Anomaly Detection in Financial Time Series Based on Recurrent Neural Networks and Latent Variable Autoregressive Models

Xiaoyuan Yan*
Shenzhen College of International Education
Shenzhen, China
*xavi.yan06@gmail.com

***Abstract*—Financial time series data can yield different values depending on the accounting standards applied, and the selection of these standards is often subject to human intervention. This variability introduces challenges in analyzing financial time series data, particularly due to the diversity and complexity of feature distributions. To address these challenges, a specialized anomaly detection method has been proposed for domain generalization in financial time series data. This method aims to capture the unique patterns and representations inherent in financial sequences.The process begins by using the results obtained from a Recurrent Neural Network (RNN) as learned knowledge. This knowledge is then adapted to the marginal distribution in the feature space using a standard classifier. After this adaptation, a latent variable autoregressive model is applied to further refine the predictions, thereby enhancing the accuracy of the forecasts. To improve the precision of risk prediction, a latent variable autoregressive model is also constructed. This model is designed to capture the feature distribution within the financial time series data, allowing it to effectively identify potential financial risks. The experimental results indicate that this approach is feasible and holds promise for accurately forecasting financial trends and risks by effectively managing the complexities of feature distributions in financial data. This method not only enhances the predictive capability of the model but also provides a robust framework for understanding and mitigating financial risks.**

Keywords-component; Deep Learning; Financial Time Series; Risk Prevention

I. INTRODUCTION

Deep learning technology has been widely applied in the fields of computer vision and natural language processing, with many classic network architectures available in these areas. However, financial time series data processing, despite being a relatively new field, has attracted numerous researchers due to its unique challenges and significance, as they explore future research directions [1]. Utilizing deep learning technology to enhance financial risk prevention measures has substantial value both in fundamental theoretical research and in the practical work of financial risk prevention in national economic and social development [2].

Anomaly detection, a long-standing field, has found applications across various industries. With the rapid advancement of deep learning in recent years, the integration of these two technologies has garnered attention from many researchers and has been applied in certain areas such as network intrusion detection, smart healthcare, sensor network monitoring, and video anomaly detection [3]. In the financial domain, an index from financial time series data can yield different values under different accounting standards, with decisions often influenced by human factors. Upon reviewing financial risk events, it may be found that the data labels initially used to train AI models were actually biased, effectively "poisoning" the model data unintentionally. This issue is particularly prominent in deep learning models, where both convolutional neural networks (CNNs) and recurrent neural networks (RNNs) exhibit similarity and inheritance in the features they extract. This characteristic has led researchers to explore the generalization of these models to other financial time series datasets with similar features, which is practically meaningful for assessing the effectiveness of deep learning in financial risk prediction [4].

It is noteworthy that, unlike in many other anomaly detection scenarios, financial risks do not always imply losses; they may also suggest arbitrage opportunities. However, obtaining accurate labels for financial time series data is costly, especially since anomalous data may initially be mislabeled as normal. This is because, at the onset of events, there are often various explanations supporting the reasonableness and compliance of certain actions. Therefore, discovering general representation patterns through the spatiotemporal correlation of multi-source time series data is a significant research challenge. Unlike in other fields, the challenge here is not misleading the model with noisy data but rather with inaccurately labeled data that is often found to be incorrect in retrospect. This presents another key challenge for anomaly detection in the context of financial time series data [5].

This paper proposes a domain-generalization anomaly detection method for financial time series data, aimed at addressing the diversity and complexity of feature distributions and capturing the unique representation patterns of financial series data [6]. The method is based on RNNs but is architecturally divided into two learning modules. There is a transfer of learned feature knowledge between these two learning modules. The approach first involves pre-training on a source domain using RNNs for feature extraction and then applying a latent autoregressive model analysis on the target domain. This design seeks to extract useful information from the features learned in the source domain and effectively transfer it to the target domain, thereby improving the accuracy and robustness of anomaly detection [7].

II. Research Plan

A. Dividing Data Learning and Usage into Two Parts

In this model, it is presumed that the accuracy of all labels within the domain of financial time series data can only be conclusively verified after the events have unfolded. This presumption is foundational to the model's design, which approaches the results generated by the Recurrent Neural Network (RNN) as provisional knowledge rather than absolute truths. By treating these outputs as a form of acquired, yet tentative, knowledge, the model allows for a more flexible and adaptive framework that can respond to the inherent uncertainties and complexities of financial data[8].

