Personalized Medical Diagnosis Recommendation Based on Neutrosophic Sets and Spectral Clustering

Mengru Dong[1], Shunmei Meng[1], Lixia Chen[2], and Jing Zhang[1(✉)]

[1] School of Computer Science and Engineering, Nanjing University of Science and Technology, Nanjing 210094, China
1850193966@qq.com, jzhang@njust.edu.cn

[2] Jiangsu Second Chinese Medicine Hospital, Nanjing 210017, China

Abstract. With the development of cloud-based services and artificial intelligence technologies, the personalized diagnosis recommender system has been a hot research topic in medical services. An effective diagnosis recommendation model could help doctors and patients make more accurate predictions in clinical diagnosis. In this paper, we propose a novel personalized diagnosis recommendation method based on neutrosophic sets, spectral clustering, and web-based medical information to offer satisfied web-based medical service. Firstly, the neutrosophic set theory is adopted to formulate the patients' personal information and the symptom features into more interpretable neutrosophic sets with uniformly normalized values. Moreover, to make more accurate predictions, the spectral clustering scheme is integrated into a neutrosophic-based prediction approach to mining the similarity relationships between the undiagnosed diseases and the history disease records. Finally, a deneutrosophication operation is applied to recommend the final fine-grain diagnoses with interpretable clinic meanings. Experimental results on four real-world medical diagnosis datasets validate the effectiveness of the proposed method.

Keywords: Personalized diagnosis recommendation · Neutrosophic sets · Healthcare service · Spectral clustering

1 Introduction

Personalized recommender systems for medical diagnosis have become a hot research topic for healthcare service in the development of modern medical technology. It has been emerged as a valuable tool in healthcare service to assist doctors in dealing with overloaded web-based medical information and making more accurate predictions on medical diagnosis, which has received full attention from both industry and academia. As we know, disease diagnosis is still full of challenges for medical professionals because not only patient's symptoms are inherently uncertain, but also the relations between patients and symptoms as well as symptoms and diseases are vague and uncertain. For

L. Qi et al. (Eds.): CloudComp 2020, LNICST 363, pp. 160–174, 2021.
https://doi.org/10.1007/978-3-030-69992-5_13

example, a small number of symptoms and laboratory indicators are not sufficient to diagnose a disease, but waiting for the occurrence of a large number of symptoms or obtaining more laboratory indicators will place a heavy burden on patients. Therefore, for young doctors who lack clinical experience, auxiliary diagnostic systems can help them to use the experience in historical clinic cases to make a more accurate diagnosis.

For auxiliary diagnosis systems, the first issue that needs to be handled is the uncertainty of the data, which often has important clinical implications in the medical field. For example, an abnormal body temperature of 38.5 °C is not a high temperature for some diseases, and it does not require special attention, but for some other diseases, it may be high temperature and needs to be treated with caution. There have been some research focusing on dealing with uncertain information in medical diagnosis. Some studies used fuzzy set methods such as Intuitionistic fuzzy Soft Sets (IFSSs) and their extensions [4, 5] to deal with the uncertain information of symptoms. Neutrosophic Set, proposed by Smarandache [6], was utilized in the medical diagnosis field [7, 8], which is a three-dimensional set while the fuzzy set has only one dimension. Thus, neutrosophic sets can convert the features of patients into a uniform scale as well as provide more meaningful semantics. However, these theories still have some shortcomings in the deneutrosophication process [19], similarity measure [9, 10], and distance measure [13, 22].

In recent years, many diagnosis recommendation models have been proposed for medical services, such as hybrid recommender systems with picture fuzzy clustering and intuitionistic fuzzy sets [14], intuitionistic fuzzy recommender systems [19], the hybrid recommendation system for heart disease diagnosis based on multiple kernel learning [1]. The principle of these systems generally combines fuzzy sets or neutrosophic sets with traditional recommendation methods. Both fuzzy sets and neutrosophic sets follow a similar way to handle the uncertainty of objects' attributes.

To broaden the usage of neutrosophic sets in medical diagnosis, many techniques in machine learning and recommender systems [16, 26] (e.g., clustering [17]) are integrated with neutrosophic set theory to provide more meaningful and accurate predictions on diagnosis recommendations. For example, Ali and Son [3] proposed a hybrid method that combines neutrosophic sets and recommender systems for medical diagnosis. Although their method has a strong mathematical basis, it still has some defects in the similarity calculation for patients and symptoms, which is the key technical points in most recommender systems.

In this paper, we propose a novel personalized diagnosis recommendation method for medical service based on neutrosophic sets, spectral clustering, and web-based medical information to offer satisfied web-based medical service. Specifically, the neutrosophic set theory is adopted to formulate the patients' personal information and the symptom features into more interpretable neutrosophic sets with uniformly normalized values. Moreover, the spectral clustering scheme is integrated into the neutrosophic-based prediction method to enhance the similarity measure in the recommendation method. Experimental results validate the effectiveness of the proposed method.

The remainder of this paper is organized as follows: Sect. 2 reviews some related work. Then, a hybrid model that incorporates the neutrosophic recommendation method for medical diagnosis and spectral clustering is proposed in Sect. 3. Section 4 presents

an empirical evaluation of the performance of the proposed method. Finally, Sect. 5 concludes the paper and provides an outlook on the continuations of our work.

