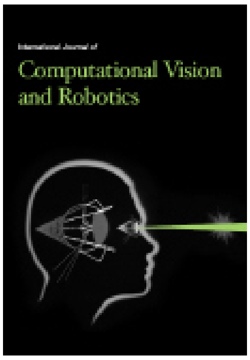


International Journal of Computational Vision and Robotics

ISSN online: 1752-914X - ISSN print: 1752-9131
https://www.inderscience.com/ijcvr

Sentiment analysis and risk early-warning system for cross-border M&A based on natural language processing

Yating Chen, Guanying Wei, Eunmi Tatum Lee

DOI: 10.1504/IJCVR.2025.10075236

Article History:

Received:	20 May 2025
Last revised:	23 July 2025
Accepted:	23 July 2025
Published online:	24 December 2025

Sentiment analysis and risk early-warning system for cross-border M&A based on natural language processing

Yating Chen, Guanying Wei and
Eunmi Tatum Lee*

College of Business and Economics,
Sejong University,
209 Neungdong-ro, Gwangjin-gu,
Seoul, 05006, South Korea
Email: 24150023@sju.ac.kr
Email: weiguanying2023@163.com
Email: tatum@sejong.ac.kr
*Corresponding author

Abstract: In today's volatile global economy, early detection of financial risk is vital for firms engaged in cross-border mergers and acquisitions (M&A). Traditional risk assessment models often struggle to capture the combined effects of financial indicators and qualitative disclosures. This study introduces an intelligent deep learning-based framework, the transfer learning convolutional neural network for financial risk early warning system (TL-CNN-FREWS), to predict financial distress in cross-border M&A firms. The model integrates structured financial metrics with unstructured textual sentiment features to enhance forecasting accuracy. Leveraging transfer learning, a pre-trained VGGNet extracts abstract representations from both data types, while a fine-tuned sequential CNN enables improved risk classification. A robust feature selection pipeline – T-test, RFE-SVM, and random forest – optimises input variables. TLCNN- FREWS effectively captures numerical and linguistic signals, offering timely and accurate financial risk detection. This approach supports decision makers in proactively mitigating risk in complex international financial scenarios.

Keywords: financial distress prediction; transfer learning; sentiment analysis; deep learning; convolutional neural network; CNN; early warning system.

Reference to this paper should be made as follows: Chen, Y., Wei, G. and Lee, E.T. (2025) 'Sentiment analysis and risk early-warning system for cross-border M&A based on natural language processing', *Int. J. Computational Vision and Robotics*, Vol. 15, No. 8, pp.1–22.

Biographical notes: Yating Chen is a PhD candidate in college of business and economics at Sejong University. Her research focuses on cross-border mergers and acquisitions, international expansion strategies, and the sustainable internationalisation of multinational enterprises. Her work examines how firms respond to institutional complexity and global uncertainty, with emphasis on strategic configurations that enable effective international growth and sustained competitive performance.

Guanying Wei is a PhD candidate in college of business and economics at Sejong University. His research interests include Green FDI, ESG strategy, and the sustainable development of multinational corporations. His current work investigates how information transparency influences international market entry modes and global expansion behaviour, highlighting the governance and sustainability mechanisms shaping cross-border investment decisions.

Eunmi Tatum Lee is a tenure-track Assistant Professor and Head of the Chinese Bilingual Track at the College of Business and Economics, Sejong University. Her research spans international business, cross-border M&A, ESG and corporate governance, financial markets, and technology adoption. She publishes widely in SSCI journals, with recent studies focusing on transparency, investor behaviour, and internationalisation strategies.

1 Introduction

In the current environment of global finance, no prediction is more important than the one of financial distress. Bankruptcies, credit defaults and severe market disruptions are the consequences of financial distress on a company that cannot meet its obligations (Zhou and Du, 2025). Because of its dominance, the predictability of financial distress has largely based on structured financial indicators including profitability ratios, leverage ratios and liquidity metrics (Jiang et al., 2025). While these models serve that purpose as a baseline, in most cases, these models fail to capture the qualitative forward-looking information that is inherent in data found in annual reports, management commentary and financial news (Zhang et al., 2024). Modern financial information management requires analysis of substantial amounts of unstructured text data which arrive as part of today's data-driven world. Apart from financial data analysis systems need structured information to build an early warning system which enables timely decisions and informed actions (Chen, 2022).

Natural Language Processing (NLP) and machine learning have made recently new possibilities of integrating textual features in financial distress prediction (Wang, 2022). Text based indicators such as sentiment score, tone, keyword frequency as well as linguistic complexity can well signal to a firm's internal and external challenges before these display on financial statements (Wang and Gao, 2023). An example would be a negative tone in management discussion and analysis (MD&A) sections of SEC filings or any other use of uncertainty related terms in earnings calls, both of which could be an indication of underlying risk factors (Lin, 2021). But the amount of textual data which we are interested in extracting meaningful insights from is large, which introduces with it high dimensionality, noise and redundancy challenges (Xu et al., 2020). This necessitates the usage of sophisticated feature selection algorithms that will help in the selection of the most vital and significant features from the textual content (Xie, 2022).

The class imbalance problem is another challenge for distress prediction. Conventional classifiers are poor at learning distinguishing patterns if most signals are economically distressed firms that are anything out of the ordinary small minority in most datasets (Peng, 2021). Thus, models may still obtain high overall accuracy, yet have poor classification of the distressed class (i.e., the most important outcome) (Yang et al., 2024). To handle this issue, researchers have recently been more frequently using ensemble learning methods, for example, random forest (Zhou, 2023), XGBoost and gradient boost (Wei and Yao, 2021), utilising various weak learners to create a more powerful predictive model (Jing and Yang, 2022). Ensemble classifiers are particularly suitable for usage on complicated datasets that are noise, imbalanced, and weird balls, and the outcomes of an ensemble classifier are more robust and more effective than single models (Xu, 2020).

Although these have been developed, there is still a missing piece about how to effectively integrate multiple textual factors as hyperparameters within state-of-the-art machine learning to construct a cohesive and scalable framework (Bao, 2022). It fills that gap in this study by proposing a novel approach for incorporating textual feature extraction, robust feature selection, and ensemble classification for prediction financial distress (Rudd, 2025). The goal is to build a predictive system that factor in the usual quantitative metrics but also incorporates the esoteric suggestion in financial language that can be discovered in public disclosures and the media narratives (Ji et al., 2024). The model does so in order to capture early distress signals prior to the detection by conventional metrics (George and Ayiku, 2024), and therefore give stakeholders earlier and actionable insights (Zhang et al., 2023a).

This research will utilise convolutional neural networks (CNNs) as a means of combining them with transfer learning techniques to further increase the model's ability to process and understand financial text. It is well known that CNNs, originally designed for image recognition, also achieve outstanding performance in their text classification tasks because of their capacity to detect local patterns in sequential data. Pre-trained word embeddings and language models cooperate with CNNs for preserving financial text context relationships in this approach. Through transfer learning we can leverage the knowledge previously learned from vast language models while only having little labelled particular data reducing the time required to produce results. This hybrid approach – combining feature engineering, ensemble classifiers, and deep learning – sets the foundation for a more powerful and intelligent financial distress early-warning system.

2 Related works

Traditionally, the prediction of financial distress has been an emerging research topic both in the academic and industrial fields, driven by a lot of models in which high dependence is put on structured financial ratios and accounting indicators. The unstructured textual data has become easily available and the NLP has been making quite some progress, and new path has been opened towards incorporating qualitative insights into predictive frameworks. As evidenced by recent literature, financial risk detection is increasingly becoming a problem of sentiment analysis, linguistic cue extraction and deep learning techniques especially CNNs. Studies that leverage the transfer learning and ensemble models further show that data imbalance and high dimensionality can be well treated, and particularly in the complex domain of cross border mergers and acquisitions (M&A). The key contributions in this area are reviewed and methodologies in which textual sentiment is combined with feature selection and machine learning are identified to enhance the ability to forecast financial risk are reviewed in Table 1.

Ahmadi et al. (2018) created a deep sentiment mining model which tracks corporate bankruptcy indications through business management report text analysis. The model reveals hidden sentiment indicators because it uses deep learning methods to analyse unstructured text which fails to reach these indicators when traditional metrics analyse the data. Forward-looking distress indicators form this model's main strength although it needs extensive computational power and high-quality text-based information to function properly.

