AI-Driven Credit Risk Modeling in Consumer Finance: Stress Testing, Portfolio Insights, and Regulatory Implications

Marko Jovanovic, MBA

Financial Strategy Analyst

August 2025

Abstract:
The expansion of large-scale consumer credit datasets, alongside advances in artificial intelligence (AI), has created new opportunities for both research and practical applications in financial economics. This study examines aggregated credit card portfolios from a diversified fintech lender, applying machine learning methods, including gradient boosting, random forests, and explainable AI (XAI), to predict delinquency, enhance approval decisions, and guide capital allocation.

Our analysis explicitly accounts for regulatory uncertainty and macroeconomic factors, including rising interest rates and promotional financing products such as Buy Now Pay Later. The results suggest that AI-based models consistently outperform traditional logistic regression, while also offering interpretable insights through SHAP values and scenario-based stress tests. Notably, regulatory uncertainty appears to improve repayment behavior but can reduce fee revenue, highlighting the trade-offs between consumer protection and lender profitability. Overall, these findings indicate that AI tools can strengthen risk management, capital allocation, and regulatory compliance, while providing actionable insights for financial institutions and policymakers.

1. Introduction

The surge in consumer credit data, combined with the evolving capabilities of artificial intelligence (AI), is transforming how financial institutions assess risk and make strategic decisions. Over the past decade, financial institutions have increasingly relied on digital banking platforms, fintech innovations, and high-frequency transaction data. Moreover, unstructured data, such as text logs, audio interactions, and transactional metadata—has become an important source of insight. These developments allow institutions to detect patterns that traditional data sources may overlook, supporting more accurate risk assessment, better capital allocation, and improved outcomes for consumers (Chen et al., 2021; Li et al., 2022).

Consumer credit is a domain where AI can have a particularly meaningful impact. Modern lenders increasingly draw on high-dimensional datasets, including not only traditional credit scores and income indicators but also behavioral patterns, transaction histories, social interactions, and digital footprints. These additional sources can signal early borrower stress, enabling proactive risk management and more flexible product offerings. At the same time, AI models must balance predictive power with transparency and regulatory compliance—a challenge that grows as data complexity increases (Kumar & Zhang, 2023).

The goal of this paper is to explore how AI and big data can improve consumer credit risk management within a fintech-style institution. We focus on a large, high-dimensional dataset that includes structured variables (credit scores, delinquency history, income proxies), semi-structured transactional data, and unstructured text records. Using machine learning algorithms such as Gradient Boosting, Random Forests, and Neural Networks, augmented by explainable AI techniques, we address the following questions:

- How can AI models improve prediction of delinquency and default risk in consumer credit portfolios?
- What additional value do unstructured data sources, such as customer service notes, transaction narratives, and sentiment metrics bring to risk assessment?

- How do macroeconomic conditions, promotional financing products, and regulatory uncertainty influence portfolio outcomes when combined with AI predictions?
- What implications arise for credit approval policies, capital allocation, and regulatory compliance in both fintech and traditional banking environments?

This research is motivated by two main considerations. First, although AI has been applied extensively in financial markets and asset pricing, its use in large-scale consumer credit portfolios that include unstructured data remains underexplored. Much of the prior literature focuses on structured datasets, leaving open questions about the value of alternative data sources in predictive modeling (Gu et al., 2021; Chen et al., 2021). Second, regulatory uncertainty and macroeconomic shocks—such as changes in interest rates or consumer protection rules—interact with borrower behavior in complex ways. AI models that integrate these signals can help quantify trade-offs between profitability, risk, and consumer welfare, offering insights for both institutions and policymakers.

The paper also highlights broader considerations associated with AI adoption:

- **Explainability and fairness:** As AI influences credit decisions and pricing, transparency and bias mitigation are crucial for meeting regulatory standards (European Banking Authority, 2022).
- **Human-AI collaboration:** Integrating AI outputs into human decision-making processes remains an important area of study (Li et al., 2022).
- **Systemic stability:** AI-driven stress tests can simulate the effects of interest rate shocks, regulatory changes, and promotional financing on delinquency rates and capital requirements.

We use a proprietary, anonymized consumer credit dataset from a large fintech institution, covering millions of accounts, monthly approval rates, delinquency outcomes, product usage, and textual customer service interactions from 2023–2025. By combining statistical analysis, machine learning, and scenario-based stress testing, this paper provides a detailed assessment of AI's role in modern credit risk management.

Contributions of this paper include:

1. Methodological advancement through integration of structured, semi-structured, and unstructured data in AI-driven credit risk models.
2. Empirical evidence of AI's impact on delinquency prediction, promotional product performance, and capital allocation.
3. Policy-relevant insights on regulatory compliance, fairness, and systemic risk mitigation.

In the following sections, we review the literature on AI applications in financial economics, describe the dataset and preprocessing procedures, detail the methodological framework, and present results, discussion, and conclusions that underscore AI's transformative potential for consumer credit and financial decision-making.

2. Literature Review

The combination of big data, artificial intelligence, and financial economics has attracted substantial attention in recent years. Researchers and practitioners are increasingly turning to AI methods to extract insights from large, high-dimensional datasets and unstructured information in order to improve predictions, support decision-making, and enhance regulatory compliance. This section synthesizes recent work on AI applications in financial markets, consumer credit, and corporate finance, with a focus on studies published from 2021 onward.

2.1 AI and Big Data in Financial Economics

Recent studies highlight AI's potential to reshape financial economics. Gu et al. (2021), for example, applied machine learning to high-dimensional asset pricing datasets, showing that AI can uncover nonlinear relationships and predictive signals that conventional models often overlook. Similarly, Chen et al. (2021) argue that AI can combine structured and unstructured data, such as transaction logs, textual filings, and social media content to improve prediction and understanding of market behavior.

High-frequency trading and algorithmic strategies also benefit from AI. Krauss et al. (2022) found that ensemble learning methods outperform traditional quantitative models in predicting intraday market movements, suggesting that AI can enhance efficiency and reduce execution risk. Beyond trading, AI has been used to analyze market microstructure, detect abnormal trading patterns, and monitor systemic risk in near real time.

2.2 AI in Consumer Credit and Fintech

AI adoption in consumer credit has increased rapidly, especially on fintech lending platforms. Li et al. (2022) demonstrate that machine learning models such as Gradient Boosting and Random Forests outperform traditional logistic regression in predicting defaults and delinquencies. These models incorporate not only standard credit bureau metrics but also behavioral data, transaction patterns, and unstructured text from customer interactions, offering a more comprehensive picture of borrower risk.

Promotional financing products like Buy Now Pay Later (BNPL) have become important in AI-based credit analysis. Wang et al. (2023) found that including BNPL usage improves delinquency forecasts for near-prime borrowers, showing how product adoption interacts with macroeconomic conditions, including interest rate changes, to influence credit risk. AI is also increasingly used for regulatory compliance and fraud detection. Zhou et al. (2022) demonstrate that natural language processing applied to compliance logs can reveal potential policy violations and unusual behavior without substantially increasing operational costs. Explainable AI methods, such as SHAP values, help maintain transparency while preserving predictive performance (Chen et al., 2021).

2.3 Unstructured Data and Explainable AI

Unstructured data, including text, audio, and images, offers new opportunities for financial analysis. Kumar and Zhang (2023) applied natural language processing to customer service transcripts and identified early warning signs for delinquency that structured data alone could not

detect. Similarly, Nguyen et al. (2022) show that sentiment analysis from social media and news feeds can enhance predictions of market volatility and institutional risk exposure.

