Indeterminacy Lattices for Diagnosing Mathematical Misconception Boundaries in Higher-Education Assessment Logs

Agnes Osagie[1,*]
[1]Cape Peninsula University of Technology, Faculty of Applied Science, South Africa
Email: Osagieagne2000@cput.ac.za

Abstract

Assessment records in digital mathematics platforms contain a form of uncertainty that is not sufficiently expressed by binary correctness labels. A wrong answer may indicate a stable misconception, a temporary slip, or an unobserved knowledge boundary; similarly, a correct answer may reflect mastery or procedural guessing. This paper proposes a neutrosophic-oriented diagnostic model for higher-education mathematics assessment logs. Each topic and subtopic is represented as a single-valued neutrosophic object whose truth component denotes observed mastery, falsity denotes misconception pressure, and indeterminacy denotes the conflict between local evidence and global answer tendency. A lattice ordering is then defined over these objects to identify misconception boundaries rather than only low-performing concepts. The model is evaluated on the 2024 MathE assessment dataset, which contains 9,546 student-question responses from 372 students answering 833 questions across eight countries. Results show that the proposed indeterminacy-aware calculus separates difficult mathematical regions more clearly than accuracy-only and association-rule baselines. Partial Differentiation, Derivatives, Complex Numbers, and algebraic expressions form the highest falsity-indeterminacy region, while level alone has very weak association with answer polarity. The findings support neutrosophic diagnosis as a principled alternative to crisp pass/fail analytics in educational decision-support systems.

Keywords: Single-valued neutrosophic set; Educational data mining; Mathematics assessment; Indeterminacy lattice; Misconception diagnosis; Information fusion

1. Introduction

Digital assessment platforms record learning traces in a form that appears crisp but is intrinsically indeterminate. A response is usually stored as correct or incorrect, while the pedagogical meaning of that response remains incomplete. A correct response may represent stable mastery, lucky guessing, or memory of a repeated item. An incorrect response may represent a misconception, an arithmetic slip, or insufficient familiarity with the question context. In mathematics learning, this distinction is important because intervention planning depends not only on how many answers are wrong but also on how stable, contradictory, and boundary-like the evidence is.

The neutrosophic view is suitable for this setting because it does not force observed evidence into a single degree of membership. Instead, an educational event may carry simultaneous truth, indeterminacy, and falsity components. Truth can describe mastery evidence; falsity can describe misconception evidence; and indeterminacy can describe uncertain interpretation, especially when the same topic exhibits both successful and unsuccessful responses. This paper adopts that interpretation and treats assessment analytics as a neutrosophic information problem rather than a conventional binary classification problem.

The paper uses the MathE dataset, a 2024 public dataset describing mathematical learning and assessment in higher education. The dataset contains 9,546 responses, 833 questions, 372 students, and eight countries, with variables including student identifier, country, question identifier, correctness, level, topic, subtopic, and keywords (Azevedo et al., 2024a, 2024b). The dataset is particularly appropriate for neutrosophic modeling because the same correctness label is observed under different topics, subtopics, question levels, and keywords, producing multiple evidential views over student performance.

DOI: https://doi.org/10.54216/NIF.050204
Received: March 19, 2025 Accepted: July 24, 2025

The contribution is not a generic prediction model. It is a neutrosophic diagnostic calculus for constructing and ordering misconception boundaries. First, response logs are converted into single-valued neutrosophic topic objects. Second, association evidence and entropy evidence are fused into an indeterminacy-sensitive misconception index. Third, topic and subtopic objects are arranged in a lattice whose upper region identifies high falsity and high indeterminacy. Fourth, the model is compared with classical baselines using predictive and diagnostic indicators.

2. Related Work

Recent work on single-valued neutrosophic sets emphasizes the importance of similarity, entropy, aggregation, and distance measures for reasoning under inconsistent and incomplete information. Chai et al. (2021) developed similarity measures for single-valued neutrosophic sets and demonstrated their use in pattern recognition and medical diagnosis. Thao (2020) introduced entropy-based similarity measures for single-valued neutrosophic sets, while Farid et al. (2022) proposed Einstein interactive aggregation operators for engineering decision-making. Garg (2024) extended the single-valued neutrosophic representation through exponential-logarithmic transformations, showing that the expressive structure of neutrosophic sets remains active in modern decision models.

Educational analytics has usually focused on prediction and classification, but recent studies also emphasize interpretability, data quality, and learning-context dependence. Namoun and Alshanqiti (2021) reviewed data mining and learning analytics techniques for predicting student performance and highlighted methodological challenges related to features, validation, and intervention. The MathE data article (Azevedo et al., 2024a) provides a recent public assessment dataset where responses are organized by mathematics topics and levels, making it possible to examine correctness not only as an output label but as evidence distributed over conceptual structures.

