Applications of Supervised Learning Methods in the Fields of Economics and Management

Yuedi Yang

School of Economics and Management, Inner Mongolia University, Hohhot 010021, China.

0231123726@mail.imu.edu.cn

Abstract. Based on a systematic literature review, this paper explores the current application status, effects, and challenges of supervised learning methods in the fields of economics and management. Grounded in literature from the past five years and utilizing multi-database retrieval and rigorous screening, the study analyzes specific applications of supervised learning across various domains. These include asset pricing and fraud detection in finance, financial fraud identification and audit automation in accounting, and corporate performance prediction and macroeconomic modeling in economics. The research finds that supervised learning, through its advantages of high-precision prediction and automated feature engineering, significantly enhances decision-making efficiency and risk management capabilities in economics and management. However, the field still faces challenges such as data privacy and model interpretability. Future research should focus on directions such as small-sample transfer learning, model transparency, and interdisciplinary integration to further promote development in this area.

Keywords: Supervised learning, Economics and management, Big data analytics, Application effectiveness, Future directions.

1. Introduction

In today's digital era, big data analytics has emerged as a pivotal force driving social progress and economic development [1, 2]. With the rapid evolution of information technology, data generation and accumulation have experienced explosive growth. Effectively mining and utilizing these massive datasets has become a central focus for both industry and academia. According to relevant research, the application of big data in business management decision-making has demonstrated immense potential and value; by leveraging big data effectively, enterprises can better understand markets and consumers, optimize resource allocation, and enhance the accuracy and timeliness of decision-making [3]. Therefore, in-depth research into big data analytics within business management is of great significance for improving corporate competitiveness and management levels [4]. Furthermore, as digital transformation accelerates, data has become a critical resource driving the development of business management, accounting, economics, and finance [5]. As a core branch of machine learning, supervised learning methods—by training on labeled data—can construct models with robust predictive and classification capabilities, providing innovative solutions and decision-support tools for the aforementioned fields [6].

The purpose of this study is to systematically review the theoretical and practical achievements of big data analytics in the fields of economics and management through a literature summary, thereby deepening the understanding of big data analytics curricula. Meanwhile, given that the application of supervised learning methods in business management, accounting, economics, and finance holds significant theoretical and practical importance, this paper will systematically explore the application effects and potential value of these methods through an analysis of relevant literature. It also examines the challenges and opportunities encountered in practical applications, aiming to provide useful references and insights for research and practice in related domains.

The structure of this paper is organized as follows: First, we will detail the importance of big data analytics and its current application status in business management to establish a foundation for subsequent research. Next, the methodology of the literature research will be explained. Subsequently, in the literature analysis and discussion section, we will conduct an in-depth analysis of the literature

to systematically summarize specific application cases and research findings of supervised learning methods in accounting, economics, and finance, analyzing their strengths and limitations. Finally, we will summarize the research results and propose directions and recommendations for future study, providing valuable references for researchers and practitioners in the field of business management.

2. Methodology

Selection Criteria and Search Methods. In this study, the formulation of literature selection criteria and search methods refers to the standardized process of systematic literature review to ensure the scientific nature and comprehensiveness of the research. The details are as follows:

When selecting literature, this study follows specific standards to ensure the quality and relevance of the selected documents. First, the titles, abstracts, or keywords of the literature must explicitly contain keywords such as “big data”, “business management”, “accounting”, “economics”, and “finance” to ensure their close connection with the theme of this study. For instance, the literature must explore specific applications of big data in accounting data analysis, economic forecasting, or financial risk management. Second, the publication timeframe is limited to the last five years (2021–2025) to guarantee the timeliness and cutting-edge nature of the research. Furthermore, the types of literature included encompass both theoretical and empirical studies to provide a comprehensive understanding of the current application status and research achievements of big data in various fields of business management. Theoretical studies provide the foundational basis for big data applications, while empirical studies demonstrate the effects and challenges in actual scenarios. Finally, regarding quality assessment, literature with unreasonable research designs, incomplete data, or insufficient argumentation is excluded to ensure the reliability and validity of the included documents.

