ISSN: 3079-9325

Intelligent Classification of Residential Property Tax Levels Based on Temporal Clustering and Federated Learning

[1] James Whitmore*, [2] Priya Mehra, [3] David R. Kim, [3] Emily Linford

[1] Department of Computer Science, University of Leeds, United Kingdom.

[2] School of Informatics, University of Edinburgh, United Kingdom.

[3] Department of Engineering Science, University of Oxford, United Kingdom.

* Corresponding author: j.t.whitmore@leeds.ac.uk

Abstract

To address the problem of inconsistent and highly sensitive property tax level classifications across multiple U.S. states, this study proposes a distributed intelligent classification model based on temporal feature clustering and federated learning. The method first uses time windows to analyze property transaction histories and tax records and constructs typical property patterns. Then, a federated framework is used to jointly train classifiers across multiple state governments, avoiding data leakage. The model achieves an average accuracy of 87.5% on validation sets from seven states and can automatically adjust tax level ranges based on land price trends.

Keywords

Property tax classification; Temporal clustering; Federated learning; Privacy protection; Intelligent assessment.

1. Introduction

In the United States, property tax is one of the main sources of local government revenue. In 2023, total property tax collected across all states exceeded $338 billion, accounting for nearly 30% of total local tax revenue [1]. However, the autonomy granted to states—and even to counties within the same state—regarding the classification of residential property tax levels has led to significant differences in tax rates, valuation methods, and classification rules [2]. For example, in 2022, the average residential property tax rate in New Jersey reached 2.23%, while in Alabama it was only 0.41%, resulting in a more than fivefold difference [3]. This highly fragmented taxation structure reduces the fairness of tax collection and presents serious challenges for cross-regional governance and policy standardization [4].

Traditional property tax classification methods mostly rely on static geographic valuation or manually defined rules [5]. These methods cannot effectively respond to fast-changing market conditions. According to Zillow’s data from the first quarter of 2024, residential prices in most major U.S. cities increased by over 6% year-on-year. In cities like Austin and Tampa, the increases reached 13.4% and 11.9%, respectively. Under such conditions, valuation approaches

ISSN: 3079-9325

based on historical averages are no longer able to accurately reflect land price trends or match appropriate tax ranges [6-7]. This can easily lead to unreasonable taxation or delays in tax rate adjustment. At the same time, property tax-related data contain sensitive information about assets, residents and transactions, which raises significant privacy risks and compliance challenges [8].

To improve the accuracy of property valuation and tax modeling, researchers have introduced various machine learning methods in recent years, including XGBoost, CatBoost and deep neural networks [9]. These methods have achieved some success in handling multi-dimensional structures and transaction features. However, most existing studies still focus on static attributes and fail to fully utilize the time-series structure embedded in property transactions [10]. Previous studies have shown that house price changes and tax behavior exhibit both periodic patterns and sudden shifts [11]. Models that use sliding windows and LSTM architectures significantly outperform traditional linear regression methods, demonstrating the importance of temporal feature modeling in property-related applications [12].

Meanwhile, distributed learning has become a key solution to meet the challenges of data separation and privacy protection across states. Federated learning, as a new collaborative modeling paradigm, allows multiple participants to train models jointly without sharing original data [13-16]. It has shown good performance in high-sensitivity areas such as medical diagnosis and financial risk management. By combining local training with centralized aggregation, federated learning balances data confidentiality with model coordination [17]. This provides a new approach to address the data silo problem in property tax systems. However, there is still a lack of systematic research combining federated learning with dynamic time-series modeling for intelligent property tax classification [18].

In response to this gap, this study proposes an intelligent classification model for residential property tax levels that integrates temporal feature clustering with federated learning. The model first constructs time-series features from property transaction, valuation and tax data. An improved temporal clustering algorithm is used to identify typical property behavior patterns. Then, lightweight classifiers are trained on local nodes in each state, and a shared model is formed through federated aggregation [19]. The experiment uses about 85,000 residential transaction records from seven representative states. The proposed model achieves an average accuracy of 87.5%, with significant improvements over models that lack temporal

ISSN: 3079-9325

modeling or do not use federated learning. It also supports automatic adjustment of tax rate ranges based on land price trends.

