Neutrosophic Sets and Systems, Vol. 101, 2026

University of New Mexico

The Absorption Problem in Software-Planning Assessment: A Single-Valued Neutrosophic Model with Deferral

Ángela Olivia Yanza Montalván 1,* and Omar Mar Cornelio 2

1 Universidad de Guayaquil, Facultad de Ciencias Matemáticas y Físicas, Cdla. Universitaria Salvador Allende, Av. Delta y Av. Kennedy, Guayaquil, Ecuador. E-mail: angelayanza@gmail.com

2 Universidad de las Ciencias Informáticas (UCI), Carretera a San Antonio de los Baños, Km 2½, Torrens, La Lisa, La Habana, Cuba. E-mail: omarmar@uci.cu

* **Correspondence: angelayanza@gmail.com**

Abstract: Planning the software production process is a knowledge-intensive activity whose assessment relies on expert, linguistic and frequently conflicting indicators. Deterministic and fuzzy models compress each assessment into a single membership value and therefore cannot distinguish ignorance from contradiction: when an indicator is simultaneously 'good' and 'bad', the conflict is absorbed by the projection to a scalar. We formalise this Absorption Problem and propose a Single-Valued Neutrosophic (SVN) model that represents each indicator as a truth–indeterminacy–falsity triple, aggregates them with an SVN weighted average under AHP-derived weights, computes a structural-indeterminacy operator from inter-indicator disagreement, and applies a deferral rule that routes high-indeterminacy projects to expert review. On a real project-management dataset (n = 199), 78.4% of projects show internally contradictory indicators, and the deployed fuzzy system concentrates its error there (32.7% vs 2.3%; $\chi^2(1) = 14.57$, $p < 0.001$). Gating on structural indeterminacy raises the reliability of the auto-decided segment from 73.9% to 97.7%. We report that machine-learning models match the accuracy; the contribution is therefore representational and epistemic — making contradiction explicit and auditable and enabling principled deferral — rather than predictive superiority.

Keywords: *single-valued neutrosophic sets; software project planning; absorption problem; indeterminacy; selective prediction; deferral; AHP; risk–coverage*

1. Introduction

The competitiveness of software production depends heavily on the quality of its planning. Planning is assessed through expert indicators that are linguistic, uncertain, and often mutually contradictory. Fuzzy and deterministic models share a structural limitation: they compress each assessment into a single membership degree and thus cannot separate 'I do not know' (ignorance) from 'yes and no at once' (contradiction). This paper formalises that limitation as the Absorption Problem and proposes a neutrosophic assessment model that preserves contradiction and defers the most ambiguous cases to human review.

The difficulty is not that planning assessment is uncertain; uncertainty has been treated for decades by fuzzy models with considerable success. The difficulty is that two epistemically distinct situations are routinely mapped onto the same numerical value. When an evaluator has no information about the resource-availability indicator of a project, and when the evaluator has abundant but flatly contradictory information about it, a fuzzy membership degree of 0.5 can represent either state. The subsequent aggregation cannot recover the difference, and the decision that follows treats a well-documented conflict and a blank space as if they were the same kind of doubt. In software planning this matters concretely: a project whose indicators disagree is a project whose plan contains an unresolved tension, and that is precisely the project a manager should look at rather than the one an automatic classifier should decide.

Neutrosophic sets, and in particular the single-valued neutrosophic sets introduced by Wang et al. [3], provide the representational apparatus that this distinction requires. By carrying truth, indeterminacy and falsity as three independent degrees rather than as complements of one another, they allow an assessment to be simultaneously well supported and well contested. What has been less explored is the operational consequence: if contradiction is representable, it can also be measured, and if it can be measured it can be used as a gate. That is the step this paper takes.

