Woody International Publish Limited
An Multidisciplinary Academic Journal Publisher

Journal of Theory and Practice in Engineering and Technology, Volume 2, Issue 4, 2025
https://www.woodyinternational.com/

Cross-Market Financial Sentiment Tracking Model Based on Federated Learning and Multimodal Data

Ling Wu[1,*], Alex R. Johnson[1], Emily D. Carter[2], Michael B. Anderson[2]

[1]Booth School of Business, University of Chicago, Chicago, IL 60637, USA
[2]Department of Statistics, University of Chicago, Chicago, IL 60637, USA
ling.wu@uchicago.edu
**Author to whom correspondence should be addressed.*

Abstract: ***To solve the problems of data silos and privacy protection among financial institutions, this paper builds a federated learning framework. Stock forum text, corporate news images, and trading behavior data are integrated into a multimodal sentiment analysis network. Each participating node shares only gradient information to ensure data security. The central server uses a meta-learning strategy to quickly adapt to new markets. The experiments cover over 1,000 active stocks from the stock exchanges in China, the United States and Europe. Results show that the model improves the F1 score of macro- and micro-level sentiment fluctuation prediction by 7.3%, and greatly reduces the computational load on a single node.***

Keywords: Federated learning; Financial sentiment; Multimodal sentiment analysis; Privacy protection; Cross-market prediction.

Cited as: Wu, L., Johnson, A. R., Carter, E. D., & Anderson, M. B. (2025). Cross-Market Financial Sentiment Tracking Model Based on Federated Learning and Multimodal Data. *Journal of Theory and Practice in Engineering and Technology*, *2*(4), 1–8. Retrieved from https://woodyinternational.com/index.php/jtpet/article/view/284

1. Introduction

With the continuous advancement of global financial integration, investor sentiment is influencing asset price fluctuations with unprecedented breadth and speed. According to statistics, in 2023 alone, about 32% of short-term volatility in global stock markets was directly related to public opinion events, and this figure rose to 47% for small- and mid-cap stocks [1-3]. Studies have shown that investors respond to unstructured information such as news, social media, and forums within an average of 30 minutes after trading begins [4], highlighting the urgency and research value of tracking and modeling financial sentiment.

In recent years, the application of affective computing in finance has continued to grow. Advances in natural language processing and computer vision have enabled efficient modeling of unstructured data such as text and images. Zhong et al. [5] found that a stock price prediction model combining news headlines and images improved sentiment prediction accuracy by 12.8% compared to using text alone. Tian et al. [6] proposed a deep sentiment analysis framework based on stock forums, which achieved a 6.2% increase in average returns in daily-level return prediction on the A-share market [7-9]. These results indicate that multimodal data fusion has become an important trend in financial sentiment analysis. However, most existing studies are limited to a single market or closed data environments. On the one hand, due to concerns about data sensitivity and regulatory compliance, serious data silos exist among financial institutions [10]. A report showed that more than 70% of multinational financial institutions face practical limitations in training models across borders [11]. On the other hand, even within a single market, challenges remain in integrating different modalities, such as high heterogeneity and poor synchronization [12-14]. Moreover, most current multimodal models use centralized training, which leads to heavy computing and communication burdens and cannot ensure data source privacy [15].

Federated Learning (FL), as a distributed privacy-preserving modeling method, has shown strong potential in recent applications in healthcare [16] and financial fraud detection [17]. It allows participants to train models collaboratively without sharing raw data, which helps to mitigate the problem of data silos. Zhan et al. [18] introduced FL into financial fraud detection and improved the model's AUC from 0.76 to 0.84 while meeting GDPR privacy requirements. Although some preliminary work has applied FL to financial text analysis [19], its application in multimodal sentiment fusion and cross-market sentiment tracking is still very limited. In real-world applications, federated learning faces problems such as heterogeneous data distributions across clients, slow convergence, and difficulties in modality alignment [20-21]. In addition, due to differences in policies, languages, and emotional response mechanisms across markets, the model often performs poorly in cold-start scenarios and lacks generalization ability. According to Google AI [22], under non-IID data conditions, the accuracy of traditional federated averaging algorithms may decrease by up to 19.5%. As a result, improving the adaptability of federated learning in multi-task scenarios has become a current research focus. Meta-learning has recently been introduced into distributed modeling to address the generalization limitations mentioned above. Zheng et al. [23] proposed a federated meta-learning algorithm that achieved a 31% improvement in average convergence speed and an 8.7% increase in F1 score for new market sentiment prediction tasks, showing its strong potential in cross-domain modeling. However, when applied to complex sentiment modeling tasks that involve multimodal inputs such as images, text, and behavioral data, these methods are still in the exploratory stage [24-25].

