AIESEC

30.05.26

MASTERCARD DATA QUEST

Team: Cashless Avengers

Izatov Imran, Khan Nataliya,
Kubysheva Fakhri, Abdulaziz Amal’

Executive Summary / Issue Diagnostic

HIDDEN COMMERCIAL ACTIVITY DETECTION

ML-based detection of undeclared business usage on consumer cards

Situation

- Kazakhstan's payments market is experiencing **rapid growth** in self-employed entrepreneurship. A significant share of micro-businesses operate **through consumer debit cards to avoid higher interchange fees** and business-account onboarding friction.

Issue

- This creates **interchange leakage** for banks and payment networks, distorts customer segmentation, and weakens risk visibility. Traditional **rule-based monitoring fails** because hidden commercial activity **increasingly mimics** normal consumer behavior.

Strategy

- We *developed an ML detection model* that **identifies hidden commercial activity** through transaction behavior patterns. *Econometric validation confirmed statistically significant differences between business and consumer cards, while a calibrated logistic regression achieved ROC-AUC = 0.98.* The strongest predictor was operational-time activity: transactions during business hours increased the probability of commercial usage **by nearly 90 percentage points.**

Impact

- The model delivers production-grade detection quality with 89% precision and 93% recall, while threshold optimization improves sensitivity to 97%. Economic modeling shows that a single hidden business card may generate approximately 75,000 KZT annual interchange leakage. At scale, 10,000 undetected cards could lead to over 751 million KZT yearly revenue loss.

External (PESTLE) Analysis

P Increasing regulatory attention toward transparency of digital payments and tax compliance among self-employed merchants.

E Rapid expansion of the gig economy and informal entrepreneurship increases interchange leakage and reduces banking monetization efficiency.

S Consumers prefer frictionless onboarding and continue using personal cards for commercial activity due to convenience and low barriers.

T Growth of digital payments and transaction data availability enables scalable ML-driven behavioral detection systems.

L Potential tightening of AML/KYC and tax-reporting requirements may force banks to improve commercial activity classification.

E Digital transaction monitoring reduces dependence on manual compliance operations and supports scalable low-paper financial infrastructure.

Issue diagnostics

SWOT

Strengths

- *High ML accuracy (ROC-AUC 0.98)*
- *Non-invasive approach (consultation instead of blocking)*
- *Preserves customer loyalty and retention*
- *Reduces interchange leakage and improves transparency*
- *Econometrically validated and statistically robust model*
- *Scalable across banks and CIS markets*

Weaknesses

- *Possibility of false positives*
- *Requires large volumes of historical transaction data*
- *Model retraining is required over time*
- *Initial integration may require significant bank resources*
- *Dependent on data quality and completeness*
- *ML decision-making may be difficult to explain to some clients*

Opportunities

- *Growth of SMEs and digital entrepreneurship*
- *Integration with SME support and tax formalization programs*
- *Cross-selling business banking products and services*
- *Improved AML and risk monitoring systems*
- *Expansion to CIS markets and fintech partnerships*
- *Contribution to economic transparency and financial inclusion*

Threats

- *Hidden businesses may adapt transaction behavior*
- *Increasing privacy and data regulation requirements*
- *Competitors may develop similar solutions*
- *Misclassification may create reputational risks*
- *Some users may switch to cash or alternative banks*
- *Economic instability may reduce SME formalization rates*

Problems and analysis

Financial and Economic Analysis

1. **Shadow Business Activity Problem**
A significant number of small entrepreneurs use consumer cards for commercial activity in order to avoid business interchange fees, taxes, and formal business registration. This creates financial leakage for banks and reduces transaction transparency within the economy.

2. **Limitations of Traditional Bank Approaches**
Most banks either fail to detect hidden business activity or respond through immediate account restrictions and blocking. Such approaches increase customer churn, reduce loyalty, and negatively affect long-term customer lifetime value (LTV).

3. **Behavioral Differences Between Consumer and Business Cards**
Our analysis of 105,000 cards showed that business-related transaction behavior differs significantly from ordinary consumer behavior. Business-like cards demonstrate:

- higher activity during working hours,
- larger and more stable turnover,
- more operational and predictable spending patterns.

