Available Online: 14 September 2026 DOI: 10.54254/2977-5701/2026.36817

Artificial intelligence driven mechanisms for reducing information asymmetry in corporate governance

Sihan Wang

The University of Sheffield, Sheffield, UK

rara481846778@gmail.com

Abstract. Agency conflict arising from information asymmetry remains a vital source of governance conflict in corporations, since the insiders usually have access to more relevant and up-to-date information than the external investors, creditors, regulators, and minority stakeholders. In this work, this study presents a new artificial intelligence-based governance intelligence solution that is capable of uncovering information asymmetry gaps, constructing governance relationship networks, and estimating the level of information transparency in each firm-year period. With Chinese A-shares firms from the time interval 2018 to 2024 as the research sample, our model draws its features from annual reports, internal control reports, board statements, regulatory inquiries, ownership structure, management compensation, analyst forecast, and stock trading. As for the methodology, the model utilizes DeBERTa-v3 to extract the governance disclosure information, Sentence-BERT to find out any inconsistency among disclosure variables, Neo4j and graph attention networks to construct the relationship-risk embedding, and LightGBM to classify firms with a higher risk of information asymmetry. From our empirical experiments, it can be concluded that the whole AI-based model yields AUC at 0.874 ± 0.018 and macro-F1 at 0.812 ± 0.021, which is an improvement of 0.116 compared to the conventional governance metrics in terms of AUC.

Keywords: artificial intelligence, corporate governance, information asymmetry, disclosure gap detection, governance transparency

1. Introduction

Information asymmetry in corporate governance is an occurrence where managers, controlling shareholders or internal decision makers have better access to information than the outsiders, which include investors, creditors, auditors, regulators, and minority shareholders. In the listed companies, information asymmetry is normally present in their annual reports, internal control, board decisions, related party transactions disclosures, executive compensation and response to regulatory inquiries. However, despite the fact that governance indicators like board independence, ownership concentration, audit opinion and executive compensation have been widely used for measurement, none of these variables is enough to determine whether the disclosure of information is clear and credible.

Recent disclosure literature shows that the mandatory disclosure can solve information asymmetry only when the information is comparable and useful [1]. At the same time, disclosures have become machine-

Copyright: © 2026 by the authors. Submitted for possible open access publication under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

readable and companies begin to adjust their language in order to facilitate positive machine processing [2]. This creates an entirely new form of governance problem because information asymmetry includes not only lack of information but also vagueness of narrative, evasive disclosure, fragmented governance incidents and hidden relationships between insiders and related parties. According to machine learning studies, there might be low predictability of traditional governance measures without an integration into wider information system [3].

AI provides a practical way to detect such implicit gaps. Text classifiers will be used to select paragraphs that are relevant to corporate governance, semantic models for comparisons between the statements and factual markers, and graph learning techniques to model the complex interrelationships of shareholders, directors, executives, auditors, subsidiaries, and other related parties. Recent literature suggests that the disclosure tone of narratives depends on corporate governance structures [4]. AI implementation would improve the quality of corporate accounting disclosures by enhancing their information processing and transparency [5]. In summary, this paper proposes developing an AI-based governance intelligence system that transforms scattered corporate governance information into transparency scores and risk forecasts. The unique contribution of this research is the integration of the above factors into one consistent AI framework.

2. Literature review

The extant literature on corporate governance and information asymmetry has mainly focused on agency problems, disclosure quality, board monitoring, ownership structure, internal controls, and market reaction. The literature on textual disclosure has shown that corporate reports carry meaningful governance messages, but these messages tend to be scattered throughout long and repetitive documents. Recent studies on AI disclosures have revealed that the processing of high-dimensional filings and corporate reports by machines is superior to their manual review [6].

In addition, machine learning has been applied to the field of governance and financial analysis. These works leverage large language models for summarization of complex disclosures, extraction of risks, and financial statement analysis [7]. This shows that artificial intelligence could save the cost of investors' information processing. But there still exist many examples where text, governance signals, and market factors are regarded separately from each other. On the other hand, a similar trend in technical studies is that more powerful language models and graph representations should be used. For instance, DeBERTaV3 improves contextualized representation with ELECTRA-style pre-training and gradient disentanglement of embedding sharing [8]. Graph attention model allows assigning weights at node level for relational structures. Therefore, hidden control paths could be discovered with graph attention models [9]. Also, recent reviews in financial large language models emphasize the necessity of using different types of data together with language models in financial risk management [10]. Based on the above studies, this paper views corporate governance disclosure as multi-source information systems.

