Neutrosophic Sets and Systems, Vol. 46, 2021
University of New Mexico

Medical Diagnosis via Distance-based Similarity Measure for Rough Neutrosophic Set

Suriana Alias[1]*, Daud Mohamad[2], Adibah Shuib[3,4], Nazhatul Sahima Mohd Yusoff[5], Norarida Abd Rhani[6] and Siti Nurul Fitriah Mohamad[7]

[1,5,6,7]Faculty of Computer and Mathematical Sciences, Universiti Teknologi MARA Cawangan Kelantan, Bukit Ilmu, 18500 Machang, Kelantan, Malaysia;
suria588@uitm.edu.my; nazha237@uitm.edu.my; norarida@uitm.edu.my; fitriah@uitm.edu.my

[2,3]Faculty of Computer and Mathematical Sciences, Kompleks Al-Khawarizmi, Universiti Teknologi MARA, 40450 Shah Alam, Selangor, Malaysia;
daud@tmsk.uitm.edu.my; adibah@tmsk.uitm.edu.my

[4]Malaysia Institute of Transport (MITRANS), Universiti Teknologi MARA (UiTM), Shah Alam, 40450 Shah Alam, Selangor, Malaysia.

*Correspondence: suria588@uitm.edu.my

Abstract: A rough neutrosophic set theory is a generalization of uncertainty set theory with a combination of upper and lower approximation and a pair of neutrosophic sets which are characterized by truth membership degree (T), indeterminacy membership degree (I), and falsity membership degree (F). This set theory is suitable for representing each criterion's relation in medical diagnoses, such as the relation of disease and symptom. This paper aims to propose a model of medical diagnosis via a distance-based similarity measure of a rough neutrosophic set. The first phase for the development model involves the roughness measure between information collected and a lower and upper approximation of rough neutrosophic set theory. Then, it is simultaneously used with extended Hausdorff distance measure to get the proper medical diagnosis. The result shows that each patient has a chest problem that contradicts the prior diagnosis. The finding shows that the roughness approximation is important to get the best result in a close distance-based similarity measure, especially for uncertainty information.

Keywords: Distance-based similarity measure, Medical diagnosis, Rough neutrosophic set, Roughness measure,

1. Introduction

The medical diagnosis contains lots of uncertainties and an increased volume of information. It becomes difficult to classify different symptoms under a single disease name. There is a possibility in some practical situations that each dimension has a different truth, indeterminacy, and falsity information. It is, therefore, important to use a more versatile method that can easily deal with unpredictable circumstances. Hence, a rough neutrosophic set (RNS) is a useful tool for dealing with uncertainty and incompleteness information for medical cases [1].

A rough neutrosophic set (RNS) is a generalization of rough set and neutrosophic set theory. Pawlak [2] introduced a rough set concept as a formal tool for modelling and processing incomplete information for information systems. The basic idea of the rough set is based upon the approximation of sets known as a lower approximation and an upper approximation of a set. Besides, the neutrosophic set proposed by Smarandache [3] is a generalization of a fuzzy set [4] and an intuitionistic fuzzy set [5]. Meanwhile, neutrosophic sets are characterized by truth membership function (T), indeterminacy membership function (I), and falsity membership function (F).

Since the rough neutrosophic set (RNS) involves a pair of approximation sets, then the roughness measure between them gives more chances for an informed decision. The study of this roughness measure is still not yet explored for RNS theory. Meanwhile, the study of distance-based similarity measures of RNS gives many measures, each representing specific properties and behavior in real-life decision making and pattern recognition works. Based on the relationship between distance and similarity measure, Pramanik et al. proposed several similarity measures: Cosine similarity measure [6] and Dice and Jaccard similarity measure [1]. Meanwhile, Pramanik et al. [7] used the Trigonometric Hamming similarity measure for multi decision-making in selecting laptops from a different company. Besides that, Mondal et al. [8] studied the similarity measure of RNS by introducing the variational coefficient for each similarity variable to solve the decision-making problem under-investment company option. Therefore, the application of RNS is widely explored, as discussed in the literature.

In this study, the roughness approximation of the rough set by Yao [10] is used to determine the roughness measure between the lower and upper approximations of RNS. Then, the extended Hausdorff distance measure is used in the first phase of implementation via medical diagnosis. Simultaneously, the roughness approximation is included in this dissimilarity measure. Therefore, this study aims to propose the new notion of roughness approximation for medical information via lower and upper approximations of RNS, and to determine the closeness of distance-based similarity measure between symptoms and diseases versus patients and symptoms for complete medical finding. The result is more accurate since the roughness of information is considered for the first term as a lower and upper approximation of RNS. For novelty, the roughness for a lower and upper approximation of RNS is not yet studied by other researchers. Following from there, medical information related to symptoms and diseases versus patients and symptoms is discussed thoroughly.

