Scholar Publishing Group
Machine Learning Theory and Practice
https://doi.org/10.38007/ML.2025.050115
ISSN 2790-0983 Vol. 5, Issue 1: 141-148

Research on Integration and Optimization Strategies of Cross-platform Machine Learning Services

Yixian Jiang

Carnegie Mellon University, Information Networking Institute, Pittsburgh, PA, 15213, USA

Keywords: Cross platform; Machine learning; integrate; Optimization strategy; Service Framework

Abstract: In diverse operating system environments, cross platform machine learning techniques are increasingly becoming the core means of enhancing data processing capabilities and model performance. This study aims to explore the integration and optimization path of cross platform machine learning technology, analyze the popular cross platform service architectures and platforms, and propose a series of integration solutions related to data interaction, model joint training, resource optimization configuration, and interface unification. At the same time, this article also explores in depth optimization measures such as improving the accuracy of machine learning algorithms, accelerating model training speed, rational resource allocation, and enhancing service robustness. By adopting these integration and optimization strategies, the performance indicators and application effectiveness of cross platform machine learning technology will be enhanced, providing theoretical basis and technical guidance for engineering applications in related fields.

1. Introduction

As a cutting-edge technology trend, cross platform machine learning technology brings together the computing power of various platforms, improving the training efficiency and application effectiveness of machine learning algorithms. With the rapid advancement of artificial intelligence and massive data, machine learning algorithms have become more advanced, relying on the collaborative work of numerous computing platforms to handle large-scale datasets and complex computing tasks. However, the differences between platforms pose significant challenges to the integration and optimization of cross platform machine learning. The current service architecture has exposed many shortcomings in data processing, model joint training, resource allocation, and interface unification, which limit the popularity and further effectiveness of this technology. Exploring the integration path and optimization methods of cross platform machine learning services in depth can enhance algorithm performance and bring more efficient and flexible solution strategies for practical applications.

2. Overview of Cross Platform Machine Learning Services

2.1. Definition of Cross Platform Machine Learning Services

Cross platform fusion of machine learning technology refers to the integration of the resource advantages of various computing platforms to complete various tasks of machine learning (including data preprocessing, feature filtering, model cultivation and inference, etc.) and collaborate and operate efficiently between different systems. The different platforms referred to here cover different computing structures, operating systems, or cloud services, each with different hardware configurations, system software, and network conditions. This cross platform service integrates various heterogeneous platforms through data exchange and sharing, unified interface standards, and distributed computing architecture, building a complex system that operates collaboratively. Such a system can handle large-scale data processing and complex computing requirements, while also enhancing the flexibility and scalability of machine learning operations.

This service involves numerous modules, including distributed training architecture, data storage solutions, computing resource control tools, etc. These modules rely on standardized interfaces or protocols to achieve interconnectivity and ensure the smooth process of machine learning. Multi platform machine learning technology is widely deployed in industries such as data science research, artificial intelligence development, and high-performance computing, promoting cross regional and cross system collaborative innovation.

2.2. Existing cross platform service frameworks and platforms

Currently, cross platform machine learning services have been applied and promoted in multiple frameworks and platforms. Mainstream deep learning frameworks such as TensorFlow and PyTorch have implemented distributed computing and model training across multiple platforms. Especially when dealing with massive data and cloud computing tasks, service platforms such as TensorFlow's TensorFlow Serving and PyTorch's TorchServe can ensure smooth integration on diverse hardware and operating systems, accelerating model deployment and inference processes. Some mainstream cloud computing platforms such as Google Cloud Platform, Amazon Web Services (AWS), and Microsoft Azure also provide users with a cross platform machine learning solution and underlying architecture. Users can conduct model training and inference work under various operating systems and hardware conditions. These platforms utilize containerization technology and standardized API interfaces to achieve smooth data transmission between different platforms, efficient model management, and effective scheduling of computing tasks, thereby improving resource utilization and overall computing efficiency.