3. Experimental methods

3.1. Data source and variable construction

The sample includes A-share listed firms in China between 2018 and 2024. Documents used for extracting data include annual reports, internal control reports, board resolutions, CEO change announcements, related party transactions announcements, significant lawsuits announcements, and regulatory inquiries from CNINFO, Shanghai Stock Exchange and Shenzhen Stock Exchange. Firm level variables include ownership

concentration, board size, proportion of independent directors, CEO duality, compensation differential, audit opinion, analyst coverage, analyst forecast dispersion, institutional holdings, bid ask spread, volatility of turnover and abnormal return of stock.

The dataset is structured in two layers: firm-year level and firm-event level. Firm-year data represents an annual governance situation which includes board characteristics, ownership structure, audit quality, internal control weaknesses, compensation system, inquiry frequency and disclosure volume. Firm-event data represents the specific governance events including inquiry letter, executives' change, related party transactions, lawsuits, pledges and abnormal stock reaction occurring in ten consecutive days after the event. After dropping financial companies and missing observations on identification variables, the final sample includes 21,486 firm-year observations, 438,720 governance disclosure sentences, 62,415 governance-related events and 2.91 million trading-day observations. The dependent variable is high information asymmetry risk, indicating the highest quartile of an index made up of bid-ask spread, analyst forecast dispersion, abnormal volatility and the presence of future regulation inquiry.

3.2. AI-based governance information gap detection

The proposed AI-based governance information gap detection system has a three-level technical architecture. The first level operates as a governance information extraction module, based on DeBERTa-v3-base. Corporate disclosures are extracted through PyMuPDF. Paragraphs of disclosures are divided using Python 3.11 and spaCy. Every paragraph is then indexed in terms of the company, year, type of disclosure, date of disclosure, heading of the section, and order of the paragraph. A classifier classifies paragraphs to one of the seven categories, namely board supervision disclosure, internal control disclosure, ownership conflict disclosure, executive incentive disclosure, related-party transaction disclosure, regulatory pressure disclosure, and risk disclosure. Each paragraph is also classified into clear disclosure, vague disclosure, evasive disclosure, and risk-warning disclosure.

The second level deals with the gap between the disclosure. Sentence-BERT embeddings are used to compare the textual claims and structured indicators. For instance, the inconsistency between statements that the company's internal control is "effective and complete", while the internal control report shows weaknesses or the auditor provides a qualified opinion is documented as a discrepancy. A disclosure claiming "standardized related-party transactions" is compared with the amounts of transactions, ownership connections, and questions from the inquiry letter. The information gap score for firm i in year t is calculated as shown in Equation (1):

$$G_{i,t} = \frac{\sum_{p=1}^{P_{i,t}} \sum_{q=1}^{Q_{i,t}} w_{p,q}[1-cos(d_{i,t,p},s_{i,t,q})]I(c_p=c_q)}{\sum_{p=1}^{P_{i,t}} \sum_{q=1}^{Q_{i,t}} w_{p,q} I(c_p=c_q)+\varepsilon} + \delta R_{i,t}^{event} \quad (1)$$

where $d_{i,t,p}$ is the embedding of governance disclosure paragraph p, $s_{i,t,q}$ is the embedding of structured evidence item q, $I(c_p = c_q)$ indicates category consistency, $w_{p,q}$ is the reliability weight of the matched evidence, $R_{i,t}^{\text{event}}$ is the event-risk penalty, and ε avoids division by zero.

The third layer is a Neo4j-based governance relationship graph. Nodes include firms, controlling shareholders, directors, executives, auditors, subsidiaries, related parties, inquiry letters, and major governance events. Edges represent shareholding, employment, audit service, transaction, control, guarantee, pledge, and regulatory inquiry relations. A graph attention network implemented in PyTorch Geometric learns relationship-risk embeddings by assigning larger weights to concentrated control links, repeated related-party transactions, auditor changes, and overlapping director-subsidiary relations.

