

Retail Banking Competition Summary

(Team: synergy-minds)

1) Objective & Scope

I competed in the Retail Banking domain, which simulates real-world analytics for **Fraud Detection (Challenge 1)** and **Credit Default Prediction (Challenge 2)**. The work integrated multi-table customer data, time-series behavior, and device/session signals to balance performance, interpretability, and reproducibility.

2) Data & Integration Strategy

Core sources:

- transactions_* (per-transaction fraud labels in train),
- customer_panel_* (weekly customer risk dynamics),
- customers_all, accounts_all, devices_all, device_sessions_all (contextual joins by CustomerID and SessionID).

Linkage:

- Fraud task: join transactions to customer/device context; derive session, location, velocity, and merchant-pattern features.
- Default task: build weekly panel features (lags, trends, utilization bands) per CustomerID, Week.

3) Modeling Approach (High-level)

- **Preprocessing:** explicit datetime parsing; categorical encoding for Channel, MCC_Group, cities/devices; robust scaling for skewed amounts; stratified CV.
- **Feature engineering:**
 - **Fraud:** recency counts (N minutes/hours), merchant diversity, amount z-scores by customer, device reuse across customers, geo/session mismatch rates, time-of-day/week patterns.
 - **Default:** rolling means/volatility of Utilisation and PaymentRatio (1–4 week windows), spikes, late-payment flags, cumulative hard inquiries; interaction of tenure \times credit score.
- **Models:** tree ensembles (XGBoost/RandomForest) with class weighting and threshold tuning for **Macro-F1**.
- **Evaluation:** Macro-F1 on validation and final submissions, aligned to competition metric.

4) Results (public leaderboard / best runs)

- **Challenge 1 (Fraud):** Macro-F1 \approx **0.5211** (best recorded run).
- **Challenge 2 (Default):** Macro-F1 \approx **0.9553** (best recorded run).
(Values reflect my best submitted/validated scores and guide the ablations below.)

5) What Worked

- Behavior-aware aggregations (short/medium windows) > raw snapshots.
- MCC & channel-conditional features (e.g., unusual amount given user–MCC history).
- Calibrated decision threshold optimized for Macro-F1, not default 0.5.
- Simple, well-regularized models with careful feature checks beat overly complex stacks.

6) What Didn't

- Overly wide time windows diluted fraud signals.
- High-cardinality IDs without smoothing increased variance.
- Plain frequency encodings underfit compared to rate/ratio features.

7) Human–AI Collaboration

I used an LLM assistant to: generate feature ideas, draft starter code, surface common datetime/encoding pitfalls, and explain trade-offs (e.g., class weights vs. focal loss). Human judgment set the modeling goals, validated features, enforced reproducibility, and performed error analysis; the AI accelerated coding, refactoring, and quick ablation scripts.

8) Reproducibility & Artifacts

- **Notebooks:**
 - synergy-minds_retail-banking_challenge1.ipynb
 - synergy-minds_retail-banking_challenge2.ipynb
- **Environment:** Python 3.11; key libs: pandas, numpy, scikit-learn, xgboost/lightgbm (pin exact versions in notebooks).
- **Determinism:** fixed seeds; saved feature lists; logged thresholds.
(Per prize rules: single executable .ipynb per challenge with package versions. Naming follows the required convention.)

9) Ethics & Risk Controls

- Fraud: prioritize precision/recall balance to limit false positives that impact customers.
- Credit risk: check for proxy bias (age, geography) and monitor drift; report model rationale and stability.