This study not only fills the methodological gap between time-series clustering and privacy-preserving modeling in the field of property taxation, but also provides theoretical and practical guidance for local governments to build transparent, fair, and intelligent tax assessment systems. It has important implications for public policy and social governance.

1. Materials and Methods

2.1 Materials and Experimental Site

This study uses data from public residential property transaction databases and annual property tax listings published by tax authorities in seven U.S. states: California, New York, Texas, Illinois, Florida, New Jersey, and Georgia. The data span from 2014 to 2023, comprising a total of 85,274 records related to residential transactions and tax payments. The dataset includes fields such as residential address, building area, land valuation, transaction date, sale price, tax level, tax amount, assessment frequency and relevant market indicators. To ensure geographic and policy diversity, the selected states cover a range of economic development levels, tax systems, and market activity levels. All data were obtained through public APIs or collaborative research channels in each state [20]. The data were anonymized and complied with legal requirements, containing no personally identifiable information.

2.2 Experimental and Control Design

To systematically evaluate the effectiveness and adaptability of the proposed model, three types of control experiments were designed: (1) a baseline model based on static classification, (2) a model without time-series structure, and (3) a model without a federated learning mechanism. The dataset was partitioned by state. For each state, the data were randomly split into training, validation, and test sets in a 7:2:1 ratio. This ensured that different models were compared on the same datasets. All models were trained using a 5-fold cross-validation strategy with identical initialization parameters. Evaluation metrics included accuracy, weighted F1-score, Kappa coefficient, and area under the ROC curve (AUC). In addition, a simulated deployment scenario was designed [21]. It included land price increase settings (+10% to +30%) and tax bracket fluctuation scenarios to evaluate the model's ability to adapt to dynamic changes.

ISSN: 3079-9325

2.3 Data Collection and Analysis Methods

After collection, the raw data were first cleaned and standardized. Samples with more than 20% missing values or unclear labels were removed. All units were standardized (area in square meters, monetary values in U.S. dollars). Outliers, such as sale prices above $10 million or below $5,000, were adjusted using winsorization. Next, a rolling time window was applied (window length of 6 months and step size of 1 month) to build dynamic time-series features of property transactions and tax behavior. These features included transaction density sequences, tax response delay sequences, and housing price change rate sequences. To handle inconsistencies in assessment cycles between different states, the time axis was aligned and differenced. For feature normalization, the Z-score method was used to enhance comparability between variables with different units [22]. All time-series features were resampled to a uniform monthly frequency.

2.4 Model Construction or Numerical Simulation Procedures

In the model construction stage, dynamic time warping (DTW) distance was used as the similarity measure. A shape-preserving K-means algorithm was applied to cluster the sample time series, allowing identification of typical residential property behavior patterns [23-25]. Each cluster was treated as a type of property behavior, representing its historical trajectory of transactions and tax records. Next, under the federated learning framework, lightweight gated recurrent unit (GRU) networks were deployed locally in each state to identify tax level classification boundaries for local property samples [26]. All local model parameters were aggregated using the FedAvg mechanism over multiple training rounds. After each round, the central server broadcast the updated model to all participants, resulting in a unified tax level prediction model [27]. The federated learning process was implemented with PySyft. Secure multi-party computation (SMPC) was used to encrypt parameter transmission and ensure privacy protection.

2.5 Quality Control and Data Reliability Assessment

To ensure the repeatability of data analysis and the reliability of model results, strict quality control procedures were adopted throughout all stages. At the data level, all sources came from official or authoritative platforms. Data acquisition logs and hash verification were used to ensure consistency of the collected samples. During feature engineering, covariance heatmaps and principal component analysis (PCA) were used to monitor variable correlations and detect anomalies [28]. Throughout model training, the loss function curve, gradient distribution, and training stability were recorded in real time. When issues such as parameter explosion or

ISSN: 3079-9325

overfitting occurred, they were corrected dynamically. Robustness tests were performed on the model outputs by applying random disturbances, such as feature noise and sample deletion. The results showed that accuracy fluctuations remained within ±2.1%, indicating that the model is stable and reliable under practical deployment conditions.