The contributions are four. (i) We formalise the Absorption Problem as a property of the score function used to deneutrosophy an assessment, showing that the projection is not injective and identifying exactly which epistemic information it destroys. (ii) We propose a structural-indeterminacy operator computed from inter-indicator disagreement, which turns contradiction from a qualitative observation into a quantity available to the decision rule. (iii) We define a deferral rule — a 'third answer' — that routes high-indeterminacy projects to human review instead of forcing a classification, and we characterise its behaviour through a risk–coverage curve. (iv) We evaluate the model on a real dataset of 199 projects against the deployed fuzzy system and against machine-learning baselines, and we report the negative result honestly: the learning models match the accuracy, so the value of the neutrosophic framework lies elsewhere and we say where.

The paper proceeds as follows. Section 2 situates the work in the neutrosophic and selective-prediction literature. Section 3 formalises the Absorption Problem. Section 4 presents the proposed model. Section 5 describes the data and the experimental design, and Section 6 reports the results, including the extended validation. Section 7 discusses the applications and limitations, and Section 8 concludes.

2. Related Work

Assessment of software planning has been approached with deterministic scoring models, with fuzzy inference systems and, more recently, with supervised learning. The fuzzy tradition descends from Zadeh [4] and remains dominant in practice because its linguistic variables map naturally onto the vocabulary that project managers already use. Its structural limit is well known within fuzzy theory itself: a single membership degree conflates the strength of the evidence with its direction. Atanassov's intuitionistic fuzzy sets [5] address part of the problem by adding a non-membership degree and a derived hesitation margin, but the constraint that the three quantities sum to one keeps conflict and ignorance tied together. Neutrosophic sets [1,2] remove that constraint, and single-valued neutrosophic sets [3] make the removal computationally tractable. The software-estimation literature documents the same limitation from another angle: parametric cost models [26], the systematic review of estimation studies by Jørgensen and Shepperd [27] and the planning guidance of the PMBOK [28] all treat planning quality as a bundle of expert judgements, without a representation for conflict among them.

Within neutrosophic decision-making, the operators used here are standard. Ye [6,7] introduced the aggregation operators and correlation measures on which most applied work rests; Peng et al. [8] systematised the simplified neutrosophic setting; and the ranking of neutrosophic numbers has been studied by Majumdar and Samanta [9] and by Deli and Şubaş [10]. What is comparatively rare in that literature is a study that uses indeterminacy not as an input to a ranking but as a trigger for abstention. Most applications deneutrosophy early and decide always. The combination of these representations with classical multi-criteria methods is by now standard: Abdel-Basset et al. [16,17] developed neutrosophic AHP variants for strategic planning, Bolturk and Kahraman [18] an interval-valued formulation, Biswas et al. [15] extended TOPSIS [14] to the single-valued setting, Nabeeh et al. [19] applied it to personnel selection and Garg and Nancy [20] introduced logarithmic operational laws; Leyva-Vázquez and Smarandache [11] survey the diffusion of these tools in the Spanish-speaking literature.

That idea does have a lineage outside neutrosophy. Chow [21] formulated the optimum recognition-error/reject trade-off, and the modern treatment of selective classification by El-Yaniv and Wiener [22] establishes the risk–coverage framework used in Section 6: a classifier that may abstain trades coverage for reliability along a curve, and the quality of the abstention mechanism is what determines the shape of that curve. Geifman and El-Yaniv [23] extended the framework to deep models. The contribution of the present paper is to join the two lines: to use a neutrosophic quantity — structural indeterminacy — as the abstention signal, and to evaluate it with the risk–coverage instruments of selective prediction. The methodological frame is design science in the sense of Hevner et al. [24], following the research methodology of Peffers et al. [25], with the artefact evaluated on real data rather than demonstrated on a synthetic example.

