Journal of Computer Technology and Software

ISSN: 2998-2383

Vol. 3, No. 8, 2024

Deep Contextual Risk Classification in Financial Policy Documents Using Transformer Architecture

Ying Qin
University of Illinois at Urbana-Champaign, Champaign, USA
yingqinuiuc@gmail.com

Abstract: This study proposes a Transformer-based method for identifying potential risks in financial policy texts. The method takes financial policy documents as input. It uses embedding layers and positional encoding to transform semantic information into learnable vector representations. Multiple layers of Transformer encoders are then applied to model deep dependencies between words. This allows the model to extract risk-related signals from policy content. To improve classification accuracy, the model introduces a nonlinear projection mechanism. It maps global semantic representations into the risk classification space. The model is optimized using the cross-entropy loss function. In terms of experimental design, a unified training framework is constructed. A publicly available financial text dataset is used to evaluate model performance. The effectiveness and stability of the model are validated through comparative experiments, hyperparameter sensitivity analysis, and attention visualization. The experimental results show that the proposed method outperforms existing mainstream models in precision, recall, and F1-score. It maintains a strong semantic understanding while effectively identifying potential risks in policy language. In addition, the study conducts further analysis on Transformer depth, choice of regularization techniques, and model adaptability across different periods. These findings provide both theoretical and empirical support for developing automated financial risk identification systems for real-world applications.

Keywords: Financial policy text; risk identification; self-attention mechanism; semantic modeling

1. Introduction

In the context of today's highly integrated global economy, the formulation and implementation of financial policies play a vital role in national macroeconomic regulation. Financial policies directly affect market liquidity, credit conditions, and investment expectations[1]. They also indirectly influence corporate operations, the behavior of financial institutions, and the overall stability of socio-economic systems. However, due to the complexity, variability, and time-lag nature of financial policies, it is often difficult to accurately identify and predict their specific impact on risk structures[2]. How to scientifically and efficiently detect financial risks that may be triggered or mitigated by financial policies has become a core issue for regulatory authorities, academic researchers, and financial institutions[3].

Traditional financial risk identification methods often rely on expert knowledge and rule-based techniques such as keyword extraction when dealing with policy texts. These approaches can capture some explicit information in the policy content[4]. However, they are often ineffective in uncovering the deeper implications of the policy. This includes the underlying attitude, adjustment direction, and future orientation implied in policy language. Moreover, with the growing volume of policy documents, manual interpretation is no longer efficient or comprehensive. Traditional methods are increasingly inadequate when processing large-scale and multidimensional policy texts. Therefore, developing an intelligent method that can fully understand the semantic content of policy texts and automatically identify risk signals is of great practical significance and research value[5].

With the rapid advancement of natural language processing technology, deep learning-based language models have made breakthroughs in text understanding and semantic modeling. In particular, the introduction of the Transformer architecture provides a new technical path for in-depth semantic analysis of policy texts. The Transformer has strong context modeling capability, a multi-head attention mechanism, and good scalability. It is particularly effective in handling complex language structures and long-text semantics. Applying the Transformer to the semantic analysis of financial policy texts helps identify sentiment tendencies and content focus. It also enables automatic identification and warning of potential financial risks by modeling correlations with historical risk events[6].

In policy interpretation, the factors influencing risk judgment are often multidimensional and dynamically changing. For example, the type of regulatory tools, frequency, tone intensity, and scope mentioned in a policy may all affect market expectations and the path of risk transmission. Rule-based or shallow models that rely only on static features are insufficient to meet the needs of intelligent analysis of financial policy texts[7]. Transformer-based risk identification methods can capture complex semantic structures and multi-level semantic relationships through training on large-scale corpora. This enables a comprehensive understanding of policy information and accurate risk classification. It improves

automation in analysis and provides a scientific and efficient decision-support tool for financial regulators[8].

From a macro perspective, building a Transformer-based method for assessing the risk impact of financial policies enhances the intelligence level of policy evaluation. It also strengthens the ability to manage potential risks in financial markets proactively. This is especially important in the context of economic cycles and frequent external shocks. The ability to identify and assess systemic financial risks in real time contributes to market stability and sustainable economic development. At the same time, this research direction expands the application of artificial intelligence in finance. It promotes deep integration between financial technology and policy analysis. It has both theoretical significance and practical value.

2. Related work and background

2.1 Financial Risk Identification Analysis

As a key component of financial risk management, risk identification aims to detect potential risk factors through the analysis and modeling of various types of financial data. Traditional risk identification methods mainly rely on classical algorithms such as statistical analysis, logistic regression, and decision trees. These methods perform well when dealing with structured financial data[9]. However, as financial markets evolve and information becomes more complex, risk factors have become more diverse, dynamic, and nonlinear. Traditional approaches are increasingly challenged in identifying deep and hidden risks. This has led researchers to explore more intelligent analytical techniques to extract valuable risk features from massive data sources[10].

In recent years, with the growing digitalization of the financial industry, unstructured data such as policy documents, public opinion, financial disclosures, and news reports have become increasingly important in risk identification. Compared with structured data, unstructured text better reflects market sentiment and hidden expectations. It is a critical source for detecting systemic financial risks[11]. As a result, natural language processing techniques have been introduced into financial risk analysis to help extract potential risk signals from policy texts, regulatory notices, and financial news. However, traditional text mining methods such as TF-IDF, topic models, and sentiment analysis often rely on keyword statistics or shallow semantic modeling. They struggle to capture contextual relationships and deep semantics in text. This limits their effectiveness in complex financial environments[12].