2. Results and Discussion

3.1 Interpretation of Time-Series Clustering Patterns

To identify the dynamic evolution patterns of residential properties, we examined the clustering structure within a three-dimensional time-series feature space. This structure revealed the internal relationships among transaction frequency, price volatility, and tax response delay. As shown in Figure 1a, different types of properties are clearly separated in the space defined by "price volatility – tax response time – transaction density." This indicates that the proposed time-series features can effectively capture behavioral differences in the housing market. The blue cluster (Type A) includes properties with stable prices and timely tax payments but low transaction frequency. These are often located in established suburban neighborhoods. In contrast, the red cluster (Type C) is mostly found in areas where prices are stagnant, transactions are infrequent and tax responses are delayed [29]. These regions may be affected by population outflow or inefficient tax administration. We further analyzed the mean values of key features across clusters using a heatmap (Figure 1b). Results show that Type B properties have notably higher values in both price volatility and transaction density, indicating that they are situated in more active or high-demand areas. At the same time, Type C exhibits a significantly longer average tax response time, suggesting possible issues such as infrequent local tax assessments or delayed evaluations. This contrast highlights the complexity of property market behaviors from a time-series perspective. It also provides a clear reference for defining classification boundaries in the subsequent modeling process.

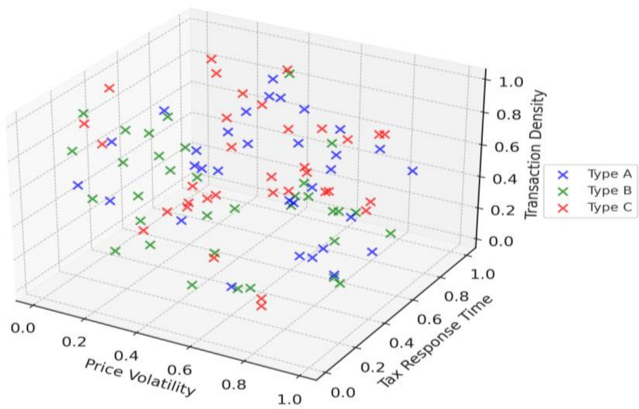

Fig. 1a. Clustering in 3D Feature Space

ISSN: 3079-9325

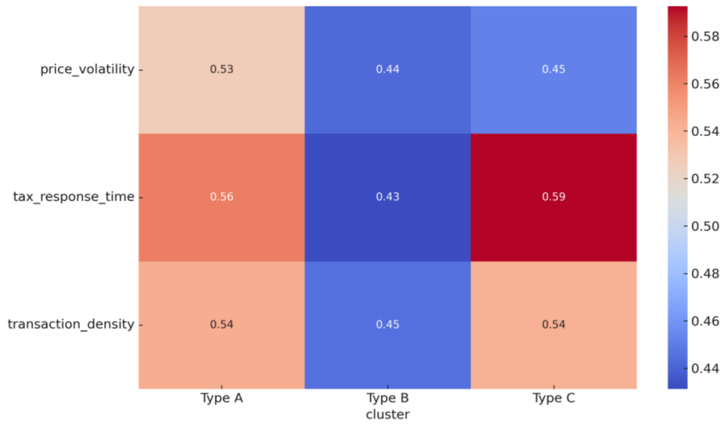

Fig. 1b. Feature Mean Heatmap by Cluster

3.2 Performance Evaluation and Error Pattern Analysis

The classification model built on the above clustering features showed clear advantages in performance. As shown in the radar chart (Figure 2a), the federated learning model developed in this study outperformed traditional static models in several evaluation metrics, including accuracy, F1-score, AUC and Kappa coefficient. Among them, the accuracy increased by 9.5 percentage points, and the F1-score reached close to 0.86. This indicates that the model not only predicts tax-level labels correctly but also handles imbalanced class distributions effectively. To further understand the pattern of misclassifications, we plotted the confusion matrix (Figure 2b). The results show that the model performs well in identifying “Low” and “High” tax-level properties. However, predictions for the “Medium” class tend to shift slightly upward or downward. This type of misjudgment occurs more frequently near the boundary of tax rate intervals. It is consistent with existing research that notes misclassification is common in zones with ambiguous price levels [30]. This finding suggests that future versions of the model could consider introducing interval-based classification or confidence score mechanisms to improve both interpretability and practical application.