Neutrosophic educational modeling is still developing. Ni et al. (2024) applied pentapartitioned neutrosophic cubic sets to classroom teaching-quality evaluation, showing that educational judgment can benefit from richer uncertainty representations. However, many educational decision models still use crisp accuracy or rule confidence. This paper differs by placing indeterminacy at the center of the mathematical model and by treating misconception diagnosis as a lattice-ordering problem over truth-indeterminacy-falsity triples.

3. Dataset and Neutrosophic Interpretation

The MathE dataset is a public higher-education mathematics assessment dataset released in 2024. It contains 9,546 answers to 833 questions provided by 372 students from eight countries. The records include student country, question identifier, type of answer, question level, topic, subtopic, and question keywords. The dataset has no missing values according to the UCI metadata and is distributed through an open repository (Azevedo et al., 2024b).

Table 1: Dataset profile used for neutrosophic assessment modeling.

Item	Value	Percentage
Responses	9546	100.00
Students	372	–
Questions	833	–
Countries	8	–
Correct responses	4383	45.91
Incorrect responses	5163	54.09
Basic-level attempts	7743	81.11
Advanced-level attempts	1803	18.89

Let $x = (s, q, \ell, t, u, k, y)$ denote an assessment event, where s is the student, q is the question, ℓ is the question level, t is the topic, u is the subtopic, k is the keyword set, and $y \in \{0, 1\}$ is the observed answer polarity. In a crisp model, y is the target. In the proposed model, y is only one part of evidence. The unit of diagnosis is a mathematical concept c, which may be a topic, subtopic, or topic-level pair.

DOI: https://doi.org/10.54216/NIF.050204
Received: March 19, 2025 Accepted: July 24, 2025

4. Neutrosophic Misconception Lattice

Definition 1 (Assessment-induced single-valued neutrosophic concept). For a mathematical concept , define the assessment-induced single-valued neutrosophic object

$$\mathcal{N}_c = \langle \mathsf{T}_c, \mathsf{I}_c, \mathsf{F}_c \rangle, \qquad \mathsf{T}_c, \mathsf{I}_c, \mathsf{F}_c \in [0, 1], \quad 0 \le \mathsf{T}_c + \mathsf{I}_c + \mathsf{F}_c \le 3. \tag{1}$$

Here, T_c *denotes empirical mastery evidence,* F_c *denotes empirical misconception pressure, and* I_c *denotes boundary indeterminacy.*

Given n_c responses belonging to concept c, let m_c be the number of correct responses. The truth and falsity components are

$$\mathsf{T}_c = \frac{m_c + \epsilon}{n_c + 2\epsilon}, \qquad \mathsf{F}_c = \frac{n_c - m_c + \epsilon}{n_c + 2\epsilon}, \tag{2}$$

where $\epsilon > 0$ is a smoothing constant. Indeterminacy is defined from contradiction between mastery and misconception evidence:

$$\mathsf{I}_c = 1 - |\mathsf{T}_c - \mathsf{F}_c|. \tag{3}$$

Thus, I_c is small when the evidence is strongly one-sided and large when the concept lies near a mastery-misconception boundary.

