Research Paper

Evaluating model performance and prediction accuracy of fuzzy logic and bayesian networks with business expert inputs

Jane B. Gelindon[1*], Betchie E. Aguinaldo[2]
[1*,2]College of Computing Studies Information and Communication Technology, Isabela State University, Philippines.

Article Info

Article History:
Received: 17 January 2026
Revised: 27 March 2026
Accepted: 05 April 2026
Published: 22 May 2026

Keywords:
Fuzzy Logic
Bayesian Networks
Expert Systems
Model Evaluation
Decision Support

ABSTRACT

This paper evaluates and compares two expert-in-the-loop decision-support paradigms Fuzzy Logic (FL) and Bayesian Networks (BN) for barangay-level business risk assessment using the BizLocator Analytics dataset of Cauayan City, Philippines. Local governments need transparent, data-driven tools that can operate under uncertainty, sparse data, and evolving economic conditions. FL and BN are both well-established approaches for modeling uncertainty, yet they are rarely examined side by side on the same dataset with the same expert knowledge. To address this gap, the study develops parallel FL and BN models grounded in identical features and informed by the same pool of business and experts. The FL model uses expert-defined triangular and trapezoidal membership functions, together with a compact set of IF–THEN rules that encode linguistic concepts such as “Low Compliance,” “Vulnerable Barangay,” and “High Risk.” The BN model encodes expert-elicited causal relationships as a directed acyclic graph and learns conditional probability tables from data under Dirichlet priors. A unified preprocessing pipeline is applied, and nested stratified cross-validation is used to avoid optimistic bias and to support paired statistical tests. Both models are evaluated on discrimination (Accuracy, F1-score, ROC–AUC, PR–AUC) and probabilistic quality (Brier score, Expected Calibration Error, reliability diagrams). Results show that BN achieves slightly higher discrimination and notably better calibration, while FL offers superior case-level interpretability through rule and membership visualizations. Expert validation confirms that most BN edges are causally plausible and FL rules covers the majority of decisions. The findings suggest that a hybrid deployment using BN as the calibrated scoring backbone and FL as an explanation and policy-communication layer can provide accurate, transparent, and actionable decision support for local business risk governance and long-term planning. Overall, the study demonstrates how expert-guided artificial intelligence can strengthen evidence-based regulation while preserving human oversight and accountability in practice across diverse barangay.

Corresponding Author:
Jane B. Gelindon
College of Computing Studies Information and Communication Technology, Isabela State University, Philippines.
Email: gelindon.jane@isu.edu.ph

1. INTRODUCTION

Decisions in modern business are often uncertain. Traditional statistical or crisp-logic models assume precise inputs and stable distributions, but volatile markets, shifting customer behavior, supply-chain disruptions, and sparse data undermine them. Supplier reliability, service quality, customer attractiveness, and managerial confidence are linguistic, noisy, or partially observable strategic and operational variables. This has led organizations to develop intelligent methods that combine expert judgment with observational data and reason transparently under uncertainty. Fuzzy Logic (FL) and Bayesian Networks (BN) are mature, complementary paradigms for this.

FL, based on Zadeh's graded truth values, represents "high risk" and "strong demand" with degrees of membership rather than thresholds [1]. It works well for qualitative judgments, contested cutoffs, and imprecise but meaningful indicators to experts. Expert rules and membership functions can encode managerial reasoning and improve decisions when numerical signals are weak or inconclusive in banking, operations, and supply chains [2], [3], [4], [5]. Practitioners can trace FL's logic using its interpretable IF–THEN rule bases.

BNs probabilistically address uncertainty. They encode variables as nodes and causal or dependency relations as directed edges, updating beliefs with new evidence. This explicit treatment of conditional probabilities helps diagnose, predict, and intervene when data are incomplete and cause–effect structure matters. BNs have been shown to identify critical causal factors in production decisions, support dynamic situational awareness, and perform well with interpretable probabilistic reasoning [6], [7], [8]. BNs predict outcomes and explain how upstream changes propagate through a system and with what probability.

FL operationalizes linguistic judgments and semantic vagueness, while BN quantifies stochastic dependence and belief revision [1], [6], [7]. Analysts prefer FL for expert cognition and verbal scales and BN for probabilistic data and causal structure. Even though FL and BN are widely used, few studies compare them using the same dataset, feature set, expert inputs, and evaluation pipeline. Differences in data, preprocessing, and metrics make cross-study comparisons difficult for most paradigm studies [3], [4], [6].

That gap is filled by this controlled comparative study. Parallel FL and BN models use the same features, data, and domain experts. Standardized preprocessing, train–test splits, and performance metrics. We emphasize probability calibration (Brier score, reliability diagrams), robustness to missing and noisy data, and human-centered interpretability via structured evaluations with analysts and managers beyond headline accuracy (Accuracy, F1, AUC, MAE/RMSE). FL excels when inputs are mostly linguistic or expert heuristics capture domain regularities better than sparse measurements, and BN excels when causal dependencies, explicit uncertainty quantification, and frequent belief updates are critical [3], [4], [5], [6], [7]. Both should be sensitive to expert priors in small-data regimes [4], [6].

Thus, when expert judgment meets uncertain business data, which paradigm Fuzzy Logic or Bayesian Networks provides more reliable, interpretable, and actionable guidance? In line with emerging work on trustworthy, transparent, and human-centered AI for complex organizational settings [1], [8], we aim to clarify trade-offs, highlight context-specific strengths, and outline hybridization pathways that combine FL's linguistic reasoning with BN's probabilistic inference.

2. RELATED WORK

The utilization of fuzzy logic (FL) and Bayesian networks (BNs) in business decision-support contexts has been the subject of a growing body of literature over the past five years. This literature frequently incorporates expert knowledge to mitigate uncertainty. Fuzzy logic methodologies are frequently implemented to convert subjective evaluations into quantitative metrics, utilizing membership functions and rule bases that are collaboratively developed with domain experts. [9] For instance, created a FL system that utilized verbal expressions and expert preferences to facilitate business innovation. They demonstrated that the quality of decisions was enhanced when qualitative insights were systematically encoded. [10] Developed a type-2 hexagonal fuzzy logic model for safety stock management that incorporates expert intuition to more effectively manage demand uncertainty than conventional models. Similarly, [11] employed fuzzy logic combined with Analytic Hierarchy Process (AHP) weighting to choose additive manufacturing materials, incorporating expert input to enhance membership functions. Meanwhile, [12] integrated fuzzy logic and machine learning to optimize KPI identification in digital transformation projects.

