EUROPEAN JOURNAL OF PURE AND APPLIED MATHEMATICS
2025, Vol. 18, Issue 4, Article Number 7168
ISSN 1307-5543 – ejpam.com
Published by New York Business Global

Correlation Measure for Quadri-Partitioned Neutrosophic Refined Sets and Its Application in Medical Diagnosis

A. A. Azzam[1], B. Alreshidi[1], M. Aldawood[1], Mohamed M. Awad[1], Abdelhalim Hasnaoui[2], Medhat Ahmed Abu-Tahon[3], Ahmad M. Abdel-Mageed[3], Arif Mehmood[4,*]

[1] *Mathematics Department, Faculty of Science and Humanities, Prince Sattam Bin Abdulaziz University, Al-Kharj 11942, Saudi Arabia*
[2] *Department of Mathematics, College of Science, Northern Border University, Arar 91431, Saudi Arabia*
[3] *Department of Biological Sciences, College of Science, Northern Border University, Arar 91431, Saudi Arabia*
[4] *Department of Mathematics, Institute of Numerical Sciences, Gomal University, Dera Ismail Khan 29050, KPK, Pakistan*

Abstract. This work introduces a new correlation measure of quadri-partitioned neutrosophic refined sets (QPNRS) to address the problem of ambiguity and bias that may exist in the medical diagnostic. The model is a patient-symptom-disease interaction model that uses four membership functions—truth, relative truth, relative falsehood, and falsehood—to illustrate its interactions. Table III concludes the results by demonstrating that there is clear disease and estimator variance and that correlations between QPNRS measures Q and R are consistently strong and positive (0.743–0.901) across four medical conditions. In particular, P_2 is more extensively distributed, while P_1 is constantly high in all environments, reaching its maximum value of 0.901 in the case of tuberculosis. These findings support the combined clinical use of Q and R by demonstrating a strong, primarily favorable reliance between them. We apply a multifaceted unsupervised learning pipeline to two distinct datasets—the conventional Iris dataset and patient disease-correlation profiles—in order to augment the correlation analysis. In both cases, the consistently strongly separated clusters support the process and imply that the detected groupings are not the result of labeling or modeling artifacts.

2020 Mathematics Subject Classifications: 03E72, 62H30, 68T37

Key Words and Phrases: Quadri-partitioned neutrosophic refined sets (QPNRS), correlations between QPNRS measures (Q and R), patient-symptom-disease interaction model, patient disease-correlation profiles

1. Introduction

Modeling human thought and handling uncertainty in complicated decision-making situations have been one of the main goals of computational intelligence. A paradigm shift occurred when Zadeh [3] created fuzzy sets, and in order to address the vagueness, the idea of partial membership was introduced. This was advanced by Atanassov's intuitionistic fuzzy sets (IFS) [4], which used a more nuanced method of expressing reluctance by adding a measure of non-membership. Additionally, the model was extended to interval-valued IFS [5], which enhances the model's ability to capture imprecise data. The Neutrosophy, as forth by Smarandache [1],

*Corresponding author.
DOI: https://doi.org/10.29020/nybg.ejpam.v18i4.7168

Email addresses: aa.azzam@psau.edu.sa (A. A. Azzam), b.alreshidi@psau.edu.sa (B. Alreshidi), m.aldawood@psau.edu.sa (M. Aldawood), m.abdelgalil@psau.edu.sa (M. M. Awad), abdllhalim.hasanawa@nbu.edu.sa (A. H. Hasnaoui), medhatahon@gmail.com (M. A. Abu-Tahon), ahmed.mohammed@nbu.edu.sa (A. M. Abdel-Mageed), mehdaniyal@gmail.com (A. Mehmood)

is a philosophical foundation for handling the indeterminacy of real-world systems and represents an even more profound generalization of these concepts. A more reliable and flexible way to describe incomplete, inconsistent, and uncertain information is to use neutrophilic sets (NS), which explicitly treat truth-membership (T), indeterminacy-membership (I), and falsity-membership (F) as distinct constituents [14]. Later developments into simpler, more practical forms like interval neutrosophic sets (INS) [12] and single-valued neutrosophic sets (SVNS) [11] have made neutrosophy extremely valuable in data science and engineering. The similarity assessment is one of the most important aspects of these sets' viability. A lot of work has gone into creating a robust similarity metric for neutrosophic situations. A groundbreaking study by Broumi and Smarandache [6] introduced a number of these measures, which Majumdar and Samanta [7] subsequently improved and examined in relation to entropy and similarity. Ye has made contributions in this area by delivering a variety of vector-based [2], interval-based [8], and cosine similarity measures [13] to multicriteria decision-making (MCDM) in interval neutrosophic and single-valued settings, even in cases where the weights have no knowledge at all [9]. These are the cornerstones of modern MCDM techniques, along with other optimization techniques as AHP's cosine maximization method [10]. Additionally, recent advancements have expanded to explore hybrid frameworks, such as the incorporation of neutrosophic sets into plithogenic logic to process multi-dimensional attribute interactions [15] and the application of such models in specific domains, such as supply chain management [17] and medical diagnosis [16].

The usage of sophisticated fuzzy and neutrosophic sets for pattern identification and decision-making has significantly increased, particularly in the medical industry. Researchers have developed more sophisticated metrics that are unique to complex contexts based on the preliminary studies on the similarity of vectors to SVNS [2] and cosines to IFS [9]. In the MCDM problems of uncertainty, for instance, the outranking relation approach proposed by Peng et al. [18] with a novel single-valued neutrosophic similarity measure has demonstrated excellent performance. In the specific instance of medical diagnosis, Abdel-Basset et al. [19] have developed a comprehensive viewpoint on medical diagnosis in a neutrophilic setting that successfully manages ambiguity and inconsistency of symptoms and illnesses, resulting in a more realistic model than traditional fuzzy or IFS approaches. Distance-based tactics were included as a result of the hunt for more effective solutions. A number of generalized similarity metrics on SVNS were introduced by Sahin and Kucuk [20], which significantly improved the discrimination capacity of applications like medical diagnostics. Their work demonstrated that the degree of similarity or distance that is employed has a significant impact on the diagnosis outcomes. Ye [21] introduced the idea of cross-entropy on SVNS, which is an effective way to parametrize the discrimination information that is crucial to distinguishing between identical medical situations, adding to the methodological toolbox. The advancements have significantly improved the way in which the uncertainties inherent in the medical diagnosis process are managed. Combining neutrosophic sets with machine learning methods to carry out validation and enhance pattern recognition is a recent and significant trend. While clustering techniques like K-Means and fuzzy C-Means (FCM) are not new, it is more recent research that has used them to support the findings of neutrosophic similarity metrics. The regular FCM algorithm [22] offered the majority of the foundation for handling uncertainty throughout the clustering process, but it was unable to handle indeterminacy, which served as the inspiration for the neutrosophic variations. Researchers have filled this gap by developing a novel neutrosophic C-means clustering algorithm that incorporates truth, indeterminacy, and falsity into the clustering process. Guo and Sengur [23] have also improved the clustering technique on noisy images. Although there is no obvious connection between this concept and medical diagnoses, it is clear that machine learning and neutrosophic logic techniques work well together. Despite their widespread usage in visualization [24], dimensionality reduction techniques like PCA and t-SNE are now employed to visually validate cluster elements inferred from neutrosophic correlation measurements. This

approach provides realistic, unmonitored validation of theoretical framework-based diagnostic classifications. Additionally, Smarandache [15] postulated the appearance of lithogenic sets, which may offer fresh perspectives. Plithageny provides a clear explanation of the relationship between the qualities' values and conflicts, which makes its use particularly suitable to complex healthcare systems. Plithogenic sets will be able to provide a more fine-grained knowledge representation, which will enable the development of more accurate diagnostic algorithms, according to some of the initial implementations created by Smarandache [25]. With machine learning techniques representing both theory and practice, our suggested QPNRS model, which combines lithogenic logic with neutrosophic principles, can be seen as a significant step toward developing strong correlation measures to medical diagnosis. The investigation of supra open soft sets and its various decompositions and applications, as well as soft set theory, are both superseded by supra soft topology, according to the sources provided. Fundamental results on decompositions of supra soft sets, together with the concept of soft continuity in this context, were provided by El-Sheikh and Abd Ellatif's basic contributions [28]. Supra soft strongly generalized closed sets [30], supra soft strongly extra generalized closed sets [31], and supra soft strongly semi* generalized closed sets [33] are only a few of the specific classes of soft sets that are now being studied in this field. Furthermore, the ideas have been crucial in defining and researching different forms of soft connectedness [35] as well as related soft separation axioms [29][32][34]. The research of supra soft compactness [37] and interactions with soft ideals [36] have also contributed to the development of the theory.

Clinicians can eliminate redundant testing, expedite routine tasks, and focus on the more convenient or informative score without sacrificing accuracy when there is a strong and consistent relationship between Q and R. It is risky to apply a single cross-condition cutoff because the strength of this link varies by measurement and disease. By evaluating these changes, additional disease-specific cutoffs of control may be established, and metrics can be made comparable across locations. In addition to generating hypotheses, quality checks, and early phenotyping before annotation, which is expensive, unsupervised algorithms that employ clean labels where none exist can also detect edge instances and natural clusterings in patient profiles. When the same clustering pipeline is run on a recognized benchmark, like Iris, and clinical data, the workflow gains external credibility and interpretability, making it easier to rely on and repeat. Generally speaking, risk reduction techniques like predictive modeling, feature cutting, and protocol standardization are considerably safer and more reproducible when they exhibit strong QR dependency and stable clusters. This also paves the way for future clinical applications.

We present the first cross, multi-metric analysis of NRS Q and R's movement. Their association across four medical disorders is measured using the three complementing measures (P_1, P_2, *and* P_3). Q and R always have a very strong and positive association (0.743-0.901), albeit this varies based on the condition and the measure, for example. P1 stays high across all conditions, but P_2 varies with disease. We also apply similar patterns to two very different datasets-the classical Iris dataset and patient disease-correlation profiles-to verify that they are not artificial. In both cases, we discover clear, distinct clusters. Combining these findings yields a more trustworthy picture that differentiates between behavior that is stable and behavior that is condition-sensitive, going beyond a single correlation value. This is a clinically significant extra detail that can help determine which metric to prioritize, reduce pointless measurements, and enable better, situation-specific diagnostic regimens.

2. Preliminaries

This section presents the foundational concepts required for the subsequent sections.

Definition 1. *[26]Let $\mathcal{M}$ be a space of points (objects), with a generic element in $\mathcal{M}$ denoted by r. A neutrosophic set $\hat{E}$ in $\mathcal{M}$ is characterized by the truth membership function $T_{\hat{E}}(r)$, an indeterminacy membership function $I_{\hat{E}}(r)$, and a falsity membership function $F_{\hat{E}}(r)$, The*

functions $T_{\hat{E}}(r)$, $I_{\hat{E}}(r)$, and $F_{\hat{E}}(r)$ are real standard or non-standard subsets of $]^{-}0,1^{+}[$; that is $T_{\hat{E}}(r):\mathcal{M}\longrightarrow]^{-}0,1^{+}[$, $I_{\hat{E}}(r):\mathcal{M}\longrightarrow]^{-}0,1^{+}[$ and $F_{\hat{E}}(r):\mathcal{M}\longrightarrow]^{-}0,1^{+}[$, with the condition $^{-}0\leq T_{\hat{E}}(r)+I_{\hat{E}}(r)+F_{\hat{E}}(r)\leq 3^{+}$. This definition of a neutrosophic set is difficult to apply in the real world in scientific and engineering field.

Therefore, the concept of SVNS, which is an instance of a neutrosophic set, was introduced by Wang et al. [27].

Definition 2. *[27]Let $\mathcal{M}$ be a space of points (objects), with a generic element in $\mathcal{M}$ denoted by r. An SVNS $\hat{E}$ in $\mathcal{M}$ is characterized by the truth membership function $T_{\hat{E}}(r)$, an indeterminacy membership function $I_{\hat{E}}(r)$, and a falsity membership function $F_{\hat{E}}(r)$. For each $r\in\mathcal{M}$, there is $T_{\hat{E}}(r)$, $I_{\hat{E}}(r)$, and $F_{\hat{E}}(r)\in[0,1]$ and $0\leq T_{\hat{E}}(r)+I_{\hat{E}}(r)+F_{\hat{E}}(r)\leq 3$. Therefore, an SVNS $\hat{E}$ can be represented by*

$$\hat{E}=\{\langle r,T_{\hat{E}}(r)+I_{\hat{E}}(r)+F_{\hat{E}}(r)\rangle : r\in\mathcal{M}\}$$

Definition 3. *[27]Let $\mathcal{M}$ be a space of points (objects), with a generic element in $\mathcal{M}$ denoted by r. For SVNS $\hat{E}$ and $\hat{G}$. $\hat{E}\subseteq\hat{G}$ if and only if $T_{\hat{E}}(r)\leq T_{\hat{G}}(r)$, $I_{\hat{E}}(r)\geq I_{\hat{G}}(r)$, $F_{\hat{E}}(r)\geq F_{\hat{G}}(r)$ for any $r\in\mathcal{M}$. Also $\hat{E}=\hat{G}$ if and only if $\hat{E}\subseteq\hat{G}$ and $\hat{G}\subseteq\hat{E}$.*

3. Quadri-Partitioned Neutrosophic Refined Sets

This study presents a new set, quadri-partitioned neutrosophic refined sets (QPNRS), which can be used as an extension of neutrosophic refined sets to express uncertainty, indeterminacy, and dual-valued information in complete decision systems. The order and coherence of their structural criteria set apart the four QPNRS membership sequences: truth, relative truth, relative false, and falsity. Basic functions like intersection, equality, union, and subset are defined as the classical set theory is expanded into the neutrosophic refined domain. Additionally, it is proposed that the similarity between QPNRS be measured using a correlation coefficient, which is symmetric, restricted, reflexive, and equal to unity only on sets of identical sets.