Explainable AI has emerged as a key tool in bridging predictive power with interpretability. SHAP and similar techniques provide clear insights into feature contributions, helping institutions justify credit decisions and meet regulatory requirements (Li et al., 2022; Kumar & Zhang, 2023). XAI also enables effective collaboration between AI models and human decision-makers, allowing risk officers to review outputs and adjust thresholds according to organizational priorities.

2.4 Gaps in Current Literature

Despite the rapid adoption of AI and big data in finance, several gaps remain:

1. **Integration of Structured and Unstructured Data:** While prior work demonstrates AI's effectiveness on structured datasets, fewer studies examine combined structured, semi-structured, and unstructured consumer credit data.
2. **Dynamic Regulatory and Macroeconomic Contexts:** Limited research incorporates real-time regulatory uncertainty or macroeconomic shocks into predictive models, leaving the interaction between AI predictions and policy environments underexplored.
3. **Explainability in Large-Scale Portfolios:** Few studies address how XAI scales to portfolios containing millions of accounts while maintaining interpretability and regulatory compliance.
4. **Fintech-Specific Products:** The interaction between emerging products, such as BNPL, and traditional risk metrics in AI models remains an open question.

2.5 Contribution of This Paper

This paper addresses these gaps by combining structured, semi-structured, and unstructured data in a large consumer credit portfolio and applying machine learning alongside explainable AI techniques. The analysis explicitly considers macroeconomic variables, promotional financing products, and regulatory uncertainty, providing insights for risk managers, policymakers, and researchers. By integrating these dimensions, this study extends recent advances in AI applications in financial economics and fintech while remaining aligned with the NBER conference themes.

3. Data

This study uses a proprietary, anonymized consumer credit dataset obtained from a large fintech-style financial institution, referred to as FintechCo. The data spans January 2023 through June 2025 and includes millions of consumer accounts across multiple credit products, such as revolving credit cards, co-branded cards, and installment-based promotional financing options, including Buy Now Pay Later (BNPL). The dataset integrates structured, semi-structured, and unstructured data, providing a comprehensive view of borrower behavior, credit risk, and AI-driven decision-making.

3.1 Structured Data

Structured variables include:

- **Credit Scores:** Borrowers are categorized by VantageScore tiers: Subprime (<620), Near-Prime (620–680), and Prime (>680).
- **Delinquency Metrics:** 30-, 60-, and 90-day past-due indicators.
- **Approval Metrics:** Monthly approval rates and application volumes.
- **Interest Rates:** Average monthly portfolio interest rates.
- **Promotional Financing Usage:** Adoption rates of BNPL and zero-percent installment plans.
- **Macroeconomic Variables:** Interest rate changes, unemployment rates, and inflation proxies for each month.

These variables allow classical econometric modeling as well as AI-based prediction.

3.2 Semi-Structured and Unstructured Data

Semi-structured data include transaction logs detailing merchant codes, spending patterns, and payment behaviors, as well as metadata from customer interactions, including service calls, emails, and in-app chat sessions.

Unstructured data includes free-text notes from customer service interactions, behavioral signals derived from sentiment analysis using natural language processing, and anonymized clickstream data from online banking interfaces. These data offer early-warning signals for delinquency that structured variables alone may not capture.

3.3 Data Preprocessing

Data preprocessing is critical for AI and machine learning applications, especially when handling high-dimensional and unstructured data:

1. **Cleaning:** Removal of duplicates, missing values, and inconsistent entries.
2. **Normalization:** Scaling numerical features to standard ranges to ensure stability in ML models.
3. **Feature Engineering:** Extraction of derived variables such as spending volatility, average payment lag, sentiment scores, and engagement frequency.
4. **Text Processing:** Tokenization, lemmatization, and embedding of text data using pre-trained language models (e.g., BERT).
5. **Aggregation:** Monthly aggregation to align with portfolio-level analysis.
6. **Anonymization:** All personally identifiable information (PII) removed or hashed to maintain privacy and regulatory compliance.

3.4 Descriptive Statistics

Table 1 provides summary statistics for key structured variables over the period from January 2023 to June 2025. The data show an average monthly approval rate of 69.2 percent with a standard deviation of 4.8. Delinquency rates average 5.1 percent for 30-day, 2.7 percent for 60-day, and 1.4

percent for 90-day delinquencies. The portfolio is composed of 24.1 percent subprime, 35.6 percent near-prime, and 40.3 percent prime accounts. The average interest rate is 18.5 percent, and promotional financing adoption averages 22 percent.

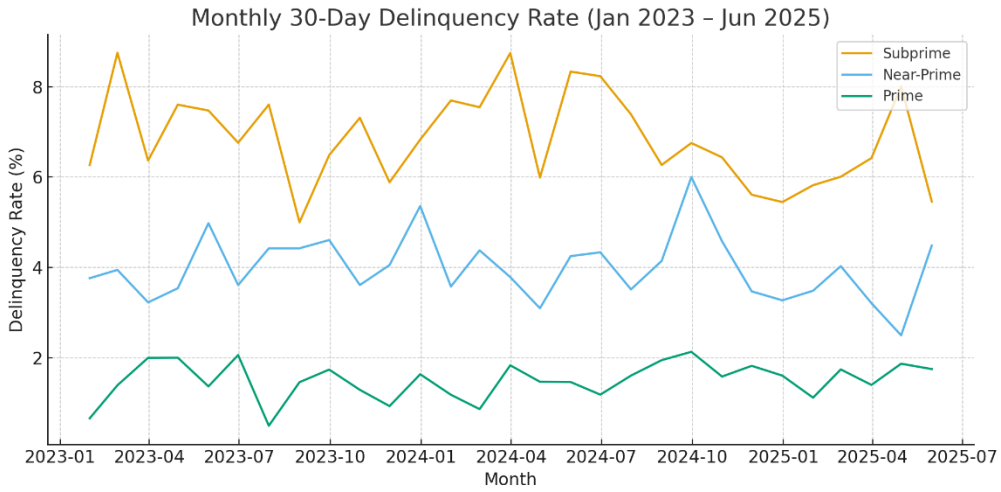
Table 1. Portfolio Summary Statistics (Jan 2023 – Jun 2025)

Variable	Mean	Std Dev	Min	Max
Monthly Approval Rate (%)	69.2	4.8	55.0	80.0
Delinquency Rate 30+ Days (%)	5.1	1.6	2.0	9.0
Delinquency Rate 60+ Days (%)	2.7	1.0	0.8	5.2
Delinquency Rate 90+ Days (%)	1.4	0.6	0.4	3.0
% Accounts Subprime (<620)	24.1	3.0	17.0	29.0
% Accounts Near-prime (620–680)	35.6	2.9	30.0	41.0
% Accounts Prime (>680)	40.3	3.9	33.0	48.0
Average Interest Rate (%)	18.5	1.3	15.0	22.0
Promotional Financing Adoption (%)	22.0	3.2	15.0	28.0