In dealing with complex risk scenarios, deep learning has shown significant potential, particularly in language modeling. Deep learning enables automatic feature extraction and hierarchical structure modeling. It helps identify multi-level semantic features hidden in financial texts and provides richer dimensions for risk detection. When facing risk signals that span time, space, or policy domains, deep models can learn from large volumes of historical data and contextual relationships. They build robust semantic mappings to support risk trend identification and transmission path analysis. This capability is especially critical today, given the frequent policy adjustments and increasing complexity of financial instruments[13].

Overall, the development of financial risk identification has moved from static analysis based on structured data to dynamic identification that integrates unstructured data, multimodal information, and deep semantic understanding. In this evolution, applying natural language processing, especially Transformer-based deep semantic modeling, has become a key direction for improving the accuracy and efficiency of risk identification. This approach expands data sources, enhances the comprehensiveness and robustness of risk modeling, and captures potential risk signals in complex language environments[14,15]. It supports financial supervision, institutional decision-making, and systemic risk prevention. It also drives the intelligent transformation of financial risk management systems.

2.2 Transformer Architecture

Since its introduction, the Transformer architecture has attracted wide attention in the field of natural language processing. Its core advantage lies in the ability to efficiently capture long-range dependencies in sequence data. Unlike traditional recurrent neural networks and convolutional neural networks, the Transformer processes information entirely based on attention mechanisms[16,17]. It eliminates recursive structures in sequential modeling, significantly improving parallel computation and training efficiency. By using multi-head self-attention, the model assigns weighted representations to each word in the input sequence. This allows it to attend to semantic information from different positions simultaneously, forming a more comprehensive contextual representation. This feature is especially important in processing financial policy texts, which often involve complex language structures and strong logical relationships[18]. It enables effective modeling of underlying semantic logic and policy direction.

The Transformer encoder consists of stacked layers of attention modules and feedforward networks. Each layer includes self-attention, residual connections, and layer normalization. These components contribute to the model's training stability and generalization capability. The multi-head attention mechanism allows the model to capture different semantic features in separate subspaces. This leads to a deeper understanding of the multiple semantic dimensions within a text. In financial policy documents, a single paragraph may contain various types of information, such as regulatory intent, applicable scope, and operational tools. Traditional models struggle to decompose and model these dimensions effectively. In contrast, the Transformer can build multi-layered semantic associations within the same input. This enhances its ability to capture the deeper meanings of the text[19].

In addition, the scalability and pretraining capability of the Transformer provide broad potential for applications in finance[20]. Pretraining on large-scale corpora allows the model to learn general language knowledge and semantic representations. Fine-tuning can then adapt the model to specific tasks in the financial domain. This enables a combination of general language understanding and domain knowledge transfer. The pretraining – fine-tuning paradigm lowers the modeling threshold for domain-specific tasks and improves performance in low-resource settings. In financial risk identification, policy texts often show imbalanced

distributions and distinct domain features. This makes the transfer learning approach especially effective and practical[21].

In summary, the Transformer has become one of the most representative models for text representation learning[22,23]. It offers multiple advantages in semantic understanding, contextual modeling, and parallel computation. Its applications in financial text analysis continue to expand. It shows strong performance in tasks such as policy semantic parsing, risk signal detection, and sentiment orientation analysis. In the future, as the volume of financial data continues to grow and policy regulation becomes more complex, the Transformer and its variants will play an increasingly important role. They will enhance the intelligence level of financial risk identification and strengthen systemic risk monitoring. This will provide solid technical support for building the next generation of financial risk warning systems.

3. Transformer architecture

This study proposes a method for identifying the impact of financial policies on risk based on the Transformer architecture, which aims to extract potential risk signals from policy texts and automatically classify them. The overall method framework includes text preprocessing, embedding representation, Transformer encoding, risk identification layer, and final risk classification output. By introducing a multi-layer self-attention mechanism, the model can capture long-distance dependencies and semantic linkages in policy language, thereby achieving high-precision risk identification modeling. The model architecture is shown in Figure 1.

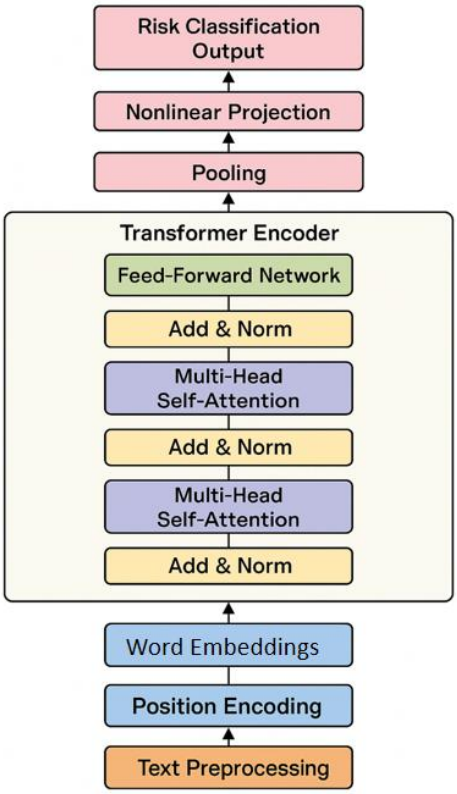

Figure 1. Overall model architecture

First, given a financial policy text sequence $X = \{x_1, x_2, ..., x_n\}$, each word x_i is mapped to a low-dimensional vector representation $e_i \in R^d$. The initial embedding representation of the entire text sequence can be defined as:

$$E = [e_1, e_2, ..., e_n] \in R^{n \times d}$$

To retain the position information in the word sequence, the position code $P \in R^{n \times d}$ is introduced, and the final input vector is:

$$H_0 = E + P$$

Next, the input sequence H_0 is passed to a multi-layer Transformer encoder, each layer of which consists of a multi-head attention mechanism and a feedforward network. For the lth layer, the h-th attention head is calculated as follows:

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

Among them, the query matrix $Q = H_l W^Q$, key matrix $K = H_l W^K$, and value matrix $V = H_l W^V, W^Q, W^K, W^V$ are learnable parameters. The outputs of all attention heads will be concatenated and mapped into a new representation:

$$MultiHead(H_l) = Concat(head_1, ..., head_h) W^O$$

The Transformer layer also includes a feed-forward network to enhance the nonlinear feature modeling capability, which is in the form of:

$$FFN(x) = \max(0, xW_1 + b_1)W_2 + b_2$$

W_1, W_2 and b_1, b_2 are trainable parameters, and ReLU is used as the activation function.

After stacking multiple layers of encoders, the final output is a high-dimensional representation $H_L \in R^{n \times d}$. We perform average pooling on this representation to obtain the global semantic representation of the entire text:

$$h_{text} = \frac{1}{n} \sum_{i=1}^{n} H_{L,i}$$

To achieve risk discrimination, a nonlinear projection layer is introduced to map the semantic representation to the risk space:

$$z = \tanh(h_{text} W_r + b_r)$$

Where $W_r \in R^{d \times k}, b_r \in R^k, k$ represents the risk label dimension.

Finally, the softmax function is used to classify the risks and obtain the predicted probability distribution:

$$\hat{y} = \text{softmax}(z)$$

The training goal is to minimize the cross entropy loss function:

$$L = -\sum_{i=1}^{k} y_i \log(\hat{y}_i)$$

Where y_i is the one-hot encoding of the true label and $\hat{y}_i$ is the model prediction probability.

This method uses the deep semantic modeling capabilities of Transformer to achieve all-around perception and abstraction of policy text information, especially in modeling the potential impact of complex policy language on financial risks. At the same time, the multi-head attention mechanism can perform weighted modeling of different semantic focuses, improving the model's sensitivity to risk signals caused by subtle changes in policy language, thereby providing a scalable technical path and computing basis for the financial risk early warning system.

4. Experimental Results

4.1 Dataset

The text dataset used in this study is the Financial PhraseBank, which consists of news and analytical sentences within a foreign financial context. It is widely used in financial natural language processing tasks and is particularly suitable for applications in risk identification, sentiment analysis, and policy text modeling. The dataset includes representative sentences drawn from various financial domains such as corporate earnings, market trends, and regulatory commentary. Due to its high semantic complexity and strong domain relevance, it serves as an effective source for financial semantic representation and classification.

Texts in the Financial PhraseBank are originally labeled with three sentiment categories: positive, neutral, and negative. To align with the objective of risk identification in policy contexts, we reinterpreted and mapped these sentiment labels into risk tendency categories. Specifically, negative sentiment is associated with high-risk indicators, as such statements often describe market losses, volatility, or regulatory pressure. Neutral sentiment is treated as medium or uncertain risk, reflecting situations with ambiguous or mixed policy impacts. Positive sentiment is considered low risk, typically corresponding to favorable financial developments or stabilizing policy actions. This mapping enables the transformation of a sentiment-annotated corpus into a risk-aware classification dataset.

Since the Financial PhraseBank is sourced from real financial news and official commentary, it exhibits high authenticity in both policy discourse and risk-related language. This enhances the dataset's applicability in identifying implicit risk signals within financial texts. Furthermore, the dataset's formal linguistic style and concise sentence structure make it particularly well-suited for deep semantic modeling using Transformer-based architectures. The use of the original English corpus ensures consistent linguistic quality, providing a reliable textual foundation for training and evaluating our policy-oriented risk identification framework in real-world financial scenarios.

4.2 Experimental setup

This study adopts a unified experimental setup for model training and evaluation. This ensures stability in parameter settings and reproducibility of results. The input text length is set to 512. A trained Transformer encoder is used for semantic extraction. A multi-layer fully connected network is applied for the risk classification task. The Adam optimizer is used with an initial learning rate of 1e-5. The cross-entropy loss function is employed for training. To improve generalization, a Dropout mechanism is introduced with a dropout rate of 0.1. All experiments are implemented on a deep learning framework with GPU acceleration. The batch size is set to 16 and the number of training epochs is 10.

During the experiments, the dataset is split into 80 percent for training, 10 percent for validation, and 10 percent for testing. Evaluation metrics include precision, recall, and F1-score. These metrics provide a comprehensive assessment of the model's performance in the task of financial policy text risk classification. Table 1 presents the key parameter settings used in this experiment.

Table 1: Experimental Configuration Parameters

Parameter	Value
Max Sequence Length	512
Batch Size	16
Learning Rate	0.001
Optimizer	AdamW
Dropout Rate	0.1
Epochs	200
Loss Function	Cross-Entropy
Evaluation Split	80/10/10

4.3 Experimental Results

1) Comparative experimental results

This paper first gives the comparative experimental results, as shown in Table 2.