4. **Data-Driven Validation and ML Detection**
Statistical testing confirmed strong differences between card types (Mann–Whitney U test, $p < 0.001$).
The ML model achieved:
• ROC-AUC = 0.98, • Precision = 89%, • Recall = 93%.
The strongest predictor was transaction activity during business hours, increasing the probability of hidden business activity by nearly 90 percentage points.

5. **Economic Impact and Business Opportunity**
Hidden business activity leads to substantial interchange leakage for banks.
Estimated annual losses:

- 75M KZT for 1,000 hidden business cards,
- 375M KZT for 5,000 cards,
- 751M KZT for 10,000 cards.

Instead of punitive blocking, our solution converts hidden entrepreneurs into legal business clients through consultation and simplified formalization pathways, preserving both revenue and customer relationships.

AIESEC

Behavioral Differences Between Consumer and Business Cards

Key Results:

- Mann-Whitney U test shows statistically significant differences across all key behavioral variables
- *p-value < 0.001 for all major indicators*

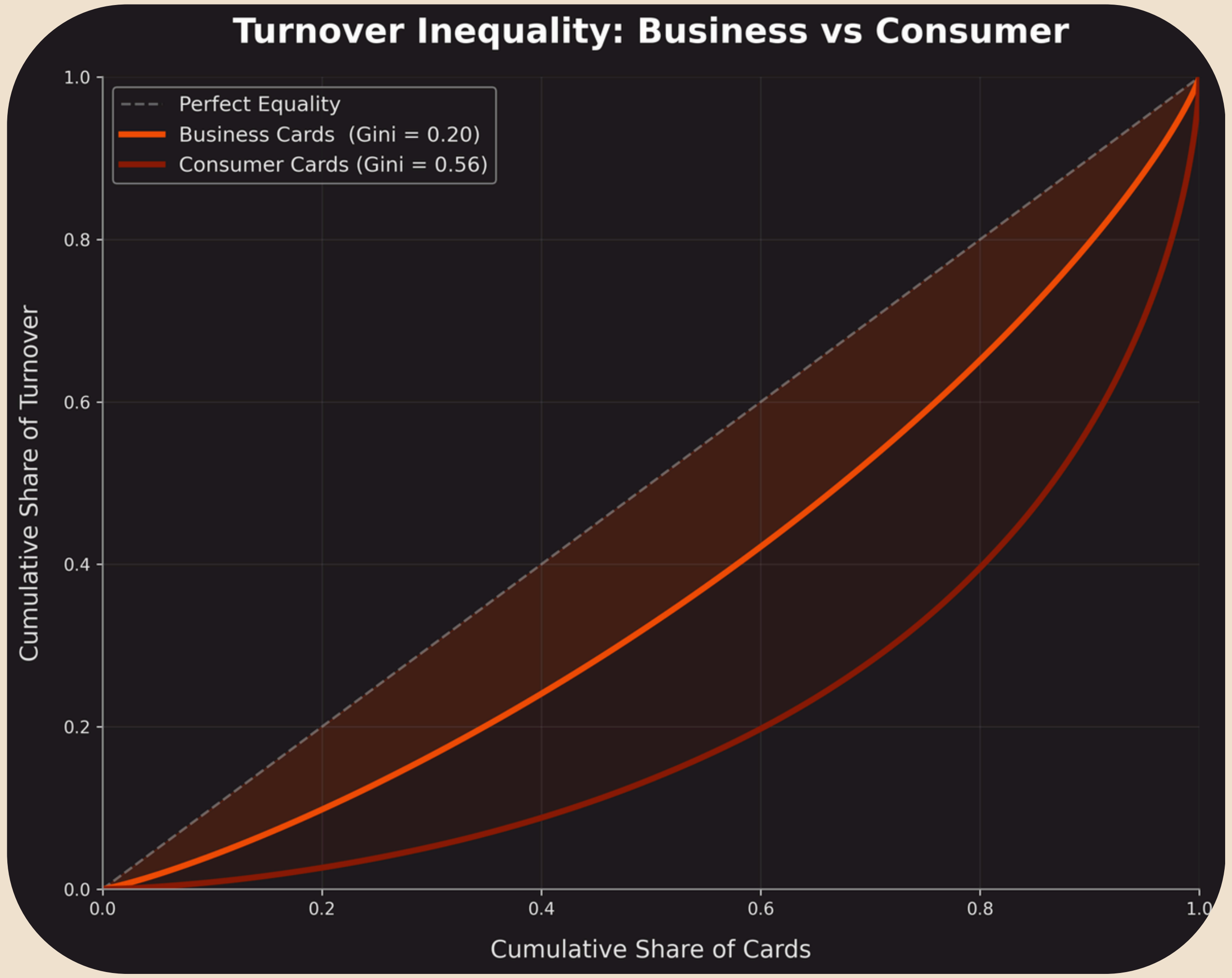
Interpretation:

- Transaction behavior of hidden business users differs significantly from regular consumer behavior. The differences are statistically proven rather than visually assumed.

Business Insight:

- This confirms that hidden commercial activity creates distinct transactional patterns that can be detected analytically.

All key behavioral variables show statistically significant differences between business and consumer cards (Mann-Whitney U test, p < 0.001).



HIDDEN BUSINESSES TRANSACT DIFFERENTLY FROM REGULAR CONSUMERS:

- higher turnover
- larger average checks
- stronger activity during business hours
- more operationally consistent spending behavior

This pattern resembles B2B operational activity rather than household consumption

Hidden Business Detection Model

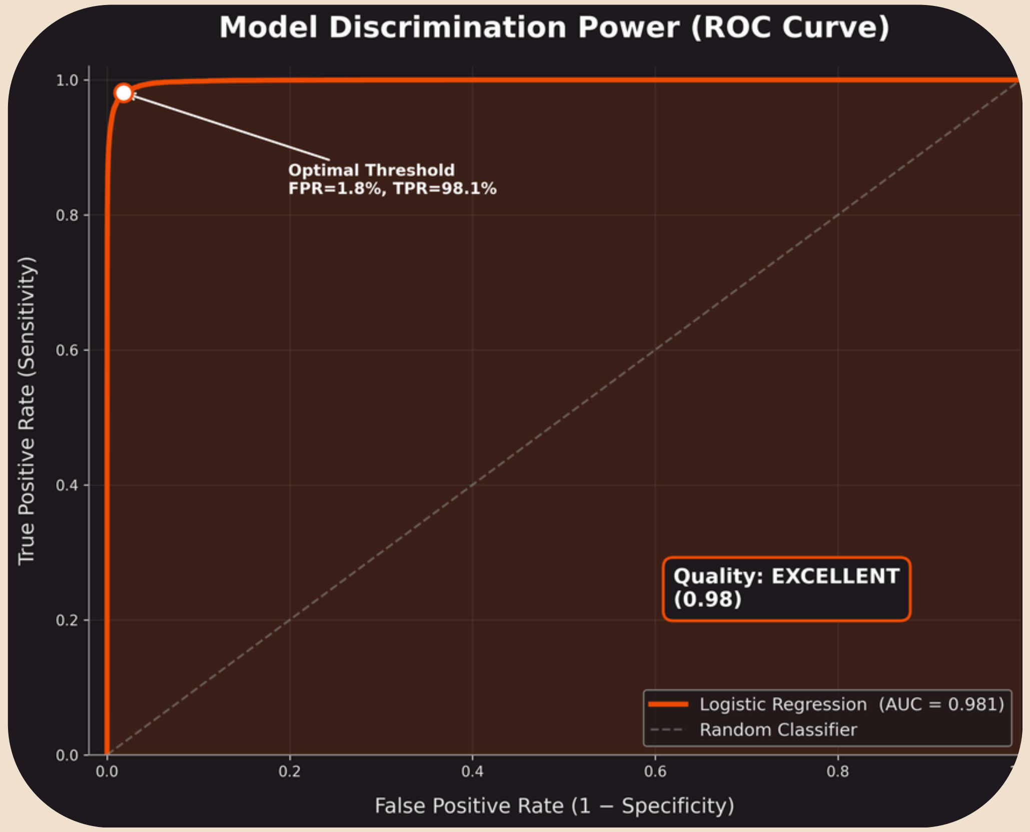
Logistic regression successfully separates hidden business and consumer cards:

- **ROC-AUC = 0.9807**
- **Precision = 89%**
- **Recall = 93%**

1. *demonstrates **extremely strong** predictive performance and accurately **identifies** suspicious business-like behavior.*

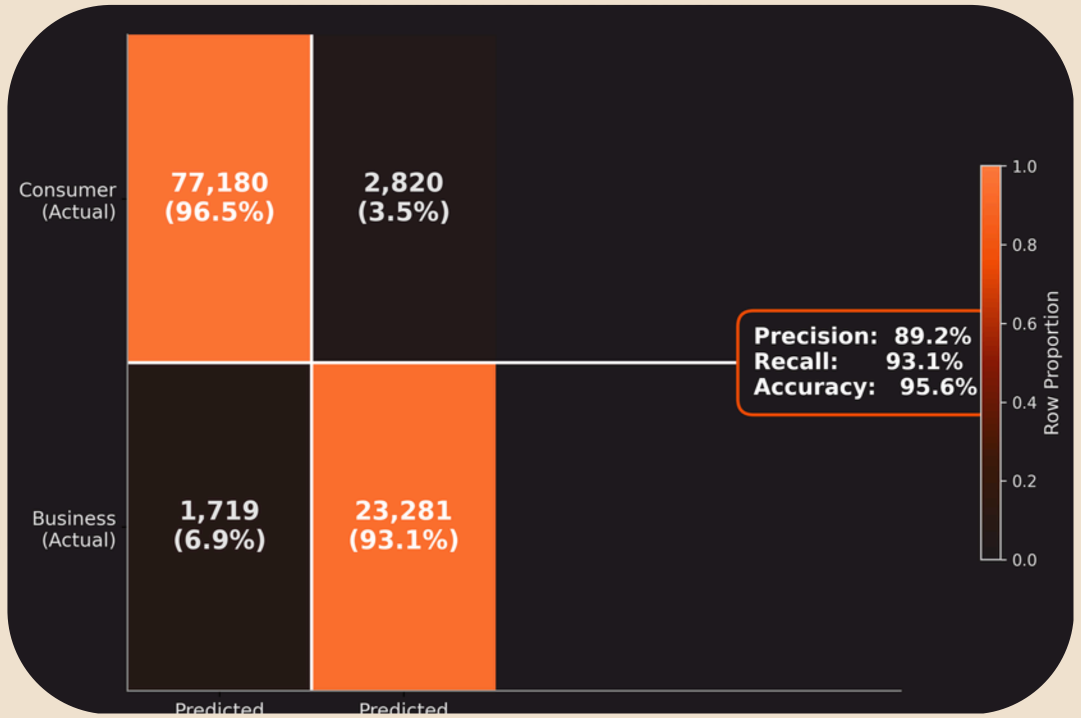
2. *can be used as an **operational risk-scoring tool** for detecting hidden business activity on consumer.*

3. *demonstrates excellent **discriminatory power with ROC-AUC = 0.98,** indicating **strong separation** between hidden business and consumer behavior.*



*The optimal classification threshold **(FPR = 1.8%, TPR = 98.1%)** minimizes false positives while maintaining high sensitivity. **The model is ready for production deployment within transaction monitoring infrastructure.***

1. ***ROC-AUC = 0.98** — excellent classification performance, indicating that the model reliably distinguishes hidden business users from regular consumers.*
2. ***Precision = 89%** — among all cards classified as business, 89% are truly commercial.*
3. ***Recall = 93%** — the model successfully identifies 93% of all hidden business cards in the sample.*



Business Hours Activity as the Strongest Predictor

Key Results:

• _hours_ratio shows the largest marginal effect
• AME = 0.8964

Interpretation:

Cards actively used during standard business hours are significantly more likely to represent hidden commercial activity.