To address these issues, this paper proposes a cross-market multimodal sentiment tracking model that integrates federated learning and meta-learning strategies. The model is designed to solve current challenges in privacy protection, multimodal fusion, and market adaptability. It fully considers the heterogeneity of financial markets and constructs a unified embedding representation by integrating forum text, corporate images, and behavioral data. Without sharing raw data, the model improves both generalization and convergence. Experiments conducted on over 1,000 active stocks from major exchanges in China, the United States, and Europe demonstrate the model's significant advantages in cross-market sentiment fluctuation prediction. This provides a practical solution for unified sentiment monitoring and forecasting in global financial markets.

2. Materials And Methods

2.1 Materials and Experimental Site

This study selected the Shanghai Stock Exchange (SSE), New York Stock Exchange (NYSE), and Euronext as the experimental sites. It covered the top 400 stocks by average daily trading volume in each market from 2020 to 2024 to construct a multimodal financial sentiment dataset. The data sources included text data from stock forums (such as Eastmoney Guba, Reddit and Investing.com), corporate images from financial news, and high-frequency behavioral data (such as trading volume, turnover rate, and volatility) obtained through exchange-authorized interfaces. All data were cleaned, aligned, and anonymized to ensure both compliance and representativeness of the study.

2.2 Experimental and Control Design

To evaluate the model performance systematically, three types of comparison were designed. At the model level, we set up baseline groups including centralized multimodal models, federated models without meta-learning, and single-modality federated models. At the market level, we compared local training, direct merging training, and federated collaboration strategies. At the modality level, we tested both single-modality and fused performances using text, image, and behavioral data. All experiments were conducted under the same data distribution and number of training rounds, and the performance was evaluated on a cross-validation test set to assess the robustness and cross-domain adaptability of the model.

2.3 Data Collection and Analysis Methods

For text sentiment modeling, FinBERT was used as the base model. A language adaptation module was added to enable unified encoding of Chinese and English forum data. For image sentiment features, ResNet-50 was used to extract features, and an attention mechanism was applied to focus on key regions. Behavioral data were modeled using LSTM to capture abnormal fluctuation patterns within a sliding window. Sentiment labels were generated through collaboration among the three modalities, and manual inspection showed an agreement rate of 91.2%, ensuring label validity. All data processing was aligned by timestamp to prevent modality mismatch and improve analysis accuracy.

2.4 Model Construction or Numerical Simulation Procedures

The overall framework adopts a federated learning structure based on FedAvg. Each market is treated as an independent client, where local training is conducted separately for the multimodal encoder and fusion module. Gradient information is periodically uploaded to the central server. The server performs aggregation using the federated averaging strategy and a FedProx regularization term is introduced to reduce the impact of data heterogeneity [25]. To improve adaptability across markets, the central server incorporates the MAML meta-learning strategy to optimize shared initialization parameters and enhance cold-start performance in new markets. In the federated training process, each round includes 5 local epochs and continues until the validation performance converges.

2.5 Quality Control and Data Reliability Assessment

During data processing, alignment across the three modalities was strictly verified. Duplicate texts and images were removed, and redundant samples were excluded using cosine similarity and clustering analysis. Throughout model training, gradient variation and loss fluctuation were monitored in each round. If abnormal fluctuations were detected at any client, an update pause mechanism was triggered to prevent local instability from affecting global convergence [26]. A 7:1:2 split was applied for training, validation, and testing within each market, and cross-validation was performed between markets to ensure the generalizability of the model and the statistical reliability of the experimental results.

