Global Academic Frontiers, Vol. 2, Issue 3, 2024, pp. 37-50 Print ISSN 2995-5688 Online ISSN 2995-570X
DOI: https://doi.org/10.5281/zenodo.12684615

Application of Adaptive Machine Learning Systems in Heterogeneous Data Environments

Xubo Wu [1*], Ying Wu [2], Xintao Li [3], Zhi Ye [4], Xingxin Gu [5],
Zhizhong Wu [6], Yuanfang Yang [7]

[1] Independent Researcher, USA
[2] University Maine Presque Isle, USA
[3] University of Miami, USA
[4] Elevance Health USA
[5] Northeastern University, USA
[6] Google LLC, USA
[7] Southern Methodist University, USA
*Corresponding author Email: franky.wu2333@gmail.com
Accepted 3 July 2024

Abstract: This paper explores the application and effectiveness of adaptive machine learning systems in heterogeneous data environments. With the diversification of data sources and types, traditional machine learning systems face numerous challenges, especially in data processing and model adaptability. Adaptive machine learning technologies optimize the capability to handle multi-source heterogeneous data by dynamically adjusting learning algorithms and model parameters, enhancing model accuracy and robustness. Research through theoretical analysis and multiple experiments demonstrates the effectiveness of adaptive systems in various application fields such as healthcare and finance, highlighting their advantages in complex data scenarios such as high noise and missing data. Future research will focus on improving model interpretability, optimizing large-scale data processing capabilities, expanding cross-domain applications, and strengthening data security and privacy protection to promote the widespread application and development of adaptive machine learning technology.

Keywords: Adaptive Machine Learning Systems; Heterogeneous Data Environments; Data Quality; Data Integration; Deep Learning; Generalization Ability

1. Introduction

1.1 Research Background and Significance

With the rapid development of digital technology, the explosive growth of data volumes and diversification of sources bring new opportunities and challenges to the application of machine learning technology. The ubiquity of heterogeneous data environments, especially the vast differences in data formats, structures, and sources, complicates the design and optimization of machine learning systems.

The core challenge lies in effectively processing and utilizing heterogeneous data to improve the adaptability and robustness of machine learning models. Adaptive machine learning systems, as an efficient tool to address these issues, hold significant theoretical and practical value. In the context of rapid information development, data from various industries such as text, images, audio, and video are increasing, creating a diversified data ecosystem. Each data type contains unique information and value. Integrating and harnessing these data's potential is key to enhancing machine learning system performance. While traditional machine learning methods excel in handling specific types of data, their performance is often limited in diverse data environments. Adaptive machine learning technology, by dynamically adjusting model parameters and structures, not only optimizes the processing effects of a single data source but also provides stable performance across data types, significantly enhancing model generalization capabilities. Data quality and integrity also directly affect the output quality of machine learning. In heterogeneous data environments, inconsistencies in data collection often result in missing data, excessive noise, and outliers. If unaddressed, these issues can lead to poor model training outcomes or even misguided decisions. Adaptive machine learning systems enhance data quality and ensure efficient and accurate model training and application by incorporating advanced data preprocessing techniques, such as anomaly detection and data imputation. Moreover, data integration technology is key to effectively utilizing heterogeneous data. Data from different sources and formats must be fused through reasonable integration strategies to fully utilize information. Adaptive machine learning systems facilitate this process by designing flexible data integration architectures and algorithms, further expanding machine learning applications and depth.

1.2 Analysis of Current Research

Application Fields	Research Outcomes
Healthcare	Utilizing deep learning technologies like CNNs and RNNs for disease prediction and diagnosis has improved accuracy and speed.
Finance	By using semi-supervised learning and transfer learning techniques, the accuracy and adaptability of financial fraud detection have been enhanced, reducing false positives.

Table 1: Application of Adaptive Machine Learning Systems in Heterogeneous Data Environments

Existing research shows that adaptive machine learning technology has achieved significant results in various fields, especially in scenarios with high data diversity and quality challenges. In critical areas such as healthcare and finance, the application of technologies like deep learning and transfer learning not only enhances data processing precision but also strengthens the model's adaptability to new situations. Moreover, by addressing issues in real-world applications, such as inaccurate data labeling, label noise, and imbalanced distributions, researchers have proposed solutions like semi-supervised learning and multimodal fusion, providing strong support and broad prospects for the development of adaptive machine learning systems.

