Development Economics of China
2025, VOL. 9, NO. 10, 71-74
DOI: 10.47297/wspdecWSP2515-797315.20250910

Big Data Analysis Methods and Practices in Enterprise Risk Management

Yuanheng Li[1,2]

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

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

ABSTRACT

In the wave of digital transformation, enterprise risk management faces challenges such as a sharp increase in data dimensions and accelerating risk transmission. Traditional risk control methods rely on human experience and sampling analysis, which fall short when it comes to real-time processing of multi-source heterogeneous data, let alone penetrating hidden risks in complex business scenarios. Big data analytics, through the integration of structured and unstructured data, constructs a framework capable of dynamically assessing risks, thereby providing robust support for the entire chain from risk identification to predictive decision-making. Current research primarily focuses on the integration of technology and business, but issues such as data governance deficiencies and organizational adaptation lags constrain the realization of its value. Therefore, there is an urgent need to explore a practical paradigm that balances efficiency and security, driving enterprise risk management towards intelligent and forward-looking upgrades.

KEYWORDS

Enterprise risk management; Big data analysis; Method; practice

1 Introduction

Corporate risk management is undergoing a paradigm shift from passive defense to proactive prediction, with the core driving force being the reshaping of risk management logic architecture by big data technology. The capability to collect massive amounts of data has broken through the limitations of traditional sampling analysis, while multi-dimensional data correlation analysis can uncover systemic risks hidden in business processes. Real-time monitoring technology is facilitating a transition from delayed response to immediate intervention in risk management. This paper focuses on strategic and financial risks as key dimensions, exploring the application mechanisms of big data analytics methods in risk quantification and pattern recognition. It also analyzes governance gaps and compliance challenges that arise during the process of technology empowerment, aiming to provide theoretical support and practical references for balancing the effectiveness of risk management with the costs of digital transformation.

2 The Theoretical Integration of Enterprise Risk Management and Big Data Analysis

2.1 The Core Dimension of Enterprise Risk Management

When companies face market liberalization, regulatory deregulation, and product innovation, the degree of change and fluctuation increases, thereby increasing operational risks. Strategic risk involves the alignment between a company's long-term development direction and its external environment. Traditional qualitative assessments struggle to penetrate the nonlinear relationship between market volatility and policy changes. Big data analysis constructs a panoramic view by integrating industry trends, public opinion dynamics, and other unstructured data, assisting decision-makers in identifying potential strategic deviations. Financial risk focuses on the stability of capital flow and asset-liability structure. Real-time monitoring of high-frequency trading data and supply chain information can enhance the accuracy of abnormal fluctuation warnings. The predictive power of algorithm models for cash flow trends strengthens the forward-looking nature of risk exposure management. Operational risk arises from technical flaws or human oversights in business process execution. Automated parsing technologies for sensor logs and operation records can accurately pinpoint process breakpoints. Deep mining of historical accident data by pattern recognition algorithms effectively extracts risk trigger thresholds, providing quantitative evidence for the iteration of standardized operating procedures [1].

2.2 Features of Big Data Analytics Technology

Big data analysis refers to the analysis of massive amounts of data, which can be summarized in five V's: volume (Volume), velocity (Velocity), variety (Variety), value (Value), and veracity (Veracity). The massive data processing module relies on a distributed computing framework, capable of deeply cleaning and integrating unstructured information such as text, images, and sensor signals. The streaming engine supports millisecond-level latency for data throughput, making real-time parsing of TB-scale information possible. The pattern recognition mechanism leverages the collaborative computation of convolutional neural networks and recurrent neural networks to capture multi-dimensional features from discrete events, particularly excelling in uncovering nonlinear relationships between transaction records and market sentiment. Weak signals that traditional statistical methods struggle to detect form visual decision-making bases after being reconstructed in feature space. The real-time prediction system integrates time series analysis with dynamic

Bayesian networks, iteratively calculating key indicators like inventory turnover rate and supply chain disruption probability every five minutes. When social media sentiment indices experience abnormal fluctuations, the predictive model completes parameter adaptive adjustments within twenty training cycles. This continuously evolving characteristic significantly enhances the timeliness and accuracy of risk warnings. The hybrid storage architecture based on Apache Hadoop ecology not only guarantees the long-term retrospective requirements of historical data, but also provides a low-latency buffer pool for real-time computing of Spark Streaming. The nested invocation of TensorFlow and Prophet framework realizes the dynamic visualization of prediction result confidence interval.