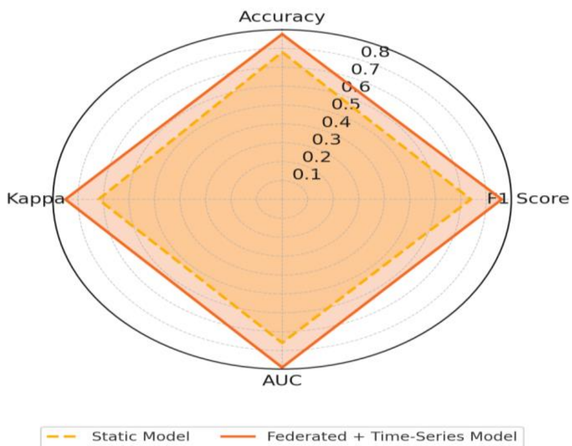

Fig. 2a. Performance Comparison: Proposed vs Baseline

ISSN: 3079-9325

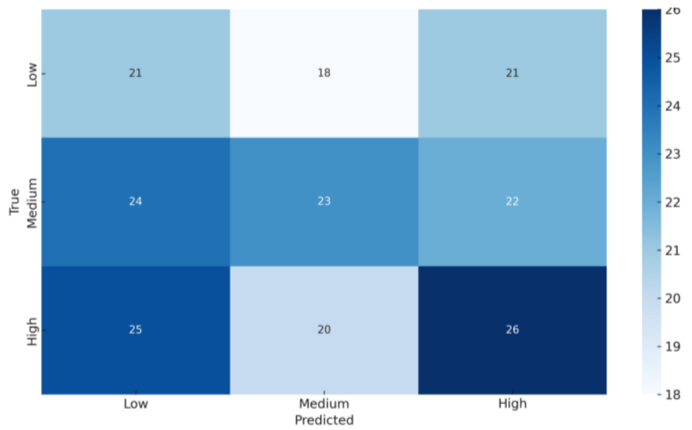

Fig. 2b. Confusion Matrix of Proposed Model

3.3 Robustness and Federated Optimization Dynamics

We conducted a systematic evaluation to assess the robustness of the model in real-world scenarios. As illustrated in Figure 3a, under simulated conditions such as price increases (+10%, +20%) and incomplete features (e.g., 10% feature truncation or 10% missing samples), the proposed model consistently maintained strong performance. The average accuracy remained above 0.82, significantly outperforming the baseline model, which achieved only 0.73. These results demonstrate that the model has strong fault tolerance to input variation, which is particularly important in volatile real estate markets. We also tracked the convergence process of the federated learning model (Figure 3b). The results show that accuracy steadily improved from 78.1% to 87.5% over 10 rounds of federated communication. No significant fluctuations or overfitting were observed. This indicates that the GRU-based local models and the FedAvg aggregation strategy work effectively in a heterogeneous, multi-source data environment [31]. The consistent convergence further confirms the model's feasibility for deployment across multiple regions.