3. The integration status of cross platform machine learning services

3.1. Cross platform integration based on data sharing

Integrating machine learning services from different platforms, especially in data exchange, plays a decisive role in enhancing system efficiency and expanding application scope. With the advancement of big data and cloud computing technology, cross platform data exchange has been achieved, promoting data exchange and integration between different computing nodes. This type of integration strategy effectively removes data barriers and ensures smooth flow of data between numerous platforms under distributed architecture. Through the linkage and sharing of cloud computing and local data centers, machine learning algorithms can process massive amounts of data in the cloud and complete specialized training or inference tasks with the help of edge servers, optimizing computing performance and accelerating response time.

Although data exchange brings a lot of convenience, it also comes with a series of challenges. Due to the varying data structures and storage modes across different platforms, there may be inefficiencies during the migration and transformation of data, which can affect the overall system's operational efficiency. For example, inconsistencies in data structure, communication protocols, and access permissions between cloud and local servers increase the cost and difficulty of data exchange. When implementing data exchange between different platforms, the issues of data privacy and security are particularly prominent, especially when dealing with sensitive information. The secure transmission of data between platforms becomes particularly critical, and this issue urgently needs to be addressed. Although the integration of cross platform data exchange can improve performance, it must overcome challenges such as standardization, data privacy protection, and resource allocation in specific implementation in order to truly unleash its potential.

3.2. Collaborative training of models on different platforms

With the continuous expansion of machine learning technology, cross platform collaborative model training has gradually become a key means to enhance model performance and accelerate training pace. By using distributed computing, the advantages of each platform can complement each other and improve the execution speed of machine learning tasks. For example, cloud platform can provide massive data storage and powerful computing power, while edge computing devices are better at processing real-time data analysis and simple reasoning tasks. In such an environment, model training can complete large-scale deep learning tasks in the cloud, with real-time feedback and model fine-tuning on edge devices, achieving efficient collaborative training modes.

We have encountered many challenges when conducting collaborative training of cross platform models. The hardware configurations and network conditions between platforms often vary greatly, which may result in delayed data transmission or uneven allocation of computing tasks, negatively affecting the effectiveness of collaborative training. For example, communication between cloud and edge devices may result in inconsistent pace of model updates due to bandwidth limitations and latency issues. In addition, the interoperability barriers between algorithms on different platforms are also a major challenge in the collaborative training process. As each platform may adopt different technical frameworks or hardware acceleration techniques, this requires the model to have good cross platform transferability and adaptability.

3.3. Multi platform resource integration

Multi platform resource integration is a key link in cross platform machine learning services, which greatly improves the execution efficiency and performance level of machine learning jobs by allocating computing, storage, and network resources between various platforms. Given the widespread use of diverse hardware platforms and operating system environments, relying solely on a single platform is no longer sufficient to meet the enormous resource demands of large-scale machine learning tasks. Therefore, the integration of cross platform resources not only optimizes resource allocation and reduces operating costs, but also improves the effective utilization of computing resources and prevents excessive consumption of resources on a single platform. The complementary characteristics between various platforms help overcome the limitations of a single platform in handling high-performance computing needs. For example, the joint application of cloud computing platform and edge computing platform can reasonably allocate data processing tasks to the most appropriate platform to ensure efficient operation of data processes and optimal execution of model reasoning.

Although resource integration brings convenience, it also comes with many challenges. The resource allocation, management strategies, and communication protocols between platforms are different, which may make data transmission and computation allocation more cumbersome. At the same time, inconsistencies between platforms may pose adaptation challenges during the integration process, thereby affecting the reliability and operational efficiency of the system. For example, the inconsistency in storage protocols between cloud computing platforms and local servers may result in more latency during data transmission.

2.4. Establish a unified interface and standard

When integrating cross platform machine learning services, the benefits of unified interfaces and specifications are first reflected in simplifying the integration process and reducing its difficulty. Developers no longer need to write corresponding code or converters to adapt to different platforms, and this universal interface greatly reduces the burden of cross platform software development and operation. At the same time, consistent standards also facilitate cooperation between the open source community and commercial entities, promote the process of technical standardization, and reduce the duplication of resources.

