Accounting and Corporate Management (2025)
Clausius Scientific Press, Canada

DOI: 10.23977/acccm.2025.070401
ISSN 2523-5788 Vol. 7 Num. 4

The Role of Big Data Analytics in Corporate Finance Shared Decision Making

Qiao Yan

Henan Weijia Automobile Trading Group Co., LTD., Zhengzhou, Henan, China

Keywords: Big Data Analytics; Corporate Finance; Shared Decision Making; Role

Abstract: Digital technology innovation has reshaped the underlying logic of enterprise financial management, and the financial sharing model is gradually leaping from process integration to data-driven. The traditional decision-making mechanism is limited by information silos and experience dependence, making it difficult to cope with the dynamic risks in the complex market environment. Big data analytics opens a new path to solve the pain points of inefficient resource allocation and lagging risk identification by integrating heterogeneous data from multiple sources and constructing a quantifiable decision support framework. This paper focuses on the intersection of algorithm design and governance system, exploring how to combine unstructured data features with dynamic optimization models to form an intelligent financial hub with real-time response capability. The breakthrough point of the research is the establishment of a decision-making system covering the whole chain of fund deployment, risk warning and budgeting, and the validation results are of demonstrative significance in promoting the transformation of the financial sharing center into a strategic value unit.

1. Introduction

Enterprise financial management is undergoing a deep transformation from decentralized operations to shared services, and the complexity of decision-making is increasing exponentially with the intensification of global competition. Traditional financial modeling relies on static reports and manual experience, and there are systematic deviations when facing non-linear variables such as supply chain fluctuations and exchange rate risks. Big data technology not only broadens the dimension of data collection, but also reconfigures the spatial and temporal boundaries of decision-making analysis - real-time cash flow monitoring, textual public opinion analysis, supply chain mapping and other technological means, which enable financial sharing centers to have a panoramic insight capability. This paper starts from the data governance system, focuses on analyzing the feature engineering method under the heterogeneous data fusion framework, and proposes a hybrid algorithm model based on dynamic planning and machine learning. A closed-loop decision-making link is formed by designing a fund allocation optimizer, a risk prediction network and an intelligent budget engine. The experimental part verifies the advantages of the algorithm in terms of efficiency improvement and error control, and the research conclusions provide a landable technical solution for enterprises to build a resilient financial system [1].

2. Financial Shared Decision Making Data System Construction

2.1 Data Source Classification and Governance

The data source classification of the financial shared decision-making system needs to take into account the collaborative governance of structured and unstructured data. Structured data is centered on the Enterprise Resource Planning (ERP) system (Figure 1), covering accounts payable details, cash flow timing records and other financial transaction information, with standardized fields and time-stamped attributes providing the basis for fund liquidity analysis. The real-time synchronization of account balances and transaction flow data through the direct interface between banks and enterprises strengthens the granularity and timeliness of fund monitoring and supports the input requirements of dynamic deployment algorithms [2]. In the dimension of unstructured data, the semantics of the terms and conditions of the contract text, the dynamics of public opinion in the upstream and downstream of the supply chain, and the compliance constraints of the industry policy documents constitute the key supplement. Contract parsing relies on natural language processing technology to extract entity relationships such as payment terms and breach of contract responsibilities, while supply chain public opinion needs to be combined with sentiment analysis models to quantify market risk signals.