Table2: Comparative Results

Method	Precision	Recall	F1-Score
LSTM+Attention[23]	83.5	80.2	81.8
CNN+Attention[24]	84.1	79.7	81.7
Bert[25]	87.3	85.6	86.4
RoBerta[26]	88.5	87.5	87.7
Ours	90.2	89.4	89.8

As shown in the comparative experimental results in Table 2, the Transformer-based risk classification model proposed in this study outperforms other methods across multiple evaluation metrics. This demonstrates its significant advantages in semantic understanding and risk identification for financial policy texts. Compared with traditional neural network models, this approach is more effective in capturing deep semantic relationships. It can handle complex expressions in financial language, including ambiguity, implicit attitudes, and strategic intentions. This provides stronger modeling capability for accurate risk detection.

Although LSTM+Attention and CNN+Attention have certain strengths in feature extraction and contextual understanding, their limited ability to model long-range dependencies restricts performance. When processing complex and lengthy financial policy texts, these models show weaker results. In policy sentences with widely distributed semantic clues, they may overlook key decision terms or layered sentiment tendencies. This leads to biased risk judgment. The method in this study uses a global attention mechanism to fully integrate semantic information. This significantly enhances the model's ability to detect implicit risk signals.

Compared with pre-trained language models such as BERT and RoBERTa, this method further incorporates structure optimization tailored to risk semantic modeling. It enables the model to inherit general semantic knowledge while focusing on risk-triggering mechanisms within financial contexts. This targeted modeling strategy improves the model's adaptability and classification accuracy in financial text scenarios. It performs especially well in cases where subtle differences in policy language may lead to changes in risk assessment.

Overall, the experimental results show that the Transformer structure exhibits strong generalization and contextual adaptability in financial policy semantic modeling. The advantages of this study's method in semantic extraction, contextual modeling, and decision signal identification contribute to overall performance improvement in risk classification tasks. These findings confirm the method's effectiveness and provide a feasible path and theoretical foundation for building intelligent financial risk identification systems for policy analysis.

2) Hyperparameter sensitivity experiment results

This paper presents the experimental results of hyperparameter sensitivity experiments. First, an adjustment experiment is conducted on the learning rate. The experimental results are shown in Table 3.

Table 3: Hyperparameter sensitivity experiments (Learning Rate)

Learning Rate	Precision	Recall	F1-Score
0.004	86.5	84.1	85.3
0.003	88.1	86.2	87.1
0.002	89.4	88.0	88.7
0.001	90.2	89.4	89.8

The results of the hyperparameter sensitivity experiments show clear performance differences under various learning rate settings. This indicates that learning rate is a critical parameter in the optimization process. It directly affects the model's ability to extract semantic features from financial policy texts. In risk classification tasks, semantic clues are often implicit and highly context-dependent. Therefore, controlling the learning rate is essential for guiding the model toward optimal convergence during training.

When the learning rate is high, the model updates quickly. However, it may overlook subtle details in policy texts, leading to insufficient capture of the logical structure behind risk signals. This is especially important in real financial policies, where risk judgment often depends on minor changes in wording or shifts in logical relations. As a result, the model shows slightly weaker performance at higher learning rates, reflecting limitations in modeling complex semantics.

As the learning rate decreases, the model gradually shows more stable and detailed semantic modeling capabilities. It becomes better at understanding and integrating multi-level information, especially in long policy texts with implicit directions or complex sentence structures. This suggests that a lower learning rate helps the model more accurately capture risk intentions and underlying logic in financial policy language. It enhances the overall ability to identify risk.

The overall results suggest that achieving high-quality risk classification requires careful control of the learning pace during training. The model needs to balance convergence speed with sensitivity to policy semantics. This tuning strategy not only confirms the flexibility of the Transformer architecture in modeling complex semantics but also strengthens the stability and robustness of the proposed method in intelligent financial risk identification scenarios.

Furthermore, the optimizer experimental results in the hyperparameter sensitivity experiment are given, as shown in Table 4.

Table 4: Hyperparameter sensitivity experiments (Optimizer)

Optimizer	Precision	Recall	F1-Score
AdaGrad	86.8	84.5	85.6
SGD	85.9	83.7	84.8
Adam	89.1	88.0	88.5
AdamW	90.2	89.4	89.8

The results of the optimizer sensitivity experiments show that different optimization strategies have a clear impact on model performance in financial risk classification tasks. The optimizer guides the direction and magnitude of parameter updates. Therefore, choosing an appropriate optimizer is critical for improving overall model performance, especially when dealing with complex semantic structures and high-dimensional text representations. Financial policy texts often contain dense semantic content and strategic language patterns. The ability of an optimizer to capture these details directly affects the precision of risk identification.

Among the optimizers compared, traditional stochastic gradient descent (SGD) shows certain advantages in simple tasks. However, it tends to suffer from gradient oscillation and slow convergence when applied to long texts and deep Transformer structures. This limits its effectiveness. AdaGrad adjusts the learning rate adaptively, but its conservative update strategy may lead to insufficient learning in later stages. This affects the model's ability to extract deep semantic features in complex policy texts.

In contrast, Adam-based optimizers combine momentum and adaptive learning rate mechanisms. They adjust parameters quickly in the early training stage and maintain stability during fine-tuning. AdamW, in particular, introduces weight decay to mitigate overfitting. This allows the model to maintain high generalization and classification accuracy even when handling

diverse and complex policy statements. This feature is especially suitable for financial policy analysis, where both precision and robustness are required.