Definition 4. *Let $\mathcal{M}$ be a universe. A quadri-partitioned neutrosophic refined set (QPNRs) $\hat{E}$ on $\mathcal{M}$ can be defined as follows:*

$$\hat{E}=\begin{pmatrix}\langle r,(T^{1}_{\hat{E}}(r),T^{2}_{\hat{E}}(r),...,T^{p}_{\hat{E}}(r)),(RT^{1}_{\hat{E}}(r),RT^{2}_{\hat{E}}(r),...,RT^{p}_{\hat{E}}(r)),\\ (RF^{1}_{\hat{E}}(r),RF^{2}_{\hat{E}}(r),...,RF^{p}_{\hat{E}}(r)),(F^{1}_{\hat{E}}(r),F^{2}_{\hat{E}}(r),...,F^{p}_{\hat{E}}(r))\rangle : r\in\mathcal{M}\end{pmatrix}$$

where $(T^{1}_{\hat{E}}(r),T^{2}_{\hat{E}}(r),...,T^{p}_{\hat{E}}(r)):\mathcal{M}\longrightarrow[0,1]$, $(RT^{1}_{\hat{E}}(r),RT^{2}_{\hat{E}}(r),...,RT^{p}_{\hat{E}}(r)):\mathcal{M}\longrightarrow[0,1]$, $(RF^{1}_{\hat{E}}(r),RF^{2}_{\hat{E}}(r),...,RF^{p}_{\hat{E}}(r)):\mathcal{M}\longrightarrow[0,1]$ and $(F^{1}_{\hat{E}}(r),F^{2}_{\hat{E}}(r),...,F^{p}_{\hat{E}}(r)):\mathcal{M}\longrightarrow[0,1]$. Such that $0\leq supT^{i}_{\hat{E}}(r)+supRT^{i}_{\hat{E}}(r)+supRF^{i}_{\hat{E}}(r)+supF^{i}_{\hat{E}}(r)\leq 3$. $(i=1,2,3,..,p)$ $T^{1}_{\hat{E}}(r)\leq T^{2}_{\hat{E}}(r)...\leq T^{p}_{\hat{E}}(r))$, $RT^{1}_{\hat{E}}(r)\leq RT^{2}_{\hat{E}}(r)...\leq RT^{p}_{\hat{E}}(r))$ for any $r\in\mathcal{M}$. $T^{1}_{\hat{E}}(r),T^{2}_{\hat{E}}(r),...,T^{p}_{\hat{E}}(r)$, $RT^{1}_{\hat{E}}(r),RT^{2}_{\hat{E}}(r),...,RT^{p}_{\hat{E}}(r)$, $RF^{1}_{\hat{E}}(r),RF^{2}_{\hat{E}}(r),...,RF^{p}_{\hat{E}}(r)$, and $F^{1}_{\hat{E}}(r),F^{2}_{\hat{E}}(r),...,F^{p}_{\hat{E}}(r)$ is the Truth-membership sequence, Relative Truth-membership sequence, Relative False-membership sequence and False-membership sequence of the element $\mathcal{M}$, respectively. Also, P is called the dimension (cardinality) of QPNRS $\hat{E}$, denoted $d(\hat{E})$. We arrange the Truth-membership, Relative Truth-membership sequence in decreasing order but the corresponding Relative False-membership and False-membership sequence may not be in decreasing or increasing order.

Definition 5. *Let $\hat{E},\hat{G}\in QPNRS$. Then, $\hat{E}$ is said to be QPNRS subset to $\hat{G}$ denoted by $\hat{E}\subseteq\hat{G}$ if $T^{i}_{\hat{E}}(r)\leq T^{i}_{\hat{G}}(r)$, $RT^{i}_{\hat{E}}(r)\leq RT^{i}_{\hat{G}}(r)$, $RF^{i}_{\hat{E}}(r)\geq RF^{i}_{\hat{G}}(r)$, $F^{i}_{\hat{E}}(r)\geq F^{i}_{\hat{G}}(r)$ for all $r\in\mathcal{M}$ and $(i=1,2,3,..,p)$.*

Definition 6. *Let $\hat{E}, \hat{G} \in QPNRS$. Then, $\hat{E}$ is said to be QPNRS equal to $\hat{G}$ denoted by $\hat{E} = \hat{G}$ if $T^i_{\hat{E}}(r) = T^i_{\hat{G}}(r)$, $RT^i_{\hat{E}}(r) = RT^i_{\hat{G}}(r)$, $RF^i_{\hat{E}}(r) = RF^i_{\hat{G}}(r)$, $F^i_{\hat{E}}(r) = F^i_{\hat{G}}(r)$ for all $r \in \mathcal{M}$ and $(i = 1, 2, 3, .., p)$.*

Definition 7. *Let $\hat{E}, \hat{G} \in QPNRS$ then the union of $\hat{E}$ and $\hat{G}$ is denoted by $\hat{E} \cup \hat{G} = C$ and is defined by*

$$C = \begin{pmatrix} \langle r, (T^1_C(r), T^2_C(r), ..., T^p_C(r)), (RT^1_C(r), RT^2_C(r), ..., RT^p_C(r)), \\ (RF^1_C(r), RF^2_C(r), ..., RF^p_C(r)), (F^1_C(r), F^2_C(r), ..., F^p_C(r))\rangle : r \in \mathcal{M} \end{pmatrix}$$

$T^i_C = T^i_{\hat{E}}(r) \vee T^i_{\hat{G}}(r)$, $RT^i_C = RT^i_{\hat{E}}(r) \vee RT^i_{\hat{G}}(r)$, $RF^i_C = RF^i_{\hat{E}}(r) \wedge RF^i_{\hat{G}}(r)$, and $F^i_C = F^i_{\hat{E}}(r) \wedge F^i_{\hat{G}}(r)$ for all $r \in \mathcal{M}$ and $(i = 1, 2, 3, .., p)$.

Definition 8. *Let $\hat{E}, \hat{G} \in QPNRS$ then the intersection of $\hat{E}$ and $\hat{G}$ is denoted by $\hat{E} \cap \hat{G} = D$ and is defined by*

$$D = \begin{pmatrix} \langle r, (T^1_D(r), T^2_D(r), ..., T^p_D(r)), (RT^1_D(r), RT^2_D(r), ..., RT^p_D(r)), \\ (RF^1_D(r), RF^2_D(r), ..., RF^p_D(r)), (F^1_D(r), F^2_D(r), ..., F^p_D(r))\rangle : r \in \mathcal{M} \end{pmatrix}$$

$T^i_D = T^i_{\hat{E}}(r) \wedge T^i_{\hat{G}}(r)$, $RT^i_D = RT^i_{\hat{E}}(r) \wedge RT^i_{\hat{G}}(r)$, $RF^i_D = RF^i_{\hat{E}}(r) \vee RF^i_{\hat{G}}(r)$, and $F^i_D = F^i_{\hat{E}}(r) \vee F^i_{\hat{G}}(r)$ for all $r \in \mathcal{M}$ and $(i = 1, 2, 3, .., p)$.

Definition 9. *Let $\{\mathcal{M} = r_1, r_2, r_3, ..., r_n\}$ be the finite universe of discourse and*

$$\hat{E} = \{\langle T^j_{\hat{E}}(r_i) + RT^j_{\hat{E}}(r_i) + RF^j_{\hat{E}}(r_i) + F^j_{\hat{E}}(r_i)\rangle | r_i \in \mathcal{M}\}$$

$$\hat{G} = \{\langle T^j_{\hat{G}}(r_i) + RT^j_{\hat{G}}(r_i) + RF^j_{\hat{G}}(r_i) + F^j_{\hat{G}}(r_i)\rangle | r_i \in \mathcal{M}\}$$

be two quadri-partitioned neutrosophic refined sets consisting of the T^j: truth-membership, RT^j: membership, RF^j: membership and F^j: membership functions. Then the correlation coefficient of $\hat{E}$ and $\hat{G}$

$$QPNRS(\hat{E}, \hat{G}) = \frac{C_{NRS}(\hat{E}, \hat{G})}{\sqrt{C_{QPRS}(\hat{E}, \hat{E}) \times C_{QPRS}(\hat{G}, \hat{G})}}$$

Where

$$C_{QPNRS}(\hat{E}, \hat{G}) = \frac{1}{p}\sum_{j=1}^{p}\sum_{i=1}^{n}\{T^j_{\hat{E}}(r_i)T^j_{\hat{G}}(r_i) + RT^j_{\hat{E}}(r_i)RT^j_{\hat{G}}(r_i) + RF^j_{\hat{E}}(r_i)RF^j_{\hat{G}}(r_i) + F^j_{\hat{E}}(r_i)F^j_{\hat{G}}(r_i)\}$$

$$C_{QPNRS}(\hat{E}, \hat{E}) = \frac{1}{p}\sum_{j=1}^{p}\sum_{i=1}^{n}\{T^j_{\hat{E}}(r_i)T^j_{\hat{E}}(r_i) + RT^j_{\hat{E}}(r_i)RT^j_{\hat{E}}(r_i) + RF^j_{\hat{E}}(r_i)RF^j_{\hat{E}}(r_i) + F^j_{\hat{E}}(r_i)F^j_{\hat{E}}(r_i)\}$$

and

$$C_{QPNRS}(\hat{G}, \hat{G}) = \frac{1}{p}\sum_{j=1}^{p}\sum_{i=1}^{n}\{T^j_{\hat{G}}(r_i)T^j_{\hat{G}}(r_i) + RT^j_{\hat{G}}(r_i)RT^j_{\hat{G}}(r_i) + RF^j_{\hat{G}}(r_i)RF^j_{\hat{G}}(r_i) + F^j_{\hat{G}}(r_i)F^j_{\hat{G}}(r_i)\}$$

Proposition 1. *The defined correlation measure between $\hat{E}$ and $\hat{G}$ satisfies the following properties:*

(i) $0 \leq S(\hat{E}, \hat{G}) \leq 1$.

(ii) $S(\hat{E}, \hat{G}) = 1$ *iff* $\hat{E} = \hat{G}$.

(iii) $S(\hat{E},\hat{G}) = S(\hat{G},\hat{E})$.

Proof.

(i) $0 \leq S(\hat{E},\hat{G}) \leq 1$. As T and RT membership, RF and F-membership functions of the QPRS lies between 0 and 1, $S(\hat{E},\hat{G})$ is also lies between 0 and 1.

