Applied Mathematical Modelling 37 (2013) 5335–5342

Contents lists available at SciVerse ScienceDirect

Applied Mathematical Modelling

journal homepage: www.elsevier.com/locate/apm

A new method for similarity measures for pattern recognition

Peter C.P. Yen [a], KuoChin Fan [a], Henry C.J. Chao [b,*]

[a] Department of Computer Science and Information Engineering, National Central University, Taiwan
[b] Department of Traffic Science, Central Police University, Taiwan

ARTICLE INFO

Article history:
Received 25 February 2012
Received in revised form 2 October 2012
Accepted 18 October 2012
Available online 27 October 2012

Keywords:
Intuitionistic fuzzy sets
Similarity measures
Pattern recognition

ABSTRACT

This paper points out three questionable areas in the realm of similarity measures and then provides a new method that will rectify the problem. The purpose of this paper is fourfold. First, we will propose a scenario where the three similarity measures proposed by Hung and Yang (2004) [1] are helpless in aiding a decision maker in deciding pattern recognition problem. Second, we will present our method for solving the dilemma. Third, we will show that our proposed similarity measures satisfy the axioms for well defined similarity measures. Fourth, we will prove that our method could solve pattern recognition problems. Our findings will help researchers handle similarity problems under intuitionistic fuzzy sets environment.

1. Introduction

Atanassov [2] pioneered to the construction of intuitionistic fuzzy sets (IFSs) that are defined by three feature functions: the degree of membership, non-membership, and hesitation. IFSs are helpful in modeling vagueness, or uncertainty such that important applications of IFSs have been developed in many diverse areas, including medical diagnosis [3], pattern recognition [4,5], machine learning [6], decision making problems [7–9], microelectronic fault analysis [10,11], drug selection [12], and weight assessment [13,14]. Even through IFSs has already been expanded and applied to many fields; there are some disputations among researchers when it comes to the justification of similarity measures. In Hung and Yang [1], they mentioned that there are drawbacks in Li and Cheng [4] and so Liang and Shi [6] and Mitchell [5] provided new similarity measures to overcome these drawbacks. Here we will specifically point out the drawback of Li and Cheng [4], proposed in previous papers of Liang and Shi [6], Mitchell [5] and Hung and Yang [1]. They proposed a pattern recognition problem with two patterns, $\tilde{A}_1$ and $\tilde{A}_2$ with one sample, $\tilde{B}$ expressed in IFSs. They found that $S_{LC}(\tilde{A}_1, \tilde{B}) = S_{LC}(\tilde{A}_2, \tilde{B})$, where S_{LC} is the similarity measure proposed by Li and Cheng [4]. Owing to $S_{LC}(\tilde{A}_1, \tilde{B}) = S_{LC}(\tilde{A}_2, \tilde{B})$, researchers cannot decide which patterns, $\tilde{A}_1$ or $\tilde{A}_2$, sample $\tilde{B}$ belonging to. Then, Liang and Shi [6], Mitchell [5] and Hung and Yang [1] criticized the similarity measure of Li and Cheng [4], respectively. In Julian et al. [15], they showed that the improvement of Mitchell [5] contained questionable results and then provided some revisions. Based on Julian et al. [15], the proposed similarity measure should be rendered invalid if it cannot decide the pattern for samples in some cases. Following this trend, the first purpose of this paper is to provide a pattern recognition problem such that the similarity measure of Hung and Yang [1] cannot decide the pattern for which the sample belongs. The second purpose is to offer a method consisting of four similarity measures to overcome the unsolvable pattern recognition problem that we proposed.

* Corresponding author.
E-mail address: una051@mail.cpu.edu.tw (H.C.J. Chao).

0307-904X/$ - see front matter
http://dx.doi.org/10.1016/j.apm.2012.10.043

2. Review of the distance proposed by Hung and Yang [1]

Hung and Yang [1] used the Hausdorff distance to measure the distance between two IFSs, $\tilde{A} = \{\langle x_i, \mu_{\tilde{A}}(x_i), \upsilon_{\tilde{A}}(x_i)\rangle | x_i \in X\}$ and $\tilde{B} = \{\langle x_i, \mu_{\tilde{B}}(x_i), \upsilon_{\tilde{B}}(x_i)\rangle | x_i \in X\}$, where $X = \{x_1, \ldots, x_n\}$ is the universe of disclose with $\mu_{\tilde{A}}$ and $\upsilon_{\tilde{A}}$ as the membership and non-membership functions for IFS $\tilde{A}$, with the restriction $\mu_{\tilde{A}}$ and $\upsilon_{\tilde{A}}$ from X to $[0,1]$ and $0 \leqslant \mu_{\tilde{A}}(x_i) + \upsilon_{\tilde{A}}(x_i) \leqslant 1$, for every $x_i \in X$, as follows,