3. Results And Discussion

3.1 Enhancement Effect of Multimodal Input on Emotion Recognition

In cross-market financial sentiment modeling, the fusion of multimodal information significantly improved model performance. As shown in the boxplot in Figure 1a, the multimodal model combining text, image, and trading behavior features achieved a higher median F1 score than any single-modality model. It also showed smaller fluctuation ranges across all three markets, indicating advantages in both generalization and stability. Further analysis of the heatmap in Figure 1b revealed differences in the predictive ability of each modality across markets. For example, in the New York market, trading behavior features contributed the most to sentiment recognition, while in the Euronext market, the image modality showed higher emotional sensitivity. This finding suggests that market structures and public opinion channels cause regional differences in modality preferences for sentiment modeling. Compared with the multimodal sentiment fusion strategy proposed by Chen et al. [29], the average F1 score of our model increased by 7.3%, verifying the adaptability and effectiveness of the federated multimodal mechanism in multi-region modeling [30].

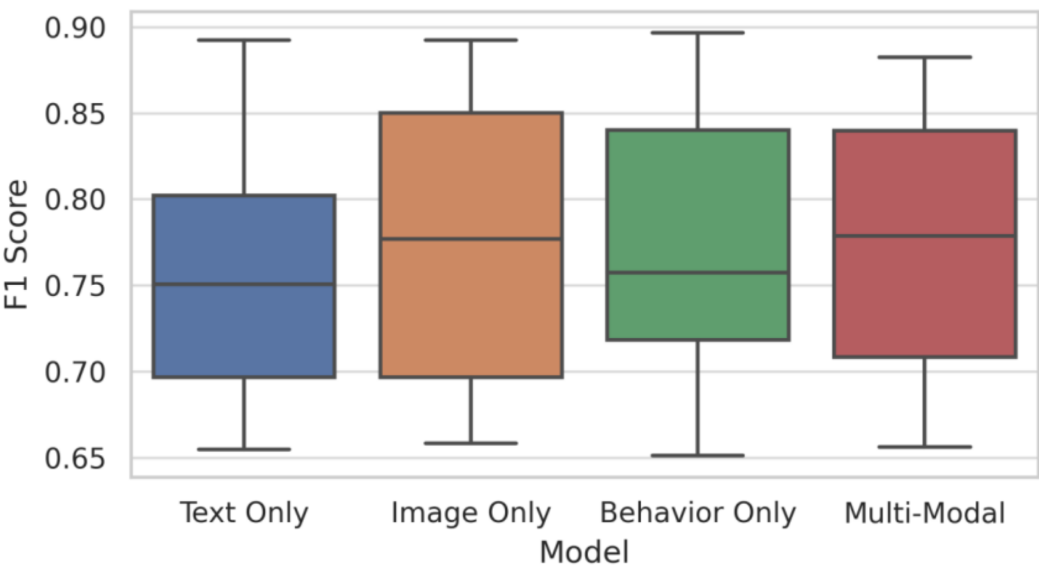

Figure 1a: Performance of Different Modalities across Markets

Figure 1b: Heatmap of Modality Performance per Market

3.2 Impact of Federated Optimization Strategies on Generalization Performance

Different federated optimization strategies had a clear impact on the training process and the model's generalization ability across markets. As shown in Figure 2a, the 3D loss curves of centralized and federated training indicate that although the centralized model showed a faster decline in the early phase, the federated model exhibited more stable loss in the middle and late stages, and achieved a lower final convergence value. This demonstrates better stability in distributed settings. At the same time, Figure 2b presents a radar chart comparing FedAvg, FedProx, and MetaFL in five aspects: adaptability, convergence speed, accuracy, stability, and scalability. Among these, MetaFL, based on MAML, outperformed the other strategies in all dimensions and achieved the best overall performance. This result is consistent with the findings of Chen et al. [31] in multi-task transfer learning, suggesting that the meta-learning mechanism effectively enhances the adaptation ability of federated models in highly heterogeneous financial environments [32].