(ii) $S(\hat{E},\hat{G}) = 1$ iff $\hat{E} = \hat{G}$.
a. Let two QPNRS $\hat{E}$ and $\hat{G}$ be equal (i.e., $\hat{E} = \hat{G}$). Hence for any $T^j_{\hat{E}}(r_i) = T^j_{\hat{G}}(r_i)$, $RT^j_{\hat{E}}(r_i) = RT^j_{\hat{G}}(r_i)$, $RF^j_{\hat{E}}(r_i) = RF^j_{\hat{G}}(r_i)$, $F^j_{\hat{E}}(r_i) = T^j_{\hat{G}}(r_i)$. Then,
$C_{QPNRS}(\hat{E},\hat{E}) = C_{QPNRS}(\hat{G},\hat{G})$

$$= \frac{1}{p}\sum_{j=1}^{p}\sum_{i=1}^{n}\{T^j_{\hat{E}}(r_i)T^j_{\hat{E}}(r_i) + RT^j_{\hat{E}}(r_i)RT^j_{\hat{E}}(r_i) + RF^j_{\hat{E}}(r_i)RF^j_{\hat{E}}(r_i) + F^j_{\hat{E}}(r_i)F^j_{\hat{E}}(r_i)\}$$

and

$$C_{QPNRS}(\hat{E},\hat{G}) = \frac{1}{p}\sum_{j=1}^{p}\sum_{i=1}^{n}\{T^j_{\hat{E}}(r_i)T^j_{\hat{G}}(r_i)+RT^j_{\hat{E}}(r_i)RT^j_{\hat{G}}(r_i)+RF^j_{\hat{E}}(r_i)RF^j_{\hat{G}}(r_i)+F^j_{\hat{E}}(r_i)F^j_{\hat{G}}(r_i)\}$$

$$= \frac{1}{p}\sum_{j=1}^{p}\sum_{i=1}^{n}\{T^j_{\hat{E}}(r_i)T^j_{\hat{E}}(r_i) + RT^j_{\hat{E}}(r_i)RT^j_{\hat{E}}(r_i) + RF^j_{\hat{E}}(r_i)RF^j_{\hat{E}}(r_i) + F^j_{\hat{E}}(r_i)F^j_{\hat{E}}(r_i)\}$$

$$= C_{QPNRS}(\hat{E},\hat{E})$$

Hence

$$S(\hat{E},\hat{G}) = \frac{C_{NRS}(\hat{E},\hat{G})}{\sqrt{C_{QPRS}(\hat{E},\hat{E}) \times C_{QPRS}(\hat{G},\hat{G})}} = \frac{C_{NRS}(\hat{E},\hat{E})}{\sqrt{C_{QPRS}(\hat{E},\hat{E}) \times C_{QPRS}(\hat{E},\hat{E})}} = 1$$

b. Let $S(\hat{E},\hat{G}) = 1$ then the unit measure is possible only if $\frac{C_{NRS}(\hat{E},\hat{G})}{\sqrt{C_{QPRS}(\hat{E},\hat{E})\times C_{QPRS}(\hat{G},\hat{G})}} = 1$. This refers that $T^j_{\hat{E}}(r_i) = T^j_{\hat{G}}(r_i)$, $RT^j_{\hat{E}}(r_i) = RT^j_{\hat{G}}(r_i)$, $RF^j_{\hat{E}}(r_i) = RF^j_{\hat{G}}(r_i)$, $F^j_{\hat{E}}(r_i) = T^j_{\hat{G}}(r_i)$ for all i, j. Hence $\hat{E} = \hat{G}$.

(iii) If $S(\hat{E},\hat{G}) = S(\hat{G},\hat{E})$, it obvious that

$$\frac{C_{NRS}(\hat{E},\hat{G})}{\sqrt{C_{QPRS}(\hat{E},\hat{E}) \times C_{QPRS}(\hat{G},\hat{G})}} = \frac{C_{NRS}(\hat{G},\hat{E})}{\sqrt{C_{QPRS}(\hat{G},\hat{G}) \times C_{QPRS}(\hat{E},\hat{E})}} = C_{QPNRS}(\hat{G},\hat{E})$$

As

$$C_{QPNRS}(\hat{E},\hat{G}) = \frac{1}{p}\sum_{j=1}^{p}\sum_{i=1}^{n}\{T^j_{\hat{E}}(r_i)T^j_{\hat{G}}(r_i)+RT^j_{\hat{E}}(r_i)RT^j_{\hat{G}}(r_i)+RF^j_{\hat{E}}(r_i)RF^j_{\hat{G}}(r_i)+F^j_{\hat{E}}(r_i)F^j_{\hat{G}}(r_i)\}$$

$$= \frac{1}{p}\sum_{j=1}^{p}\sum_{i=1}^{n}\{T^j_{\hat{G}}(r_i)T^j_{\hat{E}}(r_i) + RT^j_{\hat{E}}(r_i)RT^j_{\hat{G}}(r_i) + RF^j_{\hat{G}}(r_i)RF^j_{\hat{E}}(r_i) + F^j_{\hat{G}}(r_i)F^j_{\hat{E}}(r_i)\}$$

$$= C_{QPNRS}(\hat{G},\hat{E})$$

4. A Symptom-Disease Correlation Model based on Double-Valued Neutrosophic Set Theory

Numerous uncertainties in medical diagnosis make it challenging to unite different symptom sets under a single illness term. This difficulty is exacerbated by the increased amount of information that doctors can now access thanks to modern technologies.
Furthermore, there are instances in the real world when each component can be categorized based on its true, relative truth, relative false, and false membership. It is better to undertake a multi-time inspection (sampling a single patient at different times) in order to make a diagnosis because a single check-up could lead to an error. A quadri-partitioned neutrosophic refined set (QPNRS)-based methodology is presented in this work to improve diagnostic accuracy by modeling the patient-symptom-disease linkages. The given measure of connection between the diseases and the symptoms of the patients provides a suitable medical diagnosis. This method is distinct in that it takes into account truth membership, relative truth membership, relative false membership, and false membership functions. Three times during the day, at 8:00, 12:00, and 16:00, the symptoms of three patients (P_1, P_2, *and* P_3) were noted. Temperature variations, coughing, headaches, sore throats, and body aches were all recorded. A quadri-partitioned neutrosophic refined matrix Q, which displays the evolution of patients and symptoms over time, was created by organizing them. A second R matrix was developed to fit symptoms and four prevalent diseases: viral fever, TB, throat infection, and typhoid. The diagnostic result for each patient was determined using the highest correlation score between Q and R. The results of the study show that patient P_1 is most likely to have tuberculosis (correlation = 0.901), and patients P_2 and P_3 are most likely to have typhoid (correlations = 0.878 and 0.869, respectively). This uncertainty-friendly QPNRS-based diagnostic methodology is quite beneficial when it comes to ambiguity in symptom rating and a systematic, data-driven approach to supporting medical decision-making. This approach has been demonstrated using a medical diagnosis.

Example 1. *The framework defines three sets: Let* $S = \{Temperature, cough, throatpain, headache, bodypain\}$ *be a group of symptoms,* $P = \{P_1, P_2, P_3\}$ *be a group of patients, and* $D = \{ViralFever, Tuberculosis, Typhoid, Throatdisease\}$ *be a group of diseases. As part of the recommended remedy, patients will be observed three times a day. This temporal refinement generates unique Truth, Re-Truth, Re-False, and False membership values for each patient, resulting in a more complex dataset for diagnosis.*

Table 1: Q (The Relation between Patient and Symptoms)

Q	**Temperature**	**Cough**	**Throat Pain**	**Headache**	**Body pain**
	(0.4,0.2,0.1,0.4)	(0.5,0.2,0.2,0.4)	(0.3,0.3,0.2,0.5)	(0.5,0.2,0.1,0.4)	(0.5,0.1,0.1,0.4)
P_1	(0.3,0.2,0.2,0.6)	(0.4,0.1,0.0,0.3)	(0.2,0.3,0.3,0.4)	(0.5,0.2,0.2,0.7)	(0.2,0.2,0.1,0.5)
	(0.2,0.3.0.2,0.5)	(0.3,0.3,0.1,0.5)	(0.1,0.3,0.3,0.3)	(0.3,0.2,0.1,0.6)	(0.1,0.2,0.2,0.3)
	(0.6,0.2,0.1,0.5)	(0.6,0.2,0.1,0.7)	(0.6,0.2,0.1,0.3)	(0.6,0.2,0.1,0.1)	(0.4,0.2,0.2,0.5)
P_2	(0.5,0.3,0.2,0.2)	(0.4,0.3,0.1,0.2)	(0.3,0.3,0.2,0.4)	(0.4,0.3,0.1,0.8)	(0.3,0.1,0.2,0.7)
	(0.4,0.2,0.2,0.5)	(0.2,0.2,0.2,0.5)	(0.1,0.2,0.2,0.5)	(0.2,0.2,0.2,0.3)	(0.1,0.3,0.2,0.5)
	(0.8,0.1,0.2,0.5)	(0.5,0.3,0.2,0.3)	(0.3,0.2,0.1,0.6)	(0.6,0.1,0.1,0.5)	(0.6,0.2,0.2,0.5)
P_3	(0.7,0.3,0.2,0.4)	(0.3,0.3,0.1,0.3)	(0.2,0.3,0.2,0.7)	(0.5,0.2,0.1,0.6)	(0.3,0.2,0.1,0.4)
	(0.6,0.2,0.2,0.4)	(0.1,0.4,0.2,0.4)	(0.1,0.2,0.2,0.5)	(0.2,0.1,0.1,0.6)	(0.2,0.1,0.1,0.6)

Let the samples be taken at three different timings in a day (in 08:00, 12:00, 16: 00)

Table 2: R (The relation among Symptoms and Diseases)

R	**Viral Fever**	**Tuberculosis**	**Typhoid**	**Throat disease**
Temperature	(0.2,0.3,0,2,0.6)	(0.4,0.4,0.2,0.5)	(0.6,0.3,0.1,0.5)	(0.3,0.4,0.2,0.8)
Cough	(0.6,0.2,0.2,0.6)	(0.8,0.1,0.1,0.3)	(0.3,0.1,0.1,0.6)	(0.2,0.2,0.2,0.1)
Throat Pain	(0.5,0.1,0.1,0.3)	(0.4,0.3,0.2,0.3)	(0.4,0.3,0.2,0.5)	(0.2,0.4,0.2,0.2)
Headache	(0.6,0.5,0.3,0.2)	(0.2,0.2,0.1,0.6)	(0.1,0.3,0.3,0.3)	(0.2,0.3,0.2,0.5)
Body Pain	(0.7,0.3,0.2,0.4)	(0.2,0.2,0,1,0.4)	(0.2,0.2,0.1,0.4)	(0.2,0.1,0.1,0.3)

Table 3: The Correlation Measure between NRS Q and R

Correlation Measure	**Viral Fever**	**Tuberculosis**	**Typhoid**	**Throat disease**
P_1	0.872	0.901	0.864	0.869
P_2	0.845	0.743	0.878	0.787
P_3	0.792	0.839	0.869	0.811

The optimal correlation value in Table 3 provides the accurate medical diagnosis. Thus, although patients P_2 and P_3 have typhoid, patient P_1 has tuberculosis. The suggested QPNRS correlation measure can effectively convert difficult, multi-time, and ambiguous symptom data into a measurable and practical diagnostic tool, as demonstrated in this table.

5. Result and Discussion

The K-Means clustering figure (k=2) in Figure 1 illustrates how patient profiles ($P_1 - P_3$) group according to their correlational patterns. Point color indicates cluster membership, while the axes (PCA Components 1 and 2) provide a low-dimensional perspective of the correlation features. The centroids of the two clusters are indicated by large black "X" markings. Two distinct clusters are present: P_2 and P_3 in Cluster 0 (red dots) indicate that the correlation patterns of the two profiles for illnesses are similar. Their closeness to one another in PCA space indicates that their illness connection patterns are more comparable than P_1. Because Cluster 1 (green dot) solely contains P_1, it may be distinguished from P_2 and P_3. This divide demonstrates P_1's unique correlation behavior, especially its strong and distinctive link to tuberculosis. This separation is supported by the cluster centers, where one centroid is between P_2 and P_3, suggesting their shared structure informed by typhoid, and the other is closer to P_1, reflecting its profile saturated with tuberculosis. Generally speaking, the clustering provides an unsupervised confirmation of the earlier correlation analysis: P_2 and P_3 appear to be correlated partners primarily influenced by typhoid, while P_1 stands out due to its preponderant association with tuberculosis, which clearly separates it from the other profiles.

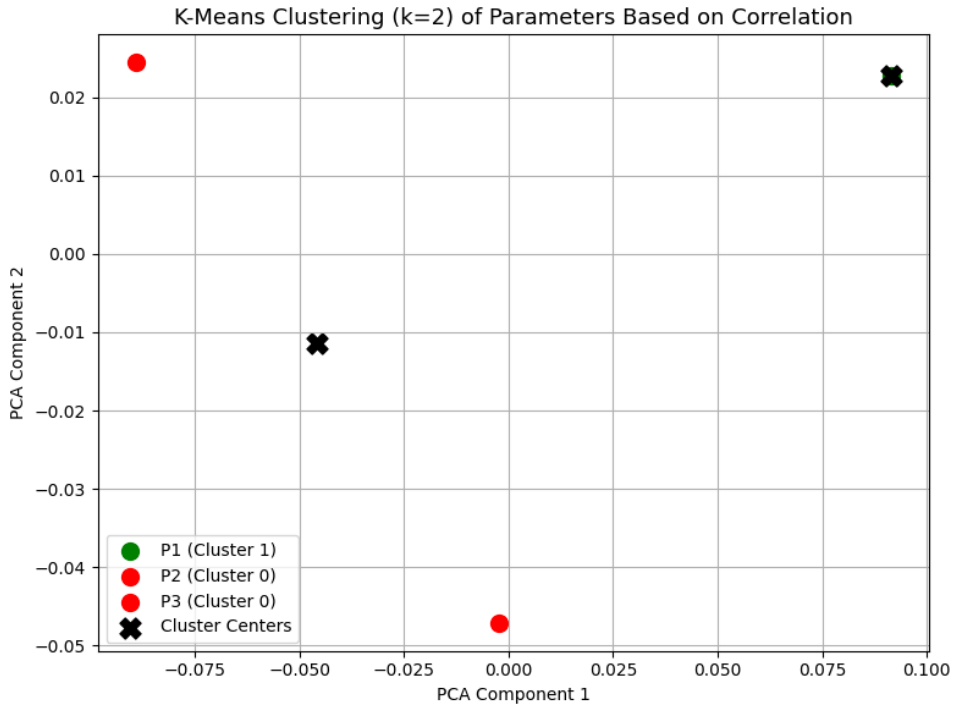

Figure 1

When projected onto two main components, the three patient profiles ($P_1 - P_3$) cluster according to their patterns of correlation, as shown in Fig. 2's K-Means++ clustering map (k=2). The centroids of the two clusters are shown by the black "X" marks, and each point represents a profile. There are two clear clusters: Cluster 1 (green point) contains only P_1, while Cluster 0 (red points) contains both P_2 and P_3. P_2 and P_3 are closely positioned, indicating that their correlation tendencies are identical, particularly because to their high association with typhoid. However, P_1's partitioning into a discrete cluster suggests a unique correlation structure, which is dominated by its significant connection with tuberculosis. These cluster centers-one with the unique pattern of P_1 and another with the shared characteristics of P_2 and P_3-support this separation. Generally speaking, the clustering finds two natural patient groups: P_1 as an outlier with a unique diagnostic profile and P_2 and P_3 as a similarity group.