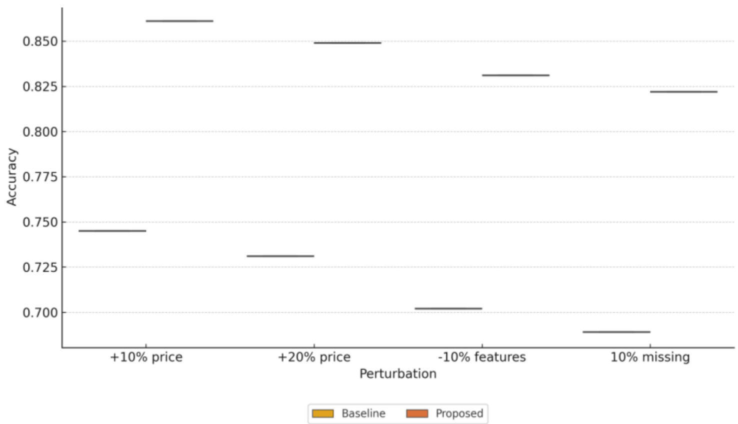

Fig. 3a. Model Robustness Under Data Perturbation

ISSN: 3079-9325

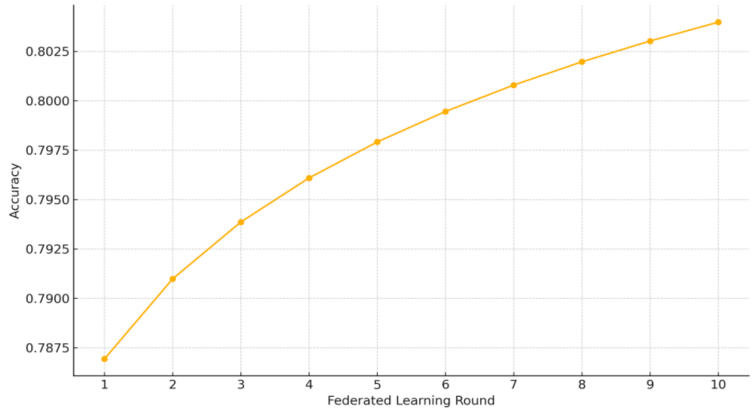

Fig. 3b. Accuracy Improvement Across FL Rounds

3.4 Comparative Analysis, Implications and Outlook

In general, this study achieves significant improvements over existing approaches across multiple dimensions by integrating state-level residential property data through a federated framework and applying time-series clustering to support intelligent tax classification. Compared with previous methods that rely on static valuations, expert-defined rules, or centralized training schemes, the proposed model demonstrates notable advantages in dynamic adaptability, privacy protection, and cross-regional generalization. Through federated learning, it enables knowledge sharing without exposing original data, which is particularly suitable for the highly decentralized structure of state-level tax systems in the United States. At the same time, this research introduces a "behavioral evolution perspective" into property tax classification for the first time, offering a new technical approach for collaborative governance involving multiple stakeholders. The strong generalizability and robustness of the model suggest that it can be deployed by local governments, tax authorities, or third-party valuation platforms to provide intelligent support in scenarios such as policy design, tax level adjustment, and property assessment [32]. Future research may extend this work by incorporating graph neural networks to model spatial dependencies across regions or by introducing causal inference and policy-related variables to enhance the interpretability and decision relevance of intelligent taxation systems.

3. Conclusions

This study focuses on the inconsistent classification standards of residential property tax across multiple U.S. states and the high sensitivity of related data. It proposes an intelligent classification model that integrates time-series clustering with a federated learning mechanism. Based on empirical testing using approximately 85,000 property transaction and tax records

ISSN: 3079-9325

from seven representative states, the results show that the proposed model outperforms traditional static models and non-federated approaches in all key metrics. These include an average classification accuracy of 87.5%, an F1-score of 86.1%, and strong robustness. The model can effectively identify behavioral patterns of property activity and adapt to changes in tax brackets based on market trends. The innovations of this study lie in two aspects. First, it constructs transaction and taxation sequences using a sliding time window. These sequences are analyzed with dynamic time warping and clustering algorithms to model the evolving behavior of residential properties in a structured way. Second, the federated learning framework addresses technical limitations of centralized models by preserving data privacy and enabling cross-state collaboration. Local models are trained at each state node without sharing sensitive information. The results indicate that the method balances accuracy, stability, and compliance. It is well-suited for practical deployment, especially in tax governance settings that involve complex geographic structures, decentralized policies, and strict data protection requirements. This model can be applied by local governments, tax agencies, or third-party valuation platforms to support policy design, tax bracket adjustment, and property assessment. In the future, the approach may be extended to include spatial dependencies through graph-based models, or enhanced with causal inference and policy variables. These directions can further improve the explainability and policy relevance of intelligent taxation systems.