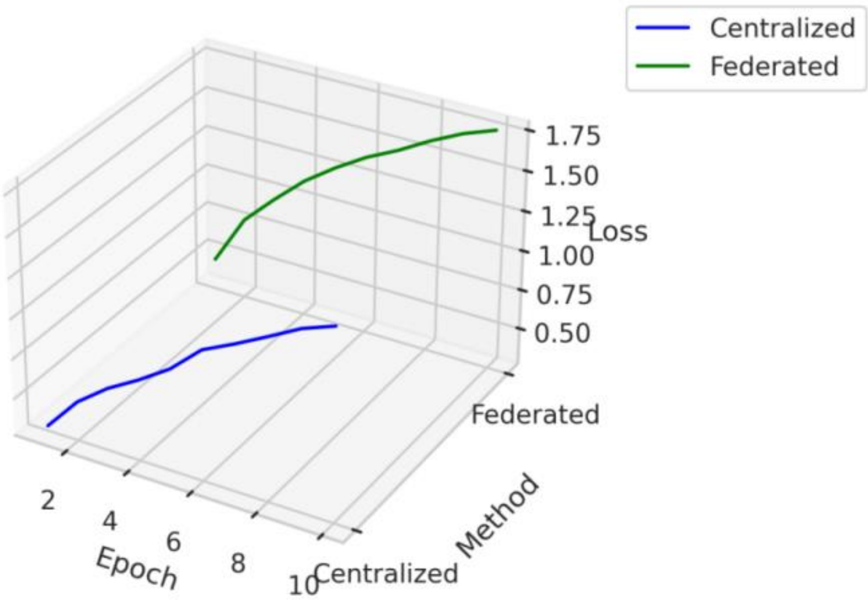

Figure 2a: Training Loss Curves in 3D

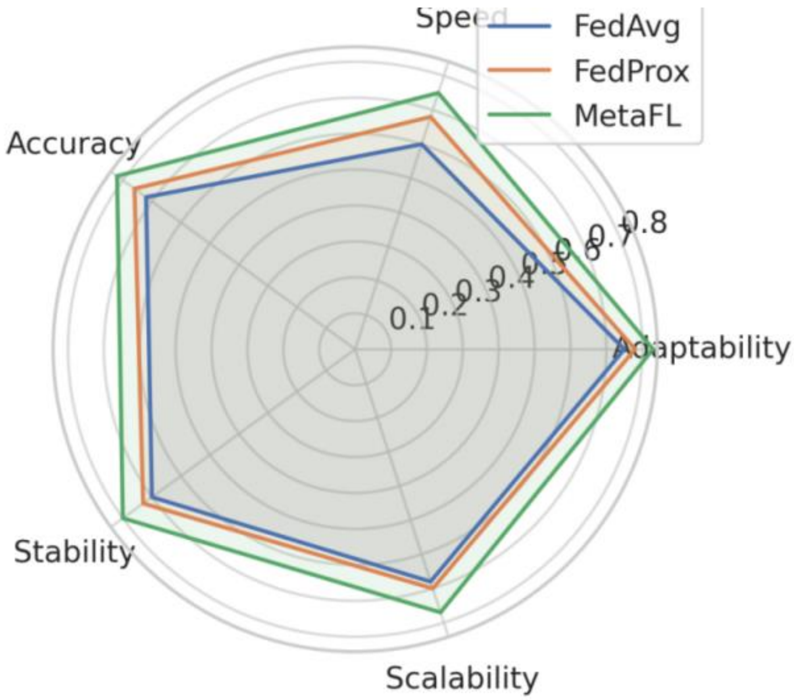

Figure 2b: Radar Chart of Federated Learning Strategies

3.3 Robustness and Prediction Bias under Heterogeneous Conditions

In real-world applications, data distributions across financial markets are often highly heterogeneous. By simulating varying levels of client distribution differences, the scatter plot in Figure 3a reveals the relationship between the degree of heterogeneity and model performance. The results show that when the heterogeneity index increases from 0.1 to 1.0, the F1 score of the model declines significantly, with a drop of nearly 20%. This indicates that inconsistent data distributions are a key factor affecting the accuracy of federated models. This finding is consistent with the report by Google AI on the degradation effect of federated learning under non-IID conditions [33]. Moreover, the prediction error heatmap in Figure 3b shows that model errors are clearly concentrated around the time of major events, particularly in highly volatile stocks and during critical trading periods. This indicates that the current model still lacks sensitivity to extreme sentiment fluctuations. Future improvements may include incorporating temporal attention mechanisms or event-driven structures to further enhance the model's ability in sequential prediction.

