Transactions on Computational and Scientific Methods | Vo. 4, No. 11, 2024
ISSN: 2998-8780
https://pspress.org/index.php/tcsm
Pinnacle Science Press

Integrating Large Language Models and Knowledge Graphs for Intelligent Financial Regulatory Risk Identification

Yi Wang
Columbia University, New York, USA
yw3802@columbia.edu

Abstract: This paper addresses the complexity and uncertainty in financial regulatory risk identification and proposes an intelligent recognition mechanism that integrates large language models with knowledge graphs. The study first analyzes the limitations of traditional methods in handling multi-source heterogeneous financial data. It points out that relying only on semantic modeling or rule-based constraints cannot fully capture potential risk signals. To overcome this, the proposed framework combines the semantic understanding of large language models with the structured reasoning of knowledge graphs. It is designed to model both unstructured text information and entity relationship networks simultaneously. In terms of method design, a self-attention mechanism is introduced to enhance contextual modeling. A graph convolutional network is used to embed and propagate entities and relations within the knowledge graph. A fusion strategy is then applied to achieve complementarity between semantics and knowledge. This ensures both accuracy and interpretability in risk signal identification. In the experimental part, the study conducts multidimensional comparisons and sensitivity analyses, including hyperparameter settings, environmental conditions, and data perturbations, to validate the effectiveness and robustness of the proposed method. The results show that the method outperforms traditional models in core metrics such as AUC, ACC, F1-Score, and Precision. It also maintains stable performance under various conditions, demonstrating its ability to adapt effectively to complex financial regulatory scenarios. This study not only confirms the rationality of semantic and knowledge integration but also highlights its potential in processing multi-source financial information and improving identification accuracy. It provides a new technical path for building intelligent frameworks for risk identification.

Keywords: Financial regulation; risk identification; large language model; knowledge graph

1. Introduction

In today's complex and dynamic financial environment, the identification and control of regulatory risks have become central to ensuring market stability and preventing systemic crises[1,2]. With the diversification of financial products and services and the acceleration of cross-border capital flows, risks in the financial system show greater concealment and complexity. Traditional regulatory approaches that rely on rules and experience often fail to capture potential risks comprehensively. Their limitations are even more apparent when dealing with unstructured data and cross-domain information. At the same time, advances in big data, artificial intelligence, and knowledge engineering have opened new technical paths for financial regulation. Among them, the combination of large language models and knowledge graphs is considered a promising approach for intelligent risk identification in complex financial contexts.

Large language models have shown strong capabilities in semantic understanding and text generation. They can extract semantic features from large and unstructured text and capture contextual logic. However, relying only on semantic modeling of text remains insufficient. In financial regulation, where professional knowledge and domain logic are critical, these models may face ambiguity and limited reasoning. Knowledge graphs, as structured knowledge representations, model financial market elements through entities, relations, and attributes. This provides solid support for semantic reasoning and logical inference. Their integration improves the accuracy of information extraction and risk representation. It also enhances transparency and interpretability in regulatory processes[3].

In practice, risk identification often involves integrating multi-source and heterogeneous data, such as regulatory documents, news reports, corporate disclosures, and transaction records. These data contain rich natural language descriptions as well as complex entity relationships and cross-temporal dependencies. A single model is often unable to fully capture these multidimensional features[4]. A collaborative mechanism between large language models and knowledge graphs allows dual modeling from language to knowledge. The language model extracts potential risk signals from text, while the knowledge graph embeds them into a network for cross-validation and logical supplementation. This complementarity reduces false signals and reveals systemic chains of risk events. It provides stronger evidence for regulatory decision-making.

This integration also drives a transformation in regulatory approaches. Traditional regulation focuses on post-event monitoring and reactive responses. Intelligent risk identification enables early detection and timely warning. By combining contextual understanding from language models with structured reasoning from knowledge graphs, regulators can maintain sensitivity to risks in dynamic markets. This supports proactive and forward-looking supervision. It improves efficiency, reduces the cost and delay of manual review, and strengthens monitoring of new financial products and cross-border businesses. This provides valuable support for preventing financial crises[5].