Transaction Timing as the Strongest Indicator of Commercial Activity

- *Average Marginal Effect (AME) analysis revealed that the share of transactions conducted during business hours increases the probability of a card being classified as a business card by* ***+89.6 percentage points****. This impact is four times greater than that of the second most important feature (logarithm of transaction volume,* ***AME = 0.22****).*

Hidden entrepreneurs predominantly conduct transactions during working hours **(9:00 AM–6:00 PM)**, while regular consumers are more active in the evening and at night. This pattern resembles **B2B operational behavior** rather than household consumption.

Economic Meaning:

Hidden businesses behave operationally like companies:
• active during working hours
• process larger operational payments
• exhibit structured transactional behavior

Operational-time transaction behavior is the strongest predictor of hidden commercial activity.

ML-Based Hidden Business Detection System

Core Idea

ThImplement the ML model into Mastercard's transaction monitoring infrastructure to automatically identify consumer cards with hidden commercial activity patterns in real time.

The model analyzes:
- transaction timing
- turnover volume
- merchant concentration
- MCC diversity
- behavioral consistency

Cards with high commercial probability are automatically flagged for review or migration to business products.

Risk Assessment and Mitigation

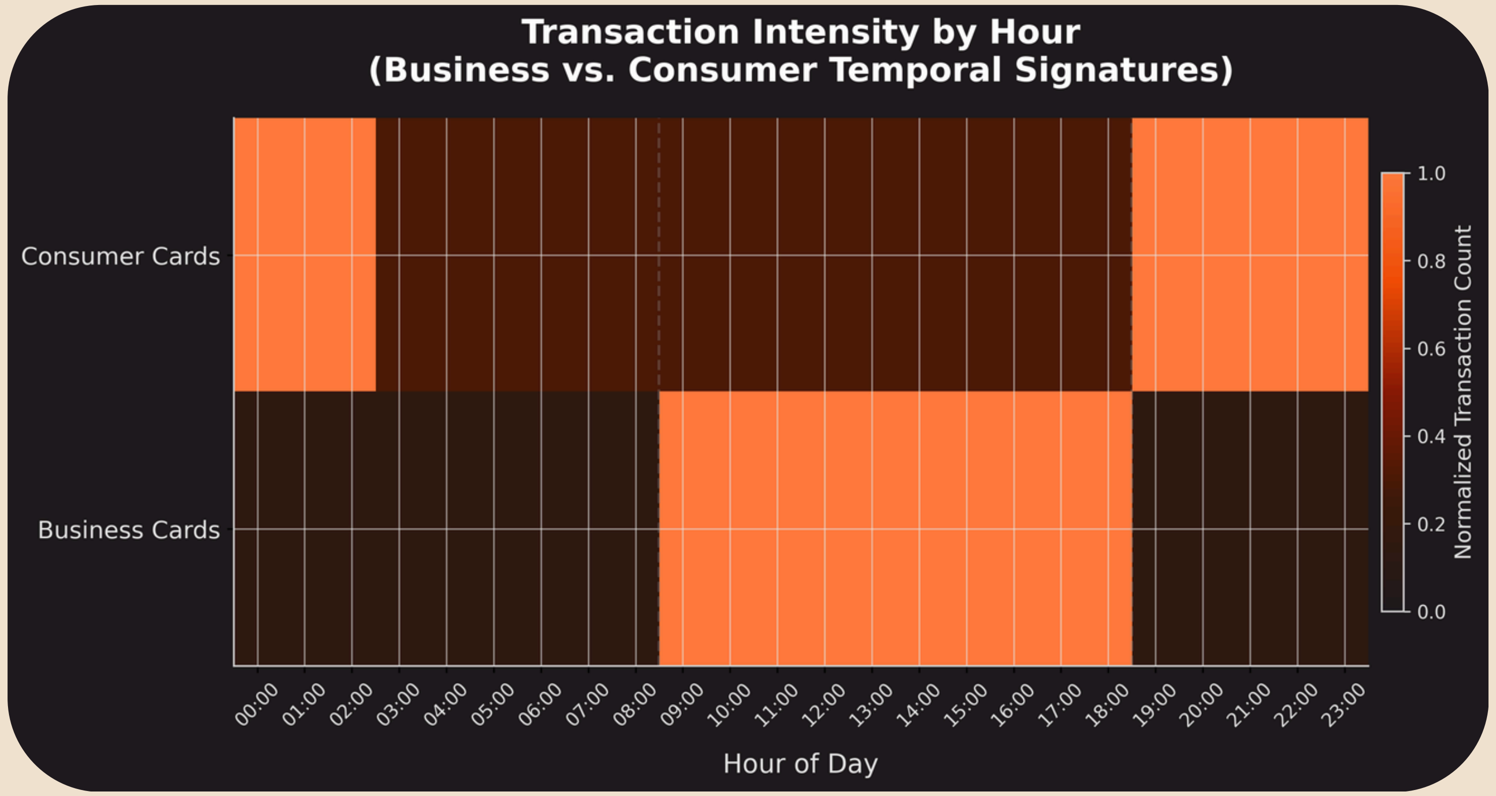
Risks	Mitigations
Consumer users may be incorrectly classified as businesses.	Use calibrated probability thresholds and human review for high-impact cases.
Clients may perceive monitoring as restrictive or invasive.	Position migration as a value-added upgrade with lower fees and business tools.
Behavioral patterns may evolve over time.	Retrain the model quarterly using updated transaction datasets.