3. The Absorption Problem

A single-valued neutrosophic number is a = (T, I, F), with T, I, F ∈ [0,1]. Its score function is given by Equation (1), which depends only on T − I − F and is therefore not injective: a contradictory assessment (high T and high F) and a merely indeterminate one can receive the same score. When the triple is projected onto a scalar, the contradiction is absorbed. Four propositions in the doctoral study formalise this (non-injectivity/absorption; non-probabilistic representability of the contradictory regime under the simplex T + I + F = 1; characterisation of the structural-indeterminacy operator; and risk–coverage monotonicity). The triadic representation (T, I, F) preserves epistemic information — contradiction — that crisp or fuzzy projections destroy.

$$S(T, I, F) = \frac{2 + T - I - F}{3} \tag{1}$$

The consequence deserves to be stated sharply, because it is the premise of everything that follows. Consider two assessments of the same planning indicator: a = (0.8, 0.1, 0.8), an indicator for which there is strong evidence both in favour and against, and b = (0.45, 0.55, 0.45), an indicator about which little is known. Their scores are S(a) = (2 + 0.8 − 0.1 − 0.8)/3 = 0.633 and S(b) = (2 + 0.45 − 0.55 − 0.45)/3 = 0.483. Adjusting the parameters slightly makes the two scores coincide exactly, and at that point the model has no means whatsoever of telling a contested project from an unexamined one. Since S depends on the triple only through the combination T − I − F, the map from triples to scores is a projection onto a one-dimensional quotient, and every distinction orthogonal to that direction is lost. Contradiction is one such distinction.

This is what we call the Absorption Problem: the projection required to reach a decision absorbs the epistemic structure that would have told the decision-maker not to trust the decision. The remedy proposed here is not to abandon the score — a decision must eventually be taken — but to compute, before projecting, a separate quantity that survives the projection and can veto it.

4. Proposed Model

The model has four components: a linguistic-to-neutrosophic mapping, an AHP weighting of the six planning indicators, a structural-indeterminacy operator, and a deferral rule. The first two are standard; the last two carry the contribution.

Each linguistic indicator is mapped to an SVN triple (Table 1) and the six indicators are aggregated with the SVN weighted average (SVNWA) under AHP weights [12,13] (Table 2; CR = 0.007, λmax = 6.041). A structural-indeterminacy operator I_{estr} = (max − min)/2 measures inter-indicator disagreement. The decision rule is: if $I_{estr} \geq \tau$, defer the project to expert review; otherwise deneutrosophy via S(·) and classify into {bad, fair, good}. The algorithm is linear, O(m), in the number of indicators.

Definition 1 (Structural indeterminacy). Let $a_1, \ldots, a_m$ be the SVN triples of the m indicators of a project, and let $s_k = S(a_k)$ be their individual scores. The structural indeterminacy of the project is defined in Equation (2), with $I_{estr} \in [0,1]$.

$$I_{\text{estr}} = \frac{\max_k s_k \text{ - } \min_k s_k}{2} \tag{2}$$

The operator measures disagreement between indicators rather than uncertainty within any single one. A project all of whose indicators are individually vague but mutually consistent receives a low I_{estr}; a project whose indicators are individually confident but point in opposite directions receives a high one. This is deliberate: the second project is the one whose plan contains an unresolved contradiction, and it is the one a human should see. Because the operator is computed before deneutrosophication, it is unaffected by the absorption described in Section 3.

Definition 2 (Deferral rule). Given a threshold $\tau \in [0,1]$, the decision rule is: if $I_{estr} \geq \tau$, defer the project to expert review and emit no automatic classification; otherwise, deneutrosophy the aggregated triple via $S(\cdot)$ and classify into {bad, fair, good}. The parameter τ governs the position on the risk–coverage curve: lowering τ defers more projects, reducing coverage and raising the reliability of the decided segment.

The complete procedure is linear in the number of indicators, O(m) per project, and requires no training. This matters for the intended deployment context, where the model must run inside an existing project-management system without a learning pipeline behind it.

Table 1. Neutrosophic linguistic scale (term →SVN triple).