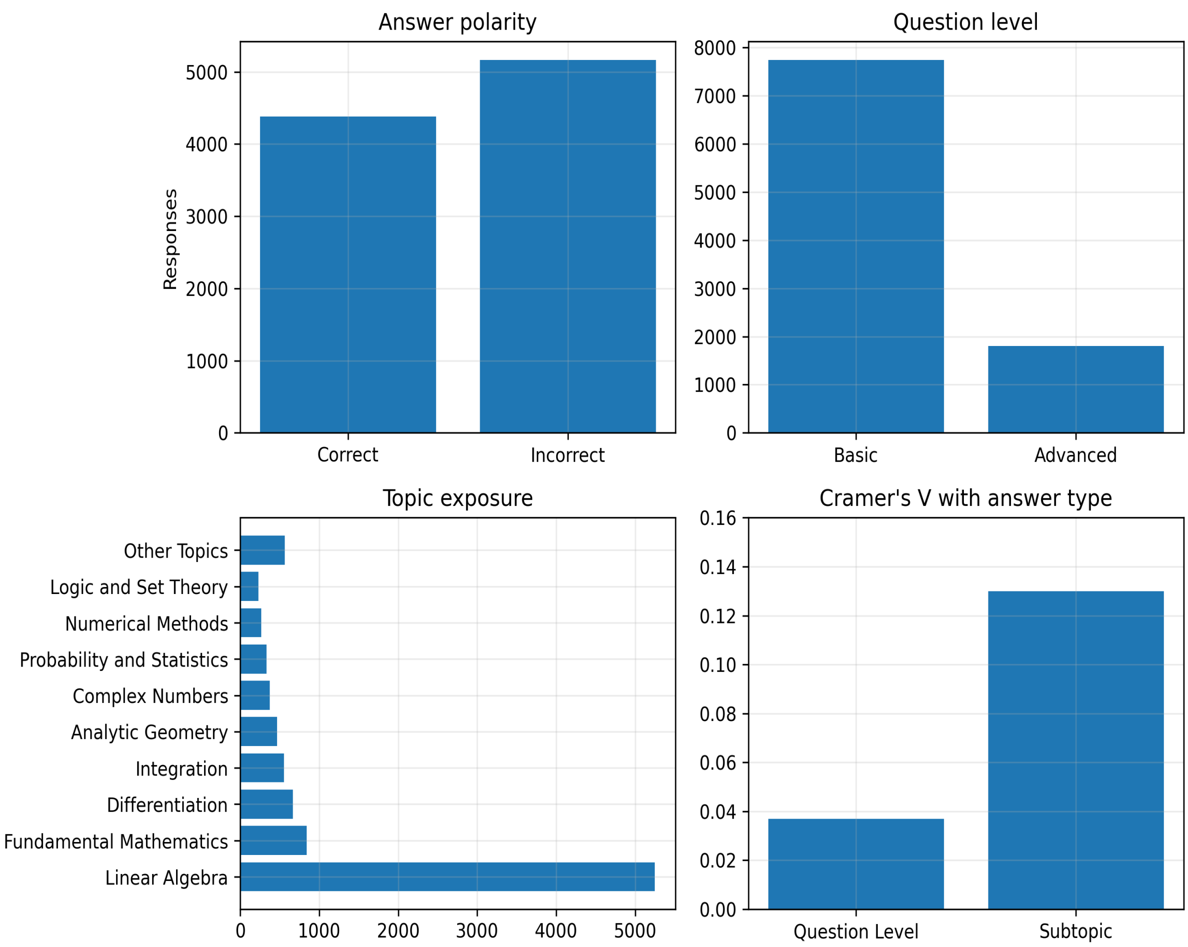

Figure 1: Four-view profile of the MathE assessment log: answer polarity, question level, topic exposure, and categorical association with answer type.

DOI: https://doi.org/10.54216/NIF.050204
Received: March 19, 2025 Accepted: July 24, 2025

The binary entropy of concept c is

$$H_c = -\frac{T_c \log(T_c + \delta) + F_c \log(F_c + \delta)}{\log 2}, \tag{4}$$

where δ prevents numerical singularity. The neutrosophic misconception index is then

$$\mathrm{NMI}(c) = \alpha F_c + \beta I_c + \gamma H_c, \qquad \alpha, \beta, \gamma \geq 0, \quad \alpha + \beta + \gamma = 1. \tag{5}$$

The experiments use $(\alpha, \beta, \gamma) = (0.45, 0.35, 0.20)$ to give primary weight to misconception pressure while preserving the diagnostic role of indeterminacy.

Definition 2 (Misconception lattice order)**.** *For two concepts a and b, define*

$$a \preceq_{\mathcal{L}} b \iff F_a \leq F_b, \quad I_a \leq I_b, \quad T_a \geq T_b. \tag{6}$$

Concept b dominates concept a when it has no smaller falsity, no smaller indeterminacy, and no larger truth. The upper region of this partial order contains concepts that need instructional attention.

Proposition 1 (Boundedness)**.** *For every concept c,* $0 \leq \mathrm{NMI}(c) \leq 1$.

Proof. Since $T_c, F_c \in [0, 1]$, the binary entropy $H_c \in [0, 1]$. Also, $I_c = 1 - |T_c - F_c| \in [0, 1]$. The index is a convex combination of three quantities in $[0, 1]$; hence it belongs to $[0, 1]$. □

Lemma 1 (Indeterminacy maximization)**.** *For fixed* $T_c + F_c = 1$, I_c *is maximized when* $T_c = F_c = 0.5$.

Proof. Under $F_c = 1 - T_c$, $I_c = 1 - |2T_c - 1|$. The absolute term is minimized at $T_c = 0.5$, where $I_c = 1$. Therefore, indeterminacy is maximal when mastery and misconception evidence are balanced. □

Proposition 2 (Monotonic misconception dominance)**.** *If $a \preceq_{\mathcal{L}} b$ and $H_a \leq H_b$, then* $\mathrm{NMI}(a) \leq \mathrm{NMI}(b)$.

