Received: xxx **Accepted:** xxx **Published:** xxx.

DOI. xxxxxxx

xxx Vol. xxx, No. xxx, (1-19)

Research Article

Open Access

Control and Optimization in Applied Mathematics - COAM

Pythagorean Fuzzy Sets for Credit Risk Assessment: A Novel Approach to Predicting Loan Default

Amal Kumar Adak[1] ✉, **Nil Kamal**[2]

[1]Department of Mathematics, Ganesh Dutt College, Begusarai, India.

[2]Department of Mathematics, Lalit Narayan Mithila University, Darbhanga.

✉ **Correspondence:**
Amal Kumar Adak

E-mail:
gdc@lnmu.ac.in

Abstract. The incorporation of Pythagorean fuzzy sets into credit risk assessment represents a relatively innovative approach for predicting loan defaults, offering a more precise and adaptable tool for financial institutions. Key customer information—such as credit history, credit mix, credit utilization, duration of credit history, income level, and employment stability—is obtained as linguistic variables. These linguistic assessments are then transformed into Pythagorean fuzzy numbers. The combined Pythagorean fuzzy information is subsequently processed using the Technique for Order of Preference by Similarity to Ideal Solution (TOPSIS). This approach employs a modified accuracy function to determine the Pythagorean fuzzy positive ideal solution and the Pythagorean fuzzy negative ideal solution. For distance calculations within the TOPSIS framework, spherical distance measurements are utilized. Alternatives are ranked based on the relative closeness coefficient and an adjusted index, collectively facilitating decision-making. The practical applicability of the proposed model is demonstrated through an illustrative numerical example.

Keywords. Pythagorean fuzzy sets, Credit risk assessment, Spherical distance measurement, TOPSIS, Revised index method.

MSC. 16Y30; 03E72.

How to Cite
Kumar Adak, A., Kamal, N. (2023). "Pythagorean fuzzy sets for credit risk assessment: A novel approach to predicting loan default", Control and Optimization in Applied Mathematics, 10(): 1-19, doi: 10.30473/coam.2025.73505.1287

https://mathco.journals.pnu.ac.ir

1 Introduction

Financial institutions rely on credit rating models to inform their decision-making processes related to lending, investment, and risk management. Traditionally, these models are built using various statistical and machine learning techniques. However, such approaches often require extensive data and hinge on assumptions about data distribution. To address these limitations, this study proposes an integrated Pythagorean fuzzy credit rating model that enhances decision accuracy under uncertainty. Accurate assessment of credit risk is essential for mitigating potential losses and ensuring the stability and robustness of lending institutions. Traditional credit risk models predominantly rely on deterministic, crisp inputs, which inadequately capture the inherent uncertainty and vagueness present in financial decision environments. Recent advances in fuzzy logic and soft computing offer promising alternatives by accommodating uncertainty and providing more nuanced predictive capabilities.

Zadeh [32] introduced fuzzy set (FS) theory as a practical and effective method for modeling imprecision in multi-criteria decision analysis (MCDA) through membership functions within the interval $[0, 1]$. Nonetheless, in real-world scenarios, the complexity of human subjective assessments often prevents FSs from adequately representing assessment information. To address this, Atanassov [4] proposed intuitionistic fuzzy sets (IFSs), which extend FSs by incorporating both membership (MG) and non-membership (NMG) degrees.

In IFSs, assuming that $0 \leq \varphi + \varrho \leq 1$, the MG and the NMG are represented by (φ, ϱ). Recent research by Yager [29] expanded this framework to Pythagorean fuzzy sets (PFSs), where the pair (φ, ϱ) must satisfy $\varphi^2 + \varrho^2 \leq 1$, allowing for a broader range of uncertainty representation. For example, with $\varphi = 0.8$ and $\varrho = 0.3$, these values form a valid PFS, whereas they do not meet the stricter sum condition $\varphi + \varrho \leq 1$ required in IFSs. PFSs thus encompass a larger scope compared to IFSs, making them more flexible for modeling complex, uncertain, and imprecise real-world situations.

Several methods have been developed to measure the distance between PFSs. Hesamian [13] proposed a distance measurement technique to fuzzy numbers, while Zhang and Xu [34] considered only the membership, non-membership, and hesitation degrees—ignoring angles and the richer structure of PFSs. Li and Zeng [18] introduced a distance measure that accounts for membership, non-membership, level and direction of commitment, with further refinement by Zeng et al. [33] to include hesitation degree. However, both approaches primarily derive from IFSs and overlook the angular component of PFSs.

More recently, Wang et al. [28] presented a bidirectional projection model in PFSs that considers both magnitude and angular distances. Yu et al. [31] proposed a distance measure based on induced ordered weighted averaging (IOWA) operators, and Adak and Kumar [1] introduced a spherical distance metric for PFSs, leveraging the triplet $(\varphi, \varrho, \varpi)$ on the unit sphere, where

$\varpi = \sqrt{1 - \varphi^2 - \varrho^2}$. This approach situates PFSs on a spherical surface, capturing their full geometric structure.

Building on this foundation, research has further explored the application of fuzzy set theory in creditworthiness evaluations. For instance, Polishchuk et al. [24] enhanced fuzzy mathematical models to assess corporate creditworthiness, integrating linguistic terms and reflective decision-maker reasoning. Setiawan and Prihatini [27] simplified loan application evaluations by considering criteria such as loan amount, income, and collateral value. Makhazhanova et al. [20] emphasized the importance of operational specifics and financial unpredictability in small business credit assessments. To tackle these complexities and uncertainties involved, Roy and Shaw [25] proposed an integrated fuzzy credit rating model utilizing fuzzy best-worst method (fuzzy-BWM) for weighting factors and fuzzy-TOPSIS-Sort-C for ranking borrowers. Similarly, Yang and Yang [30] introduced a hybrid fuzzy firefly optimization classification (FFOC) system to improve credit evaluation accuracy. Addressing issues specific to Chinese coastal cities in Pearl River Delta—such as centralized data, data forgery, and transmission delays—Zhang et al. [35] developed a CDDC model based on fuzzy sets, while Chen et al. [9] designed a credit risk assessment index for Chinese real estate firms using hesitant fuzzy linguistic word sets and the PROMETHEE method.