3.3. Transparency score and risk prediction model

The output of the AI detection model results in a firm-year measure of information transparency of firms, measured on a scale of 0 to 100. These measures consist of four components, namely disclosure clarity, disclosure consistency, relationship transparency, and market information gap. Disclosure clarity is based on the distribution of probability generated by the DeBERTa classifier, where a greater share of disclosure increases the scores while vague and evasive language as well as risk-warning language decreases it. Disclosure consistency is based on the semantic difference between the textual claim and its supporting structured evidence. Relationship transparency is calculated using graph attention embeddings, which means that the existence of hidden control, related party linkages, and repeated transaction edges decrease transparency.

The final score is defined as shown in Equation (2):

$$T_{i,t} = 100 \times \sigma(\alpha_0 + \alpha_1 C_{i,t}^{\text{clarity}} + \alpha_2 C_{i,t}^{\text{consistency}} + \alpha_3 C_{i,t}^{\text{relation}} - \alpha_4 M_{i,t}^{\text{gap}} - \alpha_5 G_{i,t}) \quad (2)$$

where $T_{i,t}$ is the transparency score, $C_{i,t}^{\text{clarity}}$ is disclosure clarity, $C_{i,t}^{\text{consistency}}$ is disclosure consistency, $C_{i,t}^{\text{relation}}$ is relationship transparency, $M_{i,t}^{\text{gap}}$ is the market information gap, and $G_{i,t}$ is the AI-detected disclosure gap. A lower transparency score indicates stronger information asymmetry.

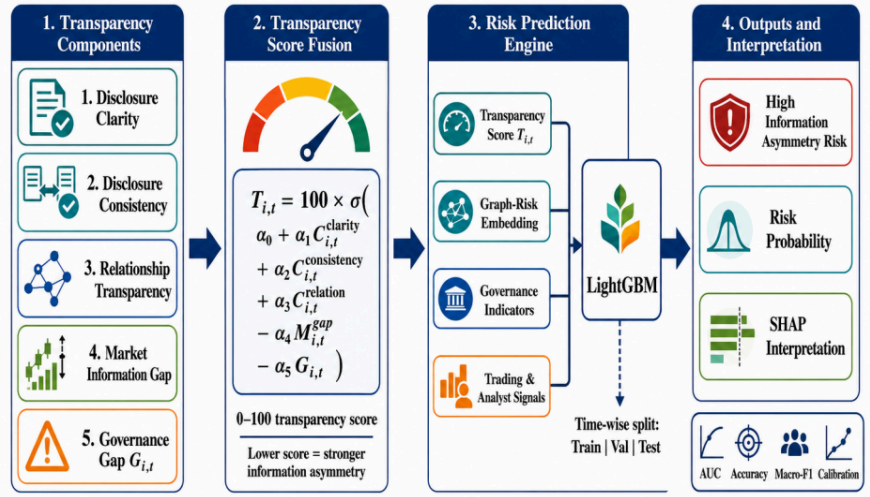

Figure 1. Transparency score fusion and AI-based information asymmetry risk prediction framework

The process involving score fusion, feature combination, and risk evaluation using LightGBM is shown in Figure 1. LightGBM is then used to predict high risk due to information asymmetry. The features vector consists of the transparency score, graph-risk embedding, board, ownership, and compensation-related variables, the audit opinion, inquiry rate, consensus forecast dispersion, bid-ask spread, abnormal turnover, and year-industry dummies. The model is set with 600 trees, tree depth of 6, a learning rate of 0.025, subsampling rate of 0.85, and column sampling ratio of 0.80. The data is split into train, validate, and test sets by date such that the samples prior to 2021 are included in training, those from 2022 in validation, and those from 2023-2024 in testing.

4. Experimental process

4.1. Document parsing and governance data alignment

Every single document is converted to structured text through the tool PyMuPDF. The text blocks are indexed based on firm id, fiscal year, document type, announcement date, heading name, page number, and paragraph sequence. Redundant boilerplate text, disclaimer text, invalid snippets, noise from tables, and duplicates are filtered out before modeling. Paragraphs related to governance are extracted using the DeBERTa classifier and saved to PostgreSQL.