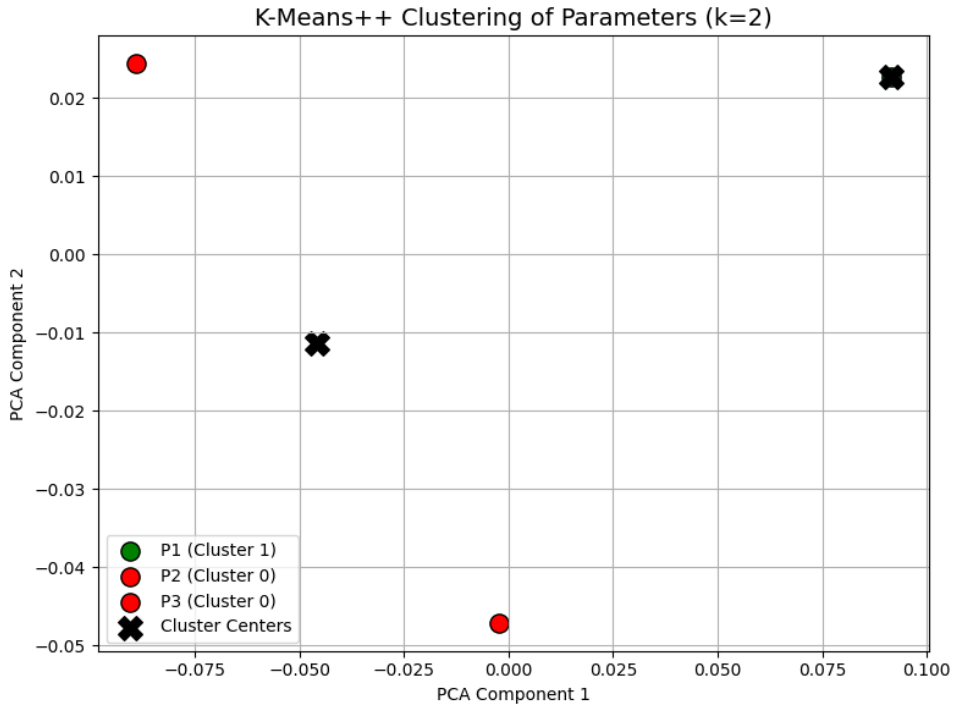

Figure 2

The three patient profiles ($P_1 - P_3$) are grouped based on their correlation patterns after normalizing and dimensionality reduction into two principal components, as shown in Fig. 3, the K-Means++ clustering diagram (k=2) on normalized data. The cluster centroids are indicated by the black "X" markers, and each point represents a profile. Two clusters are readily apparent: P_2 and P_3 are found in Cluster 0 (red dots), whereas P_1 is the only one found in Cluster 1 (green dot). The fact that P_2 and P_3 are located in the same cluster indicates that these two profiles have comparable normalized correlation patterns, with the strength of these correlations being dominated by their shared relationship with typhoid. Contrarily, P_1's placement in another cluster reveals another unique normalized profile with a strong Tuberculosis correlation as its core, setting it apart from the rest. The confirmed cluster centers also support this separation: one centroid matches closely with P_2 and P_3's common characteristics, while the other identifies P_1's singular pattern. In total, the clustering shows two inherent groupings of patients in normalized conditions P_2 and P_3 in a similarity pair, and P_1 as an outlier with a unique diagnostic profile.

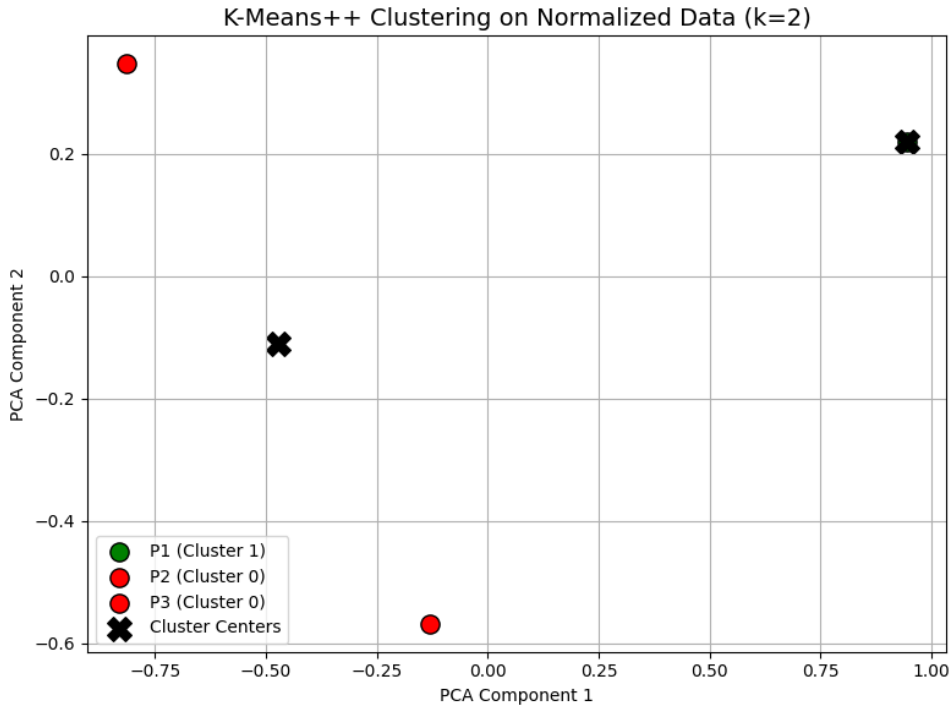

Figure 3

In Figure 4 the flower samples are grouped into three groups based on feature similarities is shown by the K-Means clustering plot on the PCA-reduced Iris data. Each point is a single sample that is projected onto the two major components for viewing, and the titles of the groups are represented by the enormous black dots with a black X. There are three distinct clusters found. Cluster 1 (green points) is tightly packed on the left, forming a tight, well-spaced cluster associated with a single species of iris (likely Iris setosa). Cluster 2, which has an intermediate nature (often Iris versicolor), partially overlaps Cluster 0 (red points), which is in the center of the field and includes more widely distributed samples. The blue dots in Cluster 2, which is located on the right, are dispersed around the cluster's centroid and partially overlap with Cluster 0, indicating both a degree of separation and resemblance (typically Iris virginica). The cluster centers that have one centroid inside the thick green cluster, another centroid lying at the center of the red samples, and a third centroid balancing the blue cluster serve as validation for these groupings. While the green cluster is an autonomous cluster with few cross-sections, the red and blue cross-sections highlight the dataset's transitional patterns. Intersected, each clearly unique. Overall, the Iris dataset's nature can be effectively depicted using PCA's reduced clustering: one species is represented by a tight and distinct clustering, while the other two exhibit more gradual boundary parting with some overlap, which is normal given their biological similarity.

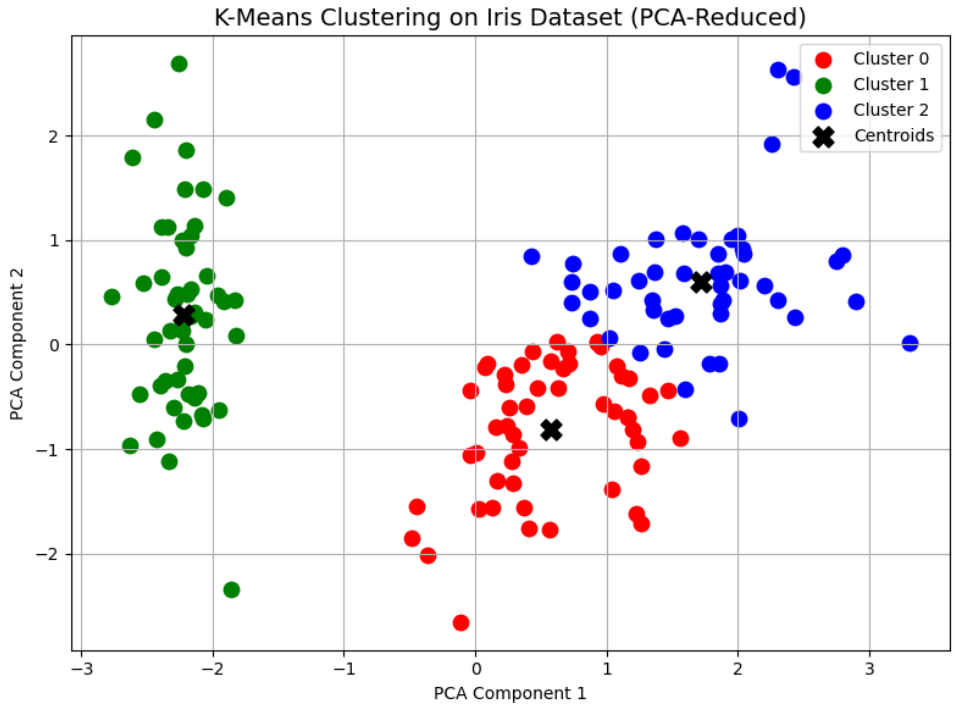

Figure 4

The normalized and PCA-reduced Iris data's K-Means++ clustering plot is shown in Fig. 5, which also shows how the samples are grouped into three distinct clusters based on how similar they are on the normalized features. Each flower is described by the points, and the cluster centroids are indicated by the enormous black X of the points. The presence of three clusters is clear. Cluster 1 (green points), which represents a single species (often Iris setosa), is located on the far left and is compact and dense with little overlap. Cluster 2 (blue points) lies halfway between the two and partially overlaps with Cluster 0, while Cluster 0 (red points) is on the far right and has a more extensive distribution.Given the biological resemblances between Iris versicolor and Iris virginica, this overlap suggests characteristics of transition between the two families. The cluster centers support this distribution: the centroid in the green cluster is firmly integrated, the centroid in the red cluster is at the center, and the third centroid is at the center of the blue group. With the green cluster becoming even more noticeable and the red and blue clusters' partial boundary overlap still discernible, the normalized scaling improves the separation.

Overall, the normalized data's K-Means++ clustering provides a great and well-structured view of the Iris dataset. Two of them exhibit a similarity gradient, while one is compact and highly segregated. These examples demonstrate how normalization can highlight natural species splits while preserving overlaps where features are convergent.