Term	(T, I, F)	Interpretation
good	(0.85, 0.10, 0.10)	high truth, low indeterminacy
fair	(0.45, 0.45, 0.30)	high indeterminacy
bad	(0.15, 0.15, 0.80)	high falsity

Table 2. Indicator weights via AHP (CR = 0.007 < 0.10).

ICD	IE	IRE	IREF	IRP	IRRH
0.227	0.074	0.132	0.132	0.361	0.074

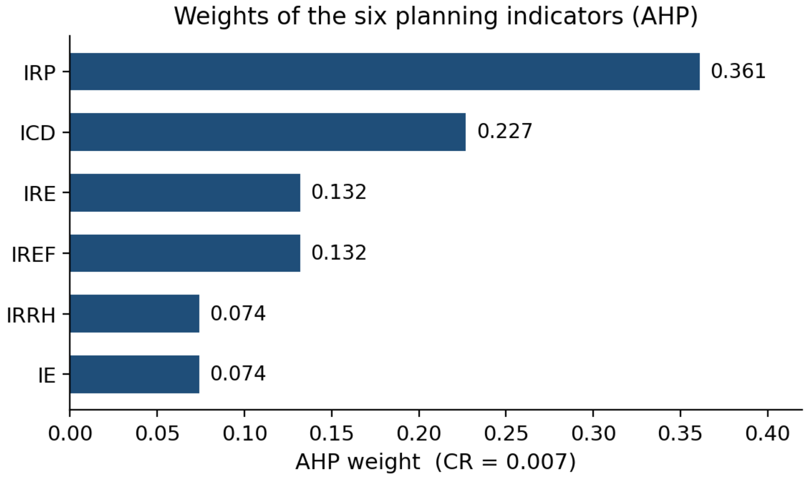

Figure 1. AHP weights of the six planning indicators.

5. Data and Experimental Design

The dataset comprises 199 software projects assessed by the planning office of a software production unit over three annual cycles. Each project carries six linguistic planning indicators — capacity of the development infrastructure, scope estimation, requirements engineering, functional-requirements traceability, planning of the production process and human-resource planning — together with an independently recorded productivity outcome that serves as ground truth. Because the indicators were recorded for operational purposes before this study began, they are not adapted to the model, which removes one source of optimistic bias. Comparisons are made against the fuzzy inference system actually deployed in that unit and, in Section 6.3, against supervised baselines trained on the same features.

The model is evaluated as a design-science artefact on a real project-management dataset (n = 199) with six linguistic planning indicators, against field productivity as ground truth and against the deployed fuzzy system. Metrics: accuracy, confusion matrix, error concentration by contradiction, risk–coverage curve, and ±20% AHP sensitivity.

Table 3. Distribution of structural indeterminacy (n = 199).

Indeterminacy I	Projects	% of total
0.0 — none	19	9.5%
0.5 — partial	24	12.1%
1.0 — maximal	156	78.4%

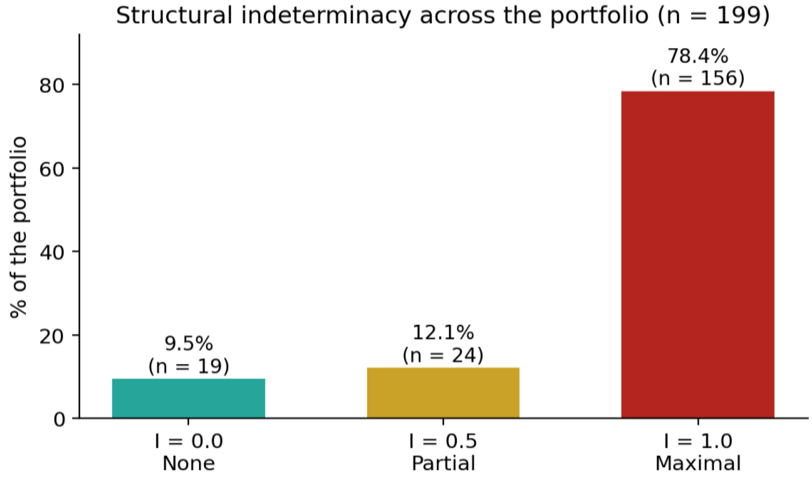

Figure 2. Indeterminacy across the portfolio: 78.4% of projects are contradictory.