$$d_H(\tilde{A}, \tilde{B}) = \frac{1}{n}\sum_{i=1}^{n} H(I_{\tilde{A}}(x_i), I_{\tilde{B}}(x_i)) = \frac{1}{n}\sum_{i=1}^{n} \max\{|\mu_{\tilde{A}}(x_i) - \mu_{\tilde{B}}(x_i)|, |\upsilon_{\tilde{A}}(x_i) - \upsilon_{\tilde{B}}(x_i)|\}. \quad (1)$$

Hung and Yang [1] demonstrated their similarity measure by the following pattern recognition problem with the following two patterns,

$$\tilde{A}_1 = \{(x_1, 0.1, 0.1), (x_2, 0.5, 0.1), (x_3, 0.1, 0.9)\} \quad (2)$$

and

$$\tilde{A}_2 = \{(x_1, 0.5, 0.5), (x_2, 0.7, 0.3), (x_3, 0.0, 0.8)\}, \quad (3)$$

with a sample, $\tilde{B} = \{(x_1, 0.4, 0.4), (x_2, 0.6, 0.2), (x_3, 0.0, 0.8)\}$ to find that

$$S_d^1(\tilde{A}_1, \tilde{B}) = S_d^1(\tilde{A}_2, \tilde{B}) = 1, \quad (4)$$

where S_d^p is the similarity measure of Li and Cheng [4]

$$S_d^p(\tilde{A}, \tilde{B}) = 1 - \frac{1}{\sqrt[p]{n}} \sqrt[p]{\sum_{i=1}^{n} |m_{\tilde{A}}(i) - m_{\tilde{B}}(i)|^p}, \quad (5)$$

where $m_{\tilde{A}}(i) = \frac{1}{2}(\mu_{\tilde{A}}(x_i) + 1 - \upsilon_{\tilde{A}}(x_i))$ and $m_{\tilde{B}}(i) = \frac{1}{2}(\mu_{\tilde{B}}(x_i) + 1 - \upsilon_{\tilde{B}}(x_i))$ for two IFSs $\tilde{A}$ and $\tilde{B}$, with $p = 1$. Based on Eq. (4), Hung and Yang [1] criticized that Li and Cheng's similarity measure cannot be used to classify this sample.

To overcome the dilemma of Li and Cheng [4], based on their distance $d_H(\tilde{A}, \tilde{B})$ for two IFSs $\tilde{A}$ and $\tilde{B}$, Hung and Yang [1] provided three similarity measures as follows,

$$S_l(\tilde{A}, \tilde{B}) = 1 - d_H(\tilde{A}, \tilde{B}), \quad (6)$$

$$S_e(\tilde{A}, \tilde{B}) = \frac{e^{-d_H(\tilde{A}, \tilde{B})} - e^{-1}}{1 - e^{-1}} \quad (7)$$

and

$$S_c(\tilde{A}, \tilde{B}) = \frac{1 - d_H(\tilde{A}, \tilde{B})}{1 + d_H(\tilde{A}, \tilde{B})}. \quad (8)$$

3. A counter example for Hung and Yang's similarity measures

Based on the pattern recognition problem of Section 4, we proposed a new sample,

$$\tilde{C} = \{(x_1, 0.3, 0.3), (x_2, 0.6, 0.2), (x_3, 0.05, 0.85)\}, \quad (9)$$

such that we derived

$$S_l(\tilde{A}_1, \tilde{C}) = S_l(\tilde{A}_2, \tilde{C}) = \frac{53}{60} = 0.8833, \quad (10)$$

$$S_e(\tilde{A}_1, \tilde{C}) = S_e(\tilde{A}_2, \tilde{C}) = \frac{e^{\frac{-7}{60}} - e^{-1}}{1 - e^{-1}} = 0.8258 \quad (11)$$

and

$$S_c(\tilde{A}_1, \tilde{C}) = S_c(\tilde{A}_2, \tilde{C}) = \frac{53}{67} = 0.7910. \quad (12)$$

From Eqs. (10)–(12), researchers cannot decide whether sample $\tilde{C}$ belongs to pattern $\tilde{A}_1$ or $\tilde{A}_2$.

This means that the three similarity measures of Hung and Yang [1] sometimes cannot be used to solve pattern recognition. We will claim that the three similarity measures of Hung and Yang [1] all have the dilemma that they challenged Li and Cheng [4] with.