Development Economics of China
2025, VOL. 9, NO. 10, 45-49
DOI: 10.47297/wspdecWSP2515-797309.20250910

Research on the Application of Big Data in Financial Forecasting and Budget Management

Suosheng Ma[1], Yuanheng Li[1,2]

[1]Lanzhou Vocational and Technical College, Lanzhou, Gansu, 730070, China

[2]University of Wolverhampton, England, WV11AD, UK

ABSTRACT

In the context of the digital economy era, the scale of data generated from corporate financial activities is experiencing exponential growth. Traditional financial forecasting and budget management models are thus facing an urgent need for systemic transformation. Integrating and processing massive heterogeneous data sources to effectively enhance prediction accuracy in dynamic market environments, as well as achieving real-time responses to budget resource allocation issues, have become critical bottlenecks constraining the effectiveness of financial decision-making. The distributed computing capabilities and intelligent algorithms of big data technology offer a viable technical path for building a financial management system that can adapt to complex business environments. This paper focuses on the integration mechanism between big data and financial management, delving into the evolution logic of financial decision-making paradigms under technological empowerment, thereby providing theoretical support for the digital transformation of enterprises.

KEYWORDS

Big data; Financial forecasting; Budget management; Applied research

1 Introduction

The current highly uncertain business environment and the fragmented nature of data resources have made it extremely difficult for traditional financial models to accurately capture market fluctuation patterns. The contradiction between rigid budget execution constraints and dynamic operational needs is also becoming increasingly prominent. By constructing a comprehensive data collection network and an intelligent analysis platform, big data technology can overcome the limitations of traditional sampling analysis, achieving deep correlation and dynamic simulation effects in financial elements. This paper employs an analytical method that combines systems dynamics and machine learning to reveal the value transformation mechanism of multi-source heterogeneous data in financial modeling, exploring the construction path of a data-driven budget control system, providing innovative ideas for enhancing corporate financial agility.

2 Problems of Big Data in Financial Forecasting and Budget Management

2.2 Data Integration is Difficult

The complexity of integrating corporate financial data stems from the long-term parallel operation of multi-source heterogeneous systems. Different business modules have developed independent data storage architectures and coding rules over their historical development. There are significant differences in data structure and update frequency between customer behavior logs on the sales side and material flow information on the supply chain side. When extracting cross-system data, the finance department often needs to handle redundant field cleaning and time series alignment issues. The phased nature of the electronic process of original vouchers has led to incomplete unification of metadata standards for paper ticket images and electronic ledgers. Non-structured text and semi-structured reports embedded in fragmented data lakes are difficult for traditional relational databases to parse directly. The uneven degree of digital transformation across business units results in lower-than-expected integration efficiency between real-time data streams and offline batch processing data.

2.2 The Prediction Model is Not Adaptive Enough

The existing big data models have obvious limitations in parsing complex business scenarios, especially when dealing with nonlinear correlations and dynamic market environments, making it difficult to accurately capture the interactions between key variables. Traditional algorithms rely on static features of historical data, lacking the ability to dynamically adjust when faced with emerging industries or sudden economic fluctuations, leading to discrepancies between predicted outcomes and actual operational needs. Some models overly simplify industry characteristics and regional differences, ignoring the long-term impact of corporate strategic adjustments on financial indicators, causing budgeting and resource allocation to deviate from expected goals. Unstructured information such as policy directions and supply chain disruptions in the external environment has not been fully integrated, weakening the timeliness of risk warnings [1]. On the technical side, issues include vague data cleaning standards and low efficiency in integrating multi-source heterogeneous data, preventing real-time data from critical business nodes from being effectively converted into decision-making support, indirectly affecting the iterative optimization process of predictive models.

2.3 Limited Real-time Analysis and Dynamic Adjustment Capabilities

The current financial management system faces the challenge of balancing time sensitivity and computational load when processing high-frequency business data. Traditional batch processing models struggle to capture instantaneous value signals generated by market fluctuations. Incremental information produced through cross-departmental collaboration is prone to timestamp discrepancies during transmission. Historical benchmark data relied upon for budget adjustment decisions fails to timely reflect the impact of sudden events on the capital chain. The asynchronous update mechanism of heterogeneous data sources leads to lagging biases in input parameters for cash flow forecasting models. Deep learning algorithms find it difficult to achieve an ideal balance between training cycles and real-time inference efficiency. Static budget control rules lack adaptive mechanisms that embed dynamic risk weights. Differences in time resolution at various nodes in the data pipeline diminish the analytical value of multi-dimensional operating indicators. The edge computing demands arising from rapid iteration of business scenarios exceed the design capacity of the existing system architecture.