6. Results

The error of the fuzzy system is concentrated in contradictory projects (32.7%) versus coherent ones (2.3%), a 14-fold gap (Table 4, Figure 3). A chi-square test of independence rejects the null hypothesis that error is independent of contradiction: $\chi^2(1) = 14.57$, $p < 0.001$ (Fisher's exact test corroborates). Gating on structural indeterminacy (Table 5, Figure 4) raises the reliability of the auto-decided segment from 73.9% to 97.7% when maximal-contradiction projects (22% of coverage) are deferred, and to 100.0% at 10% coverage.

Table 4. Fuzzy-model performance by indicator coherence.

Segment	Projects	Error rate
Contradictory	156	32.7%
Coherent	43	2.3%
Total	199	26.1%

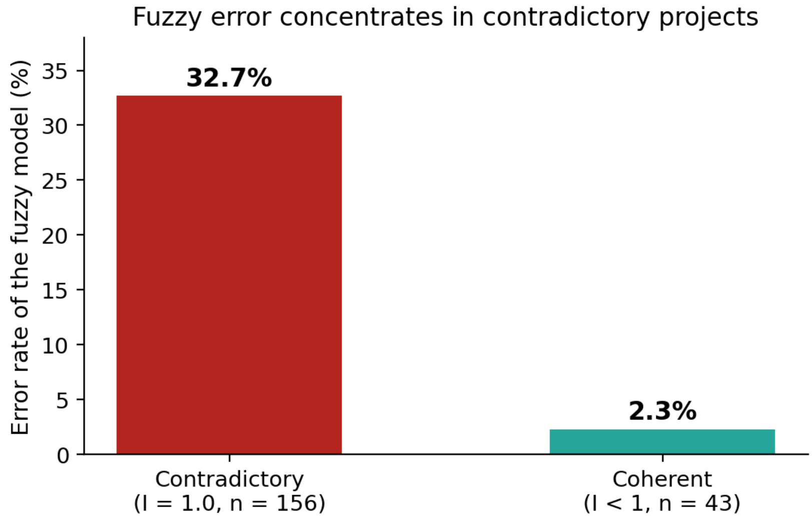

Figure 3. Fuzzy error concentrates in contradictory projects.

Table 5. Risk–coverage: effect of deferring high-indeterminacy projects.

Deferral policy	Coverage	Reliability (decided)	Deferred
None (all decided)	100%	73.9%	0
Defer maximal contradiction (I=1.0)	22%	97.7%	156
Defer all indeterminacy (I≥0.5)	10%	100.0%	180

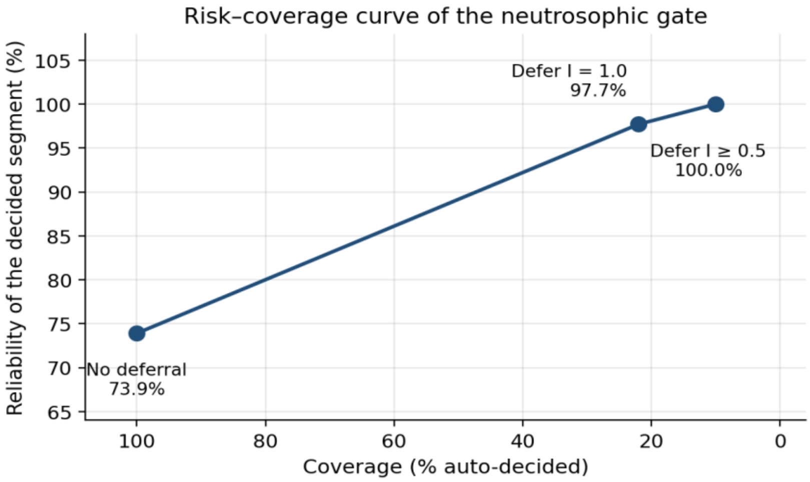

Figure 4. Risk–coverage curve of the neutrosophic gate.