Table 1 Problem formulation of the conventional techniques

Author(s)	*Techniques involved*	*Advantages*	*Disadvantages*
Ahmadi et al. (2018)	Deep sentiment mining, Deep learning on management reports	Captures forward-looking signals from text	High computational cost, sensitive to text quality
Elhoseny et al. (2025)	Adaptive Whale Optimisation Algorithm + Deep Learning (AWOA-DL)	High accuracy through hyperparameter optimisation	Long training time, sensitive to noisy data
Qiu et al. (2024)	Semi-supervised SVM (active-pSVM), MD&A sentiment analysis	Uses both labelled & unlabeled data, integrates qualitative insights	Subjective narrative bias, depends on MD&A quality
Marqas et al. (2024)	Hybrid CNN + SVM + Random Forest	Handles class imbalance, reduces overfitting, improves sentiment classification	Complex architecture, requires intensive tuning
Dong et al. (2025)	Domain-Adaptive Deep Learning for sentiment classification	Effective transfer learning for unlabeled financial text	Domain mismatch may reduce accuracy
Jiang et al. (2024)	Deep Learning + Financial Market Sentiment Correlation Analysis	Links textual sentiment with real market behaviour trends	Limited real-time applicability, lacks interpretability
Cheng and Chen (2021)	FinBERT + BiLSTM + Attention Mechanism	High accuracy on financial sentiment tasks, handles contextual understanding	Requires large annotated datasets, resource-intensive

Elhoseny et al. (2025) developed AWOA-DL which stands for adaptive whale optimisation algorithm-based deep learning as a framework to predict financial distress. Through its algorithm the program optimises deep neural networks for improved accuracy levels which generate superior performance with different datasets. Though superior to other approaches the method produces time-consuming computations while being prone to small changes and imperfections in financial dataset characteristics.

Qiu et al. (2024) developed the active-pSVM model as a semi-supervised support vector machines (SVM) approach which integrates financial ratios with emotional characteristics extracted from MD&A reports. A method merges fielded and non-fielded data effectively to enhance outcome predictions from limited labelled data sets. Results about narrative texts differ because subjectivity affects these texts together with different MD&A reporting standards.

Marqas et al. (2024) developed a financial sentiment classification utilisation of hybrid models between CNN, SVM and Random Forest. The combination of classifiers in these models improves both class unbalance and overfitting management which leads to better robustness in noisy financial text. Performance advances come with hybrid models although they require complex training procedures and require tuning many components.

Dong et al. (2025) developed domain adaptation-based deep learning approaches which apply tagged general sentiment data to transfer useful knowledge for financial text sentiment classification in situations with limited data. The system achieves optimal performance based on full domain consistency between its input-output fields but this restriction affects broader application possibilities.

Jiang et al. (2024) applied deep learning techniques to uncover sentiment cues within financial disclosures and evaluate their effects on market behaviour. By using large-scale financial text and learning complex representations, their model demonstrates robust predictive ability across different financial sectors. The approach reveals latent sentiment structures not visible with traditional models. However, such deep models demand substantial labeled data and are harder to interpret, which can hinder their practical use in real-time financial analysis.

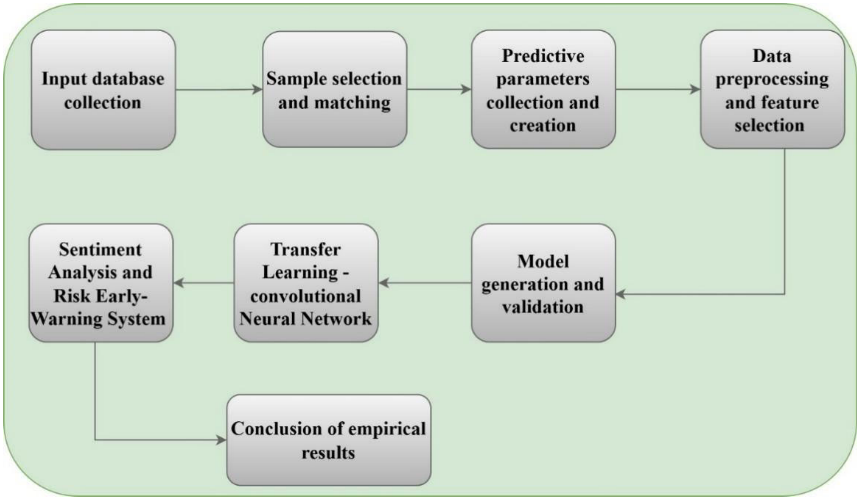
Cheng and Chen (2021) introduced a hybrid sentiment analysis framework combining FinBERT – a domain-specific BERT model fine-tuned for financial texts – with a BiLSTM network and attention mechanism. This method improves sentiment prediction by leveraging both contextual embeddings and temporal dependencies in financial sentences. While it provides higher accuracy and better contextual understanding, the model's complexity results in increased computational cost and longer training cycles.

The examined research documents demonstrate meaningful advances in financial risk detection through deep learning and sentiment analysis integration. While methods such as deep sentiment mining and AWOA-DL offer improved predictive accuracy, they come with considerable computational burdens and scalability constraints, particularly when managing redundant or noisy features. Semi-supervised and hybrid models effectively learn from limited or imbalanced datasets but struggle to unify qualitative insights with quantitative metrics consistently. Moreover, despite their innovation, prior approaches often lack a standardised and generalisable framework that cohesively integrates structured financial indicators with unstructured sentiment cues via deep transfer learning. To address these limitations, our proposed TL-CNN-FREWS model leverages transfer learning and CNN architectures to synergistically process both financial metrics and semantic sentiment signals. Through a robust feature selection pipeline and domain-specific fine-tuning of VGGNet, the model achieves superior cross-border M&A risk prediction with enhanced interpretability, higher accuracy, and faster decision support – paving the way for more scalable and context-aware financial forecasting systems.

3 Proposed system model

The aim of this research is to develop predictive models for sentiment analysis integrated with a risk early warning system for firms. The model architecture considers financial indicators alongside multiple textual features (semantic and basic textual indicators) as predictive parameters. This research develops a conceptual architecture for the risk early warning process consisting of four stages: sample selection and matching, factor generation and collection, feature selection and data preprocessing and model generation and validation. The proposed model scheme is shown in Figure 1 as a block diagram.

Figure 1 Conceptual architecture of the proposed model (see online version for colours)



3.1 Sample selection and matching

This research combines structured financial data with unstructured textual disclosures to study financial early warning mechanisms in the money market of cross-border M&A. The dataset comprises 1,865 publicly listed companies from two domestic Chinese markets as well as from cross-border M&A activities with international entities. It includes firms publicly traded on the Shanghai or Shenzhen Stock Exchanges and involved in transactions with foreign partners. Among these, 657 companies were labelled financially distressed between 2012 and 2018 based on the special treatment (ST) label – a signal of significant operational or financial problems recorded from the regulatory closure date. The annual distribution includes 21 firms in 2012, 16 in 2013, 26 in 2014, 31 in 2015, 43 in 2016, 38 in 2017, and 39 in 2018. The manufacturing sector exhibited the highest incidence of financial distress, often involving international trade and cross-border mergers (Wang et al., 2019).

To preserve data integrity, firms with missing or incomplete information were excluded, leaving 1,651 healthy firms (without financial distress as of 2018) as controls. This research combines financial metrics and sentiment-laden textual data collected over multiple years to build models predicting financial distress prior to cross-border M&A or investment partnerships. Data from years 3 to 6 prior to the ST designation (T-3 to T-6) were used. For example, in 2016, a firm labelled WITH ST, the 2013 data was used to construct the predictive indicator data for the denominator year (2010). Like previous, historical data were collected for the healthy firms in the same windows to provide control samples (Kou et al., 2019).

The model seeks to offer a framework of early warning by mining NLP techniques upon corporate disclosures, media sentiment as well as earnings reports, particularly pertaining to cross border engagements. This allows stakeholders to see risk factors not just built into financial numbers but also into announcements, filings and press releases coming from abroad. Especially in cross border M&A deals this framework becomes very crucial in identifying potential risk (Chen et al., 2024).