2.4 Data Security and Privacy Protection Risks

The process of collecting and storing massive amounts of data faces the risk of unauthorized access. Some companies, due to outdated technical architectures, fail to effectively isolate sensitive financial information from public data pools, exposing customer privacy or trade secrets to potential threats. In data sharing mechanisms, permission divisions are often vague, leading to overly open interfaces in cross-departmental collaborations. The lack of transparent tracking for data invocation records on third-party platforms exacerbates the complexity of information leaks. On the compliance front, existing legal frameworks have yet to detail regulations governing cross-border data flows. Differences in privacy protection standards across regions complicate the construction of financial forecasting models by multinational corporations, leading to compliance disputes. Technical vulnerabilities combined with human errors, such as outdated encryption algorithms or internal personnel inadvertently uploading critical budget parameters to cloud public environments, further amplify the likelihood of data tampering or misuse [2].

3 Application of Big Data in Financial Forecasting

3.1 Data Acquisition and Preprocessing

Financial forecasting involves making scientific estimates and calculations of future financial activities and outcomes based on historical data of financial activities, taking into account current requirements and conditions. During the data collection and preprocessing phase of financial forecasting, companies typically establish cross-system data pipelines to integrate raw transaction records scattered across ERP, CRM, and supply chain management platforms. The data team must design unified cleaning rules to eliminate issues such as duplicate sales order entries or inconsistent inventory data points. After cleaning, the dataset undergoes standardized mapping and transformation, converting indicators like raw material consumption and customer credit periods from different units of measurement into comparable dimensions for structured input in subsequent modeling. To address the unstructured nature of external macroeconomic data and industry intelligence, a natural language processing engine is deployed at the technical architecture layer to extract keywords related to economic policy changes and competitive market dynamics. Through semantic analysis, these are converted into quantitative parameters and integrated into the prediction variable pool. The preprocessing phase also builds a dynamic feature library, optimizing variable combinations based on historical prediction accuracy feedback, removing environmental factors with correlation below the threshold to main business revenue, and retaining strong explanatory indicators such as equipment utilization rates and supplier delivery punctuality. The data quality monitoring module continuously scans for abnormal value fluctuations; when it detects that sales data in a certain area deviates from seasonal patterns for three consecutive days, it automatically triggers an manual review process to prevent erroneous data from entering the model training phase.

3.2 Construction of Financial Forecasting Model

The construction of financial forecasting models relies on feature engineering for the in-depth analysis of multi-dimensional business variables. The model architecture needs to transform the associated features between customer payment cycles and supplier settlement terms into quantifiable time series factors. When integrating neural networks with ensemble learning techniques, the algorithm framework must balance the autocorrelation characteristics of time series data with the nonlinear impact of external shocks. Data scientists must address the co-integration relationship between historical account period fluctuations and industry sentiment indices during parameter tuning to avoid overfitting. In the model validation phase, adversarial testing mechanisms are introduced to test the robustness of accounts receivable forecasts against sudden changes in interest rate policies. The explainability module design should map abstract feature weights to credit risk transmission paths that are understandable to business personnel. Continuous learning mechanisms maintain the model's predictive performance under new tax policy environments through dynamically updating the training set's time window. Strategies for allocating heterogeneous computing resources directly affect the parallel computational efficiency of feature extraction from transactional data at the scale of millions.

3.3 Multi-dimensional Data Analysis and Prediction

Companies need to cross-model historical financial data with market trends, customer behavior, and operational costs to capture potential patterns. The finance department breaks down core indicators such as cash flow and inventory turnover based on business scenarios, forming a multi-dimensional analysis matrix. The analytical platform uses clustering algorithms to identify data coupling relationships between different business units and dynamically adjusts weight allocation strategies. During the training of predictive models, it is crucial to address lag effects and variable collinearity in time series data. Algorithm engineers use feature interaction techniques to uncover nonlinear associations between sales cycles and supply chain efficiency, while introducing sliding window mechanisms to smooth out seasonal fluctuations, reducing the risk of overfitting [3]. The real-time data update module must be embedded at key nodes in the business process to capture the latest market signals. A distributed computing framework can handle multi-source heterogeneous data streams in parallel and synchronize the updating of prediction parameters. Decision-makers continuously refine budget allocation plans based on dynamic analysis results, forming a feedback loop. The data visualization interface transforms complex analytical conclusions into actionable recommendations, helping business lines quickly respond to market changes.