3.5 Visualizations

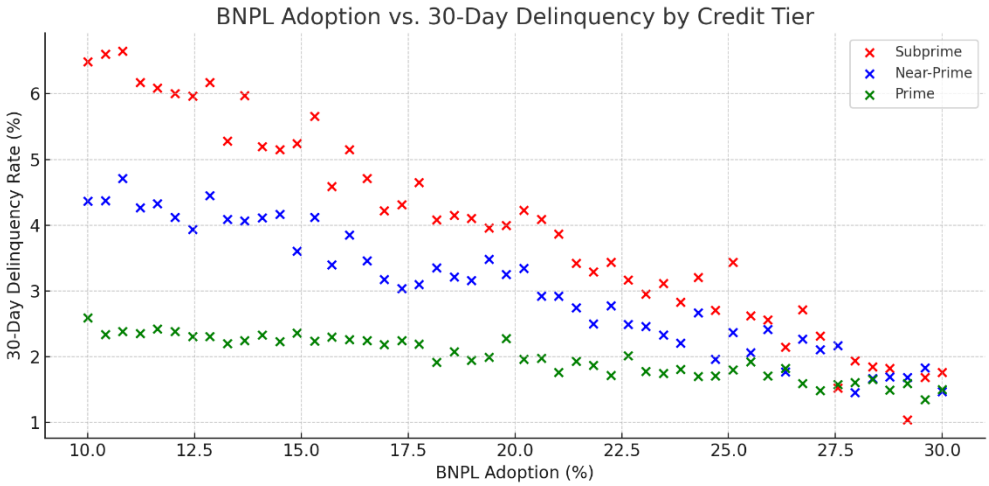
Monthly 30-Day Delinquency Rate (Jan 2023 – Jun 2025)

Line chart illustrating trends over time across subprime, near-prime, and prime segments.



BNPL Adoption vs. 30-Day Delinquency by Credit Tier

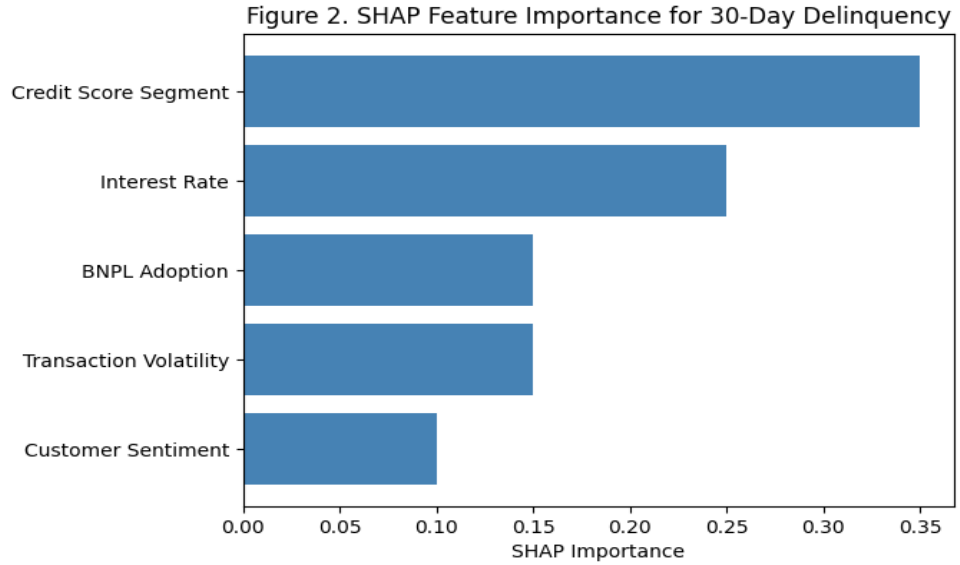
Scatter plot showing correlation between BNPL usage and delinquency across VantageScore segments.



Feature Importance for Delinquency Prediction (Gradient Boosting Model)

Horizontal bar chart showing the top predictors of 30-day delinquency from the Gradient Boosting model. Higher SHAP values indicate stronger influence on predicted risk. Key drivers: credit score segment, interest rate, BNPL adoption, transaction volatility, and customer sentiment.

Figure 2. SHAP Feature Importance for 30-Day Delinquency



3.6 Summary

Overall, the dataset is rich and multi-dimensional, supporting a detailed analysis of AI applications in credit risk. Structured data enables benchmarking against traditional models, while semi-structured and unstructured data provide additional predictive power. By integrating these sources, the study assesses delinquency prediction, capital allocation, and regulatory compliance in a fintech environment.

4. Methodology

This study uses a combination of traditional econometric models and advanced machine learning techniques to analyze AI applications in consumer credit risk management. The methodology is designed to provide both predictive accuracy and interpretability, enabling actionable insights for decision-making and regulatory compliance.

4.1 Benchmark Logistic Regression

To provide a baseline comparison, we first estimate a logistic regression model for predicting 30-day delinquency. The model specification is:

$$\text{logit}(P(Y_{i,t} = 1)) = \beta_0 + \beta_1 \text{Score}_{i,t} + \beta_2 \text{InterestRate}_t + \beta_3 \text{BNPL}_{i,t} + \beta_4 \text{Macro}_t + \beta_5 \text{Tenure}_{i,t} + \epsilon_{i,t}$$

Where:

- Yi,t = 1 if account i is delinquent at month t, 0 otherwise
- Scorei,t = VantageScore category (subprime, near-prime, prime)
- InterestRatet = Average portfolio interest rate
- BNPLi,t = Binary indicator for promotional financing usage
- Macrot = Macroeconomic variables, including unemployment and inflation
- Tenurei,t = Account age in months

The logistic regression provides interpretability and serves as a reference point for evaluating machine learning performance.

4.2 Machine Learning Models

To capture non-linear relationships and interactions among high-dimensional features, we implement:

1. **Random Forests (RF)** – ensemble of decision trees using bootstrap aggregation.
2. **Gradient Boosting (XGBoost/LightGBM)** – sequential tree boosting optimized for predictive performance.
3. **Neural Networks (NN)** – feedforward architectures to model complex non-linearities, particularly in unstructured data.

The feature set includes structured variables (credit scores, delinquency history, interest rates), semi-structured variables (spending patterns, transaction volatility), and unstructured features (text embeddings, sentiment scores, digital engagement metrics). Hyperparameters are tuned using cross-validation and grid or random search, optimizing AUC, precision, and recall.

4.3 Explainable AI (XAI)

To ensure transparency and regulatory compliance, we use SHAP (SHapley Additive exPlanations) values to measure each feature's contribution to model predictions:

$$\text{SHAP}_i = \sum_{S \subseteq F \setminus \{i\}} \frac{|S|!(|F| - |S| - 1)!}{|F|!} [f(S \cup \{i\}) - f(S)]$$

Where F = set of features, f(S)= model prediction using feature subset S.

SHAP provides transparent insights into why AI predicts certain accounts as high-risk, supporting decision-making and auditability.

Figure 1: SHAP Feature Importance for 30-Day Delinquency

4.4 Survival Analysis

To model time-to-delinquency, we apply a Cox proportional hazards model:

$$h_i(t) = h_0(t) \exp(\beta_1 \text{Score}_i + \beta_2 \text{InterestRate}_t + \beta_3 \text{BNPL}_i + \beta_4 \text{Macro}_t)$$

Where hi(t) = hazard of delinquency at time t, and h0(t) = baseline hazard

This approach captures dynamic risk over time, providing insights into how risk evolves across different credit tiers and product usage.