RESULTS

Organizational Changes

Create a transaction intelligence monitoring layer
- Integrate ML scoring into fraud/risk infrastructure
- Establish collaboration between analytics, compliance, and SME banking teams

Expected Impact
- Up to 97% detection sensitivity
- Significant reduction in interchange leakage
- Improved segmentation accuracy
- Increased migration from consumer to SME products
- Potential recovery of hundreds of millions KZT annually



Behavioral Financial Intelligence, SME Transition Platform

CORE IDEA

Develop an AI-powered **financial intelligence** platform that analyzes **transactional behavior**, **relationship networks**, and operational patterns to detect hidden commercial activity, **fraud risks**, and **informal business ecosystems**, while guiding users toward compliant and optimized financial solutions instead of immediate punitive actions.

ADAPTIVE RECOMMENDATION ENGINE

The AI model predicts the likely motivation behind suspicious behavior patterns.

CASES	RECOMMENDATIONS
Hidden SME Activity. A user consistently receives operational payments through a personal card.	Offer simplified migration into SME banking with lower onboarding friction, invoicing tools, and optimized business tariffs.
Tax Optimization Behavior. The system detects fragmentation of payments across relatives or affiliated accounts.	Provide legal tax optimization pathways, SME registration guidance, and compliant financial structuring solutions.
Fraud Risk Escalation Behavioral anomalies indicate potential coordinated fraud activity.	Increase monitoring intensity, request additional verification, and gradually restrict risky transaction flows.

PLATFORM LOGIC:

The system builds a dynamic behavioral graph of payment ecosystem participants by analyzing:

- Transaction frequency and turnover
- Merchant concentration and MCC behavior
- Shared geolocation and device patterns
- Relationship proximity (family, business partners, repeated counterparties)
- Temporal transaction behavior (business-hour concentration)
- Network transaction flows between connected individuals and entities

The platform identifies clusters of coordinated financial behavior and estimates the probability of:

- Hidden commercial activity
- Fraudulent transaction schemes
- Informal SME ecosystems
- Tax inefficiencies and regulatory risks