2 Related Work

In recent years, a large number of web-based personalized healthcare applications have emerged dramatically [28], among which different types of recommender systems for medical diagnoses have also been proposed, such as neutrosophic recommender systems [3], hybrid recommender systems with picture fuzzy clustering and intuitionistic fuzzy sets [14], intuitionistic fuzzy recommender systems [19], and the hybrid recommendation system for heart disease diagnosis based on multiple kernel learning [1]. The principle of these systems generally combines fuzzy sets or neutrosophic sets with traditional recommendation methods. Both fuzzy sets and neutrosophic sets follow a similar way to handle the uncertainty of attributes. They extend an attribute to multidimensional ones accompanied by membership values. Moreover, the core of a recommender system is the similarity measure of items, which also provide users with information about predictive rating or preference. For example, Ye et al. [11] proposed a method that uses distance-based similarity measures of single-valued neutrosophic multisets for medical diagnosis. Davis et al. [26] proposed a recommendation engine that combines collaborative filtering with clustering to predict patients' most possible disease according to patients' symptoms and historical medical materials. Since neutrosophic sets have been widely used in medical diagnosis [7–10, 13, 21], our proposed recommendation method is also built on the neutrosophic sets.

Clustering has been widely used in recommender systems [18, 25] for computing similarity measures. Compared with some traditional clustering methods like K-means and K-medoids, spectral clustering [20], which is based on graph partition theory, has strong adaptability to the variety of data distributions and can generate more reasonable clusters. Therefore, many recommendation methods tend to employ spectral clustering for better performance. Li et al. [2] proposed a recommendation method that uses spectral clustering to group users and items in the original rating matrix. Xu et al. [12] proposed a spectral clustering based on intuitionistic fuzzy sets. In our study, spectral clustering is used for the calculation of the similarity between patients' demographic and symptom features.

3 The Proposed Method

3.1 Recommender System for Medical Diagnosis

A medical diagnosis recommender system reads the demographical features and clinical symptoms of a patient and then predicts potential diseases, which can help doctors make decisions in their clinical diagnosis. In the recommender system, we use a vector $\boldsymbol{r} = [r_1, ..., r_M] \in R^M$ to represent M demographical features of patients, a vector $\boldsymbol{s} = [s_1, ..., s_N] \in S^N$ to represent N clinical symptoms, and a set $D = \{ d_1, ..., d_K \}$ to represent K diseases. Each element of $\boldsymbol{r}$, $\boldsymbol{s}$, or D is a positive real number or zero when the element is not specified.

Definition 1: (A medical diagnosis recommender system) uses a utility function $\Re$ to map patients and their symptoms onto diseases, i.e., $\Re : (R \times S) \rightarrow D$.

The utility function $\Re$ determines an optimal mapping from $R \times S$ to D by calculating some measure (usually similarity) on a diagnosis database. A diagnosis database has the form $B = \{[\boldsymbol{r}^{(i)}, \boldsymbol{s}^{(i)}, d^{(i)}]\}_{i=1}^{I}$, where each entry is called a *diagnosis*.

Definition 2: (A diagnosis) is a triple tuple $[\boldsymbol{r}, \boldsymbol{s}, d]$, where $\boldsymbol{r}$ and $\boldsymbol{s}$ together comprise the features of a patient, and the non-zero d is a disease. If d is zero, we call it an *undecided* diagnosis.

Therefore, in the system, each diagnosis is a non-negative real-number vector. A recommendation is to predict a d to substitute 0 in an undecided diagnosis $[\boldsymbol{r}, \boldsymbol{s}, 0]$. Typically, this goal is achieved by calculating the similarity among the diagnoses (including undecided diagnosis). Obviously, elements in a diagnosis vector have different scales and types. For example, "age" is in the range of [0, 120] and usually an integer, while "temperature" is in the range of [35.0, 42.0] and usually a real number. Thus, directly computing similarity on these elements (also called "attribute" in this study) may result in poor performance. To address this issue, we use neutrosophic sets, which can convert attributes into a more meaningful space with a uniform scale.

3.2 Neutrosophic Set, Neutrosophication, and Deneutrosophication [15]

For each positive real number attribute x, we can define a set of linguistic labels $L^{(x)} = \{l_1, ..., l_C\}$ on it to represent some concepts. For example, if x is the age of a patient, we can define $L^{(x)} = \{infant, child, adolescent, youth, middle, senior\}$. We can use function $L(x, c)$, $(1 \leq c \leq C)$ to retrieve the c-th linguistic label of attribute x. Without ambiguity in the context, we can directly use c to represent its c-th label. For attribute x, we can define its neutrosophic set.

Definition 3: (A neutrosophic set) of the c-th linguistic label of attribute x is a triple tuple $[T_c(x), I_c(x), F_c(x)]$, where $T_c(x)$, $I_c(x)$, and $F_c(x)$ are membership functions that measure the degrees of the truth membership, indeterminate membership, and false membership of attribute x belonging to the c-th concept, respectively.

Definition 4: (Membership functions T, I, and F) are defined on the partitions of the real number attributes on the linguistic label (concept) l_c. For each membership function, the range of attribute x is divided into three intervals, saying that the range of x is divided into $[\alpha_1^{(c)}, \alpha_2^{(c)})$, $[\alpha_2^{(c)}, \alpha_3^{(c)})$, $[\alpha_3^{(c)}, \alpha_4^{(c)})$ for T, $[\beta_1^{(c)}, \beta_2^{(c)})$, $[\beta_2^{(c)}, \beta_3^{(c)})$, $[\beta_3^{(c)}, \beta_4^{(c)})$ for I, and $[\gamma_1^{(c)}, \gamma_2^{(c)})$, $[\gamma_2^{(c)}, \gamma_3^{(c)})$, $[\gamma_3^{(c)}, \gamma_4^{(c)})$ for F, membership functions T, I and F are defined as follows:

$$T_c(x) = \begin{cases} \frac{x-\alpha_1^{(c)}}{\alpha_2^{(c)}-\alpha_1^{(c)}}, & x \in [\alpha_1^{(c)}, \alpha_2^{(c)}) \\ \frac{\alpha_3^{(c)}-x}{\alpha_3^{(c)}-\alpha_2^{(c)}}, & x \in [\alpha_2^{(c)}, \alpha_3^{(c)}) \\ \frac{x-\alpha_3^{(c)}}{\alpha_4^{(c)}-\alpha_3^{(c)}}, & x \in [\alpha_3^{(c)}, \alpha_4^{(c)}) \\ 0, & \text{otherwise} \end{cases} \tag{1}$$

$$I_c(x) = \begin{cases} \frac{\beta_2^{(c)} - x}{\beta_2^{(c)} - \beta_1^{(c)}}, & x \in [\beta_1^{(c)}, \beta_2^{(c)}) \\ \frac{x - \beta_2^{(c)}}{\beta_3^{(c)} - \beta_2^{(c)}}, & x \in [\beta_2^{(c)}, \beta_3^{(c)}) \\ \frac{\beta_3^{(c)} + \beta_4^{(c)} - x}{\beta_4^{(c)} - \beta_3^{(c)}}, & x \in [\beta_3^{(c)}, \beta_4^{(c)}) \\ 1, & \text{otherwise} \end{cases} \tag{2}$$

$$F_c(x) = \begin{cases} \frac{\gamma_2^{(c)} - x}{\gamma_2^{(c)} - \gamma_1^{(c)}}, & x \in [\gamma_1^{(c)}, \gamma_2^{(c)}) \\ \frac{x}{\gamma_3^{(c)}}, & x \in [\gamma_2^{(c)}, \gamma_3^{(c)}) \\ \frac{\gamma_4^{(c)} + \gamma_3^{(c)} - x}{\gamma_4^{(c)} - \gamma_3^{(c)}}, & x \in [\gamma_3^{(c)}, \gamma_4^{(c)}) \\ 1, & \text{otherwise} \end{cases} \tag{3}$$

Here, the value of each membership function is in the range of [0, 1].

Definition 5: (Neutrosophication operation) Ψ extends each attribute x to $3C$ attributes, given its linguistic labels $L^{(x)} = \{l_1, ..., l_C\}$. That is,

$$\Psi(x) = [T_1(x), I_1(x), F_1(x), ..., T_C(x), I_C(x), F_C(x)]. \tag{4}$$

For a M-dimensional vector $\boldsymbol{x}$ with different linguistic labels $L^{(x_1)}$, $L^{(x_2)}$, …, $L^{(x_M)}$, sized C_1, C_2,…, C_M on each element, we have:

$$\begin{aligned} \Psi(\boldsymbol{x}) = [&T_{L(x_1,1)}(x_1), I_{L(x_1,1)}(x_1), F_{L(x_1,1)}(x_1), T_{L(x_1,2)}(x_1), \\ &I_{L(x_1,2)}(x_1), F_{L(x_1,2)}(x_1), ..., T_{L(x_M,C_M)}(x_M), \\ &I_{L(x_M,C_M)}(x_M), F_{L(x_M,C_M)}(x_M)] \end{aligned} \tag{5}$$

Definition 6: (Deneutrosophication operation) $\overline{\Psi}$ first transforms the membership functions of neutrosophic set into the membership function of a fuzzy set A:

$$\mu_A(x) = \kappa T_c(x) + \tau \frac{F_c(x)}{4} + \upsilon \frac{I_c(x)}{2}, \tag{6}$$

where $\kappa, \tau, \upsilon \in [0, 1]$ and $\kappa + \tau + \upsilon = 1$. Then, a typical deneutrosophicated value $den(\mu_A(x))$ can be calculated by the centroid or center of gravity method below:

$$den(\mu_A(x)) = \frac{\int_x \mu_A(x) x dy}{\int_x \mu_A(x) dy}. \tag{7}$$

3.3 Neutrosophic Recommender System for Medical Diagnosis

Our neutrosophic recommender system directly extends each attribute in diagnoses to three membership functions. On these new attributes, the system calculates the membership functions of diseases in those undecided diagnoses. Then, it uses deneutrosophication to recover the original value of the predicted diseases.

Definition 7: (A neutrosophic recommender system) defines a utility function $\Re$ on neutrosophication operations, i.e., $\Re : \Psi(R) \times \Psi(S) \rightarrow \Psi(D)$.

Therefore, two essential functions of a neutrosophic recommender system are [3]:

(1) Prediction: given $[\Psi(\boldsymbol{r}), \Psi(\boldsymbol{s}), 0]$, computing $\Psi(d), d \in D$;
(2) Recommendation: given d, find a most meaningful clinical interpretation by

$$l_c = \arg\max_{l_c \in L^{(d)}} \{T_c(d) + T_c(d)(3 - T_c(d) - I_c(d) - F_c(d))\} \tag{8}$$

Finally, we can use a deneutrosophication operation to obtain the specific value of d.

3.4 Recommendation Algorithm

The goal of the recommendation algorithm is to calculate the similarity between diagnoses and compensate missed d in those undecided diagnoses.

