

# Enhancing Collaborative Recommendations with Federated Learning in E-Commerce

Baek Jeong

Kyung Hee University, Seoul, Korea

**Objectives:** This study explores the enhancement of collaborative recommendation systems in e-commerce platforms using federated learning (FL) integrated with glocalization strategies, focusing on data privacy preservation. The aim is to leverage natural language processing (NLP)-based models to capture real-world product relationships without sharing raw data across different platforms.

**Methods:** An NLP-based recommendation system employing a transformer-based architecture forms the core of this study. The system processes tokenized product names to determine item relationships, encapsulating the complexities of consumer behavior patterns. The glocal federated learning (glocal FL) approach is implemented by selectively sharing encoder and decoder parameters, thus allowing certain layers to remain private. The study is conducted across distinct datasets, specifically those from retailer and grocery sectors, evaluated under both Homogeneous and Heterogeneous settings. This methodological approach facilitates a thorough investigation into the system's adaptability to varying data environments.

**Results:** The findings indicate that glocal FL significantly improves HR@20 scores in Homogeneous environments, surpassing local models by up to 10%. However, performance declines in Heterogeneous environments due to data variability, highlighting the need for refined strategies to manage cross-domain complexities.

**Conclusions:** FL with glocalization offers a promising approach to enhance e-commerce recommendations by balancing global learning with local customization. This strategy effectively addresses data heterogeneity and supports privacy-preserving business communication. While proving effective in homogeneous settings, further adaptation is needed to optimize performance in diverse environments. This study underscores the potential of glocal FL to improve recommendation accuracy and maintain consumer data privacy in e-commerce.

**Key Words:** Federated Learning, Natural Language Processing (NLP)-Based Recommendation, E-Commerce, Data Privacy, Business Communication

---

<sup>†</sup>This paper is based on the first author's doctoral dissertation "NLP-based Recommendation Approach for Diverse Service Generation" (Kyung Hee University, 2024) with modifications.

Received: Oct 20, 2024 Revised: Nov 26, 2024 Accepted: Dec 9, 2024

Corresponding author: Baek Jeong

Department of Big Data Analytics, Kyung Hee University, Dongdaemun-gu, Seoul 02447, Korea

Tel: +82-10-3227-9950, E-mail: ylbaek@khu.ac.kr

This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/4.0/>) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

Copyright © 2024 Korean Association for Business Communication.

## Introduction

The expansion of data-driven businesses underscores the importance of collaboration among e-commerce platforms. Specifically, data-driven applications such as product recommendations, consumer behavior prediction, and personalized marketing strategies further expand the potential for inter-platform collaboration (Chen et al., 2019; Sun et al., 2019). However, concerns over data privacy and security act as major barriers to data sharing between companies, thereby impeding business collaboration (Mothukuri et al., 2021; Yang, Liu, Chen, & Tong, 2019). In this context, Federated Learning (FL) has emerged as an innovative approach that enables the exchange of model training parameters without sharing raw data (Kairouz et al., 2021; McMahan, Moore, Ramage, Hampson, & y Arcas, 2017). FL allows participating platforms to train models locally on their own data and then aggregate the learned parameters to create a global model. This approach enables each platform to maintain data independence while improving performance through collaborative learning (Chen et al., 2019; Yang et al., 2019).