However, the underlying differences between different platforms and architectures still exist, which may result in standardized interfaces not being fully compatible in certain specific application scenarios. Some cloud computing platforms may rely on specific hardware acceleration features that may conflict with standard interfaces, leading to performance degradation or functional limitations. In addition, the development of standards requires balancing the interests of all parties, which may sometimes mean compromising on some technical issues, thus failing to achieve the best performance pursued by certain platforms.

4. Optimization strategies for cross platform machine learning services

4.1. Improving machine learning algorithms to enhance their accuracy

Improving the accuracy of machine learning algorithms is one of the core optimization methods in enhancing the performance of cross platform machine learning services. Conventional machine learning techniques often face dual challenges of computational efficiency and accuracy when dealing with large amounts of data and different architecture platforms. In order to overcome the shortcomings of these traditional algorithms, researchers have developed numerous innovative optimization solutions. For example, in the field of deep learning, techniques such as adjusting neural network architecture, adopting adaptive learning rates, and optimizing activation functions can be used to enhance the model's ability to handle complex datasets. Especially for ensemble learning strategies such as random forest and gradient boosting, they enhance prediction accuracy by integrating multiple basic learning units, which is particularly critical for cross platform model training under diverse hardware conditions.

Conventional Convolutional Neural Networks (CNNs) encounter performance limitations in image recognition tasks on specific cross platform machine learning platforms. In order to overcome this limitation, researchers have adopted a new improvement method that integrates multiple CNN models with different structures and performs weighted integration. This strategy improves recognition accuracy while maintaining computational efficiency. This strategy adjusts the weights of different models during the integration process, so that each model contributes different prediction results based on its performance on a specific dataset. The optimized objective function can be expressed as:

$$L(\theta) = \sum_{i=1}^{N} w_i \cdot L_i(\theta) \tag{1}$$

In formula (1), Li (θ) is the loss function of the i-th model, wi is the weight of the corresponding model, and the optimized objective function maximizes overall accuracy by weighted averaging the losses of each model. After adopting this improvement strategy, cross platform machine learning services can achieve significant performance improvements on various hardware platforms.

4.2. Improving Model Training Speed

Accelerating the pace of model training is particularly crucial in improving the efficiency of cross platform machine learning services. Machine learning models such as deep learning often need to deal with massive datasets and high difficulty computational tasks, resulting in long training cycles. In order to shorten the training time, the industry often implements strategies such as parallel processing, model simplification, and leveraging efficient hardware. Parallel processing technology reduces training time by mobilizing multi-core processors or building distributed computing networks to achieve multitasking synchronous operations. The model compression method achieves a leap in training speed by simplifying network parameters and reducing computational burden, while ensuring model accuracy. For example, when handling text classification tasks for cross platform machine learning services, GPU acceleration and model pruning techniques are used to speed up the training process. By comparing the time of GPU accelerated training with traditional CPU training, the acceleration effect is clear at a glance. The following table shows the time comparison of training the same model on different hardware platforms:

Table 1 Comparison of Model Training

Training Platform	Training time (hours)	Accuracy (precision)
CPU	12	89.3%
GPU	4	89.5%
GPU+pruning	2	89.5%

According to Table 1, after GPU acceleration processing, the training cycle of the model has been reduced from the original 12 hours to only 4 hours. After further introducing pruning techniques, the training time was further reduced to 2 hours, while the accuracy of the model remained unchanged. This achievement proves that the fusion of parallel processing and hardware acceleration can improve the efficiency of model training without affecting its accuracy, which is particularly crucial for meeting the fast response requirements of cross platform machine learning services.