For each diagnosis $[\boldsymbol{r}, \boldsymbol{s}, d]$ in the dataset, we create its neutrosophic set as a vector $\boldsymbol{w} = [\Psi(\boldsymbol{r}), \Psi(\boldsymbol{s}), \Psi(d)]$. Because some ds are unknown, we define the features of a diagnosis as $\boldsymbol{v} = [\Psi(\boldsymbol{r}), \Psi(\boldsymbol{s})]$. For two diagnoses $\boldsymbol{v}^{(i)}$ and $\boldsymbol{v}^{(j)}$, we calculate their similarity as follows:

$$s(\boldsymbol{v}^{(i)}, \boldsymbol{v}^{(j)}) = \frac{\sum_{m=1}^{M} SR_m^{(ij)} + \sum_{n=1}^{N} SS_n^{(ij)}}{M + N}, \tag{9}$$

where SR_m and SS_n are similarities of a demographic feature and a symptom, respectively. That is, on the neutrosophic set space, they can be calculated through the similarity of each pair of corresponding attributes xs in two diagnoses as follows:

$$\begin{aligned} s(x^{(i)}, x^{(j)}) = \frac{1}{2C} \sum_{c=1}^{C} \max\Big\{ & \Big|T_c(x^{(i)}) - T_c(x^{(j)})\Big|, \\ & \Big|I_c(x^{(i)}) - I_c(x^{(j)})\Big|, \Big|F_c(x^{(i)}) - F_c(x^{(j)})\Big| \Big\}, \end{aligned} \tag{10}$$

where $x^{(i)}, x^{(j)} \in R \cup S$. Then, we can construct a similarity matrix $P^{I \times I}$ for all I diagnoses on their features as follows:

$$P^{I \times I} = \begin{bmatrix} s(\boldsymbol{v}^{(1)}, \boldsymbol{v}^{(1)}) & \cdots & s(\boldsymbol{v}^{(1)}, \boldsymbol{v}^{(n)}) \\ \vdots & \ddots & \vdots \\ s(\boldsymbol{v}^{(n)}, \boldsymbol{v}^{(1)}) & \cdots & s(\boldsymbol{v}^{(I)}, \boldsymbol{v}^{(I)}) \end{bmatrix}. \tag{11}$$

Although matrix P can measure the similarity of each pair of diagnoses, it is not good enough because all diagnoses have the same weights. That is, a large variance on a few features may result in considerable changes in similarity measure, leading to inaccurate predictions. To address this issue, we resort to spectral clustering [12].

The spectral clustering treats all the data as points in a uniform space, where related points are connected by edges. The edge weight between the two distant points is low,

and that between the two close points is high. The algorithm cuts the graphs composed of all the data points, making the weights between different subgraphs after cutting are as low as possible, and the weights within subgraphs are as high as possible. Thus, a spectral clustering algorithm needs to construct a similarity matrix of data points. That is, a spectral clustering algorithm can have four types of input data, which are *data points*, *similarity matrix*, *base clustering algorithm*, and *the number of clusters*. Here, the data points are diagnosis features $[\boldsymbol{v}_1; ...; \boldsymbol{v}_I]^T$ and the similarity matrix is P.

Suppose the centroids of clusters are $\{\boldsymbol{\sigma}_1, ..., \boldsymbol{\sigma}_U\}$. We define the counter-similarity of a diagnosis $\boldsymbol{v}$ and a centroid $\boldsymbol{\sigma}_u$ as follows:

$$cs(\boldsymbol{v}, \boldsymbol{\sigma}_u) = 1 - \frac{\boldsymbol{v} \cdot \sigma_u}{\|\boldsymbol{v}\| \|\boldsymbol{\sigma}_u\|}. \tag{12}$$

The probability of a diagnosis $\boldsymbol{v}$ belonging to cluster u is:

$$p(\boldsymbol{v}, u) = 1 - \frac{cs(\boldsymbol{v}, \boldsymbol{\sigma}_u)}{\max\{cs(\boldsymbol{v}^{(i)}, \boldsymbol{\sigma}_u)\}_{i=1}^{I}} \tag{13}$$

Then, we construct a similarity measure between two diagnoses based on their positions in the clusters as follows:

$$s'(\boldsymbol{v}_i, \boldsymbol{v}_j) = \frac{1}{U} \sum_{u=1}^{U} |p(\boldsymbol{v}_i, u) - \overline{p}(\boldsymbol{v}_i)| \left| p(\boldsymbol{v}_j, u) - \overline{p}(\boldsymbol{v}_j) \right|, \tag{14}$$

where we have $\overline{p}(\boldsymbol{v}) = \sum_{u=1}^{U} p(\boldsymbol{v}, u) \Big/ U$. We make a linear combination of two similarity measures as follows:

$$s''(\boldsymbol{v}_i, \boldsymbol{v}_j) = \lambda s(\boldsymbol{v}_i, \boldsymbol{v}_j) + (1 - \lambda) s'(\boldsymbol{v}_i, \boldsymbol{v}_j), \tag{15}$$

where $\lambda \in [0, 1]$ is an adjustable coefficient.

Finally, we take the similarity of diseases of two diagnoses into account and obtain the final similarity measure of two diagnoses as follows:

$$s_{final}(\boldsymbol{v}_i, \boldsymbol{v}_j) = s''(\boldsymbol{v}_i, \boldsymbol{v}_j) + SD^{(ij)} - s''(\boldsymbol{v}_i, \boldsymbol{v}_j) \cdot SD^{(ij)}, \tag{16}$$

where $SD^{(ij)}$ can be calculated by Eq. (8). Also, we can construct a similarity matrix $Q^{I \times I}$ for all I diagnoses as $Q^{(ij)} = s_{final}(\boldsymbol{v}_i, \boldsymbol{v}_j)$.For an undecided diagnosis $\boldsymbol{w}^{(i)}$, its neutrosophic set of the disease $d^{(i)}$ on the linguistic label c is calculated as following equation:

$$T_c(d^{(i)}) = \frac{\sum_{j=1}^{I} Q^{(ij)} T_c(d^{(j)})}{\sum_{j=1}^{I} Q^{(ij)}}, \tag{17}$$

$$I_c(d^{(i)}) = T_c(d^{(i)}) + \frac{\sum_{j=1}^{I} Q^{(ij)} I_c(d^{(j)})}{\sum_{j=1}^{I} Q^{(ij)}}, \tag{18}$$

$$F_c(d^{(i)}) = I_c(d^{(i)}) + \frac{\sum_{j=1}^{I} Q^{(ij)} F_c(d^{(j)})}{\sum_{j=1}^{I} Q^{(ij)}}. \tag{19}$$