To ensure the comprehensiveness and accuracy of the literature search, the following search methods are adopted: First, multiple authoritative databases are selected for literature retrieval, including Web of Science, Scopus, CNKI, and Wanfang Data. Second, in terms of determining search terms, primary search terms such as “big data”, “business management”, “accounting”, “economics,”, and “finance” are identified through direct selection, mining implicit terms, and specifying general concepts. These are expanded with synonyms and related terms; for example, “big data” is expanded to data analytics, and “business management” to corporate governance. Regarding the search scope, “Subject” is the preferred scope, combined with multi-dimensional searches of “Keywords” and “Title” to improve retrieval accuracy. Finally, the search process is recorded in detail, including database names, search dates, time ranges, terms used, languages, and the number of documents retrieved, ensuring the transparency and reproducibility of the research.

Framework and Steps of Literature Analysis. The literature analysis framework and steps refer to the standard procedures of systematic literature review.

Planning Phase. The specific theme determined for this study is "The Application of Big Data in Business Management." This choice is based on the high level of attention and application demand for big data technology in the current business management field. The study first summarizes the theories and concepts of big data analytics—including data mining, machine learning, and text analysis—and their applications in accounting, economics, and finance. On this basis, specific research questions are clarified, such as: "What are the specific applications of big data analytics in accounting?" "How does big data influence the research paradigm of economics?" and "What is the role of big data in financial risk management?" These questions help focus the research direction and improve its relevance.

Literature Retrieval. Following the established search methods, a comprehensive search was conducted across databases like *Web of Science* and *CNKI* to collect literature related to “The Application of Big Data in Business Management”. Subsequently, a preliminary screening was performed to eliminate documents that clearly did not meet the research theme or selection criteria, such as those unrelated to big data technology or outside the scope of business management.

Assessing Literature Quality. To ensure the reliability of the analysis, explicit inclusion and exclusion criteria were established. Key elements such as research design, data sources, and methodologies were strictly evaluated for each document. For instance, literature with unclear methodologies or unreliable data sources was excluded. Additionally, data extraction forms were used to systematically record and quantify key information like research purposes, methods, and results, enhancing the objectivity of the analysis.

Determining Inclusion/Exclusion. Based on the preliminary screening, the study further conducted in-depth reading of titles, abstracts, introductions, and conclusions to remove documents irrelevant to the research questions. For the literature that met the requirements, full-text reading was conducted to extract critical information, such as specific case studies and research outcomes, ensuring the depth and breadth of the research data.

Analysis and Summary. The included literature was categorized by field, such as accounting, economics, and finance. Through comprehensive analysis and meta-analysis, the data and viewpoints were integrated. For example, in accounting, big data applications are concentrated in financial risk warning and audit data analysis; research by Xu et al.[7] analyzes common risks and provides targeted prevention strategies starting from the characteristics of corporate accounting informatization in the big data era. In economics, big data is primarily applied to macroeconomic forecasting and market trend analysis. For instance, while traditional Vector Autoregressive (VAR) models are prone to overfitting, Bayesian methods can effectively improve forecasting performance through shrinkage priors. Recent research by Gruber and Kastner[8] proposes a "semi-global prior" framework to further enhance model flexibility and predictive power. In finance, big data is mainly applied to portfolio optimization and credit risk assessment; for example, Wang et al.[9] achieved precise measurement of supply chain finance credit risk using Logistic regression models combined with big data analytics.

Based on these results, the current status of big data analytics in business management is summarized, highlighting its advantages—such as improving decision-making efficiency and optimizing resource allocation—while revealing its deficiencies, such as data privacy issues and high technical barriers. Finally, future research directions and recommendations are proposed, such as strengthening research on data privacy protection technologies.