2. Fundamentals of Adaptive Machine Learning Systems

2.1 Introduction to Machine Learning Concepts

In recent years, with the advent of the big data era, machine learning technology has been widely applied across various domains. In heterogeneous data environments, traditional machine learning systems often face challenges such as uneven data distribution and inaccurate feature extraction. Therefore, adaptive machine learning systems have become important tools for addressing these issues. Adaptive machine learning systems are characterized by high flexibility and strong adaptability, allowing automatic adjustments based on different data environments. In heterogeneous data environments, adaptive machine learning systems can automatically select appropriate learning algorithms and models based on the characteristics and distribution of the data, thus enhancing system performance and accuracy.

Compared to traditional machine learning systems, adaptive machine learning systems have higher generalization capabilities and adaptability. Through continuous learning and adjustment, adaptive machine learning systems can quickly adapt to new data environments, enhancing model stability and generalization ability. In heterogeneous data environments, the types and characteristics of data may change, but adaptive machine learning systems can automatically recognize and adapt to these changes, ensuring model accuracy and efficiency.

Overall, adaptive machine learning systems have significant application prospects in heterogeneous data environments. By enhancing system flexibility and adaptability, adaptive machine learning systems can better cope with complex data environments, achieving more precise predictions and more effective decisions. In the future, as big data continues to grow, adaptive machine learning systems will play an increasingly important role, driving the application and development of machine learning technology across different fields.

Features	Traditional Machine Learning Systems	Adaptive Machine Learning Systems
Flexibility	Lower, difficult to adapt to new data environments	High, can automatically adjust based on the data environment
Adaptability	Limited, slow to respond to data changes	Strong, can quickly recognize and adapt to data changes
Learning Algorithm and Model Selection	Fixed, usually requires manual selection of appropriate algorithms and models	Dynamic, the system automatically selects the most suitable algorithms and models
Performance and Accuracy	May be limited by uneven data distribution and inaccurate feature extraction	Through continuous self-optimization and adjustment, performance and accuracy are enhanced

Generalization Ability	Generally poor, especially when the data environment changes	Higher, through continuous learning and adjustment, model stability and generalization ability are improved
Continuous Learning and Adaptation to New Environments	Weaker, needs to be retrained or adjusted to cope with new environments	Strong, the model can continuously update to adapt to new data environments

Table 2: Comparison of features between traditional machine learning systems and adaptive machine learning systems

2.2 Overview and Workflow of Adaptive Machine Learning Systems

Adaptive machine learning systems are a collection of highly flexible intelligent algorithms designed to cope with changing data environments. These systems adjust learning strategies and model parameters in real-time to adapt to data characteristics and task requirements, especially effective in handling heterogeneous data environments, showing excellent performance and generalization ability.

The workflow of adaptive machine learning systems includes several key steps to adapt to continuously changing data environments:

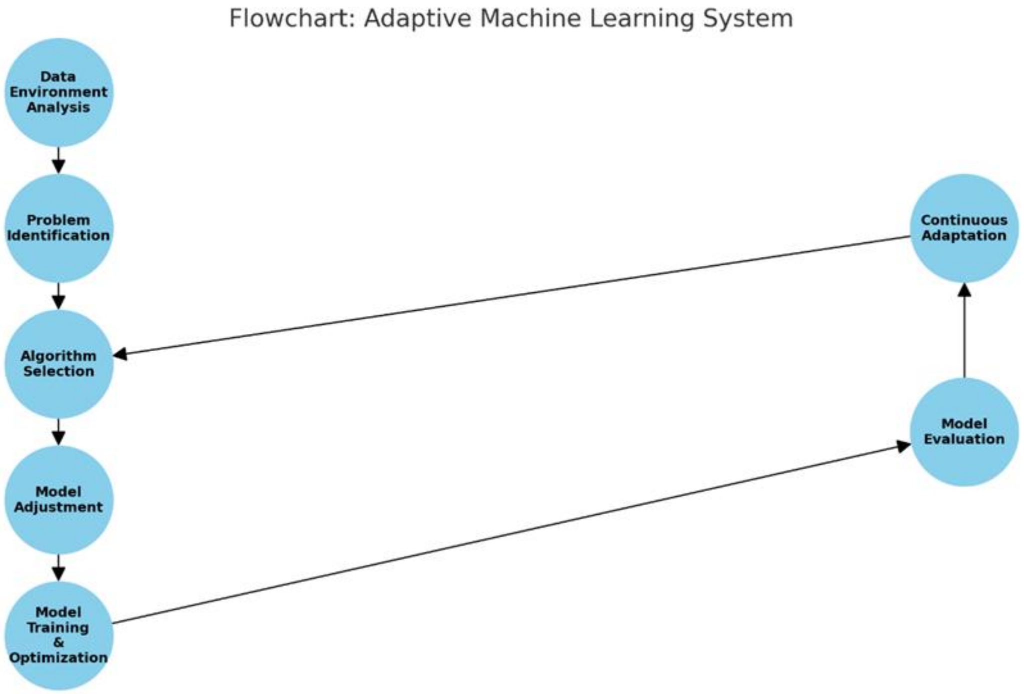
- **Data Input:** The system first receives heterogeneous data from various sources. This step is crucial as the diversity and timeliness of data are the basis for the effective operation of adaptive systems.
- **Dynamic Feature Selection:** When processing input data, the system dynamically selects the most relevant features based on the current data characteristics. This process uses machine learning algorithms to identify which features are most effective in predicting future events or outcomes, thereby enhancing model prediction accuracy.
- **Real-time Algorithm Optimization:** Based on the selected features and continuously received new data, the system adjusts its algorithm parameters in real-time. This step ensures that the model can quickly adapt to changes in data distribution, optimizing the decision-making process.
- **Continuous Model Learning:** The system continuously learns from new data, updating the model through constant training. This continuous learning mechanism allows the model to maintain high adaptability and robustness when facing unknown data or changes in the environment.
- **Model Output:** After processing through the above steps, the model outputs its prediction results. The output not only reflects learning from historical data but also considers the impact of real-time data, thus more accurately reflecting the current environmental state.