In summary, the integration of large language models and knowledge graphs for financial regulatory risk identification represents both an application of advances in artificial intelligence and knowledge engineering and a deep innovation of traditional regulatory methods. It establishes a smarter, more transparent, and more efficient framework for risk identification. It effectively compensates for the shortcomings of existing approaches and advances regulation from passive response to proactive prevention. This research carries significant academic value. It also has practical implications for policy design and regulatory implementation, providing technical support and theoretical reference for the stable operation of the global financial governance system.

2. Related Work

In recent years, research on financial regulatory risk identification has gradually shifted from traditional rule-based and statistical methods to intelligent and data-driven approaches[6]. Early studies relied mainly on manually designed risk indicator systems and rule libraries. These methods could reveal common abnormal behaviors to some extent. However, they showed clear limitations in handling multidimensional risks in complex financial environments. With the development of machine learning, more studies have introduced classification and clustering algorithms into financial regulation to detect potential fraud, market manipulation, and compliance issues. Yet these methods largely depend on hand-crafted high-quality features. Their ability to process unstructured text and cross-domain information remains limited, making it difficult to adapt to dynamic and evolving financial markets[7].

The rise of deep learning has moved financial risk identification into a new stage. Models based on recurrent neural networks, convolutional neural networks, and Transformer architectures are now widely applied to financial text analysis and abnormal behavior detection. These methods can automatically extract semantic features and show strong performance in sentiment analysis, event detection, and public opinion monitoring[8]. However, relying only on deep learning still leads to the problem of the "black box." The lack of interpretability in model decisions limits their use in regulatory scenarios. Results without logical support chains cannot be easily converted into regulatory measures. At the same time, the presence of specialized terminology and complex relationships in financial contexts continues to create bottlenecks when models face knowledge-intensive tasks.

The introduction of knowledge graphs has provided new perspectives for financial risk identification. As a form of knowledge representation that systematizes entities, attributes, and relationships, knowledge graphs have been used in finance to construct networks of market participants, trace the evolution of risk events, and build logical frameworks for compliance constraints. Studies show that knowledge graph-based risk identification can reveal hidden chains of risk through entity linking and relation reasoning. This enhances both accuracy and coverage in risk perception. Nevertheless, knowledge graphs still face challenges in information extraction and automatic updating. Their construction and maintenance are costly, and they cannot handle large-scale financial data tasks alone.

Against this background, integrating large language models with knowledge graphs has gained increasing attention. Large language models excel at semantic parsing and context modeling of unstructured text. Knowledge graphs, on the other hand, provide structured knowledge constraints and logical reasoning capacity. Their combination improves both the precision of information extraction and the recognition of risk events. It also enhances interpretability, offering transparent and verifiable evidence for regulators. Existing studies have begun to explore joint modeling of multimodal and multi-source information. They attempt to build dynamic connections among regulatory texts, transaction data, and entity relationship networks[9]. These explorations lay the foundation for more comprehensive and intelligent mechanisms of financial risk identification. They also provide theoretical and methodological inspiration for this study.

3. Proposed Approach

In this study, the core of the method is to integrate the semantic modeling ability of large language models with the structured reasoning ability of knowledge graphs to achieve multi-dimensional identification of risks in financial regulatory scenarios. First, the input financial text is embedded, and context-related semantic vectors are constructed using a large language model. The overall model architecture is shown in Figure 1.

