World Journal of Management Science
Print ISSN: 2959-9628
Online ISSN: 2959-9636
DOI: https://doi.org/10.61784/wms3083

BIG DATA IN FINANCIAL SERVICES: A SURVEY OF APPLICATIONS AND VALUE CREATION

RongHua Li*, HongFan Li
School of Finance and Economics, Hainan Vocational University of Science and Technology, Haikou 571126, Hainan, China.
Corresponding Author: RongHua Li

Abstract: Big data technology is profoundly transforming the business models and value creation methods of financial services. This paper employs a systematic literature review methodology to investigate the applications and value mechanisms of big data across key financial domains. The study identifies four primary value creation pathways: First, in credit scoring and risk management, integration of alternative data sources enhances model accuracy and extends financial access to underserved populations. Second, in fraud detection, machine learning-based real-time analytics reduce false positive rates by approximately 30% while improving detection precision. Third, personalized services enabled by multi-dimensional customer profiling generate revenue increases of 10-15% for institutions implementing data-driven strategies. Fourth, process automation and real-time decision support enhance operational efficiency, reducing costs and improving service quality. This research demonstrates that big data has become a strategic asset for financial institutions, while acknowledging ongoing challenges in data governance, model interpretability, and regulatory compliance.
Keywords: Big data; Financial services; Credit scoring; Fraud detection

1 INTRODUCTION

The financial services industry is undergoing a profound data-centric transformation. With the digitization of transaction systems, increased customer interactions, and the proliferation of Internet of Things (IoT) devices, financial institutions now process data volumes growing at exponential rates [1,2]. The emergence of big data analytics tools—including distributed computing frameworks, machine learning algorithms, and real-time processing capabilities—enables these institutions to extract valuable insights from vast and diverse datasets, thereby revolutionizing their business models and service delivery [3,4].

The integration of big data analytics in financial services represents a paradigm shift from traditional decision-making approaches to a data-driven paradigm. Unlike conventional methods that rely on limited datasets and static assessment models, contemporary approaches leverage diverse data sources—including transactional records, social media activity, and behavioral patterns—for more precise predictions and comprehensive insights [3,5]. This transition impacts multiple facets of financial services, from risk assessment and customer experience to operational efficiency and product innovation. Furthermore, regulatory developments such as the European Payment Services Directive 2 (PSD2) and Open Banking initiatives have facilitated data sharing among financial organizations, accelerating innovation and enabling new business models [2,6].

This paper systematically analyzes how big data creates value in financial services by examining four key application domains: credit scoring and risk management, fraud detection and prevention, personalization of financial services, and operational efficiency optimization. The research objectives are threefold: first, to identify the specific mechanisms through which big data enhances decision-making in each domain; second, to evaluate the quantifiable value outcomes achieved through big data implementation; and third, to examine the challenges that persist in realizing big data's full potential. The remainder of this paper is organized as follows: Section 2 establishes the conceptual framework; Section 3 examines applications across the four domains; and Section 4 concludes with implications and future directions.

2 CONCEPTUAL FRAMEWORK OF BIG DATA IN FINANCE

2.1 Defining Big Data in the Context of Digital Finance

In the context of digital finance, big data refers to the vast and diverse information generated from financial transactions, customer interactions, market activities, and external sources. These data are commonly characterized by the "Four V's": Volume (the immense quantity of data), Velocity (the rate of data generation and processing), Variety (multiple forms of structured and unstructured data), and Veracity (the reliability and quality of the data) [2,4]. Financial institutions collect and analyze data from both traditional sources—such as account information, transaction records, and credit histories—and alternative sources, including social media activity, mobile phone usage, browsing behavior, and IoT device data [3,7]. The integration of these diverse data sources provides a more comprehensive understanding of financial behavior and market dynamics, enabling more sophisticated analysis and decision-making.