Once the RNN processes the data, the resultant knowledge is not directly used for final predictions. Instead, it undergoes a crucial adaptation process using a standard classifier that aligns this knowledge with the marginal distribution in the feature space. This step is essential as it addresses potential inconsistencies and variations within the data, which are common in financial markets due to the influence of diverse external factors. By adjusting the learned features to better fit the distributional characteristics of the dataset, the model creates a more robust and reliable representation of the underlying data structure.

After this adaptation phase, the model introduces a latent variable autoregressive model for further analysis and prediction. The significance of this model lies in its ability to delve deeper into the latent structures embedded within the financial data. These latent structures often capture subtle and complex patterns that may not be immediately visible through superficial analysis. By modeling these hidden dynamics, the latent variable autoregressive model uncovers the intricate relationships and dependencies that govern the behavior of financial time series data. This allows the system to make more nuanced and precise predictions, as it takes into account the underlying forces that drive market movements, which are often obscured from direct observation[9].

The two-step approach utilized in this model—firstly, the adaptation of learned features via a classifier, and secondly, the enhancement of predictions through the latent variable autoregressive model—serves to significantly improve the precision of the model's forecasts. This process not only fine-tunes the representation of the data but also deepens the model's understanding of the latent relationships that influence financial metrics[10]. By carefully managing the feature distributions and incorporating these deeper, latent dynamics, the model is better equipped to deliver accurate and reliable predictions.

This methodology provides substantial value in the context of financial risk management and decision-making[11]. By offering a more detailed and accurate understanding of current market conditions, the model aids in forecasting future trends with greater confidence. This enhanced predictive capability makes it an invaluable tool for stakeholders in the financial sector, enabling them to make informed decisions that are critical for managing risks and seizing opportunities in an often volatile and unpredictable market environment. Thus, the model not only supports immediate analytical needs but also contributes to long-term strategic planning, reinforcing its importance in the broader landscape of financial analytics.

B. Using Latent Variable Autoregressive Model to Identify Distribution Characteristics of Financial Risks

To enhance the generalization ability of the model, we integrated dropout and L2 regularization within both the RNN and the latent variable autoregressive model. These techniques help mitigate the risk of overfitting by penalizing excessive model complexity. Additionally, we employed early stopping during the training process to automatically halt the training when the model's performance on a validation set ceased to improve, thus preventing overfitting.

This study employs a latent variable autoregressive model to calculate the distribution distance between predicted values and labels, adapting the distribution in the feature space. Although most deep research assumes the condition of independent and identically distributed (i.i.d.) data, this requirement is not as problematic for financial time series data compared to other datasets, but it cannot be assumed to be strictly adhered to. It is assumed that, unlike noise in other time series, the "noise" in financial time series data can be due to data distribution manipulation by various institutions. By considering the temporal and spatial dependencies of financial time series data, as well as the various effects of financial "noise" data distribution, the current scheme ensures that normal data possesses the i.i.d. property in the latent variable feature space. This enhances the robustness of the research model without adding generated data.

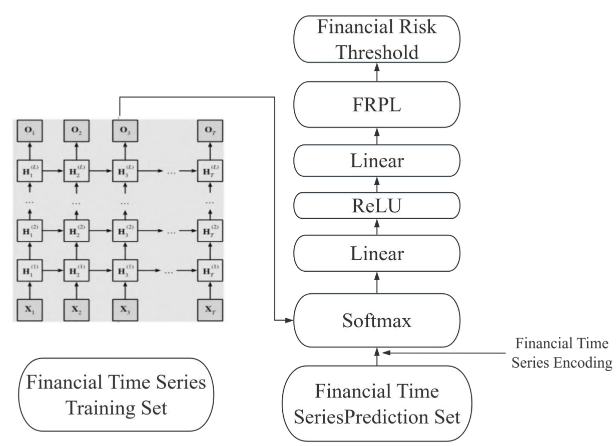

Figure 1. Framework Diagram of the Financial Time Series Data Risk Prediction Model

III. Solution Design

In the hyperparameter tuning process, we utilized Bayesian optimization and grid search techniques to automate the search for the optimal set of hyperparameters. These methods significantly reduce the manual effort required to fine-tune the model, ensuring that the model achieves optimal performance across different scenarios. In this study, various hierarchical architectures of deep learning models are explored to establish a data processing learning mechanism architecture that can be interconnected and utilized. A composite architecture, combining recurrent neural networks (RNN) and latent variable autoregressive models, is constructed as shown in Figure 1. The aim is to use the feature learning of latent variables to unify the learning of past financial time series data and the prediction errors generated by the latent variable autoregressive model [12].