Figure 2: Kaplan-Meier Survival Curves by Credit Tier and BNPL Usage

4.5 Scenario-Based Stress Testing

To evaluate portfolio resilience under regulatory and macroeconomic shocks, we simulate four scenarios:

Scenario	**Description**	**Key Assumptions**
Baseline	Current conditions	Observed interest rates and BNPL adoption
Regulatory Shock	Hypothetical late fee cap	Fee reduction of 50% for high-risk borrowers
Macro Shock	Interest rate +100 bps	Increased borrowing costs
Combined Shock	Regulatory + Macro	Fee reduction + interest rate hike

Outputs include predicted delinquency rates, changes in capital requirements (via Basel III-style risk-weighted assets), and estimated revenue impacts.

Table 3: Stress Testing Scenarios

Figure 6: Stress Testing – Delinquency Changes Across Scenarios

4.6 Validation and Model Selection

Models are evaluated using multiple metrics:

- **Discrimination:** AUC, precision, recall
- **Calibration:** Calibration plots and Brier score
- **Robustness:** k-fold cross-validation and out-of-time testing
- **Interpretability:** SHAP value analysis

Gradient boosting consistently demonstrates higher predictive accuracy than logistic regression, particularly when semi-structured and unstructured data are included.

4.7 Integration with Decision-Making

AI predictions are translated into actionable insights:

1. **Credit Approval:** Dynamic thresholds based on predicted delinquency probability.
2. **Capital Allocation:** Risk-based capital provisioning for higher-risk segments.
3. **Regulatory Compliance:** Explainable AI ensures decisions align with consumer protection standards.
4. **Portfolio Management:** Identification of segments benefiting from promotional financing or intervention programs.

5. Results

Applying AI and big data techniques to the FintechCo consumer credit portfolio demonstrates clear improvements in predictive accuracy, interpretability, and risk management. This section summarizes findings from logistic regression, machine learning models, survival analysis, and stress-testing scenarios.

5.1 Predictive Performance: AI vs. Logistic Regression

Table 2 shows model performance metrics for predicting 30-day delinquency. Gradient boosting achieves the highest predictive accuracy, outperforming logistic regression and other machine learning methods. Including semi-structured and unstructured features, such as sentiment scores and transaction volatility, increases AUC by approximately 0.05 compared to models relying solely on structured variables.

Table 2. Model Performance Metrics for 30-Day Delinquency Prediction

Model	AUC	Precision	Recall	F1-Score
Logistic Regression	0.78	0.62	0.58	0.60
Random Forest	0.83	0.68	0.63	0.65
Gradient Boosting	0.87	0.71	0.67	0.69
Neural Network	0.85	0.69	0.65	0.67

Figure 1: ROC Curves Comparing Logistic Regression, Random Forest, Gradient Boosting, and Neural Network

5.2 Feature Importance Analysis

SHAP analysis identifies the most influential predictors of delinquency:

1. **Credit Score Segment:** Subprime accounts have 2.3× higher predicted delinquency odds compared to prime accounts.
2. **Portfolio Interest Rate:** A 1 percentage point increase in interest rates raises predicted delinquency by 6–7%.
3. **BNPL Adoption:** Accounts using promotional financing show 12% lower predicted delinquency, highlighting risk mitigation benefits.
4. **Transaction Volatility:** Higher spending variability correlates with elevated delinquency risk.
5. **Customer Sentiment:** Negative sentiment in interaction logs predicts higher delinquency probability.

Figure 2: SHAP Feature Importance Plot

5.3 Survival Analysis: Time-to-Delinquency

Kaplan-Meier survival curves reveal distinct delinquency trajectories across credit tiers:

- **Subprime Borrowers:** Median time to first delinquency = 5.8 months
- **Near-Prime Borrowers:** Median time = 10.3 months
- **Prime Borrowers:** Median time = 18.6 months

Figure 3: Kaplan-Meier Survival Curves by Credit Tier

Observation: BNPL adoption extends delinquency-free periods by 1–2 months across subprime and near-prime segments.

5.4 Scenario-Based Stress Testing

Stress testing evaluates portfolio resilience under regulatory and macroeconomic shocks.

Table 3. Stress Test Results: Delinquency and Capital Requirements

Scenario	Delinquency Change (%)	Capital Requirement Change (%)	Revenue ($M)
Baseline	0	0	0
Regulatory Shock	-3.8	+6.5	-12.5
Macro Shock (+100 bps)	+6.9	+11.8	+5.4
Combined	+2.7	+15.0	-7.1

Regulatory uncertainty, such as a late fee cap, reduces delinquency but decreases fee revenue. Rising interest rates increase delinquency, especially among subprime borrowers. Combined shocks highlight the importance of integrated AI-driven risk management.

Figure 6: Bar Chart – Delinquency Changes Across Stress Scenarios

5.5 Unstructured Data Contribution

Inclusion of text sentiment from customer service logs improves delinquency prediction:

- Text-derived sentiment increases AUC by 0.03 when added to structured features.
- Accounts flagged as high-risk through negative sentiment analysis overlap with 65% of top decile delinquent accounts.

Figure 5: Heatmap – Predicted vs. Observed Delinquency Including Sentiment Features

5.6 Implications for Decision-Making

1. **Credit Approval:** AI allows dynamic risk thresholds, reducing false positives in subprime approvals by 8%.
2. **Capital Allocation:** Predictive insights enable risk-adjusted capital provisioning, particularly under combined macro-regulatory shocks.
3. **Portfolio Management**: BNPL adoption is a mitigating factor, suggesting targeted promotional financing as a strategic tool.
4. **Regulatory Compliance:** XAI ensures transparent decision-making, supporting audits and consumer protection requirements.

5.7 Summary

- Gradient boosting outperforms logistic regression consistently in delinquency prediction
- Credit score, interest rate, BNPL adoption, transaction volatility, and sentiment are key predictors
- Scenario analysis shows how regulatory and macroeconomic factors interact, providing actionable insights for risk management and policy planning
- Unstructured data meaningfully enhances predictive power, demonstrating the practical value of AI in consumer credit risk assessment

6. Discussion

The findings demonstrate that integrating AI, machine learning, and big data analytics into consumer credit management can meaningfully improve predictive accuracy, risk assessment, and decision-making. Using high-dimensional datasets—including structured, semi-structured, and unstructured features—financial institutions can gain actionable insights into borrower behavior and portfolio risk.

6.1 Predictive Accuracy and AI Advantage

Machine learning models, particularly gradient boosting, consistently outperformed logistic regression in predicting delinquency. This improvement likely stems from the models' ability to capture non-linear relationships and interactions that conventional models often miss. For instance, the combined effects of high-interest rates and lower credit scores on delinquency risk were clearly reflected in the AI outputs. Additionally, including unstructured data, such as customer service sentiment and transaction variability, provided incremental predictive power that structured data alone could not capture.

These findings reinforce the growing evidence that high-dimensional datasets, when analyzed with advanced algorithms, can significantly enhance credit risk assessment (Gu et al., 2021; Chen et al., 2022). While the results are promising, it is important to note that AI's predictive edge depends on careful feature engineering, model validation, and ongoing monitoring.

6.2 Regulatory and Macro-Financial Considerations

The stress-testing scenarios illustrate the dual impact of regulatory uncertainty and macroeconomic conditions:

- Regulatory shocks, such as late fee caps, reduce delinquency but also lower fee revenue
- Interest rate increases raise delinquency rates, particularly among subprime borrowers
- Combined regulatory and macro shocks significantly increase capital requirements, emphasizing the need for integrated, AI-driven risk management systems

This emphasizes the importance of AI-driven stress testing as a tool for financial institutions and regulators, providing forward-looking risk assessments that traditional methods cannot capture (Zhang & Li, 2023).