The structured governance data are paired up by firm id and fiscal year. The event data are paired up by announcement date and matched to stock trading data within a ten trading day window. Continuous variables are winsorized by the 1st and 99th percentile. Governance indicator variables with less than 5% missing rate are imputed by the industry-year median value. In order to avoid look-ahead bias, future inquiry letters and future abnormal volatility are not used in score construction but kept for predicting risk labels.

4.2. Model training and score generation

The governance paragraph classifier is trained using manually labeled data. The label dataset consists of 18,000 paragraphs obtained through stratified sampling by industry, year, and document type. Each of three trained annotators classifies a paragraph into one of the following classes: clear disclosure, vague disclosure, evasive disclosure, and risk-warning disclosure. If a paragraph contains a particular governance action, an actor that takes responsibility for it, a specified time frame, and some evidence, it is classified as clear. If the paragraph uses general governance statements without any verifiable information, it is considered vague. If a paragraph provides some response to a governance issue but does not address the primary point, it is considered evasive. If a paragraph demonstrates any risk warning signals related to regulatory, control, litigation, ownership, and transaction risks, it is classified as risk-warning. The last agreement between annotators gives Cohen's kappa equal to 0.826.

DeBERTa-v3-base is fine-tuned on one NVIDIA A100 GPU with batch size = 16, learning rate = 2×10^{-5}, maximum sequence length = 384, warm-up ratio = 0.10, weight decay = 0.01, and 4 training epochs. To overcome class imbalance in the data, the model uses focal loss and weighted sampling since evasive disclosure and risk-warning disclosure classes are less frequent than clear disclosure. DeBERTa-v3-base generates paragraph-level probabilities. Disclosure clarity is calculated as the aggregation of these probabilities at firm-year level.

Graph attention layers are used to train the relationship graph, wherein firms have their own governance features, persons have their roles and tenures as features, and events have their types, amounts, and regulatory severities as features. The output from the graph attention layer results in the generation of a 64-dimensional relationship-risk embedding for each firm-year relationship. The embedding is further combined with the features of disclosure clarity, disclosure consistency, and market information gap. The ultimate transparency score is computed via weighted combination and logistic normalization to be in the range of [0, 100].

4.3. Comparative experiment and robustness test

The research design involves a comparative experiment that utilizes three configurations of features. First, there is the use of pure governance variables, including board independence, ownership concentration, audit opinion, compensation gap, and CEO duality. Second, there is the addition of the AI-extracted textual signals of governance. Third, the complete transparency score, as well as the graph risk embeddings, are added to the

second configuration. Evaluation of the models is done using AUC, accuracy, macro-F1, Brier, and expected calibration error metrics.

The robustness checks involve subgroup analyses for industries, comparison of state-owned and private businesses, removing of financial companies, and yearly evaluations. The additional tests include the model predictions of future inquiry letters, prospective abnormal volatility, and analyst forecast dispersion. The error samples are manually analyzed to see whether the mistakes of predictions are caused by uncertain disclosures, lack of relational information, unexpected governance events, or noise in the market. SHAP analysis is used for the interpretation of the importance of transparency score components.

5. Experimental results

5.1. Governance signal extraction performance

The AI-based extraction model performs better than dictionary matching and BERT baseline models. The improvement is strongest for evasive disclosure and risk-warning disclosure because DeBERTa captures contextual governance meaning rather than isolated keywords (see Table 1). The final classifier achieves 0.891 ± 0.014 accuracy and 0.846 ± 0.018 macro-F1 on the test set.