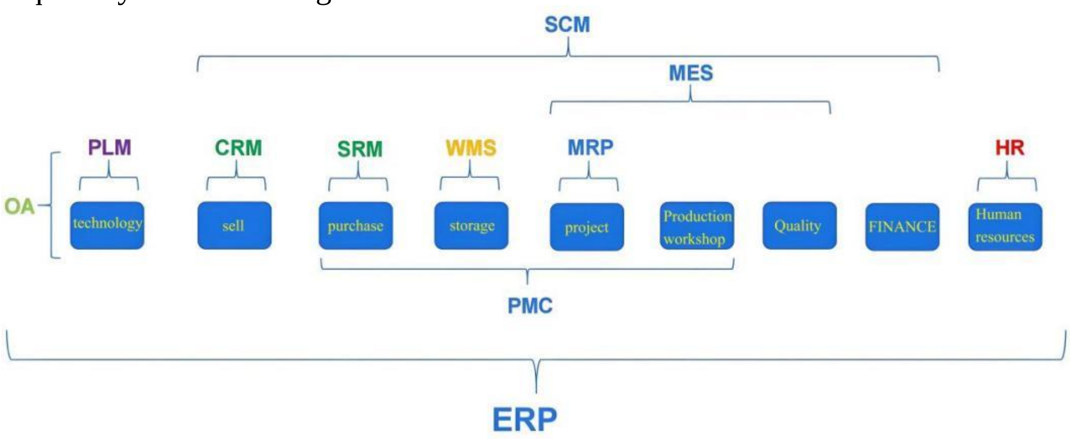

Figure 1 Enterprise Resource Planning (ERP) Systems

The technical difficulty of the governance system lies in balancing data quality and computational efficiency. The high-frequency updating of bank-enterprise direct data requires the design of a streaming processing framework to avoid the decision lag caused by batch processing; the noise filtering of supply chain public opinion relies on graph neural networks to mine key nodes in the propagation path to enhance the confidence of risk signals. The cross-validation mechanism of structured and unstructured data further optimizes the governance effect, for example, the deviation analysis of the contractual payment cycle and the actual payment record of ERP, which can reverse correct the entity extraction logic of the text parsing model [3]. The framework lays the technical foundation for enterprises to build a highly credible financial shared data base.

2.2 Characterization Project

2.2.1 Financial Risk Indicators

Characterization engineering of financial risk indicators aims to quantify the solvency and operational stability of a firm.The Z-Score model constructs bankruptcy early warning signals by linearly combining key financial ratios, which can be formulated as (1):

$$Z=\sum_{i=1}^{n} w_i X_i \quad (1)$$

Where X_i denotes standardized variables such as working capital ratio and retained earnings share, and w_i is the industry-calibrated weighting coefficient. Cash flow volatility is calculated using conditional variance model as in equation (2):

$$\sigma_t^2=\alpha_0+\alpha_1 \in_{t-1}^2+\beta\sigma_{t-1}^2 \quad (2)$$

The Z-Score model is used to capture the heteroskedasticity of cash flow time-series data, and exogenous variables are constructed to enhance the interpretability by combining with supply chain public opinion events [4].

In the technical implementation, the Z-Score model needs to solve the problem of cross-system data standardization, and the accounts payable turnover rate of ERP system needs to be aligned with the cash flow data of direct linkage between banks and enterprises in order to avoid the distortion of features caused by the mismatch of accounting cycles. Cash flow volatility feature extraction relies on sliding window statistics and frequency domain analysis, such as the use of wavelet transform to separate seasonal trends and sudden fluctuations. Unstructured data are supplemented with risk signals through semantic parsing, e.g., default clauses on account period in the contract text can be mapped as correction factors of Z-Score model.

Indicator robustness in dynamic environments is critical. Industry policy changes may alter the validity of Z-Score weighting coefficients, and an online learning mechanism needs to be introduced to update wi parameters. The cash flow volatility model needs to integrate causal inference techniques to identify the transmission path between supply chain disruption events and cash flow breaks [5].

2.2.2 Unstructured data characteristics

Feature engineering for unstructured data needs to address semantic understanding and risk quantification challenges. Risk feature extraction of contract text relies on topic modeling techniques, LDA (Latent Dirichlet Distribution) identifies potential topics by analyzing the co-occurrence probability of lexical items, as shown in Fig. 2, and combines with customized dictionaries to strengthen the semantic weight of key terms. Sentiment analysis of supply chain public opinion uses the pre-trained language model BERT, whose bidirectional attention mechanism captures the contextual associations of negative supplier events in the news text, outputs the sentiment polarity score and calculates the risk impact value, e.g., the attenuation effect of the negative report on the supplier's credit rating.

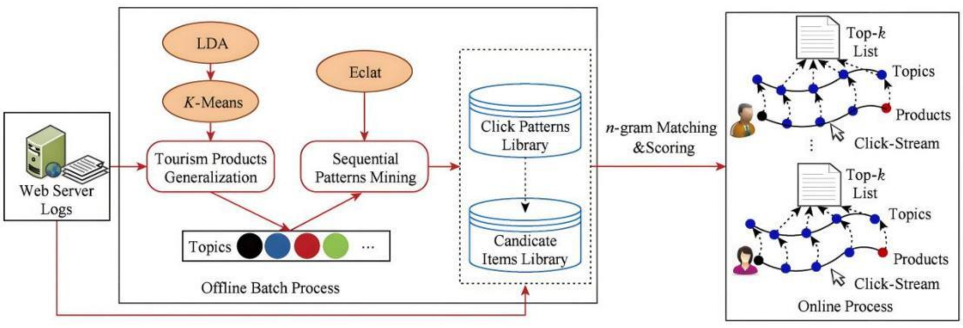

Figure 2 LDA theme model

In the technical implementation, the entity relationship of the contract text should be aligned with the structured financial data, and the contract risk labels generated by the LDA model should be mapped to the accounts payable records in the ERP system to verify the consistency of the performance status of the terms and conditions with the risk themes. Knowledge mapping technology is used to assist in the construction of the “supplier-contractual terms-opinion events” association network to identify the common characteristics of high-risk suppliers, and industry-specific corpus is introduced in the fine-tuning stage of the BERT model to improve the parsing accuracy of supply chain finance terms, as shown in Figure 3.