2.2 The Evolution of Big Data Analytics in Finance

The financial sector has historically been data-driven, yet the scale and complexity of data analysis have evolved significantly with technological advancements. Traditional analytics relied primarily on structured data and descriptive statistics, focusing on historical reporting and basic trend analysis. The advent of big data technologies enables financial institutions to process massive volumes of unstructured data, apply advanced machine learning algorithms, and derive predictive and prescriptive insights [2,4]. This evolution has been facilitated by several technological developments: distributed computing frameworks such as Hadoop and Spark, cloud computing platforms enabling scalable storage and processing, advanced machine learning algorithms, and real-time analytics capabilities [7]. Recent regulatory changes, particularly Open Banking initiatives, have further accelerated this transformation by promoting data sharing and interoperability within the financial ecosystem [6]. Collaborative efforts such as the EU-funded INFINITECH project, which brings together multiple organizations to develop innovative big data and AI solutions for digital finance, exemplify the industry's recognition of big data's transformative potential [8].

3 APPLICATIONS OF BIG DATA IN FINANCIAL SERVICES

3.1 Credit Scoring and Risk Management

Big data has transformed credit risk assessment by enabling financial institutions to develop more comprehensive and accurate models for evaluating creditworthiness. Traditional credit scoring systems rely heavily on limited datasets—such as credit history, income verification, and debt-to-income ratios—which often exclude individuals with limited or no credit records, commonly referred to as "thin-file" or "no-file" consumers [3]. Contemporary big data approaches integrate diverse data points to create more nuanced risk profiles, thereby expanding credit access while maintaining risk management standards [5,7].

Financial institutions now incorporate alternative data sources—such as utility payment histories, mobile phone usage patterns, online transaction behavior, and in some cases, social media activity—to assess creditworthiness [3][9]. This expanded data scope is particularly valuable for evaluating applicants who lack traditional credit history. For instance, fintech lenders in emerging markets have successfully utilized mobile phone data and digital payment records to extend credit to previously underserved populations [10]. Research indicates that alternative data can improve predictive accuracy for thin-file borrowers, with some studies reporting Area Under the Curve (AUC) improvements from 0.65 to 0.75 or higher when alternative data is incorporated [3,11]. These improvements demonstrate that non-traditional data sources can effectively complement conventional credit bureau data, enhancing both financial inclusion and risk management outcomes [11].

Beyond initial credit assessment, predictive analytics enhances ongoing risk management by enabling dynamic, real-time risk monitoring rather than periodic reviews. Technologies such as Apache Kafka and Apache Flink facilitate the continuous processing of risk indicators, allowing institutions to proactively identify emerging risks and adjust credit limits or terms accordingly [7]. This capability is especially valuable in volatile economic environments where traditional static models may fail to capture rapidly changing risk profiles.

Machine learning algorithms have improved the accuracy of credit risk models by analyzing complex patterns across diverse datasets and identifying subtle correlations that traditional statistical methods might overlook [5]. These advanced models enable more precise risk-based pricing, allowing financial institutions to offer interest rates that better reflect individual risk profiles. This differentiation benefits both lenders—through optimized risk-return tradeoffs—and creditworthy borrowers—through access to more favorable terms [3,5].

Despite these advancements, challenges remain in implementing big data-driven credit scoring. Key concerns include ensuring model interpretability to meet regulatory requirements, addressing potential algorithmic bias that may disproportionately affect certain demographic groups, and maintaining data quality across diverse sources [3,4]. Financial institutions must balance the pursuit of predictive accuracy with fairness, transparency, and compliance obligations.

3.2 Fraud Detection and Prevention

Financial fraud is increasingly sophisticated, with global fraud losses in the financial sector reaching billions of dollars annually. Big data analytics has emerged as a critical tool for detection and prevention, enabling financial institutions to identify fraudulent patterns, enhance risk assessment models, and implement real-time detection mechanisms [12,13].

Big data-driven fraud detection systems analyze massive volumes of structured and unstructured data from various sources, including transaction logs, user behavior patterns, device fingerprints, and external threat intelligence feeds [12]. By applying advanced machine learning algorithms—such as random forests, neural networks, and gradient boosting models—these systems can detect anomalies and suspicious activities that may indicate fraud [13]. Unlike traditional rule-based methods that rely on predefined parameters, machine learning models can identify complex and previously unknown fraud patterns, adapting to evolving fraud tactics [7,12].

Real-time analytics represents a significant advancement in fraud detection. By leveraging stream data processing technologies, financial institutions can continuously monitor transactions and flag suspicious activities within milliseconds of occurrence [12,14]. This capability is critical for minimizing potential losses, with some institutions reporting fraud loss reductions of 20-40% following implementation of real-time detection systems [13]. Real-time monitoring enables proactive risk identification and immediate response, contrasting with traditional batch-processing systems that rely on historical data and periodic reporting [14].