The rest of the paper is organized as follows. Section two is preliminaries for some important definition, while Section three introduces a new definition of the distance-based similarity measure. Section four presents the methodology involved in the medical diagnosis process, while Section five is the main implementation of medical findings. Lastly, Section six concludes the paper.

2. Preliminaries

This section recalled some important definitions of the rough neutrosophic set, extended Hausdorff distance of neutrosophic set, roughness approximation, and distance-based similarity measure. All the proof of the propositions may be referred to in [10 - 13].

2.1. Rough Neutrosophic Set

Definition 2.1.1 [11]**.** Let U be a non-null set and R be an equivalence relation on U. Let A be neutrosophic set in U with the truth membership function T_A, indeterminacy function I_A, and non-membership function F_A. The lower and the upper approximations of A in the approximation (U, R) denoted by $\underline{N}(A)$ and $\overline{N}(A)$ are respectively defined as follows:

$$\underline{N}(A) = \left(\langle x_j, T_{\underline{N}(A)}(x_j), I_{\underline{N}(A)}(x_j), F_{\underline{N}(A)}(x_j)\rangle \mid y \in [x_j]_R, j \in \mathbb{Z}^+, x_j \in U\right), \text{ and}$$
$$\overline{N}(A) = \left(\langle x_j, T_{\overline{N}(A)}(x_j), I_{\overline{N}(A)}(x_j), F_{\overline{N}(A)}(x_j)\rangle \mid y \in [x_j]_R, j \in \mathbb{Z}^+, x_j \in U\right)$$

where;

$j = 1, 2, \ldots, q$ is a positive integer, $T_{\underline{N}(A)}(x_j) = \wedge_{y \in [x_j]_R} T_A(y_j)$, $I_{\underline{N}(A)}(x_j) = \vee_{y \in [x_j]_R} I_A(y_j)$, $F_{\underline{N}(A)}(x_j) = \vee_{y \in [x_j]_R} F_A(y_j)$, $T_{\overline{N}(A)}(x_j) = \vee_{y \in [x_j]_R} T_A(y_j)$, $I_{\overline{N}(A)}(x_j) = \wedge_{y \in [x_j]_R} I_A(y_j)$, and $F_{\overline{N}(A)}(x_j) = \wedge_{y \in [x_j]_R} F_A(y_j)$.

Here $\wedge$ and $\vee$ denote "min" and "max" operators respectively and $[x_j]_R$ is the equivalence class of the x_j. The $T_A(y_j)$, $I_A(y_j)$ and $F_A(y_j)$ are the truth membership, indeterminacy membership, and falsity membership of y concerning A.

Since $\underline{N}(A)$ and $\overline{N}(A)$ are two neutrosophic sets in U, thus the neutrosophic set mappings $\underline{N}, \overline{N}: N(U) \to N(U)$ are respectively referred to as lower and upper rough neutrosophic set approximation operators, while the pair of $\left(\underline{N}(A), \overline{N}(A)\right)$ is called the rough neutrosophic set in (U, R), respectively. The rough neutrosophic set is denoted by:

$$N(A) = \left(\underline{N}(A), \overline{N}(A)\right) = \left(\langle x_j, \begin{pmatrix} [T_{\underline{N}(A)}(x_j), I_{\underline{N}(A)}(x_j), F_{\underline{N}(A)}(x_j)], \\ [T_{\overline{N}(A)}(x_j), I_{\overline{N}(A)}(x_j), F_{\overline{N}(A)}(x_j)] \end{pmatrix} \rangle \mid y \in [x_j]_R, j \in \mathbb{Z}^+, x_j \in U \right) \tag{1}$$

The truth membership set $[T_{\underline{N}(A)}(x_j), T_{\overline{N}(A)}(x_j)]$, indeterminacy membership set $[I_{\underline{N}(A)}(x_j), I_{\overline{N}(A)}(x_j)]$, and falsity membership $[F_{\underline{N}(A)}(x_j), F_{\overline{N}(A)}(x_j)]$ for lower and upper approximation of RNS may be in decreasing or increasing order.

Definition 2.1.2 [11]. If $N(A)$ is a rough neutrosophic set in (U, R), the rough complement of $N(A)$ is the rough neutrosophic set donated by $\sim N(A) = \left((\underline{N}(A))^c, (\overline{N}(A))^c\right)$, where $(\underline{N}(A))^c$ and $(\overline{N}(A))^c$ are the complements of neutrosophic set $\left(\underline{N}(A), \overline{N}(A)\right)$, respectively, given by

$$\sim N(A) = \left((\underline{N}(A))^c, (\overline{N}(A))^c\right) = \left\{ \langle x_j, \begin{pmatrix} [F_{\underline{N}(A)}(x_j), 1 - I_{\underline{N}(A)}(x_j), T_{\underline{N}(A)}(x_j)], \\ [F_{\overline{N}(A)}(x_j), 1 - I_{\overline{N}(A)}(x_j), T_{\overline{N}(A)}(x_j)] \end{pmatrix} \rangle \middle| \; x_j \in U \right\} \tag{2}$$