Sartova et al. [26] proposed a creditworthiness evaluation model based on the Mamdani fuzzy inference approach, which performs well for uncertain, incomplete , and qualitative data. Similarly, Astuti et al. [3] applied the Mamdani fuzzy method to assess creditworthiness under the Kredit Usaha Rakyat (KUR) program as implemented by Bank Rakyat Indonesia (BRI). Biryukov et al. [7] introduced neural network models for managing bank loan portfolios, offering risk mitigation strategies even during dynamic financial changes and borrower instability. Lastly, Krasavtseva [16] proposed innovative credit risk assessment methods, classifying specialized lending projects using logic- and language-based algorithms for classifying lending projects, enhancing risk analysis.

Incorporating Pythagorean fuzzy sets into the TOPSIS method facilitates the handling of inherent uncertainty and imprecision when evaluating credit risk. When making a choice that takes into account more than one criterion, the TOPSIS technique might be helpful since it compares potential solutions to the best and worst case scenarios. This paper aims to develop a Pythagorean fuzzy TOPSIS framework for credit risk assessment, providing a robust decision-making tool that accounts for multiple criteria and their associated uncertainties. To demonstrate its practical utility, a case study is presented; comparing results obtained via relative proximity and updated index methods.

The remainder of this paper is organized as follows: Section 2 discusses relevant operations on Pythagorean fuzzy sets and various score functions. Section 3 elaborates on the mathematical formulation of spherical and normalized spherical distances between PFSs and PFSNs. Section 4 presents the credit risk assessment case study, analyzing factors such as credit history, credit mix credit utilization, credit history length, , income and employment stability etc., those are important for creditworthiness. Section 5 integrate Pythagorean fuzzy sets with the TOPSIS technique to address uncertainty in credit evaluation. Section 6 illustrates the application of the proposed model through a numerical example. Finally, Section 7 concludes with summary remarks.

2 Preliminaries and Definitions

This section provides an overview of the fundamental concepts of intuitionistic and Pythagorean fuzzy sets. Additionally, we introduce key score functions for Pythagorean fuzzy numbers that are essential for the developments discussed in this paper.

Definition 1. [4] An intuitionistic fuzzy set (IFS) I in a universal set X is expressed as

$$I = \{\langle \varsigma, \varphi_I(\varsigma), \varrho_I(\varsigma)\rangle : \varsigma \in X\},$$

where $\varphi_I : X \to [0,1]$, $\varrho_I : X \to [0,1]$ denote the membership grade (MG) and non-membership grade (NMG), respectively. They satisfy the condition $0 \leq \varphi_I(\varsigma) + \varrho_I(\varsigma) \leq 1$, for all $\varsigma \in X$. The degree of Indeterminacy is given by $\pi_I(\varsigma) = 1 - \varphi_I(\varsigma) - \varrho_I(\varsigma)$.

Definition 2. [29] A Pythagorean fuzzy set (PFS) P in X is defined as

$$P = \{\langle \varsigma, \varphi_P(\varsigma), \varrho_P(\varsigma)\rangle | \varsigma \in X\},$$

where $\varphi_P(\varsigma) : X \to [0,1]$ and $\varrho_P(\varsigma) : X \to [0,1]$ denote MG and NMG, respectively. $These satisfy the Pythagorean condition 0 \leq (\varphi_P(\varsigma))^2 + (\varrho_P(\varsigma))^2 \leq 1$. The indeterminacy is $\varpi_P(\varsigma) = \sqrt{1 - (\varphi_P(\varsigma))^2 - (\varrho_P(\varsigma))^2}$, and the order pair (φ, ϱ) represents a Pythagorean fuzzy number (PFN).

A visual representation of the spaces occupied by IFSs and PFSs is provided in Figure 1.

Yager introduced an alternative formulation for PFNs using a pair $p = \langle r, d\rangle$, where, $r \in [0,1]$ indicates the commitment strength, with higher values reflecting greater certainty and less ambiguity. d represents the direction of commitment, linked to the degree of hesitancy. In this representation, the components φ and ϱ relate to r and an angle θ via:

$$\varphi = r\cos\theta, \quad \varrho = r\sin\theta,$$

with the angle d given by

$$d = 1 - \frac{2}{\pi}\theta.$$

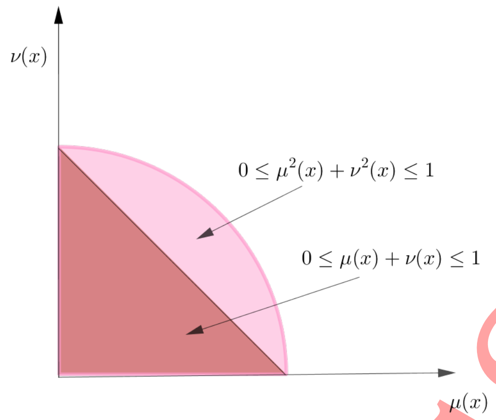

Figure 1: Spaces for IFSs and PFSs.

Example 1. Consider PFN $p = \langle 0.6, 0.5 \rangle$. Then, the parameters are computed as follows:

$$\varphi_p = 0.6, \varrho_p = 0.5, \varpi_p = \sqrt{1 - (0.6)^2 - (0.5)^2} = 0.39,$$

$$r = \sqrt{(0.6)^2 + (0.5)^2} = 0.61, \theta = \arctan\left(\frac{0.5}{0.6}\right) = 0.695,$$

and

$$d = 1 - 2 \times \frac{0.695}{\pi} = 0.557.$$

2.1 Operations on PFNs

Let $p = \langle \varphi, \varrho \rangle$, $p_1 = \langle \varphi_1, \varrho_1 \rangle$, and $p_2 = \langle \varphi_2, \varrho_2 \rangle$ be three PFNs. The following operations are defined:

i. $\bar{p} = \langle \varrho, \varphi \rangle$,

ii. $p_1 \cup p_2 = \langle \max\{\varphi_1, \varphi_2\}, \min\{\varrho_1, \varrho_2\} \rangle$,

iii. $p_1 \cap p_2 = \langle \min\{\varphi_1, \varphi_2\}, \max\{\varrho_1, \varrho_2\} \rangle$.

In decision-making applications, ranking PFNs based on their MG and NMG is essential. Several functions serve this purpose:

Definition 3. For $p = \langle \varphi, \varrho \rangle$, the score function $s(p)$ is defined as:

$$s(p) = (\varphi)^2 - (\varrho)^2, \tag{1}$$

where $s(p) \in [-1, 1]$.