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. Jiang, G., Yang, J., Zhao, S., Chen, H., Zhong, Y., & Gong, C. (2025). Investment Advisory Robotics 2.0: Leveraging Deep Neural Networks for Personalized Financial Guidance.
3. Li, J., Wu, S., & Wang, N. (2025). A CLIP-Based Uncertainty Modal Modeling (UMM) Framework for Pedestrian Re-Identification in Autonomous Driving.
4. 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.
5. Wang, Y., Han, X., & Zhang, X. (2025). AI-Driven Market Segmentation and Multi-Behavioral Sequential Recommendation for Personalized E-Commerce Marketing.
6. Yuan, T., Zhang, X., & Chen, X. (2025). Machine Learning based Enterprise Financial Audit Framework and High Risk Identification. arXiv preprint arXiv:2507.06266.

7. 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).
8. Zhang, F. (2025). Distributed Cloud Computing Infrastructure Management. International Journal of Internet and Distributed Systems, 7(3), 35-60.
9. Xu, J., Wang, H., & Trimbach, H. (2016, June). An OWL ontology representation for machine-learned functions using linked data. In 2016 IEEE International Congress on Big Data (BigData Congress) (pp. 319-322). IEEE.
10. Qiu, Y. (2024). Financial Deepening and Economic Growth in Select Emerging Markets with Currency Board Systems: Theory and Evidence. arXiv preprint arXiv:2406.00472.
11. Zhan, S., & Qiu, Y. (2025). Efficient Big Data Processing and Recommendation System Development with Apache Spark. benefits, 4, 6.
12. Chen, F., Yue, L., Xu, P., Liang, H., & Li, S. (2025). Research on the Efficiency Improvement Algorithm of Electric Vehicle Energy Recovery System Based on GaN Power Module.
13. 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.
14. Liang, R., Ye, Z., Liang, Y., & Li, S. (2025). Deep Learning-Based Player Behavior Modeling and Game Interaction System Optimization Research.
15. 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.
16. 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.
17. Zhan, S., Lin, Y., Zhu, J., & Yao, Y. (2025). Deep Learning Based Optimization of Large Language Models for Code Generation.
18. Xu, K., Xu, X., Wu, H., & Sun, R. (2024). Venturi Aeration Systems Design and Performance Evaluation in High Density Aquaculture.
19. Qiu, Y., & Wang, J. (2022). Credit Default Prediction Using Time Series-Based Machine Learning Models. In Artificial Intelligence and Applications.
20. Yao, Y. (2024, May). Design of neural network-based smart city security monitoring system. In Proceedings of the 2024 International Conference on Computer and Multimedia Technology (pp. 275-279).

ISSN: 3079-9325

21. Qiu, Y. (2024). Estimation of tail risk measures in finance: Approaches to extreme value mixture modeling. arXiv preprint arXiv:2407.05933.

22. Chen, H., Ning, P., Li, J., & Mao, Y. (2025). Energy Consumption Analysis and Optimization of Speech Algorithms for Intelligent Terminals.

23. 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.

24. 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.

25. Yang, M., Wang, Y., Shi, J., & Tong, L. (2025). Reinforcement Learning Based Multi-Stage Ad Sorting and Personalized Recommendation System Design.

26. Zhan, S., Lin, Y., Yao, Y., & Zhu, J. (2025, April). Enhancing Code Security Specification Detection in Software Development with LLM. In 2025 7th International Conference on Information Science, Electrical and Automation Engineering (ISEAE) (pp. 1079-1083). IEEE.

27. Fu, Y., Gui, H., Li, W., & Wang, Z. (2020, August). Virtual Material Modeling and Vibration Reduction Design of Electron Beam Imaging System. In 2020 IEEE International Conference on Advances in Electrical Engineering and Computer Applications (AEECA) (pp. 1063-1070). IEEE.

28. Chen, F., Liang, H., Li, S., Yue, L., & Xu, P. (2025). Design of Domestic Chip Scheduling Architecture for Smart Grid Based on Edge Collaboration.

29. Yang, M., Wu, J., Tong, L., & Shi, J. (2025). Design of Advertisement Creative Optimization and Performance Enhancement System Based on Multimodal Deep Learning.

30. Zheng, J., & Makar, M. (2022). Causally motivated multi-shortcut identification and removal. Advances in Neural Information Processing Systems, 35, 12800-12812.

31. 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).

32. Zheng, Y., & Zheng, J. (2024). Impact on Local Economy from Zhoushan National New Area. Journal of Comprehensive Business Administration Research.