity measure and applied the measure to pattern recognition. A similarity measure of IFSs based on transformation technique has been studied and applied to pattern recognition [29]. In [30, 31], some similarity measures of IFSs were introduced and applied to medical diagnostic reasoning. In addition, some similarity measures based on dice and Jaccard approaches have been studied on expected intervals of trapezoidal neutrosophic fuzzy numbers with application to MCDM [32]. Burillo and Bustince [33] initiated the concept of distances for IFSs and interval-valued fuzzy sets. Szmidt and Kacprzyk [34] modified the distances in [33], and showed that all the three parameters describing IFSs should be taken into account while calculating distances between IFSs. Hatzimichailidis et al. [35] introduced a novel distance measure between IFSs with application to cases of pattern recognition. Wang and Xin [36] proposed a novel distance measure between IFSs and its weighted version with application to the solution of pattern recognition problem. Davvaz and Sadrabadi [37] revised some existing distance measures and applied them to medical diagnostic process, and other applications of distance measures between IFSs have been studied [38, 39, 40, 41, 42].

Among the distances between IFSs studied in literature, distances in [34] are prominent for their reliable interpretations. Albeit, these distances show some limitations which needed to be strengthened to enhance reliable outputs. Although Szmidt and Kacprzyk [34] modified the intuitionistic fuzzy distance measures in [33] with better rating, they do not take account the number of the considered parameters, but just added the hesitation margins to the methods introduced in [33]. This setback adversely influence the performance rating of Szmidt and Kacprzyk’s distances [34]. The motivation for this work is to propose modified Szmidt and Kacprzyk’s distances which have better performance indexes compare to Szmidt and Kacprzyk’s distances taking into account the number of the considered parameters to forestall inaccurate outputs. Specifically, the objectives of this paper are to (i) revisit Szmidt and Kacprzyk’s distances between IFSs, (ii) propose modified versions of the distances between IFSs in [34], (iii) apply the modified distances between IFSs to determine pattern recognition and disease diagnosis, and (iv) present comparison analyses of the modified distances with the Szmidt and Kacprzyk’s distances in intuitionistic fuzzy domain. The paper is outlined as follow; Section 2 presents the concept of IFSs and discusses the Szmidt and Kacprzyk’s distances between IFSs [34]. Section 3 introduces the modified Szmidt and Kacprzyk’s distances between IFSs, Section 4 discusses the applications of Szmidt and Kacprzyk’s distances and their modified versions in pattern recognition and disease diagnosis, and Section 5 concludes the paper with recommendations for future studies.

2. Preliminaries

This section presents the concept of IFSs and discusses the Szmidt and Kacprzyk’s distances between IFSs.

2.1. Intuitionistic fuzzy sets

Numerous works on IFSs have been carried out [2, 43, 3]. Here, some basic concepts of IFSs are presented. Let us assume that \( Y \) is a non-empty set throughout this paper.

Definition 2.1. [43] An intuitionistic fuzzy set \( C \) of \( Y \) is defined by \( C = \{(y, \nu_C(y), \upsilon_C(y)) : y \in Y\} \), where the functions \( \nu_C, \upsilon_C : Y \rightarrow [0, 1] \) are the membership and non-membership degrees of \( y \in Y \), and \( 0 \leq \nu_C(y) + \upsilon_C(y) \leq 1 \). For a IFS \( C \) in \( Y \), \( \sigma_C(y) \in [0, 1] = 1 - \nu_C(y) - \upsilon_C(y) \) is the intuitionistic fuzzy index or hesitation margin of \( C \).

Definition 2.2. [3] Suppose \( C \) and \( D \) are IFSs in \( Y \), then for all \( y \in Y \) we have

(i) \( C = D \) iff \( \nu_C(y) = \nu_D(y), \upsilon_C(y) = \upsilon_D(y) \).
(ii) \( C \subseteq D \) iff \( \nu_C(y) \leq \nu_D(y), \upsilon_C(y) \geq \upsilon_D(y) \).
(iii) \( C = \{(y, \nu_C(y), \upsilon_C(y)) : y \in Y\} \).
(iv) \( C \cup D = \{(y, \max(\nu_C(y), \nu_D(y)), \min(\nu_C(y), \nu_D(y))) : y \in Y\} \).
(v) \( C \cap D = \{(y, \min(\nu_C(y), \nu_D(y)), \max(\nu_C(y), \nu_D(y))) : y \in Y\} \).