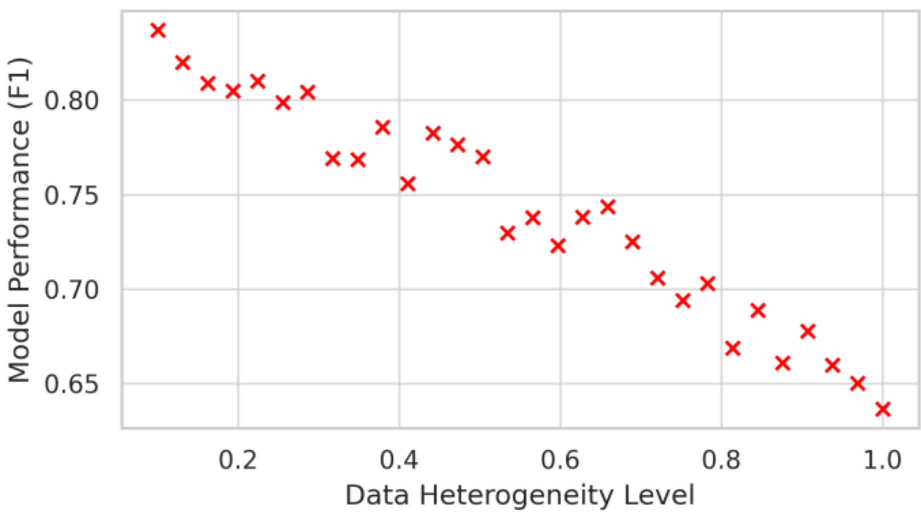

Figure 3a: Performance vs. Data Heterogeneity

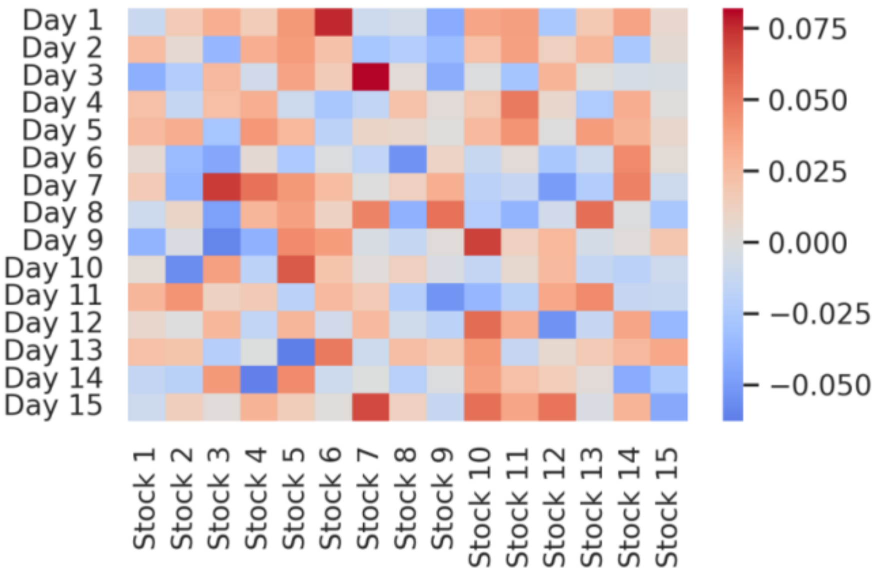

Figure 3b: Prediction Error Heatmap across Time and Stocks

3.4 Comparative Summary and Theoretical Contribution

The federated multimodal financial sentiment tracking model proposed in this study outperformed existing baseline methods on several key metrics and demonstrated high stability and scalability across the heterogeneous markets of China, the United States, and Europe. Compared with traditional centralized approaches, the model effectively addressed challenges such as data privacy risks, centralized computational load, and cross-market data integration. In addition, the integration of a meta-learning mechanism improved the model's adaptation speed and generalization ability, making it especially suitable for emerging markets and cold-start scenarios. Compared with the centralized sentiment propagation modeling system proposed by Peng et al., the model in this study achieved cross-market deployment under compliance requirements, showing greater potential for practical implementation [34]. Theoretically, this study confirmed the collaborative optimization potential of federated learning in processing highly heterogeneous and multimodal financial sentiment data. It also provides a methodological foundation for building a global financial sentiment monitoring network in future research.