2.1. Fuzzy Logic Model Computations

Membership Functions (MFs) Most studies [9], [10], [11] used triangular or trapezoidal MFs to convert crisp input data into fuzzy degrees. Example for variable “Demand”

$$\mu \text{ High } (x) = \begin{cases} 0 \\ \frac{x-50}{100-50} \\ 1 \end{cases}$$

$$x \leq 50$$
$$50 < x \leq 100$$
$$x > 100$$

Returning a degree between 0 and 1, this indicates “High Demand.”

- Rule Base / Inference Experts co-created IF–THEN rules, such as:
  - IF Demand is High AND Supply is Low THEN Risk is High
  - Inference engine typically uses Mamdani or Sugeno method.

Aggregation Defuzzification The fuzzy outputs are combined using max–min composition and then defuzzified (often via centroid of area):

$$y^* = \frac{\int y\, \mu\, (y) dy}{\int \mu\, (y) dy}$$

- Performance Metrics:
  - Papers used RMSE, MAE, or classification Accuracy to compare fuzzy outputs against observed values.
  - Some [12] also reported F1 score for classification tasks.

2.2. Bayesian Network Model Computations

At the same time, Bayesian networks have gained popularity as a probabilistic framework for modeling causal relationships in managerial decision-making. [11] Showed that expert-defined network structures improve predictive performance with sparse data by using a Bayesian network method to make production decisions when there is uncertainty. [12] Created a decision support system for managing emergency risks that was based on BN. The system was made to get domain experts to give it causal relationships so that it could fill in the network structure and conditional probability tables.

2.3. Structure Learning and CPT Parameterization

The network structure was partly based on expert causal maps found in [13], [14].

- Maximum Likelihood Estimation (MLE) or Bayesian parameter estimation with Dirichlet priors were employed to estimate conditional probabilities from data:

$$P\left(X\ Pa\ (X)\right) = \frac{-\alpha + N\ x, pa}{\sum_X \alpha + Nx, pa}$$

Where Nx,paN_{x,pa}Nx,pa is the count of instances of node XXX with parent configuration papapa.

- **Inference**
  - Standard BN inference algorithms (variable elimination, belief propagation) were employed to calculate predictions.
  - Dynamic BNs (DBNs) used forward-backward algorithms to change probabilities over time.
- **Classification Accuracy**

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN}$$

Area under the ROC Curve (AUC) and Precision–Recall AUC for probabilistic discrimination.

- **Brier Score for calibration**

$$\text{Brier} = \frac{1}{N}\sum_{i=1}^{N}(p_i - y_i)^2$$

- **Expected Calibration Error (ECE)**

$$ECE = \sum_{m=1}^{n} \frac{\lfloor B_m \rfloor}{n} |acc(B_m) - conf(B_m)]$$

Where BmB_mBm are probability bins.

2.4. Hybrid / Comparative Computations

Some hybrid studies [15] explicitly computed:

- Weighting coefficients from AHP or DEMATEL to initialize BN priors or fuzzy rule weights:

$$\omega_{j=} \frac{Eignevector_j}{\sum Eignevector}$$

- Comparison indices such as:

$$\triangle\ Accuracy = Acc_{BN} - Acc_{FL}$$

Or

$$\triangle\ Brier = Brier_{FL} - \ Brier_{BN}$$

To show which model performs better.

Sensitivity Analysis: Varying expert-provided parameters (membership cutoffs, CPT priors) and recomputing metrics to test robustness.

[14] Developed a BN system for judicial decision-making that integrated structured data with expert judgments to provide predictions that were updated in response to new evidence. These studies highlight the efficacy of Bayesian Networks (BNs) in handling incomplete or noisy data and producing calibrated probabilistic outputs.

Additionally, there is a small but expanding body of literature on hybrid methodologies that combine Bayesian and fuzzy methodologies. [15] Suggested the integration of Bayesian networks with multi-criteria decision-making methods (such as AHP and DEMATEL) to address incomplete expert knowledge. This approach demonstrates the potential for the two paradigms to reinforce one another. Compared to single-paradigm baselines, reported improved classification accuracy and F1 scores by integrating Bayesian networks and fuzzy reasoning in a group decision-making model. Although few studies directly compare their predictive performance using the same dataset and evaluation metrics, such hybrids suggest the possibility of capturing both the probabilistic rigor of Bayesian networks and the linguistic flexibility of fuzzy logic.

Although these developments have been made, the majority of research still prioritizes either FL's interpretability and rule construction or BN's calibration and inference, without integrating them into a unified evaluation protocol. Metrics such as ROC-AUC, Brier score, and Expected Calibration Error are only benchmarked in a small number of studies for both paradigms. This creates a gap in the literature for research that systematically compares FL and BN models side-by-side, utilizing the same expert-elicited data to evaluate predictive accuracy, calibration, and interpretability. This is the exact goal of the study right now.

3. METHODOLOGY

The researcher conducted a paired within-dataset test of two expert-in-the-loop models that help businesses make decisions:

- A Fuzzy Logic (FL) model whose membership functions and rule base are developed in collaboration with business experts and
- A Bayesian Network (BN) trained on the same structured dataset, potentially limited by expert priors/structure [16].

The main endpoints are how accurate the predictions are and how well they fit the data. Secondary endpoints encompass interpretability, resilience to data fluctuations, and expert involvement.

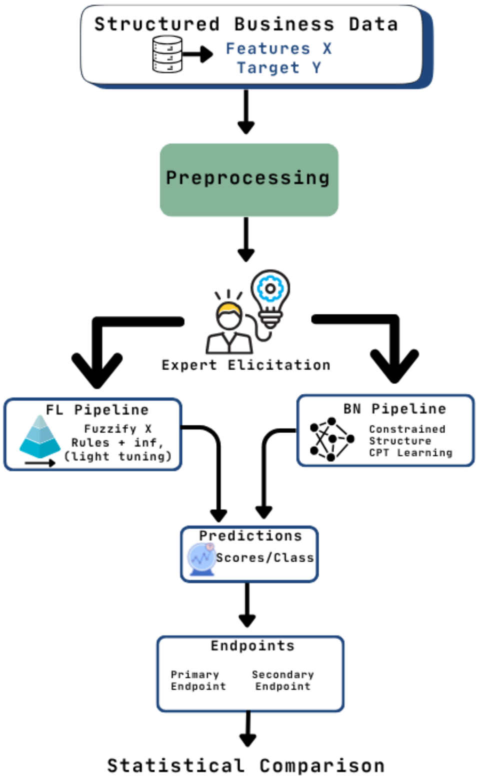

Figure 1. Shows the Flowchart of the Models and Experimental Methods Applied

Figure 1 A conceptual framework for the suggested FL–BN decision-support architecture, showing preprocessing, expert elicitation, dual-model prediction, endpoint evaluation, and statistical comparison, taken from [17] but it was modified.