2.2. Distance-based Similarity Measure

Definition 2.2.1 [12]. An extended Hausdorff Distance $d_N^{EH}(A, B)$ operator between neutrosophic set A and B is defined as follows:

$$d_N^{EH}(A, B) = \frac{1}{n} \sum_{i=1}^{n} max\{|T_A(x_i) - T_B(x_i)|, |I_A(x_i) - I_A(x_i)|, |F_A(x_i) - F_B(x_i)|\} \tag{3}$$

Definition 2.2.2 [13]. It is well known that similarity measures can be generated from distance measures. Therefore, the distance-based similarity measure based on extended Hausdorff distance between neutrosophic set A and B is defined as follows:

$$S_N(A, B) = 1 - d_N^{EH}(A, B) \tag{4}$$

where $d_N^{EH}(A, B)$ represents the extended Hausdorff distance between neutrosophic set A and B.

Proposition 1. The similarity measure $S_N(A, B)$ for neutrosophic set A and B satisfies the following properties:

- (S1) $0 \le S_N(A, B) \le 1$;
- (S2) $S_N(A, B) = 1$ if and only if $A = B$;
- (S3) $S_N(A, B) = S_N(B, A)$;
- (S4) $S_N(A, C) \le S_N(A, B)$ and $S_N(A, C) \le S_N(B, C)$ if C is neutrosophic set in X and $A \subseteq B \subseteq C$.

All proofs for these properties were discussed in [13] and [12].

2.3. Accuracy and roughness approximation

Definition 2.3.1 [10]**.** For a subset of object $X \subseteq U$, the accuracy measure is defined as:

$$\alpha_E(X) = \frac{\left|\underline{apr}_E(X)\right| + \left|(\overline{apr}_E(X))^c\right|}{|U|} \tag{5}$$

where X is a non-empty set, $E \in X$, $\underline{apr}_E(X)$ is a lower approximation of set E, $\overline{apr}_E(X)$ is an upper approximation of set E, $|.|$ denotes the cardinality of a set E, and $0 \leq \alpha_E(X) \leq 1$. Based on the accuracy measure, the roughness measure is defined by:

$$\rho_E(X) = 1 - \alpha_E(X) \tag{6}$$

3. Distance-based similarity measure with roughness approximation

This section introduces a distance-based similarity measure with roughness approximation, where the roughness approximation as in Equations 5 and 6 is defined simultaneously with an extended Hausdorff distance measure. The determination of the roughness measure is defined between a lower and upper approximation of rough neutrosophic set theory instead of the average measurement between them.

3.1. Distance-based Similarity Measure

Assume that A and B be any two rough neutrosophic sets in the universe of discourse U as follows:

$$A = \langle x_j, \begin{matrix} [T_{\underline{N}(A)}(x_j), I_{\underline{N}(A)}(x_j), F_{\underline{N}(A)}(x_j)], \\ [T_{\overline{N}(A)}(x_j), I_{\overline{N}(A)}(x_j), F_{\overline{N}(A)}(x_j)] \end{matrix} | x_j \in U\rangle \text{ and}$$

$$B = \langle x_j, \begin{matrix} [T_{\underline{N}(B)}(x_j), I_{\underline{N}(B)}(x_j), F_{\underline{N}(B)}(x_j)], \\ [T_{\overline{N}(B)}(x_j), I_{\overline{N}(B)}(x_j), F_{\overline{N}(B)}(x_j)] \end{matrix} | x_j \in U\rangle$$

Then, the distance-based similarity measure for RNS A and B is defined as:

Definition 3.1.1: Extended Hausdorff distance with roughness operator is given by

$$d_{RNS}^{EH}(A,B) = \frac{1}{k}\sum_{j=1}^{k} max \begin{Bmatrix} |\Delta T_{N(A)}(x_j) - \Delta T_{N(B)}(x_j)|, |\Delta I_{N(A)}(x_j) - \Delta I_{N(B)}(x_j)|, \\ |\Delta F_{N(A)}(x_j) - \Delta F_{N(B)}(x_j)| \end{Bmatrix} \tag{7}$$

where;

$$\Delta T_{N(A)}(x_j) = 1 - \left(\frac{T_{\underline{N}(A)}(x_j) + \left(T_{\overline{N}(A)}(x_j)\right)^c}{|X|}\right), \ \Delta T_{N(B)}(x_j) = 1 - \left(\frac{T_{\underline{N}(B)}(x_j) + \left(T_{\overline{N}(B)}(x_j)\right)^c}{|X|}\right),$$

$$\Delta I_{N(A)}(x_j) = 1 - \left(\frac{I_{\underline{N}(A)}(x_j) + \left(I_{\overline{N}(A)}(x_j)\right)^c}{|X|}\right), \ \Delta I_{N(B)}(x_j) = 1 - \left(\frac{I_{\underline{N}(B)}(x_j) + \left(I_{\overline{N}(B)}(x_j)\right)^c}{|X|}\right),$$

$$\Delta F_{N(A)}(x_j) = 1 - \left(\frac{F_{\underline{N}(A)}(x_j) + \left(F_{\overline{N}(A)}(x_j)\right)^c}{|X|}\right), \text{ and } \Delta F_{N(B)}(x_j) = 1 - \left(\frac{F_{\underline{N}(B)}(x_j) + \left(F_{\overline{N}(B)}(x_j)\right)^c}{|X|}\right).$$

Here, Δ denotes the "roughness approximation" operator by rough approximation between the lower and upper approximation of RNS, while $|.|$ is the cardinality of the universal X.