4.3. Reasonable allocation of resources

Scientific allocation of resources is crucial for enhancing the operational efficiency and economy of machine learning services with diverse operating systems. Machine learning algorithms often consume a significant amount of computing power during the training and prediction stages, such as processors, graphics processors, random access memory, and storage devices. In order to maximize the efficiency of resource utilization, it is necessary to allocate resources reasonably according to different tasks and device characteristics. In the context of heterogeneous platform coexistence, resource allocation for cross platform services becomes more complex and difficult. Proper resource allocation can not only improve computing efficiency, but also reduce energy consumption and hardware load. Taking a cross platform machine learning service as an example, when performing large-scale image recognition tasks, the system dynamically allocates graphics processor resources according to actual needs, and can automatically adjust the number of graphics processors and memory capacity as the model training progresses. The following table shows the performance of model training under different resource configurations:

Table 2. Performance of Model Training

Resource allocation	Number of GPUs	Memory (GB)	Training time (hours)	Accuracy (precision)
Configuration 1	1	16	10	92.1%
Configuration 2	2	32	6	92.3%
Configuration 3	4	64	3	92.4%

The resource allocation strategy of the system can be represented by the following formula:

$$T = \frac{C}{R} + D \tag{2}$$

In formula (2), T is the training time, C is the complexity of the computational task, R is the investment of resources (such as the combination of GPU quantity and memory capacity), and D is the delay of data transmission and preprocessing. By adjusting the resource allocation R, the training time T can be reduced to improve system efficiency. After careful resource optimization and allocation, the training period has been effectively compressed, and the accuracy of the model has slightly improved. This achievement demonstrates that efficient resource allocation plays an indispensable role in improving the performance of cross platform machine learning services.

4.4. Enhance the fault tolerance of cross platform machine learning services

In the optimization of cross platform machine learning services, enhancing fault tolerance is an important strategy to ensure system stability and reliability. Due to the involvement of multiple hardware and software platforms in a cross platform environment, the system is prone to various interference factors during operation, such as hardware damage, network connectivity issues, or task allocation errors, all of which may cause the termination of the training process or data loss. Therefore, it is particularly important to enhance the system's fault response level and ensure that it can automatically repair or take effective measures when problems occur. With the help of fault response mechanisms, the system can automatically switch to a backup node or restart training tasks from the most recently saved state point in the event of a fault, thereby preventing task failure caused by a single point of failure. This measure enhances the reliability of the system and reduces damage in the event of a malfunction.

In the process of building a machine learning recommendation system involving multiple platforms, a distributed training architecture was deployed and a periodic data backup mechanism was set up for each task. Once a node fails, the system will automatically detect and switch to a backup node to continue training tasks without causing any loss of work.

5. Conclusion

This article conducts in-depth research on the integration and optimization strategies of cross platform machine learning services, provides a detailed review of the current cross platform service architecture, and proposes rich improvement measures. At the integration level, practical and feasible solutions have been developed for core issues such as data interoperability, model joint training, and resource collaboration. At the level of performance improvement, a series of strategies were discussed to enhance algorithm accuracy, accelerate model training speed, optimize resource allocation, and improve service stability. These research results not only provide theoretical basis for the application of hybrid platform machine learning services, but also build a solid foundation for the continuous progress of related technologies. Despite this, there are still some difficulties in integrating and improving the performance of cross platform machine learning services, such as platform heterogeneity, data privacy security issues, etc. Future research can focus on these difficulties in order to further improve the efficiency and scalability of hybrid platform machine learning services.

Reference

[1] Mufti A, Fejzi N. Advancing Veterinary Epidemiology by Integration of Machine Learning: Current Status and Future Perspectives. Veterinaria, 2024, 73(2).

[2] Polo-Triana S, Camilo Gutierrez J, Leon-Becerra J. Integration of Machine Learning in the Supply Chain for Decision Making: A Systematic Literature Review. Journal of Industrial Engineering & Management, 2024, 17(2).

[3] Hwang J H, emailprotected, Emailprotected E M, et al. Synergistic Integration of Machine Learning with Microstructure/Composition-Designed SnO 2 and WO 3 Breath Sensors. ACS Sensors, 2024, 9(1):182-194.