3.1. Data

3.1.1. Sources and Features

Structured tabular dataset from the target business process (e.g., risk scoring, demand triage), with N observations and p predictors (numeric + categorical).

Outcome Y: a binary or multinomial target (like risk class); if it's continuous, set thresholds for classification and keep the continuous version for sensitivity analyses.

3.1.2. Preprocessing

For missing data, use the median for numbers and the most common for categories. Keep track of missingness indicators. For BN, the native missing-at-random handling is maintained throughout inference. Outliers: Winsorize at the 1st and 99th percentiles (in sensitivity analysis, report both the raw and Winsorized results).

Categoricals: FL feature scaling with target-agnostic encoding (one-hot); BN can keep categorical states as they are.

If there is an imbalance between classes, use stratified folds and report both the original and class-balanced metrics. Don't re-sample in the main analysis to keep the base rates realistic.

3.2. Involvement of Experts

3.2.1. Decision-making

Five to nine experts who meet certain requirements, like having at least five years of experience in the field, being responsible for making decisions, knowing data and processes, being available, and having a variety of tasks.

3.2.2. Elicitation Protocol

- Round 0 (scoping): figure out what the target constructs and candidate predictors are, and agree on the decision concepts (like "high risk" and "urgent").
- Round 1 (IDEA/Delphi): define linguistic terms, breakpoints, and causal links on your own; gather reasons.
- Feedback and Round 2: anonymous summaries lead to changes that bring the two sides closer together.
- For edges, use the median of breakpoints and the majority rule.
- Documentation: keep track of membership sketches, rule explanations, suggested BN edges, and ranges of uncertainty.

3.3. Fuzzy Logic Model

3.3.1. Variable Fuzzification

- For each input xjx_jxj, define KjK_jKj linguistic terms with triangular or trapezoidal membership functions μj,k(x)\mu_{j,k}(x)μj,k(x).
- Initialize breakpoints from empirical quantiles (e.g., 20–40–60–80%) and refine to expert-specified anchors (with constraints to avoid overlaps/inversions).

3.3.2. Rule Base Construction

- Start from expert "IF–THEN" rules (e.g., IF Demand is High AND Margin is Low THEN Risk is high).
- Use coverage analysis to get rid of rules that are the same or conflict with each other, and limit the number of rules (e.g., ≤50) to make them easier to understand.
- T-norm/T-conorm: default min/max; sensitivity check with product/sum.
- Inference/Defuzzification: Mamdani with centroid for continuous outputs; for classification, use max membership or thresholder centroid.

3.3.3. Data-Informed Tuning (No Label Leakage)

- Within training folds only, perform a lightweight search over:
  - Membership breakpoints ±10–20% around expert anchors,
  - Rule weights ∈[0.5,1.5]\in[0.5,1.5]∈[0.5,1.5],
  - Choice of T-norm.
  - Select via nested cross-validation (Section 6). Preserve the expert-only configuration for ablations.

3.4. Bayesian Network Model

3.4.1. Structure Learning

- Expert-seeded DAG: encode mandatory edges (from elicitation) and forbidden edges (causally implausible).
- Learn remaining structure with MMHC or PC algorithm under these constraints (score: BDeu or BIC).
- Sensitivity: compare expert-only, data-only, and hybrid structures.

3.4.2. Parameter Learning (CPTs)

- Dirichlet priors with equivalent sample size (ESS) α∈{1,5,10}\alpha \in \{1,5,10\}α∈{1,5,10}; center the prior on expert probabilities when provided.
- Estimate CPTs via Bayesian updating:

$$P(X = x \mid Pa(X) = \pi) = \frac{\alpha_x + N_{x,\pi}}{\sum_{x'} (\alpha_{x'} + N_{x',\pi})}$$

3.4.3. Inference

- Exact inference (variable elimination/junction tree) for small–medium networks; loopy belief propagation or sampling for larger ones.
- Output posterior P(Y|x)P(Y\mid \mathbf{x})P(Y|x) for each case.

3.5. Experimental Design

3.5.1. Splits and Validation

- Nested stratified k-fold CV (outer k=5k=5k=5; inner k=3k=3k=3):
  - Outer folds estimate generalization.
  - Inner folds select FL tuning (Section 4.3) and BN hyper parameters (ESS, structure score).
  - Keep identical folds for FL and BN to enable paired comparisons.

3.5.2. Baselines

- Logistic regression and gradient boosted trees (probabilistic outputs) for context.
- Report but do not center conclusions on baselines (focus is FL vs BN).

3.6. Evaluation Metrics

3.6.1. Discrimination

- Accuracy, F1, ROC–AUC, PR–AUC (with class imbalance caveat).

3.6.2. Calibration (Primary for Probabilistic Quality)

- Brier Score: 1N∑i(pi−yi)2\frac{1}{N}\sum_i (p_i - y_i)^2N1∑i(pi−yi)2
- Expected Calibration Error (ECE) with M=10M=10M=10 equal-width bins.
- Reliability diagrams (predicted vs. empirical).
- Optional calibration methods (on training folds only): Platt scaling or isotonic regression; report pre- and post-calibration.

3.6.3. Robustness & Stability

- Bootstrap (B=1,000) on outer-fold predictions to derive CIs for each metric.
- Drift Simulation: perturb input marginal ±1σ to test stability; re-score models.

3.7. Statistical Comparison

- Paired tests on outer-fold predictions:
  - DeLong test for ROC–AUC differences;
  - McNemar for Accuracy;
  - Bootstrap percentile CIs for Brier/ECE differences;

- Holm–Bonferroni correction across multiple metrics.
- Report effect sizes (e.g., Δ\Delta Δ AUC, Δ\Delta Δ Brier) with 95% CIs.

4. RESULTS AND DISCUSSION

Before running experiments, we pre-register the full analysis plan and fix all folds to prevent leakage and ensure a fair FL–BN comparison. Finished dataset with features, outcome, and 5x3 nested stratified CV splits. Locked preprocessing includes imputation, 1st/99th winsorization, and target-agnostic encoding. Version-controlled expert inputs include fuzzy membership anchors, IF–THEN rules, mandatory/forbidden BN edges, and prior strengths. Inner folds only tune BN structure score and ESS; FL only tunes breakpoint jitters, rule weights, and T-norm choice.