Definition 4. The accuracy function $a(p)$ for $p = \langle \varphi, \varrho \rangle$ is given by:

$$a(p) = (\varphi)^2 + (\varrho)^2, \tag{2}$$

where $h(p) \in [0, 1]$.

Definition 5. The modified accuracy function $v(p)$ is expressed as:

$$v(p) = \frac{1}{2} + r\left(d - \frac{1}{2}\right) = \frac{1}{2} + r\left(\frac{1}{2} - \frac{2\theta}{\pi}\right). \tag{3}$$

3 Distance Measurement Methods for PFNs

In this section, we introduce various distance measurement techniques—namely, spherical, normalized spherical, and weighted spherical distances— for PFSs and PFNs. These distance measure are integral to the proposed integrated TOPSIS framework.

3.1 Spherical Distance Measurement Method for PFNs

Let $p = \langle \varphi, \varrho \rangle$ be a PFN satisfying the condition $0 \leq \varphi^2 + \varrho^2 \leq 1$, with the hesitancy degree $\varpi = \sqrt{1 - \varphi^2 - \varrho^2}$. Consequently, the relation $\varphi^2 + \varrho^2 + \varpi^2 = 1$ holds, placing the triplet $(\varphi, \varrho, \varpi)$ on the surface of a unit sphere centered at the origin.

Assuming these parameters delineate a point on the spherical surface of unit radius, the spherical distance between two points on this surface can be defined as follows:

Definition 6. [1] The spherical distance between two points $A = (x_1, y_1, z_1)$ and $C = (x_2, y_2, z_2)$ on the same spherical surface is given by:

$$D_{SP}(A, C) = \arccos\left\{1 - \frac{1}{2}\left[(x_1 - x_2)^2 + (y_1 - y_2)^2 + (z_1 - z_2)^2\right]\right\}. \tag{4}$$

Incorporating this expression, the spherical distance between two PFNs $p_1 = \langle \varphi_1, \varrho_1 \rangle$ and $p_2 = \langle \varphi_2, \varrho_2 \rangle$ (with associated hesitancy degrees ϖ_1 and ϖ_2) is defined as:

Definition 7. [1] The spherical distance calculated as:

$$D_{SP}(p_1, p_2) = \frac{2}{\pi} \arccos\left\{1 - \frac{1}{2}[(\varphi_1 - \varphi_2)^2 + (\varrho_1 - \varrho_2)^2 + (\varpi_1 - \varpi_2)^2]\right\}. \tag{5}$$

The factor $\frac{2}{\pi}$ is introduced to obtain the distance value within the range $[0, 1]$.
Given that $\varphi_i^2 + \varrho_i^2 + \varpi_i^2 = 1$ for $i = 1, 2$, this expression can be simplified to:

$$D_{SP}(p_1, p_2) = \frac{2}{\pi} \arccos\left[\varphi_1\varphi_2 + \varrho_1\varrho_2 + \varpi_1\varpi_2\right]. \tag{6}$$

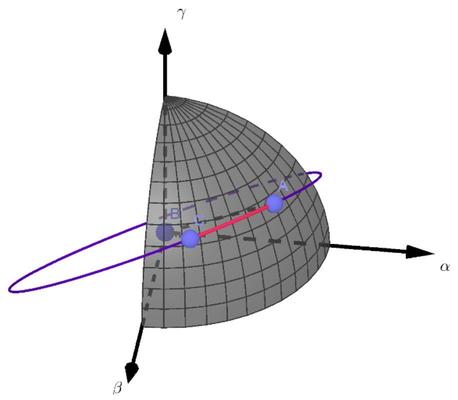

Figure 2: Spherical distance illustration.

Furthermore, when comparing sets of PFNs over the universe of discourse $X = \varsigma_1, \varsigma_2, \ldots, \varsigma_n$, the set-level spherical distances are defined as follows:

Definition 8. [1] For two PFN sets

$$P = \{\varsigma_i, \langle \varphi_P(\varsigma_i), \varrho_P(\varsigma_i)\rangle : \varsigma_i \in X\},$$

and

$$Q = \{\varsigma_i, \langle \varphi_Q(\varsigma_i), \varrho_Q(\varsigma_i)\rangle : \varsigma_i \in X\},$$

within the universe of discourse $X = \{\varsigma_1, \varsigma_2, \ldots, \varsigma_n\}$. The spherical and normalized spherical distances are given by:

- Spherical Distance:

$$D_{SP}(P,Q) = \frac{2}{\pi} \sum_{i=1}^{n} \arccos\left[\varphi_P(\varsigma_i)\varphi_Q(\varsigma_i) + \varrho_P(\varsigma_i)\varrho_Q(\varsigma_i) + \varpi_P(\varsigma_i)\varpi_Q(\varsigma_i)\right], \quad (7)$$

  where $0 \leq D_{SP}(P,Q) \leq n$.

- Normalized Spherical Distance:

$$D_{NSP}(P,Q) = \frac{2}{n\pi} \sum_{i=1}^{n} \arccos\left[\varphi_P(\varsigma_i)\varphi_Q(\varsigma_i) + \varrho_P(\varsigma_i)\varrho_Q(\varsigma_i) + \varpi_P(\varsigma_i)\varpi_Q(\varsigma_i)\right], \quad (8)$$

  where $0 \leq D_{NSP}(P,Q) \leq 1$.