Definition 2.3. [3] Intuitionistic fuzzy pair (IFP) is characterized by the form \( \langle c, d \rangle \) such that \( c + d \leq 1 \) where \( c, d \in [0, 1] \). IFP evaluate the IFS for which the components \( c \) and \( d \) are interpreted as membership and non-membership degrees.

2.2. Distances between intuitionistic fuzzy sets

Distance measure is a soft computing tool used in the applications of IFSs. The definition of distance measure between IFSs is as follows.

Definition 2.4. [34] If \( C \) and \( D \) are IFSs of \( Y \), then the distance between \( C \) and \( D \) denoted by \( \phi(C, D) \) is a function \( \phi : IFS \times IFS \rightarrow [0, 1] \) which satisfies

(i) \( 0 \leq \phi(C, D) \leq 1 \).
(ii) \( \phi(C, D) = 0 \) if \( C = D \)
(iii) \( \phi(C, D) = \phi(D, C) \)
(iv) \( \phi(C, E) \leq \phi(C, D) + \phi(D, E) \), where \( E \) is also an IFS of \( Y \).

When \( \phi(C, D) \) reaches 0, it shows that \( C \) and \( D \) are more close or related. Again, if \( \phi(C, D) \) reaches 1 then \( C \) and \( D \) are not related or close. For any two IFSs \( C \) and \( D \) in \( Y = \{y_1, \ldots, y_n\} \), we present the following distances between them.

### 2.2.1. Burillo and Bustince’s distances between IFSs

By extending the distances between fuzzy sets as presented in [44], Burillo and Bustince [33] proposed the following distances under intuitionistic fuzzy environment:

\[
\phi_1(C, D) = \frac{1}{2} \sum_{i=1}^{n} \left( |v_C(y_i) - v_D(y_i)| + |v_C(y_i) - \nu_D(y_i)| \right) \tag{1}
\]

\[
\phi_2(C, D) = \left( \frac{1}{2} \sum_{i=1}^{n} ((v_C(y_i) - v_D(y_i))^2 + (v_C(y_i) - \nu_D(y_i))^2) \right)^{\frac{1}{2}} \tag{2}
\]

\[
\phi_3(C, D) = \frac{1}{2n} \sum_{i=1}^{n} \left( |v_C(y_i) - v_D(y_i)| + |v_C(y_i) - \nu_D(y_i)| \right) \tag{3}
\]

\[
\phi_4(C, D) = \left( \frac{1}{2n} \sum_{i=1}^{n} ((v_C(y_i) - v_D(y_i))^2 + (v_C(y_i) - \nu_D(y_i))^2) \right)^{\frac{1}{2}} \tag{4}
\]

The denominator in Eqs. (1-4) depicts the number of parameters of IFSs considered similar to the approach in [44] where the denominator is unity because fuzzy set considers only membership function. The limitation of these approaches [33] is that the hesitation margin is not considered in the computations.

### 2.2.2. Szmidt and Kacprzyk’s distances between IFSs

Because of the limitation in [33], Szmidt and Kacprzyk [34] proposed the same distances by incorporating hesitation margin of the considered IFSs. For simplicity sake, let \( v_C(y_i) = v_C \), \( v_C(y_i) = v_C \), \( \sigma_C(y_i) = \sigma_C \), \( \nu_D(y_i) = v_D \), \( \nu_D(y_i) = \nu_D \), \( \omega_D(y_i) = \omega_D \). The distances are as follow:

\[
\hat{\phi}_1(C, D) = \frac{1}{2} \sum_{i=1}^{n} \left( |v_C - v_D| + |v_C - \nu_D| + |\sigma_C - \omega_D| \right) \tag{5}
\]

\[
\hat{\phi}_2(C, D) = \left( \frac{1}{2} \sum_{i=1}^{n} ((v_C - v_D)^2 + (v_C - \nu_D)^2 + (\sigma_C - \omega_D)^2) \right)^{\frac{1}{2}} \tag{6}
\]