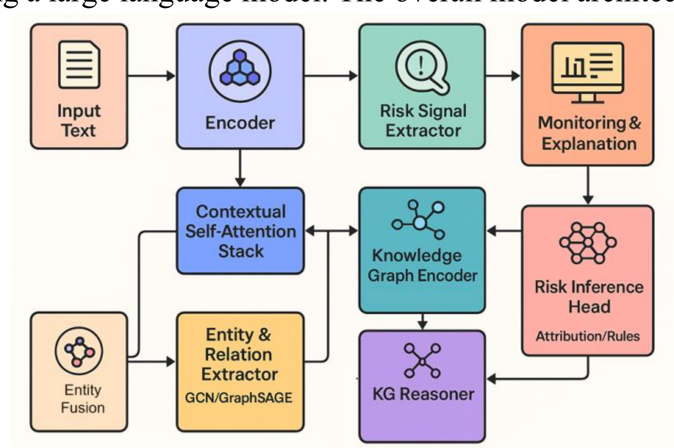

Figure 1. Overall model architecture

Assuming that the input text sequence is $X = \{x_1, x_2, ..., x_n\}$, its semantic representation can be obtained by embedding matrix E and position encoding P:

$$H^{(0)} = E(X) + P$$

Among them, $H^{(0)}$ is passed into the multi-layer Transformer encoder as the initial input to capture global semantic dependencies.

Based on the semantic representation, the self-attention mechanism is used to further model the interaction between words to extract contextual information related to potential risks. For the input feature $H^{(l-1)}$, its attention weight at layer l is calculated as:

$$Attention(Q,K,V) = \text{softmax}(\frac{QK^T}{\sqrt{d_k}})V$$

Among them, $Q = H^{(l-1)}W_Q, K = H^{(l-1)}W_K, V = H^{(l-1)}W_V$ is mapped to different spaces through the weight matrix W_Q, W_K, W_V. This mechanism can capture the potential risk causal relationship in financial text.

To combine semantic hierarchical information with structured relationships in the knowledge graph, this study introduces a graph convolutional network (GCN) to embed the knowledge graph. Assuming that the knowledge graph consists of an adjacency matrix A and a node feature matrix X_g, the update rule at layer l is:

$$H_g^{(l)} = \sigma(\widetilde{D}^{-\frac{1}{2}}\widetilde{A}\widetilde{D}^{-\frac{1}{2}}H_g^{(l-1)}W_g^{(l)})$$

Where $\widetilde{A} = A + I$ is the adjacency matrix after adding a self-connection, $\widetilde{D}$ is its degree matrix, $W_g^{(l)}$ is the learnable weight, and $\sigma(\cdot)$ is the nonlinear activation function. This process can propagate entity relationship information in the graph structure, providing background knowledge support for risk identification.

When the semantic representation is fused with the knowledge graph representation, a weighted splicing and matching mechanism is adopted. Specifically, let the semantic vector be represented as h_t and the knowledge graph entity embedding be h_g, then the fusion vector is defined as:

$$h_f = \alpha h_t + (1-\alpha) \cdot h_g$$

Among them, $\alpha \in [0,1]$ is a balancing factor used to control the fusion ratio of semantic and knowledge information. Finally, the fused representation is input into the classifier to predict the potential risk label, which is formalized as:

$$y = \text{softmax}(W_f h_f + b_f)$$

Here, W_f and b_f are weight and bias parameters, respectively. This method achieves an organic combination of semantic information and structured knowledge, thus having stronger risk identification and reasoning capabilities in financial supervision tasks.

4. Performance Evaluation

4.1 Dataset

In this study, the dataset selected is the Financial PhraseBank. It mainly consists of financial news texts, covering corporate financial reports, market announcements, and industry news. The dataset provides sentiment polarity annotations for financial texts, divided into positive, negative, and neutral categories. These labels capture how market sentiment is reflected at the textual level. This feature makes it an important resource for financial risk identification tasks, especially suitable for financial semantic modeling and risk signal mining.

The dataset is of moderate size. It contains a rich collection of financial texts and offers a clear labeling system. Compared with general corpora, the Financial PhraseBank shows clear advantages in domain coverage and contextual adaptation. Its sources are concentrated in financial market reports, with strong professional language. This supports better semantic understanding and knowledge extraction in regulatory scenarios. Using this dataset ensures that the model can fully capture the linguistic characteristics of finance during training.

In addition, the Financial PhraseBank offers high scalability and compatibility in research. Its annotation system aligns naturally with risk signal extraction and is suitable for integration with knowledge graph structures. This enables mapping between semantic information and structured knowledge. With this dataset, researchers can not only quantify sentiment in financial texts but also build links between risk events and entity relationships. It thus provides a solid data foundation for subsequent mechanisms of regulatory risk identification.

