Modern American Journal of Engineering, Technology, and Innovation
ISSN(E): 3067-7939
Volume 2, Issue 1, January, 2026
Website: usajournals.org

ARCHITECTURE OF A DATA-DRIVEN PLATFORM FOR FORECASTING THE FINANCIAL SUSTAINABILITY OF US SMALL BUSINESSES

Ali Mustapa
Entrepreneur, USA

Abstract
This article is devoted to the development of a conceptual architecture for a data-driven platform for forecasting the financial sustainability of small businesses in the United States. The paper presents the historical evolution of bankruptcy probability assessment methods, starting with

The paper explores the range of approaches from classical financial ratio models, including the approach proposed by Edward Altman, to modern machine learning methods. The need for a shift toward integrating big data, streaming analytics, and explainable artificial intelligence tools in risk assessment is substantiated.

The institutional and regulatory environment in the United States is analyzed, including legal requirements for financial data processing and the role of government institutions such as the US Small Business Administration and the Federal Reserve System. A multi-layered platform architecture is proposed, including a data source, integration, and storage layer, an analytical layer, a service layer, and a user interface. Particular attention is paid to scalability, security, model interpretability, and the implementation of MLOps practices.

The results of the study demonstrate that the use of a comprehensive data-driven architecture can improve the accuracy of financial risk forecasting, ensure adaptability to changing macroeconomic conditions, and create a foundation for sustainable small business development in the US economy.

Modern American Journal of Engineering, Technology, and Innovation
ISSN(E): 3067-7939
Volume 2, Issue 1, January, 2026
Website: usajournals.org

Keywords: Financial stability, US small business, bankruptcy prediction, data-driven platform, machine learning, big data, risk management, credit scoring, MLOps, US institutional environment.

Introduction

Scientific Novelty. The scientific novelty of this study lies in the development of an integrated architectural model for a data-driven platform focused on forecasting the financial sustainability of US small businesses, taking into account institutional, regulatory, and technological factors. Unlike traditional approaches based primarily on static analysis of financial statements and classical models, including the model developed by Edward Altman, this paper proposes a multi-layered architecture combining data lake infrastructure, streaming data processing, and machine learning methods with explainable artificial intelligence mechanisms.

The innovation also lies in the integration of alternative small business data sources, including transaction flows, behavioral characteristics, and regional macroeconomic indicators, into a unified analytical environment. The need to implement MLOps practices to ensure continuous model quality monitoring and manage data drift in conditions of high market volatility is substantiated. Additionally, principles for the architectural adaptation of the platform to the requirements of US financial regulations and information security standards are developed.

Research Objective. The objective of the study is to develop a conceptual and technologically sound architecture for a data-driven platform for forecasting the financial stability of US small businesses, ensuring high-accuracy risk assessment, scalability for processing large volumes of data, interpretability of modeling results, and compliance with regulatory requirements.

Introduction

The development of methods for forecasting the financial stability of enterprises is closely linked to the evolution of financial science, statistics, and computing technologies. As early as the early twentieth century, economists attempted to

Modern American Journal of Engineering, Technology, and Innovation
ISSN(E): 3067-7939
Volume 2, Issue 1, January, 2026
Website: usajournals.org

formalize indicators of impending bankruptcy based on accounting ratios. In the 1930s, the first empirical studies comparing the financial indicators of successful and failed companies appeared in the United States. A significant breakthrough occurred in 1968, when Edward Altman developed the Z-score model, based on multiple discriminant analysis. This model became a classic tool for assessing the probability of bankruptcy and set the standard for subsequent research.

In subsequent decades, advances in econometrics, probability theory, and computer science enabled the application of logistic regression, neural networks, and machine learning methods to credit scoring and default risk assessment. The advent of big data and cloud computing in the early twenty-first century transformed the approach to financial stability analysis. Instead of working with limited financial statement samples, it became possible to use transaction data, customer behavioral characteristics, macroeconomic indicators, and industry metrics in near real time.