\[
\hat{\phi}_3(C, D) = \frac{1}{2n} \sum_{i=1}^{n} \left( |v_C - v_D| + |v_C - \nu_D| + |\sigma_C - \omega_D| \right) \tag{7}
\]

\[
\hat{\phi}_4(C, D) = \left( \frac{1}{2n} \sum_{i=1}^{n} ((v_C - v_D)^2 + (v_C - \nu_D)^2 + (\sigma_C - \omega_D)^2) \right)^{\frac{1}{2}} \tag{8}
\]

Though the distances in [34] captured the three parameters of IFSs, the denominators in each equations do not depict the number of parameters of the considered IFSs. These omissions will lead to information loss and thus, affect the interpretations.

### 3. Modified Szmidt and Kacprzyk’s Distances between IFSs

To enhance reliable outputs and avoid information loss, we modified the Szmidt and Kacprzyk’s distances between IFSs [34] using the same hypotheses in Subsection 2.2.2 as follow:

\[
\tilde{\phi}^*(C, D) = \left( \frac{1}{3} \sum_{i=1}^{n} ((v_C - v_D)^2 + (v_C - \nu_D)^2 + (\sigma_C - \omega_D)^2) \right)^{\frac{1}{2}} \tag{9}
\]

\[
\tilde{\phi}^{**}(C, D) = \left( \frac{1}{3n} \sum_{i=1}^{n} ((v_C - v_D)^2 + (v_C - \nu_D)^2 + (\sigma_C - \omega_D)^2) \right)^{\frac{1}{2}} \tag{10}
\]

for \( r \leq 2 \). If \( r = 1 \), we get

\[
\tilde{\phi}_1(C, D) = \frac{1}{3} \sum_{i=1}^{n} |v_C - v_D| + |v_C - \nu_D| + |\sigma_C - \omega_D| \tag{11}
\]

\[
\tilde{\phi}_2(C, D) = \frac{1}{3n} \sum_{i=1}^{n} |v_C - v_D| + |v_C - \nu_D| + |\sigma_C - \omega_D| \tag{12}
\]

If \( r = 2 \), we get

\[
\tilde{\phi}_3(C, D) = \left( \frac{1}{3} \sum_{i=1}^{n} ((v_C - v_D)^2 + (v_C - \nu_D)^2 + (\sigma_C - \omega_D)^2) \right)^{\frac{1}{2}} \tag{13}
\]

\[
\tilde{\phi}_4(C, D) = \left( \frac{1}{3n} \sum_{i=1}^{n} ((v_C - v_D)^2 + (v_C - \nu_D)^2 + (\sigma_C - \omega_D)^2) \right)^{\frac{1}{2}} \tag{14}
\]

**Proposition 3.1.** Suppose \( C \) and \( D \) are IFSs of \( Y = \{y_1, \ldots, y_n\} \), then we have (i) \( \tilde{\phi}_1(C, D) = n \tilde{\phi}_2(C, D) \) (ii) \( \tilde{\phi}_3(C, D) = \sqrt{n} \tilde{\phi}_4(C, D) \).

**Proof.** Given that \( \tilde{\phi}_2(C, D) = \frac{1}{3n} \sum_{i=1}^{n} |v_C - v_D| + |v_C - \nu_D| + |\sigma_C - \omega_D| \). Then it implies that

\[
\tilde{\phi}_2(C, D) = \frac{1}{3n} \sum_{i=1}^{n} |v_C - v_D| + |v_C - \nu_D| + |\sigma_C - \omega_D|
\]

\[
= \frac{\tilde{\phi}_1(C, D)}{n}.
\]

Hence (i) holds. Similarly,

\[
\tilde{\phi}_4(C, D) = \left( \frac{1}{3n} \sum_{i=1}^{n} ((v_C - v_D)^2 + (v_C - \nu_D)^2 + (\sigma_C - \omega_D)^2) \right)^{\frac{1}{2}}
\]

\[
= \frac{\tilde{\phi}_3(C, D)}{\sqrt{n}}
\]

and so (ii) holds. \( \square \)