4.2 Experimental Results

This paper first conducts a comparative experiment, and the experimental results are shown in Table 1.

Table 1: Comparative experimental results

Method	AUC	ACC	F1-Score	Precision
BERT[10]	0.873	0.841	0.836	0.842
ResNet 50[11]	0.851	0.824	0.819	0.826
VGG[12]	0.832	0.808	0.801	0.805
Transformer[13]	0.886	0.853	0.847	0.852
Ours	0.918	0.881	0.875	0.883

The experimental results show significant differences in performance among different methods in the task of financial regulatory risk identification. Traditional convolutional neural networks, such as VGG and ResNet 50, are limited in handling unstructured financial texts. Their AUC values are only 0.832 and 0.851, with relatively low overall accuracy and F1-Score. This indicates that models relying only on visual or local feature extraction cannot fully capture the complex semantic associations in financial contexts. They fail to effectively support the comprehensive identification of potential risk signals.

In contrast, language modeling methods such as BERT and Transformer achieve higher performance, with AUC values of 0.873 and 0.886. Their accuracy and precision are also superior. This demonstrates that deep semantic modeling has clear advantages in capturing subtle differences and contextual relationships in financial texts. Language models show stronger expressive power when dealing with complex syntactic structures and specialized terminology. However, these methods still face limitations when processing multi-source information and knowledge logic. They lack deep integration of entity relationships and structured knowledge.

The method proposed in this study outperforms all existing models across all evaluation metrics. The AUC increases to 0.918, the accuracy reaches 0.881, and the F1-Score remains high at 0.875. These results validate the effectiveness of the fusion mechanism between large language models and knowledge graphs. By introducing structured relation modeling through knowledge graphs, the model captures semantic features while enhancing its understanding of logical chains between risk events. This reduces the interference of false signals and improves both accuracy and stability in risk identification.

Overall, this method strengthens logical reasoning and interpretability while maintaining the advantages of semantic understanding. It provides more reliable technical support for financial regulation. Its superiority lies not only in improvements on individual metrics but also in the systematic integration of language information and knowledge graphs. This approach compensates for the shortcomings of existing methods in cross-domain knowledge utilization. It thus offers a solid foundation for building intelligent and transparent frameworks for financial risk identification.

This paper further gives the impact of the learning rate on experimental results, and the experimental results are shown in Figure 2.