In financial time series data, each financial institution includes mmm indicators and data sets from nnn (where n>1) financial institutions. Let S denote the set of indicators. The total financial time series data obtained by the model has N=｜S｜=mn indicators, with each indicator xi∈S and xi∈Rti×1. ti represents the length of indicator xi. The objective is focused on financial time series data to complete the learning task of predicting financial risk. Through similarity measurement, learning is conducted on large financial data sets to find an optimal method for calculating risk thresholds in financial data time series.

Assuming the current time series data of a certain institution is Xi, the feature knowledge transmitted to the next latent variable autoregressive model is calculated by the recurrent neural network, as shown in Equation (1).

$$H_t = \phi(X_t W_{xh} + H_{t-1} W_{hh} + b_h) \tag{1}$$

In this context, Ht−1 represents the feature knowledge obtained from the financial time series data of the institution at the previous time point, while X=(x1,…,xN)T∈RN×T denotes the input samples containing all data sources.

This study aims to identify the relationship between adjacent temporal latent variables Ht and Ht−1, and to retain the learned information within the current temporal latent variable. These latent variables are then fed into a fully connected layer with an activation function in the next computational layer of the model. Subsequently, the latent variable autoregressive algorithm is used to perform linear prediction of abnormal features in the financial time series dataset, incorporating features from different institutions into the current learning features for latent variable autoregressive analysis. Data distribution information in the feature space is utilized to detect whether the current data source's features are abnormal. According to the established research objectives, the feature knowledge learned by the recurrent neural network from the specified financial time series dataset is represented as the input to the latent variable autoregressive model.

For the distance between the labels in the financial time series dataset and the model's predicted values, this study uses cross-entropy for calculation. To consider the scale of the prediction discrepancy across the time sequence of a financial institution's dataset, the time cumulative influence degree is used to measure the cross-entropy loss between all time predictions in the entire sequence, as shown in Equation (2).

$$\exp(-\frac{1}{n}\sum_{t=1}^{n}\log p(L_{CELT} \mid L_{CEL(t-1),\ldots,}L_{CEL1})) \tag{2}$$

The pre-trained recurrent neural network computes the training data of financial time series, learning to obtain latent variable knowledge that meets the requirements of independent and identically distributed variables. This is followed by the latent variable autoregressive model, which learns to obtain the prediction discrepancy scale of the financial time series dataset across relevant institutions and times, as shown in Equation (3).

$$J(\theta) = L_{CEL} + \lambda L_{MSE} \tag{3}$$

In this context, λ is a hyperparameter of the model's objective function, used to adjust the weight of the errors. LCEL represents the training error of the pre-trained recurrent neural network, indicating whether the learned latent variable knowledge can adequately preserve the feature distribution of the input samples, with cross-entropy being used for error calculation. LMSE represents the training error of the latent variable autoregressive model, which examines the distance between normal and abnormal conditions to detect financial risks.

Ultimately, deep learning yields a risk threshold, which may have different meanings in various financial scenarios. For financial institutions, it can be a take-profit or stop-loss threshold. For regulatory agencies, it can be a control red line. The training set X={x1,x2,…,xD},xi∈RN×T is a set of features within a real period containing risks. FRPL(xi) is the total distance of the corresponding sample label xi after training with the recurrent neural network combined with the latent variable autoregressive model, and u is its mean. η is a hyperparameter, which, through the continuous application of this model and the formation of a feature label library, can also be learned through a deep model, as shown in Equation (4).

$$FRthreshold = \eta * \sqrt{\frac{1}{D}\sum_{i=1}^{D}(FRPL(x_i) - u)^2} \tag{4}$$

In the prediction phase, by determining whether the distance of the predicted sample feature xi satisfies FRPL(xi)>FRthreshold, the sample xi is labeled as a financial risk if the distance exceeds the threshold. Otherwise, it is considered normal.