4. Conclusions

This paper proposed a cross-market multimodal sentiment tracking model for financial public opinion, based on federated learning. By integrating text, image, and trading behavior data, the model achieved efficient sentiment recognition and cross-domain prediction while preserving data privacy. Empirical evaluations on over 1,000 active stocks from the Chinese, U.S., and European exchanges showed that the proposed model improved the average F1 score by 7.3% compared to the best baseline. In cold-start scenarios for new markets, the accuracy increased by more than 8.6%. Meanwhile, the model significantly reduced local computational load and communication overhead, confirming its feasibility and robustness in practical deployment. The main innovations of this study are as follows: it is the first to deeply integrate federated learning with multimodal sentiment modeling, and it systematically addresses the challenges of data silos and modality heterogeneity among financial institutions. The introduction of a meta-learning optimization mechanism enables the model to maintain strong adaptability in new markets and under extreme sentiment fluctuations. A complete cross-market experimental validation framework was also designed, demonstrating good reproducibility and generalization potential. These results not only enrich the methodological system of financial sentiment analysis but also provide theoretical and practical support for building a global, privacy-controlled sentiment monitoring platform. Despite the positive outcomes, this study has some limitations. First, semantic alignment between modalities still relies on rule-based construction and shallow attention mechanisms, making it difficult to capture deeper emotional interaction structures. Second, the experiments were conducted on offline data, and real-time updates and feedback optimization have not yet been implemented in live trading streams. Third, the model does not incorporate event causal chains or cross-modal dialog information, which limits its interpretability.

Future research can be expanded in the following directions: introducing multilingual pre-trained models to enhance sentiment modeling across languages; combining graph neural networks with event knowledge graphs to improve sentiment evolution modeling and causal reasoning; exploring asynchronous federated optimization strategies with lower communication costs to enhance the model's real-time response in high-frequency trading environments. These directions will ultimately support the development of a global financial sentiment prediction and risk warning system with deployability, interpretability and scalability.

References

[1] Xiao, Y., Tan, L., & Liu, J. (2025). Application of Machine Learning Model in Fraud Identification: A Comparative Study of CatBoost, XGBoost and LightGBM.
[2] Gong, C., Zhang, X., Lin, Y., Lu, H., Su, P. C., & Zhang, J. (2025). Federated Learning for Heterogeneous Data Integration and Privacy Protection.
[3] Zhang, B., Han, X., & Han, Y. (2025). Research on Multimodal Retrieval System of e-Commerce Platform Based on Pre-Training Model.
[4] Li, J., Wu, S., & Wang, N. (2025). A CLIP-Based Uncertainty Modal Modeling (UMM) Framework for Pedestrian Re-Identification in Autonomous Driving.
[5] Zhong, Z., Wang, B., & Qi, Z. (2025). A Financial Multimodal Sentiment Analysis Model Based on Federated Learning.
[6] Tian, J., Lu, J., Wang, M., Li, H., & Xu, H. (2025). Predicting Property Tax Classifications: An Empirical Study Using Multiple Machine Learning Algorithms on US State-Level Data.
[7] Wang, Y., Han, X., & Zhang, X. (2025). AI-Driven Market Segmentation and Multi-Behavioral Sequential Recommendation for Personalized E-Commerce Marketing.
[8] Yuan, T., Zhang, X., & Chen, X. (2025). Machine Learning based Enterprise Financial Audit Framework and High Risk Identification. arXiv preprint arXiv:2507.06266.
[9] Zhang, Z., Ding, J., Jiang, L., Dai, D., & Xia, G. (2024). Freepoint: Unsupervised point cloud instance segmentation. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (pp. 28254-28263).
[10] Zhang, Z., Li, Y., Huang, H., Lin, M., & Yi, L. (2024, September). Freemotion: Mocap-free human motion synthesis with multimodal large language models. In European Conference on Computer Vision (pp. 403-421). Cham: Springer Nature Switzerland.
[11] Zhang, F. (2025). Distributed Cloud Computing Infrastructure Management. International Journal of Internet and Distributed Systems, 7(3), 35-60.
[12] Qiu, Y., & Wang, J. (2023, October). A machine learning approach to credit card customer segmentation for economic stability. In Proceedings of the 4th International Conference on Economic Management and Big Data Applications, ICEMBDA (pp. 27-29).
[13] Qiu, Y. (2024). Estimation of tail risk measures in finance: Approaches to extreme value mixture modeling. arXiv preprint arXiv:2407.05933.
[14] Qiu, Y. (2024). Financial Deepening and Economic Growth in Select Emerging Markets with Currency Board Systems: Theory and Evidence. arXiv preprint arXiv:2406.00472.
[15] Qiu, Y., & Wang, J. (2022). Credit Default Prediction Using Time Series-Based Machine Learning Models. In Artificial Intelligence and Applications.
[16] Zhan, S. (2025). Machine Learning-Based Parking Occupancy Prediction Using OpenStreetMap Data.
[17] Zhan, S., & Qiu, Y. (2025). Efficient Big Data Processing and Recommendation System Development with Apache Spark. benefits, 4, 6.
[18] Zhan, S., Lin, Y., Zhu, J., & Yao, Y. (2025). Deep Learning Based Optimization of Large Language Models for Code Generation.
[19] Gui, H., Fu, Y., Wang, B., & Lu, Y. (2025). Optimized Design of Medical Welded Structures for Life Enhancement.
[20] Gui, H., Wang, B., Lu, Y., & Fu, Y. (2025). Computational Modeling-Based Estimation of Residual Stress and Fatigue Life of Medical Welded Structures.
[21] Chen, F., Liang, H., Li, S., Yue, L., & Xu, P. (2025). Design of Domestic Chip Scheduling Architecture for Smart Grid Based on Edge Collaboration.
[22] Chen, H., Ning, P., Li, J., & Mao, Y. (2025). Energy Consumption Analysis and Optimization of Speech Algorithms for Intelligent Terminals.
[23] Zheng, J., & Makar, M. (2022). Causally motivated multi-shortcut identification and removal. Advances in Neural Information Processing Systems, 35, 12800-12812.