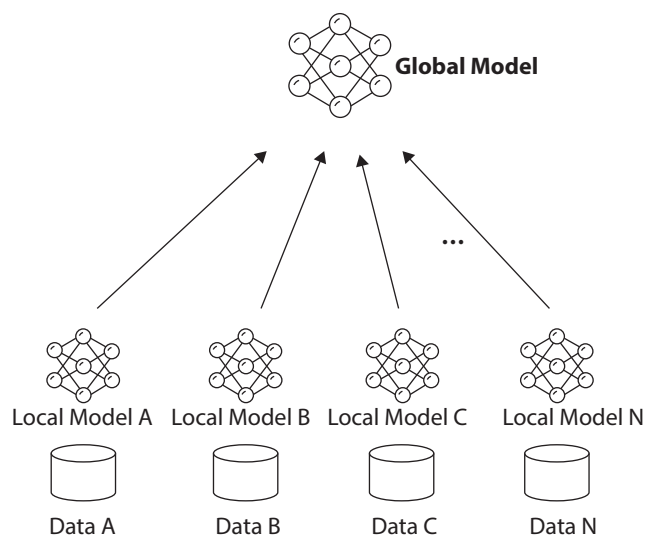
Previous studies have demonstrated the effectiveness of FL in enhancing data protection while providing personalized services (Shalom, Roitman, & Kouki, 2021). For example, Google's Gboard employs FL to enhance the performance of its global predictive model while maintaining the privacy of individual user data (McMahan et al., 2017). However, traditional FL approaches face significant challenges in managing data heterogeneity across platforms, particularly with respect to domain-specific characteristics and linguistic variations that are prevalent in e-commerce environments (Lee et al., 2024; Yang et al., 2019). To overcome these limitations, glocal federated learning (glocal FL) has been innovatively designed. glocal FL strategically enhances the traditional model by selectively sharing global parameters that benefit from broader aggregation while concurrently preserving local parameters that cater to unique platform-specific needs. This dual strategy allows glocal FL to optimize collaborative learning in natural language processing (NLP)-based recommendation systems by efficiently handling diverse vocabulary sizes and distinct data characteristics across domains (Bae, Jeong, Hwangbo, & Lee, 2022).

This study leverages NLP-based recommendation models as the foundation for FL to explore how e-commerce platforms can collaborate on business communication without sharing data. NLP-based recommendation systems tokenize product names to learn relationships between items, thereby reflecting the true semantic meaning of the data (Jeong & Lee, 2024). Unlike conventional recommendation systems, this approach preserves the authentic essence of e-commerce data, contributing

to more informed business decision-making. Specifically, this study empirically examines how glocal FL can enhance collaborative learning across platforms without sharing data. Figure 1 illustrates the fundamental structure of FL. The global model is created by aggregating parameters learned from local models, with each local model safeguarding its own data while participating in collaborative learning (McMahan et al., 2017). This structure is particularly valuable in data-sensitive environments, and the glocal FL strategy enhances such collaboration more effectively (Li et al., 2024; Yang et al., 2019). This study aims to make the following contributions. First, it proposes a method to foster collaboration among e-commerce platforms by integrating NLP-based recommendation systems with glocal FL strategies (Jeong & Lee, 2024). Second, it empirically evaluates how glocal FL improves the performance of collaborative learning in diverse domain environments. Lastly, it explores the potential of new approaches to strengthen business collaboration without data sharing (Chen & Liu, 2017).

## Methods

This study employs an NLP-based recommendation model within a FL framework to facilitate collaboration among e-commerce platforms while safeguarding the privacy of business data (Yang et al., 2019). Traditional FL approaches often experience degraded learning performance when there are significant disparities in dataset size or domain characteristics across platforms (McMahan et al., 2017). To address this issue, the glocal FL strategy is adopted (Bae et al., 2022). By selectively sharing global parameters while retaining platform-specific local parameters, glocal FL

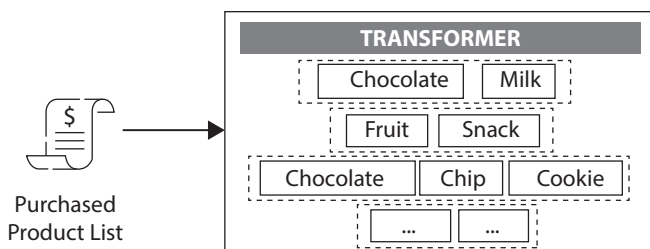


**Figure 1.** Schematic representation of federated learning.

mitigates the challenges posed by data heterogeneity. The proposed model tokenizes product names as input data, processes user purchase history, and leverages a transformer architecture to learn purchase patterns (Vaswani et al., 2017). This allows the model to generate a recommendation list of products likely to be purchased next. This tokenization approach not only captures the relationships between products but also reflects the semantic meaning of the data, enhancing the model's interpretability. The process contributes to improving the learning performance of the model by representing the actual meaning of the data more effectively (Arya, 2023; Basu, 2020; Jeong & Lee, 2024).