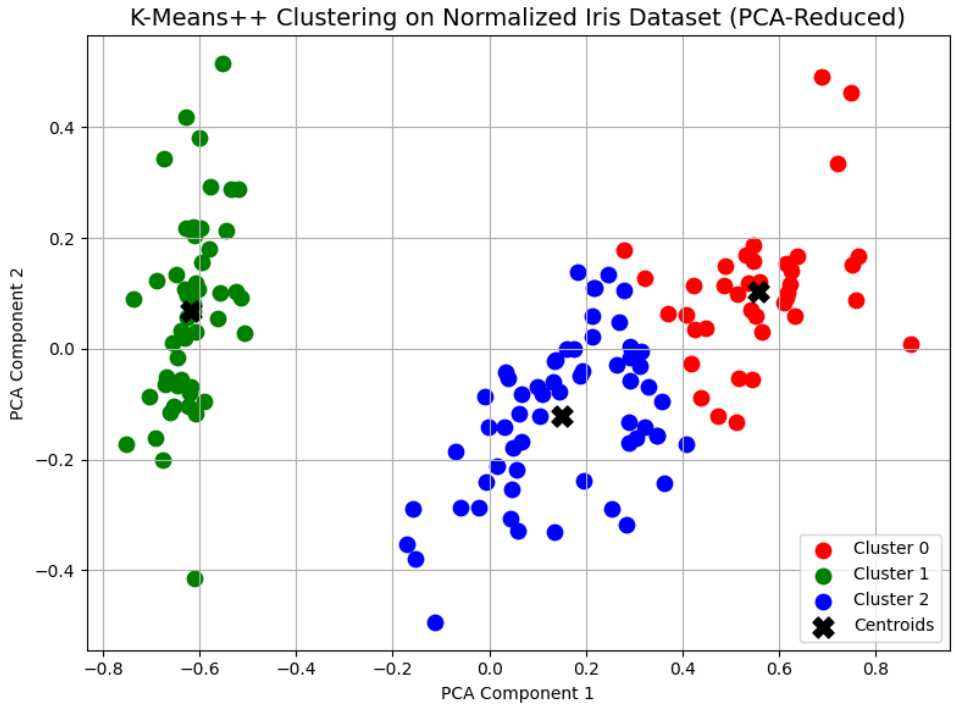

Figure 5

The distribution of the three species (setosa, versicolor, and virginica) in the lower dimensions (TSNE-1, TSNE-2, and TSNE-3) is shown by the parallel coordinates plot of the 3D t-SNE visualization of the Iris data in Fig. 6. Each colored line's value corresponds to a single flower sample and connects the values of the three components. It is possible to observe particular grouping patterns. The upper band of the TSNE-2 values is where the red lines (setosa) are mostly concentrated, creating a fairly uniform and continuous band. This demonstrates Setosa's well-known distinctiveness from the other two species. Versicolor, the green lines' intermediate band, is scattered across TSNE-1 and TSNE-3 before descending in TSNE-2. The transitional traits between setosa and virginica are depicted by their placement. The lower regions of TSNE-1 and TSNE-2 exhibit the dominance of the blue lines (virginica), while TSNE-3 shows the lines going upward to the higher side. Compared to setosa, this wider distribution is more erratic. These two species are not as easily separated as setosa because of their biological similarities, which are highlighted by the junction of the green (versicolor) and blue (virginica) lines.
The setosa red lines, on the other hand, are clearly distinguished, and the dataset clearly shows that it has a unique characteristic. Generally speaking, the parallel coordinates map provides a multivariate view of the Iris dataset's separability: versicolor and virginica are more broadly distributed, suggesting a stronger link in the feature space, while setosa is a small, discrete cluster.

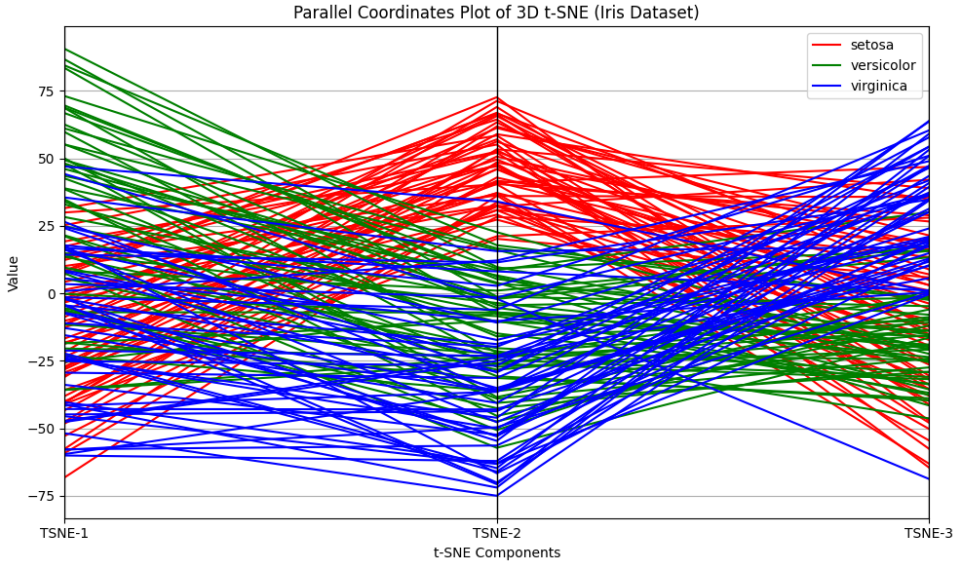

Figure 6

The distribution of the three species (setosa, versicolor, and virginica) in the normalized components (TSNE-1, TSNE-2, and TSNE-3) of the data is depicted in the parallel coordinates plot of the normalized 3D t-SNE representation of the iris data in Fig. 7. Each line represents opposition to a sample, and normalization conditionalizes the values of 0 and 1, allowing direct comparisons between species samples. Various patterns of structure are created. In a strong band around 1.0, the red lines (setosa) are concentrated in the upper part of the TSNE-2 values. This indicates that setosa, a unique feature space of the other species, is stable during separation. They are unique on the dimensions because the green lines (virginica) are more common on the lower values of TSNE-1 and rise strongly on TSNE-3, producing an upward line. The blue (versicolor) lines that separate the setosa and virginica extremes are situated in the middle of each of the three axes. The range between the mid-range values of TSNE-1 and TSNE-2 is where the overlap between the blue (versicolor) and green (virginica) curves is most noticeable. Their biological proximity, where feature features are less clearly divided, is the basis for this. The red (setosa) lines, on the other hand, are comparatively well separated, demonstrating its individuality. Because the scales are aligned, normalized parallel coordinate plots are generally easier to comprehend and make cross-species comparisons more visually appealing. It emphasizes that whereas versicolor and virginica have transitional and overlapping features in the t-SNE components and partially overlap, setosa always constitute a separate cluster.

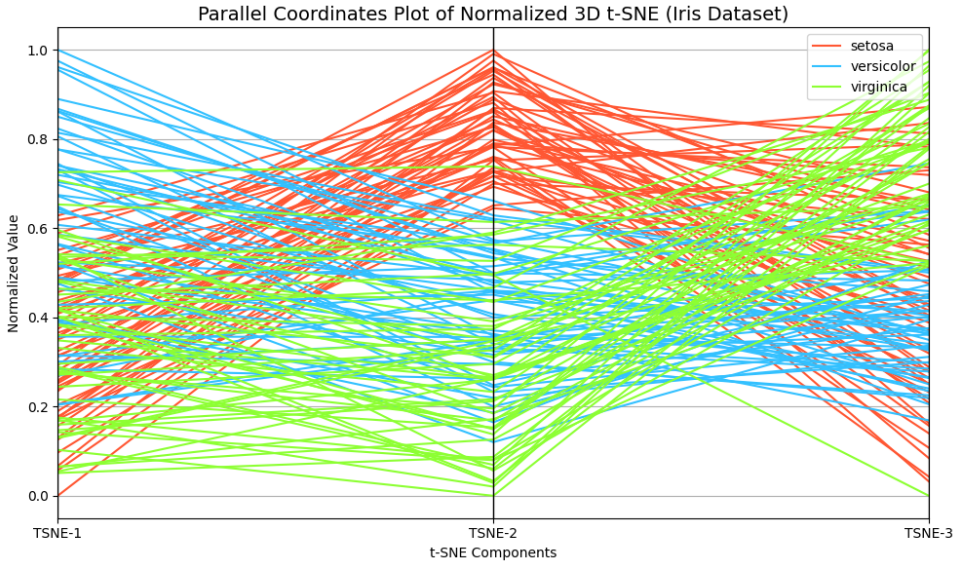

Figure 7

The iris data is shown in 3-D t-SNE format in Fig. 8, which converts the high-dimensional flower data to three dimensions (t-SNE1, t-SNE2, t-SNE3) so that the species may be distinguished in two dimensions. Each point represents a single sample, which is colored green (Virginica), blue (Versicolor), and red (Setosa). The species' numerical range of distribution across the embedding is shown. Because setosa (red) is concentrated on the positive end, versicolor (blue) is distributed between -40 and +80, and virginica (green) is concentrated at the negative end (-60 to 20), the t-SNE1 values vary from +90 to -60. Setosa is more concentrated on the upper half (0 to +60), virginica is more focused on the middle range (-20 to +40), and versicolor is more focused on the lower half (-60 to +30) of the t-SNE2 axis, which covers a region of roughly -70 to +70. Versicolor is concentrated at the lower end (-60 to 0), virginica is in the middle (-20 to +40), and setosa reemerges at a higher level (10 to +60) along the t-SNE3 axis, which runs between -70 and +65. Cluster separation is visible but not perfect. A very compact cluster, t-SNE2 (= +20 - +60) and t-SNE3 (= 10 - +60), which are unique to its morphological features, are also taken by Setosa. With wide extension over low values of t-SNE3 ($\sim$ -60 to 0) and negative and positive values of t-SNE1 ($\sim$ -40 to +80), Versicolor is partially overlapping Virginica. Virginica shares with setosa and versicolor the central areas, covering t-SNE1 (c. t -60 t +20) and t-SNE3 (c. t -20 t +40) in the middle. On the whole, the 3D t-SNE plot proves that the most separable species is setosa, which makes a distinct cluster in t-SNE2 and t-SNE3. Conversely, versicolor and virginica also overlap in ranges of t-SNE1 -40 t-SNE1 20 and t-SNE3 -20 t-SNE3 40 as is expected given the biological similarity between them. The numerical frequencies represent the distribution of the dataset in the embedding with setosa being separate whereas the other two species have boundaries which are shared.

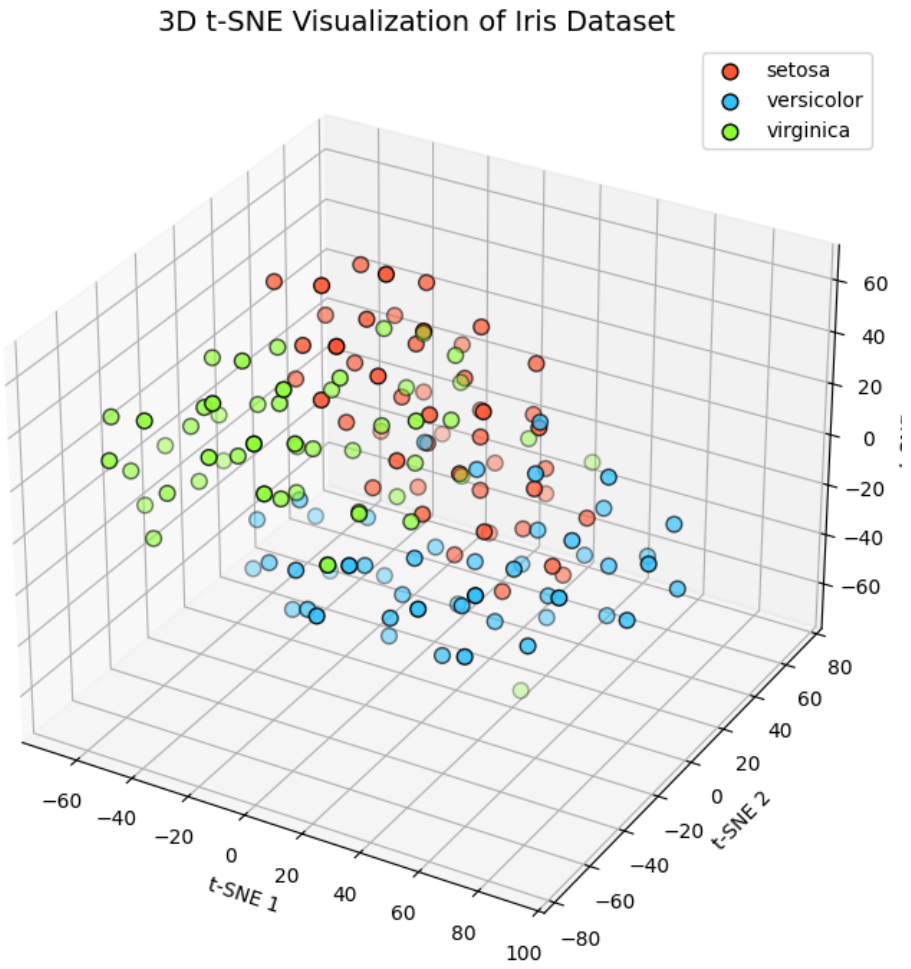

Figure 8

In order to demonstrate the separation at the species level, Fig. 9 presents the t-SNE visualization of the Iris data set in 3D, which comprises the high-dimensional floral features into three reduced components (t-SNE1, t-SNE2, and t-SNE3). Per point, the species-specific flowers are green (virginica), blue (versicolor), and red (setosa).
Numerical ranges are used to disperse clusters. Along the t-SNE1 axis, the values fall between negative and positive -70 and +90. The concentration curves of versicolor (blue) are widely distributed (between -40 and +80), which overlap with both others, setosa (red) are more concentrated in the central-to-positive range (0 to +70), and virginica (green) are concentrated on the negative side (-70 to 0). Setosa scores are higher (0 to +70), virginia scores are in the center (-20 to +40), and versicolor scores are lower (-60 to +30) along the t-SNE2 spectrum, which has scores ranging from -70 to +80. The t-SNE3 axis, which spans a range of roughly -70 to +65, provides an additional separation of the species, with virginica between the two (-20 to +40), versicolor on the negative (-60 to 0), and setosa at the top (10 to +65). According to cluster patterns, setosa produces the clearest cluster with less overlap and higher t-SNE2 and t-SNE3 bands. On the other hand, versicolor and virginica exhibit significant overlap, especially in the middle regions of t-SNE1 (-40 +20) and t-SNE3 (-20 +40).Given that they share more physical traits than setosa, this relationship shows that they are biologically comparable. In general, 3D t-SNE projection shows a layered picture of separability: virginica and versicolor are more in contact with each other and form intermediate zones in the embedding space, whereas setosa appears as a small, distinct cluster. In the reduced dimensions, each species can be found in distinct subregions that partially overlap, as shown by the numerical ranges.

