Large Language Models for Corporate Financial Distress Prediction: Overview and Exploration

Weiya Fu *

Business School, Sichuan University, Chengdu, China

* Corresponding Author Email: fuweiya@stu.scu.edu.cn

Abstract. With the increasing complexity of the business environment and the evolution of information disclosure tools, financial distress prediction (FDP) is gradually transforming from structured data-driven to semantic information fusion. Traditional models rely on financial ratios and statistical indicators, which make it difficult to capture risk propensity in “soft signals” such as management tone and textual metaphors. And the existing large language models (LLMs) provide a new perspective for the text-driven FDP system by the excellent semantic modeling and inference generation capabilities. This paper systematically sorts out the application path of LLMs in FDP by focusing on variable construction and model construction. Three types of representative text features, namely, emotional tone, semantic embedding, and generative variables, are summarized. The modeling mechanism analyzes LLMs as categorical predictive models and their fusion patterns in multimodal integrated systems. In addition, this work points out that there are still challenges such as scarce data labels, non-interpretable models, high cost of system deployment and lack of compliance mechanisms in existing studies, which urgently requires the evolution towards an intelligent early warning system with high credibility, transparency and adaptability under the synergistic promotion of multidisciplinary efforts. This work will provide a cutting-edge reference for constructing intelligent risk control systems and developing financial regulatory technology.

Keywords: Financial distress prediction; big language models; textual features; semantic variable construction; modeling mechanisms; application challenges.

1. Introduction

Against the backdrop of an increasingly complex business environment and rising external uncertainty, financial distress prediction (FDP) is becoming a key issue in corporate governance, capital allocation, and regulatory intervention. Zhao et al. point out that this topic has received a great deal of attention in the industry since the beginning of the 20th century, and the results have a direct impact on the allocation of credit resources, the construction of confidence in the market, and the efficiency of the policy response [1]. Consequently, FDP modeling methods have evolved, and their influence in the core framework of FinTech and RegTech has gradually increased [2].

The traditional FDP methods rely on structured financial data, static assessment of corporate status through ratio indicators, and multivariate statistical models, which makes it difficult to cope with the increasingly hidden financial manipulation and information asymmetry. To compensate for the prediction blindness of structured data, researchers have tried to explore the incremental value of management's language in risk identification by introducing “soft information” such as the text of annual reports and the wording of public announcements [3, 4]. The results exemplify the important complementary value of textual data in the task of financial early warning, with particular advantages in scenarios where traditional variables fail or where risks are more insidious. In addition, large language models (LLMs) based on the Transformer architecture have brought paradigm innovation to financial text analysis. Their ability in semantic understanding, contextual memory, and inference generation enables deeper mining of unstructured information, which becomes an important tool for building FDP systems based on “soft information” [5, 6].

However, there is a lack of systematic sorting and inductive analysis in the academic community on the innovative mechanisms of LLMs in variable construction, the integration paths in predictive modeling, and the applicability and challenges in actual deployment, which makes the potential

capabilities of LLMs in FDP tasks not yet fully utilized. Therefore, this paper will focus on the following three key issues:

(1) At the level of variable construction, how do LLMs support the leap from traditional textual features to higher-order semantic variables?

(2) In terms of modeling mechanism, how can LLMs be embedded into the existing FDP modeling system and realize the fusion application with structured data?

(3) In practical application, what are the key technical and institutional bottlenecks in applying LLMs to FDP missions?

To systematically respond to the above research issues, the structure of this paper is organized as follows: Section 2 reviews the basis of FDP and LLM; Section 3 sorts out the application paths and technical practices of LLM in FDP; Section 4 analyzes existing challenges and future directions; Section 5 concludes the paper.

2. Review of developments

2.1. Theoretical and methodological evolution of FDP

FDP has many definitions, but in recent years, FDP issues are usually treated as binary classification tasks, such as insolvent/non-insolvent or healthy/ distressed. In terms of methodological paths, early studies mostly used financial ratio indicators as the main feature variables [7, 8]; after entering the 21st century, the models were gradually transformed from statistical learning to machine learning, which significantly improved the nonlinear fitting and automatic feature extraction capabilities [9, 10].

However, there are still three core problems in the existing research: first, the theoretical foundation is weak, and the definition of “financial distress” lacks uniformity, which affects the consistency of the model construction; second, the selection of variables lacks the basis, mostly relying on empirical judgments, which leads to the results of the model do not have the comparability and generalization ability; third, the model and the data limitations are obvious. As most of the existing research adopts a static and linear structure and relies on financial statement data that is easy to manipulate, making it difficult to reflect the real dynamic changes of enterprises [11].