### Application of Natural Language Processing (NLP)-Based Recommendation Model

This study builds upon the methodology developed in previous research (Jeong & Lee, 2024), with a key feature being the tokenization of product names at the word level. For instance, "Chocolate Milk" and "Strawberry Milk" are connected through the common token "Milk," which helps the model learn similarities and relationships between products. This tokenization method better captures the actual meaning of the data and significantly improves the accuracy of recommendation models in e-commerce environments, where product names often carry critical information (Park, Lee, Jang, & Jung, 2020). The NLP-based recommendation model used in this study employs a transformer architecture with an encoder-decoder structure to predict products likely to be purchased next based on users' purchase histories (Vaswani et al., 2017). The transformer processes the input purchase data, enabling the model to understand the relationships between products and uncover users' purchasing patterns. Figure 2 visually represents the transformer structure utilized in this study. This structure takes a list of purchased products as input, processes it through multiple layers, and analyzes the relationships between products. For example, if "Chocolate" and "Milk" frequently co-occur, they are grouped into a related category to predict other products that are likely to be purchased next. Compared to non-NLP-based recommendation systems, this model better reflects



**Figure 2.** Natural language processing (NLP)-based recommendation system using transformer architecture.

the semantic meaning of data and offers superior recommendation accuracy (Vaswani et al., 2017).

This study demonstrates that the NLP-based recommendation model provides a robust foundation for efficient learning and enhanced collaboration in FL environments without data sharing. This approach validates its potential not only as a methodology for improving collaboration among e-commerce platforms but also as a means of safeguarding data privacy.