Example 2. Let $p_1 = \langle 0.9, 0.2\rangle$ and $p_2 = \langle 0.7, 0.3\rangle$ be two PFNs. Their spherical distance is computed as:

Definition 9. Given two sets of PFNs: $P_1 = (\varphi_{1j}, \varrho_{1j})$, $P_2 = (\varphi_{2j}, \varrho_{2j})$, $j = 1, 2, \ldots, n$, be two sets of PFNs, with criteria weights $w = (w_1, w_2, \ldots, w_n)^T$ such that $0 \leq w_i \leq 1$ and $\sum_{i=1}^{n} w_i = 1$, the weighted normalized spherical distance is defined as:

$$D'_{NSP}(P_1, P_2) = \frac{2}{n\pi} \sum_{j=1}^{n} w_j \arccos\left[\varphi_{1j}\varphi_{2j} + \varrho_{1j}\varrho_{2j} + \varpi_{1j}\varpi_{2j}\right]. \tag{9}$$

Example 3. Suppose

$$P_1 = \{\langle 0.6, 0.3\rangle, \langle 0.8, 0.2\rangle, \langle 0.5, 0.4\rangle\},$$

and

$$P_2 = \{\langle 0.7, 0.2\rangle, \langle 0.7, 0.3\rangle, \langle 0.9, 0.1\rangle\},$$

with weights $w = \{0.2, 0.5, 0.3\}$. The weighted distance computes as:

$$\begin{aligned}
& D'_{NSP}(P_1, P_2) = \\
& \frac{2}{3\pi}\Big[0.2 \times \arccos\Big((0.6 \times 0.7) + (0.3 \times 0.2) + (\sqrt{1 - 0.6^2 - 0.3^2} \times \sqrt{1 - 0.7^2 - 0.2^2})\Big) \\
+ \;& 0.5 \times \arccos\Big((0.8 \times 0.7) + (0.2 \times 0.3) + (\sqrt{1 - 0.8^2 - 0.2^2} \times \sqrt{1 - 0.7^2 - 0.3^2})\Big) \\
+ \;& 0.3 \times \arccos\Big((0.5 \times 0.9) + (0.4 \times 0.1) + (\sqrt{1 - 0.5^2 - 0.4^2} \times \sqrt{1 - 0.9^2 - 0.1^2})\Big)\Big] \\
= \;& 0.0499.
\end{aligned}$$

4 Case Study

Creditworthiness refers to the capacity of a borrower to fulfill debts and loanobligations punctually. It serves as an indicator of the individual's credibility and reliability in managing credit-related responsibilities. The assessment of creditworthiness involves analyzing multiple factors, including:

i. *Payment history* reflects an individual's record of settling credit accounts, loans, and other debts. It constitutes a fundamental element of credit reports and credit scores, accounting for approximately 35% of the overall score. This indicator encompasses data such as:

   - Bankruptcies and foreclosures,
   - Missed or late payments (including frequency and duration),
   - Accounts transferred to collections,
   - Consistent on-time payments.

A positive payment history is characterized by timely payments, absence of late or missed payments, and no collections or bankruptcies. Conversely, a poor record involves frequent delays or missed payments, collections, and negative credit events such as bankruptcies. To uphold a strong payment history, borrowers are advised to make punctual payments, utilize reminders or automate transactions, and communicate proactively with lenders regarding any financial difficulties. An exemplary payment history demonstrates financial responsibility and can enhance credit scores, while a negative record may adversely impact future credit access.

ii. *Credit utilization* or the utilization ratio, measures the proportion of credit relative to the total available credit line. It is computed as:

$$\text{Credit Utilization Ratio} = \left(\frac{\text{Total Outstanding Credit Balance}}{\text{Total Credit Limit}}\right) \times 100.$$

This metric is pivotal in credit scoring models, indicating the borrower's management of credit. Optimal utilization levels are generally considered below 30%, categorized as good, while ratios between 30% and 50% are fair, and exceeding 50% is deemed poor. Maintaining a low utilization ratio demonstrates responsible credit behavior and can positively influence credit scores.

iii. *Credit history length*:
The duration of an individual's credit activity, often referred to as the length of credit history, accounts for approximately 15% of the credit score. It considers factors such as:

- The average age of all credit accounts,
- The age of the oldest account,
- T he overall length of credit history.

A more extended credit history typically suggests experience in managing credit responsibly, indicating stability. Shorter histories may imply limited experience and are often viewed less favorably by lenders due to perceived higher risk. Longer credit histories tend to correlate with greater financial stability and better credit scores.

iv. *Credit mix*, also known as credit diversity, refers to the variety of credit types held by an individual, including:

- Credit cards,
- Installment loans (e.g., personal loans, mortgages, car loans),
- Revolving credit (e.g., lines of credit, home equity loans),
- Open credit accounts (e.g., utility bills, rent).

A diversified credit profile demonstrates the borrower's ability to handle different types of credit responsibly, reducing perceived risk. Limited diversity might indicate inexperience with various credit forms, which can negatively influence credit evaluations. A well-managed, diverse credit portfolio supports a higher credit score.

v. *Credit Inquiries*:
Inquiries refer to the instances where lenders or creditors access a borrower's credit report in the past two years. They are categorized as:

- Soft inquiries: Do not impact credit scores and include checks made by borrowers themselves, pre-approval assessments, and employment background verifications.
- Hard inquiries: Can temporarily reduce credit scores and occur when applying for new credit products, such as loans, credit cards, or mortgages.

While inquiries, especially hard ones, can influence credit scores slightly, their impact is generally minor compared to factors like payment history and utilization.

vi. *Income and Employment Stability:*
Income levels and employment stability are critical in credit risk evaluation, as they reflect the borrower's capacity to meet ongoing obligations. Factors considered include:

- Income amount and consistency over time,
- Income growth or decline,
- Duration of current employment,
- Employment stability and industry stability,
- Job history and frequency of transitions.

Demonstrating consistent employment and sufficient income bolsters confidence in the borrower's repayment ability.

vii. *Debt-to-income ratio:*
The Debt-to-Income (DTI) ratio compares monthly debt obligations to gross monthly income and is a significant indicator of financial health. It is calculated as:

$$\text{DTI ratio} = \frac{\text{Total Monthly Debt Payments}}{\text{Gross Monthly Income}}.$$

Lower DTI ratios are favored, as they suggest better capacity to service debt, thereby improving the likelihood of loan approval.

4.1 Linguistic Variables in Terms of PFNs

Decision-makers have translated linguistic assessments into PFNs to quantitatively estimate customer performance across various criteria.