2.2. LLMs capability for FDP

Financial information disclosed by companies often contains more ambiguity or subjective judgments, making it more difficult to identify risky language [12]. Traditional natural language processing methods, such as bag-of-words models and shallow neural networks, suffer from limited feature representation capabilities, such as contextual fragmentation and word order neglect, making it difficult to effectively capture contextual relationships and implicit semantics [13]. In this context, LLMs have become a new tool in financial text analysis and risk identification by virtue of their superiority in semantic understanding, context capture, and automatic feature extraction [14]. In the financial domain, vertical models such as FinBERT [15] and FinSoSent [16], have the advantage of financial domain-specific semantics.

Empirical studies have shown that thanks to the large-scale parameter training and command fine-tuning mechanism of LLMs, they can accurately recognize complex linguistic features such as negative intonation, fuzzy phrasing, and context jumps. LLMs are able to accurately recognize complex linguistic features and therefore show unique advantages in discovering potential risk expressions and constructing FDP systems based on 'soft information' [17].

3. Pathways for LLMs in FDP

3.1. Variable construction paths: a reconstruction of text-driven variable systems

With the contextual semantic depth modeling capability of LLMs, the variable construction path of financial texts is shifting from shallow lexical matching to deep semantic representation and generative reasoning. The current mainstream methods can be roughly summarized into three types of variable systems: emotional tone, semantic embedding, and generation.

(1) Emotional tone: This type is one of the earliest financial text modeling features, and the introduction of LLMs significantly improves their contextual understanding [18]. Negative sentiment regulation indicators [19], information asymmetry indicators [20], etc., constructed based on LLMs are highly correlated with post-announcement market volatility, reflecting the feasibility of LLMs in the construction of emotional tone features.

(2) Semantic embedding: based on the text encoding (word2vector) technology, unstructured text can be transformed into structured semantic vectors, providing input features with rich semantics for FDP. The constructed semantic vectors can be used for enterprise classification [21] and innovation intensity indicator construction [22] to analyze the future performance of enterprises.

(3) Generation: This type relies on the reasoning ability of LLMs to directly output quantifiable expected results. Related studies based on LLMs to generate market and macro-expectation indicators [23], inflation expectations [24], etc., can accurately identify the future direction of the market, providing a reasonable and effective reference.

Existing studies have preliminarily verified the effectiveness of LLMs in variable construction, and the constructed text features can effectively make up for the identification blindness of structured financial data in terms of subjective signals, soft information, and fuzzy expressions, and provide a modeling basis for FDP with stronger explanatory power.

3.2. Model building paths: from predictive classification models to multi-modal integrated systems

(1) Predictive classification model: Based on LLMs' own reasoning and command response capabilities, it can be used to construct judgment criteria or directly perform classification. It outperforms lexicon and shallow models in phishing email determination [25], financial fraud identification [26], and shows stronger generalization ability in semantic discrimination [27]. It is also suitable for high complexity and low label density financial distress classification scenarios.

(2) Multi-modal integrated system: existing single modeling paths are often difficult to achieve semantic understanding and structured data computation at the same time. LLMs can not only process textual information, but also synergistically model with time-series, images, or structured financial data [28], enhancing the comprehensive ability of FDP in capturing multi-source risk factors, identifying interactions among variables, and generating forward-looking judgments.

4. Challenges and Future Directions

4.1. Application bottlenecks and challenges

Although LLMs show great potential in FDP tasks, practical applications still face challenges, mainly in four aspects: data quality, modeling mechanism, system deployment, and compliance risk.

(1) Data quality: Financial texts generally have semantic ambiguity, irregular structure, and scarce annotation, which seriously affect the training efficiency and generalization ability of supervised learning models. In addition, since the existing data is not widely sampled, it will lead to hidden preferences in the model.

(2) Modeling mechanisms: The reasoning process of LLMs lacks transparency [29] and is often regarded as a “black box”. Although the application effect is great, the unexplained operation

mechanism still lacks trust, so it will be treated with caution in the highly sensitive areas of finance, and it is difficult to be widely promoted.

(3) System deployment: Existing LLMs are counting on a larger number of parameters to achieve smarter results, but the consumption of computational resources in the training process and the cost of hardware in the deployment will limit their popularization and application. Users need to strike a balance between model performance and cost, which inadvertently raises the threshold of use.

(4) Compliance risk: The existing financial data puts higher requirements on privacy protection, model bias control, and liability attribution, and the generation mechanism of LLMs has problems such as uncontrollable output and ambiguous legal liability. This may evolve into substantial compliance barriers [30], jeopardizing the financial environment.

4.2. Outlook for future development

With the rapid evolution of LLMs in the financial domain, they are also facing a shift from usable to trusted, controlled, and specialized in the FDP mission. To promote the stable application in high-risk financial scenarios, future research can explore the following four key directions.