This issue has become particularly pressing for small businesses in the United States. Small businesses account for a significant share of employment and gross domestic product. In a highly competitive environment, with market volatility and cyclical economic fluctuations, the sustainability of small businesses is becoming a critical factor in macroeconomic stability. The 2020 pandemic and the associated government support measures implemented through programs administered by the US Small Business Administration have demonstrated the need for rapid and accurate forecasting of financial risks at the company level.

A modern data-driven platform architecture for predicting small business financial stability must incorporate both historical risk assessment methods and the latest advances in big data processing, distributed computing, and artificial intelligence. The purpose of this article is to develop a conceptual architecture for such a platform, taking into account the specifics of the US institutional environment, data availability, and requirements for scalability, security, and interpretability of models.

Theoretical Foundations for Forecasting Financial Stability. In scientific literature, a company's financial stability is defined as its ability to maintain solvency and deliver positive financial results over the long term, despite

Modern American Journal of Engineering, Technology, and Innovation
ISSN(E): 3067-7939
Volume 2, Issue 1, January, 2026
Website: usajournals.org

exposure to external and internal risk factors. In the context of small businesses, this indicator depends on capital structure, liquidity, operating margins, access to credit, as well as industry and regional economic conditions.

Classic bankruptcy prediction models are based on financial ratios. These include current liquidity ratios, debt-to-equity ratios, asset turnover, and profitability. The Altman model, mentioned earlier, integrated several ratios into a single aggregate index. Logit models based on a probabilistic interpretation of the default event were later proposed.

In today's world, the development of digital technologies is shaping new organizational management patterns, where data plays a key strategic role. Data-driven approaches offer the opportunity to significantly improve forecasting accuracy, optimize business processes, personalize customer interactions, and create sustainable competitive advantages. For companies operating in a dynamically changing external environment and with limited resources, the implementation of data-driven approaches is particularly important [1].

With the development of machine learning, gradient boosting, random forests, support vector machines, and deep neural networks have come to be used in financial stability forecasting [2]. Their advantage lies in their ability to identify nonlinear relationships and account for complex interactions between features. However, with increasing model complexity, the problem of interpretability increases, which is particularly important in financial regulation.

For small businesses in the United States, it is necessary to consider the specifics of their legal forms, tax regimes, and credit histories. Many small businesses have limited financial reporting and a short operating history [3]. This requires the integration of alternative data sources, including bank transactions, payment system data, online activity, and regional economic indicators.

US institutional and regulatory environment. The platform architecture must comply with US legislation on personal data and financial information protection. Specifically, it must comply with the Gramm Leach Bliley Act and the Fair Credit Reporting Act, which regulate the processing of financial data and the use of credit information. Furthermore, state requirements, including those in California, must be taken into account when handling entrepreneurs' personal data.

Modern American Journal of Engineering, Technology, and Innovation
ISSN(E): 3067-7939
Volume 2, Issue 1, January, 2026
Website: usajournals.org

Regulatory bodies such as the Securities and Exchange Commission and the Federal Reserve System shape macroprudential policy and establish disclosure and risk management requirements. While small businesses are not always directly regulated by these institutions, credit institutions and investors using forecasting platforms are required to comply with relevant reporting and risk management standards [4].

Government support for small businesses, implemented through the US Small Business Administration, creates additional datasets on grants, loans, and guarantees. Integrating this data into the analytics platform improves forecasting accuracy by taking into account the level of government support and debt burden.

Conceptual architecture of a data-driven platform. It is advisable to build the platform architecture on the principles of modularity, scalability, and fault tolerance. Generally, the system can be represented as several functional layers: a data source layer, a data integration and processing layer, an analytics layer, a services layer, and a user interface.