Behavioral analytics plays an increasingly important role in fraud detection by establishing baseline behavior patterns for individual customers and identifying deviations that may indicate fraudulent activity. These analytical techniques integrate multiple data points, including transaction history, geographic location, device information, and activity timing [12]. This multi-dimensional approach improves detection accuracy while reducing false positive rates—which can decline by approximately 30% with advanced machine learning models compared to traditional rule-based systems [13]. Reducing false positives is critical, as excessive legitimate transaction declines erode customer trust and satisfaction.

3.3 Personalization of Financial Services

Big data has transformed how financial institutions understand and engage with their customers, enabling enhanced personalization of products and services. By analyzing customer data across multiple touchpoints—including transaction history, digital interactions, life events, and stated preferences—financial institutions can develop detailed customer profiles, predict needs, and deliver tailored offerings [4,15].

Customer behavior analysis forms the foundation of personalization efforts, providing insights into spending patterns, financial goals, risk tolerance, and service preferences. These insights enable financial institutions to recommend relevant products, offer personalized financial advice, and deliver contextually appropriate communications at optimal times. Effective personalization strategies have been shown to increase revenue by 10-15% for financial institutions through improved cross-selling effectiveness, customer retention, and increased product usage [2]. Beyond revenue impact, personalization enhances customer satisfaction and loyalty by delivering more relevant and timely services.

In the wealth and asset management sector, big data analytics enables more sophisticated personalization of investment strategies. Big data and AI technologies analyze a client's financial situation, goals, and risk tolerance to create customized investment portfolios [2]. These capabilities, once limited to high-net-worth individuals served by human advisors, are now extending to mainstream customers through robo-advisory platforms that leverage big data to provide algorithm-driven financial planning at lower costs [2,4].

Marketing and customer acquisition have been transformed by data-driven personalization. Financial institutions use data mining and analytics techniques to analyze customers' browsing histories, content preferences, and behavioral signals to deliver targeted advertising for financial products and services [10]. Advanced customer segmentation—based on demographics, behaviors, life stages, and predicted needs—enables precision-targeted marketing campaigns that achieve higher conversion rates and lower customer acquisition costs [2,10]. The European INFINITECH project demonstrates how big data and AI technologies can enhance both personalization and trust in digital finance, as effective personalization improves product relevance while strengthening customer relationships through more meaningful interactions [6].

3.4 Operational Efficiency and Process Automation

Big data analytics significantly improves the operational efficiency of financial institutions by optimizing processes, reducing costs, and enhancing decision-making capabilities. By analyzing operational data—including transaction processing times, resource utilization rates, and workflow bottlenecks—financial institutions can identify inefficiencies, optimize resource allocation, and prioritize automation opportunities [4,16].

Process automation represents a key application of big data in financial operations. By analyzing patterns in workflow data, institutions can identify high-value automation opportunities that reduce manual intervention and associated costs. Common applications include automated document processing using optical character recognition (OCR), automated compliance checking against regulatory requirements, and streamlined customer onboarding processes [4,14]. These automations not only reduce processing times—in some cases from days to minutes—but also minimize human error and improve consistency.

Real-time data processing enhances operational decision-making by providing timely insights rather than retrospective analysis. Big data analytics enables financial institutions to process and analyze diverse data sources—including transactional data, customer data, and market data—to support informed decisions across multiple operational domains [14]. This capability is particularly valuable for trading operations where millisecond-level decisions impact profitability, liquidity management where real-time cash flow monitoring prevents shortfalls, and dynamic resource allocation that responds to demand fluctuations.

Big data also improves operational risk management by identifying potential risks and vulnerabilities within internal processes. By analyzing operational data patterns, financial institutions can detect anomalies indicating potential system failures, security breaches, or process breakdowns [9,16]. Predictive maintenance models can forecast system failures before they occur, enabling preventive interventions that reduce downtime and associated costs. This proactive approach contrasts with traditional reactive risk management and contributes to more stable and reliable operations.

