Neutrosophic Sets and Systems, Vol. 67, 2024
University of New Mexico

Harnessing Pliancy Tree Soft Sets in Heart Diseases for Extracting Beneficial Rules of Association Rules

Gawaher S. Hussein [1], Khalid A. Eldrandaly [1] , Abdel Nasser H. Zaied[2] ,Samar L. Elhawy[1,*] , Mona Mohamed[3]

[1]Faculty of Computers and Informatics, Zagazig University, Sharqiyah, Zagazig, 44519, Egypt.Emails: Gawahersoliman@zu.edu.eg, Khalid_Eldrandaly@zu.edu.eg , SLAlhawy@fci.zu.edu.eg.

[2]Department of Information Systems, Misr International University, Egypt; Emails abdelnasser.riad@miuegypt.edu.eg.

[3]Higher Technological Institute, 10th of Ramadan City 44629, Egypt, Email: mona.fouad@hti.edu.eg.

Abstract

Cardiovascular diseases (CVDs)continue to be the primary cause of mortality, accounting for approximately one-third of all fatalities globally. This spawned the proposal of models in several studies. Accordingly, this study contributed to diagnosing heart disease through suggesting soft diagnosis paradigm. Various techniques have been volunteered for serving the suggested paradigm toward achieving its objective. Additionally, this study provided set of contributions. For instance, Tree Soft Technique (TrST) is applied for the first time for forming attributes and sub attributes of patients into nodes and sub-nodes of Tree to obtain relations between it. Even, the study support stakeholders to making accurate decision in mysterious circumstances and in problems with incomplete information through Collaborating the utilized techniques of entropy and Technique for order of preference by similarity to ideal solution (TOPSIS) in this study with Single Value Neutrosophic Sets (SVNSs) forked from neutrosophic uncertainty theory. As well, the relationship between sub-attributes which consider antecedent for obtaining consequent of detecting and diagnosing through collaborating TrST with association rules. Accordingly, we applied four transactions (cases) for obtaining findings of the relations in transactions as listed in Table 13 and Table14.

Keywords:

Cardiovascular diseases (CVDs), Tree Soft Technique (TrST), Single Value Neutrosophic Sets (SVNSs), association rules.

1. Introduction

1.1 Contextual Study

Indeed, according to the World Health Organization (WHO)[1] , cardiovascular diseases (CVDs) are the primary cause of mortality globally, especially in nations with the greatest poverty. The encouragement for this [2] comes from the alarming global statistics on cardiovascular disease, which indicate that by 2030, the total number of deaths per year will surpass 20 million. Overall, the expression of CVDs described in [3] as a variety of conditions that impact the heart and blood arteries. Likewise thought that CVDs or heart diseases are one of the most difficult, deadly, and life-threatening illnesses that affect people worldwide. CVDs in [4] classified into various categories as represented in Fig 1. Also, this Fig illustrates factors causing the disease and its consequences.