(1) Development of vertical domain models: General-purpose LLMs still suffer from weak semantic migration and poor predictive stability in financial scenarios, which in turn gave rise to the financial vertical model system represented by FinGPT. Domain-specific corpus, multi-language adaptation, and task-oriented fine-tuning will become the key path to enhance the generalization ability of vertical domain models [31].

(2) Interpretability and controllability mechanism construction: The “black box” nature of LLMs is still one of the core obstacles to their use in financial risk prediction. In the future, we should construct a model structure with logical traceability and behavioral stability to improve the interpretability of model output in regulatory and liability attribution scenarios [32].

(3) Multi-modal synergy and task integration: Financial information presents multi-modal coexistence characteristics such as structured data, text, images, etc., so breaking the single-modal limitation has become the key to the improvement of modeling capability. In the future, attempts can be made to make the model able to understand both charts and text content to improve the comprehensive performance of financial anomaly detection [33].

(4) Evolution of financial intelligence system: LLMs are gradually becoming the core engine for building humanized financial intelligence. The future artificial intelligence system will be centered on multimodal perception, language understanding, and strategy generation, promoting the system upgrade from information extraction to intelligent decision-making [34].

The application of LLMs in FDP modeling is expanding from the tool layer to the system layer, and future research needs to continue to go deeper in the above directions to realize the synergistic evolution under the framework of technology and system.

5. Summary

This paper reviews the development, application pathways, realistic challenges, and future directions around LLMs in FDP task modeling. The key advances of LLMs in enhancing the depth of risk identification and expanding the boundaries of semantic modeling are systematically sorted out, and their wide adaptability in financial text understanding and prediction is pointed out.

The capability of LLMs is highly compatible with what FDP task need and will greatly facilitate the development of FDP. In the specific application, on the one hand, LLMs can be used to promote the construction of new variable indicator types such as emotional tone, semantic embedding, and generation. On the other hand, LLMs can not only directly participate in dilemma classification tasks as a predictive model, but also gradually evolve to multimodal task integration and even a financial intelligent body system. At the same time, this paper also points out that LLMs are still facing multiple challenges, such as data scarcity, model opacity, high deployment cost, and unsound compliance mechanisms in practical applications, which urgently need to be synergistically promoted at the dual

levels of technical path and institutional guarantee. Further, this paper proposes four development directions: vertical domain model development, interpretability and controllability mechanism construction, multi-modal synergy and task integration, and financial intelligence system evolution, in the hope of promoting the sustainable evolution of LLMs in FDP tasks.

Therefore, the construction of an intelligent early warning system for financial distress with predictive power, interpretability and adaptability not only relies on the evolution of LLMs technology itself, but also requires the integration and innovation of multiple disciplines, such as finance, accounting, artificial intelligence and regulatory science and technology, to synergistically promote the construction of a new generation of intelligent financial risk prevention and control system.

Acknowledgements

The author would like to thank Haichuan Zhao from Sichuan University for his advice on the structure of the paper.

References

[1] Zhao Jinxian, Ouenniche Jamal, De Smedt Johannes. Survey, classification and critical analysis of the literature on corporate bankruptcy and financial distress prediction. Machine Learning with Applications, 2024, 15: 100527.

[2] El Madou Kaoutar, Marso Said, El Kharrim Moad, et al. Evolutions in machine learning technology for financial distress prediction: A comprehensive review and comparative analysis. Expert Systems, 2024, 41(2): e13485.

[3] Li Jiawang, Wang Chongren. A deep learning approach of financial distress recognition combining text. Electronic Research Archive, 2023, 31(8): 4683–4707.

[4] Qiu Yue, He Jiabei, Chen Zhensong, et al. A novel semisupervised learning method with textual information for financial distress prediction. Journal of Forecasting, 2024, 43(7): 2478–2494.

[5] Dong Mengming Michael, Stratopoulos Theophanis C., Wang Victor Xiaoqi. A scoping review of ChatGPT research in accounting and finance. International Journal of Accounting Information Systems, 2024, 55: 100715.

[6] Beckmann Lars, Beckmeyer Heiner, Filippou Ilias, et al. Unusual Financial Communication: ChatGPT, Earnings Calls, and Financial Markets. 2025.

[7] Beaver William H. Financial Ratios as Predictors of Failure. Journal of Accounting Research, 1966, 4: 71-111.

[8] Ohlson James A. Financial Ratios and the Probabilistic Prediction. Journal of Accounting Research, 1980, 18(1): 109-131.

[9] Alaka Hafiz A., Oyedele Lukumon O., Owolabi Hakeem A., et al. Systematic review of bankruptcy prediction models: Towards a framework for tool selection. Expert Systems with Applications, 2018, 94: 164–184.

[10] Qu Yi, Quan Pei, Lei Minglong, et al. Review of bankruptcy prediction using machine learning and deep learning techniques. Procedia Computer Science, 2019, 162: 895–899.