Figure 3 illustrates the relationship between evaluation criteria and their corresponding linguistic variables:

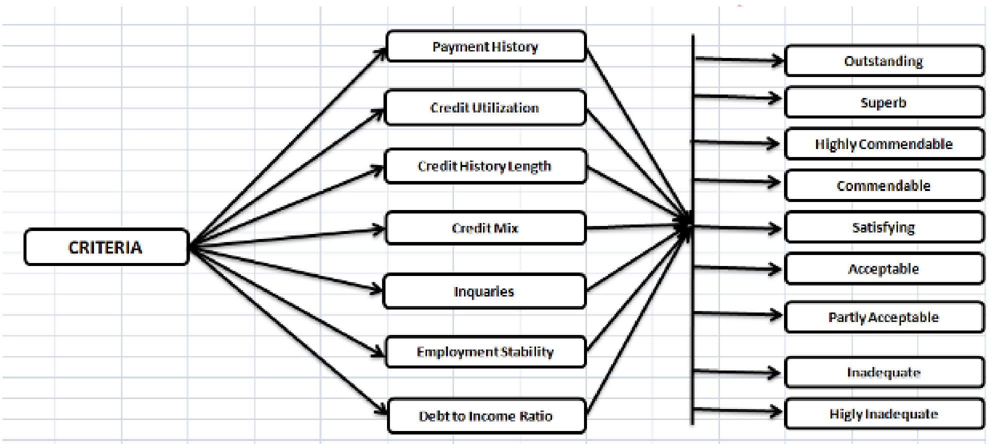

Figure 3: Criteria and linguistic variable.

Table 1 presents the linguistic concepts expressed through PFNs:

Table 1: Linguistic concepts expressed through PFNs

Linguistic Term	PFNs
Outstanding	$\langle 0.98, 0.20\rangle$
Superb	$\langle 0.87, 0.35\rangle$
Highly commendable	$\langle 0.70, 0.40\rangle$
Commendable	$\langle 0.65, 0.45\rangle$
Satisfying	$\langle 0.50, 0.55\rangle$
Acceptable	$\langle 0.40, 0.70\rangle$
Partly Acceptable	$\langle 0.36, 0.80\rangle$
Inadequate	$\langle 0.25, 0.87\rangle$
Highly Inadequate	$\langle 0.20, 0.98\rangle$

Table 1 demonstrates how decision-makers assign linguistic evaluations to each customer across different criteria, utilizing the PFNs specified above.

5 Methodology

This section presents a multi-criteria decision-making issue using information represented by PFNs and use the spherical distance measurement approach for resolution. For each criterion and alternative, construct a decision matrix using PFSs. Each entry of matrix represents membership and non-membership grade of each alternative concerning each criterion.

Let $\mathcal{A} = \{\mathcal{A}_1, \mathcal{A}_2, \ldots, \mathcal{A}_m\}$, where $m \geq 2$ and $\Gamma = \{\Gamma_1, \Gamma_2, \ldots, \Gamma_n\}$, with $n \geq 2$, denote set of alternatives and criteria respectively. Weight $w = (w_1, w_2, \ldots, w_n)^T$ with $0 \leq w_i \leq 1$ for all i, satisfies $\sum_{i=1}^{n} w_i = 1$.

Let the PFNs $\langle \varphi_{ij}, \varrho_{ij} \rangle$ represent the assessment value of i-th alternative and j-th criterion, such that $\Gamma_j(\varsigma_i) = \langle \varphi_{ij}, \varrho_{ij} \rangle$. Additionally, $R = (\Gamma_j(\varsigma_i))_{m \times n}$, where

$$R = \begin{bmatrix} \langle \varphi_{11}, \varrho_{11} \rangle & \langle \varphi_{12}, \varrho_{12} \rangle & \cdots & \langle \varphi_{1n}, \varrho_{1n} \rangle \\ \langle \varphi_{21}, \varrho_{21} \rangle & \langle \varphi_{22}, \varrho_{22} \rangle & \cdots & \langle \varphi_{2n}, \varrho_{2n} \rangle \\ \cdots & \cdots & \cdots & \cdots \\ \langle \varphi_{m1}, \varrho_{m1} \rangle & \langle \varphi_{m2}, \varrho_{m2} \rangle & \cdots & \langle \varphi_{mn}, \varrho_{mn} \rangle \end{bmatrix}.$$

5.1 Process of the Proposed Method

To address multi-criteria decision making (MCDM) problems under a Pythagorean fuzzy framework, we propose a Pythagorean fuzzy TOPSIS technique. The fundamental principle of TOPSIS is that the optimal alternative should be closest to the positive ideal solution (PIS) while simultaneously while being farthest from the negative ideal solution (NIS).

This approach involves the computation of the Pythagorean fuzzy positive ideal solution (PFPIS) and the Pythagorean fuzzy negative ideal solution (PFNIS). Let J_1 and J_2 denote the sets of benefit criteria and cost criteria respectively. The determination of PFPIS and PFNIS is performed utilizing a modified accuracy function $v(p)$. We denote the PFPIS and PFNIS as ς^+ and ς^-, respectively, calculated as follows:

$$\varsigma^+ = \begin{cases} \Gamma_j = \max_i S(\Gamma_j(\varsigma_i)) & \text{for } j \in J_1, \\ \Gamma_j = \min_i S(\Gamma_j(\varsigma_i) & \text{for } j \in J_2, \end{cases} \tag{10}$$

$$\varsigma^- = \begin{cases} \Gamma_j = \max_i S(\Gamma_j(\varsigma_i)) & \text{for } j \in J_1, \\ \Gamma_j = \min_i S(\Gamma_j(\varsigma_i)) & \text{for } j \in J_2. \end{cases} \tag{11}$$

Subsequently, the distances of each alternative ς_i from PFPIS and PFNIS are computed as:

$$D_{NSP}(\varsigma_i, \varsigma^+) \quad \text{and} \quad D_{NSP}(\varsigma_i, \varsigma^-),$$

using the formulas below, derived from the normalized Pythagorean fuzzy similarity measure:

$$\begin{aligned} D_{NSP}(\varsigma_i, \varsigma^+) &= \sum_{j=1}^{n} D_{NSP}(\Gamma_j(\varsigma_i), \Gamma_j(\varsigma^+)) \\ &= \frac{2}{n\pi} \sum_{j=1}^{n} w_j \arccos(\varphi_{ij}\varphi_j^+ + \varrho_{ij}\varrho_j^+ + \varpi_{ij}\varpi_j^+), \quad i = 1, 2, \ldots, n. \end{aligned} \tag{12}$$