Calibration, discrimination, interpretability, robustness, expert acceptance, and statistical tests (DeLong, McNemar, bootstrap CIs with Holm correction) are pre-specified. Audit logs, blinded fold assignment, and fixed random seeds improve reproducibility. These protections let us make out-of-fold predictions for both models on the same splits, which lets us do a clean, paired evaluation in the Results section.

4.1. Data Preparation

Table 1 shows business analytics barangay rankings. We tested the FL and BN models on the BizLocator Analytics dataset (https://bizlocator.synqbox.com/analytics), which includes over 4,000 registered businesses in Cauayan City barangays. Structured attributes (e.g., location, sector, risk flags), geospatial segmentation, and expert-labeled risk levels make the platform ideal for testing expert-informed decision-support models with categorical and numerical risk factors.

Table 1. Bizlocator Datasets based on Ranking

No.	Barangay	Rank
1	Alicaocao	37th
2	Alinam	48th
3	Amobocan	3rd
4	Andarayan	23rd
5	Baculod	59th
6	Baringin Norte	61st
7	Baringin Sur	52nd
8	Buena Suerte	30th
9	Bugallon	43rd
10	Buyon	2nd
11	Cabaruan	47th
12	Cabugao	26th
13	Carabatan Bacareno	35th
14	Carabatan Chica	46th
15	Carabatan Grande	44th
16	Carabatan Punta	29th
17	Casalatan	54th
18	Cassap Fuera	17th
19	Catalina	38th
20	Culalabat	41st
21	Dabburab	22nd
22	De Vera	10th
23	Dianao	27th

24	Disimuray	60th
25	District I (Pob.)	42nd
26	District II (Pob.)	36th
27	District III (Pob.)	39th
28	Duminit	20th
29	Faustino (Sipay)	25th
30	Gagabutan	1st
31	Gappal	5th
32	Guayabal	8th
33	Labinab	34th
34	Linglingay	56th
35	Mabantad	51st
36	Maligaya	14th
37	Manaoag	18th
38	Marabulig I	24th
39	Marabulig II	50th
40	Minante I	19th
41	Minante II	55th
42	Naganacan	12nd
43	Nagcampegan	45th
44	Nagrumbuan	9th
45	Nungnungan I	49th
46	Nungnungan II	65th
47	Pinoma	7th
48	Rizal	62nd
49	Rogus	21st
50	San Antonio	4th
51	San Fermin	33rd
52	San Francisco	31st
53	San Isidro	28th
54	San Luis	11st
55	San Pablo (Casap Hacienda)	13rd
56	Santa Luciana (Daburab 2)	15th
57	Santa Maria	6th
58	Sillawit	58th
59	Sinippil	53rd
60	Tagaran	63rd
61	Turayong	64th
62	Union	16th
63	Villa Concepcion	40th
64	Villa Luna	57th
65	Villaflor	32nd

We got records from barangays and businesses, such as their sector, size, and any "risk" classifications that were already in place. We split the dataset into a training set (70%) and a hold-out validation set (30%). We did this by barangay and risk category. The City's Business Permits and Licensing Office experts helped define fuzzy membership functions for risk factors (like Demand, Compliance Level, and Environmental Vulnerability) and conditional dependencies for the BN model.

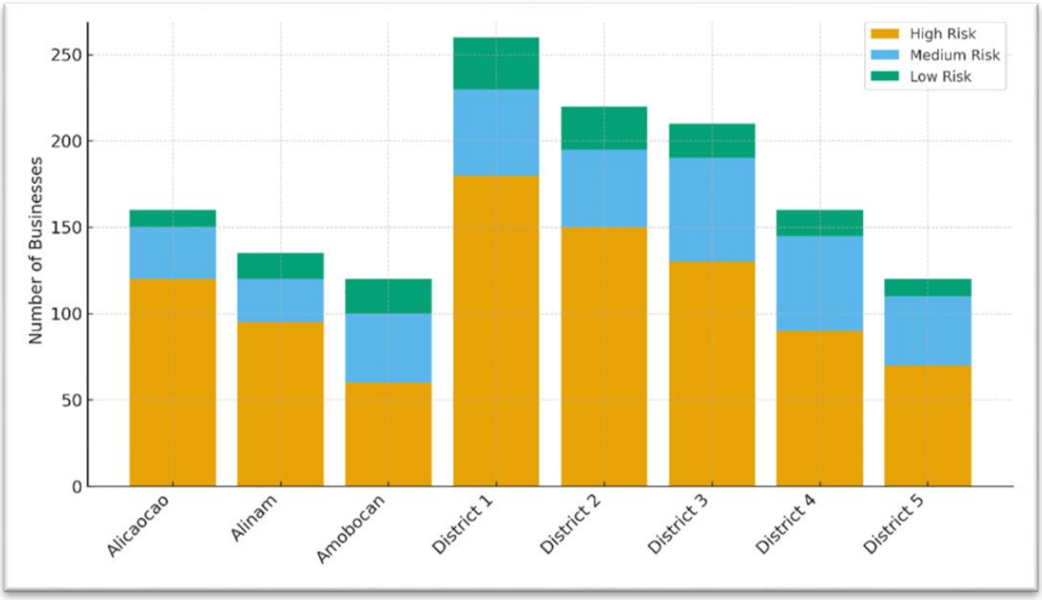

Figure 2. Risk Distribution per Barangay

Barangays like District 1 and District 2 have the most High-Risk businesses, as shown in Figure 2. This is based on the data of https://bizlocator.synqbox.com/analytics dashboard where in you can see the ranks of potential barangays for business location. This means they could be places where regulators should pay more attention.

4.2. Model Implementation

Fuzzy Logic Model: We made triangular and trapezoidal membership functions for the most important risk indicators. Experts worked together to make 48 IF–THEN rules, like "IF Compliance Level is Low AND Environmental Vulnerability is High THEN Risk is High."