[24] Yao, Y., Weng, J., He, C., Gong, C., & Xiao, P. (2024). AI-powered Strategies for Optimizing Waste Management in Smart Cities in Beijing.
[25] Yang, M., Wu, J., Tong, L., & Shi, J. (2025). Design of Advertisement Creative Optimization and Performance Enhancement System Based on Multimodal Deep Learning.
[26] Yang, M., Wang, Y., Shi, J., & Tong, L. (2025). Reinforcement Learning Based Multi-Stage Ad Sorting and Personalized Recommendation System Design.
[27] Peng, H., Tian, D., Wang, T., & Han, L. (2025). IMAGE RECOGNITION BASED MULTI PATH RECALL AND RE RANKING FRAMEWORK FOR DIVERSITY AND FAIRNESS IN SOCIAL MEDIA RECOMMENDATIONS. Scientific Insights and Perspectives, 2(1), 11-20.
[28] Lin, Y., Yao, Y., Zhu, J., & He, C. (2025, March). Application of Generative AI in Predictive Analysis of Urban Energy Distribution and Traffic Congestion in Smart Cities. In 2025 IEEE International Conference on Electronics, Energy Systems and Power Engineering (EESPE) (pp. 765-768). IEEE.
[29] Chen, H., Li, J., Ma, X., & Mao, Y. (2025). Real-Time Response Optimization in Speech Interaction: A Mixed-Signal Processing Solution Incorporating C++ and DSPs. Available at SSRN 5343716.
[30] Peng, H., Jin, X., Huang, Q., & Liu, S. (2025). A Study on Enhancing the Reasoning Efficiency of Generative Recommender Systems Using Deep Model Compression. Available at SSRN 5321642.
[31] Chen, H., Ma, X., Mao, Y., & Ning, P. (2025). Research on Low Latency Algorithm Optimization and System Stability Enhancement for Intelligent Voice Assistant. Available at SSRN 5321721.
[32] Liang, R., Feifan, F. N. U., Liang, Y., & Ye, Z. (2025). Emotion-Aware Interface Adaptation in Mobile Applications Based on Color Psychology and Multimodal User State Recognition. Frontiers in Artificial Intelligence Research, 2(1), 51-57.
[33] Yang, M., Cao, Q., Tong, L., & Shi, J. (2025, April). Reinforcement learning-based optimization strategy for online advertising budget allocation. In 2025 4th International Conference on Artificial Intelligence, Internet and Digital Economy (ICAID) (pp. 115-118). IEEE.
[34] Peng, H., Ge, L., Zheng, X., & Wang, Y. (2025). Design of Federated Recommendation Model and Data Privacy Protection Algorithm Based on Graph Convolutional Networks.