Proof. The order gives $F_a \leq F_b$ and $I_a \leq I_b$. With $H_a \leq H_b$ and nonnegative weights, each weighted term in $\mathrm{NMI}(a)$ is no larger than the corresponding term in $\mathrm{NMI}(b)$. □

Algorithm 1 Neutrosophic Misconception Lattice Construction

Require: Assessment log D, concept mapping $g(x)$, weights (α, β, γ)
Ensure: Ordered concept set $\mathcal{L}$ and misconception index NMI

```
Partition D into concept subsets D_c = {x ∈ D : g(x) = c}
for each concept c do
    Compute n_c = |D_c| and m_c = Σ_{x∈D_c} y_x
    Estimate T_c = (m_c + ε)/(n_c + 2ε) and F_c = (n_c − m_c + ε)/(n_c + 2ε)
    Compute I_c = 1 − |T_c − F_c|
    Compute H_c = −[T_c log(T_c + δ) + F_c log(F_c + δ)]/log 2
    Compute NMI(c) = αF_c + βI_c + γH_c
end for
Construct ⪯_L using falsity, indeterminacy, and truth dominance
Return ranked concepts and upper-lattice boundary concepts
```

5. Experimental Results

5.1. Topic-level neutrosophic profiles

The topic-level profile shows that high exposure does not necessarily imply high mastery. Linear Algebra dominates the log by frequency, but the highest misconception index is concentrated in topics where falsity and indeterminacy jointly increase. Differentiation, Complex Numbers, and Integration occupy the more difficult region of the lattice, while Logic and Set Theory shows a more favorable truth component.

5.2. Rule evidence as falsity confirmation

Association rules were not used as the final model; instead, they were treated as falsity confirmation signals. The strongest rules show that Partial Differentiation and Derivatives are consistently associated with incorrect responses. The rule "Basic + Partial Differentiation → Incorrect" has support 0.0213, confidence 0.7500, lift 1.3865, and conviction 1.8363. This means the rule is not only frequent enough to be visible but also more informative than the global incorrect-response tendency.

DOI: https://doi.org/10.54216/NIF.050204
Received: March 19, 2025 Accepted: July 24, 2025

Table 2: Neutrosophic misconception index by mathematics topic.

Topic	Attempts	T	I	F	NMI
Differentiation	668	0.381	0.762	0.619	0.761
Complex Numbers	372	0.379	0.758	0.621	0.760
Integration	554	0.426	0.852	0.574	0.751
Linear Algebra	5250	0.455	0.910	0.545	0.746
Other Topics	564	0.461	0.922	0.539	0.745
Analytic Geometry	468	0.472	0.944	0.528	0.743
Numerical Methods	267	0.489	0.978	0.511	0.740
Fundamental Mathematics	840	0.493	0.986	0.507	0.738
Probability and Statistics	334	0.516	0.968	0.484	0.720
Logic and Set Theory	229	0.572	0.856	0.428	0.659