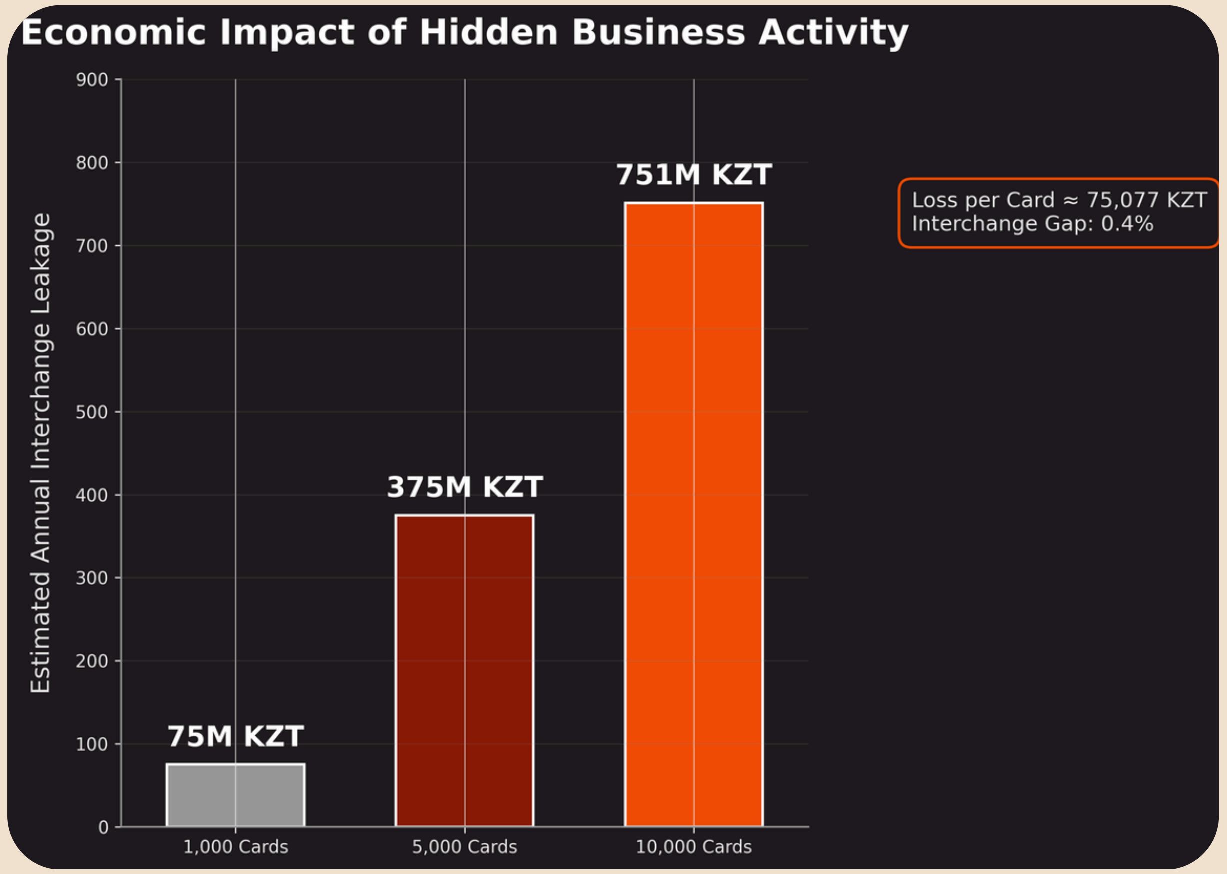
Instead of instantly blocking suspicious accounts, the system applies adaptive intervention logic.