Neutrosophic Recommendation with Spectral Clustering (NRSC). Algorithm NRSC presents the main steps of the proposed method. The input of algorithm 1 contains $B = \{[\boldsymbol{r}^{(i)}, \boldsymbol{s}^{(i)}, d^{(i)}]\}_{i=1}^{I}$, Δ, base_cluster_algo, U, λ. Here, $B = \{[\boldsymbol{r}^{(i)}, \boldsymbol{s}^{(i)}, d^{(i)}]\}_{i=1}^{I}$ represents the diagnosis database. Neutrosophication parameters Δ (such as $L^{(x)}$, $\alpha_1^{(c)}, ..., \alpha_4^{(c)}, \beta_1^{(c)} ..., \beta_4^{(4)}, \gamma_1^{(c)}, ..., \gamma_4^{(c)}, \kappa, \tau, \upsilon$) for each attribute x that will be used in Eqs. (1)–(7) are required a Neutrosophication Settings. Parameter U is the number of clusters. Parameter $\lambda \in [0, 1]$ is the adjustable coefficient that is used to calculate the final similarity measure.

First, the process of neutrosophication is that calculating the *T, I, F* values for each element including different linguistic labels to form $\boldsymbol{w}^{(i)}$ and $\boldsymbol{v}^{(i)}$ (Line 1). Then, constructing the neutrosophic similarity matrix P for all I diagnoses on their features according to each element's *T, I, F* values that are required to the neutrosophic algebraic operations (Line 2). Third, although P can measure the similarity of each pair of diagnoses, it is not good enough. Therefore, a spectral clustering algorithm is required to cluster the data points, and a base clustering algorithm (base_cluster_algo) such as K-means, which is chosen for spectral clustering (Line 3). After clustering, making a combination of two similarity measures and taking the similarity of diseases of two diagnoses into account and obtain the final similarity measure of two diagnoses (Line 4). Moreover, to predict the disease, the final similarity matrix Q is needed to calculate the neutrosophic set of the disease (Line 5), then find the most meaningful clinical interpretation from linguistic labels (Line 6). Finally, calculating the final result of disease according to the process of deneutrosophication (Line 7). Based on Algorithm 1, the predicted result is generated, which can be compared with other algorithms in our experiments.

Algorithm 1: Neutrosophic Recommendation with Spectral Clustering (NRSC)

Input: $B = \{[\boldsymbol{r}^{(i)}, \boldsymbol{s}^{(i)}, d^{(i)}]\}_{i=1}^{I}$, Δ, base_cluster_algo, U, λ
Output: d for an undecided disease
1: **For each** diagnosis $[\boldsymbol{r}^{(i)}, \boldsymbol{s}^{(i)}, d^{(i)}], i = 1, ..., I$
Do neutrosophication to form $\boldsymbol{w}^{(i)}$ and $\boldsymbol{v}^{(i)}$ by Eq. (5)
2: Calculate similarity matrix P by Eqs. (9)-(11)
3: **Call** Spectral_Clustering($[\boldsymbol{v}^{(i)}]_{i=1}^{I}$, base_cluster_algo, U)
4: Calculate the final similarity matrix Q by Eqs. (12)-(16)
5: For undecided diagnosis $\boldsymbol{w}^{(i)}$, calculate its neutrosophic set by Eqs. (17-19)
6: Recommend a linguistic label $L(d^{(i)}, c)$ by Eq. (8)
7: Predict $d^{(i)}$ by deneutrosophication using Eqs. (6-7)

4 Experiment

In this section, we first introduce the experimental settings and the methods in comparison. Then, we focus on discussing experimental results.

4.1 Experimental Settings

Our experiments were conducted on four medical diagnosis datasets *Heart*, *Diabetes*, *RHC,* and *DMD*. Dataset *Heart* is from the UCI Machine Learning Repository [23] and consists of 270 medical records characterized by 13 attributes, such as chest pain type, resting blood pressure, fasting blood sugar and so on. Datasets *Diabetes*, *RHC*, *DMD* are all taken from [24]. Dataset *Diabetes* consists of 403 medical records characterized by 19 attributes and is often associated most strongly with obesity and hypertension. Dataset *RHC* consists of 5735 medical records characterized by 62 attributes and is associated with Mean Blood Pressure, white blood cell, heart rate, and so on. Dataset *DMD* consists of 209 medical records characterized by 9 attributes and is associated with Hemopexin, Pyruvate Kinase, Lactate Dehydrogenase and so on.

We implemented the proposed method based on the settings in [3] using MATLAB, where neutrosophication parameters are provided for the above datasets. We directly used their settings and conducted our spectral clustering only on the symptoms. In our spectral clustering implementation, we use K-means method as the base clustering algorithm.

We compared the proposed method (NRSC) with five state-of-the-art methods PFS, ICSM, NR, CARE, and DSM.

- PFS [27] presented an improved max-min-max composite relation using Pythagorean fuzzy sets for medical diagnosis.
- ICSM [9] presented an improved method based on the cosine similarity measures to solve medical diagnosis problems with simplified neutrosophic information.
- NR [3] designed a new hybrid method that combines the neutrosophic sets and traditional recommender systems for medical diagnosis.
- CARE [26] presented a recommendation engine that combines collaborative filtering with clustering to predict patients' most possible diseases.
- DSM [21] presented the dice similarity measure which was applied to medical diagnosis to deal with indeterminate and inconsistent information.