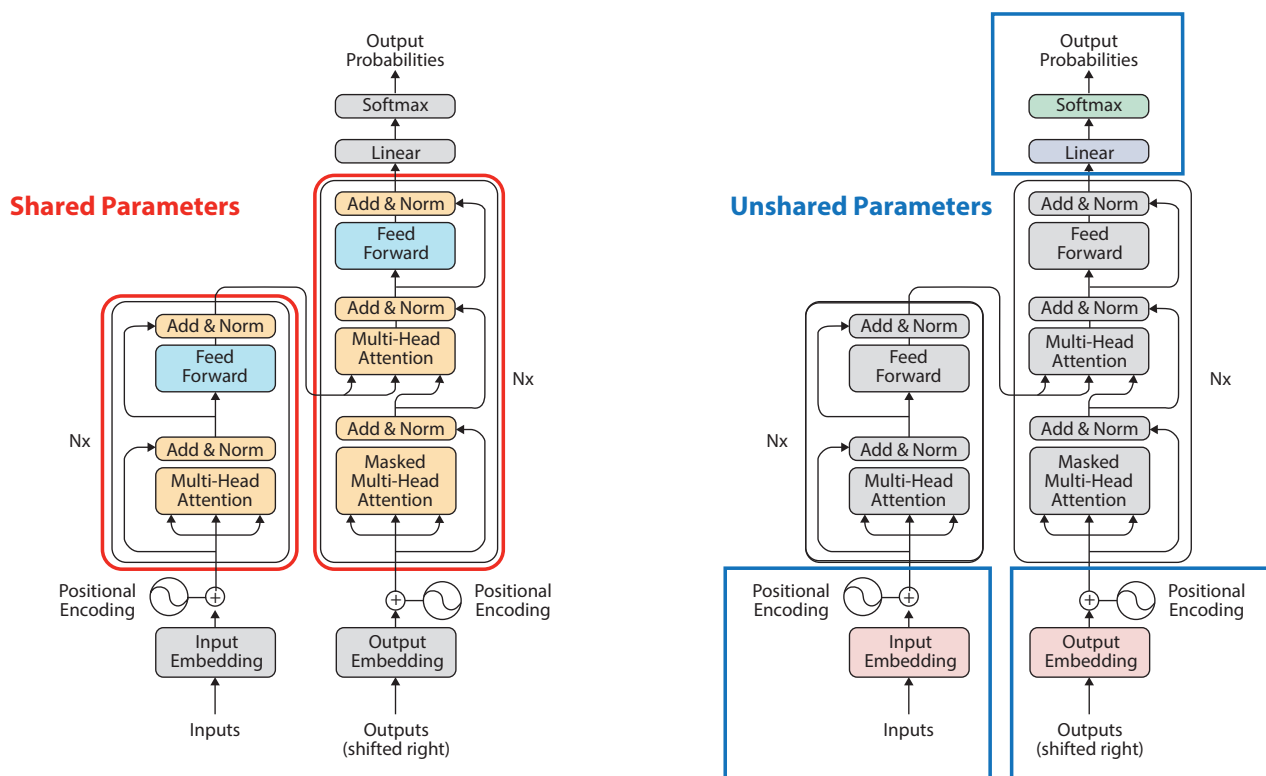
### Federated Learning Environment

This study applies the glocal FL strategy to the FL environment, drawing on insights from previous research. Glocal FL combines the strengths of global learning and local learning, effectively enhancing collaboration among platforms. In particular, the globalization strategy proposed by Bae et al. (2022) has been shown to facilitate collaborative learning while preserving the uniqueness of data by separating global and local parameters in FL models. This approach allows both collaboration and performance improvements in situations where data sharing is not feasible within the FL environment.

The core of glocal FL lies in distinguishing between global and local parameters during training. Global parameters are shared across all participating platforms, reflecting common characteristics to foster collaboration among them. In contrast, local parameters retain platform-specific information and remain confined to individual platforms. This separation enables the preservation of data uniqueness across platforms while harmonizing with the global model. Such a strategy makes collaborative learning feasible through the exchange of model parameters without the need to share raw data.

Glocalization is particularly promising in the context of NLP research. In NLP environments, vocabulary sizes can vary significantly across datasets. To account for this, embedding layers and output layers are excluded from parameter sharing, while identical layers in the encoder and decoder are shared. This approach maintains dataset specificity while maximizing the efficiency of global model training.

In this study, globalization strategies are implemented in a transformer-based model architecture by distinguishing between shared and unshared parameters. Figure 3 illustrates the structure of the transformer model, highlighting the distinction between shared parameters and unshared parameters. Shared parameters include the attention and feed-forward layers of the encoder and decoder, which facilitate global learning. Conversely, embedding layers and output layers are not shared, as their parameter sizes differ depending on the unique vocabulary sizes of each platform. This configuration alleviates issues arising from dataset heterogeneity in FL environments and con-



**Figure 3.** Shared and unshared parameters in transformer model architecture.

tributes to improving the performance of collaborative learning.

Figure 4 demonstrates how glocal FL is applied to an NLP-based transformer architecture, detailing the implementation of shared and unshared parameters. This figure underscores the adaptive application of glocal FL, optimizing the transformer model for varied linguistic and data characteristics specific to each platform. By applying glocal FL principles, the model effectively integrates the benefits of shared global knowledge with the necessity of maintaining local specificities, thereby enhancing both precision and relevance of NLP tasks across different data-sensitive environments.

As highlighted in prior research, the glocalization strategy offers flexibility to maximize collaboration across diverse domain characteristics in FL environments. This study extends that approach by applying glocalization to NLP-based recommendation models, demonstrating the feasibility of achieving high learning performance and collaboration in FL environments without data sharing.