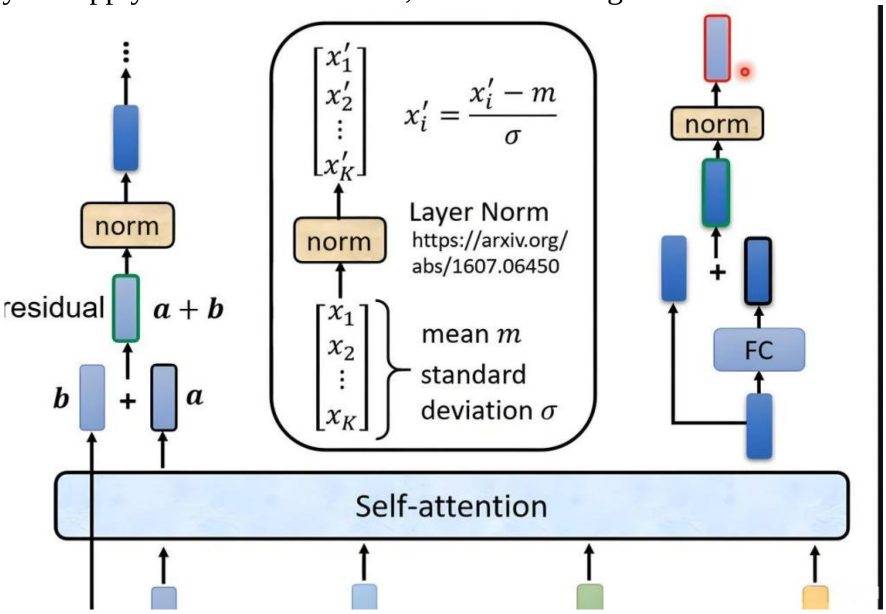

Figure 3 BERT model

The feature fusion session faces cross-modal data alignment challenges. Contractual risk themes need to establish dynamic correlation rules with financial metrics of the Z-Score model, such as the synergistic trend of the frequency of occurrence of specific risk keywords and the working capital ratio. Public opinion sentiment scores need to be time-series aligned with cash flow volatility features to capture the lagged effect of negative events on financial chain stress. Experiments show that the introduction of unstructured features can enhance the recall of the financial risk early warning model, especially in the scenarios of sudden changes in industry policies or sudden supply chain crises, which exhibit stronger robustness [6].

3. Core Algorithm Design

3.1 Funds dynamic deployment optimization algorithm

The core of the dynamic fund allocation algorithm is to solve the problem of balancing the utilization rate and risk of funds under the scenario of multi-subsidiary collaboration. The problem modeling takes the group capital pool as the object, and the objective function needs to synchronously optimize the minimization of idle capital cost and the maximization of investment return, and at the same time satisfy the minimum retained capital threshold required for subsidiary operation, the foreign exchange control policy of cross-border capital flow and other hard constraints. Among the constraints, the minimum threshold of retained funds of subsidiaries needs to be dynamically calibrated according to the historical cash flow volatility, and the foreign exchange control rules need to be embedded in the country-specific policy knowledge base to realize real-time compliance calibration.

The algorithm design adopts the improved genetic algorithm NSGA-II, whose non-dominated sorting mechanism and congestion comparison operator can effectively handle the Pareto frontier search for multi-objective optimization problems, as shown in Fig. 4. The technical realization needs to be deeply coupled with real-time data flow [7]. The cash flow time series data provided by the direct interface between banks and enterprises triggers the iterative updating of the algorithm, and the sliding window mechanism ensures that the fund allocation scheme is adapted to the fluctuation of market interest rates and exchange rate changes. Supply chain public opinion events are transformed into risk coefficients through the BERT sentiment analysis model, which dynamically adjusts the slackness of the minimum retained funds constraint of subsidiaries.