4 CONCLUSION

This paper has systematically examined the applications and value creation mechanisms of big data in financial services across four key domains: credit scoring and risk management, fraud detection and prevention, personalization of financial services, and operational efficiency optimization. The analysis reveals that big data has transformed financial services through multiple value pathways. In credit risk assessment, integration of alternative data sources enhances model accuracy while extending financial access to underserved populations. In fraud detection, machine learning-based real-time analytics reduce false positive rates by approximately 30% while improving detection precision. Personalization strategies enabled by multi-dimensional customer profiling generate revenue increases of 10-15% for institutions implementing data-driven approaches. Process automation and real-time decision support enhance operational efficiency, reducing costs and improving service quality.

Despite these substantial benefits, effective implementation of big data in financial services faces ongoing challenges. Technical challenges include ensuring data quality across diverse sources and maintaining model interpretability to meet regulatory requirements. Organizational challenges involve developing analytical capabilities and fostering data-driven cultures. Regulatory challenges center on balancing innovation with data privacy protection, algorithmic fairness, and compliance with evolving regulations such as GDPR and Open Banking directives.

Looking forward, the convergence of big data with emerging technologies—including generative artificial intelligence, federated learning for privacy-preserving analytics, and blockchain for secure data sharing—promises to unlock new value creation opportunities. Financial institutions that successfully navigate the technical, organizational, and regulatory challenges will be well-positioned to leverage big data as a strategic asset, driving innovation, enhancing competitiveness, and creating value for both institutions and customers in an increasingly data-driven financial ecosystem.

COMPETING INTERESTS

The authors have no relevant financial or non-financial interests to disclose.

FUNDING

This article is the teaching reform research project of Hainan Vocational University of Science and Technology in 2023: Exploration and Practice of the Application of Blockchain Technology in the Cultivation of Composite Talents in the Context of the Free Trade Port (Project No. HKJG2023-18).

REFERENCES

[1] Mhlanga D. The role of big data in financial technology toward financial inclusion. Frontiers in Big Data, 2024, 7: 1184444.
[2] Soldatos John, Kyriazis Dimosthenis. Big Data and artificial intelligence in digital finance: Increasing personalization and trust in digital finance using Big Data and AI. Springer Nature, 2022.
[3] Abraham F, Schmukler S L, Tessada J. Using big data to expand financial services: benefits and risks. World Bank Research and Policy Briefs, 2019: 143463.
[4] Sharma S. Big data in finance: A systematic literature review. AIP Conference Proceedings, 2023, 2909(1).
[5] Shakoory S. Next-Gen Credit Scoring: The Impact of Big Data & Predictive Analytics. LinkedIn, 2024, 06.
[6] European data. The Advances in Big Data and AI in digital finance. data.europa.eu, 2023, 05. https://data.europa.eu/en/news-events/news/advances-big-data-and-ai-digital-finance.
[7] Goyal N. Using Big Data for Credit Risk Management. Banking Exchange, 2024, 10. https://www.bankingexchange.com/news-feed/item/10138-using-big-data-for-credit-risk-management.
[8] Beaumont P. Digital finance: Big data, start-ups, and the future of financial services. Routledge, 2019.
[9] ZenGRC Team. Applying Big Data to Risk Management. zengrc.com, 2021, 06. https://www.zengrc.com/blog/applying-big-data-to-risk-management/.
[10] Mail V E. Re: Request for Information and Comment on Financial Institutions' Use of Artificial Intelligence, Including Machine Learning. 2021, 07.
[11] Djeundje V B, Crook J, Calabrese R, et al. Enhancing credit scoring with alternative data. Expert Systems with Applications, 2021, 163: 113766.
[12] Collins O O, Kleopas E G, Kenechukwu O P. The role of big data analytics in improving financial risk management and decision-making processes. World J Adv Res Rev, 2024, 22(1): 1743-1752.
[13] Shoetan P O, Oyewole A T, Okoye C C, et al. Reviewing the role of big data analytics in financial fraud detection. Finance & Accounting Research Journal, 2024, 6(3): 384-394.
[14] FM Contributors. The Role of Big Data Analytics in Risk Management for Financial Institutions. Finance Magnates, 2023, 05.
[15] Trigyn Insights. The Role of Big Data in the Financial Services Industry. Trigyn, 2024, 10.
[16] Doyle Megan. Benefits of Using Big Data in Risk Management. American Express, 2021, 10. https://www.americanexpress.com/en-us/business/trends-and-insights/articles/benefits-of-using-big-data-in-risk-management/.