Table 1. Governance signal extraction performance

Model setting	Accuracy	Macro-F1	Risk-warning recall	Evasive-disclosure recall	Calibration error
Dictionary matching	0.683 ± 0.032	0.594 ± 0.037	0.512 ± 0.044	0.486 ± 0.041	0.118 ± 0.016
BERT-base classifier	0.812 ± 0.021**	0.764 ± 0.024**	0.728 ± 0.031**	0.701 ± 0.034**	0.076 ± 0.011**
RoBERTa classifier	0.847 ± 0.019***	0.801 ± 0.022***	0.769 ± 0.028***	0.742 ± 0.029***	0.061 ± 0.009***
DeBERTa-v3 governance classifier	0.891 ± 0.014***	0.846 ± 0.018***	0.824 ± 0.023***	0.793 ± 0.025***	0.043 ± 0.007***

Note: *** and ** indicate statistical significance at the 0.1% and 1% levels respectively.

5.2. Information asymmetry risk prediction results

Adding AI-extracted textual signals improves the prediction of high information asymmetry risk. The full model with transparency score and graph-risk embeddings achieves the best performance, with AUC of 0.874 ± 0.018 and macro-F1 of 0.812 ± 0.021. The Brier score declines from 0.186 ± 0.013 to 0.128 ± 0.009, indicating better probability calibration.

5.3. Shap-based interpretation of transparency scores

SHAP analysis shows that disclosure consistency is the strongest transparency component, followed by relationship transparency, inquiry frequency, analyst forecast dispersion, and bid-ask spread. When the transparency score falls below 45.00, predicted high-asymmetry risk increases by 0.136 ± 0.029. Firms with dense related-party edges and low disclosure consistency show an additional risk contribution of 0.084 ± 0.017*. The results suggest that information asymmetry is not only a market microstructure problem, but also a governance communication and relationship-structure problem.

6. Conclusion

This study constructs an artificial intelligence-driven governance intelligence system for reducing information asymmetry in corporate governance. By integrating textual disclosure extraction, structured indicator alignment, governance relationship graphs, transparency scoring, and risk prediction, the study transforms fragmented corporate governance information into measurable firm-year transparency indicators. The experimental results show that DeBERTa-v3 improves governance signal extraction, while graph attention embeddings capture hidden relationship-risk structures that traditional governance variables cannot fully represent. The full model provides stronger prediction of high information asymmetry risk than models based only on board, ownership, audit, and compensation indicators. The findings suggest that AI can support investors, creditors, auditors, regulators, and minority shareholders by identifying vague disclosure, inconsistent governance claims, opaque related-party structures, and market-level information gaps.

References

[1] Christensen, H. B., Hail, L., & Leuz, C. (2021). Mandatory CSR and sustainability reporting: Economic analysis and literature review. *Review of Accounting Studies, 26*(3), 1176–1248.

[2] Cao, S., Jiang, W., Yang, B., & Zhang, A. L. (2023). How to talk when a machine is listening: Corporate disclosure in the age of AI. *The Review of Financial Studies, 36*(9), 3603–3642.

[3] Gow, I. D., Larcker, D. F., & Zakolyukina, A. A. (2023). How important is corporate governance? Evidence from machine learning. SSRN Working Paper.

[4] Hasan, A., Sufi, U., Elmarzouky, M., & Hussainey, K. (2024). The impact of corporate governance on narrative disclosure tone: A machine learning approach. *Journal of Applied Accounting Research, 26*(3), 577–604.

[5] Yang, T., & Zhou, N. (2025). Artificial intelligence and the quality of corporate accounting information disclosure. *Finance Research Letters, 85*, 108136.

[6] Cao, S., Jiang, W., Wang, J., & Yang, B. (2024). From man vs. machine to man + machine: The art and AI of stock analyses. *Journal of Financial Economics, 160*, 103910.

[7] Kim, A. G., Muhn, M., & Nikolaev, V. V. (2024). Bloated disclosures: Can ChatGPT help investors process information? *SSRN Working Paper*.

[8] He, P., Gao, J., & Chen, W. (2021). DeBERTaV3: Improving DeBERTa using ELECTRA-style pre-training with gradient-disentangled embedding sharing. *arXiv* preprint arXiv:2111.09543.

[9] Brody, S., Alon, U., & Yahav, E. (2022). How attentive are graph attention networks? *International Conference on Learning Representations*.

[10] Lee, J., Stevens, N., Han, S. C., & Song, M. (2024). A survey of large language models in finance. *arXiv* preprint arXiv:2402.02315.