3.4 Risk Assessment and Uncertainty Analysis

The risk assessment framework needs to quantify the intrinsic coupling mechanism between macroeconomic fluctuations and corporate cash flows. The dynamic monitoring of credit risk exposure relies on cross-verification of transaction patterns and public opinion data from upstream and downstream enterprises in the supply chain. In the process of uncertainty modeling, it is necessary to distinguish the transmission thresholds between systemic market risks and non-operational financial risks. The design of stress test scenarios must cover the combined impact effects of raw material price volatility and exchange rate fluctuations. Probability distribution fitting techniques are used to analyze the hidden tail risk characteristics in historical bad debt rate data. The visualization of risk transmission paths requires converting the association pattern between supplier rating changes and accounts receivable turnover into a directed graph structure. The Monte Carlo simulation engine continuously adjusts the joint distribution parameters of customer order cancellation probabilities and inventory turnover cycles during iterations. The dynamic calibration mechanism for risk preference parameters needs to integrate the semantic parsing results of board strategic resolution texts and industry risk benchmark values. Orthogonalization of heterogeneous risk factors helps to strip away the cumulative influence of policy regulatory signals on operating cash flow forecasts.

4 Application of Big Data in Budget Management

4.1 Real-time Monitoring and Dynamic Adjustment of Budget Implementation

Budget management refers to the strategic goal-oriented activities of enterprises, which involve comprehensive forecasting and planning of business operations and corresponding financial outcomes over a certain future period. It scientifically and reasonably allocates both financial and non-financial resources, monitors and analyzes the execution process, evaluates and provides feedback on the results, guiding improvements and adjustments in business activities to ultimately drive the achievement of corporate strategic goals. Enterprises use IoT devices and business system interfaces to automatically obtain raw data streams from production, sales, inventory, and other stages, breaking down information silos between departments to form a unified data warehouse. This infrastructure enables budget execution progress to be updated at a minute-level frequency. Based on real-time operational data, the finance department develops early warning models, identifying budget variances through preset cost thresholds and cash flow fluctuation ranges. When the system detects that raw material consumption rates exceed the quarterly budget by fifteen percent or that marketing expenses surge significantly in a single day, it automatically triggers a tiered early warning mechanism to send verification instructions to responsible departments [4]. Management initiates the budget adjustment process based on deviation analysis reports generated by the system, focusing on reviewing the business rationality and fund utilization efficiency of overspent items. For research and development projects or equipment procurement needs that require additional budgeting, the finance team collaborates with supply chain and production departments to estimate the adjusted funding gap and return cycle. After cross-departmental discussions, a revised plan is submitted for approval by the decision-making level.

4.2 Budget Performance Evaluation and Feedback Mechanism

The design of the budget performance evaluation system requires establishing a dynamic mapping relationship between business activities and financial indicators. The analysis engine must identify the nonlinear correlation characteristics between fluctuations in marketing expenses and the effectiveness of sales channel expansion when parsing departmental budget execution data. A multi-dimensional evaluation system will conduct cross-period matching analysis on the technology conversion rate of R&D investment and product lifecycle cash flows. The optimization of the feedback mechanism demands real-time tracking of the lagged effects of budget consumption trajectories and changes in market share for strategic investment projects. The deviation diagnosis module generates attribution analysis charts by comparing the degree of deviation between budget preparation assumptions and actual operating environment variables. The intelligent correction algorithm needs to balance the inertia effect of historical execution deviations with changes in resource allocation priorities due to organizational changes when adjusting the next period's budget allocation coefficients. The knowledge base update mechanism continuously integrates experience decision-making patterns from budget adjustment approval records, encoding the risk preference of the budget committee into constraints to inject into the optimization model. The cross-year rolling forecast function enhances the precision of budget alignment through iterative corrections of technical depreciation rates and human cost growth rates. The adaptive calibration mechanism for abnormal fluctuation warning thresholds relies on the coordinated changes in supplier payment cycles and customer account periods for dynamic updates.