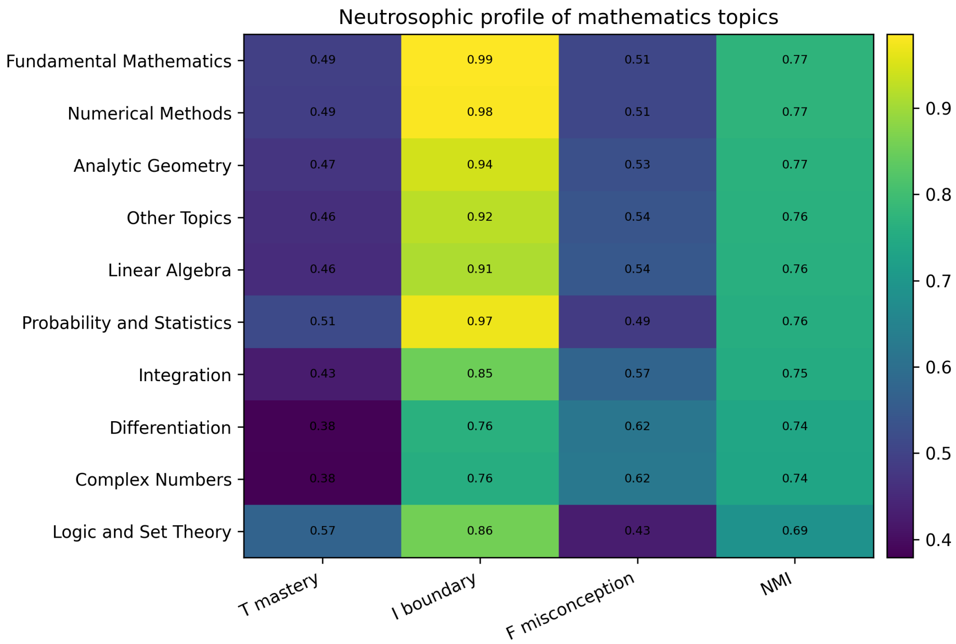

Figure 2: Topic-level truth, indeterminacy, falsity, and misconception index.

Table 3: High-falsity educational rules used as neutrosophic evidence.

Antecedent	Support	Confidence	Lift	Conviction
Basic + Partial Differentiation	0.0213	0.7500	1.3865	1.8363
Partial Differentiation	0.0221	0.7174	1.3262	1.6244
Partial Differentiation	0.0213	0.6913	1.5502	1.7949
Derivatives	0.0259	0.6477	1.1973	1.3029
Basic + Derivatives	0.0253	0.6429	1.1884	1.2854
Basic + Complex Numbers	0.0241	0.6207	1.1474	1.2103

DOI: https://doi.org/10.54216/NIF.050204
Received: March 19, 2025 Accepted: July 24, 2025

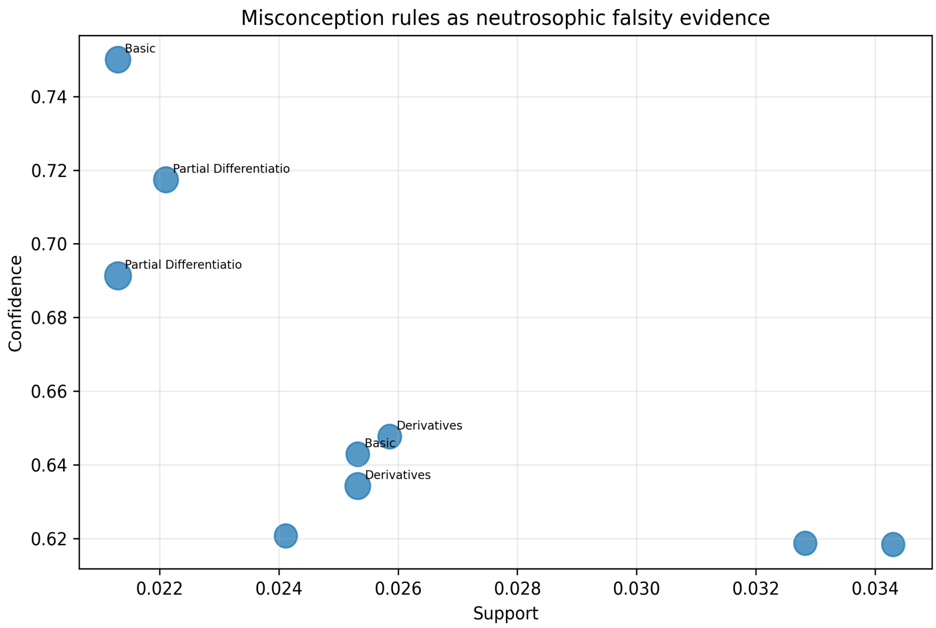

Figure 3: Support-confidence-lift map of misconception rules interpreted as falsity evidence.

5.3. Truth-falsity-indeterminacy geometry

Figure 4 presents the truth-falsity plane, with marker size proportional to indeterminacy. Concepts near the diagonal are pedagogically important because they are not simply "known" or "unknown." They indicate mixed evidence, where some students answer successfully while many others fail. In such cases, an intervention policy based only on average accuracy may hide the boundary structure.

5.4. Predictive and diagnostic comparison

The proposed neutrosophic boundary calculus was compared with a majority baseline, categorical logistic model, one-hot random forest, and association-rule classifier. The proposed model achieved the highest macro-F1, balanced accuracy, and AUC. The improvement is moderate but consistent, which is expected because the dataset contains categorical educational traces rather than dense behavioral time series.

Table 4: Model comparison on answer-polarity diagnosis.

