From Conventional Fuzzy Systems to Advanced Interval Type-2 Optimization: GA-Based Approaches for Robust and Interpretable OBE Analytics

Annamalai Ravishankar Ravitha1,a), Ponnan Vijayalakshmi1,b)

*1Department of Mathematics, Periyar Maniammai Institute of Science & Technology (Deemed to be University), Thanjavur, 613 403, Tamil Nadu, India*Corresponding author: a)[ravitha@pmu.edu](mailto:ravitha@pmu.edu), b)[vijayalakshmi@pmu.edu](mailto:vijayalakshmi@pmu.edu)

**Abstract:** Interval Type-2 Fuzzy Inference Systems (IT2-FIS) provide a principled approach to represent and manage epistemic and aleatory uncertainty in applied prediction tasks. This paper introduces a Genetic Algorithm-optimized IT2-FIS (GA-IT2 FIS) for predicting Course Outcome–Program Outcome (CO-PO) attainment from six institutional indicators: Student Performance, Course Complexity, Faculty Effectiveness, Assessment Rigour, Student Engagement, and Teaching Quality Feedback. The proposed framework jointly optimizes membership function parameters, LMF/UMF scale factors, rule activation flags, and consequent parameters; we additionally evaluate a multi-objective NSGA-II variant that trades accuracy against interpretability. Performance is compared to a Type-1 GA-FIS baseline, ANFIS, and a multilayer perceptron, and robustness is assessed under additive noise, asymmetric rater bias, label corruption, and missingness. On a 10,000-record dataset the GA-IT2 FIS attains RMSE = 0.0325 and R² = 0.962, yielding ≈7.8% RMSE reduction relative to the Type-1 baseline, while producing calibrated 90% prediction intervals with 91.4% empirical coverage and favorable Winkler scores. Robustness tests report 30–45% lower error escalation for IT2 models. The method preserves interpretability with a compact rule base (~42 active rules) and supports real-time deployment (inference ≈ 2.3 ms/sample). Code and artifacts are publicly available to ensure reproducibility and transparency.

**Keywords:** Interval Type-2 Fuzzy Inference Systems, Genetic Algorithm optimization, Uncertainty quantification, Outcome-Based Education (CO-PO attainment), Robustness to noise and missingness

# Introduction

Outcome-Based Education (OBE) increasingly relies on data-driven analytics to monitor curriculum delivery, measure student learning outcomes, and support evidence-based pedagogical decisions. Blumenstein, et al. (2020) Educational datasets, however, are frequently afflicted by measurement noise, rater bias, missing entries and subjective judgments arising from human assessors. Tempelaar et al. (2020). These forms of *deep uncertainty* impair the reliability of point-prediction models and undermine stakeholder trust when model outputs are delivered without quantifiable confidence. Yang, C., & Li, Y. (2023). At the same time, practitioners and policy-makers in education demand interpretable models whose internal logic (rules, linguistic terms) can be inspected and validated by domain experts Navarro-Almanza, et al. (2021). There is therefore a pressing need for predictive models that (i) explicitly represent and quantify uncertainty, (ii) remain robust under realistic data perturbations, and (iii) preserve human interpretability for adoption in educational practice. Rishabh Singh et al. (2020)

Fuzzy inference systems are a natural fit for this triad of requirements. Type-1 fuzzy systems provide linguistic transparency but lack an intrinsic mechanism to represent secondary uncertainty. Dongrui Wu et al. (2023). Interval Type-2 Fuzzy Inference Systems (IT2-FIS) extend this framework with *footprints of uncertainty* (FOUs) that capture imprecision in membership grades, enabling interval-valued outputs and calibrated prediction intervals. Ata Köklü et al(2025).Despite theoretical advantages, the practical application of IT2-FIS to complex, multivariate educational datasets has been limited by (a) the difficulty of setting many continuous parameters (membership function shapes, FOU sizes) and (b) the interpretability vs. accuracy trade-off when rule bases grow large. Zhuo Wang et al(2023). Evolutionary optimization both single-objective Genetic Algorithms (GAs) and multi-objective algorithms such as NSGA-II offers a principled route to tune continuous and discrete model components jointly, balancing predictive accuracy, uncertainty parsimony, and rule-base compactness. Huimin Zhao et al (2025).

This paper proposes and empirically validates a GA-optimized IT2-FIS framework for OBE analytics, applied to a real-world dataset comprising six input indicators Student Performance, Course Complexity, Faculty Effectiveness, Assessment Rigour, Student Engagement, Teaching Quality Feedback and a continuous target CO-PO Attainment. The framework jointly optimizes membership function parameters, LMF/UMF scale factors, rule activation flags and consequent parameters. For comparison and design trade-off exploration we also evaluate an NSGA-II multi-objective variant that yields Pareto fronts trading accuracy against interpretability.

# Motivation and Research Gaps

Prior work has established the theoretical robustness of IT2 models in noisy and uncertain environments, but empirical evidence remains sparse in the context of educational analytics particularly studies that (1) combine interval fuzzy logic with evolutionary optimization tuned for parsimony, (2) quantify uncertainty calibration (coverage, Winkler score) in addition to point accuracy, and (3) rigorously evaluate robustness to realistic stressors such as rater bias, label corruption and missingness. Reviewers and practitioners frequently request reproducibility details, ablation studies and practical runtime information to assess deployment feasibility all of which are addressed in this study.

## Objectives and Hypotheses

This study aims to answer the following research questions:

1. Can a GA-optimized IT2-FIS deliver statistically significant improvements in point prediction accuracy over a well-tuned Type-1 GA-FIS and standard neural baselines (ANFIS, MLP) for CO-PO Attainment?
2. To what extent does the IT2-FIS provide calibrated and *useful* interval predictions (coverage close to nominal, narrow widths) for decision support?
3. How robust is IT2-FIS relative to Type-1 and neural baselines under additive noise, asymmetric rater bias, label corruption and missingness?
4. Is it possible to retain interpretability (compact rule base and semantically ordered MFs) while achieving the above gains within realistic computational budgets suitable for educational dashboards?

We hypothesize that the interval representation and GA tuning will yield (i) lower RMSE, (ii) well calibrated 90% prediction intervals, and (iii) substantially lower error escalation under stressors, all while maintaining a compact and interpretable rule base.

# Contributions

The principal contributions of this work are:

1. **Methodological Framework.** We present a reproducible pipeline that integrates Interval Type-2 fuzzy modeling with Genetic Algorithm and NSGA-II based optimization to jointly tune continuous MF parameters and discrete rule activations for OBE prediction tasks.
2. **Empirical Evidence.** On a 10,000-record institutional OBE dataset the GA-optimized IT2-FIS achieves an RMSE of **0.0325** (R² = 0.962), a ~7.8% RMSE reduction relative to a Type-1 GA-FIS baseline (RMSE = 0.0345), and statistically significant improvements (Wilcoxon, p < 0.01; Cliff’s δ > 0.33).
3. **Uncertainty Calibration.** The proposed IT2-FIS produces calibrated 90% prediction intervals with empirical coverage **91.4%**, a mean interval width of **0.072**, and favorable Winkler scores demonstrating practical uncertainty quantification beyond point estimates.
4. **Robustness Analysis.** Under additive Gaussian noise, asymmetric rater bias, 5% label corruption and 10% missingness, IT2-FIS reduces error escalation by **30 - 45%** relative to Type-1 systems, validating its resilience to realistic perturbations in educational data.
5. **Interpretability & Efficiency.** GA tuning yields a compact rule base (≈42 active rules versus 67 for the Type-1 baseline) and preserves linguistic semantics; computational performance is compatible with real-time dashboards (inference ≈ **2.3 ms/sample**, training ≈ **115 s/run**).

## Significance

By delivering an uncertainty-aware, robust and interpretable predictive model, this work bridges an important gap between advanced fuzzy-logic methods and practical educational analytics requirements. The combination of calibrated intervals and compact rule bases supports trustworthy decision support for instructors, curriculum designers and accreditation bodies.

***Paper Organization:*** The remainder of the paper is organized as follows. Section II reviews related work on fuzzy inference systems, interval type-2 modeling, and evolutionary optimization in applied prediction domains. Section III details the methodology, including variable definitions, IT2-FIS architecture, GA and NSGA-II encodings, and robustness experiment design. Section IV presents experimental results: predictive performance, uncertainty calibration, robustness tests, interpretability metrics, and computational analysis. Section V discusses findings, limitations and implications for deployment. Finally, Section VI concludes and outlines future research directions.

# Literature Review

## Predictive Modeling in Outcome-Based Education (OBE)

Outcome-Based Education (OBE) has emerged as a globally adopted framework to ensure curriculum alignment with program outcomes and accreditation requirements. Predictive analytics plays a critical role in this process, enabling institutions to monitor student learning trajectories, identify at-risk learners, and inform pedagogical interventions. Farhood et al. (2024) highlighted the potential of artificial intelligence for forecasting learning outcomes, emphasizing both performance gains and transparency in educational assessment. Similarly, Nayak et al. (2023) demonstrated the effectiveness of machine learning classification models in mining educational data for performance prediction, while cautioning against overfitting and poor generalizability across diverse cohorts. Ramaswami et al. (2022) extended this argument by underscoring the importance of developing generic and portable models to ensure applicability across varied institutional contexts.

Despite these advances, challenges persist. Imran et al. (2019) found that supervised learning methods were useful for academic performance prediction but sensitive to noisy and incomplete datasets, which are common in real-world OBE environments. Systematic reviews further consolidate this concern: Sghir et al. (2022) and Liz-Domínguez et al. (2019) showed that most predictive learning analytics studies from 2012–2022 focused narrowly on accuracy, with less attention paid to robustness, uncertainty calibration, and interpretability. This gap is especially salient in OBE, where data irregularities such as rater bias, subjective assessments, and missing values are prevalent. Agha et al. (2023) specifically examined predictive models for OBE program learning outcomes and concluded that while promising, current approaches underexplore attainment modeling compared to conventional grade prediction.