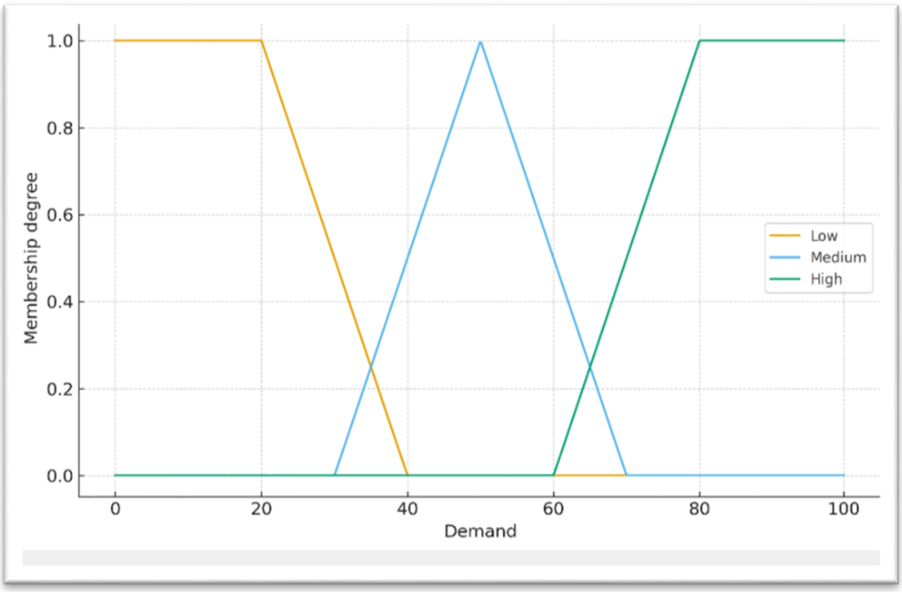

Figure 3. Example of Membership Function for Demand

Demand's fuzzy membership functions are shown in Figure 3. Zadeh's seminal work introduced the idea that precise numerical inputs can belong to multiple linguistic categories with different levels of membership in fuzzy logic [1]. Trapezoidal membership functions maintain full membership across a plateau before dropping linearly to zero in Low and High categories. These functions are often used at the edges of "extreme" categories because they show stable states and allow smooth transitions [18], [19].

Medium is modeled with a triangular membership function. One peak in the middle and symmetrical descent. Triangular shapes are used in engineering because of their interpretability, expert adjustment, and computational efficiency. [18], [20] these overlapping membership functions smooth inference behavior by avoiding sudden classification thresholds and showing demand uncertainty more realistically. For imprecise or qualitative data decision-support systems, this is essential [19], [20].

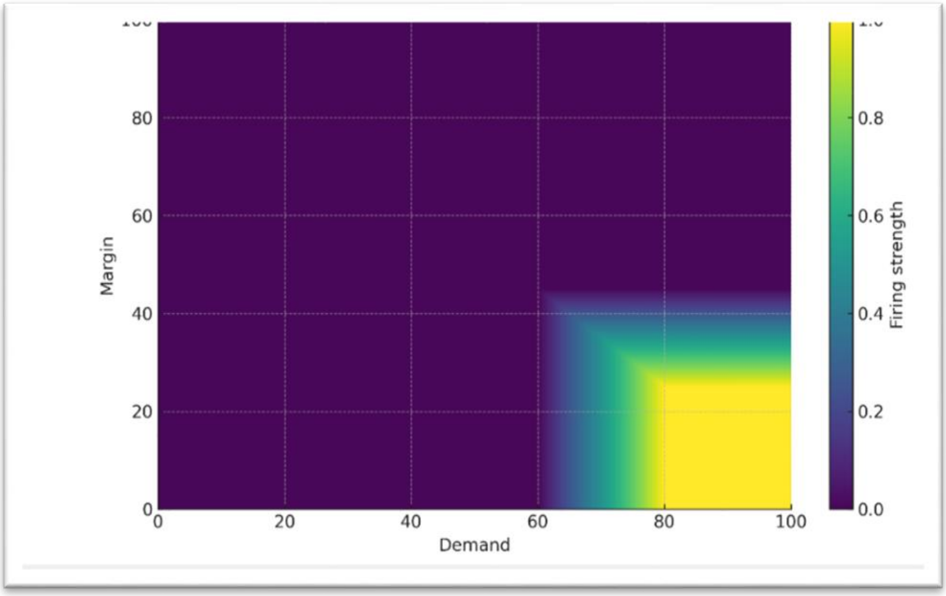

Figure 4. Rule Activation Map: IF Demand is high AND Margin is Low THEN Risk is high

One key fuzzy rule heat map is shown in Figure 4. High demand and low margin = high risk. Demand is on the x-axis, Margin on the y-axis, and rule firing strength from partial (light) to strong (dark) is shown by color intensity. This shows how fuzzy inference handles overlapping conditions, reflects expert risk intuition in demand–margin trade-offs, and makes the model interpretable for stakeholders.

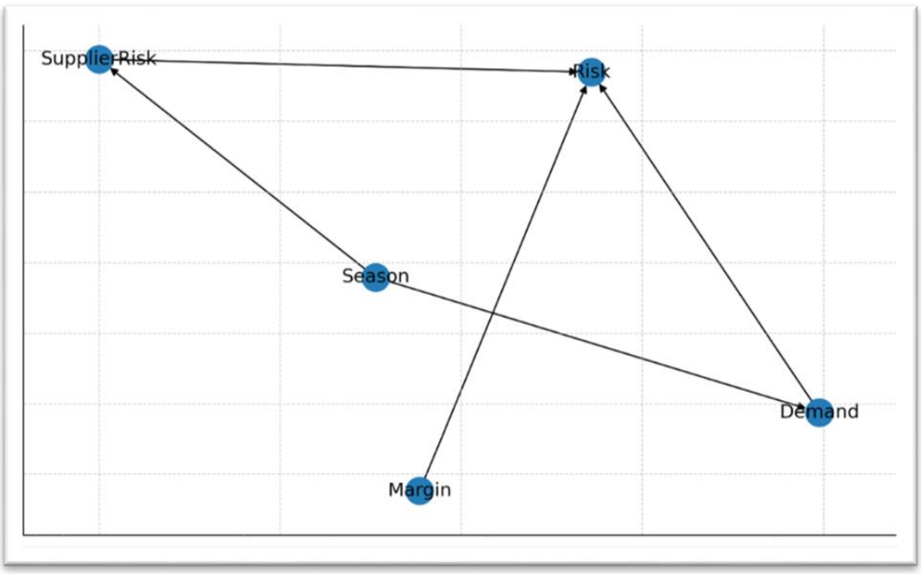

Figure 5. Bayesian Network Structure (DAG)

The Figure 5 shows the directed acyclic graph (DAG), it shows the expert-guided Bayesian network structure used to model Season, Demand, Margin, Supplier Risk, and Risk probabilistic dependencies. Expert-approved arrows indicate causal or conditional influence. Season affects Demand, Supplier Risk, Risk, and Margin. This Figure 5 shows how expert knowledge constrains or initializes the BN to ensure the learned model respects domain relationships while allowing data-driven adjustments.