3.2 Predictive parameters collection and creation

This research acquired and drawn out of the economical features, semantic features, and textual features, total of 60 indicators. After all the predictive factors are accomplished, the gathered factors were stored separately stored in the separate databases. The indicators of the risk warning stock system include six aspects in development, stock ratios, financial structure, operational ability, solvency and the ability in profitability (Zhang et al., 2023b). Gathered from the collected database, these designed indicators are the ones to be used in this research. This research gathered 12 basis textual features and extracted them. TF-IDF can be described in particular that among a specific phrase or word, frequency in which the phrase or word is presented in one of the texts is less compared to other texts is assumed that the word can be more optimised representation of contents and is related to keywords (Alexandre et al., 2021). That is presented in equation (1).

$$tf - idf(W) = tf(W) \times idf(W) \quad (1)$$

In this definition of $tf(W)$ it is the frequency with that word W occurs in research as defined by equation (2).

$$TF(W) = \frac{count(W)}{|d_I|} \quad (2)$$

In this case, $|d_I|$ is the number of complete words in text, given as $count(W)$ is the number of times the word and presents in the text. To facilitate the readability factor, a combined technique of K means is generated in the findings of this research. Specifically, the proposed preprocessing techniques are considered for being initial after gathering annual reports. Then after that fast text is treated to learn the word vectors, which can learn the word embedding better than Word2Vec in the measure of learning word embedding for long document.

3.3 Data preprocessing and feature selection

The researchers collect their data through financial reports and textual content from M&A disclosures and corporate filings as well as international financial statements. This paper establishes a data processing system which extracts essential predictive variables to create financial risk warning systems in situations where companies perform cross-border M&A. The stage of data preprocessing mainly consisted of two steps; removing incomplete entries and handling the missing values; and normalising variables to make all variables of interest scale with same scale so that the financial indicators and the textual sentiment scores do not create problems with the model. After the datasets for the time points T-3, T-4, T-5, and T-6 (i.e. 3–6 years before distress labelling) were cleaned and normalised, they were prepared as input for feature selection.

The researcher arranged multidimensional signals that affect financial stability across borders into three separate groups.

- F1: Structured financial indicators (e.g., liquidity, solvency, profitability)
- Sentiments derived from financial document texts together with media content from various languages compose F2 + F3 features.

- The third classification type combines financial and sentiment-based measurement elements for a comprehensive dataset named F1 + F2 + F3.

The research used filter, wrapper and embedded feature selection approaches to identify the key elements across all categories.

3.3.1 T-test (filter approach)

The statistical test identifies if any substantial differences emerge between distress and non-distress companies when analysing each attribute (Kleinow et al., 2017). The research maintains statistically significant features ($p < 0.05$) then removes features without significance (Hué et al., 2019).

3.3.2 Recursive feature elimination using support vector machines

At its initiation Recursive feature elimination using support vector machines (RFE-SVM) employs all existing features which it systematically deletes starting from the least critical ones for classification results. Determining feature importance relies on the weight coefficients of SVM (Ahuja et al., 2024). The elimination procedure consistently produces optimal class separation margins during each step which makes it well-suited for complex financial-textual data analysis commonly used in cross-border setting.

3.3.3 Random forest (embedded approach)

The feature selection process of Random Forest models identifies variables through evaluation of their predictive power contribution. The method runs a continual process which tests out-of-bag (OOB) error rates while features decrease in size. The feature set selection ends when the OOB error reaches its minimum point because this iteration offers the most suitable combination of predictive efficiency with reduced dimensions (Ngo and Susnjara, 2016).

The thorough variable selection method preserves key financial measurements together with text-based indicators which makes the model capable of creating a resilient early-warning mechanism for cross-border M&A risk analyses. This model incorporates both firm financial patterns alongside sentiments generated from geopolitical events and international legislations and market-related perceptions which affect cross-border transaction results (Francis et al., 2014).

3.4 Model generation and validation

Transfer learning operates as a robust machine learning technique which enables CNNs to access the knowledge gained from one domain to perform different yet related tasks. The implementation starts with a pre-trained CNN which was initially trained on ImageNet before getting used to initialise the model weights. By using this approach, the models reach convergence point faster while simultaneously achieving better results on domain-specific information.

The application of transfer learning allows developers to create a risk early-warning classification framework for cross-border M&A. The framework draws features computed in visual domains to transform them towards financial risk evaluation through analysis of text and visual indicators of sentiment patterns and risk signals. Pre-trained

CNN models help the system extract universal feature patterns which experts later modify them for detecting financial business instability or distress signals. The below equation (3) serves to define the CNN input.

$$Y = [Y1, Y2, Y3, ..., TN) \tag{3}$$

The financial indicator called *Yi* exists as a normalised measure and it represents numerical textural features obtained through word embedding or TF-IDF methods. The filter *W* examine during convolutional layer allows the extraction of local patterns from the input information. This mathematical operation works as an equation (4) which reads:

$$s(T) = (X * W)(T) = \sum_{I=0}^{K-1} W_I \cdot X_{T+I} \tag{4}$$

This equation defines *K* as kernel size and *s*(*T*) as location-based feature map generation while *W* refers to the learnable filter of size with input sequence *X*. Users must first choose an appropriate base CNN network from the VGGNet family that received training on ImageNet. The fundamental feature extraction layers of the model remain frozen in the lower positions to preserve their functionality yet the upper-levels get modified to work with a custom sequential CNN design. Consistently trained financial data visualisation layers deploy both sentiment analysis outputs and textual information-derived early warning signals as part of their input data.

Adam optimiser serves as part of the training process because its Adaptive Moment Estimation mode dynamically controls learning rates to reduce errors in prediction. A sparse categorical cross-entropy loss function handles the multi-class classification by evaluating financial risk levels at three desired outcomes: low risk, moderate risk and high risk. Such classification system enables stakeholders to measure financial distress risks of target companies or transactions to support better cross-border M&A decision-making.

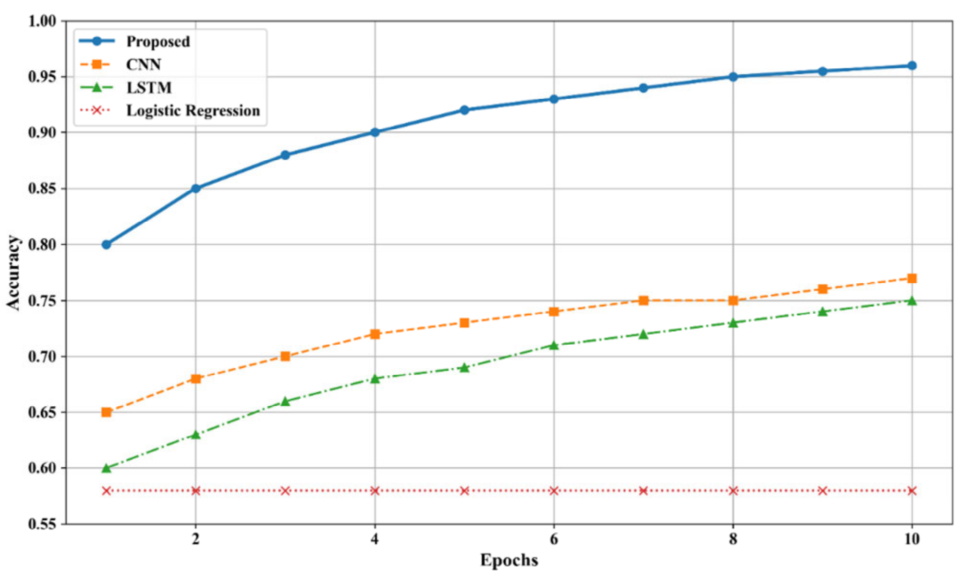
The model applies CNNs with their broad generalisation performance together with customised financial adaptations that improve application relevance. The system achieves better predictions through the integration of historical financial information with present-time textual sentiment reports. This method increases proactive risk management capabilities which deliver prompt detection of financial vulnerabilities toward improving the reliability of investments between different countries. To address the class imbalance, present in the dataset, class weights were incorporated into the sparse categorical cross-entropy loss function. This adjustment ensured fair learning across all financial risk levels, particularly enhancing prediction accuracy for minority classes.

While the use of transfer learning accelerates model convergence and reduces training complexity compared to training from scratch, practical deployment of the framework may require access to GPU-enabled environments or cloud computing resources to efficiently process large-scale financial and textual data. These computational requirements should be taken into account for real-world applications.