Behavioral Financial Intelligence, SME Transition Platform

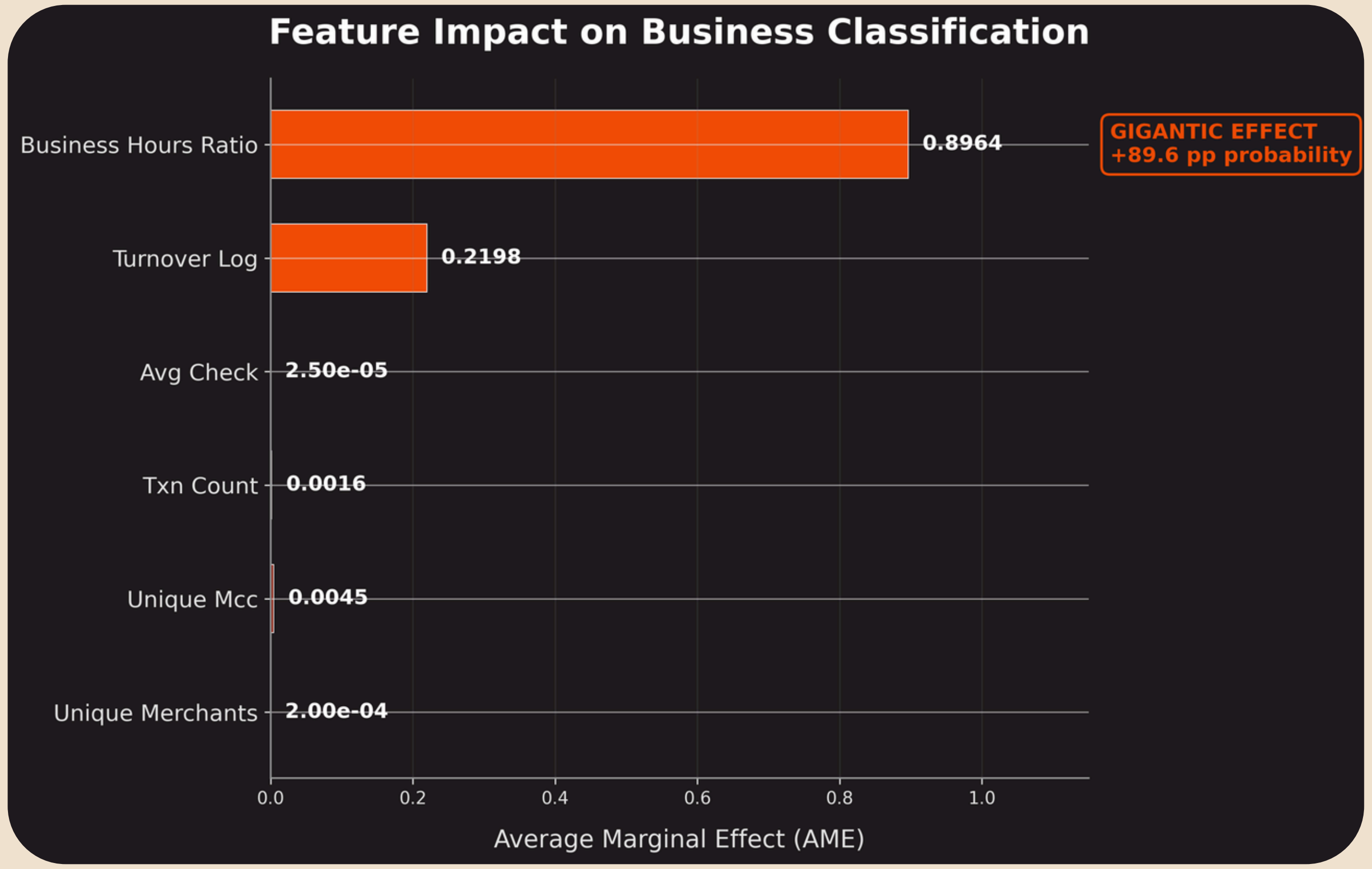
Risks	Mitigations	RESULT
Privacy Concerns	Use anonymized behavioral signals, regulatory-compliant data governance, and transparent consent policies.	Earlier detection of hidden commercial ecosystems
False Relationship Mapping	Apply probabilistic scoring instead of deterministic assumptions and require multi-factor validation.	Reduced interchange leakage and fraud losses
Advanced behavioral monitoring may trigger legal scrutiny.	Ensure compliance with AML/KYC standards and maintain explainable AI auditability.	Increased SME formalization rates
Aggressive intervention could damage brand perception.	Prioritize educational and incentive-based interventions before enforcement actions.	Higher ecosystem monetization efficiency
		Stronger long-term trust between financial institutions and entrepreneurs

BUSINESS IMPACT

SCENARIO	ESTIMATED ANNUAL LEAKAGE
1,000 cards	75M KZT
5,000 cards	375M KZT
10,000 cards	751M KZT



Even under conservative assumptions, hidden business activity may generate hundreds of millions KZT in annual interchange leakage



Hidden business activity is not a niche phenomenon but a material revenue issue.
- **The estimated leakage scales rapidly as the number of hidden business users increases.**
- **Behavioral analytics can transform revenue leakage into a customer acquisition opportunity.**
- **Instead of penalizing users, banks can proactively convert hidden businesses into SME clients through targeted onboarding offers.**