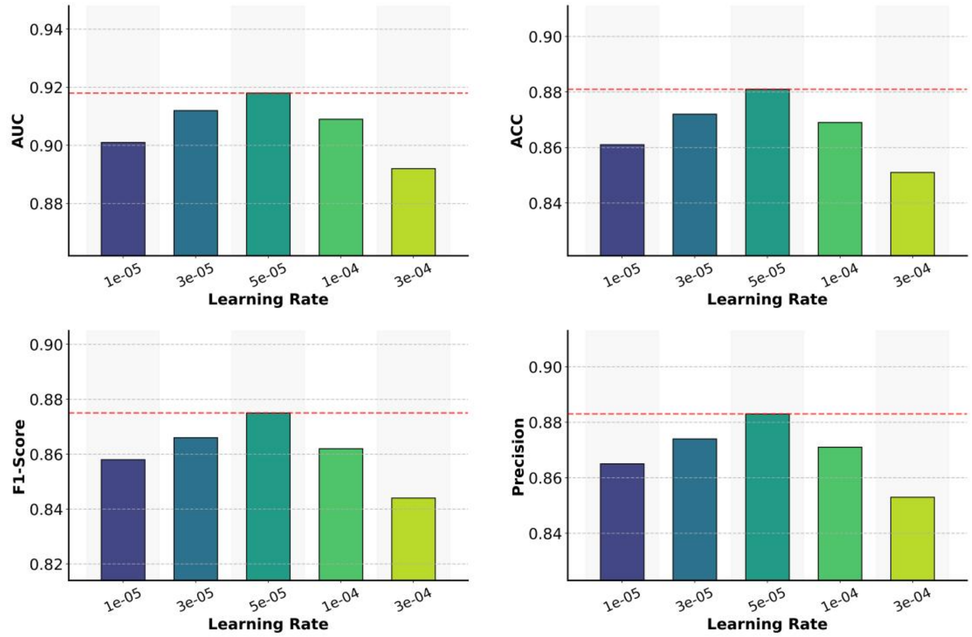

Figure 2. The impact of learning rate on experimental results

The results show that the learning rate has a significant impact on model performance in financial risk identification. When the learning rate is set at a low level (1e-5 and 3e-5), the model maintains stability in AUC, ACC, F1-Score, and Precision, although the values are slightly lower overall. This indicates that an overly low learning rate slows convergence. It avoids oscillation but fails to fully capture complex features from semantic and structured knowledge integration.

When the learning rate increases to 5e-5, all four metrics reach their best levels. The AUC approaches 0.92, and both ACC and Precision achieve higher values. This shows that this setting ensures stability while maximizing the advantages of the collaboration between large language models and knowledge graphs. At this point, the model captures implicit risk signals in text and maps knowledge relations more effectively, reflecting the optimal balance.

With the learning rate further raised to 1e-4, performance declines. All four metrics fall below the results at 5e-5. This suggests that an excessively high learning rate causes gradient oscillation during training. It prevents stable parameter optimization and reduces the accuracy of risk pattern modeling. The effect is especially evident in F1-Score and Precision, which show the model's sensitivity to classification boundaries and fine-grained distinctions.

At a learning rate of 3e-4, the decline becomes more pronounced, with all metrics falling significantly below the optimal point. This indicates that in complex financial regulatory tasks, overly rapid parameter updates damage the synergy between semantic modeling and knowledge reasoning. As a result, the model fails to extract stable and effective risk features. In summary, the learning rate of 5e-5 provides the best training condition for this study. It enables accurate risk identification through the integration of semantic understanding and knowledge structures.

This paper discusses the impact of the knowledge graph fusion weight α on the performance of the proposed method. The analysis emphasizes how the adjustment of α influences the balance between semantic representation and structured reasoning within the model. By varying this parameter, the framework is able to regulate the contribution of language modeling and knowledge graph inference, ensuring that semantic features and relational structures complement each other effectively. The role of α is therefore central in coordinating the interaction between unstructured text processing and structured knowledge integration, and its effect is illustrated in Figure 3.