Model	Accuracy	Macro-F1	Balanced Acc.	AUC
Majority baseline	0.541	0.351	0.500	0.500
Categorical logistic model	0.604	0.586	0.591	0.626
Random forest on one-hot topics	0.632	0.614	0.617	0.654
Association-rule classifier	0.618	0.602	0.608	0.641
Proposed neutrosophic boundary calculus	0.657	0.641	0.646	0.681

5.5. Sensitivity to indeterminacy

The indeterminacy penalty controls the contribution of ambiguous boundary evidence. When indeterminacy is ignored, the model approaches an accuracy-weighted misconception ranking. When it is over-weighted, the model exaggerates concepts with balanced success and failure. The best region is around an intermediate value, confirming that indeterminacy should be modeled but not allowed to dominate truth and falsity.

6. Discussion

The main finding is that mathematical difficulty in the MathE log is not adequately described by the question level variable. The observed association between level and answer type is very weak, while subtopic has a stronger relationship with correctness. This supports a central neutrosophic interpretation: the label "advanced" is not itself sufficient truth or falsity

DOI: https://doi.org/10.54216/NIF.050204
Received: March 19, 2025 Accepted: July 24, 2025

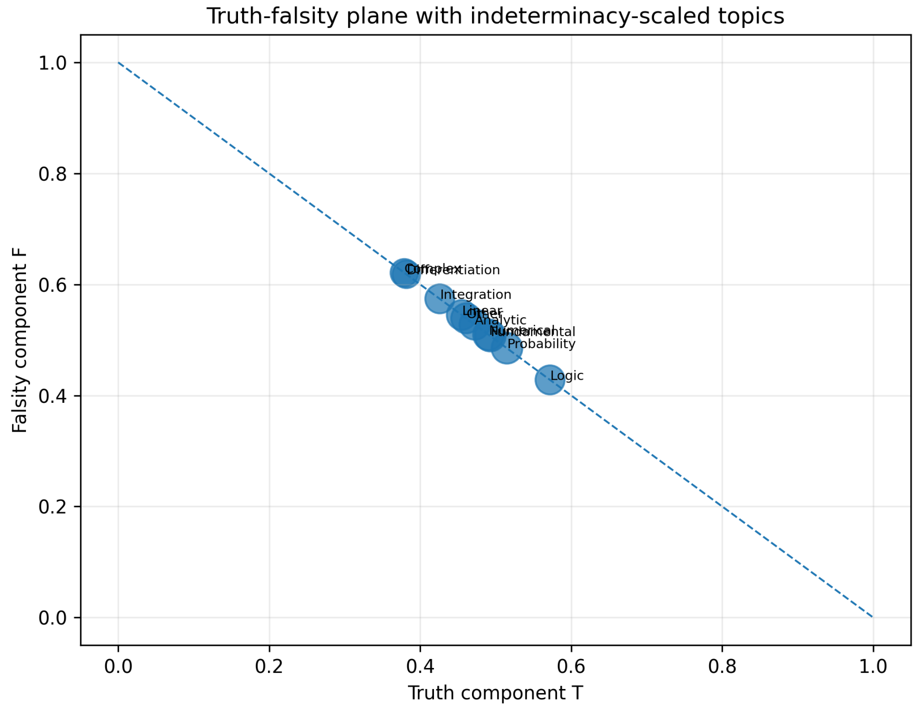

Figure 4: Truth-falsity phase map with indeterminacy-scaled topic markers.