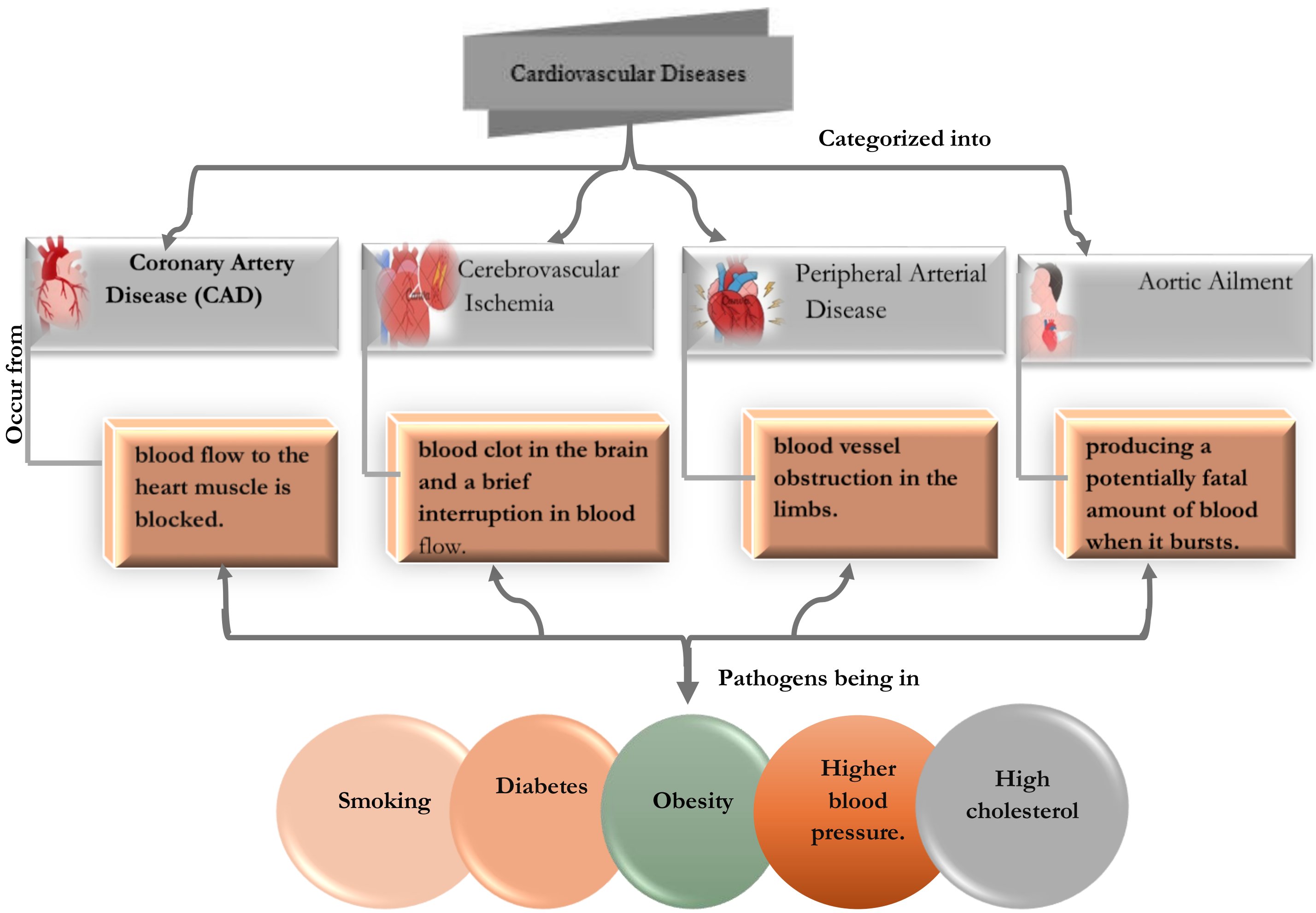

Fig 1. Cardiovascular Diseases Categories

Even though stressing that heart disease is among the most common diseases in humans, [5]also acknowledged that the diagnosis of heart disease can be complicated and delay the right diagnosis decision due to several factors, including symptoms of the disease and the relationship between the disease's pathological and functional manifestations and human organs other than the heart. In this context [6] indicated that improving clinical outcomes and preventing significant adverse cardiac events is possible with early detection of CVDs, which also enables the escalation of guideline-directed medical treatment. Arguably early identification of cardiovascular disorders is imperative. From the perspective of [7], forecast with precision is a challenge, particularly in developing and Asian nations where resources, technology, and peripheral devices are few. In the same context [8] , was revealed that the capacity for controlling CVD is currently inadequate. Individuals lack knowledge about CVDs and possible detrimental practices that contribute to the disease. Another challenge stated in [9] where a wealth of information on heart disease in the healthcare sector needs to be analyzed to aid with decision-making. The mentioned challenges are catalysts for the notion of diagnosing CVDs early to lessen the risk embraced by [10] through using appropriate and precise diagnostic techniques. Moreover, earlier studies tackle the difficulties associated with CVDs and diagnostics by developing a variety of diagnostic methods for the techniques-based prediction of heart disease [11]. For instance [12] where medical practitioners are integrating their experiences as physicians with artificial intelligence (AI) techniques to automate the diagnosing process. Recently, several machine learning (ML)techniques of AI have been constructed to improve the prediction of cardiovascular diseases. Utilizing Extreme Gradient Boosting (XGBoost) classifier [13] to predicate CVDs. On other hand [14] adapted uncertainty theory of Complex intuitionistic fuzzy set (CIFS) for choosing the best method for diagnosing cardiovascular diseases.

1.2 Novelty of Study: Rendered Contributions

Surveys for earlier research that were connected to our study as [15] indicated that doctors frequently make their judgments not on the knowledge-rich material concealed in the database, but rather on their experience and intuition. This results in unintentional prejudices, mistakes, and exorbitant medical expenses, and these issues have an impact on the standard of care given to patients.

The study discussed some of the challenges that retard the process of CVDs diagnosis. As well, this study attempts to avoid these handicaps by constructing an innovative soft diagnostic paradigm. The bedrock of this paradigm is harnessing the various soft computing techniques where neutrosophic theory plays a vital role as vagueness theory. This theory was proposed by Smarandache and can treat dynamic and uncertain environment[16]. As well Tree Soft technique is utilized for forming association between determined attributes which contribute to the diagnosis process. Table 1 shows the contribution of constructing an innovative soft diagnostic paradigm by covering the set of aspects.

Table1. Study Contributions

Aspects	Challenges	Study's Contribution
Theoretically	- From the perspective of [17] CVDs are difficult to diagnose as a Scarcity of diagnostic capabilities, as well as fewer physicians and other healthcare providers. This will affect CVDs patients' optimal prognosis.	- Soft diagnostic paradigm is constructed for diagnosing cases of CVDs to guarantee the accuracy of the diagnosis
Practically	- Due to ambiguity and uncertainty in many complicated health situations as well as limited information regarding the patient's medical status[18], the medical field—including experienced doctors—faces challenges in making accurate diagnoses.	- Set of techniques has been harnessed in this study for constructing soft diagnostic paradigm. - Each technique is responsible for a certain role. i. Tree Soft Technique (TrST) is utilized for forming a patient's attributes into nodes and sub-nodes of a tree. This technique can illustrate the associative rules between determined patients' attributes. ii. Soft computing technique of neutrosophic is utilized for analyzing attributes formed into TrST. This technique known as the uncertainty technique can treat mysterious circumstances.
Credibility	It takes time and experience to identify CVD early and treat patients more successfully, which endangers the patient's life[6].	- To guarantee the paradigm's validity, it is applied to real case studies to guarantee patient safety and life.

2. Methodology of Soft Diagnostic Paradigm

The objective of this section is exhibiting through cover the following points (P_n):

P1: The Preliminaries of techniques that contributed to constructing the proposed paradigm.

P2: The role of each technique is exhibited by presenting the proposed model in the form of steps in the proposed paradigm.

P3: The problem is formed into tree form by adopting TrST. The main attributes related to diagnosis for patients are represented into nodes and its sub-attributes are represented into sub-nodes.

P4: Employing soft computing (SC) technique is neutrosophic which is harnessed in TrST for treating uncertainty problems related to attributes for diagnosing the patients.
As a result, this section is divided into two sub-sections. Each one is responsible for a certain function for serving the previously listed points.

2.1 Preliminaries

Herein, the basic concepts and fundamentals of utilized techniques of the proposed paradigm are clarified.

2.1.1 Tree Soft Technique

The technique of TrST is suggested by Smarandache [19] who is the founder of uncertainty theory is neutrosophic. The objective of TrST is to illustrate the relationship between attributes and sub-attributes of patients and SVDs. Hence, the technique's basic aspects and relationship formed according to [20] as:

- Assum $\mathfrak{H}$ be a universe of discourse, and $\mathcal{H}$ a non-empty subset of $\mathfrak{H}$, with the powerset of $\mathcal{H}$ $P(\mathcal{H})$.
- Suppose ∂ be a set of attributes for main nodes as $\partial = \{ \partial_1, \partial_2, \ldots, \partial_n\}$ where $n \geq 1$ and considering attributes of ∂ resident at the first level.
- Accordingly, sub-attributes of the main attributes are located in the second level as sub-nodes ∂_1 symbolled as $\{ \partial_{1-1}, \partial_{1-2}, \ldots, \partial_{1-n}\}$ also, sub-attributes of ∂_2 expressed as $\{ \partial_{2-1}, \partial_{2-2}, \ldots, \partial_{2-n}\}$.
- Considering ∂ is root and located at level zero, sequentially nodes of level 1, level 2, up to level n are inherent of ∂. Moreover, Then Tree Soft is expressed as F: $P(\mathrm{Tree}(\partial)) \rightarrow P(\mathcal{H})$.

2.1.2 The notion of Tree Soft Technique in Association Rule

various studies as [21] described association rule as a technique deployed for identifying intriguing and recurring patterns in transactional. Therefore, the objective of this technique [22] is discovering recurrent patterns, relationships, correlations, or chains of causality among sets of items in transactional and relational databases where called item sets[23].Suppose that number of transactions [24]expressed as $\mathrm{D} = \{\mathrm{T}_1, \mathrm{T}_2, \ldots \mathrm{T}_\mathrm{n}\}$ which are including items are symbolled as $\mathrm{I} = \{\mathrm{i}_1, \mathrm{i}_2, \ldots \mathrm{i}_\mathrm{n}\}$.Generally, association rule consists of two parts are antecedent leads to consequent of the rule. Also, there are important factors that should take not considerations:(i) support ($\wp$)is characterized as proportion of transactions $\mathrm{T}_\mathrm{i} \in \mathrm{D}$ and $\wp \subseteq \mathrm{T}_\mathrm{i}$.(ii) Let $\wp \rightarrow \vartheta$ and confidence of ($\wp \rightarrow \vartheta$)refers to support $(\wp \cup \vartheta)/\mathrm{support}\,(\wp)$.(iii)moreover, lift indicated to confidence$(\wp \rightarrow \vartheta)/\mathrm{support}\,(\wp)$.