3. Literature Analysis and Discussion

This section collects and analyzes literature regarding the application effects and potential value of supervised learning in the fields of business and management. Supervised Learning, as one of the core paradigms of machine learning, has demonstrated extensive application potential in the fields of business and management in recent years due to its ability to automatically extract patterns from labeled data. The following is an in-depth analysis of its application effects and potential value based on the reviewed literature.

Core Advantages and Application Logic. The essence of supervised learning lies in constructing predictive models through training data (input features and corresponding labels) [10] . The advantages of this method are primarily reflected in three aspects: First, supervised learning can capture complex non-linear relationships (such as interaction effects) between variables, thereby outperforming traditional linear regression methods in predictive accuracy [11] . Second, it possesses automated feature engineering capabilities; for instance, the LASSO algorithm automatically selects key variables through sparsity constraints, thereby reducing manual intervention [11] . Finally, supervised learning offers strong scalability, making it applicable to high-dimensional data (such as multi-factor models in financial markets) and heterogeneous data (such as mixed text-numerical accounting reports) [12, 13] .

Application Effects by Field. ***Finance: Asset Pricing and Risk Management.*** In finance, big data analytics optimizes portfolios and assesses credit risk. A review by Gupta et al. [14] indicates that text-based big data is applied in investor sentiment analysis and market impact studies.

The application effects of supervised learning methods are particularly significant in this domain. Financial institutions utilize supervised learning models to optimize portfolios and evaluate credit risk, thereby improving risk management capabilities and investment returns. For instance, Support Vector Machines and Neural Networks are used for credit risk assessment, enabling more accurate predictions of borrower default probabilities [15]. Additionally, supervised learning methods are applied to market trend forecasting and investment decision support, helping financial institutions maintain competitiveness in complex and volatile markets [14, 16].

In asset pricing, Gu et al. [12] found that non-linear models like Gradient Boosting Decision Trees (GBDT) significantly improve risk premium prediction accuracy. For instance, introducing interaction factors into momentum strategies can double returns. In fraud detection, while unsupervised learning is common for anomaly detection [17], supervised learning improves classification accuracy through historical fraud samples [18]. However, Sun et al. [18] noted a "prediction-trade gap" in quantitative trading, suggesting the integration of reinforcement learning to optimize action mapping.

Accounting: Financial Analysis and Auditing. In accounting, big data improves the efficiency and accuracy of processing financial data. Wamba et al.[19] noted that big data provides precise decision support for identifying financial risks. Kumar et al.[20] mentioned that accounting information systems in the digital economy must leverage big data for financial transformation.

In the field of accounting, the application prospects of supervised learning methods are equally broad. Traditional accounting auditing relies on manual sampling and empirical judgment, which is inefficient and prone to bias. Supervised learning methods can construct fraud detection models by learning from large volumes of historical financial data, automatically identifying potential financial fraud and improving auditing efficiency and accuracy [6]. For example, algorithms such as Logistic Regression and Decision Trees have been used to analyze financial statement data and predict corporate financial risks.

By training classifiers (such as SVM) on labeled financial anomaly data—such as profit manipulation indicators—supervised learning can detect risks earlier than traditional statistical methods [11], thereby achieving the identification of financial fraud. It has also been applied in cost control prediction; for instance, Liaras et al.[21] verified the impact of learning patterns (labels) on cost control through regression models, though modern machine learning methods were not introduced. Future research could attempt to use Random Forest to enhance explanatory power. In the domain of audit automation, weak supervision can be utilized to extract numerical claims (e.g., "profit increased by 15%") from financial report texts, thereby reducing manual review costs [13].

Economics: Corporate Governance and Management. Big data provides richer sources for macroeconomic analysis. Gangwani & Zhu[22] pointed out that machine learning is changing the research paradigm of economics. Gruber & Kastner[8] summarized applications in monetary policy communication and macro-prudential regulation.