$$\begin{aligned} D_{NSP}(\varsigma_i, \varsigma^-) &= \sum_{j=1}^{n} D_{NSP}(\Gamma_j(\varsigma_i), \Gamma_j(\varsigma^-)) \\ &= \frac{2}{n\pi} \sum_{j=1}^{n} w_j \arccos(\varphi_{ij}\varphi_j^- + \varrho_{ij}\varrho_j^- + \varpi_{ij}\varpi_j^-), \quad i = 1, 2, \ldots, n. \end{aligned} \tag{13}$$

where a smaller value of $D_{NSP}(\varsigma_i, \varsigma^+)$ indicates a preferable alternative, and conversely, a larger $D_{NSP}(\varsigma_i, \varsigma^-)$ is desirable for the negative ideal. We define the minimum and maximum distances across all alternatives as:

$$D_{\text{min}}(\varsigma_i, \varsigma^+) = \text{min}\left\{D_{NSP}(\varsigma_i, \varsigma^+) : \quad i = 1, 2, \ldots, n\right\},$$

$$D_{\text{max}}(\varsigma_i, \varsigma^+) = \text{max}\, D_{NSP}\left\{(\varsigma_i, \varsigma^+) : \quad i = 1, 2, \ldots, n\right\}.$$

The relative closeness coefficient, $RC(\varsigma_i)$, of each alternative with respect to PFPIS and PFNIS is then computed as per the classical TOPSIS approach:

$$RC(\varsigma_i) = \frac{D_{NSP}(\varsigma_i, \varsigma^-)}{D_{NSP}(\varsigma_i, \varsigma^+) + D_{NSP}(\varsigma_i, \varsigma^-)}. \tag{14}$$

A higher value of $RC(\varsigma_i)$ signifies a preferable alternative, as it indicates greater proximity to the PIS and distance from the NIS.

To further refine the ranking, Zhang and Xu [34] proposed a modified index $\xi(\varsigma_i)$, defined as:

$$\xi(\varsigma_i) = \frac{D_{NSP}(\varsigma_i, \varsigma^-)}{D_{\text{max}}(\varsigma_i, \varsigma^-)} - \frac{D_{NSP}(\varsigma_i, \varsigma^+)}{D_{\text{min}}(\varsigma_i, \varsigma^+)}. \tag{15}$$

Based on either $RC(\varsigma_i)$ or $\xi(\varsigma_i)$, the alternatives are ranked, with the optimal choice being the one that maximizes these values.

5.2 Algorithm for Proposed Method

The traditional TOPSIS method, as introduced by Hwang and Yoon [14], serves as a foundational and effective approach for addressing MCDM problems involving precise numerical data. Building upon this, Zhang and Xu [34] proposed an enhanced version of TOPSIS tailored to handle MCDM challenges involving the Pythagorean fuzzy data. The approach involves the following key steps:

Step 1. In addressing an MCDM problem involving PFNs, the first step is to construct the decision matrix $R = (\Gamma_j(\varsigma_i))_{m\times n}$. Here, $\Gamma_j(\varsigma_i)$ for $i = 1, 2, \ldots, m$ and $j = 1, 2, \ldots, n$, represents the evaluation of ς_i in relation to criterion Γ_j.

Step 2. A new scoring function is employed to ascertain determine the PFPIS denoted by (ς^+) and the PFNIS denoted by (ς^-).

Step 3. Utilizing Equations (12) and (13), compute the weighted spherical distances of each alternative ς_i relative to both the PFPIS (ς^+) and the PFNIS (ς^-).

Step 4. Apply Equations (14) and (15) to calculate the relative closeness $RC(\varsigma_i)$ and the modified closeness measure $\xi(\varsigma_i)$ for each alternative ς_i.

Step 5. Rank the options based on the descending order of the relative proximity $RC(\varsigma_i)$ and $\xi(\varsigma_i)$. The highest value indicates the most preferable alternative, with greater $RC(\varsigma_i)$ corresponding to a more favorable ς_i where $i = 1, 2, \ldots, m$.

6 Illustrative Example

For a practical application, suppose you are evaluating several loan applicants based on criteria such aspayment history (Γ_1), credit utilization (Γ_2), credit history length (Γ_3), credit mix (Γ_4), inquiries (Γ_5), income and employment stability (Γ_6), debt to income ratio (Γ_7). By applying the TOPSIS method with Pythagorean fuzzy sets, rank these applicants more effectively, accounting for the inherent uncertainty in their financial profiles and improving the accuracy of the credit risk assessment.

	Γ_1	Γ_2	Γ_3	Γ_4	Γ_5	Γ_6	Γ_7
ς_1	$\langle .8, .3\rangle$	$\langle .4, .7\rangle$	$\langle .7, .5\rangle$	$\langle .6, .4\rangle$	$\langle .3, .7\rangle$	$\langle .8, .4\rangle$	$\langle .5, .6\rangle$
ς_2	$\langle .7, .4\rangle$	$\langle .5, .6\rangle$	$\langle .9, .3\rangle$	$\langle .8, .4\rangle$	$\langle .5, .8\rangle$	$\langle .7, .6\rangle$	$\langle .6, .5\rangle$
ς_3	$\langle .9, .2\rangle$	$\langle .3, .6\rangle$	$\langle .8, .5\rangle$	$\langle .7, .4\rangle$	$\langle .5, .5\rangle$	$\langle .8, .5\rangle$	$\langle .4, .7\rangle$
ς_4	$\langle .8, .4\rangle$	$\langle .4, .8\rangle$	$\langle .6, .4\rangle$	$\langle .9, .3\rangle$	$\langle .3, .8\rangle$	$\langle .6, .5\rangle$	$\langle .6, .6\rangle$
ς_5	$\langle .7, .5\rangle$	$\langle .3, .9\rangle$	$\langle .7, .5\rangle$	$\langle .8, .4\rangle$	$\langle .4, .7\rangle$	$\langle .7, .5\rangle$	$\langle .4, .8\rangle$

where for ς_1 and criterion Γ_1 the membership degree is $.8$ and the non-membership degree is $.3$.

Considering that payment history, credit history length, credit mix, income and employment stability as benefit criteria $J_1 = \{\Gamma_1, \Gamma_3, \Gamma_4, \Gamma_6\}$ and credit utilization, inquiries, debt to income ratio are as the cost criteria $J_2 = \{\Gamma_2, \Gamma_5, \Gamma_7\}$.