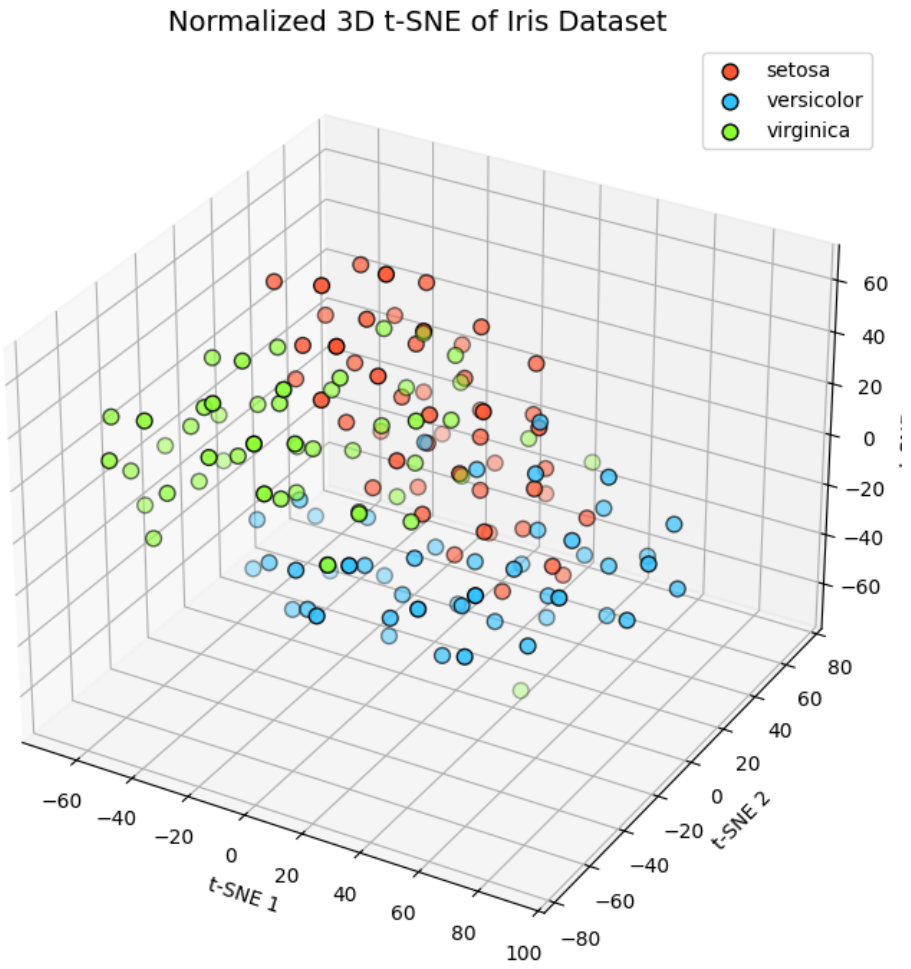

Figure 9

The three patient profiles (P_1, P_2, and P_3) are projected into two dimensions in Fig. 10, which displays the t-SNE view of patient disease correlations. Each point in the plot represents the overall correlation structure of a patient disease-wise. The abstract t-SNE dimensions, shown by the horizontal axis (Component 1) and the vertical axis (Component 2), show the similarity in relative terms; the closer the distances, the more similar the correlation patterns. The numerical positions are used to highlight the grouping. P_1 (blue), which is located at the far left at around (-460, 140), indicates that this correlation profile differs greatly from others. P_2 (orange), which is located at roughly (190, -160) on the far right, forms a unique cluster with a low vertical position. Positioned in the upper-right quadrant, around (140, 550), and significantly above P_2 and P1, P_3 (green) shows a totally different correlation pattern. The large spatial dispersion of the three dots indicates that patients' patterns of correlation are less comparable to one another. While P_2 and P_3 are significantly different on Component 2 (about 700 units), P_1 and P_2 are separated on Component 1 (approximately 650 units). P_1 and P_3 have a wide (almost 750 units) diagonal spacing, which emphasizes the fact that the two patients are not closely packed. Overall, each patient has a unique correlation signature across diseases, according to the t-SNE projection. This graphic shows pure divergence as opposed to the previous clustering trends that included overlap or pairings (like P_2 and P_3). P_1, P_2, and P_3 are separate, non-overlapping points in the reduced space, showing highly differentiated disease-correlation patterns.

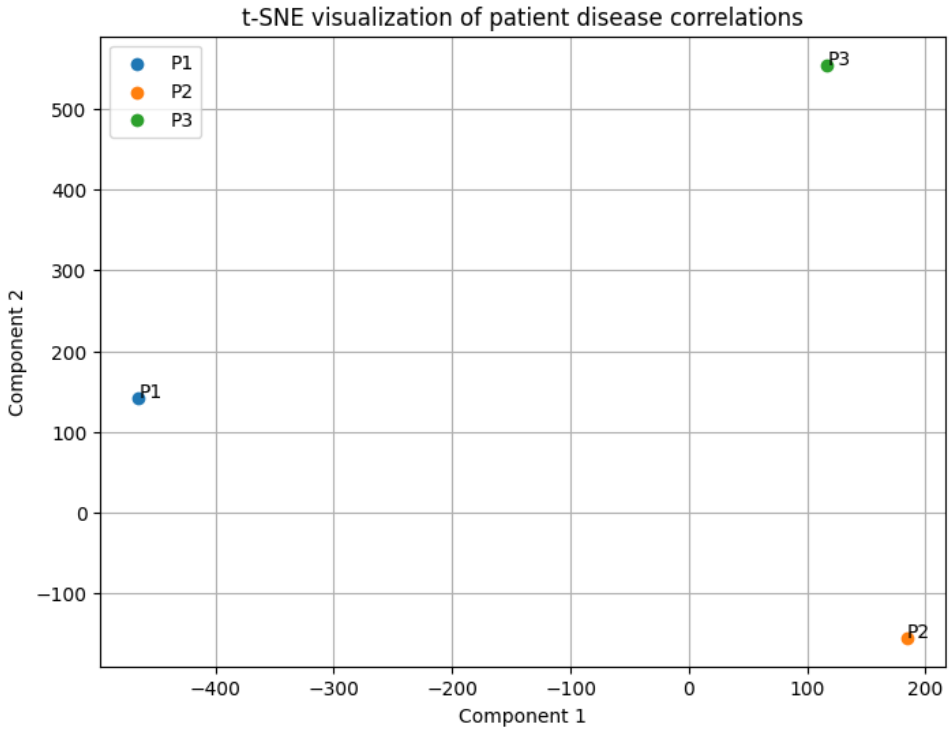

Figure 10

The three patient profiles (P_1, P_2, and P_3) are projected into a two-dimensional space by the first two principal components, which indicate the most diversity in their correlation structures, in the PCA plot of the correlation measurements in Fig. 11. Every point represents a single patient's total disease-correlation pattern in reduced dimensions. Defined differentiation is indicated by numerical position. With a value of around (2.3, 0.55), P_1 is located in the upper-right quadrant, suggesting that Component 2 has a mediocrely positive connotation and Principal Component 1 has a strong positive implication.
With Component 1 showing a strong negative indication and Component 2 showing the maximum positive signal, P_2 is located in the first quadrant, around (-2.0, 0.8). P_3 has a mildly negative bias for Component 1 and a strong negative bias for Component 2, and it is situated in the lower-center region of roughly (-0.3, -1.35). The distances between points show the structural differences. The fact that P_1 and P_2 are big in Component 1 (with an average separation of roughly 4.3 units) suggests that their correlation structure is opposing in this key direction. While P_1 and P_3 are considerably differed on both axes (about 2.6 units along PC1 and 1.9 units along PC2), P_2 and P_3 are different based on Component 2 (a difference of about 2.1 units). The PCA projection generally shows that the three profiles are quite distinct, with P1 having strong positive PC1 values (because to tuberculosis), P_2 having strong positive PC2 values (due to typhoid), and P_3 having negative PC2 values (due to neither). This split suggests that there is no meaningful grouping of the patients in the condensed space and that each patient's correlation pattern is unique.

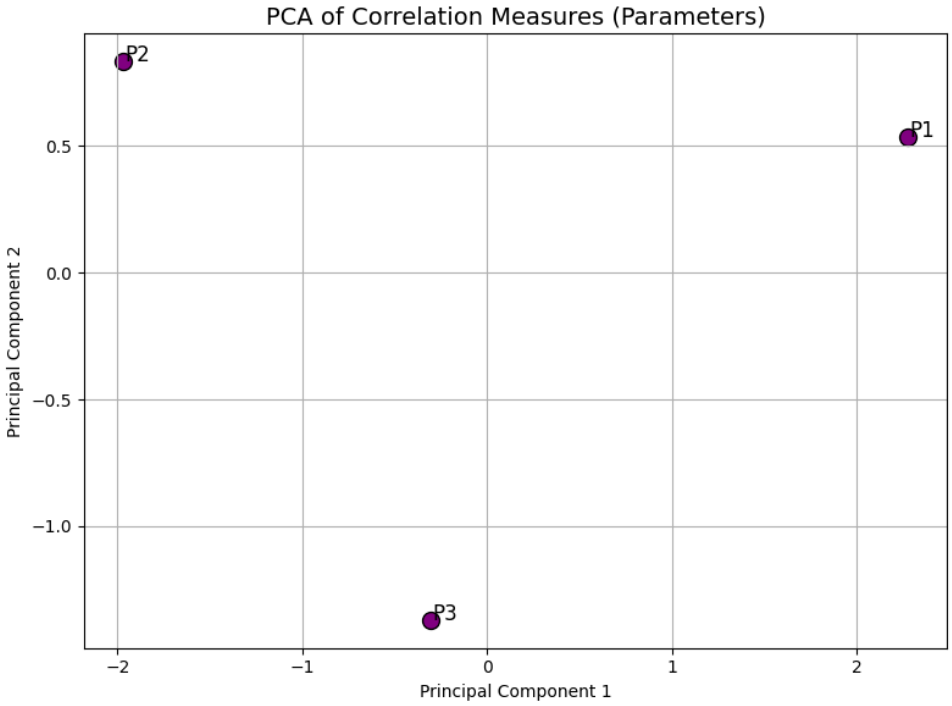

Figure 11

Plotting the three patient profiles (P_1, P_2, and P_3) in three principal components, which show the greatest variation in the disease-correlation structure of the three patients, is how the 3D PCA plot of correlation measurements is shown in Fig. 12. With an emphasis on decreased spatial separability, all three axes (Principal Components 1, 2, and 3) are weighted sums of the original variables. Numerical places exhibit distinct clustering. At around (2.2, 0.6, -9.0), P_1 is on the right extreme of Principal Component 1, meaning that PC 1 has a positive impact of about 1 and PC 2 has a modest impact. There is a significant negative PC1 and a positive PC2, as shown by P_2, which is at the far left at roughly (-2.0, 0.5, -8.6). It is marginally higher on PC3 than the others. P_3's primary feature is the negative PC 2 value and its higher location on PC 3, which places it even closer to the origin on PC1 and PC2 at roughly (-0.3, -1.2, -9.3). There are notable structural variations in the dissemination. With a difference of almost 4.2 units, the P1 and P_2 are the most different on PC1 and tend to show opposite correlations. Additionally, P_1 and P_3 are separated along PC1 (about 2.5 units) and PC2 (about 1.8 units), while P_2 and P_3 are sharp along PC2 (about 1.7 units). The three lie close to one other on PC3 (around -9) which suggests that this component is less important in their separation than PC1 and PC2. The 3D PCA often confirms that each of the three profiles is unique. P_2 is a high PC2 loading with negative PC1 (also typical of typhoid-related), P_3 is a negative PC2 loading (different in both situations), and P1 is a strong positive PC1 loading (also typical of its TB-dominated connection). The fact that the first two variables determine the main separation and that P_1, P_2, and P_3 have different signatures of association with disease is highlighted by the modest format of the distribution along PC3.