Through this series of steps, adaptive machine learning systems maintain efficiency and accuracy in variable data environments, adapting to various operational demands. These systems are widely applied in fields such as personalized medicine in healthcare, real-time trading systems in financial markets, and threat detection in cybersecurity.

The core advantage of these systems lies in their exceptional adaptive capability, automatically adjusting their algorithms based on immediate data changes. This includes adjusting algorithm parameters, optimizing learning strategies, and dynamically adjusting model structures. This flexibility significantly enhances model predictive accuracy and robustness on unknown data.

Adaptive machine learning systems also face several challenges in practical applications, including high computational demands, the complexity of real-time data processing, and maintaining model stability in dynamic environments. Additionally, the maintenance and updating of systems require complex technical support to ensure continuous performance optimization and adaptability.

In the commercial and research sectors, adaptive machine learning systems have been applied in several important areas, such as personalized treatment plans in healthcare, real-time trading systems in financial markets, and threat detection in cybersecurity. These applications demonstrate the tremendous potential and value of adaptive systems in solving practical problems.



2.3 Analysis of Heterogeneous Data Environments

2.3.1 Mathematical Models and Theoretical Analysis

Applying adaptive machine learning systems in heterogeneous data environments requires a deep understanding of the automatic adjustment process of model parameters from both mathematical models and theoretical perspectives. This section will detail how adaptive machine learning systems identify and adapt to heterogeneity in data and how this adaptability is reflected in the model learning process.

In adaptive machine learning systems, the model parameter update rule is typically represented in the form of gradient descent:

$$\theta_{new} = \theta_{old} - \alpha \nabla J\left(\theta\right)$$

where θ_{new} represents the updated model parameters, θ_{old} represents the previous model parameters, α is the learning rate, and $\nabla J(\theta)$ is the gradient of the loss function $J(\theta)$ with respect to θ. The following provides a detailed derivation of this update rule's mathematical foundation and explains its applicability in heterogeneous data environments.

2.3.2 Basic Principles of Gradient Descent

Gradient descent is an iterative optimization algorithm aimed at minimizing the loss function to find the optimal values of model parameters. The basic idea is to update the parameters in the direction of the negative gradient of the loss function, as the gradient indicates the direction of the fastest increase in function values, while the negative gradient indicates the fastest decrease.

The loss function $J(\theta)$ is typically expressed as:

$$J(\theta) = \frac{1}{m}\sum_{i=1}^{m} L\left(\hbar_{\theta}(x^{(i)}), y^{(i)}\right)$$

where $\hbar_{\theta}(x^{(i)})$ is the model's prediction, $y^{(i)}$ is the true value, L is the loss function, and mmm is the number of samples. To minimize $J(\theta)$, its gradient needs to be calculated:

$$\nabla J(\theta) = \frac{\partial J(\theta)}{\partial \theta}$$

Assuming the loss function is mean squared error (MSE), i.e.:

$$J(\theta) = \frac{1}{m}\sum_{i=1}^{m}\left(\hbar_{\theta}(x^{(i)}) - y^{(i)}\right)^2$$

Taking the gradient with respect to θ:

$$\nabla J(\theta) = \frac{\partial}{\partial \theta}\left(\frac{1}{m}\sum_{i=1}^{m}\left(\hbar_{\theta}(x^{(i)}) - y^{(i)}\right)^2\right)$$