[11] Laitinen Erkki K., Camacho-Miñano María-del-Mar, Muñoz-Izquierdo Nora. A review of the limitations of financial failure prediction research: Revisión de las limitaciones de la investigación sobre predicción de quiebras financieras. Revista de Contabilidad, 2023, 26(2): 255–273.

[12] Bushee Brian J., Gow Ian D., Taylor Daniel J. Linguistic Complexity in Firm Disclosures: Obfuscation or Information? Journal of Accounting Research, 2018, 56(1): 85–121.

[13] Xia Bolun (Namir), Rawte Vipula, Gupta Aparna, et al. FETILDA: Evaluation Framework for Effective Representations of Long Financial Documents. ACM Transactions on Knowledge Discovery from Data, 2024, 18(7): 1–27.

[14] Hajek Petr, Munk Michal. Speech emotion recognition and text sentiment analysis for financial distress prediction. Neural Computing and Applications, 2023, 35(29): 21463–21477.

[15] Huang Allen H., Wang Hui, Yang Yi. FINBERT : A Large Language Model for Extracting Information from Financial Text*. Contemporary Accounting Research, 2023, 40(2): 806–841.

[16] Delgadillo Josiel, Kinyua Johnson, Mutigwe Charles. FinSoSent: Advancing Financial Market Sentiment Analysis through Pretrained Large Language Models. Big Data and Cognitive Computing, 2024, 8(8): 87.

[17] Zhang Boyu, Yang Hongyang, Zhou Tianyu, et al. Enhancing Financial Sentiment Analysis via Retrieval Augmented Large Language Models. 4th ACM International Conference on AI in Finance, 2023, 349–356.

[18] Chen Yifei, Kelly Bryan, Xiu Dacheng. Expected Returns and Large Language Models. 2023.

[19] Cao Sean, Jiang Wei, Yang Baozhong, et al. How to Talk When a Machine Is Listening: Corporate Disclosure in the Age of AI. The Review of Financial Studies, 2023, 36(9): 3603–3642.

[20] Kim Alex G., Muhn Maximilian, Nikolaev Valeri V. Bloated Disclosures: Can ChatGPT Help Investors Process Information? SSRN Electronic Journal, 2023.

[21] Vamvourellis Dimitrios, Toth Máté, Bhagat Snigdha, et al. Company Similarity using Large Language Models. arXiv, 2023.

[22] Yang Stephen. Predictive Patentomics: Forecasting Innovation Success and Valuation with ChatGPT. 2023.

[23] Bybee J Leland. The Ghost in the Machine: Generating Beliefs with Large Language Models.

[24] Zarifhonarvar Ali. Experimental Evidence on Large Language Models. SSRN Electronic Journal, 2024.

[25] De Rosa Sara, Gringoli Francesco, Bellicini Gabriele. Hey ChatGPT, Is This Message Phishing? 22nd Mediterranean Communication and Computer Networking Conference (MedComNet), 2024, 1–10.

[26] Zou Yi, Shi Mengying, and Chen Zhongjie, et al. ESGReveal: An LLM-based approach for extracting structured data from ESG reports. Journal of Cleaner Production, 2025, 489: 144572.

[27] Hansen Stephen, Lambert Peter John, Bloom Nicholas, et al. NBER WORKING PAPER SERIES. 2023.

[28] Gao Chen, Lan Xiaochong, Li Nian, et al. Large Language Models Empowered Agent-based Modeling and Simulation: A Survey and Perspectives. arXiv, 2023.

[29] Zheng Zifan, Wang Yezhaohui, Huang Yuxin, et al. Attention heads of large language models. Patterns, 2025, 6(2): 101176.

[30] Yao Yifan, Duan Jinhao, Xu Kaidi, et al. A survey on large language model (LLM) security and privacy: The Good, The Bad, and The Ugly. High-Confidence Computing, 2024, 4(2): 100211.

[31] Pahune Saurabh, Chandrasekharan Manoj. Several Categories of Large Language Models (LLMs): A Short Survey. International Journal for Research in Applied Science and Engineering Technology, 2023, 11(7): 615–633.

[32] Sohail Shahab Saquib, Farhat Faiza, Himeur Yassine, et al. Decoding ChatGPT: A Taxonomy of Existing Research, Current Challenges, and Possible Future Directions. Journal of King Saud University - Computer and Information Sciences, 2023, 35(8): 101675.

[33] Feng Zifeng, Hu Gangqing, Li Bingxin, et al. unleashing the power of ChatGPT in finance research: opportunities and challenges. Financial Innovation, 2025, 11(1): 93.

[34] Desai Akshar Prabhu, Mallya Ganesh Satish, Luqman Mohammad, et al. Opportunities and Challenges of Generative-AI in Finance. In 2024 IEEE International Conference on Big Data, 2024, 4913–4920.