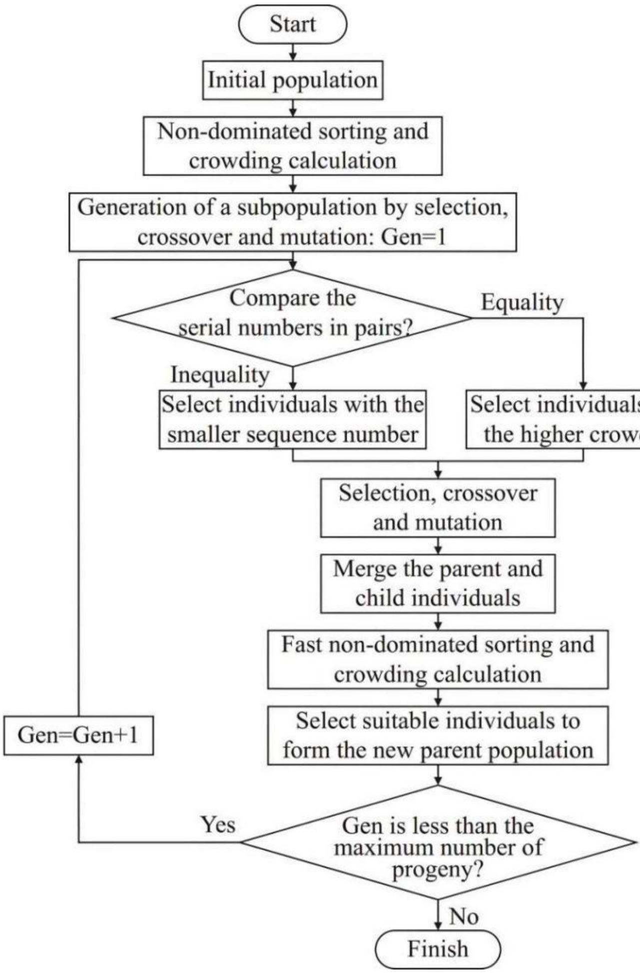

Figure 4 Improved genetic algorithm NSGA-II flow

3.2 Financial risk prediction algorithm

The financial risk prediction algorithm needs to integrate multimodal data to capture the early signals of financial chain breakage. The input layer integrates historical financial data, contractual risk features and supply chain public opinion indices, where contractual risk features are represented by a vector of "probability of default" keyword weights extracted from the LDA topic model, and the public opinion indices are quantified by the impact values of negative events output from the BERT sentiment analysis model. The model architecture is designed as a two-channel structure:

Graph Convolutional Network (GCN) modeling subsidiary affiliation transaction network, node features are defined as financial indicators of each subsidiary, neighbor matrix is constructed based on the transaction amount and frequency, and the propagation formula is (3), which is used to capture the risk diffusion path within the group.

$$H^{(l+1)} = \sigma(\tilde{D}^{-\frac{1}{2}}\tilde{A}\tilde{D}^{-\frac{1}{2}}H^{(l)}W^{(l)}) \quad (3)$$

The LSTM-Attention module (Fig. 5) handles time-series financial data and unexpected events, the time-step feature vector is composed of cash flow volatility, public opinion index splicing, and the attention weight is calculated as formula (4), which screens the key event nodes that affect the stability of the capital chain [8].