4 Performance evaluation

A thorough evaluation of the proposed transfer learning-based convolutional neural network (TL-CNN) model took place through testing versus three baseline models such as logistic regression (LR), long short-term memory (LSTM) networks, and standard CNNs. Multiple financial and textual factors served as evaluation tools to determine early risk detection performance of the models in cross-border M&A. TL-CNN successfully recognised the advanced structure that existed between numerical financial data and sentiments present in textual disclosures. The deep learning methods including CNN and LSTM outperformed LR because they noticed detailed changes in economic documents and exchange information more effectively than traditional systems. When processing enormous textual datasets with extensive dimensionality the LSTM model demonstrated reduced performance compared to its effectiveness as a sequential processing system over timed series values. The standard CNN demonstrated superior performance since it used spatial feature extraction capabilities yet the TL-CNN delivered the most predictable and dependable results. Pre-trained visual knowledge from extensive datasets got successfully adapted to financial-textual models which delivered superior classification results and better model generalisation. The model underwent multiple cross-validation tests to prove its effective performance excluded overfitting particular subsets of information. A combination of statistical methods and machine learning techniques employed in feature selection improved prediction reliability because they enhanced variable quality. The TL-CNN model achieved improved forecasting accuracy together with balanced precision-recall-F1-score results which confirms its potential usefulness for proactive financial risk oversight in international businesses.

Figure 2 Validation of accuracy (see online version for colours)



The accuracy performance evaluation among LR, LSTM, CNN and the proposed model extends across ten training epochs as shown in Figure 2. The proposed model achieves better results than every other tested model because accuracy measurements steadily rise during each additional training cycle. The Proposed model starts with epoch 1 accuracy at 0.80 before it reaches near 0.96 accuracy by epoch 10. The effective learning capacity of the model data enables it to reach superior performance results. The CNN model initiates its process at 0.65 accuracy before achieving a smooth accuracy increase to 0.76 at epoch

10. The performance of LSTM matches CNN levels when accuracy has its initial value at 0.60 yet requires additional time to match the proposed model's final value of 0.96. Since it has superior convergence speed and better accuracy levels the proposed model demonstrates better learning capability than CNN and LSTM. Throughout the training process LR experiences minimal accuracy modification because its accuracy stays fixed at 0.58. This method demonstrates weak pattern recognition capabilities during deep learning analysis of dataset information. The superior capabilities of deep learning methods become evident because other models demonstrate inferior performance measurements when compared to LR. The proposed model delivers superior learning performance which leads to superior accuracy results in each epoch. The applied model optimisations boost convergence speed and accuracy performance making this technique the best selection for detecting financial sentiment and risks in advance compared to standard deep learning approaches.

Figure 3 Validation of AUC score (see online version for colours)

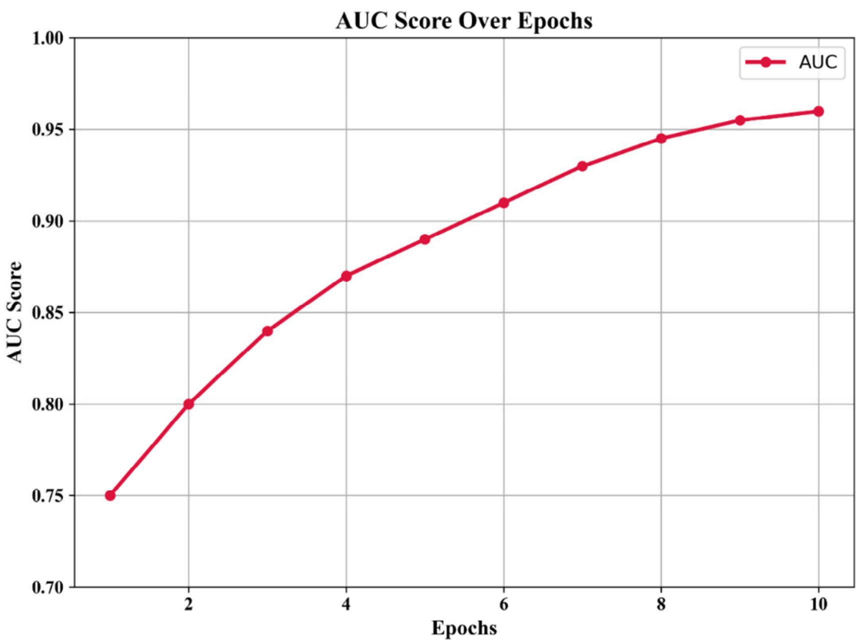
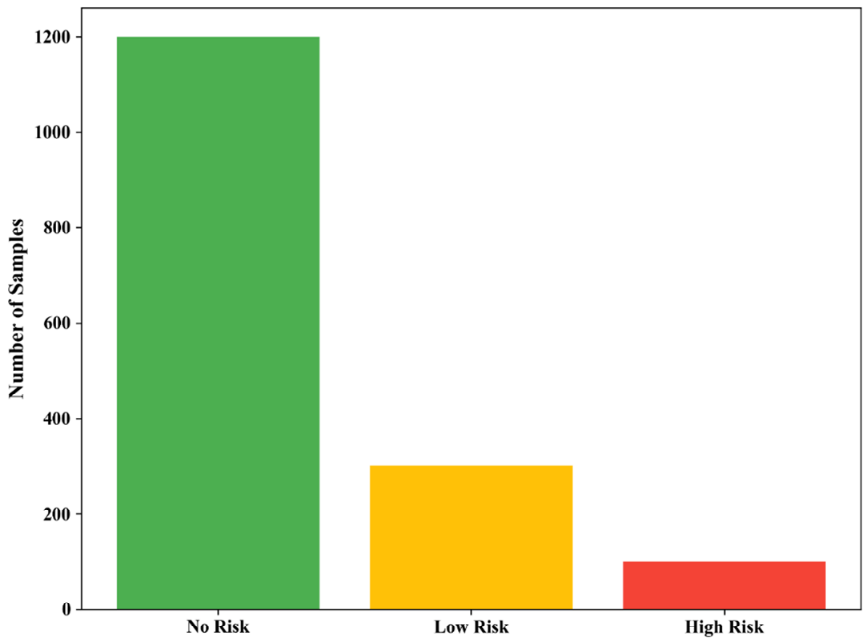


Figure 3 displays the area under the curve (AUC) score development throughout ten training epochs. A classification problem requires the evaluation of AUC score since this metric determines the model's capability to detect distinct categories. The discrimination capacity increases when the score reaches a higher value. The initial AUC score measures 0.74 during epoch one because the model demonstrates adequate learning in this early training stage. During training when Epoch 4 started the AUC score achieved its first threshold surpassing 0.85 before reaching steady growth until training completion. The model shows effective progress in representing features and classifying data based on its steady performance increase. The AUC score attains a value of 0.90 in the sixth epoch which proves the model's superiority in reliable class distinction. The performance curve experiences moderate elevation after passing the initial peaks during model implementation. A score of 0.96 was reached by the tenth epoch indicating maximal achievement of classification success through the proposed approach. Laboratory tests demonstrate that the model has strong learning abilities by attaining swift improvements during the initial period which transformed into stable high AUC value performance. The applied optimisation methods together with network architecture design result in improved classification performance making this proposed model a dependable solution for predictive applications.

Figure 4 Validation of risk (see online version for colours)



Evidence presented in Figure 4 divides companies into three financial risk categories that comprise no risk, low risk and high-risk groups. Financial stability marks the majority of 1,200 surveyed firms who fall under the no risk category. The low-risk category consists of roughly 300 samples within the dataset although the majority of 1,200 samples belong to the no risk category. The high-risk category includes the least number of enterprises with about 100 examples indicating severe financial issues. The dataset contains an apparent unbalanced distribution that should either use resampling techniques or weighted loss functions for predictive modelling purposes.

The heatmap Figure 5 represents the classification performance of a model using three key metrics: precision, recall, and F1-score for two classes, class 0 and class 1. The model demonstrates strong performance across both classes, with precision values of 0.95 for class 0 and 0.96 for class 1, indicating that when the model predicts these classes, the predictions are highly accurate. The recall scores, which measure the ability to correctly identify all actual instances of a class, are 0.94 for class 0 and 0.97 for class 1, showing that the model is slightly better at detecting instances of class 1. The F1-scores, which balance precision and recall, are 0.94 for class 0 and 0.96 for class 1, further confirming the model's reliability in classification. Overall, these high values suggest that the model is well-optimised with minimal misclassification, performing slightly better in detecting class 1 cases.