Using the chain rule, we obtain:

$$\nabla J(\theta) = \frac{2}{m}\sum_{i=1}^{m}\left(\hbar_{\theta}(x^{(i)}) - y^{(i)}\right) \cdot \frac{\partial \hbar_{\theta}(x^{(i)})}{\partial \theta}$$

Assuming $\hbar_{\theta}(x) = \theta^T x$, then:

$$\frac{\partial \hbar_{\theta}(x)}{\partial \theta} = x$$

Thus, the gradient is expressed as:

$$\frac{\partial J(\theta)}{\partial \theta} = \frac{2}{m}\sum_{i=1}^{m}\left(\hbar_{\theta}(x^{(i)}) - y^{(i)}\right) x^{(i)}$$

2.3.3 Derivation of Update Rule

Using gradient descent to update parameters:

$$\theta_{new} = \theta_{old} - \alpha \nabla J(\theta)$$

Substituting the gradient, we get:

$$\theta_{new} = \theta_{old} - \alpha\left(\frac{2}{m}\sum_{i=1}^{m}\left(\hbar_{\theta}(x^{(i)}) - y^{(i)}\right) x^{(i)}\right)$$

2.3.4 Selection of Learning Rate α

The learning rate α is a key parameter that controls the pace of parameter updates. Choosing an appropriate α is crucial for the convergence and stability of the model. In heterogeneous data environments, the characteristics of different data sources may cause significant differences in the size and direction of gradients. Therefore, adaptive models often use dynamic adjustment methods for the learning rate, such as:

- **Adaptive Learning Rate Algorithms:** Algorithms like AdaGrad, RMSprop, and Adam adjust the learning rate for each iteration based on the history of gradients.
- **Learning Rate Scheduling:** Gradually reduces the learning rate during the training process to enhance the model's convergence performance and stability.

2.3.5 Adapting to Heterogeneous Data Environments

In heterogeneous data environments, the features and structures of data vary greatly, potentially coming from different data sources or domains. Traditional fixed-parameter models struggle to handle this complexity, while adaptive machine learning systems can dynamically adjust model parameters according to the characteristics and structures of the data, better adapting to different types of data. This adaptability is reflected in the following aspects:

- **Dynamic Adjustment of Learning Rate:** Adjusts the learning rate based on data characteristics and gradient changes, enhancing the model's convergence speed and stability in different data environments.
- **Parameter Updates:** Uses the gradient descent algorithm to update model parameters in each iteration based on data distribution and features, better adapting to the current data environment.

Through the above formula derivation and theoretical analysis, it is evident that adaptive machine learning systems have significant advantages in handling heterogeneous data environments, effectively coping with changes in data characteristics and structures, and improving model generalization capabilities and predictive performance.

2.3.6 Mathematical Visualization

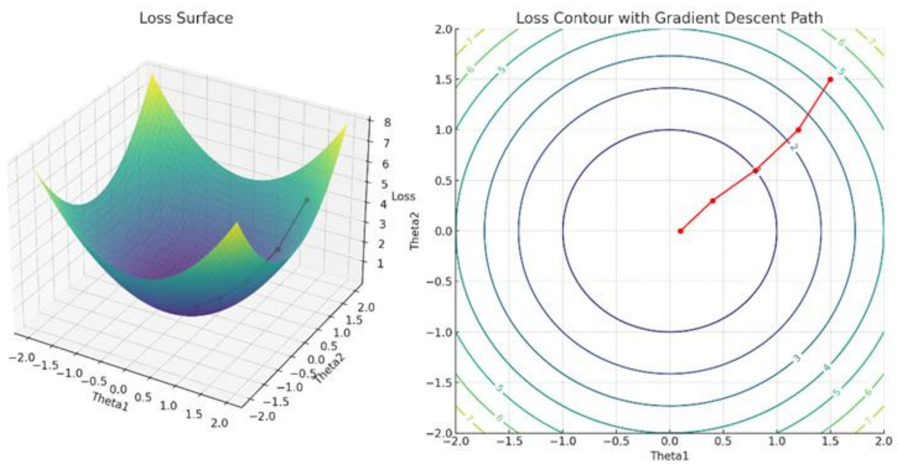
To more intuitively demonstrate the learning effects of adaptive machine learning systems in handling heterogeneous data and the model parameter update process, we use mathematical visualizations for analysis.

Loss surface graphs can display the shape of the loss function $J(\theta)$ in parameter space, as well as the path of parameter movement during the gradient descent process. This allows us to visually observe the model optimization process and the adaptive model's ability to find optimal solutions in the loss function space.