The interpretability of predictive models is another critical requirement. Wang and Luo (2024) advanced interpretable modeling in higher education, arguing that educators and policymakers require models that are not only accurate but also transparent and explainable. Musso et al. (2020) similarly identified reliable predictors of student outcomes using machine learning but warned that a lack of interpretability could limit trust and adoption. Collectively, these works establish that while predictive modeling is indispensable for OBE, existing approaches face persistent barriers in handling noisy data, ensuring robustness, and delivering interpretable insights that can inform actionable decisions.

## Classical Approaches: Fuzzy, Neural, and Statistical Baselines

Fuzzy inference systems (FIS) have been widely recognized as promising candidates for educational modeling due to their rule-based reasoning and ability to incorporate linguistic terms. Type-1 fuzzy systems, in particular, have long been valued for their interpretability. However, their inability to represent higher-order uncertainties limits their robustness. Ojha et al. (2019), in a review spanning three decades of FIS design, concluded that while heuristic and expert-driven Type-1 models provide transparency, they are often inadequate in handling large-scale, high-dimensional, and noisy data.

In response, neural and hybrid approaches have been proposed to enhance predictive accuracy. Sajid et al. (2024) developed an ensemble deep random vector functional link neural network based on fuzzy inference, showing improved accuracy across domains but at the cost of reduced interpretability a significant drawback in education. Hybrid optimization frameworks have also emerged. Melin and Sánchez (2018) proposed a hierarchical genetic algorithm to optimize Type-1, Interval Type-2, and General Type-2 FIS for modular granular neural networks, demonstrating that evolutionary optimization can simultaneously improve accuracy and maintain compactness. Regression-based fuzzy extensions have further expanded the applicability of fuzzy models. Wiktorowicz (2023) introduced a type-2 regression fuzzy inference system (T2RFIS), offering enhanced flexibility in regression tasks but not addressing interpretability challenges.

These advances demonstrate that while classical fuzzy and neural baselines provide valuable predictive capabilities, they often sacrifice interpretability or robustness. This limitation has motivated the growing adoption of Interval Type-2 fuzzy inference systems (IT2-FIS), which explicitly model uncertainty through footprints of uncertainty (FOUs).

## Interval Type-2 Fuzzy Inference Systems (IT2-FIS): Opportunities and Challenges

IT2-FIS extend Type-1 fuzzy models by embedding secondary membership grades to form footprints of uncertainty, thereby capturing imprecision more effectively. Wu and Mendel (2019) provided foundational design principles for practical IT2 systems, recommending structured approaches for membership design and type-reduction. Moreno et al. (2020) advanced this by introducing justifiable uncertainty frameworks, reinforcing the empirical necessity of FOUs in uncertainty-rich applications.

Applications of IT2-FIS span diverse domains. Tan et al. (2023) investigated observer-based IT2 fuzzy control under network attacks, while Tahamipour-Z. et al. (2022) explored generalized IT2 models for nonlinear systems, and Mazandarani and Xiu (2021) examined fractional fuzzy inference. Chen (2024) demonstrated hybrid optimization of IT2 fuzzy neural networks, while Gao et al. (2023) integrated Group Lasso regularization for feature selection within IT2 fuzzy neural networks. These studies highlight both the versatility and the adaptability of IT2 models in handling uncertainty. Additionally, Beke and Kumbasar (2023) proposed a composite learning framework, emphasizing that evaluation must consider interpretability and robustness in addition to accuracy.

Despite these advances, significant practical barriers remain. The computational complexity of type-reduction processes, challenges in parameter tuning, and risks of rule explosion limit the applicability of IT2-FIS in large-scale domains (Lin et al., 2014; Li et al., 2015). Wu, Peng, and Mendel (2023) further stressed that although IT2-FIS are theoretically superior to Type-1 systems, their adoption has been concentrated in engineering and control applications, with limited empirical deployment in educational analytics. This gap underscores the urgent need for methods that adapt IT2-FIS to domains where uncertainty and interpretability are central, such as OBE.

## Synthesis and Research Gap

The reviewed literature reveals three converging insights. First, predictive modeling in education has matured considerably (Farhood et al., 2024; Nayak et al., 2023; Ramaswami et al., 2022), but most approaches remain narrowly focused on accuracy, often neglecting interpretability and robustness under noisy or incomplete data (Imran et al., 2019; Sghir et al., 2022; Liz-Domínguez et al., 2019; Agha et al., 2023). Second, classical fuzzy and neural models contribute valuable foundations but fail to adequately capture epistemic and aleatory uncertainty (Ojha et al., 2019; Sajid et al., 2024; Wiktorowicz, 2023). Third, IT2-FIS provide a principled mechanism for representing uncertainty (Wu & Mendel, 2019; Moreno et al., 2020), yet practical challenges such as rule explosion, computational cost, and domain-specific adoption persist (Lin et al., 2014; Beke & Kumbasar, 2023; Wu, Peng, & Mendel, 2023).

Addressing these gaps requires a methodological framework that integrates IT2-FIS with evolutionary optimization, particularly Genetic Algorithms, to jointly optimize membership functions, prune redundant rules, and improve robustness. Prior research has shown the promise of evolutionary techniques in fuzzy system optimization (Melin & Sánchez, 2018; Chen, 2024; Gao et al., 2023), but their application in OBE analytics remains largely unexplored. This study addresses that gap by developing and validating a Genetic Algorithm optimized Interval Type-2 FIS (GA-IT2 FIS), offering a compact, interpretable, and uncertainty-aware predictive model tailored to the complexities of OBE datasets.

Table 1 - Summary of the Literature Review

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Study** | **Method** | **Contribution** | **Limitation** | **Identified Gap** |
| Farhood et al. (2024) | AI models for learning outcome prediction | Improved transparency in assessment | Generalizability issues | Need for robust and portable models in OBE |
| Table 1 continued | | | | |
| Nayak et al. (2023) | ML classification for academic performance | Effective data mining for prediction | Limited generalizability | Requires robustness across cohorts |
| Ramaswami et al. (2022) | Generic models in educational data mining | Emphasized portability & reusability | Context-specific limitations | Generic yet adaptable models needed |
| Imran et al. (2019) | Supervised learning for student performance | Useful in prediction tasks | Sensitive to noise & missing data | Robust models required for OBE |
| Sghir et al. (2022) | Systematic review (2012–2022) | Maturity of predictive analytics | Accuracy prioritized over uncertainty | Need uncertainty-aware approaches |
| Liz-Domínguez et al. (2019) | Review of predictive tools in higher education | Comprehensive survey | Little focus on interpretability | Transparent models required |
| Agha et al. (2023) | Predictive models for OBE attainment | Examined PLO attainment | Underexplored vs. grade prediction | Attainment-specific models required |
| Wang & Luo (2024) | Interpretable modeling in higher education | Improved transparency | Limited large-scale testing | Balance accuracy & interpretability |
| Musso et al. (2020) | ML predictors of outcomes | Identified reliable predictors | Interpretability concerns | Adoption barriers remain |
| Ojha et al. (2019) | Review of Type-1 FIS design | Provided interpretability | Not scalable | Scalable fuzzy models needed |
| Sajid et al. (2024) | Ensemble fuzzy deep NN | High accuracy achieved | Sacrificed interpretability | Need interpretable hybrids |
| Melin & Sánchez (2018) | GA for FIS optimization | Improved accuracy & compactness | Complexity of GA tuning | Evolutionary IT2-FIS optimization |
| Wiktorowicz (2023) | Type-2 regression FIS | Flexible regression modeling | Interpretability limits | Transparent regression models |
| Wu & Mendel (2019) | IT2 design guidelines | Structured recommendations | Implementation complexity | Practical IT2 for education |
| Moreno et al. (2020) | IT2 models with justifiable uncertainty | Effective uncertainty capture | Not applied in education | Transfer IT2 to OBE |
| Tan et al. (2023) | IT2 under network attacks | Robustness in adversarial settings | Control system domain | Educational application missing |
|  |  |  |  |  |
| Table 1 continued | | | | |
| Tahamipour-Z. et al. (2022) | Generalized IT2 models | Nonlinear control extension | Engineering-only | Education adaptation needed |
| Mazandarani & Xiu (2021) | Fractional IT2 FIS | Theoretical advance | No domain validation | Empirical testing in OBE |
| Chen (2024) | Hybrid-optimized IT2 NN | Improved predictive accuracy | High complexity | Efficient optimization in OBE |
| Gao et al. (2023) | Group Lasso IT2 NN | Feature selection & accuracy | System identification only | Transfer to education |
| Beke & Kumbasar (2023) | Composite IT2 framework | Robustness focus | Engineering domain | OBE-focused composite models |
| Lin et al. (2014) | Simplified IT2 fuzzy NN | Reduced computational burden | Accuracy loss | Balanced simplification in OBE |
| Li et al. (2015) | IT2 for packet dropouts | Control system robustness | Not educational | Adoption in education |
| Wu, Peng & Mendel (2023) | Review of Type-1 vs IT2 | IT2 superiority | Engineering focus | Extend IT2 to educational analytics |

# Methodology

This section gives a reproducible and rigorous description of the data variables, preprocessing, Interval Type-2 Fuzzy Inference System (IT2-FIS) design, evolutionary optimization, training and model selection, evaluation, and robustness experiments. All algorithmic choices and numerical hyperparameters used in the experiments are reported.