Herein, we will gain from TrST by deploying it in the association rule technique especially, multi-level as a type of association rule [25] to discover relationships between attributes and sub-attributes that are resident in nodes and sub-nodes in form of tree.

Example 1: suppose that the universal set is U consists of elements of houses as U={h_1, h_2,h_3,h_4,h_5,h_6} as in node 0 as in Fig 2.

- Nodes 1,2,4 are main attributes which are inherent of Node 0.
- Node 1 has sub nodes (sub- attributes) are N_{1-1},..N_{1-n}.Also, Node 2 has N_{2-1},..N_{2-n} . Similarly, Node 3 has N_{3-1},..N_{3-n} .
- Considering the mapping between attributes/nodes and powerset of U as: F(Ns) → P(U).
- Table 2 is representing the mapping based on Boolean-valued[26].

$$(F,N) = \begin{cases} \{N_{1-1}, N_{2-2}, N_{3-2}\} \to \{h_1, h_5\} \\ \{N_{1-2}, N_{2-3}, N_{3-1}\} \to \{h_2, h_6\} \\ \{N_{1-3}, N_{2-1}, N_{3-3}\} \to \{h_{3,}\ h_4\} \end{cases}$$

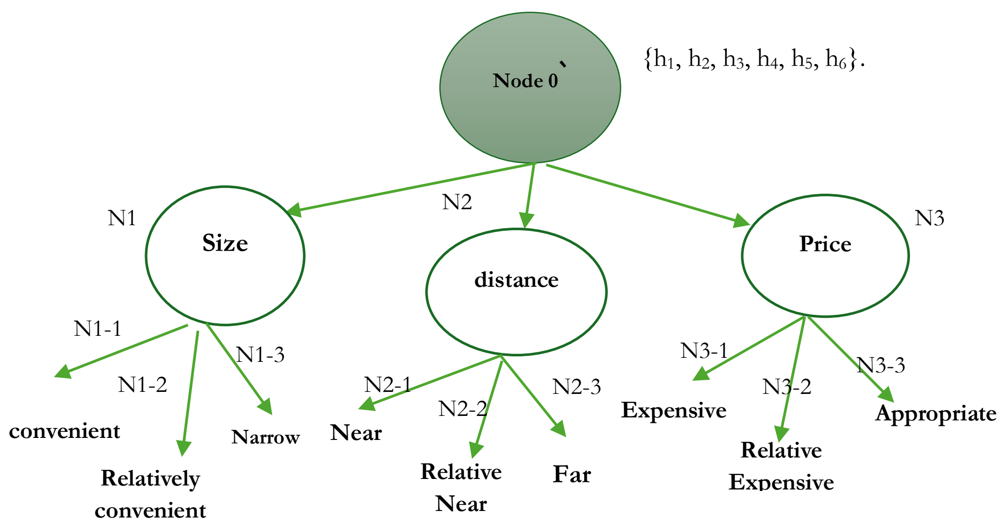

Fig 2. Mapping attributes and sub-attributes in Tree Soft Technique

Table 2. Mapping based on Boolean-valued.

U / Sub-Nodes	N_{1-1}	N_{1-2}	N_{1-3}	N_{2-1}	N_{2-2}	N_{2-3}	N_{3-1}	N_{3-2}	N_{3-3}
h_1	1	0	0	0	1	0	0	1	0
h_2	0	1	0	0	0	1	1	0	0
h_3	0	0	1	1	0	0	0	0	1
h_4	0	0	1	1	0	0	0	0	1

h_5	1	0	0	0	1	0	0	1	0
h_6	0	1	0	0	0	1	1	0	0

2.2 Methodology of Extracting Knowledge Toward CVDs Diagnosis

Herein, we are achieving the study's objective through constructing soft diagnostic paradigm. Thereby, this section is illustrating the needed steps which contributing to construct paradigm as following:

2.2.1Forming the most influencing heart diseases' attributes based on TrST

- The group of patients as $p = \{p_1, p_2, \dots p_n\}$ are volunteered in this study to detect CVDs.
- The most influencing heart diseases' attributes are determined for group of patients through utilizing Cleveland heart disease dataset from the UCI repository.
- Forming the main attributes into initial nodes moreover, the sub-attributes are represented into sub-nodes.
- The group of decision makers (DMs) who are related to the medical field is formed to evaluate the patient's medical condition.

2.2.2Weighting TrST's attributes of heart diseases

DMs are evaluating patient's medical condition through analyzing attributes and sub-attributes encoded in TrST which related to heart diseases. The evaluation process is conducted through harnessing Single Value Neutrosophic Sets(SVNSs) as branch of uncertainty technique for evaluating patient's medical condition. SVNSs are leveraged as guide for DMs during evaluation process in hazy situations. The purpose of this phase is generating weights for encoded attributes in TrST. Hence, SVNSs based entropy are utilized for generating weights through following steps:

- Various Neutrosophic decision matrices are constructed based on evaluation of each DM.
- Eq.(1) utilized in constructed neutrosophic decision matrices for transforming these matrices into crisp matrices.

$$\boldsymbol{s}(\partial_{\mathrm{ij}}) = \frac{(2+g-q-\delta)}{3} \qquad (1)$$

Where:

g, q, δ refers to truth, false, and indeterminacy respectively.

- Crisp matrices are aggregated into an aggregated decision matrix based on Eq.(2).

$$Q_{ij} = \frac{(\sum_{\mathrm{j=1}}^{\mathrm{N}} \partial_{\mathrm{ij}})}{T} \qquad (2)$$

Where:

∂_{ij} refers to value of criterion in matrix, T refers to number of decision makers.

- Eq.(3) employed for normalizing the aggregated matrix

$$D_{ij}=\frac{Q_{ij}}{\sum_{j=1}^{n} Q_{ij}} \tag{3}$$

Where:

$\sum_{j=1}^{n} Q_{ij}$ represents sum of each criterion in aggregated matrix per column

- normalized matrix computes its entropy by Eq. (4).

$$e_{j}=-h\sum_{i=1}^{n} D_{ij} \ln D_{ij} \tag{4}$$

Where:

$$h=\frac{1}{\ln(Ps)} \tag{5}$$

Ps refers to number of alternatives of patients.

- weight vectors are generated through Eq.(6).

$$w_{j}=\frac{1-e_j}{\sum_{j=1}^{n}(1-e_j)} \tag{6}$$

2.2.3 Technique for order of preference by similarity to ideal solution (TOPSIS)

Through collaborating TrST with TOPSIS based on SVNSs, a set of sub-nodes or sub-attributes are inherent in 9 of the main attributes to diagnose the patients. Hence, sub-attributes (N_{n-m})={N_{1-3},N_{2-3},N_{3-2},N_{4-2},N_{5-3},N_{6-1},N_{7-2},N_{8-3},N_{9-3}} are employed in this step to detect CVDs for patients through implementing the following steps.