In this study, we use Root Mean Square Error (RMSE) and Mean Absolute Error (MAE) as evaluation metrics. MSE is the expected root of the square of the difference between the estimated value and the true value of the parameter and can evaluate the change degree of data. The smaller the value of MSE is, the better the accuracy of the prediction model to describe the experimental data is. MAE is the mean of absolute errors, which is not sensitive towards outliers and can better reflect the actual situation of the predicted value errors.

4.2 Experimental Results

1) Comparison of Proposed Method NRSC with NR, PFS, ICSM, DSM and CARE in RMSE and MAE

The comparison results of six state-of-the-art algorithms on the four datasets in terms of RMSE and MAE are shown in Fig. 1. From Fig. 1(a), we can find that RMSE of the proposed method (NRSC) is 0.225 on the *Heart* dataset, while that of PFS, NR, CARE, ICSM, and DSM are 0.243, 0.238, 0.250, 0.340, and 0.341, respectively. On the other hand, the MAE of NRSC is still smaller than the other methods. Therefore, NRSC outperforms all the previously above-mentioned methods.

From Fig. 1(b) and Fig. 1(d), we can find that NRSC consistently outperforms the other five methods in both evaluation metrices on both datasets *Diabetes* and *DMD*. Obviously, from Fig. 1(b), we can find that on dataset *Diabetes* the RMSE of NRSC is 0.167, while that of PFS, NR, ICSM, DSM and CARE are 0.179, 0.195, 0.329, 0.33, and 0.354, respectively. The MAE of NRSC is 0.14, while that of PFS, NR, ICSM, DSM and CARE are 0.153, 0.154, 0.289, 0.289, and 0.312, respectively. Similarly, from Fig. 1(d), we can find that on dataset *DMD*, the RMSE of NRSC is 0.215, while that of PFS, NR, ICSM, DSM and CARE are 0.231, 0.25, 0.258, 0.358 and 0.243, respectively. The MAE of NRSC is 0.315, while that of PFS, NR, ICSM, DSM and CARE are 0.371, 0.5, 0.358, 0.358, and 0.346, respectively. Therefore, NRSC still outperforms the other five methods.

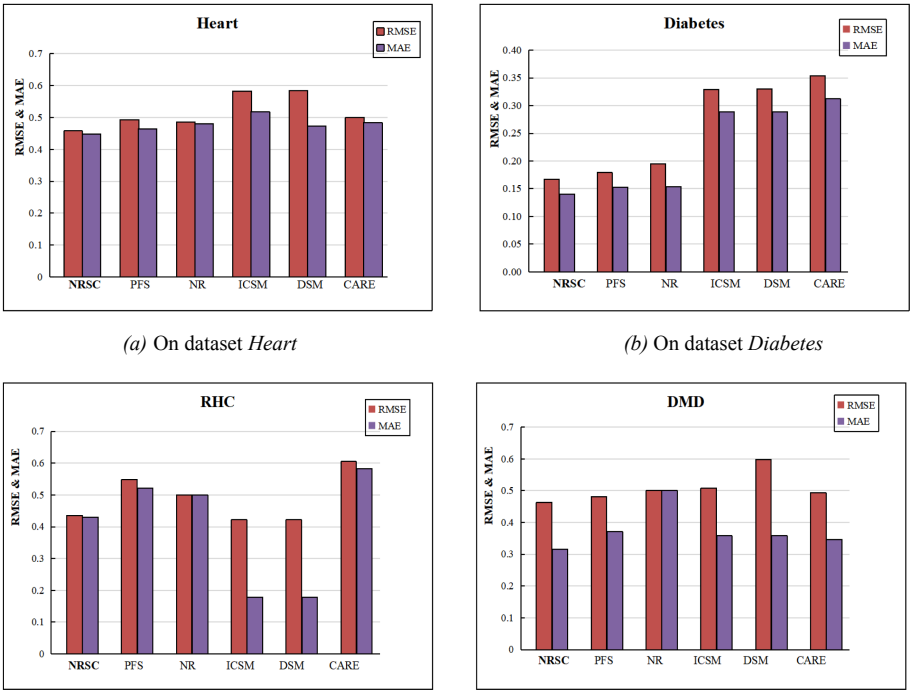
From Fig. 1(c), on dataset *RHC*, we can find that the RMSE of the proposed NRSC (0.435) is slightly worse than ICSM (0.422) and DSM (0.422) in both evaluation metrics. However, differences between them are not statistically significant. Compared with the methods PFS, NR and CARE, the performance of NRSC is still significantly better, which suggests that the spectral clustering mechanism is still of effectiveness.

To sum up, in terms of both RMSE and MAE, the proposed NRSC achieves a better overall performance on the four datasets against all five state-of-the-art algorithms.

2) The performance of NRSC by tuning the parameters

Since there are several parameters in our proposed method, we need to investigate the robustness of the algorithm when tuning these parameters. Table 1 and Table 2 shows the RMSEs and MAEs, and running time of the proposed method (NRSC) when the parameters U, and λ are set to different values on four datasets. Parameter U is the number of clusters and parameter $\lambda \in [0, 1]$ is the adjustable coefficient that is used to calculate the final similarity measure. We find that the similarity degrees between patients derived from spectral clustering are supplemented into neutrosophic recommender similarity matrix to obtain the final similarity between items with similar demographic information and symptoms.

In the experiment, we can find that the values of RMSE and MAE change almost at the same time on each dataset as the parameters change. The running time also varies when we changed the values of parameters.