**Proposition 3.2.** If \( C \) and \( D \) are IFSs in \( Y \), then the following hold:

(i) \( \tilde{\phi}^*(C, D) = \tilde{\phi}^*(D, C) \)
(ii) \( \tilde{\phi}^*(C, D) = \tilde{\phi}^*(\overline{C}, \overline{D}) \).
Proposition 3.3. Suppose C and D are IFSs in Y, then the following hold:
(i) $\tilde{\Phi}^{**}(C, D) = \tilde{\Phi}^{**}(D, C)$
(ii) $\tilde{\Phi}^{**}(C, D) = \tilde{\Phi}^{**}(\overline{C}, \overline{D})$.

Theorem 3.4. Suppose C, D and E are IFSs in Y, then the function $\tilde{\Phi}(C, D)$ satisfies
(i) $0 \leq \tilde{\Phi}^{*}(C, D) \leq 1$
(ii) $\tilde{\Phi}^{*}(C, D) = 0$ if and only if $C = D$
(iii) $\tilde{\Phi}^{*}(C, D) = \tilde{\Phi}^{*}(D, C)$.

Proof. The proof of (i) is straightforward. Recall that
$$
\tilde{\Phi}^{*}(C, D) = \left(\frac{1}{3} \sum_{i=1}^{n}[(v_C - v_D)^{\tau} + v_C - v_D] + |\sigma_C - \sigma_D|)^{\tau}\right)
$$

To proof (ii), suppose that $\tilde{\Phi}^{*}(C, D) = 0$. Then
$$
|v_C - v_D|^{\tau} = 0, \quad |v_C - v_D|^{\tau} = 0, \quad |\sigma_C - \sigma_D|^{\tau} = 0,
$$
and thus,
$$
v_C = v_D, \quad v_C = v_D \quad \text{and} \quad \sigma_C = \sigma_D.
$$

4. Experimental Examples

In this section, we apply both the Szmidt and Kacprzyk’s distances and their modified versions to problems of pattern recognition and medical diagnosis to determine which of the approaches are better in terms of performance indexes.

4.1. Case I

Pattern recognition is the process of identifying patterns by using machine learning method. The idea of pattern recognition is important because of its application potential in diverse areas. Assume there are three patterns $P_1$, $P_2$, $P_3$ denoted with IFPs in $Y = \{y_1, y_2, y_3\}$. If there is an unknown pattern $Q$ denoted with IFP in the same feature space $Y$. The intuitionistic fuzzy representations of these patterns are in Table 1.

<p>| Table 1: Intuitionistic fuzzy representations of patterns |
|---------------------------------|-----------|-----------|-----------|</p>
<table>
<thead>
<tr>
<th>Feature space</th>
<th>$y_1$</th>
<th>$y_2$</th>
<th>$y_3$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\nu_{P_1}$</td>
<td>0.1000</td>
<td>0.5000</td>
<td>0.1000</td>
</tr>
<tr>
<td>$\nu_{P_2}$</td>
<td>0.1000</td>
<td>0.1000</td>
<td>0.9000</td>
</tr>
<tr>
<td>$\nu_{P_3}$</td>
<td>0.8000</td>
<td>0.4000</td>
<td>0.0000</td>
</tr>
<tr>
<td>$\sigma_{P_1}$</td>
<td>0.1000</td>
<td>0.4000</td>
<td>0.5000</td>
</tr>
<tr>
<td>$\sigma_{P_2}$</td>
<td>0.5000</td>
<td>0.3000</td>
<td>0.8000</td>
</tr>
<tr>
<td>$\sigma_{P_3}$</td>
<td>0.7000</td>
<td>0.1000</td>
<td>0.2000</td>
</tr>
<tr>
<td>$\nu_{Q}$</td>
<td>0.2000</td>
<td>0.8000</td>
<td>0.4000</td>
</tr>
<tr>
<td>$\sigma_{Q}$</td>
<td>0.2000</td>
<td>0.6000</td>
<td>0.8000</td>
</tr>
</tbody>
</table>

Take SKD as Szmidt and Kacprzyk’s distance and MSK as Modified Szmidt and Kacprzyk’s distance.