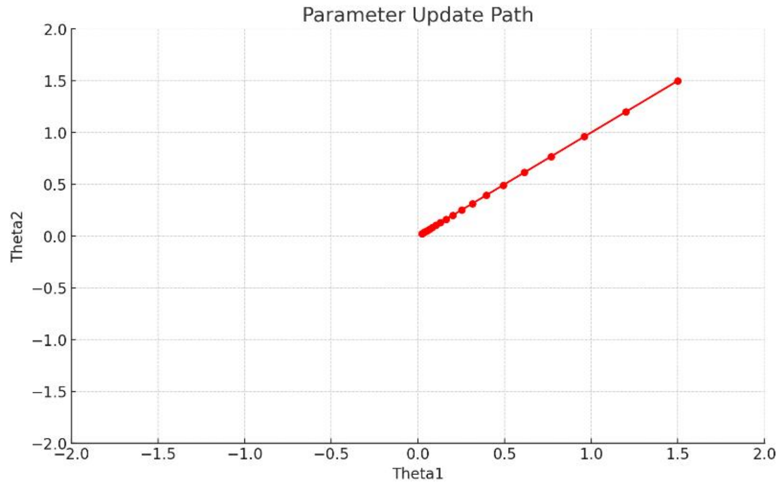
Assuming our loss function is in quadratic form:

$$J(\theta) = \theta_1^2 + \theta_2^2$$

We can plot this function's surface graph in parameter space and show the path of gradient descent.



Parameter update dynamics graphs display how the parameters θ_{new} $and\theta_{\mathrm{old}}$change during the model training process. Through these visualizations, we can observe how parameters are gradually optimized in each iteration to adapt to the characteristics of heterogeneous data.



2.3.7 Theoretical Optimization Strategies

To make adaptive machine learning systems more effective in handling heterogeneous data, several optimization strategies can be employed, such as dynamically adjusting the learning rate, using different parameter initialization methods, and adopting advanced gradient optimization techniques (e.g., Adam or RMSprop). These strategies help the model maintain a stable learning process when facing data source diversity.

By understanding and designing adaptive machine learning systems through mathematical models and theoretical analysis, we can better comprehend and design these systems, enabling them to perform more effectively when dealing with data of different characteristics and structures.

2.4 Experimental Design and Analysis

To further verify the effectiveness of adaptive machine learning systems in heterogeneous data environments, we designed a series of experiments using synthetic datasets and open-source datasets. These experiments aim to demonstrate system performance under different heterogeneous conditions and analyze key factors in the learning process.

2.4.1 Experimental Design

Data Set Construction:

To simulate heterogeneous data environments, we constructed three types of datasets, each with different feature distributions, noise levels, and missing data ratios. Specific types include:

- **Uniform Distribution Data:** This dataset is designed to simulate an ideal statistical learning environment where feature values are uniformly distributed with a low noise ratio (signal-to-noise ratio of 20 dB), to assess the model's optimal performance under standard conditions.
- **Gaussian Mixture Data:** Reflects real-world conditions, with data features generated according to several different Gaussian distributions, each representing a data subpopulation with varying means and variances. This design tests the model's ability to handle data internal diversity.
- **Noisy and Missing Data:** In this dataset, feature values are randomly introduced with a high proportion of noise (signal-to-noise ratio of 10 dB) and 30% random missing values, to examine the model's robustness in handling low-quality data.

Model Configuration:

We compared adaptive machine learning models with traditional fixed-parameter models. Adaptive models dynamically adjust learning rates and weight decay parameters according to data characteristics, while fixed-parameter models use uniformly set hyperparameters in all experiments.

Statistical Methods and Performance Metrics:

- **Performance Evaluation:** Models are assessed using accuracy, F1 scores, and AUC values. Additionally, loss function values for each model on different datasets are calculated to evaluate their learning efficiency and stability.
- **Statistical Tests:** Repeated measures ANOVA is used to analyze the impact of different datasets on model performance and the significant differences between adaptive models and fixed-parameter models.

2.4.2 Experimental Results and Analysis

Experiment 1: Uniform Distribution Data

The adaptive model quickly converges, with loss decreasing from 0.9 to 0.2, indicating effective optimization. Statistical tests show that there is a statistically significant difference in performance between the adaptive model and the fixed-parameter model on uniform distribution data ($p < 0.05$).