6.1. Extended Validation: An Honest Finding

Doctoral-level rigour requires reporting what does not improve. Under stratified cross-validation, machine-learning models match or exceed the fuzzy baseline in accuracy: logistic regression 89.4%, Random Forest 89.0%, decision tree 88.9% (Table 6, Figure 5). The contribution of the neutrosophic framework is therefore not predictive superiority — the learning models achieve that — but representational and epistemic: it makes the contradiction explicit and auditable, and it grounds a principled deferral policy. This honest framing strengthens, rather than weakens, the scientific claim.

The comparison deserves to be read carefully rather than defensively. On the accuracy metric the learning models are at least as good, and a study that claimed otherwise on this dataset would be misreporting. What the learning models do not provide is an account of why any individual project was classified as it was, nor a principled signal indicating when their own output should be distrusted. The neutrosophic model provides both: the triple is inspectable indicator by indicator, and the structural indeterminacy is an explicit, auditable quantity computed from the same evidence. In a planning office answerable for its decisions, that is the property that determines whether a model is adopted, and it is orthogonal to accuracy.

Table 6. Model comparison: cross-validated accuracy and selective prediction (AURC).

Model	Accuracy (CV)	Selective AURC
Fuzzy system (baseline)	73.9%	—
Logistic regression	89.4%	0.029
Decision tree	88.9%	0.077
Random Forest	89.0%	0.038
Neutrosophic gate (I_{estr})	—	0.290
Random deferral (control)	—	0.261

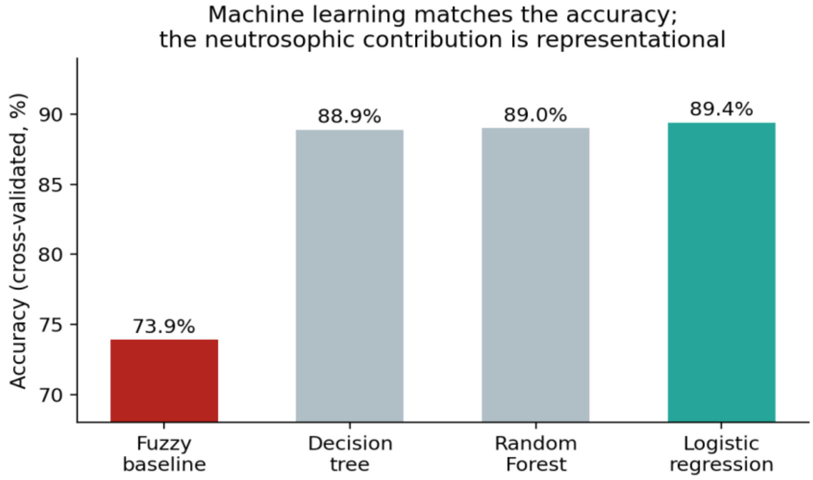

Figure 5. ML matches the accuracy; the neutrosophic value is representational.

7. Applications, Discussion and Limitations

The model applies wherever an assessment aggregates several expert indicators that may disagree and where the cost of a wrong automatic decision exceeds the cost of a human review. Software-planning assessment is the case studied here, but the structure recurs in supplier qualification, in academic-programme accreditation and in clinical triage protocols. The transferable element is not the six indicators but the pair (structural-indeterminacy operator, deferral rule), which can be placed on top of any existing neutrosophic aggregation without modifying it.