$$\alpha_t = soft\max(v^T \tanh(W_h h_t + W_s s)) \quad (4)$$

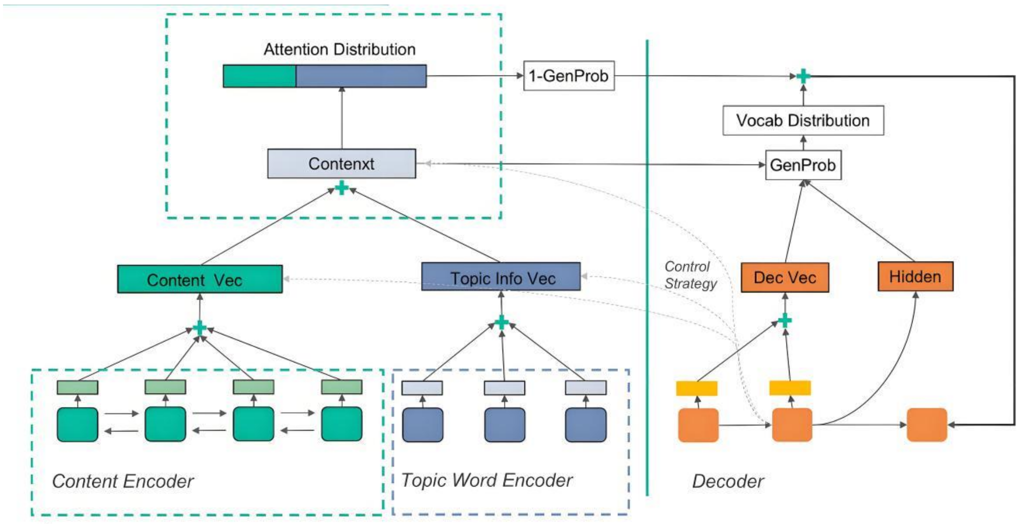

Figure 5 Lstm-attention module

The edge weights of GCN need to be dynamically updated to adjust the risk transmission intensity based on the health of the financial reports of the related counterparties. The LSTM hidden states and GCN node embeddings are fused through a gating mechanism to generate a joint feature representation. The output layer adopts a sigmoid function to map to the probability of capital chain breakage in the next 6 months, and the loss function is defined as weighted cross-entropy, which strengthens the penalty weight of high-risk samples. Adversarial sample enhancement technique is introduced in the training phase to simulate the extreme scenario of exchange rate fluctuation and supply chain disruption to improve the model robustness.

4. Enterprise Financial Management Insights and Application Recommendations

4.1 Technology Implementation Path

The construction of enterprise-level financial big data platform needs to address the integration of heterogeneous data sources and real-time analysis needs. The technical architecture adopts Lambda layered design, and the batch-flow integrated processing engine synchronizes the transactional data of ERP system, structured approval records of OA process and unstructured text flow of public opinion system. Data silos are bridged by relying on a standardized data governance framework, establishing a unified metadata catalog based on the DCMM (Data Management Capability Maturity) model, and realizing semantic alignment between multiple systems at the field level through an API gateway, e.g., mapping “Purchase Requisition No.” in the OA system to

"Accounts Payable Voucher No." in the ERP system. The "accounts payable voucher number" in the OA system is mapped to the "accounts payable voucher number" in the ERP system. The data cleaning phase introduces a dynamic quality rule engine to automatically repair missing fields and formatting anomalies, and flag low-confidence data for manual review [9].

Visual decision dashboard development needs to be deeply coupled with the underlying algorithms. The Power BI embedded analytics module calls the Pareto frontier solution set of the funds dynamic allocation optimization algorithm to show the funds allocation scheme under different risk preferences in the form of a heat map, as shown in Figure 6. the Tableau Early Warning Watchdog integrates the financial risk prediction algorithm's LSTM-Attention weights to visualize the risk transmission paths and key influencing factors. In the technical realization, the front-end components and distributed computing engine using microservice communication protocols, real-time rendering of data updates to enable cache degradation strategy to protect the response speed.