The experimental results further show that optimizer selection is not merely a matter of parameter tuning. It is directly related to model performance in semantic modeling, context understanding, and risk reasoning. For financial policy-related risk identification tasks, choosing the right optimization strategy not only improves model performance but also lays a foundation for building stable, controllable, and interpretable financial intelligence systems.

3) The impact of different Transformer layers on risk discrimination

This paper further gives the impact of different numbers of Transformer layers on risk identification, and the experimental results are shown in Figure 2.

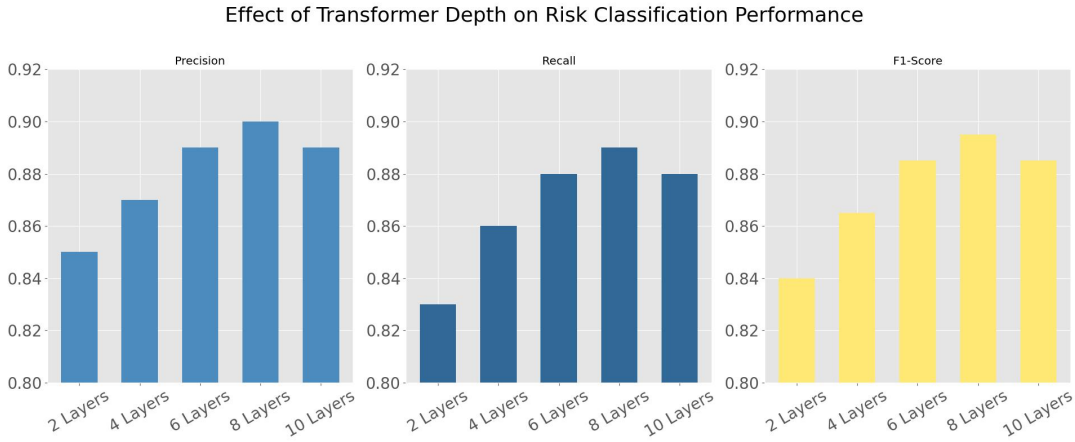

Figure 2. The impact of different numbers of Transformer layers on risk identification

As shown in the experimental results in Figure 2, the number of Transformer layers has a significant impact on the performance of the financial risk identification model. Across different evaluation metrics, the overall performance increases with more layers at first, then stabilizes, and in some cases slightly declines. This trend suggests that increasing model depth can enhance semantic modeling. However, excessive stacking may introduce noise or cause gradient vanishing, reducing the model's sensitivity to key risk signals.

Deep semantic understanding is particularly important when processing financial policy texts. Risks are often embedded in complex sentence structures and logical relations. Shallow models fail to capture such semantic hierarchies and cross-sentence dependencies, resulting in lower precision and recall. As the number of Transformer layers increases, the model's ability to model context and integrate semantics improves. This allows it to better uncover potential risk cues in policy language, making the classification results more consistent and accurate.

However, when the number of layers continues to increase to a high level, the performance metrics begin to plateau or slightly decline. This indicates that deeper structures do not always bring further semantic benefits. Instead, they may lead to model redundancy or training instability, which can harm the final output. In financial texts, the information density and linguistic variation have practical limits. Over-modeling may disperse attention and weaken the model's ability to detect local risk factors.

In summary, the performance of the Transformer in policy semantic modeling is significantly influenced by layer depth. Choosing an appropriate configuration enhances the model's ability to represent policy risk tendencies. It also helps avoid overfitting and efficiency loss. This provides structural optimization guidance for building accurate, efficient, and interpretable risk identification systems.

4) Robustness analysis experiment of the model on policy texts in different periods

This paper further presents a robustness analysis experiment of the model on policy texts of different periods, and the experimental results are shown in Figure 3.

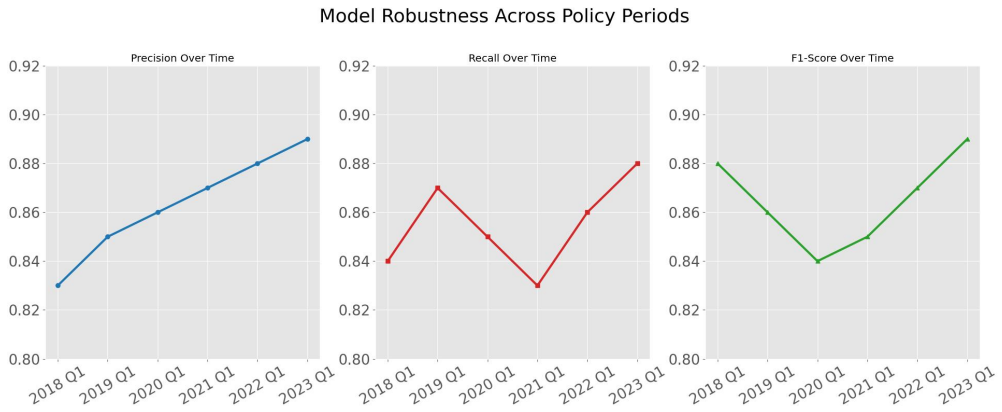

Figure 3. Robustness analysis experiment of the model on policy texts in different periods

As shown in the experimental results in Figure 3, the model demonstrates overall stability across policy texts from different periods. However, the three key evaluation metrics show varied trends across stages. This indicates that the evolution of policy semantics has some impact on the model's ability in semantic modeling and risk judgment. The continuous rise in precision suggests that the model's accuracy in identifying risk-related language has improved over time. This may be due to the increasing standardization of policy language in recent years, making it easier for the model to capture explicit risk indicators.