Proposition 2. The extended Hausdorff distance $d_{RNS}^{EH}(A,B)$ between rough neutrosophic A and B satisfies the following properties:

(D1) $d_{RNS}^{EH}(A,B) \geq 0$. (non-negative)

(D2) $d_{RNS}^{EH}(A,B) = 0$ if and only if $A = B$, for all $A, B \in RNS$. (definiteness)

(D3) $d_{RNS}^{EH}(A,B) = d_{RNS}^{EH}(B,A)$. (symmetry)

(D4) If $A \subseteq B \subseteq C$, for $A, B, C \in RNS$, then $d_{RNS}^{EH}(A,C) \geq d_{RNS}^{EH}(A,B)$ and $d_{RNS}^{EH}(A,C) \geq d_{RNS}^{EH}(B,C)$. (triangle inequality)

Proof:

(D1) $d_{RNS}^{EH}(A,B) \geq 0$.

As $\Delta T_{N(A)}(x_j), \Delta I_{N(A)}(x_j), \Delta F_{N(A)}(x_j) \in [0,1]$, $\Delta T_{N(B)}(x_j), \Delta I_{N(B)}(x_j), \Delta F_{N(B)}(x_j) \in [0,1]$ for all $A, B \in RNS$, the distance measurement based on these functions also lies between $[0,1]$.

(D2) $d_{RNS}^{EH}(A,B) = 0$ if and only if $A = B$, for all $A, B \in RNS$.

For any two RNS A and B, if $A = B$, then the following relations hold for any $\Delta T_{N(A)}(x_j) = \Delta T_{N(B)}(x_j)$, $\Delta I_{N(A)}(x_j) = \Delta I_{N(B)}(x_j)$, $\Delta F_{N(A)}(x_j) = \Delta F_{N(B)}(x_j)$ which states that $|\Delta T_{N(A)}(x_j) - \Delta T_{N(B)}(x_j)| = 0$, $|\Delta I_{N(A)}(x_j) - \Delta I_{N(B)}(x_j)| = 0$, and $|\Delta F_{N(A)}(x_j) - \Delta F_{N(B)}(x_j)| = 0$. Thus, $d_{RNS}^{EH}(A,B) = 0$. Conversely, if $d_{RNS}^{EH}(A,B) = 0$, then the zero distance measure is possible only if $|\Delta T_{N(A)}(x_j) - \Delta T_{N(B)}(x_j)| = 0$, $|\Delta I_{N(A)}(x_j) - \Delta I_{N(B)}(x_j)| = 0$, and $|\Delta F_{N(A)}(x_j) - \Delta F_{N(B)}(x_j)| = 0$. This resulted from $\Delta T_{N(A)}(x_j) = \Delta T_{N(B)}(x_j)$, $\Delta I_{N(A)}(x_j) = \Delta I_{N(B)}(x_j)$, and $\Delta F_{N(A)}(x_j) = \Delta F_{N(B)}(x_j)$ for all i, j values. Hence $A = B$.

(D3) $d_{RNS}^{EH}(A,B) = d_{RNS}^{EH}(B,A)$. The proof is obvious.

(D4) If $A \subseteq B \subseteq C$, for $A, B, C \in RNS$, then $d_{RNS}^{EH}(A,C) \geq d_{RNS}^{EH}(A,B)$ and $d_{RNS}^{EH}(A,C) \geq d_{RNS}^{EH}(B,C)$.

Let $A \subseteq B \subseteq C$, which implies that; $\Delta T_{N(A)}(x_j) \leq \Delta T_{N(B)}(x_j) \leq \Delta T_{N(C)}(x_j)$, $\Delta I_{N(A)}(x_j) \geq \Delta I_{N(B)}(x_j) \geq \Delta I_{N(C)}(x_j)$, $\Delta F_{N(A)}(x_j) \geq \Delta F_{N(B)}(x_j) \geq \Delta F_{N(C)}(x_j)$ for every $x_j \in X$;

