Fuzzy-Monotonic LightGBM for Explainable Credit Default Prediction

A hybrid approach combining ML performance with regulatory compliance.

The Challenge

Addressing the complexities of credit default prediction in a highly regulated environment:

Regulatory Demands

Credit default prediction is high-stakes with Basel II/III and SR-11-7 regulatory requirements demanding explainability.

Black Box Models

Pure ML models like LightGBM perform well but are black boxes that banks reject in deployment.
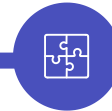
Weak Traditional Models

Traditional scorecards are explainable but weak in performance.

Explainability Gaps

Post-hoc SHAP explainability doesn't guarantee directionally correct behavior.

The Hybrid Need

Need for a hybrid solution that is regulator-aligned, economically interpretable, AND competitive in predictive strength.

Our Proposed Solution

Introducing the Fuzzy-Monotonic LightGBM framework, a hybrid approach designed to combine superior predictive performance with essential regulatory explainability:

Behavioral Features

Unified Solution

Fuzzy Rule Reasoning

Monotonic LightGBM

This framework integrates three key components into a single, cohesive model:

- **Engineered credit behavioral features:** Leveraging advanced data processing for comprehensive borrower profiles.
- **Fuzzy linguistic rule reasoning:** Incorporating human-understandable rules for behavior membership (e.g., Low/Medium/High delinquency).
- **Monotonic constraints inside LightGBM:** Ensuring economically interpretable predictions where, for example, higher delinquency ALWAYS leads to higher risk.

For validation, we utilized two distinct datasets:

- **Taiwan dataset (30,000 records):** Primarily used for demonstrating predictive performance and model robustness.
- **German dataset (1,000 records):** Employed to showcase interpretability and explainability aspects of the model.

Made with GAMMA

Implementation Steps

01

Feature engineering: utilization ratio, bill average, repayment trend, delinquency intensity etc.

02

Fuzzy membership functions created on numeric variables (Low/Medium/High semantics)

03

Fuzzy rule activations appended as model features

04

Monotonic constraints applied inside LightGBM so direction is economically consistent

05

Ablation validation proved hybrid contributes meaningful improvement

Performance Results

Our ablation study demonstrates that explainability does NOT sacrifice performance. The proposed Fuzzy-Monotonic LightGBM model not only meets regulatory requirements but also achieves superior predictive accuracy.
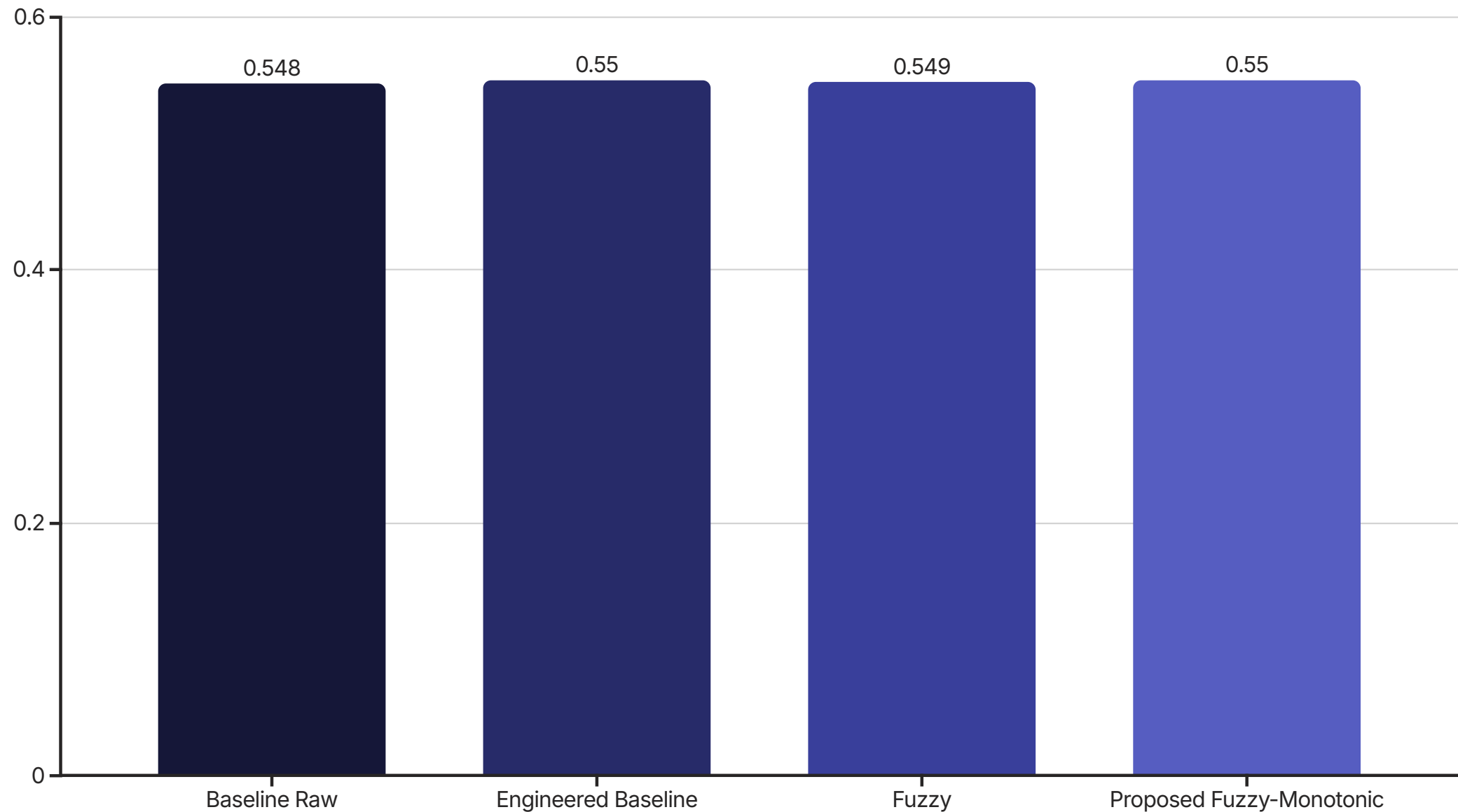
The "Proposed Fuzzy-Monotonic" model achieves PR-AUC score of **0.55**, representing the best-calibrated and best-ranking performance among all tested variants.

Future Scope

Add Bureau Graph + Transaction Embeddings

Enhancing data inputs with sophisticated bureau graph analysis and transaction embeddings for richer feature representation.

Macro Drift Adaptive Recalibration

Implementing a dynamic recalibration mechanism to adapt to macroeconomic shifts and maintain model accuracy over time.

Causal Monotonic Structures Instead of Statistical Monotonic

Moving towards causally-driven monotonic constraints to ensure more robust and economically sound interpretability.

Production Competitive NBFC Scoring Engine Deployable Online

Developing a robust and competitive scoring engine specifically for Non-Banking Financial Companies, ready for online deployment.

Key Takeaways

Our Fuzzy-Monotonic LightGBM framework delivers a breakthrough solution for credit default prediction, effectively addressing critical industry challenges:

Closing the Industry Gap

Our hybrid method closes the long-standing industry gap between accuracy (ML) and interpretability (fuzzy rules + monotonic constraints).

Regulator-Aligned & Superior Performance

This approach is regulator-aligned, deployable, economically interpretable and outperforms the raw baseline in PR-AUC + calibration quality.

Co-existence of Performance & Explainability

Explainability and performance can co-exist.