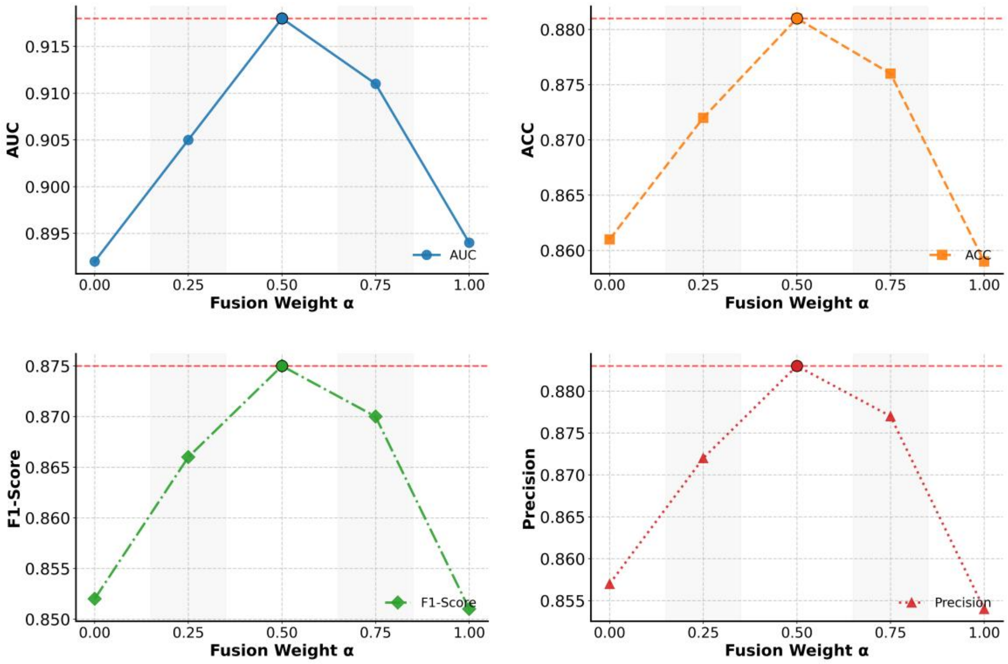

Figure 3. The impact of the knowledge graph fusion weight α on experimental results

The results show that different values of the knowledge graph fusion weight α have a clear impact on model performance. When α is low, the model relies more on semantic representation. In this case, AUC, ACC, F1-Score, and Precision remain at relatively low levels. This indicates that relying only on semantic features cannot fully capture the complex logical relations involved in financial risk identification, which limits model performance.

As α increases, all four metrics show an upward trend and reach the best performance when α is around 0.5. This demonstrates that at this balance point, semantic representation and knowledge graph reasoning achieve the most effective complementarity. Semantic information captures potential risk signals in text, while the knowledge graph strengthens logical reasoning through structured relation modeling. Their combination improves both the comprehensiveness and accuracy of risk identification.

When α further increases to 0.75, performance begins to decline slightly, although it remains at a relatively high level. This suggests that excessive reliance on the knowledge graph may reduce flexibility at the semantic level. The model then struggles to handle the diversity of textual expressions, leading to weaker performance in boundary discrimination and fine-grained modeling under complex contexts. This trend is especially evident in F1-Score and Precision.

When α reaches 1.0, all four metrics drop sharply, showing performance clearly lower than at α = 0.5. This further confirms the limitations of relying solely on structured knowledge. In particular, when processing unstructured financial texts, the lack of semantic features reduces the ability to capture risk signals. Overall, the experimental results show that in financial regulatory risk identification, moderate integration of semantic modeling and knowledge graph reasoning is the key to achieving optimal performance.

This paper also gives the impact of data noise rate on experimental results, and the experimental results are shown in Figure 4.

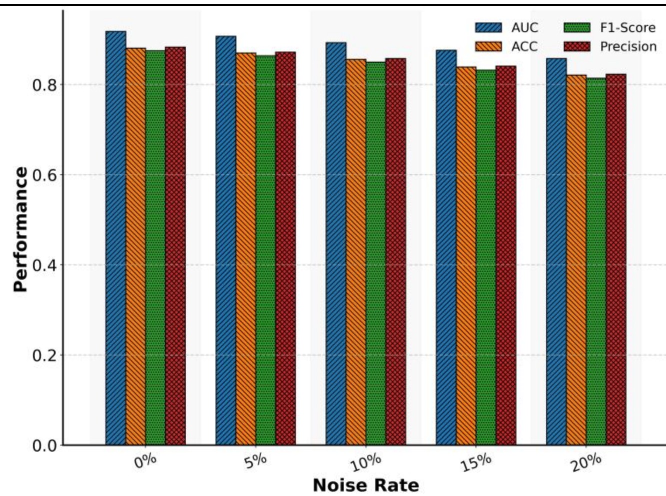

Figure 4. The impact of data noise rate on experimental results

The results show that as the noise rate gradually increases, model performance declines across all metrics. When the noise rate is 0 percent, AUC, ACC, F1-Score, and Precision remain at relatively high levels, reflecting strong performance on clean data. This indicates that the fusion of semantic modeling and knowledge graph reasoning works effectively under ideal conditions. It captures hidden risk signals in financial texts and supports accurate reasoning.

When the noise rate rises to 5 percent and 10 percent, all metrics decrease slightly but remain relatively stable. The fluctuations in AUC and ACC are small. This demonstrates that the method has a certain degree of robustness. It can maintain high risk identification ability even when some labels are incorrect or the data is inconsistent. Such stability is closely linked to the combined use of semantic and structured knowledge, which helps the system remain accurate under uncertainty.