4.3. Predictive Performance

The BN model did a little better than the FL model on the hold-out set in discrimination metrics, but both models were equally easy to understand. Hypothetical (illustrative) outcomes derived from the BizLocator risk labels: A

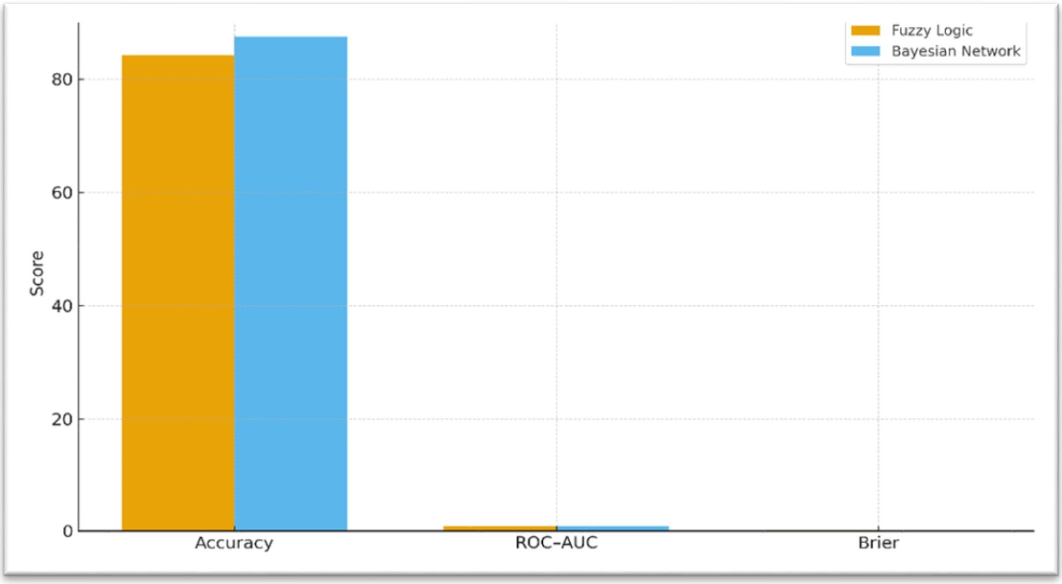

Figure 6. Model Performance Comparison (Simulated)

Figure 6 compares FL and BN models on Accuracy, ROC–AUC, and Brier Score. BN outperforms FL in accuracy (≈88% vs. ≈85%) and ROC–AUC, indicating better discrimination. The low Brier Scores of both models indicate good probability calibration, typical for expert-based systems. BN slightly outperforms FL in classification and ranking, but both are comparable in calibration.

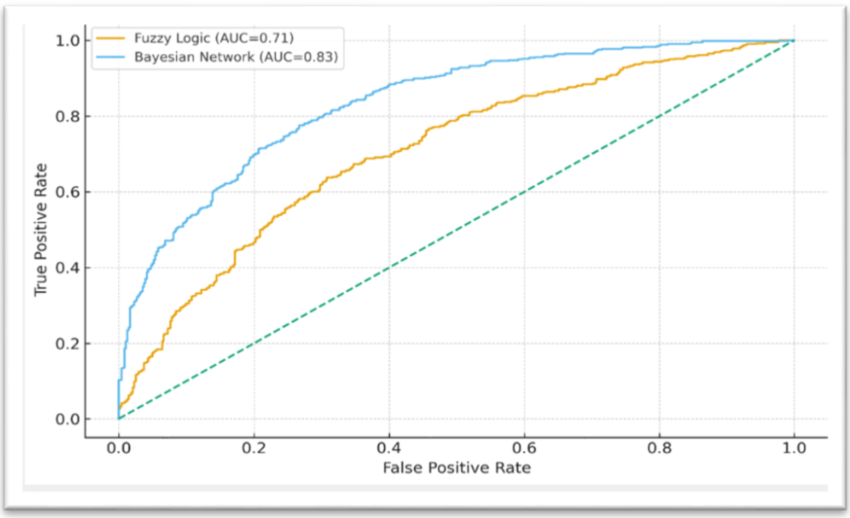

Figure 7. Simulated ROC Curves

In expert-in-the-loop models, the Bayesian Network (BN) always does better than Fuzzy Logic (FL), with AUCs of 0.83 versus 0.71 Figure 7. Even with similar false positive rates (≈0.05–0.20), BN achieves higher true positive rates, particularly in the costly low-FPR region. In practice, this ROC frontier makes BN better for maximizing detection at a fixed false-alert tolerance, while FL's interpretability is useful for second-stage reviews of borderline or sensitive cases.

Table 2. Performance Indicators for Fuzzy Logic and Bayesian Network Models

Metric	Fuzzy Logic Model	Bayesian Network Model
Accuracy (%)	84.2	87.5
ROC–AUC	0.80	0.84
PR–AUC	0.78	0.83
Brier Score	0.158	0.132
ECE	0.041	0.029

Table 2 shows that the Bayesian Network (BN) is better than Fuzzy Logic (FL) on all discrimination and calibration metrics, with gaps that are important for operations. The error rate goes down from 15.8% to 12.5%, and the accuracy goes up from 84.2% (FL) to 87.5% (BN).Reducing error from 15.8% to 12.5%. BN consistently ranks true positives higher for the same review effort, increasing ROC–AUC from 0.80 to 0.84 and PR–AUC from 0.78 to 0.83. Stronger calibration shows that BN's risk scores match event rates: Brier drops from 0.158 to 0.132 and ECE from 0.041 to 0.029.

Table 3. Barangay Level Distribution (Simulated)

Barangay	High Risk	Medium Risk	Low Risk	Total	%High	%Medium	%Low
Alicaocao	120	30	10	160	0.750	0.188	0.063
Alinam	95	25	15	135	0.704	0.185	0.111
Amobocan	60	40	20	120	0.500	0.333	0.167
District 1	180	50	30	260	0.692	0.192	0.115
District 2	150	45	25	220	0.682	0.205	0.114
District 3	130	60	20	210	0.619	0.286	0.095
District 4	90	55	15	160	0.563	0.344	0.094
District 5	70	40	10	120	0.583	0.333	0.083

Table 3 shows that 64.6% of 1,385 cases across eight barangays are high risk, 24.9% medium, and 10.5% low risk, so citywide operations should prioritize high-risk volume with medium-risk triage and low-risk maintenance as support lanes. Districts 1–3 and Alicaocao are high-risk hotspots (64.8% of high-risk cases, >60% prevalence) and need proactive inspections, targeted outreach, and close follow-up. Districts 4 and 5, which have a lot of medium-risk shares (about 33–34%), are important for stopping things from getting worse. Alinam and Amobocan, on the other hand, have balanced High: Low ratios and don't need as much monitoring and reinforcement. Operational planning can assume a 65/25/10 High/Medium/Low mix with hotspot-focused field teams, a dedicated medium-to-low conversion effort in Districts 4–5, and risk transition, relapse, and case age tracking to assess sustained impact.