3 Key Issues of Big Data Driven Enterprise Risk Management

3.1 Data Quality Issues

Data fragmentation between business systems leads to gaps in the risk indicator traceability chain. The differences in collection standards and update frequencies across various data sources exacerbate the cumulative bias in cross-dimensional analysis. The heterogeneity of internal and external data sources makes it difficult to align unstructured text, sensor logs, and financial statements within a unified semantic framework. Key attribute loss or noise interference during data cleaning directly impacts the sensitivity of subsequent risk modeling. The timeliness requirements of real-time data streams and the completeness of historical data archiving create dual pressures. High-frequency operational equipment status information, if not filtered and de-noised, can easily mask truly predictive risk signals. Traditional data governance frameworks lag behind dynamic business scenarios, and the weak data lineage tracing mechanism makes it challenging for risk analysts to assess statistical significance over time for specific datasets. Some off-line batch processing-derived indicators lack contextual markers, casting doubt on their credibility in real-time decision-making [2].

3.2 Technology Application Bottleneck

Technical implementation faces challenges of compatibility with heterogeneous systems. Structured data and unstructured logs scattered across various nodes in the supply chain have format differences, and fragmented storage architectures create substantial barriers to cross-departmental information sharing. Data cleaning consumes over sixty percent of engineering resources, while anomaly detection algorithms lack sensitivity to concept drift in time series data. Feature engineering stages often overlook risk factors with weak correlations. Streaming data processing frameworks can incur millisecond-level latency in scenarios with millions of concurrent operations, potentially missing critical decision-making windows. The black-box nature of deep learning models hinders explainability when tracing risk causes, and regulatory requirements for model logic conflict with existing technical approaches. In dynamic business environments, algorithm iteration lags behind market changes, and static assumptions in historical training sets struggle to adapt to non-linear risk propagation patterns driven by sudden events. Domain adaptation errors during transfer learning applications can weaken cross-industry risk warning effectiveness.

3.3 Organizational Coordination Problems

There are significant differences in the perception and application goals of risk data among different functional departments within an enterprise. Business units focus on short-term operational metrics, while technical teams emphasize the stability of data architecture. This goal gap leads to information loss during the cross-system transmission of risk analysis requirements. Under the traditional hierarchical structure, the segmented data permissions result in a lack of a holistic perspective for building risk profiles. Supply chain data and customer behavior logs, due to their different management systems, struggle to form coherent risk assessment inputs. The barrier between technical language and business terminology increases communication costs. Misunderstandings by algorithm engineers about business scenarios can cause risk warning models to deviate from actual management needs, while the limited understanding of data science methods by business personnel affects their critical application of analytical results [3]. The departmental performance evaluation system does not include a dimension for assessing collaborative efficiency, and the absence of a data sharing incentive mechanism causes key risk clues to remain trapped in local systems. Differences in decision-making rhythms across management levels make it difficult to quickly translate real-time risk insights into cross-departmental action plans.

3.4 Compliance and Privacy Risks

The integrity of the data governance framework influences the legal boundaries of information collection and application. Some companies, in the process of integrating multi-source heterogeneous data, may lack dynamic tracking of regulatory policies, leading to potential issues such as ambiguous data sovereignty or conflicts in cross-border transmission rules. Design flaws in privacy protection mechanisms can result in risks of informed consent failure when analyzing user profiles and behavioral trajectories, especially during the processing of sensitive information like biometric data and health records. Over-reliance on algorithmic black boxes may weaken the control of data subjects. If the iteration speed of dynamic risk assessment models cannot keep up with the expansion pace of data scale, it can easily create blind spots in data lifecycle management. The differentiated legislation on personal information protection across different jurisdictions further complicates the confirmation and sharing of data assets. The tension between technical ethics and

commercial interests requires companies to consider social trust costs and legal accountability mechanisms simultaneously when deploying analytical tools.