In the field of economics, supervised learning methods provide new insights for macroeconomic forecasting and market trend analysis. Compared with traditional econometric methods, supervised learning can handle high-dimensional data and complex non-linear relationships, thereby predicting economic indicators and market trends more accurately [11]. For example, algorithms such as LASSO Regression and Random Forest are used in macroeconomic forecasting, significantly improving the prediction accuracy of key indicators such as inflation, consumption growth, and GDP growth rate. Furthermore, supervised learning methods are also applied to sentiment measurement of text data, such as extracting sentiment features from analyst research reports to support economic decision-making [23].

For corporate performance prediction, Athey and Imbens [11] noted that Random Forest handles data sparsity effectively. Sidki et al. [24] found that management experience significantly affects performance, suggesting that feature selection must incorporate domain knowledge. It also aids macroeconomic modeling by capturing dynamic non-linear relationships in production functions [25].

Comprehensive Applications in Business Management. In the field of business management, supervised learning methods have been widely applied to variable measurement and decision support. For example, scholars have utilized Support Vector Machines (SVM) to analyze the spoken text of CEOs to measure their personality traits, further investigating the impact of these traits on mergers and acquisitions (M&A) and corporate governance decisions. These applications not only enhance the scientific rigor of decision-making but also offer new perspectives for corporate management.

In customer churn prediction, classification algorithms (such as Logistic Regression)—by utilizing user behavior labels—predict churn rates more accurately than traditional RFM models [22]. Meanwhile, regression models can predict demand fluctuations to achieve supply chain optimization, though they must address the issue of label noise (e.g., errors in historical sales data) [26].

Potential Value and Future Directions. Despite the vast application potential of supervised learning methods across various domains, practical implementation still faces numerous challenges that necessitate further research. Key factors for realizing their potential value include data privacy and security, model interpretability, and generalization capabilities [6]. To overcome these challenges and fully leverage the advantages of supervised learning, the following aspects should be addressed:

Efficiency Enhancement in Decision-Making: Supervised learning can enhance decision-making efficiency through dynamic adjustment and interpretability strengthening. For instance, real-time model updates (such as online learning) allow models to adapt to market fluctuations, thereby improving the timeliness of decisions [18]. Meanwhile, utilizing tools such as SHAP values can enhance the interpretability of "black-box" models (e.g., neural networks), ensuring that their decision logic meets auditing and regulatory requirements [17].

Interdisciplinary Integration: The application scope of supervised learning can be expanded through integration with other disciplines. For example, combining it with causal inference methods, such as Difference-in-Differences (DID), helps distinguish correlation from causation, providing a more reliable basis for policymaking and business decisions [11]. Furthermore, multimodal data processing is a critical developmental direction. By integrating financial report texts (Natural Language Processing, NLP), panel data, and image information, comprehensive prediction systems can be constructed to capture information from data more holistically [13].

Risks and Challenges: Nevertheless, supervised learning faces certain risks in practical applications. First, there are high requirements for data quality; labeling errors can lead to model bias, such as the common imbalance issue in accounting fraud samples. Second, model overfitting is a frequent problem, requiring the control of model complexity through methods like regularization (e.g., Elastic Net) or cross-validation [10]. Additionally, in the financial sector, it is essential to ensure that models comply with regulatory requirements, such as Basel III risk control standards [18].

Summary. The application of supervised learning in various sub-fields of business management has gradually transitioned from auxiliary tools to core decision-support systems. Its value is primarily manifested in three dimensions: First, at the methodological level, supervised learning breaks through the limitations of traditional statistical assumptions and is capable of handling complex non-linear relationships and high-dimensional data. Second, at the practical level, it significantly improves prediction accuracy and automation, reduces manual intervention, and enhances decision-making efficiency. Finally, at the strategic level, supervised learning promotes innovation in data-driven management models, helping enterprises better adapt to market changes and optimize resource allocation.