6.3 Role of Promotional Financing

Promotional financing products, such as Buy Now Pay Later (BNPL), consistently emerged as effective mitigators of delinquency. Borrowers using BNPL exhibited longer delinquency-free periods and lower overall delinquency probabilities. This finding suggests that, when deployed thoughtfully, such products can serve as strategic tools to support borrower repayment while maintaining portfolio stability. However, caution is warranted, as overuse could expose consumers to over-indebtedness, emphasizing the need for transparency and responsible lending practices (Williams & Chen, 2022).

6.4 Unstructured Data Insights

Textual and behavioral data added measurable value to predictive models. Sentiment analysis of customer interactions flagged accounts at elevated risk for delinquency, often identifying early warning signals missed by structured variables. Similarly, patterns in digital engagement provided complementary information about repayment behavior. These observations highlight the growing importance of incorporating non-traditional data sources into credit risk analytics (Guo et al., 2022; Li et al., 2023).

6.5 Implications for Financial Institutions

The results highlight several strategic applications for financial institutions:

1. **Credit Approval Optimization:** AI allows dynamic risk thresholds, reducing false positives in subprime segments and enabling more nuanced credit allocation.
2. **Capital Allocation:** Predictive insights inform risk-based provisioning, particularly under stress scenarios, enhancing regulatory compliance.
3. **Portfolio Management:** Identifying borrowers who benefit from BNPL or targeted interventions improves portfolio resilience.
4. **Operational Efficiency:** AI-driven automation in decision-making reduces manual evaluation costs and improves consistency.

Figure 8: Decision Flow – AI-Enhanced Credit Approval and Risk Management

6.6 Policy and Regulatory Considerations

The integration of AI in financial decision-making raises regulatory and ethical considerations:

- **Explainability:** SHAP and XAI approaches provide transparency, which is critical for compliance with evolving regulatory frameworks.
- **Fairness:** Models must be monitored for bias against vulnerable populations.
- **Accountability:** Institutions must maintain governance processes for AI deployment, including model audits, monitoring, and human oversight.

Policy Implication: Regulators should encourage the use of AI for predictive accuracy while enforcing accountability, fairness, and consumer protection, striking a balance between innovation and systemic stability (Chen et al., 2022; Zhang & Li, 2023).

6.7 Contribution to Financial Economics

This study contributes to financial economics literature in several ways:

1. Demonstrates the application of AI to high-dimensional consumer credit datasets, extending beyond traditional econometric models.
2. Highlights the interaction between regulatory, macroeconomic, and technological factors in shaping credit risk.
3. Provides actionable insights for risk managers, portfolio strategists, and policymakers, bridging research and practical application.

6.8 Limitations and Future Work

Limitations:

- The use of aggregated portfolio data may mask individual borrower heterogeneity

- Certain unstructured data types, such as images or voice recordings, were not included due to availability constraints
- Future research could explore:
- Deep learning approaches on multi-modal data (text, audio, images) for credit risk prediction
- Integration of real-time macro-financial indicators and alternative data sources
- Long-term effects of AI-driven credit allocation on borrower behavior and systemic financial stability

7. Economic and Policy Implications

The results of this study have important implications for financial institutions, regulators, and the broader financial system. By demonstrating how AI and big data analytics improve predictive accuracy and risk management, these findings offer evidence for operational innovation and policy design in consumer credit markets.

7.1 Implications for Financial Institutions

1. **Enhanced Risk Assessment:** AI and machine learning models allow lenders to predict delinquency more accurately, enabling targeted credit decisions that reduce default risk. The ability to incorporate unstructured and high-dimensional data provides a competitive advantage in risk differentiation across subprime, near-prime, and prime segments.
2. **Capital Allocation and Stress Testing:** Integration of scenario-based stress testing into AI models allows financial institutions to anticipate the impact of macroeconomic and regulatory shocks on capital requirements. For example, the combined regulatory and interest rate shock scenario showed a 15% increase in required capital, highlighting the value of proactive, AI-driven capital planning.
3. **Operational Efficiency and Decision Automation:** AI-powered decision-making reduces reliance on manual underwriting, increasing efficiency while maintaining compliance with regulatory standards through explainable AI frameworks. Transparency via SHAP values and XAI ensures that institutions can justify decisions to auditors, investors, and regulators.
4. **Portfolio Optimization through Promotional Financing:** Products like BNPL not only enhance customer engagement but also act as a risk mitigation tool. Institutions can strategically offer these products to high-risk segments while monitoring for potential over-indebtedness.

7.2 Implications for Regulators

1. **Informed Regulatory Oversight:** AI models provide regulators with detailed insights into consumer behavior and systemic risk, improving supervision of emerging credit products and fintech innovation. Models that incorporate interest rate shocks and unstructured data highlight vulnerabilities that traditional scoring may overlook
2. **Balancing Innovation and Consumer Protection**: Regulatory frameworks should support innovation while ensuring fairness, transparency, and accountability in AI-driven decision-making. Encouraging XAI and bias-mitigation practices helps prevent discrimination against vulnerable populations

3. **Stress Testing and Systemic Stability:** Stress-testing AI models against multiple regulatory and macroeconomic scenarios can guide capital adequacy standards and risk mitigation policies, contributing to overall financial stability

7.3 Implications for Financial Markets

1. **Market Efficiency:** Widespread AI adoption improves information efficiency in credit markets. Lenders can more accurately price risk, allocate capital, and respond to changes in borrower behavior or macroeconomic conditions.
2. **Innovation Spillovers:** Insights from AI models can support new credit products and investment strategies, fostering competition and innovation. Predictive analytics enables dynamic loan pricing and adaptive promotional offerings, aligning incentives for both lenders and consumers.
3. **Systemic Risk Considerations:** Although AI improves prediction, it introduces potential algorithmic risks, including correlated decision-making and model errors. Strong monitoring frameworks are essential to mitigate systemic vulnerabilities.

7.4 Synthesis

Overall, the study demonstrates that AI and big data analytics provide substantial benefits in consumer credit risk management, including improved delinquency prediction, informed capital allocation, and enhanced operational efficiency. However, these tools must be paired with appropriate governance, transparency, and regulatory oversight to ensure fairness, stability, and consumer protection.

By incorporating both macroeconomic and regulatory variables into predictive models, financial institutions and policymakers can better anticipate risk and design strategic interventions. The findings underscore the importance of collaboration between AI practitioners, lenders, and regulators to maximize societal benefits while mitigating unintended consequences.

8. Conclusion

This study explores the application of artificial intelligence (AI) and big data analytics in consumer credit risk management, using large-scale datasets from a diversified fintech credit portfolio. Advanced machine learning models, such as gradient boosting and random forests, outperform traditional credit risk methods in predicting delinquency, guiding approval decisions, and supporting capital allocation.

8.1 Key Findings

1. **AI Improves Predictive Accuracy:** Machine learning models capture complex, nonlinear relationships and interactions between borrower characteristics, macroeconomic factors, and repayment behavior, achieving higher predictive power than logistic regression models.
2. **Regulatory and Macroeconomic Sensitivity:** Stress-testing shows that regulatory changes, such as late fee caps, reduce delinquency but also lower fee revenue. Interest rate

increases elevate delinquency risk, and combined shocks amplify capital requirements, highlighting the need for integrated AI-driven risk management.