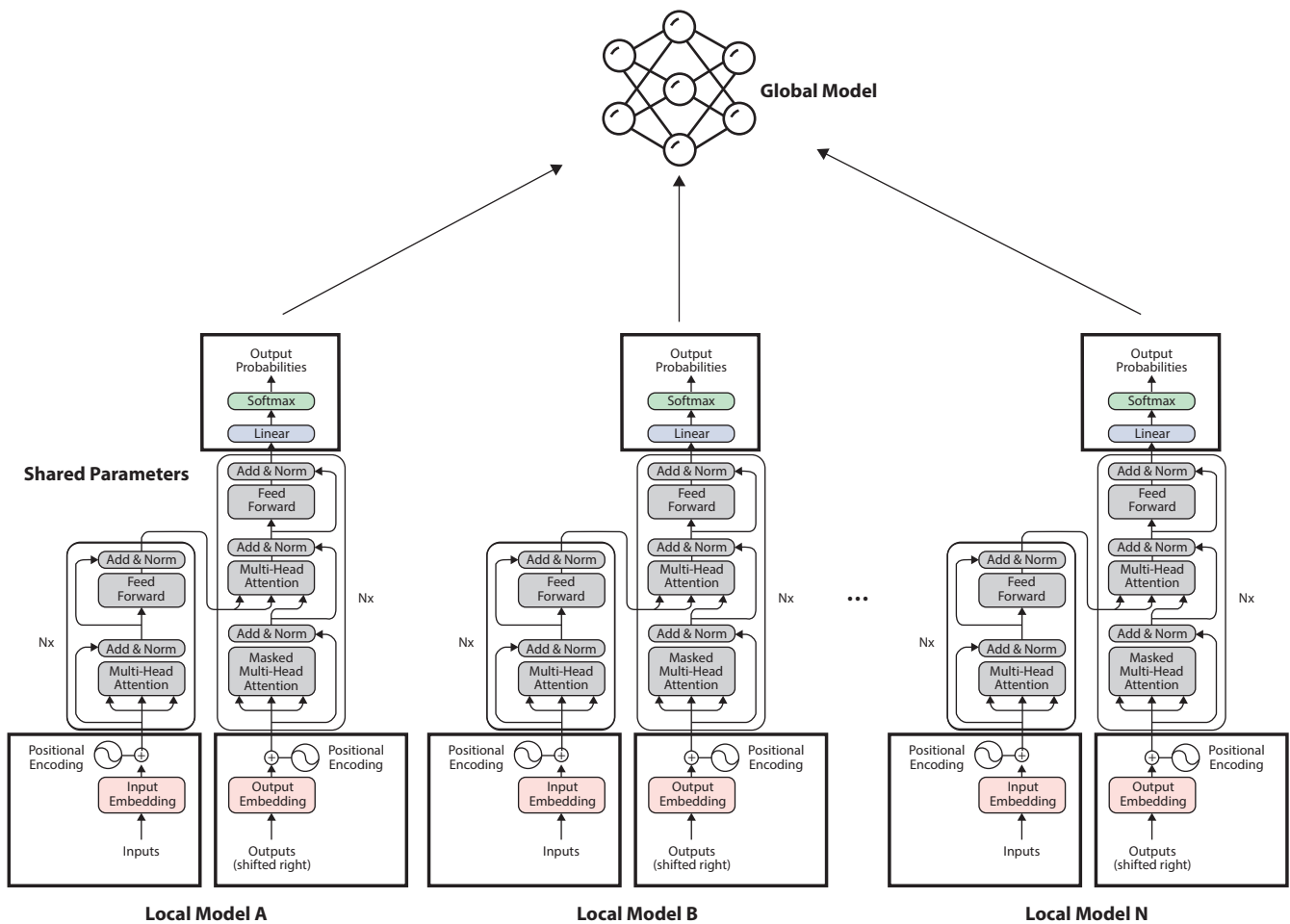
### Dataset Composition and Preprocessing

This study employs two major datasets, UK e-commerce and Instacart, to validate the performance of the glocal FL strategy applied to an NLP-based recommendation system. Each dataset underwent preprocessing and partitioning to reflect the charac-

teristics of domain-specific data and to evaluate model performance in an FL environment without data sharing.

The UK e-commerce dataset comprises transaction records from a UK-based online retail platform, including a total of 210,367 transaction records, 2,920 users, and 3,741 product names. This dataset encompasses a variety of product categories, such as groceries, household items, and electronics, and was used to generate training data based on user purchase histories and product names. The instacart dataset, derived from the U.S.-based Instacart platform, includes 33,819,106 transaction records, 206,209 users, and 49,688 product names. It features detailed purchase histories, such as product purchase cycles and timestamps, making it suitable for training the NLP-based recommendation model.

The preprocessing phase involved structuring training and label data based on user purchase histories. Product names in each dataset were sequentially listed to create individual user purchase histories, which were then split into training and label data. For example, a purchase sequence of [Product A, Product B, Product C, Product D] was used as training data, with [Product E] set as the label. This method was designed to train the model to learn user purchase patterns and evaluate recommendation performance effectively. Due to the large size of the instacart dataset, only 1% of the data was sampled to maintain a



**Figure 4.** Application of glocal federated learning to natural language processing (NLP)-based recommendation models.

feasible scale for experiments.