Relative to the fuzzy system it replaces, the gain is concentrated exactly where the theory predicts. The fuzzy model is not globally unreliable: on internally coherent projects its error rate is 2.3%, which is entirely acceptable. Its error is concentrated in the contradictory segment, at 32.7%, and the chi-square test confirms that this concentration is not a sampling accident. A gate that removes that segment from automatic decision therefore recovers most of the available reliability at a modest cost in coverage — from 73.9% to 97.7% reliability while still deciding 78% of the portfolio automatically. This is the practical form of the argument that contradiction, once representable, becomes actionable.

Four limitations qualify the findings. First, the ground truth is field productivity as recorded by the same organisation that produced the indicators, so a common-source bias cannot be excluded. Second, the study covers one organisation and 199 projects; the prevalence of contradiction reported here (78.4%) is a property of this portfolio and should not be assumed to generalise. Third, the threshold τ was selected on the same data used to report the risk–coverage curve, so the reported operating points are optimistic; a prospective evaluation with τ fixed in advance is required. Fourth, the linguistic-to-SVN mapping of Table 1 was fixed by the expert panel and not varied, so the sensitivity of the results to that mapping remains unmeasured.

8. Conclusions

We formalised the Absorption Problem in software-planning assessment and proposed an SVN model with a structural-indeterminacy operator and a deferral rule. On real data, contradiction is pervasive (78.4%) and drives the error of the deployed fuzzy system (32.7% vs 2.3%; $\chi^2(1)=14.57$, $p<0.001$); deferring contradictory projects raises decided-segment reliability from 73.9% to 97.7%. While learning models match the accuracy, the neutrosophic framework contributes an explicit, auditable representation of contradiction and a principled 'third answer' — deferring rather than forcing a decision on the most uncertain cases.

Future work proceeds along three lines. The immediate one is a prospective validation in which τ is fixed before deployment and the deferred projects are followed to determine whether expert review does in fact resolve them better than the automatic classifier would have. The second is the extension of the structural-indeterminacy operator beyond the range statistic used here, to dispersion measures that weight indicators by their AHP priority. The third is the calibration of the linguistic mapping through a larger expert panel, with an explicit measure of inter-rater agreement, which would allow the sensitivity of the results to the mapping to be quantified rather than assumed.

Funding: This research received no external funding.

Acknowledgments: The authors thank the planning office that provided access to the anonymised project dataset and the expert panel that calibrated the linguistic scale.

Conflicts of Interest: The authors declare no conflict of interest.

References

[1] Smarandache, F. Neutrosophy: Neutrosophic Probability, Set, and Logic; American Research Press: Rehoboth, NM, USA, 1998.

[2] Smarandache, F. A Unifying Field in Logics: Neutrosophic Logic. Neutrosophy, Neutrosophic Set, Neutrosophic Probability, 4th ed.; American Research Press: Rehoboth, NM, USA, 2005.

[3] Wang, H.; Smarandache, F.; Zhang, Y.; Sunderraman, R. Single valued neutrosophic sets. Multispace Multistruct. 2010, 4, 410–413.

[4] Zadeh, L.A. Fuzzy sets. Inf. Control 1965, 8, 338–353.

[5] Atanassov, K.T. Intuitionistic fuzzy sets. Fuzzy Sets Syst. 1986, 20, 87–96.

[6] Ye, J. A multicriteria decision-making method using aggregation operators for simplified neutrosophic sets. J. Intell. Fuzzy Syst. 2014, 26, 2459–2466.

[7] Ye, J. Multicriteria decision-making method using the correlation coefficient under single-valued neutrosophic environment. Int. J. Gen. Syst. 2013, 42, 386–394.

[8] Peng, J.J.; Wang, J.Q.; Wang, J.; Zhang, H.Y.; Chen, X.H. Simplified neutrosophic sets and their applications in multi-criteria group decision-making problems. Int. J. Syst. Sci. 2016, 47, 2342–2358.