Then, we obtain the following relation:

a) $|\Delta T_{N(A)}(x_j) - \Delta T_{N(B)}(x_j)| \leq |\Delta T_{N(A)}(x_j) - \Delta T_{N(C)}(x_j)|$,
$|\Delta T_{N(B)}(x_j) - \Delta T_{N(C)}(x_j)| \leq |\Delta T_{N(A)}(x_j) - \Delta T_{N(C)}(x_j)|$,

b) $|\Delta I_{N(A)}(x_j) - \Delta I_{N(B)}(x_j)| \leq |\Delta I_{N(A)}(x_j) - \Delta I_{N(C)}(x_j)|$,
$|\Delta I_{N(B)}(x_j) - \Delta I_{N(C)}(x_j)| \leq |\Delta I_{N(A)}(x_j) - \Delta I_{N(C)}(x_j)|$,

c) $|\Delta F_{N(A)}(x_j) - \Delta F_{N(B)}(x_j)| \leq |\Delta F_{N(A)}(x_j) - \Delta F_{N(C)}(x_j)|$,
$|\Delta F_{N(B)}(x_j) - \Delta F_{N(C)}(x_j)| \leq |\Delta F_{N(A)}(x_j) - \Delta F_{N(C)}(x_j)|$,

Combining a), b), and (c), we obtain

$\frac{1}{k}\sum_{j=1}^{k} max\{|\Delta T_{N(A)}(x_j) - \Delta T_{N(B)}(x_j)|, |\Delta I_{N(A)}(x_j) - \Delta I_{N(B)}(x_j)|, |\Delta F_{N(A)}(x_j) - \Delta F_{N(B)}(x_j)|\} \leq$
$\frac{1}{k}\sum_{j=1}^{k} max\{|\Delta T_{N(A)}(x_j) - \Delta T_{N(C)}(x_j)|, |I_{N(A)}(x_j) - \Delta I_{N(C)}(x_j)|, |\Delta F_{N(A)}(x_j) - \Delta F_{N(C)}(x_j)|\}$ and
$\frac{1}{k}\sum_{j=1}^{k} max\{|\Delta T_{N(B)}(x_j) - \Delta T_{N(C)}(x_j)|, |\Delta I_{N(B)}(x_j) - \Delta I_{N(C)}(x_j)|, |\Delta F_{N(B)}(x_j) - \Delta F_{N(C)}(x_j)|\} \leq$
$\frac{1}{k}\sum_{j=1}^{k} max\{|\Delta T_{N(A)}(x_j) - \Delta T_{N(C)}(x_j)|, |\Delta I_{N(A)}(x_j) - \Delta I_{N(C)}(x_j)|, |\Delta F_{N(A)}(x_j) - \Delta F_{N(C)}(x_j)|\}$. This implies that $d_{RNS}^{EH}(A,B) \leq d_{RNS}^{EH}(A,C)$ and $d_{RNS}^{EH}(B,C) \leq d_{RNS}^{EH}(A,C)$. Thus, the property (D4) is satisfied.

This completes the proof. ∎

4. Methodology

In this study, there are four phases to complete the medical diagnosis findings via distance-based similarity measure of a rough neutrosophic set (RNS).

Phase 1: Collection of data involving the information regarding the symptoms and diseases versus patients and symptoms from the medical report.

In this phase, the data collected on the relationship between symptoms and diseases as well as patients and symptoms are collected from the medical personnel.

Phase 2: Construct the RNS-set for the medical report.

The data collection is converted to RNS-set by using Definitions 2.1.1, as in Equation (1).

Phase 3: The determination of roughness approximation simultaneously with the distance-based similarity measure of RNS-set for medical findings.

RNS-set is used to determine the distance-based similarity measure of the relationship between symptoms and diseases as well as patients and symptoms using Definition 3.1.1 and Equations (2) and (7).

Phase 4: Discussion of a complete medical report.

Lastly, the complete medical report can be written to determine which patient's symptoms and diseases are related. If the distance measure is closer to zero, the conclusion is that the patient possibly suffers from the disease. Meanwhile, for similarity measure, if the measurement is greater than 0.5, then the conclusion is that the patient possibly suffers from the disease. On the other hand, if the similarity measure is less than 0.5, then the conclusion is that the patient may not possibly suffer from the disease.

5. Case Study: Implementation in Medical Diagnosis

In this section, the relationship between symptoms and diseases as well as patients and symptoms are considered in the same equivalence relation. Table 1 shows an example of the medical findings of patients represented in a tabular form. For diagnosis purpose, the patient is kept under supervision for a one-time inspection.

Table 1. Example of a medical finding of a patient

Temperature	Headache	Stomach pain	Cough	Chest pain
High	Yes (moderate)	Yes (moderate)	Yes (high)	Yes (high)

The main feature of this study is to consider the degree of truth membership, indeterminacy membership, and falsity membership for each element between two approximations. The data is adapted from Pramanik and Mondal [6]. Let $P = \{ p_1, p_2, p_3\}$ be a set of patients, $D = \{ d_1, d_2, d_3, d_4\}$ be a set of diseases, and $S = \{x_1, x_2, x_3, x_4, x_5\}$ be a set of symptoms. The relation between patients and symptoms (see Table 2) and the relation between symptoms and diseases (see Table 3) are considered in the same equivalence relation.