The datasets were divided by domain and size to simulate various user scenarios in an FL environment. The UK e-commerce dataset was categorized as retailer, while the instacart dataset was labeled as grocery. Each domain was further split into small and large datasets, resulting in four experimental groups. The number of rows in each group is shown in Table 1.

This dataset partitioning was designed to analyze the impact of dataset size and domain characteristics on learning performance in FL environments. It allowed the evaluation of FL model performance across heterogeneous domain environments, such as retailer and grocery, as well as differences between Small and Large data groups.

By leveraging these datasets, the study experimentally verified the performance of FL models in various scenarios, demonstrating that the glocalization strategy effectively addresses dataset heterogeneity. This approach achieved improved collaborative learning and recommendation performance, even in environments with restricted data sharing.

**Table 1.** Dataset for federated learning experiments

Dataset	Retailer S	Retailer L	Grocery S	Grocery L
#Rows	28,482	42,723	66,576	99,864

## Results

This study evaluated the performance of an NLP-based recommendation model in a FL environment using the glocalization strategy. The glocal FL strategy was meticulously designed to reflect the characteristics of each dataset, focusing on how the encoder and decoder layers' parameters were shared or unshared based on the dataset-specific requirements. This selective parameter-sharing approach balances the advantages of global learning with the preservation of dataset-specific uniqueness.

The performance of the model was assessed using the Hit Rate@20 (HR@20) metric. HR@20 measures the proportion of recommended items in the top 20 that were actually purchased by users. As a widely recognized metric for recommendation



accuracy, HR@20 provides a straightforward and clear measure of recommendation quality.

To further elaborate on the experimental setup, four distinct experimental groups were created based on dataset size and parameter configuration requirements to better understand the effectiveness of glocal FL under varying conditions. Each group was tailored to simulate different aspects of data heterogeneity and learning dynamics, as outlined in Table 1.

As shown in Table 2, the retailer S and retailer L datasets had vocabulary sizes of 2,138 and 2,160, respectively, with a total of approximately 1.7 million parameters. In contrast, the grocery S and grocery L datasets had larger vocabulary sizes and included over 4 million parameters each. The shared parameters, totaling 879,360, consisted of encoder and decoder layers, which facilitated the learning of common patterns across datasets through global learning. The unshared parameters captured the unique characteristics of each dataset and were handled separately in the embedding and output layers.

Table 3 compares the HR@20 performance of the recommendation model under different FL configurations. In homogeneous environments, the glocal FL strategy demonstrated performance improvements over the local model. For example, in the retailer datasets, the FL glocal AB configuration increased

HR@20 values to 0.111 and 0.119, respectively, while in the grocery datasets, the FL glocal CD configuration improved HR@20 values to 0.063 and 0.067. These results indicate that parameter sharing was effective in environments where datasets had similar domain characteristics.

Conversely, in heterogeneous environments, the FL glocal ABCD configuration resulted in performance declines. For the retailer datasets, HR@20 values dropped to 0.076 and 0.084, while performance degradation was also observed in the grocery datasets. This suggests that in cases of high data heterogeneity, parameter sharing may negatively impact model performance.

Overall, the results demonstrate that the glocal FL strategy effectively improves the performance of recommendation models in homogeneous environments, while performance varies depending on dataset characteristics and learning conditions. The HR@20 metric provided a clear evaluation of recommendation accuracy, serving as the basis for analyzing the potential and limitations of glocal FL. These findings offer significant insights into the feasibility of collaborative learning across diverse domain environments.