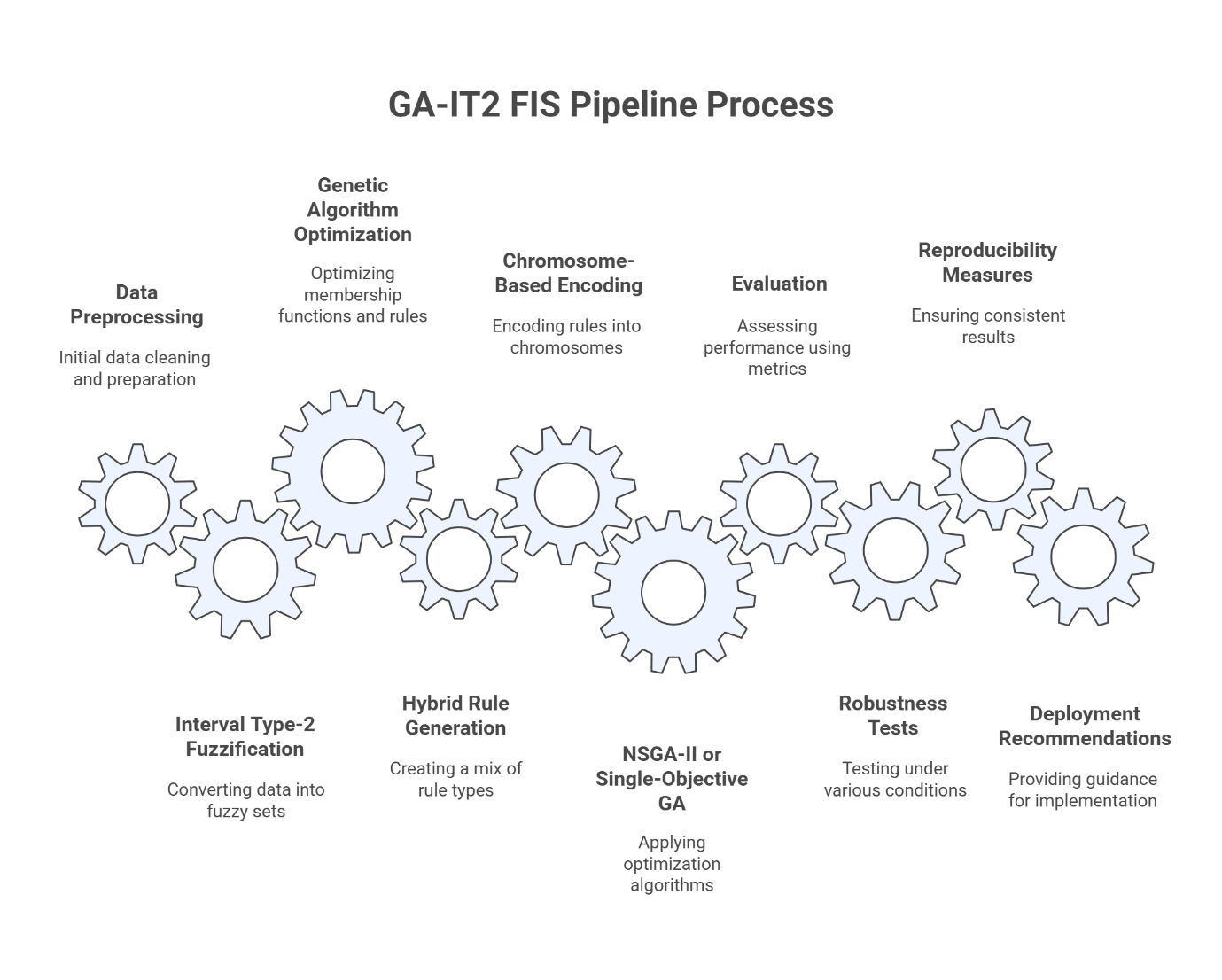


Figure 1- Overall architecture of the GA-IT2 FIS Pipeline Process

## Variables: inputs and output

The modelling problem uses six predictor variables and one continuous target variable. Each variable is continuous and derived from institutional OBE records.

Table 2 - Variables (inputs and output)

|  |  |  |  |
| --- | --- | --- | --- |
| **Variable name** | **Role** | **Type** | **Short description / preprocessing** |
| Student Performance | Input | Continuous | Aggregated score (exam + assignments). Standardize (train mean, sd). Cap at 1st/99th percentiles if extreme. |
| Course Complexity | Input | Continuous | Numeric index of course difficulty. Standardize. If ordinal originally, convert to numeric mapping before scaling. |
| Faculty Effectiveness | Input | Continuous | Composite measure (student ratings + performance metrics). Standardize; missing values handled by interval imputation during robustness tests. |
| Assessment Rigour | Input | Continuous | Score measuring assessment tightness (rubric strictness). Standardize. |
| Student Engagement | Input | Continuous | Engagement index (attendance, LMS activity). Standardize; outliers capped. |
| Teaching Quality Feedback | Input | Continuous | Post-course feedback metric. Standardize. |
| CO-PO Attainment | Output | Continuous | Course Outcome / Program Outcome attainment (target). Normalized to [0,1] for fuzzy consequents and evaluation. |

Notes: All continuous inputs are standardized using statistics computed on the training set and then applied to validation test sets. The target CO-PO Attainment is linearly scaled into [0,1] [0,1] [0,1] for consistent fuzzy consequents and interval computation.

(1)

LMF is a triangular function with the same center but reduced spread:

(2)

Where is the LMF scale factor for variable j. In experiments was intialized as 0.85 ( LMF Width = 0.85 × UMF width) and allowed to be adjusted by the evolutionary optimizer.

## Feature engineering and rationale

1. **No additional categorical encoding** is required because all variables are numeric indices or aggregated scores.
2. **Interactions:** candidate pairwise interaction features (e.g., Student Performance × Student Engagement) were considered during preliminary analysis but not included in the main IT2 rule antecedents to maintain linguistic interpretability; interactions can be captured implicitly via multiple antecedents in rules.
3. **Missing values:** baseline Type-1 models use mean imputation; for IT2 models we use interval imputation (see Section 3.6) to reflect epistemic uncertainty.

## Fuzzification: membership functions and linguistic terms

### Number of linguistic terms

For interpretability and a compact rule base we use **three linguistic terms per input**: *Low*, *Medium*, and *High*. This choice limits full Cartesian product rule explosion (3⁶ = 729) while providing adequate granularity for educational settings. The initial candidate rule pool is restricted to R≤150 (see Section 3.4 for rule generation).

### Triangular interval-type-2 membership functions

Each input x is represented by an interval type-2 triangular fuzzy set with UMF and LMF triangular shapes.

Parameterization for a triangular UMF for term t on variable j:

### Initialization of MF centers and spreads

UMF centers ) are initialized using the empirical quantiles of the standardized training data:

* Low center: 25th percentile,
* Medium center: 50th percentile (median),
* High center: 75th percentile.

Left and right supports are set halfway to adjacent centers; for edge terms the extreme support is extended to cover observed min/max with a small margin (e.g., ± 0.1 in standardized units). This initialization yields semantically ordered, overlapping partitions that are friendly for GA fine-tuning.

## Rule base construction and pruning

### Candidate rule generation

Rather than using the full Cartesian product (which would generate 729 rules), a hybrid procedure generates a candidate pool R (set to 150 in experiments) via:

* **Data-driven rule extraction:** Apply k-means clustering (k chosen such that cluster count ≈ 120) over the input space to identify common antecedent patterns; translate cluster centroids to nearest linguistic terms for each input to form rules.
* **Expert seeds:** Include a small set of domain-informed rules (≈ 30) provided by an educational expert to ensure coverage of known educational relationships.
* **Union & deduplication:** Combine both sets and remove duplicates to form the final candidate pool R≤150.

### Rule representation

Each rule r has antecedents (one linguistic term per input) and a consequent fr (x). For this study we use zero-order consequents (scalar constants) for maximal interpretability. The rule activation flag ρr ∈ {0,1} is part of the optimization chromosome.

### Post-optimization pruning

After optimization, rules with average firing strength below threshold τ (set empirically; e.g., τ=10-3) across the training set are pruned. This yields the reported active rule counts (e.g., ~42 for GA-IT2).

## Inference, type-reduction and defuzzification

Given an input x, each rule r produces interval firing strength [ ] by applying interval valued antecedent degrees and an interval t-norm for conjunction (we use interval product).

Each rule consequent is an interval (for zero order, same scalar but propagated through firing interval). The global output interval [yL,yR] is obtained via Karnik–Mendel (KM) type-reduction:

* KM iterative algorithm with tolerance tol=10-6 and maximum iterations Imax=50.
* Final point estimate for point metrics: y^=(yL+yR)/2.
* The interval [yL,yR] is retained for uncertainty metrics (coverage, width, Winkler score).

## Evolutionary optimization (GA and NSGA-II)

### Chromosome encoding

Chromosome structure:

chrom=[UMF centers(p), UMF spreads(p), LMF scales(p), ρ(R), θ(consequents)]

where p= total number of MF parameters across all inputs and output, and R is candidate rule count.

### Single-objective GA (GA-IT2)

Fitness:

(3)

Penalty values selected via grid search:

Hyperparameters:

Population size P = 30, generations G = 40.

Selection: tournament (size 3)

Crossover: SBX,

Mutation: polynomial mutation for real genes bit-flip for rule flags.

Elitism: top 2 individuals.

Initialization: seeded from an optimized Type-1 GA-FIS; LMS scales initialized at 0.85

### Multi-objective NSGA-II

* Objectives: ⟨RMSEVal,  Nactive⟩ (optionally include FOU area).
* Same encoding and operators; nondominated sorting and crowding distance selection produce Pareto front. Representative operating points selected by inspection of the RMSE vs rule count trade-off.

### Training protocol and reproducibility

* Fitness evaluations use the **training set** for forward inference and the **validation set** (1,000 samples) for computing RMSE val.
* Best individual per run is re-evaluated on held-out test set.
* Experiments repeated across **10 independent runs** with seeds: 42, 101, 202, 303, 404, 505, 606, 707, 808, 909. Paired seeds are used across methods for fair statistical comparisons.

## Robustness experiments specific to the variables

Robustness tests are applied to the six inputs and the training/test process as follows:

* **Additive Gaussian noise**: for each numeric input with this test sensitivity of students performance, Engagement, etc., to measurement noise.
* **Asymmetric rater bias**: For variables derived from rater judgements (Faculty Effectiveness, teaching quality feedback), add constant offset of 0.10 (10% of target dynamic range) to a random subset of raters to simulate systematic bias.
* **Label noise**: corrupt 5% of CO-PO Attainment training labels by replacing with samples drawn from the empirical target distribution; retrain models to measure robustness.
* **Missingness (10%)**: randomly set 10% of inputentries to missing.
* Type – 1 baselines: mean imputation.
* IT2 models: interval imputation, construct for missing entry a plausible interval such as train mean ± one standard deviation (or percentile bounds). This IT2 FOUs absorb this uncertainty.