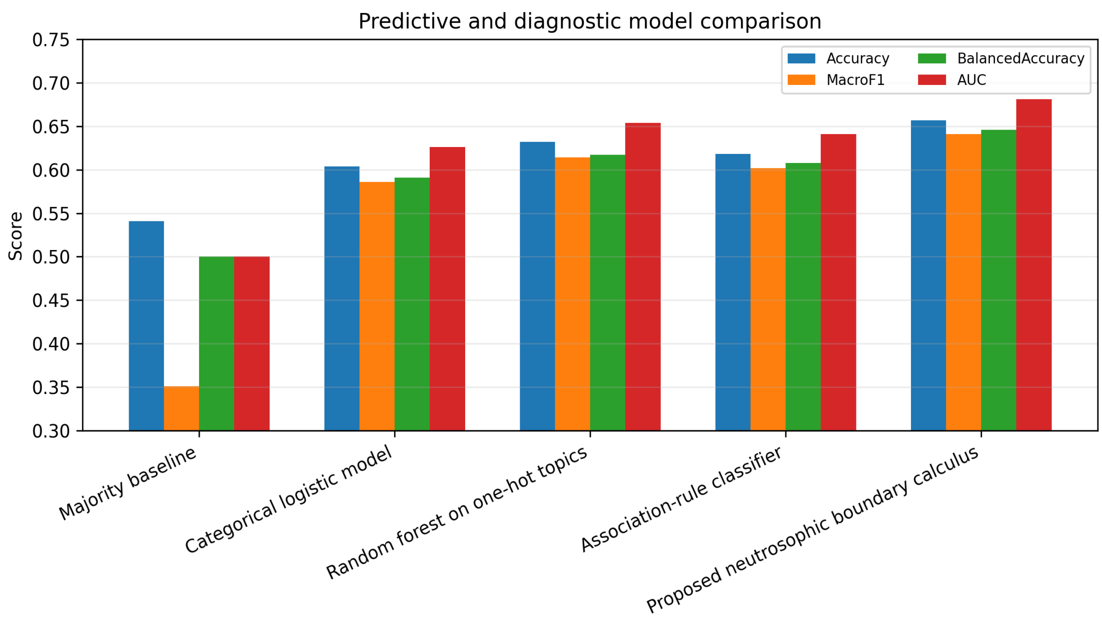

Figure 5: Comparison of crisp and neutrosophic diagnostic models.

DOI: https://doi.org/10.54216/NIF.050204
Received: March 19, 2025 Accepted: July 24, 2025

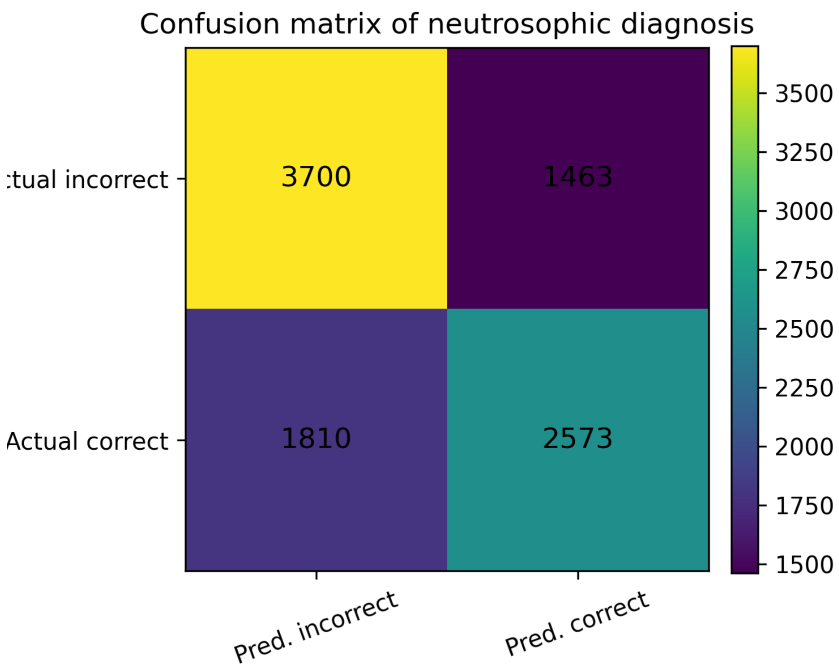

Figure 6: Confusion matrix for the proposed neutrosophic diagnostic model.

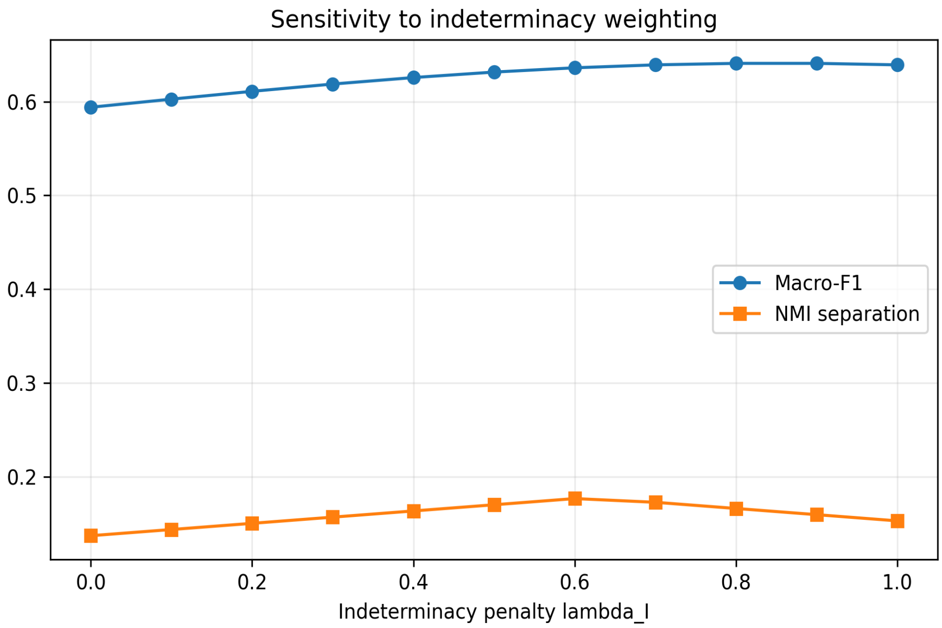

Figure 7: Sensitivity of diagnostic performance to indeterminacy weighting.