- The Neutrosophic decision matrices are constructed and deploying Eqs.(1),(2) to deneutrosphic matrices and aggregated it into an aggregated matrix
- The aggregated decision matrix is normalized according to based on the following Eq.(7).

$$\aleph_{ij}=\frac{x_{ij}}{\sqrt{\sum_{j=1}^{m}(x^2{}_{ij})}} \tag{7}$$

- the weighted decision matrix is generalized through Eq. (8).

$$wz_{ij}=weight_j * \aleph_{ij} \tag{8}$$

- positive ideal solution and negative ideal solution are computing based on Eqs (9,10) respectively.

$$ð^{*}=(wz_1{}^{*},wz_2{}^{*},\dots.wz_n{}^{*}),wz_j{}^{*}=max_i\{wz_{ij}\} \tag{9}$$

$$\sigma^{-}=(wz_1{}^{-},wz_2{}^{-},\dots.wz_n{}^{-}),wz_j{}^{-}=min_i\{wz_{ij}\} \tag{10}$$

Where:

$wz_1{}^* \dots wz_n{}^*, wz_1{}^- \dots wz_n{}^-$ are max and min values of weighted normalized criteria per column respectively.

- Eqs (11,12) deployed for computing the distance between the positive ideal solution and negative ideal solution to each patient.

$$d_i{}^* = \sum_{j=1}^{n} d(\ wz_{ij}, wz_j{}^*) \tag{11}$$

$$d_i{}^- = \sum_{j=1}^{n} d(\ wz_{ij}, wz_j{}^-) \tag{12}$$

- The diagnosing for patients is determining based on the values of $\mathrm{CC_i}$ in Eq.(13). After that Eq.(14) deployed to determine heart patients [27].

$$CC_i = \frac{d_i{}^-}{d_i{}^* + d_i{}^-} \tag{13}$$

$$\text{diagnosis} = \begin{cases} \mathrm{CVD} = 0, & \mathrm{CC_i} < 0.5 \\ \mathrm{CVD} = 1, & \mathrm{CC_i} \geq 0.5 \end{cases} \tag{14}$$

2.2.4 Collaboration of TrST -Association rules through applying various cases for extracting the relationships between sub-attributes which resulting in decision of diagnosis

3. Application of Paradigm

In this section, we implement the constructed soft diagnostic paradigm in case study.

3.1 Problem Description

According to Cleveland Heart Disease, which is sourced from the UCI repository, we volunteered attributes and its descriptions in our study. Fig 3 illustrates the utilized attributes encoded into TrST for diagnosis CVDs. The process of diagnosing is conducted by volunteering five patients and evaluating their medical conditions based on nine attributes and 27 sub-attributes in Fig 3.

3.2 SVNSs based Entropy

The attributes' weights are generated through utilizing entropy with the support of SVNSs as following:

- Three Neutrosophic matrices are constructed for each DM based on SVN scale listed in Table 3.
- The constructed matrices transformed to crisp matrices based on Eq.(1) and aggregated based on Eq.(2) into an aggregated decision matrix as in Table 4.
- Eq.(3) utilized in aggregated decision matrix to generate normalized matrix as in Table 5.

- After employing Eq.(4) to compute entropy, the attributes' (nodes') weights are obtained based on Eq.(6) and illustrated in Fig 4. Based on this Fig, attribute 9 is superior to other attributes In contrast to attribute 4.