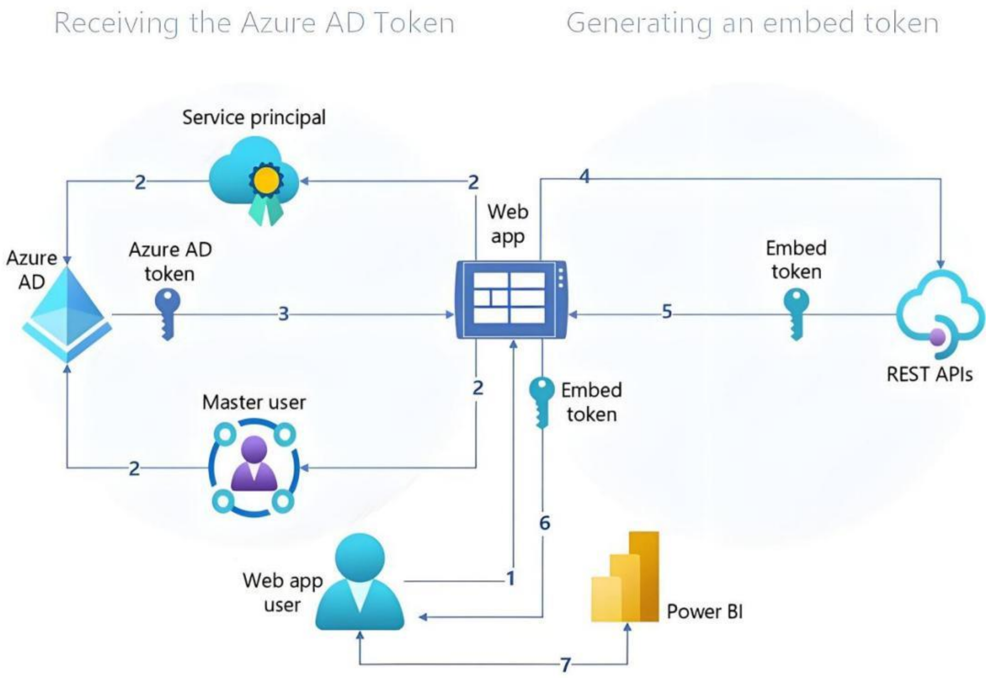

Figure 6 Power BI Embedded Analytics Module

A progressive iterative mechanism needs to be established in the implementation stage. At the initial stage, focusing on the data access of core business units, the containerized deployment mode is used to quickly verify the validity of the ETL process; at the mid-term, it is expanded to complex scenarios such as supply chain finance, and the time zone differences in cross-border capital flows are dealt with through the Flink state management function; and at the later stage, it builds a self-service analytical portal, embeds a natural language query module, and supports the generation of dynamic reports by non-technical personnel through semantic parsing. The security system adopts zero-trust architecture, based on RBAC model to control data column access rights, combined with homomorphic encryption technology to ensure the computing security of sensitive financial data in the cloud.

4.2 Organizational change challenges

The transformation of financial personnel's capabilities requires the reconstruction of the traditional accounting skills system and the establishment of a dual-core capability model of "data interpretation - strategy optimization". The ability assessment formula can be quantified as formula (5):

$$C = \alpha \cdot Data\quad Literacy + \beta \cdot Strategic\quad Alignment \quad (5)$$

Among them, Data Literacy covers SQL query and machine learning model interpretation capabilities, and Strategic Alignment is assessed through the contribution of business goal disassembly and budget optimization. In the technical realization, the embedded learning platform integrates the Jupyter Lab interactive environment, provides sandbox training modules for cash flow forecasting, risk scoring and other scenarios, and strengthens the causal reasoning ability between algorithm results and business decisions [10].

The data governance committee needs to build a cross-departmental collaboration mechanism, and its core function consists of data quality standard definition and permission rule design. The data quality scoring model is equation (6):

$$Q = \sum_{i=1}^{n} w_i \cdot Compliance_i \cdot (1 - NullRate_i) \quad (6)$$

The weights w_i be dynamically adjusted based on the business importance of the fields, and the compliance indicator verifies whether the data complies with IFRS or local accounting standards. Permission rules use the Attribute Based Access Control (ABAC) model to encapsulate roles, data sensitivity levels, and business scenarios into policy decision points, e.g., Subsidiary Finance only has access to the Linked Transactions data subgraph.

5. Experimental Design and Validation

5.1 Experiment 1: Validation of capital deployment optimization

The experimental design is based on six years of fund flow data from 12 subsidiaries of a multinational group to verify the effectiveness of the multi-objective optimization algorithm in the dynamic deployment of funds. Three types of features are extracted in the data preprocessing stage: transaction amount, account balance, and payment timeliness, and the nodes are defined as subsidiary accounts when constructing the fund flow mapping, and the edge weights are jointly determined by the average daily transaction amount and the urgency coefficient. Comparison experiments are set up with two baseline methods: traditional linear programming method (the objective function is liquidity maximization), and manual empirical provisioning (relying on historical rule base), and the optimization algorithm adopts the improved NSGA-II, whose fitness function is defined as Equation (7):

$$F = \lambda_1 \cdot ROI - \lambda_2 \cdot RiskExposure + \lambda_3 \cdot LiquidityScore \quad (7)$$

The weighting coefficients λ_1, λ_2, and λ_3 are dynamically optimized by Pareto frontier analysis.

The experimental results are shown in Table 1.The NSGA-II algorithm significantly outperforms the baseline method in the three dimensions of capital utilization, annualized return and response timeliness. The Pareto solution set generated by the algorithm reveals the equilibrium relationship between return and risk in the fund allocation scheme, and the decision maker can choose the nondominated solution according to the risk preference. Anomalous scenario tests show that NSGA-II completes cross-region fund dispatch within 2 hours through topology reconfiguration when a subsidiary has a sudden large amount of payment demand, while the traditional method has an average delay of 8 hours due to the fixed allocation ratio.

Table 1 Comparative results of optimization of funds deployment

Comparative Indicators	Traditional Linear Programming	Manual Empirical Blending	NSGA-II algorithm
Capital Utilization Rate	68%	62%	90%
Annualized return (million yuan)	6500	5800	8300
Response time (hours)	6.8	8.2	2.1

Experiments verify the practicality of the multi-objective optimization algorithm in complex funding networks. The constraint violation rate analysis shows that the feasible solution share of NSGA-II reaches 98.7%, which is 23 percentage points higher than the linear programming method. The algorithm needs to be deployed with a real-time monitoring module embedded, which triggers the model retraining when the sudden change of the structure of the money flow mapping exceeds the threshold value, and guarantees the dynamic adaptability.