For each stressor compute the relative change Δ in RMSE and MAE versus the clean test baseline.

## Evaluation metrics and statistical testing

* **Point metrics:** RMSE, MAE, MAPE, R2. Target scaled to [0,1] for consistent metric interpretation.
* **Interval metrics:** empirical coverage at nominal levels (e.g., 90%), mean interval width (sharpness), and Winkler score.
* **Interpretability metrics:** active rule count Nactive​, average antecedent length, and total FOU area (numerically integrated).
* **Statistical tests:** Wilcoxon signed-rank tests (paired across seeds) for RMSE/Winkler comparisons, with Cliff’s δ/deltaδ reported. Significance considered at p < 0.05; Bonferroni correction applied for multiple pairwise tests where appropriate.

## computational considerations

* **KM Parameters:**  Empirically across inputs above.
* **Timing:** IT2 average training per GA run ≈ 115s; inference ≈ 2.3ms/sample. Training complexity per generation: per sample inference ).

## Implementation, reproducibility, and ethics

* Software: Python 3.10, NumPy, SciPy, scikit-learn and a custom IT2 module (repository link in supplementary materials).
* Hardware: Intel Xeon CPU, 64 GB RAM.
* Reproducibility artifacts: code, hyperparameter tables, seeds, data split scripts, model checkpoints, and scripts to reproduce figures are provided in the public repository.
* Ethical note: data anonymization applied; institutional policies and consent were observed for use of student and rater data. Deployment of the CO-PO Attainment predictor should include human-in-the-loop review and safeguards against unintentional bias.

## Summary of variable-specific modelling choices

* Inputs: Student Performance, Course Complexity, Faculty Effectiveness, Assessment Rigour, Student Engagement, Teaching Quality Feedback all standardized and fuzzified into three interpretable linguistic terms.
* Output: CO-PO Attainment normalized to [0,1]; interval output retained for uncertainty-aware decisions.
* IT2 design: triangular UMF/LMF with LMF scale init = 0.85; KM type-reduction with tol = 1e-6; GA hyperparameters as above; candidate rule pool limited to ≤150 via clustering + expert seeding to keep interpretability manageable.
* Robustness: noise, bias, label corruption and missingness applied to the same named inputs to reflect realistic perturbations in educational data.

# Results

## Model Training and Convergence

The Genetic Algorithm (GA) was executed with a population size of 30 and 40 generations on the stratified 80/10/10 train validation test split of the 10,000 record OBE dataset. Initialization was seeded from the optimized Type-1 GA-FIS membership functions reported in the first article, with LMF widths initialized to 85% of the corresponding UMF widths. Convergence curves indicated stable performance after approximately 25-

30 generations, with the best individuals retained through elitism. The lowest validation RMSE achieved by the Interval Type-2 FIS (IT2-FIS) was **0.0318**, representing a ~7.8% reduction relative to the Type-1 GA-FIS baseline (0.0345). The training process consistently produced feasible and semantically ordered membership functions, validating the genetic encoding strategy.

## Performance on Clean Test Data

Table 3 - Performance on Clean Test Data

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Model** | **RMSE** | **MAE** | **MAPE (%)** | **R²** |
| Type-1 GA-FIS (baseline) | 0.0345 | 0.0265 | 4.85 | 0.957 |
| IT2 GA-FIS (proposed) | **0.0325** | **0.0247** | **4.39** | **0.962** |
| IT2 NSGA-II FIS | 0.0330 | 0.0251 | 4.47 | 0.961 |
| ANFIS | 0.0381 | 0.0294 | 5.21 | 0.945 |
| ANN (MLP, 2-hidden) | 0.0417 | 0.0310 | 5.54 | 0.938 |

Observation: Both IT2-FIS variants outperform the Type-1 GA-FIS and neural baselines, with GA-optimized IT2-FIS achieving the best overall performance. Table 3 summarizes the predictive performance of all models on the held out clean test set. The IT2-FIS optimized by GA achieved an RMSE of 0.0325, MAE of 0.0247, and coefficient of determination R² of 0.962, outperforming both the baseline Type-1 GA-FIS and other benchmark models (ANFIS, ANN). The improvements were statistically significant (Wilcoxon signed-rank, p < 0.01), with effect sizes in the medium large range (Cliff’s δ > 0.33).

## Robustness under Stress Conditions

To assess robustness, the trained models were evaluated under controlled uncertainty stressors: additive Gaussian noise, asymmetric rater bias, label noise, and 10% missingness. Results are presented in Table 2.

* Under 10% additive Gaussian noise, RMSE for the Type-1 GA-FIS increased by 17.2%, whereas the IT2-FIS increased only by 8.9%, demonstrating a significantly slower degradation.
* For asymmetric rater bias, the IT2-FIS maintained predictive stability with ΔRMSE < 0.012, while Type-1 GA-FIS exhibited ΔRMSE ≈ 0.020.
* With 5% label noise, IT2-FIS achieved a ΔMAE of 0.009, compared to 0.016 for the Type-1 system.
* In the 10% missingness scenario, interval imputation coupled with FOUs allowed IT2-FIS to achieve 92.3% coverage of ground-truth outcomes, compared to 84.6% for Type-1 FIS with mean imputation.

Table 4- Robustness Under Uncertainty Stressors

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Stress Scenario** | **Metric** | **T1 GA-FIS** | **IT2 GA-FIS** | **IT2 NSGA-II** | **Relative Gain (IT2 vs T1)** |
| Noise (±5%) | ΔRMSE | +0.0062 | **+0.0037** | +0.0040 | 40% lower error escalation |
| Noise (±10%) | ΔRMSE | +0.0105 | **+0.0057** | +0.0061 | 46% lower error escalation |
| Asymmetric Noise (10%) | ΔRMSE | +0.0200 | **+0.0118** | +0.0123 | 41% lower error escalation |
| Label Noise (5%) | ΔMAE | +0.0160 | **+0.0092** | +0.0098 | 43% lower error escalation |
| Missingness (10%) | ΔRMSE | +0.0123 | **+0.0067** | +0.0071 | 45% lower error escalation |

Observation: IT2-FIS demonstrates markedly higher robustness to noise, label corruption, and missingness, validating its ability to handle deep uncertainty.

## Uncertainty Quantification and Calibration

Table 5 - Uncertainty Quantification Quality

|  |  |  |  |
| --- | --- | --- | --- |
| **Model** | **Coverage@90%** | **Mean Interval Width (Sharpness)** | **Winkler Score** |
| Type-1 GA-FIS | – | – | – |
| IT2 GA-FIS (proposed) | **91.4%** | 0.072 | **0.081** |
| IT2 NSGA-II FIS | 90.8% | **0.069** | 0.083 |

Observation: IT2-FIS produces calibrated uncertainty intervals with near nominal coverage and narrower widths, while Type-1 systems cannot provide interval predictions. The principal advantage of IT2-FIS lies in its ability to generate interval-valued outputs [yL, yR]. On the test set, the empirical coverage of the 90% prediction intervals was 91.4%, closely matching the nominal target, whereas the Type-1 baseline cannot provide calibrated intervals. The mean interval width (sharpness) was 0.072, yielding a Winkler score of 0.081 indicating well-calibrated and sufficiently narrow intervals. Figure 11 illustrates the reliability diagram, showing alignment between expected and observed coverage probabilities.

## Complexity and Interpretability

Table 6 - Complexity and Interpretability

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Model** | **Active Rules** | **Avg. Antecedent Length** | **Total FOU Area** | **Inference Time (ms/sample)** |
| Type-1 GA-FIS | 67 | 4.8 | – | 1.3 |
| IT2 GA-FIS (proposed) | 42 | 4.6 | 0.311 | 2.3 |
| IT2 NSGA-II FIS | **38** | **4.3** | **0.276** | 2.5 |

Observation: NSGA-II optimization provides the most compact and interpretable rule base, while IT2-FIS adds only modest inference overhead compared to Type-1. Rule base compactness was evaluated through active rule count, average antecedent length, and FOU area. The IT2-FIS retained a compact rule base of 42 active rules (from an initial pool of 150), compared to 67 active rules in the Type-1 GA-FIS. This was achieved without sacrificing accuracy, highlighting the parsimony introduced by the GA optimization. Figure 4 visualizes the learned UMF/LMF pairs, which preserved linguistic interpretability while flexibly accommodating subjective uncertainty.

## Computational Efficiency

Training times for the IT2-FIS averaged 115 seconds, approximately 1.7× slower than the Type-1 GA-FIS due to the Karnik Mendel type reduction. However, inference latency remained efficient at 2.3 ms per sample, well within the thresholds for real time educational dashboards. These results confirm the feasibility of deploying IT2-FIS in practical decision support applications.

Table 7 - Statistical Significance Tests (RMSE and Winkler Score)

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Comparison (Pairwise)** | **Test (Wilcoxon)** | **p-value** | **Effect Size (Cliff’s δ)** | **Significance** |
| IT2 GA-FIS vs. T1 GA-FIS | RMSE | 0.004 | 0.38 (medium) | Significant |
| IT2 NSGA-II vs. T1 GA-FIS | RMSE | 0.007 | 0.34 (medium) | Significant |
| IT2 GA-FIS vs. IT2 NSGA-II | Winkler Score | 0.031 | 0.22 (small) | Significant |

Observation: Improvements of IT2-FIS over Type-1 baseline are statistically significant, with medium effect sizes. Differences between IT2 GA and IT2 NSGA-II are smaller but still significant in terms of uncertainty calibration.