(a) On dataset *Heart*

(b) On dataset *Diabetes*

(c) On dataset *RHC*

(d) On dataset *DMD*

Fig. 1. Comparison of six state-of-the-art methods on four datasets in terms of RMSE and MAE

From the achieved results on Table 1, the value of λ should be larger than 0.5, ideally in the range of [0.6, 0.8] as expressed on four datasets. Besides, from Table 2, we find that the number of clusters is larger, the running time will be longer when we set $\lambda = 0.6$ and change the number of clusters U. Although the number of clusters U affects the running time, the increment is small and linear to U. At last, it is necessary to select appropriate parameters on four datasets. When we set the number of clusters $U = 3$ and $\lambda = 0.6$, the proposed method can consistently achieve its mostly best performance on the four medical datasets.

We also compared the proposed NRSC (setting $U = 3$ and $\lambda = 0.6$) with the five existing methods in running time, whose results are shown in Table 3. Compared with those simple methods, such as ICMS and DSM, the running time of our NRSC only slightly increases because of introducing spectral clustering. However, our method is still better than PFS and NR. Especially, when the size of the dataset is large (RHC), our method is as fast as ICSM and DSM. To sum up, the running time of the proposed NRSC is at least not worse than the state-of-the-art methods and sometimes better than some of the other methods, which is acceptable in a real-world environment.

Table 1. Performance of NRSC when tuning parameters on four datasets.

U	λ	Heart		Diabetes		RHC		DMD	
		RMSE	MAE	RMSE	MAE	RMSE	MAE	RMSE	MAE
2	0.2	0.233	0.480	0.037	0.157	0.201	0.440	0.218	0.321
2	0.4	0.226	0.470	0.034	0.151	0.198	0.440	0.216	0.317
2	0.5	0.225	0.469	0.032	0.150	0.196	0.435	0.215	0.315
2	0.6	**0.226**	**0.468**	**0.031**	**0.149**	0.196	0.434	**0.215**	**0.315**
2	0.8	0.226	0.470	0.032	0.149	**0.195**	**0.430**	0.216	0.315
3	0.2	0.230	0.475	0.035	0.154	0.198	0.438	0.218	0.318
3	0.4	0.225	0.468	0.033	0.150	0.197	0.435	0.215	0.314
3	0.5	0.226	0.468	0.033	0.149	0.195	0.435	0.215	0.315
3	0.6	0.224	0.464	**0.030**	**0.148**	**0.193**	**0.430**	**0.214**	**0.310**
3	0.8	**0.223**	**0.462**	0.033	0.149	0.195	0.433	0.214	0.311
4	0.2	0.230	0.473	0.038	0.154	0.198	0.443	0.221	0.321
4	0.4	0.227	0.469	0.036	0.152	0.196	0.438	**0.215**	**0.315**
4	0.5	0.227	0.470	0.037	0.152	0.195	0.436	0.217	0.318
4	0.6	0.230	0.470	**0.034**	**0.151**	**0.195**	**0.433**	0.216	0.316
4	0.8	**0.226**	**0.469**	0.035	0.151	0.195	0.433	0.217	0.315
5	0.2	0.235	0.475	0.040	0.160	0.214	0.452	0.225	0.323
5	0.4	0.232	0.471	0.038	0.157	0.201	0.446	0.220	0.320
5	0.5	0.228	0.470	0.038	0.155	0.198	0.440	0.218	0.320
5	0.6	**0.228**	**0.469**	0.036	0.153	0.197	0.439	0.217	0.318
5	0.8	0.228	0.469	**0.035**	**0.153**	**0.197**	**0.435**	**0.216**	**0.315**

Table 2. Running time (sec) of NRSC when tuning parameter U on four datasets.

U	2	3	4	5
Heart	1.651	2.173	2.673	3.152
Diabetes	3.621	3.788	4.963	5.324
RHC	175.282	253.667	282.376	302.231
DMD	1.104	1.37	1.620	1.962

Table 3. Comparison in running time (sec) on four datasets ($U = 3$ and $\lambda = 0.6$ for NRSC).

Algorithms	NRSC	PFS	NR	ICSM	DSM	CARE
Heart	0.173	1.050	0.218	0.126	0.154	0.185
Diabetes	0.802	1.638	0.753	0.289	0.362	0.553
RHC	253.667	420.379	380.456	247.189	254.254	280.821
DMD	1.071	2.842	1.157	0.934	1.191	1.329

5 Conclusion

In this paper, we concentrated on improving the performance of prediction in medical diagnosis. To address the problem that the features of patients' demographic information and symptoms have different scales and types, we resort to neutrosophic sets to convert these features into uniform-scaled values described by three membership functions. Then, we proposed a novel recommendation method that combines the neutrosophic recommendation method and spectral clustering. The spectral clustering can better group the items with similar demographic information and symptoms together and makes the similarity measures between items more accurate, which results in more accurate prediction and recommendation of diseases. And we compared the proposed method with four state-of-the-art methods on four medical diagnosis datasets. Experimental results demonstrate that the proposed methods outperform the others in terms of Root Mean Square Error and Mean Average Error.

This study is still preliminary. We only focus on predicting one disease. However, in real-world circumstances, relevant diseases usually happen together. Thus, the system should be able to predict a set of relevant diseases simultaneously. Moreover, constructing neutrosophic sets for the complicated datasets with large dimensions of features is time-consuming but essential to the development of the recommendation methods. In the future, we will keep our studies along with these directions.

Acknowledgments. This work has been supported by the National Natural Science Foundation of China under grants 62076130 and 91846104, and the National Key Research and Development Program of China under grant 2018AAA0102002.