## Discussion

This study evaluated the performance of an NLP-based rec-

**Table 2.** Number of (un)shared parameters of each model

	(A) Retailer S	(B) Retailer L	(C) Grocery S	(D) Grocery L
Vocabulary size	2,138	2,160	8,765	9,521
#Total parameters	1,702,490	1,710,960	4,253,885	4,544,945
#Shared parameters		879,360		
Encoder layer		406,528		
Decoder layer		472,832		
#Unshared parameters	823,130	831,600	3,374,525	3,665,585
Encoder embedding layer	273,664	276,480	1,121,920	1,218,688
Decoder embedding layer	273,664	276,480	1,121,920	1,218,688
Output layer	275,802	278,640	1,130,685	1,228,209

**Table 3.** Comparison of FL with glocalization performance

HR@20	Proprietary	Homogeneous		Heterogeneous	
	Local	FL glocal AB (retailer data)	Data-shared AB (retailer data)	FL glocal ABCD (all data)	Data-shared ABCD (all data)
(A) Retailer S	0.106	0.111	0.124	0.076	0.105
(B) Retailer L	0.112	0.119	0.129	0.084	0.111
	Local	FL glocal CD (grocery data)	Data-shared CD (grocery data)	FL glocal ABCD (all data)	Data-shared ABCD (all data)
(C) Grocery S	0.051	0.063	0.076	0.055	0.072
(D) Grocery L	0.062	0.067	0.086	0.065	0.080

ommendation model within a FL environment, leveraging the glocalization strategy to explore the potential for collaborative learning among e-commerce platforms without data sharing. The experimental results show that in homogeneous environments, the glocal FL strategy improves recommendation performance compared to local models. This improvement indicates the effectiveness of selectively sharing global and local parameters under the glocalization strategy when datasets exhibit high similarity.

In contrast, the performance declines in heterogeneous environments illustrate that parameter sharing can complicate learning processes when datasets are highly diverse. For example, in datasets with significant domain differences such as those between retailers and groceries, global parameters intended to learn common patterns may inadvertently impair model performance. These findings suggest that while the glocal FL strategy is effective in homogeneous environments, there is a necessity for additional learning mechanisms that are tailored to the compatibility of datasets in heterogeneous environments.

Another significant finding from this study is the impact of balancing shared and unshared parameters on the performance of recommendation models. By limiting shared parameters to encoder and decoder layers and excluding embedding and output layers, the unique characteristics of each dataset were preserved. This selective sharing approach not only reflects the vocabulary size and linguistic characteristics of the datasets but also contributes to maintaining the authenticity of e-commerce data.

The use of the HR@20 metric facilitated a straightforward and reliable analysis of recommendation quality. The results convincingly demonstrated that the glocal FL strategy could enhance recommendation accuracy in homogeneous environments, thereby highlighting its significance as a method for improving collaboration among e-commerce platforms through a blend of global and local learning.

Despite validating the efficacy of the glocal FL strategy in fostering collaborative learning without data sharing, this study also acknowledges certain limitations. Firstly, additional mechanisms for learning or data preprocessing strategies are necessary to offset performance declines in heterogeneous environments. Secondly, the reliance on HR@20 as the sole performance metric may not fully capture the broader user experience, suggesting that incorporating more diverse evaluation metrics could yield a more comprehensive analysis. Lastly, considering the application of the glocal FL strategy in sectors beyond e-commerce could provide valuable insights and demonstrate the strategy's broader applicability.

The findings underscore the potential of the glocal FL strategy to enable collaborative learning among e-commerce platforms without the need for data sharing while also highlighting

challenges that need addressing in practical implementations. This study serves as a foundational exploration into the application of glocal FL, demonstrating that it is a potent method for enhancing both data privacy and recommendation accuracy and providing direction for future research.

## Conclusion

This study proposed and evaluated a method for enabling collaborative learning among e-commerce platforms without data sharing by applying the glocalization strategy in a FL environment with an NLP-based recommendation model. The glocal FL strategy combines the strengths of global and local learning, demonstrating its effectiveness in homogeneous environments where it enhances recommendation accuracy significantly.

However, the challenges in heterogeneous environments where the performance declines suggest a need for adaptive strategies that better manage the compatibility between diverse datasets. This insight is critical for business applications where e-commerce platforms frequently deal with varied product types and customer demographics.

The contributions of this study are twofold: firstly, it shows that glocal FL can significantly improve the efficacy of recommendation systems in settings with similar data characteristics. Secondly, it identifies critical areas where Glocal FL may require enhancements to better serve diverse environments, thereby providing a roadmap for future research in applying FL to business applications.