Table 2. The relation between patients and symptoms

Relation, A	Temperature (x_1)	Headache (x_2)	Stomach pain (x_3)	Cough (x_4)	Chest pain (x_5)
Patient (p_1)	$\langle (0.6, 0.4, 0.3), (0.8, 0.2, 0.1) \rangle$	$\langle (0.4, 0.4, 0.4), (0.6, 0.2, 0.2) \rangle$	$\langle (0.5, 0.3, 0.2), (0.7, 0.1, 0.2) \rangle$	$\langle (0.6, 0.2, 0.4), (0.8, 0.0, 0.2) \rangle$	$\langle (0.4, 0.4, 0.4), (0.6, 0.2, 0.2) \rangle$
Patient (p_2)	$\langle (0.5, 0.3, 0.4), (0.7, 0.3, 0.2) \rangle$	$\langle (0.5, 0.5, 0.3), (0.7, 0.3, 0.3) \rangle$	$\langle (0.5, 0.3, 0.4), (0.7, 0.1, 0.4) \rangle$	$\langle (0.5, 0.3, 0.3), (0.9, 0.1, 0.3) \rangle$	$\langle (0.5, 0.3, 0.3), (0.7, 0.1, 0.3) \rangle$
Patient (p_3)	$\langle (0.6, 0.4, 0.4), (0.8, 0.2, 0.2) \rangle$	$\langle (0.5, 0.2, 0.3), (0.7, 0.0, 0.1) \rangle$	$\langle (0.4, 0.3, 0.4), (0.8, 0.1, 0.2) \rangle$	$\langle (0.6, 0.1, 0.4), (0.8, 0.1, 0.2) \rangle$	$\langle (0.5, 0.3, 0.3), (0.7, 0.1, 0.1) \rangle$

Table 3. The relation between symptoms and diseases

Relation, B	Temperature (x_1)	Headache (x_2)	Stomach pain (x_3)	Cough (x_4)	Chest pain (x_5)
Viral fever (d_1)	$\langle(0.6, 0.5, 0.4), (0.8, 0.3, 0.2)\rangle$	$\langle(0.5, 0.3, 0.4), (0.7, 0.3, 0.2)\rangle$	$\langle(0.2, 0.3, 0.4), (0.4, 0.3, 0.2)\rangle$	$\langle(0.4, 0.3, 0.3), (0.6, 0.1, 0.1)\rangle$	$\langle(0.2, 0.4, 0.4), (0.4, 0.2, 0.2)\rangle$
Malaria (d_2)	$\langle(0.1, 0.4, 0.4), (0.5, 0.2, 0.2)\rangle$	$\langle(0.2, 0.3, 0.4), (0.6, 0.3, 0.2)\rangle$	$\langle(0.1, 0.4, 0.4), (0.3, 0.2, 0.2)\rangle$	$\langle(0.3, 0.3, 0.3), (0.5, 0.1, 0.3)\rangle$	$\langle(0.1, 0.3, 0.3), (0.3, 0.1, 0.1)\rangle$
Stomach problem (d_3)	$\langle(0.3, 0.4, 0.4), (0.5, 0.2, 0.2)\rangle$	$\langle(0.2, 0.3, 0.3), (0.4, 0.1, 0.1)\rangle$	$\langle(0.4, 0.3, 0.4), (0.6, 0.1, 0.2)\rangle$	$\langle(0.1, 0.6, 0.6), (0.3, 0.4, 0.4)\rangle$	$\langle(0.1, 0.4, 0.4), (0.3, 0.2, 0.2)\rangle$
Chest problem (d_4)	$\langle(0.2, 0.4, 0.6), (0.4, 0.4, 0.4)\rangle$	$\langle(0.1, 0.5, 0.5), (0.5, 0.3, 0.3)\rangle$	$\langle(0.1, 0.4, 0.6), (0.3, 0.2, 0.4)\rangle$	$\langle(0.5, 0.3, 0.4), (0.7, 0.3, 0.2)\rangle$	$\langle(0.4, 0.4, 0.4), (0.6, 0.2, 0.2)\rangle$

Based on Pawlak [2], the lower approximation explains that the element set surely belongs to the object, while the upper approximation possibly belongs to the object. For example, based on the data collected in Table 2, the truth membership degree for temperature (x_1) that surely belongs to patient 1 (p_1) is equal to 0.6 and which possibly belongs to patient 1 (p_1) is equal to 0.8. The indeterminacy membership degree for temperature (x_1) that surely belongs to patient 1 (p_1) is equal to 0.4 and which possibly belongs to patient 1 (p_1) is equal to 0.2. Meanwhile, the falsity membership degree for temperature (x_1) which surely belongs to patient 1 (p_1) is equal to 0.3 and which possibly belongs to patient 1 (p_1) is equal to 0.2. The same description is indicated for each data in Table 2 and Table 3.