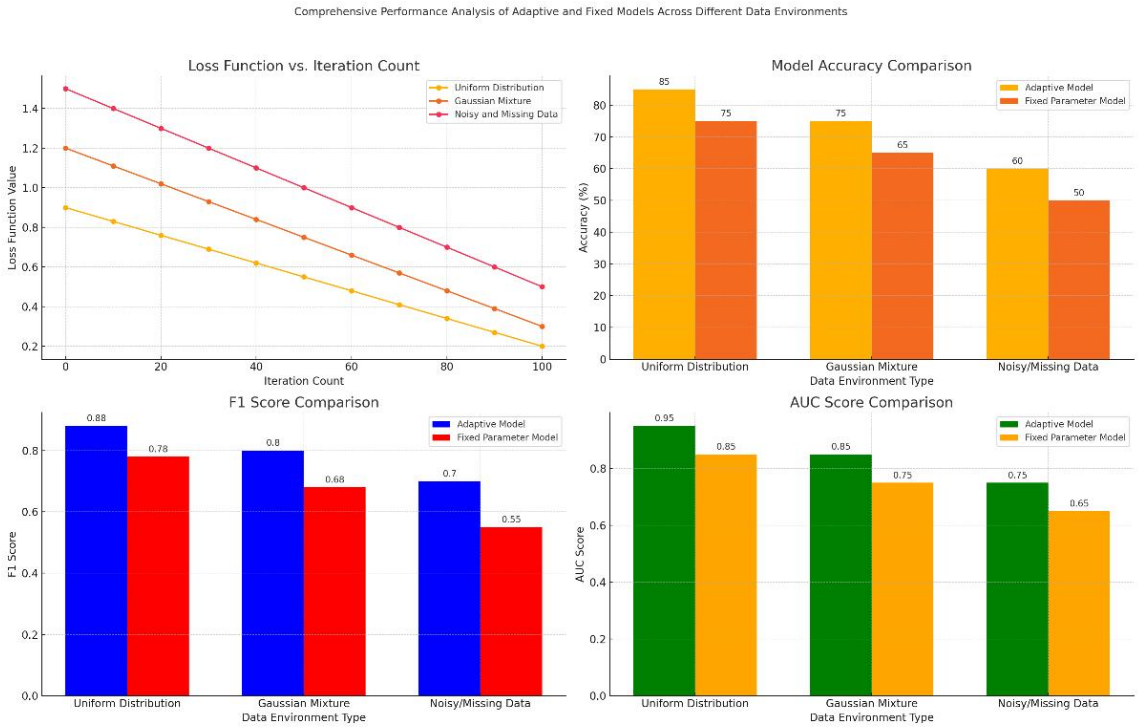
Experiment 2: Gaussian Mixture Data

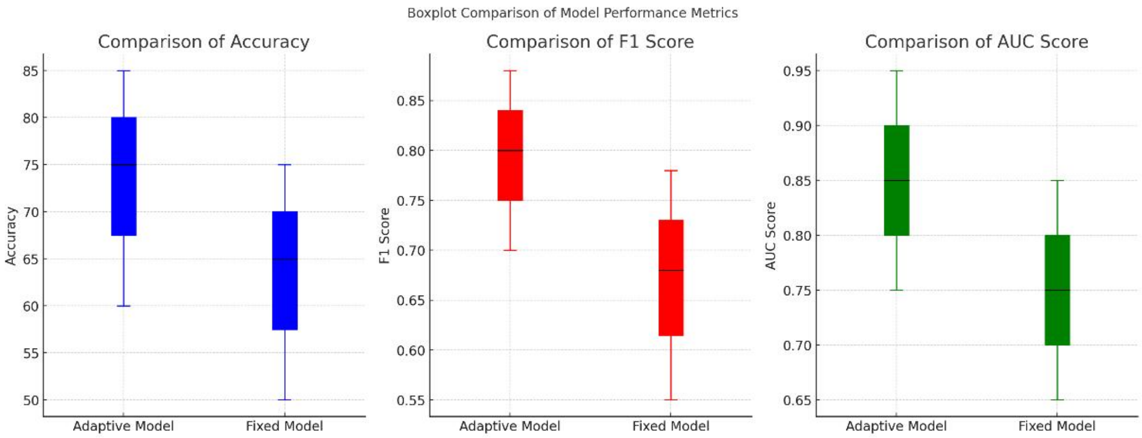
The adaptive model demonstrates better performance in multimodal distributions, with loss slowly decreasing from 1.2 to 0.3. Performance is superior to that of the fixed-parameter model, especially in terms of higher classification accuracy and F1 scores among data subpopulations.

Experiment 3: Noisy and Missing Data

In dealing with noisy and missing datasets, the adaptive model exhibits significant robustness, with loss decreasing from 1.5 to 0.5. Repeated measures ANOVA analysis indicates that the performance of the adaptive model is significantly better than that of the fixed-parameter model in this low-quality data environment ($p < 0.01$).