Future research directions proposed include developing new learning mechanisms to better handle dataset heterogeneity and incorporating more comprehensive evaluation metrics that extend beyond HR@20 to consider user experience more holistically. Expanding the application of Glocal FL strategies to industries beyond e-commerce is also seen as a valuable exploration path.

## References

- Arya, V. (2023). *The future of personalized recommendations: Federated learning and privacy*. Taichung, Taiwan: Insights2Techinfo.
- Bae, S., Jeong, B., Hwangbo, Y., & Lee, K. J. (2022). AI sharing for glocalization: Federated learning experiments. *Paper Presented at the 24th Convergence Conference of the Korea Academic Society for Business Administration*. Yeosu, Korea.
- Basu, R. (2020). *Privacy-preserving recommendation system using federated learning* (Master's these). New Jersey Institute of Technology, Newark, NJ.
- Chen, Q., Zhao, H., Li, W., Huang, P., & Ou, W. (2019). Behavior sequence transformer for e-commerce recommendation in Alib-

- aba. *DLP-KDD '19: Proceedings of the 1st International Workshop on Deep Learning Practice for High-Dimensional Sparse Data* (pp. 1-4). Anchorage, AK.
- Jeong, B., & Lee, K. J. (2024). NLP-based recommendation approach for diverse service generation. *IEEE Access*, *12*, 14260-14274.
- Kairouz, P., McMahan, H. B., Avent, B., Bellet, A., Bennis, M., Bhagoji, A. N., ... Zhao, S. (2021). Advances and open problems in federated learning. *Foundations and Trends® in Machine Learning*, *14*(1-2), 1-210.
- Lee, K. J., Jeong, B., Kim, S., Kim, D., & Park, D. (2024). General commerce intelligence: Glocally federated NLP-based engine for privacy-preserving and sustainable personalized services of multi-merchants. *Proceedings of the Thirty-Eighth AAAI Conference on Artificial Intelligence* (pp. 22752-22760). Vancouver, BC.
- Li, Z., Wu, X., Pan, W., Ding, Y., Wu, Z., Tan, S., ... Ming, Z. (2024). FedCORE: Federated learning for cross-organization recommendation ecosystem. *IEEE Transactions on Knowledge and Data Engineering*, *36*(8), 3817-3831.
- McMahan, B., Moore, E., Ramage, D., Hampson, S., & y Arcas, B. A. (2017). Communication-efficient learning of deep networks from decentralized data. *Proceedings of the 20th International Conference on Artificial Intelligence and Statistics (AISTATS)* (pp. 1273-1282). Fort Lauderdale, FL.
- Mothukuri, V., Parizi, R. M., Pouriyeh, S., Huang, Y., Dehghan-tanha, A., & Srivastava, G. (2021). A survey on security and privacy of federated learning. *Future Generation Computer Systems*, *115*, 619-640.
- Park, K., Lee, J., Jang, S., & Jung, D. (2020). An empirical study of tokenization strategies for various Korean NLP tasks. *Proceedings of the 1st Conference of the Asia-Pacific Chapter of the Association for Computational Linguistics and the 10th International Joint Conference on Natural Language Processing* (pp. 133-142). Suzhou, China.
- Shalom, O. S., Roitman, H., & Kouki, P. (2021). Natural language processing for recommender systems. In F. Ricci, L. Rokach, & B. Shapira (Eds.), *Recommender systems handbook* (pp. 447-483). New York, NY: Springer.
- Sun, F., Liu, J., Wu, J., Pei, C., Lin, X., Ou, W., ... Jiang, P. (2019). BERT4Rec: Sequential recommendation with bidirectional encoder representations from transformer. *CIKM '19: Proceedings of the 28th ACM International Conference on Information and Knowledge Management* (pp. 1441-1450). Beijing, China.
- Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., ... Polosukhin, I. (2017). Attention is all you need. *Proceedings of the 31st Conference on Neural Information Processing Systems (NIPS 2017)* (pp. 6000-6010). Long Beach, CA.
- Yang, Q., Liu, Y., Chen, T., & Tong, Y. (2019). Federated machine learning: Concept and applications. *ACM Transactions on Intelligent Systems and Technology*, *10*(2), 1-19.