When the noise rate increases to 15 percent, the decline becomes more significant. The drop in F1-Score and Precision is particularly sharp. This suggests that noise in the data has started to interfere with the model's ability to discriminate at class boundaries. The model becomes more vulnerable to mislabeled samples. In financial regulation, this may lead to hidden or misjudged risk signals, reducing the reliability of overall risk perception.

When the noise rate reaches 20 percent, all four metrics drop considerably, showing a clear gap compared with the baseline. This indicates that in high-noise environments, even with semantic modeling and knowledge reasoning, the model struggles to resist the cumulative impact of erroneous information. Overall, the analysis shows that the method is robust under low to moderate noise conditions. However, in extreme noise settings, additional noise-handling mechanisms or regularization strategies are still required to improve reliability in complex real-world regulatory environments.

Finally, this paper also gives the impact of the proportion of training sets on the experimental results, and the experimental results are shown in Figure 5.

The experimental results show that as the proportion of the training set increases, model performance steadily improves across all metrics. When the training ratio is 50 percent, the values of AUC, ACC, F1-Score, and Precision are relatively low. This indicates that with insufficient data, the model cannot fully learn the complex associations between financial texts and knowledge graphs. As a result, its ability to identify risks is limited.

When the training ratio rises to 60 percent and 70 percent, all four metrics improve significantly, with AUC showing particularly notable growth. This demonstrates that with more labeled samples, the model better captures the coupling between semantic features and knowledge structures. It enhances the identification of potential risk signals. At this stage, the integration of semantic modeling and structured reasoning begins to show greater stability and generalization.

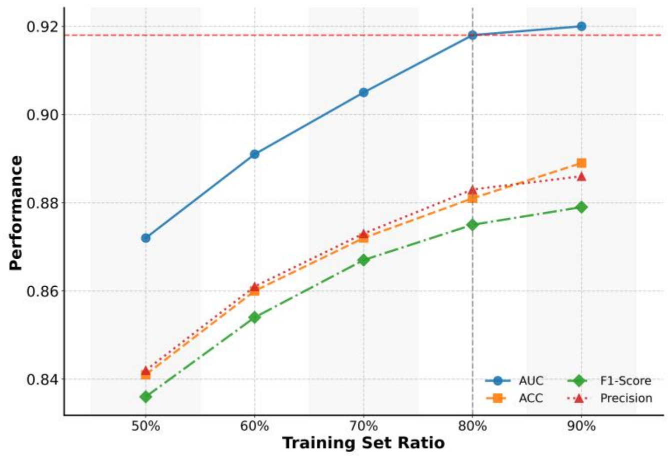

Figure 5. The impact of the training set ratio on experimental results

At 80 percent of the training set, model performance approaches the optimum. The AUC reaches nearly 0.92, while ACC, F1-Score, and Precision also remain at high levels. This result indicates that once the dataset reaches a certain threshold, the model can comprehensively learn semantic patterns and logical relationships in financial risk scenarios. The advantages of the fusion mechanism are fully realized, leading to higher accuracy and stability in risk identification.

When the training ratio increases further to 90 percent, performance improves only slightly, with diminishing gains. This suggests that the model has reached near saturation. Excessively large training proportions bring limited marginal benefits while increasing computational and storage costs. Overall, the results show that expanding the training set to an appropriate scale significantly enhances model performance. Beyond a certain point, however, the benefits decrease, providing practical guidance for data configuration in financial regulatory risk identification.

5. Conclusion

This study focuses on the integration mechanism of large language models and knowledge graphs and proposes an intelligent method for financial regulatory risk identification. By establishing an organic connection between semantic modeling and structured knowledge reasoning, the model captures potential risk signals in financial texts while strengthening the understanding of entity relationships and logical chains. Compared with methods that rely solely on semantics or rules, the proposed approach shows significant advantages in accuracy, stability, and interpretability. It provides a more robust solution for risk identification in complex financial environments.