Modified accuracy function $v(p)$ is used to calculate score type PFPIS (ς^+) and PFNIS (ς^-), we utilize formula (10) and (11). The values are

$$\varsigma^+ = \{\langle .5, .8\rangle, \langle .7, .2\rangle, \langle .6, .2\rangle, \langle .9, .2\rangle\}$$
$$\varsigma^- = \{\langle .7, .3\rangle, \langle .5, .8\rangle, \langle .5, .4\rangle, \langle .5, .3\rangle\}$$

Next, utilize equation (12) and D_{NSP} of each alternatives ς_i from PFPIS and PFNIS as

	$D_{NSP}(\varsigma_i, \varsigma^+)$	$D_{NSP}(\varsigma_i, \varsigma^-)$
ς_1	.0590	.0554
ς_2	.0593	.0475
ς_3	.0561	.0475
ς_4	.0712	.0359
ς_5	.0290	.0742

Equation (14) and (15) used to calculate $RC(\varsigma_i)$ and $\xi(\varsigma_i)$ listed bellow:

	$RC(\varsigma_i)$ (Rank)	$\xi(\varsigma_i)$(Rank)
ς_1	.4842(2)	−1.2878(3)
ς_2	.4447(4)	−1.4046(4)
ς_3	.4642(3)	−1.2794(2)
ς_4	.3352(5)	−1.9713(5)
ς_5	.7189(1)	0(1)

Based on $RC(\varsigma_i)$ rank of the alternatives are $\varsigma_5 \succ \varsigma_1 \succ \varsigma_3 \succ \varsigma_2 \succ \varsigma_4$ and ς_5 is the best alternative. With respect to $\xi(\varsigma_i)$ ranking of the alternatives are $\varsigma_5 \succ \varsigma_3 \succ \varsigma_1 \succ \varsigma_2 \succ \varsigma_4$. Here also, the best alternative is ς_5.

7 Conclusion

The proposed Pythagorean fuzzy set (PFS)-based model offers several significant advancements over traditional credit risk assessment methods. Primarily, it enhances predictive accuracy by providing a more flexible and sophisticated framework for handling uncertainty, accommodating the nuanced nature of financial judgments. Furthermore, it bolsters the robustness of the credit scoring process by integrating a wider spectrum of information, thereby more effectively capturing the complexity inherent in real-world financial Environments. The incorporation of Pythagorean fuzzy sets into credit risk evaluation not only advances theoretical understanding but also yields tangible benefits for practical financial decision-making. Future research should aim to further refining this model and exploring its applicability across diverse financial sectors and datasets. Additionally, exploring hybrid approaches—combining PFS with other machine learning techniques—may present avenues for achieving even greater improvements in the accuracy and reliability of credit risk predictions.

Declarations

Availability of Supporting Data

All data generated or analyzed during this study are included in this published paper.

Funding

The authors conducted this research without any funding, grants, or support.

Competing Interests

The authors declare that they have no competing interests relevant to the content of this paper.

Authors' Contributions

A.K. contributed to visualization, validation, supervision, investigation, and conceptualization, while N.K. was responsible for writing the original draft, methodology, and formal analysis.

References

[1] Adak, A.K., Kumar, D. (2023). "Spherical distance measurement method for solving MCDM problems under Pythagorean fuzzy environment", *Journal of Fuzzy Extension and Applications*, 4(1), doi:10.22105/jfea.2022.351677.1224.

[2] Adak, A.K., Kumar, G., Bhowmik, M. (2023). "Pythagorean fuzzy semi-prime ideals of ordered semi-groups", *International Journal of Computer Applications*, 185(5), doi:10.5120/ijca2023922661.

[3] Astuti, I.F., Faizah, L., Khairina, D.M., Cahyadi, D. (2021). "A fuzzy Mamdani approach on community business loan feasibility assessment", *In 3rd 2021 East Indonesia Conference on Computer and Information Technology, EIConCIT 2021 (pp. 438–442). Institute of Electrical and Electronics Engineers Inc.*, doi:10.1109/EIConCIT50028.2021.9431899.

[4] Atanassov, K.T. (1989). "More on intuitionistic fuzzy sets", *Fuzzy Sets and Systems*, 33(1), doi:10.1016/0165-0114(89)90215-7.

[5] Bazmara, A., Donighi, S.S. (2014). "Bank customer credit scoring by using fuzzy expert system", *International Journal of Intelligent Systems and Applications in Engineering*, 11, 29-35, doi:10.5815/ijisa.2014.11.04.

[6] Bazmara, A., Sarkar, B. (2019). "Pythagorean fuzzy TOPSIS for multicriteria group decision-making with unknow weight information through entropy measure", *International Journal of Intelligent Systems*, 34, 1108-1128, doi:10.1002/int.22088.

[7] Biryukov, A., Murzagalina, G., Kagarmanova, A., Kochetkova, S. (2022). "Concepts of improving fuzzy and neural network methods for simulating bankruptcy in risk management by the bank's loan portfolio", *In Lecture Notes in Networks and Systems (vol. 381 LNNS, 513–524). Springer Science and Business Media Deutschland GmbH.*, doi:10.1007/978-3-030-93677-8_45

[8] Carvalho, J.P., Tome, J.A.B. (2009). "Rule based fuzzy cognitive map in socio-economic system", *In Proceedings of the 2009 International Fuzzy Systems Association World Congress and 2009 European Society for Fuzzy Logic and Technology Conference*, Lisbon, Portugal, 1821-1826.

[9] Chen, Z.S., Zhou, J., Zhu, C.Y., Wang, Z.J., Xiong, S.H., Rodríguez, R.M., Skibniewski, M.J. (2023). "Prioritizing real estate enterprises based on credit risk assessment: An integrated multi-criteria group decision support framework", *Financial Innovation*, 9(1), doi:10.1186/s40854-023-00517-y.

[10] Chourmouziadis, K., Chatzoglou, P.D. (2016). "An intelligent short term stock trading fuzzy system for assisting investors in portfolio management", *Expert Systems with Applications*, 43, 298-311, doi:10.1016/j.eswa.2015.07.063.

[11] Garg, H. (2016). "A novel accuracy function under interval-valued Pythagorean fuzzy environment for solving multicriteria decision making problem", *Journal of Intelligent & Fuzzy Systems*, 31(1), 529-540, doi:10.3233/IFS-162165.