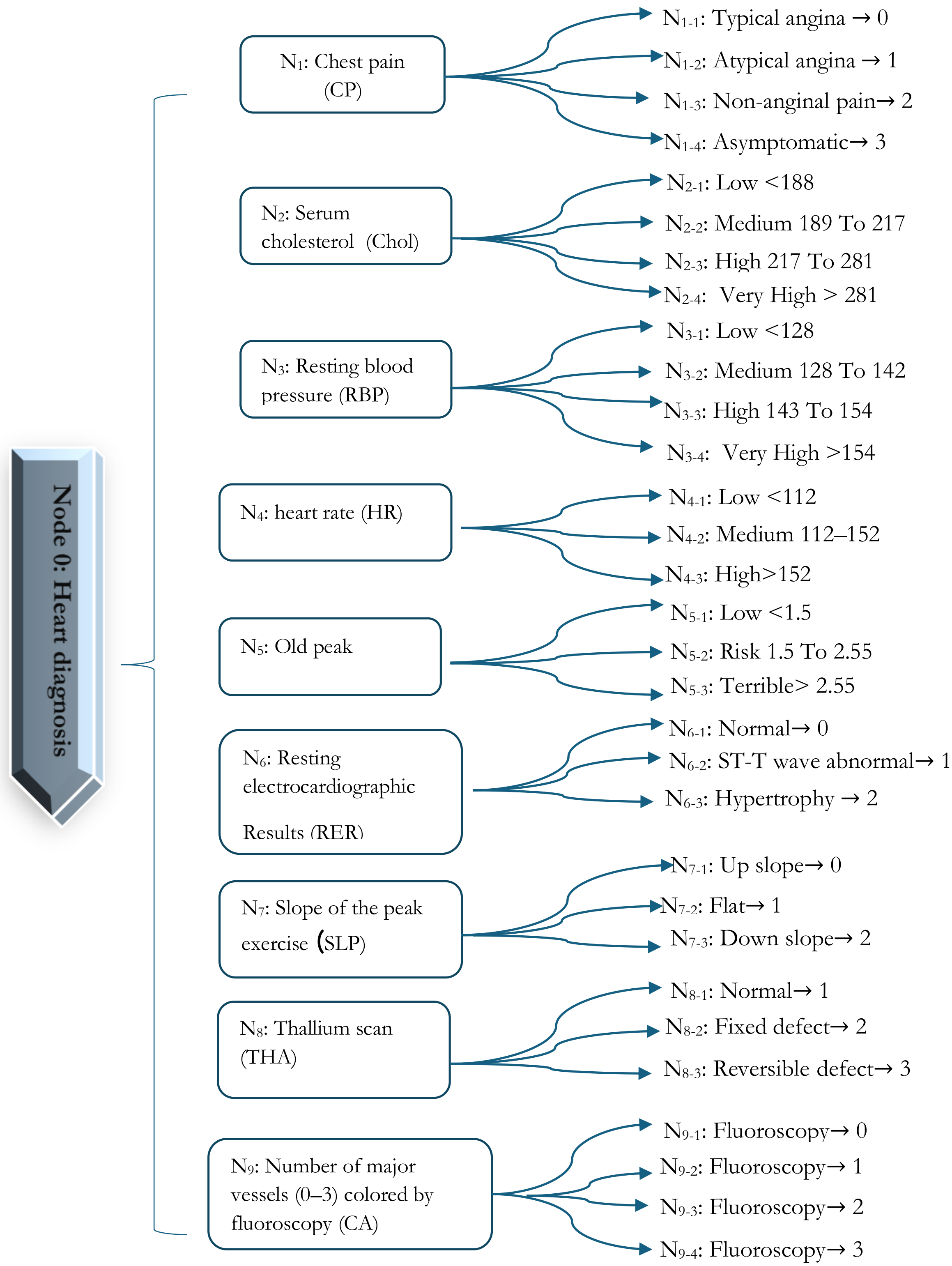

Fig 3. Attributes of Cleveland heart disease encoded into Tree Soft Tree

Table 3. SVN Scale

Synonmy	Acronym	Scale		
		T	I	F
Extremly Weak	EW	0.00	1.00	1.00
Absolutely Weak	AW	0.10	0.90	0.90
Very Weak	VW	0.20	0.85	0.80
Weak	W	0.30	0.75	0.70
Fairly Weak	FW	0.40	0.65	0.60
Fairly	F	0.50	0.50	0.50
Fairly Well	FW	0.60	0.35	0.40
Well	W	0.70	0.25	0.30
Very Well	VW	0.80	0.15	0.20
Absolutely Well	AW	0.90	0.10	0.10
Extremly Well	EW	1.00	0.00	0.00

	N_1	N_2	N_3	N_4	N_5	N_6	N_7	N_8	N_9
P_1	0.7467	0.78	0.7067	0.6133	0.7533	0.79	0.58	0.6133	0.46
P_2	0.6533	0.78	0.787	0.54	0.747	0.653	0.54	0.42	0.82
P_3	0.607	0.66	0.507	0.7	0.4267	0.5	0.82	0.607	0.46
P_4	0.38	0.3533	0.46	0.46	0.58	0.75	0.347	0.58	0.313
P_5	0.46	0.42	0.393	0.5067	0.3133	0.813	0.82	0.78	0.3533

Table 4. Aggregated decision Matrix

Table 5. Normalized decision Matrix

	N1	**N2**	**N3**	**N4**	**N5**	**N6**	**N7**	**N8**	**N9**
P1	0.262295082	0.2605791	0.24766355	0.21749409	0.2671395	0.224762	0.186695	0.204444	0.191136
P2	0.229508197	0.2605791	0.27570093	0.191489362	0.2647754	0.186667	0.17382	0.14	0.34072
P3	0.213114754	0.22049	0.17757009	0.24822695	0.1513002	0.142857	0.263948	0.202222	0.191136
P4	0.133489461	0.1180401	0.16121495	0.163120567	0.2056738	0.213333	0.111588	0.193333	0.130194
P5	0.161592506	0.1403118	0.13785047	0.179669031	0.1111111	0.232381	0.263948	0.26	0.146814

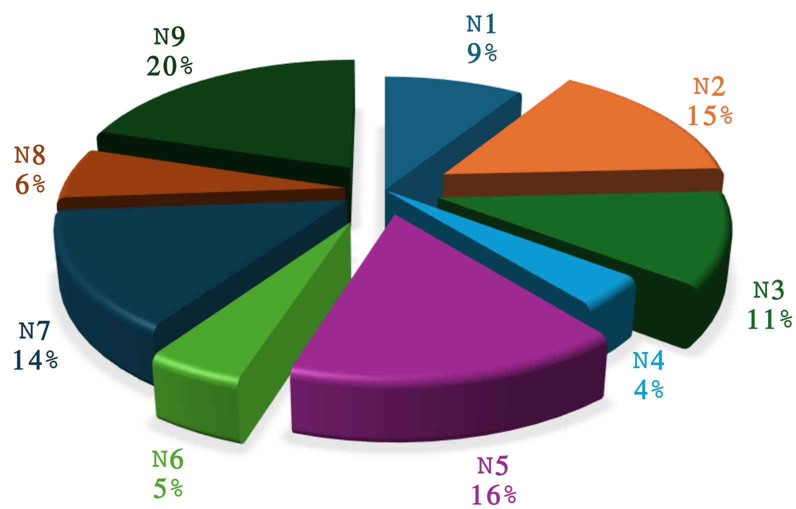

Fig 4 weighs of attributes/nodes encoded into Tree Soft Technique

3.3 Toward Diagnosis: SVNSs based TOPSIS

Herein, the heart patient is detecting based on implementing TOPSIS under authority of SVNSs as in the following steps.

- The aggregated matrix is obtained as in Table 6 after employing Eq.s (1,2) to deneutrosophic the matrices and aggregated it into single matrix.
- The aggregated matrix normalized through Eq.(7) as listed in Table 7.
- -the normalized matrix leveraged to construct weighted decision matrix in Table 8 based on Eq.(8).
- According to Eq.s(11,12) the distance between positive and negative ideal solution is calculated and consider important step for obtaining CC_i as obtained in Table 9.

- According to Values of CC_i in Table 9 which contributed to Eq.(14) for diagnosing the patient condition where P1,P4 classify absence of heart disease whereas P2,P3,P5 classify to presence of heart disease.

Table 6. Aggregated matrix

	N_{1-3}	N_{2-3}	N_{3-2}	N_{4-2}	N_{5-3}	N_{6-1}	N_{7-2}	N_{8-3}	N_{9-3}
P_1	0.845556	0.54	0.52	0.31333	0.53333	0.53333	0.39333	0.72	0.813333
P_2	0.5	0.74667	0.78	0.8333	0.9	0.62	0.667	0.94	0.553333
P_3	0.39333	0.78	0.87333	0.54	0.54	0.81333	0.70667	0.593333	0.54
P_4	0.72	0.5	0.5333	0.52	0.43333	0.526667	0.39333	0.5	0.653333
P_5	0.42667	0.6867	0.84	0.7533	0.43333	0.81333	0.71333	0.7466	0.346667

Table 7. Normalized matrix

	N_{1-3}	N_{2-3}	N_{3-2}	N_{4-2}	N_{5-3}	N_{6-1}	N_{7-2}	N_{8-3}	N_{9-3}
P_1	0.626746859	0.365816197	0.320551	0.22600333	0.401832516	0.354044692	0.296310367	0.449863344	0.605163
P_2	0.370612466	0.505819926	0.4808264	0.60107269	0.678092372	0.411576954	0.502220961	0.587321587	0.411709
P_3	0.291548473	0.528401173	0.5383612	0.38949511	0.406855423	0.539918155	0.532354218	0.370720718	0.401788
P_4	0.533681951	0.338718701	0.3287702	0.37506936	0.32648892	0.349619133	0.296310367	0.3124051	0.486114
P_5	0.316255971	0.465173682	0.5178131	0.54336972	0.32648892	0.539918155	0.537376428	0.466524949	0.257938