4 Implementation Suggestions for Enterprise Big Data Risk Management Optimization

4.1 Build an Enterprise-level Risk Data Governance System

In the construction of an enterprise-level risk data governance system, top-level design must clarify the framework for responsibility and authority allocation between the risk management department and business units. Data classification standards and risk preference statements should be embedded into business processes to avoid redundant decision-making chains or delayed responses due to overlapping functions. The routine operation of cross-departmental collaboration mechanisms relies on a unified risk terminology database and information sharing platform, thereby reducing communication costs and enhancing the granularity of risk identification. As a core component of the governance system, data quality control must establish metadata management rules covering the entire process from data collection, cleaning, to storage. Dynamic validation algorithms should be designed to address semantic parsing discrepancies in unstructured data, while third-party audit institutions should be introduced to independently verify the completeness and consistency of data sources, ensuring that input parameters for risk modeling are traceable. Optimization of the technical architecture should focus on the security computing capabilities of distributed storage systems. Under the premise of meeting real-time risk monitoring requirements, homomorphic encryption and differential privacy technologies should balance data utilization efficiency with the protection of sensitive information. The permission grading control module should dynamically adjust access thresholds based on job sensitivity to prevent secondary risks caused by unauthorized operations [4]. The iteration of compliance audit processes should align with industry regulatory trends and internal risk exposure changes, regularly assessing the adaptability of data lifecycle management strategies. Emergency plans should be developed for complex scenarios such as cross-border data flows, and risk governance performance should be incorporated into the organizational evaluation system to drive continuous improvement. Legal advisory teams should deeply participate in the revision process of governance rules to ensure that technical solutions keep pace with judicial practices.

4.2 Develop Lightweight Analysis Tools and Models that are Suitable for Business Scenarios

Companies need to develop lightweight analytical tools based on the actual risk touchpoints of business processes as the starting point for development. They should design modular functional components tailored to the differentiated monitoring needs of procurement, production, and sales stages, while ensuring the accuracy of core algorithms and stripping away redundant computational layers to reduce deployment costs. Model architects should select appropriate machine learning frameworks based on the update frequency and granularity characteristics of business data. For risk scenarios with significant temporal features, such as supply chains, a sliding window mechanism should be used to dynamically adjust feature weights. For non-structured data processing requirements like public opinion monitoring, a semantic understanding layer must be embedded to improve entity recognition accuracy. Business teams need to collaborate with data scientists to define risk warning thresholds and model iteration trigger conditions. In the development of inventory turnover rate warning tools, historical stockout rates and market fluctuation parameters should be converted into adaptive adjustment variables in the model, enabling the tool to autonomously optimize decision boundaries in response to changes in corporate strategies. Operations teams need to establish a model performance degradation monitoring system, dynamically calibrating algorithm parameters by comparing the deviation between the risk prediction results output by the tool and actual events. For businesses with frequent adjustments to pricing strategies at the sales end, a quarterly model reconstruction mechanism should be set up to prevent outdated logic from affecting the effectiveness of decisions.