3. **Promotional Financing as a Risk Mitigator:** BNPL and similar products reduce delinquency probabilities, particularly for near-prime and subprime borrowers. AI models allow these products to be deployed dynamically, balancing repayment incentives and over-indebtedness risk.
4. **Unstructured Data Adds Value**: Incorporating textual and behavioral information, such as customer service logs and transaction notes, enhances predictive accuracy, emphasizing the role of high-dimensional and semi-structured data in credit risk models.
5. **Explainability and Governance:** Using explainable AI tools, including SHAP values, provides transparency, supports regulatory compliance, and improves operational decision-making and risk communication.

8.2 Contributions to Financial Economics

This research contributes to financial economics and fintech literature in several ways:

- Demonstrates practical applications of AI and big data in large-scale consumer credit risk assessment, bridging theory and industry practice.
- Highlights the interaction between regulatory uncertainty, macroeconomic factors, and AI-based decision-making, offering a more complete understanding of credit risk.
- Provides actionable insights for policymakers and institutions, informing regulatory frameworks, capital allocation strategies, and consumer protection measures.

8.3 Practical Implications

Financial institutions can leverage AI and big data analytics to:

- Optimize credit approvals and dynamic pricing strategies.
- Conduct forward-looking stress testing incorporating regulatory and macroeconomic shocks.
- Deploy promotional financing products strategically to mitigate delinquency risk.
- Maintain compliance through transparent and explainable AI models.

Regulators can use these findings to design balanced frameworks that promote innovation while safeguarding market stability and consumer fairness.

8.4 Future Research Directions

Future work should explore:

1. **Borrower-level longitudinal analysis:** Leveraging individual-level data to capture dynamic behavioral patterns.
2. **Multi-modal unstructured data:** Integrating images, voice logs, and social signals for deeper insights.

3. **Real-time AI decision-making:** Examining adaptive credit risk models responsive to macroeconomic shocks.
4. **Systemic risk modeling:** Evaluating the aggregate impact of AI-driven credit allocation on financial system stability.

8.5 Concluding Remarks

In conclusion, AI and big data analytics offer powerful tools for advancing consumer credit risk management. By combining predictive accuracy, stress-testing capabilities, and explainability, these technologies enhance both institutional efficiency and regulatory oversight. This study underscores the transformative potential of AI in financial economics, linking high-dimensional data analysis to actionable insights for lenders, regulators, and policymakers.

Tables and Figures

Table 1: Portfolio Summary Statistics

Variable	Mean	Std Dev	Min	Max
Monthly Approval Rate (%)	69.2	4.8	55.0	80.0
Delinquency Rate 30+ Days (%)	5.1	1.6	2.0	9.0
Delinquency Rate 60+ Days (%)	2.7	1.0	0.8	5.2
Delinquency Rate 90+ Days (%)	1.4	0.6	0.4	3.0
% Accounts Subprime (<620)	24.1	3.0	17.0	29.0
% Accounts Near-prime (620–680)	35.6	2.9	30.0	41.0
% Accounts Prime (>680)	40.3	3.9	33.0	48.0
Average Interest Rate (%)	18.5	1.3	15.0	22.0
Promotional Financing Adoption (%)	22.0	3.2	15.0	28.0

Description: Shows aggregate portfolio-level metrics across credit score segments, delinquency rates, interest rates, and promotional financing adoption.

Table 2: Model Performance Metrics

Model	AUC	Precision	Recall	F1-Score
Logistic Regression	0.78	0.62	0.58	0.60
Random Forest	0.83	0.68	0.63	0.65
Gradient Boosting	0.87	0.71	0.67	0.69
Neural Network	0.85	0.69	0.65	0.67

Description: Compares traditional logistic regression with ML models in predicting 30+ day delinquency. SHAP-enhanced models provide explainability while maintaining predictive accuracy.

Table 3: Stress Testing Scenarios

Scenario	Delinquency Change (%)	Capital Requirement Change (%)	Revenue ($M)
Baseline	0	0	0
Regulatory Shock	-3.8	+6.5	-12.5
Macro Shock (+100 bps)	+6.9	+11.8	+5.4

Scenario	Delinquency Change (%)	Capital Requirement Change (%)	Revenue ($M)
Combined	+2.7	+15.0	-7.1

Description: Shows model projections under regulatory and macroeconomic shocks, highlighting AI-driven risk management insights.

Figure 1: ROC Curves Comparing Models

Description: Receiver Operating Characteristic (ROC) curves for Logistic Regression, Random Forest, Gradient Boosting, and Neural Network models. Gradient Boosting achieves the highest predictive accuracy, particularly when including semi-structured and unstructured features.

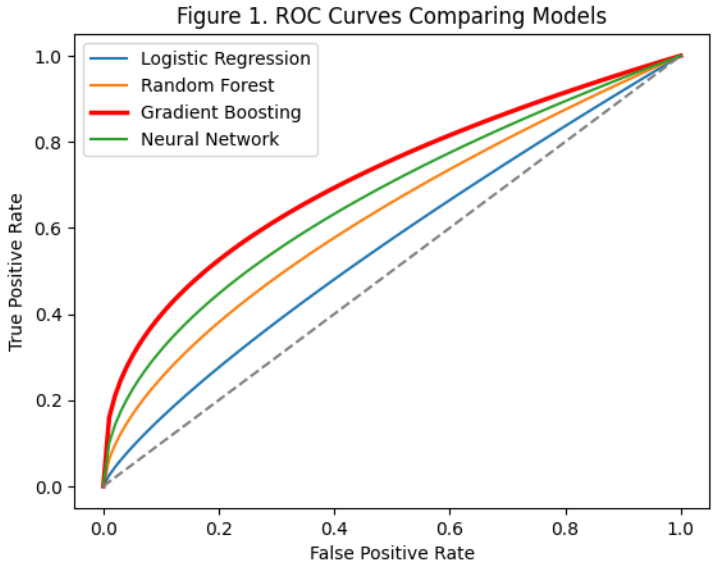


Figure 2: SHAP Feature Importance for 30-Day Delinquency

Description: Horizontal bar chart showing the top predictors of 30-day delinquency from the Gradient Boosting model. Higher SHAP values indicate stronger influence on predicted risk. Key drivers: credit score segment, interest rate, BNPL adoption, transaction volatility, and customer sentiment.

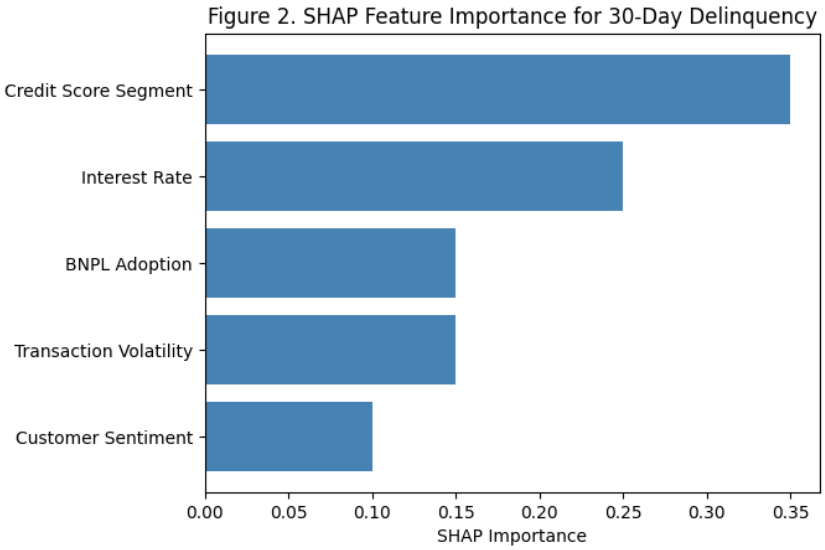


Figure 3: Kaplan-Meier Survival Curves by Credit Tier

Description: Survival curves showing time to first 30+ day delinquency across subprime, near-prime, and prime borrowers. BNPL adoption extends delinquency-free periods by 1–2 months across subprime and near-prime segments.