Table 8. Weighted decision Matrix

	N_{1-3}	N_{2-3}	N_{3-2}	N_{4-2}	N_{5-3}	N_{6-1}	N_{7-2}	N_{8-3}	N_{9-3}
P_1	0.117162978	0.009118738	0.05824655	0.038864036	0.014126615	0.034807326	0.043939085	0.027528672	0.055381574
P_2	0.069281656	0.012608625	0.087369824	0.103361798	0.023838664	0.040463516	0.074473025	0.035940211	0.037677628
P_3	0.054501569	0.01317151	0.097824333	0.066978445	0.014303198	0.053081171	0.078941406	0.022685665	0.036769734
P_4	0.099765585	0.008443276	0.059740051	0.064497762	0.011477875	0.034372234	0.043939085	0.019117133	0.044486838
P_5	0.059120346	0.011595432	0.09409058	0.093439066	0.011477875	0.053081171	0.079686136	0.028548253	0.023605261

Table 9. Diagnosing the patient based on **CCi**

	d*	d-	Cci	Diagnosing
P_1	0.086412	0.065964	0.432905	0
P_2	0.053804	0.081547	0.602488	1
P_3	0.076572	0.063048	0.52	1
P_4	0.07359	0.052609	0.416872	0
P_5	0.068568	0.078642	0.534218	1

4. Analysis and discussion

Herein, we discuss the proposed paradigm's findings. Thus, this section divides into two sub-sections. Generally speaking, in the constructed paradigm there are set of steps have been conducted for diagnosing the heart conditions for various patients. Hence, the paradigm's notion is based on soft set and uncertainty techniques; for bolstering the constructed paradigm in diagnosing vague patients' cases. Moreover, TrST and neutrosophic are collaborating with entropy and TOPSIS in diagnosing the cases.

4.1 Soft Diagnosis Paradigm's Findings

The constructed paradigm generated group of finding:

- Firstly, implementing SVNSs based entropy in attributes and sub-attributes encoded into Tree Soft which illustrated in Fig 3 to generate attributes' weights indicated that attribute 9 (CA) is superior to other attributes and attribute 4 (HR) is worst as mentioned in Fig 4.
- Secondly, SVNSs based TOPSIS are working in sub-attributes which encoded into Tree Soft to diagnosing five patients over the determined sub-attributes/sub-nodes.
- The sub-attributes are determining and nominating through TrST where $F: N_1 \times N_2 \times N_3 \times N_4 \times N_5 \times N_6 \times N_7 \times N_8 \times N_9 \rightarrow P(P)$.Hence, sub-attributes ($N_{n-m}$) ={$N_{1-3}$,$N_{2-3}$,$N_{3-2}$,$N_{4-2}$,$N_{5-3}$,$N_{6-1}$,$N_{7-2}$,$N_{8-3}$,$N_{9-3}$} are employed in SVNSs based TOPSIS doe detecting and diagnosing five patients . The findings indicated that two patients of P1, P4 classify absence of heart disease whereas P2, P3, P5 classify to presence of heart disease as listed in Table 9.

4.2 Association Rule and Soft Paradigm Collaboration Findings

The objective of association rules is to extract and discover the relation between sub-attributes/sub-nodes in Fig 3 as antecedent to diagnose the patient cases as consequent. Thereby, we implemented various cases which consider our problem as transactions for obtaining relation between sun-nodes/sub-attributes which consider items toward diagnosing patients. Thus, D={ T_1,T_2,T_{3-},T_4 } is transactions for items of sub-attributes I={ I_{1-n},I_{2-n},I_{3-n-},I_{4-n},I_{5-n},I_{6-n},I_{7-n-},I_{8-n}, I_{9-n} } .

4.2.1 Case 2 (T_2): Let sub-attributes (N_{n-m}) ={N_{1-2},N_{2-1},N_{3-1},N_{4-1},N_{5-2},N_{6-2},N_{7-3},N_{8-2},N_{9-1}} are employed in this step to detect CVDs.

- According to values of $\mathrm{CC_i}$ in Table 10 which contributed to Eq.(14) for diagnosing the patient condition where P4,P5 classify absence of heart disease whereas P2,P3,P1 classify to presence of heart disease.

Table 10. Diagnosing the patient based on **CCi** in case 2

	d*	**d-**	**Cci**	**Diagnosing**
P_1	0.051868007	0.064332043	0.55	1
P_2	0.053525933	0.059938779	0.53	1
P_3	0.049792455	0.060048848	0.54	1
P_4	0.076682931	0.029550656	0.3	0
P_5	0.059870647	0.05437008	0.4	0

4.2.2 Case 3 (T_3): Let sub-attributes ($N_{n\text{-}m}$) ={$N_{1\text{-}3}$,$N_{2\text{-}1}$,$N_{3\text{-}3}$,$N_{4\text{-}3}$,$N_{5\text{-}1}$,$N_{6\text{-}1}$,$N_{7\text{-}3}$,$N_{8\text{-}1}$,$N_{9\text{-}1}$} are employed in this step to detect CVDs.

- According to values of CC_i in Table 11 which contributed to Eq.(14) for diagnosing the patient condition where P4,P5 classify absence of heart disease whereas P2,P3,P1 classify to presence of heart disease.

Table 11. Diagnosing the patient based on **CCi** in case 3

	d*	**d-**	**Cci**	**Diagnosing**
P_1	0.127123265	0.111984744	0.5	1
P_2	0.052719105	0.187088011	0.78	1
P_3	0.013877717	0.237815505	0.95	1
P_4	0.237783628	0.008110786	0.033	0
P_5	0.21115882	0.028586828	0.12	0

4.2.3 Case 4 (T_4): Let sub-attributes ($N_{n\text{-}m}$) ={$N_{1\text{-}1}$,$N_{2\text{-}3}$,$N_{3\text{-}1}$,$N_{4\text{-}3}$,$N_{5\text{-}1}$,$N_{6\text{-}3}$,$N_{7\text{-}1}$,$N_{8\text{-}1}$,$N_{9\text{-}2}$} are employed in this step to detect CVDs.

- According to values of CC_i in Table 12 which contributed to Eq.(14) for diagnosing the patient condition where P4,P5, P1 classify absence of heart disease whereas P2,P3 classify to presence of heart disease

Table 12. Diagnosing the patient based on **CCi** in case 4

	d*	**d-**	**Cci**	**Diagnosing**
P_1	0.92568692	0.372601434	0.287	0
P_2	0.494731035	0.891295757	0.643	1
P_3	0.668434479	0.801959546	0.545	1
P_4	1.064228169	0.08716001	0.0758	0
P_5	0.955166414	0.3815659	0.2855	0

Finally, Table 13 aggregated the set of transactions for set of items for discover and diagnose the patients according to Boolean value[26].Based on transaction ID(TID)in Table 13, support and confidence for items are computing as listed in Table 14.