Future research directions should prioritize three key areas: exploring transfer learning in small-sample scenarios to effectively address model training challenges when labeled data is limited, enhancing model transparency to both satisfy stringent regulatory requirements and improve the interpretability of complex models, and pursuing theoretical integration with traditional econometrics to further strengthen the scientific rigor and practical utility of these models.

4. Conclusion

Main Findings of Literature Analysis. Through a systematic literature review, this paper provides an in-depth analysis of the current application status, advantages, and challenges of big data analytics in the fields of business and management. The research finds that big data analytics demonstrates significant application value in accounting, economics, and finance. In accounting, big data technology—utilizing data mining and supervised learning methods—effectively improves the efficiency of financial data processing, accurately identifies financial risks, and optimizes auditing processes [19]. In the field of economics, big data provides richer data sources and more precise analytical tools for macroeconomic forecasting and market trend analysis, significantly enhancing predictive accuracy [22]. In finance, big data analytics optimizes portfolio management, credit risk assessment, and market trend forecasting, thereby improving risk management capabilities and investment returns for financial institutions [14].

Furthermore, supervised learning methods, as critical tools in big data analytics, play a key role in various sub-fields of business management. Their advantages—including high-precision prediction, automated feature engineering, and scalability—enable them to handle complex high-dimensional data and non-linear relationships, providing robust support for decision-making [11]. However, practical applications still face challenges such as data privacy and security, as well as model interpretability and generalization capabilities [21].

Importance of Big Data Analytics in Business Management. Big data analytics has become an indispensable tool in the field of business management, and its importance is primarily reflected in several key aspects. First, through real-time data updates and dynamic model adjustments, big data analytics enables a rapid response to market changes, providing enterprises with precise decision support and significantly enhancing decision-making efficiency. Second, big data analytics helps enterprises better understand market demands and consumer behavior, thereby optimizing resource allocation and improving operational efficiency. Additionally, in the finance and accounting sectors, big data analytics significantly strengthens risk management capabilities through fraud detection and credit risk assessment. Finally, big data analytics drives innovation in data-driven management models, offering new research directions and practical pathways for both enterprises and the academic community.

Future Research Directions and Recommendations. Future research should prioritize several key directions. First, investigating how to enhance model generalization through transfer learning in scenarios with limited data remains a vital subject for future study. Second, to satisfy regulatory requirements, further research is needed to improve the interpretability of complex models (such as neural networks) to achieve model transparency. Finally, by combining causal inference methods with multimodal data processing technologies, more comprehensive prediction systems can be constructed to enhance both the accuracy and practical utility of these models.

In summary, big data analytics possesses profound theoretical and practical value in the fields of economics and management. Through continued research and application, it is expected to create even greater benefits for both enterprises and society at large.

References

[1] S.F. Wamba, S. Akter, A. Edwards, G. Chopin, D. Gnanzou, How ‘big data’ can make big impact: Findings from a systematic review and a longitudinal case study, Int. J. Prod. Econ. 165 (2015) 234-246.

[2] G. George, M.R. Haas, A. Pentland, Big data and management, Acad. Manage. J. 57 (2014) 321-326.

[3] A. Merendino, S. Dibb, M. Meadows, L. Quinn, D. Wilson, L. Simkin, A. Canhoto, Big data, big decisions: The impact of big data on board level decision-making, J. Bus. Res. 93 (2018) 67-78.

[4] P. Mikalef, M. Boura, G. Lekakos, J. Krogstie, Big data analytics and firm performance: Findings from a mixed-method approach, J. Bus. Res. 98 (2019) 261-276.

[5] G. Vial, Understanding digital transformation: A review and a research agenda, in: Managing Digital Transformation, 2021, pp. 13-66.

[6] S. Mullainathan, J. Spiess, Machine learning: An applied econometric approach, J. Econ. Perspect. 31 (2017) 87-106.

[7] C. Xu, X. Pang, J. Hu, K. He, Exploring the Accounting Informatization in the "Internet +" Era from the Financial Perspective, Front. Bus. Econ. Manag. 5 (2022) 85-88.