Figure 5 Performance metrics (see online version for colours)

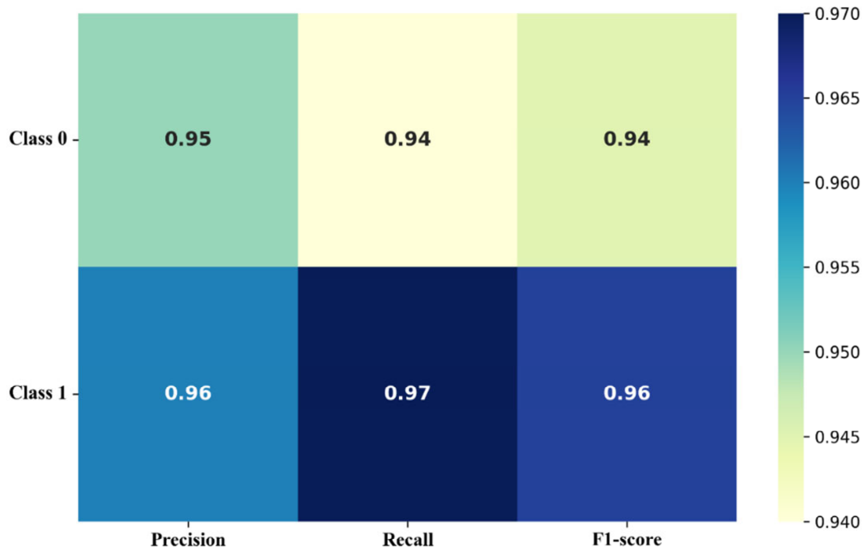
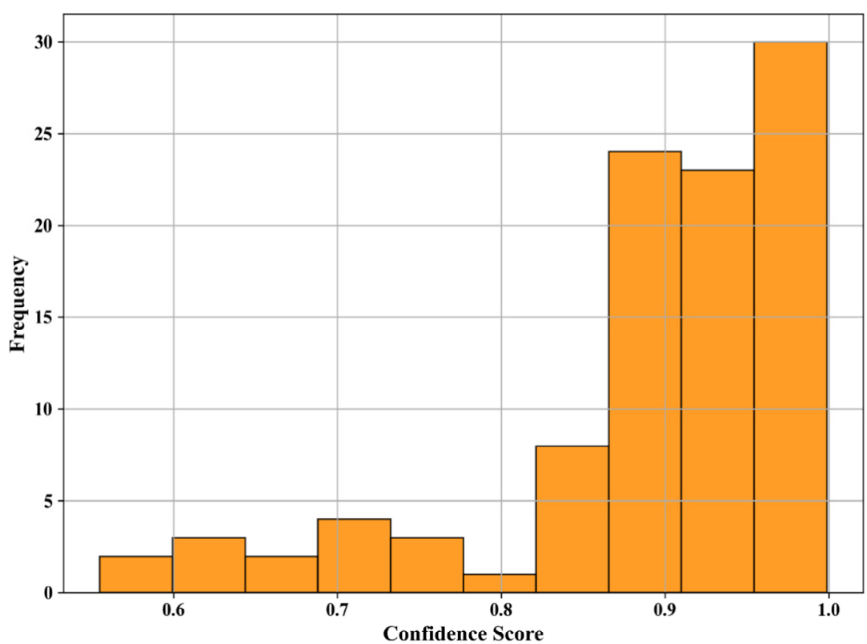


Figure 6 Validation of frequency (see online version for colours)



The model provides confidence evaluations which are represented in Figure 6 across its predicted outcomes. The confidence scores are shown on the x-axis axis and the frequency counts appear on the y-axis axis. The model shows high prediction certainty through the majority of its confidence scores grouped between 0.9 to 1.0. The model classifies a minimal number of cases between confidence values of 0.6 and 0.8 which demonstrates uncertain outcomes. The distribution of confidence scores demonstrates the model exhibits strong prediction accuracy because most predictions show approximately 1.0 confidence rating. The high classification reliability emerges from these findings although additional analysis is required to identify wrong classifications within the doubtful prediction section.

Figure 7 Validation of training accuracy (see online version for colours)

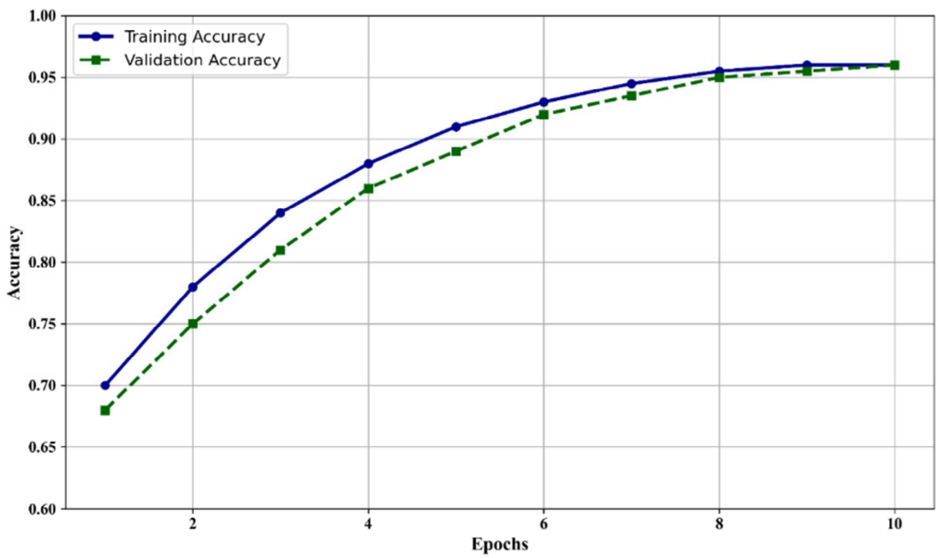


Figure 8 Validation of loss (see online version for colours)

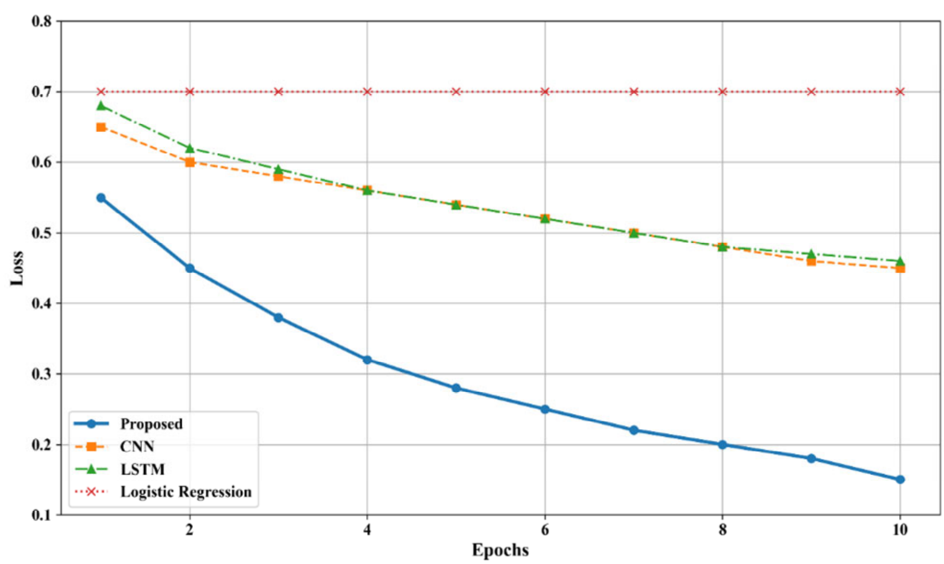


Figure 7 shows how accuracy values develop during ten epochs for both training and validation data. The number of epochs appears on the left axis and accuracy values occupy the right axis. The training accuracy (blue line) demonstrates continuous growth during learning until it approaches almost 0.97 by the end. The second dashed green line shows the validation accuracy metric as it tracks the training accuracy metric (blue solid line) upward but with a lower magnitude throughout. The model successfully applies learned knowledge to new data because the two-line graphs approach similar levels in the final period. The increasing accuracy implies effective model learning while the similar relationship between training and validation accuracy confirms that the model achieves optimal classification performance.