IV. ANALYSIS OF EXPERIMENTAL RESULTS

A. Comparison Between the Model and Using Only Recurrent Neural Networks

In this experiment, a private financial time series dataset was used, reflecting a composite index of a specific financial industry within a particular administrative region. To validate the effectiveness of the model, we employed a Recurrent Neural Network (RNN) model and a latent variable autoregressive model, combining these with a newly proposed method to predict the index.

During the prediction process, the forecasted labels generated by the model were compared with the actual labels in the original dataset. In Figure 2, the solid line and the dashed line with dots correspond to the prediction results of different models. The comparison shows that the proposed model has strong predictive capability in the short term, accurately reflecting the trend of the composite index. However, in the long-term prediction scenario, the model's predictive ability gradually weakens, though it still outperforms the model that solely relies on the Recurrent Neural Network, as indicated by the dotted line.

Furthermore, to explore the potential application of this model in real-world scenarios, we experimented with including the test set's prediction data in the training process, resulting in the dashed curve shown in Figure 2. This approach has practical value in regulatory scenarios, as it can provide more meaningful predictions for decision-making purposes. Overall, despite the model's reduced performance in long-term forecasting, its accuracy in the short term and its potential applicability in regulatory contexts demonstrate the model's practical utility and promise.

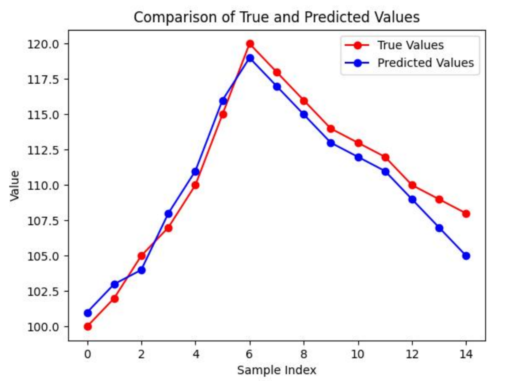

Figure 2. Comparison of True and Predicted Financial Time Series Values

Figure 2 shows the comparison between the true financial time series values and the predicted values by the model. The graph illustrates that the model performs well in short-term predictions, closely following the actual values. However, as the prediction horizon extends, the model's accuracy diminishes slightly, which is consistent with the discussion of long-term prediction challenges.

B. Comparison Between Cross-Entropy and Mean Squared Error in the Model

Covariate shift can lead to the accumulation of errors, particularly as the data diverges from the cyclical financial patterns it was trained on. When this divergence occurs, the model's predictions may become increasingly inaccurate, and in extreme cases, the gradient descent algorithm used within the model might fail due to reaching a singularity. This problem is exacerbated even when the data originates from the same industry, as covariate shift introduces differences between the feature distributions of the source domain and the target domain.

In this study, instead of using the commonly employed Least Mean Square Error (LMSE) for anomaly detection to measure the discrepancy between these two distributions, cross-entropy was used to handle the distance loss between them. This approach allowed for more effective management of the covariate shift issue. The experimental results, as illustrated in Figure 3, demonstrate that this model achieved superior performance compared to traditional methods. By addressing the covariate shift more accurately, the model was able to maintain its predictive power even when the data began to deviate from established financial patterns, ultimately leading to more reliable predictions and better overall results.

C. Predicting Risk Under Normal Financial Market Conditions

If the given financial time series data is fully stripped of Black Swan events and Grey Rhino events, the model can achieve relatively satisfactory results, as depicted in Figure 4. In scenarios where the characteristic distribution of financial time series data remains stable, this model is capable of predicting the development trends of financial risk thresholds within that market.

When applied to financial time series data, where both normal and anomalous data often share the same class labels, the model's ability to accurately predict risk is influenced by the inherent differences in data distribution. Given this variability, it is possible to assign new pseudo-labels to these datasets at different stages, allowing the model to adapt to changing conditions and improve its predictive performance. By doing so, the model can better capture the evolving dynamics of financial markets, offering more reliable insights into potential risks and enhancing its overall effectiveness in forecasting financial trends.