[8] L. Gruber, G. Kastner, Forecasting macroeconomic data with Bayesian VARs: Sparse or dense? It depends!, Int. J. Forecast. (2025).

[9] Z. Wang, Y. Li, Z. Cui, W. Zheng, T. Wang, A machine learning-based study of credit risk in supply chain finance of listed service-oriented enterprises in China, Pac. Basin Finance J. (2025) 103043.

[10] F.J. Bargagli-Stoffi, J. Niederreiter, M. Riccaboni, Supervised learning for the prediction of firm dynamics, in: Data Science for Economics and Finance, Springer International Publishing, Cham, 2021, pp. 19-41.

[11] S. Athey, G.W. Imbens, Machine learning methods that economists should know about, Annu. Rev. Econ. 11 (2019) 685-725.

[12] S. Gu, B. Kelly, D. Xiu, Empirical asset pricing via machine learning, Rev. Financ. Stud. 33 (2020) 2223-2273.

[13] A. Shah, A. Hiray, P. Shah, A. Banerjee, A. Singh, D. Eidnani, S. Chava, Numerical Claim Detection in Finance: A New Financial Dataset, Weak-Supervision Model, and Market Analysis, arXiv preprint arXiv:2402.11728 (2024).

[14] A. Gupta, V. Dengre, H.A. Kheruwala, M. Shah, Comprehensive review of text-mining applications in finance, Financ. Innov. 6 (2020) 39.

[15] S. Lessmann, B. Baesens, H.V. Seow, L.C. Thomas, Benchmarking state-of-the-art classification algorithms for credit scoring: An update of research, Eur. J. Oper. Res. 247 (2015) 124-136.

[16] T. Fischer, C. Krauss, Deep learning with long short-term memory networks for financial market predictions, Eur. J. Oper. Res. 270 (2018) 654-669.

[17] N. Husnaningtyas, T. Dewayanto, Financial fraud detection and machine learning algorithm (unsupervised learning): systematic literature review, J. Riset Akuntansi dan Bisnis Airlangga 8 (2023).

[18] S. Sun, R. Wang, B. An, Reinforcement learning for quantitative trading, ACM Trans. Intell. Syst. Technol. 14 (2023) 1-29.

[19] S.F. Wamba, A. Gunasekaran, S. Akter, S.J. Ren, R. Dubey, S.J. Childe, Big data analytics and firm performance: Effects of dynamic capabilities, J. Bus. Res. 70 (2017) 356-365.

[20] S. Kumar, M. Marrone, Q. Liu, N. Pandey, Twenty years of the International Journal of Accounting Information Systems: A bibliometric analysis, Int. J. Account. Inf. Syst. 39 (2020) 100488.

[21] E. Liaras, M. Nerantzidis, A. Alexandridis, Machine learning in accounting and finance research: a literature review, Rev. Quant. Finance Account. 63 (2024) 1431-1471.

[22] D. Gangwani, X. Zhu, Modeling and prediction of business success: a survey, Artif. Intell. Rev. 57 (2024) 44.

[23] A.H. Huang, A.Y. Zang, R. Zheng, Evidence on the information content of text in analyst reports, Account. Rev. 89 (2014) 2151-2180.

[24] M. Sidki, L. Boerger, D. Boll, The effect of board members' education and experience on the financial performance of German state-owned enterprises, J. Manag. Gov. 28 (2024) 445-482.

[25] V. Babenko, A. Panchyshyn, L. Zomchak, M. Nehrey, Z. Artym-Drohomyretska, T. Lahotskyi, Classical machine learning methods in economics research: Macro and micro level example, WSEAS Trans. Bus. Econ. 18 (2021) 209-217.

[26] D. Hoang, K. Wiegratz, Supply Chain: Optimize the Production Cost Using Machine Learning, Int. J. Prof. Bus. Rev. 8 (2023) e03756.