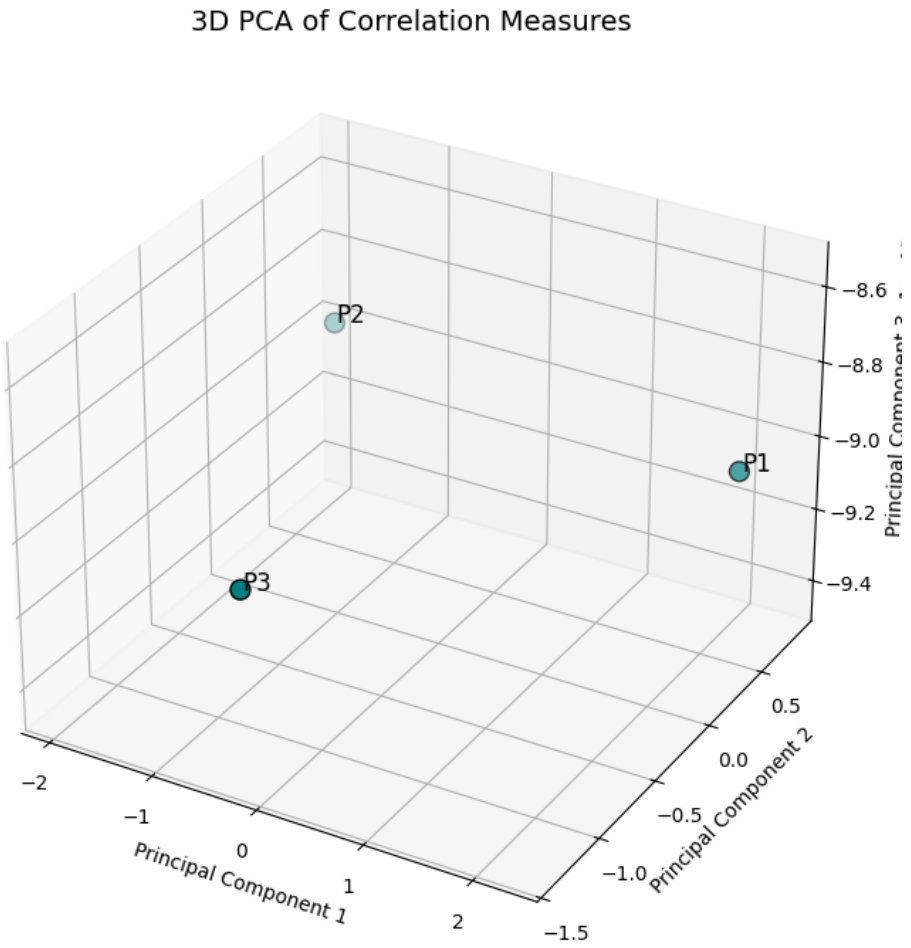

Figure 12

The three patient profiles (P_1, P_2, and P_3) are projected into three non-linear components, similarity patterns in their correlation patterns, in Fig. 13, which displays a 3D t-SNE depiction of patient disease correlations. Individual patients are represented by the points, and the points' positions show how similar or different the correlation patterns between the patients and the diseases are. Numbers are used to highlight separation. P_1 (blue) indicates a very strong divergent correlation profile in that direction because it is on the extreme negative side of Component 1, about between -1050 and -250 and 8.4.Since P_2 (orange) is at the other extreme, somewhere between 850 and 900 and 8.6, it is fundamentally distinct from P_1. P_3 (green), which is identified by its Component 2 value, which is the most negative, is situated in a central but isolated location close to (300, -1200, 8.5). Each patient profile is unique, as shown by the axes' dispersion. The correlation patterns of P_1 and P_2 are the least similar, as evidenced by the biggest distance between them-nearly 1900 units. Another line of separation is highlighted on Component 2, where P_2 and P_3 are separated by an average of 2100 units. Additionally, Component 3 values are maintained very close to all three ($\approx$ 8.4 -8.6), indicating that the third axis is not experiencing separation. The 3D t-SNE plot generally shows a large number of unique patient signatures that do not overlap on one another. P_3 is distinguished by its strong negative position on Component 2, P_2 dominates in the positive of Components 1 and 2, and P_1 is kept on the negative of Component 1. This validates earlier PCA and 2D t-SNE results of individual profiles and demonstrates that each patient's disease-correlation structure is distinct.

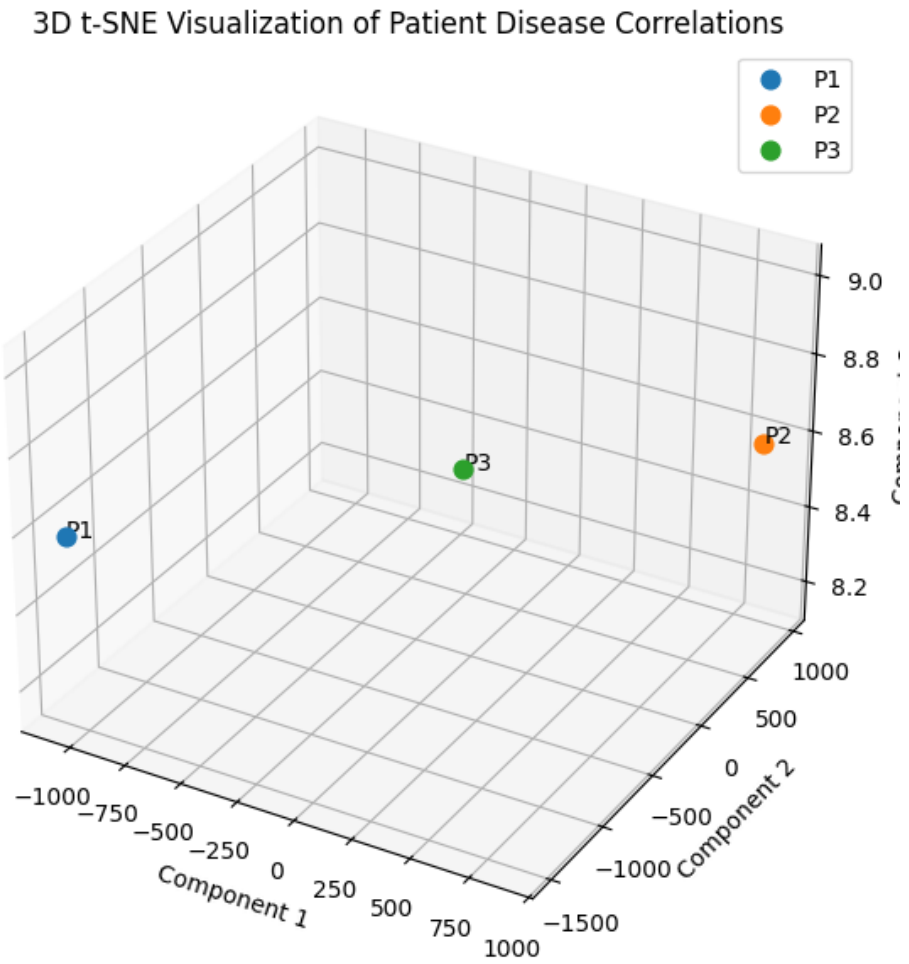

Figure 13

The three patient profiles (P_1, P_2, and P_3) are mapped into reduced space by the 3D t-SNE representation of normalized disease correlations in Fig. 14, where normalization ensures that values from various patients are comparable. Each point represents the general correlation structure of a single patient, while the plot's separation represents the variation in each patient's normalized structure. The use of numbers highlights distinct and well-expressed grouping.
P_1 (blue), which is located at roughly 1050, -1800, and 1.02 on the far right, shows a significant divergence on Components 1 and 2. On the left, P_2 (orange) has a value of roughly (-600, 700, 1.11), and P_3 (green) is close to (200, 150, 1.08). Component 3's vertical axis is incredibly narrow, measuring 1.02 to 1.11, suggesting that Components 1 and 2 account for the majority of the significant variance. The geographical distances highlight relative similarity. With a distance of roughly 800 units on Component 1 and 550 on Component 2, P_2 and P_3 are closer to one another, suggesting that their normalized correlation structure is rather comparable. However, P_1 is much farther away: P_1 and P_2 are separated by over 1600 units while on Component 2 and about 1950 when on Component 2. Because of this, P_1 is the most distinctive profile, whereas P_2 and P_3 are a fairly close pair. Overall, the 3D data's normalized t-SNE depiction supports earlier findings: While P_2 and P_3 have more comparable, though not identical, structures, P_1 has a unique disease-correlation signature. In order to concentrate interpreterability on Components 1 and 2 as the primary axes of separation, normalization removes variation in Component 3.

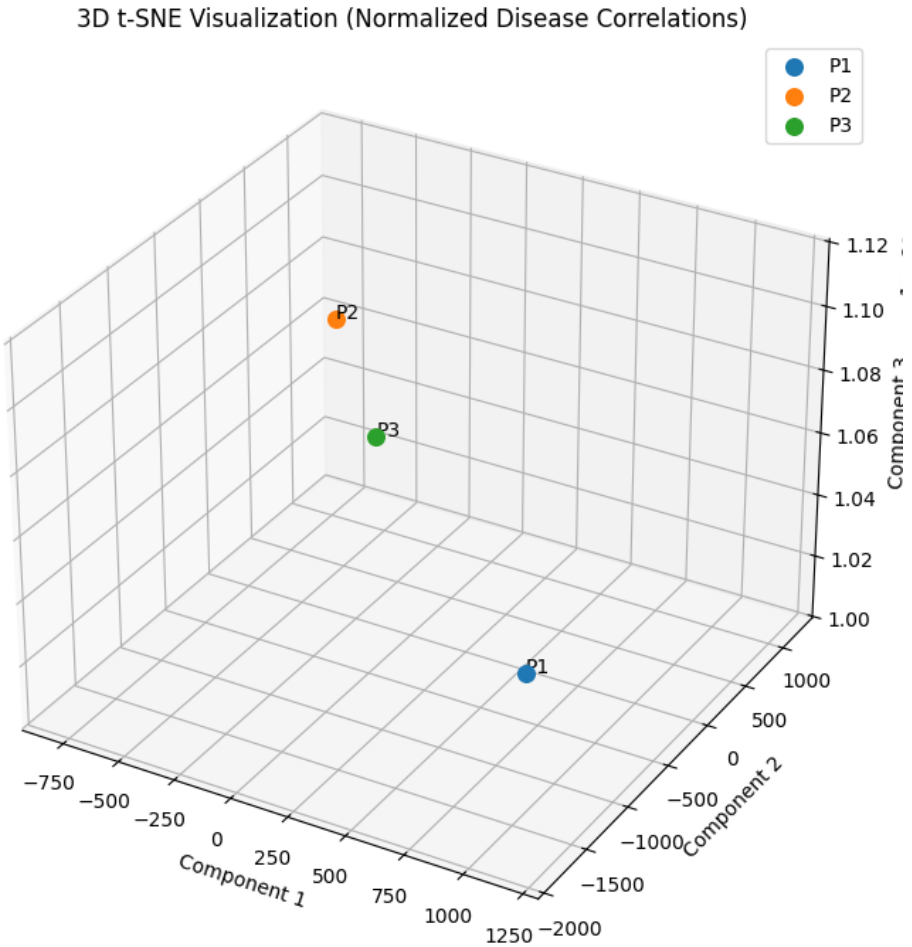

Figure 14

6. Analysis of Patient Disease Correlation and Clustering Results

The effectiveness of a QPNRS correlation metric in medical diagnostics is demonstrated by this discussion. According to Table III Analysis, the approach can diagnose P_1 with tuberculosis and P_2 and P_3 with typhoid. Additionally, P_1 is always isolated whereas P_2 and P_3 are always clustered by the clustering techniques (K-Means and K-Means++), confirming the commonalities in patient profiles. Dimensionality reduction (PCA and t-SNE) is used to visualize these particular groupings, while 3D projections provide nuanced information. The use of correlation measurements in patient diagnosis and categorization is firmly based on this multifaceted study.

6.1. Aspect-based Analysis

The diagnostic strength of the correlation measures and the underlying patient profile groupings are the two main sections of the analysis that can be distinguished.
Diagnostic Accuracy: As Table 3 demonstrates, the primary outcome should be the successful application of the QPNRS correlation measure for diagnosis. The disease with the highest correlation rate for each patient is selected to provide an unambiguous diagnosis: TB in P_1 (0.901), typhoid in P_2 and P_3 (0.878 and 0.869, respectively). This demonstrates how the technique can be applied to de-identify (de-ambiguate) complex symptom data into a particular, measurable diagnostic tool. Profile Similarity and Distinction: Internal structures are indicated in part by the correlation table itself. Typhoid has the highest values, but the correlation patterns of P_2 and P_3 are quite similar across all disease cases. In contrast to the other two profiles, P_1

is particularly connected with tuberculosis, which makes it stand out. The further clustering techniques validate the initial notion, which is this observation in terms of aspect.