The fluctuations in recall reflect that the model may have missed some potential risk information during certain stages. This is especially likely when the content of policy texts becomes more complex and risk expressions are more implicit. In such cases, the model's recall capability is challenged. This implies a need to further enhance the model's ability to detect hidden risk factors in complex semantic structures, thereby improving overall coverage.

The F1-score shows a typical V-shaped trend, indicating fluctuations in the model's balance between precision and recall. The initial decline followed by a rise may reflect the model's early adaptation difficulties with earlier policy texts. Over time, with additional training or strategy adjustments, the model gradually builds a more stable capacity for risk judgment. This trend also highlights the influence of cyclical changes in policy language on model performance.

Overall, the experiment validates the applicability and a certain degree of robustness of the proposed method across policy texts from different periods. While some metrics show short-term variations, the overall trend suggests that the model possesses strong semantic generalization ability. It is capable of adapting to changes in the financial policy context. This provides both theoretical support and empirical evidence for building a sustainable risk classification system.

5) Self-attention visualization and risk signal attention area analysis

This paper further presents an analysis focused on self-attention visualization and the attention distribution over risk-related signals within financial policy texts. By examining how the model allocates attention across different terms, the study aims to better understand the internal decision-making process of the Transformer-based architecture when identifying potential risks.

The attention visualization highlights key areas within the text that the model considers important during semantic processing. This approach offers an intuitive means to interpret the model's focus and enhances transparency in risk classification tasks. The corresponding attention maps and visual examples are provided in Figure 4 to support this interpretive analysis.

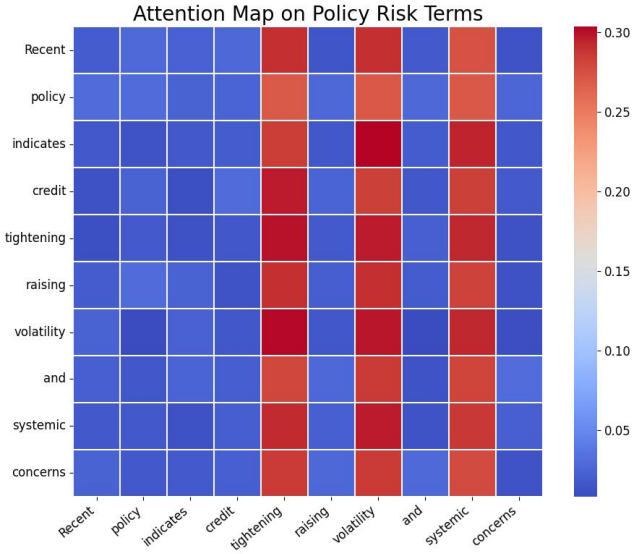

Figure 4. Self-attention visualization and risk signal attention area analysis

As shown in the attention visualization results in Figure 4, the model can focus on core terms closely related to financial risk, such as "tightening," "volatility," and "systemic." These terms appear as high-weight regions in the attention heatmap. This indicates that the model successfully identifies potential risk-triggering words during the encoding phase. It reflects strong semantic sensitivity and contextual awareness.

The model's high attention to these keywords shows its ability to recognize key semantic structures in financial policy texts. In complex policy expressions, risk signals are often

hidden in technical terms or indirect phrases. The attention distribution suggests that the model is not distracted by surface language. Instead, it concentrates on vocabulary likely to influence market volatility and systemic risk. This is important for building stable and effective financial risk identification systems.

The visualization also shows that the model maintains moderate attention on words such as "credit," "policy," and "concerns." These words provide contextual guidance. The multi-level attention mechanism enables the model to capture not only explicit risk terms but also supporting elements that form semantic chains. This enhances the overall sentence-level modeling. Such capability is especially important when dealing with long texts and varied sentence structures in policy documents.

In summary, this attention visualization experiment further confirms the model's interpretability and focus on identifying risk signals. It highlights the strengths of the Transformer architecture in handling texts with high semantic density. By visually presenting the model's attention path, the experiment increases trust in the financial risk classification model. It also provides useful support for attention-guided risk tracing and early warning applications.

6) Effects of different regularization methods on suppressing model overfitting

This paper further gives the inhibitory effect of different regularization methods on model overfitting, and the experimental results are shown in Figure 5.

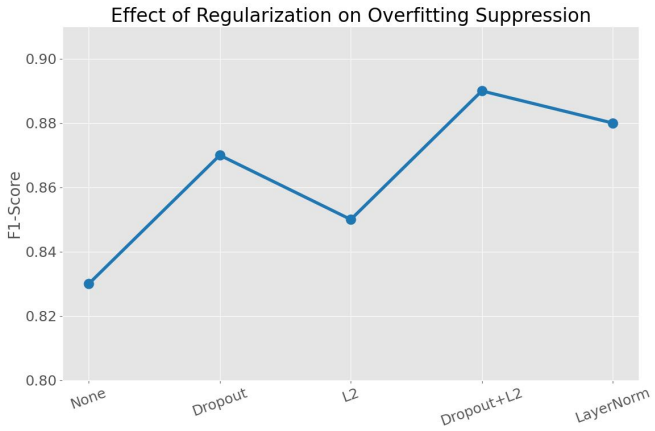

Figure 5. Effects of different regularization methods on suppressing model overfitting

As shown in the experimental results in Figure 5, different regularization strategies have clear effects on reducing model overfitting. The model without regularization learns the training data quickly but performs less stably during testing. This indicates a certain degree of overfitting. In the task of financial policy text risk classification, where semantic expressions are highly uncertain and complex, model generalization is especially important.

