IJARSCT

ISSN (Online) 2581-9429

International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)

International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal

Impact Factor: 7.53

Volume 4, Issue 1, April 2024

FEDRETAIL: A Framework for Distributed Retail Data Analysis and Learning Toward E-commerce 5.0

Sumit Chongder
Department of Computer Science and Engineering
Maharashtra Institute of Technology - Art, Design and Technology University, Pune, India

Abstract*: Retail data analysis has been identified as a crucial component in the pursuit of E-commerce 5.0. The recent rapid development of Information and Communication Technology (ICT) has revolutionized retail data analysis by providing advanced technologies such as big data analysis and machine learning. However, the privacy of customers has become a significant concern, making retailers hesitant to share their customer data. This reluctance forms isolated data islands, hindering the realization of comprehensive retail data analysis. I propose a federated learning-based retail data analysis framework, FEDRETAIL, to address this challenge. This framework allows retail data analysis federations to be formed by several retailers. None of these retailers need to exchange their customer data with each other directly, and they always keep the data in their place to ensure their customers' privacy. I apply the FEDRETAIL framework to analyze a retail dataset via different federated learning paradigms. The experimental results show that our framework not only guarantees the customers' privacy but also effectively breaks the borders of data islands by achieving higher analysis quality. FEDRETAIL framework closely approaches the performance of centralized analysis, which requires data collection in a commonplace, posing a risk of privacy exposure.*

Keywords: Retail Data Analytics; Federated learning; Machine Learning; E-commerce 5.0.

REFERENCES

[1] S. Akter and S. F. Wamba, 'Big data analytics in E-commerce: a systematic review and agenda for future research', *Electronic Markets*, vol. 26, no. 2, pp. 173–194, May 2016, doi: 10.1007/s12525-016-0219-0.

[2] J. Park and H. Lim, 'Privacy-Preserving Federated Learning Using Homomorphic Encryption', *Applied Sciences (Switzerland)*, vol. 12, no. 2, Jan. 2022, doi: 10.3390/app12020734.

[3] H. Xiao, K. Rasul, and R. Vollgraf, 'Fashion-MNIST: a Novel Image Dataset for Benchmarking Machine Learning Algorithms', Aug. 2017, [Online]. Available: http://arxiv.org/abs/1708.07747

[4] L. Eglite and I. Birzniece, 'Retail Sales Forecasting Using Deep Learning: Systematic Literature Review', *Complex Systems Informatics and Modeling Quarterly*, vol. 2022, no. 30, pp. 53–62, 2022, doi: 10.7250/csimq.2022-30.03.

[5] K. J. Ferreira, H. A. Lee, and D. Simchi-Levi, 'Analytics for an Online Retailer: Demand Forecasting and Price Optimization'. [Online]. Available: www.ruelala.com

[6] G. Chaubey, P. R. Gavhane, D. Bisen, and S. K. Arjaria, 'Customer purchasing behavior prediction using machine learning classification techniques', *J Ambient IntellHumanizComput*, vol. 14, no. 12, pp. 16133–16157, Dec. 2023, doi: 10.1007/s12652-022-03837-6.

[7] Q. Wang, 'E-commerce Sites Search Results Relevance Prediction Based on Ensemble Approach', 2017.

[8] B. Venkatesh and J. Anuradha, 'A review of Feature Selection and its methods', *Cybernetics and Information Technologies*, vol. 19, no. 1, pp. 3–26, 2019, doi: 10.2478/CAIT-2019-0001.

[9] H. C. S. C. Lima, F. E. B. Otero, L. H. C. Merschmann, and M. J. F. Souza, 'A Novel Hybrid Feature Selection Algorithm for Hierarchical Classification', *IEEE Access*, vol. 9, pp. 127278–127292, 2021, doi: 10.1109/ACCESS.2021.3112396.

[10] S. Raizada and J. R. Saini, 'Comparative Analysis of Supervised Machine Learning Techniques for Sales Forecasting', *International Journal of Advanced Computer Science and Applications*, vol. 12, no. 11, pp. 102–110, 2021, doi: 10.14569/IJACSA.2021.0121112.

[11] I. Ullah, B. Raza, A. K. Malik, M. Imran, S. U. Islam, and S. W. Kim, 'A Churn Prediction Model Using Random Forest: Analysis of Machine Learning Techniques for Churn Prediction and Factor Identification in Telecom Sector', *IEEE Access*, vol. 7, pp. 60134–60149, 2019, doi: 10.1109/ACCESS.2019.2914999.

[12] P. Lalwani, M. K. Mishra, J. S. Chadha, and P. Sethi, 'Customer churn prediction system: a machine learning approach', *Computing*, vol. 104, no. 2, pp. 271–294, Feb. 2022, doi: 10.1007/s00607-021-00908-y.

[13] Y. Liu, T. Fan, T. Chen, Q. Xu, and Q. Yang, 'FATE: An Industrial Grade Platform for Collaborative Learning With Data Protection', 2021. [Online]. Available: https://www.fedai.org.

[14] I. Kholod *et al.*, 'Open-source federated learning frameworks for IoT: A comparative review and analysis', *Sensors (Switzerland)*, vol. 21, no. 1, pp. 1–22, Jan. 2021, doi: 10.3390/s21010167.

[15] H. R. Roth *et al.*, 'Empowering Federated Learning for Massive Models with NVIDIA FLARE', Feb. 2024, [Online]. Available: http://arxiv.org/abs/2402.07792

[16] M. Arafeh, A. Hammoud, H. Otrok, A. Mourad, C. Talhi, and Z. Dziong, 'Independent and Identically Distributed (IID) Data Assessment in Federated Learning', in *Proceedings - IEEE Global Communications Conference, GLOBECOM*, 2022, pp. 293–298. doi: 10.1109/GLOBECOM48099.2022.10001718.

[17] H. Zhu, J. Xu, S. Liu, and Y. Jin, 'Federated learning on non-IID data: A survey', *Neurocomputing*, vol. 465, pp. 371–390, Nov. 2021, doi: 10.1016/j.neucom.2021.07.098.