Table 1 - Functional layers of the data-driven platform for forecasting the financial stability of US small businesses

Functional platform layer	Main characteristics and tasks
Data source layer	Integration of financial statements, transaction data, credit histories, industry statistics and macroeconomic indicators
Data integration and storage layer	Implementation of ETL and ELT processes, data cleansing and transformation, creation of a centralized warehouse and data lake
Analytical layer	Feature generation, machine learning model training, default probability estimation, results interpretation
MLOps and monitoring layer	Model quality control, version control, data drift monitoring, and automated retraining
Service layer and API	Providing scoring assessments, scenario analysis and analytical reports to external and internal users
User interface	Visualization of key risk indicators, creation of dashboards and provision of differentiated user access

Modern American Journal of Engineering, Technology, and Innovation
ISSN(E): 3067-7939
Volume 2, Issue 1, January, 2026
Website: usajournals.org

The table provides a structured description of the key functional layers of the data-driven platform architecture. It reflects their primary functions and role in ensuring forecasting accuracy, data processing scalability, and compliance with US regulatory requirements.

Data source layer. This layer includes internal and external sources of information. Internal sources include the company's financial statements, management accounting data, bank statements, and information on accounts receivable and accounts payable. External sources include credit bureaus, industry databases, macroeconomic indicators, regional employment data, and consumer activity [5].
For small businesses in the US, credit reports generated by private agencies and tax filings are important data sources. Open data from government portals, including industry-specific and regional statistics, can also be used.

Data Integration and Storage Layer. This layer implements ETL and ELT processes that extract, cleanse, transform, and load data into a centralized repository. For large volumes of data, it makes sense to use a data lake architecture in combination with a structured data warehouse. A data lake allows for the storage of both structured and unstructured data, including text documents and transaction logs.
To ensure scalability and fault tolerance, it is recommended to use distributed file systems and cloud infrastructures. Streaming data can be processed using stream processing technologies, allowing forecasts to be updated in near real time.

Analytical layer. The analytical layer includes tools for feature generation, model training, and validation. Feature engineering plays a key role in improving forecast accuracy. In addition to traditional financial ratios, dynamic features are generated that reflect revenue trends, seasonality, cash flow volatility, and changes in the liability structure.
Prediction models may include logistic regression for basic interpretable risk assessment, gradient boosting for improved accuracy, and neural networks for

analyzing complex patterns. A key element is a model quality monitoring system, including AUC, precision, recall, and probability calibration metrics.
To ensure interpretability, it is recommended to use explainable artificial intelligence methods, such as SHAP values. This allows credit analysts to understand the contribution of individual factors to the final forecast and justify their decisions.

Services and API layer. This layer provides analytical results to external and internal users. Through the API, credit institutions, investment funds, and government agencies can obtain predictive financial stability assessments for specific companies. Services may include scoring reports, risk ranking, and scenario analysis.
Scenario analysis allows one to model the impact of changes in interest rates, inflation, or industry demand on a company's financial stability. This is especially relevant in conditions of macroeconomic uncertainty.

User interface and visualization. The platform interface should provide intuitive visualization of key indicators. Dashboards can include default probability dynamics, risk factor structure, comparative analysis with industry peers, and risk mitigation recommendations. Different levels of access and information presentation should be implemented for different user categories, including analysts, managers, and regulators.

Ensuring security and confidentiality. Given the sensitivity of financial data, it is necessary to implement a multi-layered information security system. This includes data encryption during transmission and storage, user authentication and authorization, and access auditing. Using zero-trust architecture principles increases the system's resilience to cyber threats.
Additionally, mechanisms for anonymization and pseudonymization of data should be provided when using it for research purposes. This is especially important when complying with US federal and state law.

Modern American Journal of Engineering, Technology, and Innovation
ISSN(E): 3067-7939
Volume 2, Issue 1, January, 2026
Website: usajournals.org

Platform Performance Assessment: The ability to collect, store, and analyze large amounts of data allows businesses to improve product quality and tailor products to consumers, increasing their value [6]. The platform's effectiveness is determined not only by the accuracy of forecasts but also by the overall economic impact for all user categories [7]. For credit institutions, this translates into a reduced share of problem assets, improved loan portfolio quality, and optimized provisioning processes in accordance with regulatory requirements, including the Federal Reserve System. A more accurate assessment of the probability of default allows for more accurate calculations of expected credit losses, reducing over-provisioning while simultaneously minimizing under-provisioning, which positively impacts capital adequacy ratios and return on assets.