After introducing the Dropout mechanism, the model's F1 score improves. This suggests that randomly deactivating neurons enhances robustness by reducing reliance on specific semantic paths. It helps the model identify implicit but inconsistently distributed risk semantics. This leads to better performance consistency across various text structures. L2 regularization also shows some effect, but it does not outperform Dropout in this task. This may be due to its rigid constraint on overall parameter magnitude, which limits flexibility in modeling locally important features in financial texts.

Notably, the combination of Dropout and L2 regularization yields the most significant improvement. This suggests the two methods are complementary in mitigating overfitting. Dropout introduces structural randomness, while L2 imposes continuous constraints on parameters. Together, they help prevent the model from relying too heavily on a few dominant features during training. In addition, LayerNorm, as an internal normalization strategy, shows good stability and generalization. It plays a positive role in maintaining gradient stability and balanced feature distribution during deep semantic modeling.

In summary, regularization strategies play an important role in enhancing the stability and generalization of Transformer models in financial policy semantic modeling. Proper regularization not only suppresses overfitting but also improves the model's ability to detect complex, ambiguous, or latent risk information in text. This provides strong technical support for building long-term, reliable policy risk identification systems.

7) Loss function changes with Epoch

This paper further gives a graph of the loss function changing with Epoch, and the experimental results are shown in Figure 6.

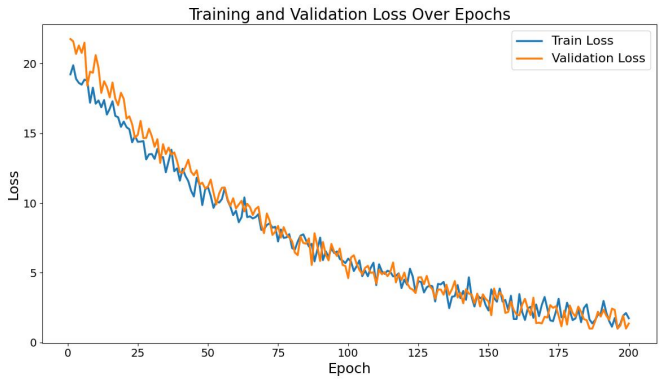

Figure 6. Loss function changes with Epoch

As shown in the loss curve in Figure 6, the model exhibits a stable convergence trend during training. With more training epochs, both training loss and validation loss show a clear downward trend. This indicates that the model gradually learns effective semantic representations for the financial policy text risk classification task. In the early stages, the loss decreases rapidly. This reflects the model's strong ability to learn basic semantic structures and quickly capture key information in the text.

In the middle and later stages, the loss decline becomes more gradual and shows slight fluctuations. These fluctuations suggest that the model is sensitive when handling high-complexity semantics or boundary samples. At the same time, the stable pattern indicates robustness during training, without signs of instability or overfitting. This is particularly important for modeling financial policy texts, which often contain vague or indirect expressions of risk. The model needs to maintain learning stability while responding to unusual semantic cues.

It is worth noting that the gap between training loss and validation loss remains small throughout. This suggests that the model generalizes well between training and validation data. Since financial policy texts may vary significantly across periods and topics, the model's ability to maintain consistent learning across diverse data further validates its adaptability in risk identification tasks.

In summary, this experiment demonstrates the model's training stability and convergence in handling complex semantic modeling. It confirms the effectiveness of the proposed architecture and training strategy for financial text risk classification. These findings provide foundational support for building high-performance and interpretable financial policy risk warning models.

5. Conclusion

This paper addresses the problem of risk classification in financial policy texts. It proposes a semantic modeling method based on the Transformer architecture. The method effectively integrates contextual relations and deep semantic features of policy language. A risk identification framework with strong expressive power is constructed. By introducing multi-head attention and deep encoding structures, the model performs well in handling long sentences and complex semantic patterns. It can extract potential risk signals from unstructured policy texts. This enables accurate risk recognition and classification. The overall approach balances accuracy, robustness, and interpretability. It provides a new technical path for policy semantic analysis. The study conducts a series of detailed experiments to evaluate the model's stability and generalization under different hyperparameter settings, data distributions, and regularization strategies. The results show that the Transformer architecture is naturally suited to the financial domain. It is especially effective in modeling policy texts that are semantically dense and logically complex. This finding offers strong data support and methodological tools for real-world applications such as financial regulation, risk warning, and policy evaluation. It also highlights the practical value of the proposed method. From an application perspective, this approach can be widely applied to the identification and quantification of potential risks in government regulatory documents, banking reports, and financial disclosures. It can provide forward-looking analytical support for financial regulators. At the same time, enterprises and investment institutions can use the model to build internal policy interpretation systems. This enables faster responses to macroeconomic policy changes and improves risk management and strategic decision-making. As financial policies become increasingly complex and strategically worded, automated and intelligent policy analysis tools will become an important research direction in financial technology.

6. Future work

Future research can be expanded in two directions. First, combining external financial knowledge graphs or event graphs may help integrate policy texts with structured background information. This would enhance the model's understanding of policy impact pathways. Second, exploring multimodal fusion methods could bring in news, public opinion, and market data to build a more comprehensive risk identification system. In practical deployment, the interpretability and transparency of the model should also be strengthened. This is essential to meet the traceability and compliance requirements of financial regulation and auditing. In summary, this study provides theoretical support for the semantic modeling of financial texts and lays a practical foundation for the development of related application systems. It has broad potential for future development.