References

1. Manogaran, G., Varatharajan, R., Priyan, M.K.: Hybrid recommendation system for heart disease diagnosis based on multiple kernel learning with adaptive neuro-fuzzy inference system. Multimedia Tools Appl. **77**(4), 4379–4399 (2017). https://doi.org/10.1007/s11042-017-5515-y
2. Li, X., Wang, Z.: A new recommendation algorithm combined with spectral clustering and transfer learning. Cluster Comput. **22**(1), 1151–1167 (2017). https://doi.org/10.1007/s10586-017-1161-4

3. Ali, M., Son, L.H., Thanh, N.D., Minh, N.V.: A neutrosophic recommender system for medical diagnosis based on algebraic neutrosophic measure. Appl. Soft Comput. **71**, 1054–1071 (2018)
4. Hu, J., Pan, L., Yang, Y., Chen, H.: A group medical diagnosis model based on intuitionistic fuzzy soft sets. Appl. Soft Comput. J. **77**, 453–466 (2019)
5. Khalil, S.M.: Decision making using algebraic operations on soft effect matrix as new category of similarity measures and study their application in medical diagnosis problems. J. Intell. Fuzzy Syst. **37**(2), 1865–1877 (2019)
6. Smarandache, F.: A Unifying Field in Logics. Neutrosophy: Neutrosophic Probability, Set and Logic, Rehoboth. American Research Press, Santa Fe (1998)
7. Guo, Y., et al.: Automated iterative neutrosophic lung segmentation for image analysis in thoracic computed tomography. Med. Phys. **40**(8), 081912 (2013)
8. Zhang, M., Zhang, L., Cheng, H.D.: Segmentation of ultrasound breast images based on a neutrosophic method. Opt. Eng. **49**(11), 117001 (2010)
9. Ye, J.: Improved cosine similarity measures of simplifield neutrosophic sets for medical diagnoses. Artif. Intell. Med. **63**(3), 171–179 (2015)
10. Ye, J., Fu, J.: Multi-period medical diagnosis method using a single valued neutrosophic similarity measure based on tangent function. Comput. Methods Programs Biomed. **123**, 142–149 (2016)
11. Ye, S., Fu, J., Ye, J.: Medical diagnosis using distance-based similarity measures of single valued neutrosophic multisets. Neutrosophic Sets Syst. **7**, 47–54 (2015)
12. Xu, D., Xu, Z., Liu, S., Zhao, H.: A spectral clustering algorithm based on intuitionistic fuzzy information. Knowl.-Based Syst. **53**, 20–26 (2013)
13. Broumi, S., Smarandache, S.: Extended hausdorff distance and similarity measure of refined neutrosophic sets and their application in medical diagnosis. J. New Theory **11**(7), 64–78 (2015)
14. Thong, N.T., Son, L.H.: HIFCF: an effective hybrid model between picture fuzzy clustering and intuitionistic fuzzy recommender systems for medical diagnosis. Exp. Syst. Appl. **42**(7), 3682–3701 (2015)
15. Wang, H., Smarandache, F., Zhang, Y.Q., Sunderraman, R.: Single valued neutrosophic sets. Multispace Multistruct. **4**, 410–413 (2010)
16. Duan, L., Street, W.N., Xu, E.: Healthcare information systems: data mining methods in the creation of a clinical recommender system. Enterp. Inf. Syst. **5**(2), 169–218 (2011)
17. Mahabaleshwar, S.: Kabbur : an efficient multiclass medical image CBIR system based on classification and clustering. J. Intell. Syst. **27**(2), 275–290 (2018)
18. Thanh, N.D., Ali, M., Son, L.H.: A novel clustering algorithm in a neutrosophic recommender system for medical diagnosis. Cogn. Comput. **9**(4), 526–544 (2017). https://doi.org/10.1007/s12559-017-9462-8
19. Son, L.H., Thong, N.T.: Intuitionistic fuzzy recommender systems: an effective tool for medical diagnosis. Knowl.-Based Syst. **74**, 133–150 (2015)
20. Yang, Y., Ma, Z., Yang, Y.: Multitask spectral clustering by exploring intertask correlation. IEEE Trans. Cybern. **45**(5), 1083–1094 (2015). https://doi.org/10.1109/TCYB.2014.2344015
21. Ye, S., Ye, J.: Dice similarity measure between single valued neutrosophic multisets and its application in medical diagnosis. Neutrosophic Sets Syst. **6**, 48–53 (2014)
22. Broumi, S.: F, Smarandache: serveral similarity measures of neutrosophic sets. Neutrosophic Sets Syst. **1**, 55 (2013)
23. UCI Repository of Machine Learning Databases, University of California (2007). https://archive.ics.uc.edu/ml
24. Department of Biostatistics. Vanderbilt University (2016). https://biostat.mc.vanderbilt.edu/DataSets

25. Lee, J., Noh, G., Oh, H., Kim, C.K.: Trustor clustering with an improved recommender system based on social relationships. Inf. Syst. **77**, 118–128 (2018)
26. Davis, D.A., Chawla, N.V., Blumm, N., Christakis, N., Barabási, A.L.: predicting individual disease risk based on medical history. In: Proceedings of the 17th ACM Conference on Information and Knowledge Management, pp. 769–778 (2008)
27. Ejegwa, P.A.: Improved composite relation for pythagorean fuzzy sets and its application to medical diagnosis. Granular Comput. **5**(2), 277–286 (2019). https://doi.org/10.1007/s41066-019-00156-8
28. He, S., Cheng, B., Huang, Y., Duan, L., Chen, J.: Proactive personalized services in large-scale IoT-based healthcare application. In: 2017 IEEE International Conference on Web Services (ICWS), pp. 808–813 (2017)