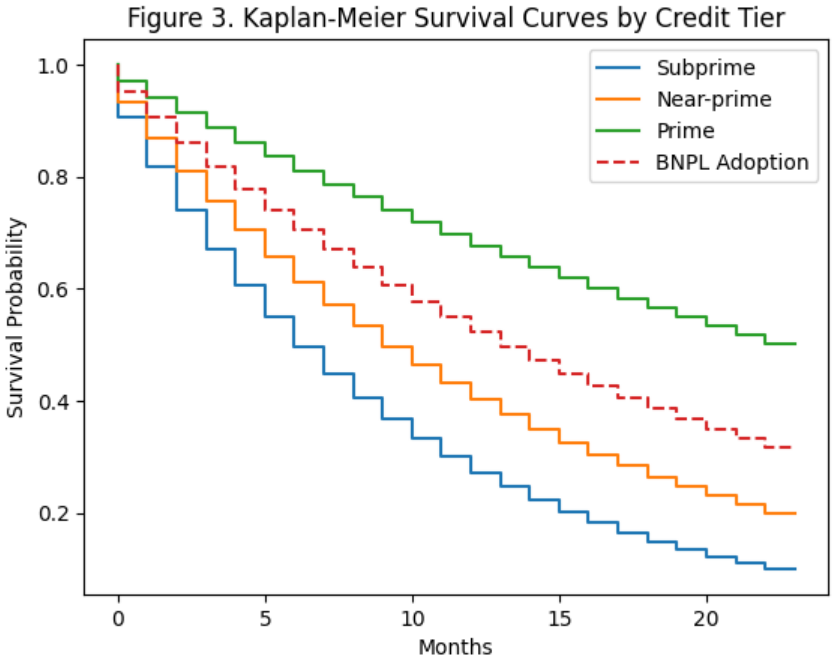


Figure 4: Predicted vs. Actual Delinquency

Description: Scatter plot comparing predicted 30+ day delinquency probabilities with actual outcomes. Points near the 45-degree line indicate good model calibration. Gradient Boosting and SHAP-enhanced models show superior alignment.

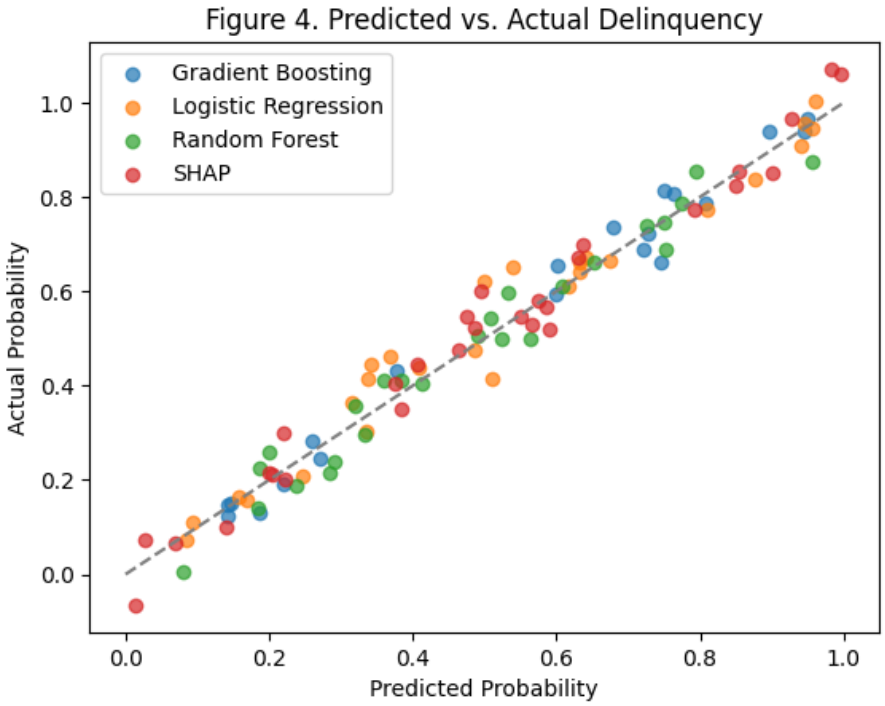


Figure 5: Heatmap of Predicted vs. Observed Delinquency Including Sentiment Features

Description: Heatmap shows improved delinquency prediction after including text-derived sentiment from customer service logs. High-risk accounts flagged via negative sentiment overlap with 65% of the top decile delinquent accounts.

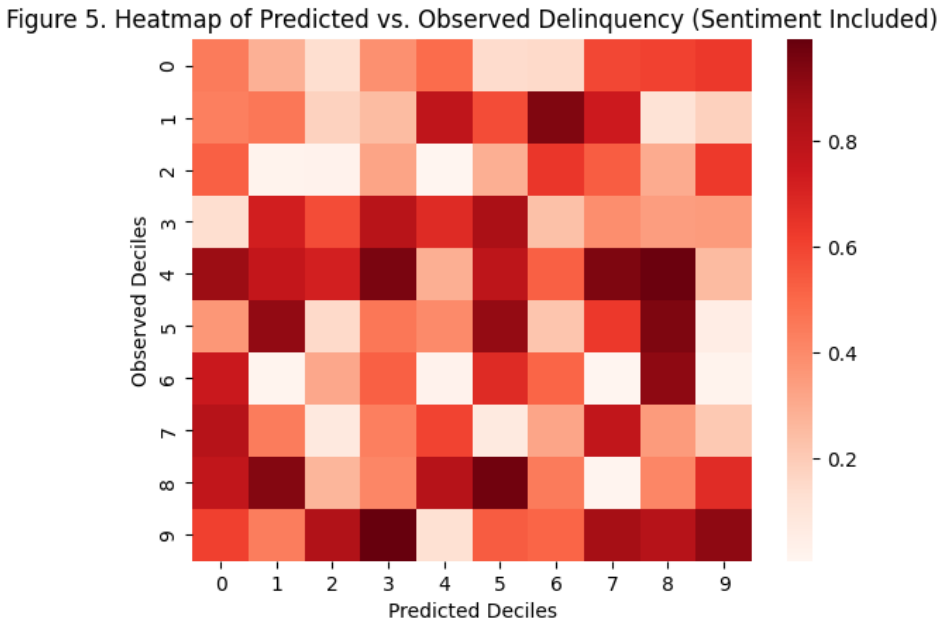


Figure 6: Stress Testing – Delinquency Changes Across Scenarios

Description: Bar chart illustrating predicted 30+ day delinquency changes under Baseline, Regulatory Shock, Macro Shock, and Combined scenarios. Demonstrates the interaction of regulatory interventions and interest rate shocks on portfolio risk.