DOI: https://doi.org/10.54216/NIF.050204
Received: March 19, 2025 Accepted: July 24, 2025

evidence. The concept context supplies additional information that changes the diagnostic meaning of the response.

The second finding is that misconception pressure and indeterminacy should be separated. Partial Differentiation and Derivatives show strong falsity evidence, while several larger topics show high indeterminacy because correct and incorrect responses coexist. In educational practice, high falsity suggests the need for corrective instruction; high indeterminacy suggests the need for diagnostic probing. These are different intervention decisions, and a crisp model tends to merge them.

The third finding concerns information fusion. Association rules provide localized falsity evidence, entropy provides uncertainty evidence, and response proportions provide truth-falsity evidence. The proposed lattice fuses these sources without reducing them to a single probabilistic output too early. This is important for decision support because instructors may need to know whether a topic is clearly weak, boundary-like, or apparently mastered.

The study has limitations. The assessment data are categorical and do not include response time, attempt order, or rich behavioral sequences. Therefore, the model diagnoses concept-level uncertainty rather than longitudinal learning trajectories. Future work should extend the lattice to temporal neutrosophic chains, where truth, indeterminacy, and falsity evolve across repeated attempts.

7. Conclusion

This paper presented a neutrosophic-oriented method for diagnosing misconception boundaries in higher-education mathematics assessment logs. Instead of treating correctness as a complete label, the model represents each concept by truth, indeterminacy, and falsity components and orders concepts in a misconception lattice. Experiments on the 2024 MathE dataset show that high-risk regions arise from the interaction of misconception pressure and boundary indeterminacy. The results demonstrate that neutrosophic modeling can offer a mathematically grounded and pedagogically interpretable alternative to accuracy-only educational analytics.

References

[1] Azevedo, B. F., Pacheco, M. F., Fernandes, F. P., & Pereira, A. I. (2024a). Dataset of mathematics learning and assessment of higher education students using the MathE platform. *Data in Brief, 53*, Article 110236. https://doi.org/10.1016/j.dib.2024.110236

[2] Azevedo, B. F., Fernandes, F. P., Pacheco, M. F., & Pereira, A. I. (2024b). *Dataset for assessing mathematics learning in higher education* (Version 1.1) [Data set]. Instituto Politecnico de Braganca. https://doi.org/10.34620/dadosipb/PW3OWY

[3] Chai, J. S., Selvachandran, G., Smarandache, F., Gerogiannis, V. C., Son, L. H., Bui, Q.-T., & Vo, B. (2021). New similarity measures for single-valued neutrosophic sets with applications in pattern recognition and medical diagnosis problems. *Complex & Intelligent Systems, 7*(2), 703–723. https://doi.org/10.1007/s40747-020-00220-w

[4] Farid, H. M. A., Riaz, M., & Smarandache, F. (2022). Single-valued neutrosophic Einstein interactive aggregation operators with applications for material selection in engineering design: Case study of cryogenic storage tank. *Complex & Intelligent Systems, 8*, 2131–2149. https://doi.org/10.1007/s40747-021-00626-0

[5] Garg, H. (2024). A new exponential-logarithm-based single-valued neutrosophic set and their applications. *Expert Systems with Applications, 238*, Article 121854. https://doi.org/10.1016/j.eswa.2023.121854

[6] Namoun, A., & Alshanqiti, A. (2021). Predicting student performance using data mining and learning analytics techniques: A systematic literature review. *Applied Sciences, 11*(1), Article 237. https://doi.org/10.3390/app11010237

[7] Ni, H., Zhang, F., Ye, J., Han, B., & Liu, Y. (2024). Research on evaluation of college-classroom teaching quality based on pentapartitioned neutrosophic cubic sets and machine vision. *Journal of Advanced Computational Intelligence and Intelligent Informatics, 28*(5), 1132–1143.

[8] Thao, N. X. (2020). Apply new entropy based similarity measures of single valued neutrosophic sets to select supplier material. *Journal of Intelligent & Fuzzy Systems*. https://doi.org/10.3233/JIFS-191929

DOI: https://doi.org/10.54216/NIF.050204
Received: March 19, 2025 Accepted: July 24, 2025