[4] Wang Y, Guo F, Chen S, et al. Quantitative study of rainfall lag effects and integration of machine learning methods for groundwater level prediction modelling. Hydrological processes. 2024(5):38.

[5] Alabdulatif A, Thilakarathne N N, Aashiq M. Machine Learning Enabled Novel Real-Time IoT Targeted DoS/DDoS Cyber Attack Detection System. Computers, Materials & Continua, 2024, 80(3):3655-3683.

[6] Liu, Y. (2025). Use SQL and Python to Advance the Effect Analysis of Financial Data Automation. Financial Economics Insights, 2(1), 110-117.

[7] Ye, J. (2025). Optimization of Neural Motor Control Model Based on EMG Signals. International Journal of Engineering Advances, 2(4), 1-8.

[8] Lu, C. (2025). Application of Multi-Source Remote Sensing Data and Lidar Data Fusion Technology in Agricultural Monitoring. Journal of Computer, Signal, and System Research, 2(7), 1-6.

[9] K. Zhang. Optimization and Performance Analysis of Personalized Sequence Recommendation Algorithm Based on Knowledge Graph and Long Short Term Memory Network. 2025 2nd International Conference on Intelligent Algorithms for Computational Intelligence Systems (IACIS), Hassan, India, 2025, pp. 1-6.

[10] Su H, Luo W, Mehdad Y, et al. Llm-friendly knowledge representation for customer support[C]//Proceedings of the 31st International Conference on Computational Linguistics: Industry Track. 2025: 496-504.

[11] B. Li. Research on the Spatial Durbin Model Based on Big Data and Machine Learning for Predicting and Evaluating the Carbon Reduction Potential of Clean Energy. 2025 International Conference on Intelligent Communication Networks and Computational Techniques (ICICNCT), Bidar, India, 2025, pp. 1-5.

[12] Q. Xu. Implementation of Intelligent Chatbot Model for Social Media Based on the Combination of Retrieval and Generation. 2025 2nd International Conference on Intelligent Algorithms for Computational Intelligence Systems (IACIS), Hassan, India, 2025, pp. 1-7.

[13] Y. Zou. Research on the Construction and Optimization Algorithm of Cybersecurity Knowledge Graphs Combining Open Information Extraction with Graph Convolutional Networks. 2025 2nd International Conference on Intelligent Algorithms for Computational Intelligence Systems (IACIS), Hassan, India, 2025, pp. 1-5.

[14] M. Zhang. Research on Joint Optimization Algorithm for Image Enhancement and Denoising Based on the Combination of Deep Learning and Variational Models. 2025 International

Conference on Intelligent Communication Networks and Computational Techniques (ICICNCT), Bidar, India, 2025, pp. 1-5,
[15] F. Liu. Architecture and Algorithm Optimization of Realtime User Behavior Analysis System for Ecommerce Based on Distributed Stream Computing. 2025 International Conference on Intelligent Communication Networks and Computational Techniques (ICICNCT), Bidar, India, 2025, pp. 1-8.
[16] Mingjie Chen. (2025). Exploration of the Application of the LINDDUN Model in Privacy Protection for Electric Vehicle Users. Engineering Advances, 5(4), 160-165.
[17] Hu, Q. (2025). The Practice and Challenges of Tax Technology Optimization in the Government Tax System. Financial Economics Insights, 2(1), 118-124.
[18] Sheng, C. (2025). Analysis of the Application of Fintech in Corporate Financial Decision-Making and Its Development Prospects. Financial Economics Insights, 2(1), 125-130.
[19] Wei, X. (2025). Deployment of Natural Language Processing Technology as a Service and Front-End Visualization. International Journal of Engineering Advances, 2(3), 117-123.
[20] Ding, J. (2025). Research On CODP Localization Decision Model Of Automotive Supply Chain Based On Delayed Manufacturing Strategy. arXiv preprint arXiv:2511. 05899.