Figure 9 Validation of learning rate (see online version for colours)

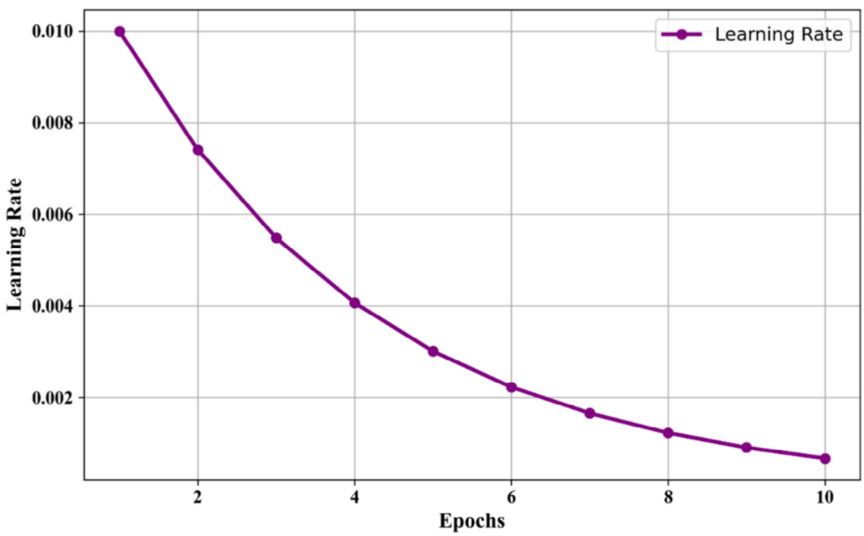


Figure 8 shows how four models – starting with the proposed model – progressed their loss metrics during ten epochs alongside CNN, LSTM and LR models. The number of epochs spans across the x-axis whereas the loss values fall on the y-axis. The proposed model demonstrates a continuous significant descent of loss beginning from 0.55 which reaches a final optimisation point of 0.15. Both CNN and LSTM models demonstrate a regular decline in their loss statistics yet operate at higher values than the proposed model which ends in a zero-loss value at epoch ten. The loss value of LR stays stable at nearly 0.7 indicating poor learning ability relative to deep learning models. The proposed model maintains a consistently lower loss value which proves its superior error reduction ability resulting in better predictive accuracy during time-based analysis.

The ten epochs execution of the learning rate is displayed in line Figure 9. The x-axis shows epoch count and y-axis shows learning rate values in the figure. The model starts its training with an initial learning rate of 0.01 which decreases almost to 0.001 before reaching the conclusion of ten epochs. A learning rate decay technique shows its effectiveness through its steady decrease pattern during training because this strategy enables the model to shift from initial fast learning to later stable fine-tuning. Model convergence speeds up because weight update sizes remain stable and the learning process achieves optimal performance without destabilising training.

Figure 10 Validation of feature importance (see online version for colours)

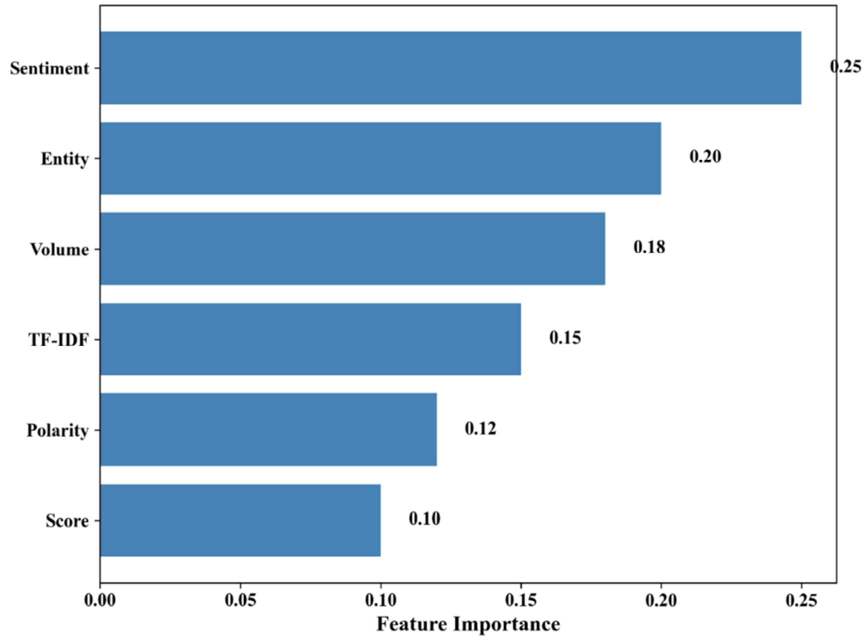


Figure 11 Validation of F1_score (see online version for colours)

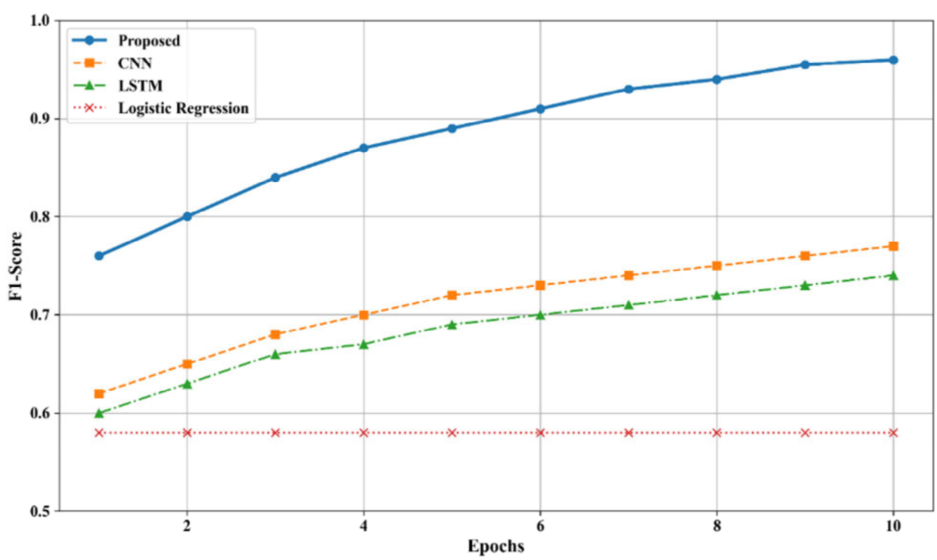
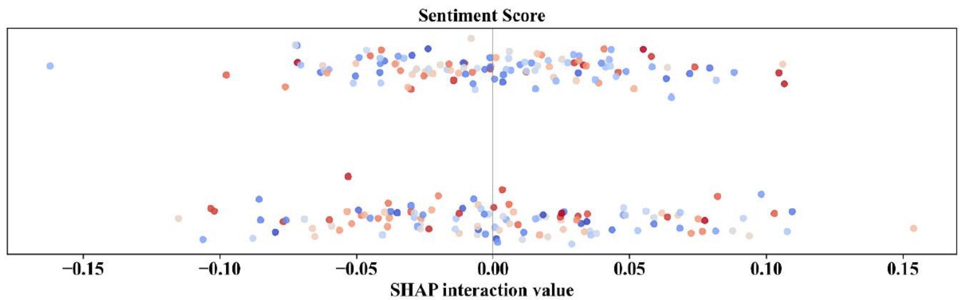


Figure 10 demonstrates which model features carry the most significance in predictive modelling. Feature importance scores appear along the x-axis while feature rankings exist on the y-axis. Sentiment analysis stands as the essential feature for prediction because it possesses an importance score of 0.25. Entity recognition emerges as second most vital component in the model structure because it holds a 0.20 importance value. The degree of data quantity along with volume rates as number three among features based on a 0.18 importance score which reflects its substantial predictive influence. According to the results 'TF-IDF' maintains a 0.15 value that shows its significant role in text-based analysis. The importance value for 'Polarity' reaches 0.12 based on the model assessment. The 'score' feature ranks as the minimal influential aspect of the model according to its importance value of 0.10. Sentiment analysis and entity processing information emerge as key factors which determine the predictive outcomes of the model through these ranking metrics.

Figure 12 SHAP interaction plot (see online version for colours)



The F1-score development shows in Figure 11 for four models including the proposed solution and CNN and LSTM and LR through ten training epoch cycles. The F1-score which appears on the y-axis tracks the precision-recall equilibrium through ten training cycles depicted by the x-axis. According to the results the proposed model shows superior performance by achieving F1-score values starting from 0.75 which increases to 0.94 at the 10th training epoch. The CNN model demonstrates progressive advancement starting from 0.62 before it reaches an F1-score of 0.75. The performance of the LSTM model advances steadily from initial 0.60 towards 0.72 throughout the epochs. The F1-score result of LR remains static at 0.58 throughout the test period due to its restricted learning capabilities compared to deep learning models. The proposed model exhibits superiority through its significant performance difference against other models in precision and recall optimisation which leads to better classification results. The proposed framework is promising, but its generalisability across different financial contexts and scalability to real-time analysis warrant further exploration. Future studies should validate its adaptability to varied datasets and larger-scale deployments.