4.3 Resource Optimization and Cost Control

The company builds resource consumption characteristic models based on historical operational data accumulated from its business systems. It creates heat maps of resource flow for procurement, production, and warehousing processes to identify business nodes with high idle equipment rates or excessively long raw material turnover cycles. The finance team designs dynamic resource allocation plans in conjunction with process standards and market supply-demand fluctuations, aiming to reduce redundant inventory to a safe threshold while ensuring delivery schedules. In terms of cost control, the focus is on establishing a multi-dimensional expense attribution system. This upgrades the traditional departmental accounting model to a more refined one, dividing costs into units by product line, customer group, and project team. Machine learning algorithms are used to uncover hidden cost drivers, such as fuel loss due to flawed transportation route planning or increased manual intervention caused by unreasonable production line layouts. Based on these insights, differentiated cost control strategies are formulated. A real-time cost tracking module is embedded in the budget execution process. When actual expenditures deviate from preset standards, it automatically triggers root cause analysis, prompting the procurement department to reassess supplier quotes and coordinating with R&D teams to optimize product design, thereby shifting cost control from post-event accounting to pre-event prediction [5].

4.4 Intelligent Decision Support for Budget Management

The construction of an intelligent decision support system relies on the deep integration and processing capabilities of multi-source heterogeneous data. The real-time synchronization mechanism between business flows and cash flows aligns the execution progress of purchase orders with accounts payable forecasts in a spatial dimension. When handling unstructured meeting minutes, the algorithm architecture must extract potential impacts of changes in strategic priorities on budget allocation strategies. Knowledge graph technology transforms implicit decision-making logic from historical budget adjustment records into quantifiable association rule libraries. Reinforcement learning models need to balance short-term operational goal achievement rates with the diminishing marginal benefits of long-term R&D investment when simulating budget allocation schemes. A dynamic rule engine automatically generates boundary conditions for flexible budget adjustments by parsing the semantic characteristics of board risk preferences. A distributed computing framework maintains global optimal solution search capabilities while processing cross-regional subsidiary budget collaboration needs under data sovereignty isolation. The data governance module continuously cleans up vague descriptions in budget applications submitted by business departments, converting them into standardized resource requirement feature vectors. Digital twin technology constructs virtual operating environments to test the impact gradients of different budget reduction schemes on production line balance rates.

5 Conclusion

The underlying logic of financial forecasting and budget management has been restructured through big data. By leveraging the operation of data assets and the iteration of intelligent algorithms, a new management model has emerged where risks can be quantified and decisions traceable. In practical implementation, it is crucial to address the lack of data governance standards and the issue of algorithmic black boxes, establishing cross-departmental data sharing mechanisms and interpretive frameworks for models. For enterprises, it is recommended to build an integrated application system that includes data quality assessment, dynamic model calibration, and privacy protection technologies, facilitating the transformation of data insights into budget allocation efficiency. Future research could focus on the application of edge computing and federated learning in distributed financial scenarios, exploring the ethical boundaries and value balance mechanisms of human-machine collaborative decision-making.

Funding

Research Topic for 2024 at Lanzhou Vocational and Technical College: "Research on the Training Model for Financial Planning and Risk Management Abilities of Vocational Students in an Entrepreneurial Context"(NO: 2024-XY67)

About the Author

Suosheng Ma, male, Han ethnicity, from Lanzhou, Gansu Province. Postgraduate degree, professor. Research direction: financial management.

Yuanheng Li, male, Han ethnicity, from Lanzhou, Gansu Province. Doctoral candidate, lecturer. Research direction: business management.

References

[1] Chen Junhua. Financial Risk Management and Control of Enterprises under the Background of Big Data [J]. Vitality, 2024,42(24):106-108.

[2] Song Li. Research on Comprehensive Budget Management of Enterprises in the Era of Big Data [J]. Quality and Market, 2022, (07):73-75.

[3] Gaudi. Innovation and Transformation Analysis of State-Owned Enterprise Financial Management in the Era of Big Data [J]. Finance Guest, 2024, (04):90-92.

[4] Bai Ru Jun. Transformation of Enterprise Financial Management in the Big Data Environment [J]. China Foreign Investment, 2021, (04):56-57.

[5] Gaudi. Innovation and Transformation Analysis of State-Owned Enterprise Financial Management in the Era of Big Data [J]. Finance Guest, 2024, (04):90-92.