5.2 Experiment 2: Financial risk prediction accuracy test

The experiment verifies the effectiveness of the fusion model of graph convolutional network and long and short-term memory network in the early warning of the risk of capital chain breakage. Positive samples cover 50 enterprises with the history of financial chain breaks, negative samples are selected from 200 enterprises with healthy financial status, and feature engineering extracts key indicators such as cash flow volatility, related party transaction ratio, and interest coverage multiple. The model compares logistic regression, random forest and pure LSTM baseline, and the evaluation dimensions contain precision rate, recall rate and F1-score.

Table 2 shows the results of the model performance comparison: the F1-score of the GCN-LSTM fusion model is 22% higher than the optimal traditional model, and the false alarm rate is controlled within the industry benchmark. The case retrospective validation shows that the model identifies the risk signal of a group subsidiary 3 months in advance, and when the early warning is triggered, the accounts payable turnover days of the enterprise is already significantly higher than the industry average, and a substantial debt default occurs 3 months later.

Table 2 Comparison of financial risk prediction model performance

Model	Precision Rate	Recall rate	F1-score	ROC-AUC
logistic regression	0.71	0.68	0.69	0.73
Random Forest	0.82	0.74	0.78	0.85
Pure LSTM	0.76	0.80	0.78	0.83
GCN-LSTM fusion model	0.89	0.93	0.91	0.94

The validation process introduces SHAP value analysis, and supplier concentration and quick ratio have the highest contribution to the model decision. During the deployment phase, a dynamic feature monitoring module is built to trigger model retraining when the structure of the enterprise's network or financial indicators change abruptly, ensuring the stability of the prediction. The algorithm output should be integrated with the audit system to automatically generate risk response plans for management decision-making.

5.3 Experiment 3: Smart budgeting efficiency assessment

The experiment was designed to validate the improved efficiency and quality of the smart budgeting algorithm. The control group adopts manual preparation mode, and the 10-member team completes the budget preparation based on historical templates and spreadsheet tools; the experimental group deploys the smart algorithm to generate the first draft, and the 2-member team performs the calibration of strategic priorities and fine-tuning of constraints. The evaluation indexes introduce budget deviation rate and strategy matching degree, the former measures the degree of deviation between budget allocation and actual resource consumption, and the latter quantifies the consistency between the budget program and strategic plan through expert scoring.

The smart budget algorithm integrates Prophet time series prediction and deep reinforcement learning strategies. The budget deviation rate is calculated using the standardized formula (8):

$$\Delta = \frac{1}{N}\sum_{i=1}^{N}\frac{\left|B_{alloc}^{i} - B_{actual}^{i}\right|}{B_{actual}^{i}} \quad (8)$$

Where B_{alloc}^{i} is the budget allocated by the algorithm and B_{actual}^{i} is the actual expenditure after audit. The strategy matching score parses the strategy document keywords through the semantic similarity model and generates quantitative indicators by combining the departmental budget share. The manual fine-tuning phase deploys an interactive decision support system to visually display the Pareto frontier solution set of the algorithmic recommendation scheme with risk sensitivity analysis results.

Figure 7 presents the core results of the experiment. The budget deviation rate of the intelligent algorithm group is 61% lower than that of the manual group, and the strategy matching degree is 19.4% higher. The efficiency improvement stems from the fact that the algorithm automatically generates 80% of the base budget framework, and the manual team focuses on dealing with resource tradeoffs for strategic-level projects. Anomalous case tests showed that when the market environment changed suddenly, resulting in a revenue forecast error of more than 15%, the algorithm reallocated the budget within 4 hours through an online reinforcement learning module, while the manual team had to restart the whole process of revisions.