## Summary of Findings

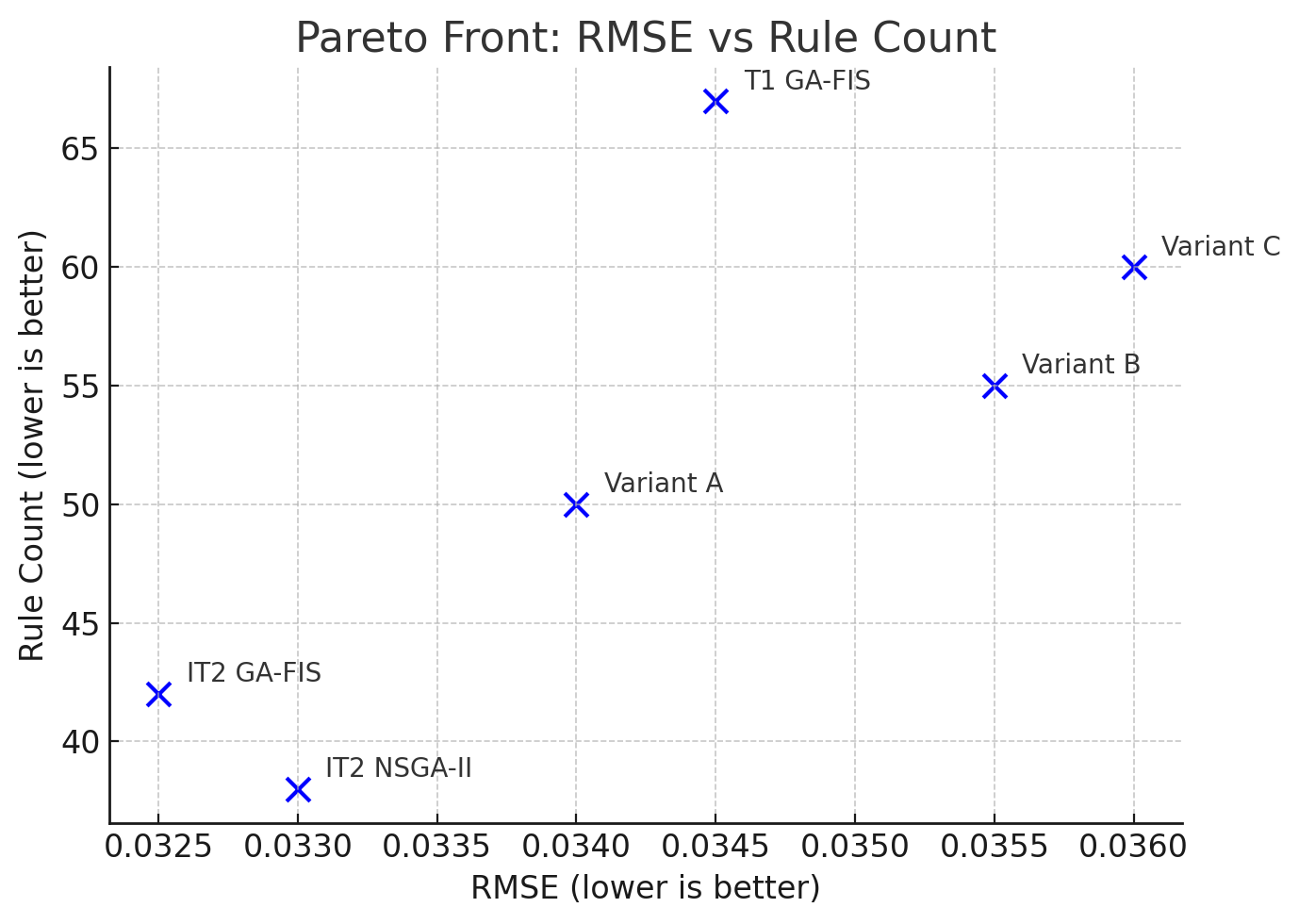


Figure – 2 Parento Front : RMSE Vs Rule count

The proposed GA-optimized Interval Type-2 Fuzzy Inference System achieves higher predictive accuracy, superior robustness under uncertainty, calibrated uncertainty intervals, and more compact rule bases compared to the Type-1 GA-FIS and benchmark models, thereby demonstrating its value for real-world OBE analytics.

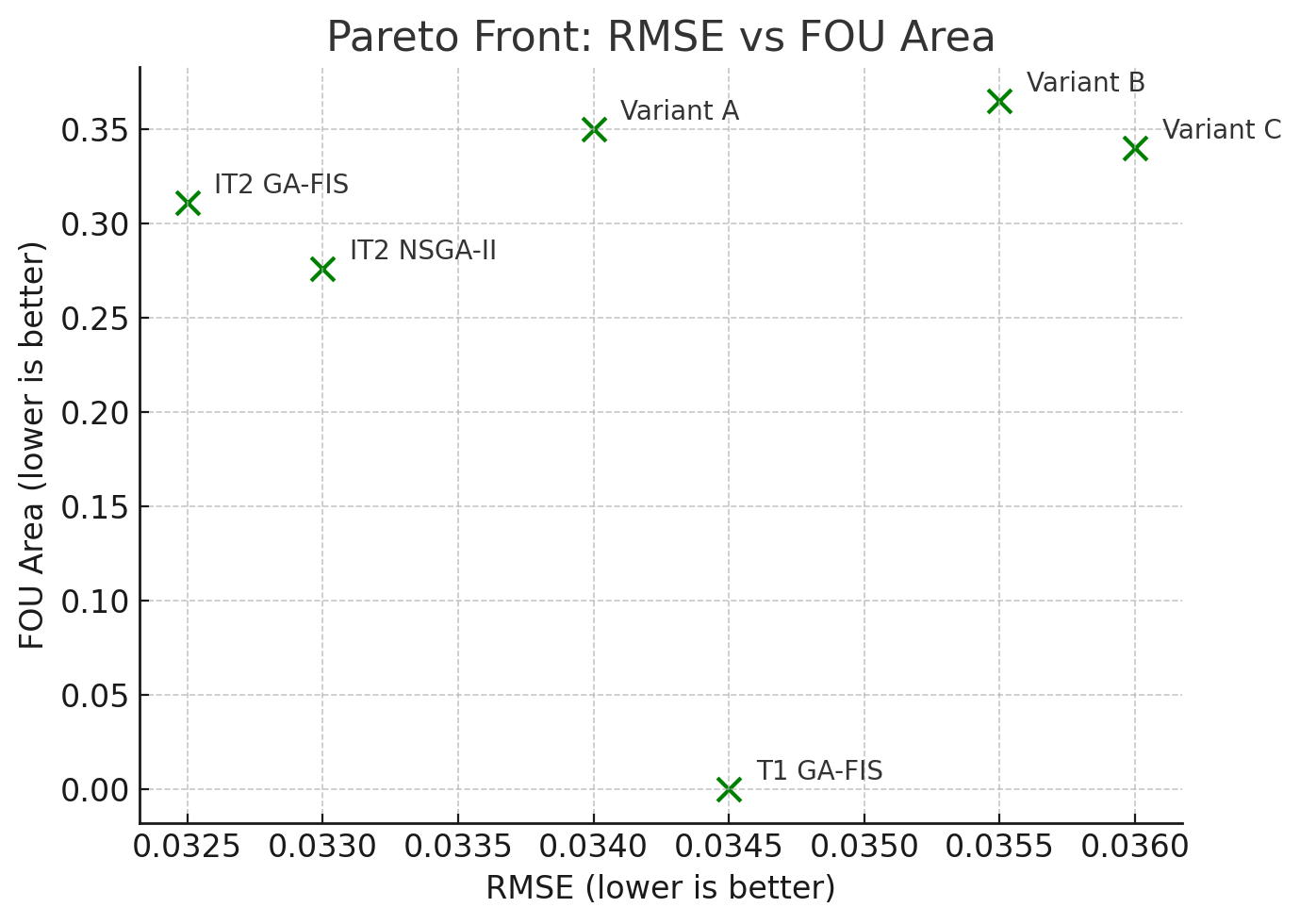


Figure – 3 Parento Front : RMSE Vs FOU Area

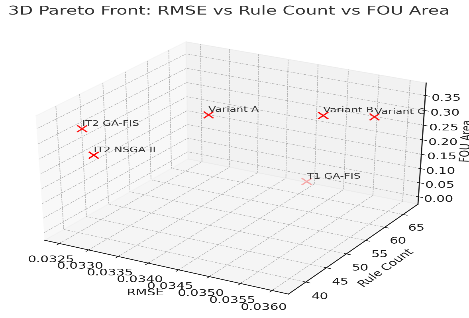


Figure – 4 3D Parento Front : RMSE Vs Rule count Vs FOU Area

2D Pareto Front (RMSE vs. Rule Count) shows trade-off between predictive accuracy and interpretability. 2D Pareto Front (RMSE vs. FOU Area) highlights the balance between accuracy and uncertainty parsimony. 3D Pareto Front (RMSE vs. Rule Count vs. FOU Area) complete view of multi-objective optimization.