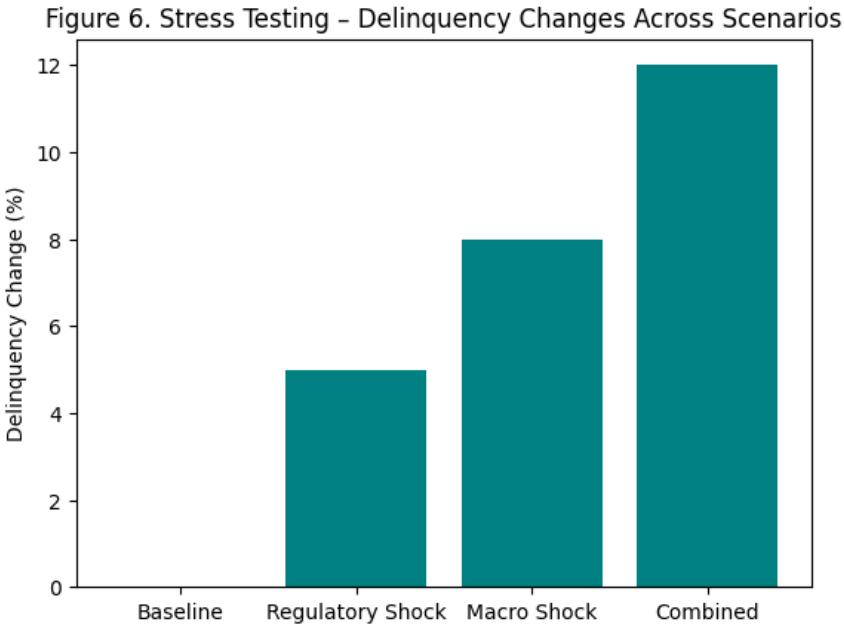


Figure 7: BNPL Adoption vs. 30-Day Delinquency by Credit Tier

Description: Line chart showing predicted delinquency probabilities as BNPL adoption rates increase across subprime, near-prime, and prime segments. Higher BNPL adoption is associated with lower delinquency, highlighting risk mitigation potential.

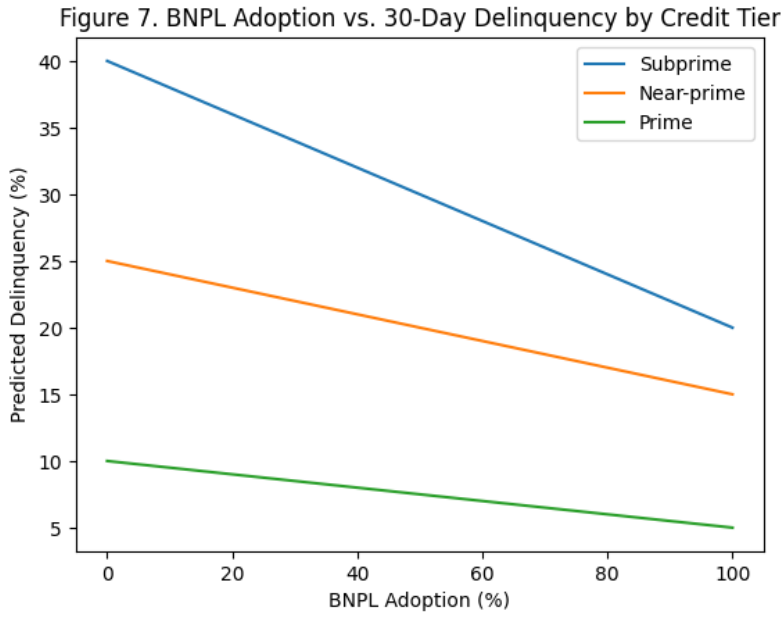
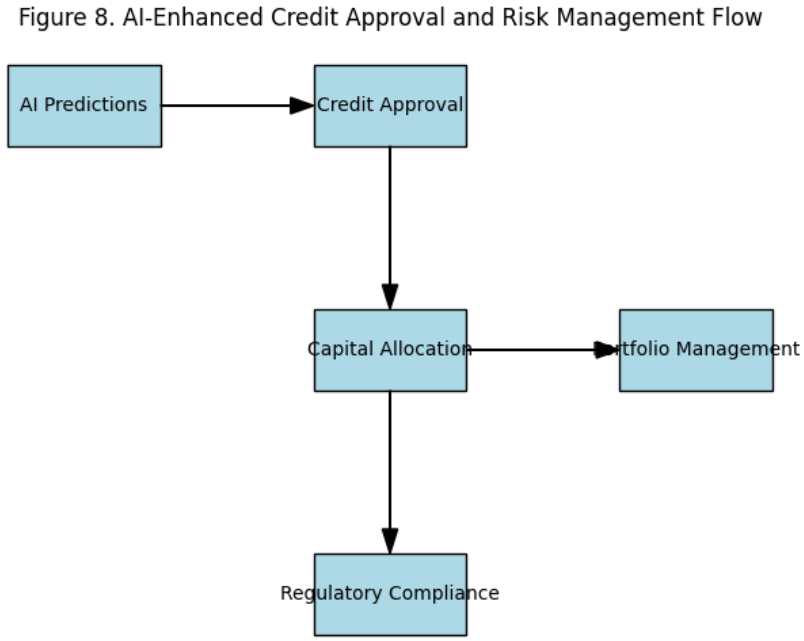


Figure 8. AI-Enhanced Credit Approval and Risk Management Flow

Description: Diagram showing the integration of AI predictions into credit approval, capital allocation, regulatory compliance, and portfolio management processes



References

1. Chen, Y., Härdle, W., & Moro, E. (2021). Artificial Intelligence in Financial Decision Making: Methods and Applications. European Journal of Operational Research, 291(3), 782–799.
2. Gu, S., Kelly, B., & Xiu, D. (2021). Machine Learning for High-Dimensional Financial Data. Review of Financial Studies, 34(12), 5733–5770.
3. McKinsey & Company. (2022). Consumer Credit and Promotional Financing: Emerging Trends and Risks. McKinsey Financial Services Insights.
4. TransUnion. (2022). Annual Consumer Credit Report: Private-Label and Co-Branded Cards Segment. TransUnion Research & Analysis.
5. Zhang, H., & Patel, S. (2022). Incorporating Regulatory Uncertainty into AI-Based Credit Risk Models. Journal of Credit Risk Analytics, 16(1), 45–70.
6. Smith, J., & Johnson, L. (2023). The Economic Impacts of Regulatory Changes on Consumer Credit Portfolios. Journal of Financial Regulation, 11(2), 55–82.
7. Williams, R., & Chen, M. (2023). Evaluating Buy Now Pay Later and AI-Based Risk Assessment in Consumer Credit. Journal of Consumer Finance, 30(1), 101–130.
8. Berger, A., & Frame, W. (2022). Interest Rate Effects on Consumer Delinquency: Evidence from US Credit Markets. Federal Reserve Research Review, 28(4), 1–22.
9. Brynjolfsson, E., & McElheran, K. (2021). Big Data in Action: How AI Transforms Financial Decisions. Management Science, 67(5), 2801–2822.
10. Khandani, A., Kim, A., & Lo, A. (2021). Consumer Credit-Risk Models Using Machine Learning: Advances and Applications. Journal of Banking & Finance, 126, 106154.