Next, the determination of roughness approximation simultaneously with the distance-based similarity measurement by extended Hausdorff distance is used to determine the proper medical diagnosis for model RNS for each patient. By using an Equation (2) and roughness operator in Definition 3.1.1, the truth roughness measure for relation A for patient (p_1) is calculated as follows:

$$\Delta T_{N(A)}(x_1) = 1 - \left(\frac{T_{\underline{N}(A)}(x_1) + \left(T_{\overline{N}(A)}(x_1)\right)^C}{|X|}\right) = 1 - \left(\frac{0.6+0.1}{|5|}\right) = 0.86.$$

Then, by using the same equation and definition, the roughness measure for all membership function for each relation A and relation B for patient (p_1), is presented as follows:

$\Delta T_{N(A)}(x_2) = 0.88$, $\Delta T_{N(A)}(x_3) = 0.86, \Delta T_{N(A)}(x_4) = 0.84$, and $\Delta T_{N(A)}(x_5) = 0.88$;
$\Delta I_{N(A)}(x_1) = 0.76, \Delta I_{N(A)}(x_2) = 0.76$, $\Delta I_{N(A)}(x_3) = 0.76, \Delta I_{N(A)}(x_4) = 0.76$, and $\Delta I_{N(A)}(x_5) = 0.76$;
$\Delta F_{N(A)}(x_1) = 0.78, \Delta F_{N(A)}(x_2) = 0.8$, $\Delta F_{N(A)}(x_3) = 0.82, \Delta F_{N(A)}(x_4) = 0.76$, and $\Delta F_{N(A)}(x_5) = 0.8$;
$\Delta T_{N(B)}(x_1) = 0.84, \Delta T_{N(B)}(x_2) = 0.86$, $\Delta T_{N(B)}(x_3) = 0.92, \Delta T_{N(B)}(x_4) = 0.9$, and
$\Delta T_{N(B)}(x_5) = 0.92$; $\Delta I_{N(B)}(x_1) = 0.76, \Delta I_{N(B)}(x_2) = 0.8$, $\Delta I_{N(B)}(x_3) = 0.8, \Delta I_{N(B)}(x_4) = 0.76$, and $\Delta I_{N(B)}(x_5) = 0.76$; $\Delta F_{N(B)}(x_1) = 0.76, \Delta F_{N(B)}(x_2) = 0.8$, $\Delta F_{N(B)}(x_3) = 0.84, \Delta F_{N(B)}(x_4) = 0.82$, and $\Delta F_{N(B)}(x_5) = 0.84$;

Then, simultaneously using Equation (7), the extended Hausdorff distance for medical diagnosis of patient 1 (p_1) with viral fever (d_1) symptom is calculated as follows:

$$d_{RNS}^{EH}(A,B) = \frac{1}{5}\sum_{j=1}^{5} max \begin{Bmatrix} \left|\Delta T_{N(A)}(x_j) - \Delta T_{N(B)}(x_j)\right|, \left|\Delta I_{N(A)}(x_j) - \Delta I_{N(B)}(x_j)\right|, \\ \left|\Delta F_{N(A)}(x_j) - \Delta F_{N(B)}(x_j)\right| \end{Bmatrix}$$

$= \frac{1}{5}(max\{|0.86 - 0.84|, |0.76 - 0.76|, |0.78 - 0.76|\} + max\{|0.88 - 0.86|, |0.76 - 0.8|, |0.8 - 0.8|\} + max\{|0.86 - 0.92|, |0.76 - 0.8|, |0.82 - 0.84|\} + max\{|0.84 - 0.9|, |0.76 - 0.76|, |0.76 - 0.82|\} + max\{|0.88 - 0.92|, |0.76 - 0.76|, |0.8 - 0.84|\}) = \frac{1}{5}(0.02 + 0.04 + 0.06 + 0.06 + 0.04) = \frac{1}{5}(0.22) = 0.044.$

Therefore, the extended Hausdorff distance for patient 1 (p_1) with viral fever (d_1) symptoms are 0.044. Then, a similar calculation will be repeated to obtain the result of medial finding for each patient by employing extended Hausdorff distance. The summary result for the proposed extended Hausdorff distance measure with roughness approximation is represented in Table 4.

Table 4. The proposed extended Hausdorff distance measure with roughness approximation

Proposed extended Hausdorff distance	**Viral fever (d_1)**	**Malaria (d_2)**	**Stomach problem (d_3)**	**Chest problem (d_4)**
Patient (p_1)	0.0440	0.0720	0.0560	**0.0280**
Patient (p_2)	0.0640	0.0960	0.0620	**0.0400**
Patient (p_3)	0.0440	0.0800	0.0600	**0.0360**

According to the result, all the proposed distance measure is close to zero. Here, the closest value to zero indicates the result is possibly "more suffering". Therefore, it shows that all patients are suffering from a chest problem.