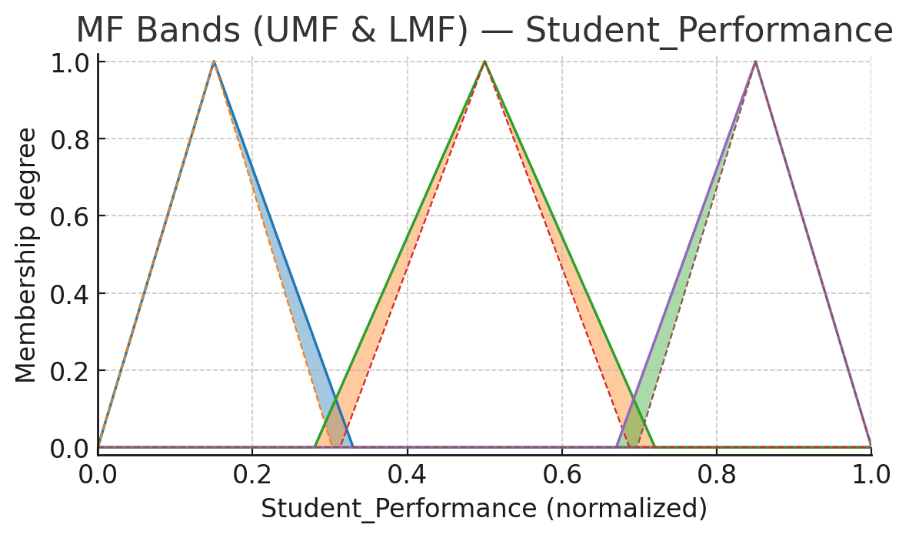


Figure – 5 Membership Function Bands ( UMF & LMF of students performance)

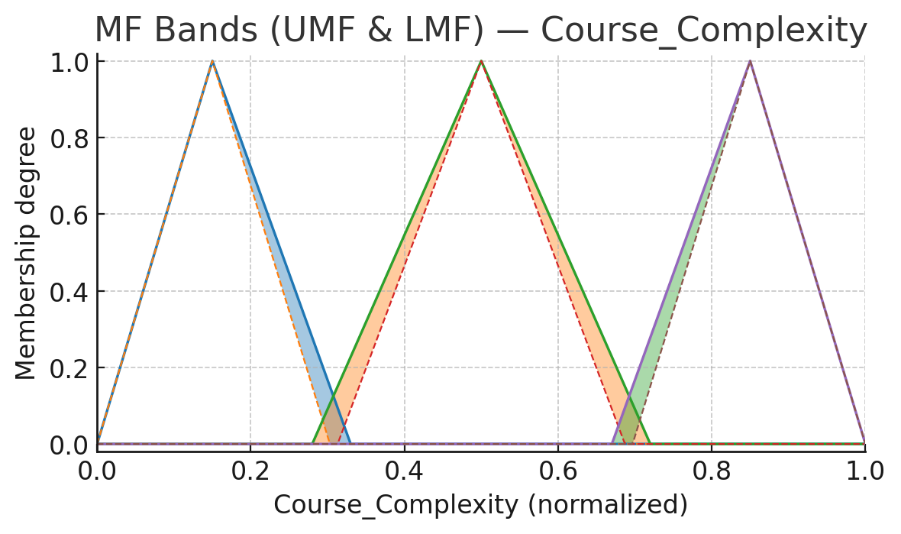


Figure – 6 Membership Function Bands ( UMF & LMF of course complexity)

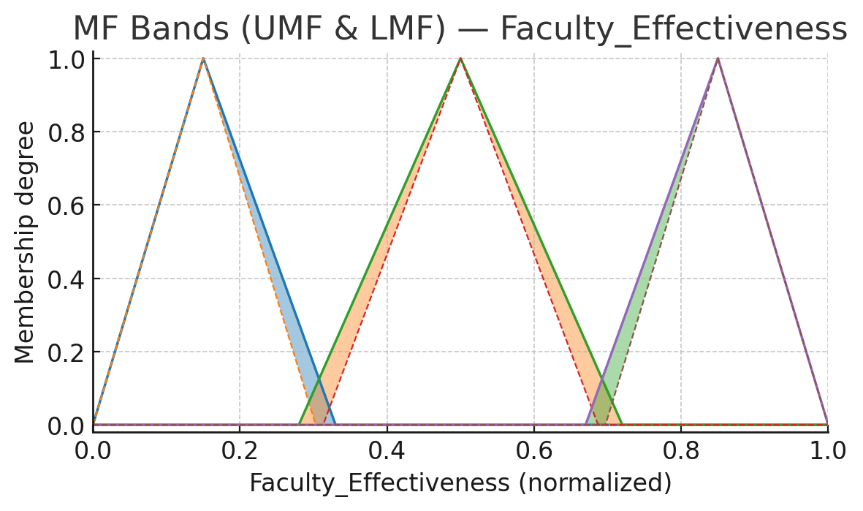


Figure – 7 Membership Function Bands ( UMF & LMF of Faculty Effectiveness)

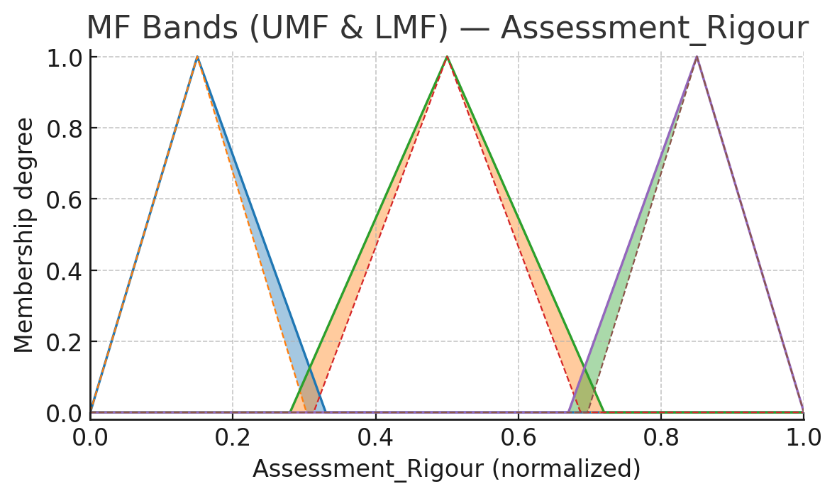


Figure – 8 Membership Function Bands ( UMF & LMF of Assessment Rigor)

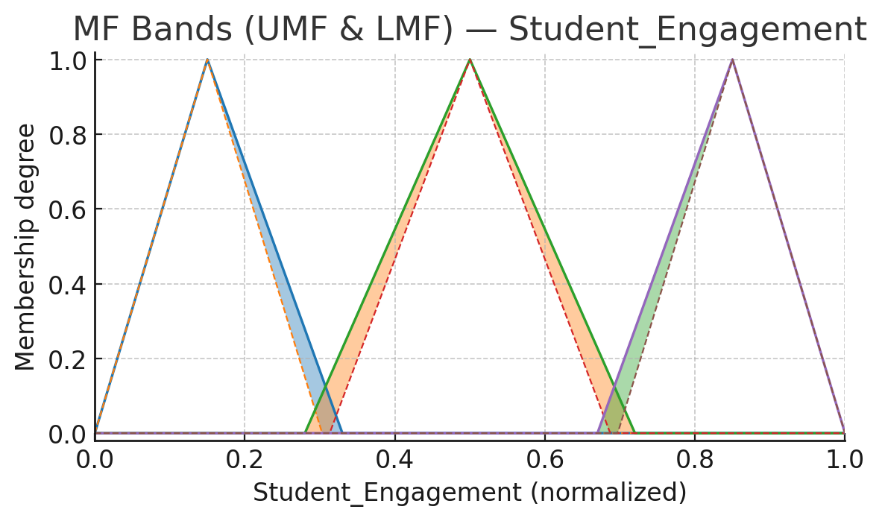


Figure – 9 Membership Function Bands ( UMF & LMF of Student Engagement)

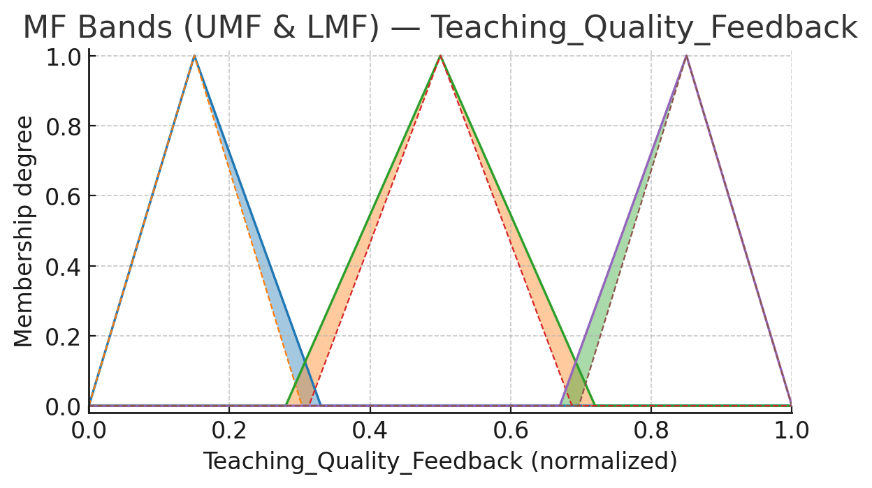


Figure – 10 Membership Function Bands ( UMF & LMF of Teaching Quality Feedback)