References

[1] Clapham B, Bender M, Lausen J, et al. Policy making in the financial industry: a framework for regulatory impact analysis using textual analysis[J]. Journal of Business Economics, 2023, 93(9): 1463-1514.

[2] Ying H, Chen L, Zhao X. Application of text mining in identifying the factors of supply chain financing risk management[J]. Industrial Management & Data Systems, 2021, 121(2): 498-518.

[3] Wei L, Deng Y, Huang J, et al. Identification and analysis of financial technology risk factors based on textual risk disclosures[J]. Journal of Theoretical and Applied Electronic Commerce Research, 2022, 17(2): 590-612.

[4] Zhu X, Wang Y, Li J. What drives reputational risk? Evidence from textual risk disclosures in financial statements[J]. Humanities and Social Sciences Communications, 2022, 9(1): 1-15.

[5] Huang A H, Wang H, Yang Y. FinBERT: A large language model for extracting information from financial text[J]. Contemporary Accounting Research, 2023, 40(2): 806-841.

[6] Bua G, Kapp D, Ramella F, et al. Transition versus physical climate risk pricing in European financial markets: A text-based approach[J]. The European Journal of Finance, 2024, 30(17): 2076-2110.

[7] Tian, Y., & Liu, G. (2023). Transaction fraud detection via spatial-temporal-aware graph transformer. arXiv preprint arXiv:2307.05121.

[8] Ronyastra I M, Saw L H, Low F S. Monte Carlo simulation-based financial risk identification for industrial estate as post-mining land usage in Indonesia[J]. Resources Policy, 2024, 89: 104639.

[9] Wang B. A financial risk identification model based on artificial intelligence[J]. Wireless Networks, 2024, 30(5): 4157-4165.

[10] Hajimahmud V A, Khang A, Ali R N, et al. Managing financial risks in the Gig economy[M]//Synergy of AI and Fintech in the Digital Gig Economy. CRC Press, 2024: 125-134.

[11] Andaloro A, Salvalai G, Fregonese G, et al. De-risking the energy efficient renovation of commercial office buildings through technical-financial risk assessment[J]. Sustainability, 2022, 14(2): 1011.

[12] Qiao Y, Jin J, Ni F, et al. Application of machine learning in financial risk early warning and regional prevention and control: A systematic analysis based on shap[J]. WORLD TRENDS, REALITIES AND ACCOMPANYING PROBLEMS OF DEVELOPMENT, 2023, 331.

[13] Odinet C K. Fintech credit and the financial risk of AI[J]. Cambridge Handbook of AI and Law (Kristin Johnson & Carla Reyes eds., forthcoming 2022)., U Iowa Legal Studies Research Paper, 2021 (2021-39).

[14] Murugan K, Selvakumar V, Venkatesh P, et al. The big data analytics and its effectiveness on bank financial risk management[C]//2023 6th

International Conference on Recent Trends in Advance Computing (ICRTAC). IEEE, 2023: 313-316.

[15] Khunger A. DEEP Learning for financial stress testing: A data-driven approach to risk management[J]. International Journal of Innovation Studies, 2022.

[16] Oyedokun O, Ewim S E, Oyeyemi O P. Leveraging advanced financial analytics for predictive risk management and strategic decision-making in global markets[J]. Global Journal of Research in Multidisciplinary Studies, 2024, 2(02): 016-026.

[17] Wei Y, Xu K, Yao J, et al. Financial Risk Analysis Using Integrated Data and Transformer-Based Deep Learning[J]. Journal of Computer Science and Software Applications, 2024, 4(7): 1-8.

[18] Wu Y. Enterprise financial sharing and risk identification model combining recurrent neural networks with transformer model supported by blockchain[J]. Heliyon, 2024, 10(12).

[19] Shishehgarkhaneh M B, Moehler R C, Fang Y, et al. Transformer-based named entity recognition in construction supply chain risk management in Australia[J]. IEEE Access, 2024, 12: 41829-41851.

[20] Song Y, Du H, Piao T, et al. Research on financial risk intelligent monitoring and early warning model based on LSTM, transformer, and deep learning[J]. Journal of Organizational and End User Computing (JOEUC), 2024, 36(1): 1-24.

[21] Liu, Y., Wu, H., Wang, J., & Long, M. (2022). Non-stationary transformers: Exploring the stationarity in time series forecasting. Advances in neural information processing systems, 35, 9881-9893.

[22] Mishra A K, Renganathan J, Gupta A. Volatility forecasting and assessing risk of financial markets using multi-transformer neural network based architecture[J]. Engineering Applications of Artificial Intelligence, 2024, 133: 108223.

[23] Ouyang Z, Yang X, Lai Y. Systemic financial risk early warning of financial market in China using Attention-LSTM model[J]. The North American Journal of Economics and Finance, 2021, 56: 101383.

[24] Li J, Xu C, Feng B, et al. Credit risk prediction model for listed companies based on CNN-LSTM and attention mechanism[J]. Electronics, 2023, 12(7): 1643.

[25] Zhao L, Li L, Zheng X, et al. A BERT based sentiment analysis and key entity detection approach for online financial texts[C]//2021 IEEE 24th International conference on computer supported cooperative work in design (CSCWD). IEEE, 2021: 1233-1238.

[26] Liao J, Shi H. Research on joint extraction model of financial product opinion and entities based on roberta[J]. Electronics, 2022, 11(9): 1345.