By comparing the similarity measure as in Equation (4) with the previous result by Pramanik and Mondal [6] shown in Table 5, we can see that previously all patients were diagnosed with a viral fever. Therefore, a different diagnose result is determined for this study. However, all the similarity measure values are greater than 0.5, indicating that the patients possibly suffer from the disease. The closest similarity value to one indicates the highest possibility of diseases.

Table 5. The Cosine similarity measure and proposed distance-based similarity measure

Cosine similarity measure	**Viral fever (d_1)**	**Malaria (d_2)**	**Stomach problem (d_3)**	**Chest problem (d_4)**
Patient (p_1)	**0.9595**	0.9114	0.8498	0.8743
Patient (p_2)	**0.9624**	0.9320	0.8935	0.8307
Patient (p_3)	**0.9405**	0.8873	0.8487	0.8372
A proposed distance-based similarity measure				
Patient (p_1)	0.9560	0.9280	0.9440	**0.9720**
Patient (p_2)	0.9360	0.9040	0.9380	**0.9600**
Patient (p_3)	0.9560	0.9200	0.9400	**0.9640**

However, the proposed distance-based similarity result is more accurate since the roughness between the lower and upper approximations of RNS is considered simultaneously with the extended Hausdorff distance instead of only the mean operator between the lower and upper approximation of RNS. Even the other similarity measures led to the same final answer but extended Hausdorff distance shows the simplest and easiest way. Therefore, the chances to obtain the wrong answer are less than other similarity measures.

6. Conclusions

The complete medical diagnosis covered all the relation between the collection of medical information, such as the relationship between patients and symptoms as well as symptoms and diseases. This study successfully examines all the factors needed to complete the medical diagnosis where the distance-based similarity between the medical information is taken over for the first phase. The new notion of roughness approximation for medical information via lower and upper approximations of a rough neutrosophic set is successfully presented. In future work, it is valid to use the same method that involved data with upper and lower approximations. Besides that, distance-based similarity measures by extended Hausdorff distance can be applied in other fields such as the distance for a spatial object in Geographical Information Science (GIS), object recognition for multimedia application, and others.

Funding: "This research received no external funding"

Acknowledgments: Thank you to Universiti Teknologi MARA, Kelantan, for providing financial support. Thank you to our group members for sharing information and giving cooperation to complete this study. We are also grateful to the anonymous reviewer for their insightful comments and suggestions in improving the paper.

Conflicts of Interest: "The authors declare no conflict of interest."

References

1. Pramanik, S.; Mondal, K. Some rough neutrosophic similarity measures and their application to multi-attribute decision making, *Global Jo*, **2015**, *2(7)*, 61–74.
2. Pawlak, Z. Rough sets, *International Journal of Computer & Information Sciences*, **1982**, *11(5)*, 341–356.
3. Smarandache, F. A Unifying Field in Logics: Neutrosophic Logic, Neutrosophy, Neutrosophic Set, Neutrosophic Probability, In *American Research Press*, **1998**.
4. Zadeh, L. A. Fuzzy sets, *Information and Control,* **1965**, *8(3),* 338-353.
5. Atanassov, K. T. Intuitionistic fuzzy set, *Fuzzy Sets and Systems*, **1983**, *20*, 87–96.
6. Pramanik, S.; Mondal, K. Cosine similarity measure of rough neutrosophic sets and its application in medical diagnosis, *Global Journal of Advanced Research*, **2015**, *2(1)*, 212–220.
7. Pramanik, S.; Roy, R.; Roy, T. K.; Smarandache, F. Multi-attribute decision making based on several trigonometric hamming similarity measures under interval rough neutrosophic environment, *Neutrosophic Sets and Systems Neutrosophic*, **2018**, *19*, 110–118.
8. Mondal, K.; Pramanik, S.; Smarandache, F. Multi-attribute decision making based on rough neutrosophic variational coefficient similarity measure, *Neutrosophic Sets and Systems*, **2016**, *13*, 3–17.
9. Majumdar, P.; Samanta, S. K. On similarity and entropy of neutrosophic sets, *Journal of Intelligent and Fuzzy Systems*, **2013**, *I*, 1–13.
10. Yao, Y.-Y. Notes on rough set approximations and associated measures, *J. Zhejiang Ocean Univ. (Natur. Sci.),* **2010**, *29(5)*, 399–410.
11. Broumi, S.; Smarandache, F.; Dhar, M. Rough neutrosophic sets, *Italian Journal of Pure and Applied Mathematics*, **2014**, *32*, 493–502.
12. Broumi, S.; Smarandache, F. Several similarity measures of neutrosophic sets, *Neutrosophic Sets and Systems*, **2013**, *1(1)*, 54–62.
13. Broumi, S.; Smarandache, F. Extended Hausdorff distance and similarity measures for neutrosophic refined sets and their application in medical diagnosis, *Journal of New Theory*, **2015**, *7*, 64–78.

Received: May 12, 2021. Accepted: October 6, 2021