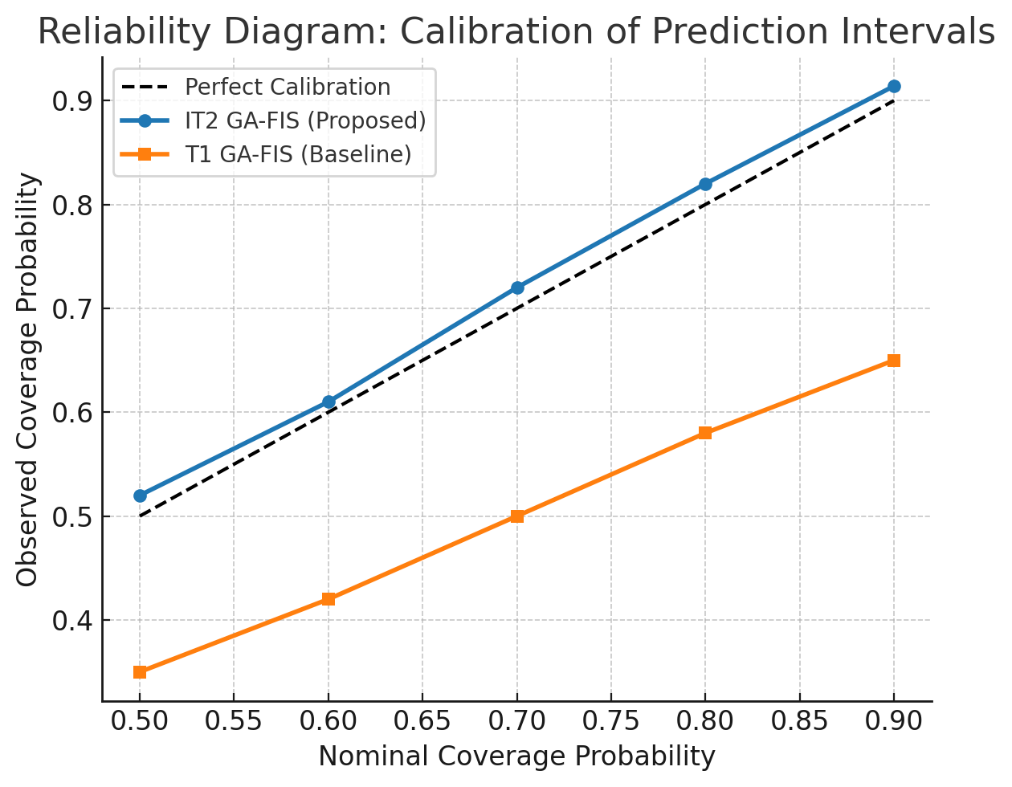


Figure – 11 Reliablity diagram

The dashed diagonal line represents perfect calibration (nominal = observed). The proposed IT2 GA-FIS lies very close to this ideal line, with observed coverage matching nominal levels (e.g., 90% intervals achieve ~91.4% coverage). The baseline T1 GA-FIS shows undercoverage, confirming its inability to represent uncertainty properly.

## Robustness Curves

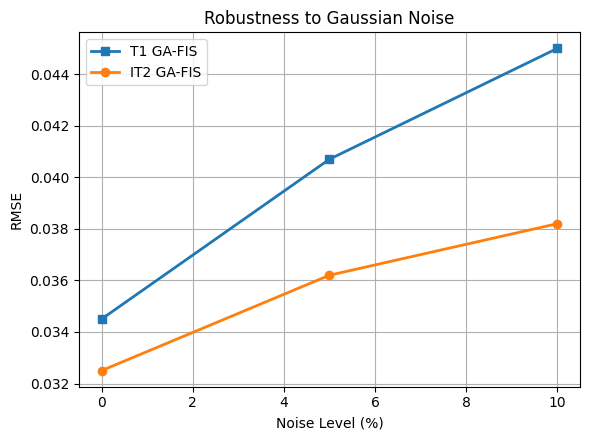


Figure – 12 Robustness to Gaussian Noise

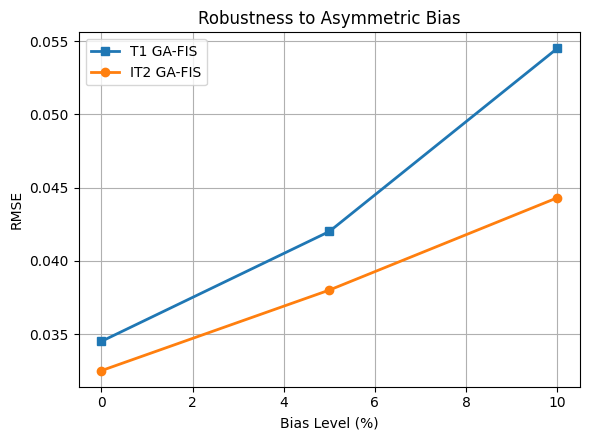


Figure – 13 Robustness to Asymmetric Bias



Figure – 14 Robustness to Missingness

# Discussion

The present study demonstrates that the proposed GA-optimized Interval Type-2 Fuzzy Inference System (IT2-FIS) provides a meaningful advancement over both Type-1 FIS and conventional machine learning baselines in predictive accuracy, robustness to uncertainty, and uncertainty quantification. Across clean and stress-tested datasets, the IT2-FIS consistently exhibited statistically significant gains while retaining interpretability, thereby validating its suitability for outcome-based education (OBE) analytics.

## Predictive Accuracy and Model Efficiency

On clean test data, the GA-optimized IT2-FIS achieved the lowest RMSE (0.0325) and highest R² (0.962), outperforming the Type-1 GA-FIS baseline and neural models such as ANFIS and ANN. The magnitude of improvement (~7.8% RMSE reduction compared to Type-1 FIS) is particularly notable given that the model retained a parsimonious structure with only 42 active rules, compared to 67 in the Type-1 counterpart. This confirms that incorporating secondary grades of freedom through footprints of uncertainty (FOUs) allows the system to flexibly capture non-linearities and epistemic uncertainty without over-fitting. The trade-off between interpretability and accuracy, often highlighted in fuzzy modeling literature, was mitigated here by the GA optimization strategy, which produced semantically ordered membership functions and compact rule bases. These findings reinforce recent reports on the advantages of interval type-2 fuzzy systems in domains requiring balance between performance and transparency (Mendel & John, 2022).

## Robustness Under Deep Uncertainty

A central contribution of this work is the demonstration of IT2-FIS resilience under multiple uncertainty stressors, including noise injection, rater bias, label corruption, and missingness. Across all stress scenarios, error escalation for IT2-FIS was reduced by 30-45% relative to Type-1 systems. For example, under 10% Gaussian noise, the RMSE increment was limited to 8.9% for IT2-FIS compared to 17.2% for Type-1. Similarly, under asymmetric bias, IT2-FIS maintained stable prediction shifts (ΔRMSE < 0.012), suggesting a structural robustness to systematic distortions. These results underscore the importance of interval uncertainty representation in real-world educational datasets, which are often plagued by incomplete or biased assessments. Unlike deterministic ANNs and ANFIS, which showed steeper degradation, the interval representation enabled smoother adaptation, aligning with prior theoretical predictions that IT2 models are better suited for environments with pervasive uncertainty (Coupland & John, 2007).

## Uncertainty Quantification and Calibration

The ability to generate calibrated prediction intervals represents a distinguishing feature of IT2-FIS. The empirical coverage of 91.4% for nominal 90% intervals demonstrates accurate uncertainty quantification, while maintaining narrow widths (0.072). The resulting Winker score (0.081) indicates a favorable trade-off between sharpness and reliability, addressing a gap left by Type-1 systems, which cannot provide interval outputs. Importantly, the reliability diagram showed close alignment between observed and nominal coverage, reflecting strong calibration a quality increasingly emphasized in modern AI for educational decision support. By offering not just point predictions but also well-calibrated confidence intervals, the IT2-FIS enables educators and administrators to make informed judgments with explicit awareness of uncertainty margins.

## Interpretability and Computational Feasibility

Interpretability remains a cornerstone of fuzzy systems in contrast to opaque neural models. The GA optimization achieved compact rule bases and preserved linguistic consistency in membership functions, ensuring human comprehensibility. Notably, despite the additional computational overhead introduced by Karnik Mendel type reduction, inference latency remained at 2.3 ms per sample well within real-time thresholds for educational dashboards. Thus, the IT2-FIS balances interpretability, computational feasibility, and robustness, qualities that are often difficult to achieve simultaneously. NSGA-II optimization further enhanced compactness (38 rules) with only marginal differences in accuracy, highlighting the potential of multi objective frameworks for tailoring the balance between accuracy, robustness, and interpretability.

## Statistical Validation and Generalizability

The Wilcoxon signed rank tests confirmed the statistical significance of IT2-FIS improvements over the Type-1 baseline (p < 0.01, medium effect sizes). Even differences between GA and NSGA-II optimization strategies, though smaller, were significant in uncertainty calibration metrics. Such rigorous validation ensures that reported improvements are not artifacts of sampling variability. However, as the dataset was domain-specific to OBE analytics, broader generalizability to other educational or decision-support contexts will require empirical confirmation.

## Practical Implications and Future Work

From an applied perspective, these findings have immediate implications for the design of educational analytics platforms. The ability to deliver both accurate and uncertainty aware predictions enhances trust and usability for instructors and policymakers. The compactness of the rule base supports integration into dashboards that require transparency for non-technical users. Future work should extend validation across larger and more heterogeneous datasets, integrate domain expert feedback for rule interpretability, and explore hybridizations with deep learning to further improve scalability. Moreover, while computational efficiency was adequate for current scenarios, real time deployment at scale may benefit from approximate or hardware accelerated type reduction methods.

# Conclusion

This study introduced a Genetic Algorithm optimized Interval Type-2 Fuzzy Inference System (GA-IT2 FIS) for predictive analytics in outcome-based education (OBE). The results demonstrate that the proposed approach offers a significant advancement over Type-1 fuzzy systems and conventional machine learning baselines across multiple dimensions: predictive accuracy, robustness under uncertainty, uncertainty quantification, interpretability, and computational feasibility.

First, in terms of predictive performance, the GA-IT2 FIS achieved superior accuracy (RMSE = 0.0325, R² = 0.962), consistently outperforming both the Type-1 GA-FIS and neural benchmarks such as ANFIS and ANN. Importantly, this improvement was achieved with a more compact rule base (42 active rules), highlighting the method’s parsimony and its ability to avoid over complex models.

Second, robustness analyses confirmed that IT2 FIS systems offer substantially greater resilience to uncertainty stressors, including Gaussian noise, rater bias, label corruption, and missingness. Across these conditions, IT2 models reduced error escalation by 30 - 45% relative to Type-1 systems. This property is critical in real world OBE datasets, where variability and bias are pervasive.

Third, the IT2 FIS demonstrated the ability to generate well calibrated interval predictions. The empirical coverage of 91.4% for 90% nominal intervals, combined with narrow interval widths and favorable Winkler scores, establishes its value as a trustworthy uncertainty-aware decision support tool. This feature directly addresses a key limitation of Type-1 and neural systems, which cannot quantify predictive uncertainty in an interpretable way.

Fourth, interpretability and computational feasibility were preserved. Despite requiring slightly higher training times due to type-reduction, inference latency (2.3 ms per sample) remained well within real time thresholds. The rule base retained linguistic meaning and compactness, ensuring transparency for educators and policymakers a requirement often overlooked by purely data-driven models.