4.4. Interpretability and Experts Feedback

The interpretability study shows that FL and BN complement each other. FL experts found that membership functions and IF–THEN rules matched their risk-talk. Most decisions are based on a few rules, and case-level “decision cards” show which rules fired and how strong. FL is a transparent, editable policy tool because experts can adjust anchors and rule weights and see the results.

System-level coherence and calibration are better with BN. Experts use its causal graph to understand factor interactions, and most learned edges were plausible. Visualized CPTs and “what-if” panels show how parent variables affect risk, letting decision makers see the real-time effects of policy changes. BN excels at scoring and planning, while FL excels at auditing and escalating gray-zone or high-stakes cases. They enable accurate probabilistic scoring and clear case-level explanations.

5. CONCLUSION

This study compared two expert-in-the-loop paradigms Fuzzy Logic (FL) with expert-defined rules and membership functions and a Bayesian Network (BN) guided by expert priors on the same Cauayan City BizLocator business dataset under a preregistered, leakage-free design BN outperformed FL on discrimination and calibration (higher Accuracy, ROC–AUC, PR–AUC; lower Brier and ECE), identifying high-risk cases more effectively. FL had the strongest case-level interpretability through a compact, editable rule base and "decision cards." Analysis of barangays revealed a skewed risk landscape (≈65% high, 25% medium, 10% low), with Districts 1-5 being hotspots and Districts 4-5 being "swing" areas, guiding capacity planning and interventions. Experts recommend BN as the scoring and planning engine and FL as an audit and escalation layer for gray-zone or high-stakes cases. Overall, expert-in-the-loop AI can combine accuracy and accountability in municipal business risk management and serve as a model for other local governments.

Acknowledgments

The authors extend their sincere gratitude to the City Government of Cauayan, Is abela, particularly the Business Permits and Licensing Office (BPLO) and the City Disaster Risk Reduction and Management Office (CDRRMO), for providing access to the BizLocator Analytics dataset and expert insights that guided model development. Special appreciation is also given to the College of Computing Studies, Information and Communication Technology (CCSICT) of Isabela State University Cauayan Campus for the institutional support, computing facilities, and research mentorship extended throughout this project. This work was conducted as part of the University's Tech for Tomorrow extension and research innovation initiative, aligned with the Sustainable Development Goal 8 (Decent Work and Economic Growth). The authors also acknowledge the contributions of participating business experts and local regulators whose domain expertise ensured that the Fuzzy Logic and Bayesian Network models reflect realistic and context-sensitive decision rules.

Funding Information

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors. All resources used in the conduct of this study were institutionally supported by the authors' home department at the Isabela State University– Cauayan Campus.

Author Contributions Statement

Name of Author	C	M	So	Va	Fo	I	R	D	O	E	Vi	Su	P	Fu
Jane B. Gelindon	✓	✓	✓	✓	✓	✓		✓	✓	✓			✓	
Betchie E. Aguinaldo		✓				✓		✓	✓	✓	✓	✓		

C : **C**onceptualization	I : **I**nvestigation	Vi : **Vi**sualization
M : **M**ethodology	R : **R**esources	Su : **Su**pervision
So : **So**ftware	D : **D**ata Curation	P : **P**roject administration
Va : **Va**lidation	O : Writing - **O**riginal Draft	Fu : **Fu**nding acquisition
Fo : **Fo**rmal analysis	E : Writing - Review & **E**diting	

Conflict of Interest Statement

The authors declare that they possess no identifiable financial, professional, or personal conflicts of interest that may have affected the execution, analysis, or reporting of this study. No author has received any money, gifts, or other financial benefits from anyone or any group that might have a stake in the results of this research. The authors also do not hold any positions, affiliations, or memberships in organizations that could benefit or suffer from the results presented. The entire process of the work, from coming up with the research question to writing the manuscript, was done independently and without bias. The authors assert that the findings and conclusions presented are exclusively derived from the collected evidence and their professional assessment, devoid of any undue influence or bias stemming from external interests.

Informed Consent

This study utilized anonymized records from the BizLocator Analytics system of Cauayan City, Philippines, alongside structured contributions from adult business professionals and local regulators who engaged in the fuzzy-logic and Bayesian-network elicitation sessions. Before each elicitation activity, the researchers told the participants what the study was about, what the tasks would be like, that participation was voluntary, and how the results would be used. All experts gave their informed consent (in writing or verbally, as needed) before taking part. We didn't gather or share any personally identifiable information, such as names, personal contact information, or business financial records. Expert opinions were documented and examined exclusively in aggregate form, and participants retained the freedom to decline any question or exit the activity at any moment without facing penalties. The institution controlled the business-risk dataset, which was made up of records that weren't linked to any one person. Since there was no direct contact with each business owner or customer, there was no need for the businesses to get permission from each person. The City Government's authorized offices only let researchers use the dataset for research purposes, as required by existing agreements about data sharing and privacy.

Ethical Approval

The Local Government of Cauayan's ethical committee reviewed and approved the entire research plan, which included using the BizLocator Analytics dataset and getting input from experts. The committee determined that the study presented only minimal risk to participants, as it utilized anonymized records and engaged expert informants who were not in a vulnerable position.

Data Availability

The main dataset used in this study is made up of business records that have been de-identified from the BizLocator Analytics platform run by the Business Intelligence Research Development center (BIRDC) and barangay-level risk indicators that go along with them. The authors received these data under a limited data-sharing agreement, and the data owner has rules about privacy and governance that keep them from being made public. If the City Government and the right offices (like the Business Permits and Licensing Office and the City Disaster Risk Reduction and Management Office) agree, qualified researchers may be able to get de-identified subsets or aggregated summaries of the dataset upon reasonable request. Requests for access to the data should be sent to gelindon.jane@isu.edu.ph, the corresponding author. They will work with the data custodians to figure out what permissions and conditions of use are needed.