For investment funds and fintech companies, the platform creates a tool for quantitatively assessing counterparty and partner risk, which improves the validity of investment decisions and reduces information asymmetries. For companies themselves, the platform facilitates the early detection of financial imbalances, including deteriorating liquidity, increasing debt burden, and declining operating margins. This enables timely management decisions, such as restructuring liabilities, adjusting the business model, or optimizing costs.

Performance evaluation should include both quantitative and qualitative indicators. Quantitative indicators include a reduction in the portfolio default rate, increased forecast accuracy measured through AUC and Brier score metrics, reduced application processing time, and improved decision-making speed. Qualitative indicators include increased risk management transparency, improved internal controls, and increased trust from partners and regulators.

Continuously training models on new data is critical, as the ability to process massive amounts of information, identify non-obvious patterns, and provide informed recommendations regarding optimal decisions is key to improving the speed and quality of financial flow management [8]. Given the highly volatile US macroeconomic environment and structural changes in the small business sector, models are susceptible to conceptual drift and changes in feature distribution. Implementing MLOps practices automates model updating, data drift monitoring, and version control. This architecture includes the implementation of continuous

Modern American Journal of Engineering, Technology, and Innovation
ISSN(E): 3067-7939
Volume 2, Issue 1, January, 2026
Website: usajournals.org

integration and continuous deployment pipelines for models, automated testing on validation samples, and regular hyperparameter reassessment.

Additionally, the system implements feature stability monitoring, input data quality control, and model decision auditing using explainable artificial intelligence tools. This not only maintains high predictive accuracy but also ensures compliance with regulatory transparency requirements and internal corporate risk management standards. Together, these elements form a sustainable framework for managing the platform's performance and ensure its long-term economic viability.

Table 2 - Performance indicators for the data-driven financial stability forecasting platform

Group of indicators	Specific evaluation metrics	Practical effect for users
Forecasting accuracy	AUC, Brier score, precision, recall, probability calibration level	Improving the quality of risk assessment and reducing the likelihood of erroneous decisions
Portfolio credit risk	The share of problem assets, the default rate, and expected credit losses	Reducing losses and increasing the stability of the loan portfolio
Financial stability of enterprises	Dynamics of liquidity ratios, debt burden and profitability	Early detection of financial imbalances and strategy adjustments
Operational efficiency	Application processing time, decision-making speed, level of automation	Reducing transaction costs and accelerating business processes
Stability of models	Data drift rates, overfitting rate, feature stability	Maintaining forecasts up-to-date in a changing market environment
Regulatory compliance	Availability of interpretability, completeness of decision audit, transparency of algorithms	Reducing regulatory risks and increasing trust from regulatory authorities

Modern American Journal of Engineering, Technology, and Innovation
ISSN(E): 3067-7939
Volume 2, Issue 1, January, 2026
Website: usajournals.org

Table 2 systematizes the key areas of performance evaluation for a data-driven platform for forecasting the financial stability of small businesses and demonstrates the relationship between the measured metrics and their practical significance. The table is structured within three analytical blocks: a group of indicators, quantitative assessment metrics, and the practical impact for users. The first block reflects the methodological basis for measuring performance, including model accuracy, credit risk level, operational efficiency, and the resilience of algorithms to data drift. The second block specifies the metrics used, such as AUC, Brier score, share of non-performing assets, expected credit losses, and feature stability indicators. The third block demonstrates the economic and managerial significance of the obtained results for credit institutions and small businesses.

Thus, the table provides a comprehensive understanding of the multi-level nature of the platform's performance, encompassing financial results, technological reliability, model interpretability, and compliance with regulatory requirements.

Conclusion

The architecture of the data-driven platform for forecasting the financial stability of US small businesses is a complex, multi-layered system [9] integrating diverse data sources, modern machine learning methods, and visualization tools. Taking into account the historical development of bankruptcy models, the specifics of the US institutional environment, and data security requirements, it enables the creation of a sustainable and scalable infrastructure.