The task is to determine which of the patterns can the sample $Q$ be associated with. Table 2 contains the distances using the Szmidt and Kacprzyk’s distances while Table 3 contains the distances based on the modified Szmidt and Kacprzyk’s distances.

From Tables 2 and 3, the sample $Q$ can be classified with pattern $P_2$ since the distances between them are the smallest. Tables 2 and 3 can be represented by the following figures to show the superiority of the modified Szmidt and Kacprzyk’s distances over the Szmidt and Kacprzyk’s distances.
4.2. Case II

Diagnosis of diseases is challenging due to embedded fuzziness in the processes. Here, we present a scenario of mathematical approach of diagnosing a patient medical status using the Szmidt and Kacprzyk’s distance and the modified Szmidt and

Kacprzyk’s distance, where the symptoms or clinical manifestations of the diseases are represented as IFPs using hypothetical case.

Suppose we have a set of diseases namely; viral fever (V), malaria (M), typhoid fever (T), Stomach ulcer (S) and chest...
problem (C) represented by IFPs, and a set of symptoms $S = \{s_1, s_2, s_3, s_4, s_5\}$ where $s_1 =$ temperature, $s_2 =$ headache, $s_3 =$ stomach pain, $s_4 =$ cough, $s_5 =$ chest pain. These symptoms are the clinical manifestations of the mentioned diseases. Assume a patient $P$ manifests symptoms as mentioned above, represented by IFPs. Table 4 contains intuitionistic fuzzy information of the diseases and patient $P$ with respect to the symptoms.

Now, we find which of the diseases has the smallest distance with the patient with respect to the symptoms by deploying the Szmidt and Kacprzyk’s distances and their modifications. Tables 5 and 6 contain the results.

From Tables 5 and 6, it is inferred that the patient is suffering from viral fever since the distance between the patient and viral fever is the smallest. Tables 5 and 6 can be represented by the following figures to show the supremacy of the modified Szmidt and Kacprzyk’s distances over the Szmidt and Kacprzyk’s distances.

From Figs. 1–8, it is observed that the modified Szmidt and Kacprzyk’s distances outperformed the Szmidt and Kacprzyk’s distances in terms of accuracy because while the modified Szmidt
and Kacprzyk’s distances take cognizance of the average of the differences among the three parameters of IFSs, the Szmidt and Kacprzyk’s distances do not consider the average of the differences. Since distance measure lies between 0 and 1, we conclude that $\phi_1$, $\phi_2$, $\hat{\phi}_1$ and $\hat{\phi}_2$ are not reliable distances.

5. Conclusion

In this paper, we have studied the Szmidt and Kacprzyk’s distances between IFSs and noticed a setback with the distance measures. Because of this setback, modifications of the Szmidt and Kacprzyk’s distances between IFSs were proposed to enhance accuracy of measure. It was verified mathematically that the modified Szmidt and Kacprzyk’s distances between IFSs satisfied the conditions for distance measure. Both the Szmidt and Kacprzyk’s distances and their modified versions were applied to determine pattern recognition and diagnostic medical reasoning where information were represented as IFPs. From the work, it is observed that the modified Szmidt and Kacprzyk’s distances outperformed the Szmidt and Kacprzyk’s distances.
in terms of accuracy because while the modified Szmidt and Kacprzyk’s distances take account of the number of the considered parameters, the Szmidt and Kacprzyk’s distances just added the hesitation margins to the methods introduced in [33]. In nutshell, the contributions of the paper includes; (i) modifications of the distance measures in [34] for better performance rating (ii) characterizations of the novel distance measures (iii) comparative analysis to show the edge of the novel approaches over the approaches in [34] (iv) applications of the distances in pattern recognition and disease diagnosis. The modified Szmidt and Kacprzyk’s distances could be extended to TOPSIS method as viable information measures to tackle multi-attributes decision-making problems. These novel distance measures of IFSs can be studied in other variants of fuzzy sets with minimal modifications.

Acknowledgements

We thank the referees for the positive enlightening comments and suggestions, which have greatly helped us in making improvements to this paper.

References