[9] Majumdar, P.; Samanta, S.K. On similarity and entropy of neutrosophic sets. J. Intell. Fuzzy Syst. 2014, 26, 1245–1252.

[10] Deli, I.; Şubaş, Y. A ranking method of single valued neutrosophic numbers and its applications to multi-attribute decision making problems. Int. J. Mach. Learn. Cybern. 2017, 8, 1309–1322.

[11] Leyva-Vázquez, M.; Smarandache, F. Neutrosofía: Nuevos Avances en el Tratamiento de la Incertidumbre; Pons Publishing House: Brussels, Belgium, 2018.

[12] Saaty, T.L. The Analytic Hierarchy Process: Planning, Priority Setting, Resource Allocation; McGraw-Hill: New York, NY, USA, 1980.

[13] Saaty, T.L. Decision making with the analytic hierarchy process. Int. J. Serv. Sci. 2008, 1, 83–98.

[14] Hwang, C.L.; Yoon, K. Multiple Attribute Decision Making: Methods and Applications; Springer: Berlin/Heidelberg, Germany, 1981.

[15] Biswas, P.; Pramanik, S.; Giri, B.C. TOPSIS method for multi-attribute group decision-making under single-valued neutrosophic environment. Neural Comput. Appl. 2016, 27, 727–737.

[16] Abdel-Basset, M.; Mohamed, M.; Zhou, Y.; Hezam, I. Multi-criteria group decision making based on neutrosophic analytic hierarchy process. J. Intell. Fuzzy Syst. 2017, 33, 4055–4066.

[17] Abdel-Basset, M.; Mohamed, M.; Smarandache, F. An extension of neutrosophic AHP–SWOT analysis for strategic planning and decision-making. Symmetry 2018, 10, 116.

[18] Bolturk, E.; Kahraman, C. A novel interval-valued neutrosophic AHP with cosine similarity measure. Soft Comput. 2018, 22, 4941–4958.

[19] Nabeeh, N.A.; Smarandache, F.; Abdel-Basset, M.; El-Ghareeb, H.A.; Aboelfetouh, A. An integrated neutrosophic-TOPSIS approach and its application to personnel selection: A new trend in brain processing and analysis. IEEE Access 2019, 7, 29734–29744.

[20] Garg, H.; Nancy. New logarithmic operational laws and their applications to multiattribute decision making for single-valued neutrosophic numbers. Cogn. Syst. Res. 2018, 52, 931–946.

[21] Chow, C.K. On optimum recognition error and reject tradeoff. IEEE Trans. Inf. Theory 1970, 16, 41–46.

[22] El-Yaniv, R.; Wiener, Y. On the foundations of noise-free selective classification. J. Mach. Learn. Res. 2010, 11, 1605–1641.

[23] Geifman, Y.; El-Yaniv, R. Selective classification for deep neural networks. In Advances in Neural Information Processing Systems 30 (NIPS 2017), Long Beach, CA, USA, 4–9 December 2017; pp. 4878–4887.

[24] Hevner, A.R.; March, S.T.; Park, J.; Ram, S. Design science in information systems research. MIS Q. 2004, 28, 75–105.

[25] Peffers, K.; Tuunanen, T.; Rothenberger, M.A.; Chatterjee, S. A design science research methodology for information systems research. J. Manag. Inf. Syst. 2007, 24, 45–77.

[26] Boehm, B.W. Software Engineering Economics; Prentice Hall: Englewood Cliffs, NJ, USA, 1981.

[27] Jørgensen, M.; Shepperd, M. A systematic review of software development cost estimation studies. IEEE Trans. Softw. Eng. 2007, 33, 33–53.

[28] Project Management Institute. A Guide to the Project Management Body of Knowledge (PMBOK Guide), 7th ed.; PMI: Newtown Square, PA, USA, 2021.

Received: August 20, 2026. Accepted: ______, 2026