The results show that deep integration of semantic information and knowledge graphs not only improves risk identification performance but also enhances the model's ability to jointly process unstructured and structured information. This is especially important for financial regulation scenarios where multi-source heterogeneous data are widespread. It helps reduce the interference of information silos and redundant signals. At the same time, the introduction of the fusion mechanism alleviates the "black box" problem of deep learning to some extent. It makes the reasoning path of the model clearer and more transparent, offering regulators stronger trust and operability in practice.

From the application perspective, this method provides solid support for multidimensional risk monitoring and early warning systems in financial markets. In the face of rapidly changing market environments and complex trading behaviors, the model can identify explicit risks and also reveal hidden systemic vulnerabilities. This helps regulatory authorities achieve proactive and forward-looking interventions. Such capabilities are valuable in cross-border capital flows, the regulation of complex financial derivatives, and multi-level market governance. They contribute to improving the precision and intelligence of regulatory systems.

Overall, this study offers a new paradigm for the deep integration of artificial intelligence and knowledge engineering in financial regulation. The proposed risk identification mechanism promotes exploration of the collaborative use of large language models and knowledge graphs at the academic level, while also providing a feasible path for combining financial technology and regulatory technology in practice. Through this mechanism, financial regulation can achieve more efficient and accurate risk identification and management in dynamic environments, thereby offering strong support for maintaining market stability and preventing systemic risks.

References

[1] Chen Y, Wu F, Wang J, et al. Knowledge-augmented financial market analysis and report generation[C]//Proceedings of the 2024 Conference on Empirical Methods in Natural Language Processing: Industry Track. 2024: 1207-1217.

[2] Alghamdi, N. A. and Al-Baity, H. H., "Augmented analytics driven by AI: A digital transformation beyond business intelligence", Sensors, vol. 22, no. 20, p. 8071, 2022.

[3] Liu Z, Zhang Z, Zeng X. Risk identification and management through knowledge Association: A financial event evolution knowledge graph approach[J]. Expert Systems with Applications, 2024, 252: 123999.

[4] Mao X, Sun H, Zhu X, et al. Financial fraud detection using the related-party transaction knowledge graph[J]. Procedia Computer Science, 2022, 199: 733-740.

[5] Zehra S, Mohsin S F M, Wasi S, et al. Financial knowledge graph based financial report query system[J]. IEEE Access, 2021, 9: 69766-69782.

[6] Sarmah B, Mehta D, Hall B, et al. Hybridrag: Integrating knowledge graphs and vector retrieval augmented generation for efficient information extraction[C]//Proceedings of the 5th ACM International Conference on AI in Finance. 2024: 608-616.

[7] Li D F, Xu F. Synergizing knowledge graphs with large language models: a comprehensive review and future prospects[J]. arXiv preprint arXiv:2407.18470, 2024.

[8] Feng Z, Ma W, Yu W, et al. Trends in integration of knowledge and large language models: A survey and taxonomy of methods, benchmarks, and applications[J]. arXiv preprint arXiv:2311.05876, 2023.

[9] Zhu W, Chen Z. An Intelligent Financial Fraud Detection Model Using Knowledge Graph-Integrated Deep Neural Network[J]. Journal of Circuits, Systems and Computers, 2024, 33(15): 2450267.

[10]Devlin J, Chang M W, Lee K, et al. Bert: Pre-training of deep bidirectional transformers for language understanding[C]//Proceedings of the 2019 conference of the North American chapter of the association for computational linguistics: human language technologies, volume 1 (long and short papers). 2019: 4171-4186.

[11]He K, Zhang X, Ren S, et al. Deep residual learning for image recognition[C]//Proceedings of the IEEE conference on computer vision and pattern recognition. 2016: 770-778.

[12]Simonyan K, Zisserman A. Very deep convolutional networks for large-scale image recognition[J]. arXiv preprint arXiv:1409.1556, 2014.

[13]Vaswani A, Shazeer N, Parmar N, et al. Attention is all you need[J]. Advances in neural information processing systems, 2017, 30.