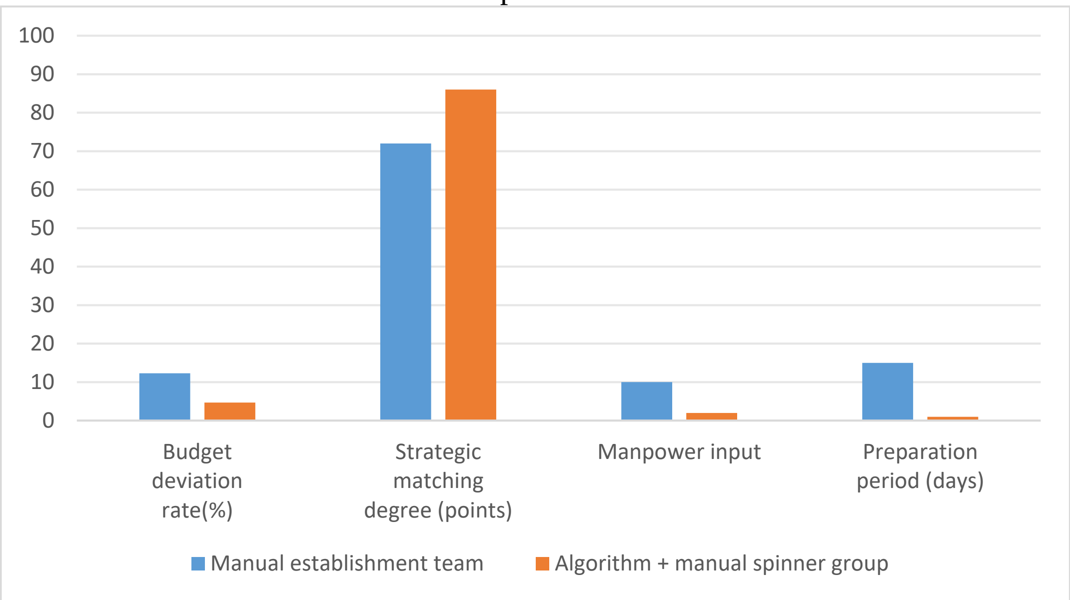

Figure 7 Smart budgeting efficiency comparison

The validation process uses a combination of cross-validation and stress testing. The dataset was divided to reserve 20% as a test set for unexpected events, simulating scenarios such as supply chain disruptions and policy adjustments. The expert scoring team consists of strategy department, finance department, and business line leaders, and the scoring consistency is verified by Krippendorff's alpha coefficient up to 0.82. Algorithm deployment needs to be accompanied by revision of the approval workflow, and manual fine-tuning of the operation traces of the blockchain deposits in order to meet the audit compliance requirements.

6. Conclusion

The deep embedding of big data analytics into financial shared decision-making system marks the paradigm shift of financial management from rearview mirror recording to navigational ritual prognosis. Research confirms that feature engineering's ability to extract value from unstructured data directly affects the sensitivity of risk early warning models; dynamic deployment algorithms show resilience advantages beyond traditional rule engines in fund liquidity management. The technology implementation path needs to take into account the maturity of data governance and organizational change tolerance to avoid falling into the misconception of technological determinism. The experimental validation session reveals the adaptation law of algorithm iteration and actual business scenarios, providing differentiated solution references for enterprises of different sizes. With the fusion of edge computing and privacy computing technology, the future financial decision-making system will evolve in the direction of distributed intelligence, which will have a far-reaching impact on reconfiguring the enterprise value creation chain.

References

[1] Ren S. Optimization of Enterprise Financial Management and Decision-Making Systems Based on Big Data[J]. Journal of Mathematics, 2022, 2022(1): 1708506.

[2] Merendino A, Dibb S, Meadows M, et al. Big data, big decisions: The impact of big data on board level decision-making[J]. Journal of Business Research, 2018, 93: 67-78.

[3] Ghasemaghaei M. Does data analytics use improve firm decision making quality? The role of knowledge sharing and data analytics competency[J]. Decision Support Systems, 2019, 120: 14-24.

[4] Torre C, Guazzo G M, Çekani V, et al. The relationship between big data and decision making. A systematic literature review[J]. Journal of Service Science and Management, 2022, 15(2): 89-107.

[5] Jeble S, Kumari S, Patil Y. Role of big data in decision making[J]. Operations and Supply Chain Management: An International Journal, 2017, 11(1): 36-44.

[6] Yu W, Wong C Y, Chavez R, et al. Integrating big data analytics into supply chain finance: The roles of information processing and data-driven culture[J]. International journal of production economics, 2021, 236: 108135.

[7] Megeid A, Sobhy N. The role of big data analytics in supply chain "3Fs": Financial reporting, financial decision making and financial performance "An Applied Study"[J].Accounting concept, 2022, 26(2): 207-268.

[8] Wang L, Wu Y, Huang Z, et al. Big data application and corporate investment decisions: Evidence from A-share listed companies in China[J]. International Review of Financial Analysis, 2024, 94: 103331.

[9] Wirawan P. Leveraging Predictive Analytics in Financing Decision-Making for Comparative Analysis and Optimization[J]. Advances in Management & Financial Reporting, 2023, 1(3): 157-169.

[10] Pattnaik M, Shah T R. Role of big data to boost corporate decision making[C]//2023 2nd International Conference on Edge Computing and Applications (ICECAA). IEEE, 2023: 105-111.