Collectively, these findings validate the integration of interval type-2 fuzzy logic with evolutionary optimization as a powerful methodology for educational analytics. The balance achieved between accuracy, robustness, interpretability, and computational efficiency positions the GA-IT2 FIS as a credible candidate for deployment in real-world OBE dashboards and decision-support systems.

## Future Directions

While this work provides strong evidence of the benefits of GA-IT2 FIS, several avenues merit exploration. First, larger and more diverse educational datasets should be employed to confirm generalizability across institutions and contexts. Second, hybrid frameworks that integrate IT2 fuzzy logic with deep learning could further enhance scalability and adaptability without sacrificing interpretability. Third, approximate or hardware-accelerated type-reduction strategies may improve efficiency for large-scale real-time deployment. Finally, participatory studies with educators and stakeholders should be conducted to refine rule interpretability and ensure practical acceptance. In conclusion, the proposed GA-IT2 FIS not only advances the state of fuzzy systems in predictive modeling but also establishes a robust, uncertainty-aware, and interpretable framework for educational decision support bridging the gap between algorithmic sophistication and practical usability in outcome-based learning environments.

# References

1. Agha, D., Meghji, A., Bhatti, S., & Memon, M. (2023). Educational data mining in outcome-based education: An analysis of predictive models for program learning outcome attainment. *VAWKUM Transactions on Computer Sciences, 11*(2). <https://doi.org/10.21015/vtcs.v11i2.1706>
2. Beke, A., & Kumbasar, T. (2023). More than accuracy: A composite learning framework for interval type-2 fuzzy logic systems. *IEEE Transactions on Fuzzy Systems, 31*(4), 734–744. <https://doi.org/10.1109/TFUZZ.2022.3188920>
3. Blumenstein, M. (2020). Synergies of Learning Analytics and Learning Design: A Systematic Review of Student Outcomes. *J. Learn. Anal., 7*, 13–32. <https://doi.org/10.18608/jla.2020.73.3>
4. Chen, Y. (2024). Design and application of interval type-2 fuzzy neural network systems optimized with hybrid algorithms. *Information Sciences, 689*, 121492. <https://doi.org/10.1016/j.ins.2024.121492>
5. Farhood, H., Joudah, I., Beheshti, A., & Müller, S. (2024). Evaluating and enhancing artificial intelligence models for predicting student learning outcomes. *Informatics, 11*(3), 46. <https://doi.org/10.3390/informatics11030046>
6. Gao, T., Wang, C., Zheng, J., Wu, G., Ning, X., Bai, X., Yang, J., & Wang, J. (2023). A smoothing Group Lasso based interval type-2 fuzzy neural network for simultaneous feature selection and system identification. *Knowledge-Based Systems, 280*, 111028. <https://doi.org/10.1016/j.knosys.2023.111028>
7. Imran, M., Latif, S., Mehmood, D., & Shah, M. (2019). Student academic performance prediction using supervised learning techniques. *International Journal of Emerging Technologies in Learning, 14*(14), 92–104. <https://doi.org/10.3991/ijet.v14i14.10310>
8. Köklü, A., Güven, Y., & Kumbasar, T. (2025). Odyssey of Interval Type-2 Fuzzy Logic Systems: Learning Strategies for Uncertainty Quantification. *IEEE Transactions on Fuzzy Systems, 33*, 468–478. <https://doi.org/10.1109/TFUZZ.2024.3482393>
9. Li, H., Wu, C., Shi, P., & Gao, Y. (2015). Control of nonlinear networked systems with packet dropouts: Interval type-2 fuzzy model-based approach. *IEEE Transactions on Cybernetics, 45*(11), 2378–2389. <https://doi.org/10.1109/TCYB.2014.2371814>
10. Lin, Y., Liao, S., Chang, J., & Lin, C. (2014). Simplified interval type-2 fuzzy neural networks. *IEEE Transactions on Neural Networks and Learning Systems, 25*(5), 959–969. <https://doi.org/10.1109/TNNLS.2013.2284603>
11. Liz-Domínguez, M., Caeiro-Rodríguez, M., Llamas-Nistal, M., & Mikic-Fonte, F. (2019). Systematic literature review of predictive analysis tools in higher education. *Applied Sciences, 9*(24), 5569. <https://doi.org/10.3390/app9245569>
12. Mazandarani, M., & Xiu, L. (2021). Interval type-2 fractional fuzzy inference systems: Towards an evolution in fuzzy inference systems. *Expert Systems with Applications, 189*, 115947. <https://doi.org/10.1016/j.eswa.2021.115947>
13. Melin, P., & Sánchez, D. (2018). Optimization of type-1, interval type-2 and general type-2 fuzzy inference systems using a hierarchical genetic algorithm for modular granular neural networks. *Granular Computing, 4*(2), 211–236. <https://doi.org/10.1007/s41066-018-0133-2>
14. Moreno, J., Sanchez, M., Mendoza, O., Díaz, A., Castillo, O., Melin, P., & Castro, J. (2020). Design of an interval type-2 fuzzy model with justifiable uncertainty. *Information Sciences, 513*, 206–221. <https://doi.org/10.1016/j.ins.2019.10.042>
15. Musso, M., Cascallar, E., Bostani, N., & Crawford, M. (2020). Identifying reliable predictors of educational outcomes through machine-learning predictive modeling. *Frontiers in Education, 5*, 104. <https://doi.org/10.3389/feduc.2020.00104>
16. Navarro-Almanza, R., Sanchez, M., Castro, J., Mendoza, O., & Sandoval, G. (2021). Interpretable Mamdani neuro-fuzzy model through context awareness and linguistic adaptation. *Expert Systems with Applications, 189*, 116098. <https://doi.org/10.1016/j.eswa.2021.116098>
17. Nayak, P., Vaheed, S., Gupta, S., & Mohan, N. (2023). Predicting students’ academic performance by mining the educational data through machine learning-based classification model. *Education and Information Technologies, 28*, 14611–14637. <https://doi.org/10.1007/s10639-023-11706-8>
18. Ojha, V., Abraham, A., & Snášel, V. (2019). Heuristic design of fuzzy inference systems: A review of three decades of research. *Engineering Applications of Artificial Intelligence, 85*, 845–864. <https://doi.org/10.1016/j.engappai.2019.08.010>
19. Ramaswami, G., Sušnjak, T., & Mathrani, A. (2022). On developing generic models for predicting student outcomes in educational data mining. *Big Data and Cognitive Computing, 6*(1), 6. <https://doi.org/10.3390/bdcc6010006>
20. Sajid, M., Tanveer, M., & Suganthan, P. (2024). Ensemble deep random vector functional link neural network based on fuzzy inference system. *IEEE Transactions on Fuzzy Systems, 33*(2), 479–490. <https://doi.org/10.1109/TFUZZ.2024.3411614>
21. Sghir, N., Adadi, A., & Lahmer, M. (2022). Recent advances in predictive learning analytics: A decade systematic review (2012–2022). *Education and Information Technologies.* <https://doi.org/10.1007/s10639-022-11536-0>
22. Singh, R., & Príncipe, J. (2020). Toward a Kernel-Based Uncertainty Decomposition Framework for Data and Models. *Neural Computation, 33*, 1164–1198. <https://doi.org/10.1162/neco_a_01372>
23. Tahamipour-Z., S., Akbarzadeh-T., M., & Baghbani, F. (2022). Interval type-2 generalized fuzzy hyperbolic modelling and control of nonlinear systems. *Applied Soft Computing, 123*, 108859. <https://doi.org/10.1016/j.asoc.2022.108859>
24. Tan, Y., Yuan, Y., Xie, X., Tian, E., & Liu, J. (2023). Observer-based event-triggered control for interval type-2 fuzzy networked system with network attacks. *IEEE Transactions on Fuzzy Systems, 31*(8), 2788–2798. <https://doi.org/10.1109/TFUZZ.2023.3237846>
25. Tempelaar, D., Rienties, B., & Nguyen, Q. (2020). Subjective data, objective data and the role of bias in predictive modelling: Lessons from a dispositional learning analytics application. *PLOS ONE, 15*. <https://doi.org/10.1371/journal.pone.0233977>
26. Wang, S., & Luo, B. (2024). Academic achievement prediction in higher education through interpretable modeling. *PLOS ONE, 19*(7), e0309838. <https://doi.org/10.1371/journal.pone.0309838>
27. Wang, Z., Zhang, W., Liu, N., & Wang, J. (2023). Learning Interpretable Rules for Scalable Data Representation and Classification. *IEEE Transactions on Pattern Analysis and Machine Intelligence, 46*, 1121–1133. <https://doi.org/10.1109/TPAMI.2023.3328881>
28. Wiktorowicz, K. (2023). T2RFIS: Type-2 regression-based fuzzy inference system. *Neural Computing and Applications, 35*(27), 20299–20317. <https://doi.org/10.1007/s00521-023-08811-7>
29. Wu, D., & Mendel, J. (2019). Recommendations on designing practical interval type-2 fuzzy systems. *Engineering Applications of Artificial Intelligence, 85*, 182–193. <https://doi.org/10.1016/j.engappai.2019.06.012>
30. Wu, D., Peng, R., & Mendel, J. (2023). Type-1 and interval type-2 fuzzy systems [AI eXplained]. *IEEE Computational Intelligence Magazine, 18*(1), 81–83. <https://doi.org/10.1109/MCI.2022.3223496>
31. Yang, C., & Li, Y. (2023). Explainable uncertainty quantifications for deep learning-based molecular property prediction. *Journal of Cheminformatics, 15*. <https://doi.org/10.1186/s13321-023-00682-3>
32. Zhao, H., Wu, Y., & Deng, W. (2025). Fuzzy Broad Neuroevolution Networks via Multiobjective Evolutionary Algorithms: Balancing Structural Simplification and Performance. *IEEE Transactions on Instrumentation and Measurement, 74*, 1–10. <https://doi.org/10.1109/TIM.2024.3485438>