In the context of increasing digitalization of the economy and intensifying competition, the role of such platforms will only increase [10]. Their implementation contributes to increased transparency of financial risks, strengthens trust between market participants, and promotes the sustainable development of small businesses as a key segment of the US economy.

References

1. Chernov Igor Evgenievich Implementation of data-driven approaches in Russian companies: cases, foreign maturity models and a conceptual Russian maturity model // Financial markets and banks. 2025. No. 10. URL:

https://cyberleninka.ru/article/n/vnedrenie-data-driven-podhodov-v-rossiyskih-kompaniyah-keysy-zarubezhnye-modeli-zrelosti-i-kontseptualnaya-rossiyskaya-model

2. Vladimir Viktorovich Kudryavtsev, IMPROVING ORGANIZATIONAL MANAGEMENT IN THE CONTEXT OF DIGITAL TRANSFORMATION // Human Progress. 2025. No. 2. URL: https://cyberleninka.ru/article/n/sovershenstvovanie-upravleniya-organizatsiey-v-kontekste-tsifrovoy-transformatsii
3. Wang Zhong, "The Impact of Digital Transformation on Organizational Performance: A Comparative Study of Small and Large Enterprises," Human Progress, 2025, no. 3, available at: https://cyberleninka.ru/article/n/vliyanie-tsifrovoy-transformatsii-na-effektivnost-rabota-organizatsii-sravnitelnoe-issledovanie-malyh-i-krupnyh-predpriyatiy
4. Seryan G.N., Adamyan A.V. TRANSFORMATION OF CORPORATE FINANCIAL STRATEGIES IN THE ERA OF DIGITAL ECOSYSTEMS // Professional Bulletin: Economics and Management. 2025. No. 1. URL: https://cyberleninka.ru/article/n/transformation-of-corporate-financial-strategies-in-the-era-of-digital-ecosystems
5. P. Yu. Motorygin, A. D. Vetrova, "Using Big Data for Forecasting Financial Markets," Informatics. Economics. Management, 2024, no. 4, available at: https://cyberleninka.ru/article/n/primenenie-bolshih-dannyh-dlya-prognozirovaniya-finansovyh-rynkov
6. Makarov M. Yu. INFLUENCE OF ARTIFICIAL INTELLIGENCE ON LABOR PRODUCTIVITY // Economics and Management. 2020. No. 5 (175). URL: https://cyberleninka.ru/article/n/vliyanie-iskusstvennogo-intellekta-na-proizvoditelnost-truda
7. Ivanovsky Boris Georgievich ECONOMIC EFFECTS FROM THE IMPLEMENTATION OF ARTIFICIAL INTELLIGENCE TECHNOLOGIES // Social Innovations and Social Sciences. 2021. No. 2 (4). URL: https://cyberleninka.ru/article/n/ekonomicheskie-effekty-ot-vnedreniya-tehnologiy-iskusstvennogo-intellekta
8. Pershin, Ya. R. Digitalization of financial management of an organization: integration of distributed technologies

registry and artificial intelligence in financial management processes / Ya. R. Pershin, E. S. Budkina //
Bulletin of Eurasian Science. — 2024. — Vol. 16. — No. 6. — URL: https://esj.today/PDF/105ECVN624.pdf
9. I. D. Teplova, I. A. Zhuravlev, A. A. Karmanova. Data-driven solutions for creating a comfortable and safe environment on city streets. Ecology of urbanized territories. 2022. No. 1. URL: https://cyberleninka.ru/article/n/data-driven-resheniya-dlya-organizatsii-komfortnoy-i-bezopasnoy-sredy-gorodskih-ulits
10. Evgeniya Mikhailovna Izotova, Digital Solutions for Event Personalization: Methodology, Metrics, and Practice Applications // Universum: Technical Sciences. 2025. No. 1 (130). Available at: https://cyberleninka.ru/article/n/tsifrovye-resheniya-dlya-personalizatsii-sobytiy-metodologiya-metriki-i-prakticheskoe-primenenie.