4.3 Establish a "Business-technology-risk Control" Collaborative Management Mechanism

The business department should proactively convert risk tolerance indicators into quantifiable business rules, working with the technical team to define the business relevance of data collection dimensions. This avoids delays or misjudgments in risk signal identification due to misaligned requirements. The design of risk control models must be embedded at core decision-making nodes in business scenarios, such as dynamically linking credit assessment thresholds with supply chain transaction frequencies. The technical team needs to establish scalable data interface standards to facilitate real-time interaction between business systems and risk databases. For high-frequency trading scenarios, a streaming computing framework should be designed to reduce risk response latency. The risk control department adjusts monitoring frequency thresholds based on business fluctuation patterns to ensure that anomaly detection algorithms can capture long-tail risks without disrupting normal operations. The routine operation of cross-functional collaborative teams relies on a standardized risk assessment language system. Business experts need to transform industry experience into a risk label rule library, while the technical team optimizes feature engineering logic based on rule confidence. Risk analysts validate model prediction results against actual business losses through A/B testing. Regularly, the three parties calibrate the data weight allocation scheme [5]. The establishment of a joint decision-making mechanism requires business

units to submit risk exposure calculation reports when formulating market strategies. The technical team predicts stability bottlenecks in data links based on risk transmission paths, and the risk control department designs stress test parameters using historical event databases. Major decisions require tripartite approval to balance efficiency and robustness. Post-event reviews must cross-verify the dynamic relationship between business benefits and risk management costs, forming a closed-loop optimization path that spirals upward.

4.4 Strengthen Data Security Compliance and Ethical Risk Control

In the construction of data security compliance, companies need to embed privacy protection design principles throughout the entire lifecycle of data collection and processing. They should implement tiered and categorized management for sensitive data such as customer information and trade secrets, dynamically adjusting encryption strength and access permission granularity based on exposure risks along the data flow path. The technical team should build an adaptive security protection architecture, deploying real-time desensitization engines at the data warehouse level to contextualize query results. When external partners retrieve supply chain data, unnecessary fields should be automatically stripped and watermark tracking identifiers added. The compliance department must establish a regulatory rule knowledge graph, converting personal information protection regulations from different regions and industry supervision requirements into executable data operation prohibitions. For example, in cross-border data transmission scenarios, geofencing should be preset to block API calls that violate data sovereignty laws. The audit committee should develop a data usage ethics review process, regularly testing the bias of training sets for algorithm models. If customer segmentation profiles involve sensitive dimensions such as race or gender, a human review mechanism should be triggered to prevent discriminatory decisions. The data governance platform should integrate risk event traceability functions, generating tamper-proof operation logs for any data access involving user privacy. Combined with blockchain technology, this provides verifiable evidence chains for compliance audits. Business departments should conduct data impact assessments when developing new product features, predicting potential misuse risks associated with newly added data fields and preemptively deploying access control policies to avoid compliance disputes arising from innovative business expansion.

5 Conclusion

Big data technology has built a three-dimensional perception network and decision-making model for enterprise risk management. Its value lies in successfully transforming vague risk perceptions into quantifiable and traceable management objects. Research has found that the accuracy of risk warnings is directly influenced by the completeness of the data governance system, while the marginal benefits of technology application are determined by the compatibility between business scenarios and algorithm models. In the future, companies should embed data-driven risk control genes into their organizational structures. On one hand, they need to cultivate interdisciplinary talents to overcome the barriers to technology application; on the other hand, they must establish flexible iteration mechanisms to adapt to rapidly changing regulatory environments. Subsequent research should focus on the restructuring effects brought about by emerging technologies such as edge computing and privacy computing on risk management models, and further explore the ethical boundaries and application paths of human-machine collaborative decision-making in complex business environments.

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

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] Zhou Shenglai. Research on Enterprise Supply Chain Risk Management Strategies under Big Data [J]. China Chief Accountant, 2024, (12): 119-121.

[3] Zhao Jiaxin, Yao Shuzi, Wang Junming. Research on Enterprise Financial Risk Management from the Perspective of Big Data [J]. Business Exhibition Economy, 2024,(24):101-104.

[4] Shan Linhua. Exploration of the Path to Improve the Quality and Efficiency of Financial Accounting Management in Enterprises in the Era of Big Data [J]. Administrative Affairs and Finance, 2024, (24):124-126.

[5] Sun Zhiqi. How to Optimize Corporate Tax Risk Management Strategies in the Era of Big Data [J]. China Business, 2024, (12):120-121.