$$(\mathrm{F},\mathrm{N}) = \begin{cases} \{\mathrm{N}_{1-3}, \mathrm{N}_{2-3}, \mathrm{N}_{3-2}, \mathrm{N}_{4-2}, \mathrm{N}_{5-3}, \mathrm{N}_{6-1}, \mathrm{N}_{7-2}, \mathrm{N}_{8-3}, \mathrm{N}_{9-3}\} \rightarrow \{(\mathrm{p}_2, \mathrm{p}_3, \mathrm{p}_5) = 1\} \\ \{\mathrm{N}_{1-2}, \mathrm{N}_{2-1}, \mathrm{N}_{3-1}, \mathrm{N}_{4-1}, \mathrm{N}_{5-2}, \mathrm{N}_{6-2}, \mathrm{N}_{7-3}, \mathrm{N}_{8-2}, \mathrm{N}_{9-1}\} \rightarrow \{(\mathrm{p}_2, \mathrm{p}_3, \mathrm{p}_1) = 1\} \\ \{\mathrm{N}_{1-3}, \mathrm{N}_{2-1}, \mathrm{N}_{3-3}, \mathrm{N}_{4-3}, \mathrm{N}_{5-1}, \mathrm{N}_{6-1}, \mathrm{N}_{7-3}, \mathrm{N}_{8-1}, \mathrm{N}_{9-1}\} \rightarrow \{(\mathrm{p}_2, \mathrm{p}_3, \mathrm{p}_1) = 1\} \\ \{\mathrm{N}_{1-1}, \mathrm{N}_{2-3}, \mathrm{N}_{3-1}, \mathrm{N}_{4-3}, \mathrm{N}_{5-1}, \mathrm{N}_{6-3}, \mathrm{N}_{7-1}, \mathrm{N}_{8-1}, \mathrm{N}_{9-2}\} \rightarrow \{(\mathrm{p}_2, \mathrm{p}_3) = 1\} \end{cases}$$

Table 13. Total Transactions (D) based on items (I_n)

TID	Items (sub-nodes)																									
	N_{1-1}	**N_{1-2}**	**N_{1-3}**	**N_{2-1}**	**N_{2-3}**	**N_{3-1}**	**N_{3-2}**	**N_{3-3}**	**N_{4-1}**	**N_{4-2}**	**N_{4-3}**	**N_{5-1}**	**N_{5-2}**	**N_{5-3}**	**N_{6-1}**	**N_{6-2}**	**N_{6-3}**	**N_{7-1}**	**N_{7-2}**	**N_{7-3}**	**N_{8-1}**	**N_{8-2}**	**N_{8-3}**	**N_{9-1}**	**N_{9-2}**	**N_{9-3}**
T_1	0	0	1	0	1	0	0	0	0	1	0	0	0	1	1	0	0	0	1	0	0	0	1	0	0	1
T_2	0	1	0	1	0	1	0	0	1	0	0	0	1	0	0	1	0	0	0	1	0	1	0	1	0	0
T_3	0	0	1	1	0	0	0	1	0	0	1	1	0	0	1	0	0	0	0	1	1	0	0	1	0	0
T_4	1	0	0	0	1	1	0	0	0	0	1	1	0	0	0	0	1	1	0	0	1	0	0	0	1	0

Table 14. Support and confidence of items

Items	Support	Confidence
N_{1-1}	1	¼=25%
N_{1-2}	1	¼=25%
N_{1-3}	2	2/4=50%
N_{2-1}	2	2/4=50%
N_{2-3}	2	2/4=50%
N_{3-1}	1	¼=25%
N_{3-2}	1	¼=25%
N_{3-3}	1	¼=25%
N_{4-1}	1	¼=25%
N_{4-2}	1	¼=25%
N_{4-3}	2	2/4=50%
N_{5-1}	2	2/4=50%
N_{5-2}	1	¼=25%
N_{5-3}	1	¼=25%
N_{6-1}	1	¼=25%
N_{6-2}	1	¼=25%
N_{6-3}	1	¼=25%
N_{7-1}	1	¼=25%
N_{7-2}	1	¼=25%
N_{7-3}	2	2/4=50%
N_{8-1}	2	2/4=50%
N_{8-2}	1	¼=25%
N_{8-3}	1	¼=25%
N_{9-1}	2	2/4=50%
N_{9-2}	1	¼=25%
N_{9-3}	1	¼=25%
$\mathbf{N_{1-3}, N_{2-3}, N_{3-2}, N_{4-2}, N_{5-3}, N_{6-1}, N_{7-2}, N_{8-3}, N_{9-3}}$	1	¼=25%
$\mathbf{N_{1-2}, N_{2-1}, N_{3-1}, N_{4-1}, N_{5-2}, N_{6-2}, N_{7-3}, N_{8-2}, N_{9-1}}$	1	¼=25%
$\mathbf{N_{1-3}, N_{2-1}, N_{3-3}, N_{4-3}, N_{5-1}, N_{6-1}, N_{7-3}, N_{8-1}, N_{9-1}}$	1	¼=25%
$\mathbf{N_{1-1}, N_{2-3}, N_{3-1}, N_{4-3}, N_{5-1}, N_{6-3}, N_{7-1}, N_{8-1}, N_{9-2}}$	1	¼=25%

5. Conclusion

Cardiovascular disease (CVD) is one of the leading causes of death worldwide. It affects not only the heart and blood vessels but also heart failure, blood vessel disorders, stroke, arrhythmia, and myocardial infarction. To intervene with the patient promptly, it is essential to identify the critical risk factors.

Hence, the process of early detecting and diagnosing this disease is important. Scientifically, there are many studies that embraced the process of diagnosing CVDs. Moreover, various techniques and models are suggested for serving this process. Therefore, herein we attempted to treat such problem through constructing soft diagnosis paradigm. For constructing this paradigm, we are harassed various techniques each one responsible for vital role. Firstly,
TrST is employed in this problem for the first time where this technique is based on the notion of soft sets. We leveraged this technique for representing attributes and sub-attributes of patients into Tree soft form. Secondly, SVNSs based entropy are implementing in the formed Tree for analyzing the main nodes of attributes and generating weights for it to showcase the most influenced attribute and the least influenced attribute.

Thirdly, SVNSs based TOPSIS is responsible for detecting the diagnosis for five patients based on the value of CC_i and according to Eq.(14).The findings of soft diagnosis paradigm indicated that P2,P3,P5 are belongs to presence of heart disease otherwise, P1,P4 belongs to absence of heart disease. Fourthly, we exploited the notion of soft set in TrST through collaborating TrST with association rule the relationship between sub-attributes through applying set of transactions (cases) for exhibiting relationships of sub-attributes(items) which considering antecedent that lead to consequent of detecting diagnosis of medical conditions of patients as mentioned in Table 13 and Table 14.

Reference

[1] Stephen, M., & Felix, A. (2023). Fuzzy AHP point factored inference system for detection of cardiovascular disease. *Journal of intelligent & fuzzy systems*, *44*(4), 6655–6684.