Figure 12 illustrates the SHAP interaction values for Feature B, showing how it interacts with other features in influencing the model's output. The plot displays individual SHAP interaction values along the x-axis, with colour gradients representing the values of Feature A that interact with Feature B. Points clustered around zero suggest minimal interaction effects, while deviations toward the extremes (positive or negative) highlight instances where the combination of Feature B and another feature significantly affects the prediction. The clear separation of clusters along the y-axis suggests stratified effects, indicating distinct interaction patterns based on data subgroups. This visualisation provides valuable insights into feature dependencies and complex nonlinear effects within the model.

Figure 13 presents the prediction error plot, which visualises the relationship between the actual and predicted values of the proposed model. The blue markers represent individual prediction instances, while the red dashed line signifies the line of perfect prediction (i.e., where predicted values exactly equal actual values). The closer the blue dots are to this red line, the more accurate the model's predictions are. As shown, most of the points cluster tightly around the line, indicating that the proposed model achieves high accuracy with minimal prediction error. This reinforces the model's reliability and performance consistency across varying input scenarios.

Figure 13 Prediction error plot (see online version for colours)

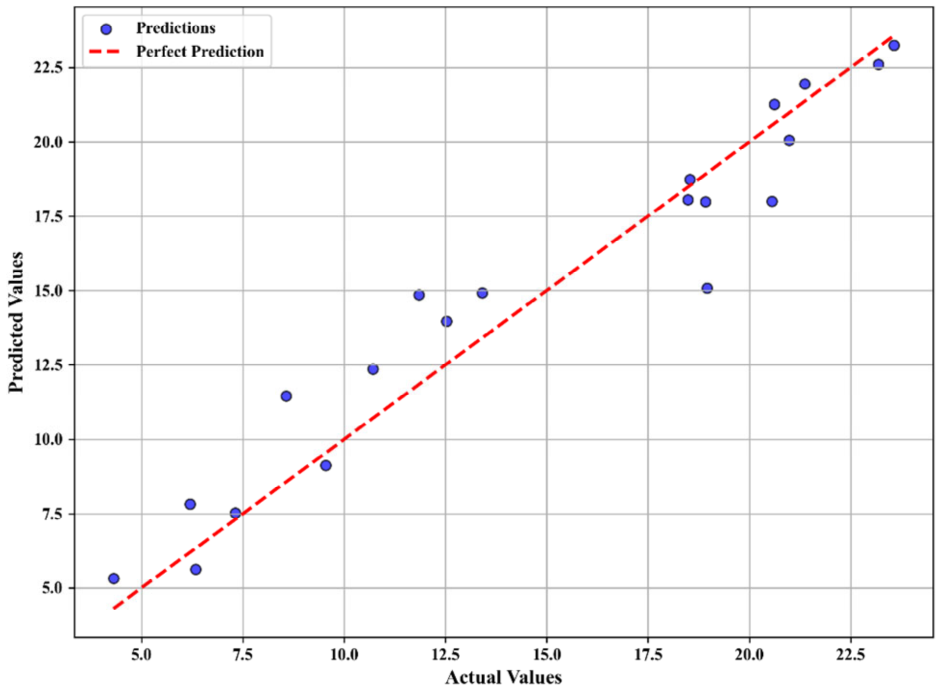


Table 2 Ablation study

Model variant	*Accuracy*	*F1-score*
Full model	87.2%	0.84
Without sentiment score	83.1%	0.79
Without optimisation step	81.4%	0.77

A brief sensitivity analysis was conducted to assess the contribution of key components to overall model performance. As shown in Table 2, removing the sentiment score module led to a noticeable drop in both accuracy and F1-score (from 87.2% to 83.1% and from 0.84 to 0.79, respectively), indicating its strong influence on prediction capability. Similarly, omitting the optimisation step further decreased performance, highlighting the importance of hyperparameter tuning for robust sentiment-based financial forecasting.

- *Real-world application scenario:* The proposed framework can be applied to MD&A reports of financially distressed firms to detect early warning signs of failure. For instance, a gradual decline in sentiment indicators across consecutive financial quarters could serve as a red flag for investors and regulators. This illustrates the framework's potential use in real-world financial risk monitoring and early intervention systems.

- *Generalisability and limitations:* While the proposed framework demonstrates strong predictive performance on the current dataset, it is important to acknowledge that the dataset consists solely of Chinese listed firms. Therefore, its generalisability to other countries, markets, or financial systems remains to be fully validated. Different regulatory environments, reporting standards, and market behaviours may affect model effectiveness. Future research should aim to test and adapt the model across diverse geographic regions and financial contexts to ensure wider applicability and robustness.
- *Ethical implications of sentiment-based forecasting:* While the integration of sentiment analysis into financial risk modelling offers significant predictive advantages, it also presents ethical challenges. These include potential biases embedded in training data that may influence model outcomes, privacy concerns related to the use of publicly sourced sentiment data (such as social media), and the unintended consequences of algorithm-driven financial decisions, especially in high-stakes cross-border contexts. Ensuring transparency, fairness, and accountability in model design is essential to avoid discriminatory impacts and promote responsible AI use in financial applications.

5 Conclusions

The proposed system includes an intelligent framework for financial distress risk assessment in cross-border M&A by combining structured financial elements and unstructured textual assessment such as semantic sentiment analytics and basic textual metrics. The research performs financial instability predictor extraction with data collection and feature engineering along with preprocessing and advanced feature selection techniques T-test and RFE-SVM and Random Forest which lead to effective high-impact predictor retrieval. A predictive model performance upgrade came from the implementation of transfer learning convolutional neural network (TL-CNN). This model benefits from pre-trained VGGNet network strength by adapting it with specialised financial and textual representation training layers. The combination between Adaptive Moment Estimation (Adam) optimiser and sparse categorical cross-entropy loss provides exact multi-class financial risk level predictions. When historian data patterns integrate real-time sentiment analysis through the TL-CNN model it produces superior performance to traditional machine learning for detecting financial trouble indicators. The merged method utilises real-time textual analysis with conventional risk methodologies which fills numerical data gaps and extends capabilities to capture global financial forces that involve country dependencies and market moods. The proposed framework delivers substantial value to financial risk prediction because it provides an intelligent scalable data-driven prediction framework. The system delivers imminent actionable information to corporate stakeholders together with investors and policymakers who thus make superior decisions to protect against international financial losses. Future work may involve extending this framework to multilingual and multi-market environments, incorporating real-time news streams and social media data, and applying advanced transformer-based models to further improve cross-domain generalisability and prediction accuracy.

Declarations

All authors have made substantial contributions to the conception, methodology, analysis, and writing of the manuscript. All authors have read and approved the final submitted version and agree with its content.

The authors declare that there are no actual or potential competing interests – financial, professional, or personal – that could be perceived as influencing the submitted work. Any future relevant interests will be disclosed in a timely and transparent manner.

The submitted work is entirely original and has not been previously published, nor is it under consideration elsewhere. Proper attribution has been provided for all cited sources, and the manuscript contains no plagiarism, data falsification, or redundant publication.

The study did not involve human participants, human data, or animal subjects; therefore, informed consent and institutional ethics approval were not required. The authors affirm their adherence to the ethical publishing standards applicable to scholarly research, including the principles outlined by Inderscience Publishers and the Committee on Publication Ethics (COPE).

No generative artificial intelligence tools were used in the writing, editing, or generation of any part of this manuscript.

All authors declare that they have no conflicts of interest.

References

Ahmadi, Z. et al. (2018) *Towards Bankruptcy Prediction: Deep Sentiment Mining to Detect Financial Distress from Business Management Reports* [online] https://dblp.org/pid/37/10813 (accessed 5 April 2025).

Ahuja, S., Singh, S. and Yadav, S.S. (2024) 'Informativeness of announcement returns in cross-border acquisition completion: a perspective from emerging economies', *Journal of Emerging Market Finance*, Vol. 23, No. 4, pp.424–449, DOI: 10.1177/09726527241247975.

Alexandre, M., Silva, T.C., Connaughton, C. and Rodrigues, F.A. (2021) 'The drivers of systemic risk in financial networks: a data-driven machine learning analysis', *Chaos, Solitons & Fractals*, Vol. 153, p.111588, DOI: 10.1016/j.chaos.2021.111588.

Bao, X. (2022) 'Construction of financial management system model based on internet technology', *Wireless Communications and Mobile Computing*, No. 1, p.7487770, DOI: 10.1155/2022/7487770.

Chen, L. (2022) 'Security modeling of overseas asset management information system under international trade and international capital flow in the era of big data', in *2022 International Conference on Applied Artificial Intelligence and Computing (ICAAIC)*, pp.775–778, IEEE, DOI: 10.1109/ICAAIC53929.2022.9797663.