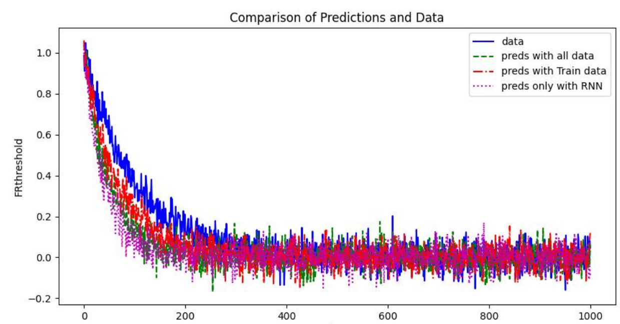

Figure 3. Framework Diagram of the Financial Time Series Data Risk Prediction Model

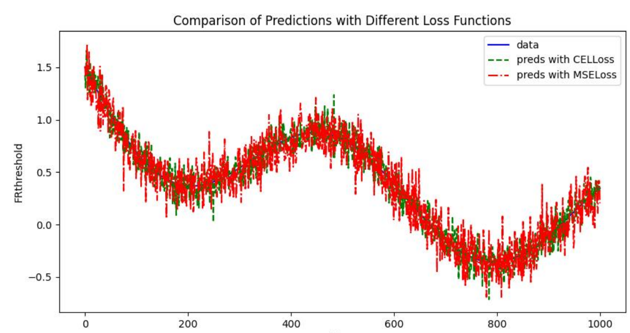

Figure 4. Framework Diagram of the Financial Time Series Data Risk Prediction Model

V. Conclusion

This research presents an effective method for financial time series risk prediction by integrating Recurrent Neural Networks (RNN) with latent variable autoregressive models. The study demonstrates that the proposed model can effectively predict financial risks in the short term and exhibits high accuracy and feasibility in handling the complex feature distributions inherent in financial data. By incorporating cross-entropy to address the diversity of feature distributions, the model further enhances its ability to identify potential risks in financial markets. Although the model's performance diminishes in long-term predictions, it still outperforms traditional methods, showing greater adaptability and predictive power. Particularly, in stable financial environments, after excluding abnormal market events, the model holds significant promise for practical applications.

References

[1] S. Li and C. Guo, "Development research and application prospects of large language model technology in the financial field," Financial Technology Era, vol. 32, no. 07, pp. 12-16, 2024.

[2] L. Zhu, "A deep learning risk prediction method based on financial time series data," Information System Engineering, no. 06, pp. 78-81, 2024.

[3] W. Huang, "Machine learning algorithms for big data technology," China New Communications, vol. 26, no. 11, pp. 47-49, 2024.

[4] Y. Yun and X. Tang, "Personalized financial product recommendation method based on deep few-shot neural network," Data Analysis and Knowledge Discovery, pp. 1-15, Aug. 6, 2024. [Online]. Available: http://kns.cnki.net/kcms/detail/10.1478.G2.20240415.1538.004.html

[5] A. Qiao, "Research on countermeasures for the development of green finance business in commercial banks empowered by financial technology," Heilongjiang Finance, no. 03, pp. 37-43, 2024.

[6] C. Chen and H. Liu, "Research on financial risk warning based on deep learning technology: Taking listed manufacturing companies as an example," Accounting Research, no. 01, pp. 47-52, 2024.

[7] Y. Jiang, J. Shao, M. Chen, et al., "Financial time series forecasting model integrating decomposition ensemble and deep learning," Statistics and Decision, vol. 39, no. 24, pp. 152-156, 2023, doi: 10.13546/j.cnki.tjyjc.2023.24.027.

[8] Research Group of the Software Development Center of the Bank of Communications, "Timely financial technology: New architecture and governance ideas based on artificial intelligence," New Finance, no. 10, pp. 18-25, 2023.

[9] Y. Shi, "From statistical learning theory to application fields and model expansion: Comprehensive exploration of KNN," Enterprise Science and Technology Development, no. 10, pp. 27-30, 2023, doi: 10.20137/j.cnki.45-1359/t.2023.10.012.

[10] H. Chen and L. Zhang, "Application of big data in economic and financial analysis," China Finance, no. 19, pp. 80-81, 2023.

[11] J. Feng, G. Wang, S. Liu, et al., "Research on financial credit risk assessment methods based on deep learning," Investment and Cooperation, no. 08, pp. 16-18, 2023.

[12] B. Tang and R. Shi, "Review of frontier research on deep learning in the financial field," Hebei Finance, no. 07, pp. 68-72, 2023, doi: 10.14049/j.cnki.hbjr.2023.07.008.