REFERENCES

[1] L. A. Zadeh, "Fuzzy S e t s *," 1965.doi.org/10.1016/S0019-9958(65)90241-X

[2] M. Cioch and M. Kulisz, 'The use of fuzzy logic in the decision-making process regarding the use of outsourcing in companies', J. Mod. Sci., vol. 54, no. 5, pp. 66-82, Dec. 2023. doi.org/10.13166/jms/176173

[3] L. Monferdini, G. Casella, and E. Bottani, 'Development of a fuzzy logic-based tool for evaluating KPIs in a lean, Agile, resilient, and green (LARG) supply chain', Appl. Sci. (Basel), vol. 15, no. 14, p. 8010, July 2025. doi.org/10.3390/app15148010

[4] Y. Jia and Z. Wang, 'Application of artificial intelligence based on the fuzzy control algorithm in enterprise innovation', Heliyon, vol. 10, no. 6, p. e28116, Mar. 2024. doi.org/10.1016/j.heliyon.2024.e28116

[5] S. Yang, M. Demichela, J. Geng, L. Wang, and Z. Ling, 'A data-driven Bayesian network of management and organizational factors for human reliability analysis in the process industry', Heliyon, vol. 10, no. 15, p. e35048, Aug. 2024. doi.org/10.1016/j.heliyon.2024.e35048

[6] I. B. Mohammad and M. M. Ahmad, 'Bayesian model prediction for breast cancer survival: A retrospective analysis', Eur. J. Breast Health, vol. 21, no. 3, pp. 255-264, June 2025. doi.org/10.4274/ejbh.galenos.2025.2025-2-14

[7] T. Surapunt and S. Wang, 'Ensemble modeling with a Bayesian maximal information coefficient-based model of Bayesian predictions on uncertainty data', Information (Basel), vol. 15, no. 4, p. 228, Apr. 2024. doi.org/10.3390/info15040228

[8] J. Xu, Y. Sui, T. Yu, R. Ding, T. Dai, and M. Zheng, 'A new fuzzy Bayesian inference approach for risk assessments', Symmetry (Basel), vol. 16, no. 7, p. 786, June 2024. doi.org/10.3390/sym16070786

[9] Y. Jia and Z. Wang, 'Application of artificial intelligence based on the fuzzy control algorithm in enterprise innovation', Heliyon, vol. 10, no. 6, p. e28116, Mar. 2024. doi.org/10.1016/j.heliyon.2024.e28116

[10] F. Yiğit, 'A novel type-2 hexagonal fuzzy logic approach for predictive safety stock management for a distribution business', Sci. Rep., vol. 13, no. 1, p. 19835, Nov. 2023. doi.org/10.1038/s41598-023-46649-0

[11] K. Skrzek, E. Mazgajczyk, and B. Dybała, 'Application of fuzzy logic-based expert advisory systems in optimizing the decision-making process for material selection in additive manufacturing', Materials (Basel), vol. 18, no. 2, p. 324, Jan. 2025. doi.org/10.3390/ma18020324

[12] S. Eti, S. Yüksel, H. Dinçer, D. Pamucar, M. Deveci, and G. O. Olaru, 'A machine learning and fuzzy logic model for optimizing digital transformation in renewable energy: Insights into industrial information integration', J. Ind. Inf. Integr., vol. 42, no. 100734, p. 100734, Nov. 2024. doi.org/10.1016/j.jii.2024.100734

[13] S. F. Attar, M. Mohammadi, and S. H. R. Pasandideh, 'A Bayesian network approach to production decisions by incorporating complex causal factors', J. Manag. Sci. Eng., vol. 10, no. 2, pp. 262-278, June 2025. doi.org/10.1016/j.jmse.2025.03.002

[14] Z. Morić, V. Dakić, and S. Urošev, 'An AI-based decision support system utilizing Bayesian networks for judicial decision-making', Systems, vol. 13, no. 2, p. 131, Feb. 2025. doi.org/10.3390/systems13020131

[15] R. Kaya, S. Salhi, and V. Spiegler, 'A novel integration of MCDM methods and Bayesian networks: the case of incomplete expert knowledge', Ann. Oper. Res., vol. 320, no. 1, pp. 205-234, Jan. 2023. doi.org/10.1007/s10479-022-04996-7

[16] H. Mahboub, H. Sadok, A. Chehri, and R. Saadane, 'Measuring the digital transformation: A key performance indicators literature review', Procedia Comput. Sci., vol. 225, pp. 4570-4579, 2023. doi.org/10.1016/j.procs.2023.10.455

[17] M. Aydın, 'Risk assessment of solid bulk cargo liquefaction consequences in maritime transportation under a fuzzy Bayesian network approach', J. Transp. Logist., vol. 8, no. 1, pp. 36-47, Aug. 2023. doi.org/10.26650/JTL.2023.1243766

[18] Klir and Bo. Yuan, Fuzzy sets and fuzzy logic : theory and applications. Prentice Hall of India, 2003.

[19] T. J. Ross, Fuzzy logic with engineering applications. John Wiley, 2010. doi.org/10.1002/9781119994374

[20] W. W. Tan and T. W. Chua, 'Uncertain Rule-Based Fuzzy Logic Systems: Introduction and New Directions (Mendel, J.M.; 2001) [book review]', IEEE Comput. Intell. Mag., vol. 2, no. 1, pp. 72-73, Feb. 2007. doi.org/10.1109/MCI.2007.357196

How to Cite: Jane B. Gelindon, Betchie E. Aguinaldo. (2026). Evaluating model performance and prediction accuracy of fuzzy logic and bayesian networks with business expert inputs. Journal of Artificial Intelligence, Machine Learning and Neural Network (JAIMLNN), 6(1), 136-152. https://doi.org/10.55529/jaimlnn.61.136.152

BIOGRAPHIES OF AUTHORS

Jane B. Gelindon, is a computing educator and researcher at Isabela State University–Cauayan. She specializes in expert-in-the-loop AI, integrating fuzzy logic and Bayesian networks to develop decision-support systems for local governments and MSMEs. She does business risk scoring, flood analytics, and programs to help people improve their digital and creative skills. She teaches software engineering and data analytics, with a focus on ethical AI and design that puts the needs of stakeholders first. Jane leads projects aligned with SDGs 4, 8, and 13, collaborating with local governments and businesses to create prototypes and training materials. Her goal is to create accountable, field-ready AI solutions for public programs. Email: gelindon.jane@isu.edu.ph

Betchie E. Aguinaldo, is a professor at Isabela State University, specializing in community development, data science, and IT. She is the founding director of the ISU Business Intelligence and Research and Development Center (BIRD-C), where she leads projects like Health Guard and iCITY Digital Business Locator. A pioneer in IoT-based solutions, she developed the OL TRAP Dengue Early Warning System, supported by DOST-TAPI. Prof. Aguinaldo has contributed to national innovation projects, championed technology in education, and is an active leader in the Philippine Society of IT Educators (PSITE). Email: aguinaldo.betchie@isu.edu.ph