Chen, Y., Wang, G-J., Zhu, Y., Xie, C. and Uddin, G.S. (2024) 'Identifying systemic risk drivers of FinTech and traditional financial institutions: machine learning-based prediction and interpretation', *The European Journal of Finance*, Vol. 30, No. 18, pp.2157–2190, DOI: 10.1080/1351847X.2023.2228760.

Cheng, W. and Chen, S. (2021) 'Sentiment analysis of financial texts based on attention mechanism of FinBERT and BiLSTM', in *2021 International Conference on Computer Engineering and Application (ICCEA)*, June, IEEE, pp.73–78, DOI: 10.1109/ICCEA53728.2021.00022.

Dong, S. et al. (2025) 'Sentiment classification for financial texts based on deep learning', *Journal of Advanced Computational Intelligence* [online] https://onlinelibrary.wiley.com/doi/10.1155/2021/9524705.

Elhoseny, M. et al. (2025) '*A Deep Learning-Based Model for Financial Crisis Prediction*, pp.1–19, Research Square, DOI: https://doi.org/10.21203/rs.3.rs-5829931/v1.

Francis, B.B., Hasan, I., Sun, X. and Waisman, M. (2014) 'Can firms learn by observing? Evidence from cross-border M&As', *Journal of Corporate Finance*, Vol. 25, pp.202–215, DOI: 10.1016/j.jcorpfin.2013.11.018.

George, M.B. and Ayiku, E.O. (2024) 'AI-driven fire risk indices integrating climate, fuel, and terrain for wildfire prediction and management', *Int. J. Eng. Technol. Res. Manag.*, Vol. 8, No. 2, p.67, DOI: 10.5281/zenodo.15043141.

Hué, S., Lucotte, Y. and Tokpavi, S. (2019) 'Measuring network systemic risk contributions: a leave-one-out approach', *Journal of Economic Dynamics and Control*, Vol. 100, pp.86–114, DOI: 10.1016/j.jedc.2018.12.001.

Ji, X., Li, J. and Hao, J. (2024) 'A bibliometric analysis of risks for cross-border business and data movement', *Procedia Computer Science*, Vol. 242, pp.1315–1322, DOI: 10.1016/j.procs.2024.04.189.

Jiang, F., Liu, Y., Meng, L. and Zhang, H. (2024) 'Deep learning, textual sentiment, and financial market', *Information Technology and Management*, pp.1–25, https://doi.org/10.1007/s10799-024-00428-z.

Jiang, L., Cao, X., Wang, Z., Zhan, Y., Zhang, J. and Chen, S. (2025) 'Promoting energy saving and emission reduction benefits in small and medium-sized enterprises supply chains through green finance – evidence based on artificial intelligence intervention', *International Review of Financial Analysis*, p.104112, DOI: 10.1016/j.irfa.2024.104112.

Jing, Y. and Yang, P. (2022) 'Research on the risk prevention of cross-border e-commerce logistics in China by applying big data technology', *Procedia Computer Science*, Vol. 208, pp.407–413, DOI: 10.1016/j.procs.2022.10.189.

Kleinow, J., Moreira, F., Strobl, S. and Vähämaa, S. (2017) 'Measuring systemic risk: a comparison of alternative market-based approaches', *Finance Research Letters*, Vol. 21, pp.40–46, DOI: 10.1016/j.frl.2017.01.003.

Kou, G., Chao, X., Peng, Y., Alsaadi, F.E. and Viedma, E.H. (2019) 'Machine learning methods for systemic risk analysis in financial sectors', DOI: 10.3846/tede.2019.8740.

Lin, R. (2021) 'Risk analysis and development research of cross-border e-commerce payment under the background of internet finance', in *International Conference on Applications and Techniques in Cyber Security and Intelligence*, pp.72–79, Springer International Publishing, Cham., DOI: 10.1007/978-3-030-79129-7_10.

Marqas, R.B. et al. (2024) *Innovative Hybrid Deep Learning Models for Financial Sentiment Analysis* [online] https://www.researchgate.net/publication/387393551 (accessed 5 April 2025).

Ngo, T. and Susnjara, J. (2016) 'Hostility and deal completion likelihood in international acquisitions: the moderating effect of information leakage', *Global Finance Journal*, Vol. 31, pp.42–56, DOI: 10.1016/j.gfj.2016.04.002.

Peng, H. (2021) 'Research on credit risk identification of internet financial enterprises based on big data', *Mobile Information Systems*, No. 1, p.1034803, DOI: 10.1155/2021/1034803.

Qiu, Y. et al. (2024) 'A novel semisupervised learning method with textual information for financial distress prediction', *International Journal of Finance & Economics* [online] https://onlinelibrary.wiley.com/doi/10.1002/for.3136.

Rudd, M.A. (2025) 'Identification of important 2024 Bitcoin research needs using an AI-assisted'key questions' approach', Available at SSRN 5146515, DOI: 10.2139/ssrn. 5146515.

Wang, H. (2022) 'Enterprise financial asset risk measurement based on embedded microprocessor security analysis', *Wireless Communications and Mobile Computing*, No. 1, p.8382504, DOI: 10.1155/2022/8382504.

Wang, L. and Gao, X. (2023) 'Construction of cross-border e-commerce financial risk analysis system based on support vector machine', *Journal of Industrial Integration and Management*, Vol. 8, No. 1, pp.25–38, DOI: 10.1142/S2424862223500028.

Wang, R., Chen, S., Huang, Z. and You, J. (2019) *Review on the Status Quo of China's Cross-border E-commerce Research Based on Bibliometrics*, Association for Information Systems AIS Electronic Library (AISeL), DOI: https://aisel.aisnet.org/whiceb2019/17.

Wei, R. and Yao, S. (2021) 'Enterprise financial risk identification and information security management and control in big data environment', *Mobile Information Systems*, No. 1, p.7188327, DOI: 10.1155/2021/7188327.

Xie, Z. (2022) 'A risk feature recognition method of cross-border financial derivatives' transaction based on fuzzy support vector machine', *Mobile Information Systems*, No. 1, p.3495504, DOI: 10.1155/2022/3495504.

Xu, X. (2020) 'The architecture of China's cross-border e-commerce platform based on the data mining technology', in *Big Data Analytics for Cyber-Physical System in Smart City: BDCPS*, 28–29 December, Shenyang, China, Springer Singapore, pp.1607–1613, DOI: 10.1007/978-981-15-2449-2_220.

Xu, Z., Cheng, X., Wang, K. and Yang, S. (2020) 'Analysis of the environmental trend of network finance and its influence on traditional commercial banks', *Journal of Computational and Applied Mathematics*, Vol. 379, p.112907, DOI: 10.1016/j.cam.2020.112907.

Yang, Y., Lv, X., Huang, Y., Yu, X., Feng, Y., Dou, P., Duan, R., Liu, Y. and Zhao, R. (2024) 'Optimization and prediction of risk factors in centralized financial management system based on prediction algorithm', *Advances in Transdisciplinary Engineering*, Vol. 56, pp.229–239.

Zhang, L., Ling, J. and Lin, M. (2023a) 'Risk management research in East Asia: a bibliometric analysis', *International Journal of Intelligent Computing and Cybernetics*, Vol. 16, No. 3, pp.574–594, DOI: 10.1108/IJICC-10-2022-0276.

Zhang, X., Zhang, X., Lee, C-C. and Zhao, Y. (2023b) 'Measurement and prediction of systemic risk in China's banking industry', *Research in International Business and Finance*, Vol. 64, p.101874, DOI: 10.1016/j.ribaf.2022.101874.

Zhang, Y., Qin, X., Tian, Y. and Xiao, S. (2024) 'Research on classification and grading of cross-border data in free trade port construction', in *2024 IEEE 9th International Conference on Data Science in Cyberspace (DSC)*, IEEE, pp.533–540, DOI: 10.1109/DSC60073 .2024.00085.

Zhou, K. (2023) 'Financial model construction of a cross-border e-commerce platform based on machine learning', *Neural Computing and Applications*, Vol. 35, No. 36, pp.25189–25199, DOI: 10.1007/s00521-023-08315-6.

Zhou, Z.F. and Du, Y.K. (2025) 'Risk warning method of cross-border flow of international business management data based on integrated classification algorithm', *International Journal of High-Speed Electronics and Systems*, p.2540139, DOI: 10.1142/S0129156425401394.