6.2. Method-based Analysis

To verify the diagnostic findings and familiarize themselves with the data structure, a variety of unsupervised learner types were employed. Figures 1-3 show K-Means Clustering and K-Means++ Clustering: Two strong clusters, such as one of P_2, P_3, and another of P_1, were consistently found in both raw and adjusted data by both non-optimized and optimized K-Means algorithms. This provides excellent unsupervised evidence, showing that P_2 and P_3 are substantially similar to each other and fundamentally different from P_1. The strength of this classification is demonstrated by the consistency of the algorithms and the data pre-processing. PCA (Figs. 11, 12): t-SNE (Figs. 10, 13, 14): It is clear from the way the patient profiles occupy distinct and diverse positions in the component space that the PCA provided a linear projection to maximize variance. It was discovered that PC1 was significantly impacted by the TB-Typhoid contrast (between P_1 and P_2), whereas PC2 was impacted by another component of the correlation structure (between P_3 and P_2). A non-linear projection that placed emphasis on local similarities was provided by t-SNE. It showed that each of the three patients was completely different in 2D (Fig. 10). However, the 3D normalized t-SNE (Fig. 14) gave a more detailed picture, showing that P_2 and P_3 were more closely associated to one another than to P_1, which validated the clusters' findings.

6.3. Dimensional-based Analysis

The inquiry has used a range of perspectives in order to completely understand the results. Table 3-2D Tabular Data: The patient vs. disease table in two dimensions was the subject of the initial analysis. Although this is simple (and directly diagnosable), it is less successful in illustrating the global pattern of patient similarity. Figures 1, 2, 10, and 11 show 2D projections: For the data to be shown, it had to be reduced to two dimensions. High-dimensional correlations are easily explained by intuition, and it was possible to show the main separation ($P_1/P_2/P_3$ in clustering; the three are discrete in PCA/t-SNE). Figure 12, Figure 13, Figure 14: 3D projections. The inclusion of a third dimension added greater detail. It was clear from 3D PCA (Fig. 12) that the third component (PC3) did not significantly increase patient separation, confirming that the first two components provide the primary discriminatory information. The richest to examine were the 3D t-SNE plots (Figs. 13, 14), and it was intriguing to see how normalization (Fig. 14) might highlight the relationship between P_2 and P_3 and how, in the absence of the outlier P_1, they obviously constitute a cohesive subgroup of their own.

7. Conclusion

In summary, our multi-dimensional analysis has successfully demonstrated that the proposed QPNRS correlation measure is a useful tool for medical diagnosis since it may transform ambiguous symptoms into specific diagnostic findings. Furthermore, K-Means clustering has shown reliable and consistent findings in a variety of settings, providing strong support for unsupervised patient stratification. In addition to confirming the distinct nature of the patient P_1 diagnostic profile, the 2D and 3D visualizations created by integrating PCA and t-SNE offered a more comprehensive and multifaceted understanding of the relative similarities and differences among the three patient profiles. A logical and validated analytical pathway is produced by the entire analysis process, which includes clustering, correlation calculation, and visualization.
Despite the positive results, this study had the following limitations: Small sample size: Three patient profiles (P_1, P_2, and P_3) are used in the analysis. Producing statistical significance or drawing broad inferences about the method's performance at the population level is insufficient.

Limited Disease scope: Only four diseases were used in the testing of the diagnostic model. It has not yet been proven to be accurate or applicable in a clinical practice setting when dozens or even hundreds of conditions need to be taken into account. Dependency on Correlation Thresholds: The diagnostic rule would just employ the highest correlation value and be basic in nature. It ignores situations when the two largest correlations are quite similar, which could indicate comorbidity or a rare case. Interpretability of Reduced Dimensions: Principal components and t-SNE dimensions are abstract and cannot be readily converted into specific, clinically actionable symptoms or traits, even when PCA and t-SNE plots demonstrate clear separations.

The following approaches are recommended for use in subsequent study in order to get around the constraints and advance this research: Validation on an Extensive Data Set: Testing this method on a sizable real-world data set of patient records is the next crucial step in ensuring its external validity, strength, and diagnostic quality. Illness Taxonomy: To assess the model's scalability and discriminatory power in challenging diagnostic situations, it must be compared to a more comprehensive and realistic illness collection. Enhancement of Advanced Diagnostic Rules: To improve the diagnostic outcome, future research should investigate more complex decision rules, such as those that apply machine learning classifiers to the correlation data, incorporate a confidence interval, or take into account the shape of the correlation vector. Clinical Integration and Interpretation: Research on how to improve the model's interpretability for clinicians should be conducted. This could involve attempting to map high component loadings onto the initial symptoms or developing a hybrid system that combines expert knowledge with this quantitative approach. Other Methods Investigation: To see if they provide any new information about patient subgroups, different clustering techniques (like DBSCAN and Hierarchical Clustering) and dimensionality reduction techniques (like UMAP) may be investigated. For future work, neutrosophic soft sets could model the indeterminacy in non-Hermitian topological phase transitions and the robustness of Majorana modes [38].

Acknowledgements

The authors extend their appreciation to Prince Sattam bin Abdulaziz University for funding this research work through the project number (PSAU/2025/01/34111).

Authors' Contributions: All authors read and approved the final manuscript.

Funding: Prince Sattam bin Abdulaziz University through the project number (PSAU/2025/01/34111).

Conflicts of Interest: The authors declare that they have no conflicts of interest to report regarding the present study.

References

[1] F. Smarandache, A Unifying Field in Logics: Neutrosophy: Neutrosophic Probability, Set and Logic. Rehoboth, MA, *USA: American Research Press*, 1999.

[2] J. Ye, "Vector similarity measures of simplified neutrosophic sets and their application in multicriteria decision making," *Int. J. Fuzzy Syst.*, vol. 16, no. 2, pp. 204-211, 2014.

[3] L. A. Zadeh, "Fuzzy sets," *Inf. Control*, vol. 8, no. 3, pp. 338-353, 1965.

[4] K. Atanassov, "Intuitionistic fuzzy sets," *Fuzzy Sets Syst.*, vol. 20, no. 1, pp. 87-96, 1986.

[5] K. Atanassov and G. Gargov, "Interval valued intuitionistic fuzzy sets," *Fuzzy Sets Syst.*, vol. 31, no. 3, pp. 343-349, 1989.

[6] S. Broumi and F. Smarandache, "Several similarity measures of neutrosophic sets," *Neutrosophic Sets Syst.*, vol. 1, no. 1, pp. 54-62, 2013.

[7] P. Majumdar and S. K. Samanta, "On similarity and entropy of neutrosophic sets," *J. Intell. Fuzzy Syst.*, vol. 26, no. 3, pp. 1245-1252, 2014.

[8] J. Ye, "Similarity measures between interval neutrosophic sets and their applications in multicriteria decision-making," *J. Intell. Fuzzy Syst.*, vol. 26, no. 1, pp. 165-172, 2014.
[9] J. Ye, "Multiple attribute group decision-making method with completely unknown weights based on similarity measures under single valued neutrosophic environment," *J. Intell. Fuzzy Syst.*, vol. 27, no. 1, pp. 2927-2935, 2014.
[10] G. Kou and C. Lin, "A cosine maximization method for the priority vector derivation in AHP," *Eur. J. Oper. Res.*, vol. 235, no. 1, pp. 225-232, 2014.
[11] H. Wang, F. Smarandache, Y. Zhang, and R. Sunderraman, "Single valued neutrosophic sets," *Multispace Multistruct.*, vol. 4, pp. 410-413, 2010.
[12] H. Wang, F. Smarandache, Y. Zhang, and R. Sunderraman, Interval Neutrosophic Sets and Logic: Theory and Applications in Computing. *Phoenix, AZ, USA: Hexis*, 2005.
[13] J. Ye, "Cosine similarity measures for intuitionistic fuzzy sets and their applications," *Math. Comput. Model.*, vol. 53, no. 1-2, pp. 91-97, 2011.
[14] F. Smarandache, "Neutrosophic set, a generalisation of the intuitionistic fuzzy sets," *Int. J. Pure Appl. Math.*, vol. 24, no. 3, pp. 287-297, 2005.
[15] F. Smarandache, Plithogeny, Plithogenic Set, Logic, Probability, and Statistics. *Brussels, Belgium: Pons Editions*, 2018.
[16] A. A. Salama and S. A. Alblowi, "Neutrosophic Set and Neutrosophic Topological Spaces," *IOSR J. Math.*, vol. 3, no. 4, pp. 31-35, 2012.
[17] J. Ye, "Trapezoidal neutrosophic set and its application to multiple attribute decision-making," *Neural Comput. Appl.*, vol. 26, no. 5, pp. 1157-1166, 2015.
[18] J. J. Peng, J. Q. Wang, X. H. Wu, J. Wang, and X. H. Chen, "A novel similarity measure in interval-neutrosophic set and its application in decision-making," *Cogn. Comput.*, vol. 7, no. 4, pp. 437-447, 2015.
[19] M. Abdel-Basset, M. Gunasekaran, M. Mohamed, and N. Chilamkurti, "A framework for risk assessment, management and evaluation: Economic tool for quantifying risks in supply chain," *Future Gener. Comput. Syst.*, vol. 90, pp. 489-502, 2019.
[20] R. Sahin and A. Kucuk, "Generalized similarity measures for interval neutrosophic sets and their application to multicriteria decision-making," *Neutrosophic Sets Syst.*, vol. 12, pp. 42-52, 2016.
[21] J. Ye, "Single-valued neutrosophic cross-entropy for multicriteria decision making problems," *Appl. Math. Model.*, vol. 38, no. 3, pp. 1170-1175, 2014.
[22] J. C. Bezdek, R. Ehrlich, and W. Full, "FCM: The fuzzy c-means clustering algorithm," *Comput. Geosci.*, vol. 10, no. 2-3, pp. 191-203, 1984.
[23] Y. Guo and A. Sengur, "A novel image segmentation algorithm based on neutrosophic similarity clustering," *Appl. Soft Comput.*, vol. 25, pp. 391-398, 2014.
[24] L. van der Maaten and G. Hinton, "Visualizing Data using t-SNE," *J. Mach. Learn. Res.*, vol. 9, no. Nov, pp. 2579-2605, 2008.
[25] F. Smarandache, "Plithogenic Probability & Statistics are generalizations of Multi Variate Probability & Statistics," *Neutrosophic Sets Syst.*, vol. 43, pp. 280-289, 2021.
[26] F. Smarandache,. A Unifying Field in Logics: Neutrosophic Logic, Neutrosophy, Neutrosophic Set, Probability, and Statistics. *American Research Press*, 2000.
[27] H. Wang, F. Smarandache, Y. Zhang, & R. Sunderraman. Single-valued neutrosophic sets. *Review*, vol. 17, pp. 10-14, 2010.
[28] S. A. El-Sheikh, & A. M. Abd El-Latif. Decompositions of some types of supra soft sets and soft continuity. *Int. J. Math. Trends Technol.*, vol. 9(1), pp. 37-56, 2014.
[29] A. M. Abd El-Latif, & R. A. Hosny. Supra open soft sets and associated soft separation axioms. *Int. J. Adv. Math.*, vol. 6, pp. 68-81, 2017.
[30] A. M. Abd El-Latif. Soft supra strongly generalized closed sets. *J. Intell. Fuzzy Systems*, vol. 31(3), pp. 1311-1317, 2016.
[31] A. M. Abd El-Latif, & R. A. Hosny. Soft supra extra strongly generalized closed sets. *An.*

Univ. Oradea Fasc. Mat., vol. 24(1), pp. 103-112, 2017.

[32] A. M. Abd El-Latif, & R. A. Hosny. Supra soft separation axioms and supra irresoluteness based on supra b-soft sets. *Gazi Univ. J. Sci.*, vol. 29(4), pp. 845-854, 2016.

[33] A. M. Abd El-Latif. Soft supra strongly *semi*$*$ generalized closed sets. *Ann. Fuzzy Math. Inform.*, vol. 13(1), pp. 63-71, 2017.

[34] A. M. Abd El-Latif, & R. A. Hosny. Supra semi open soft sets and associated soft separation axioms. *Appl. Math. Inf. Sci.*, vol. 10(6), pp. 2207-2215, 2016.

[35] A. M. Abd El-Latif. Supra soft b-connectedness II: Some types of supra soft b-connectedness. *Creat. Math. Inform.*, vol. 26(1), pp. 1-8, 2017.

[36] A. M. Abd El-Latif, & S. Karatas. Soft supra strongly generalized closed sets via soft ideals. *Appl. Math. Inf. Sci. Lett.*, vol. 5(1), pp. 21-26, 2016.

[37] A. M. Abd El-Latif. On soft supra compactness in supra soft topological spaces. *Tbil. Math. J.*, vol. 11(1), pp. 169-178, 2018.

[38] X. Ji, H. Geng, N. Akhtar, and X. Yang, Floquet engineering of point-gapped topological superconductors, *Phys. Rev. B*, vol. 111(19), p. 195419, 2025.