[2] Stajić, D., & Đonović, N. (2016). Cardiovascular diseases – risk factors. *Medicinski casopis*, *50*(2), 43–48. DOI:10.5937/mckg50-11761

[3] Gaidai, O., Cao, Y., & Loginov, S. (2023). Global Cardiovascular Diseases Death Rate Prediction. *Current problems in cardiology*, *48*(5), 101622. DOI:10.1016/j.cpcardiol.2023.101622

[4] Swathy, M., & Saruladha, K. (2022). A comparative study of classification and prediction of Cardio-Vascular Diseases (CVD) using Machine Learning and Deep Learning techniques. *ICT express*, *8*(1), 109–116.

[5] Hassan, D., Hussein, H. I., & Hassan, M. M. (2023). Heart disease prediction based on pre-trained deep neural networks combined with principal component analysis. *Biomedical signal processing and control*, *79*, 104019.

[6] Khozeimeh, F., Sharifrazi, D., Izadi, N. H., Joloudari, J. H., Shoeibi, A., Alizadehsani, R., … Islam, S. M. S. (2022). RF-CNN-F: random forest with convolutional neural network features for coronary artery disease diagnosis based on cardiac magnetic resonance. *Scientific reports*, *12*(1), 1–12. DOI:10.1038/s41598-022-15374-5

[7] Rustam, F., Ishaq, A., Munir, K., Almutairi, M., Aslam, N., & Ashraf, I. (2022). Incorporating cnn features for optimizing performance of ensemble classifier for cardiovascular disease prediction. *Diagnostics*, *12*(6), 1474.

[8] Wei, X., Rao, C., Xiao, X., Chen, L., & Goh, M. (2023). Risk assessment of cardiovascular disease based on SOLSSA-CatBoost model. *Expert systems with applications*, *219*(February), 119648. DOI:10.1016/j.eswa.2023.119648

[9] Azmi, J., Arif, M., Nafis, M. T., Alam, M. A., Tanweer, S., & Wang, G. (2022). A systematic review on machine learning approaches for cardiovascular disease prediction using medical big data. *Medical engineering & physics*, *105*, 103825.

[10] Ashraf, M., Salal, Y. K., Abdullaev, S. M., Zaman, M., & Bhut, M. A. (2022). Introduction of feature selection and leading-edge technologies viz. tensorflow, pytorch, and keras: an empirical study to improve prediction accuracy of cardiovascular disease. *International conference on innovative computing and communications: proceedings of icicc 2021, volume 3* (pp. 19–31). Springer.

[11] Ali, L., Niamat, A., Khan, J. A., Golilarz, N. A., Xingzhong, X., Noor, A., … Bukhari, S. A. C. (2019). An optimized stacked support vector machines based expert system for the effective prediction of heart failure. *IEEE access*, *7*, 54007–54014.

[12] Arumugam, K., Naved, M., Shinde, P. P., Leiva-Chauca, O., Huaman-Osorio, A., & Gonzales-Yanac, T. (2023). Multiple disease prediction using Machine learning algorithms. *Materials today: proceedings*, *80*, 3682–3685.

[13] Budholiya, K., Shrivastava, S. K., & Sharma, V. (2022). An optimized XGBoost based diagnostic system for effective prediction of heart disease. *Journal of king saud university-computer and information sciences*, *34*(7), 4514–4523.

[14] Alghazzawi, D., Liaqat, M., Razaq, A., Alolaiyan, H., Shuaib, U., & Liu, J. B. (2023). Selection of Optimal Approach for Cardiovascular Disease Diagnosis under Complex Intuitionistic Fuzzy Dynamic Environment. *Mathematics*, *11*(22). DOI:10.3390/math11224616

[15] Hossain, M. E., Khan, A., Moni, M. A., & Uddin, S. (2019). Use of electronic health data for disease prediction: A comprehensive literature review. *IEEE/acm transactions on computational biology and bioinformatics*, *18*(2), 745–758.

[16] Hussein, G. S., Zaied, A. N. H., & Mohamed, M. (2023). ADM: Appraiser Decision Model for Empowering Industry 5.0-Driven Manufacturers toward Sustainability and Optimization: A Case Study. *Neutrosophic systems with applications*, *11*, 22–30.

[17] Balamurugan, R., Ratheesh, S., & Venila, Y. M. (2022). Classification of heart disease using adaptive Harris hawk optimization-based clustering algorithm and enhanced deep genetic algorithm. *Soft computing*, *26*(5), 2357–2373. DOI:10.1007/s00500-021-06536-0

[18] Mariadoss, S., & Augustin, F. (2023). Enhanced sugeno fuzzy inference system with fuzzy

AHP and coefficient of variation to diagnose cardiovascular disease during pregnancy. *Journal of king saud university-computer and information sciences*, *35*(8), 101659.

[19] Smarandache, F. (2023). New Types of Soft Sets" HyperSoft Set, IndetermSoft Set, IndetermHyperSoft Set, and TreeSoft Set": An Improved Version. Neutrosophic Systems With Applications, 8, 35-41. https://doi.org/10.61356/j.nswa.2023.41.

[20] AL-baker, S. F., El-henawy, I., & Mohamed, M. (2024). Pairing New Approach of Tree Soft with MCDM Techniques: Toward Advisory an Outstanding Web Service Provider Based on QoS Levels. *Neutrosophic systems with applications*, *14*, 17–29.

[21] Amoozad Mahdiraji, H., Tavana, M., Mahdiani, P., & Abbasi Kamardi, A. A. (2022). A multi-attribute data mining model for rule extraction and service operations benchmarking. *Benchmarking: an international journal*, *29*(2), 456–495.

[22] Buxton, E. K., Vohra, S., Guo, Y., Fogleman, A., & Patel, R. (2019). Pediatric population health analysis of southern and central Illinois region: A cross sectional retrospective study using association rule mining and multiple logistic regression. *Computer methods and programs in biomedicine*, *178*, 145–153.

[23] Shah, J. J., & Gagnani, L. P. (2015). Introduction to Multilevel Association Rule and Its Methods, *3*(10), 531–534.

[24] Sarıyer, G., & Öcal Taşar, C. (2020). Highlighting the rules between diagnosis types and laboratory diagnostic tests for patients of an emergency department: Use of association rule mining. *Health informatics journal*, *26*(2), 1177–1193. DOI:10.1177/1460458219871135

[25] Hameed, N., Shabut, A. M., Ghosh, M. K., & Hossain, M. A. (2020). Multi-class multi-level classification algorithm for skin lesions classification using machine learning techniques. *Expert systems with applications*, *141*, 112961.

[26] Herawan, T., & Deris, M. M. (2011). A soft set approach for association rules mining. *Knowledge-based systems*, *24*(1), 186–195. DOI:10.1016/j.knosys.2010.08.005

[27] Kishore, A. H. N., & Jayanthi, V. E. (2018). Multi Criteria Decision Making Methods to Predict the Prevalence of Coronary Artery Disease. *Journal of medical imaging and health informatics*, *8*(4), 719–726. DOI:10.1166/jmihi.2018.2357

Received: Feb 5, 2024. Accepted: April 28, 2024