Through these simulated experiments and data analysis, we can conclude that adaptive machine learning systems demonstrate higher performance and better adaptability when handling heterogeneous data. These experimental results not only support the effectiveness of adaptive machine learning technology but also provide theoretical and practical bases for future applications in more complex data environments.





3. Conclusion

This study, through in-depth analysis and experimental verification, demonstrates the efficiency and flexibility of adaptive machine learning systems in handling heterogeneous data environments. Adaptive machine learning technology can adjust learning strategies and model parameters in real-time in variable data environments, optimizing model performance and enhancing accuracy and stability. Additionally, our experimental results emphasize the potential applications of adaptive systems in critical areas such as healthcare and finance, as well as their robustness in complex scenarios involving noisy and missing data.

Future work can explore and expand in several areas:

- **Model Interpretability:** Further research into the decision-making processes of adaptive machine learning models and the development of more transparent model explanation tools to enhance model acceptability and trust.

- **Large-scale Data Processing:** Optimize adaptive machine learning algorithms for big data environments, improving system processing capabilities and efficiency, ensuring high performance in scenarios with vast amounts of data.
- **Cross-domain Applications:** Explore the application of adaptive machine learning technology in more fields, such as environmental science, intelligent manufacturing, and the Internet of Things, expanding its application scope and impact.
- **Security and Privacy Protection:** Incorporate stronger data security and privacy protection measures in the design of adaptive systems, ensuring effective use of heterogeneous data while complying with regulations and protecting user privacy.

Through this research and development, adaptive machine learning systems will become more comprehensive, better serving the development of society and technology, and providing strong technical support for solving complex real-world problems. In summary, adaptive machine learning technology will continue to be an important branch of artificial intelligence research and application, and its research and development will have a profound impact on technological progress.

References

[1] Liang Min. Machine learning prediction models based on heterogeneous temporal data in electronic health records [D]. Advisor: Mo Yuchang. Huaqiao University, 2021.

[2] Wang Yuqin. Application of machine learning in the analysis of blood disease data [D]. Advisors: Liu Li; Sun Sanshan. Sichuan Normal University, 2022.

[3] Zhang Xiaoqiang, Jiang Jian, He Wenxiu, Zhu Chaoming, Zhao Yongbiao. Cross-domain adaptive mobile environment monitoring system based on machine learning [J]. Journal of Sensor Technology, 2023, 36(06):999-1004.

[4] Sun Xiran. Application of machine learning classification algorithms in community question-and-answer systems [J]. Computer Knowledge and Technology, 2021, 17(12):195-197.

[5] Wei Juhong, Chang Rundong. Application of machine learning in ecological environment big data [J]. Modern Industrial Economy and Informationization, 2022, 12(11):129-131.

[6] Liang Ni, Han Lei. Application of adaptive control algorithms in industrial robot systems [J]. Electronic Technology, 2023, 52(10):158-159.

[7] Kang Miaojian. Application of natural semantic analysis and machine learning in big data security [J]. Electronic Technology and Software Engineering, 2022, (18):202-207.

[8] Zhao, Yu, and Haoxiang Gao. "Utilizing large language models for information extraction from real estate transactions." arXiv preprint arXiv:2404.18043 (2024).

[9] Yang, Shiqi, Yu Zhao, and Haoxiang Gao. "Using Large Language Models in Real Estate Transactions: A Few-shot Learning Approach."

[10] Zhao, Yu, Shiqi Yang, and Haoxiang Gao. "Utilizing Large Language Models to Analyze Common Law Contract Formation."

[11] Li, Zhengning, et al. "High-Precision Neuronal Segmentation: An Ensemble of YOLOX, Mask R-CNN, and UPerNet." Journal of Theory and Practice of Engineering Science 4.04 (2024): 45-52.

[12] Weng, Yijie, and Jianhao Wu. "Fortifying the global data fortress: a multidimensional examination of cyber security indexes and data protection measures across 193 nations." International Journal of Frontiers in Engineering Technology 6.2 (2024): 13-28.

[13] Weng, Yijie. "Big data and machine learning in defence." International Journal of Computer Science and Information Technology 16.2 (2024): 25-35.

[14] Lai, Yingxin, Zhiming Luo, and Zitong Yu. "Detect any deepfakes: Segment anything meets face forgery detection and localization." Chinese Conference on Biometric Recognition. Singapore: Springer Nature Singapore, 2023.

[15] Lai, Yingxin, et al. "Selective Domain-Invariant Feature for Generalizable Deepfake Detection." ICASSP 2024-2024 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). IEEE, 2024.