[12] He, X.X. , Li, Y.F., Qin, K.Y., Meng, D. (2019). "Distance measures on intuitionistic fuzzy based on intuitionistic fuzzy dissimilarity functions", *Soft Computing.* 24(1), 523-541, doi:10.1007/s00500-019-03932-5.

[13] Hesamian, G. (2020). "A fuzzy distance measure for fuzzy numbers", *Control and Optimization in Applied Mathematics*, 5(1), 29-39, doi:10.30473/coam.2021.44290.1105.

[14] Hwang, C.L., Yoon, K. (1981). "Multiple attribute decision methods and applications: A state of the art survey", *Springer Verlag, New York*, doi:10.1007/978-3-642-48318-9.

[15] Karimi, S.S. , Sohrabi, T., Bayat Tork, A. (2023). "Blockchain technology in optimizing logistics information security in business process technology transfer management", *Control and Optimization in Applied Mathematics*, 8(2), 63-84, doi:10.30473/coam.2023.66693.1224.

[16] Krasavtseva, A. (2022). "Logical-linguistic method for assessing the risk of specialized lending (on the example of project financing)", *Economics and the Mathematical Methods*, 58(4), 83, doi:10.31857/s042473880020295-3.

[17] Lai, K.K., Yu, L., Wang, S., Zhou, L. (2006). "Neural Network Metalearning for Credit Scoring", *In: Huang, DS., Li, K., Irwin, G.W. (eds) Intelligent Computing. ICIC 2006. Lecture Notes in Computer Science*, 4113. Springer, Berlin, Heidelberg, doi:10.1007/11816157_47.

[18] Li, D.Q., Zeng, W.Y. (2018). "Distance measure of Pythagorean fuzzy sets", *International Journal of Intelligent Systems*, 33, 348-361, doi:10.1002/int.21934.

[19] Malhotra, R., Malhotra, D.K. (2002). "Differentiating between good credits and bad credits using neuro-Fuzzy systems", *European Journal of Operational Research*, 136, 190-211, doi:10.1016/S0377-2217(01)00052-2.

[20] Makhazhanova, U., Kerimkhulle, S., Mukhanova, A., Bayegizova, A., Aitkozha, Z., Mukhiyadin, A., Azieva, G. (2022). "The evaluation of creditworthiness of trade and enterprises of service using the method based on fuzzy logic", *Applied Sciences (Switzerland)*, 12(22), doi:10.3390/app122211515.

[21] Nosratabadi, H.E., Nadali, A., Pourdarab, S. (2012). "Credit assessment of bank customers by a fuzzy expert system. based on rules extracted from association rules", *International Journal of Machine Learning and Computing*, 2, 662-666, doi:10.7763/IJMLC.2012.V2.210.

[22] Peng, X.D., Li, W.Q. (2019). "Algorithms for interval-valued Pythagorean fuzzy sets in emergency decision making based on multi-parametric similarity measures and WSBA", *IEEE Access*, 7, 7419-7441, doi:10.1109/ACCESS.2018.2890097.

[23] Peng, X., Yang, Y. (2015). "Some results for Pythagorean fuzzy sets", International Journal of Intelligent Systems, 30, 1133-1160, doi:10.1002/int.21738.

[24] Polishchuk, V., Kelemen, M., Povkhan, I., Kelemen, M., Liakh, I. (2021). "Fuzzy model for assessing the creditworthiness of Ukrainian coal industry enterprises", *Acta Montanistica Slovaca*, 26(3), 444-454, doi:10.46544/AMS.v26i3.05.

[25] Roy, P.K., Shaw, K. (2023). "An integrated fuzzy credit rating model using fuzzy-BWM and new fuzzy-TOPSIS-Sort-C", *Complex and Intelligent Systems*, 9(4), 3581-3600, doi:10.1007/s40747-022-00823-5.

[26] Sartova, R., Mussina, A., Uakhitova, A. (2023). "Fuzzy logic application for credit risk assessment". *In AIP conference proceedings (vol. 2948), American Institute of Physics Inc.*, doi:10.1063/5.0165250.

[27] Setiawan, T.H., Prihatini, L. (2023). "Tsukamoto fuzzy in optimizing the creditworthiness assessment process at savings and loan cooperatives", *Barekeng: Jurnal Ilmu Matematika Dan Terapan*, 17(2), 0775-0786, doi:10.30598/barekengvol17iss2pp0775-0786.

[28] Wang, H.D., He, S.F., Pan, X. H. (2018). "A new bi-directional projection model based on Pythagorean uncertain linguistic variable", *Information*, 9, 23-34, doi:10.3390/info9050104.

[29] Yager, R.R. (2016). "Properties and applications of Pythagorean fuzzy sets. In: Angelov, P., Sotirov, S. (eds) Imprecision and uncertainty in information representation and processing", *Studies in Fuzziness and Soft Computing*, 332, Springer, Cham, doi:10.1007/978-3-319-26302-1_9.

[30] Yang, H., Yang, S. (2023). "Empirical research on credit system management for Chinese vocational college students based on personal mobile terminals", *International Journal of Intelligent Systems and Applications in Engineering*, 12(6s), 386-400.

[31] Yu, L.P., Zeng, S.Z., Merigo, J.M., Zhang, C.H. (2019). "A new distance measure based on the weighted induced method and its application to Pythagorean fuzzy multiple attribute group decision making", *International Journal of Intelligent Systems*, 34, 1440-1454, doi:10.1002/int.22102.

[32] Zadeh, L.A. (1965). "Fuzzy sets", *Information and Control*, 8, 338-353.

[33] Zeng, W.Y., Li, D.Q., Yin, Q. (2018). "Distance and similarity measures of Pythagorean fuzzy sets and their applications to multiple criteria group decision making", *International Journal of Intelligent Systems*, 33, 2236-2254, doi:10.1002/int.22027.

[34] Zhang, X.L., Xu, Z.S. (2014). "Extension of TOPSIS to multiple criteria decision making with Pythagorean fuzzy sets", *International Journal of Intelligent Systems* 29, 1016-1078, doi:10.1002/int.21676.

[35] Zhang, J., Guo, L., Lyu, T. (2021). "An enhanced personal credit identification coin-day destruction model based on blockchain technology fuzzy sets for region of China pearl river delta", *Journal of Intelligent and Fuzzy Systems*, 41(3), 4519-4525, doi:10.3233/JIFS-189712.