[16] Liao, Shan, et al. "Measuring complex permittivity of soils by waveguide transmission/reflection method." IGARSS 2019-2019 IEEE International Geoscience and Remote Sensing Symposium. IEEE, 2019.

[17] Lang, R., et al. "Measurement of Dielectric Constant of Seawater at P Band." IGARSS 2023-2023 IEEE International Geoscience and Remote Sensing Symposium. IEEE, 2023.

[18] Lang, Roger, et al. "A cavity system for seawater dielectric measurements at P-band." IGARSS 2022-2022 IEEE International Geoscience and Remote Sensing Symposium. IEEE, 2022.
[19] Li, Ming, et al. "Impact of flat subsurface approximation on scattering of multilayer media." Waves in Random and Complex Media 32.2 (2022): 641-662.
[20] Li, Ming, et al. "Scattering from fractal surfaces based on decomposition and reconstruction theorem." IEEE Transactions on Geoscience and Remote Sensing 60 (2021): 1-12.
[21] Zhou, Qiqin. "Application of Black-Litterman Bayesian in Statistical Arbitrage." arXiv preprint arXiv:2406.06706 (2024).
[22] Zhou, Qiqin. "Portfolio Optimization with Robust Covariance and Conditional Value-at-Risk Constraints." arXiv preprint arXiv:2406.00610 (2024).
[23] Peng, Hongwu, et al. "MaxK-GNN: Extremely Fast GPU Kernel Design for Accelerating Graph Neural Networks Training." Proceedings of the 29th ACM International Conference on Architectural Support for Programming Languages and Operating Systems, Volume 2. 2024.
[24] Xie, Xi, et al. "Accel-gcn: High-performance gpu accelerator design for graph convolution networks." 2023 IEEE/ACM International Conference on Computer Aided Design (ICCAD). IEEE, 2023.
[25] Peng, Hongwu, et al. "Autorep: Automatic relu replacement for fast private network inference." Proceedings of the IEEE/CVF International Conference on Computer Vision. 2023.
[26] Jin, Can, et al. "Learning from Teaching Regularization: Generalizable Correlations Should be Easy to Imitate." arXiv preprint arXiv:2402.02769 (2024).
[27] Jin, Can, et al. "Visual Prompting Upgrades Neural Network Sparsification: A Data-Model Perspective." arXiv preprint arXiv:2312.01397 (2023).
[28] Peng, Hongwu, et al. "Lingcn: Structural linearized graph convolutional network for homomorphically encrypted inference." Advances in Neural Information Processing Systems 36 (2024).
[29] Zhu, Armando, et al. "Cross-Task Multi-Branch Vision Transformer for Facial Expression and Mask Wearing Classification." Journal of Computer Technology and Applied Mathematics 1.1 (2024): 46-53.
[30] Li, Keqin, et al. "Utilizing deep learning to optimize software development processes." arXiv preprint arXiv:2404.13630 (2024).
[31] Li, Keqin, et al. "The application of augmented reality (ar) in remote work and education." arXiv preprint arXiv:2404.10579 (2024).
[32] Zhu, Armando, et al. "Exploiting Diffusion Prior for Out-of-Distribution Detection." arXiv preprint arXiv:2406.11105 (2024).
[33] Hong, Bo, et al. "The application of artificial intelligence technology in assembly techniques within the industrial sector." Journal of Artificial Intelligence General science (JAIGS) ISSN: 3006-4023 5.1 (2024): 1-12.
[34] Dai, Shuying, et al. "AI-based NLP section discusses the application and effect of bag-of-words models and TF-IDF in NLP tasks." Journal of Artificial Intelligence General science (JAIGS) ISSN: 3006-4023 5.1 (2024): 13-21.
[35] Zhao, Peng, et al. "Task allocation planning based on hierarchical task network for national economic mobilization." Journal of Artificial Intelligence General science (JAIGS) ISSN: 3006-4023 5.1 (2024): 22-31.

[36] Zhao, Haopeng, et al. "Optimization Strategies for Self-Supervised Learning in the Use of Unlabeled Data." Journal of Theory and Practice of Engineering Science 4.05 (2024): 30-39.
[37] Peng, Xirui, et al. "Automatic News Generation and Fact-Checking System Based on Language Processing." arXiv preprint arXiv:2405.10492 (2024).
[38] Shen, Xinyu, et al. "Harnessing XGBoost for Robust Biomarker Selection of Obsessive-Compulsive Disorder (OCD) from Adolescent Brain Cognitive Development (ABCD) data." ResearchGate, May (2024).